Structured Electronics Design - A conceptual approach to amplifier design - 3rd ed.

Anton J.M. Montagne

- basic design theory
- design of application-specific amplifiers
  - in CMOS, Bipolar, and BiCMOS technology
  - in PCA technology
- using SLiCAP (python) and SPICE
- summary of background knowledge

TU Delft OPEN

Structured Electronics Design

A conceptual approach to amplifier design

Structured Electronics Design

A conceptual approach to amplifier design

- basic design theory
- design of application-specific amplifiers
  - in CMOS , Bipolar and BiCMOS technology
  - in PCA technology
- using SLiCAP (python) and Spice
- summary of background knowledge

Anton J.M. Montagne

1st edition 2018
2nd edition 2021

3rd edition
ISBN 978-94-6366-712-8 (Ebook)
ISBN 978-94-6366-711-1 (Paperback)
DOI: https://doi.org/10.59490/t.2023.005

Keywords: electronics, design

Scientific theories deal with concepts, never with reality. All theoretical results are derived from certain axioms by deductive logic. In physical sciences the theories are so formulated as to correspond in some useful sense to the real world, whatever that may mean. However, this correspondence is approximate, and the physical justification of all theoretical conclusions is based on some form of inductive reasoning.

A. Papoulis [Papoulis1965][1]

[1] Athanasios Papoulis. *Probability Random Variables and Stochastic Processes*. McGraw-Hill, New York, 1965

To Ernst Nordholt

Contents

Preface

Structured Electronics Design

Structured Electronics Design offers strategies, methods, and techniques for electronic circuit design.

The complexity of analog electronics

Many people consider analog electronic circuit design complex. This is because designers can achieve the desired performance of a circuit in many ways. Together, theoretical concepts, circuit topologies, electronic devices, their operating conditions, and the system's physical construction constitute an enormous design space in which it is easy to get lost. For this reason, analog electronics often is regarded as an art rather than a solid discipline.

Experienced designers and intuitive design

At first glance, there doesn't seem to be a straightforward way to design analog circuits. In daily practice, experienced designers intuitively use all these degrees of freedom to modify and combine known solutions into new ones. However, intuition is knowledge of which the origin has become unclear. It results from a personal internalization process and cannot be shared with novices who have yet to acquire this knowledge. Therefore, intuition cannot and should not be a basis for design education.

The design of electronic circuits can be taught, shared, and understood if presented in a structured way with a distinct formulation of design goals, strategies, methods, and techniques. Moreover, it requires a clear distinction between theoretical concepts and their physical implementations.

Structured Electronics Design

Rather than taking numerous existing circuits as a starting point, it is much more effective to start a new design with a clear understanding of

1. The application and its environment
2. The required functionality
3. The implementation technology and physical operating mechanisms
4. Theoretical concepts, design strategies, methods, and techniques

All of this needs to be combined to achieve an acceptable performance-cost ratio.

Structured Electronics Design is a systems engineering approach to the design of analog electronics. It places analog electronics design in the perspective of information processing and provides a top-down design method with a bottom-up awareness. The hierarchical design process has a similar structure at each hierarchical level.

Important questions driving the design process are

1. How is the information present in the electrical signals?

2. Which physical signals present in the application environment may degrade the quality of the signal processing?

3. In which way and to what extent is the application sensitive to all kinds of information processing errors?

4. Which physical operating principles exist for implementing[2] the information processing functions, and to what extent are they available in the desired technology?

5. Do the fundamental physical limitations of information processing and technological or economic limitations cause unacceptable design risks?

6. Which design methods and techniques are available to maximize the performance to costs ratio of the design?

[2] also: *materializing*.

The first three questions emphasize the top-down approach, the next two illustrate bottom-up awareness, and the last shows the conceptual design approach.

Structured Electronics Design

1. Defines a step-by-step design process based on the above questions, and indicates which *design question* is answered at each design step.

2. Properly orders and orthogonalizes the design steps, based on solid principles from physics, signal processing, control theory, and network theory.

3. Helps circuit designers efficiently and effectively obtain the desired performance to costs ratio.

4. Provides a solid foundation for circuit design education and automation.

5. Has been developed and educated since the 1980s at the Delft University of Technology.

About the author

Anton Montagne (Leiden, 1953) received his master's degree in electrical engineering in 1984 at the Delft University of Technology. In 1983, he joined Philips Semiconductors in Nijmegen where he designed analog integrated circuits for audio and video applications. At Philips, he also set up training courses on analog electronics. In 1986, he cofounded the product development company Product Partners where he carried out many analog designs in the field of instrumentation. In 1989, together with Catena Microelectronics, Delft University of Technology, and the Institute of Microelectronics in Stuttgart, he cooperated in the development of an intensive training course, covering many topics of analog information processing. Since 1997, he has worked as an independent consultant, trainer, and designer in the field of analog electronics. Over the past 38 years, he developed analog electronics for instrumentation and communication systems for the industry and carried out many training courses on analog electronics at, amongst others, Catena Microelectronics, Philips Semiconductors, Philips Medical Systems, NXP, Ericsson, Plessey, Texas Instruments, ASML, TNO, Bruco IC design, 3T, Carl Zeiss SMT, TMC and ASMPT.

Anton Montagne is the inventor and coinventor of patents in the fields of position sensors, imaging, charge-coupled devices, and high-stability crystal oscillators.

Since 2016 Anton Montagne is coaching students and giving lectures and masterclasses "Structured Electronics Design" at the Delft University of Technology.

Acknowledgments

This book is the spin-off of workshops I have given and the design projects I have participated in over the past 35 years. Writing such a book is never the result of one person's effort. I am indebted to many people I cooperated with and the many students I have taught. They taught and inspired me, and they have contributed in their very own way to its creation.

The design approach for analog electronic signal processing systems, presented in this book, was suggested around 1987 by Dr. Ir. Ernst H. Nordholt. At that time, we both cooperated in educating circuit designers around Europe. Ernst suggested that analog electronic signal processing systems could be designed with a limited number of basic information-processing functions and references. As a result of the fundamental physical limitations of information processing, and due to technological limitations, their performance would deviate from that of their idealized function concepts. The number of signal processing errors, however, could be reduced through the application of a limited number of so-called error-reduction techniques. Limitation of both the number of basic functions and the number of techniques that can be exploited for their improvement helps to solve day-to-day design problems in a structured way and facilitates circuit design education. Moreover, it can be a basis for partial automation of the complex design process.

In the subsequent years, developed design courses based on this approach. In our day-to-day design work, we applied and improved this design method. I am very grateful to have cooperated with Ernst and to be inspired by him. Without this cooperation, this book would not have been written.

Despite my interest in signal processing, during my education in analog electronics, I was primarily focused on all kinds of aspects of circuit design. I am therefore very grateful to have cooperated with Dr. Ir. Huib Dane. His ability to explain complex topics from statistical signal processing helped me with my professional development and inspired me in developing training courses for the design education of professionals.

It has been a pleasure and honor to work together with Catena Microelectronics in Delft. Since 1999 this company offered me the possibility to educate their novice analog IC designers, and in this way, it contributed very much to the development of this material.

I am also very grateful to cooperate with The High-Tech Institute and Ir. Hans Vink. Since 2010. They organize my training courses, which allows me to spend my time teaching, serving my customers, and writing this book.

A special word of gratitude is reserved for my friend Dr. Ir. Jan Deiman. I am indebted to him for the pleasant conditions he offered me in his guest house in Bali, which facilitated the writing of this book in a most inspiring way.

I also want to thank Ir. Dr. Chris Verhoeven from the TU Delft for offering me the opportunity to coach students during the execution of their master thesis and for giving guest lectures and masterclasses at the TU Delft. Cooperation with him very much encouraged me to complete this work. Thanks to him, this book is now used for teaching *Structured Electronics Design* at the Delft University of Technology and, I have the pleasure of working with him. Our conversations about the structured design method, orthogonality, basic functions, and error reduction techniques, were very inspiring.

I also had many fruitful discussions about this material with Ir. Jeroen van Duivenbode, a fellow at the TU Eindhoven. He inspired me in writing the MATLAB version of SLiCAP and contributed the name *servo function* used in this book. I would like to thank Jeroen for this contribution and our pleasant and inspiring cooperation.

While working with this book, many students gave me valuable feedback. I am very grateful for their contributions and encourage everyone to inform me about errata and future improvements.

Edition 3, June 2023
Delft, The Netherlands
Anton Montagne

1 Introduction

1.1 Electronics

The invention of the *Audion* by De Forest, patent [Forest1907][1], can be regarded as the birth of electronics. Since then, around the world, thousands of amateurs, scientists, and engineers have contributed to its development. Nowadays, concepts developed in the first decennia of its existence are continuously being adapted for application to new technologies, while new ones are still being developed.

[1] L. De Forest. Wireless Telegraphy, Januari 1907

1.1.1 Analog Electronics

Although signal processing is mainly implemented digitally nowadays, analog electronics still play a significant role in all kinds of electronic equipment. On the one hand, fast digital processing and strict EMC requirements require knowledge about the continuous nature of electrical signals. On the other hand, interfacing with sensors, actuators, transmission lines, or communication channels often requires information embedded in analog or multilevel digital signals.

Electronics deals with the *manipulation of electrical currents with nonlinear electrical devices*. According to this definition, electronics does not deal with the meaning of electrical signals. However, when designing electronic circuits, one cannot ignore their information processing tasks. This is a consequence of the fundamental physical and technological limitations of information processing:

> The amount of information that can be processed by any physical system is limited. This is due to the addition of noise, the limited availability of power and the limitation of the rate of change of signals.

In addition, the physical principles for the implementation of information processing functions are usually imperfect and contribute to information processing errors. For example, the nonlinear behavior of semiconductor devices and vacuum tubes introduces technology-dependent errors into intended linear systems.

Hence, we first need definitions for the *intentional behavior* of information processing systems.[2] Secondly, we need knowledge about the manifestation of information processing errors due to physical and technological limitations. These manifestations depend on how the information is present in the signal. For example, binary digital signals will be less sensitive to nonlinearity than multi-valued digital signals or analog signals.

[2] The intentional behavior is also referred to as: the ideal behavior, or the conceptual behavior.

Usually, information processing happens in an environment where different noise sources adversely affect the signal quality.

Another aspect of information processing is its costs. Electronic information processing requires matter, space, and electrical energy. It also contaminates its environment with temperature rise and EMI.[3]

[3] EMI: Electromagnetic interference can be considered as noise.

Apparently, during the design of electronic information processing systems, the designer has to deal with many aspects.

At first glance, there does not seem to be a straightforward way to design such systems. However, the complexity of its design is not unique for analog electronics. Designing cars, airplanes, production equipment, test equipment, and many other modern products is highly complex. Proper structuring of complex design processes is a proven key to their success!

1.1.2 Structured design

There exists considerable literature about structured design methods. It is suggested that interested readers study *systems engineering*, for example [BlanchardEA1998][4] and [KossiakoffEA2003][5]. Particularly those who fear a conflict between creativity and structured design methods are encouraged to read [Terninko1998][6] and [Altshuller1999][7] about TRIZ, a Russian acronym for *Theory of Inventive Problem Solving*. TRIZ is a generalized design theory that offers discipline and technology-independent design methods and techniques.

In this book, we will present a structured approach to the design of negative feedback amplifiers, based on clear concepts from systems engineering, information processing, network theory and control theory. The design approach is also *inspired* by TRIZ.

This book elaborates the work presented in [Nordholt1983][8] and the work presented in [Verhoeven2003][9]. It uses SLiCAP for deriving and solving design equations.[10] By doing so, it provides a solid basis for (partial) automation of amplifier design.

1.1.3 This chapter

In this chapter, we will introduce and define the basic concepts of *Structured Electronics Design*. In section 1.2, we will summarize basic concepts and techniques from *Systems Engineering*, and in section 1.3, we will review some concepts from electronic information processing. In section 1.4, we will combine these concepts and techniques and outline the principles of *Structured Electronics Design*. In section 1.5, we will summarize what you will know after studying this book, briefly discuss its contents, and suggest how it can be organized in courses.

1.2 Selected topics from systems engineering

In essence, *Structured Electronics Design* is a systems engineering approach to analog electronics design. There exist many different definitions of systems engineering, the one below is taken from `https://en.wikipedia.org/wiki/Systems_engineering`:

> **Systems engineering** is an interdisciplinary field of engineering and engineering management that focuses on how to design, integrate, and manage complex systems over their life cycles. At its core, systems engineering utilizes systems thinking principles to organize this body of knowledge. The individual outcome of such efforts, an **engineered system**, can be defined as a combination of components that work in synergy to collectively perform a useful function.

In this section, we will briefly summarize some topics from systems engineering. In section 1.2.1, we will introduce some basic concepts, such as innovation, development, research, the invention, the product life cycle, the hierarchical organization of the design process, and the idea of considering a design process from a risk management point of view.

In section 1.2.2 we will present a basic design process that can be used at any hierarchical level of the design.

All sub-processes of this design process result in various kinds of data. Design documents are views upon this data, captured over time. In section 1.2.3, we will describe these results in more detail.

[4] Benjamin S. Blanchard and Wolter J. Fabrycky. *Systems engineering and analysis*. Prentice-Hall, Inc, New York, 1981. ISBN: 013-135047-1

[5] Alexander Kossiakoff and William N. Sweet. *Systems engineering principles and practice*. John Wiley and Sons, Inc., New Jersey, 2003. ISBN: 0-471-23443-5

[6] Terninko et al. *Systematic Innovation, An introduction to TRIZ*. CRC Press LLC, USA, 1998

[7] Altshuller. *The Innovation Algorithm, TRIZ, systematic innovation and technical creativity, Translated by Lev Shulyak and Steven Rodman*. Technical Innovation Center, Inc., USA, 1999. ISBN: 0-9640740-4-4

[8] Ernst H. Nordholt. *Design of High-Performance Negative Feedback Amplifiers*. Delft Academic Press / VSSD, 1 edition, 1983-2006. ISBN: 9789040712470

[9] C.J.M. Verhoeven, A. van Staveren, G.L.E. Monna, M.H.L. Kouwenhoven and E. Yildiz. *Structured Electronic Design*. Kluwer Academic Publishers, Boston - Dordrecht - London, 2003. ISBN: 1-4020-7590-1

[10] SLiCAP; **S**ymbolic **Li**near **C**ircuit **A**nalysis **P**rogram, written by the author (free, open-source, Python-based).

1.2.1 Basic concepts

Innovation

Innovation can be defined as a process in which the implementation of new ideas results in new markets, new technologies or new products.

Product, market and technology development

Mature companies innovate by developing only one of these three at a time:

- Market development:

  Development of new markets for existing product-technology combinations

- Technology development:

  Development of new technologies for existing product-market combinations

- Product development:

  Development of new products for existing market-technology combinations

Simultaneous development of combinations of the above is called diversification. Diversification is considered risky. In practice, it has a very low success rate.

Innovation plans are often presented in road maps. Such maps illustrate the successive development of products, markets, and technologies over time.

Research

Research is a process in which one acquires knowledge about something. It must be separated from innovation. Knowledge about products, markets, and technologies, is the basis for their development, and missing knowledge must be identified and acquired before starting a development that needs it.

The invention

The innovation process in start-ups usually differs from that in mature companies. The latter ones have a well-defined product-market-technology portfolio and innovate by improving or extending products, markets, or technologies. Start-ups often start with a basic idea based on a discovery or an invention, while they don't yet have a well-defined product-market-technology portfolio. The source of the invention can be a dream, an annoyance, something that somehow pops up in the mind of an engineer, but in many cases also: the result of a structured design process!

Product life-cycle processes

In this book, we confine ourselves to the design of application-specific amplifiers. One important aspect is to determine the requirements of such amplifiers. In general, the product requirements follow from the interests of the stakeholders of the product life cycle processes. Figure 1.1 shows an IDEFo model of the product life cycle processes.

The IDEFo drawing convention for processes is shown in Figure 1.2. A process, is a collection of activities that generates output (right) from its input (left). The process is controlled by its control input (top), and the activities are based on resources, methods, or mechanisms (bottom); see `https://www.idef.com/`.

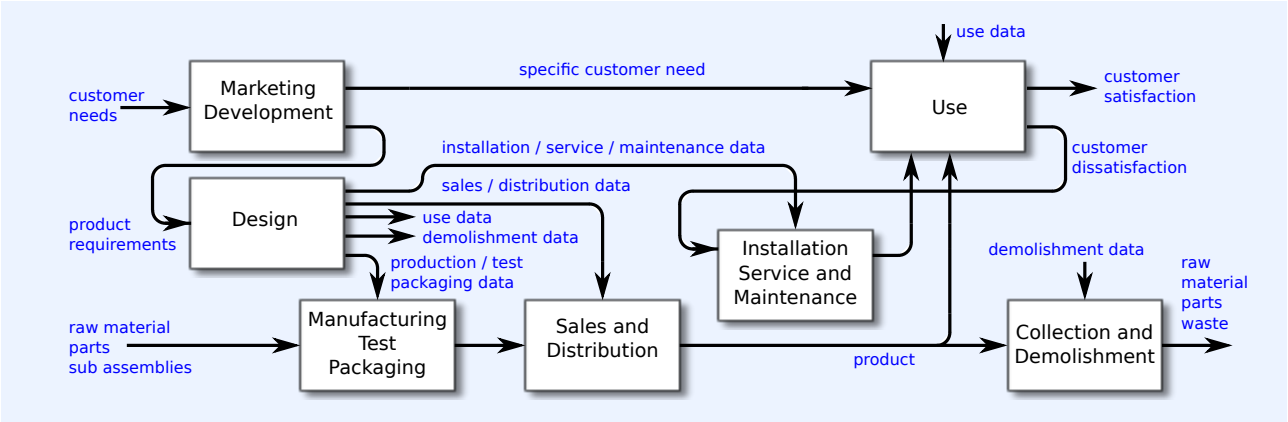


Figure 1.1: IDEFo product life cycle process model. The viewpoint of this model is the data that needs to be generated during design.

The life cycle processes shown in Figure 1.1 are:

1. Marketing
2. Design
3. Manufacturing/test/packaging
4. Sales and distribution
5. Installation service and maintenance
6. Use
7. Collection and demolishment

The viewpoint of the model from Figure 1.1 is the data to be generated during the design process. The figure clearly shows that during the design process, data is generated for almost all other life cycle processes. Therefore, the stakeholders of almost all life cycle processes contribute to the product requirements in one way or another.

Fortunately, many of their requirements are covered by regulations and standards. Standardization of components, materials, production methods, and many other aspects enable a short time-to-market and facilitate global mass production.

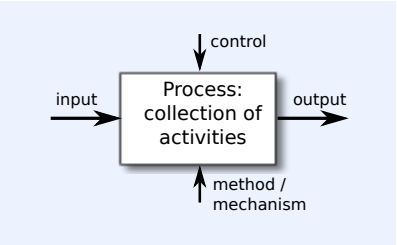


Figure 1.2: Business process model according to the IDEFo standard.

Hierarchically organized design processes

In general, engineers solve complex problems by breaking them down into less complex. In this way, they create a hierarchical structure in which the level of detailing increases at each level of hierarchy. This process ends when all product parts can be purchased or manufactured. Hence, designing is a top-down, hierarchically structured process.

The materialization of the design, however, proceeds bottom-up. Hierarchical levels in this phase correspond with those in the design phase. This way of working makes it possible to perform predefined integration tests of physical (sub)systems at each hierarchical level.

A design process with an identical structure at each hierarchical level facilitates the management of design projects. In section 1.2.2 we will present such a design process.

Risk management

We have seen that the design of a product is a top-down process, while its materialization is a bottom-up process. As a result, the feasibility of a product

is proven only when all components and parts can be purchased or manufactured. Therefore, potential design risks may not manifest themselves until the end of the design process. This, of course, is unacceptable because it causes loops in the design process.

There are several ways to avoid design loops resulting from unforeseen risks:

1. Keep the innovation level low:
   (a) Separate research from development
   (b) Avoid diversification
   (c) Identify and justify assumptions
2. Start design projects with a risk analysis
3. Solve design risks at the hierarchical level of their appearance.[11]

[11] Don't let them unresolved so they appear at the next hierarchical level.

From the above, it is clear that product development can be regarded as risk management. Risks, with the greatest product of their assumed probability of occurrence and their assumed impact, should always be addressed first.

From a risk management point of view, one should not start a design detailing familiar parts of the product. Design risks that will manifest themselves later while designing the unfamiliar parts may force the designer to reconsider the entire structure of the product. As a result, requirements of these previously detailed parts may change drastically, or worse, these parts may no longer be required.[12]

[12] Low-hanging fruit may not be consumed first.

Show stoppers

In this book, we will pay attention to the early identification of *show stoppers*.

> A show stopper is something that stops or could stop the progress of the design process, such as a performance-cost ratio that cannot be achieved.

Experienced designers are often intuitively aware of the consequences of economic and technological constraints on the design. Therefore they usually intuitively account for them at an early stage of the design process. A novice designer, who is not yet fully aware of the impact of such limitations, may encounter a show stopper for a design proposal at a later stage and may then be forced to reconsider the earlier selection of this proposal.

In general, design risks increase with increasing levels of innovation. Conducting research is the way to reduce those risks.

Figure Of Merit (FOM)

Throughout the design process, designers need to select the most promising solution from a set of possible solutions. Such selections are generally based on the *performance-cost* ratio of the different solutions. A *Figure Of Merit* (FOM) is the most compact way to represent the performance-cost ratio. It is the ratio of the weighted product of the performance parameters and the weighted product of the cost factors of a solution:

$$\text{FOM} = \frac{\text{product of weighted performance measures}}{\text{product of weighted cost factors}} \tag{1.1}$$

1.2.2 Basic design process

In this section, we will introduce a basic design process to be used at any hierarchical level. The underlying idea is to consider the product itself, as well as all of its parts, as objects. These objects or combinations thereof perform functions in a physical environment and at the expense of resources, such as matter, energy, time, and space.

This basic design process is depicted in Figure 1.3. It consists of a series of activities that together define a composition of objects that fulfills the requirements of the main object. The output at the lowest hierarchical level is purchase or production data for the object. At higher levels, the outputs are initial object performance requirement specifications of subsystems. These initial specifications are the inputs at the next hierarchical level. The input at the highest hierarchical level, is the initial specification of the product.

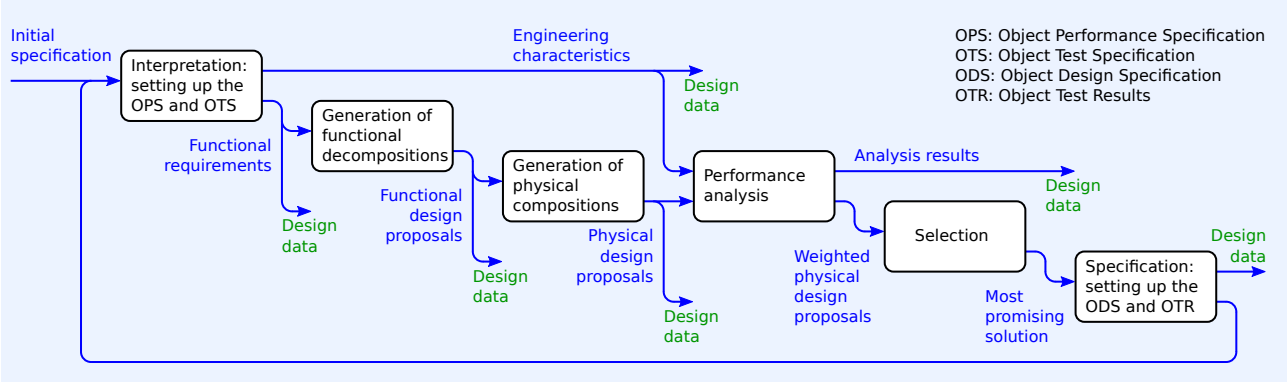


Figure 1.3: This Functional Flow Block Diagram shows the basic object design process applied at any hierarchical level of the design process.

The different activities of this basic design process are elucidated below.

Object Performance Specification (OPS) and Object Test Specification (OTS)

The first activity in the basic design process is the translation of the incoming initial specification into an Object Performance Specification (OPS) and an Object Test Specification (OTS). The OPS contains a variety of requirements and conditions arising from life cycle processes. The OTS specifies the test methods for these requirements. Below is a list of topics that are included in the OPS.

1. Functional requirements

   Specification of the functions that have to be performed by the object.

2. Reliability requirements

   Specification of the required reliability level of the performance. This is usually expressed in parameters, such as MTTF: mean time to failure, MTTR: mean time to repair, MTBF: mean time between failures, etc.

   Failure Mode Effect Analysis (see: [McDermottEA1996][13]) can be used to specify the probability of occurrence, the detectability, and the action to be taken at specific failure modes.

   Reliability requirements can be accounted for in the FOM by defining them as performance parameters.

3. Safety requirements

   Specification of the required safety level of the performance. This is usually done by referring to applicable safety standards.

   Like reliability requirements, safety requirements can be accounted for in the FOM by defining them as performance parameters.

[13] Robin E. McDermott, Raymond J. Mikulak, Michael R. Beauregard. *The basics of FMEA*. Productivity Inc., Portland, USA, 1996. ISBN: 0-527-76320-9

4. Performance requirements

   Specification of the required quality level of the performance. This is done by defining the desired values of performance parameters.

The Object Performance Specification

1. Introduction

1.1. Primary function

A description of the main function of the product.

1.2. Application description

A description of the way in which the product should be applied.
- External equipment that has to be used
- Interaction with its environment
- etc.

1.3. Overview of functions and interfaces

A block diagram of the functions and interfaces that should be included.
These functions and interfaces (ports) will be described in more detail in a separate section of the document.

2. Overall requirements

2.1. Environmental conditions

Life-cycle process	Manufacturing	Testing	Certification	Transport	Storage	Installation	Operation	Service	Demolishment
Temperature									
Shock									
Vibration									
Humidity									
Air pressure									
Cleanroom standard									
Optical									
ESD									
EMI									

whatever is relevant | during any of the life-cycle processes | Give values or refer to standards

2.2. Reliability

MTTD	MTTR	MTTF
diagnose	repair	normal operation
MTBF		

2.3. Safety

2.3.1. Human safety
- Injury
- Accessibility
- Fire
- Polution

Refer to standards: CE, UL

2.3.2. Product safety
- Short circuit protection
- Overload
- Over voltage
- Mechanical damage

2.3.3. FMEA: **F**ailure **M**ode **E**ffect **A**nalysis

RPN: **R**isk **P**riority **N**umber (> 300 Critical, < 125 No action required)

Severity x Occurence x Detection

S = 1 ... 10: 1: no problem, 10: dangerous
O = 1 ... 10: 1: never, 10: very often
D = 1 ... 10: 1: sure, 10: impossible

2.3.4. Single-point failure:
- Failure does not propagate (cause more failures)
- Failure does not cause a hazardous situation

2.4. Cost factors

Life-cycle process	Design	Manufacturing	Testing	Certification	Transport	Storage	Installation	Operation	Service	Demolishment
2.4.1. Methods										
2.4.2. Means										

Methods/tools required | during any of the life-cycle processes | Give values or refer to standards | Examples: Test tools, IPC standards, PCB classes, tools for reparing and servics, etc.

2.4.3. Dimensions Dimensions of the product, referring to a mechanical drawing

2.4.4. Mass Mass of the product

2.4.5. Output to the environment Might be relevant during different life-cycle processes
- Power dissipation
- Emission (EMI, light, radiation, etc.)
- Waste

2.4.6. Price Target cost price of the product

3. Physical interface requirements

3.1. Mechanical drawing with physical interfaces

- Connectors
- Mounting holes
- Other interfaces to the environment, such as , optical indicators

3.2. Description of physical interfaces

3.2.1. Description of the connectors
- Type
- Part number
- BCB mount / panel mount
- Description of the pins
  - Number
  - Name
  - Schematic reference RefDes
  - Signal type

3.2.2. Description of display/indicators
- LED indicators
- Displays

3.2.3. Description of other interfaces
- Jumpers
- DIP switches
- Whatever is relevant

3.3. Grounding and shielding concept

Drawing and description of the way of interconnection of the signal reference, the shields and the safety ground with other system parts.

4. Power supply requirements

Description of the required power suppl(y)(ies)
- Voltage range
- Noise and ripple voltage
- Over voltage
- Source/sink current rating
- Short circuit current
- Isolation voltage w.r.t. safety ground
- Additional requirements

5. Functional requirements

Functions that are listed in the introduction will be described in separate sub sections.

5.x. Specification of function x

5.x.1. Description of the function
- A diagram with text explaining the desired behavior of the function.
- Expressions for input-output relations of the function or of sub functions
- Performance measures and test methods with pass/fail criteria of the function or of sub functions

5.x.2. Interfaces connected to this function

5.x.2.x. Different interfaces can be described in separate sub sections
- References to interfaces to this function that have been described in (3).
- Specification of the signals to these interfaces:
  - Voltage / current / impedance / noise level / bandwidth / rise time / fall time / equivalent network / test conditions
  - For indicator use specific descriptions: LED: color, display: rows / columns / menu, etc

Figure 1.4: Example of a structure of an object performance specification of a Printed Circuit Board Assembly (PCA).

5. Environmental conditions

   Specification of operating conditions. These are the conditions under which the object must perform its function with the specified quality, safety, and reliability.

   If relevant, environmental conditions during other life cycle processes, such as storage and transport should also be specified.

6. Resources

   Specification of resources (also called *cost factors*) that are needed or

available during the different life cycle processes. Operational resources are those necessary for the operation of the product. Examples are matter, time, and space.

Sometimes, the means or the lack of means during other life-cycle processes also must be specified.

7. Interaction with the environment
   Every physical product influences its environment. Examples of such influences are temperature rise and pollution. Allowed and forbidden influences have to be specified.
8. Interfacing with the environment
   Specification of interfaces with other products and operators, such as mechanical interfaces, electrical interfaces, and user interfaces.

The collective noun for all measurable quantities from 2-8 is: *engineering characteristics* (see Figure 1.3).

Generation of functional decompositions

As stated before, engineers solve complex problems by breaking them down into less complex. This way of working results in the presented hierarchical design method. *Functional decomposition* is a collective term for techniques that create a set of sub-functions that together perform the complete functionality of an object. Functional decomposition techniques are numerous. Brainstorming sessions, literature study, patent study, or techniques, such as *Objectives Tree Analysis* as described by Cross [Cross1989][14], can all be used for this purpose.

[14] Nigel Cross. *Engineering Design Methods, Strategies for Product Design*. John Wiley & Sons Ltd., England, 1989. ISBN: 0-471-94228-6

Generation of physical compositions

The functions defined during functional decomposition must be performed by physical objects. The first step is to find a physical mechanism or *operating principle* to perform the function. This mechanism must then be embodied in a physical system that consists of objects realized in some technology and supplied with energy.

Performance analysis and selection

To determine the FOM of the different solutions, the designer must analyze their performance parameters and cost factors. Such analyses can be carried out through modeling and computer-aided symbolic and/or numerical analysis. If necessary, these analyses can be supported by testing and measuring the performance and costs of so-called *functional models*. A functional model, or *Fumo* for short, is a physical model of an object that contains only the parts needed to determine specific performance aspects or cost factors. In this way, a Fumo differs from a prototype, which is a fully manufactured object that is used for the evaluation of performance and cost.

After the designer has a complete picture of the performance and the costs of all design proposals, the most promising one[15] is selected for further engineering.

[15] The one with the best performance-cost ratio.

Object Design Specification (ODS) and Object Test Results (OTS)

The Object Design Specification is a report that summarizes the object's design process. It discusses the functional decomposition, possible physical implementations, the performance and cost analysis, the figure of merit, and the considerations regarding the selection of the design proposal.

The Object Test Results gives the results of the tests defined in the Object Test Specification.

1.2.3 Design data and documents

Throughout the design process, all kinds of data are generated. Examples are:

- Calculations
- Simulation models
- Technical drawings and schematics
- Purchase data
- Production data
- etc.

Parts of this data often need to be included or elucidated in documents, such as the Object Performance Specification and the Object Design Specification. Documents can be considered a specific view on the design data at some moment. These documents are intended for project management, and for securing the acquired knowledge.

1.3 Electronic information processing

The structured design approach presented in this book puts the design of analog electronics in the perspective of information processing. To fully appreciate this approach, one must first understand some basic concepts from information processing.

Information processing systems

Figure 1.5 shows the architecture of a simple information processing system. The information processing system processes the information acquired by some source and passes the results to an observer. All takes place in an environment in which noise is present.

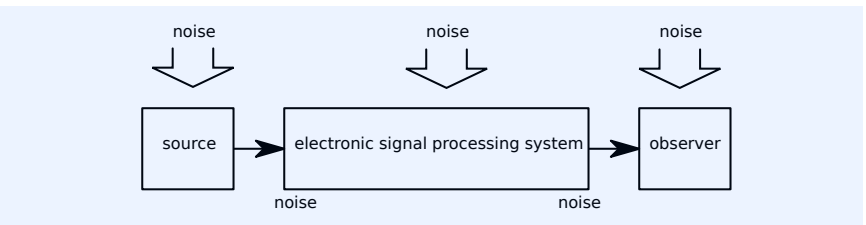


Figure 1.5: Architecture of a data acquisition system.

In the above description, we used some terms that need further explanation:

- What is a signal?
- What is data?
- What is information?
- What is noise?

Figure 1.6 shows a more complex situation. The system depicted in this figure also controls the process. Hereto, it provides control signals to actuators that influence the progress of the process.

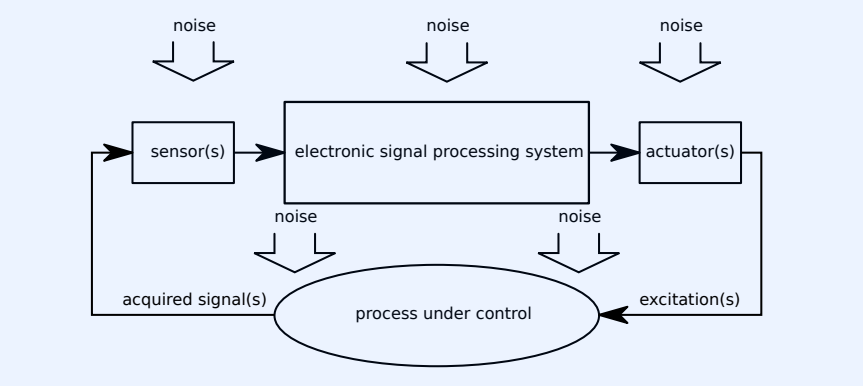


Figure 1.6: Architecture of a system that acquires information from a process, processes the information, and controls the process.

In the section 1.3.1, we will define important concepts from information processing.

In section 1.3.2 we will show that the amount of information that can be processed by a physical system is limited, and introduce the fundamental physical, technological and economical limitations.

1.3.1 Basic concepts

- Signal

  A signal is a physical quantity that contains meaningful data. In electronic systems, we use electrical signals: the time-varying values of electrical current or electrical voltage.

- Data

  Data comes from the Latin word datum, which means: something given. Here, data are properties or details of the signal that may contain information.

- Information

  Information is the meaning of the data. One could say: if the data provides an answer to a question, it contains information.

- Interpretation

  Interpretation is the process of extracting the information from the data.

- Environment

  The environment is the total of surroundings in which this all takes place. Each environment imposes specific limitations to environmental conditions, such as temperature, shock, vibration, humidity, sensitivity to radiation of any kind, etc.

- Noise

  Noise is a physical quantity whose data is meaningless. If added to a signal, it compromises the retrieval of the information from the signals. Noise can be of any kind: thermal, optical, acoustic, electromagnetic, etc.

- Interference

  Interference is considered noise of which specific properties are known. This knowledge can be used to compensate for the adverse effects of interference on the retrieval of information.

- Signal source

  The signal source provides a signal.

- Observer

  The observer is the device or the person that needs the information.

- Excitation

  Excitation is a signal with specific properties that lets a process generate responses that provide information about the process.

- Transducer

  A transducer is a device that transfers the signal from one physical domain to another.

- Sensor

  A sensor is a transducer that converts a non-electrical signal into an electrical signal.

- Actuator

  An actuator is a transducer that converts an electrical signal into a non-electrical signal.

1.3.2 Limitations of information processing

The amount of information that can be processed by a physical system is limited. We will distinguish fundamental physical, technological and economical limitations.

Fundamental physical limitations

According to Shannon [ShannonWeaver1963][16] the amount of information processing errors can be arbitrarily low if the amount of information transported over a linear channel, perturbed by white noise is limited to the *Channel Capacity* C:

$$C = B \log_2 \frac{P+N}{N} \left[\frac{\text{bits}}{\text{s}}\right], \tag{1.2}$$

where B [Hz] is the bandwidth of the channel, P [W] the signal power, and N [W] the noise power. Any physical system adds noise and suffers from power and speed limitations. Hence, the addition of noise, the limitation of the signal power, and the limitation of the rate of change of signals are regarded as fundamental physical limitations of information processing.

[16] Claude E. Shannon and Warren Weaver. *The Mathematical Theory of Communication*. The University of Illinois Press, Ubana, 1 edition, 1963. ISBN: 0-88179-205-5

Technological limitations

The technology not only determines the strength of the above fundamental physical limitations. It also introduces processing errors that result from the non-ideal implementation of the physical operating mechanism. An example is the distortion of a signal in amplifiers due to their non-linear behavior.

Economical limitations

A shortage of the required resources, such as matter, energy, time, and space, can also lead to errors in information processing. Like technological limitations, they determine the strength of the manifestation of the fundamental physical limitations.

1.4 Structured Electronics Design

In sections 1.2 and 1.3, we have summarized topics from systems engineering and information processing. Together with statistical signal processing, control theory and network theory, these disciplines are the basis for *Structured Electronics Design*.

The idea behind *Structured Electronics Design* is the creation of a design language. In the next section, we will give an outline of such a design language.

1.4.1 Outline of the design language

Like any language, the design language must have a set of words and rules[17], and we want to create an unlimited amount of stories using only relatively small sets of words and rules. The stories we want to write with the design language are electronic information processing systems.

[17] The grammar of the language.

The words of the design language are the so-called *basic functions*. We state:

> The functional behavior of any electronic information processing system can be obtained from a combination of *basic functions*.

The *materialization* of a function in an *object* requires the implementation of one or more *physical operating principles* in some technology. Therefore, physical operating principles are part of the grammar of the design language.

Due to the physical, technological, and economic limitations, the performance of these objects may be insufficient or the cost to achieve their desired performance may be too high. The application of techniques and methods from network theory, control theory, and statistical signal processing can improve the performance-cost ratio. Hence, these so-called *error reduction techniques* are also part of the grammar of the design language. Application of error reduction techniques changes the functionality or optimizes the implementation of one or more operating principles.

We now have the design language with its words and its grammar:

> With relatively small sets of basic functions, physical principles, and error reduction techniques, we can create an unlimited amount of information processing systems.

Figure 1.7 illustrates the design of a physical composition from a functional decomposition.

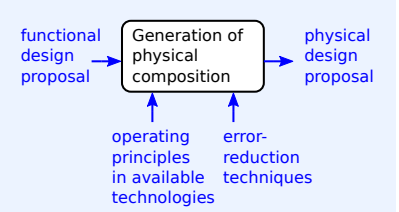


Figure 1.7: Generation of a physical composition from a functional decomposition according to Structured Electronics Design.

At this moment, the above may sound rather abstract, and the concepts introduced require further explanation. In section 1.4.2, we will introduce a set of basic functions and discuss their materialization in basic objects. Materialization or implementation is the application of one or more physical operating principles in available technologies. In electronic systems, the implementation of functions is based upon operating mechanisms of electronic devices.

The reader is assumed to be familiar with elementary physical principles from electrostatics and electrodynamics. The operation of semiconductor devices is summarized in Chapter 4.

In section 1.4.3 we will summarize important performance aspects and costs factors of electronic information processing systems and give examples of Figures Of Merits.

In section 1.4.4, we will give some examples of error reduction techniques. In this book, we will apply:

1. Balancing techniques
2. Negative feedback techniques

3. Frequency compensation techniques

In the concluding section 1.4.5 we will discuss the differences between *Structured Electronics Design* and the traditional approach to the design of electronics.

1.4.2 Basic functions and basic objects

Top-down definition of basic functions

One way of defining basic functions is to consider elementary mathematical operations required for information processing. Information processing deals with performing operations on time-varying physical quantities.

Below, is a list with basic information processing and reference functions for the functional decomposition of information processing systems:

1. Addition of signals
2. Subtraction of signals
3. Multiplication of signals
4. Integration of signals
5. Differentiation of signals
6. Selection of signals

   Mathematical selection is simply comparing two variables. The physical nature of the variables is not of interest. From the viewpoint of information processing, it is useful to distinguish comparison in different physical domains:

   (a) Selecting of signal levels
   (b) Selecting of signal frequencies
   (c) Selecting of signal time intervals
   (d) Selection of signal locations

   Selecting also requires a reference, which brings up the group of basic references:

7. References

   (a) Level reference
   (b) Frequency reference
   (c) Time reference
   (d) Location reference

8. Shifting of signals

   Mathematically shifting of a signal is adding a constant to it. From the viewpoint of information processing, the physical domain is of interest:

   (a) Level shift (adding a constant)
   (b) Frequency shift (modulation)
   (c) Time-shift (memorization)
   (d) Location shift (transportation)

9. Change the power of a signal

   In the context of information processing, it is also meaningful to consider the two types of variables associated with physical signals: across variables and through variables. Examples of across variables are voltage, force, and pressure. Examples of through variables are current, velocity, and flow. The product of an across variable and its associated through variable at a certain time instant is the instantaneous signal power. The ratio of an across variable and its associated across variable shows a relation with the impedance of the signal source or load. For maximum power transfer from a source to a load, the load impedance needs to be the complex conjugate of the source impedance. The maximum power that can be delivered by a signal source is called the available power of that source.

   Basic functions related to the power transfer are:

   (a) Impedance transformation (optimization of power transfer)

   (b) Attenuation (reducing the available power)

   (c) Amplification (increasing the available power)

10. Solving

    Solving an equation is considered an essential mathematical operation often applied in control theory. In network theory, the nullor provides this function. A nullor consists of a nullator and a norator. The nullator sets a condition and a norator that provides the dependent variable to be solved.

The set of functions listed above is not a minimum set of orthogonal functions. As an example, consider frequency shift. This function can be resolved in multiplication of the signal with a frequency reference[18] and selection in the frequency domain. Another example is the level shift. This function can be replaced with addition and a level reference.

[18] A periodic signal

Bottom-up definition of basic functions

A *basic object* is a physical implementation or the materialization of a *basic function*. There exists not necessarily a one-to-one mapping of basic functions onto basic objects. The materialization of basic functions in basic objects strongly depends on the physical operating principles in the available technologies. Moreover, a physical operation principle in some technology may contribute an attractive function not listed above. Bipolar transistors, for example, exhibit an exponential relation between their base-emitter voltage and their collector current. This relation holds over many decades. Resolving the functionality of a system into $\exp(x)$ functions may then result in a straightforward mapping of the functional decomposition on the physical composition in bipolar technology. For example, the design of integrated circuit analog multipliers is based upon the application of the exponential function and its inverse. In this example, multiplication is decomposed into lin-log conversion, addition, and lin-exp conversion.

Implementation of basic functions in basic objects

There are many ways to materialize basic functions. As an example, consider the implementation of impedance transformation. Both transformers and impedance matching networks can be used for this purpose. These two implementations use different operating principles. A transformer uses the principle of electromagnetic induction, while the operation of a matching network is based upon resonance.

Another example is the selection of time intervals. This function requires a switch. Switches can be realized with nonlinear electronic devices or with

electromechanical devices, such as relays. These two implementations use different physical operating principles and technologies.

In this book, we mainly focus on the design of application-specific amplifiers. The amplification function increases the available power of a signal. Its embodiment requires a power source and a mechanism to control the power transfer from this source to the load with the input signal.

1.4.3 Performance, costs, and FOM

As discussed in section 1.2.2, the performance and the costs of the object are specified in the *Object Performance Specification* (OPS). An in-depth treatment of the structure of the Object Performance Specification for electronic information processing systems is beyond the scope of this book. Companies usually use their own templates for it. Figure 1.4 shows an example of the contents of an Object Performance Specification for a *Printed Circuit board Assembly* (PCA).

In section 1.2.2 we also introduced the Figure Of Merit as the most compact representation of the performance-cost ratio. Throughout the design process, design decisions can be based on this FOM.

An example of a FOM for electronic information processing systems is the amount of information that can be processed per unit of energy and per Euro:

$$FOM = \frac{\text{bits}}{\text{Joule} \cdot €} \tag{1.3}$$

In this book, we confine ourselves to the circuit design of amplifiers. It is impossible to define a figure of merit for amplifier design because the relevant performance aspects and the relevant cost factors strongly depend on the application and the technology.

1.4.4 Error-reduction techniques

In most cases, the materialization of a function in an object proceeds stepwise. This is because the performance-cost ratio of a first physical design proposal is not as required and needs improvement. The performance-cost ratio can be improved by applying error-reduction techniques. We distinguish two groups of error-reduction techniques.

1. Improvement of the performance-cost ratio while maintaining the way the information is coded in the signal.

   Techniques that belong to this group are:

   (a) Optimization techniques

   (b) Compensation techniques

   Examples of compensation techniques are:

   i. Balancing techniques

   ii. Frequency compensation techniques

   iii. Error feedforward techniques

   (c) Negative feedback techniques

2. Changing the way the information is coded in the signal.

   Examples of techniques in this group are:

   (a) Sampling

   (b) Quantization

   (c) Modulation

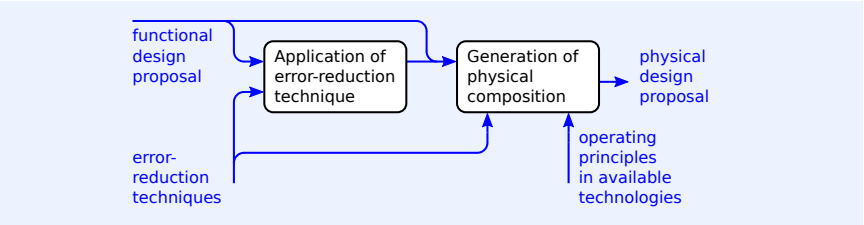


Figure 1.8: Detailed process model of the implementation of functions in objects according to Structured Electronics Design.

(d) Coding

Figure 1.8 shows a more elaborated model of the physical design. It shows that the application of an error-reduction technique may change the functional decomposition.

1.4.5 Differences with traditional analog design

Traditionally, the design of analog electronics predominantly consists of analysis and improvement of the behavior of known circuits. These improvements, in turn, usually consist of applying commonly known changes, and the impact of such changes on the cost factors often remains unclear. In addition, these circuits and their improvements often carry the name of their inventor or their topology. As a result, it is difficult or impossible to recognize the intended functionality, the applied operating principles, and the applied design methods with their intended effects. Therefore, the traditional design approach is a rather heuristic process.

A clear separation between functions, the applied operating principles, error-reduction techniques, their intended effect, and their technological implementation is the distinguishing difference of *Structured Electronics Design*. Moreover, *Structured Electronics Design* uses circuit analysis for setting-up design equations for well-defined performance parameters and cost factors, and numeric simulation is considered a verification method rather than a design method.

The advantages of *Structured Electronics Design* for analog design automation and design education are evident, and presented below.

1. There exist two different approaches to design automation:

   (a) A big-data oriented approach

   This approach searches a circuit database circuits for viable circuits that solve a set of design problems. However, new problems cannot always be solved with known solutions. In many cases, these solutions require modifications. Modification of existing solutions without knowledge of the underlying principles is risky. It may result in show stoppers, and it requires extensive simulations to prove the viability under all possible conditions. The above can only partly be solved using parameterized designs.

   (b) An algorithmic approach

   Structured Electronics Design essentially is an algorithmic approach. It is very well suited for design automation. This approach identifies show stoppers at an early design stage and requires less extensive simulations because it considers the viability of a solution throughout the design process. Moreover, an algorithmic approach can invent new solutions because it does not start with an existing solution but builds up from well-defined concepts.

2. Design education requires an approach that facilitates the internalization of newly acquired knowledge. Therefore, the structured design approach is well suited for design education. Each step of the design process can be motivated because it uses principles from physics, signal processing, control theory or network theory, and last but not least: principles familiar to the student.

1.5 This book

The approach to education in analog electronics in this book strongly differs from the traditional approach that focuses on circuit analysis rather than synthesis. This book gives the synthesis of analog electronics a solid scientific foundation: it gives clear motivations for design steps to be taken. Knowledge of physics, signal processing, network theory, and control theory is indispensable for taking design decisions, while knowledge of mathematics is required for deriving design equations.

This book is primarily intended for educating students in Electrical Engineering but is also very well suited for designers educated in applied physics, mechatronics, or control theory. It is divided into two parts:

1. Design of application-specific amplifiers
2. Background knowledge

1.5.1 What you will know after studying this book

1. You will know the characteristic properties of ideal(ized) amplifiers, and you will be able to derive the functional requirements for amplifiers from their application:
   (a) The input and output impedance
   (b) The source-to-load transfer
   (c) The port isolation
2. You will be able to characterize the non-ideal behavior of amplifiers, and you will know to derive performance requirements from the application description:
   (a) The small-signal noise behavior
   (b) The small-signal dynamic behavior
   (c) The instantaneous nonlinear behavior
   (d) The dynamic nonlinear behavior
   (e) The influence of temperature and ageing
3. You will come to understand other relevant design aspects of amplifiers, such as:
   (a) Environmental conditions
   (b) Cost factors
4. You will understand the operating principle of amplification
   (a) You will be able to evaluate the available power gain of a two-port
   (b) You will understand the concept of biasing
5. You will be understand the operation and modeling of active devices

(a) BJT: Gummel-Poon model: relate the parameters of small-signal and noise model to device parameters, geometry and operating point

(b) JFET: Shichman and Hodges model: relate the parameters of the small-signal and noise model to the device parameters, the geometry and the operating point

(c) MOS: Shichman and Hodges model: relate the parameters of the small-signal and noise model to the device parameters, the geometry and the operating point

(d) MOS: Meyer capacitance model: relate the small-signal capacitances to the device parameters, the geometry and the operating point

(e) MOS: Ward-Dutton capacitance model: relate the small-signal capacitances to the device parameters, the geometry and the operating point

(f) MOS: EKV model: relate the parameters of the small-signal and noise model to the device parameters, the geometry and the operating point

6. You can design a basic CS or CE stage amplifier stage

(a) You can select a transistor or design the transistor geometry considering its noise contribution and source impedance

(b) You can select a transistor or design its geometry considering its load and its required voltage and current drive capability

(c) You can design the operating point and bias the basic stage

(d) You can determine the source to load transfer, the input impedance and the output impedance of a CE and CS stage for resistive and/or R∥C type source and load impedances

7. You can apply balancing techniques

(a) You will understand the concepts of additive compensation and balancing

(b) You know the behavioral modifications resulting from anti-series connection

(c) You can apply this to evaluate the behavior of a differential pair

(d) You can bias a differential pair using only common-mode current sources

(e) You know the behavioral modifications resulting from complementary-parallel connection

(f) You can apply this to evaluate the behavior of a push-pull stage

(g) You can bias a push-pull stage using only common-mode voltage sources

8. You will be able to design low-noise and power efficient amplifier structures for arbitrary port impedance and port isolation requirements with the aid of feedback techniques, balancing techniques and isolation techniques:

(a) Direct feedback and indirect (model-based) feedback

(b) Nonenergic, passive and active feedback

(c) Balancing and port isolation

9. You will be able to relate the properties of the components in the feedback network to important performance aspects and costs factors of the amplifier:

(a) Inaccuracy
(b) Noise
(c) Nonlinearity
(d) Power efficiency

10. You will be able to model individual performance aspects of voltage-feedback and current-feedback operational amplifiers:

(a) Noise behavior
(b) Gain and input and output impedances, including their dynamic behavior
(c) Offset and bias quantities
(d) PSRR and CMRR

And you will become familiar with other relevant performance aspects, such as:

(a) Input voltage range
(b) Output voltage and current drive capability
(c) Voltage slew rate

11. You will know in which way and to what extent the equivalent input noise sources of an operational amplifier affect the noise performance of the negative feedback amplifier.

12. You will be able to apply the asymptotic gain negative feedback model to derive budgets for properties of the operational amplifiers and the passive components of the negative feedback amplifier

13. You will be able to design the dynamic behavior of a negative feedback amplifier with the aid of frequency compensation techniques:

(a) Phantom-zero compensation
(b) Pole-splitting by means of capacitive feedback
(c) Pole-splitting by means of pole-zero canceling
(d) Resistive broadbanding
(e) Bandwidth reduction
(f) Nested loops

And you will qualitatively know in which way these high-frequency compensation techniques interact with other performance aspects:

(a) Noise behavior
(b) Accuracy
(c) Distortion
(d) Overdrive recovery

14. You will know in which way and to what extent the temperature behavior of an operational amplifier affects the operating point of a negative feedback amplifier, and you will be able to derive requirements for the temperature behavior of the operational amplifier from the performance requirements of its application.

You will be able to apply techniques to reduce this influence:

(a) AC coupling

(b) Negative-feedback biasing

(c) Auto-zero techniques

15. You can apply negative feedback to a CS or CE stage

(a) You can design single-transistor feedback stages

(b) You will be able to relate the properties of local feedback stages to those of the CE or CS stage by considering the behavioral modifications resulting from application of negative feedback

(c) You will understand that the CD and CC stage can be considered as non-inverting, unity-gain, negative feedback voltage amplifiers

(d) You will understand that the CG and CB stage can be considered as non-inverting, unity-gain, negative feedback current amplifiers

(e) You will understand that the current mirror can be considered as an inverting indirect feedback current amplifier

(f) You can apply balancing techniques to local feedback amplifier stages

(g) You can design cascode stages and the CC-CB and CD-CG cascade stages

16. You can design multiple-stage negative feedback amplifiers

(a) You will be able to define the type of input stage, its geometry and operating point on ground of its noise performance

(b) You will be able to define the type of output stage, its geometry and operating point on ground of its current and voltage drive capability

(c) You will be able to define the number of stages on grounds of the static accuracy, the low-pass and high-pass cut-off frequencies and the distortion of the amplifier

(d) You will know how to combine multiple stages in a controller

(e) You will be able to apply common-mode feedback biasing techniques and optimize the biasing concept of a multiple-stage controller

17. You can design the bias sources

(a) You will be able to define the specifications for the bias sources considering their influence on the noise performance, the dynamic performance, the static accuracy and the temperature stability

(b) You will be able to design the bias sources

1.5.2 Part 1: design of application-specific amplifiers

The design method presented in this book has been introduced by Nordholt [Nordholt1983][19]. It is further developed at the TU-Delft by Verhoeven, van Staveren, Monna, Kouwenhoven and Yildiz [Verhoeven2003][20]. This book builds further on this material. It is based upon course material for post-graduate education in *Structured Electronics Design*, developed by the author.

The first part of this book presents the design method for application-specific, negative feedback amplifiers.

[19] Ernst H. Nordholt. *Design of High-Performance Negative Feedback Amplifiers*. Delft Academic Press / VSSD, 1 edition, 1983-2006. ISBN: 9789040712470

[20] C.J.M. Verhoeven, A. van Staveren, G.L.E. Monna, M.H.L. Kouwenhoven and E. Yildiz. *Structured Electronic Design*. Kluwer Academic Publishers, Boston - Dordrecht - London, 2003. ISBN: 1-4020-7590-1

Application, modeling and characterization of amplifiers

In Chapter 2, we will give applications of amplifiers, define the amplification function and discuss the ideal behavior of amplifiers. We will then discuss the nature of information processing errors of practical amplifiers and introduce related performance aspects.

In this chapter, we will also discuss some physical appearances of amplifiers. The construction of amplifiers results from trade-offs between performance and costs made during the design process. Cost factors of amplifiers can be numerous, and some examples of cost factors will be given. For design decisions, it is convenient to have a figure of merit at one's disposal, and some suggestions will be given.

We will then discuss the modeling and the characterization of the ideal behavior of amplifiers.

The fundamental physical limitations of information processing and the technological limitations cause the behavior of practical amplifiers to deviate from their desired behavior. How the observer experiences these deviations depends on how the information is present in the signal. This will be elucidated with some examples.

Principle of amplification

In Chapter 3, we will study how the principle of amplification is materialized in electronic amplifiers. We will show that signal amplification can be obtained through the proper interconnection of electrical power sources and passive, electronic devices. Examples of such devices are MOS transistors, bipolar transistors, and vacuum tubes. It will be shown that these devices, when used in combination with power sources, can provide an available power gain larger than unity; which is a distinguishing property of amplifiers. For this reason, these devices are often called *active devices*. Combining power sources with such devices is usually called *biasing*. This chapter concludes with a conceptual approach to biasing. We will show that the quiescent operating conditions of a correctly biased amplifier stage do not depend on the stage's drive and termination resistances.

Modeling of semiconductor devices

Knowledge of the operation and modeling of modern semiconductor devices is indispensable when designing analog electronic circuits. However, in-depth treatment of semiconductor physics and modeling techniques is beyond the scope of this book. Chapter 4 briefly describes the construction, operation, and modeling of BJTs, JFETs, and MOS transistors. The main goal is to provide a basic understanding of the construction and operation of the devices as a basis for modeling the device's performance during various stages of the design. The Gummel-Poon model for BJTs [GummelPoon1970][21] and the Shichman and Hodges model for JFETs [ShichmanHodges1968][22] will be presented and simple models for hand calculations will be derived from them. For MOS devices, the basic Shichman and Hodges model, the Meyer capacitance model [Meyer1971][23] and the Ward and Dutton capacitance model [Ward1998][24] will be discussed. The latter model is used in the BSIM3 small-signal model [BSIM3-1995][25], often used in SPICE.

With the EKV model [EnzVittoz2006][26], the small-signal parameters can be written expressed in the device geometry, the drain current, and the drain-source voltage. These expressions are valid from weak to strong inversion, including short-channel effects. In this way, it is possible to design the small-signal dynamic transfer and the noise performance independent from the biasing circuitry. SLiCAP has built-in small-signal models for CMOS18 devices whose parameters depend on the channel width and length, the opera-

[21] H. K. Gummel and H. C. Poon. An integral charge control model of bipolar transistors. *Bell Syst. Tech. J.*, 49(5):827–852, May-June 1970

[22] H. Shichman and D. Hodges. Modeling and simulation of insulated-gate field-effect transistor switching circuits. *IEEE J. Solid-State Circuits*, 3(3):285–289, September 1968

[23] J. Meyer. MOS models and circuit simulation. *RCA Review*, 32:42–63, 1971

[24] D. Ward and R. Dutton. A Charge-Oriented Model for MOS Transistor Capacitances. *IEEE Solid-state Circuits*, sc-13(5):703–708, October 1978

[25] Cheng, Y. et al. *BSIM3 Version 3.0 Manual*. University of California/Berkeley, Electronics Research Laboratory, 1995

[26] Christian C. Enz, and Eric A. Vittoz. *Charge-based MOS Transistor Modeling*. John Wiley & Sons Inc., 2006. ISBN: 978-0-470-85541-6

tion current and voltage, and only a few EKV model parameters. This way of working facilitates the design of CMOS circuits using the inversion coefficient or the transconductance efficiency, as described by Binkley[Binkley2008][27].

[27] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

Basic amplification with CS stage

In Chapter 5, we will study the performance limitations and the design considerations for the common-source (CS) stage, which can be considered as the basic MOS transistor amplifier stage. At a later stage, we will show that other MOS amplifier stages can be derived from the CS stage through the application of error reduction techniques such as compensation or negative feedback.

The common-emitter (CE) stage can be regarded as the basic BJT amplifier stage. Performance limitations and design considerations for this stage will be added in a future version of this book.

For these basic amplifier stages, we will discuss the way in which their performance can be altered by design. We will see that the designer does not have many degrees of freedom to optimize the performance of such elementary amplifier stages. The operating conditions, the fabrication technology, and the geometry or the device type are the only design variables at the disposal of the designer to optimize their performance-to-costs ratio. Moreover, the various performance aspects of single-transistor amplifiers cannot be designed independently[28] and compromises between performance aspects often need to be made.

[28] This is called "orthogonal design".

Application of balancing techniques: differential pair and push-pull stages

If the desired performance-cost ratio of an amplifier cannot be achieved with basic amplifier stages, error-reduction techniques can be applied for its improvement. In Chapter 6, we will study the application of balancing techniques and their impact on the performance-cost ratio of an amplifier. Two particular applications of balancing techniques will be discussed in more detail: anti-series connection of equal devices, and parallel connection of complementary devices.[29]

[29] Shortly: anti-series connection and complementary-parallel connection.

Anti-series connection of basic amplifier stages provides a four-terminal stage with an odd transfer characteristic and improved isolation between the input port and the output port. The behavioral modifications that are a result of series, complementary-series, and anti-series connection, will be investigated. It will be shown that the properties of the MOS and the BJT differential pair can easily be related to those of the CS and the CE stage, only by considering such behavioral modifications. We will see that, when applied in a truly balanced environment, the small-signal transfer and the noise performance of the differential pair can equal those of the basic CE or CS stage at the costs of four times the area and four times the operating current. As a result of the anti-series connection, offset voltages are canceled and the bias sources change from differential-mode to common-mode. The large-signal transfer of these anti-series stages has an odd characteristic with current saturation.

The complementary-parallel connection of amplifier stages split an input signal into a push and a pull current. These currents can be much larger than the quiescent operating current of the stage and can be used as high-efficiency amplifier stages. The CMOS inverter is a complementary parallel stage.

The behavioral modifications resulting from parallel, anti-parallel, and complementary-parallel connection of amplifiers or amplifier stages, will be investigated. The properties of the MOS and the BJT push-pull stage will be related to those of their constituting CS or CE stage.

Not all performance aspects of an amplifier or amplifier stage can be improved with balancing. The (gain) accuracy, the dynamic transfer, and their temperature dependencies of balanced stages equal those of their unbalanced version. Improvement of those aspects requires the application of techniques with better error-reduction capabilities.

Design of negative feedback amplifier configurations

Negative feedback is very a powerful error-reduction technique. The characteristics of negative feedback amplifiers are primarily fixed with *reference networks* or *feedback networks*. These feedback networks create an accurate copy of the source signal from the load signal. They can be passive or nonenergic elements of which the electrical behavior is accurately defined. A high-gain *controller* or *error amplifier* minimizes the difference between the source signal and this copy. In this way, the properties of the feedback amplifier are predominantly determined by its feedback networks; the error amplifier provides the available power gain but does not define the source-to-load transfer.

The design of application-specific negative feedback amplifiers is discussed in Chapter 7. In this chapter, we will discuss the design of negative feedback amplifier configurations for specific sources and loads. We will see that the amplifier types, introduced in Chapter 2 can be synthesized by combining voltage and or current sensing at the load with voltage and or current comparison at the signal source. During the conceptual design of negative feedback amplifiers, nullors are used as ideal controllers. The nullor can be considered a network element with an infinite available power gain and no speed limitation. At a later stage of the design, these nullors need to be replaced with error amplifiers, implemented with amplifier stages, or with operational amplifiers.

In this chapter, we will show that the amplifier types from Chapter 2 can all be designed using nonenergic feedback elements. Nonenergic feedback elements are lossless and have no energy storage. In addition, two of them, the ideal transformer and gyrator behave like natural two ports. Unfortunately, transformers are not ideal and suffer from energy storage and losses, while physical operating mechanisms for gyrators have not (yet) been discovered. The use of more practical passive feedback elements limits the number of negative feedback configurations. If this is the case, the application of *active feedback*, *balancing*, or *indirect feedback* may be considered.

The operational amplifier as controller in feedback amplifiers

Operational amplifiers are intended as controllers for negative feedback amplifiers. With their high voltage gain, high common-mode and differential-mode input impedance, high common-mode rejection ratio and low output impedance early types were versatile building blocks for negative feedback voltage amplifiers. Nowadays, the behavior of current-feedback and rail-to-rail output operational amplifiers strongly deviates from this behavior, which complicates their application.

Chapter 8 deals with the modeling of operational amplifiers. Aside from modeling all behavioral aspects with so-called macro models, attention will be paid to the modeling of individual performance aspects, which is considered to be more useful for deriving budgets for different performance limitations and for taking early-stage design decisions.

Another aspect that is limiting the application of operational amplifiers as universal controller is the fact that their output port, which is usually a high-efficiency push-pull output stage, has a split return path connected to both supply terminals. This imposes difficulties to the implementation of grounded current sensing techniques and limits the number of amplifier configurations that can be realized using solely operational amplifiers as con-

trollers. This limitation as well as various ways to deal with it will also be discussed in this chapter.

Introduction to biasing

An introduction to the biasing of negative feedback amplifiers will be presented in Chapter 9. In this book, we will advocate a strict separation between the design of the signal transfer and the design of the biasing of amplifiers and amplifier stages. At an earlier stage we have already shown that the principle of amplification requires the application of properly interconnected power sources and passive nonlinear electronic devices. Biasing refers to the derivation of all these power sources from the power supply source(s).

Biasing of (cascaded) amplifier stages will only be presented after we have discussed the design of the signal processing properties of an amplifier. The reason for this is that biasing of stages and of interconnected stages only needs to be done if the signal processing by the conceptually biased stages is adequate. Biasing of a configuration of which the signal processing is not according to the requirements is meaningless and regarded as a lost of valuable design time. In this introductory chapter, we will only discuss the consequences of errors that are a result of imperfect biasing of controllers. Such errors occur due to device tolerances and temperature deviations. These errors are usually modeled with the aid of equivalent-input offset and bias currents and voltages. Statistical description methods will be given and error reduction techniques to minimize their effects, will be discussed. Examples will be given for negative feedback amplifiers equipped with operational amplifiers, but the theory is not limited to these cases. Compensation, *AC coupling* and *negative feedback biasing* will be introduced as methods for the reduction of biasing errors. The latter two can only be applied if frequency components of the signal differ from those of temperature changes. These techniques establish a high-pass character of the signal transfer and design criteria for the high-pass cut-off frequency will be given. Proper high-pass filter characteristic can be established using frequency compensation techniques.

Modeling of negative feedback circuits

After we are able to design all kinds of application-specific amplifier configurations with nullors as controllers, we need to find specifications for practical controllers. To this end, we need a way of feedback modeling that facilitates a two-step design:

1. Design of the ideal transfer which is fixed by the feedback network
2. Design of an acceptable deviation from this ideal behavior caused by the nonideal controller.

The widely used feedback model introduced by Black [Black1934][30] provides accurate performance analysis of negative feedback amplifiers only under limited conditions. It does not account for the so-called direct transfer from the source to the load, and it assumes unilateral transfer and ideal sensing at the load and comparison at the source. As a result of these limitations, it is suited for the analysis of negative-feedback systems rather than for the two-step design of negative-feedback amplifier circuits. The feedback theory introduced by Bode [Bode1945][31] in 1945 and described by Chen [Chen1991][32] gives a method to analyze the stability of a feedback loop. Middlebrook [Middlebrook1975][33] introduced the double injection theory to measure the loop gain. However, all these models focus on stability analysis, rather than on a two-step design of a well-defined dynamic behavior of the feedback amplifier.

[30] Harold S. Black. Stabilized feed-back amplifiers. *Electrical Engineering*, 53(1):114–120, January 1934

[31] H.W. Bode. *Network Analysis and Feedback Amplifier Design*. Van Nostrand, New York, 1945

[32] W.K. Chen. *Active Network Analysis*. World Scientific Publishing Co. Pte. Ltd., Singapore, 1991. ISBN: 9971-50-912-1

[33] R.D. Middlebrook. Measurement of loop gain in feedback systems. *Int. J. Electronics*, 38:485–512, April 2001

[34] Solomon Rosenstark. A Simplified Method of Feedback Amplifier Analysis. *Transactions on education*, E-17(4):192–198, November 1974

The only feedback model that facilitates the two-step design is the *asymptotic gain model* as described by Rosenstark [Rosenstark1974][34]. This model shows that the design of a negative feedback amplifier can be performed in the two subsequent and independent steps that have been mentioned above. Moreover, the source-to-load transfer obtained from this model equals the one found from network analysis. The asymptotic gain model will be discussed in Chapter 10.

Setting up controller performance specifications

With the aid of the asymptotic gain model, we are able to relate performance aspects of the controller to those of the negative feedback amplifier. This enables use to derive budgets for the performance aspects of the controller, which is a minimum requirement for the two-step design approach described above. In Chapter 11 this will be done for the static accuracy, the nonlinearity and the bandwidth of the amplifier. We will find that:

1. The static error of a feedback amplifier sets a requirement for the controller's contribution to the static or DC loop gain.
2. The low-pass cut-off frequency of a feedback amplifier sets a requirement for the contribution of the controller to the gain-poles product of the dominant poles of the loop gain.
3. The static differential-gain error of the negative-feedback amplifier sets a requirement for the contribution of the controller to the static differential error to gain ratio of the DC loop gain.

Design conclusions for other performance aspects such as the high-pass cut-off frequency will also be derived. The derivation of budgets for noise and power losses of the controller has already been dealt with in Chapter 7.

In this chapter, we will also introduce techniques for the evaluation of the stability of negative feedback amplifiers. Techniques known from control theory, such as the Nyquist stability criterion, the Routh array analysis method and the root-locus technique will be summarized and elucidated with examples. Frequency compensation techniques for establishing the desired filter characteristics will be discussed at a later stage.

Frequency compensation

Frequency compensation techniques for establishing proper high-pass or low-pass filter characteristics will be extensively discussed in Chapter 12. Concepts and strategies for frequency compensation will be introduced and implementation examples will be given. Special attention will be paid to the impact that frequency compensation may have on other performance aspects such as bandwidth, linearity, overdrive recovery and noise. It will be shown that frequency compensation with the aid of phantom zeros is the most powerful method because it has the lowest interaction with other performance aspects. Implementation of both active and passive phantom zeros will be discussed and illustrated with examples. Other techniques such as pole-splitting techniques, resistive broadbanding and nested control will be discussed as well.

Design of local feedback stages

Amplifier stages that use a single, unbalanced or balanced CE or CS stage as controller are called *local feedback stages*. Local feedback stages can be used as single-stage amplifiers, or as stages in negative feedback amplifiers. In Chapter 13, we will discuss the design of local feedback stages. We will

show that the well-known CD stage or *source follower* as well as its bipolar version, known as the *emitter follower* or CC stage, can be considered as unity-gain voltage amplifiers that exploit nonenergic feedback and that have the CS or CE stage as controller, respectively. Similarly, the common-gate (CG) or the common-base (CB) stage can be considered as nonenergic negative feedback current followers. The advantage of such an description method is evident. If those stages are feedback versions of the basic CS or CE stages, then the loop gain, which can be regarded as a measure for the amount of negative feedback, indicates the extent to which their behavior deviates from that of the CE or CS stage. It will then become clear that commonly known properties, such as the low output impedance of the CC stage, are only true if the stage is driven from an impedance that establishes a relatively large loop gain. If such a stage is driven from a current source, accurate input voltage comparison cannot be performed, the loop gain will be low and the output impedance does not differ from that of a CS stage. In addition, since the CS and CE stages are nonenergic negative feedback amplifiers, they inherit the properties of nonenergic feedback amplifiers. Without further analysis it then becomes clear that the equivalent input noise sources of a CD or the CC stage equal those of the CS or the CE stage of which they are constituted, respectively. Similar things can be said about their power efficiency.

Aside from the design of CD, CC, CG and CB stages, the design of other basic local feedback stages, such as the series, the shunt stage and some dual-loop local feedback stages, as well as the application of balancing techniques will be discussed as well. A separate section will be devoted to the so-called *cascode* stage. This stage consists of a cascade connection of the CS and CG stage (MOS version) or a CE and a CB stage (BJT version). Its interesting properties makes it an ideal inverting amplifier stage in multiple-stage negative feedback amplifiers that ensures low interaction between stages. It will be shown that the CD-CG cascode and its bipolar version, the CC-CB cascode, can similarly be regarded as basic non-inverting amplifier stages.

Design of multiple-stage negative feedback amplifiers

High-gain amplifiers may be constructed from a cascade connection of amplifier stages. High-performance negative feedback amplifiers, however, require controllers that comprise multiple amplifier stages. The CS or the CE stage, the cascode stages, the local feedback stages as well as the balanced versions of all these stages may be candidates for amplifier stages in such a multiple-stage controller. In Chapter 14 we will discuss the design of multiple-stage controllers. We will show that by selecting a CS or CE (cascode) stage, or their balanced version, the controller will have the best possible noise performance if the noise performance of this stage is optimized for the given source impedance and feedback network(s). Similarly, by selecting a CE or CS (cascode) stage or its balanced version the contribution to the differential error to gain ratio of the loop gain will be as low as possible.

Design criteria for the number of stages and the preferred type(s) for intermediate stages will also be given. It is important to have a rough estimate for the number of stages at an early stage of the design. If the number of stages is more that two or three, frequency compensation may become difficult and one may consider to construct the amplifier from a number of multiple-stage feedback amplifiers. Nested feedback techniques will also be discussed and illustrated with examples.

The motivation of the type of stages inside the controller may also be driven from practical limitations such as the power supply voltage and the complexity of the biasing. In modern analog CMOS design, the low power supply voltage may put a serious constraint to the architecture of the controller. Since biasing considerations may seriously influence the design of

amplifiers, they need to be accounted for during all stages of the design process. However, this does not change the design approach for amplifiers. If the signal processing performance of an optimally designed signal path does not leave room for any degradation possibly resulting from biasing, the detailed design of the bias sources is of no use. Hence, the design of the noise performance, the bandwidth, the linearity and the frequency compensation should always be done before implementing the bias sources.

Biasing

The biasing and design of the biasing elements is performed in four steps:

1. Simplification of the biasing scheme.

   During the design of the *signal path* of the amplifier we use *conceptually biased* amplifier stages as introduced in Chapter 3. Such stages use four bias sources. During this design step, this biasing scheme will be simplified and the remaining bias sources will be replaced with the power supply and nonlinear resistive elements that exhibit a voltage or current source character. This step will be elucidated in Chapter 15.

2. Setting up specifications for the resulting bias sources.

   After a biasing scheme has been developed, the performance requirements for the bias sources need to be derived from error budgets for noise, bandwidth and nonlinearity. This step will also be elucidated in Chapter 15.

3. Design of the bias sources.

   The design of bias sources will be added to a future edition of this book.

4. Application of error reduction techniques for minimization of biasing errors resulting from device tolerances and temperature changes.

   This has been discussed in Chapter 9.

1.5.3 Part 2: background knowledge

Part 4 summarizes background knowledge and places it in the context of the design method.

1. Selected topics from signal, data and information modeling can be found in Chapter 16.

2. Selected topics from system modeling can be found in Chapter 17.

3. Selected topics from network theory are included in Chapter 18.

   This summary requires knowledge of linear algebra. Four topics are usually not found in other books, but helpful for a better understanding amplifier design:

   (a) The time constant matrix and its eigenvalues

   (b) Estimation of poles and zeros of a transfer function by network inspection

   (c) Decomposition of balanced circuits into differential-mode and common-mode equivalent circuits

   (d) Two-port conditions

4. Physical mechanisms, modeling and characterization of noise in electronic circuits is summarized in Chapter 19.

1.5.4 How to use this book

This book is organized in such a way that it can be used in three *subsequent* courses:[35]

[35] Each course requires the knowledge of its preceeding course.

1. An introduction course

   The introduction course summarizes the background knowledge that is required to study the BSc and MSc level courses. It comprises deterministic and random modeling of signals and systems, network theory, basic knowledge about noise in electronic circuits and hands-on experience with PYTHON, SLICAP and SPICE.

2. A BSc level course

   At the end of the BSc level course the students are able to specify and design an application-specific negative feedback amplifier using an operational amplifier as controller. This course is intended for board designers who need to design signal conditioning amplifiers between sensors and analog-to-digital converters and amplifiers that convert output signals from digital-to-analog converters to actuators. A chapter about the design of class D amplifiers will be added in future versions of this book.

3. An MSc level course

At the end of the MSc level course, the students are able to specify and design an application-specific amplifier using in Bipolar, BiCMOS or CMOS technology. This course is intended for board designers who want to construct amplifiers using both operational amplifiers and discrete transistors, and for IC designers. Table 1.1 gives an overview of the chapters for each class.

		Courses		
Chapter	Description	Intro	BSc	MSc
1	Introduction to structured electronic design		✓	
2	Amplifiers: application, classification, modeling & characterization		✓	
3	Principle of amplification			✓
4	Modeling of active devices			✓
5	Basic amplification: CS stage			✓
6	Balancing techniques			✓
7	Design of negative feedback amplifier configurations		✓	
8	Application and modeling of operational amplifiers		✓	
9	Introduction to biasing		✓	
10	Modeling of negative feedback circuits		✓	
11	Deriving controller requirements from amplifier specifications		✓	
12	Frequency compensation		✓	
13	Design of local feedback amplifier stages			✓
14	Design of multiple-stage negative-feedback amplifiers			✓
15	Design of controller biasing concepts			✓
16	Signal modeling (selected topics)	✓		
17	System modeling (selected topics)	✓		
18	Network theory (selected topics)	✓		
19	Noise in electronic systems (selected topics)	✓		

Table 1.1: Subsequent course programs: Introduction, BSc level and MSc level.

Design of application-specific amplifiers

2
Modeling and specification of amplifiers

2.1 Introduction to amplifier design

Amplification is the most important basic electronic information processing function. Electronic amplifiers are the physical objects that embody this function.

Application areas for electronic amplifiers are numerous. Amplifiers are applied in radio receivers to raise the level of an antenna signal to such an extent that digitization and/or demodulation can be carried out. In all kinds of electronic equipment, amplifiers are applied to adapt the level of sensor signals to the input range of analog to digital converters, or to adapt the output level of a digital to analog converter to the driving level of actuators. Amplification is also implemented in other basic functions, such as voltage and current references, oscillators, active filters, comparators, limiters and memory circuits. As a matter of fact, the principle of amplification is applied in basic digital cells, such as inverters, gates and digital memory circuits. Amplification of electrical signals is thus indissolubly connected with electronic information processing, and amplifiers can be found in many manifestations.

Figure 2.1: The 'Solo' is a high-end mono audio power amplifier, designed by the author.

2.1.1 Functionality

Intuitively, one might say that an amplifier provides its load with an enlarged copy of the source signal. According to this assumption, amplifiers would need one input port and one output port to which the signal source and load are connected, respectively. This, however, is not true. The signal power that an amplifier can deliver exceeds the *available power* of the signal source. Amplifiers therefore, have an additional port to which a power source needs to be connected. Due to the power from this source and the amplification mechanism embodied in the amplifying devices, the amplifier's so-called *available power gain* can exceed unity.[1] This property makes amplifiers fundamentally different from transformers and lossless electrical networks, such as, matching networks. Transformers and matching networks can have either a voltage or a current gain that exceeds unity, but their available power gain is limited to unity.

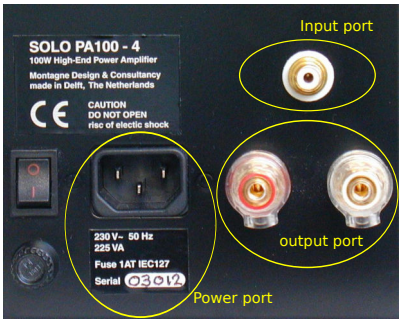


Figure 2.2: The back side of the 'Solo' clearly shows its input, output port and power port.

[1] The definition of the available power gain will be given in section 2.3.4.

The amplification mechanism itself is embodied in the *biased amplifying devices* from which the amplifier is constructed. Biased amplifying devices are combinations of power sources and devices such as MOS transistors, bipolar transistors and/or vacuum tubes. The principle of amplification will be elucidated in Chapter 3.

Figure 2.3 shows a model of an amplifier with its three ports: the input port, the output port and the power port. It gives a simple representation of the amplification mechanism: the source signal modulates the power transfer from the amplifier's power source to the load.

Ideally, the information at the input port will be copied to the output port and the input and the output signals will not be affected by variations of the power supply. For this reason, amplifiers are functionally modeled as two-ports. In a functional description, the power port is simply omitted.

In most cases, we do not want any reverse signal transfer from the output port to the input port. Amplifiers with this property are called *unilateral*.

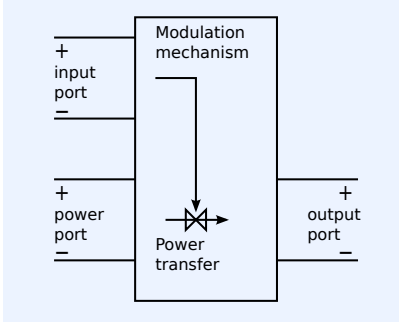


Figure 2.3: Amplifier with input port, output port and power port.

2.1.2 Definition

According to the previous description, we can define an amplifier as follows:

An amplifier is a physical object having at least three electrical ports:

1. An input port to which the source signal will be connected (referred to as the "input")

2. An output port to which the load will be connected (referred to as the ”output”)
3. An input port to which the power supply will be connected (referred to as the ”supply”)

An amplifier provides the load with an accurate copy of the source signal, and its available power gain exceeds unity.

This definition is a good starting point for a discussion of the quality of information transfer by the amplifier.

2.1.3 Information-processing quality

The quality of information processing tells us something about the amount of correspondence between the load signal and the source signal. By their physical nature, electronic amplifiers generate noise and are sensitive to electromagnetic fields (EMI[2]). These and other effects will adversely affect the correspondence between the source signal and the load signal. Fortunately, a unique correspondence between signal values at the load and signal values at the source at any time instant is not required. The correspondence between the load and the source signal has to be such that an observer can retrieve the relevant information from the load signal. Errors due to the fundamental physical limitation of speed, power and noise and those due to technological limitations are thus allowed, as long as the information can be retrieved. The manifestation of these information processing errors also depends on the physical construction of the amplifier. This physical construction, in its turn, shows a strong interaction with the cost price of the amplifier. Hence, the amplifier's information processing characteristics, its physical construction and the economical constraints[3] show a strong interaction.

[2] EMI: Electro Magnetical Interference.

[3] Also called cost factors.

2.1.4 Physical appearance

Audio amplifiers, antenna amplifiers and operational amplifiers are names for different objects that have at least one thing in common: they are all amplifiers. The meaning of the word ”amplifier” thus depends on one's perspective. A designer of audio power integrated amplifier circuits might use the term audio amplifier for a single IC, while the user of an audio amplifier refers to a complete product with housing and its user interface.

Generally, electronic amplifiers can have the following physical appearance:

1. A complete product, including housing, power supply and user interface
2. A hard wired electromechanical assembly
3. A printed circuit board assembly (PCA)
4. A thick or thin film assembly
5. An integrated circuit (IC)
6. A part of an integrated circuit.

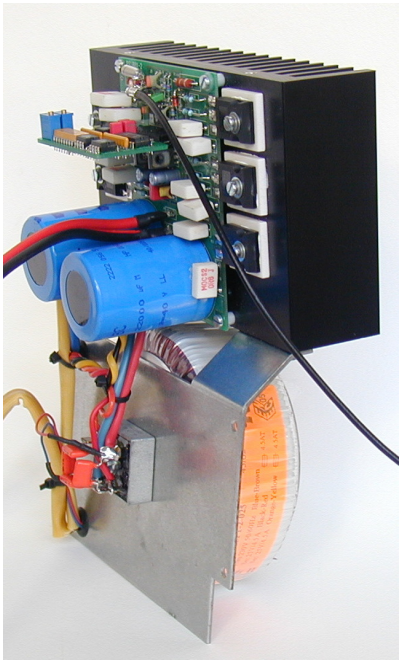


Figure 2.4: The inside of the 'Solo' clearly shows the PCAs comprising other amplifier circuits and biased active components that together constitute the main amplification function.

These physical realizations of amplifiers can be nested in a hierarchical way. A complete audio amplifier, for example, may consist of several printed circuit board assemblies and other parts that are mechanically and electrically interconnected. One or more of these printed circuit board assemblies can be an electronic amplifier itself. Such an amplifier, in its turn, may be built with discrete components, as well as with integrated circuit operational amplifiers. These integrated circuits also comprise interconnected electronic devices. The ability to amplify signals that finds its origin in the biased amplifying devices, is used in these different forms of amplifiers.

2.1.5 Cost factors and environmental conditions

The processing of information by amplifiers is not free of cost. It requires physical resources such as power, matter and space. The amplifier also generates heat and electromagnetic interference signals. Among others, these are called *costs factors*.

The quality of information processing should not be affected by normal changes in the operating conditions. The *environmental operating conditions*, or shortly *environmental conditions*, such as, interference signals, temperature range, humidity, shock and vibration should be specified.

The aim of a good design is to provide the required information processing capability for an acceptable level of cost factors under specific environmental conditions.

2.1.6 Figure of merit

In order to compare the performance-to-cost ratio of various amplifiers, it would be nice to have some figure of merit (FOM). Moreover, such a figure of merit, would help the designer in selecting solutions for specific design problems. There are various kinds of figure of merit. They usually relate the most important performance aspects to the most important cost factors, the degrees of importance usually being defined by the user. Some examples are listed below.

- Output power versus dimensions or weight
- Output power versus consumed power
- Output power versus product costs
- Noise figure versus power consumption
- Dynamic range versus power consumption
- Bandwidth versus power consumption
- etc.

2.1.7 This chapter

In section 2.2, we will start with the description of the functional behavior of an amplifier. At the start of the design, questions that need to be answered follow from the information that needs to be processed, the character of the source and the character of the load of the amplifier:

- Which electrical quantity (current or voltage) should be selected as the electrical input quantity of the amplifier?
- Which electrical quantity should be delivered to the load?
- Is the signal source electrically connected to the power supply or is it floating with respect to the power supply?
- Is the load electrically connected to the power supply or is it floating with respect to the power supply?

Based on this knowledge, we will make an inventory of amplifier types (or concepts) and model their (ideal) behavior with the aid of a two-port. This will be done in section 2.3.

Due to fundamental physical and technological limitations, the behavior of practical amplifiers will deviate from that of their idealized function concept

and information processing errors are inevitable. However, as long as the load signal shows a sufficiently large correspondence with the source signal, the observer will be able to retrieve the relevant information. Predictable errors do not necessarily degrade the information handling capability of the amplifier. Known errors can be compensated for in pre- or post-processing functions, thereby restoring the correspondence between the source signal and the load signal. Reproducible errors due to nonlinearity, for example, can be compensated for by adding opposite pre- or post-distortion to input or output signal, respectively. This also holds for reproducible errors that result from bandwidth limitation. However, we will consider the *conceptual amplifier* to behave *instantaneously*, *linearly* and *time-invariantly*, and model it accordingly.

In section 2.4, we will introduce performance parameters that describe the non-ideal behavior of amplifiers. We will discuss:

1. Errors arising from imperfect isolation between the input port and the output port, and between the power supply port and the two signal ports
2. Errors due to a nonzero transfer from the power port to the input port or vice versa, and from the power port to the output port or vice versa
3. Signal processing errors due to fundamental physical and technological limitations. We will study different kinds of errors resulting from:
   (a) The addition of noise and interference
   (b) Limitation of the current and voltage handling capability due to the physical limitation of power and due to the nonlinear signal transfer of the characteristics of biased amplifying devices
   (c) Small-signal bandwidth limitation and slew rate limitations due to the physical limitation of the rate of change of electrical currents and voltages
   (d) Ageing due to changes in the properties of electronic devices over time
   (e) Temperature changes

In section 2.5, we will describe the performance aspects of cascaded amplifiers and learn about error propagation in cascaded amplifiers. This helps us to relate properties of individual amplifiers to those of a chain of amplifiers.

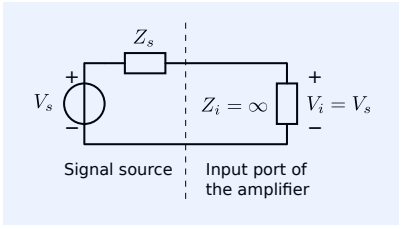


Figure 2.5: Interfacing of the signal source with the input port of the amplifier in cases in which the voltage of the signal source accurately represents the information, while the source impedance may be partly unpredictable.

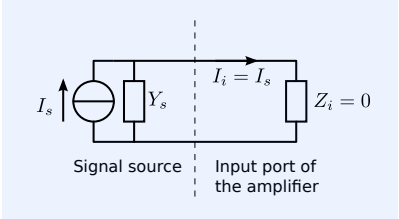


Figure 2.6: Interfacing of the signal source with the input port of the amplifier in cases in which the current of the signal source accurately represents the information, while the source admittance may be partly unpredictable.

2.2 Amplifier port requirements

The first step in amplifier design is to determine which electrical quantities must be used at the amplifier's input and output ports. If the source and the load of an amplifier consist of transducers, we have to select the electrical port quantity that shows the best correspondence to the physical input or output signal of the transducer.

1. If the open circuit voltage of a signal source is accurately related to the primary physical information, we will model the signal source as a voltage source in series with its source impedance. This impedance may be complex and depend on the signal level and on the temperature. For accurate (unity-gain) transfer of the information from the signal source to the input port of the amplifier, the impedance of the input port of the amplifier should be infinite. This has been elucidated in Figure 2.5.
2. If the short-circuit current of a signal source is accurately related to the primary physical information, we will model the signal source as a current source in parallel with its source impedance. The source impedance

may be complex and depend on the signal level and on temperature. For accurate (unity-gain) transfer of the information from the signal source to the input port of the amplifier, the input impedance of the amplifier should be zero. This has been shown in Figure 2.6.

3. If the voltage across the load is accurately related to the information provided by the load and the load impedance is inaccurately known, the output port of the amplifier should have zero output impedance to ensure that the information transfer does not depend on the load impedance. This is illustrated in Figure 2.7.

4. If the current through the load is accurately related to the information provided by the load and the load impedance is inaccurately known, the output impedance of the amplifier should be infinite to ensure that the information transfer does not depend on the load impedance. This has been shown in Figure 2.8.

5. If the source or the load has an accurately known linear impedance, there is no preference for current or voltage as the electrical quantity at the input port or at the output port of the amplifier, respectively. In such cases, the current to voltage conversion or the voltage to current conversion by the source or the load impedance does not impose an error on the information to be processed.

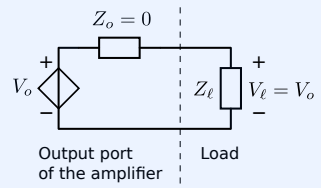


Figure 2.7: Interfacing of the load with the output port of the amplifier in cases in which the voltage across the load accurately represents the information, while the load impedance may be partly unpredictable.

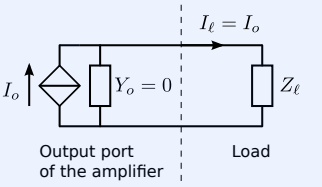


Figure 2.8: Interfacing of the load with the output port of the amplifierin cases in which the current through the load accurately represents the information, while the load impedance may be partly unpredictable.

In all other cases, the transducer mechanism determines whether current or voltage shows the best correspondence to the information acquired by the sensor or delivered by the actuator.

The interfacing with some popular transducers has been described in the examples below.

Example 2.1

A PIN diode can be used as a sensor that converts optical power into electrical current. Study of the transducer mechanism shows that the short-circuit current of a PIN diode is accurately (linearly and instantaneously) related to the intensity of the light on the diode. The open circuit voltage across the PIN diode terminals shows both a logarithmic and a dynamic relation to the light intensity. This light intensity to voltage relation also strongly depends on temperature and is only partly reproducible due to fabrication tolerances. An accurate and reproducible conversion from the optical input quantity to the electrical output quantity is thus obtained when the short-circuit current of the PIN diode is taken as the electrical quantity.

Example 2.2

A microphone converts acoustical power into electrical power. The microphone is designed such that its open-circuit output voltage shows the best correspondence with the sound pressure.

Figure 2.9: A microphone is designed such that its open-circuit output voltage shows the best correspondence with the sound pressure.

Example 2.3

An electric motor can be used as an actuator converting electrical signals into mechanical signals. Investigation of the transducer mechanism shows an accurate relation between the electrical current driving the motor and the torque it delivers. The electrical voltage shows the best relation with the angular speed.

Example 2.4

Piezoelectric transducers can be applied for the conversion of mechanical signals into electrical signals and vice versa. Study of the transducer mechanism shows the mechanical force is converted into electrical charge and vice versa. This charge can be measured by taking the open-circuit voltage at the output of the

sensor, or by taking the time integral of the short-circuit current. Investigation of the transducer mechanism shows the latter method is more accurate, since the influence of the nonlinear Q-V relation of the transducer will then be negligible.

Example 2.5

Transmission lines and filters have to be driven from and/or terminated with accurate and linear impedances. There is no preference for either voltage or current as electrical input and output quantity.

Often, it is not directly clear which electrical quantity best reproduces the non-electrical quantity of the transducer and a study of the operation of the transducer is required to obtain this knowledge. Usually modeling, of sensors and actuators with the focus on information transfer is required to define proper interfacing with the amplifier and to define possible pre- or post-processing functions that compensate for predictable signal processing errors.

Based on the requirements for the port impedances of the amplifier, we will first define nine different amplifier types in section 2.2.1.

2.2.1 Amplifier types

In the previous section, we investigated the selection of the electrical quantity at the input port and at the output port of the amplifier. We found three possibilities for each port. For the input port, we have one of the following cases:

1. Sensing of the source current. The input port must behave as a short-circuit in which the current is sensed. The amplifier's input impedance should be zero: $Z_i = 0$.

2. Sensing of the source voltage. The input port must behave as an open-circuit and the voltage across the input port terminals must be sensed. In this case, the amplifier's input impedance should be infinite: $Z_i = \infty$.

3. Sensing of current or voltage and termination of the source with an accurate linear impedance. The amplifier's input impedance should now be accurately fixed to some characteristic value Z_i. The voltage across this impedance or the current through this impedance can be used as the electrical input quantity of the amplifier.

For the output port, we have one of the following situations:

1. The load must be driven from a voltage source. The output port behaves as an ideal controlled voltage source. The output impedance of the amplifier should be zero: $Z_o = 0$.

2. The load must be driven from a current source. The output port behaves as an ideal controlled current source. In this case, the output impedance of the amplifier should be infinite: $Z_o = \infty$.

3. The load must be driven from an accurate linear impedance Z_o. The output behaves as a controlled voltage source with an accurate and linear impedance in series, or as a controlled current source with an accurate and linear impedance in parallel.

Based on the impedance requirements for both ports, nine amplifier types can now be defined. If they have zero reverse transfer (from the output port to the input port) they are called unilateral. The nine types are listed in Table 2.1.

no	amplifier type	source quantity	load quantity	Z_i	Z_o
1	Voltage amplifier	voltage	voltage	∞	0
2	Transadmittance	voltage	current	∞	∞
3	Voltage to V/I	voltage	voltage or current	∞	Z_o
4	Transimpedance	current	voltage	0	0
5	Current amplifier	current	current	0	∞
6	Current to V/I	current	voltage or current	0	Z_o
7	V/I to voltage	voltage or current	voltage	Z_i	0
8	V/I to current	voltage or current	current	Z_i	∞
9	V/I to V/I	voltage or current	voltage or current	Z_i	Z_o

Table 2.1: Amplifier types and their input and output port impedances.

nt and instantaneous.

According to the above, all amplifiers can be characterized by their transfer characteristics and their port characteristics. For an ideal[4] voltage amplifier, these characteristics are shown in Figure 2.10.

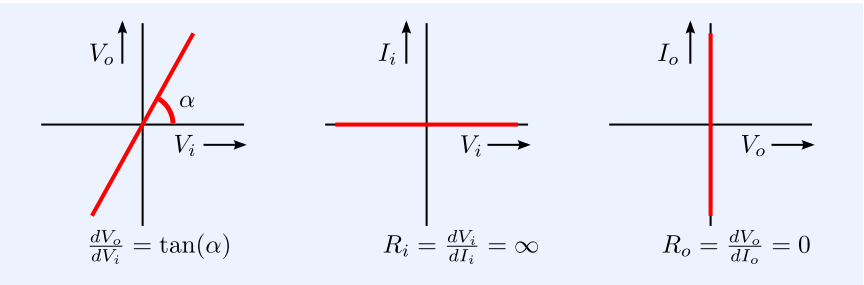


Figure 2.10: Transfer and port characteristics of an ideal voltage amplifier with a voltage gain that equals $tan(\alpha)$.

In section 2.2.4 we will find more arguments for selecting the proper amplifier type for a specific application. First, we first need to introduce the concept *ground* and study possible design consequences that follow from interconnections between the source, the load and the power supply. The ground concept will be introduced in section 2.2.2, and design consequences that follow from interconnections between the amplifier ports will be discussed in section 2.2.3.

2.2.2 Ground

In electrical systems, we often use the concept of the *electrical ground*, also simply know as the *ground node* or the *ground*.

In network theory, the ground is the reference node of a network. The voltage at other nodes is given with respect to the ground potential (see Chapter 18). The absolute potential of the ground node itself cannot be defined and is simply assumed to be zero.

In real-world electrical systems, the electrical connection with the largest dimensions is usually taken to be the ground node. In most cases, this is a power supply terminal.

2.2.3 Port configurations

Until now, we did not consider whether one of the terminals of the source and/or load may be connected to the ground, or if both terminals should be floating with respect to the ground. At a first glance, its seems easy and straightforward to share the ground node; this is the case for example in cars. The ground is the metal chassis, and all devices like batteries, lamps, radio, etc. have one terminal connected to it. This simplifies the wiring: the

metal chassis can be used as a return wire for all devices. However, in many situations, it is not possible nor desirable to use the ground as the common electrical connection for source, amplifier and load. Safety regulations in medical equipment or ground noise due to the physical dimensions of the ground connection and currents flowing through it may force us to use so-called *floating* input or output ports that do not share one of their terminals with the ground.

Regarding this aspect, we can define five different versions for each of the nine amplifier types listed in Table 2.1. Figure 2.11 shows the five different port configurations with their commonly used names.

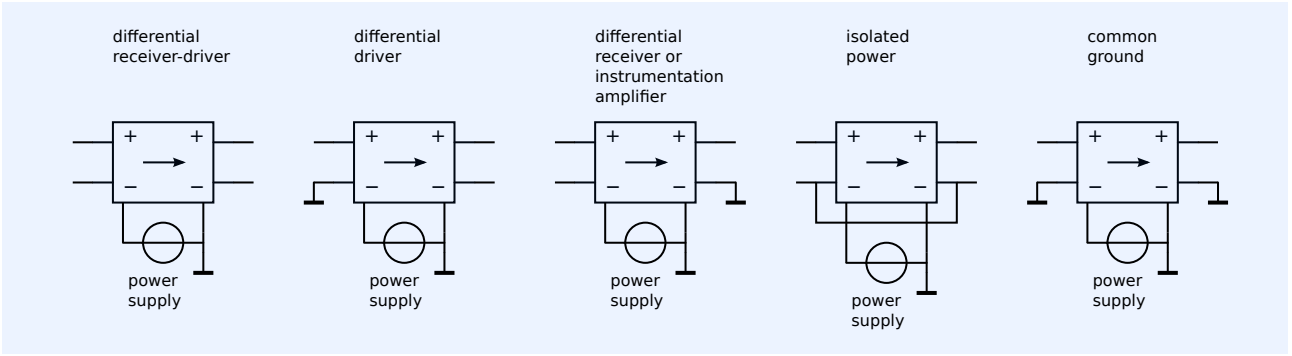


Figure 2.11: Five different versions for each amplifier type.

Common-mode and differential-mode signals

The electrical behavior of ports that are floating with respect to the ground is not uniquely defined by the impedance between the two port terminals.

Let us, for example, consider the application depicted in Figure 2.12A. There, a signal voltage source V_s with a source impedance Z_s has been connected to the floating input port of an amplifier.

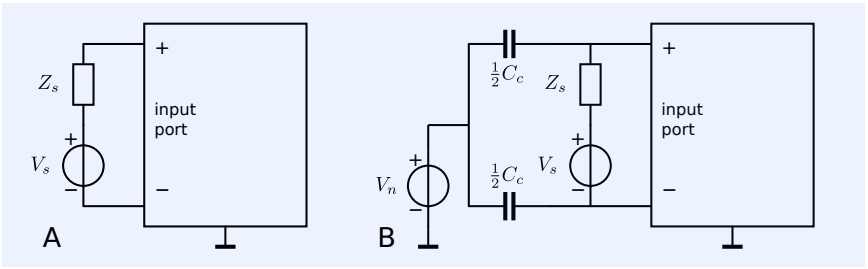


Figure 2.12: A signal source which is floating with respect to ground has been connected to the amplifier's input port. Both terminals of the input port exhibit a equal, weak capacitive coupling to a noise voltage source.

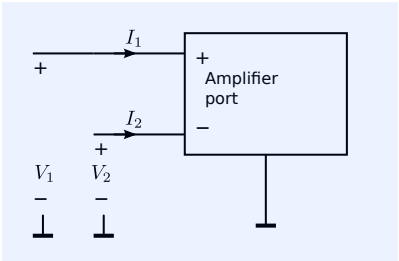


Figure 2.13: Port definition with nodal voltages at the port terminals and currents flowing into the port terminals.

Let us now assume that the signal source exhibits a capacitive coupling to a noise voltage source V_n, which is referred to the ground. The coupling capacitance between the signal source and the noise source is C_c. In Figure 2.12B, this coupling is modeled by means of two capacitors, each with a value of $\frac{C_c}{2}$, between the noise voltage source V_n and both terminals of the input port of the amplifier.

If both terminals of the input port exhibit an infinite impedance to the ground, the voltage of this noise source is found at both port terminals. Such a voltage is then called a *common-mode* voltage: it is common for both inputs. In this case, we speak of a common-mode noise voltage at the input port, introduced by V_n. Although this common-mode noise does not necessarily affect the so-called *differential-mode* signal voltage between the two terminals, it may hamper the signal processing performed by the amplifier. This will, for example, be the case if this common-mode voltage becomes too large.

In such cases, the amplifier will not be able to process the differential-mode input signal because the common-mode noise voltage drives the input port of the amplifier out of its linear operating range. A low port impedance would then be beneficial: it would attenuate the common-mode voltage of the port. At a later stage, we will show that common-mode and differential-mode port impedances can be designed independently.

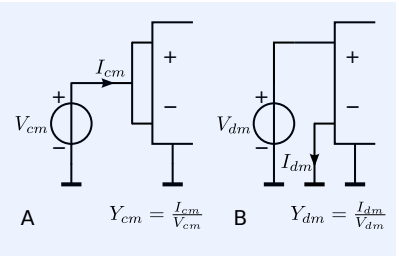


Figure 2.14: Measurement setup for determination of:
A: The common-mode port admittance
B: The differential-mode port admittance
The port can be an input port or an output port of a grounded circuit.

Definitions of common-mode and differential-mode quantities

The definitions of the common-mode quantities and the differential-mode quantities below refer to the circuit in Figure 2.13.

The common-mode voltage V_{cm} of a port is defined as the sum of the two terminal voltages of that port divided by two:

$$V_{cm} = \frac{V_1 + V_2}{2}. \tag{2.1}$$

The common-mode current I_{cm} that flows into a port is defined as the sum of the currents that flow into the two-port terminals:

$$I_{cm} = I_1 + I_2. \tag{2.2}$$

The differential-mode voltage V_{dm} of a port is defined as the difference between the voltages at the two terminals:

$$V_{dm} = V_1 - V_2. \tag{2.3}$$

The differential-mode current I_{dm} is defined as the difference between the currents that flow into the two-port terminals, divided by two:

$$I_{dm} = \frac{I_1 - I_2}{2}. \tag{2.4}$$

Figure 2.14A and B show the test setup for determination of the common-mode and the differential-mode input admittances, Y_{cm} and Y_{dm}, respectively.

The common-mode impedance Z_{cm} is the reciprocal value of Y_{cm}, it is defined as:

$$Z_{cm} = \frac{dV_{cm}}{dI_{cm}}. \tag{2.5}$$

The differential-mode impedance Z_{dm} is the reciprocal value of Y_{dm}, it is defined as:

$$Z_{dm} = \frac{dV_{dm}}{dI_{dm}}. \tag{2.6}$$

The port impedances from Table 2.1 refer to Z_{dm}.

In section 2.4.1, we will introduce techniques for modeling the common-mode and the differential-mode behavior of the amplifier ports.

In Chapter 7, we will present methods for designing amplifiers with specific port isolation properties and we will show the way in which the common-mode impedance and the differential-mode impedance of a port can be designed independently.

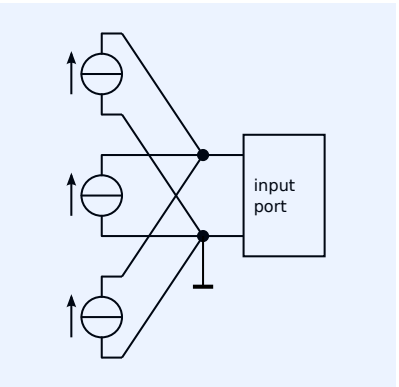


Figure 2.15: Addition of currents requires parallel connections. All sources may be connected to ground.

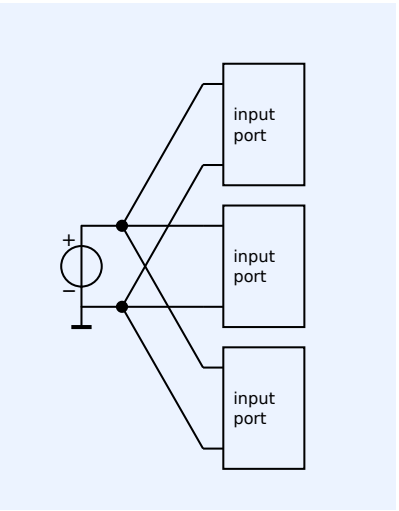


Figure 2.16: Distribution of voltages requires parallel connections. All receiving inputs may be connected to ground.

2.2.4 Summing and distributing signals

The need for addition (combination) and/or distribution of signals may also provide arguments for the selection of the information carrying electrical quantity. Figure 2.15 shows the addition of currents in a common-ground system.[5] Addition of voltages requires voltage sources to be connected in series, which is not possible in a common-ground system. Similarly, distribution of voltages in a common-ground input-output system is much easier than the distribution of currents (see Figure 2.16).

[5] A common-ground system is a system in which all signal sources and loads share the ground terminal.

2.3 Modeling of the ideal behavior

Although the amplification mechanism requires a power source, it is usually omitted in a functional description of the amplifier. The functional model of the amplifier only describes the signal transfer from the source to the load. The load power is simply assumed to be delivered by a voltage or a current source, which is controlled by the signal source. The simplest network model of an ideal(ized) amplifier thus has two ports:

1. An input port, which is modeled as an open-circuit, a short circuit or as an accurate and linear impedance.
2. An output port, which is modeled as a controlled current source or a controlled voltage source, which is controlled by the input quantity. A linear impedance in series with a controlled voltage source, or in parallel with a controlled current source, can be added if a finite nonzero output impedance is required.

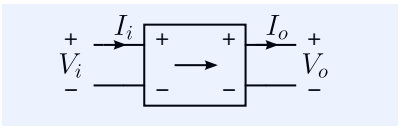


Figure 2.17: Two-port definition.

The symbol and the sign conventions of such a two-port model are shown in Figure 2.17. The sign of the current in the output port is opposite to what is customary in network theory. This to prevent minus signs in the transfer of cascaded two-ports.

A two-port representation of a network only yields a complete description if the two-port constraints apply:

1. A current that flows into one terminal of a port, flows out of the opposite terminal of that port.
2. A voltage difference between the input port and the output port does not affect any of the two-port currents.

The two-port constraints are always valid if the ports of a two-port are terminated with one-ports. Two-ports for which the port constraints are always satisfied, regardless of the electrical network to which it is connected, are called *natural two-ports*. Examples of natural two-ports are *ideal transformers*, *ideal gyrators* and *ideal controlled sources*. If the two-port constraints are not valid, the four-terminal network requires a 3×3 matrix representation. For more information on two-ports and the two-port constraints, see Chapter 18.6.

2.3.1 Transmission matrix-1 representation

The transmission-1 matrix representation will be used as the two-port model during amplifier design. This anti-causal two-port model relates the input port quantities to the output port quantities.[6] It will be shown that all amplifier types can be described by this model, using finite values[7] for the model parameters only.

The transmission-1 matrix equation for the two-port depicted in Figure 2.17 is:

$$\begin{pmatrix} V_i \\ I_i \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_o \\ I_o \end{pmatrix}. \tag{2.7}$$

The parameters are defined and measured as shown in Figure 2.18. The reciprocal values of A, B, C and D are called the *voltage gain factor* μ, the *transadmittance factor* γ, the *transimpedance factor* ζ and the *current gain factor* α, respectively. The word *factor* indicates that it refers to a property of the two-port itself: a transfer from a port input quantity to a port output quantity, rather than the transfer from the source to the load.

The transmission-1 two-port parameters A, B, C and D can be obtained through measurement of the gain factors μ, γ, ζ and α. The parameters A

[6] Anti-causal: The input quantities of the two-port are described as a function of the output quantities, while the physical operation is opposite: the input signals cause the output signals.
[7] Including zero.

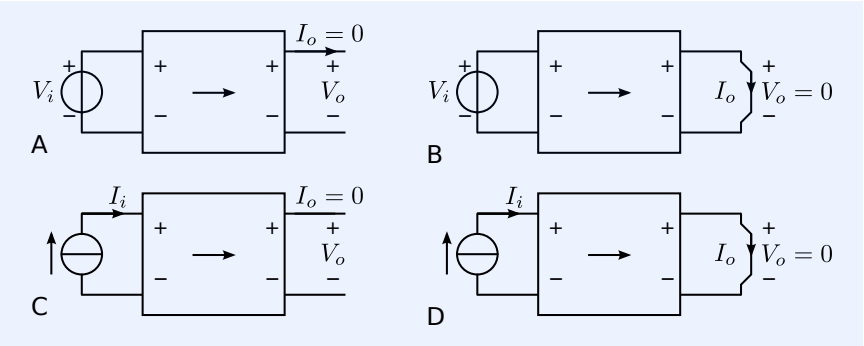


Figure 2.18: Test benches for determination of the Transmission-1 two-port parameters:
A: $A = \frac{1}{\mu}$
B: $B = \frac{1}{\gamma}$
C: $C = \frac{1}{\zeta}$
D: $D = \frac{1}{\alpha}$

and C are determined when the output port is left open ($I_o = 0$), while B and D are determined with the output port shorted ($V_o = 0$):

$$A = \frac{1}{\mu} = \left.\frac{V_i}{V_o}\right|_{I_o=0}, \tag{2.8}$$

$$B = \frac{1}{\gamma} = \left.\frac{V_i}{I_o}\right|_{V_o=0}, \tag{2.9}$$

$$C = \frac{1}{\zeta} = \left.\frac{I_i}{V_o}\right|_{I_o=0}, \tag{2.10}$$

$$D = \frac{1}{\alpha} = \left.\frac{I_i}{I_o}\right|_{V_o=0}. \tag{2.11}$$

2.3.2 Source-to-load transfer

We can express the source-to-load transfer in terms of the transmission parameters of the two-port and the source and the load impedance. From these expressions, we can draw some important design conclusions.

The test setup for determination of the voltage gain A_v and the transadmitance A_y is shown in Figure 2.19A.

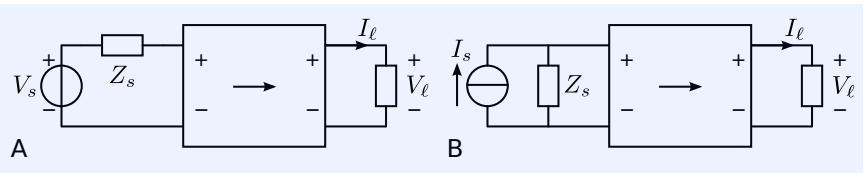


Figure 2.19: Measurement setup for the source-load relations:
(a) The voltage to voltage transfer and the voltage to current transfer, and
(b) The current to current and current to voltage transfer.

The voltage gain A_v can be written as:[8]

$$A_v = \frac{V_\ell}{V_s} = \frac{1}{A + B\frac{1}{Z_\ell} + CZ_s + D\frac{Z_s}{Z_\ell}}. \tag{2.12}$$

The transadmittance A_y is found as:

$$A_y = \frac{I_\ell}{V_s} = \frac{1}{AZ_\ell + B + CZ_\ell Z_s + DZ_s}. \tag{2.13}$$

The test setup for the transimpedance A_z and for the current gain A_i is depicted in Figure 2.19B. The transimpedance A_z is found to be:

$$A_z = \frac{V_\ell}{I_s} = \frac{1}{A\frac{1}{Z_s} + B\frac{1}{Z_s Z_\ell} + C + D\frac{1}{Z_\ell}}. \tag{2.14}$$

[8] Please notice it is our intention to show these transfer functions and draw design conclusions rather than to derive these expressions.

The current gain A_i can be obatined as:

$$A_i = \frac{I_\ell}{I_s} = \frac{1}{A\frac{Z_\ell}{Z_s} + B\frac{1}{Z_s} + CZ_\ell + D}. \tag{2.15}$$

At this stage, we can already draw some design conclusions from expressions (2.12) through (2.15). Consider, for example, a situation in which we need to design a voltage amplifier of which the source-to-load transfer should not depend on the source impedance Z_s and on the load impedance Z_ℓ. Expression (2.12) shows us that we need to assign a nonzero value to the transmission parameter A, while all other parameters should be zero. Only for this situation, the source-to-load transfer will not depend on the source and load impedances. This conclusion can already be seen from the T_1 matrix equation itself. With B, C and D zero, we have $V_i = AV_o$, and $I_i = 0$ for all values of I_o. Hence, V_i will be equal to V_s and V_ℓ will be equal to V_o. In the next section, we will study the expressions for the port impedances in a similar way, and derive design conclusions for all nine amplifier types from Table 2.1.

2.3.3 Input and output impedance

The input impedance and the output impedance of the amplifier can also be expressed in terms of the transmission parameters and of the load and the source impedance.

Input impedance

The test setup for the input impedance is shown in Figure 2.20A.

The input impedance Z_i of a two-port terminated with a load impedance Z_ℓ can be written as:

$$Z_i = \frac{V_i}{I_i} = \frac{AZ_\ell + B}{CZ_\ell + D}. \tag{2.16}$$

Please notice that the input impedance Z_i does not depend on the load impedance Z_ℓ if $\frac{A}{B} = \frac{C}{D}$, which is equivalent to: $AD = BC$.

Output impedance

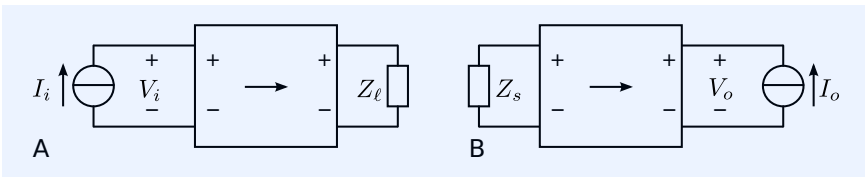


Figure 2.20: Measurement setup for:
(A) The input impedance
(B) The output impedance.

The test setup for the output impedance is depicted in Figure 2.20B. The output impedance Z_o of the two-port driven from Z_s can be written as:

$$Z_o = \frac{V_o}{I_o} = \frac{DZ_s + B}{CZ_s + A}. \tag{2.17}$$

Please notice that the output impedance Z_o does not depend on the source impedance Z_s if $\frac{D}{B} = \frac{C}{A}$, which is equivalent to: $AD = BC$.

Unilateral behavior

Amplifiers of which the input impedance does not depend on the load impedance are called *unilateral*.[9] The reverse transfer of unilateral amplifiers is zero.[10]

[9] In a unilateral amplifier, the input impedance does not depend on the load impedance and the output impedance does not depend on the source impedance.

[10] The reverse transfer of an amplifier is the transfer from its output to its input.

The condition for unilateral behavior is:

$$AD = BC. \tag{2.18}$$

Design conclusions

In the following example, we will demonstrate how to derive design conclusions from the above expressions.

Example 2.6

Let us assume we want to design an amplifier that needs to drive a cable from its characteristic impedance. This cable driver itself should be driven from a voltage source of which the internal impedance Z_s is not accurately known. The input impedance of the cable driver thus needs to be infinite and the output impedance must equal the characteristic impedance of the cable. This type of amplifier is listed under line 3 in Table 2.1. An infinite input impedance is obtained if $C = 0$ and $D = 0$. A finite output impedance that does not depend on the source impedance requires either $A = 0$ and $B = 0$, or $C = 0$ and $D = 0$. If we combine both requirements: an infinite input impedance and a finite accurate output impedance, we need $C = 0$ and $D = 0$, as well as finite, nonzero values for A and B. The output impedance of this amplifier then equals $\frac{B}{A}$.

The parameter values for the nine amplifier types from Table 2.1 are listed in Table 2.2.

no	amplifier type	$\mathbf{Z}_i$	$\mathbf{Z}_o$	**A**	**B**	**C**	**D**
1	Voltage amplifier	∞	0	A	0	0	0
2	Transadmittance amplifier	∞	∞	0	B	0	0
3	Voltage input, finite nonzero output impedance	∞	Z_o	A	B	0	0
4	Transimpedance amplifier	0	0	0	0	C	0
5	Current amplifier	0	∞	0	0	0	D
6	Current input, finite nonzero output impedance	0	Z_o	0	0	C	D
7	Finite nonzero input impedance, voltage output	Z_i	0	A	0	C	0
8	Finite nonzero input impedance, current output	Z_i	∞	0	B	0	D
9	Finite nonzero input and output impedance	Z_i	Z_o	A	B	C	D

Table 2.2: Transmission parameters of the nine amplifier types from Table 2.1.

As we have seen, the transfer and the port impedances of the nine types of amplifiers that are listed in Table 2.2 can be characterized with the amplifier's transmission-1 matrix parameters A, B, C and D. The desired values of these parameters are either zero or finite.

The first eight amplifier types from Table 2.2 are unilateral (see section 2.3.3). The last amplifier is unilateral if $AD = BC$.

In total, there exist sixteen possible combinations of zero and finite nonzero transmission parameters. The other seven other amplifier types are:

- The nullor. This is a network element that has all its transmission-1 parameters equal to zero ($A = B = C = D = 0$). See Chapter 18.3 for the formal definition of the nullor. Although the input impedance and the output impedance of the nullor are undefined, the nullor is a unilateral two-port: $AD = BC$. The nullor will be applied as the *ideal controller* in negative-feedback amplifiers.[11]
- The transformer-like configuration (A and D nonzero and B and C zero). This amplifier is not unilateral.

  The name *transformer-like configuration* stems from the similarity of the transmission-1 matrix of this configuration to that of the transformer. An ideal transformer has transmission-1 parameters: $B = C = 0$ and $AD = 1$. See Chapter 18 for the formal definition of this network element.
- The gyrator-like configuration (A and D zero and B and C nonzero). This amplifier is not unilateral.

[11] This will be discussed in Chapter 7.

The name *gyrator-like configuration* stems from the similarity of the transmission-1 matrix of this configuration with that of the gyrator. A gyrator has transmission-1 parameters: $A = D = 0$ and $BC = 1$. See Chapter 18 for the formal definition of this network element.

- Four amplifier configurations that have three out of four nonzero parameters. These amplifiers are not unilateral because $AD \neq BC$.

2.3.4 Available power gain

We have seen that amplifiers can be distinguished from passive two-ports by their available power gain. The available power gain of amplifiers exceeds unity, while that of passive two-ports will always be smaller than unity. In this section, we will give the definitions of the available power and the available power gain. We will also express the available power gain of a two-port in its transmission-1 parameters. We will do this for the simple case of a resistive two-port driven from a resistive source and terminated with a resistor.[12]

[12] A, B, C, D, Z_s and Z_ℓ are real.

Available power of a source

The available power of a source is defined as the maximum power it can deliver to a load.

The available power P_s of a signal source that delivers an RMS voltage V_s and that has an internal impedance Z_s equals

$$P_s = \frac{V_s^2}{4\,\mathrm{Re}\,(Z_s)}. \tag{2.19}$$

It is the power which is delivered to a load with an impedance Z_ℓ that is the complex conjugated of Z_s: $Z_\ell = Z_s^*$.

Please notice that the available power of a source has nothing to do with the actual power delivered by that source: the available power is a property of the source itself. If we leave a source open ($Z_\ell = \infty$), or if we short the source ($Z_\ell = 0$), the power delivered to the load equals zero because the load current equals zero if $Z_\ell = \infty$, and the load voltage equals zero if $Z_\ell = 0$. However, the available power of that source, for all source terminations, is given by (2.19).

Available power at the output of a two-port

Similarly, the available power at the output of a two-port is the maximum power that can be delivered by that port. In this case, the output port is considered a signal source. The available power P_p of a port with an port impedance Z_o that delivers an open circuit port voltage V_o thus equals:

$$P_p = \frac{V_o^2}{4\,\mathrm{Re}\,(Z_o)}. \tag{2.20}$$

Again, this power is delivered to a load impedance which is the complex conjugated of the port impedance: $Z_\ell = Z_o^*$.

Available power gain of a two-port

The available power gain G_p of a two-port is defined as the ratio of the available power at the output port and the available power of the source connected to its input.

With the aid of (2.19) and (2.20), it can be shown that:

$$G_p = A_v^2 \frac{\mathrm{Re}\,(Z_s)}{\mathrm{Re}\,(Z_o)}, \tag{2.21}$$

where A_v is the voltage gain factor of the two-port, which is the reciprocal value of the T1 matrix parameter A.

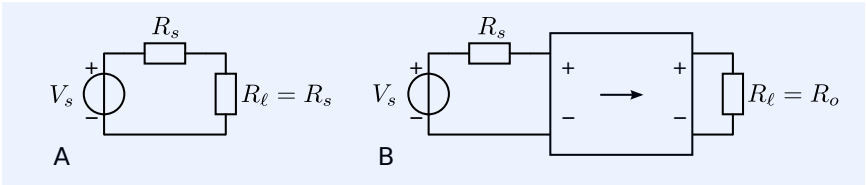


Figure 2.21: Available power at optimum port termination:
(A) of a source,
(B) of a two-port.

We will now express the available power gain of a two-port in its transmission-1 matrix parameters and in the impedance of the source connected to its input port. For the sake of simplicity, we will do this for the simple case of a resistive two-port (A, B, C, D are real) and a resistive source and load, with resistances R_s and R_ℓ, respectively.

We will evaluate the available power of a signal source as shown in Figure 2.21A and the available power at the output of the two-port as shown in Figure 2.21B.

The available power P_s of the signal source is found to be:

$$P_s = \frac{V_s^2}{4R_s}. \tag{2.22}$$

The test setup for the evaluation of the available power of the two-port is shown in Figure 2.21B. The transmission-1 parameters of the amplifier are A, B, C and D.

The amplifier will deliver its maximum power if $R_\ell = R_o$. According to (2.17), this is the case if:

$$R_\ell = \frac{DR_s + B}{CR_s + A}. \tag{2.23}$$

This maximum power is the available power of the two-port P_p. It can be obtained as:

$$P_p = \frac{V_o^2}{4R_o}, \tag{2.24}$$

where V_o is the open circuit voltage at the amplifier output with the input port being driven from the signal source.

With the aid of (2.12), we obtain

$$V_o = \frac{V_s}{A + CR_s}. \tag{2.25}$$

With the aid of expressions (2.23) through (2.25), we may express the available power of the amplifier in terms of the signal source voltage and the transmission-1 parameters of the amplifier:

$$P_p = \frac{V_s^2}{4\,(DR_s + B)\,(CR_s + A)}. \tag{2.26}$$

The available power gain G_p of an instantaneous two-port with (real) transmission parameters A, B, C, D, driven from a resistive source with resistance R_s, is then obtained as

$$G_p = \frac{1}{AD + AB/R_s + BC + CDR_s}. \tag{2.27}$$

Expression (2.27) shows that a large available power gain is obtained when

all transmission parameters approximate zero. A two-port that has all its transmission parameters equal to zero is called a nullor. In Chapter 7 we will use the nullor as an ideal controller when designing negative feedback amplifier configurations. The operational amplifier (OpAmp) can be viewed as a particular implementation (and approximation) of the nullor.

Expression (2.27) also shows that for a given two-port with transmission parameters A, B, C, D, the available power gain has maximum value $G_{p,\max}$ at some optimum source resistance $R_{s,opt}$. This optimum source resistance can be found as:

$$R_{s,opt} = \sqrt{\frac{AB}{CD}}. \tag{2.28}$$

After substitution of (2.28) in (2.27) we obtain an expression for the maximum available power gain of a two-port:

$$G_{p,\max} = \frac{1}{\left(\sqrt{AD} + \sqrt{BC}\right)^2}. \tag{2.29}$$

If the amplifier is unilateral, we have $AD = BC$, which simplifies (2.29) to:

$$G_{p,\max} = \frac{1}{4AD} = \frac{1}{4}\mu\alpha. \tag{2.30}$$

Hence, the maximum available power gain of a unilateral resistive two-port is a quarter of the product of the voltage gain factor μ and the current gain factor α of that two-port.

Expressions (2.29) and (2.30) can be used to evaluate the amplifying capabilities of a two-port.

Available power gain, power gain and transducer power gain

The *power gain* or *operating power gain* of an amplifier is defined as the ratio of the power delivered to the load and the power delivered by the source. It depends on the properties of the amplifier, and of the source impedance Z_s and the load impedance Z_ℓ. It will be clear that if $\mathrm{Re}(Z_\ell) = \infty$ or $\mathrm{Re}(Z_\ell) = 0$, the power gain equals zero, because in that case, no power is delivered to the load. The *available power gain* of the amplifier, however, can still be much larger than unity.

The *transducer power gain* of an amplifier is defined as the ratio of the load power and the available power of the source.

2.3.5 Idealized amplifier models

In the previous sections, we have discussed the modeling of the ideal behavior of amplifiers with the aid of a two-port and the T1 matrix representation. In order to evaluate the idealized behavior of an amplifier with simulation programs, it is convenient to have two-port models at our disposal that can be constructed from basic network elements, available to the simulator.[13]

[13] Basic two-port elements are the controlled sources :
- VCVS: voltage-controlled voltage source
- VCCS: voltage-controlled current source
- CCVS: current-controlled voltage source
- CCCS: current-controlled current source

Basic models

Figure 2.24 shows four high-level network models of amplifiers that have at least one nonzero coefficient in their T1 matrix. At a later stage, we will introduce one high-level model for all amplifier types.

In the models from Figure 2.24, the power delivered to the input port equals the total power dissipated in the controlled sources at the input port. The power delivered to the load is the total power delivered by the controlled sources at the output port. Similar things can be said about the energy stored in the two-port. If the sum of the powers dissipated all the controlled sources

equals zero and the total energy storage equals zero, we speak of non-energic two-ports. Those two-ports are not amplifiers, because the output port delivers as much power to the load, as the input port takes from the source. Transformers ($AD = 1$, $B = 0$, $C = 0$) and gyrators ($BC = 1$, $A = 0$, $D = 0$) are non-energic two-ports.

In the following example we will present two models for transformers that can be derived from the models presented in Figure 2.24.

Example 2.7

Let us assume we need a network model of an ideal transformer such as the one in Figure 2.22A. The transmission-1 parameters of an ideal transformer are: $AD = 1$, $B = 0$ and $C = 0$. Hence, we can use the two-port models for $A \neq 0$ and $D \neq 0$.

The model for $A \neq 0$ with $A = \frac{1}{n}$ is shown in Figure 2.22B.

The model for $D \neq 0$ with $D = n$ is shown in Figure 2.22C.

It can be seen that both models are equal: one can be derived from another after interchanging the input port and the output port, while replacing the turns ratio n with its reciprocal value.

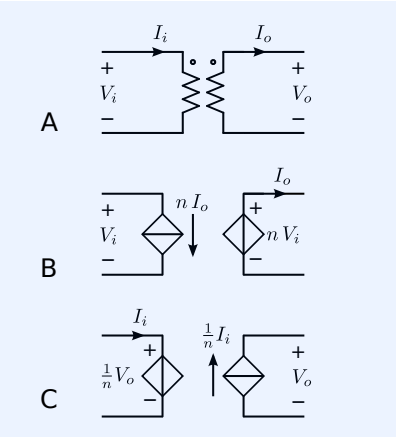


Figure 2.22: Network models of ideal transformers.
A. Symbol of ideal transformer
B. Network model for voltage-controlled output port
C. Network model for current controlled output port

Transformers and gyrators have an available power gain of unity; this directly follows from (2.27).

Two-ports of which the sum of the powers dissipated in the controlled sources is negative, deliver more power to the load than they take from the source. They are called active two-ports.

Practical implementations can be constructed from passive devices and power sources.[14]

[14] In the strict sense, transistors are passive devices.

If a two-port delivers less power to the load than it takes from the source, it can still be an amplifier. As discussed before, all amplifiers have an available power gain larger than unity, but the actual power gain in an application may be much less.

The high-level models presented here do not tell us anything about the power losses in the amplifier. They only model the power delivered to the load and the power taken from the source. Hence, if there are losses in an amplifier we have to add them to the model.

So, if an amplifier has an output impedance of 50Ω and it delivers a voltage of 1V into a 50Ω load, it does not mean that there is a 50Ω resistor in series or in parallel with the load that reduces the power efficiency of the amplifier to 50%. At a later stage, we will discuss the design of accurate port impedances without degrading the power efficiency of the amplifier.

Similar things can be said about the relation between the input impedance and the noise performance. If an amplifier has an input impedance of 50Ω, and the input port is connected to a source with an internal impedance of 50Ω it does not mean that this amplifier contributes as much noise as the source. At a later stage, we will discuss the design of accurate port impedances without degrading the noise performance of the amplifier.

General network model

Figure 2.23 shows a general model that can be applied in all situations. The model comprises a nullor: a network element that consists of a nullator and a norator. The nullator sets the network condition: zero voltage across its terminals with no current flow through it. The norator provides an extra degree of freedom to satisfy this condition by introducing an unknown port current I_o. In this way, we simply implemented the equations:

$$\begin{aligned} V_3 - V_4 &= A(V_1 - V_2) + BI_o, \\ I_i &= C(V_1 - V_2) + DI_o, \end{aligned} \tag{2.31}$$

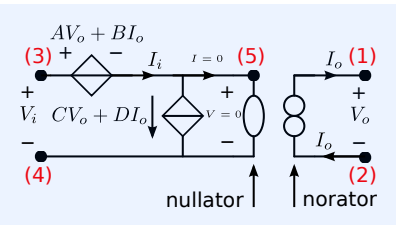


Figure 2.23: Universal 4-terminal network model of a two-port with T1 parameters.

where $V_3 - V_4 = V_i$ and $V_1 - V_2 = V_o$.

Readers who are not familiar with the concept of the nullor are invited to study Chapter 18.

The MNA matrix stamp of this two-port is:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ C & -C & 0 & 0 & 0 & D & 1 \\ -C & C & 0 & 0 & 0 & -D & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ -A & A & 1 & 0 & -1 & -B & 0 \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ I_o \\ I_i \end{pmatrix}, \tag{2.32}$$

where V_n is the voltage at node (n) with respect to that of the reference node (0). For more information about modified nodal analysis and matrix stamps, please refer to Chapter 18.

The simpler models from Figure 2.24 require at least one nonzero coefficient in their T1 matrix.

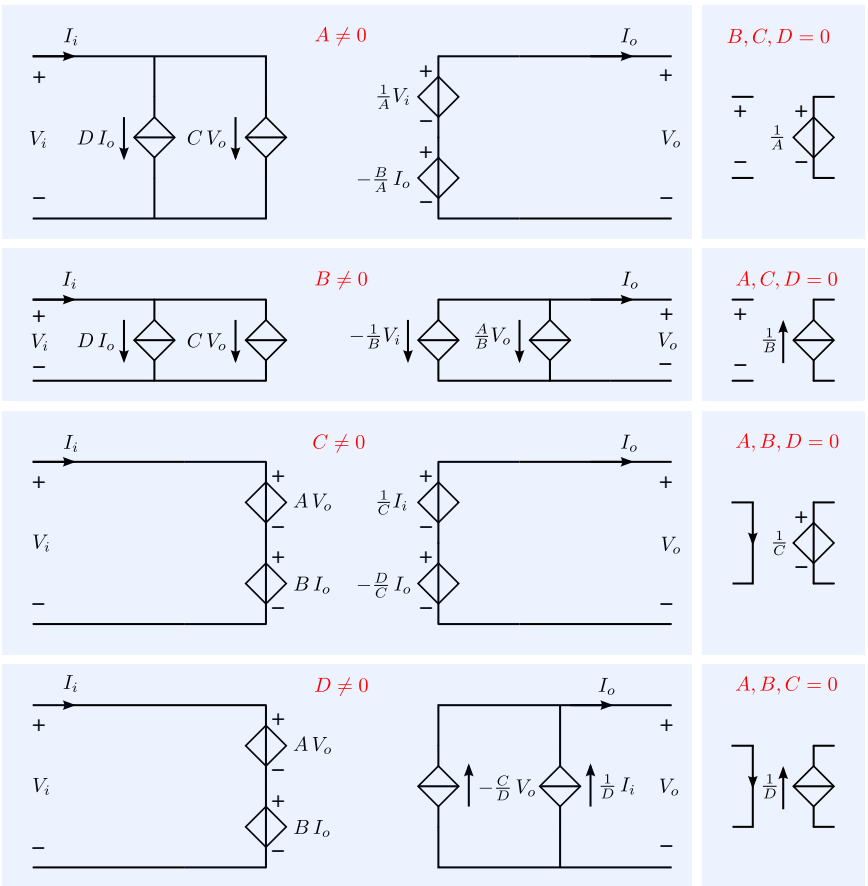


Figure 2.24: Network models of amplifiers that have at least one nonzero transmission-1 matrix parameter.

From top to bottom:

1. Left: High-level model for amplifiers having $A \neq 0$.
1. Right: Voltage-controlled voltage source (VCVS): $B = C = D = 0$.
2. Left: High-level model for amplifiers having $B \neq 0$.
2. Right: Voltage-controlled current source (VCCS): $A = C = D = 0$.
3. Left: High-level model for amplifiers having $C \neq 0$.
3. Right: Current-controlled voltage source (CCVS): $A = B = D = 0$.
4. Left: High-level model for amplifiers having $D \neq 0$.
4. Right: Current-controlled current source (CCCS): $A = B = C = 0$.

2.4 Modeling of the non-ideal behavior

In this section, we will discuss the modeling of errors that adversely affect the information processing by the amplifier.

The amount of information that can be processed by any information processing system is limited by the three fundamental physical limitations. Any

physical system will:

1. Add noise to the signal
2. Suffer from speed limitations
3. Suffer from power limitations.

The influence of these limitations on the quality of the information transfer strongly depends on:

1. The way in which the information is embedded in the signal
2. The applied technology
3. The way in which the observer perceives these errors

The influence of a particular physical effect that causes information processing errors can often be characterized in various ways. There are many ways, for example, to specify the amount of nonlinearity that can be accepted. In audio applications, we often use the amount of harmonic distortion, while for radio applications, the amount of gain compression and third-order intermodulation distortion are more useful measures.

Since the design method for amplifiers as presented in this book can be used for a wide range of applications, we will use the most common description methods, but we will not discuss the perception of errors for all kinds of applications. The designer of amplifiers should use error descriptions that closely match the error perception in the particular field of application.

Before we begin with the description of the errors that result from the fundamental physical limitations, we will first pay attention to the non-ideal isolation between the three amplifier ports. The non-ideal port isolation[15] adds unwanted functionality to the amplifier. This undesired functionality results in a sensitivity of the amplifier to power supply noise and common-mode noise.

We will then discuss the modeling of non-ideal behavior due to the influence of the fundamental physical and the technological limitations. We will discuss the modeling of the noise behavior, the dynamic behavior and the nonlinear behavior, and introduce specific amplifier classes related to their power efficiency.

[15] The input port, the output port and the power supply port, see Figure 2.11.

2.4.1 Modeling of the source and load isolation

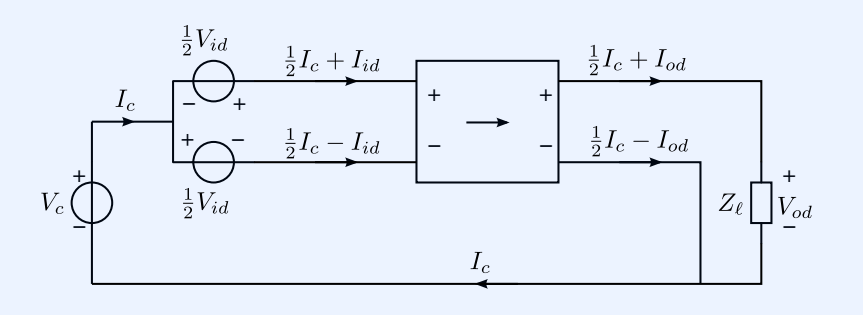


Figure 2.25: Measurement setup for determination of the parameters of the two-port with non-ideal port isolation

Modeling of the non-ideal source-to-load isolation requires a four terminal network description of the amplifier instead of a two-port model. This simply follows from network theory: the behavior of an electrical network with four nodes can be fully described with three nodal equations (a 3×3 matrix; see Chapter 18.2). A convenient representation that closely matches the previous description of the amplifier's behavior would use the transmission

parameters A, B, C and D, plus five parameters that describe the non-ideal port isolation. Figure 2.25 shows the model of an amplifier with a floating input port and its output port referenced to the ground. The sum of the currents that flow into the input port is defined as the common-mode current I_c. The information is embedded in the differential-mode quantities V_{id}, I_{id}, V_{od} and I_{od}.

Descriptions for amplifiers with 3×3 matrices are rarely found. In practice, far simpler and, unfortunately, often incomplete description methods for the imperfect port isolation are used. These simplified methods are only valid for particular port terminations that are assumed during simplification.

It is common practice to describe the influence of the imperfect port isolation with the *common-mode rejection ratio CMRR*, and the *common-mode input impedance* Z_{cm} only. Moreover, the $CMRR$ is usually only defined for the voltage transfer. It is defined as the ratio of a common-mode voltage V_c at the input of the amplifier and the differential-mode voltage V_{id} which is required to compensate for the change in the output voltage V_o due to this common-mode voltage:

$$CMRR = \left.\frac{V_c}{V_{id}}\right|_{V_{od}=0}. \tag{2.33}$$

Alternatively, it can be defined as the ratio of the differential-mode voltage gain $A_{dm} = \frac{V_{od}}{V_{id}}$ and the common-mode to differential-mode voltage conversion gain $A_{cd} = \frac{V_{od}}{V_c}$:

$$CMRR = \frac{A_{dm}}{A_{cd}} = \frac{\frac{V_{od}}{V_{id}}}{\frac{V_{od}}{V_c}} = \frac{V_c}{V_{id}}. \tag{2.34}$$

The common-mode input impedance is defined as the ratio of the common-mode input voltage and the common-mode input current:

$$Z_{cm} = \frac{V_c}{I_c}. \tag{2.35}$$

It will be clear that a description of the non-ideal port isolation with only two extra parameters can never be complete. Different drive and terminating conditions for each input port terminal, generally result in different values of the $CMRR$. This compels us to specify the conditions under which the $CMRR$ has been measured.

2.4.2 Modeling of the power supply isolation

An even more complex situation occurs if both the input port and the output port are floating with respect to the power supply port and if there exists a nonzero transfer from the power supply port to the input port or to the output port, or vice versa. Figure 2.26 shows a network model for this situation. The voltages at the port terminals with respect to the ground are denoted as $V_1 \cdots V_4$. The currents that flow into the input port terminals are I_1 and I_2, and the currents that flow out of the output port terminals are I_3 and I_4.

The network model of the amplifier in this configuration requires five nodal equations. Usually one of the power supply terminals is selected as the reference node. The definitions of the common-mode and the differential-mode quantities are listed in Table 2.3.

The coefficients of the 5×5 matrix can be divided into nine groups:

1. Four coefficients that describe the differential-mode transfer from input to output

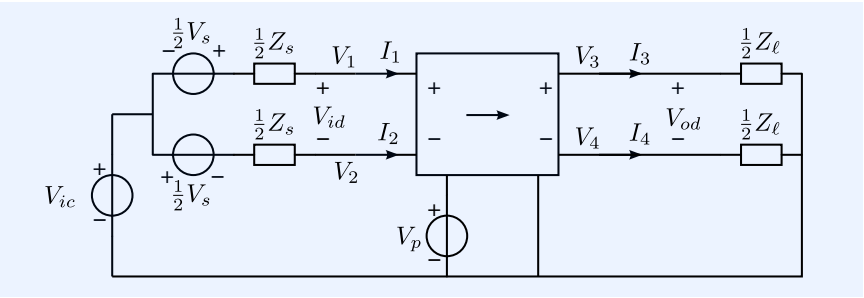


Figure 2.26: Definition of input and output quantities of a voltage amplifier with floating ports.

Symbol	Value	Description
V_{id}	$V_1 - V_2$	differential-mode input voltage
I_{id}	$\frac{1}{2}(I_1 - I_2)$	differential-mode input current
V_{ic}	$\frac{1}{2}(V_1 + V_2)$	common-mode input voltage
I_{ic}	$I_1 + I_2$	common-mode input current
V_{od}	$V_3 - V_4$	differential-mode output voltage
I_{od}	$\frac{1}{2}(I_3 - I_4)$	differential-mode output current
V_{oc}	$\frac{1}{2}(V_3 + V_4)$	common-mode output voltage
I_{oc}	$I_3 + I_4$	common-mode output current
V_p	V_p	power supply voltage
I_p	$I(V_p)$	current flow through V_p

Table 2.3: Definition of the common-mode and differential-mode quantities for the differential amplifier with power supply from Figure 2.26

2. Four coefficients that describe the common-mode input to differential-mode output transfer

3. Four coefficients that describe the differential-mode input to common-mode output transfer

4. Four coefficients that describe the common-mode input to common-mode output transfer

5. Two coefficients that describe transfer from the power supply voltage to the common-mode input quantities

6. Two coefficients that describe transfer from the power supply voltage to the common-mode output quantities

7. Two coefficients that describe transfer from the power supply voltage to the differential-mode input quantities

8. Two coefficients that describe transfer from the power supply voltage to the differential-mode output quantities

9. One coefficient that describes the power supply voltage to power supply current transfer.

In situations in which either the input or the output port has one terminal in common with the power supply port, a 4×4 matrix representation is sufficient.

However, the port isolation is usually only described by a few parameters. As stated earlier, such a description can only be accurate if the connection of the amplifier to its electrical environment is completely specified.

The most commonly used characteristics are:

1. The *CMRR* (see definition in expression 2.33).

2. The Power Supply Rejection Ratio: *PSRR*.

   Like the CMRR, the PSRR is usually defined for voltages only. It is defined as the ratio between the change in power supply voltage and the input voltage that must be applied to compensate for its effect at the output:

$$PSRR = \left.\frac{V_p}{V_{id}}\right|_{V_{od}=0}. \tag{2.36}$$

3. The common-mode input impedance Z_{cm}.

4. The rejection factor F for voltages (for voltage amplifiers with floating input and output). It is defined as the ratio of the differential-mode voltage gain and the common-mode voltage gain:

$$F = \frac{V_{od}}{V_{id}}\frac{V_{ic}}{V_{oc}}. \tag{2.37}$$

   A large rejection factor, implies a relatively small common-mode transfer, which, in a case of cascaded differential amplifiers, reduces the influence of the next stage's common-mode to differential-mode conversion.

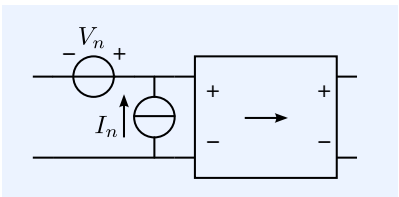


Figure 2.27: Two-port noise model consisting of two noise sources connected to a noise-free two-port.

2.4.3 Modeling of the noise behavior

The general aspects of noise modeling are discussed in Chapter 19. Readers who are unfamiliar with noise in electronic circuits are advised to study this chapter before continuing. They may also want to study the summary on signal modeling to understand the concepts of *probability density functions* and *spectral density*.

In section 19.2, it has been shown that the noise behavior of a two-port can be modeled with the aid of two equivalent noise sources. Figure 2.27 shows a noise model with those two equivalent noise sources at the input port: a noise current source in parallel with the input port and a noise voltage source in series with the input port of the amplifier. With the aid of the Thévenin or Norton transformations, these two sources can be represented by one equivalent input noise source of the same type as the signal source (voltage or current source). This makes it possible to model the degradation of the signal-to-noise ratio due to these two equivalent sources. The noise of this source is usually called the *total equivalent input noise* or the *total source-referred noise*.

Its evaluation will be elucidated in the following example.

Example 2.8

We will study the influence of the equivalent noise sources on the signal-to-noise ratio at the output of the amplifier. Since the amplifier processes noise and signals in the same way, this can be done by relating the RMS value of the source voltage or current to the total equivalent input RMS noise voltage or current. respectively.

Figure 2.28A shows a situation in which the source signal is a voltage. The signal-to-noise ratio then needs to be expressed in terms of voltage ratios, hence, we need to evaluate the total source-referred noise voltage.

In Figure 2.28A, the amplifier's noise behavior is modeled with the aid of an equivalent input voltage noise source V_n and an equivalent input current noise source I_n. The noise associated with the signal source V_s is modeled by V_{ns}.

The total equivalent input noise voltage can now be obtained by transforming the equivalent input current noise source I_n into an equivalent input voltage noise source using the Thévenin equivalent of I_n and Z_s, as shown Figure 2.28B. The result with one equivalent input voltage noise source is shown in Figure 2.28C. If

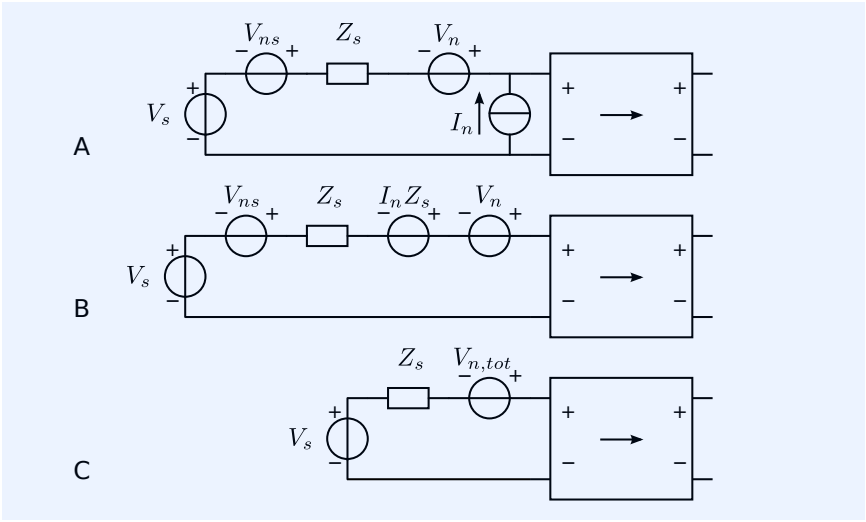


Figure 2.28: Determination of the total equivalent input noise voltage for an amplifier driven from a voltage source.

we assume no correlation between I_n, V_n and V_{ns}, the spectral density $S_{Vn,tot}$ of the total equivalent input voltage noise $V_{n,tot}$ can be obtained as:

$$S_{Vn,tot} = S_{V_{ns}} + S_{V_n} + S_{I_n} |Z_s|^2 , \tag{2.38}$$

where $S_{V_{ns}}$ [V²/Hz] is the spectral density of V_{ns}, S_{V_n} [V²/Hz] is the spectral density of V_n and S_{I_n} [A²/Hz] is the spectral density of I_n.

Noise figure

The total source-referred noise can thus be evaluated with the aid of its two equivalent input noise sources, the source impedance and the noise associated with the source signal. The noise figure F is a measure for the deterioration of the signal-to-noise ratio due to noise added by the amplifier.[16] It can now alternatively be defined as:

[16] The noise figure F of a system has been defined in section 19.3.

$$F = \frac{\text{(weighted) total equivalent input noise power}}{\text{(weighted) source noise power}}. \tag{2.39}$$

According to this definition, we obtain:

$$F = \frac{\int_0^\infty S_{V_{n,tot}} |W(f)|^2 \, df}{\int_0^\infty S_{V_{ns}} |W(f)|^2 \, df} \tag{2.40}$$

In which $W(f)$ is a weighting function that models the frequency dependent noise sensitivity of the observer.

Determination of the equivalent input noise sources

Figure 2.27 shows an amplifier, modeled as a noise-free two-port with two equivalent noise sources at the input. The spectral densities of these voltage and current noise sources accurately model the amplifier's noise behavior for any input port termination.

In fact, these two noise sources model the contributions of all physical noise sources that exists in the amplifier. They can be obtained by multiplying the internal noise sources of the circuit with the reciprocal values of their associated transfer functions to the input current and voltage of the circuit. As a consequence, the equivalent input noise current and noise voltage

sources may be partly correlated.

Transformation of internal noise sources to equivalent input noise sources can be performed with the aid of two-port models. This will be elucidated in the following example.

Example 2.9

We will calculate the total equivalent input noise of an amplifier that is modeled as a noise-free two-port with two uncorrelated equivalent output noise sources, as shown in Figure 2.29.

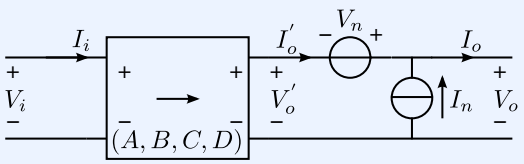


Figure 2.29: Noisy amplifier modeled as noise-free two-port with two equivalent output noise sources.

We have the following network equations:

$$\begin{pmatrix} V_i \\ I_i \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_o' \\ I_o' \end{pmatrix}, \tag{2.41}$$

$$V_o = V_o' + V_n, \tag{2.42}$$

$$I_o = I_o' + I_n, \tag{2.43}$$

where V_n and I_n denote the equivalent output noise sources. If we substitute the expressions for V_o and I_o in the two-port equations, we obtain

$$\begin{pmatrix} V_i \\ I_i \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_o - V_n \\ I_o - I_n \end{pmatrix}. \tag{2.44}$$

The output noise sources can now be transferred to the input:

$$\begin{pmatrix} V_i + AV_n + BI_n \\ I_i + CV_n + DI_n \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_o \\ I_o \end{pmatrix}. \tag{2.45}$$

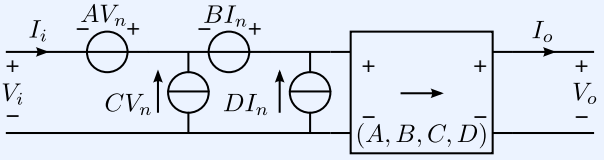


Figure 2.30: Circuit equivalent to that of Figure 2.29, but now with equivalent input noise sources.

Expression 2.45 and Figure 2.30 clearly show correlation between the voltage noise and the current noise. Both the total noise voltage and the total noise current consist of contributions of the two originally uncorrelated sources V_n and I_n. In the next example, we will demonstrate how to evaluate the total equivalent input noise for partly correlated input noise sources.

Example 2.10

Let us now connect a signal voltage source V_s with internal impedance Z_s to the input of the amplifier from Figure 2.30, and calculate the total equivalent input voltage noise.

Figure 2.31A shows the initial model. The spectral density of the voltage noise

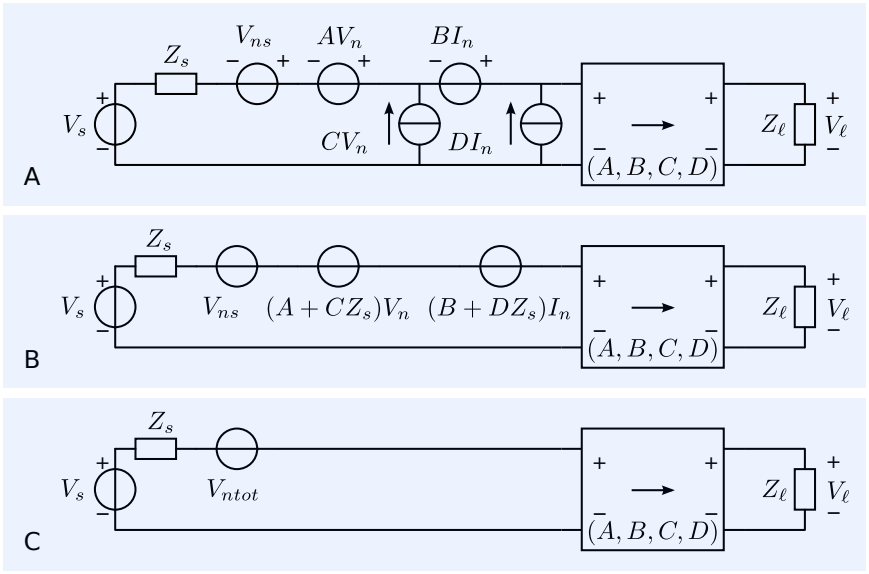


Figure 2.31:
A. Amplifier from Figure 2.30 connected to a signal source V_s with source impedance Z_s.
B. Equivalent noise model.

associated with the source is V_{ns} is given as:

$$S_{V_{ns}} = 4kT \operatorname{Re}\{Z_s\}. \tag{2.46}$$

The noise V_{ns} is assumed to be uncorrelated with I_n and V_n.

With the aid of the Thévenin transformation, we can convert the current noise sources into equivalent voltage noise sources. Figure 2.31B shows the result of this transformation. Figure 2.31C shows the final model in which the total source referred noise V_{ntot} has been represented with one voltage noise source. The power spectral density of V_{neq} of the total noise is obtained as:

$$S_{V_{ntot}} = 4kT \operatorname{Re}(Z_s) + |A + CZ_s|^2 S_{V_n} + |B + DZ_s|^2 S_{I_n}. \tag{2.47}$$

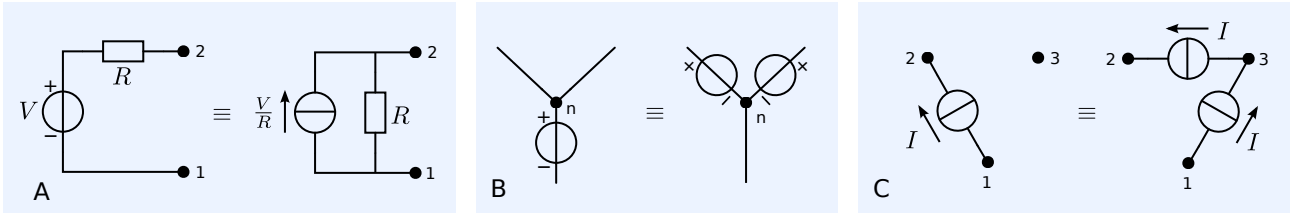


Figure 2.32: Network transformation:
(A) Thévenin and Norton equivalent networks
(B) voltage shift theorem
(C) current split theorem.

Transformation of noise sources can sometimes be simplified using network transformation techniques. The use of Thévenin and Norton equivalent networks has already been demonstrated in the previous examples. Figure 2.32 shows the Thévenin and Norton transformation, as well as Blakesley's voltage shift theorem (see [Blakesley1994][17]) and its dual version, the current split (redirecetion) theorem.

The current split theorem states that the network solutions remains unchanged if a current flowing from node (1) into node (2) is redirected via node (3). The voltage shift theorem states that the network solutions remain unchanged if a voltage source is shifted from one branch through a node and inserted into all the other branches connected to that node.

Example 2.11

In this example, we will evaluate the influence of impedance in series and in parallel with the signal path on the noise performance using the above network

[17] T. A. Blakesley. A New Electrical Theorem. *Proc. Phys. Soc. London*, 13:65–67, 1994

transformations.

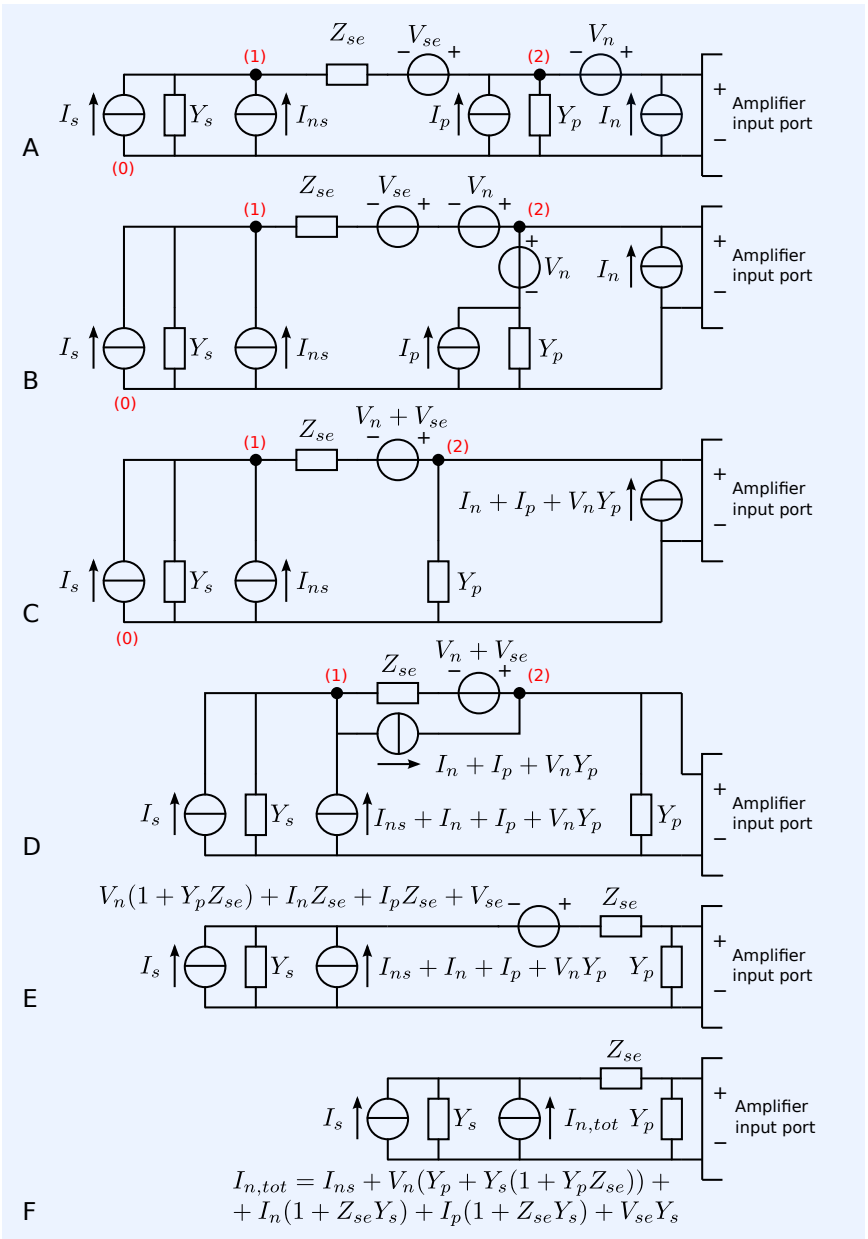


Figure 2.33:
The noise sources in this figure are uncorrelated. The plus signs in the figure indicate the summation of the Fourier Transforms of noise voltages or currents, rather than the summation of RMS values. A: Amplifier with equivalent input noise sources, signal (current) source and impedances in series and parallel with the signal path.
B: V_n is shifted through node (2).
C: The network between node (2) and (0) has been replaced with its Norton equivalent circuit: the currents in parallel with Y_p and the voltages in series with Z_{se} have been totalized.
D: The total noise current in parallel with the input port of the amplifier has been redirected over node (1): the noise currents in parallel with the signal source have been totalized.
E: The circuit between node (1) and node (2) has been replaced with its Thévenin equivalent circuit.
F: The series connection of the voltage of this Thévenin equivalent circuit and the source adtmittance Y_s has been replaced with its Norton equivalent circuit; the currents in parallel with the signal source have been totalized.

Let us consider the circuit from Figure 2.33A. It shows a signal current source I_s with a source admittance Y_s, which has been connected to an amplifier through a passive network that consists of an admittance Y_p in parallel with the signal path and an impedance Z_{se} in series with the signal path. The uncorrelated noise sources in Figure 2.33A are:

- *I_{ns}: current noise associated with Y_s. It has a spectral density $S_{I_{ns}} = 4kT\,\mathrm{Re}(Y_s)$ A^2/Hz.*
- *I_n: equivalent input current noise of the amplifier. It has a spectral density S_{I_n} A^2/Hz.*
- *V_n: equivalent input voltage noise of the amplifier. It has a spectral density S_{V_n} V^2/Hz.*

- *I_p: current noise associated with Y_p. It has a spectral density $S_{I_p} = 4kT\,\mathrm{Re}(Y_p)$ A^2/Hz.*

- *V_{se}: voltage noise associated with Z_{se}. It has a spectral density $S_{V_{se}} = 4kT\,\mathrm{Re}(Z_{se})$ V^2/Hz.*

The noise transformation steps have been listed with the figure. Please notice that all the above noise sources are assumed uncorrelated. Hence, the plus signs in the figure indicate the summation of the Fourier Transforms of noise voltages or currents, rather than the summation of RMS values.

The spectral density $S_{I_{n,tot}}$ of the total equivalent input noise $I_{n,tot}$ is obtained as:

$$S_{i_{n,tot}} = S_{I_{ns}} + S_{I_n}\left|1 + Z_{se}Y_s\right|^2 + S_{V_n}\left|Y_p + Y_s\left(1 + Z_{se}Y_p\right)\right|^2 + S_{I_p}\left|1 + Z_{se}Y_s\right|^2 + S_{V_p}\left|Y_s\right|^2. \tag{2.48}$$

Evaluation of the noise performance with the aid of network transformations, as illustrated in the previous example is not always the fastest or the most straightforward method. It is also difficult to automate this procedure.

Clear and straightforward noise analysis can be performed with the aid of the modified nodal analysis technique. This method is used by SPICE-like simulators and by SLiCAP.

The procedure is as follows:

1. Set up the matrix equations of the network and use independent sources for uncorrelated noise sources and for the signal source. Correlation between noise sources needs to be modeled with the aid of controlled sources.

2. Define the load quantity: a nodal voltage or a current through a voltage source.

3. Define the source quantity: the voltage or current of an independent source.

4. For each noise source:

   (a) Evaluate the squared magnitude of the transfer function from that noise source to the load quantity.

   (b) Evaluate the contribution to the load-referred noise by multiplying the spectral density of that noise source with the result from the previous step.

5. The total spectral density of the load-referred noise is found by adding all of the above contributions.

6. The total spectral density of the source-referred noise is found by dividing the above result by the squared magnitude of the source-to-load transfer.

This procedure will be elucidated in the example below.

Example 2.12

Figure 2.34 shows a signal source represented by a current source I_s in parallel with a capacitance C_s. This signal source is connected to the input of a noisy two-port. An attenuator that consists of a series impedance Z_s and a parallel admittance Y_p has been placed between the current source and the input port of the noisy two-port. The noisy two-port is modeled as a noise-free two-port with an equivalent-input noise voltage source V_e and an equivalent-input noise current source I_e. Noise sources associated with Z_s and Y_p have been included. The spectral densities of the noise sources have been shown in the figure.

We will determine the spectral density of the total source-referred current noise using Modified Nodal Analysis. We will use the voltage at node 4 (V_4) as the output variable. By doing so, we do not need to include the noise-free two-port in the matrix equations. This is allowed if the transfer from V_4 to the actual signal load differs from zero. For the sake of simplicity, we will assume this to be the case.

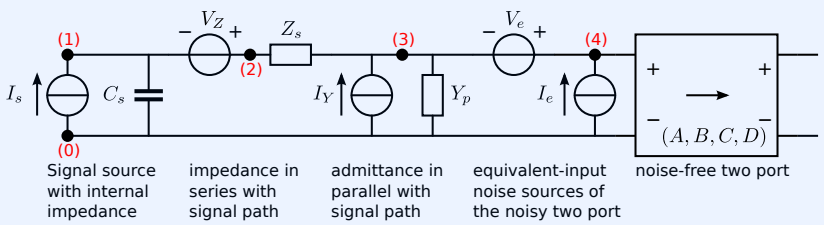


Figure 2.34: A capacitive current source cascaded with a passive network is connected to the input of a noisy two-port.

I_s: Signal source
V_Z: Noise voltage associated with Z_s. Spectral density: $S_{V_Z} = 4kT\mathrm{Re}(Z_s), V^2/Hz$
I_Y: Noise current associated with Y_p. Spectral density: $S_{I_Y} = 4kT\mathrm{Re}(Y_p), A^2/Hz$
V_e: Equivalent-input noise voltage of the two port. Spectral density: $S_{Ve}, V^2/Hz$
I_e: Equivalent-input noise current of the two port. Spectral density: $S_{Ie}, A^2/Hz$

The MNA matrix equation for the circuit, excluding the noise-free two-port can be written as:

$$\mathbf{I} = \mathbf{MV}. \tag{2.49}$$

In this equation, $\mathbf{I}$ is the vector with the Fourier Transforms of the independent noise and uncorrelated noise voltages and currents:

$$\mathbf{I} = \begin{pmatrix} 0 & 0 & I_Y & I_e & V_Z & V_e \end{pmatrix}^T. \tag{2.50}$$

The matrix $\mathbf{M}$ is the MNA matrix (see Chapter 18):

$$\mathbf{M} = \begin{pmatrix} j\omega C_s & 0 & 0 & 0 & -1 & 0 \\ 0 & \frac{1}{Z_s} & -\frac{1}{Z_s} & 0 & 1 & 0 \\ 0 & -\frac{1}{Z_s} & \frac{1}{Z_s} + Y_p & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \end{pmatrix}, \tag{2.51}$$

and $\mathbf{V}$ is the vector the Fourier Transforms of the dependent variables:

$$\mathbf{V} = \begin{pmatrix} V_1 & V_2 & V_3 & V_4 & I_{V_Z} & I_{V_e} \end{pmatrix}. \tag{2.52}$$

The voltage V_4 can be obtained with the aid of Cramer's rule:

$$V_4 = \frac{\det \mathbf{M}'}{\det \mathbf{M}}, \tag{2.53}$$

where:

$$\mathbf{M}' = \begin{pmatrix} j\omega C_s & 0 & 0 & 0 & -1 & 0 \\ 0 & \frac{1}{Z_s} & -\frac{1}{Z_s} & 0 & 1 & 0 \\ 0 & -\frac{1}{Z_s} & \frac{1}{Z_s} + Y_p & I_Y & 0 & -1 \\ 0 & 0 & 0 & I_e & 0 & 1 \\ -1 & 1 & 0 & V_Z & 0 & 0 \\ 0 & 0 & -1 & V_e & 0 & 0 \end{pmatrix}. \tag{2.54}$$

In order to evaluate the source referred noise, we need to know the transfer

from the source to V_4. This transfer can be obtained as:

$$\frac{V_4}{I_s} = \frac{\text{cofactor}(\mathbf{M}, 1, 4)}{\det \mathbf{M}}\,, \tag{2.55}$$

where:

- $\text{cofactor}(\mathbf{M}, 1, 4) = (-1)^{(1+4)}\,\text{minor}(\mathbf{M}, 1, 4)$
- $\text{minor}(\mathbf{M}, 1, 4)$: *determinant of matrix* $\mathbf{M}$*, after deleting row* 1 *and column* 4 :

$$\text{minor}(\mathbf{M}, 1, 4) = \det \begin{pmatrix} 0 & \frac{1}{Z_s} & -\frac{1}{Z_s} & 1 & 0 \\ 0 & -\frac{1}{Z_s} & \frac{1}{Z_s} + Y_p & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix}. \tag{2.56}$$

The source-referred noise I_{ns} can thus be obtained as:

$$I_{ntot} = \frac{V_4}{\frac{V_4}{I_s}} = \frac{\det \mathbf{M}'}{\text{cofactor}(\mathbf{M}, 1, 4)}. \tag{2.57}$$

After collecting the terms for each (uncorrelated) noise source, the result can be written as:

$$\begin{aligned} I_{ntot} = {} & I_e \left(1 + j\omega C_s Z_s\right) \\ & + I_Y \left(1 + j\omega C_s Z_s\right) \\ & + V_e \left(Y_p + j\omega C_s \left(1 + Y_p Z_s\right)\right) \\ & + V_Z j\omega C_s. \end{aligned} \tag{2.58}$$

In this way, we found the transfer from each noise source to its source-referred contribution. The spectral density of the total source-referred noise current is found as the sum of the spectral densities of the individual noise sources, each multiplied by the squared magnitude of their corresponding transfer:

$$\begin{aligned} S_{I_{ns}} = {} & S_{I_e} \left|1 + j\omega C_s Z_s\right|^2 \\ & + S_{I_Y} \left|1 + j\omega C_s Z_s\right|^2 \\ & + S_{V_e} \left|Y_p + j\omega C_s \left(1 + Y_p Z_s\right)\right|^2 \\ & + S_{V_Z} \omega^2 C_s^2. \end{aligned} \tag{2.59}$$

Determination of the equivalent input noise sources by measurements

If we want to determine the equivalent noise sources through measurements, we connect a spectrum analyzer to the output of an amplifier and measure the output voltage noise spectrum S_{Vm} [V²/Hz] for two different terminations for the input port :

1. The input terminals of the amplifier are shorted
2. The input terminals of the amplifier are left open.

Figure 2.35 shows the noise measurement setup. The amplifier is modeled as a noisy two-port with transmission-1 parameters A, B, C, D. It has a finite input impedance Z_m.

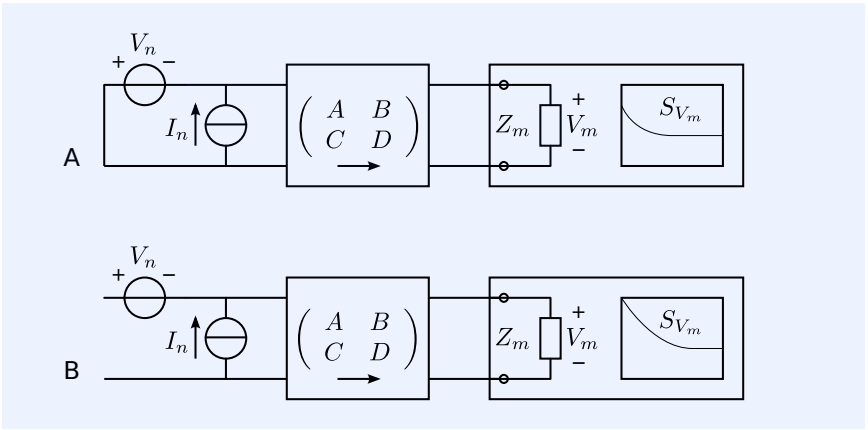


Figure 2.35: Determination of the equivalent input noise sources with the aid of noise measurements.

(A) The equivalent input noise voltage is determined from noise measurements with the input port shorted.

(B) The equivalent input noise current is determined from noise measurements with the input port left open.

With shorted input port terminals (see Figure 2.35A), the equivalent input noise current I_n flows through the short. The measured output noise must then be caused by the equivalent input noise voltage source V_n. The spectral density S_{V_n} [V²/Hz] of the input noise voltage V_n can thus be found with the aid of expression (2.12) as:

$$S_{V_n} = \left| A + \frac{B}{Z_m} \right|^2 S_{Vm} \text{ [V}^2\text{/Hz]}. \tag{2.60}$$

When the input port is left open (see Figure 2.35B), the equivalent input noise voltage V_n is floating, thus the measured output voltage noise can only be caused by the equivalent input noise current source I_n. The spectral density S_{I_n} [A²/Hz] of the input noise current I_n can then be found with the aid of expression (2.14) as:

$$S_{I_n} = \left| C + \frac{D}{Z_m} \right|^2 S_{Vm} \text{ [A}^2\text{/Hz]}. \tag{2.61}$$

Design conclusions

Expression 2.48 clearly shows the result of inserting impedances in series or in parallel with the signal path between the source and the input port of an amplifier:

1. If an impedance in series with the source has a nonzero real part, it adds noise.
2. Any nonzero impedance in series with the source increases the contribution of the equivalent input noise current of the amplifier.
3. If an admittance in parallel with the source has a nonzero real part, it adds noise.
4. Any admittance in parallel with the source increases the contribution of the equivalent input voltage noise of the amplifier.

From this, we obtain an important design conclusion:

Insertion of impedances in series or in parallel with the signal path at the input of the amplifier should be avoided!

The deterioration of the signal-to-noise ratio due to the insertion of impedances in series or in parallel with the signal path between the source and the input of the amplifier can be explained as follows.

In general, those impedances reduce the available signal power at the input port of the amplifier. Hence, if we first reduce the available power of the source and then add noise, it will become clear that the signal-to-noise ratio will be lower compared with the situation in which no attenuator was inserted.

Only in narrow-band applications, the insertion of impedances in parallel or in series with the signal path between the source and the input port of the amplifier may result in an improvement of the signal-to-noise ratio in the frequency band of interest. Such situations will be discussed at a later stage.

2.4.4 Modeling of the power efficiency

The power efficiency η of an amplifier is defined as the ratio of the load power P_ℓ and the power P_s delivered by the power supply sources of the amplifier:

$$\eta = \frac{P_\ell}{P_s}. \tag{2.62}$$

The ideal amplifier has a power efficiency of at least unity. Real-world amplifiers have a power efficiency less than unity.

Two-port model

We have modeled the ideal amplifier as a two-port. The output port of this two-port can be modeled as a controlled source or, in the case of negative feedback, as a norator.[18] This element delivers the load power in case there are no impedances placed in series or in parallel with the load. In real world amplifiers, the power that needs to be delivered by the power supply exceeds the load power. This is the result of power losses in the amplifier.

[18] This will be discussed in Chapter 7.

Power losses (dissipation) in amplifiers arise from:

1. Quiescent losses

   Electronic circuits may consume power even in the case in which no power or energy is transferred to the load. These quiescent losses are measured in the absence of a source signal.

   Quiescent losses can be modeled with the aid of a voltage source V_Q and a nonlinear resistor R_Q. This is shown in Figure 2.36. If V_Q is equal to the supply voltage, the $v - i$ characteristic of this nonlinear resistor should equal that of the power supply port of the amplifier. The quiescent losses then equal the power dissipated in this nonlinear resistor.

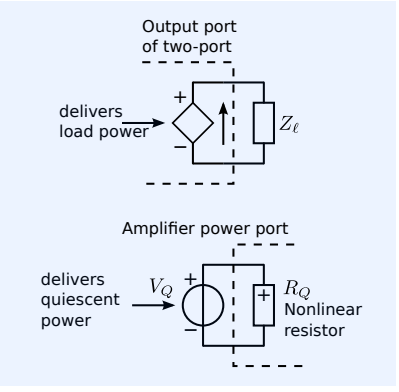


Figure 2.36: Load power delivered by the output port and quiescent power delivered by V_Q.

2. Signal-dependent losses and energy storage

   Voltage drop across devices, as well as current flow through devices, may result in energy storage or power dissipation in those devices. Energy stored in devices that, under dynamic signal conditions, is not fully exchanged with the power supply, contributes to power losses.

   Signal-dependent power dissipation and energy storage can be modeled with the aid of two controlled sources: a voltage-controlled source and a current-controlled source.

   Figure 2.37 shows the application of a linear voltage-controlled voltage source and a linear current-controlled current source for this purpose. In this figure, the impedance Z_p models the frequency-dependent and signal dependent energy storage and power dissipation in the case in which the amplifier output terminals have been left open. The impedance Z_{se} models

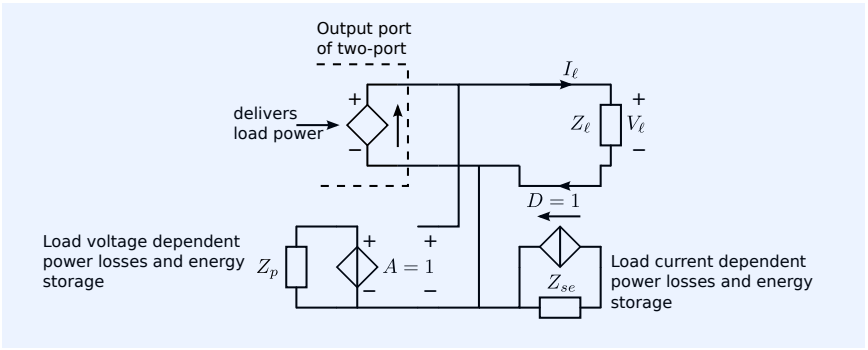


Figure 2.37: Load power delivered by the output port and load voltage-dependent and load current-dependent power losses and energy storage, modeled with the aid of controlled sources and linear impedances.

the frequency-dependent and signal dependent energy storage and power losses in the case in which the amplifier terminals have been shorted.

However, modeling of the power losses and of the energy storage as described above is seldom performed. This is because:

(a) Linear models for power losses and energy storage seldom apply to real world amplifiers in which those effects usually show a strong nonlinear relationship with the load voltage and the load current.

(b) The above model does not provide design information other than minimization of energy storage and power losses in elements that carry signal.

Impedances in the signal path

We have seen that insertion of impedances in series or in parallel with the signal path adversely affects the signal-to-noise ratio. Similarly, insertion of such impedances also reduces the power efficiency. In properly designed amplifiers, the largest signal excursions occur at the load. For this reason, care should be taken with insertion of impedances in series or in parallel with the load:

1. If an impedance in series with the load has a nonzero real part, it decreases the power efficiency because it dissipates power.
2. If an impedance in series with the load has a nonzero imaginary part, it increases the energy storage in the amplifier. If, under dynamic signal conditions, this energy is not recovered, it results in an increase of the power dissipation of the amplifier and reduces its power efficiency.
3. If an admittance in parallel with the load has a nonzero real part, it decreases the power efficiency because it dissipates power.
4. If an admittance in parallel with the load has a nonzero imaginary part, it increases the energy storage in the amplifier. If, under dynamic signal conditions, this energy is not recovered, it results in an increase of the power dissipation of the amplifier and reduces its power efficiency.

In narrow-band amplifiers, the power efficiency of the amplifier does not necessarily degrade due to energy storage. In such amplifiers, the power losses can be minimized with the aid of resonant circuits.

2.4.5 Power losses and amplifier classes

More accurate modeling of the power losses requires knowledge about the amplifier's internal structure, especially about its output stage.

Many amplifiers have to provide bipolar currents to their loads.[19] The electronic devices from which they are constructed, however, are operating in unipolar mode[20]. As a consequence, devices in output stages can either *source* a current to a load, or *sink* a current from a load as shown in Figure 2.38.

One method for delivering bipolar currents is to add a bias current to a device such that it will conduct for both positive and negative load currents. Amplifiers that use this principle are referred to as *class A* amplifiers. Their output stage is said to operate in class A mode: *the output device conducts during the source and the sink phase*. Class A output stages suffer from a low power efficiency. Different classes of output stages have been developed to increase the power efficiency of amplifier output stages. A short overview is given below.

[19] Bipolar currents: currents that can flow in positive and negative directions.
Unipolar currents: currents that flow in either a positive or negative direction.
[20] Single-quadrant operation in the $v - i$ plane.

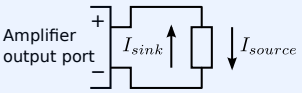


Figure 2.38: The amplifier sources current to its load if the current flows out of its positive output terminal. The amplifier sinks current from its load if the current flows into its positive output terminal.

1. Class A: a sourcing output device and a sinking output device both conduct during the entire source phase and during the entire sink phase.

   Class A amplifiers suffer from a low power efficiency.

2. Class B: a sourcing output device conducts exclusively during the entire source phase, while a sinking output device conducts exclusively during the entire sink phase.

   A class B amplifier has no overlap between the source and the sink phase. In practice, this will never be the case: either there will be an overlap (class AB operation) or there will be a dead zone (class C operation). Class B, therefore, is a concept without implementations.

3. Class AB: a sourcing output device conducts predominantly during the source phase, while a sinking output device conducts predominantly during the sink phase.

   All operational amplifiers and almost all low-frequency non-switching power amplifiers have a class AB output.

   The power efficiency of class AB amplifiers exceeds that of class A amplifiers.

4. Class C: a sourcing and/or a sinking output device conduct only during a part of the source phase and/or the sink phase, respectively.

   Class C output stages produce a large amount of distortion due to a dead zone in their transfer. In narrow-band applications the resulting distortion components can be attenuated with band pass filters. Class C amplifiers are often applied as narrow-band RF output stages with a relatively high power efficiency.

5. Class D: switching output stage.

   Class D amplifiers have a switching output stage. The information at the output of the switching stage is the low-frequency contents of a PWM[21] signal. This low-pass contents is passed to the load through a low-pass filter. High-frequency components of the PWM signal are reflected into the power supply. This gives the class D amplifier a high power efficiency.

6. Class E: tuned (narrow-band) resonant switching output stage.

   Class E amplifiers are resonant switching amplifiers: a switch closes when the voltage across it equals zero. A narrow-band and slightly detuned series resonator is placed between the switch output and the load. Class E amplifiers are applied as RF transmitter amplifiers. They exhibit a better power efficiency compared with class C amplifiers.

[21] Pulse Width Modulated

7. Class F: tuned (narrow-band) switching output stage with stacked output filters

   Class F amplifiers are switching amplifiers with a more complex output filter. This filter is designed such that the voltage at the output of the switches approximates a square wave, while the load voltage is sinusoidal. These amplifiers are used in narrow-band applications.

8. Class G: class AB with step-wise adaption of the power supply voltages to the load voltage

   Class G amplifiers have class AB output stages of which the power supply voltages are step-wise adapted to the load voltage.

   This technique reduces the power dissipation of the class AB amplifier.

9. Class H: class AB with continuous adaption of the power supply voltages to the load voltage

   Class H amplifiers have class AB output stages of which the power supply voltages are continuously adapted to the load voltage.

   With this technique, a low and constant voltage drop across the output devices minimizes the power dissipation.

It is common practice to speak of *linear amplifiers* (class A, B, AB and C) versus *switching amplifiers* (class D, E and F). Unfortunately, this is rather confusing and conceptually wrong. All amplifiers are intended to be linear, and switching is not the opposite of linear. It would be better to speak of amplifiers with switching output stages and those with non-switching output stages. Alternatively, one could speak of baseband output stages and modulating output stages.

In the following example, we will evaluate the power losses of an amplifier that periodically charges and discharges a capacitive load.

Example 2.13 *The amplifier has a positive supply voltage source with a value V_p and a negative supply voltage source with a value V_n. In the absence of an input signal, the amplifier draws a negligible quiescent current. Hence we may assume class B operation.*

When charging the load with a charge $q(t)$, the positive supply delivers the source current $i_p(t)$:

$$i_p(t) = \frac{d}{dt} q(t). \tag{2.63}$$

The momentary source power $P_+(t)$ delivered by the positive supply source can be calculated as:

$$P_+(t) = i_p(t) V_{p.} \tag{2.64}$$

The average supply power $\overline{P_c}$ over period T delivered during the charging phase can be calculated as:

$$\begin{aligned}
\overline{P_c} &= \frac{1}{T} \int_0^T P_+(t) dt, && (2.65)\\
&= \frac{1}{T} \int_0^T i_p(t) V_p dt, && (2.66)\\
&= \frac{1}{T} \int_0^T \frac{d}{dt} q(t) V_p dt, && (2.67)\\
&= \frac{1}{T} \left(q(T) - q(0) \right) V_p, && (2.68)
\end{aligned}$$

where $q(T)$ is the load charge at the end of the charging phase, say, Q_p, and $q(0)$ is the load charge at the start of the charging phase, say, Q_n, we can write:

$$P_c = \frac{1}{T} \left(Q_p - Q_n \right) V_p. \tag{2.69}$$

If the load conserves the charge, the power dissipation in the load is zero and

the power delivered by the positive supply is dissipated in the amplifier.

During a discharging phase, the negative voltage supply delivers the sink current. The average supply power $\overline{P_d}$ over the period T delivered during the discharging phase can be similarly obtained:

$$\overline{P_d} = \frac{1}{T}\left(Q_n - Q_p\right) V_n. \tag{2.70}$$

Again, if the load conserves the charge, this power is dissipated in the amplifier.

Hence, if the amplifier periodically charges and discharges a capacitive load, the average power dissipation in the amplifier $\overline{P_A}$ can be obtained as:

$$P_A = \frac{1}{T}\left(Q_p - Q_n\right)\left(V_p - V_n\right), \tag{2.71}$$

or, alternatively,

$$P_A = fQ_{pp}\left(V_p - V_n\right), \tag{2.72}$$

where Q_{pp} is the peak to peak charge.

If the load is a linear capacitor with a capacitance C_ℓ, we may write:

$$Q_{pp} = C_\ell V_{pp}, \tag{2.73}$$

where V_{pp} is the peak to peak load voltage. After substitution of (2.72) in (2.73), we obtain:

$$P_A = fC_\ell\left(V_p - V_n\right) V_{pp}. \tag{2.74}$$

From this we see that although the amplifier is loaded with a lossless element, the periodic change of the energy storage in it results in amplifier losses.

2.4.6 Modeling of the small-signal dynamic behavior

The ideal amplifier is an instantaneous, linear time-invariant system. However, due to the fundamental limitation of speed, all real world amplifiers will behave like dynamic systems. In fact, they will behave like nonlinear dynamic systems. The specification and design of amplifiers, however, is strongly facilitated by separating performance aspects such as nonlinearity and dynamic behavior. For ease of design, we consider them either as linear(ized) dynamic systems or as nonlinear instantaneous systems. At relatively low rates of change of the signals, such an approach is almost always justified. Hence, the amplifier's behavior for small signals up to high frequencies may be considered linear dynamic, and its behavior for low speed signals up to large signal excursions may be considered as nonlinear instantaneous. Behavior at relatively high rates of change of the signals is often referred to as *large signal dynamic behavior*. This will be discussed later.

The reader is assumed to be familiar with the modeling of linear time-invariant dynamic systems. Those who are not, or those who want to refresh their knowledge, can find a summary on modeling of linear time invariant-dynamic systems in Chapter 17.4.

The small-signal dynamic behavior of an amplifier can be characterized with:

1. Time-domain responses:

   (a) Impulse response $h(t)$. The unit impulse response is only of theoretical use.

   (b) The step response $a(t) = \int_{-\infty}^{t} h(\tau)d\tau$ is often used for time domain characterization. As a matter of fact, a periodic square wave signal is often used instead of a single step. The response to such a periodic signal can easily be observed on an oscilloscope (see Figure 2.39).

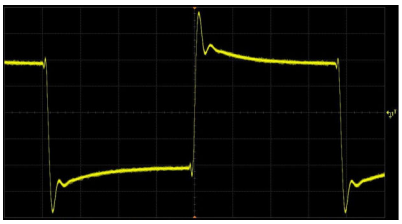


Figure 2.39: Example of a small-signal step response of an amplifier.

Figure 2.40: Characterization of a step response.

Characterization of the step response is done in terms of:

i. Settling time: the time it takes before the difference between the signal value and the final value stays within a predefined error budget
ii. Final value: the value after settling
iii. Ringing, overshoot and undershoot: peak errors with respect to settling value
iv. Delay time: the time it takes to reach 50% of the settling value
v. Rise time: the time it takes to rise from 10% to 90% of the settling value
vi. Fall time: the time it takes to fall from 90% to 10% of the settling value
vii. Droop or tilt: rate of change of a relatively slow signal change after settling

Figure 2.40 shows a step response with the definition of the most important parameters.

2. Frequency domain description: the transfer function $H(j\omega)$ is the Fourier transform of its unit impulse response $h(t)$. It can be characterized by:

   (a) Bode plots. These plots consist of a magnitude versus frequency plot $|H(j\omega)|$ and a phase versus frequency plot $\phi(j\omega) = \arg(H(j\omega))$
   (b) The transfer bandwidth B of $H(j\omega)$, which is characterized by means of the low frequency -3dB frequency and the high frequency -3dB frequency and the order of the low-frequency and high-frequency roll-off
   (c) The group delay $\tau(\omega)$ versus frequency: $\tau(\omega) = -d\phi(\omega)/d\omega$
   (d) The real part and the imaginary part of $H(j\omega)$ versus frequency

   This can be a useful presentation for the small-signal port impedances. If the real part of a port impedances is positive for all frequencies, the amplifier is stable for all termination impedances of that port. In other words the reflected power at that port is always less than the power transmitted to that port.

3. Complex frequency domain description: the system function $H(s)$ is the Laplace transform of its unit impulse response $h(t)$

   (a) The amplifier's system function $H(s)$ is often characterized by means of its poles and zeros.

   As we will see later, this representation method is often used during design. Performance evaluation by means of measurements is always done with time domain of frequency domain measurements.

   (b) Poles and zeros are real or complex conjugated.
   (c) A system is stable if all system poles have a negative real part.

All these description methods can be translated into each other. The reader is assumed to be familiar with relations between pole-zero diagrams, Bode plots and step responses.

Figure 2.41 shows various graphic representations of the small-signal dynamic behavior of a second order low-pass transfer function with an MFM characteristic.[22]

[22] MFM: Maximally Flat Magnitude.

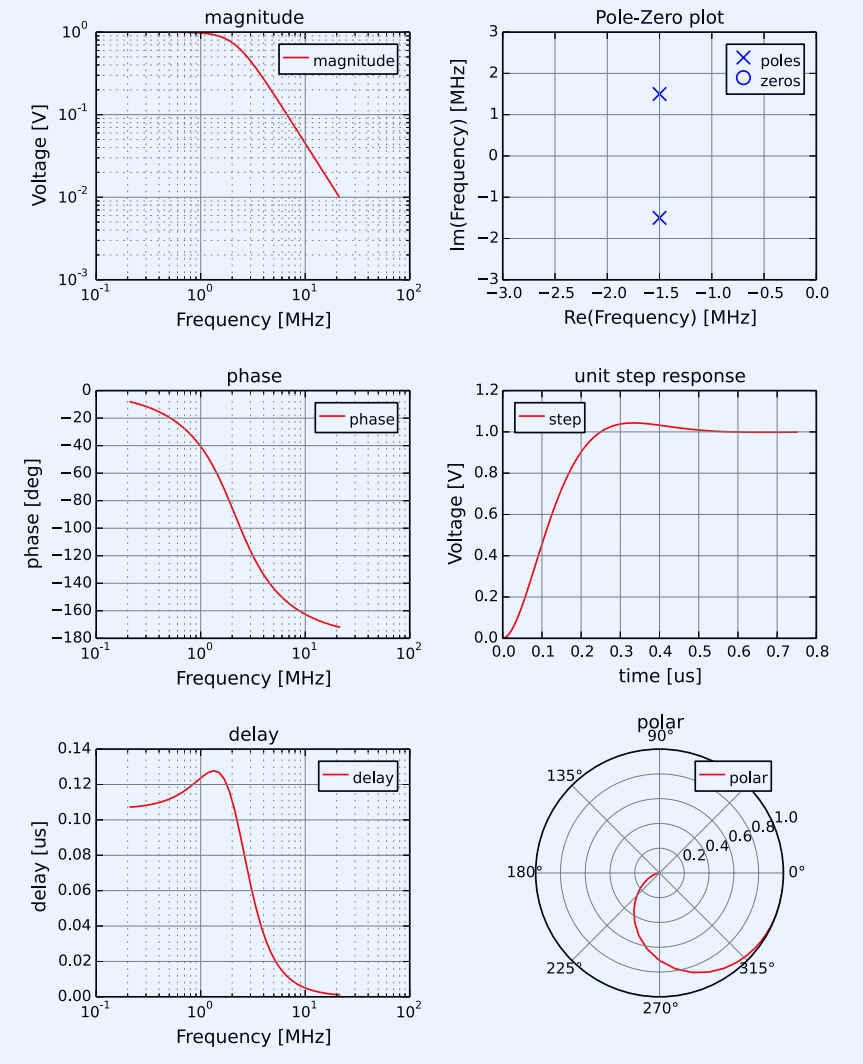


Figure 2.41: Various representations of the small-signal transfer of a linear stationary dynamic system.

2.4.7 Modeling of the static nonlinear behavior

In this section, we will discuss deviations from the linear behavior of amplifiers. General techniques and for describing stationary instantaneous nonlinear behavior have been summarized in section 17.5. Readers who are not familiar with Taylor series and with the terms 'offset', 'nonlinearity', 'differential gain', 'harmonic distortion' and 'intermodulation distortion' are referred to this section.

Instantaneous, DC and AC behavior

Description methods for the static or instantaneous behavior of amplifiers are valid for amplifiers that behave instantaneously. In practice, this does not necessarily mean that the amplifier shows no dynamic behavior. It simply means that dynamic effects are not of interest, or simply neglected because they are too small. This may be the case if one is only interested in the DC solution of a network, or if the frequencies of the poles and zeros are outside the frequency range of interest.[23]

In electronics, it is common practice to speak of *DC behavior* and of *AC*

[23] DC: direct current
AC: alternating current

behavior. The DC behavior of a circuit describes the behavior of a circuit all the capacitors of which have been replaced with open circuits and all inductors with short circuits, and in which all independent sources have their DC values, as described in section 18.3.2. The modified circuit thus shows no dynamic elements and behaves instantaneously. The AC behavior of a circuit describes the *small-signal dynamic behavior* at an operating point. The circuit is linearized at an operating point, obtained with a DC analysis. This small-signal model is also valid at zero frequency. The term AC is thus somewhat misleading.

The *operating point* of a circuit at a time instant t is the network solution of the circuit that instant in time. This solution accounts for flux in inductors and charge on capacitors and is generally not the same as the DC operating point in which such initial conditions are usually ignored.

In the network simulation program SPICE, these two operating points are known as:

- The DC operating point
- The operating point using initial conditions (UIC).

Example 2.14

An amplifier with a frequency-independent transfer between 10*Hz to* 10*MHz shows an approximate instantaneous behavior over this frequency range. The DC transfer of this amplifier may, however, strongly differ from the AC transfer over this frequency range. An audio amplifier, for example, should behave instantaneously at audible frequencies, however, a nonzero DC transfer is neither necessary nor desirable.*

Amplifiers that need to transfer signals with zero frequency are usually called DC amplifiers.

Bias and offset quantities

In section 2.2, we have characterized the instantaneous behavior of amplifiers with three curves:

1. The input port's $v - i$ characteristic
2. The input port to output port transfer characteristic
3. The output port's $v - i$ characteristic.

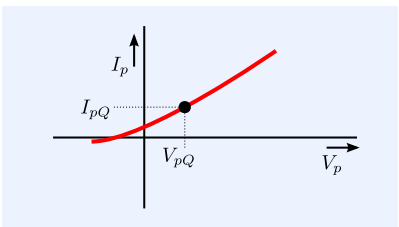


Figure 2.42: Example of a static $I_p(V_p)$ relation of an amplifier port.

For ideal amplifiers, these curves are straight lines through the origin, but in practice, these curves will be nonlinear and not pass the origin.

Figure 2.42 shows an example of a static nonlinear $v - i$ characteristic of a port. If we can select a point $Q = (V_{pQ}, I_{pQ})$ of this curve as the desired quiescent operating point of that port, this means that, in the absence of a signal, the port should operate in Q. This can be achieved by inserting a *bias voltage source* V_{pQ} in series with the port and a *bias current source* I_{pQ} in parallel with the port, as shown in Figure 2.43. By doing so, the quiescent operating point of the port becomes (V_{pQ}, I_{pQ}), while that of the termination (source or load) is $(0, 0)$. Hence, in this way, the poert operates in its desired operating point, which does not depend on the DC characteristics of a passive, nonlinear termination.

Application of the correct operating point to electronic devices is called *biasing*. In practice, after the devices have been biased, small *offset* errors may remain due to device tolerances, supply voltage variations, temperature variations and other causes. The influence of those small offset errors can be evaluated in a similar way as noise: small output offsets may be converted

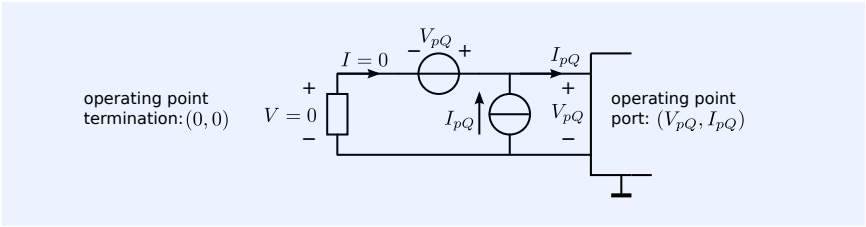


Figure 2.43: Insertion of a bias voltage source V_{pQ} in series with the port and a bias current source I_{pQ} in parallel with the port, shifts the operating point of the port with $v - i$ characteristics depicted in Figure 2.42 from $(0,0)$ to (V_{pQ}, I_{pQ}).

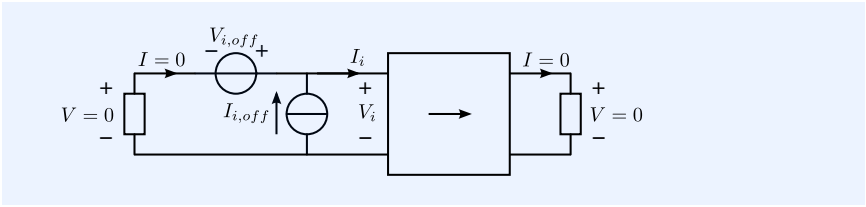


Figure 2.44: The influence of small input port and output port offset voltages and currents can be evaluated in a similar way as has been done with noise sources.

into equivalent input offsets, or vice versa, using a linearized model for the two-port.

Figure 2.44 shows a two-port of which those offset errors have been modeled with the aid of equivalent input current and voltage offset sources. As with noise, two equivalent offset sources are required for correct modeling for all DC port terminations.

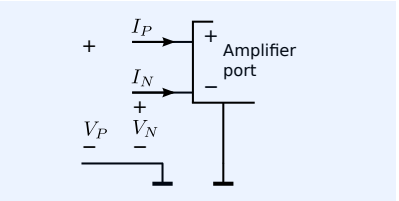


Figure 2.45: The common-mode bias voltage of an amplifier port is defined as: $V_{bias} = \frac{V_P + V_N}{2}$.
The port terminal bias currents are I_P and I_N for the '+' and the '-' port terminal, respectively.

Common-mode port bias quantities

Although we do not (yet) want to discuss the detailed design of electronic amplifiers, we should realize ourselves that practical amplifiers always exhibit limitations that directly depend on the properties of the electronic devices from which they have been constructed.

One of those limitations is that the operating voltage of electronic devices is limited by the power supply voltage and by the breakdown voltages of these devices. This poses limitations to the amplifier's terminal voltages, which is of particular interest for amplifiers with floating ports. In those cases, proper (common-mode) operating conditions for the floating port need to be provided by external circuitry.

- The common-mode bias voltage of an amplifier port is defined as half the sum of the port terminal voltages that need to be applied for correct port operation.

- The bias current is defined as the current that needs to flow into a port terminal for proper port biasing.

These definitions have been elucidated in Figure 2.45.

Clipping and voltage and current drive capability

The current and voltage drive capability of amplifiers is limited by the power supply and by voltage and current limiting mechanisms in the amplifier itself. If an amplifier delivers its maximum drive capability, we speak of *clipping* or *hard-limiting* of the output signal. It is usually characterized by the maximum voltage the amplifier can deliver for a given output current or vice versa.

Weak nonlinearity

When operating in the normal operating region (no clipping), the source to load transfer is intended to be linear. However, imperfections in the operating mechanism of the amplifier usually cause so-called weakly nonlinear behavior. The perception of signal processing errors due to weakly nonlinear behavior strongly depends on the way in which the information is embedded in the signal and the way in which the error is perceived by the observer. For this reason, there exist many different description methods for nonlinear behavior. Harmonic distortion, intermodulation distortion, gain compression and differential gain all express a specific perception of such behavior. In general, terms that accurately describe the observer's error perception should be used. In radio applications, intermodulation distortion and gain compression often closely describe the phenomena observed. Nonlinear behavior of amplifiers in control systems can better be characterized by their differential gain.

Harmonic distortion

Sinusoidal signals retain their shape in linear stationary dynamic systems. The nonlinearity of a stationary[24], nonlinear system can thus be characterized by the amount of distortion in the response to a sinusoidal excitation. The total harmonic distortion (THD) is defined as

[24] Also: time-invariant or fixed.

$$THD = \frac{1}{d_1}\sqrt{\sum_{n=2}^{n=\infty} d_n^2}, \tag{2.75}$$

where d_n is the amplitude of the $n-th$ harmonic in the output signal.

Even order nonlinearity causes even harmonic distortion as well as signal-dependent offset. This can easily be seen from the following expression:

$$A\cos^2 x = \frac{A}{2} + \frac{A}{2}\cos 2x. \tag{2.76}$$

This phenomena is known as "operating point shift" or "bias point shift".

Intermodulation distortion

If the excitation of a stationary non-linear system consists of two sinusoidal components with different frequencies ω_1 and ω_2, nonlinearity will give rise to output signal components at multiples of ω_1 and ω_2 (known as harmonic distortion) and at $m\omega_1 \pm n\omega_2$ (m and n are integers). The latter effect is called intermodulation distortion. The amplitudes of the signal components at these frequencies will be denoted as $A_{m\omega_1 \pm n\omega_2}$.

The second-order intermodulation distortion IM_2 is defined as the *relative* magnitude of the component in the output signal with $m = 1$ and $n = 1$, when the components at ω_1 and ω_2 have equal amplitudes:

$$IM_2 = \frac{|A_{\omega_1 \pm \omega_2}|}{|A_{\omega_1}|}. \tag{2.77}$$

The third-order intermodulation distortion IM_3 is defined as the relative magnitude of the component with $m = 2$ and $n = 1$, or $m = 1$ and $n = 2$, when both the components at ω_1 and ω_2 have equal amplitudes:

$$IM_3 = \frac{|A_{\omega_1 \pm 2\omega_2}|}{|A_{\omega_1}|} = \frac{|A_{2\omega_1 \pm \omega_2}|}{|A_{\omega_1}|}. \tag{2.78}$$

2.4.8 Modeling of the dynamic nonlinear dynamic behavior

The characterization of the large signal dynamic behavior, or the nonlinear dynamic behavior of amplifiers, also strongly depends on the perception of errors by the observer. In transceiver amplifiers, intermodulation distortion and gain compression are effects that are used to characterize nonlinear behavior. For video amplifiers and amplifiers in negative feedback control loops, differential gain and differential phase are often used. Other performance parameters for characterizing nonlinear dynamic behavior, such as overdrive recovery, slew rate limitation and hysteresis will be introduced in this section.

Harmonic distortion

Harmonic distortion can also be used to characterize the large signal dynamic behavior of amplifiers. In nonlinear dynamic systems, the relation between harmonic distortion and differential gain, as found in 17.35 for instantaneous nonlinear systems, is not longer valid.

Intermodulation intercept points

In Chapter 17 is has been shown that at low distortion levels, the amplitude of the IM_2 component is proportional to the squared amplitude of the input signal, while the amplitude of the IM_3 component is proportional to the third power of the amplitude of the input signal. Hence, if we plot the output power P_o versus the input power P_i on a double logarithmic scale, we can predict the IM_2 and IM_3 distortion levels at any input signal level from the *intermodulation intercept points*. This is shown in Figure 2.46.

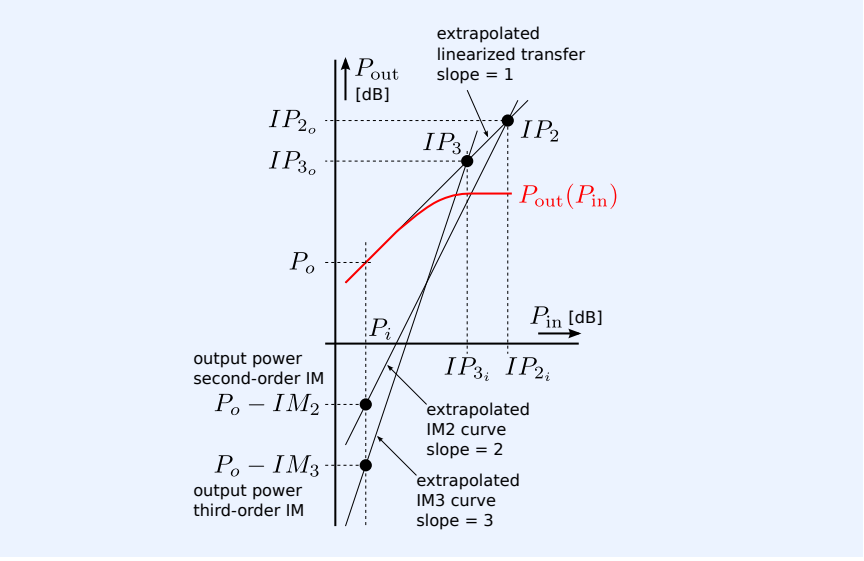


Figure 2.46: Definition of the intermodulation intercept points IP2 and IP3.

The IP_2 (second order IM intercept point) is the intersection point of the extrapolated output level line $P_{out}(P_{in})$ and the IM_2 level line. The IP_3 (third order IM intercept point) is the intersection point of the extrapolated output level line and the IM_3 level line.

The coordinates (IP_{2_i}, IP_{2_o}) and (IP_{3_i}, IP_{3_o}) of these intercept points can be obtained from the second and third-order intermodulation distortion levels:

$$(IP_{2_i},\ IP_{2_o}) = (P_i + |IM_2|\,,\ P_o + |IM_2|)\,, \tag{2.79}$$

$$\left(IP_{3_i},\ IP_{3_o}\right) = \left(P_i + \left|\frac{IM_3}{2}\right|,\ P_o + \left|\frac{IM_3}{2}\right|\right), \tag{2.80}$$

where P_i and P_o are the input power and the output power in dB at one frequency component applied for the intermodulation distortion measurement, respectively. IM_2 is the power of the second order intermodulation distortion component in dB relative to the power of the fundamental, and IM_3 is the power of the third order intermodulation distortion component in dB relative to the power of the fundamental. Input power levels should be at a level where quadratic and qubic extrapolation of the IM_2 and the IM_3 are justified, respectively.

1dB compression point

At high input levels, the gain usually drops due to clipping of the load signal. This *gain compression* is shown in Figure 2.46. The 1 dB compression point is defined for a sinusoidal input signal, as the input level at which the (large-signal) gain drops 1 dB with respect to the small-signal gain. Gain compression is a consequence of odd nonlinearity. For systems with that exhibit a weak nonlinearity[25], the relation between the 1 dB compression input level $X_{-1\text{dB}}$, and the IP_3 input level X_{IP3}, can be determined as:

[25] Weakly nonlinear systems exhibit a smooth transition from linear to clipping.

$$20\log_{10}\left(\frac{X_{-1\text{dB}}}{X_{IP3}}\right) = -9.6\text{ dB} \tag{2.81}$$

Differential gain and differential phase

In dynamic nonlinear systems, the differential gain depends on the frequency. Since the small-signal gain in dynamic systems is a complex quantity, the differential gain is a complex quantity as well, with both a magnitude and a phase (differential gain and differential phase).

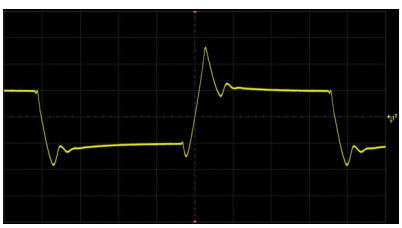


Figure 2.47: Response of an amplifier to a square wave. This response suffers from slew-rate limitation and ringing. The amplifier behaves as a nonlinear dynamic system.

Slew rate limitation

In electronic amplifiers, clipping of amplifier stages may give rise to limitation of the rate of change of the output signsl. This *slew rate limitation* generally arises from two effects:

1. Limitation of the current through capacitors

   Let $I_{\max}$ be the maximum current for charging a linear capacitor with capacitance C. The maximum rate of change of the capacitor voltage is then obtained as

   $$\left.\frac{dV}{dt}\right|_{\max} = \frac{I_{\max}}{C}. \tag{2.82}$$

2. Limitation of the voltage across inductors

   Let $V_{\max}$ be the maximum voltage that can be applied across a linear inductor with inductance L. The maximum rate of change of the inductor current is then obtained as:

   $$\left.\frac{dI}{dt}\right|_{\max} = \frac{V_{\max}}{L}. \tag{2.83}$$

Full-power bandwidth

The full-power bandwidth is defined as the maximum frequency for which the system can deliver its maximum peak to peak sinusoidal output signal without distortion due to slew rate effects. Let y_p be the maximum amplitude

of a sinusoidal signal delivered by a system at a frequency for which it can be considered to be instantaneous. If we increase the frequency, the rate of change of the output signal will increase. The maximum frequency f_{fp} (full power frequency) at which the system with a slew rate SR can deliver a sinusoidal output signal with an amplitude y_p can be determined as

$$f_{fp} = \frac{SR}{2\pi y_p}. \tag{2.84}$$

Hysteresis

Hysteresis is a nonlinear memory effect. A system with hysteresis shows a different behavior for rising and falling slopes of the input signal. Such a behavior can be quasi-static, which means that is does not depend on the rate of change of the signal. In such cases we speak of *rate-independent hysteresis*.

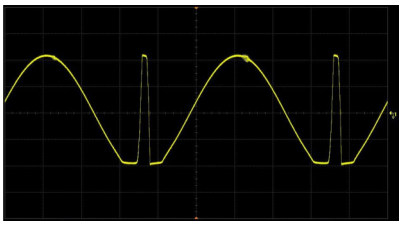


Figure 2.48: Response to a sinusoidal signal of an amplifier that has its input driven outside its specified operating range. This effect is called *phase reversal*.

Overdrive recovery

The signal excursions at the amplifier's output are limited by the power supply voltage. When the maximum excursion is achieved, any increase of an input signal generally causes an increase of energy storage in the amplifier. After the excessive drive signal has been removed, it will generally take some time before the amplifier returns to its normal state. This phenomenon is called *overdrive recovery*. The overdrive recovery time specifies the time the amplifier needs to return to its linear operating range. Figure 2.48 shows so-called *phase-reversal*, an even more harsh effect when an amplifier is driven outside its specified operating range.

2.4.9 Modeling of temperature effects

The characteristics of electronic devices all depend on temperature. As a result, the characteristics of amplifiers will depend on temperature. Hence, all characteristics that have been discussed in this chapter depend on temperature. A change of a performance parameter over time is often referred to as *drift*. Since temperature slowly varies with time, we speak of temperature drift that results in offset drift, bandwidth drift, etc.

Due to power dissipation in the amplifier, temperature drift may become signal-dependent. Gain and offset drift may then be observed differently, for example, as a slow droop or tilt in the step response.

2.4.10 Ageing

The change of amplifier characteristics over time is called ageing. On a time scale, parameter changes due to ageing are usually much slower than parameter changes due to temperature drift.

2.5 Cascaded Amplifiers

The available power gain of amplifiers or amplifier stages can be increased by using cascaded amplifiers or amplifier stages. In this section, we will discuss the behavior of cascade connections of amplifiers. Figure 2.49 shows a cascade connection of two two-ports. The output of the first two-port is connected to the input port of the second.

If we want to design a system that consists of two or more cascaded subsystems, we need error distribution methods for specification of the performance of each subsystem. In this section, we will give expressions for the

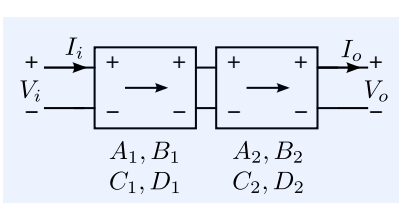


Figure 2.49: A cascade connection of two two-ports.

evaluation of the total error due to the limitations of noise, speed and power. These expressions form the basis for error distribution in designing cascaded systems.

The transmission parameters of a two-port that consists of the cascade connection of two two-ports, as shown in Figure 2.49 are:

$$\begin{pmatrix} V_i \\ I_i \end{pmatrix} = \begin{pmatrix} A_1A_2 + B_1C_2 & A_1B_2 + B_1D_2 \\ A_2C_1 + C_2D_1 & B_2C_1 + D_1D_2 \end{pmatrix} \begin{pmatrix} V_o \\ I_o \end{pmatrix}, \tag{2.85}$$

where A_1, B_1, C_1, D_1 and A_2, B_2, C_2, D_2 are the transmission parameters of the first and the second two-port of the cascade connection, respectively.

2.5.1 Port isolation

In section 2.4.2, we have seen that the non-ideal port isolation of an amplifier is usually described with a few parameters. Under well-specified conditions for source, load and power supply, such simplified descriptions can be meaningful. Figure 2.50 shows two cascaded differential voltage amplifiers that have been connected to one power supply. Under appropriate drive conditions, the port isolation properties of the amplifier chain can be estimated from the port isolation properties of the individual amplifiers. We will give some expressions in the following paragraphs.

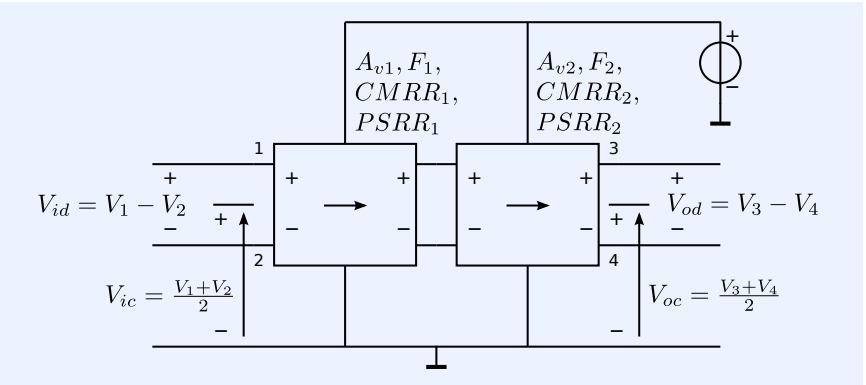


Figure 2.50: Cascaded differential voltage amplifiers.

Rejection Factor

If n amplifiers form a cascaded amplifier chain, the rejection factor of the chain F_{tot} can be obtained from the rejection factors F_i of the amplifiers that constitute the chain:

$$F_{tot} = F_1F_2...F_n. \tag{2.86}$$

Please consider the remarks about the completeness of such a description set down in section 2.4.2.

Common-Mode Rejection Ratio

If n amplifiers form a cascaded amplifier chain, the common-mode rejection ratio of the chain $CMRR_{tot}$ can be obtained from the common-mode rejection ratios $CMRR_i$ and the rejection factors F_i of the amplifiers that constitute the chain:

$$\frac{1}{CMRR_{tot}} = \frac{1}{CMRR_1} + \frac{1}{F_1CMRR_2} + ... + \frac{1}{F_1F_2F_{n-1}CMRR_n}. \tag{2.87}$$

The index number i refers to the amplifier's position in the chain; $i = 1$ refers to the first amplifier and $i = n$ to the last amplifier of the chain.

Please consider the remarks about the completeness of such a description as set down in section 2.4.2.

If the rejection factor of all amplifiers is much larger than unity, and their common-mode rejection ratios are in the same order of magnitude, then the total common-mode rejection ratio approximates that of the first amplifier in the chain.

Power Supply Rejection Ratio

If n amplifiers in a cascaded amplifier chain are connected to the same power supply source, the power supply rejection ratio of the chain $PSRR_{tot}$ can be obtained from the power supply rejection ratios $PSRR_i$ and the gains A_i of the individual amplifiers. The index number i refers to amplifier's position in the chain; $i = 1$ refers to the first amplifier and $i = n$ to the last amplifier of the chain.

We will give the expression for voltage amplifiers (A_{vi} is the voltage gain of the $i - th$ voltage amplifier):

$$\frac{1}{PSRR_{tot}} = \frac{1}{PSRR_1} + \frac{1}{A_{v1}PSRR_2} + ... + \frac{1}{\left(A_{v1}...A_{v(n-1)}\right)PSRR_n}. \quad (2.88)$$

Please consider the remarks about the completeness of such a description as they have been made in section 2.4.2.

If the gain of all amplifiers is much larger than unity, and their power supply rejection ratios are in the same order of magnitude, then the total power supply rejection ratio equals that of the first amplifier in the chain.

2.5.2 Noise behavior

The noise figure of a system that consists of cascaded amplifiers that have a finite nonzero available power gain, can be calculated from the noise figures and the available powers gains of the individual amplifiers. This was shown by Friis (see [Friis1944][26]).

[26] H.T. Friis. Noise Figures of Radio Receivers. *Proceedings of the IRE*, 32:419–422, February 1944

For n cascaded subsystems from which the input of the first system is connected to the signal source and the input of each following subsystem is connected to the output of the previous subsystem, the total noise factor NF_{tot} can be expressed as:

$$NF_{tot} = NF_1 + \frac{NF_2 - 1}{A_{p1}} + ... + \frac{NF_n - 1}{A_{p1}...A_{p(n-1)}}, \quad (2.89)$$

where NF_i [-] is the noise factor of subsystem i, calculated with respect to the output impedance of its driving subsystem, and A_{pi} [-] the available power gain of subsystem i.

Given a unilateral voltage amplifier with voltage gain A_v, an input resistance R_i and an output resistance R_o, and driven from a source impedance R_s, the available power gain can be written as

$$A_p = \left(\frac{R_i}{R_s + R_i}\right)^2 A_v^2 \frac{R_s}{R_o}. \quad (2.90)$$

Expression 2.89 is only useful for situations in which the cascaded amplifiers have a nonzero, finite output resistance. This is often the case in so-called characteristic impedance systems. If this is not the case, the noise of cascaded amplifiers can be evaluated with the aid of the techniques described in section 2.4.3.

2.5.3 Small-signal dynamic behavior

If n amplifiers form a cascaded amplifier chain, the small-signal transfer function $H_{tot}(j\omega)$ can be obtained as the product of the transfer functions $H_i(j\omega)$ of all amplifiers in the chain:

$$H_{tot}(j\omega) = H_1(j\omega)H_2(j\omega)...H_n(j\omega). \tag{2.91}$$

The pole-zero patterns of all constituting amplifiers thus have to be added.

2.5.4 Static nonlinear behavior

The small-signal gain of a chain of n cascaded amplifiers that all operate at their quiescent operating point is the product of the small-signal gains of the individual amplifiers. This simply follows from expression 2.91. If the amplifiers show nonlinear behavior, the gain at an excursion from the quiescent point differs from that at the quiescent operating point. This can be expressed with the aid of the differential gain error. Let A_i be the gain at the quiescent operating point of an amplifier located at position i in the amplifier chain, and let $\epsilon_i(y_i)$ be the differential gain error of that amplifier at an output signal excursion y_i from the quiescent operating point. Then, if the amplifier's nonlinearity is small, the small-signal gain $A_i(y_i)$ at output excursion y_i can be approximated by

$$A_i(y_i) = A_i\left(1 + \epsilon_i(y_i)\right). \tag{2.92}$$

If the signal excursion from the quiescent point at the output of the last amplifier in the chain equals y, the excursion at the output of the $i-th$ amplifier in the chain can be approximated by

$$y_i = \frac{y}{A_{i+1}A_{i+2}...A_n}. \tag{2.93}$$

We then obtain the total differential gain ϵ_{tot} of the amplifier chain as

$$\epsilon_{tot} = \epsilon_1\left(\frac{y}{A_2A_3...A_n}\right) + \epsilon_2\left(\frac{y}{A_3A_4...A_n}\right) + ... \tag{2.94}$$

2.6 Amplifier requirement specification

In this chapter, we have discussed the modeling and characterization of amplifiers. We have found description methods for their ideal behavior and for their non-ideal behavior. Now, we will use this knowledge for the specification of amplifiers. In section 2.6.1, we will discuss the so-called *operational requirements* of application-specific amplifiers. Operational requirements describe the desired functionality, the performance measures, the cost factors and the environmental conditions that apply during the process of operation. This process, however, is not the only process that introduces design constraints. Generally, all *life-cycle processes* (see section 1.2.1) generate various requirements that need to be accounted for during design. As a matter of fact, even the requirements of the design process itself may seriously limit the set of possible solutions. A few remarks on relevant requirements that follow from those life-cycle processes will be made in section 2.6.2.

2.6.1 Operational requirements

A complete list of performance aspects for application-specific amplifiers cannot be given. The definition of relevant performance aspects and limitations,

as well as the definition of available resources and environmental conditions, needs to be extracted from a description of the application of the amplifier to be designed. This task is usually performed by experienced system architects. However, based on the knowledge acquired in this chapter, we can make a list of groups of items that usually need to be described.

As mentioned above, setting up the requirements for a design usually starts with describing the application of the amplifier. This comprises a description of:

1. The signal processing task to be performed by the amplifier (also: fucntional specification)
2. The quality level of the execution of these tasks (also: performance measures)
3. The environment in which the signal processing takes place (also: environmental conditions)
4. The resources that are available for the performance of this task (also: cost factors).

The description of the application needs to be translated into a clear list of measurable properties of the amplifier, including their test methods and test conditions. During the design of the amplifier, the performance measures of various design alternatives with their specific cost factors can be compared. Design choices can then be made on grounds of the performance-to-cost ratio.

Performance requirements

The performance requirements describe the electrical properties of the amplifier that are required for its proper operation in the application, as well as their test methods. The following list of specification items is usually required (but not complete!).

1. Input port requirement specification

   Relevant performance aspects for the input port are:

   (a) Input port configuration (grounded, floating)
   (b) Input impedance
   (c) Input signal specification

   In many cases, these specifications need to be extracted from the source specification:

   (a) Source configuration (grounded, floating)
   (b) Source impedance
   (c) Source signal specification (current or voltage , frequency spectrum, rate of change, peak values, etc.)

2. Output port requirement specification

   Relevant performance aspects for the output port are:

   (a) Output port configuration (grounded, floating)
   (b) Output impedance
   (c) Output signal specification

   In many cases, these specifications must be found from the load specification:

(a) Load configuration (grounded, floating)

(b) Load impedance

(c) Load signal specification (current or voltage, frequency spectrum, rate of change, peak values, ...)

3. Signal transfer specification (type, value and error budgets)

   Relevant performance aspects for the signal transfer are usually derived from the source and the load specification and from the specification of the environmental conditions. Usually, error budgets should be given for:

   (a) Imperfect port isolation (common-mode port impedances, CMRR and PSRR)

   (b) Noise addition

   (c) Small-signal dynamic behavior (frequency range and filter characteristics)

   (d) Static nonlinear behavior (offset, gain error, nonlinearity, voltage and current clipping)

   (e) Nonlinear dynamic behavior (slew rate, overdrive recovery)

Operating conditions

The operating conditions represent the environmental conditions under which the amplifier should perform according to its requirements. All environmental conditions that affect the operation of the amplifier should be specified. Amongst others, the operating conditions listed below are relevant to the functioning of electronics in general.

1. Temperature (affects electrical properties of all electronic devices and thermal noise)
2. Humidity (may cause parasitic current paths between devices)
3. Shock and vibration (may cause defects in connections and in electronic devices)
4. Electro Magnetic Interference (EMI: may cause degradation of signal quality or failure during operation)
5. Power supply noise
6. Electro Static Discharge (ESD: may cause degradation of the performance of devices or failure of devices)

Operating cost factors

The operating cost factors are the resources that are required for the operation of the amplifier. Typical resources are:

1. Power supply voltage(s) and current(s)
2. The amount of space
3. The maximum mass or weight

Reliability requirements

Some examples of reliability requirements are:

1. Mean Time To Failure (MTTF)
2. Mean Time To Repair (MTTR)
3. Mean Time Between Failures (MTBF)

Safety requirements

Safety regulations for electronic products strongly depend on the application domain (consumer, automotive, industrial, medical, space) and on the nationality or region of use (e.g. CE compliance and UL compliance).

2.6.2 Requirements from other life cycle processes

Although the operating process is the most important life-cycle process of the amplifier, other life-cycle processes may introduce serious design constraints. *Life-cycle design* is the name for a design process that accounts for requirements from all life-cycle processes. We will restrict ourselves to a few remarks on certain life-cycle processes.

Design

The availability of design resources such as device models, device samples, design tools and design verification tools may seriously limit the solution space.

Production

Electronic circuits are usually produced in standardized production processes. Such processes introduce design constraints that have to be accounted for. Design For Production (DFP) indicates that production aspects have been accounted for during design.

Test

Test processes and the availability of test tools also introduce design constraints that have to be accounted for. Design For Testability (DFT) indicates that test aspects have been accounted for during design.

Transport

Environmental conditions during transportation, as well as cost factors, may strongly differ from those during operation. This has to be accounted for during design.

Installation

The availability of tools for installation may put specific constraints on the design of an amplifier.

Service

The availability of tools, as well as the service environment, may introduce specific constraints that have to be accounted for during design

3
Amplification Mechanism

3.1 Introduction

In Chapter 2, we discussed the characterization of amplifiers. We concluded that an amplifier should provide its load with an accurate copy of its source signal, while the power available to the load should exceed the available power of the source. In this chapter, we will discuss the principle of amplification, as it is based upon the application of nonlinear, passive devices and power sources. We will discuss under which conditions the available small-signal power gain of a network comprising nonlinear, passive devices and bias sources can exceed unity. To this end, a formal introduction of the concepts *operating point*, *biasing* and *available power gain* will be presented.

3.1.1 Active devices

Electronic devices with *amplifying capabilities*, such as, MOSFETs, Junction FETs, Bipolar Junction Transistors and Vacuum Tubes are often referred to as *active devices*. However, the adjective *active* is somewhat misleading. These devices themselves are passive because they do not provide any electrical power. The term *active* means that when these devices are properly combined with power sources it appears as if they can provide electrical power themselves.[1] From now on, we will also use the term active devices, while bearing in mind the fact that we actually refer to passive devices that require power sources to behave as such.

[1] According to network theory, a network delivers power if the sum of the products of its branch currents and the corresponding branch voltages is negative.

3.1.2 This chapter

In this chapter, we will show the way in which passive devices have to be combined with power sources, such as to obtain amplifying capabilities. In section 3.2 we will do this for two-terminal, passive, resistive elements. We will also give a formal definition of *complementary elements* and of the concepts *operating point*, *biasing* and *available power gain*.

In section 3.3 we will discuss the deployment of the amplifying capabilities of biased, passive, resistive, multi-terminal elements and especially of biased, passive, resistive, two-ports and biased, passive, resistive three-terminal elements.

In section 3.4 we will give a formal approach to the biasing of electronic devices. We will discuss the selection of independent and dependent bias sources and briefly introduce methods to derive those sources from the power supply sources. We will also discuss the application of a biased, passive, three-terminal element as basic amplifier stage and elucidate the amplification mechanism.

3.2 Two-terminal resistive elements

Two-terminal, resistive elements are two-terminal network elements whose behavior can be described with an instantaneous $v - i$ relation. Hence, the shape of their $v - i$ plot does not depend on the rate of change of *the*voltage or the current across the resistive element. Passive, resistive, two-terminal elements have their $v - i$ characteristic pass through the origin: they do not carry any current when shorted.[2] Figure 3.1 shows the schematic symbol of a two-terminal, passive, nonlinear, resistive element with associated signs for the voltage across and the current through it.

[2] Or, alternatively, the voltage across them equals zero if they are left open,

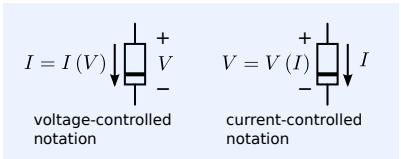


Figure 3.1: Schematic symbol of a two-terminal, non-linear resistor with reference directions for voltage and current.

3.2.1 Voltage-controlled and current-controlled notation

There are two representation methods for the $v-i$ characteristic of these elements: the *voltage-controlled* and the *current-controlled* representation. Two-terminal, resistive elements of which the branch current is uniquely defined by the branch voltage, are called voltage-controlled. Two-terminal, resistive elements of which the branch voltage is uniquely defined by the branch current, are called current-controlled elements.

Monotonously increasing or decreasing $v-i$ relations can be expressed in both voltage-controlled and current-controlled notation.

Examples of voltage-controlled and current-controlled, nonlinear resistors with a non-monotonic relation between voltage and current are shown in Figure 3.2.

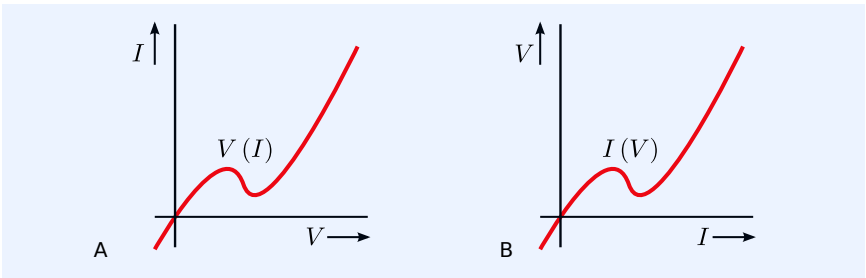


Figure 3.2:
A. $i(v)$ function of a voltage-controlled, nonlinear resistor
B. $v(i)$ function of a current-controlled, nonlinear resistor.

A voltage-controlled representation of a $v-i$ characteristic is shown in Figure 3.2A. A voltage-controlled $v-i$ relation is written in the form

$$I = I(V). \tag{3.1}$$

In words: the branch current is uniquely defined by the branch voltage.

An example of a current-controlled representation of the characteristic of a nonlinear, two-terminal, resistive element is shown in Figure 3.2B. We speak of a current-controlled $v-i$ relation if the branch voltage is uniquely defined by the branch current:

$$V = V(I). \tag{3.2}$$

3.2.2 Resistive two-terminal elements

Some examples of nonlinear, resistive elements and their characteristics are shown in Figure 3.3. The $v-i$ relations of these elements are given table 3.1.

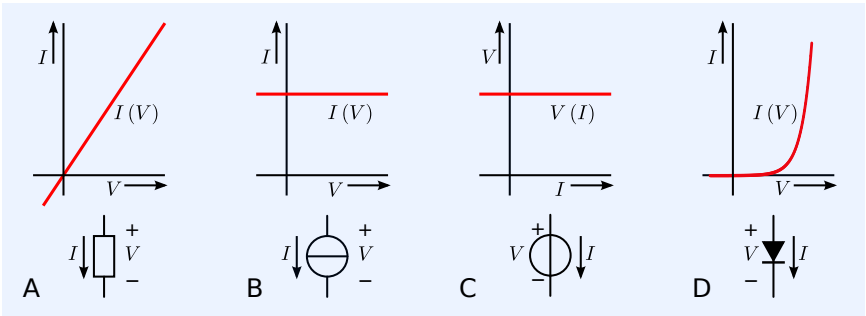


Figure 3.3: Examples of two-terminal, resistive elements:
A. Linear resistor
B. Independent current source
C. Independent voltage source
D. Ideal diode.

Table 3.1: Examples of two-terminal nonlinear resistive elements

element	parameters	current-controlled	voltage controlled
linear resistor	R	$V(I) = I\mathrm{R}$	$I(V) = \frac{V}{\mathrm{R}}$
voltage source	V	$V(I) = \mathrm{V}$	$I(V) =$ undefined
current source	I	$V(I) =$ undefined	$I(V) = \mathrm{I}$
ideal diode	IS	$V(I) = \frac{kT}{q} \ln\left(1 + \frac{I}{\mathrm{IS}}\right)$ valid for $I > -\mathrm{IS}$	$I(V) = \mathrm{IS}\left(\exp\left(\frac{qV}{kT}\right) - 1\right)$

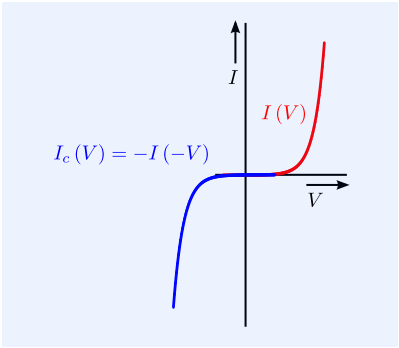


Figure 3.4: $v - i$ Characteristics of a nonlinear, two-terminal resistor and its complementary device.

3.2.3 Complementary devices

Let us consider two current-controlled, nonlinear, resistive elements of which the behavior is described by their respective $v - i$ relations $V(I)$ and $V_c(I)$. These elements are said to be *complementary* if:

$$V_c(I) = -V(-I). \tag{3.3}$$

Similarly, two nonlinear, resistive elements with their respective voltage-controlled notations $I(V)$ and $I_c(V)$ are said to be complementary if:

$$I_c(V) = -I(-V). \tag{3.4}$$

A graphical interpretation of this relation is shown in Figure 3.4. It illustrates that the $v - i$ relation of a complementary, two-terminal element is obtained from is normal version by rotating it over 180 degrees. Note that a two-terminal, resistive element is always complementary to its reversely connected element.

3.2.4 Operating point

The *quiescent operating point*, or shortly the *operating point* of a device is a point on the $v - i$ curve of the device at which it is operating in the absence of a signal. At a later stage, we will show that if the (small-signal) resistance of a two-terminal resistive element in its operating point is negative, the element exhibits amplifying capabilities. As a matter of fact, we will see that the amplifying capabilities of active devices strongly depend on the device's operating point. Selection and design of proper operating conditions of active devices is an important activity of analog circuit design engineers.

Fixing the operating point

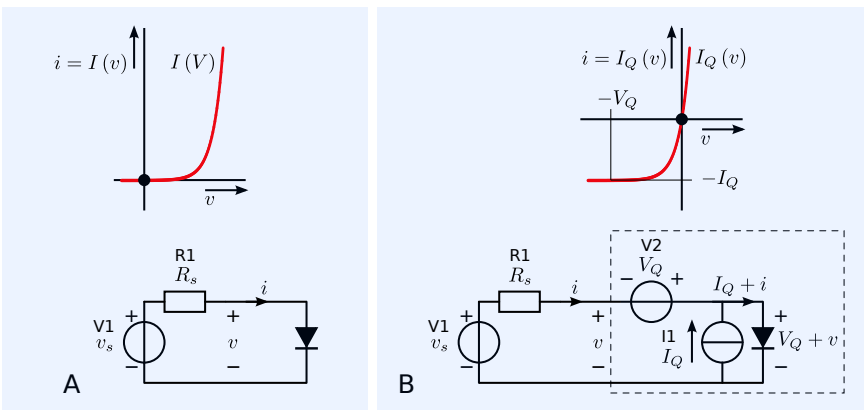


Figure 3.5:
A. $v - i$ characteristic of a PN diode
B. Definition of the operating point and the application of bias sources.
C. $v - i$ characteristic of the biased diode.

The operating point of a two-terminal device can be altered by placing a voltage source in series with the element and a current source in parallel with

it. This process is called *biasing*, and the added voltage and current sources are called *bias sources*.

Figure 3.5 illustrates the process of biasing a diode. Figure 3.5A shows the unbiased diode connected to a signal voltage source `V1` with source resistance R_s. If the signal voltage v equals zero, the diode operates in the origin of its $v-i$ characteristic. Hence, the quiescent operating point is $(0,0)$. If the signal voltage deviates from zero, a signal current i will flow through the source.

Let us now change the operating point of the diode, *while maintaining the quiescent operating conditions of the signal source*. In other words: we still want no current through `V1` if $v=0$.

Figure 3.5B shows the way in which this can be achieved. The operating point of the diode has been changed to (V_Q, I_Q) through the addition of a bias voltage V_Q in series with the diode and a bias current I_Q in parallel with the diode. The bias voltage source `V2` provides the bias voltage V_Q and the bias current source `I1` provides the bias current I_Q. In the quiescent state, the bias current I_Q causes a voltage V_Q across the diode. The bias voltage source `V2` compensates for this voltage, such that the quiescent operating conditions of the signal source have not changed. Please notice that this way of biasing makes the operating point insensitive for variations in R_s. In fact, any change in a passive, resistive termination of the biased diode will not alter its operating point.

Dependent and independent bias sources

Either the bias voltage, or the bias current can be selected by design, while the other follows from the one selected and the $V-I$ characteristic of the device. If the current I_Q is selected by design, the value of the voltage source should equal the voltage across the diode, with the current I_Q flowing through it, hence: $V_Q = V(I_Q)$. If the voltage V_Q is selected by design, the value of the current source equals $I_Q = I(V_Q)$.

Biasing errors

Since the real-world $v-i$ characteristic of an element to be biased is usually not fully known and temperature-dependent, biasing errors will be inevitable and error reduction may be required to improve the biasing accuracy and temperature stability. At a later stage we will present techniques for the reduction of biasing errors.

Biased device $v-i$ characteristic

The $v-i$ characteristic of the biased device can be found from the $V(I)$ function of the unbiased device and the operating point (V_Q, I_Q).

The voltage-to-current transfer of the biased device can be written in terms a modified nonlinear transfer $I_Q(v)$:

$$i = I_Q(v), \tag{3.5}$$

where v and i are the voltage and the current excursion from the operating point, respectively.

The function $I_Q(v)$ can be obtained from the original nonlinear device characteristic $I(V)$ as

$$I_Q(v) = I\left(V_Q + v\right). \tag{3.6}$$

Similarly, but now in current-controlled notation, we may write

$$v = V_Q(i), \tag{3.7}$$

where

$$V_Q(i) = V\left(I_Q + i\right). \tag{3.8}$$

The modified functions $I_Q(v)$ and $V_Q(i)$ now pass through the origin, while the quiescent operating point of the nonlinear device itself is (V_Q, I_Q).

3.2.5 Linearization and available power gain

For small excursions from the operating point the $v-i$ relation can be linearized. The small-signal conductance and resistance in the operating point in voltage-controlled or current-controlled notation, are:

$$i = g_Q v; \qquad g_Q = \left.\frac{I_Q(v)}{dv}\right|_{v=0}, \tag{3.9}$$

$$v = r_Q i; \qquad r_Q = \left.\frac{V_Q(i)}{di}\right|_{i=0}, \tag{3.10}$$

respectively.

We will now study the conditions under which nonlinear resistive two-terminal elements can provide an available power gain larger than unity. As discussed earlier, an available power gain that exceeds unity is a distinguishing property of amplifiers.

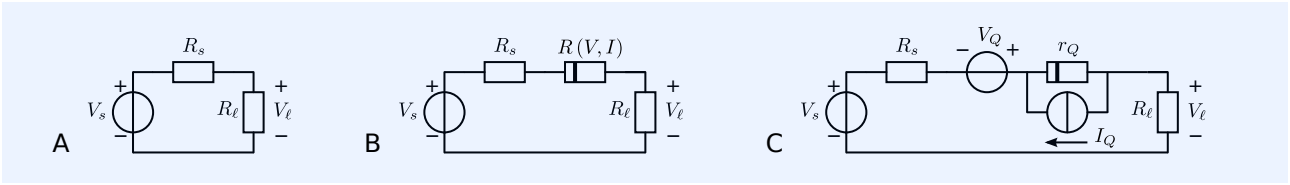


Figure 3.6:
A. Resistive load connected to a voltage source with source resistance R_s.
B. A nonlinear resistor is placed between the source and the load.
C. The nonlinear resistor is biased in an operating point (v_Q, i_Q). In this operating point its small-signal equivalent resistance equals r_Q.

Let us hereto consider a signal source that consists of a voltage source with a nonzero source resistance, as shown in Figure 3.6A. The source voltage equals V_s and the source resistance equals R_s. The source is connected to a load that has a resistance R_ℓ.

The available power of the source equals

$$P_{\text{av}} = \frac{V^2}{4R_s}. \tag{3.11}$$

Let us now place a biased two-terminal nonlinear device $R(v,i)$ between the source and the load as shown in Figure 3.6B. The device is biased in an operating point Q and its small-signal equivalent resistance in this operating point is assumed r_Q. The maximum power will be delivered to a load with a resistance $R = R_s + r_Q$. The available power P'_{av} of the series connection of the source and the biased nonlinear device equals

$$P'_{\text{av}} = \frac{V^2}{4(R_s + r_Q)}. \tag{3.12}$$

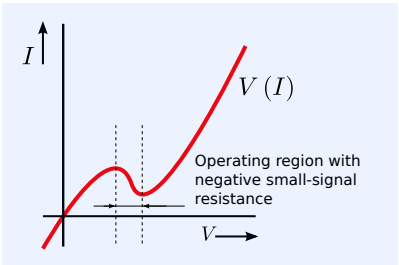


Figure 3.7: Typical $v-i$ characteristic of a tunnel diode.

The available power gain G_P of the two-terminal device R is obtained as the ratio of P'_{av} and P_{av}:

$$G_P = \frac{P'_{\text{av}}}{P_{\text{av}}} = \frac{R_s}{R_s + r_Q}. \tag{3.13}$$

It appears that the available power gain will exceed unity if the small-signal resistance r_Q of the two-terminal device in the operating point Q has a negative value. The tunnel diode is an example of a two-terminal device that has a negative small-signal resistance in a certain operating region. The $v-i$ characteristic of a tunnel diode is shown in Figure 3.7.

3.3 Multi-terminal resistive elements

Most active electronic devices are three-terminal elements. The behavior of a multi-terminal resistive element with n terminals can be described by $n-1$ relations that have all $n-1$ branch voltages and $n-1$ branch currents as arguments. The general form of such a set of equations is

$$f_{1...n-1}\left(V_1...V_{n-1}, I_1...I_{n-1}\right) = 0. \tag{3.14}$$

3.3.1 Complementary multi-terminal elements

Let us consider two n-terminal resistive elements. One is described by a set of multi-variate functions $f_{1...n-1}\left(V_1...V_{n-1}, I_1...I_{n-1}\right)$ and the other by $g_{1...n-1}\left(V_1...V_{n-1}, I_1...I_{n-1}\right)$. These two n-terminal elements are complementary if

$$f_{1...n}\left(V_1...V_{n-1}, I_1...I_{n-1}\right) = -g_{1...n-1}\left(-V_1... - V_{n-1}, -I_1... - I_{n-1}\right). \tag{3.15}$$

NMOS and PMOS transistors are complementary if the above is true. Similar holds for NPN and PNP bipolar transistors.

3.3.2 Resistive two-ports

Nonlinear resistive two-ports can be represented by a pair of nonlinear functions:

$$f_{1,2}\left(V_i, V_o, I_i, I_o\right) = 0. \tag{3.16}$$

The arguments of the functions are the port voltages V_i and V_o and the port currents I_i and I_o. By selecting two independent variables and two dependent variables out of the four variables, we obtain six different representation methods. These representation methods have been listed in table 3.2.

Voltage-controlled representation:	$I_i = I_i\left(V_i, V_o\right)$	$I_o = I_o\left(V_i, V_o\right)$
Current-controlled representation:	$V_i = V_i\left(I_i, I_o\right)$	$V_o = V_o\left(I_i, I_o\right)$
Hybrid 1 representation:	$I_i = I_i\left(V_i, I_o\right)$	$V_o = V_o\left(V_i, I_o\right)$
Hybrid 2 representation:	$V_i = V_i\left(I_i, V_o\right)$	$I_o = I_o\left(I_i, V_o\right)$
Transmission 1:	$I_o = I_o\left(V_i, I_i\right)$	$V_o = V_o\left(V_i, I_i\right)$
Transmission 2:	$V_i = V_i\left(v_o, I_o\right)$	$I_i = I_i\left(V_o, I_o\right)$

Table 3.2: Six representation methods for nonlinear resistive two-ports

3.3.3 Complementary two-ports

Let us consider two resistive two-ports that are described by their respective multi-variate functions $f_{1,2}\left(V_i, V_o, I_i, I_o\right)$ and $g_{1,2}\left(V_i, V_o, I_i, I_o\right)$. These two-ports are complementary if

$$f_{1,2}\left(V_i, V_o, I_i, I_o\right) = -g_{1,2}\left(-V_i, -V_o, -I_i, -I_o\right). \tag{3.17}$$

3.3.4 Operating point

Nonlinear resistive two-ports may, under specific operating conditions, exhibit amplifying capabilities. Similar as with two-terminal devices, such operating conditions may be established by adding bias sources. Biasing of a port requires the insertion of a voltage sources in series with the port and a current sources in parallel with the port. This is illustrated in Figure 3.8. Because the output port quantities depend on the input port quantities, two of those bias sources may be selected by design, while the remaining two follow from the ones selected and the device equations. The two selected by design will be referred to as independent bias sources and the two remaining as dependent bias sources.

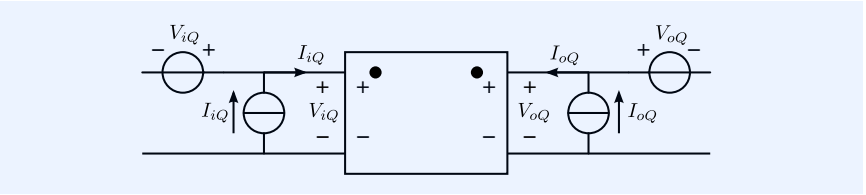


Figure 3.8: Biased non-linear two-port.

Let us consider a nonlinear two-port described by two voltage controlled relations:

$$I_i = I_i\left(V_i, V_o\right), \tag{3.18}$$
$$I_o = I_o\left(V_i, V_o\right). \tag{3.19}$$

After biasing the input port in the operating point (V_{iQ}, I_{iQ}) and the output port in the operating point (V_{oQ}, I_{oQ}), we obtain the modified two-port equations for deviations from the operating point as

$$i_i = I_{iQ}\left(v_i, v_o\right), \tag{3.20}$$
$$i_o = I_{oQ}\left(v_i, v_o\right), \tag{3.21}$$

where v_i, i_i and v_o, i_o represent the voltage and current excursions from the quiescent operating point of the input port and the output port, respectively. The modified functions $I_{iQ}\left(v_i, v_o\right)$ and $I_{oQ}\left(v_i, v_o\right)$ can be obtained from the original two-port functions $I_i\left(V_i, V_o\right)$ and $I_o\left(V_i, V_o\right)$, and the operating points (V_{iQ}, I_{iQ}) and (V_{oQ}, I_{oQ}) of the input port and the output port, respectively:

$$I_{iQ}\left(v_i, v_o\right) = I_i\left(V_{iQ} + v_i, V_{oQ} + v_o\right), \tag{3.22}$$
$$I_{oQ}\left(v_i, v_o\right) = I_o\left(V_{iQ} + v_i, V_{oQ} + v_o\right). \tag{3.23}$$

Such description methods can also be given for other two-port representation methods.

3.3.5 Linearization and available power gain

For small excursions from the operating point a linearized two-port model can be used. In amplifier design, the anti-causal transmission-1 matrix representation will often be used (see Chapter 2 and Chapter 2).

The transmission-1 matrix parameters can be obtained from $v - i$ relations of the biased device as

$$A = \left.\frac{\partial V_{iQ}\left(v_o, i_o\right)}{\partial v_o}\right|_{i_o=0}, \tag{3.24}$$
$$B = -\left.\frac{\partial V_{iQ}\left(v_o, i_o\right)}{\partial i_o}\right|_{v_o=0}, \tag{3.25}$$
$$C = \left.\frac{\partial I_{iQ}\left(v_o, i_o\right)}{\partial v_o}\right|_{i_o=0}, \tag{3.26}$$
$$D = -\left.\frac{\partial I_{iQ}\left(v_o, i_o\right)}{\partial i_o}\right|_{v_o=0}. \tag{3.27}$$

Because the direction of the positive output current of the two-port differ from the positive current direction in the elements, minus signs appear in the expressions for B and D.

The available power gain of a two-port can be expressed in the transmission-1 matrix parameters and the source impedance; see expressions (2.27), (2.29)

and (2.30). A biased, nonlinear, multi-terminal device can have an available power gain that exceeds unity, a property that will be exploited in amplifiers.

3.4 Introduction to biasing

In the previous sections we have discussed under which conditions a nonlinear resistive element can be used as an amplifying device. First of all, there needs to be an operating point at which the available power of a signal source can be increased. Secondly, we must apply bias voltage and current sources that let the device operate at such an operating point. In order not to adversely affect the signal processing, we may place only ideal voltage sources in series with the signal path and ideal current sources in parallel with the signal path.

However, the bias sources to be applied cannot be chosen freely. For a two-terminal element we may assign an arbitrary value to either the voltage source or the current source. The other one relates to the selected one through the element's $v - i$ relation. In general, in a resistive network with n terminals we may define $n - 1$ bias sources freely while the values of another $n - 1$ bias sources relate to ones selected.

In practice, we will also have to face numerous implementation difficulties. Even if we carefully select the desired operating point and design the required bias sources, the device tolerances as well as temperature dependency of the device characteristics cause temperature-dependent biasing errors that may be too large to maintain the proper operating conditions.

Another practical limitation is imposed by the requirement of a single power supply source. All bias voltage and current sources should in some way be derived from a single source or a limited number of power sources. If we have multiple sources at our disposal, these sources usually share a common terminal. This will impose serious limitations to the implementation of floating bias sources.

In this section we will only give a brief introduction to biasing and not addresses all these topics. A more detailed treatment of biasing will be given in Chapter 15.

3.4.1 Independent and dependent bias sources

In the introduction we have already stated that not all operating point variables can be chosen independently. The first question we need to answer is which one(s) do we want to fix by design and which one(s) as dependent variable(s). To this end, let us consider a biased two-port with amplifying capabilities. Such a two-port will exhibit a large available power gain if its transmission-1 small-signal parameters in the operating point are close to zero. This can be seen from (2.27). In such cases, the voltage and current excursions at the output port exceed those at the input port. If we take the input bias quantities as independent and fixed, then the output bias current and bias voltage will depend on the device characteristics and on the temperature. In order to prevent from clipping during large signal excursions at the output port, we may need to apply much larger bias sources than strictly required on grounds of the signal handling requirements. This is because we have to account for device tolerances and temperature changes. Hence, it will be clear that selecting the output port bias quantities on grounds of the signal handling requirements, while adapting the values of the input bias sources to device tolerances and temperature changes, will result in a better power efficiency. Moreover, in later chapters we will see that many performance aspects, such as the noise behavior, the bandwidth and the linearity of biased amplifying devices show a direct relation with the operating voltage

and current of the output port.

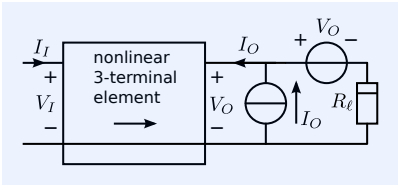


Figure 3.9: A nonlinear two-port with its output port biased at (V_O, I_O). In order to achieve zero load voltage and current at a given temperature, the input port needs to be biased at (V_I, I_I).

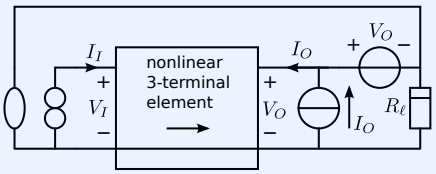
3.4.2 Biasing of 3-terminal elements

Amplifying devices such as MOSFETs, JFETs and BJTs can be used as 3-terminal active devices. Operation and modeling of these devices will be discussed in Chapter 4. Here, we will discuss generalized biasing methods for these devices.

Let us assume that a passive nonlinear resistive 3-terminal element can deliver a sufficiently large load signal if its output port is biased at a voltage of V_O and at a current I_O. Let us also assume that, at a given temperature, the biasing of its output port requires a bias voltage V_I and a bias current I_I at its input port. This situation is shown in Figure 3.9.

The question is how to obtain the required values of the input port bias quantities V_I and I_I. One way is to use the device equations and obtain analytical expressions for the input port quantities. With strongly simplified device equations this may yield useful estimations for taking early stage design decisions.

Figure 3.10: Determination of the input port biasing quantities (V_I, I_I) with the aid of a nullor.



If we want to create an accurately biased two-port that can be used during simulation, we can use the approach sketched in Figure 3.10. In this configuration a nullator sets the zero load condition and the norator provides (V_I, I_I) that satisfies this condition. Such a setup works if this network has a unique DC solution at the temperature of interest. This will usually be the case with amplifier stages. If not, (V_I, I_I) should be limited to a range in which a unique solution exists.

Figure 3.11: Biased 3-terminal element.

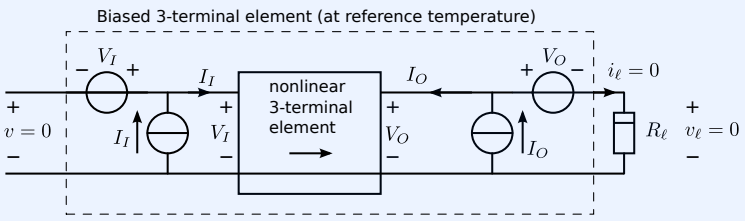


Figure 3.11 shows the biasing result. A bias current source and a bias voltage source are connected to the input port to create zero load conditions for all passive, DC port terminations, while having its output port biased at (V_O, I_{IO}).

3.4.3 Amplification mechanism

Figure 3.12A shows an amplifier stage with a biased 3-terminal element, driven from a voltage source with a finite, nonzero source resistance. The stage is biased as described in section 3.4.2. We assume positive nonzero values for all the bias sources. In Chapter 4 we will see that the commonly used

(N-type) amplifying devices require such biasing conditions. The amplification mechanism is explained in the caption of this figure.

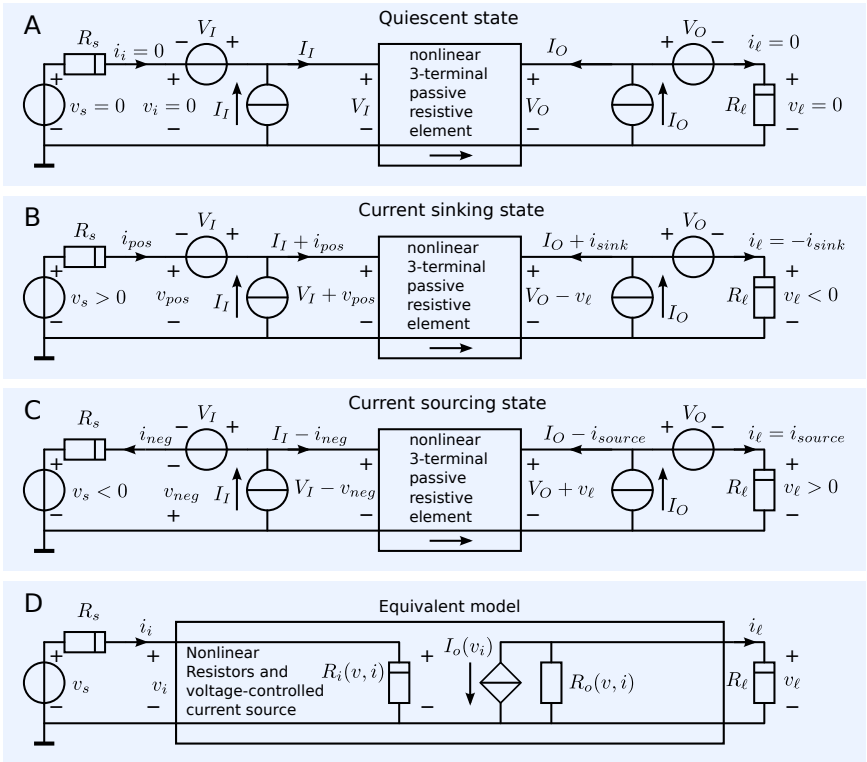


Figure 3.12: Voltage-driven amplifier stage with the three terminal, passive, resistive element, which is biased according to Figure 3.11.

A: Biased 3-terminal element in the quiescent state. For proper operation of this device, all bias voltages and currents are positive (NMOS, NPN, vavuum tube).

B: Sink state: the stage is driven from a positive voltage. The output current of the 3-terminal device increases. The extra (sink) current is delivered by the output bias voltage source. The output voltage of the 3-terminal device drops below V_O and the load voltage v_ℓ drops below zero.

C: Source state: the stage is driven from a negative voltage. The output current of the 3-terminal device decreases. The excess current delivered by the output bias current source flows through the load. The output voltage of the 3-terminal device rises above V_O and the load voltage v_ℓ becomes positive.

D: Compact model of the biased 3-terminal element. This model comprises an active element: a voltage-controlled voltage source. For the sake of simplicity, a possible reverse transfer is not modeled. Under the conditions that no breakdown or saturation effects are present, this stage operates in Class A.

This Figure 3.12D shows that a combination of a passive, resistive, three-terminal amplifying device and properly configured bias sources, can be modeled with the aid of a controlled source, which is an active network element since it delivers power to its load.

3.4.4 Deriving the bias sources from the power supply

The biasing of a 3-terminal element with four bias sources is conceptually correct, but not very practical. Generally we want the amplifier to be supplied from a few power sources that share the reference terminal. This means that the bias voltages and currents in some way have to be derived from the power supply source(s). The general procedure for this is first to minimize the number of bias sources and then derive them from the power supply with passive, (nonlinear) resistive elements. Such elements should exhibit either a voltage source character when operating at a non-zero current, or a current source character when operating at a non-zero voltage.[3] For bias voltage sources we will use elements with a voltage-source character and for bias current sources we will use elements with a current source character.

Figure 3.13 shows an example of the $V-I$ characteristic of a nonlinear resistive two-terminal element that exhibits a current source character when biased at a voltage between saturation and breakdown: $V_{sat} < V < V_{br}$. Outside this range we speak of voltage limiting or clipping.

Chapter 15 is devoted to this topic. There, we will also discuss biasing techniques that use AC coupling. The concept of AC coupling will be introduced in Chapter 9. A chapter about the design of the bias sources themselves has not yet been included.

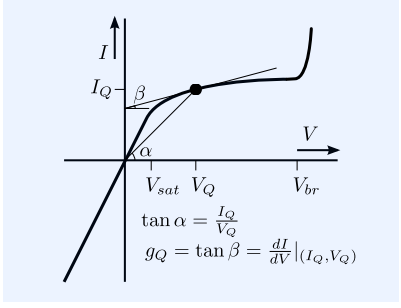


Figure 3.13: $V-I$ characteristic of a two-terminal resistive elements that exhibits a current source character when biased at a sufficiently large voltage.

[3] A two-terminal resistive element exhibits a voltage source character if, at an operating point (V_Q, I_Q), its small-signal resistance is smaller than $\frac{V_Q}{I_Q}$.

A two-terminal resistive element exhibits a current source character if, at an operating point (V_Q, I_Q), its small-signal conductance is smaller than $\frac{I_Q}{V_Q}$.

3.5 Conclusions

In this chapter, we have discussed in which way the amplification function can be implemented with the aid of electronic devices and power sources.

We have seen that a two-terminal, nonlinear, passive device, in conjunction with power sources can provide an available gain larger than unity, if such a device is biased in an operating point in which its small-signal resistance is negative. Biasing of such a device is performed by placing a current source in parallel with the device and a voltage source in series with the device, in such a way that the operating point of the device does not depend on the $v - i$ characteristic of a passive, DC termination. Either the bias voltage or the bias current can be selected freely, while the other one follows from the one selected and from the device's $v - i$ characteristic and the temperature.

The implementation of amplification with three-terminal devices proceeds similarly. Three-terminal active devices such as MOS transistors, BJTs or vacuum tubes can be regarded as nonlinear two-ports of which the input port and the output port share one terminal. When properly biased, these devices can also provide an available power gain that exceeds unity. To this end, we need to select a proper operating point and bias each port with a current source and a voltage source in such a way that the biasing is independent of the passive DC termination resistance at both ports. Two of the four bias sources can be designed freely, while the other two follow from the device characteristics. As we will see later, important performance aspects, such as the noise performance, the dynamic behavior and the nonlinearity show a direct relation with the biasing quantities at the output port. For this reason, the bias current and the bias voltage of the output port are usually selected by design, while the bias voltage and current of the input port have to be adjusted such that the device's operating point does not depend on the temperature and on the passive DC termination resistance at both ports.

3.5.1 Generalized biased active device

In the following chapters, we will design amplifiers using biased active devices. Since the design theory itself is technology-independent, we will use one symbol for a generalized biased device. This symbol is shown in Figure 3.14A. It refers to any type of biased active device, as indicated in Figure 3.14B.

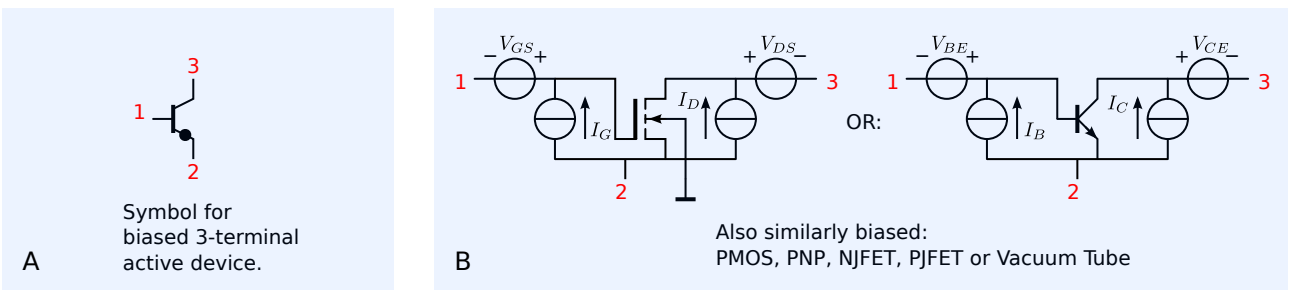


Figure 3.14: Symbol for a generalized biased 3-terminal active device.

4
Active Devices

4.1 Introduction

Since the invention of the vacuum tube, the design of electronic circuits has indissolubly been connected with the development of electronic devices. In fact, the design of electronic circuits cannot be performed without knowledge of the physical operation of the devices from which they will be constructed. However, the *approach* to the design of application-specific amplifiers, as presented in this book, does not depend on the technology in which the circuits finally will be realized. The design approach, the concepts and the techniques presented can be used for designing vacuum tube amplifiers, amplifiers realized with discrete components or amplifiers realized in modern CMOS IC processes. Different technologies, simply introduce different design constraints and offer different possibilities for the implementation of the concepts. The designer's task is to deal with the limitations and maximally exploit the specific implementation possibilities of the selected technology. For this reason, knowledge of the operation and modeling of electronic devices is regarded indispensable for designers of electronic circuits. However, physics and modeling of electronic devices has become a field of knowledge on its own. A full treatment would require a bookshelf full of material and is considered outside the scope of this work. The reader is assumed to have basic knowledge of the operation and modeling of discrete semiconductor devices, of passive devices and of CMOS and BiCMOS IC technology.

During circuit design, the designer must know in which way the performance aspects of the circuit are related to those of the electronic devices and how they can be affected by design. In general, the performance aspects of electronic devices depend on the device type and on its operating conditions. In IC technology, device scaling may also be used for performance optimization. Hence, our aim is to briefly summarize topics related to device selection, device scaling and selection of an operating point. This will be done for so-called active semiconductor devices only, hence for semiconductor devices that, when applied in conjunction with power sources, exhibit amplifying capabilities. We will confine ourselves to Bipolar Junction Transistors (BJTs), Junction Field Effect Transistors (JFETs) and Metal Oxide Semiconductor Field Effect Transistors (MOSFETs). We will describe their basic operation and device models.

In this book, we will advocate the use of different device models at different stages of the design process. At an early stage, we usually want to use relatively simple models that describe performance aspects of interest accurate enough to motivate early stage design decisions. Based on the analysis of relevant performance aspects, we will be able to select the device and its operating region. Parameters for these simplified models can be found from simulation with more complete models, or by estimating them from the parameters of those models or graphs. At a later stage, numeric simulation with more elaborate models is required to give an accurate prediction of the circuit operation. Nowadays, automated circuit optimization is used to obtain the best possible performance with the lowest cost factors and sensitivity for production process variations. Automated circuit optimization, however, is outside the scope of this book.

4.1.1 Design equations and symbolic circuit analysis

Setting up design equations for specific performance aspects, such as, noise behavior, small-signal dynamic behavior and temperature stability requires symbolic circuit analysis. SLiCAP (**S**ymbolic **L**inear **C**ircuit **A**nalysis **P**rogram) is an open-source symbolic simulator that can be used for this purpose.[1] It uses linearized (small-signal) device models of which the parameters can be related to the technology, the geometry and the operating point of a device.

[1] Based on Python and Maxima Computer Algebra Systems.

SLiCAP will be used throughout this text book.

4.1.2 Numeric circuit analysis

The development of device models and the development of circuit simulators is an ongoing process that helps to increase the predictability of the behavior of electronic circuits. The first version of Berkeley SPICE (*Simulation Program with Integrated Circuit Emphasis*) was released in 1973. The program was written in Fortran by Nagel (see [Nagel-M382][2] and [Nagel-M520][3]). Within a few years, SPICE2 was released and widely accepted by circuit manufacturers. SPICE3 was written in C by Quarles (see [Quarles-M89/42][4]), it was released in 1989. Nowadays many SPICE-like simulators have been integrated in CAD software packages.

Most SPICE-like simulators can perform DC, AC + Noise and Transient analysis. The DC analysis uses nonlinear instantaneous device descriptions. During DC analysis all dynamic effects are discarded. The AC analysis uses linearized dynamic device descriptions. The operating point for the linearization is found from a DC analysis that always precedes the AC analysis. The AC analysis is in fact a *small-signal* frequency-domain analysis.

The influence of stationary noise sources can be evaluated with a small-signal noise analysis.

The transient analysis performs a time-domain analysis thereby using nonlinear dynamic device descriptions.

Some SPICE-like simulators are also capable of performing time domain noise analysis.

In this chapter, we will briefly describe basic models for SPICE3X versions. For a more detailed study of device modeling for SPICE, the reader is referred to literature: [SPICE1988][5]. From these SPICE models, we will derive simplified descriptions that can be used for hand calculations for estimating dominant behavioral properties of the device. Some SPICE versions as well as other simulators support more elaborate models.

Most simulators also provide operating point information that can be used for the analysis of linearized circuits. If not, design engineers have to use IC design manuals or data sheets and extract these parameters from measured device characteristics. Alternatively, they can be determined from simulations with more complex models.

[2] Laurence W. Nagel and D.O. Pederson. Technical Report UCB/ERL M382, EECS Department, University of California, Berkeley, April 1973

[3] Laurence W. Nagel. *SPICE2: A Computer Program to Simulate Semiconductor Circuits*. PhD thesis, EECS Department, University of California, Berkeley, 1975

[4] Thomas L. Quarles. *Analysis of Performance and Convergence Issues for Circuit Simulation*. PhD thesis, EECS Department, University of California, Berkeley, 1989

[5] P. Antognetti and G. Massobrio. *Semiconductor Device Modeling with SPICE*. McGraw-Hill, 1988

4.1.3 Simulation accuracy

The validity of the simulation results is limited by the level of detailing in the device models, the accuracy of the model parameters and the numerical routines that have been implemented in the simulator. Both the numerical routines and the device models are continuously improved. The validity of the model parameters, however, is not always obvious; particularly when using discrete components. Many CAD packages include device libraries with parameter values that are represented by more than three digits. However, the suggested accuracy may be far beyond reality. On the one hand, physical reproducibility does not justify the suggested accuracy, while on the other hand the characteristics obtained from simulation with typical parameter values may not correspond to those obtained from typical measurement data. Parameter extraction is a critical process and the designer should be aware of the validity of the models and the parameters. In case of any doubts the designer should check whether results obtained with measurements given in data sheets, can be reproduced with simulations.

4.1.4 This chapter

Basic models for the BJT, the JFET and the MOSFET will be discussed in sections 4.2, 4.3 and 4.4, respectively. For all these devices will start with a brief description of the available device types, their symbols and their simplified physical structure. We will then focus on the device models that are implemented in SPICE and derive simplified models for hand calculations.

Section 4.5 is devoted to the implementation of these device models in SLiCAP.

4.2 Bipolar transistors

In December 1947, John Bardeen, Walter Brattain and William Shockley discovered the transistor effect and developed the first device. The patent [Shockley1952][6] was issued on September 25, 1951. For this invention, the inventors obtained the Nobel Prize in Physics in 1956.

[6] William Shockley. Circuit element utilizing semiconductive material, June 1948

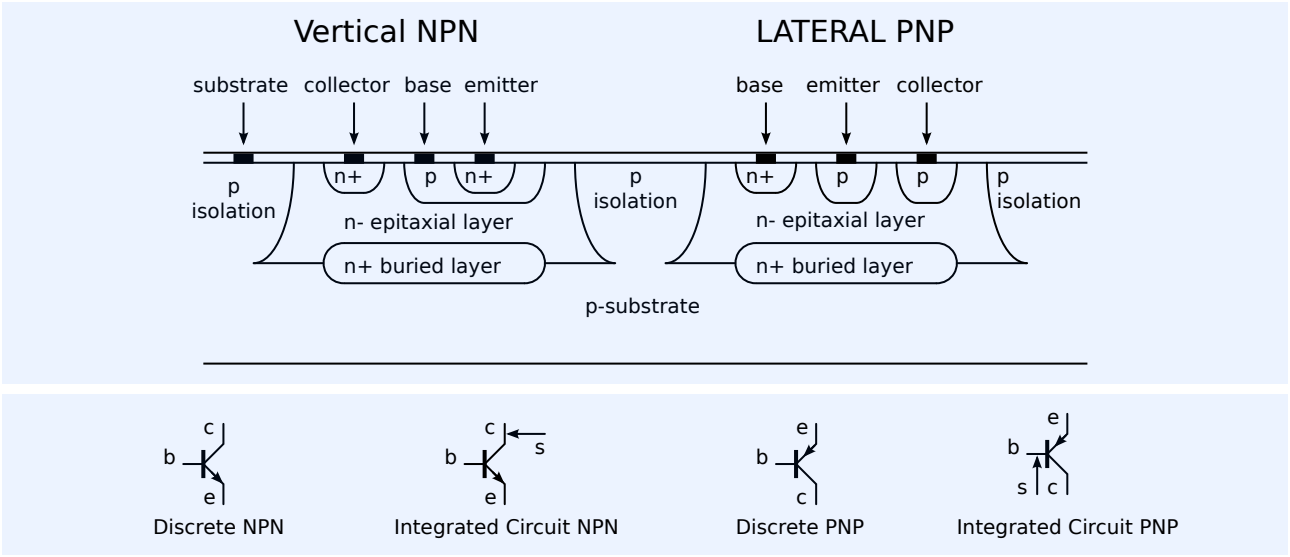


Figure 4.1: Cross-section of an integrated circuit vertical NPN transistor and a lateral PNP transistor and their schematic symbols.

Bipolar junction transistors (BJTs) are available as discrete components or as integrated circuit devices. Figure 4.1 shows the cross sections of a vertical NPN transistor and a lateral PNP transistor, realized in a simple junction-isolated bipolar integrated circuit process. Both vertical NPN and PNP bipolar transistors are available in so-called *complementary bipolar* IC processes. In IC technology, the transistors are realized in separate islands, isolated either by depleted PN junctions or by SiO_2. Figure 4.1 also shows the symbols that are used for both transistors with and without substrate connection.

4.2.1 Operation

In the so-called forward active region, the emitter base junction is forward biased and the base collector junction is reverse biased. In an NPN transistor electrons are then injected from the emitter into the base.[7] Due to this injection the concentration in of electrons in the base at the emitter-base junction increases. This causes a *diffusion current* of electrons in the base towards the collector. At the base-collector junction the concentration of electrons is zero because the drift field in the reverse biased base-collector junction accelerates

[7] Minorities in the base.

them towards the collector terminal. In modern transistors, the recombination of the minorities (electrons) in the base is very low and almost all electrons injected by the emitter will reach the collector terminal. A small reverse injection of holes from the base into the emitter causes a so-called *ideal base current*. This current can be kept low by keeping the doping level of the base much lower than that of the emitter.

The width of the base-collector depletion layer depends on the collector-to-base voltage. Since the base doping level usually exceeds the collector doping level, the depletion layer extends mostly in the collector region. However, an increase the reverse collector-to-base voltage causes a small decrease of the base width, thereby increasing the collector current. The resulting collector-to-base voltage dependency of the collector current is called the *Early effect*, named after its discoverer James M. Early[8] (see [Early1952][9]).

When the excess minority density in the base exceeds the doping level, we speak of *high injection*. At high injection levels the majority carrier concentration increases with the excess minority concentration, resulting in a reduction of the emitter efficiency by a factor two.

[8] James M. Early (July 25, 1922 - January 12, 2004) was an American engineer, best known for his work on transistors and charge-coupled device imagers. He is also known as Jim Early. Early was the first to make a transistor that would oscillate faster than "a thousand megacycles" (1 GHz), circa 1952, for which feat he won a bottle of Scotch whisky from John Robinson Pierce. He also developed the transistors for America's first commercial communications satellite, the Telstar I.[2] In the early 1970s, Early led research for Fairchild Semiconductor, where he invented the vertical anti-blooming drain for CCD image sensors.

[9] J.M. Early. Effects of space-charge layer widening in junction transistors. *Proceedings IRE*, 40:1401–1406, 1952

4.2.2 Gummel-Poon model

The Gummel-Poon model (see [GummelPoon1970][10]) is a charge-control model of the bipolar transistor that is used in SPICE-like simulation programs. It describes the behavior of bipolar transistors outside the breakdown or strong saturation regions. The operation of parasitic substrate transistors is also not modeled. Since, at high frequencies the substrate current can generally not be neglected, a voltage-dependent junction capacitance is added to the three-terminal transistor model. Four-terminal NPN transistors that are fabricated in standard bipolar processes, have their substrate capacitance connected to the collector. SPICE also has a model for lateral PNP transistors. These transistors have their substrate capacitance connected to the base.

In this section, we will give a short description of this charge-control model and its implementation in SPICE. The Gummel-Poon model for four-terminal NPN transistors is shown in Figure 4.2.

[10] H. K. Gummel and H. C. Poon. An integral charge control model of bipolar transistors. *Bell Syst. Tech. J.*, 49(5):827–852, May-June 1970

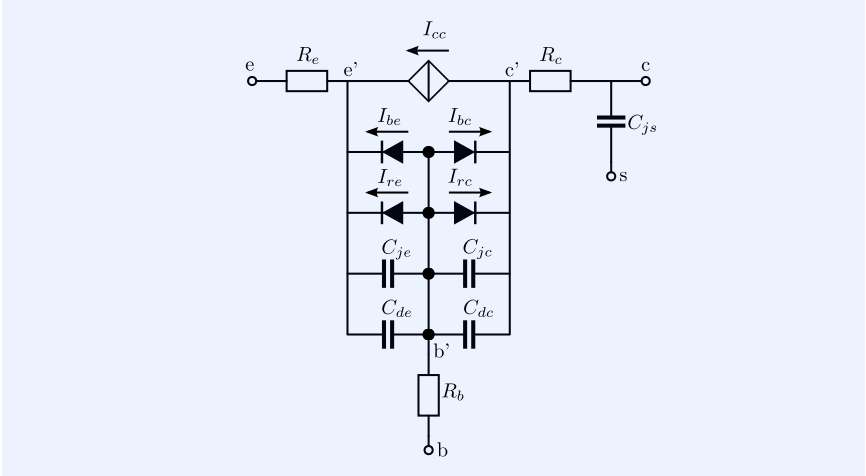


Figure 4.2: SPICE Gummel-Poon model of the four-terminal vertical BJT.

The charge based Gummel-Poon model is an enhancement of the older Ebers-Moll model that describes the transistor as two merged PN diodes (see [EbersMoll1954][11]).

[11] J.J. Ebers and J.L Moll. Large-signal behavior of junction transistors. *Proceedings of the IRE*, 42(12):1761–1772, 1954

Static (DC) operation

For static operation, the transistor is modeled as a nonlinear resistive network that consist of the intrinsic transistor[12] with added bulk resistances R_e, R_b and R_c. The intrinsic transistor is modeled with a number of nonlinear voltage-controlled elements. In the expressions, the Gummel-Poon model parameters are written in SMALL CAPITALS. An overview of all the model parameters is given in section 4.2.3.

[12] The intrinsic transistor is the transistor without the bulk resistors R_e, R_b and R_c. its connections to the external circuit are e', b'and c'.

1. The ideal base currents consists of injected majorities from the base into the emitter, or from the base into the collector (in reverse operation). They are described by I_{be} and I_{bc}:

$$I_{be} = \frac{I_s}{\beta_F}\left(\exp\left(\frac{V_{b'e'}}{\text{NF}U_T}\right) - 1\right), \tag{4.1}$$

$$I_{bc} = \frac{I_s}{\beta_R}\left(\exp\left(\frac{V_{b'c'}}{\text{NR}U_T}\right) - 1\right), \tag{4.2}$$

where c', b' and e' are the collector, base and emitter of the intrinsic transistor, respectively. The thermal voltage U_T equals $\frac{kT}{q}$, in which k represents the Boltzmann constant $\left(1.381 \times 10^{-23}\text{J/K}\right)$ and q the elementary charge $\left(1.602 \times 10^{-19}\,\text{C}\right)$. At room temperature ($T = 293$K), U_T equals 25.26mV.

The forward current gain β_F and the reverse current gain β_R both depend on temperature. Their values at reference temperature T_0 are given by the parameters BF and BR, respectively. Their temperature behavior is modeled with the parameter XTB as

$$\beta_F(T) = \text{BF}\left(\frac{T}{T_0}\right)^{\text{XTB}}, \tag{4.3}$$

$$\beta_R(T) = \text{BR}\left(\frac{T}{T_0}\right)^{\text{XTB}}. \tag{4.4}$$

2. The non-ideal base currents model recombination of minorities in the base. For forward and reverse operation they are given by I_{re} and I_{rc}, respectively as

$$I_{re} = \text{ISE}\left(\exp\left(\frac{V_{b'e'}}{\text{NE}U_T}\right) - 1\right), \tag{4.5}$$

$$I_{rc} = \text{ISC}\left(\exp\left(\frac{V_{b'c'}}{\text{NC}U_T}\right) - 1\right). \tag{4.6}$$

3. The transport current of minorities from the emitter to the collector is represented by the transport current I_{cc}

$$I_{cc} = \frac{Q_{b0}}{Q_b}\text{IS}\left(\exp\left(\frac{V_{b'e'}}{\text{NF}U_T}\right) - \exp\left(\frac{V_{b'c'}}{\text{NR}U_T}\right)\right), \tag{4.7}$$

where Q_{b0} and Q_b are the total minority charges in the base at zero bias and applied bias $(V_{b'e'}, V_{b'c'})$, respectively. Their ratio is defined as

$$\frac{Q_b}{Q_{b0}} = \frac{1}{2}\left(1 + \frac{V_{c'b'}}{\text{VAF}} + \frac{V_{e'b'}}{\text{VAR}}\right) + \sqrt{\frac{1}{4}\left(1 + \frac{V_{c'b'}}{\text{VAF}} + \frac{V_{e'b'}}{\text{VAR}}\right)^2 + \frac{I_s}{\text{IKF}}\left(\exp\frac{V_{b'e'}}{\text{NF}U_T} - 1\right) + \frac{I_s}{\text{IKR}}\left(\exp\frac{V_{b'c'}}{\text{NR}U_T} - 1\right)}. \tag{4.8}$$

The parameters VAF and VAR represent the forward and the reverse Early voltages, respectively. The saturation current I_s is a function of temperature. Its value at the reference temperature T_0 equals IS. The temperature effects

are described by XTI and the bandgap voltage EG:

$$I_s(T) = \text{IS}\left(\frac{T}{T_0}\right)^{\text{XTI}} \exp\left(\frac{\text{EG}}{U_T}\frac{(T-T_0)}{T_0}\right). \tag{4.9}$$

The temperature dependence of the non-ideal base currents I_{se} and I_{sc} can differ for the various SPICE versions. We therefore refer to the appropriate reference manual. The bulk resistors R_c, R_b and R_e in series with the terminals of the intrinsic transistor are modeled by the parameters RB, RE and RC, respectively. Some SPICE versions support the modeling of temperature effects on these resistors.

The parameters IKF and IKR represent the current at which high injection occurs for forward and reverse operation, respectively.

Dynamic effects

Under non-equilibrium conditions, changes in the excess minority base charge and in the charge stored in the base-emitter and base-collector junction capacitances cause dynamic currents. The dynamic parts of the terminal currents I_b, I_c and I_e can be obtained as

$$I_c = I_{cc} - I_{bc} - I_{rc} + \frac{dQ_{dc}}{dt} - c_{jc}\frac{dV_{b'c'}}{dt} + c_{js}\frac{dV_{cs}}{dt}, \tag{4.10}$$

$$I_e = -I_{cc} - I_{be} - I_{re} + \frac{dQ_{de}}{dt} - c_{je}\frac{dV_{b'e'}}{dt}, \tag{4.11}$$

$$I_b = I_{be} + I_{bc} + I_{re} + I_{rc} + \frac{d}{dt}Q_b + c_{jc}\frac{dV_{b'c'}}{dt} + c_{je}\frac{dV_{b'e'}}{dt}, \tag{4.12}$$

$$I_s = -c_{js}\frac{dV_{cs}}{dt}, \tag{4.13}$$

where Q_{de} and Q_{dc} represent the excess minority charges in the base, and c_{jc} and c_{je} are the (small-signal) base-collector and base-emitter junction capacitances, respectively. The total excess minority charge depends on the minority transport current I_{cc} and the minority transit time in the base. These transit times are modeled by tf and TR for forward and reverse operation, respectively. The associated excess minority charges Q_{de} and Q_{dc} can be expressed as

$$Q_{de} = I_{cc}\tau_f\left(\exp\left(\frac{V_{b'e'}}{U_T}\right)\right), \tag{4.14}$$

$$Q_{dc} = I_{cc}\text{TR}\left(\exp\left(\frac{V_{b'c'}}{U_T}\right)\right). \tag{4.15}$$

The forward transit time τ_f is modeled as a function of both the forward ideal transport current I_{tf} and the collector to base voltage $V_{b'c'}$. At high currents τ_f increases due to the so-called base push-out or Kirk effect (see [Kirk1962][13]):

$$\tau_f = \text{TF}\left(1 + \left(\text{XTF}\frac{I_{tf}}{I_{tf}+\text{ITF}}\right)^2\right)\exp\left(\frac{V_{b'c'}}{1.44\text{VTF}}\right), \tag{4.16}$$

where the ideal forward transport current I_{tf} is defined as

$$I_{tf} = \text{IS}\exp\left(\frac{V_{b'e'}}{\text{NF}U_T}\right). \tag{4.17}$$

The amount of charge stored in the depletion capacitances of PN-junctions depends on the voltage across these junctions. Due to the voltage dependence of the depletion layer width, this charge shows a nonlinear relation with the

[13] C. T. Kirk. A theory of transistor cutoff frequency (f_t) falloff at high current densities. *IEEE Trans. Electron Devices*, ED-9(2):164–174, February 1962

voltage. In SPICE the $c_j(V)$ relation instead of the $Q_j(V)$ relation is modeled. For reverse biased junctions ($V \leq 0$) the value of the depletion capacitance is modeled as:

$$c_j(V) = \frac{C_j(0)}{\left(1 - \frac{V}{V_j}\right)^M} \tag{4.18}$$

Where $C_j(0)$ is the zero bias depletion capacitance, V_j the built-in junction barrier voltage and M the *junction grading factor*, $M = 0.33$ for linear graded junctions and $M = 0.5$ for step junctions. These expressions are also used for weakly forward biased junctions. SPICE uses this expression if the forward junction voltage V is smaller than FC times the built-in junction voltage. The SPICE parameter FC has a default value of 0.5.

The bipolar transistor model has three junction capacitances. Their values for reverse biased junctions are

$$c_{jc} = \frac{\text{CJC}}{\left(1 - \frac{V_{b'c'}}{\text{VJC}}\right)^{\text{MJC}}}, \tag{4.19}$$

$$c_{je} = \frac{\text{CJE}}{\left(1 - \frac{V_{b'e'}}{\text{VJE}}\right)^{\text{MJE}}}, \tag{4.20}$$

$$c_{js} = \frac{\text{CJS}}{\left(1 - \frac{V_{cs}}{\text{VJS}}\right)^{\text{MJS}}}. \tag{4.21}$$

Model improvements

One improvement of the enhanced Gummel-Poon model is the modeling of a current-dependent base resistance. This effect, however, is not found from noise measurements. The current-dependency of R_b is characterized by three parameters RB, RBM and IBM that represent the maximum value, the minimum value at high currents and the current where the resistor is halfway its minimum value, respectively. In some SPICE versions, the temperature dependency of the bulk resistors is modeled. Some SPICE versions have model extensions for deep saturation (strongly forward biased collector-base junction).

Small-signal dynamic model

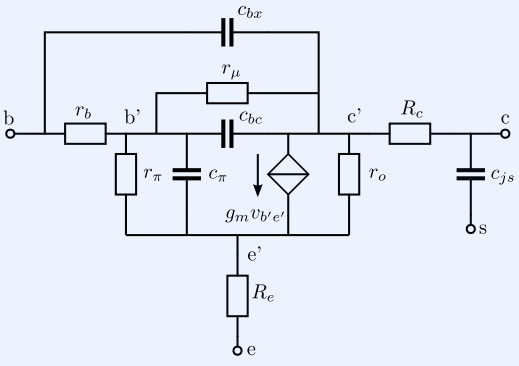


Figure 4.3: Small-signal equivalent circuit of the bipolar transistor in the active forward region. In the Gummel-Poon model, the collector base resistance r_μ is ∞.

For small excursions from an operating point Q, at which the DC operating currents and voltages are given by I_B, I_C, V_{BE} and V_{CE}, we can use a small-signal equivalent circuit. The values of the small-signal model parameters are obtained from differentiation of the $v - i$ and $q - v$ relations in the operating point Q. The small-signal equivalent model of the BJT is shown in

Figure 4.3. The Gummel-Poon model does not include r_μ. This small-signal model is usually referred to as the *hybrid-π* small-signal equivalent circuit, its parameters are

$$
\begin{aligned}
r_b &= R_b|_Q, & r_\pi &= \left.\frac{\partial V_{b'e'}}{\partial I_b}\right|_Q, & c_{bc} &= \left.\left(\frac{\partial Q_{dc}}{\partial V_{b'c'}} + c_{jc}\,\text{XCJC}\right)\right|_Q,\\
R_e &= \text{RE}, & g_m &= \left.\frac{\partial I_c}{\partial V_{b'e'}}\right|_Q, & c_{bx} &= (1-\text{XCJC})\,c_{jc}|_Q,\\
R_c &= \text{RC}, & r_o &= \left.\frac{\partial V_{c'e'}}{\partial I_c}\right|_Q, & c_\pi &= \left.\left(\frac{\partial Q_{de}}{\partial V_{b'e'}} + c_{je}\right)\right|_Q,\\
& & r_\mu &= \left.\frac{\partial V_{c'e'}}{\partial I_b}\right|_Q, & c_{js} &= c_{js}|_Q.
\end{aligned}
$$

Stationary noise model

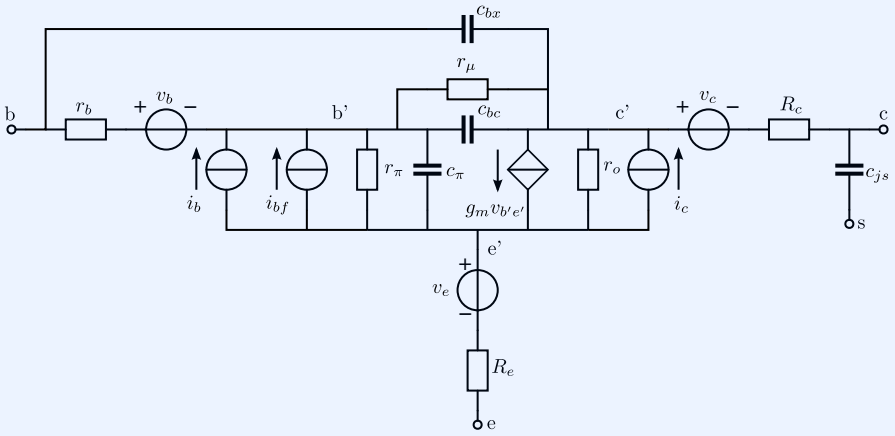


Figure 4.4: Small signal hybrid π equivalent circuit with noise sources.

In SPICE, frequency-domain noise analysis can be performed with a small-signal noise analysis. For this purpose independent and stationary noise sources are added to the small-signal hybrid-π equivalent circuit from Figure 4.3. The resulting noise model is shown in Figure 4.4. The noise sources v_b, v_e and v_c represent the thermal noise of the bulk resistors. These sources have a Gaussian amplitude distribution function and a uniform spectrum. Their spectral densities are

$$S_{vb} = 4kTr_b \quad [\text{V}^2/\text{Hz}], \tag{4.22}$$

$$S_{ve} = 4kTR_e \quad [\text{V}^2/\text{Hz}], \tag{4.23}$$

$$S_{vc} = 4kTR_c \quad [\text{V}^2/\text{Hz}]. \tag{4.24}$$

The current sources i_b and i_c represent the shot noise of the base current and the collector current in the operating point, respectively. These sources have a Gaussian amplitude distribution function and a uniform spectrum. Their spectral densities are given by

$$S_{ib} = 2qI_B \quad [\text{A}^2/\text{Hz}] \tag{4.25}$$

$$S_{ic} = 2qI_C \quad [\text{A}^2/\text{Hz}] \tag{4.26}$$

The noise source i_{bf} represents the flicker noise associated with the base current. This source has a Gaussian amplitude distribution function and a power spectrum that is inversely proportional with frequency. The spectral density S_{ibf} of this source is modeled as

$$S_{ibf} = \frac{\text{KF}\,I_B^{\text{AF}}}{f} \quad [\text{A}^2/\text{Hz}], \tag{4.27}$$

where KF and AF are the model parameters for this $1/f$ noise term.

The noise current associated with the base current is often described with the aid of the noise corner frequency f_ℓ of the $1/f$ noise:

$$S_{ibt} = 2qI_B\left(1+\frac{f_\ell}{f}\right) \quad [\text{A}^2/\text{Hz}], \tag{4.28}$$

where the corner frequency f_ℓ of the $\frac{1}{f}$ noise is obtained from (4.25) and (4.27) as

$$f_\ell = \frac{\text{KF}}{2q} I_B^{(\text{AF}-1)} \quad [\text{Hz}]. \tag{4.29}$$

Low-noise BJTs can have f_ℓ below 100Hz. This cut-off frequency strongly depends on the applied technology. In modern RF IC processes it may exceed 10kHz.

4.2.3 Device parameters

In this section, we will give an overview of the Gummel-Poon model parameters. Both NPN and PNP devices have positive valued parameters. SPICE has different BJT models for vertical and for lateral transistors. The supported BJT types are listed in Table 4.1.

Table 4.1: Types of BJTs supported in SPICE.

model	description
NPN	vertical NPN transistor with collector-substrate capacitance
PNP	vertical PNP transistor with collector-substrate capacitance
LPNP	lateral PNP transistor with base-substrate capacitance

The SPICE parameters of the Gummel-Poon model are listed in Table 4.2. There are two arguments AREA and TEMP that represent the device emitter area scaling factor and the device temperature, respectively.

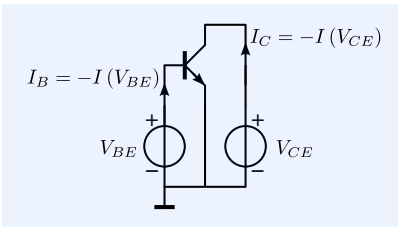


Figure 4.5: Circuit for determination of the DC characteristics of a BJT with SPICE.

4.2.4 Simulated device characteristics

Figure 4.5 shows a test bench for simulation of the device characteristics. The netlist file is shown below:

```
DCchars
* FILE: myNPN_DCchars.cir
* LTspice circuit file
VBE 1 0 0
VCE 2 0 0
Q1 2 1 0 myNPN
.model myNPN NPN
+ IS=0.5f BF=100 NF=1 IKF=100m ISE=10f NE=2 RB=10 VAF=20
+ TF=1n CJE=5p CJC=1p VJE=0.6 VJC=0.8 XTF=1 VTF=2 ITF=20m
.dc VCE 0 5 10m VBE 0.6 0.65 10m
* .dc VBE 0 0.65 10m VCE 1 5 1
.end
```

The simulation results are shown in Figure 4.7. These figures show that the BJT behaves as a voltage-controlled current source. In the Gummel-Poon model the input current I_B only depends on the input voltage. The output current is controlled by the input voltage. Over a wide range of output voltages it shows a weak dependence of the output voltage. For the forward active region, this is characterized with the forward Early voltage VAF. The value of this parameter can be found from the intersection point of the extrapolated $I_C\,(V_{CE})$ characteristic and the x-axis, as shown in Figure 4.6.

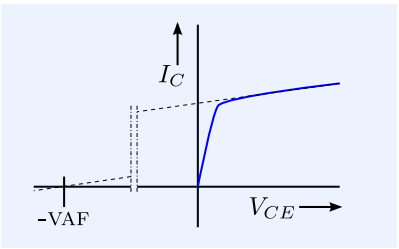


Figure 4.6: Determination of the forward Early voltage VAF from the extrapolated $I_C(V_{CE})$ curve.

Gummel plot

The $I_B\,(V_{BE})$ characteristic and the $I_C\,(V_{BE})$ characteristic plotted together on a semi-logarithmic scale is called the Gummel plot.[14] It shows the ex-

[14] Named after Hermann Karl Gummel (1923, Hannover Germany), a pioneer in the semiconductor industry.

name	description	unit	default	AREA
IS	transport saturation current	A	10^{-16}	*
NF	forward emission coefficient	-	1	
VAF	forward early voltage	V	∞	
IKF	forward high-injection knee current	A	∞	*
BF	ideal forward DC current gain factor	-	100	
ISE	base-emitter leakage saturation current	A	0	*
NE	base-emitter leakage emission coefficient	-	1.2	
EG	bandgap voltage	V	1.11	
XTI	temperature coefficient (IS)	-	3	
NR	reverse emission coefficient	-	1	
VAR	reverse early voltage	V	∞	
IKR	reverse high-injection knee current	A	∞	*
BR	ideal reverse DC current gain factor	-	1	
ISC	base-collector leakage saturation current	A	0	*
NC	base-collector leakage emission coefficient	-	2	
XTB	temperature coefficient (BF, BR, ISE, ISC)	-	0	
RB	zero-bias base resistance	Ω	0	*
IRB	current at which R_b falls halfway to RBM	A	∞	*
RBM	minimum value of R_b at high current	Ω	RB	*
RE	emitter bulk resistance	Ω	0	*
RC	collector bulk resistance	Ω	0	*
TNOM	temperature for parameter measurement	^{o}C	27	
TF	ideal forward transit time	s	0	
XTF	coefficient for bias dependence of TF	-	0	
VTF	voltage describing V_{BC} dependence of TF	V	∞	
ITF	high-current parameter for effect on TF	A	0	*
PTF	excess phase at $f = 1/(2\pi$TF$)$ [Hz]	deg	0	
TR	ideal reverse transit time	s	0	
CJE	base-emitter zero-bias depletion capacitance	F	0	*
VJE	base-emitter built-in potential	V	0.75	
MJE	base-emitter junction grading coefficient	-	0.33	
CJC	base-collector zero-bias depletion capacitance	F	0	*
VJC	base-collector built-in potential	V	0.75	
MJC	base-collector junction grading coefficient	-	0.33	
XCJC	fraction of C_{bc} connected to internal base	-	1	
CJS	collector-substrate zero-bias depletion capacitance	F	0	*
VJS	collector-substrate built-in potential	V	0.75	
MJS	collector-substrate junction grading coefficient	-	0	
FC	forward bias depletion capacitance coefficient	-	0.5	
KF	flicker noise coefficient	-	0	
AF	flicker noise exponent	-	1	

Table 4.2: BJT Gummel-Poon model parameters. An asterix (*) in the AREA column indicates scaling of the corresponding parameter with AREA.

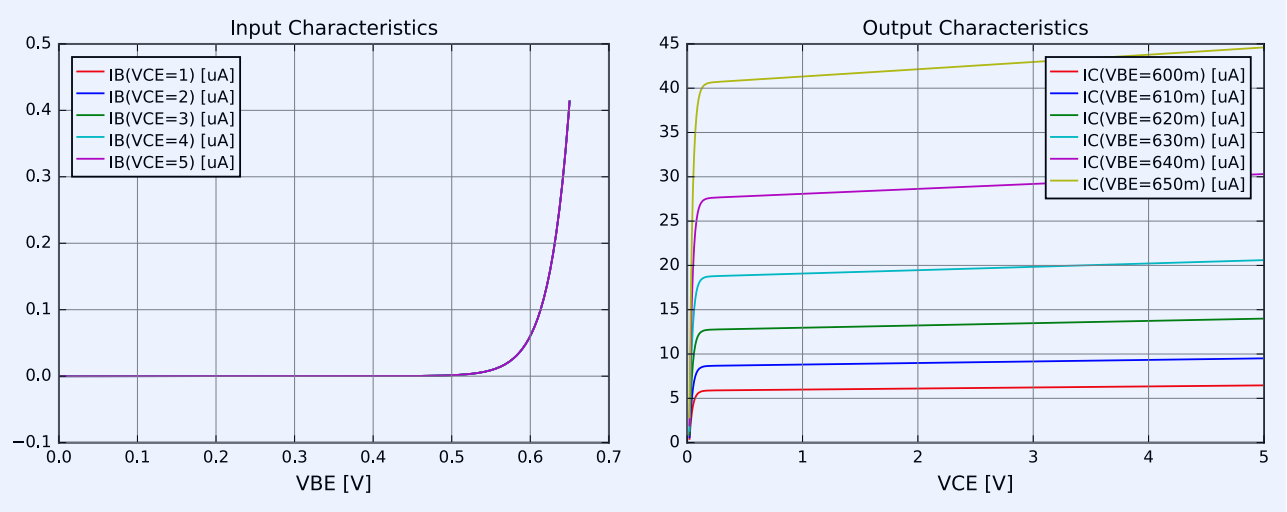


Figure 4.7: DC input and output characteristic of the transistor "myNPN", obtained from simulation with the circuit from Figure 4.5.

[15] Simulation of very small currents requires a modification of the minimum conductance simulation parameter GMIN. This parameter determines the value of the conductance that SPICE places between the nodes to enhance the convergence during determination of the operating point.

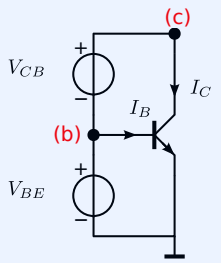


Figure 4.8: Circuit for determination of the Gummel plot with SPICE.

ponential relation between the base to emitter voltage and the base current components. Figure 4.8 shows the simulation setup for this plot. The SPICE netlist file is shown below:[15]

```
* FILE: gummelPlot.cir
* LTspice circuit file
VBE b 0 0
VCB c b 0
Q1 c b 0 myNPN
.model myNPN NPN
+ IS=0.5f BF=100 NF=1 IKF=100m ISE=10f NE=2 RB=10
.options gmin=1f
.dc VBE 0 1 10m
.end
```

Figure 4.9 shows the resulting Gummel plot. It shows the exponential relation between the base-emitter voltage and the collector current over many decades. The parameter BF models the ratio between the ideal collector current and the ideal base current. At very low current levels, the non-ideal base current models the recombination of minorities in the base. At high collector currents, high injection reduces the slope of the $I_C\,(V_{BE})$ curve by a factor two, while the voltage drop across the bulk resistors causes a deviation of the ideal exponential relation.

Definition of β_{DC}

The DC current gain β_{DC} of a bipolar transistor is defined as the ratio of the DC collector current I_C and the DC base current I_B:

$$\beta_{DC} = \frac{I_C}{I_B}.$$

The DC current gain is a function of the collector current. This can be seen from the Gummel plot. Due to the Early effect, it also increases with the collector-emitter voltage.

Definition of β_{AC}

The small-signal current gain is somewhat misleadingly denoted by β_{AC}. This small-signal quantity is defined for infinitesimally small excursions from

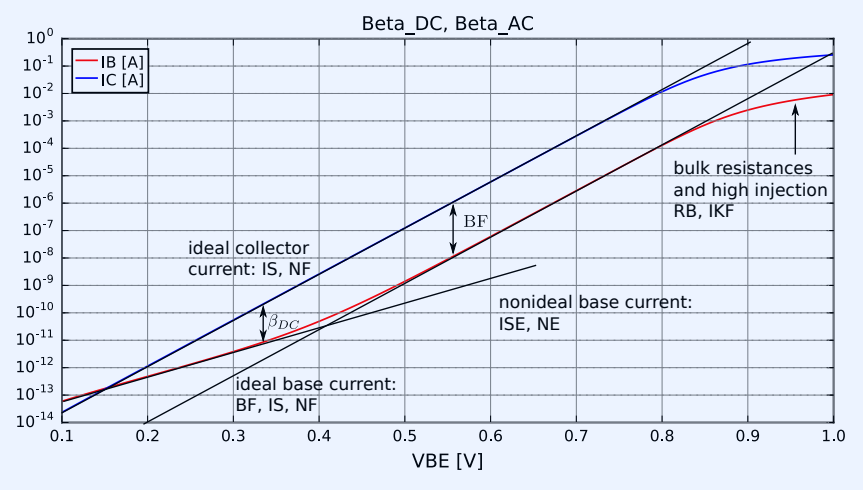


Figure 4.9: Gummel plot of the transistor "myNPN", obtained by simulation with the circuit from figure 4.8.

the operating point:

$$\beta_{AC} = \left.\frac{\partial I_c}{\partial I_b}\right|_Q, \tag{4.30}$$

alternatively we may write

$$\beta_{AC} = \left.\frac{\partial I_c}{\partial V_{b'e'}}\right|_Q \cdot \left.\frac{\partial V_{b'e'}}{\partial I_b}\right|_Q = g_m r_\pi. \tag{4.31}$$

Figure 4.10 shows the DC current gain and the static (zero-frequency) forward small-signal current gain β_{AC} versus the collector current of `myNPN` for $V_{CB} = 0$V.

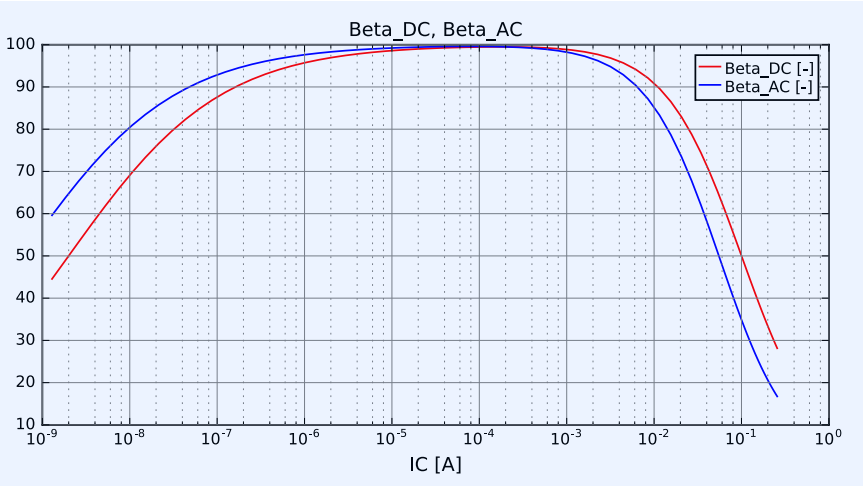


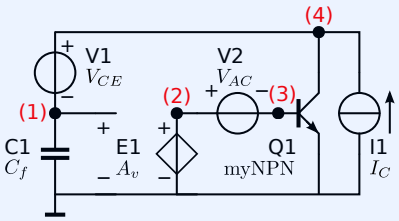
Figure 4.10: The forward DC current gain β_{DC} and the forward static small-signal current gain β_{AC} of "myNPN" as a function of the collector current.

Dynamic behavior of β_{AC}

At high frequencies the small-signal current gain β_{AC} drops below its static value. This is caused by charge storage in the base due to the finite transit time for minorities and in the depletion regions. The cut-off frequency of a bipolar transistor is defined as the frequency at which the magnitude of β_{AC} equals unity. The cut-off frequency ω_T is a figure of merit for the speed limitation of the bipolar transistor.

Figure 4.11 shows the simulation test bench for the determination of the

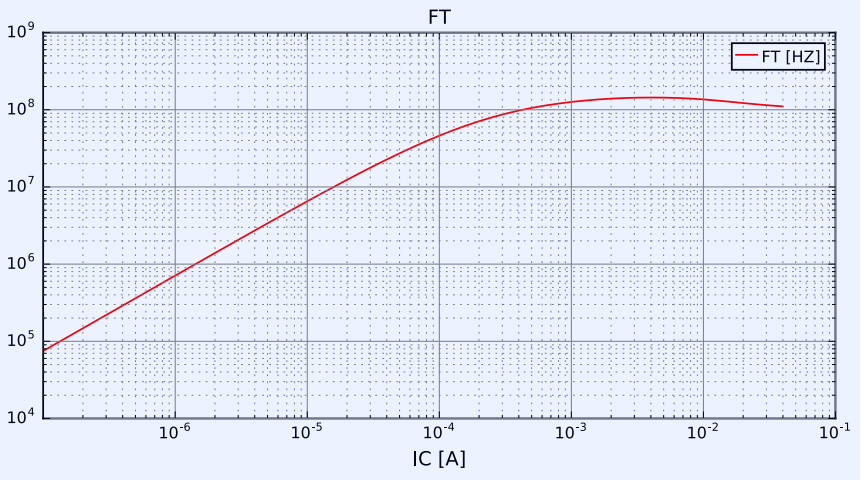
Figure 4.11: Simulation test bench for determination of the cut-off frequency f_T as a function of the DC collector current.



cut-off frequency using SPICE. The voltage amplifier E_1 provides the driving conditions for the bipolar transistor such that its operating collector current and collector-emitter voltage equal I_C and V_{CE}, respectively. The large capacitor C_1 prevents feedback at signal frequencies of interest, while it concurrently provides the condition for measuring the short circuit collector current. Below is the LTSPICE listing of the netlist file of the circuit from Figure 4.11:

```
myNPN_FT
* FILE: myNPN_FT.cir
* LTspice circuit file for plotting fT(Ic)
V1 4 1 3
V2 2 3 DC 0 AC 1
I1 0 4 {Ic}
C1 1 0 1
E1 2 0 1 0 1
Q1 4 3 0 myNPN
.model myNPN NPN
+ IS=0.5f BF=100 NF=1 IKF=100m ISE=10f NE=2 RB=10
+ TF=1n CJE=5p CJC=1p VJE=0.6 VJC=0.8 XTF=1 VTF=2 ITF=20m
.ac dec 50 1 1G
* LTspice syntax for plotting fT(Ic)
.step dec param Ic 100n 10m 10
.meas AC fT FIND Frequency WHEN dB(I(V1)/I(V2))=0 CROSS=1
* Run this netlist and press CTRL + L in the trace window
* This will bring up the output file (Spice error log) window
* In this window right-click "Plot .step'ed .meas data"
.end
```

Figure 4.12: Plot of FT versus frequency of "myNPN", obtained with the simulation test bench from Figure 4.11.



The simulation results for myNPN are shown in Figure 4.12. At low currents the cut-off frequency is dominated by the depletion capacitance, while at high frequencies it is limited by the forward transit time TF. The maximum value

$f_{T\,\max}$ of the cut-off frequency can be estimated from TF:

$$f_{T\,\max} \approx \frac{1}{2\pi \mathrm{TF}}. \tag{4.32}$$

At high collector currents, the base region tends to extend towards the collector. This base push-out or Kirk-effect (see [Kirk1962][16]) causes a reduction of the cut-off frequency at high collector currents.

[16] C. T. Kirk. A theory of transistor cut-off frequency (f_t) falloff at high current densities. *IEEE Trans. Electron Devices*, ED-9(2):164–174, February 1962

Operating point information

The current-drive capability and voltage-drive capability of a BJT depend on the DC collector current I_C and the collector-emitter voltage V_{CE}, respectively. Many other performance aspects, such as, the noise performance and the cut-off frequency also show a direct relation with the collector current. As a consequence, we usually want to fix the operating point through fixing I_C and V_{CE}.

A method for fixing the operating point of nonlinear resistive multi-terminal devices has been discussed in Chapter 3. According to the presented method, fixing the operating point of a BJT by means of I_C and V_{CE}, requires the addition of a voltage source V_{CE} between the collector and the output and a current source I_C that flows from the emitter to the collector.

In order to obtain zero output voltage and zero output current for all DC input and output terminations, a voltage source V_{BE} has to be placed in series with the base and a current source I_B has to be connected in parallel with the base-emitter junction. The values of these input sources depend on the required values of I_C and V_{CE}, on the DC characteristics of the device and on the operating temperature. They can be determined with the aid of the circuit from Figure 4.13. The nullator at the output port (collector-emitter) sets the condition for zero output voltage and zero output current, while the norator at the input port delivers the correct driving quantities to satisfy these conditions. The nullor is not available in SPICE but it can be implemented with two unity-gain voltage-controlled voltage sources as illustrated in this figure.

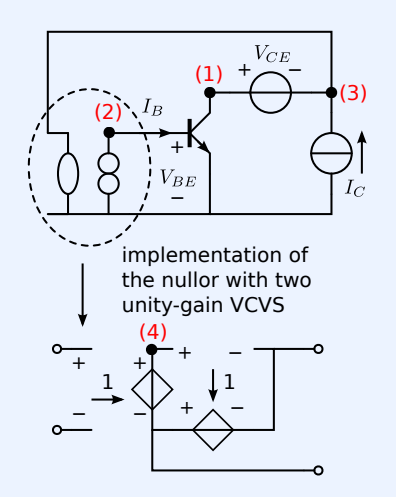


Figure 4.13: Simulation test bench for determination of the small-signal parameters.

SPICE returns small-signal parameters in a certain operating point as the result of an operating point analysis (`.op` statement). The parameters returned depend on the device model and the SPICE version. Table 4.3 gives an overview of the small-signal parameters of the Gummel-Poon model returned by different SPICE versions.

Below the listing of a circuit for determination of the small-signal parameters of myNPN for $I_C = 1\text{mA}$ and $V_{CE} = 3\text{V}$.

```
1  * FILE: myNPN_OP.cir
2  * SPICE circuit file
3  *
4  * Transistor with VCE and IC definition
5  Q1 1 2 0 myNPN
6  VCE 1 3 {V_CE}
7  IC 0 3 {I_C}
8  *
9  * nullor
10 E1 4 0 3 0 1
11 E2 2 0 2 4 1
12 *
13 .model myNPN NPN
14 + IS=0.5f BF=100 NF=1 IKF=100m ISE=10f NE=2 RB=10
15 + TF=1n CJE=5p CJC=1p VJE=0.6 VJC=0.8 XTF=1 VTF=2 ITF=20m
16 *
17 .param V_CE=3 I_C=1m
18 .op
19 .end
```

If simulators do not provide the operating point information for the hybrid-π equivalent circuit from Figure 4.3, the designer can:

1. Estimate it from the characteristics given in data sheets or design manuals
2. Obtain it from AC (small-signal) analysis in the desired operating point.

name	**description** (see Figure 4.3)	LTspice	SIMetrix	Pspice	ngspice
β_{DC}	DC current gain	BetaDC	BetaDC	BETADC	
β_{AC}	Zero frequency small-signal current gain	BetaAC	BetaAC	BETAAC	
c_{bc}	Small-signal internal base-collector capacitance	Cbc	Cjc	CBC	cmu
c_{π}	Small-signal base-emitter capacitance	Cbe	Cje	CBE	cpi
c_{cs}	Small-signal collector-substrate capacitance	Cjs	Cjs	CJS	csub
c_{bx}	Small-signal external base-collector capacitance	Cbx	Cxjc	CBX	cbx
g_m	Forward transconductance	Gm	Gm	GM	gm
I_B	DC base current	Ib	IB	IB	ib
I_C	DC collector current	Ic	IC	IC	ic
I_E	DC collector current	-	IE		ie
I_S	DC substrate current	-	IS	-	-
P	DC power dissipation	-	Power	-	-
r_b	Small-signal base resistance	Rx	Rbase	-	-
$g_b = \frac{1}{r_b}$	Small-signal base conductance	-	-	GX	gx
r_o	Small-signal output resistance	Ro	Ro	RO	-
$g_o = \frac{1}{r_o}$	Small-signal output conductance	-	-		go
r_{π}	Small-signal internal base-emitter resistance	Rpi	Rpi	RPI	-
$g_{\pi} = \frac{1}{r_{\pi}}$	Small-signal internal base-emitter conductance	-	-	-	gpi
r_{μ}	Reverse resistance	-	-	-	-
$g_{\mu} = \frac{1}{r_{\mu}}$	Reverse conductance	-	-	-	gmu
F_T	Cut-off frequency	Ft	-	FT	-
V_{BE}	DC base-emitter voltage	Vbe	Vbe	VBE	vbe
V_{CB}	DC collector-base voltage	Vbc	Vcb	VBC	vbc
V_{CE}	DC collector-emitter voltage	Vce	Vce	VCE	-

Table 4.3: Small-signal parameters of the Gummel-Poon model, returned from an operating point analysis with various simulators.

4.2.5 Other models

For modern SiGe RF IC processes the Vertical Bipolar Inter-Company (VBIC) model of the BJT is often used. The model is a public domain replacement for the Gummel-Poon model. It has improved modeling of the saturation region and includes modeling of breakdown and thermal effects. For more information, the reader is referred to literature [VBIC1996][17].

[17] Colin C. McAndrew, Jerold A. Seitchik, Derek F. Bowers, Mark Dunn, Mark Foisy, Ian Getreu, Marc McSwain, Shahrir Moinian, James Parjer, David J. Roulston, Michael Schroeter, Paul van Wijnen, Lawrence F. Wagner. VBIC95, The Vertical Bipolar Inter-Company Model. *IEEE Journal of solid-state circuits*, 31(10):1476–1483, October 1996

4.2.6 Simplified models for hand calculations

Complete SPICE models are suitable for numerical simulations but they are too complex to provide design information from analytical expressions. For this purpose we need simplified models that are suited for hand calculations. In this section, we will derive these models from the Gummel-Poon model. We will introduce a large-signal static model, a small-signal dynamic model and a noise model that can be used for analytical determination of the operating point, the dynamic small-signal transfer and the noise behavior of the bipolar transistor, respectively.

DC behavior

[18] DC behavior.

The static large-signal behavior[18] of a bipolar transistor in the *active forward*

operating region can roughly be predicted from the model shown in Figure 4.14. In this model, the bulk resistances, the high-injection effect, the non-ideal base current and reverse operation have not been modeled. The model describes port currents of a nonlinear resistive two-port as a function of the port voltages:

$$I_C = \text{IS}\left(\exp\left(\frac{V_{BE}}{U_T}\right)\right)\left(1+\frac{V_{CE}}{\text{VAF}}\right), \tag{4.33}$$

$$I_B = \frac{I_C}{\text{BF}}. \tag{4.34}$$

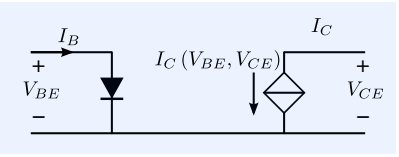


Figure 4.14: Simplified BJT model for nonlinear instantaneous (DC) behavior.

The active forward operating range is limited by

$$V_{BE} > 0, \tag{4.35}$$

$$V_{CE} > V_{CE,SAT}. \tag{4.36}$$

Where $V_{CE,SAT}$ is the collector-emitter voltage where the base-collector junction is forward biased at a current that cannot longer be neglected with respect to the current in the forward biased base-emitter junction.

A simple model for hand calculations that includes operation in the forward saturation region can be formulated as

$$I_C = \text{IS}\left(\left(\exp\frac{V_{BE}}{U_T}\right)-\left(\exp\frac{V_{BC}}{U_T}\right)\right)\left(1+\frac{V_{CE}}{\text{VAF}}\right), \tag{4.37}$$

$$I_B = \frac{\text{IS}}{\text{BETAF}}\left(\exp\frac{V_{BE}}{U_T}-1\right)-\frac{\text{IS}}{\text{BETAR}}\left(\exp\frac{V_{BC}}{U_T}-1\right). \tag{4.38}$$

Small-signal dynamic model

In most cases, the small-signal dynamic behavior of the bipolar junction transistor in the active forward operating region can sufficiently accurate be described with the simplified hybrid-π equivalent circuit from Figure 4.15. The bulk resistances R_e and R_c have been omitted. If $\text{XJC} \leq 0.5$, the base-collector capacitances c_{bc} and c_{bx} can both be connected to the internal node (b'), if not, they should both be connected to the external node (b). For both cases we may then use

$$c_\mu = c_{bx} + c_{bc}. \tag{4.39}$$

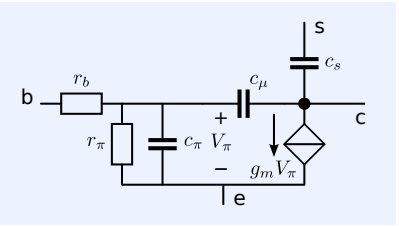


Figure 4.15: Simplified hybrid-π small-signal equivalent circuit of the BJT.

The small-signal parameters of the simplified hybrid-π equivalent circuit can be estimated from the Gummel-poon model parameters and the operating point, thereby ignoring the current dependency of the current gain, of the base resistance and of the forward transit time. In this way we obtain, for operation in the active forward region

$$\begin{array}{lll} r_b = \text{RB}, & r_o = \frac{\text{VAF}+V_{CE}}{I_C}, & c_\pi = \text{CJE} + \text{TF}\frac{I_C}{V_T}, \\ r_\pi = \text{BF}\frac{V_T}{I_C}, & g_m = \frac{I_C}{V_T}, & c_\mu = \text{CJC}. \end{array} \tag{4.40}$$

f_T and $f_{\max}$

The cut-off frequency f_T is defined as the unity-gain frequency of β_{AC}. It can be obtained as

$$f_T = \frac{g_m}{2\pi\left(c_\pi + c_\mu\right)}. \tag{4.41}$$

The cut-off frequency can be seen as a figure of merit for the amplifying capabilities of a transistor.

A better figure of merit for the amplifying capabilities of an active device, is the frequency at which the power gain of the transistor equals unity, when

is driven from and terminated with the complex conjugate of its input and output impedances, respectively. Unity power gain of the BJT is found at the frequency $f_{\max}$, were

$$f_{\max} = \sqrt{\frac{f_T}{8\pi r_b c_\mu}}. \tag{4.42}$$

Noise model

The simplified noise model is shown in Figure 4.16. The noise contributions of the bulk resistances R_e and R_c are omitted, while the remaining noise sources of the complete noise model from Figure 4.4 have been added to the simplified hybrid-π equivalent circuit from Figure 4.15.

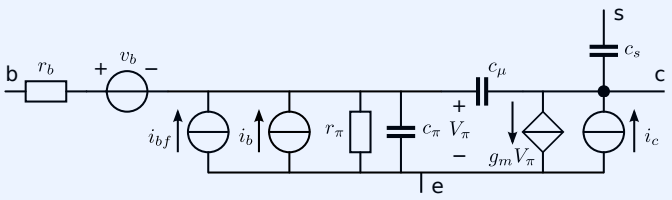


Figure 4.16: Simplified hybrid-π small-signal equivalent circuit with stationary noise sources.

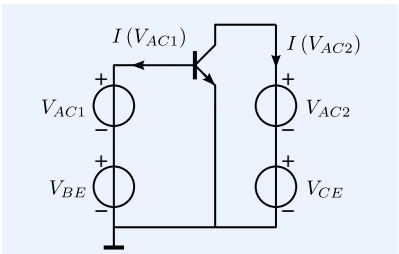


Figure 4.17: Simulation test bench for determination of the hybrid-π equivalent circuit parameters.

Determination of hybrid-π parameters from simulation

The small-signal parameters can also be obtained from simulation. Figure 4.17 shows the simulation test bench for determination of the small-signal hybrid-π parameters.

The procedure is as follows:

1. Bias the transistor in the required operating point with the aid of the *DC* voltage sources V_{BE} and V_{CE}. The value of V_{BE} can be obtained from an operating point simulation with the test bench from Figure 4.13.
2. Make $V_{AC1} = 1,0$ (magnitude 1, phase 0) and $V_{AC2} = 0$ and perform an AC analysis over the frequency range of interest
3. Obtain approximations the values of the following parameters

$$r_b + r_\pi = \frac{1}{\mathrm{Re}\{-I(V_{AC1})\}}, \tag{4.43}$$

$$g_m \frac{r_\pi}{r_b + r_\pi} = \mathrm{Re}\{-I(V_{AC2})\}, \tag{4.44}$$

$$c_\pi + c_\mu = \frac{\mathrm{Im}\{-I(V_{AC1})\}}{2\pi f}. \tag{4.45}$$

4. Make $V_{AC1} = 0$ and $V_{AC2} = 1,0$ (magnitude 1, phase 0)
5. Obtain approximations for the values of the following parameters

$$r_o = \frac{1}{\mathrm{Re}\{-I(V_{AC2})\}}, \tag{4.46}$$

$$c_\mu = \frac{\mathrm{Im}\{I(V_{AC1})\}}{2\pi f}, \tag{4.47}$$

$$r_\mu = \frac{r_\pi + r_b}{r_\pi} \frac{1}{\mathrm{Re}\{-I(V_{AC1})\}}. \tag{4.48}$$

6. The base resistance r_b can at best be determined from noise measurements. This will be explained later.

4.3 Junction Field Effect Transistors

The principle of operation of a Junction Field effect device was first patented by J. E. Lilienfeld in 1925: [Lilienfeld1925][19]. In 1952 Shockley (see [ShockleyFET1952][20]) presented the theory for this device. The first working devices were reported by Dacey and Ross in 1953: [DaceyRoss1952][21].

Junction field effect transistors (JFETs) have a gate that is isolated by a depletion layer. They are available as discrete devices and in integrated circuit technology.

Figure 4.18 shows how a P-channel JFET is realized in a bipolar integrated circuit process. N-channel and P-channel JFETs are also available as discrete devices.

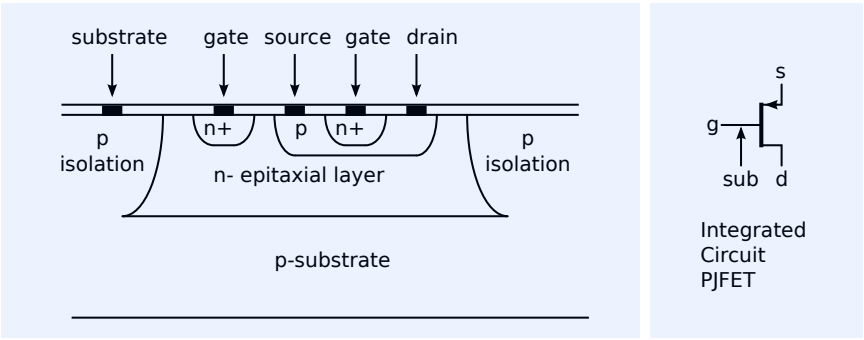


Figure 4.18: Cross section and device symbol of a p-channel JFET in (bipolar) integrated circuit technology

[19] Julius Edgar Lilienfeld. Method and apparatus for controlling electric currents, October 1925

[20] W. Shockley. A Unipolar "Field-Effect" Transistor. *Proc. of the I.R.E*, pages 1365–1376, November 1952

[21] G.C. "Dacey and I.M." Ross. Unipolar "Field-Effect" Transistor. *Proc. of the I.R.E*, pages 970–979, August 1953

4.3.1 JFET simulation model

The SPICE simulation model for Junction field effect transistors is derived from the model proposed by Shichman and Hodges [ShichmanHodges1968][22], it is shown in Figure 4.19. The intrinsic JFET with connections g, d' and s', is modeled with two diodes, two nonlinear capacitors and a voltage-controlled current source. Source and drain series resistances R_s and R_d complete the model.

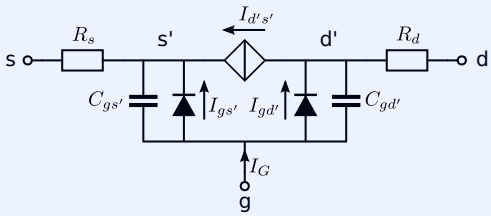


Figure 4.19: NJFET model according to Shichman and Hodges.

[22] H. Shichman and D. Hodges. Modeling and simulation of insulated-gate field-effect transistor switching circuits. *IEEE J. Solid-State Circuits*, 3(3):285–289, September 1968

Instantaneous behavior

The instantaneous behavior of the intrinsic JFET is described by the $I(V)$ relations of the two diodes and by a voltage-controlled channel current $I_{d's'}\left(V_{gs'}, V_{d's'}\right)$.

The $I(V)$ relations of the two diodes are

$$I_{gs'} = I_s \left(\exp \frac{V_{gs'}}{\text{N}U_T} - 1 \right) + \text{ISR} \left(\exp \frac{V_{gs'}}{\text{NR}U_T} - 1 \right) \left(\left(1 - \frac{V_{gs'}}{\text{PB}} \right)^2 + 0.005 \right)^{\frac{\text{M}}{2}}, \tag{4.49}$$

$$I_{gd'} = I_s \left(\exp \frac{V_{gd'}}{\text{N}U_T} - 1 \right) + \text{ISR} \left(\exp \frac{V_{gd'}}{\text{NR}U_T} - 1 \right) \left(\left(1 - \frac{V_{gs'}}{\text{PB}} \right)^2 + 0.005 \right)^{\frac{\text{M}}{2}} + I_i, \tag{4.50}$$

where I_i is the impact ionization current, which differs from zero in the forward saturation region ($V_{gs'} > V_t$ and $V_{gd'} < V_t$)

$$I_i = \text{ALPHA} \left(V_{d's'} - \left(V_{gs'} - V_t \right) \right) \exp \left(\frac{-\text{VK}}{V_{d's'} - \left(V_{gs'} - V_t \right)} \right). \tag{4.51}$$

The saturation currents I_s is obtained from the model parameter IS, the temperature coefficient XTI, the temperature T and the scaling parameter AREA as

$$I_s = \text{AREA} \cdot \text{IS} \left(\frac{T}{T_0} \right)^{\text{XTI}}. \tag{4.52}$$

The reference temperature T_0 is the temperature at which the parameters have been measured. Its default value in SPICE is 300K. The scaling factor AREA scales the device as if AREA devices are connected in parallel.

The DC gate current I_G is the sum of the two diode currents:

$$I_G = I_{gs'} + I_{gd'}. \tag{4.53}$$

The bulk resistances R_d and R_s are obtained from the model parameters RE and RS, respectively. They depend on the scaling factor AREA:

$$R_d = \frac{\text{RD}}{\text{AREA}}, \tag{4.54}$$

$$R_s = \frac{\text{RS}}{\text{AREA}}. \tag{4.55}$$

The voltage-controlled current $I_{d's'} \left(V_{gs'}, V_{d's'} \right)$ is modeled differently for different operating regions:

1. Forward mode ($V_{d's'} > 0$) cut-off region: $V_{gs'} < V_t$

   With the gate-source voltage $V_{gs'}$ below the threshold voltage V_t the device is turned off:

   $$I_{d's'} = 0. \tag{4.56}$$

2. Forward mode ($V_{d's'} > 0$) linear region: $V_{gs'} > V_t$ and $V_{gd'} > V_t$

   When both the intrinsic gate-source voltage $V_{gs'}$ and the intrinsic gate-drain voltage $V_{gd'}$ are above threshold, the device operates in the so-called linear region. In the linear region the device can be considered a voltage-controlled resistor. The resistance of this resistor also depends on the intrinsic drain-source voltage

   $$I_{d's'} = \text{AREA}.\beta \left(1 + \text{LAMBDA}.V_{d's'} \right) V_{d's'} \left(2 \left(V_{gs'} - V_t \right) - V_{d's'} \right). \tag{4.57}$$

3. Forward mode ($V_{d's'} > 0$) saturation region: $V_{gs'} > V_t$ and $V_{gd'} < V_t$

   With the intrinsic gate-source voltage $V_{gs'}$ above threshold and the intrinsic gate-drain voltage $V_{gd'}$ below threshold the device acts as a voltage-controlled current source. The drain-source current depends approxi-

mately quadratically on the so-called *effective intrinsic gate-source voltage* $V_{gs'} - V_t$:

$$I_{d's'} = \text{AREA}.\beta\,(1 + \text{LAMBDA}.V_{d's'})\left(V_{gs'} - V_t\right)^2. \tag{4.58}$$

4. For reverse mode operation ($V_{d's'} < 0$) the above conditions and equations hold after the source and the drain connections have been swapped.

The transconductance factor β [A^2/V] and its temperature dependency is modeled with two parameters BETA and BETATC:

$$\beta = \text{BETA} \times 1.01^{\text{BETATC}(T-T_0)}.$$

The threshold voltage V_t is defined by the parameter VTO, its temperature coefficient VTOTC and the temperature T:

$$V_t = \text{VTO} + (T - T_0)\,\text{VTOTC}.$$

The parameter LAMBDA models the change of the drain-source current due to a change of the channel voltage $V_{d's'}$ of the intrinsic JFET, in a similar way as the Early voltage for a bipolar transistor. The Shichman and Hodges model does not have a smooth transition between the cut-off region, the linear region and the saturation region. This results in discontinuities in the $dI(V)/dV$ characteristic, which may result in unreliable distortion simulation results with signals that show large drain-source voltage excursions.

Modeling of dynamic effects

The charge storage in JFETs is modeled by the voltage-dependent junction capacitances. The built-in junction voltage PB, the grading coefficient M and the zero bias depletion capacitances CGS, CDS and FC are associated parameters. Similar as with bipolar transistors the parameter FC is used for modeling of the capacitances in case of a forward-biased junction. The capacitances $C_{gs'}$ and $C_{gd'}$ are proportional with the scaling constant AREA. The model equations are

$$C_{gs'} = \frac{\text{AREA}.\text{CGS}}{\left(1 - \frac{V_{gs'}}{\text{PB}}\right)^{\text{M}}}; \quad V_{gs'} \le \text{FC}.\text{PB}, \tag{4.59}$$

$$C_{gs'} = \frac{\text{AREA}.\text{CGS}}{(1 - \text{FC})^{\text{M}+1}}\left(1 - \text{FC}\,(\text{M}+1) + \text{M}\frac{V_{gs'}}{\text{PB}}\right); \quad V_{gs'} > \text{FC}.\text{PB}, \tag{4.60}$$

$$C_{gd'} = \frac{\text{AREA}.\text{CDS}}{\left(1 - \frac{V_{gd'}}{\text{PB}}\right)^{\text{M}}}; \quad V_{gd'} \le \text{FC}.\text{PB}, \tag{4.61}$$

$$C_{gs'} = \frac{\text{AREA}.\text{CGS}}{(1 - \text{FC})^{\text{M}+1}}\left(1 - \text{FC}\,(\text{M}+1) + \text{M}\frac{V_{gd'}}{\text{PB}}\right); \quad V_{gd'} > \text{FC}.\text{PB}. \tag{4.62}$$

Small-signal equivalent circuit

The small-signal equivalent circuit of the JFET is shown in Figure 4.20. It is obtained from linearization of the model from Figure 4.19 in an operating point Q , defined by V_{GS}, V_{DS}, I_G and I_{DS}.

The parameter values are

$$\begin{aligned} c_{gs'} &= C_{gs'}\Big|_Q, \quad r_{ds} = \frac{\partial V_{d's'}}{\partial I_{ds}}\Big|_Q, \quad R_d = \frac{\text{RD}}{\text{AREA}}, \\ c_{gs'} &= C_{gs'}\Big|_Q, \quad g_m = \frac{\partial I_{d's'}}{\partial V_{gs'}}\Big|_Q, \quad R_s = \frac{\text{RS}}{\text{AREA}}. \end{aligned} \tag{4.63}$$

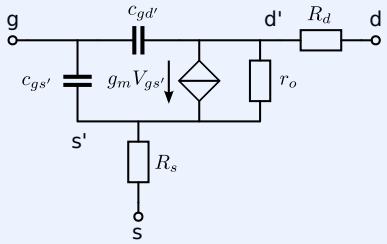


Figure 4.20: JFET small-signal equivalent circuit.

Stationary noise model

Figure 4.20 shows the small-signal noise model of the JFET. It models noise associated with the bulk resistances and with the channel current.

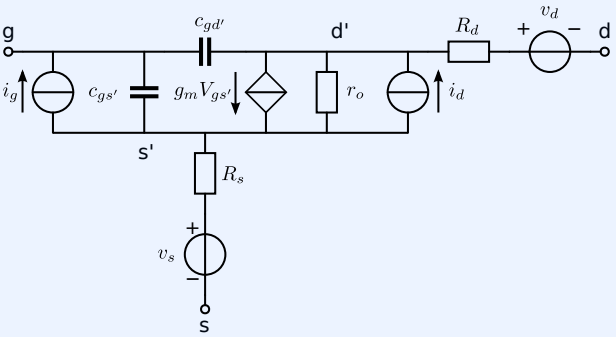


Figure 4.21: JFET small-signal equivalent circuit with stationary noise sources.

The thermal noise of the two bulk resistors is given by

$$S_{vs} = 4kT\frac{\mathrm{RS}}{\mathrm{AREA}}, \tag{4.64}$$

$$S_{vd} = 4kT\frac{\mathrm{RD}}{\mathrm{AREA}}. \tag{4.65}$$

The shot noise i_g associated with the gate current has a spectral density

$$S_{ig} = 2qI_G. \tag{4.66}$$

The thermal noise of the channel current shows has a $\frac{1}{f}$ component:

$$S_{id} = 4kTg_m\frac{2}{3} + \frac{\mathrm{KF}}{f}I_{DS}^{\mathrm{AF}}. \tag{4.67}$$

This can be denoted as

$$S_{id} = 4kTg_m\frac{2}{3}\left(1 + \frac{f_\ell}{f}\right), \tag{4.68}$$

$$f_\ell = \frac{3\mathrm{KF}}{8kTg_m}I_{DS}^{\mathrm{AF}}, \tag{4.69}$$

where f_ℓ is the corner frequency for the $\frac{1}{f}$ noise. Low noise JFETs can have f_ℓ below 100Hz. This cut-off frequency strongly depends on the technology and may as well exceed 10kHz.

4.3.2 Device parameters

In this section we will give an overview of the SPICE model parameters for junction FETs. The following model types are supported:

model	description
NJF	N-channel Junction Field Effect Transistor
PJF	P-channel Junction Field Effect Transistor

Table 4.4 gives an overview of the SPICE JFET model parameters. The scaling factor AREA scales the device as if AREA devices are connected in parallel. Some SPICE versions have additional parameters. The parameters listed in table 4.4 are supported by LTSPICE, SIMETRIX and PSPICE.

name	description	unit	default
VTO	pinch-off voltage	V	-2
BETA	transconductance factor	A/V^2	10^{-4}
LAMBDA	channel length modulation coefficient	$1/V$	0
RD	drain bulk resistance	Ω	0
RS	source bulk resistance	Ω	0
IS	gate junction saturation current	A	10^{-14}
CGS	zero-bias gate-source capacitance	F	0
CGD	zero-bias gate-drain capacitance	F	0
PB	built-in gate junction potential	V	1
FC	forward bias depl. cap. coefficient	V	0.5
AF	flicker noise exponent		1
KF	flicker-noise coefficient		0
N	gate junction emission coefficient	-	1
NR	gate junction recombination coefficient	-	2
XTI	saturation current temperature coefficient	-	3
ISR	recombination saturation current	A	0
ALPHA	ionization coefficient	$1/V$	0
VK	ionization knee voltage	V	0
M	junction grading coefficient	-	0.5
VTOTC	vto temperature coefficient	$V/^oC$	0
BETATC	beta exponential temperature coefficient	$\%/^oC$	0

Table 4.4: SPICE JFET model parameters; currents, transconductance and capacitances are proportional to AREA, resistances are inversely proportional to AREA.

4.3.3 JFET simulated device characteristics

Figure 4.22 shows a test bench for simulation of the device characteristics. The netlist file is shown below:

```
DCchars
* FILE: myNJF_DCchars.cir
* LTspice circuit file
VGS 1 0 0
VDS 2 0 0
J1 2 1 0 myNJF
.model myNJF NJF Beta=25m Betatce=-.5 Rd=1 Rs=10 Lambda=40m
+ Vto=-.6 Vtotc=-2m Is=250p Isr=1p N=1 Nr=2 Xti=3 Alpha=-1m
+ Vk=30 Cgd=5p M=.6 Pb=.5 Fc=.5 Cgs=5p Kf=50a Af=1
* Syntax of nested DC analysis depends on spice version
*.dc VDS 0 5 10m VGS -0.6 0 0.1
.dc VGS -0.6 0 10m VDS 1 5 1
.end
```

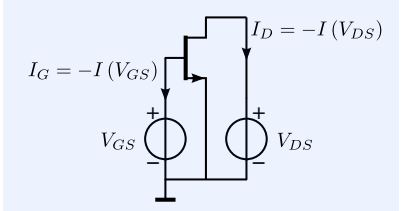


Figure 4.22: Circuit for determination of the DC characteristics of an n-channel JFET with SPICE.

Figure 4.7 shows the forward transfer characteristics and the output characteristics of the junction FET myNJF. The transfer characteristics in the for-

ward saturation region and with LAMBDA $= 0$ can be approximated by

$$I_{ds} = \text{AREA.BETA}\left(V_{gs} - \text{VTO}\right)^2. \tag{4.70}$$

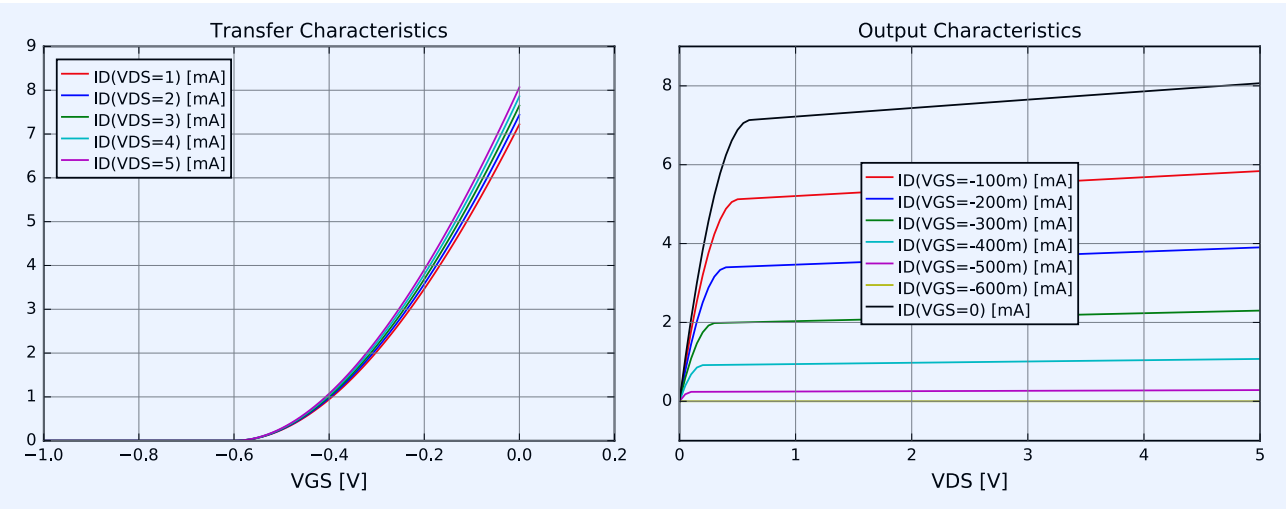


Figure 4.23: DC transfer characteristics and output characteristics of "myNJF", obtained from simulation with the circuit from Figure 4.22.

Saturation current

When $V_{gs} = 0$ the drain current equals the so-called *saturation current* I_{dss}. Since the gate-source junction of a JFET should be reversely biased, I_{dss} is the largest possible drain current; with LAMBDA $= 0$, it equals

$$I_{dss} = \text{AREA.BETA.VTO}^2. \tag{4.71}$$

Cut-off frequency

The cut-off frequency of a JFET is defined in a similar way as with bipolar transistors. The unity gain frequency for the current gain is found as

$$f_T = \frac{g_m}{2\pi\left(C_{gs} + C_{ds}\right)}, \tag{4.72}$$

$$f_T = \frac{g_m}{2\pi C_{iss}}. \tag{4.73}$$

Where C_{iss} is defined as the input capacitance with the output shorted.

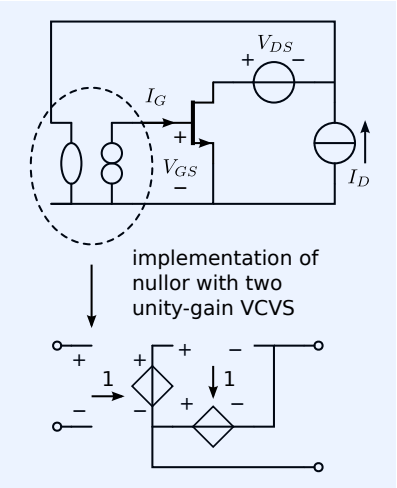


Figure 4.24: Simulation test bench for determination of the small-signal parameters.

Operating point information

Similar as with the bipolar transistor, the current drive capability and the voltage drive capability of a JFET depend on the DC drain current I_D and the drain-source voltage V_{DS}, respectively. Many other performance aspects, such as, the noise performance, and the cut-off frequency also show a direct relation with the drain current. As a consequence, we often want to fix the operating point by means of fixing I_D and V_{DS}.

A method for fixing the operating point of nonlinear resistive multi-terminal devices has been discussed in Chapter 3. According to the presented method, fixing the operating point of a JFET by means of I_D and V_{DS}, requires the addition a voltage source V_{DS} between the drain and the output, and a current source I_D that flows from the source to the drain.

In order to obtain zero output voltage and zero output current for all DC input and output terminations, a voltage source V_{GS} has to be placed in series with the gate, and a current source I_G has to be connected in parallel with the gate-source junction. The values of these sources depend on the required values of I_D and V_{DS}, on the DC characteristics of the device and on the operating temperature. They can be determined with the aid of the circuit from Figure 4.24. The nullator at the output port (drain-source) sets the condition for zero output voltage and zero output current, while the norator at the input port delivers the correct driving quantities to satisfy these conditions. Although the nullor is not available in SPICE, it can be implemented with the aid of two unity-gain voltage-controlled voltage sources as illustrated in this figure.

SPICE returns the small-signal parameters in a certain operating point as the result of an operating point analysis (`.op` statement). The parameters returned depend on the device model and the SPICE version. Table 4.5 gives an overview of the small-signal parameters of the JFET returned by different SPICE versions.

name	**description** (see Figure 4.20)	`LTspice`	`SIMetrix`	`Pspice`	`ngspice`
c_{gd}	Gate-drain capacitance	`Cgd`	`CGDt`	`CGD`	
c_{gs}	Gate-source capacitance	`Cgs`	`CGSt`	`CGS`	
g_m	Forward transconductance	`Gm`	`GM`	`GM`	`gm`
g_π	Input conductance				`ggs`
g_{ds}	Output conductance	`Gds`		`GDS`	`gds`
I_G	DC gate current		`IG`		`ig`
I_D	DC drain current	`Id`	`ID`	`ID`	`id`
I_S	DC source current	`-`	`IS`		`is`
I_{GD}	DC drain-gate current				`igd`
P	DC power dissipation	`-`	`Power`		
V_{GS}	DC gate-source voltage	`Vgs`		`VGS`	`vgs`
V_{GD}	DC gate-drain voltage				`vgd`
V_{DS}	DC drain-source	`Vds`		`VDS`	`-`

Table 4.5: Small-signal parameters of the JFET, returned from an operating point analysis with various simulators.

SIMETRIX does not output data for the output conductance g_{ds} and the operating voltages. However, the latter can be obtained from the nodal voltages. NGSPICE does not provide data for the small-signal capacitances.

4.3.4 Simplified models for hand calculation

The complete SPICE models are suitable for numerical simulations but are too complex to provide design information. For this purpose we need simplified models that are suitable for hand calculations. In this section we will derive such models. We will introduce a large signal static model, a small-signal dynamic model and a noise model that can be used for determination of the operating point, the dynamic small-signal transfer and the noise behavior of the JFET, respectively. We will only discuss the active forward saturation region.

DC behavior

The instantaneous large-signal model that is suitable for hand calculations is depicted in Figure 4.25. The diodes, the capacitances and the bulk resistances have been omitted and only the nonlinear voltage-controlled current source remains. The current of the controlled source depends on both the gate-source and the drain-source voltage. Different expressions must be used for the cut-off region, the linear region and the saturated region.

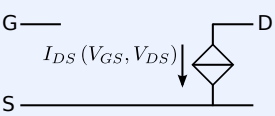


Figure 4.25: Simplified JFET DC model for hand calculations.

1. Forward mode ($V_{DS} > 0$), cut-off region: $V_{GS} - \text{VTO} < 0$

$$I_{DS} = 0. \tag{4.74}$$

2. Forward mode ($V_{DS} > 0$), linear region: $V_{DS} < V_{GS} - \text{VTO}$, or, alternatively: $V_{DG} < -\text{VTO}$

$$I_{DS} = \text{AREA.BETA}\,(1 + \text{LAMBDA}V_{DS})\,V_{DS}2\,(V_{GS} - \text{VTO}) - V_{DS}). \tag{4.75}$$

3. Forward mode ($V_{DS} > 0$), saturation region: $0 \leq V_{GS} - \text{VTO} \leq V_{DS}$, or, alternatively: $V_{DG} > -\text{VTO}$

$$I_{DS} = \text{AREA.BETA}\,(1 + \text{LAMBDA}V_{DS})\,(V_{GS} - \text{VTO})^2\,. \tag{4.76}$$

Small-signal dynamic behavior

The simplified small-signal dynamic model is shown in Figure 4.26.

Table 4.6: JFET small-signal model parameters for given device model parameters and operating point.

param.	cut-off.	linear region	saturation region
C_{gs}	AREA.CGS	see expression 4.59	see expression 4.59
C_{ds}	AREA.CDS	see expression 4.61	see expression 4.61
g_m	0	$2\text{AREA.BETA}V_{DS}$	$2\text{AREA.BETA}\,(V_{GS} - \text{VTO})$ $= 2\sqrt{\text{AREA.BETA}I_{DS}}$
g_{ds}	0	$2\text{AREA.BETA}(V_{GS} - \text{VTO})$ voltage controlled resistor; approx. for $V_{DS} = 0$	$\text{LAMBDA}I_{DS}$ (LAMBDA $\ll 1$)

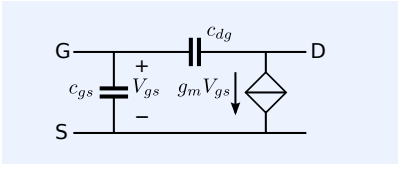


Figure 4.26: Simplified hybrid π small-signal equivalent circuit of the JFET.

The parameters can be obtained from a SPICE operating point analysis. They can also be estimated from the device parameters and the DC operating voltages and currents. The equations are shown in Table 4.6.

Noise model

A simplified noise model is shown in Figure 4.27. The noise of the bulk resistances R_s and R_d is omitted, while the remaining noise sources are added to the simplified hybrid-π equivalent circuit from Figure 4.26.

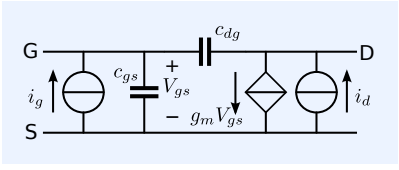


Figure 4.27: Simplified small-signal equivalent circuit of a JFET with noise sources.

Determination of hybrid-π parameters from simulation

The small-signal parameters can also be obtained from simulation. Figure 4.17 shows the simulation test bench for determination of the small-signal hybrid-π parameters.

The procedure is as follows:

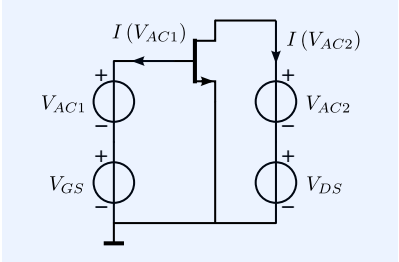


Figure 4.28: Simulation test bench for determination of the hybrid-π equivalent circuit parameters.

1. Bias the JFET in the required operating point with the aid of the *DC* voltage sources V_{GS} and V_{DS}. The value of V_{GS} can be obtained from an operating point simulation with the test bench from Figure 4.24.

2. Make $V_{AC1} = 1,0$ (magnitude 1, phase 0) and $V_{AC2} = 0$ and perform an AC analysis over the frequency range of interest

3. Obtain approximations the values of the following parameters

$$g_m = \text{Re}\{-I(V_{AC2})\}, \tag{4.77}$$

$$c_{gs} + c_{ds} = \frac{\text{Im}\{-I(V_{AC1})\}}{2\pi f}. \tag{4.78}$$

4. Make $V_{AC1} = 0$ and $V_{AC2} = 1,0$ (magnitude 1, phase 0)

5. Obtain approximations for the values of the following parameters

$$r_o = \frac{1}{\mathrm{Re}\{-I(V_{AC2})\}}, \tag{4.79}$$

$$c_{ds} = \frac{\mathrm{Im}\{I(V_{AC1})\}}{2\pi f}. \tag{4.80}$$

4.4 MOS transistors

The Metal Oxide Semiconductor Transistor was invented by Kahng and Dawon in 1959: [KahngDawon1960][23]. The first CMOS circuit has been reported in 1963 by Wanlass and Sah (see [Wanlass-Sah1963][24]).

MOS Transistors are field effect devices with their gate isolated from the channel by SiO_2. Aluminium and polysilicon are commonly used as gate materials. Depending on their application MOS transistors are fabricated in different ways:

1. Standard IC CMOS devices are lateral devices with a symmetrical structure: the source and the drain can be interchanged.

2. Modern IC Fin FET devices have a three dimensional gate structure that almost completely surrounds the channel.

3. High-voltage MOS transistors can be fabricated as lateral devices or as vertical devices. They are strongly asymmetrical: the low-voltage gate-source structure differs from the high-voltage gate-drain structure.

MOS transistors found their first application in digital circuits. Although bipolar and BiCMOS[25] IC processes are still in use, high-volume analog ICs and mixed signal ICs[26] are nowadays predominantly fabricated in CMOS technology.

Power MOS transistors are available in high-voltage IC processes and as discrete devices. They are mostly used as power switches in switched regulators and switched amplifiers.

Figure 4.29 shows a simplified cross section of P-substrate C-MOS (complementary MOS) process as well as the device symbols for NMOS and PMOS transistors.

[23] Kahng and Dawon. Electric Field Controlled Semiconductor Device, May 1960

[24] F. "Wanlass and C." Sah. Nanowatt logic using field-effect metal-oxide semiconductor triodes. *1963 IEEE International Solid-State Circuits Conference. Digest of Technical Papers*, VI:32–33, 1963

[25] Combined bipolar and CMOS IC technology.

[26] Mixed-signal: analog and digital signal processing combined in one IC.

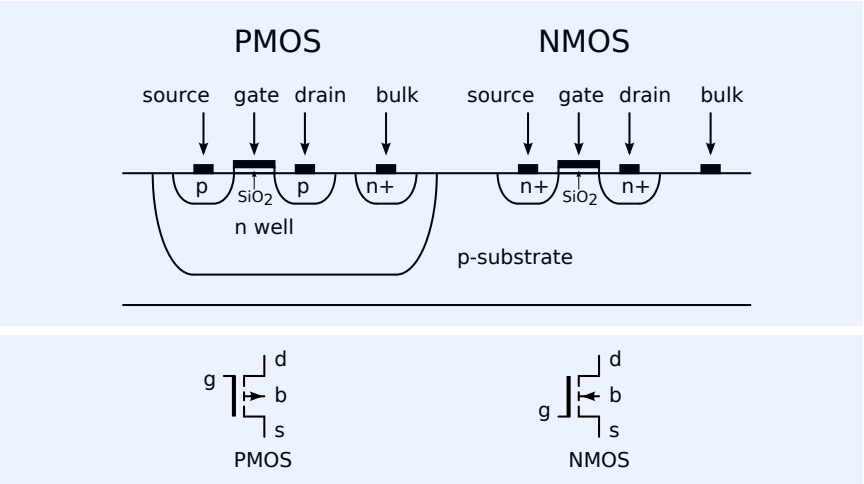


Figure 4.29: Cross section of a standard CMOS integrated circuit process.

4.4.1 Operation

We will describe the basic MOS operation for the NMOS from Figure 4.29. We will assume that the source and the bulk (substrate) of the NMOS have been connected to a reference potential (0V), while the drain has been connected to a fairly large positive voltage V_{DS}. For this set up, we will briefly describe the relation between the drain-source current I_{DS} and the gate-source voltage V_{GS}.

With $V_{GS} = 0$ the potential at the interface between the oxide and p-doped substrate is defined by the built-in junction voltage and the oxide capacitance.

At small values of V_{GS}, due to capacitive coupling with the gate, this so-called *surface potential* almost linearly increases with V_{GS} and the source starts injecting electrons in the p-region. The electrons injected in the p-region start to form an n-type channel under the gate and we speak of *weak inversion*. These electrons reach the accelerating electrical field of the drain depletion layer through diffusion. They are collected at the drain terminal. As a result the $I_{DS}\,(V_{GS})$ characteristic shows an exponential relationship comparable with that of the $I_C\,(V_{BE})$ relationship of a lateral bipolar NPN transistor.

At increasing values of V_{GS}, the channel extends further under the gate, and the surface potential does no longer increase significantly with the gate-source voltage. The drain current now increases quadratically with the difference between V_{GS} and the so-called *threshold voltage* V_T, just as with the JFET. The device now operates in *strong inversion*. The threshold voltage V_T is defined as the gate-source voltage at which the transition from weak inversion to strong inversion takes place.

As long as the *effective gate-source* voltage $V_{GS} - V_T$ is below V_{DS}, the channel does not extend to the drain. In this *pinch-off* or *saturation* region the drain-source current only shows a minor dependency of V_{DS} due to the *channel length modulation*, an effect comparable to the Early effect in bipolar transistors.

If V_{GS} increases to a level of V_T above V_{DS} the channel extends from the source to the drain and I_{DS} becomes strongly dependent on V_{DS} as well. The transistor then acts as a voltage-controlled resistor and we speak of the *linear operating* region.

Small geometry effects

The basic operation described above holds for large MOS transistors. However, with shrinking dimensions in CMOS IC processes, small geometry effects have to be taken into account. Below a listing of the most important small-geometry effects.

1. With shrinking oxide thickness, the influence of the vertical component of the electrical field in the channel on the charge carrier mobility cannot longer be ignored. This field influences the vertical current density profile. A high vertical field causes a reduction of the mobility of the electrons moving horizontally from the source to the drain.

2. A shrinking channel length also results in an increase of the lateral electrical field. At high field strengths the velocity of the charge carriers saturates.

3. The electrical field in the gate increases with shrinking oxide thickness. This causes a slow change of charge storage in the oxide over time and results in a slow change of V_T over time. This effect may limit the operational life time of CMOS ICs.

4. In large MOSFETs the channel length modulation is the main contributor to the V_{DS} dependency of I_{DS}. In short-channel MOSFETs the depletion

region of the drain reaches far and deep under the channel and reduces the channel to bulk depletion capacitance under the channel. This improves the coupling between the gate-source voltage and the surface potential. This *drain-induced barrier lowering* (DIBL) effect causes V_T to decrease with V_{DS}. As a result I_{DS} increases with V_{DS}.

4.4.2 MOSFET modeling

MOS transistors found their first application in digital circuits. Adequate simulation of these digital circuits could be preformed with relatively simple simulation models. Nowadays MOS transistors are also applied in analog integrated circuits, which requires an accurate description for all possible operating conditions.

The behavior of MOS transistors strongly depends on both their geometry and their manufacturing technology. The rapid development of new IC technologies and the reduction of dimensions continuously present new challenges to MOSFET modeling. For a complete description of MOSFET models including references to publications about underlying physical mechanisms the reader is referred to literature. Foty describes the following models(see [Foty1997][27]):

1. The Level 1 model is usually referred to the model presented by Shichman and Hodges (see [ShichmanHodges1968][28]). This is a relatively simple first generation MOS model for thick-oxide, long and wide channel MOSFETs, operating in strong inversion. It can be used if simulation speed is more important than accuracy. Due to its low complexity, this model can also be used as a basis for hand calculations.
2. The Level 2 model is also a first generation model. It is based on the Level 1 model. Corrections are in the form of additional equations that account for small-geometry effects. They make the model mathematically complex and inefficient.
3. The Level 3 model is also based on the Level 1 model, but the added equations that account for small-geometry effects have a more empirical character. They use less parameters and are more simple than those of the Level 2 model. The Level 3 model also includes basic modeling of operation in weak inversion.
4. The BSIM (Berkeley Short-Channel IGFET Model, see [SheuEtAll1987][29]) is a second generation MOSFET model. The model has its focus on fast and robust simulation rather than on a physical basis.
5. The HSPICE Level 28 model is a second generation model based on the BSIM model (see [HSPICE1993][30]). Its improvements strongly facilitated analog CMOS circuit design. It is a proprietary model developed by Meta-Software.
6. The BSIM2 model is also a second generation MOSFET model. It is an improvement of the BSIM model developed by the university of Berkeley.
7. The BSIM3 model is a third generation model (see [BSIM3-1995][31]). Second generation models are constructed with an extrinsic model on top of the intrinsic model structure. The third generation BSIM3 model has its geometry parameters included the intrinsic model.
8. The MOS Model 9 is a third generation MOS model developed at Philips Laboratories that has been made generally available (see [MOS9][32]). The modeling approach differs from the approach taken in the development of the BSIM3 model. The MOS Model 9 still distinguishes an intrinsic and an extrinsic structure, but it has improved mathematical modeling.

[27] Daniel P. Foty. *MOSFET modeling with SPICE: principles and practice*. Prentice-Hall, Inc., USA, 1997. ISBN: 0-13-227935-5

[28] H. Shichman and D. Hodges. Modeling and simulation of insulated-gate field-effect transistor switching circuits. *IEEE J. Solid-State Circuits*, 3(3):285–289, September 1968

[29] B. Sheu, D. Scharfetter, P. Ko, and M. Jeng. BSIM: Berkely Short Channel IGFET Model for MOS Transistors. *IEEE Journal of Solid State Circuits*, SC-22:558–566, 1987

[30] *HSPICE User's Manual*. Campbell, California, 1993

[31] Cheng, Y. et al. *BSIM3 Version 3.0 Manual*. University of California/Berkeley, Electronics Research Laboratory, 1995

[32] D. Velghe, R. Klaassen and F. Klaassen. MOS Model 9, 1994. Unclassified Report NL-UR 003/94

9. The EKV model (see [EnzVittoz2006][33]) has a physical base and is oriented on low-power analog design. It differs from other third generation models because it takes the bulk node as reference rather than the source node. Source and drain are treated equally, yielding a symmetrical model with smooth analytical equations for all operating regions. Hence, it is very well suited for analog design simulation.

10. The BSIM4 model is suited for analog CMOS design in technology nodes below 100nm (see [BSIM4-2017][34]).

[33] Christian C. Enz, and *Charge-based MOS Tran* Wiley & Sons Inc., 2006 85541-6

[34] Chetan Kumar Dabhi, Shivendra Singh Parihar, Harshit Agrawal, Navid Paydavosi, Tanvir Hasan Morshed, Darsen D. Lu, Wenwei(Morgan)Yang, Mohan V. Dunga, Xuemei (Jane) Xi, Jin He, Weidong Liu, Kanyu, M. Cao, Xiaodong Jin, Jeff J. Ou, Mansun Chan, Yogesh Singh Chauhan, Ali M. Niknejad, Chenming Hu. *BSIM4 4.8.1 MOSFET Model*. University of California/Berkeley, 2017

In the following sections only a selection of the above models will be discussed. Due to its relative simplicity, the Level 1 model can be used as a basis for hand calculations for devices working in strong inversion in the saturated and in the linear operating region. We will discuss the model in section 4.4.3.

The Level 1 model is not accurate enough to predict the behavior of analog CMOS circuits in modern sub micron technology. In modern IC design predominantly BSIM3, BSIM4 and MOS Model 9 are used. Because of their complexity they are not very well suited for hand calculations. For hand calculations of the voltage and current handling capabilities of MOS transistors, the Level 1 is still widely used. In section 4.4.6 we will confine ourselves to a brief discussion of small-signal models and model data obtained from simulation with some of these models.

More attention will be paid to the EKV model. The physical bases of this model with smooth analytical expressions that cover all operating conditions, makes them very well suited for deriving design information at early stages of the design. Binkley describes the application of the EKV model, and particularly, the use of the *inversion coefficient*, for taking design decisions (see [Binkley2008][35]).

[35] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

The symbolic simulator SLiCAP uses small-signal models of which the parameters are related to the device geometry and the inversion coefficient. Both the basics and the application of the EKV model will be discussed in section 4.4.7.

4.4.3 MOSFET level 1 model

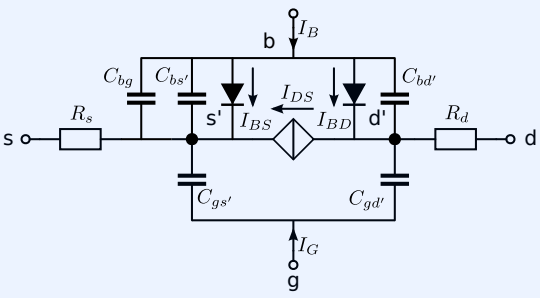


Figure 4.30: SPICE MOSFET model

Figure 4.30 shows the level 1 MOSFET model as it is implemented in SPICE. The model consists of a four-terminal intrinsic MOS transistor and two bulk resistances in series with the intrinsic transistor. Some SPICE versions support bulk resistances for the gate and the bulk connections. Some geometrical arguments can be given to the MOS model (see the SPICE reference manual). We will give an overview of these arguments and the parameters in section 4.4.4. The expressions are given for N-channel MOSFETs.

Instantaneous behavior

The gate and bulk currents for all operating conditions are given by the diode currents:

$$I_G = 0, \tag{4.81}$$
$$I_B = I_{BS} + I_{BD}, \tag{4.82}$$
$$I_{BS} = I_{bs}\left(\exp\frac{V_{b's'}}{U_T} - 1\right), \tag{4.83}$$
$$I_{BD} = I_{bd}\left(\exp\frac{V_{b'd'}}{U_T} - 1\right), \tag{4.84}$$

if JS = 0, AS = 0, or AD = 0 (AS and AD are arguments for the drain and source areas, respectively) we have

$$I_{bs} = \text{IS}, \tag{4.85}$$
$$I_{bd} = \text{IS}, \tag{4.86}$$

else:

$$I_{bs} = \text{AS.JS}, \tag{4.87}$$
$$I_{bd} = \text{AD.JS}, \tag{4.88}$$

where IS and JS represent the drain-bulk and source-bulk junction saturation current and current density, respectively. The bulk resistances can be modeled with geometrical arguments:

$$R_s = \text{RSH.NRS or } R_s = \text{RS/W}, \tag{4.89}$$
$$R_d = \text{RSH.NRD or } R_d = \text{RD/W}, \tag{4.90}$$

or without geometrical arguments:

$$R_s = \text{RS}, \tag{4.91}$$
$$R_d = \text{RD}. \tag{4.92}$$

The model equations of the controlled source I_{ds} differ for the various operating regions:

1. Normal mode ($V_{d's'} > 0$), cut-off region: $V_{gs'} - V_{to} < 0$; the cut-off voltage V_{to} depends on the bulk-source or "back-gate" voltage:

$$V_{to} = \text{VTO} + \text{GAMMA}\left(\sqrt{(\text{PHI} - V_{bs'})} - \sqrt{\text{PHI}}\right), \tag{4.93}$$

   while

$$I_{ds} = 0. \tag{4.94}$$

2. Normal mode ($V_{d's'} > 0$), linear region: $V_{d's'} < V_{gs'} - V_{to}$, or, alternatively: $V_{dg'} < -V_{to}$

$$I_{DS} = \frac{\text{W}}{\text{L}}\text{KP}\,(1 + \text{LAMBDA}.V_{d's'})\,V_{d's'}\left(\left(V_{gs'} - V_{to}\right) - \frac{V_{d's'}}{2}\right). \tag{4.95}$$

3. Normal mode ($V_{d's'} > 0$), saturation region: $0 \le V_{gs'} - V_{to} \le V_{d's'}$, or, alternatively: $V_{dg'} > -V_{to}$

$$I_{DS} = \frac{\text{W}}{\text{L}}\frac{\text{KP}}{2}\,(1 + \text{LAMBDA}.V_{d's'})\left(V_{gs'} - V_{to}\right)^2. \tag{4.96}$$

4. Reverse mode ($V_{d's'} < 0$)

For reverse mode operations the drain and the source connections are swapped.

If not given, the transconductance factor KP is obtained from geometrical input parameters and physical constants:

$$\text{KP} = \text{UO} \times 10^{-4} \frac{\epsilon_{ox}}{\text{TOX}} \text{ [A/V}^2\text{]}, \tag{4.97}$$

in which ϵ_{ox} is the permittivity of SiO_2, which is about 34×10^{-12}F/m. The factor 10^{-4} is a consequence of the units given to the surface mobility UO: cm^2/Vs. If we define C_{ox} as the oxide capacitance per unit of area:

$$C_{ox} = \frac{\epsilon_{ox} \text{W.L}}{\text{TOX}} \text{ [F]}, \tag{4.98}$$

the transconductance factor can alternatively be written as

$$\text{KP} = \text{UO} \times 10^{-4} C_{ox} \text{ [A/V}^2\text{]}. \tag{4.99}$$

In IC design manuals, the MOS transconductance factor KP is often denoted as β_{sq}. Modern low-voltage MOS processes can have β_{sq} values larger than 100×10^{-6}A/V^2.

Dynamic effects

The dynamic behavior of the MOS transistor is modeled with the capacitances from Figure 4.30. The input capacitance $C_{gs'}$ is the sum of the gate-source overlap capacitance C_{gso} and a portion of the oxide capacitance. The latter one depends on the operating mode of the MOS. In the linear operating range it is $\frac{1}{2}C_{ox}$ and in the saturation region it amounts $\frac{2}{3}C_{ox}$.

The capacitance $C_{gd'}$ also depends on the mode of operation. In the linear operating range it is $\frac{1}{2}C_{ox} + C_{gdo}$ and in the saturation range it equals the gate-drain overlap capacitance C_{gdo}. The gate-source overlap capacitance C_{gso} and the gate-drain overlap capacitance C_{gdo} depend on the device width:

$$C_{gso} = \text{CGSO.W}, \tag{4.100}$$

$$C_{gdo} = \text{CGDO.W}. \tag{4.101}$$

The gate-bulk capacitance C_{gb} also depends on the operating mode. In this model it is determined by the bulk-gate overlap capacitance:

$$C_{gb'} = \text{CGBO.L}. \tag{4.102}$$

The source-bulk and drain-bulk depletion capacitances depend on the junction voltages:

$$C_{b's'} = \frac{\text{CJ.AS}}{\left(1 - \frac{V_{bs'}}{\text{PB}}\right)^{\text{MJ}}} + \frac{\text{CJSW.PS}}{\left(1 - \frac{V_{bs'}}{\text{PB}}\right)^{\text{MJSW}}}; \quad V_{bs'} \leq \text{FC.PB}, \tag{4.103}$$

$$C_{b'd'} = \frac{\text{CJ.AD}}{\left(1 - \frac{V_{bd'}}{\text{PB}}\right)^{\text{MJ}}} + \frac{\text{CJSW.PD}}{\left(1 - \frac{V_{bd'}}{\text{PB}}\right)^{\text{MJSW}}}; \quad V_{bd'} \leq \text{FC.PB}. \tag{4.104}$$

If no geometrical arguments are given, W and L have their default values and the model equations for these capacitances are:

$$C_{b's'} = \frac{\text{CBS}}{\left(1 - \frac{V_{bs'}}{\text{PB}}\right)^{\text{MJ}}}; \quad V_{bs'} \leq \text{FC.PB}, \tag{4.105}$$

$$C_{b'd'} = \frac{\text{CBD}}{\left(1 - \frac{V_{bd'}}{\text{PB}}\right)^{\text{MJ}}}; \quad V_{bd'} \leq \text{FC.PB}. \tag{4.106}$$

Small-signal equivalent circuit

The small-signal equivalent circuit of the MOSFET is shown in Figure 4.31. It consists of the five capacitances, two linear controlled sources and the bulk resistances.

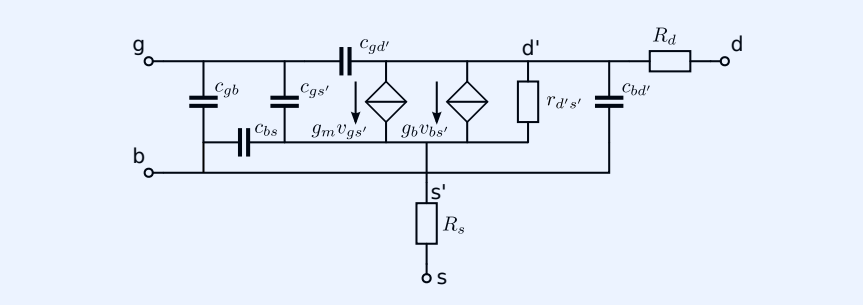


Figure 4.31: Small-signal equivalent circuit of a MOS transistor.

The transconductances of the controlled sources are defined as

$$g_m = \frac{\partial I_{DS}}{\partial V_{gs'}}, \tag{4.107}$$

$$g_{mb} = \frac{\partial I_{DS}}{\partial V_{bs'}}. \tag{4.108}$$

The capacitances are according to the definitions above. The output resistance $r_{d's'}$ is defined as

$$r_{d's'} = \frac{\partial V_{d's'}}{\partial I_{DS}}. \tag{4.109}$$

The values of the bulk resistances are obtained according to expression 4.89 to 4.92.

Stationary noise model

A frequency domain noise analysis can be performed in SPICE. For this purpose, stationary noise sources are added to the small-signal equivalent circuit from Figure 4.31. This stationary noise model is shown in Figure 4.32.

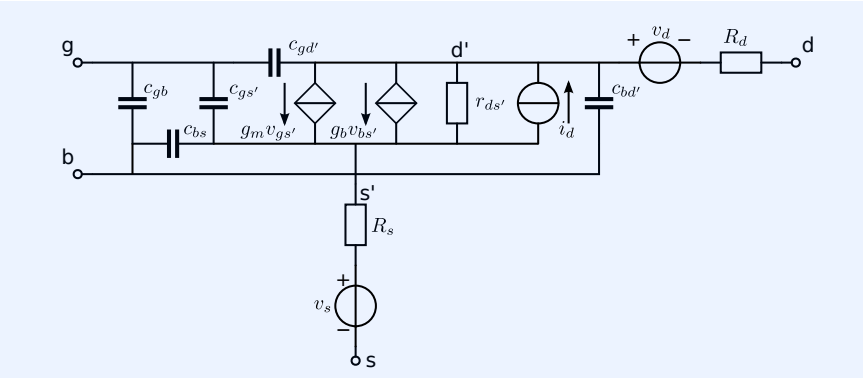


Figure 4.32: Small-signal equivalent circuit of a MOS transistor with added noise sources.

The thermal noise of the two bulk resistors R_s and R_d is given by their voltage spectral densities S_{vs} and S_{vd}, respectively:

$$S_{vs} = 4kTR_s, \tag{4.110}$$

$$S_{vd} = 4kTR_d. \tag{4.111}$$

The noise related to the channel current shows a $\frac{1}{f}$ component:

$$S_{id} = 4kTg_m\frac{2}{3} + \frac{\text{KF}}{f}I_{ds_Q}^{\text{AF}}. \tag{4.112}$$

This can also be denoted as

$$S_{id} = 4kTg_m\frac{2}{3}\left(1 + \frac{f_\ell}{f}\right), \tag{4.113}$$

$$f_\ell = \frac{3\text{KF}}{8kTg_m}I_{ds_Q}^{\text{AF}}, \tag{4.114}$$

where f_ℓ is the corner frequency for the $\frac{1}{f}$ noise. The cut-off frequency strongly depends on the technology and can exceed 1 [MHz].

4.4.4 Device parameters

Geometrical scaling factors

Table 4.7 shows the geometrical scale factors for MOS devices.

Table 4.7: MOS device model equations scale factors.

name	description	unit	default
W	channel width	m	10^{-4}
L	channel length	m	10^{-4}
AD	drain area	m^2	0
AS	source area	m^2	0
PD	perimeter of drain junction	m	0
PS	perimeter of source junction	m	0
NRD	number of squares drain diffusion	-	0
NRS	number of squares source diffusion	-	0

Model parameters instantaneous behavior

The MOS model parameters that describe the instantaneous behavior are listed in Table 4.8.

Table 4.8: MOS level 1 model parameters that describe the instantaneous behavior of the device

name	description	unit	default	note
VTO	pinch-off voltage	V	0	1
KP	transconductance factor	A/V^2	2×10^{-5}	1
LAMBDA	channel length modulation coeff.	1/V	0	
PHI	surface potential	V	0.6	1
GAMMA	bulk threshold parameter	$V^{1/2}$	0.0	1
RD	drain bulk resistance	Ω	0	1,2
RS	source bulk resistance	Ω	0	1,2
RSH	drain and source sheet resistance	$\Omega/\square$		
IS	bulk junction saturation current	A	10^{-14}	1
JS	bulk junction sat. current density	A/m^2	0	
TNOM	nominal temperature	^oC	27	
note 1	overrides results from process and geometry input			
note 2	either resistance, or resistance per unit of channel width			

Model parameters dynamic behavior and noise

Additional parameters that are required for the description of a MOSFET's dynamic behavior are listed in Table 4.9.

Process parameters

The process parameters are used to calculate the instantaneous model parameter values. For the relations between the process parameters and the instantaneous model parameters we refer to the literature. These parameters are listed in Table 4.10.

name	description	unit	default
CGSO	gate-source overlap capacitance per meter channel width	F/m	0
CGDO	gate-drain overlap capacitance per meter channel width	F/m	0
CGBO	gate-bulk overlap capacitance per meter channel length	F/m	0
CBD	zero-bias bulk-drain capacitance	F	0
CBS	zero-bias bulk-source capacitance	F	0
CJ	zero-bias bulk bottom capacitance per area	F/m^2	0
PB	built-in gate junction potential	V	0.8
MJ	bulk junction bottom grading coefficient		
MJSW	bulk junction sidewall grading coefficient		
CJSW	zero-bias bulk-junction sidewall capacitance		
FC	forward bias depletion capacitance coefficient	V	0.5
AF	flicker noise exponent	-	1
KF	flicker-noise coefficient	-	0

Table 4.9: Parameters for modeling the dynamic behavior

Cut-off frequency

The cut-off frequency f_T of a MOSFET is defined as the unity-gain frequency of the current gain factor:

$$f_T = \frac{g_m}{2\pi C_{iss}}, \tag{4.115}$$

where C_{iss} is the total input capacitance with shorted output:

$$C_{iss} = C_{gs} + C_{dg} \tag{4.116}$$

name	description	unit	default
TOX	gate oxide thickness	m	10^{-7}
NSUB	substrate doping	$1/cm^3$	0
NSS	surface state density	$1/cm^2$	0
TPG	type of gate material:	-	1
	+1: poly silicon, type opposite to substrate		
	-1: poly silicon, type same as substrate		
	0: Al gate		
LD	lateral diffusion	m	0
UO	surface mobility	cm^2/Vs	600
UTRA	transverse field coefficient	-	0

Table 4.10: MOS process parameters

Operating point information

Similar as with the BJT and the JFET, the current drive capability and the voltage drive capability of a MOS transistor depend on the DC drain current I_D and the drain-source voltage V_{DS}. Many other performance aspects, such as the noise performance and the cut-off frequency, also show a direct relation with the drain current. As a consequence, we usually want to fix the operating point by means of fixing I_D and V_{DS}.

A method for fixing the operating point of nonlinear resistive multi-terminal devices has been discussed in Chapter 3. According to the presented method,

fixing the operating point of a MOSFET by means of I_D and V_{DS} requires the addition a voltage source V_{DS} between the drain and the output and a current source I_D that flows from the source to the drain.

In order to obtain zero output voltage and zero output current for all DC input and output terminations, a voltage source V_{GS} has to be placed in series with the gate and a current source I_G has to be connected in parallel with the gate-source junction. The values of these input sources depend on the required values of I_D and V_{DS}, on the DC characteristics of the device and on the operating temperature. They can be determined with the aid of the circuit from Figure 4.33. The nullator at the output port (drain-source) sets the condition for zero output voltage and zero output current, while the norator at the input port delivers the correct driving quantities to satisfy these conditions. Although the nullor is not available in SPICE, it can be implemented with the aid of two unity-gain voltage-controlled voltage sources as illustrated in this figure.

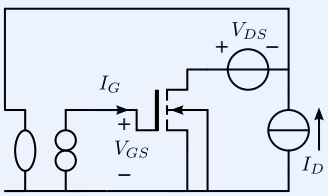


Figure 4.33: Simulation test bench for determination of the operating point information.

4.4.5 Simplified models for hand calculations

Complete SPICE models are suitable for numerical simulations but they are too complex to provide design information from analytical expressions. For this purpose we need simplified models that are suited for hand calculations. In this section we will derive such models. We will introduce a large signal static model, a small-signal dynamic model and a noise model that can be used for analytical determination of the operating point, the dynamic small-signal transfer and the noise behavior of a MOSFET. We will deal with the active forward region only.

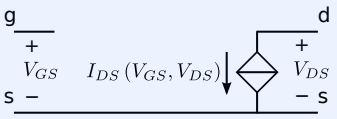


Figure 4.34: Simplified DC MOSFET model for hand calculations.

DC behavior

A DC model, suitable for hand calculations is depicted in Figure 4.34. The diodes, the capacitances and the bulk resistances have been omitted and only the nonlinear voltage-controlled current source remains. The current of this controlled source depends on both the gate-source and the drain-source voltage. The bulk is connected to the source, hence the threshold voltage V_{to} is not affected by V_{bs}. Different relations exist for the cut-off region, the linear region and the saturation region.

1. Normal mode $(V_{ds} > 0)$, cut-off region: $V_{gs} - \text{VTO} < 0$, we have

$$I_{ds} = 0. \tag{4.117}$$

2. Normal mode $(V_{ds} > 0)$, linear region: $V_{ds} < V_{gs} - \text{VTO}$, or, alternatively: $V_{dg} < -\text{VTO}$

$$I_{ds} = \frac{\text{W}}{\text{L}} \text{KP} V_{ds} \left(V_{gs} - \text{VTO}\right) \left(1 + \text{LAMBDA}.V_{ds}\right). \tag{4.118}$$

3. Normal mode $(V_{ds} > 0)$, saturation region: $0 \leq V_{gs} - \text{VTO} \leq V_{ds}$, or, alternatively: $V_{dg} > -\text{VTO}$

$$I_{ds} = \frac{\text{W}}{\text{L}} \frac{\text{KP}}{2} \left(V_{gs} - \text{VTO}\right)^2 \left(1 + \text{LAMBDA}.V_{ds}\right). \tag{4.119}$$

4. Reverse mode $(V_{ds} < 0)$

   For reverse operation, the drain and the source connections are swapped.

Small-signal dynamic behavior

The simplified small-signal dynamic model for a source-bulk connected MOSFET is shown in Figure 4.35.

The model parameters can roughly be determined from the Level 1 device model, the geometry parameters and the operating voltages and currents. These approximations are shown in Table 4.11.

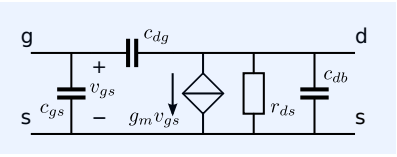


Figure 4.35: Simplified small-signal equivalent circuit for a MOSFET of which the source has been connected to the bulk.

par.	cut-off.	linear region	saturation region
C_{gs}	CGSO.W+CGBO.L+ $+\frac{W.L}{TOX}\epsilon_{ox}$ (max. value)	CGSO.W+CGBO.L+ $+\frac{1}{2}\frac{W.L}{TOX}\epsilon_{ox}$ (max. value)	CGSO.W+CGBO.L+ $+\frac{2}{3}\frac{W.L}{TOX}\epsilon_{ox}$ (max. value)
C_{gd}	CGDO.W	CGDO.W+ $+\frac{1}{2}\frac{W.L}{TOX}\epsilon_{ox}$ (max. value)	CGDO.W
C_{ds}	CJ.AD+CJSW.PD or CBD	CJ.AD+CJSW.PD or CBD	CJ.AD+CJSW.PD or CBD
g_m	0	$\frac{W}{L}$.KPV_{DS}	$\frac{W}{L}$.KP$(V_{DS} - VTO) =$ $= \sqrt{2\frac{W}{L}.KP I_{DS}}$
r_{ds}	∞	$\left[\frac{W}{L}.KP\,(V_{GS} - VTO)\right]^{-1}$ voltage controlled resistor approximation ($V_{DS} = 0$)	$1/(LAMBDA\ I_{DS})$ (LAMBDA $\ll 1$)

Table 4.11: MOSFET small-signal model parameters obtained from operating point and device model parameters

Noise model

The simplified noise model is shown in Figure 4.36. The bulk resistances R_s and R_d are omitted and the remaining noise sources are added to the simplified hybrid-π equivalent circuit from Figure 4.35.

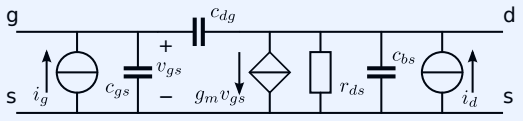


Figure 4.36: Simplified small-signal equivalent circuit with noise sources of a bulk-source connected MOSFET.

Determination of the hybrid-π parameters by a simulation

If the small-signal parameters of a MOSFET are not provided by the simulator's operating point analysis, they can be determined with the aid of a small-signal analysis. The procedure is similar to that of bipolar transistors as discussed in section 4.2.6. Figure 4.37 shows the simulation setup.

The procedure is as follows:

1. Bias the device in the required operating point with V_{GS} and V_{DS}
2. Add an AC signal (`AC 1 0`) to V_{AC1} only
3. Perform an AC analysis over the frequency range of interest
4. Obtain the following small-signal parameters

$$g_m = \mathrm{Re}\{I(V_{AC2})\}, \tag{4.120}$$

$$C_{gg} = C_{iss} = -\frac{\mathrm{Im}\{I(V_{AC1})\}}{2\pi f}, \tag{4.121}$$

$$C_{gb} = \frac{\mathrm{Im}\{I(V_{AC3})\}}{2\pi f}. \tag{4.122}$$

1. Now add the AC signal (`AC 1 0`) to V_{AC3} only

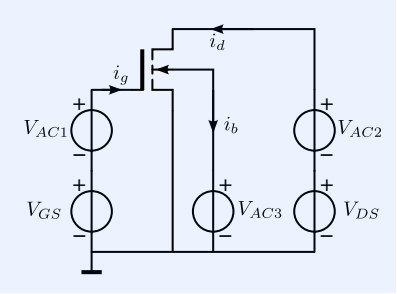


Figure 4.37: Simulation test bench for determination of the small-signal MOS parameters.

2. Perform an AC analysis over the frequency range of interest
3. Obtain the following small-signal parameters

$$g_{mb} = \mathrm{Re}\left\{I(V_{AC2})\right\}, \tag{4.123}$$

$$C_{bs} = -\frac{\mathrm{Im}\left\{I(V_{AC1})\right\}}{2\pi f}, \tag{4.124}$$

$$C_{bd} = -\frac{\mathrm{Im}\left\{I(V_{AC2})\right\}}{2\pi f}. \tag{4.125}$$

4. Now add the AC signal (`AC 1 0`) to V_{AC2} only
5. Perform an AC analysis over the frequency range of interest
6. Finally obtain the following small-signal parameters

$$g_o = \frac{1}{r_{ds}} = -\mathrm{Re}\left\{I(V_{AC2})\right\}, \tag{4.126}$$

$$C_{dg} = -\frac{\mathrm{Im}\left\{I(V_{AC1})\right\}}{2\pi f}. \tag{4.127}$$

In the following example, we will demonstrate the determination of the small-signal parameters of a biased NMOS that has its source connected to the bulk, by means of a SPICE simulation.

Example 4.1

Figure 4.38 shows the simulation test bench for the determination of the small-signal parameters of an NMOS transistor.

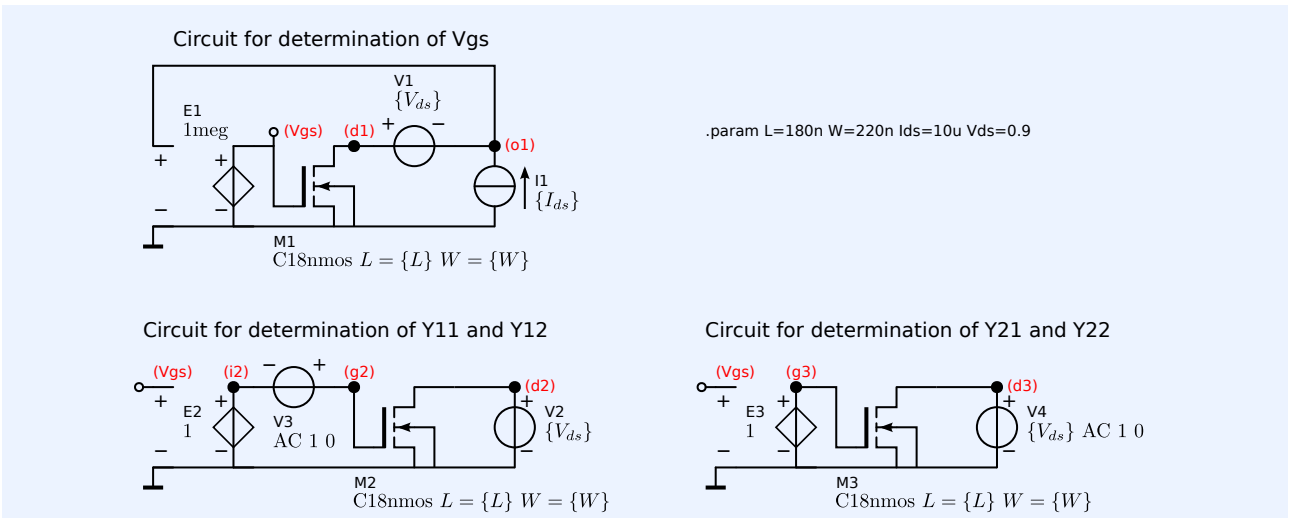


Figure 4.38: Simulation test bench for determination of the small-signal Y parameters of an NMOS transistor.

The netlist, including the LTSPICE `.measure` *statements for the determination of* g_m, g_o, c_{iss}, c_{oss}, *and* c_{dg}, *has been listed below:*

```
1  YparHyPiPar.cir
2  *LTspice netlist
3
4  *Circuit for determination of Vgs
5  M1 d1 Vgs 0 0 C18nmos L={L} W={W}
6  V1 d1 o1 {Vds}
7  I1 0  o1 {Ids}
8  E1 Vgs 0 o1 0 1k
9
10 *Circuit for determination of Y11 and Y12
```

```
M2 d2 g2 0 0 C18nmos L={L} W={W}
V2 d2 0 {Vds}
V3 g2 i2 AC 1 0
E2 i2 0 Vgs 0 1

*Circuit for determination of Y21 and Y22
M3 d3 g3 0 0 C18nmos L={L} W={W}
V4 d3 0 {Vds} AC 1 0
E3 g3 0 Vgs 0 1

.lib CMOS18TT.lib

.AC LIN 3 9.5Meg 10.5Meg

* Device parameters
.param W=220n L=180n

* Operating point
.param Vds=0.9 Ids=10u

* LTspice specific instructions for printing the small-signal parameters (at f
    =10MHz) in dB:
.meas AC g_m FIND Re(-I(V2)) AT 10MEG
.meas AC g_o FIND Re(-I(V4)) AT 10MEG
.meas AC c_iss FIND Im(-I(V3))/(2*pi*10meg) AT 10MEG
.meas AC c_oss FIND Im(-I(V4))/(2*pi*10meg) AT 10MEG
.meas AC c_dg FIND Im(I(E3))/(2*pi*10meg) AT 10MEG
```

The simulation results are printed in dB the LTspice *output file (error log). The results are listed in the table below:*

parameter	*dB value*	*value*	*units*
g_m	-80.9417	89.725	μAV^{-1}
g_o	-109.767	3.2483	μAV^{-1}
c_{iss}	-305.649	521.86	aF
c_{oss}	-316.932	142.36	aF
c_{dg}	-317.011	141.07	aF

From these values we find

$$c_{gs} = c_{iss} - c_{dg} = 380.79aF, \tag{4.128}$$

and

$$c_{ds} = c_{oss} - c_{dg} = 1.29aF. \tag{4.129}$$

4.4.6 Capacitance models

One of the challenges of the modeling of the behavior of semiconductor devices is to translate the dynamic spacial charge distribution in the physical device, into a lumped element network model in which charge is concentrated on capacitances, while the dynamic currents in the device terminals depend on the changes of these charges. By using these techniques, effects of a limited carrier velocity cannot be accounted for accurately.

For an in depth treatment of this subject, the reader is referred to literature. Foty (see [Foty1997][36]) addresses two topics related with the modeling with lumped elements: charge conservation and the reciprocity of the capacitive elements. In this section we will simply introduce two capacitance models for the intrinsic MOS transistor without discussion of the underlying physical mechanisms.

[36] Daniel P. Foty. *MOSFET modeling with SPICE: principles and practice*. Prentice-Hall, Inc., USA, 1997. ISBN: 0-13-227935-5

Meyer model

The model proposed by Meyer is commonly applied in the first generation MOS models (see [Meyer1971][37]). The model assumes reciprocal capacitances

[37] J. Meyer. MOS models and circuit simulation. *RCA Review*, 32:42–63, 1971

$\left(c_{dg} = c_{gd}\right)$. The model equations of the capacitances $c_{gs'}, c_{gd'}$ and c_{gb} differ for the linear region, the saturation region and the cut-off region.

The general implementation of the Meyer capacitance model is as follows:
In the linear region we have

$$c_{gs'} = \frac{2}{3}\text{W.L}C_{ox}\left(1 - \frac{\left(V_{gd'} - V_t\right)^2}{\left(V_{gs'} - V_t + V_{gd'} - V_t\right)^2}\right) + \text{CGSO.W}, \tag{4.130}$$

$$c_{gd'} = \frac{2}{3}\text{W.L}C_{ox}\left(1 - \frac{\left(V_{gs'} - V_t\right)^2}{\left(V_{gs'} - V_t + V_{gd'} - V_t\right)^2}\right) + \text{CGDO.W}, \tag{4.131}$$

$$c_{gb} = 0. \tag{4.132}$$

In the saturation region the model equations are:

$$c_{gs'} = \frac{2}{3}\text{W.L}C_{ox} + \text{CGSO.W}, \tag{4.133}$$

$$c_{gd'} = \text{CGDO.W}, \tag{4.134}$$

$$c_{gb} = 0. \tag{4.135}$$

In the cut-off region we have:

$$c_{gs'} = \text{CGSO.W}, \tag{4.136}$$

$$c_{gd'} = \text{CGDO.W}, \tag{4.137}$$

$$c_{gb} = \text{W.L}C_{ox} + \text{CGBO.L (maximum value at } V_{GB} = 0). \tag{4.138}$$

Ward-Dutton capacitance model

[38] D. Ward and R. Dutton. A Charge-Oriented Model for MOS Transistor Capacitances. *IEEE Solid-state Circuits*, sc-13(5):703–708, October 1978

The Ward-Dutton capacitance model (see [Ward1998][38]), uses a capacitance matrix that relates the four dynamic terminal currents of the intrinsic transistor to the nodal voltages at the corresponding terminals. The capacitances are calculated from the charge equations of the MOSFET model. The resulting capacitance matrix is not necessarily reciprocal.

The matrix equation is

$$\begin{pmatrix} i_g \\ i_d \\ i_s \\ i_b \end{pmatrix} = s \begin{pmatrix} C_{GG} & -C_{GD} & -C_{GS} & -C_{GB} \\ -C_{DG} & C_{DD} & -C_{DS} & -C_{DB} \\ -C_{SG} & -C_{SD} & C_{SS} & -C_{SB} \\ -C_{BG} & -C_{BD} & -C_{BS} & C_{BB} \end{pmatrix} \begin{pmatrix} v_g \\ v_d \\ v_s \\ v_b \end{pmatrix}, \tag{4.139}$$

where i_g, i_d, i_s and i_b are the terminal currents of the intrinsic transistor, and v_g, v_d, v_s and v_b the nodal voltages at the corresponding terminals.

If the source and the bulk are connected together, and the source-bulk connection is taken as the reference node, this model simplifies to

$$\begin{pmatrix} i_g \\ i_d \end{pmatrix} = s \begin{pmatrix} C_{GG} & -C_{GD} \\ -C_{DG} & C_{DD} \end{pmatrix} \begin{pmatrix} v_{gs} \\ v_{ds} \end{pmatrix}. \tag{4.140}$$

The resulting model can be written in a form of which the equivalent hybrid-π network can be constructed:

$$\begin{pmatrix} i_g \\ i_d \end{pmatrix} = s \begin{pmatrix} C_{GG} & -C_{GD} \\ (C_{GD} - C_{DG}) - C_{GD} & C_{DD} \end{pmatrix} \begin{pmatrix} v_{gs} \\ v_{ds} \end{pmatrix}. \tag{4.141}$$

The simplified hybrid-π equivalent circuit of the intrinsic transistor is

found after addition of the static transadmittance g_m and the static output resistance r_o. This circuit is shown in Figure 4.39.

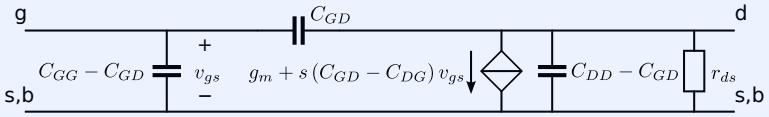


Figure 4.39: Hybrid-π equivalent circuit of the intrinsic MOSFET according to the Ward-Dutton capacitance model.

4.4.7 MOS EKV Model

In this section, we will give a brief summary of the construction and the use of the EKV MOS model. The model is developed by Enz and Vittoz: [EnzVittoz2006][39]. The application of the model and especially the use of the *inversion coefficient* as basis for early stage design decisions, is extensively described by Binkley (see [Binkley2008][40]). Expressions in this section involving the inversion coefficient have been taken from Binkley, but adapted in such a way that they can be used for estimation of the small-signal parameters of the hybrid-π equivalent model for all modes of operation.

[39] Christian C. Enz, and Eric A. Vittoz. *Charge-based MOS Transistor Modeling*. John Wiley & Sons Inc., 2006. ISBN: 978-0-470-85541-6

[40] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

Figure 4.40 illustrates the basic idea behind the static model of the intrinsic MOS: the drain-source current is the sum of a forward and a reverse component. Both components I_F and I_R have smooth expressions with a validity range that covers weak inversion through strong inversion, including velocity saturation. The bulk node has been taken as reference node.

The expressions for I_R equal those for I_F after swapping V_D and V_S. This makes the model symmetrical for the forward and reverse operation. The current equations are based upon the charge equations of the model.

In this section, we will briefly discuss the modeling of the $I_{DS}\,(V_{GS}, V_{DS})$ characteristic for the different operating regions, introduce the inversion coefficient and relate the parameters of the small-signal model to the operating point, the technology parameters and the device geometry.

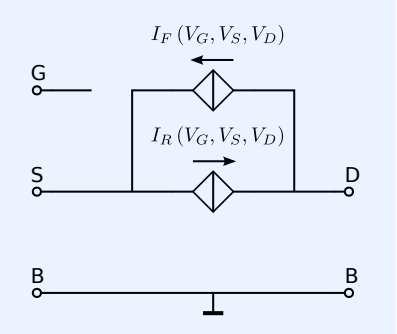


Figure 4.40: Basic static EKV model of the intrinsic MOS transistor.

The technology current

Binkley (see [Binkley2008][41]) defines the technology current I_0 as

$$I_0 \triangleq 2n\mu_0 C'_{OX} U_T^2 \text{ [A]}, \tag{4.142}$$

[41] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

where n is the substrate factor:

$$n = 1 + \frac{C'_{DEP}}{C'_{OX}} \text{ [-]}, \tag{4.143}$$

C'_{DEP} is the surface depletion capacitance and C'_{OX} is the oxide capacitance per unit of area:

$$C'_{OX} = \frac{\varepsilon_0 \varepsilon_r}{t_{ox}} \text{ [Fm}^{-2}\text{]}, \tag{4.144}$$

where ε_r is the relative permittivity of SiO_2, t_{ox} is the thickness of the gate oxide and U_T is the thermal voltage.

The reciprocal value of the substrate factor ($\frac{1}{n}$) models the coupling between the gate voltage and the surface potential at weak inversion.

The transconductance factor

The transconductance factor β_{sq} is defined as

$$\beta_{sq} = \mu_0 C'_{OX} \text{ [AV}^{-2}\text{]}, \tag{4.145}$$

where μ_0 is the low-field channel carrier mobility in [$m^2V^{-1}s^{-1}$]. The technology current can be expresses in β_{sq} as

$$I_0 = 2n\beta_{sq}U_T^2 \text{ [A]}. \tag{4.146}$$

Weak inversion

In weak inversion the drain-source current $I_{F,R}$ shows an exponential relation with the gate voltage:[42]

$$I_{F,R} = I_0 \frac{W}{L} \exp\left(\frac{\frac{V_G - V_{T0}}{n} - V_{S,D}}{U_T}\right) \text{ [A]}. \tag{4.147}$$

[42] Read this expression as folows:
$I_F = I_0 \frac{W}{L} \exp\left(\frac{\frac{V_G - V_{T0}}{n} - V_S}{U_T}\right)$
$I_R = I_0 \frac{W}{L} \exp\left(\frac{\frac{V_G - V_{T0}}{n} - V_D}{U_T}\right)$

The voltage V_{T0} is the *equilibrium threshold voltage*. For $V_S = 0$ it corresponds with the threshold voltage V_{to} of the models discussed above.

Strong inversion

When the device is operating in strong inversion, the drain-source current depends quadratically on the drive voltage:

$$I_{F,R} = \frac{W}{L}\frac{\beta_{sq}}{2n}\left(V_G - V_{T0} - nV_{S,D}\right)^2 \text{ [A]}, \tag{4.148}$$

where W and L are the effective width and length of the channel, respectively.

Weak inversion to strong inversion

With the aid of a transition function $F(x)$, the expressions for weak inversion and strong inversion can be combined into one:

$$F(x) = \left(\ln\left(1 + \exp\left(\frac{x}{2}\right)\right)\right)^2 \text{ [-]}, \tag{4.149}$$

$$= \begin{matrix} \exp(x) & \text{if } x \ll 0, \\ \left(\frac{x}{2}\right)^2 & \text{if } x \gg 0. \end{matrix} \tag{4.150}$$

This function returns the forward and the reverse inversion coefficient, IC_F and IC_R, respectively:

$$IC_{F,R} = F\left(\frac{V_G - V_{T0} - nV_{S,D}}{nU_T}\right) \text{ [-]}. \tag{4.151}$$

These coefficients are a measure for the level of inversion at which the transistor operates. An inversion coefficient much smaller than unity indicates weak inversion. An inversion coefficient much larger than unity indicates operation in strong inversion. Between weak and strong inversion we speak of moderate inversion. The actual forward and reverse current can be calculated from the technology current, the device geometry and their respective inversion coefficients as:

$$I_{F,R} = I_0 \frac{W}{L} IC_{F,R} \text{ [A]}, \tag{4.152}$$

or, alternatively:

$$I_{F,R} = 2n\beta_{sq}U_T^2 \frac{W}{L} IC_{F,R} \text{ [A]}. \tag{4.153}$$

Total drain-source current

The total drain-source current I_{DS} is the difference between the forward and the reverse current (see Figure 4.40):

$$I_{DS} = I_F - I_R \text{ [A]}. \tag{4.154}$$

When the transistor is operating in the saturation region, one of the current components I_R or I_F can be ignored with respect to the other.

Vertical field mobility reduction

The reduction of the mobility caused by the vertical field in the channel is modeled as reduction of the transconductance factor β_{sq} due to the gate-source voltage. A simple mobility reduction model uses the vertical field mobility reduction factor θ [V^{-1}]

$$\beta'_{sq} = \frac{\beta_{sq}}{1 + \theta \frac{V_G - V_{T0}}{n}} \text{ [AV}^{-2}\text{]}. \tag{4.155}$$

With this expression β'_{sq} increases below threshold and has a singularity at $\theta \frac{V_G - V_{T0}}{n} = -1$. This can be corrected by softly clipping the VFMR effect below threshold:

$$\beta'_{sq} = \frac{\beta_{sq}}{1 + 2\theta U_T \sqrt{IC_F}} \text{ [AV}^{-2}\text{]}. \tag{4.156}$$

Channel length modulation and velocity saturation

Channel length modulation and velocity saturation occur if the device is operating in the saturation region. In the SPICE EKV 2.6 model, the channel length modulation (CLM) and velocity saturation (VS) have been modeled through modification of the reverse current I_R. For a full SPICE implementation of the EKV2.6 model the reader is referred to the EKV2.6 manual (see [spiceEKV2.6][43]). Both the velocity saturation and the channel length modulation can be modeled in a relatively simple way by ignoring their absence in the linear operating region. This will provide sufficiently accurate values for taking early stage design decisions.

[43] Matthias Bucher, Christoff Lallement, Christian Enz, Fabien Théodoloz and François Krummenacher. The EPFL-EKV MOSFET Model Equations for Simulation. Technical report, Electronics Laboratory, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, July 1998

Similar as with bipolar transistors, the channel length modulation can be modeled with an Early voltage. The Early voltage V_A is assumed proportional with the length of the device

$$V_A = V_{AL} L \text{ [-]}, \tag{4.157}$$

where V_{AL} [Vm^{-1}] is the Early voltage per unit of length. After including the CLM in the expression of the drain-source current, I_{DS} changes to

$$I_{DS} = I_F \left(1 + \frac{V_D - V_S}{V_{AL} L}\right) - I_R \left(1 + \frac{V_S - V_D}{V_{AL} L}\right) \text{ [A]}. \tag{4.158}$$

Binkley (see [Binkley2008][44]) describes the modeling of velocity saturation in the forward saturation operating region through introduction of a second term in β_{sq} (4.156):

[44] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

$$\beta'_{sq} = \frac{\beta_{sq}}{1 + \left(\theta + \frac{1}{E_{CRIT} L}\right) 2U_T \sqrt{IC_F}} \text{ [AV}^{-2}\text{]}, \tag{4.159}$$

where E_{CRIT} [Vm^{-1}] is the value of the lateral field at which velocity saturation occurs.

At a later stage we will use the critical inversion coefficient IC_{CRIT}, which is defined as the inversion coefficient at which the reduction of the mobility

due to VFMR and velocity saturation sets in:

$$IC_{CRIT} \triangleq IC_{F,R}\Big|_{\left(\theta+\frac{1}{L\,E_{CRIT}}\right)\left(\frac{V_G-V_{T0}}{n}-V_{S,D}\right)=1} \text{ [-]}. \tag{4.160}$$

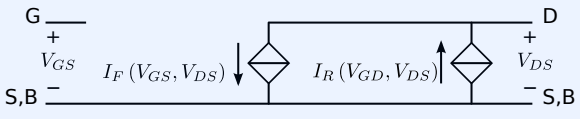
It can be approximated by

$$IC_{CRIT} \approx \frac{1}{\left(4nU_T\left(\theta+\frac{1}{L\,E_C}\right)\right)^2} \text{ [-]}. \tag{4.161}$$

Static device characteristics

Figure 4.41 shows the model with the source taken as reference node: $V_S = 0$ and $V_B = 0$.

Figure 4.41: EKV model with the Source and Bulk connected togeter and taken as reference node.



With the aid of this model we can find the $I_{DS}\,(V_{GS}, V_{DS})$ characteristic for the linear and saturation region at strong inversion and compare them with those of provided by the Level 1 model.

In the saturation region, I_R can be ignored. If we discard the velocity saturation and the channel length modulation we can simplify (4.158) to the well known quadratic relation:

$$I_{DS} = \frac{W}{L}\frac{\beta_{sq}}{2n}\left(V_G - V_{T0} - nV_S\right)^2 \text{ [A]}. \tag{4.162}$$

With $V_S = 0$, this equation simplifies to the one the Level 1 model uses for the saturation region:

$$I_{DS} = \frac{W}{L}\frac{\beta_{sq}}{2n}\left(V_{GS} - V_{T0}\right)^2 \text{ [A]}. \tag{4.163}$$

In the linear region, we cannot ignore the reverse current component and, under the above conditions, we may write

$$I_{DS} = \frac{W}{L}\frac{\beta_{sq}}{2n}\left(\left(V_G - V_{T0} - nV_S\right)^2 - \left(V_G - V_{T0} - nV_D\right)^2\right) \text{ [A]}. \tag{4.164}$$

If $V_S = 0$, this simplifies to the equation the Level 1 model uses for the linear region:

$$I_{DS} = \frac{W}{L}\frac{\beta_{sq}}{n}V_D\left(V_{GS} - V_{T0} - \frac{1}{2}V_{DS}\right) \text{ [A]}. \tag{4.165}$$

Figure 4.42 shows the static device characteristics of an NMOS transistor `myNMOS` ($W = 220$nm, $L = 180$nm) calculated according to the above (simplified) model with parameters of a standard CMOS18 process as listed by Binkley (see [Binkley2008][45]). The technology current for this device is 634nA and the critical inversion coefficient IC_{CRIT} equals about 41. Hence, velocity saturation starts to play a role at about 32μA. It should be noted that with this simple model of the channel length modulation and the velocity saturation, the characteristics are not accurate for short devices operating in the linear region.

[45] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

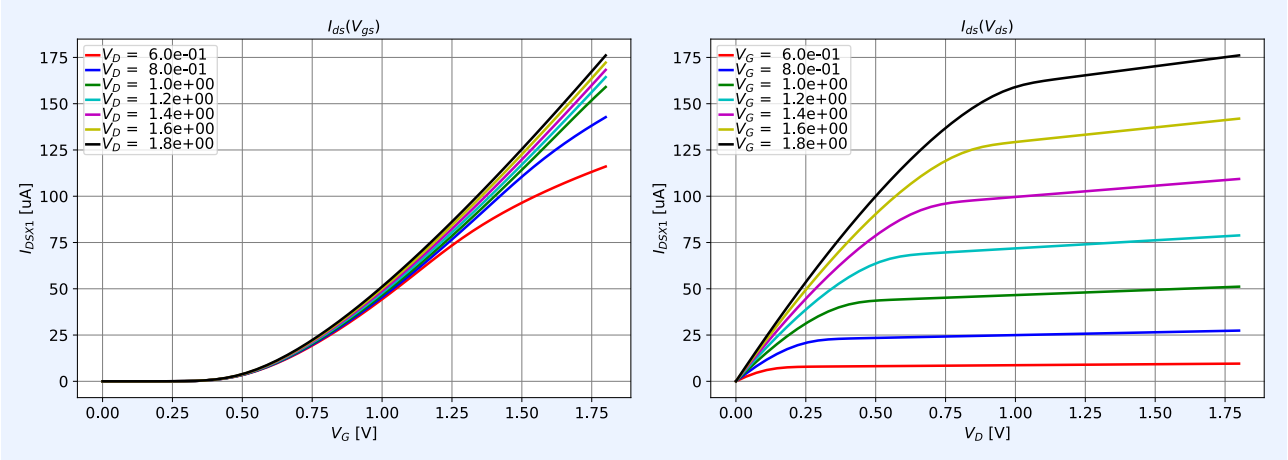


Figure 4.42: Device characteristics myN-MOS, CMOS18 process.
Left:
Transfer characteristics $I_{DS}\,(V_{GS}, V_{DS})$
Right:
Output characteristics $I_{DS}\,(V_{DS}, V_{GS})$

Linear and saturation region

The transition from the linear to the saturation region occurs at $V_{DS,sat}$:

$$V_{DS,sat} = 2U_T\sqrt{IC_F + 0.25} + 3U_T \text{ [V]}. \tag{4.166}$$

For reverse operation $V_{DS,sat}$ and IC_F should be replaced with $V_{SD,sat}$ and IC_R, respectively.

Small signal parameters

Figure 4.43 shows the small-signal model for the case that the source is taken as the reference node. Since the transconductance is defined at shorted output, we have $v_{gs} = -v_{ds}$.

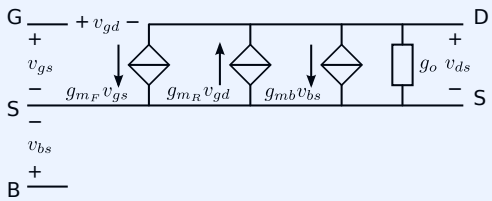


Figure 4.43: Static small-signal model of the intrinsic MOS transistor with the source taken as reference node according to the EKV model.

The forward small-signal transconductance g_m can be written as a function of the forward current and its associated inversion coefficient

$$g_{m_F} = \frac{I_F\left(1 + \frac{V_S - V_D}{V_{AL}L}\right)}{nU_T\sqrt{IC_F\left(1 + \frac{IC_F}{IC_{CRIT}}\right)} + 0.5\sqrt{IC_F\left(1 + \frac{IC_F}{IC_{CRIT}}\right)} + 1} \text{ [AV}^{-1}\text{]}. \tag{4.167}$$

The reverse transconductance g_{mR} can similarly be written as:

$$g_{m_R} = \frac{I_R\left(1 + \frac{V_D - V_S}{V_{AL}L}\right)}{nU_T\sqrt{IC_R\left(1 + \frac{IC_R}{IC_{CRIT}}\right)} + 0.5\sqrt{IC_R\left(1 + \frac{IC_R}{IC_{CRIT}}\right)} + 1} \text{ [AV}^{-1}\text{]}. \tag{4.168}$$

The total transconductance is found as the difference between the two:

$$g_m = g_{m_F} - g_{m_R} \text{ [AV}^{-1}\text{]}. \tag{4.169}$$

Figure 4.44 shows the transconductance as a function of the drain current of myNMOS.

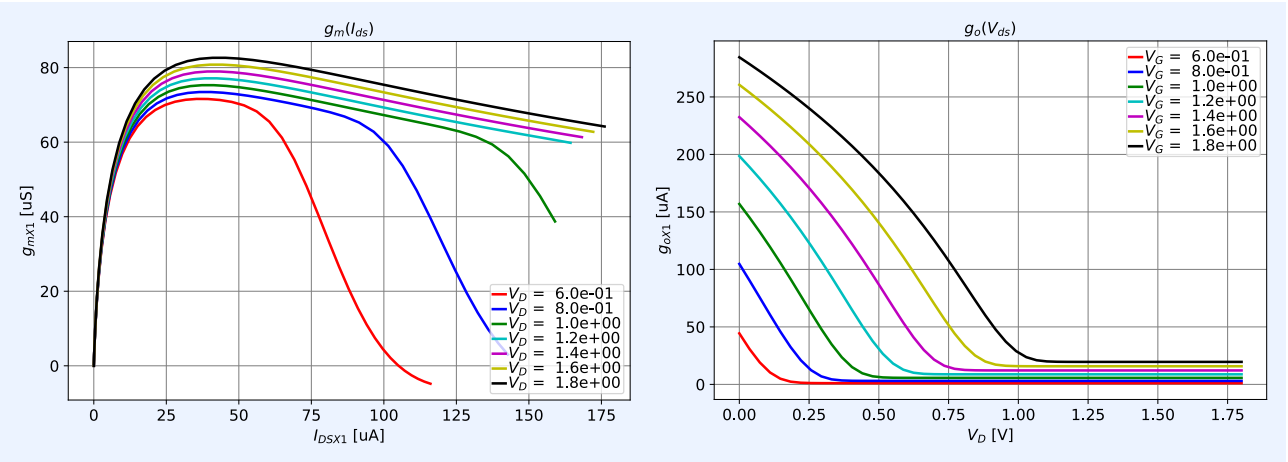


Figure 4.44: Device characteristics myNMOS, CMOS18 process.
Left: g_m calculated from I_{DS} and the inversion coefficient according to (4.169).
Right: g_ds calculated from I_{DS} and the inversion coefficient according to (4.172).

In the forward operating range, the finite small-signal output conductance g_{ds} is caused by the channel length modulation and by g_{m_R}:

$$g_{ds} = g_{o_F} + g_{m_R} \text{ [AV}^{-1}\text{]}. \tag{4.170}$$

In the saturation region, g_{mR} approximates zero and the output conductance is determined by the channel length modulation only. This part of the output conductance modeled in g_o:[46]

$$g_{o_F} = \frac{I_{DS}}{V_{DS} + V_{AL}L} \text{ [AV}^{-1}\text{]}. \tag{4.171}$$

In the linear region g_{ds} equals g_{mR}, however, expression (4.168) includes the effect of velocity saturation, which yields a too small value for the output conductance. A better estimate for g_{ds} in the forward linear operating range is:

$$g_{ds} = g_{o_F} + g_{m_R}\frac{1}{4}\sqrt{\frac{1}{4} + IC_R}. \tag{4.172}$$

Figure 4.44 also shows the output conductance of myNMOS as a function of the drain current.

The body effect simply follows from g_m and n as

$$g_{mb} = (n-1)\, g_m \text{ [AV}^{-1}\text{]} \tag{4.173}$$

[46] The effect of the drain induced barrier lowering has not been modeled.

Transconductance efficiency

The transconductance efficiency is defined as the ratio of the transconductance and the drain current. It is a measure for the transconductance produced per unit of drain current. The transconductance is used as a measure for the level of inversion. In weak inversion and in the saturation region, the transconductance efficiency has its largest value: [47]

$$\frac{g_m}{I_{DS}} = \frac{1}{nU_T} \text{ [V}^{-1}\text{]}. \tag{4.174}$$

In strong inversion (without short channel effects) and in the saturation region, it drops to:

$$\frac{g_m}{I_{DS}} = \frac{2}{V_{eff}} \text{ [V}^{-1}\text{]}, \tag{4.175}$$

[47] About 25.

where $V_{eff} = V_{GS} - V_{T0}$.

The transconductance efficiency of a MOS operating at an arbitrary inversion level in the forward saturation region, can be obtained from the inversion coefficient as:

$$\frac{g_m}{I_{DS}} = \frac{1}{nU_T\sqrt{IC_F\left(1+\frac{IC_F}{IC_{CRIT}}\right)}+0.5\sqrt{IC_F\left(1+\frac{IC_F}{IC_{CRIT}}\right)}+1} \text{ [V}^{-1}\text{]}. \quad (4.176)$$

[48] Matthias Bucher, Christoff Lallement, Christian Enz, Fabien Théodoloz and François Krummenacher. The EPFL-EKV MOSFET Model Equations for Simulation. Technical report, Electronics Laboratory, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, July 1998

Intrinsic capacitances

The EKV2.6 manual gives expressions for the intrinsic small-signal capacitances (see [spiceEKV2.6][48]). The small-signal dynamic model with these capacitances is shown in Figure 4.45. A complete small-signal model requires addition of the extrinsic capacitances and the series resistances in the drain, the source and the gate.

The intrinsic capacitances can be expressed as a part of the total oxide capacitance C_{OX}:

$$C_{OX} = WLC'_{OX} \text{ [F]} \quad (4.177)$$

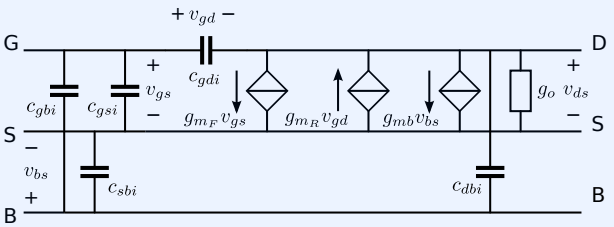


Figure 4.45: Dynamic small-signal model of the intrinsic MOS transistor with the source taken as reference node according to the EKV model.

The SPICE EKV2.6 model uses two parameters for calculation of these relative parts:

$$x_f = \sqrt{\frac{1}{4}+IC_F} \text{ [-]}, \quad (4.178)$$

$$x_r = \sqrt{\frac{1}{4}+IC_R} \text{ [-]}. \quad (4.179)$$

The intrinsic capacitances are

$$c_{gsi} = \frac{2}{3}\left(1-\frac{x_r^2+x_r+\frac{1}{2}x_f}{\left(x_r+x_f\right)^2}\right)C_{OX} \text{ [F]}, \quad (4.180)$$

$$c_{gdi} = \frac{2}{3}\left(1-\frac{x_f^2+x_f+\frac{1}{2}x_r}{\left(x_r+x_f\right)^2}\right)C_{OX} \text{ [F]}, \quad (4.181)$$

$$c_{gbi} = \frac{n-1}{n}\left(C_{OX}-c_{gsi}-c_{gdi}\right) \text{ [F]}, \quad (4.182)$$

$$c_{sbi} = (n-1)\,c_{gsi} \text{ [F]}, \quad (4.183)$$

$$c_{sdi} = (n-1)\,c_{gdi} \text{ [F]}. \quad (4.184)$$

In the linear region, c_{gbi} drops to zero because it is completely covered by the conductive channel. The value of c_{gbi} below accumulation is not modeled.[49]

[49] At very low values of V_{GS}, the intrinsic gate bulk capacitance equals C_{OX}.

The total input capacitance with shorted output c_{iss} is the sum of the three

capacitances. The cut-off frequency f_{Ti} of the intrinsic transistor is defined as:

$$f_{Ti} = \frac{g_m}{2\pi c_{iss}} \text{ [Hz]}. \tag{4.185}$$

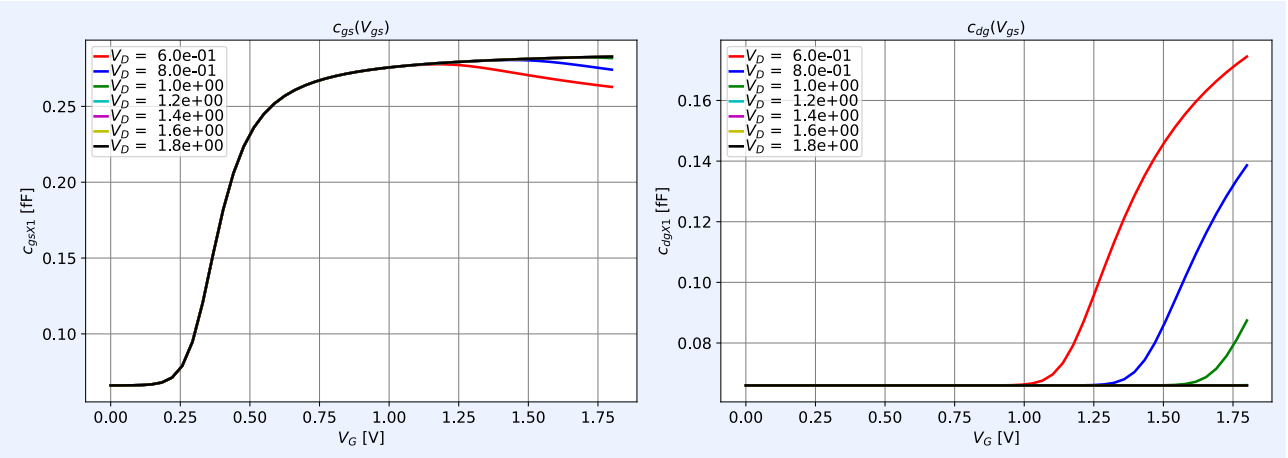


Figure 4.46: Device characteristics myNMOS, CMOS18 process.
Left: c_{gsi} plotted against V_{GS} with V_{DS} as parameter.
Right: c_{gdi} plotted against V_{GS} with V_{DS} as parameter.

Figure 4.46 shows the plots of c_{gsi} and c_{gdi} against I_{DS}, with V_{DS} as parameter. The plots
show that c_{gsi} rapidly increases from zero to $\frac{2}{3}C_{OX}$, but it drops to $\frac{1}{2}C_{OX}$, while c_{gdi} increases from zero to $\frac{1}{2}C_{OX}$ when the device starts operating in the linear region.

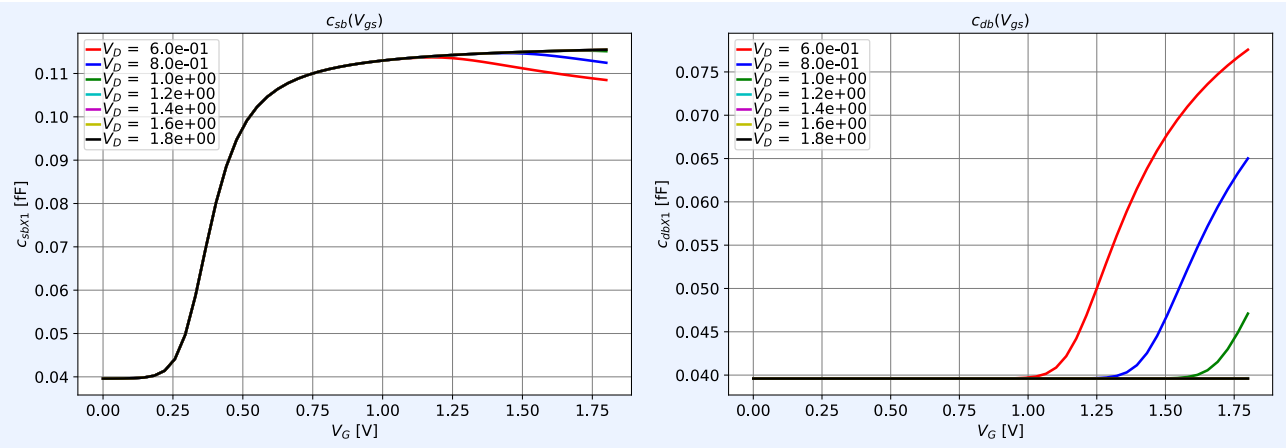


Figure 4.47: Device characteristics myNMOS, CMOS18 process.
Left: c_{sbi} plotted against V_{GS} with V_{DS} as parameter.
Right: c_{sbi} plotted against V_{GS} with V_{DS} as parameter.

Figure 4.47 shows the intrinsic source-bulk and drain-bulk capacitances, c_{sbi} and c_{dbi}, respectively. These are parts of the junction capacitances under the channel, assigned to the source and the drain terminal.
Figure 4.48 shows the plots of c_{gbi} and f_{Ti} against I_{DS} with V_{DS} as parameter.

Extrinsic capacitances

The extrinsic capacitances have to be added to this intrinsic model. They consists of the overlap capacitances, the source-bulk junction capacitance and the drain-bulk capacitances. As a first approximation they can be modeled as in the Level 1 model.

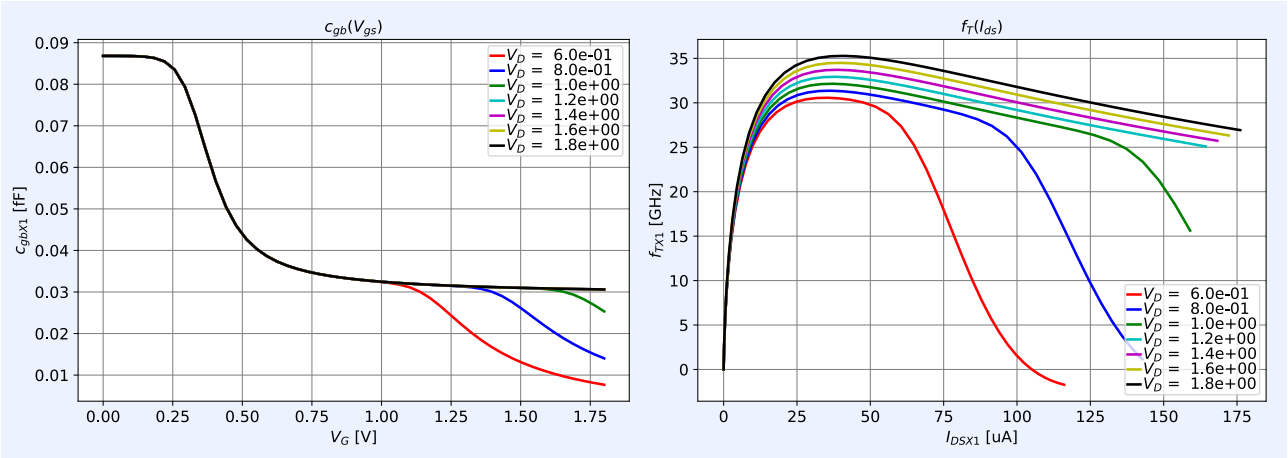


Figure 4.48: Device characteristics myN-MOS, CMOS18 process.
Left: c_{gbi} plotted against V_{GS} with V_{DS} as parameter.
Right: The cut-off frequency of the intrinsic transistor f_{ti} plotted against the drain current with V_{DS} as parameter.

Intrinsic noise model

Figure 4.49 shows the small-signal model of the intrinsic MOS transistor with added noise sources i_g and i_d.

The spectral density S_{i_g} of i_g equals

$$S_{i_g} = 2qI_G \text{ [A}^2\text{Hz}^{-1}\text{]}, \tag{4.186}$$

where I_G is the DC gate (leakage) current. Flicker noise, associated with this noise source, can be modeled with the aid of a $\frac{1}{f}$ corner frequency $f_{\ell g}$:

$$S_{i_g} = 2qI_G\left(1 + \frac{f_{\ell g}}{f}\right) \text{ [A}^2\text{Hz}^{-1}\text{]}. \tag{4.187}$$

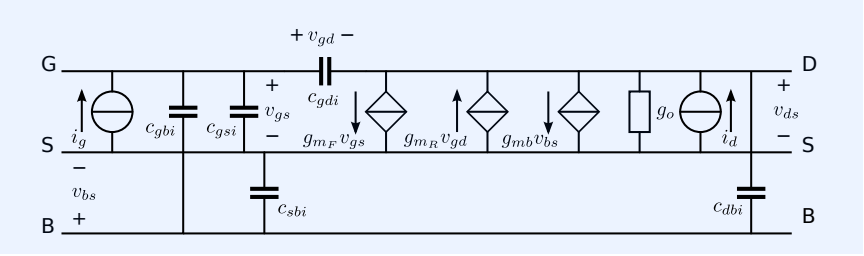


Figure 4.49: Dynamic small-signal noise model of the intrinsic MOS with the source taken as reference node according to the EKV model.

The spectral density of i_d differs for the linear and the saturation region. In the linear region ($V_{DS} < V_{DS,sat}$) the spectral density S_{id} is that of a resistor with resistance $\frac{1}{g_o}$:

$$S_{i_{d,lin}} = 4kTg_o \text{ [A}^2\text{Hz}^{-1}\text{]}, \tag{4.188}$$

where g_o can be written according to (4.172).

The spectral density S_{id} of the current noise associated with the drain current can be written as

$$S_{id} = 4kTn\Gamma g_m, \tag{4.189}$$

where Γ equals $\frac{1}{2}$ in weak inversion and $\frac{2}{3}$ in strong inversion[50,51]. Binkley (see [Binkley2008][52]) describes several expressions that model a smooth transition from weak inversion to strong inversion. The simplest function for Γ including VFMR and VS is:

$$\Gamma = \frac{\frac{1}{2} + \frac{2}{3}IC}{1 + IC} \text{ [-]}. \tag{4.190}$$

[50] Assuming no VFMR and velocity saturation.

[51] These are idealized values for Γ. In practice much larger values may be observed.

[52] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

Due to fluctuations in the carrier mobility, a flicker noise or $\frac{1}{f}$ noise component is associated with the channel current. The spectral density of this noise current is

$$S_{idf} = \frac{\text{KF}g_m^2}{C_{OX}f} \text{ [A}^2\text{Hz}^{-1}\text{]}, \tag{4.191}$$

where KF [J] is the flicker noise coefficient. The corner frequency f_ℓ is defined as the frequency at which $S_{idf} = S_{id}$. With the aid of this corner frequency, the spectral density $S_{id,tot}$ of the total noise current associated with the channel current can be written as:

$$S_{id,tot} = 4KTn\Gamma g_m \left(1 + \frac{f_\ell}{f}\right) \text{ [A}^2\text{Hz}^{-1}\text{]}, \tag{4.192}$$

The corner frequency f_ℓ is obtained as:

$$f_\ell = \frac{\text{KF}g_m}{4kTn\Gamma C_{OX}} \text{ [Hz]}. \tag{4.193}$$

If the transistor operates in the saturation region, the cut-off frequency ω_T equals the ratio of g_m and $\frac{2}{3}C_{OX}$. We then may write:

$$f_\ell = \text{KF}\frac{\pi}{3kTn\Gamma}f_T \text{ [Hz]}. \tag{4.194}$$

Hence, for a specific IC process the ratio between the flicker noise corner frequency f_ℓ and the cut-off frequency f_T is constant. In a 180nm CMOS process with the transistor operating in the saturation region we have: $n \approx 1.35$, $\Gamma = \frac{2}{3}$ and KF$\approx 4 \times 10^{-25}$ J. With these values we find $f_\ell \approx \frac{f_T}{9000}$.

Summary

Expressions for the parameters of the small-signal static model of the intrinsic MOS transistor operating in the saturation region have been described by Binkley (see [Binkley2008][53]). Extension of the validity to the linear operating range is possible by using the both the forward and the reverse inversion coefficient. Full-range expressions for the small-signal intrinsic capacitances can be found in the EKV2.6 manual [spiceEKV2.6][54].

Modeling of the small-signal parameters with the forward and the reverse current requires the use of V_{GS} and V_{DS} as independent variables for the operating point. The use of the preferred output quantities I_{DS} and V_{DS} requires an iterative calculation process for resolving I_{DS} in a forward and a reverse component.

[53] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

[54] Matthias Bucher, Christoff Lallement, Christian Enz, Fabien Théodoloz and François Krummenacher. The EPFL-EKV MOSFET Model Equations for Simulation. Technical report, Electronics Laboratory, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, July 1998

4.5 SLiCAP device models

The design method for amplifiers described in this book separates the design of the signal path from the design of the biasing. We will first design the signal path and then the biasing. Reasons for this are:

1. The signal performance with ideal biasing should be within specifications and leave extra room for degradation due to biasing errors. If the biasing is designed together with the signal path, it is hard to find the cause of possible performance limitations. The performance of the signal path might then be out of specifications, while the focus may be on improvement of the biasing. If so, the designer is improving the biasing of an unfeasible amplifier, valuable design time gets lost and possible show stoppers may appear during a relatively late design phase.
2. If the signal path of an amplifier shows unstable behavior, a numeric simulation may not be able to find the correct quiescent operating point due

to convergence problems. As a consequence, the noise performance and the small-signal performance cannot be evaluated. This makes it hard or even impossible to find the causes for instability and *trial and horror* methods may be the only way to obtain stability. Only after stable behavior has been achieved, it may become clear that the performance is not as desired and valuable design time has gone lost.

4.5.1 Signal path and biasing

Designing the signal path first does not mean that the designer shouldn't account for biasing constraints while designing the signal path. Biasing considerations should be taken into account as early as possible. Especially in low-power applications, the selection of the device type, the operating conditions and the circuit topology are to a large extent driven by the feasibility of the biasing circuits that will be designed at a later stage. However, investigation of the adverse effects of the implementation of bias sources can very well be done without implementing them in the complete circuit.

One way to achieve this is to design the amplifier's signal path using a linearized approach. In such an approach the frequency stability of an amplifier is investigated in operating points that occur during operation. Biasing circuitry can be omitted if the parameters of the small-signal models of the active devices can be related to the operating point. Operating points of interest can be estimated from the excursions and the rates of change of the signals that occur at the various nodes of the amplifier. By doing so, budgets for parasitic impedances that will be introduced by bias sources can be determined by inserting those impedances in the small-signal equivalent circuit and studying their effect independently from biasing changes.

Spice-like numeric simulators do not support this way of working. Those simulators can determine the small-signal parameters in an operating point only after a bias solution has been found. In other words: they need a biased circuit to work with. SLiCAP however, is designed for this purpose.

4.5.2 SLiCAP parametric small-signal models

SLiCAP is a symbolic simulator that can be used at an early stage of the design for:

- The design of the noise behavior
- The design of the small-signal dynamic behavior
- Deriving budgets for biasing imperfections

SLiCAP does not need to find a DC solution to determine the small-signal parameters. It has built-in parameterized sub-circuits with small-signal models of active devices. Operating conditions and device geometry parameters can be passed to these sub-circuits, while technology parameters and device equations relate them to the small-signal parameters.

In the following sections, we will give a brief description of some device models that have been included in SLiCAP. These sub-circuit definitions arebe found in the libraries supplied with SLiCAP.

SLiCAP is intended to be used to motivate early-stage design decisions and useful design information can be obtained with low complexity models.[55]

[55] Einstein: "Everything should be made as simple as it can be, but not simpler."

4.5.3 BJT forward region, no saturation

The subcircuit `BJTV4` is a parameterized small-signal model of a four-terminal vertical BJT. It can be used for simulation of the small-signal behavior under

variation of the collector current I_C and the collector-emitter voltage V_{CE}. The transconductance $g_m(I_C)$ of the device is modeled for operation in the forward active region (no saturation). The base-emitter capacitance is modeled as a function of the collector current. The base-collector capacitance is modeled as a function of the collector-emitter voltage. For the sake of simplicity, the base resistance and the substrate capacitance are both set to zero.

The sub-circuit definition is found in the SLiCAP library `SLiCAP.lib`.

```
358 * device equations EKV model
359 * See Binkley: "Tradeoffs and Optimization in Analog CMOS Design"
360 + IC_CRIT = {1/(4*(N_s_P18*U_T)*(Theta_P18+1/L/E_CRIT_P18))^2}
361 + g_m     = {-ID/(N_s_P18*U_T*sqrt(IC*(1+IC/IC_CRIT)+0.5*sqrt(IC*(1+IC/IC_CRIT
        ))+1))}
362 + g_o     = {-ID/VAL_P18/L}
363 + c_gs    = {2/3*W*L*C_OX_P18 + CGS0_P18*W}
364 + c_dg    = {CGSO_P18*W}
365 + c_gb    = {CGBO_P18*2*L+(N_s_P18-1)/N_s_P18*C_OX_P18*W*L/3}
366 + c_db    = {CJB0_P18*W*LDS_P18}
367 + IC_i    = {-ID*L/W/I_0_P18} ; Correction for initial estimate of inversion
        coefficient
368 + IC      = {IC_i*(1+IC_i/4/IC_CRIT)} ; Inversion coefficient corrected for
        short-channel effects
369 + V_GS    = {-2*N_s_P18*U_T*ln(exp(sqrt(IC))-1)+Vth_P18}
370 .ends
371
372 .subckt BJTV4 collector base emitter bulk IC={IC} VCE={VCE}
373 * small-signal GP model of vertical BJT
374 Q1 collector base emitter bulk BJTV4
375 .model BJTV4 QV
376 * device small-signal parameters
377 + gm  = {g_m}  ; transconductance
378 + go  = {g_o}  ; output conductance
379 + gpi = {g_pi} ; input conductance
380 + gbc = 0      ; not modeled in GP model
381 + cpi = {c_pi} ; base-emitter capacitance
382 + cbc = {c_bc} ; base-collector capacitance
383 + cbx = 0      ; no external collector-base capacitance
384 + cs  = 0      ; no substrate capacitance
385 + rb  = 0      ; zero base resistance
386 .param
387 * device equations GP model
388 + g_m  = {IC/U_T}
389 + g_o  = {IC/(VAF+VCE)}
```

The first line of the listing defines the name, the nodes and the parameters to be passed. The syntax is that of SPICE. Symbolic parameters that will be used in the sub-circuit are placed between curly brackets.

Lines 17 through 21 define the device equations for the small-signal parameters of this model. If desired, an equation for the substrate capacitance can be added here. Line 13 should then be changed to:

```
+ cs = {c_s}
```

Where `c_s` is the parameter name used in the equation. Other equations can be added or modified in a similar way.

Line 360 places the SLiCAP small-signal model in the sub-circuit. The small-signal model is shown in Figure 4.51. A full description of this model can be found in the SLiCAP help file. The model itself is also defined as a SPICE subcircuit, it can be found in the SLiCAP library: `SLiCAPmodels.lib`, together with the definition of global parameters and technology parameters. Please view the SLiCAP manual for details.

4.5.4 NMOS EKV forward region, saturation range

The sub-circuit `CMOS18N` is a parameterized small-signal model of an NMOS transistor in CMOS18 technology. It can be used for simulation of the small-signal behavior under variation of the channel current I_{DS}. The small-signal parameters are a function of I_{DS}, W and L. The model equations cover operation in the saturation region from weak inversion until strong inversion including velocity saturation. Capacitances of the small-signal model are

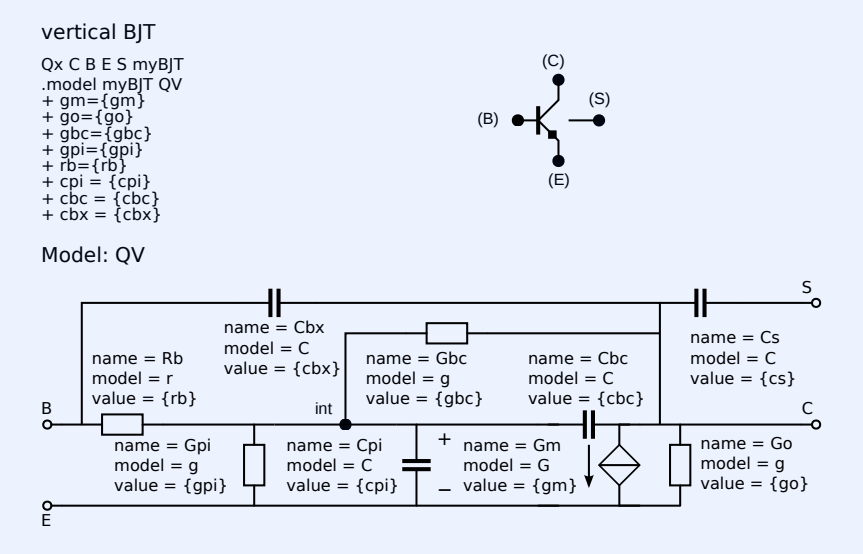


Figure 4.50: SLiCAP small-signal dynamic model of the vertical BJT.

calculated from the oxide capacitance, the overlap capacitances and the zero-voltage depletion capacitances of the drain and the source. The voltage dependency of these capacitances has not been modeled, but the model can be extended, if desired. The model can easily be adapted to another technology by changing the technology parameters.

The CMOS 18 model parameters are defined in the SLiCAP library `SLiCAPmodels.lib`.

```
"SLiCAPmodels"

* Physical constants
**********************

.param
+ q             = 1.60217662e-19    ; Electron charge in [C]
+ c             = 2.99792458e+08    ; Speed of light in [m/s]
+ mu_0          = {4*pi*1e-7}       ; Permeability of vacuum in [H/m]
+ epsilon_SiO2  = 3.9               ; Relative permittivity of SiO2 [-]
+ k             = 1.38064852e-23    ; Boltzmann constant in [J/K]
+ epsilon_0     = {1/mu_0/c^2}      ; permittivity of vacuum in [F/m]

* Temperature and thermal voltage
*********************************

.param
+ T             = 300               ; Default value of the absolute
    temperature in [K]
+ U_T           = {k*T/q}           ; Thermal voltage [V]

* CMOS18 technology parameters for EKV models (SI units)
********************************************************

.param
+ TOX_N18     = 4.1n  ; oxide thickness [m]
+ Vth_N18     = 0.36  ; threshold voltage [V]
+ N_s_N18     = 1.35  ; substrate factor [-]
+ Theta_N18   = 0.28  ; vertical field mobility reduction coefficient [1/V]
+ E_CRIT_N18  = 5.6M  ; lateral field strength for velocity saturation [V/m]
+ u_0_N18     = 42m   ; zero field carrier mobility [m^2/V/s]
+ CGBO_N18    = 1p    ; gate-bulk overlap capacitance [F/m]
+ CGSO_N18    = 300p  ; gate-source and gate-drain overlap capacitance [F/m]
+ CJB0_N18    = 1m    ; source/bulk drain/bulk capacitance [F/m^2]
+ LDS_N18     = 180n  ; length of drain and source [m]
+ VAL_N18     = 40M   ; Early voltage per unit of length [V/m]
+ KF_N18      = 2e-27 ; flicker noise (1/f noise) coefficient, zero for f_ell
    =0 [C/m^2]
+ AF_N18      = 1     ; flicker noise exponent [-]
+ C_OX_N18    = {epsilon_0 * epsilon_SiO2 / TOX_N18}; oxide capacitance per
    unit of area [F/m^2]
```

```
+ I_0_N18     = {2*N_s_N18*u_0_N18*C_OX_N18*U_T^2}  ; technology current [A]
+ V_KF_N18    = 2     ; flicker noise voltage coefficient [V]

.param
+ TOX_P18     = 4.1n  ; oxide thickness [m]
+ Vth_P18     = -0.36 ; threshold voltage [V]
+ N_s_P18     = 1.35  ; substrate factor [-]
+ Theta_P18   = 0.35  ; vertical field mobility reduction factor [1/V]
+ E_CRIT_P18  = 14M   ; lateral field strength for velocity saturation [V/m]
+ u_0_P18     = 8.92m ; zero field carrier mobility [m^2/V/s]
+ CGBO_P18    = 1p    ; gate-bulk overlap capacitance [F/m]
+ CGSO_P18    = 300p  ; gate-source and gate-drain overlap capacitance [F/m]
+ CJB0_P18    = 1m    ; source/bulk drain/bulk capacitance [F/m^2]
+ LDS_P18     = 180n  ; length of drain and source [m]
+ VAL_P18     = 40M   ; Early voltage per unit of length [V/m]
+ KF_P18      = 1e-27 ; flicker noise (1/f noise) coefficient, zero for f_ell
    =0 [C/m^2]
+ AF_P18      = 1     ; flicker noise exponent [-]
+ C_OX_P18    = {epsilon_0 * epsilon_SiO2 / TOX_P18}; oxide capacitance per
    unit of area [F/m^2]
```

The sub-circuit of the small-signal model (see Figure 4.51) is defined in the same library file:

```
.ends

* MOSFET
.subckt M d g s b gm=1m gb=0 go=0 cgs=0 cdg=0 cgb=0 cdb=0 csb=0 ; default
    values
Gm d s g s g value={gm}
Gb d s b s g value={gb}
Go d s d s g value={go}
Cgs g s {cgs}
Cdg d g {cdg}
Cgb g b {cgb}
Cdb d b {cdb}
```

Symbolic parameters that will be used in the sub-circuit are placed between curly brackets.

The definition of the EKV model based on this sub circuit is found in the SLiCAP library `SLiCAP.lib`.

```
.ends

.subckt O_dcvar 1 2 4 COMMON
+ sib={sib}
+ sio={sio}
+ svo={svo}
+ iib={iib}
Ib 1 5 I dc={iib} dcvar={sib^2}
F1 2 COMMON 5 COMMON 1
Io 1 2 I dcvar={sio^2}
Vo 1 3 V dcvar={svo^2}
N1 4 COMMON 3 2
.ends

.subckt CMOS18N drain gate source bulk W={W} L={L} ID={ID}
* EKV model of transistor without bulk resistances
* Voltage dependency of bulk capacitances not modeled
* Operating in forward saturation region
*
M1 drain gate source bulk CMOS18N

.model CMOS18N M
+ gm      = {g_m}
+ go      = {g_o}
+ gb      = {g_b}
+ cgs     = {c_gs}
+ cdg     = {c_dg}
+ cgb     = {c_gb}
+ cdb     = {c_db}
+ csb     = {c_sb}
* Parameters will be substituted if simType has been set to "numeric"
.param
* device equations EKV model
* See Binkley: "Tradeoffs and Optimization in Analog CMOS Design"
+ IC_CRIT = {1/(4*(N_s_N18*U_T)*(Theta_N18+1/L/E_CRIT_N18))^2}
```

Line 122 places the SLiCAP small-signal model in the sub-circuit. The model is shown in Figure 4.51. A full description of this model can be found in the SLiCAP help file.

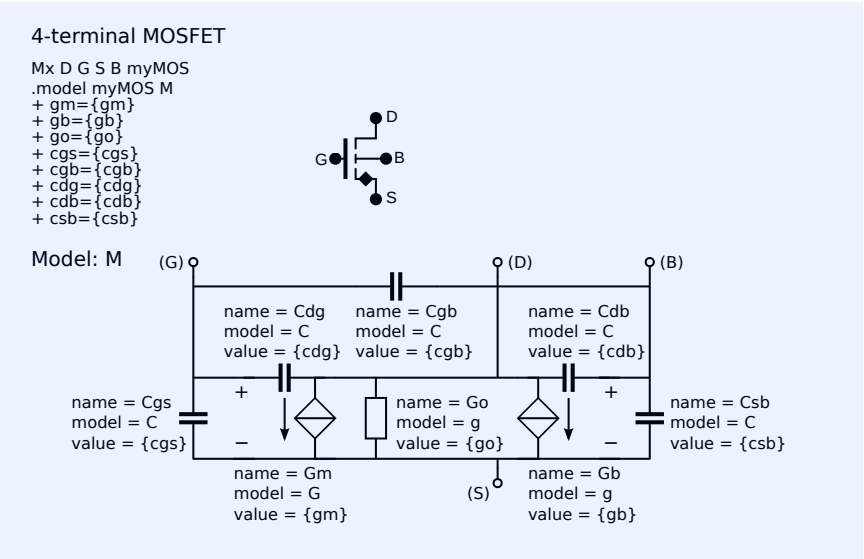


Figure 4.51: SLiCAP small-signal dynamic model of the MOS transistor.

Lines 125 through 132 define the parameter that will be used for symbolic analysis. In order to keep the symbolic expressions as compact as possible, simple single-parameter values have been used. If desired, numeric values or more complex expressions may as well be used.

Lines 135 through 150 give the device equations for the elements of the small-signal model. All intermediate parameters such as IC_{CRIT} in line 137, will be local parameters and appear in the main circuit as `PARNAME_Xnnn`, where `PARNAME` is the name of the parameter and `Xnnn` the name of the sub-circuit. In case of nested sub-circuits it will be `PARNAME_XnnnXmmm`, etc, where `Xmmm` is the sub-circuit that calls `Xnnn`. The values and expressions of these parameters can be listed in the HTML output or passed to the PYTHON workspace. Please notice that the thermal voltage $U_T = \frac{kT}{q}$, is a global parameter. See the SLiCAP help file for more information about the use of global and local parameters.

In section 4.5.6 we will demonstrate in which way the device parameters can be plot against each other. This makes it possible to compare the small-signal element values obtained with SLiCAP with those obtained from other simulators, and to adjust the SLiCAP device equations and technology parameters to obtain optimum model correspondence.

4.5.5 NMOS EKV forward region, linear and saturation range

The sub-circuit `CMOS18N_V` is a parameterized small-signal model of an NMOS transistor in CMOS18 technology. It can be used for simulation of the small-signal behavior under variation of the voltages V_{DB}, V_{GB} and V_{SB}. The small-signal parameters of the intrinsic transistor are modeled for forward operation in the linear region and in the saturation region from weak inversion until strong inversion including velocity saturation. The model equations for the elements of the small-signal model of the intrinsic transistor are those described in section 4.4.7. The extrinsic capacitances have been added as in the `CMOS18N` sub-circuit. This model can easily be adapted to another technology by changing the technology parameters.

The definition of this sub-circuit is listed below:

```
153 + g_o     = {ID/VAL_N18/L}
154 + g_b     = {(N_s_N18-1)*g_m}
155 + c_gs    = {2/3*W*L*C_OX_N18 + CGSO_N18*W}
156 + c_dg    = {CGSO_N18*W}
157 + c_gb    = {CGBO_N18*2*L+(N_s_N18-1)/N_s_N18*C_OX_N18*W*L/3}
158 + c_db    = {CJB0_N18*W*LDS_N18}
159 + c_sb    = {CJB0_N18*W*LDS_N18}
160 + IC_i    = {ID*L/W/I_0_N18} ; Initial estimate of inversion coefficient
161 + IC      = {IC_i*(1+IC_i/4/IC_CRIT)} ; Inversion coefficient corrected for
        short-channel effects
162 + V_GS    = {2*N_s_N18*U_T*ln(exp(sqrt(IC))-1)+Vth_N18}
163 + f_T     = {g_m/2/pi/c_iss}
164 + c_iss   = {c_gs+c_dg+c_gb}
165 .ends
166
167 .subckt CMOS18N_V drain gate source bulk W={W} L={L} VD={VD} VG={VG} VS={VS}
168 * EKV model of transistor without bulk resistances
169 * Voltage dependency of bulk capacitances not modeled
170 * Operating voltages with respect to the bulk
171 * Assumes forward operation: VDS > 0
172 *
173 M1 drain gate source bulk CMOS18N_V
174
175 .model CMOS18N_V M
176 + gm      = {g_m}
177 + go      = {g_o}
178 + gb      = {g_b}
179 + cgs     = {c_gs}
180 + cdg     = {c_dg}
181 + cgb     = {c_gb}
182 + cdb     = {c_db}
183 + csb     = {c_sb}
184 * Parameters will be substituted if simType has been set to "numeric"
185 .param
186 * device equations EKV model
187 * See Binkley: "Tradeoffs and Optimization in Analog CMOS Design"
188 * See EKV2.6 model manual
189 + V_A     = {VAL_N18*L}
190 + C_gso   = {CGSO_N18*W}
191 + C_gdo   = {CGSO_N18*W}
192 + C_gbo   = {CGBO_N18*2*L}
193 + C_sb0   = {CJB0_N18*W*LDS_N18}
194 + C_db0   = {CJB0_N18*W*LDS_N18}
195 + IC_CRIT = {1/((4*N_s_N18*U_T)*(Theta_N18+1/L/E_CRIT_N18))^2}
196 + V_effF  = {(VG-Vth_N18-N_s_N18*VS)}
197 + V_effR  = {(VG-Vth_N18-N_s_N18*VD)}
198 + IC_F    = {(ln(1+exp(V_effF/2/N_s_N18/U_T)))^2}
199 + IC_R    = {(ln(1+exp(V_effR/2/N_s_N18/U_T)))^2}
200 + x_f     = {sqrt(1/4+IC_F)}
201 + x_r     = {sqrt(1/4+IC_R)}
202 + I_DSF   = {W/L*I_0_N18*(IC_F)*(1+(VD-VS)/V_A)/(1+(Theta_N18+1/L/E_CRIT_N18)
        *2*U_T*sqrt(IC_F))}
203 + I_DSR   = {W/L*I_0_N18*(IC_R)*(1+(VS-VD)/V_A)/(1+(Theta_N18+1/L/E_CRIT_N18)
        *2*U_T*sqrt(IC_R))}
204 + I_DS    = {I_DSF-I_DSR}
205 + g_mF    = {I_DSF/(N_s_N18*U_T*sqrt(IC_F*(1+IC_F/IC_CRIT)+0.5*sqrt(IC_F*(1+
        IC_F/IC_CRIT))+1))}
```

The first line of the listing defines the name, the nodes and the parameters to be passed. Line 159 places the SLiCAP small-signal model in the sub-circuit. The model is shown in Figure 4.51. Line 162 through 169 define the parameter values or equations that will be used for symbolic analysis. They are the same as in the `CMOS18N` sub-circuit.

Lines 172 through 204 give the device equations for the elements of the small-signal model. All intermediate parameters such as V_A in line 175, will be local parameters and appear in the main circuit as `PARNAME_Xnnn`, where `PARNAME` is the name of the parameter and `Xnnn` the name of the sub-circuit. In case of nested sub-circuits it will be `PARNAME_XnnnXmmm`, etc. Where `Xmmm` is the sub-circuit that calls `Xnnn`. The values and expressions of these parameters can be listed in the HTML output or passed to the PYTHON workspace. Please notice that the thermal voltage $U_T = \frac{kT}{q}$, as well as ε_0 and ε_{SiO_2}, are global parameters. Please see the SLiCAP help file for more information about the use of global and local parameters.

The technology parameters are taken from the SLiCAP library `SLiCAPmodels.lib` (see description above). In section 4.5.6 we will demonstrate in which way the device parameters can be plot against each other. This makes it possible to compare the small-signal element values obtained with SLiCAP with those obtained from other simulators, and to adjust the SLiCAP device equations and technology parameters to obtain optimum model correspondence.

4.5.6 SLiCAP MOS device characteristics

The simulation results obtained with the SLiCAP models can be adjusted to those obtained from measurements or from other simulators. This can be done by comparing graphs obtained from measurement or simulation with similar graphs obtained from SLiCAP and adjust the SLiCAP model parameters or model equations to reduce differences to an acceptable level.

In this section we will demonstrate how to plot the device characteristics with SLiCAP.

Below is the listing of a SLiCAP file that can been used for this purpose:

```
#!/usr/bin/env python2
# -*- coding: utf-8 -*-
"""
Created on Sun Jul  5 19:15:00 2020

@author: anton
"""

from SLiCAP import *
t1 = time()

prj = initProject('NMOS EKV plots') # Creates the SLiCAP libraries and the
                                    # project HTML index page

fileName = 'mosEKVplotsN_V.cir'
i1 = instruction()           # Creates an instance of an instruction object
i1.setCircuit(fileName)      # Checks and defines the local circuit object and
                             # sets the index page to the circuit index page
htmlPage('Circuit data')
netlist2html(fileName)
elementData2html(i1.circuit)
params2html(i1.circuit)

# Put the plots on a page
htmlPage('CMOS18 EKV model plots')

i1.setDataType('params')
i1.defPar('V_G', 1.8)
i1.defPar('V_D', 1.8)

i1.stepOn()
i1.setStepVar('V_G')
i1.setStepStart(0.6)
i1.setStepStop(1.8)
i1.setStepNum(6)
i1.setStepMethod('lin')

result = i1.execute()

fig_Ids_Vds = plotSweep('IdsVds', '$I_{ds}(V_{ds})$', result, 0, 1.8, 50,
     sweepVar= 'V_D', xUnits = 'V', yVar = 'I_DS_X1', yScale = 'u', yUnits = '
     A', funcType = 'param', show = True)
fig2html(fig_Ids_Vds, 600)

i1.setStepVar('V_D')
result = i1.execute()

fig_Ids_Vgs = plotSweep('IdsVgs', '$I_{ds}(V_{gs})$', result, 0, 1.8, 50,
     axisType = 'lin', sweepVar= 'V_G', xUnits = 'V', yVar = 'I_DS_X1', yScale
      = 'u', yUnits = 'A', funcType = 'param', show = True)
fig2html(fig_Ids_Vgs, 600)

fig_Ids_Vgs_Log = plotSweep('IdsVgsLog', '$I_{ds}(V_{gs})$', result, 0.01,
     1.8, 50, axisType = 'semilogy', sweepVar= 'V_G', xUnits = 'V', yVar = '
     I_DS_X1', yScale = 'u', yUnits = 'A', funcType = 'param', show = True)
```

```
50 fig2html(fig_Ids_Vgs_Log, 600)
51
52 fig_gm_Ids  = plotSweep('gmIds', '$g_m(I_{ds})$', result, 0, 1.8, 50, sweepVar
       = 'V_G', xVar = 'I_DS_X1', xScale = 'u', xUnits = 'A', yVar = 'g_m_X1',
       yScale = 'u', yUnits = 'S', funcType = 'param', show = True)
53 fig2html(fig_gm_Ids, 600)
54
55 fig_fT_Ids  = plotSweep('fTIds', '$f_{T}(I_{ds})$', result, 0, 1.8, 50,
       sweepVar= 'V_G', xVar = 'I_DS_X1', xScale = 'u', xUnits = 'A', yVar = '
       f_T_X1', yScale = 'G', yUnits = 'Hz', funcType = 'param', show = True)
56 fig2html(fig_fT_Ids, 600)
57
58 fig_CissVg  = plotSweep('CissVg', '$c_{iss}(V_{gs})$', result, 0, 1.8, 50,
       sweepVar= 'V_G', xScale = '', xUnits = 'V', yVar = 'c_iss_X1', yScale = '
       f', yUnits = 'F', funcType = 'param', show = True)
59 fig2html(fig_CissVg, 600)
60
61 t2=time()
62 print(t2-t1,'s')
63
64 LTspiceTraces =  LTspiceData2Traces('nmosChar.txt')
65 traces2fig(LTspiceTraces, fig_Ids_Vgs)
66 traces2fig(LTspiceTraces, fig_Ids_Vgs_Log)
67 fig_Ids_Vgs.plot()
68 fig_Ids_Vgs_Log.plot()
69
70 figLT = plot('LTspiceIdsVgs', 'LTspice $I_{ds}(V_{gs})$', 'lin', LTspiceTraces
       , xName = '$V_{gs}$', xUnits = 'V', yName = '$I_{ds}$', yUnits = 'A',
       yScale = 'u', show = True)
```

SLiCAP needs a circuit file to work with. Checking of a circuit creates an internal data structure with circuit parameters, which is needed for calculating parameter values. Checking of the circuit is performed by the instruction in line 1.

The circuit file itself is a very simple netlist comprising the transistor and the inclusion of the library with the model and initial parameter definitions. The listing of this circuit file is shown below:

```
1 mosEKVplots
2 * SLiCAP netlist file
3 .include C18.lib
4 X1 d g s 0 CMOS18N_V W={W} L={L} VD={V_D} VG={V_G} VS={V_S}
5 .param V_D=1.8 V_G=0.5 V_S=0 W=220n L=180n
6 .end
```

Line 4 of the circuit file assigns numeric values to the parameters to be passed to the sub-circuit. These values can be changed from within the PYTHON environment.

The instructions in lines 13 until 16 create an HTML page that shows the circuit data after the circuit has been checked. Figure 4.53 and 4.54 show this HTML page. The parameter listing section shows the symbolic expression and the numeric value[56] of each parameter. Thanks to the math rendering they are easier to read than those in the sub-circuit definition in the library file. Please note that all local parameters of the sub-circuit `X1` have `X1` added to the subscript of their parameter name.

The graphs have been shown in section 4.4.7.

[56] After recursive substitution of all parameters.

4.5.7 PMOS EKV models

Similar as for NMOS devices, SLiCAP has two sub-circuits for PMOS EKV models:

1. `CMOS18P`

   The only difference with the NMOS sub-circuit `CMOS18N` is the technology section in the model definition. The direction of the drain current is not accounted for; the parameter ID should be given a negative value.

2. `CMOS18P_V`

Apart from the technology definition and the signs of the voltages, this sub-circuit is identical to its NMOS version: `CMOS18N_V`. Please notice that all voltages should be given a positive value with respect to the bulk voltage, similar as with the NMOS devices.

4.6 Conclusions

In this chapter, we have briefly studied the operation and modeling of modern active devices. We have seen that those devices can be considered as nonlinear two-ports, of which the input port and the output port share one terminal. When biased properly, they can provide an available power gain that exceeds unity.

Although the fabrication, the operation and the modeling of active devices differs for each device, they have a lot in common:

1. The physical operation mechanism of all these active devices is that the current in the output port is controlled by the voltage across the input port, while building up this voltage requires a nonzero charge. This charge controls the current in the output port.[57]
2. The static intrinsic noise sources of all devices are the noise currents associated with the static (DC) port currents. We have studied in which way they can be converted into equivalent-input noise sources.
3. When properly biased, the active devices, together with their bias sources, can provide an available power gain that exceeds unity. In the following chapters we will design amplifiers using biased active devices. Since the design theory itself is technology-independent, we will often use generalized biased three-terminal devices from Figure 3.14. Figure 4.52 shows a simple small-signal of such a device that can be used during early stages of the design. It models the small-signal behavior of an intrinsic device. For BJTs we have $g_i = g_m/\beta_{AC}$. For field effect devices g_i equals zero.

[57] A current will flow in the output port only if both the input port voltage and the output port voltage differ from zero.

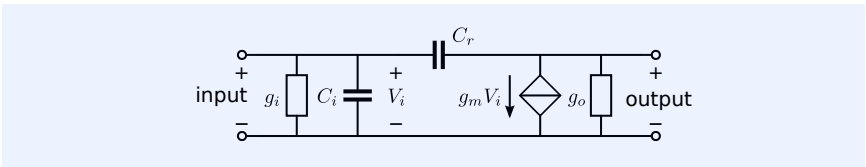


Figure 4.52: Generalized small-signal model of the active devices.

Circuit data

Netlist: mosEKVplotsN_V.cir

```
mosEKVplots
* SLiCAP netlist file
X1 d g s 0 CMOS18N_V W={W} L={L} VD={V_D} VG={V_G} VS={V_S}
.param V_D=1.8 V_G=0.5 V_S=0 W=220n L=180n
.end
```

Table: Element data of expanded netlist 'mosEKVplots'

RefDes	Nodes	Refs	Model	Param	Symbolic	Numeric
Cdb_M1_X1	d 0		C	value	c_{dbX1}	$3.96 \cdot 10^{-17}$
Cdg_M1_X1	d g		C	value	c_{dgX1}	$6.6 \cdot 10^{-17}$
Cgb_M1_X1	g 0		C	value	c_{gbX1}	$4.388 \cdot 10^{-17}$
Cgs_M1_X1	g s		C	value	c_{gsX1}	$2.316 \cdot 10^{-16}$
Csb_M1_X1	s 0		C	value	c_{sbX1}	$9.758 \cdot 10^{-17}$
Gb_M1_X1	d s 0 s		g	value	g_{bX1}	$1.439 \cdot 10^{-5}$
Gm_M1_X1	d s g s		g	value	g_{mX1}	$4.112 \cdot 10^{-5}$
Go_M1_X1	d s d s		g	value	g_{oX1}	$4.319 \cdot 10^{-7}$

Table: Parameter definitions in 'mosEKVplots'.

Name	Symbolic	Numeric
$CGBO_{N18}$	$1.0 \cdot 10^{-12}$	$1.0 \cdot 10^{-12}$
$CGSO_{N18}$	$3.0 \cdot 10^{-10}$	$3.0 \cdot 10^{-10}$
CJB_{0N18}	0.001	0.001
C_{OXN18}	$\frac{\epsilon_0 \epsilon_{SiO2}}{TOX_{N18}}$	0.008422
$E_{CRITN18}$	$5.6 \cdot 10^{6}$	$5.6 \cdot 10^{6}$
I_{0N18}	$2C_{OXN18}N_{sN18}U_T^2 u_{0N18}$	$6.383 \cdot 10^{-7}$
L	$1.8 \cdot 10^{-7}$	$1.8 \cdot 10^{-7}$
LDS_{N18}	$1.8 \cdot 10^{-7}$	$1.8 \cdot 10^{-7}$
N_{sN18}	1.35	1.35
T	300	300.0
TOX_{N18}	$4.1 \cdot 10^{-9}$	$4.1 \cdot 10^{-9}$
Θ_{N18}	0.28	0.28
U_T	$\frac{Tk}{q}$	0.02585
VAL_{N18}	$4.0 \cdot 10^{7}$	$4.0 \cdot 10^{7}$
V_D	1.8	1.8
V_G	0.5	0.5
V_S	0	0
Vth_{N18}	0.36	0.36
W	$2.2 \cdot 10^{-7}$	$2.2 \cdot 10^{-7}$
c	$2.998 \cdot 10^{8}$	$2.998 \cdot 10^{8}$
ϵ_0	$\frac{1}{c^2 \mu_0}$	$8.854 \cdot 10^{-12}$
ϵ_{SiO2}	3.9	3.9
k	$1.381 \cdot 10^{-23}$	$1.381 \cdot 10^{-23}$
μ_0	$4.0 \cdot 10^{-7} \pi$	$1.257 \cdot 10^{-6}$
q	$1.602 \cdot 10^{-19}$	$1.602 \cdot 10^{-19}$
u_{0N18}	0.042	0.042
C_{db0X1}	$CJB_{0N18}LDS_{N18}W$	$3.96 \cdot 10^{-17}$
C_{gboX1}	$2CGBO_{N18}L$	$3.6 \cdot 10^{-19}$
C_{gdoX1}	$CGSO_{N18}W$	$6.6 \cdot 10^{-17}$

Figure 4.53: SLiCAP html page with circuit data.

C_{gsoX1}	$CGSO_{N18}W$	$6.6 \cdot 10^{-17}$
C_{sb0X1}	$CJB_{0N18}LDS_{N18}W$	$3.96 \cdot 10^{-17}$
IC_{CRITX1}	$\frac{1}{16N_{sN18}^2U_T^2\left(\Theta_{N18}+\frac{1}{E_{CRITN18}L}\right)^2}$	31.71
IC_{FX1}	$\log\left(e^{\frac{V_{effFX1}}{2N_{sN18}U_T}}+1\right)^2$	4.545
IC_{RX1}	$\log\left(e^{\frac{V_{effRX1}}{2N_{sN18}U_T}}+1\right)^2$	$3.081 \cdot 10^{-29}$
I_{DSFX1}	$\frac{IC_{FX1}I_{0N18}W\left(1+\frac{V_D-V_S}{V_{AX1}}\right)}{L\left(\sqrt{IC_{FX1}}U_T\left(2\Theta_{N18}+\frac{2}{E_{CRITN18}L}\right)+1\right)}$	$3.887 \cdot 10^{-6}$
I_{DSRX1}	$\frac{IC_{RX1}I_{0N18}W\left(1+\frac{-V_D+V_S}{V_{AX1}}\right)}{L\left(\sqrt{IC_{RX1}}U_T\left(2\Theta_{N18}+\frac{2}{E_{CRITN18}L}\right)+1\right)}$	$1.803 \cdot 10^{-35}$
I_{DSX1}	$I_{DSFX1}-I_{DSRX1}$	$3.887 \cdot 10^{-6}$
V_{AX1}	$LVAL_{N18}$	7.2
V_{effFX1}	$-N_{sN18}V_S+V_G-Vth_{N18}$	0.14
V_{effRX1}	$-N_{sN18}V_D+V_G-Vth_{N18}$	−2.29
c_{dbX1}	$C_{db0X1}+c_{dgiX1}\left(N_{sN18}-1\right)$	$3.96 \cdot 10^{-17}$
c_{dgX1}	$C_{gdoX1}+c_{dgiX1}$	$6.6 \cdot 10^{-17}$
c_{dgiX1}	$C_{OXN18}LW\left(\frac{2}{3}-\frac{2\left(x_{fX1}^2+x_{fX1}+0.5x_{rX1}\right)}{3\left(x_{fX1}+x_{rX1}\right)^2}\right)$	0
c_{gbX1}	$C_{gboX1}+\frac{(N_{sN18}-1)(C_{OXN18}LW-c_{dgiX1}-c_{gsiX1})}{N_{sN18}}$	$4.388 \cdot 10^{-17}$
c_{gsX1}	$C_{gsoX1}+c_{gsiX1}$	$2.316 \cdot 10^{-16}$
c_{gsiX1}	$C_{OXN18}LW\left(\frac{2}{3}-\frac{2\left(0.5x_{fX1}+x_{rX1}^2+x_{rX1}\right)}{3\left(x_{fX1}+x_{rX1}\right)^2}\right)$	$1.657 \cdot 10^{-16}$
c_{issX1}	$c_{dgX1}+c_{gbX1}+c_{gsX1}$	$3.415 \cdot 10^{-16}$
c_{sbX1}	$C_{sb0X1}+c_{gsiX1}\left(N_{sN18}-1\right)$	$9.758 \cdot 10^{-17}$
f_{TX1}	$\frac{g_{mX1}}{2\pi c_{issX1}}$	$1.916 \cdot 10^{10}$
g_{bX1}	$g_{mX1}\left(N_{sN18}-1\right)$	$1.439 \cdot 10^{-5}$
g_{mFX1}	$\frac{I_{DSFX1}}{N_{sN18}U_T\sqrt{IC_{FX1}\left(1+\frac{IC_{FX1}}{IC_{CRITX1}}\right)}+0.5\sqrt{IC_{FX1}\left(1+\frac{IC_{FX1}}{IC_{CRITX1}}\right)}+1}$	$4.112 \cdot 10^{-5}$
g_{mRX1}	$\frac{I_{DSRX1}}{N_{sN18}U_T\sqrt{IC_{RX1}\left(1+\frac{IC_{RX1}}{IC_{CRITX1}}\right)}+0.5\sqrt{IC_{RX1}\left(1+\frac{IC_{RX1}}{IC_{CRITX1}}\right)}+1}$	$5.166 \cdot 10^{-34}$
g_{mX1}	$g_{mFX1}-g_{mRX1}$	$4.112 \cdot 10^{-5}$
g_{oX1}	$\frac{I_{DSFX1}}{V_{AX1}+V_D-V_S}+\frac{g_{mRX1}x_{rX1}}{4}$	$4.319 \cdot 10^{-7}$
x_{fX1}	$\sqrt{IC_{FX1}+\frac{1}{4}}$	2.19
x_{rX1}	$\sqrt{IC_{RX1}+\frac{1}{4}}$	0.5

Go to mosEKVplots_index

Last project update: 2021-10-31 21:09:13

Figure 4.54: SLiCAP html page with circuit data.

5 Basic amplification: CS stage

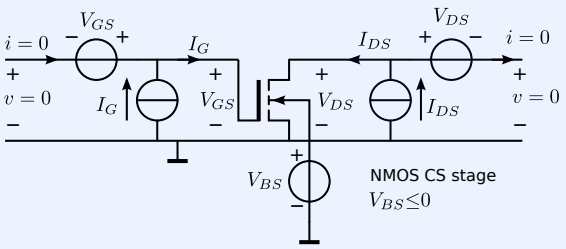
5.1 Introduction

In this chapter, we will study the amplifying capabilities of a MOS transistor. We will do this for the MOS in common-source (CS) configuration. A common-source, or CS amplifier stage is a biased MOS transistor of which the source has been taken as common terminal for both the input port and the output port. At a later stage it will become clear that all other MOS stages can be derived from the CS stage with the aid of error reduction techniques. At this stage, we will simply assume that the CS stage can be considered as the basic MOS amplifier stage.

5.1.1 The CS stage

Figure 5.1 shows the MOS CS stage biased as proposed in Chapter 3.4.

Figure 5.1: Biased NMOS CS stage.



The biased gate-source is the input port and the biased drain-source is the output port of the stage. In *this* CS configuration the source is assumed to carry no signal with respect to the bulk, however, the bias voltage of the source may differ from that of the bulk.[1] For an NMOS device $V_{BS} \leq 0$. For a PMOS device all polarities of the bias sources are opposite to those of an NMOS device.

[1] In floating CS stages, the source may also carry signal with respect to the bulk.

5.1.2 This chapter

The ideally biased CS stage from Figure 5.1 will be referred to as the *intrinsic CS stage*. An intrinsic CS stage has no source and load connected to it. Its behavior will be studied in section 5.2. There, we will evaluate its transmission-1 matrix parameters A, B, C and D. Studying the behavior of the intrinsic CS stage is helpful in understanding the source-to-load signal transfer of a CS stage driven from and terminated with finite non-zero impedances. The behavior of the CS stage between source and load will be discussed in section 5.3. The primary focus will be on source and load impedances that consists of parallel RC networks. This is because such drive and load conditions approximate those present in many applications. However, the theory can as well be applied for other source and load types.

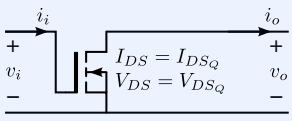


Figure 5.2: Simplified circuit diagram of the intrinsic CS stage.

5.2 The intrinsic CS stage

The intrinsic CS stage is a CS stage biased with ideal sources. For the sake of simplicity we leave the bias sources out of the schematics. The fact that the CS stage is based in a quiescent operating point Q will be indicated by denoting the relevant operating point quantities in the schematics. This has been shown in Figure 5.2, where the quiescent operating voltage and current of the output port have been selected as operating quantities that somehow need to be fixed.

In the following example, we will demonstrate the biasing of a CS stage according to the method discussed in Chapter 3.

Example 5.1

In this example, we will determine the input port bias quantities of a CS stage for biasing its output port at: $I_{DS_Q} = 10\mu A$, $V_{DS_Q} = 0.9V$ *and* $V_{SB_Q} = 0V$. *The transistor is an NMOS with* $W = 220nm$ *and* $L = 180nm$, *fabricated in a typical CMOS18 process.*

Figure 5.3 gives the circuit for determination of V_{GS_Q} *and* I_{GS_Q}. *Its* SPICE *netlist is shown below:*

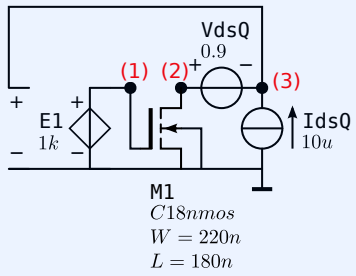


Figure 5.3: Circuit for determination of the input bias sources of a C18 nmos transistor with $W = 220$nm and $L = 180$nm, biased at $I_{DS_Q} = 10\mu$A, $V_{DS_Q} = 0.9$V and $V_{SB} = 0$V.

```
CSbias0_9V-10uA
* file: CSbias0_9V-10uA
* Spice circuit file
.include CMOS18TT.lib
M1 2 1 0 0 C18nmos W=220n L=180n
VdsQ 2 3 0.9
IdsQ 0 3 10u
E1 1 0 3 0 1k
.op
.end
```

After running SPICE, *the operation point information can be found in the output file. It shows:* $V_{GS_Q} = 643.709mV$ *and* $I_{GS_Q} = 0$. *Since the gain of* `E1` *equals* 1000, *a small error voltage of* $V_{GS_Q}/1000$ *is expected at the output port. This error voltage can be reduced by increasing the gain of* `E1`.

In the next example, we will apply the above results to the biasing of the CS stage.

Example 5.2

Figure 5.4 shows the biased stage. A resistor has been added to prevent from having a floating node at the input port. Its value is large enough not to affect the properties of the stage in practical situations. The resulting operating voltages at the input port the output port have been shown in the circuit. They approximate zero; as expected.

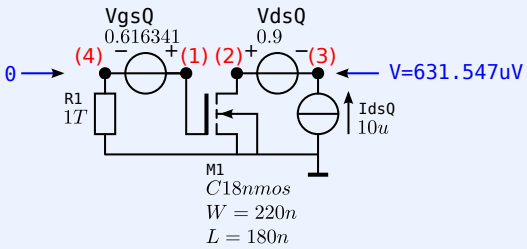


Figure 5.4: Biased CS stage with a C18 nmos transistor with $W = 220$nm and $L = 180$nm, biased at $I_{DS_Q} = 10\mu$A, $V_{DS_Q} = 0.9$V and $V_{SB} = 0$V.

The SPICE *netlist file of this circuit has been shown below:*

```
CSbiased0_9V-10uA
* file: CSbiased0_9V-10uA
* Spice circuit file
.include CMOS18TT.lib
M1 2 1 0 0 C18nmos W=220n L=180n
VdsQ 2 3 0.9
IdsQ 0 3 10u
VgsQ 1 4 643.709m
R1 4 0 1T
.op
.end
```

In the following sections, we will discuss various performance aspects of the intrinsic CS stage. In section 5.2.1, we will discuss the instantaneous behavior of the intrinsic CS stage. In section 5.2.2, we will discuss its small-

signal dynamic behavior and in section 5.2.3, its large-signal dynamic behavior.

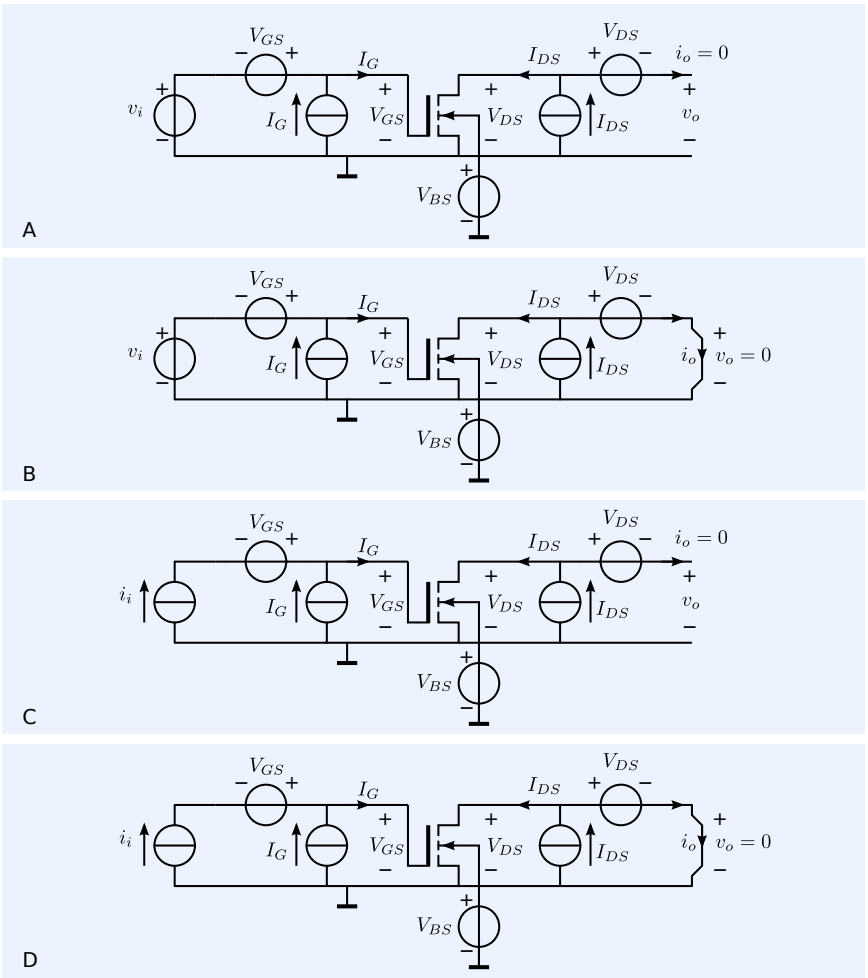


Figure 5.5: Simulation test benches for determination of the transfer characteristics:
A. Circuit for determination of $v_o = v_{oQ}(v_i)$
B. Circuit for determination of $i_o = i_{oQ}(v_i)$
C. Circuit for determination of $v_o = v_{oQ}(i_i)$
D. Circuit for determination of $i_o = i_{oQ}(i_i)$

In those sections, the following transfer characteristics have our interest:

1. Voltage transfer $v_o = v_{oQ}(v_i)$ (voltage driven input, open output)
2. Voltage-to-current transfer $i_o = i_{oQ}(v_i)$ (voltage driven input, shorted output)
3. Current-to-voltage transfer $v_o = v_{oQ}(i_i)$ (current driven input, open output)
4. Current transfer $i_o = i_{oQ}(i_i)$ (current driven input, shorted output)

Simulation test benches for determination of these characteristics are shown in Figure 5.5.

The voltage transfer, the current-to-voltage transfer, the voltage-to-current transfer and the current transfer can be determined with the aid of the circuits from Figure 5.5A through D. The four characteristics together fully describe the behavior of the intrinsic CS stage. The nonlinear dynamic input and output characteristics of this stage, that also may have our interest, can be derived from these characteristics. As an example, the input $v_i(i_i)$ or $i_i(v_i)$ characteristic with shorted output can be derived from the $i_o(i_i)$ transfer and the $i_o(v_i)$ transfer. They may as well be determined with dedicated simulation test benches.

In practice, only the instantaneous transfer characteristics can be measured using ideal drive and termination conditions.[2] At high frequencies, such ideal drive and termination conditions can never be met. For this reason, the measurement of the dynamic transfer parameters is usually performed with S-parameter test equipment, using 50Ω drive and termination impedances.

In section 5.2.4 we will study the noise behavior of the intrinsic CS stage.

[2] Ideal drive and termination conditions: $Z_s = 0$ or $Z_s = \infty$ and $Z_\ell = 0$ or $Z_\ell = \infty$

5.2.1 Instantaneous behavior

In this section, we will discuss the instantaneous nonlinear transfer of a CS stage. We will confine ourselves to the following transfer characteristics:

1. Voltage transfer $v_o = v_{oQ}(v_i)$ (voltage driven input, open output)
2. Voltage-to-current transfer $i_o = i_{oQ}(v_i)$ (voltage driven input, shorted output)
3. Output $v - i$ characteristic $v_o = v_{oQ}(i_o)$

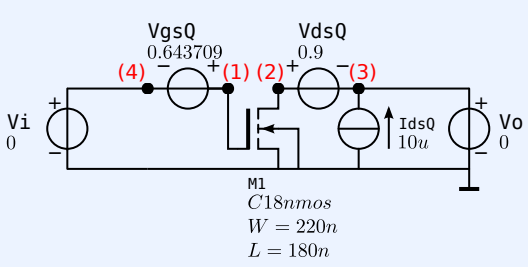
   In case of a nonzero instantaneous (DC) reverse transfer, the output $v - i$ characteristic will depend on the resistive termination of the input port. The intrinsic MOS transistor in CS configuration has zero instantaneous reverse transfer because both the DC gate current and the DC gate voltage do not depend on the DC drain voltage and the DC drain current. Parasitic DC feedback from the output port to the input port, however, may introduce a non-zero reverse transfer. Such feedback can for example be caused by a non-zero resistance of the source, or by a gate leakage current that depends on the drain voltage. In most cases such effects can be neglected for taking early stage design decisions.

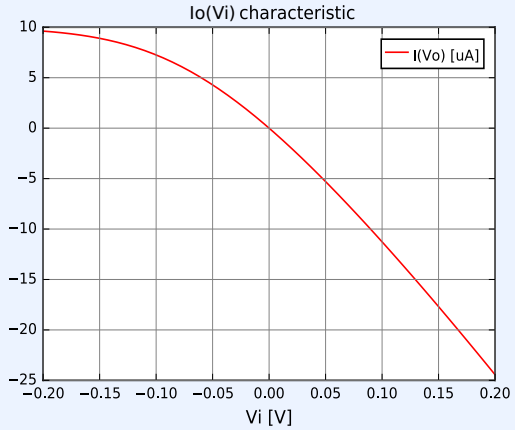
Due to the absence of a DC gate current, the DC input resistance, the DC current to voltage transfer as well as the DC current gain of the CS stage are infinite. Hence, they need no further study here.

The study of the instantaneous behavior of the intrinsic CS stage will be performed with the aid of the biased CS stage from Figure 5.4.

Voltage-to-current transfer

Figure 5.6 shows the SPICE simulation test bench for the determination of the voltage-to-current transfer of the intrinsic CS stage from Figure 5.4.





The graph shows the simulation results for $v_i = -0.2 \cdots 0.2$V. It clearly shows that the $i_o\,(v_i)$ characteristic of the properly biased transistor passes

Figure 5.6: Left: Circuit for determination of the $i_o(v_i)$ characteristic of the biased CS stage from Figure 5.4
Right: $I_o(V_i)$ characersitic.

through the origin. For negative input voltage excursions the current through the MOS decreases and the portion of `IdsQ` that flows through the shorted output increases until its maximum value of 10μA is achieved. The transconductance drops with the current through the MOS transistor. For positive excursions the transconductance does not increase noticeably with I_{DS} due to the VFMR and the velocity saturation. The maximum current that can be delivered to the shorted output is now limited by `VdsQ` and the channel resistance of the MOS transistor.

The netlist of this circuit is:

```
CSbiased0_9V-10uAViIo
* file: CSbiased0_9V-10uAViIo
* Spice circuit file
.include CMOS18TT.lib
M1 2 1 0 0 C18nmos W=220n L=180n
VdsQ 2 3 0.9
IdsQ 0 3 10u
VgsQ 1 4 643.709m
Vi   4 0 0
Vo   3 0 0
.dc Vi -0.1 0.1 10m
.end
.graph "Vo#p"
```

Voltage transfer

Figure 5.7 shows the SPICE simulation test bench for the determination of the voltage transfer of the intrinsic CS stage from Figure 5.4. The graph shows the simulation results for $v_i = -0.1 \cdots 0.1$V. It clearly shows that the $v_o\,(v_i)$ characteristic of the properly biased transistor passes through the origin. For small voltage excursions: $v_i = -0.02 \cdots 0.02$V, the voltage transfer is approximately linear. At large negative voltage excursions both DIBL and increased leakage of the drain bulk diode limit the positive voltage excursion. At large positive voltage excursions, the transistor operates in the linear region and the output voltage is limited to $-V_{DS_Q} = -0.9$V.

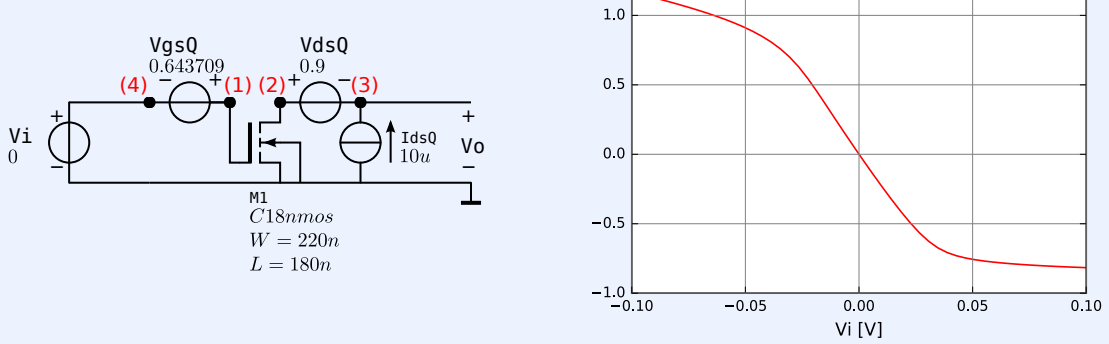


Figure 5.7: Left: Circuit for determination of the $v_o(v_i)$ characteristic of the biased CS stage from Figure 5.4
Right: $V_o(V_i)$ charactersitic.

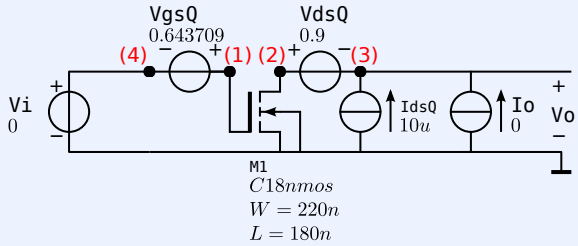
The netlist of this circuit is:

```
CSbiased0_9V-10uAViIo
* file: CSbiased0_9V-10uAViIo
* Spice circuit file
.include CMOS18TT.lib
M1 2 1 0 0 C18nmos W=220n L=180n
VdsQ 2 3 0.9
IdsQ 0 3 10u
VgsQ 1 4 643.709m
```

```
Vi   4 0 0
.dc Vi -0.05 0.05 1m
.end
.graph "3"
```

Output characteristics

Figure 5.8 shows the SPICE simulation test bench for the determination of the output $v-i$ characteristic at shorted input of the intrinsic CS stage from Figure 5.4. The graph shows the simulation results for $i_i = -20\mu\text{A}\cdots 20\mu\text{A}$. It clearly shows that the $v_o\,(i_o)$ characteristic of the properly biased transistor passes through the origin. For small current excursions: $i_o = -3\mu\text{A}\cdots 3\mu\text{A}$, the output $v-i$ characteristic is approximately linear. At large positive voltage excursions both DIBL and increased leakage of the drain bulk diode limit the positive voltage excursion. At large negative voltage excursions the transistor operates in the linear region and the output voltage is limited to $-V_{DS_Q} = -0.9\text{V}$.



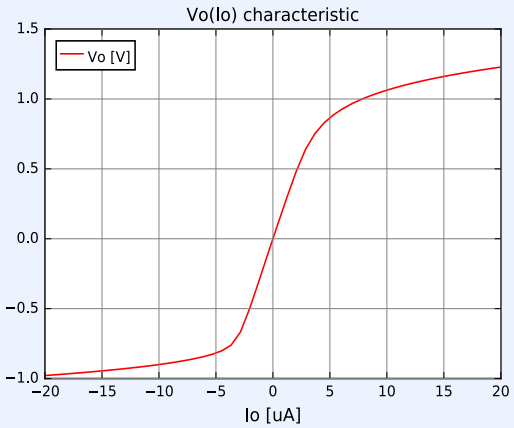


Figure 5.8: Left: Circuit for determination of the $v_o(i_o)$ characteristic of the biased CS stage from Figure 5.4
Right: $V_o(I_o)$ characersitic.

The netlist of this circuit is:

```
CSbiased0_9V-10uAVoIo
* file: CSbiased0_9V-10uAVoIo
* Spice circuit file
.include CMOS18TT.lib
M1 2 1 0 0 C18nmos W=220n L=180n
VdsQ 2 3 0.9
IdsQ 0 3 10u
VgsQ 1 4 0.643709
Vi   4 0 0
Io   0 3 0
.dc Io -20u 20u 1u
.end
```

The output $v-i$ characteristic at open input is obtained after replacing the voltage source `Vi` with a high-value resistor that prevents from having a floating node at the input port (see Figure 5.3). For the given `C18nmos` model this characteristic equals the one obtained with shorted input.

Conclusions

At this stage we may draw the following conclusions concerning the large signal instantaneous behavior of the biased CS stage:

1. When properly biased, the transfer characteristics and the $v-i$ characteristics of the input port and the output port all pass through the origin.

2. The maximum positive current that can be delivered to the output port equals I_{DS_Q}. This current is called the maximum source current: the maximum current that can be sourced or provided to the load.

3. The maximum negative current that can be delivered to the output port depends on V_{DS_Q} and the channel resistance of the MOS transistor. This current is called the maximum sink current: the maximum current that can be taken from the load.

4. The maximum positive voltage that can be delivered to the load is limited by the breakdown voltage of the source-bulk diode at I_{DS_Q} or by the DIBL effect and I_{DS_Q}.

5. The maximum negative voltage that can be delivered to the load equals $-V_{DS_Q}$.

5.2.2 Small-signal dynamic behavior

In this section, we will study the small-signal dynamic behavior of the intrinsic CS stage. We will derive expressions for the transmission-1 matrix parameters A, B, C and D. Expressions for the input impedance with shorted and open output and expressions for the output impedance with open and shorted input can be derived from these matrix parameters. They can also directly be obtained using network analysis and we will give circuits to do so.

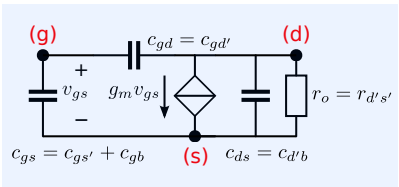


Figure 5.9: Hybrid-pi equivalent circuit of the intrinsic CS stage that can be used if the source carries no signal with respect to the bulk. The model is a simplified version of the one shown in Figure 4.31: the bulk resistances have been omitted.

The small-signal dynamic behavior of the intrinsic CS stage can be obtained from network analysis using the hybrid-pi equivalent circuit of the MOSFET. The small-signal model of the MOS has been shown in Figure 4.31. If we neglect the bulk resistances and if the source carries no signal with respect to the bulk, the model can be simplified to the one shown in Figure 5.9. Such simplifications are usually justified for taking early stage design decisions.

Voltage gain factor

Figure 5.10 shows the network for the determination of the small-signal voltage gain factor μ of the CS stage.

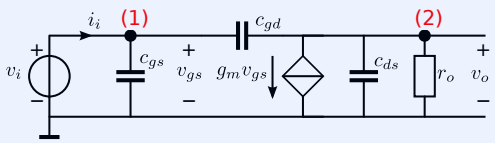


Figure 5.10: Circuit for determination of the small-signal voltage gain factor of the biased CS stage.

The voltage gain factor is obtained as

$$\mu = \frac{1}{A} = \left.\frac{v_o}{v_i}\right|_{i_o=0} = -g_m r_o \frac{1 - s\frac{c_{gd}}{g_m}}{1 + s r_o \left(c_{gd} + c_{ds}\right)}. \tag{5.1}$$

It has a right half plane zero at $s = \frac{g_m}{c_{gd}}$, and a pole at $s = -\frac{1}{r_o\left(c_{gd}+c_{ds}\right)}$. If the device is operating in the saturation region, the frequency of the zero exceeds that of the pole.

Transadmittance factor

Figure 5.11 shows the network for the determination of the small-signal transadmittance factor γ of the CS stage.

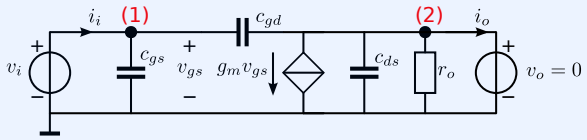


Figure 5.11: Circuit for determination of the small-signal transadmittance factor of the biased CS stage.

The transadmittance factor is obtained as

$$\gamma = \frac{1}{B} = \left.\frac{i_o}{v_i}\right|_{v_o=0} = -g_m\left(1 - s\frac{c_{gd}}{g_m}\right). \tag{5.2}$$

It has a right half plane zero at $s = \frac{g_m}{c_{gd}}$.

Transimpedance factor

Figure 5.12 shows the network for the determination of the small-signal transimpedance factor ζ of the CS stage.

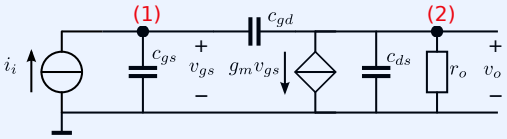


Figure 5.12: Circuit for determination of the small-signal transimpedance factor of the biased CS stage.

The transimpedance factor is obtained as

$$\zeta = \frac{1}{C} = \left.\frac{v_o}{i_i}\right|_{i_o=0}; \tag{5.3}$$

$$= -g_m r_o \frac{1 - s\frac{c_{gd}}{g_m}}{s\left(c_{gs} + (1 + g_m r_o)\, c_{gd}\right)\left(1 + s\frac{r_o\left(c_{gs}c_{ds} + c_{gs}c_{gd} + c_{ds}c_{gd}\right)}{c_{gs} + (1 + g_m r_o)c_{gd}}\right)}. \tag{5.4}$$

It has a right half plane zero at $s = \frac{g_m}{c_{gd}}$, and two poles. One pole is found in the origin of the s-plane; at: $s = 0$ and a second pole is found at

$$s = -\frac{c_{gs} + (1 + g_m r_o)\, c_{gd}}{r_o\left(c_{gs}c_{ds} + c_{gs}c_{gd} + c_{ds}c_{gd}\right)}; \tag{5.5}$$

Current gain factor

Figure 5.13 shows the network for the determination of the small-signal current gain factor α of the CS stage.

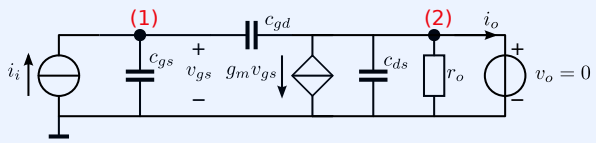


Figure 5.13: Circuit for determination of the small-signal current gain factor of the biased CS stage.

The current gain factor is obtained as

$$\alpha = \frac{1}{D} = \left.\frac{i_o}{i_i}\right|_{v_o=0} = -\frac{g_m\left(1 - s\frac{c_{gd}}{g_m}\right)}{s\left(c_{gs} + c_{gd}\right)}. \tag{5.6}$$

It has a right half plane zero at $s = \frac{g_m}{c_{gd}}$, and a pole in the origin. In the saturation region the capacitance c_{gs} is much larger than c_{gd}. The unity-gain frequency ω_T of the current gain is then obtained as

$$\omega_T = \frac{g_m}{c_{gs} + c_{gd}}. \tag{5.7}$$

This frequency is called the cut-off frequency of the transistor.

Input impedance

Figure 5.14A and Figure 5.14B show the circuits for determination of the input impedance with open output and shorted output, respectively.

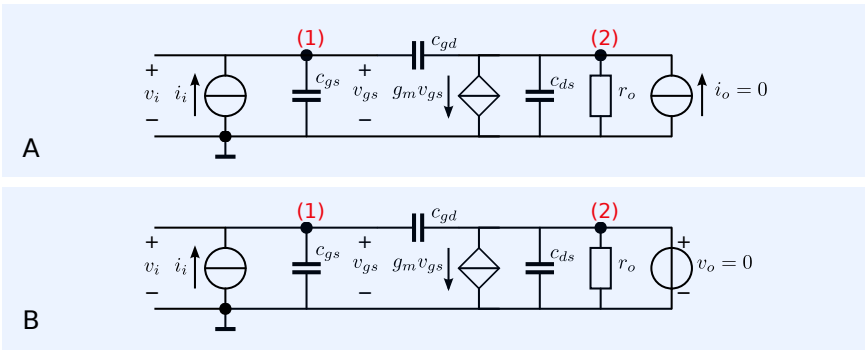


Figure 5.14: Left: Circuits for determination of the input impedance of the biased CS stage:
A. With open output
B. With shorted output.

The input impedance with *open output* can be obtained as

$$z_i|_{i_o=0} = \left.\frac{v_i}{i_i}\right|_{i_o=0} = \frac{A}{C}, \tag{5.8}$$

which yields:

$$z_i|_{i_o=0} = \frac{1 + s r_o \left(c_{gd} + c_{ds}\right)}{s \left(c_{gs} + (1 + g_m r_o)\, c_{gd}\right) \left(1 + s \frac{r_o \left(c_{gs} c_{ds} + c_{gs} c_{gd} + c_{ds} c_{gd}\right)}{c_{gs} + (1 + g_m r_o) c_{gd}}\right)}. \tag{5.9}$$

The poles equal those of the transimpedance factor. It can be shown that the zero is located between the two poles (see section 5.3.5). Thus, at frequencies below that of the zero, the input is capacitive. At frequencies between the frequency of the zero and the frequency of the second pole, the input impedance is resistive, and at frequencies above those of the second pole the input impedance is capacitive.

The input impedance with *shorted output* can be obtained as

$$z_i|_{v_o=0} = \frac{B}{D} = \frac{1}{s \left(c_{gs} + c_{gd}\right)}. \tag{5.10}$$

It consists of the parallel connection of c_{gs} and c_{gd}. This input capacitance at shorted output is often referred to as c_{iss}:

$$c_{iss} = c_{gs} + c_{gd}. \tag{5.11}$$

Output impedance

Figure 5.15A and Figure 5.15B show the circuits for determination of the output impedance with shorted input and open input, respectively.

The output impedance with *shorted input* is obtained as the parallel con-

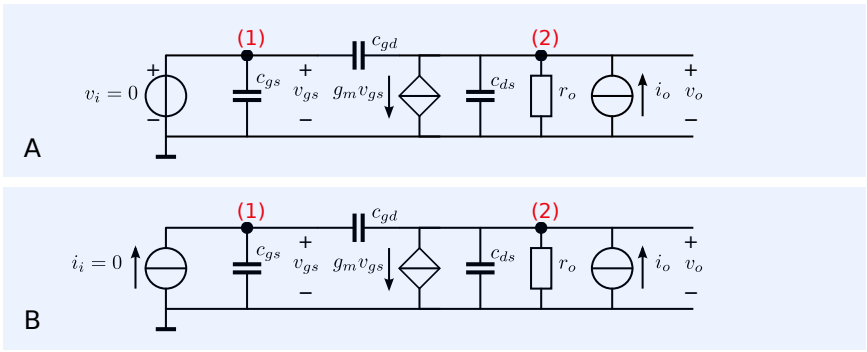


Figure 5.15: Left: Circuits for determination of the output impedance of the biased CS stage:
A. With shorted input
B. With open input.

nection of r_o and the output capacitance at shorted input c_{oss}:

$$z_o|_{v_i=0} = \frac{B}{A} = \frac{r_o}{1 + s r_o c_{oss}}, \tag{5.12}$$

where

$$c_{oss} = c_{gd} + c_{ds}. \tag{5.13}$$

The output impedance with *open input* can be obtained as

$$z_o|_{i_i=0} = \frac{D}{C}, \tag{5.14}$$

which yields

$$z_o|_{i_i=0} = r_o \frac{c_{gs} + c_{gd}}{c_{gs} + (1 + g_m r_o)\, c_{gd}} \frac{1}{1 + s \frac{r_o (c_{gs} c_{ds} + c_{gs} c_{gd} + c_{ds} c_{gd})}{c_{gs} + (1 + g_m r_o) c_{gd}}}. \tag{5.15}$$

It consists of a parallel RC network with

$$R = r_o \frac{1}{1 + g_m r_o \frac{c_{gd}}{c_{gs} + c_{gd}}}, \tag{5.16}$$

and

$$C = \frac{c_{gs} c_{ds} + c_{gs} c_{gd} + c_{ds} c_{gd}}{c_{gs} + c_{gd}}. \tag{5.17}$$

If the transistor is operating in the saturation, and $c_{gs} \gg c_{gd}$, the output resistance R with open input approximates

$$R = \frac{1}{g_m \frac{c_{gd}}{c_{gs}} + \frac{1}{r_o}}. \tag{5.18}$$

Under these conditions, the output capacitance C approximates c_{oss}.

Conclusions

At this stage we may draw the following conclusions from the small-signal analysis above:

1. Right half plane zeros
   (a) The four gain factors μ, γ, ζ and α are inverting and have a right half plane zero.
   (b) This zero is the consequence of the inverting transconductance g_m and the transfer through the gate-drain capacitance c_{gd}.

(c) If the transistor is operating in the saturation region the frequency of this zero exceeds ω_T and its influence on the magnitude characteristics for frequencies below ω_T can often be ignored.

2. Poles

   (a) The voltage gain has one pole at $s = -\frac{1}{r_o\left(c_{gd}+c_{ds}\right)}$

   (b) The transadmittance factor can be considered instantaneous up to the frequency of the right half plane zero.

   (c) The transimpedance factor has two poles, of which one is found at $s = 0$.

   (d) The current gain has one pole at $s = 0$. The unity-gain frequency ω_T of the current gain factor is called the cut-off frequency of the transistor.

3. Input impedance

   (a) The input impedance of a shorted CS stage is purely capacitive.

   (b) The input impedance of a CS stage of which the output has been left open has two poles. These poles equal those of the transimpedance factor.

4. Output impedance

   (a) The output impedance of a CS stage of which the input is left open (current-driven) is resistive up to high frequencies. If the MOS transistor operates in the saturation region, the output resistance of a current-driven CS stage approximates $1/g_m$ up to frequencies above ω_T.

   (b) The output impedance of a CS stage with shorted input consists of r_o in parallel with $c_{ds} + c_{gd}$.

5.2.3 Large-signal dynamic behavior

In this section, we will briefly discuss some important aspects of the large signal dynamic behavior of the intrinsic CS stage. We will only give a brief summary because the large-signal behavior of a system cannot easily be analyzed using simple hand calculations. Modeling of the nonlinear dynamic behavior of a system requires the use of nonlinear differential equations. Such equations can usually only be solved numerically. The transient analysis technique in SPICE can be used for this purpose.

In some special cases it is possible to obtain clear design information for the nonlinear dynamic behavior.

Slew rate limitation

Limitation of the rate of change of a voltage or current is called slew rate limitation.

1. Voltage limiting across inductors causes current slew rate limitation.

   If a voltage change across an inductor with inductance L is limited to $V_{\max}$, the rate of change of the current $\frac{dI_L}{dt}$ through that inductor is limited to $\frac{V_{\max}}{L}$.

   - In section 5.2.1, we have studied voltage limiting mechanisms at the output of the CS stage.

2. Current limitation through capacitors causes voltage slew rate limitation.

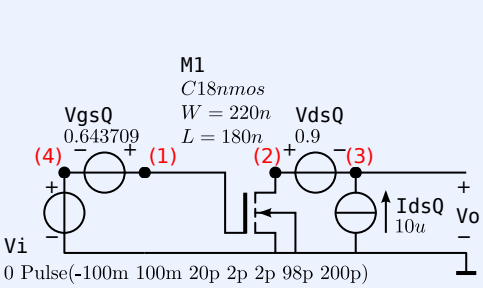
   If a current change through a capacitor with capacitance C is limited to $I_{\max}$, the rate of change of the voltage $\frac{dV_C}{dt}$ across that capacitor is limited to $\frac{I_{\max}}{C}$.

   - In section 5.2.1, we have studied current limiting mechanisms at the output port. Since the output source current is limited by the bias current I_{DS_Q} and the CS stage has a nonzero output capacitance we may expect slew rate limitation for rising output voltages.
   - In section 5.2.2, we have seen that the input impedance of a CS stage has a pole at $s = 0$. For low frequencies the input port thus behaves capacitively. If such a stage is driven from a current source with limited drive capabilities, input voltage slew rate limitation will occur.

In the following example we will determine the slew rate limitation of the output voltage of the biased CS stage from Figure 5.4.

Example 5.3

In this example we will study the large-signal pulse response of the voltage transfer of the biased CS stage from Figure 5.4.



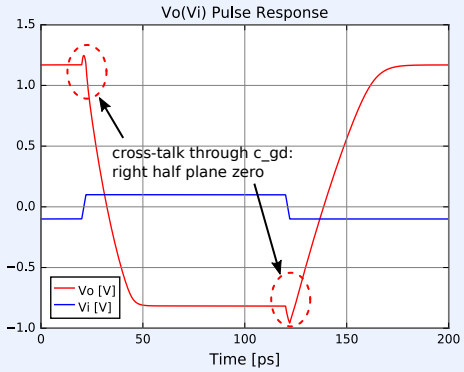


Figure 5.16: Left: Circuit for determination of the voltage pulse reponse of the biased CS stage from Figure 5.4
Right: Input voltage and output voltage of the biased CS stage.

Figure 5.16 shows the circuit. A square wave input voltage of $\pm 100mV$ with a rise and fall time of 2ps and a period of 200ps has been applied to the input of the biased CS stage. Both the input and the output voltage of the biased CS stage have been shown in the graph. The figure clearly shows that the rise time exceeds the fall time. This is because the current sinking capability of the biased CS stage is larger than its current sourcing capability. The latter one is limited to the bias current of $10\mu A$. This causes slew rate limitation. The graph also shows the effect of the right half plane zero caused by forward transfer through c_{gd}. This transfer is caused by the coupling between the input and the output port. It is often referred to as crosstalk *between input and output port. This crosstalk differs for positive output and negative output voltage. This is because when the output voltage is negative, the transistor operates in the linear region, while at positive output it operates in the saturation region. In the linear region c_{gd} is larger than in the saturation region, while g_m is smaller.*

The netlist of the circuit is:

```
CSbiased0_9V-10uAViVoPulse
* file: CSbiased0_9V-10uAViVoPulse
* Spice circuit file
.include CMOS18TT.lib
M1 2 1 0 0 C18nmos W=220n L=180n
VdsQ 2 3 0.9
IdsQ 0 3 10u
VgsQ 1 4 0.643.709
Vi   4 0 0 pulse(-100m 100m 20p 2p 2p 98p 200p)
```

```
.tran 0 200p 0 0.2p
.end
```

Operating point shift

Even order nonlinearity gives rise to operating point shift. This effect can be observed in the graph shown in Figure 5.16: the mean value of the input voltage is zero, while the mean value of the output voltage is about 58mV. Since the properties of the CS stage strongly depend on its bias point, operating point shift should be kept as small as possible. At a later stage we will introduce techniques for creating transfer functions with predominantly odd order nonlinearity.

5.2.4 Noise behavior

In this section, we will discuss the modeling of the noise behavior of the CS stage. The small-signal noise model of the MOS transistor has already been introduced in Figure 4.32. Figure 5.17 shows a simplified small-signal noise model for the CS stage. In this model, the source is assumed to carry no signal with respect to the bulk and the drain and source series resistances have been omitted. These simplifications are usually justified for taking early stage design decisions.

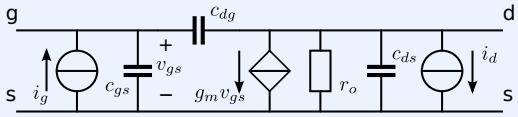


Figure 5.17: Small signal noise model of the intrinsic CS stage.

A more design friendly model would be one that has the noise sources modeled at the input port of the CS stage. With such a model we can find design equations and show stopper values for the noise contribution of the CS stage to the total source-referred noise.

The equivalent input noise sources of the CS stage can be found by multiplying the output noise current with the small-signal transmission parameters B and D of the CS stage (see Chapter 2). Multiplication with B yields a contribution to the equivalent input voltage noise, while multiplication with D yields a contribution to the equivalent input current noise. The transmission parameters B and D of the intrinsic CS stage have been evaluated in section 5.2.2. If we ignore the influence of the right half plane zero in the transmission parameters, we may write

$$B = -\frac{1}{g_m}, \tag{5.19}$$

$$D = -\frac{s}{\omega_T}. \tag{5.20}$$

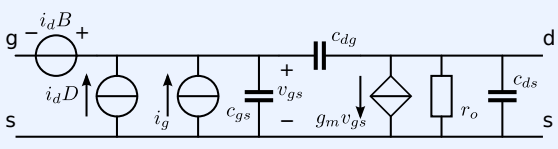


Figure 5.18: Small signal noise model of the intrinsic CS stage with equivalent input noise sources.

The small-signal noise model with equivalent input noise sources is shown in Figure 5.18.

At a later stage, we will show that, if designed properly, the noise performance of an amplifier of which the first stage is equipped with a CS stage is determined by the noise contribution of this CS stage. If the amplifier is a negative feedback amplifier, also the noise contributions of the feedback elements as well as their influence on the noise contributions of the CS stage to the source-referred noise have to be considered. Hence the noise performance of an amplifier can be designed with the aid of a noise model of the CS stage that comprises its equivalent input noise sources only.

SLiCAP MOS EKV noise model

Figure 5.19 shows a simplified EKV noise model. This model is available as a SLiCAP sub-circuit. It can be used for determination of the device geometry and the drain current of a MOS transistor for minimum noise contribution.

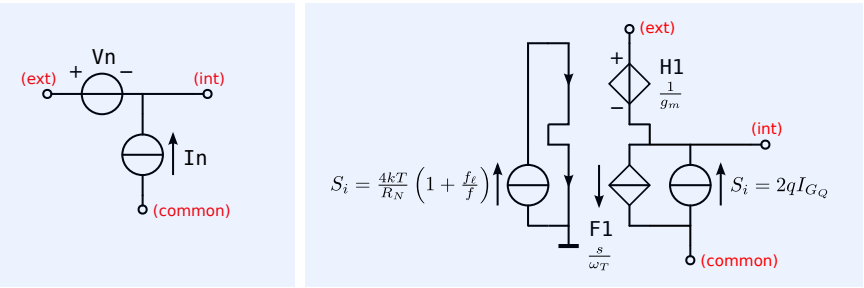


Figure 5.19: SLiCAP equivalent input noise sources of the CS stage.

Below netlist of the sub circuit `NM18_noise` that models the noise of an NMOS in a typical CMOS18 process as a function of the device width W and length L, the gate leakage current I_G and the drain bias current I_D. It can be found in the SLiCAP library file `SLiCAP.lib`.

```
457 + c_pi = {tau_F*g_m + CJE}
458 + c_bc = {CJC*(1+((VCE-VBE)/VJC)^-MJC)} ; Divide by zero error if VBE-VCE=VJC
459 + VBE  = {U_T*ln(IC/Is)}
460 * technology parameters (SI units)
461 + Is      = 1e-16
462 + VAF     = 50
463 + beta_AC = 100
464 + VJC     = 0.7
465 + MJC     = 0.33
466 + CJC     = 1p
467 + CJE     = 5p
468 + tau_F   = 1n
469 .ends
470
471 .subckt NM18_noise ext comm int ID={ID} IG={IG} W={W} L={L}
472 * intrinsic noise sources, gate resistance should be added externally
473 * CMOS18 N device: copy and modify this model for other devices
474 I1 0 1 I value=0 noise={4*k*T/R_N*(1+(f_ell/f)^AF_N18)} ; channel noise
        current
475 H1 ext int 1 10 {1/g_m}                    ; equivalent-input voltage
        noise
476 F1 ext comm 10 0 {s/2/pi/f_T}              ; gate-induced noise
477 I2 ext comm I value=0 noise={2*q*IG}       ; gate shot noise
478 .param
```

Lines 460 through 463 describe the netlist with circuit elements, lines 465 through 477 list the EKV model noise equations. The CMOS 18 technology parameters and global constants are defined in the SLiCAP library file `SLiCAPmodels.lib`. The device equations and the technology parameters have been adapted from Binkley [Binkley2008][3]. They have been discussed in Chapter 4.

Figure 5.20 shows a test bench for determination of the equivalent input voltage noise of an NMOS according to this model. The effect of the channel length on the corner frequency of the flicker noise can clearly be observed by maintaining the drain current and the aspect ratio $\frac{W}{L}$, while sweeping W.

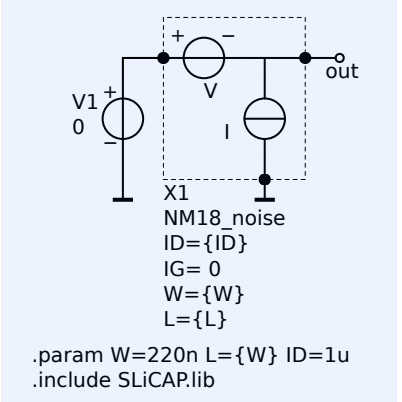


Figure 5.20: SLiCAP test bench for determination of the input voltage noise of a MOS transistor. The circuit uses the subcircuit `NM18_noise` from the SLiCAP library.

[3] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

The SLiCAP file for plotting the input voltage noise is listed below:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
N18noise.py
"""
from SLiCAP import *
prj = initProject('N18noise');
makeNetlist('N18noise.asc', 'CS stage noise');
i1 = instruction();
i1.setCircuit('N18noise.cir')
#i1.defPar('AF_N18', 1)
htmlPage('Circuit data')
img2html('N18noise.svg', 500)
netlist2html('N18noise.cir')
elementData2html(i1.circuit)
params2html(i1.circuit)
i1.setSimType('numeric');
i1.setGainType('vi');
i1.setDataType('noise');
i1.setSource('V1');
i1.setDetector('V_out');
htmlPage('Voltage noise NMOS')
# Plot S_vi for ID=10uA, step W and L, while keeping W/L=1
i1.defPar('ID','10u')
i1.setStepVar('W')
i1.setStepMethod('list')
i1.setStepList([0.2e-6, 0.5e-6, 1e-6, 2e-6, 5e-6, 10e-6, 20e-6, 50e-6])
i1.stepOn()
noiseResult = i1.execute()
SvN18mos = plotSweep('SvN18mos', 'Svi V/rt(Hz) NMOS W/L=1 ID=10uA',
                     noiseResult, 10, 100e6, 200, funcType='inoise', show=True)
fig2html(SvN18mos, 800)
# Another method is to calculate S_vi(f, W), define it as a circuit parameter
# and plot this parameter.
# Calculate the function and define it as a circuit parameter
i1.delPar('W') # Delete the definition of W, this keeps it a symbolic variable
i1.stepOff()   # Disable parameter stepping to obtain a single expression
inoise_f_W = sp.N(i1.execute().inoise)    # calculate S_vi(f,W)
i1.defPar('Si_f_W', inoise_f_W) # define a circuit parameter for this function
i1.defPar('W', '1u')  # Redefine the parameter W otherwise it cannot be
    stepped
i1.defPar('f', 1)     # Define the parameter f otherwise it cannot be swept
i1.stepOn()           # Enable parameter stepping
i1.setDataType('params')
result = i1.execute()
S_vi_f_W = plotSweep('S_vi_f_W', 'Svi V/rt(Hz) noise NMOS W/L=1 ID=10uA',
                     result, 10, 100e6, 200, funcType='param', axisType='log',
                     sweepVar='f', xUnits='Hz', yVar='Si_f_W', yUnits='$V^2/
                         Hz$',
                     show=True)
fig2html(S_vi_f_W, 800)
# Plot f_T and f_l versus W for W/L=1 and ID=10uA (dataType = 'params')
i1.setDataType('params')
i1.stepOff()
result = i1.execute()
f_T_f_L = plotSweep('f_T_f_L', '$f_T,\, f_{\ell}, NMOS W/L=1 ID=10uA$', result
    ,
                    0.2e-6, 50e-6, 200, funcType='param', axisType='log',
                    sweepVar='W', xUnits='m', xScale='u',
                    yVar=['f_T_XU1', 'f_ell_XU1'], yScale='G', yUnits='Hz',
                    show=True)
fig2html(f_T_f_L, 800)
i1.delPar('f')      # Remove the numeric definition of the frequency
# Plot S_vi for W = L = 50u while stepping ID
i1.defPar('W', '50u')
i1.setDataType('noise');
i1.setStepVar('ID')
i1.stepOn()
i1.setStepMethod('list')
i1.setStepList([10e-6, 20e-6, 50e-6, 100e-6, 1e-3, 2e-3, 5e-3])
i1.stepOn()
noiseResult = i1.execute()
SvN18mos50u = plotSweep('SvN18mos50u', 'Svi V/rt(Hz) noise W=50u', noiseResult
    ,
                        1, 1e6, 200, funcType='inoise', show=True)
fig2html(SvN18mos50u, 800)
```

Figure 5.21 shows the simulation results.

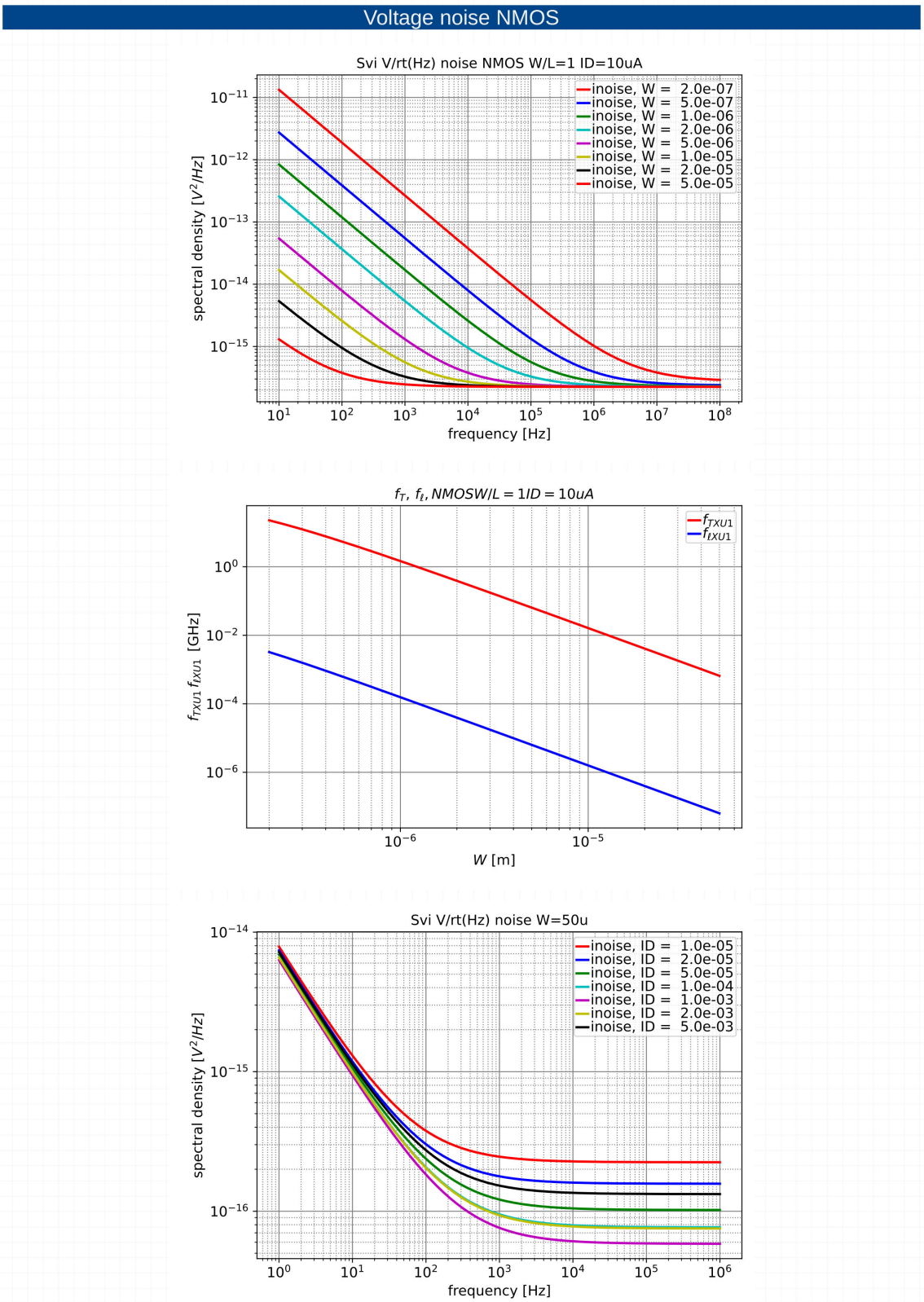


Figure 5.21: SLiCAP noise simulation results of a CMOS18 NMOS.

The upper figure shows the equivalent input voltage noise spectral density for a transistor operating at 10μA. The length and the width of this transistor are stepped simultaneously from 200nm to 50μm. By doing so both f_T and f_ℓ drop with increasing dimensions, as shown in the middle figure. The middle figure clearly shows the fixed ratio vetween f_T and f_ℓ, as derived in (4.194). Since the drain current and the ratio $\frac{W}{L}$ are kept constant, the floor noise does not change, it is determined by g_m and weakly depends on the inversion coefficient. The lower figure shows the way in which the floor noise depends on the current for a transistor that has $W = L = 50\mu$m. At high currents, g_m the transconductance efficiency drops and an increase of the channel current does not result in a significant decrease of the floor noise.

SLiCAP generic MOS and JFET noise model

SLiCAP also has a noise model for a MOS transistor that operates in strong inversion in the saturation region. At a later stage we will see that this operating mode is preferred for low-noise operation at high frequencies. This model does not include effects due to VFMR and velocity saturation and can also be applied for JFETs. The listing of the sub circuit `J_noise` is shown below.

```
+ R_N     = {(1+IC)/(1/2 + 2/3*IC)/N_s_P18/g_m}
+ IC_CRIT = {1/((4*N_s_P18*U_T)*(Theta_P18+1/L/E_CRIT_P18))^2}
+ g_m     = {-ID/(N_s_P18*U_T*sqrt(IC*(1+IC/IC_CRIT)+0.5*sqrt(IC*(1+IC/IC_CRIT
    ))+1))}
+ c_gs    = {2/3*W*L*C_OX_P18 + CGSO_P18*W}
+ c_dg    = {CGSO_P18*W}
+ c_gb    = {CGBO_P18*2*L+(N_s_P18-1)/N_s_P18*C_OX_P18*W*L/3}
+ c_db    = {CJB0_P18*W*LDS_P18}
+ f_T     = {g_m/2/pi/c_iss}
+ c_iss   = {c_gs+c_dg+c_gb}
+ IC_i    = {-ID*L/W/I_0_P18} ; Initial estimate of inversion coefficient
+ IC      = {IC_i*(1+IC_i/4/IC_CRIT)} ; Inversion coefficient corrected for
    short-channel effects
+ f_ell   = {3*KF_P18*g_m*(1+2*N_s_P18*U_T*sqrt(IC)/V_KF_P18)/(8*k*T*W*L*
    N_s_P18*C_OX_P18^2)}
.ends

.subckt J_noise ext comm int ID={ID} IG={IG} W={W} L={L}
* simplified intrinsic noise sources, gate resistance should be added
    externally
* MOS or JFET: copy and modify this model for other devices
I1 0 1 I value=0 noise={8*k*T*g_m/3*(1+f_ell/f)}     ; channel noise current
```

5.3 Small-signal behavior of CS stage between source and load

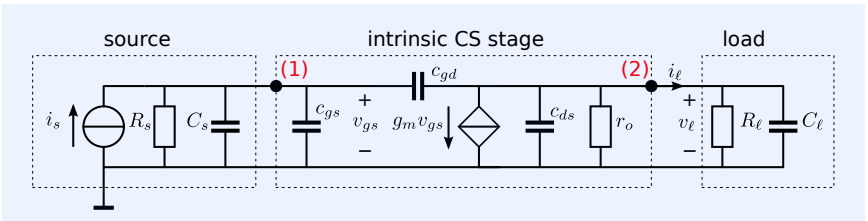


Figure 5.22: CS stage driven from a current source with $R_s \,/\!/\, C_s$ source impedance and loaded with $R_\ell \,/\!/\, C_\ell$.

Now that we have studied the behavior of the intrinsic CS stage, we study the small-signal behavior of a CS stage placed between a source and a load. We will only consider source and load impedances that can be modeled as parallel *RC* networks. This is because such situations usually occur in practice. However, the analysis methods used in this study can easily be applied

for other types of source and load impedances.

Figure 5.22 shows a CS stage driven from a current source with an RC type source impedance and loaded with a parallel RC network. We will study the transimpedance and the current gain of this stage.

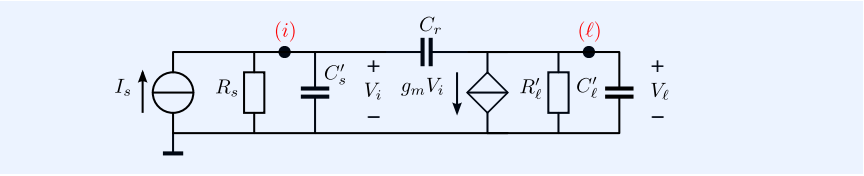


Figure 5.23: Simplified network from Figure 5.22.

In order to obtain relatively simple expressions, we will combine the elements of the circuit as shown in Figure 5.23.

The nodal equations for this network are

$$\begin{pmatrix} I_s \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{R_s} + s(C'_s + C_r) & -sC_r \\ g_m - sC_r & \frac{1}{R'_\ell} + s(C'_\ell + C_r) \end{pmatrix} \begin{pmatrix} V_i \\ V_\ell \end{pmatrix}, \tag{5.21}$$

where

$$C'_s = C_s + c_{gs}, \tag{5.22}$$
$$C'_\ell = C_\ell + c_{ds}, \tag{5.23}$$
$$C_r = c_{gd}, \tag{5.24}$$
$$R'_\ell = \frac{R_\ell r_o}{R_\ell + r_o}. \tag{5.25}$$

5.3.1 Transimpedance

These equations can be solved to obtain the transimpedance $Z_t = \frac{V_\ell}{I_s}$; it is found as

$$Z_t = \frac{-g_m R_s R'_\ell \left(1 - s\frac{C_r}{g_m}\right)}{1 + s\left(C'_s R_s + C'_\ell R'_\ell + C_r\left(R_s + R'_\ell + R_s R'_\ell g_m\right)\right) + s^2 R_s R'_\ell \left(C'_s C'_\ell + C'_s C_r + C'_\ell C_r\right)}. \tag{5.26}$$

The transimpedance has one right-plane zero, which is caused by the transfer through C_r. The two poles of the transimpedance can be found from their product and their sum.

The product of the two poles is found from the coefficient of s^2 in the denominator of (5.26):

$$p_1 p_2 = \frac{1}{R_s R'_\ell \left(C'_s C'_\ell + C'_s C_r + C'_\ell C_r\right)}. \tag{5.27}$$

The sum of the two poles is found from the coefficient of s and from the product of the poles:

$$p_1 + p_2 = -\frac{C'_s R_s + C'_\ell R'_\ell + C_r\left(R_s + R'_\ell + R_s R'_\ell g_m\right)}{R_s R'_\ell \left(C'_s C'_\ell + C'_s C_r + C'_\ell C_r\right)}. \tag{5.28}$$

We will study these expressions later.

5.3.2 Current gain

The current gain A_i of this CS stage can be obtained as

$$A_i = \frac{Z_t}{Z_\ell} = Z_t \frac{1 + sR_\ell C_\ell}{R_\ell}. \tag{5.29}$$

The current gain has a zero at $s = -\frac{1}{R_\ell C_\ell}$, at this complex frequency the load impedance is infinite, hence no current can flow through it. The current gain has the same poles as the transimpedance. This is because they are determined with the same network.

5.3.3 Qualitative description of the dynamic behavior

In the previous sections, we have seen that the transimpedance and the current gain of a CS stage driven from and $R /\!\!/ C$ source and loaded with an $R /\!\!/ C$ load both have two poles. These poles are the solutions of the characteristic equation of the network from Figure 5.23. We will first discuss the dynamic behavior of the CS stage in a qualitative way. This helps us to find conditions for simplifying the characteristic equation and derive design conclusions. We will then elucidate these conclusions with the aid of some examples.

The small-signal dynamic behavior of the CS stage which is driven from and loaded with an $R /\!\!/ C$ network can easily be understood if $C_r = 0$. In this case the network consists of two current-driven $R /\!\!/ C$ networks that show no interaction. This network has two independent capacitor voltages and thus two poles. The poles of the network with $C_r = 0$ are

$$p_1 = -\frac{1}{R_s C_s'},\ p_2 = -\frac{1}{R_\ell' C_\ell'}. \tag{5.30}$$

Since $C_r = 0$, the zero in the right half plane is not present.

A non-zero value of C_r adds a right half plane zero to the transfer, but it does not add a pole because it introduces a loop of capacitors rather than an independent capacitor voltage. The presence of a non-zero capacitance C_r, however, does affect the positions of the two poles. The influence of C_r on the poles strongly depends on the low-frequency voltage gain $\frac{V_\ell}{V_i}$ of the stage.

If the voltage gain of the stage is low, we have $|V_\ell| \ll |V_i|$ and the CS stage can be considered shorted. In that case, the voltage across C_r approcimates V_i, hence, C_r appears to be in parallel with C_s'. If the MOS is operating in the saturation region, we have $c_{dg} \ll c_{ds}$ and thus $C_r \ll C_s'$, and the effect of C_r on the pole positions can usually be neglected.

The situation becomes different if $|V_\ell| \gg |V_i|$ and the CS stage cannot longer be considered as shorted. In this case the voltage across C_r amounts $(1 - A_v)\, V_i$, where $A_v = \frac{V_\ell}{V_i}$. At low frequencies the voltage gain A_v can be approximated as:

$$A_v = -g_m R_\ell' \tag{5.31}$$

The voltage across C_r causes a current I_{C_r} through it:

$$I_{C_r} = V_i \left(1 + g_m R_\ell'\right) sC_r. \tag{5.32}$$

From this expression, it can be seen that it appears as if the voltage V_i causes a current through an input capacitor with capacitance $C_i = \left(1 + g_m R_\ell'\right) C_r$. The increase of the input capacitance due to a non-zero value of C_r by a factor $(1 - A_v)$ is known as the *Miller effect*, named to its discoverer James Miller.[Miller1920][4]

[4] John M. Miller. Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit. *Scientific Papers of the Bureau of Standards*, 15(351):367–385, 1920

5.3.4 Quantitative description of the dynamic behavior

We will now study the above effects in a quantitative way. We will first study the expression for the product of the poles in more detail. If C_r is small with

respect to both C'_ℓ and and C'_s, expression (5.27) can be approximated as

$$p_1 p_2 = \frac{1}{R_s R'_\ell C'_s C'_\ell}. \tag{5.33}$$

Hence, under the above assumptions, the product of the poles does not depend on C_r.

Let us now study the influence of C_r on the sum of the poles as described in (5.28).

If $C_r = 0$ (5.28) changes to:

$$p_1 + p_2 = -\frac{C'_s R_s + C'_\ell R'_\ell}{R_s R'_\ell C'_s C'_\ell} = -\frac{1}{R_s C'_s} - \frac{1}{R'_\ell C'_s}. \tag{5.34}$$

We then have two separate poles, one at $s = -\frac{1}{R_s C'_s}$ and another at $s = -\frac{1}{R'_\ell C'_s}$, which confirms our expectations from the above qualitative description.

If C_r is larger than zero, but small with respect to both C'_ℓ and and C'_s, the sum of the poles can be approximated by

$$p_1 + p_2 = -\frac{1}{R_s C'_s} - \frac{1}{R'_\ell C'_\ell} - \frac{C_r}{C'_\ell}\frac{\frac{R_s + R'_\ell}{R_s R'_\ell} + g_m}{C'_s}, \tag{5.35}$$

which can be written as

$$p_1 + p_2 = -\frac{1}{R_s C'_s} - \frac{1}{R'_\ell C'_\ell} - \frac{C_r}{C'_\ell}\frac{1}{R_p C'_s}, \tag{5.36}$$

where

$$R_p = \frac{1}{\frac{1}{R'_\ell} + \frac{1}{R_s} + g_m}. \tag{5.37}$$

The last term of (5.36) shows that $|p_1 + p_2|$ increases with C_r. Since the product of the poles does not change with C_r, one pole must move to a higher frequency while the other one moves to a lower frequency. This effect is called *pole-splitting by means of capacitive feedback*. It occurs if the last term in (5.36) dominates, while C_r is still small with respect to both C'_ℓ and C'_s. Hence, if in (5.28)

$$C_r \left(R_s + R'_\ell + R_s R'_\ell g_m\right) \gg C'_s R_s + C'_\ell R'_\ell. \tag{5.38}$$

If $R_s \gg \frac{1}{g_m}$ the above condition can be simplified to

$$C_r \left(1 + g_m R'_\ell\right) > C'_s \left(1 + \frac{R'_\ell C'_\ell}{R_s C'_s}\right). \tag{5.39}$$

If $R_s \to \infty$ this condition changes to

$$C_r \left(1 + g_m R'_\ell\right) > C'_s. \tag{5.40}$$

This describes a situation in which, as a result of the Miller effect, the input capacitance at low frequencies is predominantly determined by C_r.

Pole-splitting can be used to manipulate the poles of a negative feedback amplifier into their desired positions, without changing the bandwidth of the amplifier. However, undesired pole-splitting in a negative feedback amplifier may move one of the poles out of the dominant group, which reduces the achievable bandwidth of the amplifier (see Chapter 12). At a later stage, we will introduce techniques for removing undesired pole-splitting. Without pole-splitting, a CS stage always contributes maximally to the loop gain poles product in a negative feedback amplifier.

In the following example, we will investigate the presence of pole-splitting in a CS stage.

Example 5.4

Let us consider a CS stage which is driven from a current source with a source impedance that can be represented by a resistor of $100k\Omega$ *in parallel with a capacitor of* $0.2pF$*. The source is loaded with a parallel RC network with a resistance of* $100k\Omega$ *in parallel with a capacitance of* $0.5pF$*. The small-signal parameters of the MOS in the operating point are:* $g_m = 0.5mA/V$*,* $g_o = 20\mu A/V$*,* $c_{gs} = 1.25fF$*,* $c_{gd} = 0.3fF$ *and* $c_{ds} = 0.2fF$*.*

We want to investigate whether pole-splitting occurs in this stage.

The values of the components of the simplified equivalent circuit according to Figure 5.23 are: $R_s = 100k\Omega$*,* $C'_s = 0.20125pF$*,* $R'_\ell = 33.33k\Omega$*,* $C'_\ell = 0.5002pF$*,* $C_r = 0.3fF$ *and* $g_m = 0.5mA/V$*. Since* $g_m R_s \gg 1$ *we may use expression 5.39 to determine whether pole-splitting occurs in this stage. We obtain*

$$C_r\left(1 + g_m R'_\ell\right) = 0.3 \cdot 10^{-15}\left(1 + 0.5 \cdot 10^{-3} \times 33.33 \cdot 10^{3}\right) = 5.3fF. \qquad (5.41)$$

Since $C'_s = 0.20125pF$*, we may conclude that no significant pole-splitting occurs in this stage.*

In the following example we will check the above results with SLiCAP.

Example 5.5

Figure 5.24 shows the circuit file of the CS stage driven from and terminated with an RC network.

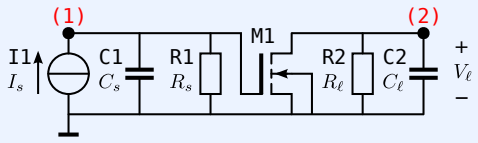


Figure 5.24: Circuit for demonstration of poles splitting with SLiCAP.

The netlist of this circuit is listed below:

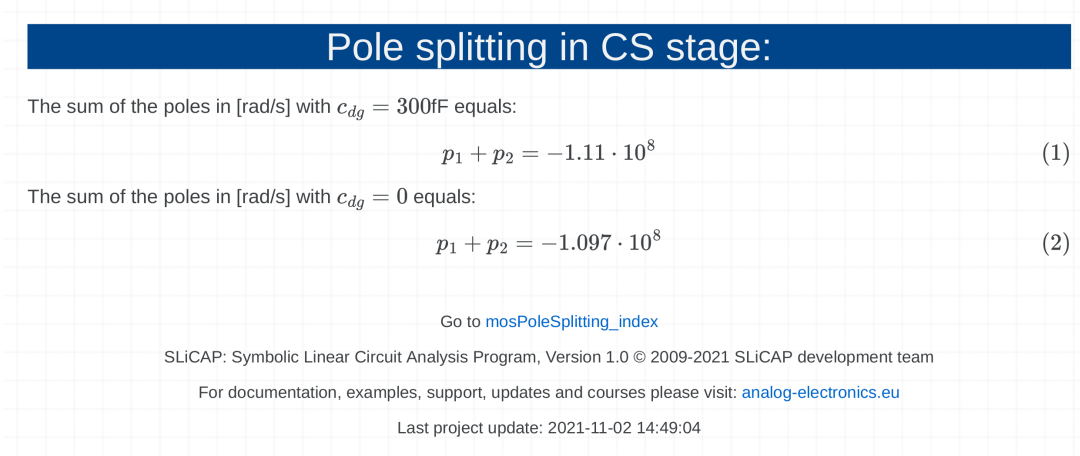
```
1  mosPoleSplitting
2  * SLiCAP netlist file
3  I1 0 1 {I_s}
4  C1 1 0 {C_s}
5  R1 1 0 {R_s}
6  R2 2 0 {R_ell}
7  C2 2 0 {C_ell}
8  .include C18.lib
9  M1 2 1 0 0 M gm=0.5m cgs=1.25f cdg={c_dg} cdb=0.2f go=20u
10 .param c_dg=0.3f I_s=1 C_s=0.2p R_s=100k R_ell=100k C_ell=0.5p
11 .end
```

Lines 1 through 23 of the script `mosPolesSplitting.py` *determine the sum of the poles for* $c_{dg} = 0.3fF$ *and* $c_{dg} = 0$*.*

```
1  #!/usr/bin/env python3
2  # -*- coding: utf-8 -*-
3  """
4  mosPoleSplitting.py
5  """
6  from SLiCAP import *
7  prj = initProject('mosPoleSplitting')
8  i1 = instruction()
9  i1.setCircuit('mosPoleSplitting.cir')
10 htmlPage('Circuit data')
11 netlist2html('mosPoleSplitting.cir')
12 elementData2html(i1.circuit)
13 params2html(i1.circuit)
14 i1.setSimType('numeric')
15 i1.setSource('I1')
```

```
i1.setDetector('V_2')
i1.setGainType('gain')
i1.setDataType('poles')
htmlPage('Pole splitting in CS stage:')
result = i1.execute()
poles = result.poles
sp.symbols('p_1 p_2')
sumOfPoles = poles[0] + poles[1]
text2html('The sum of the poles in [rad/s] with $c_{dg}=300$fF equals:')
eqn2html('p_1+p_2', sumOfPoles)
i1.defPar('c_dg', 0);
result =i1. execute()
poles = result.poles
sumOfPoles = poles[0] + poles[1]
text2html('The sum of the poles in [rad/s] with $c_{dg}=0$ equals:')
eqn2html('p_1+p_2', sumOfPoles)
htmlPage('Pole splitting with $c_{dg}$')
```

The results are displayed on the HTML page shown in Figure 5.25. They confirm our conclusion from the previous example: with $c_{dg} = 0.3fF$ we have no significant pole-splitting.



Pole splitting in CS stage:

The sum of the poles in [rad/s] with $c_{dg} = 300$fF equals:

$$p_1 + p_2 = -1.11 \cdot 10^8 \tag{1}$$

The sum of the poles in [rad/s] with $c_{dg} = 0$ equals:

$$p_1 + p_2 = -1.097 \cdot 10^8 \tag{2}$$

Go to mosPoleSplitting_index

SLiCAP: Symbolic Linear Circuit Analysis Program, Version 1.0 © 2009-2021 SLiCAP development team

For documentation, examples, support, updates and courses please visit: analog-electronics.eu

Last project update: 2021-11-02 14:49:04

Figure 5.25: SLiCAP simulation results

In the next example, we will determine the value of a capacitor that needs to be placed in parallel with c_{gd} to increase the sum of the poles by a factor 2.

Example 5.6

The sum of the poles of the CS stage from the previous example can be calculated using (5.34). We obtain

$$\begin{aligned} p_1 + p_2 &= -\frac{1}{100 \cdot 10^3 \times 0.20125 \cdot 10^{-12}} - \frac{1}{33.33 \cdot 10^3 \times 0.5002 \cdot 10^{-12}}, \\ &= -1.0967 \cdot 10^8 rad/s. \end{aligned} \tag{5.42}$$

If we need to increase this by a factor two we require (see expressions (5.36) and (5.37))

$$-\frac{C_r}{C'_\ell}\frac{1}{R_p C'_s} = -1.0967 \cdot 10^8, \tag{5.43}$$

from which we obtain

$$\begin{aligned} C_r &= 1.0967 \cdot 10^8 \times 0.20125 \cdot 10^{-12} \times 0.5002 \cdot 10^{-12} \times \\ &\times \frac{1}{\frac{1}{33.33 \cdot 10^3} + \frac{1}{100 \cdot 10^3} + 500 \cdot 10^{-6}}, \end{aligned} \tag{5.44}$$

$$= 20.44fF. \tag{5.45}$$

Since $c_{gd} = 0.3fF$, we need to place an external capacitance of $20.14fF$ in parallel with c_{gd}.

In the following example, we will check these results with SLiCAP.

Example 5.7

The circuit and its netlist have already been shown in example 3.5.5.

Lines 33 - 38 of the script file `mosPolesSplitting.py` *perform the calculation of the sum of the poles with* $c_{dg} = 20.44fF$. *Lines 39 through 46 create a root locus plot with* c_{dg} *stepping 0 to 50fF and display this plot on the HTML page.*

```
33 i1.defPar('c_dg', 20.44e-15);
34 result = i1.execute();
35 poles = result.poles
36 sumOfPoles = poles[0] + poles[1]
37 text2html('The sum of the poles in [rad/s] with $c_{dg}=20.44$fF equals:')
38 eqn2html('p_1+p_2', sumOfPoles)
39 i1.setStepVar('c_dg')
40 i1.setStepStart(0)
41 i1.setStepStop('50f')
42 i1.setStepNum(100)
43 i1.setStepMethod('lin')
44 i1.stepOn()
45 figPZ = plotPZ('poleSplitting', 'Poles vs $c_{d_g}$', i1.execute(), show=True)
46 fig2html(figPZ, 500)
```

This page is shown in Figure 5.26. It shows that the sum of the poles with $c_{dg} = 20.44fF$ *is approximately two times the sum of the poles with* $c_{dg} = 0$. *The root locus plot shows the pole-splitting for increasing values of* c_{dg}.

Pole splitting with c_{dg}

The sum of the poles in [rad/s] with $c_{dg} = 20.44$fF equals:

$$p_1 + p_2 = -1.92 \cdot 10^8 \tag{1}$$

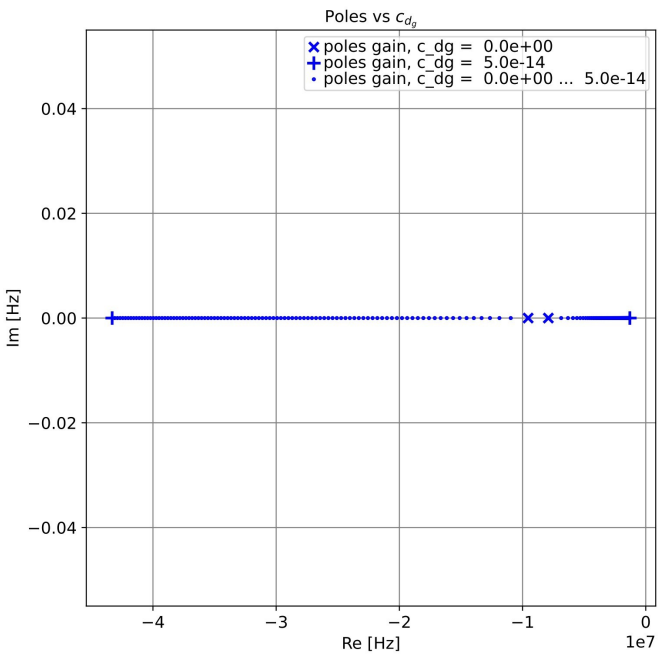


Figure 5.26: SLiCAP simulation results

5.3.5 Input impedance

Knowledge of the input impedance of an *RC* loaded CS stage, is useful for the determination of the dynamic behavior of cascaded stages. If the input

impedance of an RC loaded CS stage can be described by a parallel RC network, then the poles of cascaded CS stages could be estimated as described above.

In section 5.2.2 we have seen that the input impedance of a shorted CS stage can be described by a parallel RC network. The input impedance of a non-shorted CS stage has two poles and one zero (see expression (5.9)). For the circuit from Figure 5.27, this impedance can be written as

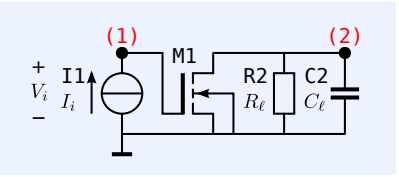


Figure 5.27: SLiCAP test circuit for determination of the input impedance of an RC loaded CS stage.

$$Z_i = \frac{1 + sR'_\ell\left(c_{gd} + C'_\ell\right)}{s\left(c_{gs} + \left(1 + g_m R'_\ell\right)c_{gd}\right)\left(1 + s\frac{R'_\ell\left(c_{gs}C'_\ell + c_{gs}c_{gd} + C'_\ell c_{gd}\right)}{c_{gs} + \left(1 + g_m R'_\ell\right)c_{gd}}\right)}. \tag{5.46}$$

Below the frequency of the zero and the second pole, the input impedance is capacitive:

$$Z_i = \frac{1}{sC_{i,\ell}} = \frac{1}{s\left(c_{gs} + \left(1 + g_m R'_\ell\right)c_{gd}\right)}. \tag{5.47}$$

The low-frequency input capacitance C_i is found as

$$C_i = c_{gs} + \left(1 + g_m R'_\ell\right)c_{gd}.$$

This is the value predicted by the Miller effect. At high frequencies, the character of the input impedance depends on the relative position of the pole and the zero. We may write

$$Z_i = \frac{1}{sC_i}\frac{1 + s\tau_z}{1 + s\tau_p}, \tag{5.48}$$

and obtain the ratio of the frequency of the pole and the zero from $\frac{\tau_z}{\tau_p}$ as

$$\frac{\tau_z}{\tau_p} = \frac{\left(R'_\ell c_{gd} + R'_\ell C'_\ell\right)\left(c_{gs} + c_{gd} + g_m R'_\ell c_{gd}\right)}{R'_\ell\left(c_{gs}C'_\ell + c_{gs}c_{gd} + C'_\ell c_{gd}\right)}. \tag{5.49}$$

This can be written as

$$\frac{\tau_z}{\tau_p} = 1 + \frac{R'_\ell c_{gd}\left(c_{gd} + g_m R'_\ell\left(c_{gd} + C'_\ell\right)\right)}{R'_\ell\left(c_{gs}C'_\ell + c_{gs}c_{gd} + C'_\ell c_{gd}\right)}. \tag{5.50}$$

From which we see that the zero always has a lower frequency than the second pole. Hence, for frequencies between the frequency of the zero and the frequency of the second pole the input impedance is resistive.

Above the frequency of the second pole, the input impedance is capacitive. The input capacitance C_h at those frequencies is found as

$$C_h = c_{gs} + \frac{c_{gd}C'_\ell}{c_{gd} + C'_\ell}. \tag{5.51}$$

Figure 5.28 shows the equivalent circuit for the input impedance of a CS stage. The left branch of this network is introduced by the Miller effect. The value of the resistor in this branch follows from the frequency of the zero and the capacitance $c_{gd}g_m R'_\ell$.

The circuit shows that the Miller effect can be ignored if the circuit is shorted $g_m R'_\ell \ll 1$.

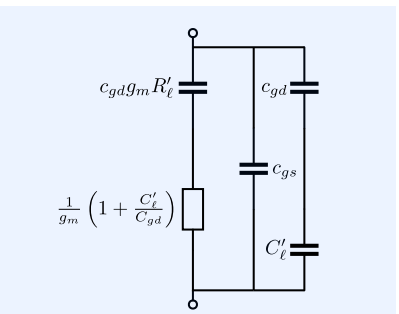


Figure 5.28: Equivalent circuit of the the input impedance of an RC loaded CS stage.

5.3.6 Influence of the gate series resistance

Until now, we have assumed that the *RC* type source impedance can simply be considered in parallel with the gate-source input of the intrinsic CS stage. In fact, we neglected any effect due to the gate series resistance. The influence of this resistance, however, cannot always be ignored. In CMOS processes with poly silicon gates, the series resistance may be relatively large and its influence on the dynamic behavior cannot always be ignored.

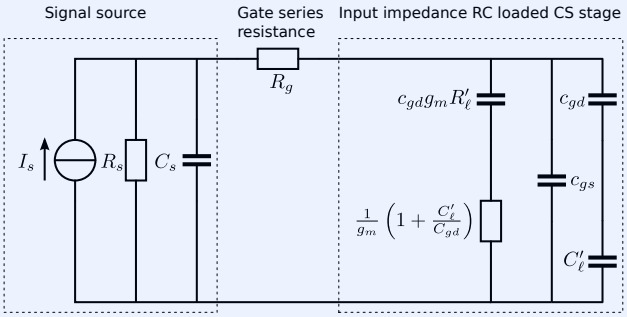


Figure 5.29: Equivalent network of the driving point impedance at the input of an *RC* loaded CS stage with a nonzero gate series resistance, and driven from an *RC* source.

The influence of the series gate resistance on the dynamic behavior of the *RC* driven and loaded CS stage can be studied by investigating the poles of the driving point impedance at the input of the CS stage. Figure 5.29 shows the circuit for evaluation of these poles.

Until now, we studied the transimpedance and the current gain of the CS stage with $R_g = 0$. In this case the R_s and C_s are simply in parallel with Z_i. First of all, the source capacitor does not introduce a new pole. Secondly, the dominant pole is no longer found at $s = 0$. It moves to $s = -\frac{1}{R_s C_t}$, where C_t is the total capacitance in parallel with R_s.

If $R_g > 0$ the source capacitance is decoupled from the input capacitance. In this case we have three independent capacitor voltages and thus three poles. The exact influence of R_g depends on the values of all other circuit elements. However, since R_g introduces an attenuation between the source and the input of the CS stage it will be clear that its influence will somehow contribute to bandwidth reduction. Hence, for high-frequency applications we need to keep R_g as small as possible!

For small values of R_g, a shorted CS stage and $c_{gs} \ll c_{gd}$, the value of p_3 can be estimated as:

$$p_3 \approx -\frac{C_s + c_{gs}}{R_g C_s c_{gs}}. \tag{5.52}$$

Hence, if $R_g > \frac{C_s + c_{gs}}{g_m C_s}$, the frequency of this pole is found below f_T of the transistor. This gives a rough design criterion for R_g in high-frequency applications.

5.4 Optimization of the noise performance of a CS stage

In this section, we will discuss the design of the noise performance of a CS stage. We will show that the noise contribution of a CS stage can often be reduced to an acceptable level by proper selection of the device geometry and the operating conditions. However, reduction or even minimization of the of the noise contribution is not without costs, and in some cases the associated cost factors, such as the device width and the operating current, may become unacceptably large.

Determination of the optimum device geometry and operating conditions is essential for evaluation of the feasibility of noise performance requirements at an early stage of the design. If the noise performance cannot be met under optimized conditions, it certainly cannot be met under sub-optimum conditions.

5.4.1 Noise design considerations

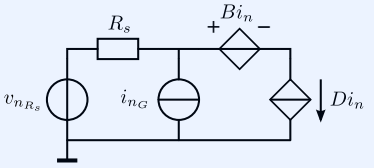
The equivalent input noise model of the intrinsic CS stage has already been shown in Figure 5.18. For a low noise contribution, both g_m and f_T need to be as large as possible. This follows from (5.19) and (5.20). To achieve this, the device needs to operate in the saturation region. Since for a given device geometry, g_m increases with the device current and f_T is proportional with g_m and inversely proportional with c_{iss}, we require the smallest possible device, operating at the highest possible level of inversion. However, above the critical inversion coefficient IC_{CRIT}, the transconductance g_m does not significantly increase with the inversion coefficient. Hence, increasing the inversion level above IC_{CRIT}, does not improve the performane-to-cost ratio.[5] So, a general design rule is to let the device operate around its critical inversion coefficient and use the device width W as design parameter to optimize the sum of the contributions of the equivalent input voltage noise source and the equivalent input current noise source, for the specified source impedance.

We will discuss the design of the noise contribution of a CS stage driven from a resistive source and a capacitive source. The design method, however, is not limited to specific source types.

[5] In this case the noise performance versus the current consumption.

5.4.2 Noise minimization for resistive source

Figure 5.30 shows the equivalent input noise model of a CS stage driven from a voltage source with source resistance R_s. This model is equal to the intrinsic model discussed in section 5.2.4, added with the resistive source and its associated noise.



$$S_{v_{n_{R_s}}} = 4kTR_s$$
$$S_{i_n} = 4kTn\Gamma g_m\left(1+\frac{f_\ell}{f}\right)$$
$$S_{i_{n_G}} = 2qI_G$$
$$B = -\frac{1}{g_m}$$
$$D = -\frac{sc_{iss}}{g_m} = -\frac{s}{2\pi f_T}$$

Figure 5.30: Equivalent input noise model of a CS stage driven from a resistive source.

The voltage spectral density of the source-referred input noise can be obtained after transformation of the current noise sources into voltage noise sources. The spectral density S_{Vn} of the total source-referred (voltage) noise is obtained as

$$S_{Vn} = 4kTR_s + 2qI_GR_s^2 + 4kTn\Gamma g_m\left(\frac{1}{g_m^2} + R_s^2\frac{f^2}{f_T^2}\right)\left(1+\frac{f_\ell}{f}\right)\ [\mathrm{V^2Hz^{-1}}]. \tag{5.53}$$

We will investigate the noise optimization process for frequencies above f_ℓ and assume $I_G = 0$. Under these conditions, equation (5.53) can be simplified to

$$S_{Vn} = 4kTR_s + \frac{4kTn\Gamma}{g_m} + 4kTR_s^2n\Gamma g_m\frac{f^2}{f_T^2}\ [\mathrm{V^2Hz^{-1}}]. \tag{5.54}$$

Qualitative description of the noise minimization

Let us now study expression (5.54) in more detail and see if we can find means to minimize the noise contribution of the CS stage. If so, we can quickly estimate the feasibility of noise requirements. If the minimum noise contribution is too large we have a *show stopper* and either the specifications have to be relieved, or a different technology needs to be considered.

Expression (5.54) has an optimum for g_m:

$$g_{m,opt} = \frac{f_T}{R_s f} \text{ [AV}^{-1}\text{]}. \tag{5.55}$$

This optimum depends on f which means that it depends on the frequency range over which the noise contribution should be minimized. This is elucidated in Figure 5.31. In this figure we assume that the device operates in strong inversion and that the effective gate source voltage is held constant while the channel width is varied. This implies that the transistor operates at a constant inversion coefficient and that g_m is proportional with W.

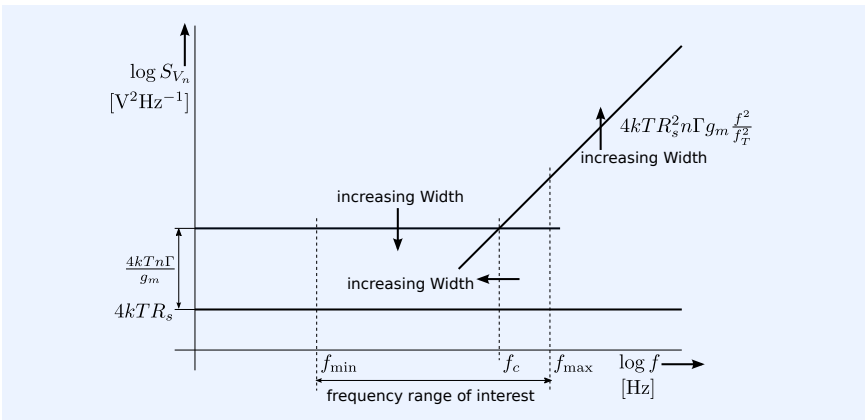


Figure 5.31: Noise optimization process for a CS stage driven from a resistive source. The contribution of the gate induced noise to the total RMS noise over the frequency range of interest is made equal to the contribution of the equivalent input voltage noise.

When g_m is below its optimum, the contribution of the frequency independent part, the second term in (5.54), to the total RMS noise over the frequency range of interest, dominates over the contribution of the third term in (5.54). The dominating part is thus the term Bi_n, usually referred to as the equivalent input voltage noise of the MOS. When g_m is above its optimum, the contribution of the third term dominates over the second term. This term originates from Di_n, usually referred as the *gate induced noise*. Hence, minimization of the noise can be achieved if g_m can be adjusted to its optimum value, without affecting n, Γ and f_T. This can be done by adjusting the channel width W of the MOS, while maintaining the effective gate-source voltage $V_{GS,eff}$. By doing so, the operating current, the transconductance factor g_m and the total input capacitance c_{iss} are proportional with the width, while the cut-off frequency f_T maintains its value. By increasing the width, the corner frequency f_c moves towards a lower value. This can be seen if we write (5.54) in the form

$$S_{Vn} = 4kT\left(R_s + \frac{n\Gamma}{g_m}\right)\left(1 + \frac{f^2}{f_c^2}\right) \text{ [V}^2\text{Hz}^{-1}\text{]}, \tag{5.56}$$

where the corner frequency f_c equals

$$f_c = \frac{1}{2\pi c_{iss} R_s}\sqrt{1 + \frac{g_m R_s}{n\Gamma}} \text{ [Hz]}. \tag{5.57}$$

Since both c_{iss} and g_m are proportional with the width, f_c is about inversely proportional with the square root of the width.

We will now derive the value of the best possible noise factor that can be

achieved with a CS stage when driven from a resistive source. This value and the conditions under which it can be achieved, help us to give a quick estimation of the feasibility of the noise performance of an amplifier driven from a resistive source that should be realized in a MOS process.

Quantitative description of the noise minimization

From the above formulated noise design considerations we may already conclude that for a low noise addition the transistor should operate in the saturation region and in strong inversion. Operation in strong inversion with velocity saturation is not very effective because in this region g_m does not significantly increase with the inversion coefficient. Hence, a good starting point is to let the device operate at $IC = IC_{CRIT}$.

The method for finding the best possible noise figure and the conditions to achieve it is as follows:

1. Express g_m in the process parameters, the operating conditions and in the geometry parameters W and L
2. Assume a brick wall (frequency) weighting function for the noise, and determine the optimum device with and the optimum noise figure as a function of the frequency range of interest

In the saturation region and at strong inversion without velocity saturation, the transconductance g_m is proportional with the effective gate-source voltage V_{eff}:

$$g_m = \frac{W\beta_{sq}V_{eff}}{Ln} \text{ [S]}. \tag{5.58}$$

After substitution of (5.58) in (5.54) we obtain

$$S_{Vn} = 4kT\left(R_s + \frac{n^2\Gamma L}{W\beta_{sq}V_{eff}}\right)\left(1 + \Gamma\frac{W}{L}\beta_{sq}V_{eff}R_s\frac{f^2}{f_T^2}\right) \text{ [V}^2\text{Hz}^{-1}\text{]}. \tag{5.59}$$

The mean square value $\overline{v_n^2}$ of the source-referred noise, over a frequency range from $f_{\min}$ to $f_{\max}$, is obtained after integration over this frequency range:

$$\overline{v_n^2} = \int_{f_{\min}}^{f_{\max}} S_{Vn}df \text{ [V}^2\text{]}. \tag{5.60}$$

For a given value of V_{eff}, the total source-referred noise has a minimum value at $W = W_{opt}$. This can be seen from (5.59) which has a term that is proportional with W and a term that is inversely proportional with W. Notice that by changing W, while keeping the effective gate-source voltage constant, the inversion coefficient does not change. Hence, the drain-source current increases linearly with the device width.

The optimum width can be found as the solution of W of

$$\frac{d}{dW}\overline{v_n^2} = 0. \tag{5.61}$$

If we assume a uniform weighting function over a frequency range from $f_{\min}$ to $f_{\max}$, $\overline{v_n^2}$ can be calculated as

$$\overline{v_n^2} = 4kTR_s\left(1 + \frac{1}{W}\frac{n^2\Gamma L}{\beta_{sq}R_sV_{eff}} + \frac{f_{\max}^3 - f_{\min}^3}{f_{\max} - f_{\min}}\left(\frac{n^2\Gamma^2}{3f_T^2} + \frac{\Gamma R_s\beta_{sq}V_{eff}}{3Lf_T^2}W\right)\right) \text{ [V}^2\text{]}. \tag{5.62}$$

A minimum value $\left(\overline{v_n^2}\right)_{\min}$ of the source-referred noise is obtained if

$W = W_{opt}$. The optimum width is found when the coefficient of W and the coefficient of W^{-1}in 5.65 are equal. This yields

$$W_{opt}^2 = \frac{3n^2L^2f_T^2}{\beta_{sq}^2R_s^2V_{eff}^2}\frac{f_{\max}-f_{\min}}{f_{\max}^3-f_{\min}^3}\ [\mathrm{m}^2]. \tag{5.63}$$

After substitution of

$$f_T = \frac{3g_m}{4\pi WLC'_{OX}} = \frac{3\beta_{sq}V_{eff}}{L^2n4\pi C'_{OX}}\ [\mathrm{Hz}], \tag{5.64}$$

we obtain

$$W_{opt} = \frac{3\sqrt{3}}{4\pi LC'_{OX}R_s}\sqrt{\frac{f_{\max}-f_{\min}}{f_{\max}^3-f_{\min}^3}}\ [\mathrm{m}]. \tag{5.65}$$

After substitution of (5.65) and (5.64) in (5.62), we obtain the minimum of the mean square value of the total source-referred noise voltage $\overline{v_n^2}$ as:

$$\overline{v_n^2} = 4kTR_s\left(1+\frac{n\Gamma}{\sqrt{3}f_T}\sqrt{\frac{f_{\max}^3-f_{\min}^3}{f_{\max}-f_{\min}}}\right)^2\ [\mathrm{V}^2]. \tag{5.66}$$

The best possible noise factor F_{opt} is obtained by dividing the total source-referred noise by the contribution of the source resistance:

$$F_{opt} = \left(1+\frac{n\Gamma}{\sqrt{3}f_T}\sqrt{\frac{f_{\max}^3-f_{\min}^3}{f_{\max}-f_{\min}}}\right)^2\ [\text{-}]. \tag{5.67}$$

If $f_{\max} \gg f_{\min}$, we can simplify the expressions for W_{opt} to

$$W_{opt} = \frac{\sqrt{3}}{2\pi f_{\max}R_sLC'_{OX}}[\mathrm{m}]. \tag{5.68}$$

The mean square value of the source-referred noise voltage under this condition becomes

$$\left(\overline{v_n^2}\right)_{\min} = 4kTR_sf_{\max}\left(1+\frac{\Gamma nf_{\max}}{f_T\sqrt{3}}\right)^2\ [\mathrm{V}^2], \tag{5.69}$$

and the noise figure is obtained as

$$F_{\min} = \left(1+\frac{\Gamma n}{\sqrt{3}}\frac{f_{\max}}{f_T}\right)^2\ [\text{-}]. \tag{5.70}$$

This noise figure is achieved with the device operating in the saturation region at strong inversion without velocity saturation and with the width of the device designed such that:

$$c_{iss} = \frac{\sqrt{3}}{2\pi f_{\max}R_s}\ [\mathrm{F}], \tag{5.71}$$

assuming

$$c_{iss} \approx \frac{2}{3}WLC'_{OX}\ [\mathrm{F}]. \tag{5.72}$$

In the following example, we will estimate the feasibility of the noise performance of a resistively driven CS stage.

Example 5.8

The voltage of a voltage source with an internal resistance of 600Ω *needs to be amplified. The information is contained in a frequency range from* $1\cdots5GHz$.

The noise figure of the amplifier should be less than 1*dB. The amplifier should be realized in a* 180*nm CMOS process. The NMOS process characteristics are:*

1. *Technology current:* $I_0 = 640nA$
2. *Critical inversion coefficient:* $IC_{CRIT} = 31.7$
3. *Effective gate-source voltage at critical inversion:* $V_{eff_{CRIT}} = 393mV$
4. *Normalized oxide capacitance:* $C'_{OX} = 0.0084Fm^{-2}$
5. *Zero field carrier mobility:* $\mu_0 = 0.042$
6. *Substrate factor:* $n = 1.35$
7. *Noise constant:* $\Gamma = 0.7$

We will check if the required noise figure can be achieved with an NMOS CS stage.

The minimum noise figure that can be achieved can be approximated by (5.67). Hence, we need to calculate f_T. *If we substitute (4.145) in the expression for* f_T *(5.64) we obtain an expression for the cut-off frequency as function of the effective gate-source voltage:*

$$f_T = \frac{\mu_0 V_{eff}}{2\pi n L^2} \; [Hz]. \tag{5.73}$$

At $IC = IC_{CRIT}$, *and with* $L = 180nm$ *this yields* $f_T = 60GHz$.

With the aid of (5.67) and $f_{\max} = 5GHz$ *we find the best possible noise figure at* $IC = IC_{CRIT}$: $F = 1.104$ *which corresponds with* 0.43*dB. Hence, a noise figure of* 1*dB can be realized with an NMOS CS stage.*

However, the estimated value of the cut-off frequency is rather optimistic. This is because at $IC = IC_{CRIT}$ *the effects of velocity saturation and vertical field mobility reduction are already noticeable, and* f_T *will not be proportional with* V_{eff}. *Although not very efficient, increasing the inversion coefficient to a level above* IC_{CRIT} *may still yield a larger* g_m *and a larger* f_T *and thus reduce the noise. The increasing of* g_m *and* f_T *with the inversion coefficient ends if the velocity saturation and vertical mobility reduction become too strong, or if* V_{eff} *exceeds* V_{ds} *and the transistor does no longer operate in the saturation region.*

Realization of this noise figure is not free of costs. In the following example we will determine the device width and the operating current of the NMOS to achieve the optimum noise figure of 0.43dB over the frequency range of interest driven from a source resistance of 600Ω.

Example 5.9

According to (5.65) we require an optimum device width of: $W_{opt} = 54.6\mu m$. *With the device operating at* $IC = IC_{CRIT}$ *we can determine the operating current* I_{DS} *and obtain*

$$I_{DS} = I_0 \frac{W}{L} IC_{CRIT} = 6.15mA. \tag{5.74}$$

In the next example, we will verify the above result with SLiCAP. We will use the built-in EKV noise model and the model for g_m according to (4.167). This will result in a lower cut-off frequency then estimated in the previous example and yields a larger noise figure.

Example 5.10

In this example we will verify the results of the previous examples with SLiCAP. *Below is the listing of the* SLiCAP *script file for determination of the source-referred noise spectrum, the total source-referred RMS noise and the noise figure versus the width at constant inversion coefficient.*

In line 20, the parameter KF *is set to zero, this sets the* $1/f$ *noise to zero. The instructions in lines 22 through 27 set the drain current to a value at which the device operates at critical inversion. The instructions in lines 29 through 44 print the device width and the drain current at critical inversion on the html page with the circuit data.*

```
1  #!/usr/bin/env python3
2  # -*- coding: utf-8 -*-
3  """
4  CSresnoise.py
5  """
6  from SLiCAP import *
7  prj = initProject('CS stage noise with resistive source')
8  fileName = 'CSresNoise'
9  i1 = instruction()
10 i1.setCircuit(fileName + '.cir')
11 htmlPage('Circuit data')
12 img2html(fileName + '.svg', 700)
13 netlist2html(fileName+'.cir')
14 # Set value of 1/f noise to zero, and I_D to critical inversion
15 i1.defPar('KF_N18', 0)
16 i1.setSimType('numeric')
17 I_D     = i1.getParValue('ID')
18 IC      = i1.getParValue('IC_X1')
19 IC_CRIT = i1.getParValue('IC_CRIT_X1')
20 I_D     = I_D*IC_CRIT/IC
21 i1.defPar('ID', I_D)
22 # Print some important noise parameters to an HTML page
23 htmlPage('Operating point parameters')
24 R_N     = i1.getParValue('R_N_X1')
25 R_s     = i1.getParValue('R_s')
26 f_T     = i1.getParValue('f_T_X1')
27 g_m     = i1.getParValue('g_m_X1')
28 Width   = i1.getParValue('W')
29 text2html('Device width:')
30 eqn2html('W', Width)
31 text2html('Dain current at critical inversion:')
32 eqn2html('I_D', I_D)
33 text2html('Effective noise resistance $R_N$:')
34 eqn2html('R_N', R_N)
35 text2html('Cut-off frequency $f_T$:')
36 eqn2html('f_T', f_T)
37 i1.setSource('V1')
38 i1.setDetector('V_out')
39 i1.setGainType('vi')
40 i1.setDataType('noise')
41 i1.setSimType('numeric')
42 noise_result   = i1.execute()
```

Figure 5.32 shows the page with the circuit data, and Figure 5.33 shows the page with the operating point information.

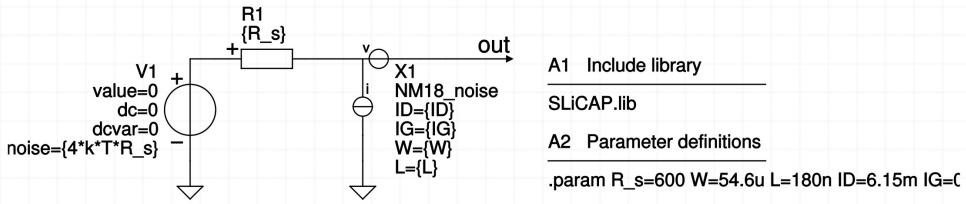
Lines 43-45 of the script plot the spectral density of the source referred noise over a frequency range from 100MHz *to* 100GHz*. Line 47 calculates the total RMS input noise over this frequency range, lines 49-51 the noise figure, and line 53 estimates the corner frequency* f_c*. The corner frequency can be estimated using expression (5.75) with the data from Figure 5.32.*

$$f_h = f_T\sqrt{\frac{1}{g_m n \Gamma R_s}} = f_T\sqrt{\frac{R_N}{R_s}} \tag{5.75}$$

These obtained results are all shown on an HTML page (lines 54 until 62).

```
43 figInoise      = plotSweep('Inoise', 'Source-referred noise spectrum',
44                            noise_result, 1e8, 1e11, 100, funcType = 'inoise',
45                            show=True)
46 # Calculate the noise figure at critical inversion and the given width
47 tot_inoise     = rmsNoise(noise_result, 'inoise', 1e9, 5e9)
48 # Calculate the noise figure
49 tot_inoise_src = rmsNoise(noise_result, 'inoise', 1e9, 5e9,
50                            source = noise_result.source)
51 NF             = 20*sp.log(tot_inoise/tot_inoise_src)/sp.log(10)
52 # Estimation of the corner frequency f_c:
53 f_c            = f_T*sp.sqrt(R_N/R_s)
54 htmlPage("Noise analysis-1")
```

Circuit data



Netlist: CSresNoise.cir

```
"CS stage resistive noise"
* gnetlist -q -g spice-noqsi -o CSresNoise.net CSresNoise.sch
* SPICE file generated by spice-noqsi version 20130710
* Send requests or bug reports to jpd@noqsi.com
.INCLUDE SLiCAP.lib
X1 1 0 out NM18_noise ID={ID} W={W} L={L} IG={IG}
R1 2 1 {R_s}
V1 2 0 V value=0 dc=0 dcvar=0 noise={4*k*T*R_s}
.param R_s=600 W=54.6u L=180n ID=6.15m IG=0
.end
```

Figure 5.32: SLiCAP html page with the circuit data.

Operating point parameters

Device width:

$$W = 5.46 \cdot 10^{-5} \tag{1}$$

Dain current at critical inversion:

$$I_D = 0.00491 \tag{2}$$

Effective noise resistance R_N:

$$R_N = 63.93 \tag{3}$$

Cut-off frequency f_T:

$$f_T = 2.932 \cdot 10^{10} \tag{4}$$

Figure 5.33: SLiCAP html page with the operating point information.

```
55 text2html("The figure below shows the spectrum of the source-referred " +
56          "voltage noise.")
57 fig2html(figInoise, 500)
58 text2html("The source-referred RMS noise voltage over this frequency range " +
59          "equals: %s [$\mu$V]."%(sp.N(1e6*tot_inoise, ini.disp)))
60 text2html("The noise figure equals: %s [dB]."%(sp.N(NF, ini.disp)))
61 text2html("The estimated conrner frequency $f_c$" +
62          ": %s [GHz]."%(sp.N(f_c*1e-9, ini.disp)))
```

The output page with the plot is shown in Figure 5.34. It clearly shows the increase of the source-referred noise at high frequencies as a result of the gate-induced noise.

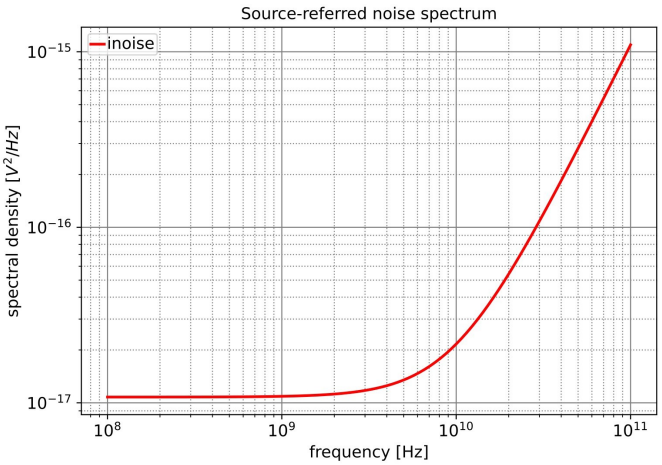
The third part of the script (lines 63-118) calculates the width for the best noise figure (at critical inversion) and shows how do do this for different frequency ranges. Inline comments in this script file elucidate the way of working.

```
63 # Calculate the width W at which we will have the best noise performance.
64 W              = sp.Symbol('W') # 'W' in the Python environment
65 i1.circuit.delPar('W')        # delete the numeric definition of the width
66 # We will keep the inversion coefficient at critical inversion, hence we scale
67 # the current with the width.
68 i1.defPar('ID', I_D*W/Width)
69 noise_w = i1.execute() # calculate the noise spectra as a function of W and f
70 # We now calculate the noise as a function of W over a frequency range
71 # 'fmin' to 'fmax':
72 f_min = sp.Symbol('f_min')
73 f_max = sp.Symbol('f_max')
74 rms_noise_w        = rmsNoise(noise_w, 'inoise', f_min, f_max)
75 rms_noise_w_source = rmsNoise(noise_w, 'inoise', f_min, f_max, noise_w.source)
76 # We now calculate the noise figure as a function of 'W', 'f_min' and 'f_max':
77 # Use the variance instead of the RMS value (simpler equation for later use)
78 NF_W               = (rms_noise_w/rms_noise_w_source)**2
79 # We now calculate the optimum width as a function of 'fmin' and 'fmax':
80 W_opt              = sp.solve(sp.diff(NF_W, W), W)
81 # The sympy solve function returns a list with solutions, we will print the
82 # positive one.
83 for w in W_opt:
84     w = sp.N(w.subs([(f_min, 1e9), (f_max, 5e9)]), ini.disp)
85     if w > 0:
86         W = w
87         print(W)
88 # Create a plot of the noise figure versus the with for different values of
89 # f_max and f_min = 1G
90 # Define the plot parameters, 'fw', 'W' and 'fmax'
91 i1.defPar('W', W)
92 i1.defPar('f_max', '10G')
93 # Define the noise figure as a function of f_max:
94 i1.defPar('NF', 10*sp.log(NF_W.subs([(f_min, 2e8)]))/sp.log(10))
95 # Define the step parameters
96 i1.setStepVar('f_max')
97 i1.setStepStart('2G')
98 i1.setStepStop('10G')
99 i1.setStepMethod('lin')
100 i1.setSimType('numeric')
101 i1.setStepNum(5)
102 i1.stepOn()
103 i1.setDataType('params')
104 result = i1.execute()
105 # Plot the function
106 fig_NF_W = plotSweep('NF_W', 'Noise Figure versus width, $f_{min}$ = 200MHz',
107                      result, 10, 200, 50, sweepVar = 'W', sweepScale = 'u',
108                      funcType = 'param', xUnits = 'm', yVar = 'NF',
109                      yUnits = 'dB', show = True)
110 # Put it all on an HTML page
111 htmlPage("Noise analysis-2")
112 text2html("The lowest noise figure over a frequency range from 1GHz to 5GHz "
      +
113          "and at critical inversion is achieved at a width " +
114          "of: %s [um]"%(sp.N(W*1e6, ini.disp)))
115 text2html("The figure below shows the noise figure as a function of the " +
116          "width and at critical inversion for diferent values of the " +
117          "maximum frequency $f_{max}$, and $f_{min}$=200MHz.")
118 fig2html(fig_NF_W, 500)
```

The results are shown in Figure 5.35.

Noise analysis-1

The figure below shows the spectrum of the source-referred voltage noise.



The source-referred RMS noise voltage over this frequency range equals: 218.2 [μV].

The noise figure equals: 0.7836 [dB].

The estimated conrner frequency f_c: 9.571 [GHz].

Figure 5.34: SLiCAP html page with the plot of the spectrum of the source-referred noise.

Noise analysis-2

The lowest noise figure over a frequency range from 1GHz to 5GHz and at critical inversion is achieved at a width of: 47.38 [um]

The figure below shows the noise figure as a function of the width and at critical inversion for diferent values of the maximum frequency f_{max}, and f_{min}=200MHz.

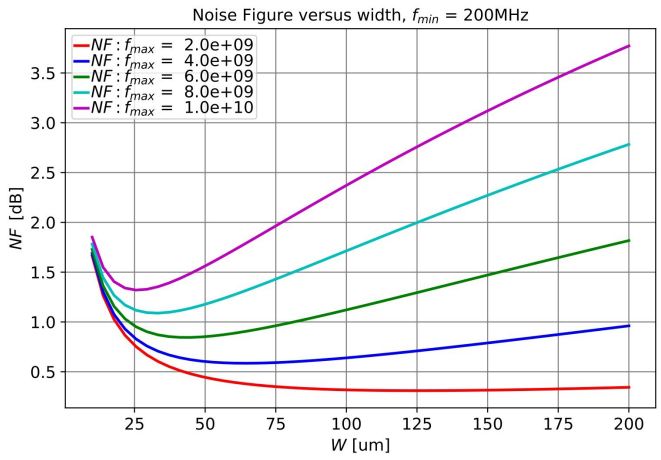


Figure 5.35: SLiCAP html page with the calculated and estimated values of the noise figure, the plot of the total noise and the noise figure versus the device widths and the calculated optimum width.

Model inaccuracies

The following model inaccuracies have to be taken into account:

1. In strong inversion the theoretical value of Γ is $\frac{2}{3}$. In most cases it will be larger. This increases the value of the lowest possible noise figure.

2. In the above calculations and simulations, the effect of the right half plane zero of the current gain on the gate induced noise has been ignored. Above the frequency of the zero the gate induced noise will no longer increase with frequency but it will introduce a noise contribution correlated with the voltage Bi_n from Figure 5.30.

3. We also ignored the contribution of the $\frac{1}{f}$ noise and the noise current associated with the gate leakage current I_G.

Conclusion

The following conclusion can be drawn:

A CS stage driven from a resistive source has the lowest noise figure if it operates in the saturation region and in strong inversion with g_m as large as possible and with its device width designed such that:

$$c_{iss} \approx \frac{\sqrt{3}}{2\pi R_s}\sqrt{\frac{f_{\max}-f_{\min}}{f_{\max}^3-f_{\min}^3}}\ [\mathsf{F}]. \tag{5.76}$$

Spice simulation

In the following example, we will evaluate the optimum noise figure and device width with SPICE and use the BSIM level 49 model of an NMOS fabricated in a 180nm CMOS process.

Example 5.11

The LTSPICE *netlist file with the instructions for plotting the source-referred noise spectrum and the noise factor as a function of the width is shown below. The netlist part can be used for other simulators, but the instruction section is for* LTSPICE *only.*

```
CS-resNoise
* file: CS-resNoise.cir
* LTspice circuit file
.lib CMOS18TT.lib
C1 1 0 1
E1 2 0 1 0 1k
V1 3 2 0
R1 3 4 600
M1 5 4 0 0 C18nmos L=180n W={W}
V2 5 6 0.9
R2 6 1 1meg
I1 0 6 {W*6.15m/54.6u}
.param W=54.6u
* instruction for plotting source referred noise spectrum
* over a frequency range from 100MHz to 100GHZ
;.noise V(6) V1 DEC 25 0.1G 100G

* instruction for plotting the noise figure as a function of the width
.noise V(6) V1 lin 100 1G 5G
.step param W 10u 100u 1u
.meas NOISE totOnoise integ V(onoise) FROM 1G TO 5G
.meas NOISE totR1noise integ V(R1) FROM 1G to 5G
.meas NOISE noiseFig PARAM 20*log10(totOnoise/totR1noise)
.end
```

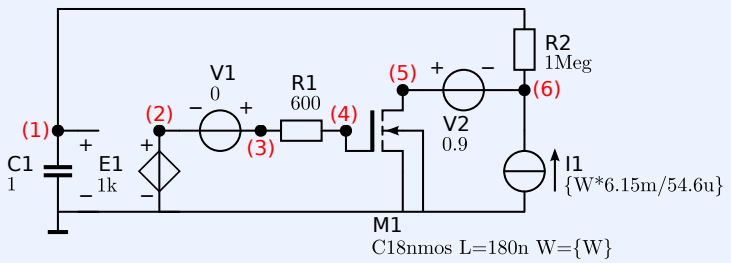


Figure 5.36: SPICE simulation circuit for determination of the optimum width at $IC = IC_{CRIT}$ for minimization of the source referred noise.

Figure 5.36 shows the corresponding circuit. The transistor operates in the saturation region, at critical inversion: $IC = IC_{CRIT}$.

The simulated source-referred spectrum and the noise figure versus the width are shown in Figure 5.37.

These results are obtained with SIMETRIX. *The noise spectrum shows the influence of the right half-plane zero. This zero was not included in the* SLICAP *simulation model nor in the hand calculations above. The lowest noise figure of about* 0.71*dB is found at a width of about* 42*µm. This optimum width is less than predicted by the hand calculations and by* SLICAP. *This is the result of differences between the capacitance models used in both simulators. The noise figure is slightly higher than predicted by* SLICAP. LTSPICE *gives a minimum noise figure of* 0.64*dB at a width of* 45*µm.*

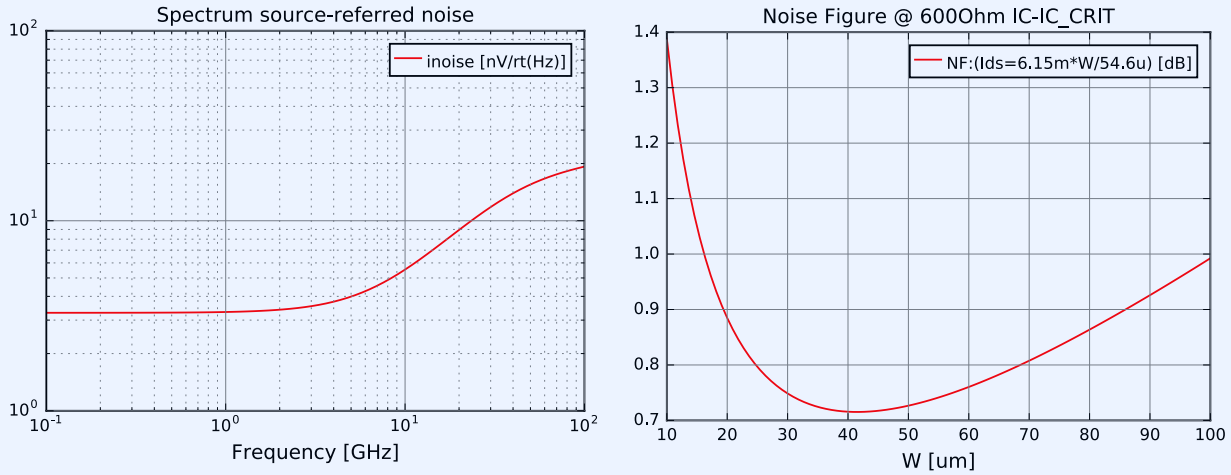


Figure 5.37: SIMETRIX noise simulation results.

5.4.3 Noise minimization for capacitive voltage source

Figure 5.38 shows the equivalent input noise model of a MOSFET which is driven from a capacitive voltage source. The signal source itself is noise free because there is no thermal noise associated with a capacitor.

If we ignore the $\frac{1}{f}$ noise and the shot noise associated with the gate leakage current, the spectrum S_{vn} of the source-referred noise can be obtained after transformation of DS_{id} into a source-referred voltage. This yields

$$S_{vn} = \left| B + D\frac{1}{j\omega C_s} \right|^2 S_{id} \ [\mathrm{V^2Hz^{-1}}]. \tag{5.77}$$

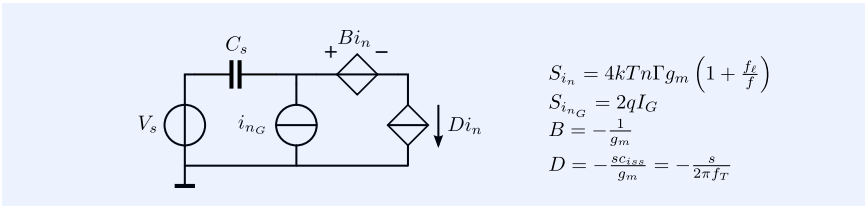


Figure 5.38: Equivalent input noise model of a CS stage driven from a capacitive voltage source.

After substitution of $B = -\frac{1}{g_m}$, $D = -\frac{j\omega}{c_{iss}}$ and $S_{id} = 4kTn\Gamma g_m$, the source-referred noise spectrum can be written as

$$S_{vn} = \frac{4kTn\Gamma}{g_m}\left(1 + \frac{c_{iss}}{C_s}\right)^2 \text{ [V²Hz⁻¹]}. \tag{5.78}$$

Similar as in the case with the resistive source, both g_m and c_{iss} can be written as a function of the device width. After substitution of (5.58) and (5.72) in (5.78) we obtain

$$S_{vn} = \frac{4kTLn^2\Gamma}{W\beta_{sq}V_{eff}}\left(1 + \frac{WLC'_{OX}}{C_s}\right)^2 \text{ [V²Hz⁻¹]}. \tag{5.79}$$

This expression can be expanded to

$$S_{vn} = \frac{4kTLn^2\Gamma}{\beta_{sq}V_{eff}}\left(\frac{1}{W} + 2\frac{LC'_{OX}}{C_s} + \frac{WL^2C'^2_{OX}}{C_s^2}\right) \text{ [V²Hz⁻¹]}. \tag{5.80}$$

This spectrum does not depend on frequency and it has a minimum if

$$\frac{W^2L^2C'^2_{OX}}{C_s^2} = 1. \tag{5.81}$$

If we use the approximation for c_{iss} given in (5.72), this condition is equivalent to:

$$c_{iss} = C_s. \tag{5.82}$$

The value of the noise spectrum at this optimum device width is then found as

$$S_{vn} = \frac{16kTn\Gamma}{g_m} \text{ [V²Hz⁻¹]}. \tag{5.83}$$

In the following example we will evaluate the feasibility of a low-noise CS stage that is driven from a capacitive voltage source.

Example 5.12

The voltage of a capacitive voltage source with an internal capacitance C_s of 100fF needs to be amplified. The information is contained in a frequency range from 100MHz$\cdots$1GHz. The spectral density of the source-referred noise voltage should be less than $5nV/\sqrt{Hz}$. The amplifier should be realized in a 180nm CMOS process. The NMOS process characteristics are:

1. *Technology current:* $I_0 = 640nA$
2. *Critical inversion coefficient:* $IC_{CRIT} = 31.7$
3. *Effective gate-source voltage at critical inversion:* $V_{eff_{CRIT}} = 393mV$
4. *Normalized oxide capacitance:* $C'_{OX} = 0.0084Fm^{-2}$
5. *Zero field carrier mobility:* $\mu_0 = 0.042$

6. *Substrate factor:* $n = 1.35$
7. *Noise constant:* $\Gamma = 0.7$

We will check if the required noise performance can be met with an NMOS CS stage.

The best possible noise performance of a capacitively driven CS stage is achieved if $c_{iss} = C_s$. *In this process, and at minimum device length* ($L = 180nm$) *this requires a device width W of about*

$$W = \frac{C_s}{LC'_{OX}} = \frac{0.1 \times 10^{-12}}{180 \times 10^{-9} \times 0.0084} = 66\mu m. \tag{5.84}$$

If we operate the CS stage in strong inversion and in the saturation region, we may calculate g_m *as*

$$g_m = \frac{W}{L}\frac{\mu_0 C'_{OX} V_{eff}}{n} = \frac{66}{0.18}\frac{0.042 \times 0.0084 \times 0.393}{1.35} = 38 \times 10^{-3}\ S. \tag{5.85}$$

This value will probably be too large because at $IC = IC_{CRIT}$ *the transconductance* g_m *will no longer be proportional with* V_{eff}.

With the estimated value of g_m, *the spectral density of the source-referred voltage noise at optimum width and* $IC = IC_{CRIT}$ *equals*

$$S_{vn} = \frac{16 \times 1.38 \times 10^{-23} \times 300 \times 1.35 \times 0.7}{38 \times 10^{-3}} = 1.65 \times 10^{-18}\ V^2Hz^{-1}. \tag{5.86}$$

Which is equivalent with $1.3nV/\sqrt{Hz}$. *Hence, a source-referred noise spectrum of* $5nV/\sqrt{Hz}$ *seems to be feasible.*

The operating current I_{DS} *required for optimum noise performance can be found from the technology current, the inversion coefficient and the device geometry*

$$I_{DS,opt} = 640 \times 10^{-9} \times \frac{66}{0.18} \times 31.7 = 7.44mA. \tag{5.87}$$

In the next example, we will verify the above results with SLiCAP. We expect to find a lower value of g_m as well as a different value of W because of a more accurate estimation of c_{iss}.

Example 5.13

Below is the script for the verification of the results from the previous example.

Lines 1-21 check the circuit and list the most relevant operating point parameters on the output page.

Lines 22-38 calculate the source-referred noise spectrum and the RMS value of the total source-referred voltage noise, and plot of the spectral density of the source-referred noise. Figures 5.39 shows the HTML output.

```
1  #!/usr/bin/env python3
2  # -*- coding: utf-8 -*-
3  """
4  CScapNoiseV.py
5  """
6  from SLiCAP import *
7  prj = initProject('CScapNoiseV')
8  i1 = instruction()
9  i1.setCircuit('CScapNoiseV.cir')
10 # create an html page for the results
11 htmlPage('Noise analysis')
12 img2html('CScapNoiseV.svg', 300)
13 # Discard 1/f noise
14 i1.defPar('KF_N18', 0)
15 # print important operating point parameters
16 text2html('The inversion coefficient $IC$ equals: ' +
17           '%s'%(sp.N(i1.getParValue('IC_X1'), ini.disp)))
18 text2html('The critical inversion coefficient $IC_{CRIT}$ equals: ' +
```

```
          '%s'%(sp.N(i1.getParValue('IC_CRIT_X1'), ini.disp)))
text2html('The transconductance $g_m$ equals: ' +
          '%s'%(sp.N(i1.getParValue('g_m_X1'), ini.disp)))
# calculate source referred noise spectrum
i1.setSource('V1')
i1.setDetector('V_out')
i1.setGainType('vi')
i1.setDataType('Noise')
i1.setSimType('numeric')
noiseResult = i1.execute()
head2html('Source referred noise')
iNoise = sp.sqrt(noiseResult.inoise)
text2html('The spectrum of the source-referred voltage noise [V/rt(Hz)] ' +
          'is: %s'%(sp.N(iNoise, ini.disp)))
text2html('The plot below shows the source-referred noise spectrum ' +
          'from  100MHz to 100GHz; as expected, it does not depend on ' +
          'the frequency.');
figSin = plotSweep('CScapNoiseVspectrum','Input noise spectrum', noiseResult,
                   1e8, 1e11, 100, funcType='inoise', show = True)
fig2html(figSin,  500)
```

The next part of the script calculates the optimum width at $IC = IC_{CRIT}$. *To this end, we need to delete the numeric definition of W given in the* `.param` *statement of the circuit definition. This is done in line 43. In order to operate at* $IC = IC_{CRIT}$ *for all values of W we let* I_{DS} *track with W by defining* $I_{DS} = \frac{W}{66\times10^{-6}}7.44 \times 10^{-3}$ *in line 45.*

If we now calculate the input noise spectrum it will consists sum of three terms:

1. *A constant*
2. *A term proportional with W*
3. *A term inversely proportional with W*

see (5.80).

This equation has a unique solution for W. The optimum values of W, I_{DS} *and* g_m *are calculated and displayed on the output page by the instructions in lines 39-65. Lines 66-79 determine theRMS value of the source-referred input noise as a function of the device width. Figure 5.40 shows the results.*

As expected, the results deviate from our estimations:

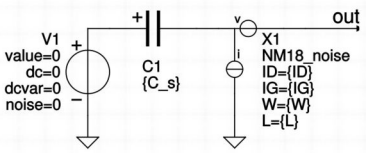
1. *The input capacitance* c_{iss} *equals the source capacitance* C_s *at a smaller width. This is because we ignored the overlap capacitance in both* c_{gs} *and* c_{dg}.
2. *Because we operate at* $IC = IC_{CRIT}$, *the operating current will be lower at a smaller width. This also reduces* g_m.
3. *At* $IC = IC_{CRIT}$ *the transconductance is no longer proportional with* V_{eff}. *With results in an even smaller value for* g_m.

These deviations result in a larger value of the source-referred noise, although it appears that $5nV/\sqrt{Hz}$ *is still feasible.*

```
# Find the width for the lowest noise
htmlPage('Noise performance optimization')
# Delete the numeric definition of W so we can calculate the optimum value
# symbolically
i1.delPar('W')
# Keep IC at IC_CRIT
i1.defPar('ID', '7.44m*W/66u');
Svi_f_W = i1.execute().inoise
W = sp.Symbol('W')
# Find optimum value of W
W_opt = sp.solve(sp.diff(Svi_f_W, W), W)
for w in W_opt:
    if w > 0:
        i1.defPar('W', w)
        print(w)
text2html('The optimum device width $W_{opt}$ is found as: ' +
          '%s'%(sp.N(i1.getParValue('W'), ini.disp)) + ' [m].')
```

```
text2html('At this width we have in input capacitance $c_{iss}$ of: ' +
          '%s'%(sp.N(i1.getParValue('c_iss_X1'), ini.disp)) + ' [F],')
text2html('a drain current $I_{DS}$ of: %s'%(sp.N(i1.getParValue('ID'),
          ini.disp)) + ' [A],')
text2html('and a transadmittance $g_m$ of: %s'%(sp.N(i1.getParValue('g_m_X1'),
          ini.disp)) + ' [S]')
text2html('The plot below shows the total source referred noise over a ' +
          'frequency range from 0.1GHZ to 1GHz as a function of the ' +
          'device width, with the inversion coefficient held constant at ' +
          '$IC_{CRIT}$.')
f_max = 1e9
f_min = 1e8
B = f_max-f_min
i1.defPar('Vni', sp.sqrt(Svi_f_W*B))
# Redefine the width so it can be used as sweep variable, any value is K
i1.defPar('W', 0)
i1.setDataType('params')
result = i1.execute()
fig_Vni_W = plotSweep('Vni_W', 'Source-referred noise voltage versus ' +
                      'width: 0.1GHz-1GHz',
                     result, 10, 200, 200, sweepVar = 'W', sweepScale = 'u',
                     funcType = 'param', xUnits = 'm', yVar = 'Vni',
                     yUnits = 'V', yScale='u', show = True)
fig2html(fig_Vni_W, 500)
```

Noise analysis



A1 Include library

SLiCAP.lib

A2 Parameter definitions

.param C_s=0.1p W=66u L=180n ID=7.44m IG=0

The inversion coefficient IC equals: 39.76

The critical inversion coefficient IC_{CRIT} equals: 31.71

The transconductance g_m equals: 0.02183

Source referred noise

The spectrum of the source-referred voltage noise [V/rt(Hz)] is: 1.771e-9

The plot below shows the source-referred noise spectrum from 100MHz to 100GHz; as expected, it does not depend on the frequency.

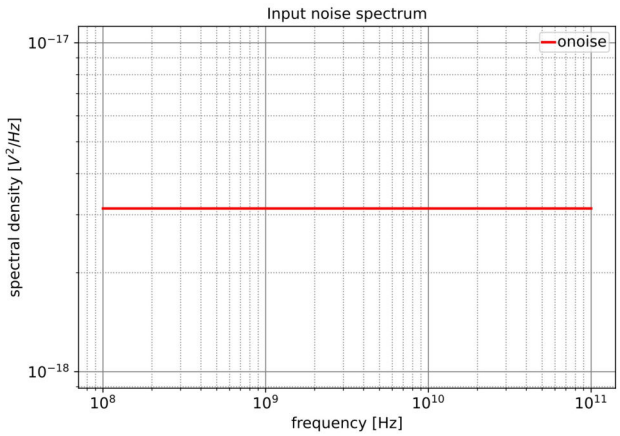


Figure 5.39: SLiCAP simulation results, page 1.

Noise performance optimization

The optimum device width W_{opt} is found as: 5.742e-5 [m].

At this width we have in input capacitance c_{iss} of: 1.000e-13 [F],

a drain current I_{DS} of: 0.006472 [A],

and a transadmittance g_m of: 0.01899 [S]

The plot below shows the total source referred noise over a frequency range from 0.1GHZ to 1GHz as a function of the device width, with the inversion coefficient held constant at IC_{CRIT}.

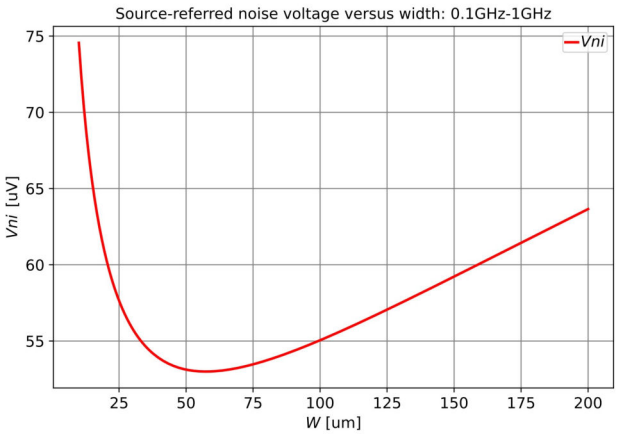


Figure 5.40: SLiCAP simulation results, page 2.

Model inaccuracies

1. Model inaccuracies are identical as those mentioned with the resistive source. The right half plane zero in the current gain causes a change in the frequency spectrum of the source-referred voltage noise. In this case it gives it a low-pass character.

2. The contributions of the $\frac{1}{f}$ noise and the noise current associated with the gate leakage current I_G have been ignored.

Conclusion

The following conclusion can be drawn:

> A CS stage driven from a capacitive voltage source, has the lowest noise figure if it operates in the saturation region and in strong inversion with g_m as large as possible and with its device width designed such that: $c_{iss} \approx C_s$. The spectral density of the total source-referred noise voltage can then be approximated by (5.83).

5.4.4 Noise minimization for capacitive current source

Figure 5.41 shows a setup in which an electrical signal is generated by a capacitive current source.

In order to determine the source-referred signal to noise ratio, the voltage source Bi_d needs to be converted into a current source. The spectral density S_{in} of the total source-referred current noise i_n is obtained as

$$S_{in} = |Bj\omega C_s + D|^2 \, S_{id} \; [\mathrm{A^2Hz^{-1}}]. \tag{5.88}$$

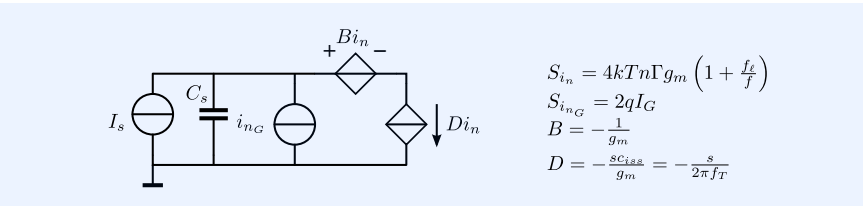


Figure 5.41: Equivalent input noise model of a CS stage driven from a capacitive current source.

After substitution of $B = -\frac{1}{g_m}$, $D = -\frac{j\omega c_{iss}}{g_m}$ and $S_{id} = 4kTn\Gamma g_m$, this can be written as

$$S_{in} = \frac{16kTn\Gamma\pi^2 f^2}{g_m} (C_s + c_{iss})^2 \ [\text{A}^2\text{Hz}^{-1}]. \tag{5.89}$$

The mean square value $\overline{i_n^2}$ of the source-referred current noise is obtained after integration of this spectrum over the frequency range of interest, say $f_{\min} \cdots f_{\max}$:

$$\overline{i_n^2} = \int_{f_{\min}}^{f_{\max}} S_{in} df = \frac{16kTn\Gamma\pi^2 (C_s + c_{iss})^2}{3g_m} \left(f_{\max}^3 - f_{\min}^3\right) \ [\text{A}^2]. \tag{5.90}$$

Similar as in the previous cases, both g_m and c_{iss} can be written as a function of the device width W.

Substitution of (5.58) and (5.72) in (5.90) yields

$$\overline{i_n^2} = \frac{16kTn^2 L\Gamma\pi^2}{3\beta_{sq} V_{eff}} \left(f_{\max}^3 - f_{\min}^3\right) \left(\frac{C_s^2}{W} + 2LC'_{OX}C_s + WL^2C'^2_{OX}\right) \ [\text{A}^2]. \tag{5.91}$$

Hence, the mean square value of the total source-referred current noise has a minimum value if

$$\frac{W^2L^2C'^2_{OX}}{C_s^2} = 1. \tag{5.92}$$

If we use the approximation for c_{iss} from (5.72) this condition is equivalent to

$$c_{iss} = C_s. \tag{5.93}$$

The minimum of the mean square value of the source-referred current noise is found after substitution of the optimum width in (5.91):

$$\overline{i_n^2} = \frac{64kTn\Gamma\pi^2 C_s^2}{3g_m} \left(f_{\max}^3 - f_{\min}^3\right) \ [\text{A}^2]. \tag{5.94}$$

At this optimum width the spectral density of the source-referred noise equals

$$S_{in} = \frac{64kTn\Gamma\pi^2 f^2 C_s^2}{g_m} \ [\text{A}^2\text{Hz}^{-1}]. \tag{5.95}$$

In the following example, we will evaluate the feasibility of a low-noise CS stage driven from a capacitive current source.

Example 5.14

The current of a capacitive current source with an internal capacitance C_s of 100fF needs to be converted into a voltage. The information is contained in a frequency range from 100MHz$\cdots$1GHz. The RMS value of the source-referred noise current over the frequency range of interest should be less than 50nA. The amplifier should be realized in a 180nm CMOS process. The NMOS process characteristics are:

1. *Technology current: $I_0 = 640nA$*

2. *Critical inversion coefficient:* $IC_{CRIT} = 31.7$
3. *Effective gate-source voltage at critical inversion:* $V_{eff_{CRIT}} = 393mV$
4. *Normalized oxide capacitance:* $C'_{OX} = 0.0084Fm^{-2}$
5. *Zero field carrier mobility:* $\mu_0 = 0.042$
6. *Substrate factor:* $n = 1.35$
7. *Noise constant:* $\Gamma = 0.7$

We will check if the required noise performance can be met with an NMOS CS stage.

The width and the operating current for optimum noise can be taken from the previous example. With SLiCAP *we found:* $W_{opt} = 57.4\mu m$ *and with* $IC = IC_{CRIT}$ *the operating current should equal* $6.47mA$*. The transconductance in this operating point equals* $22.46 \times 10^{-3} AV^{-1}$*.*

With the aid of (5.94) we find

$$\overline{i_n^2} = 4.1 \times 10^{-16}\ A^2. \tag{5.96}$$

Hence, the RMS value of the total source-referred current noise equals $20.3nA$ *and we may conclude that the required noise performance seems feasible.*

In the following example we will check the above result with SLiCAP. Since the default value of Γ in SLiCAP is $\frac{2}{3}$ the result will be slightly less than predicted by (5.96).

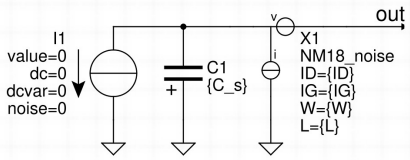
Example 5.15

The listing of the script for the determination of the spectrum and the RMS value of the total source-referred noise is shown below:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
CScapNoiseI.py
"""
from SLiCAP import *
prj = initProject('CScapNoiseI')
i1 = instruction()
i1.setCircuit('CScapNoiseI.cir')
htmlPage('Noise analysis')
img2html('CScapNoiseI.svg', 300)
# Discard 1/f noise
i1.defPar('KF_N18', 0)
i1.setSource('I1')
i1.setDetector('V_out')
i1.setGainType('vi')
i1.setDataType('noise')
i1.setSimType('numeric')
noiseResult = i1.execute()
head2html('Source referred noise');
text2html('The figure below shows the source referred noise spectrum ' +
          'from  100MHz to 100GHz for $W=W_{opt}$ and at $IC=IC_{CRIT}$.')
figSin = plotSweep('CScapNoiseIspectrum', 'Input noise spectrum',
                   noiseResult, 1e8, 1e11, 100, funcType='inoise', show=True)
fig2html(figSin,  500)
IniRMS = rmsNoise(noiseResult, 'inoise', 100e6, 1e9);
text2html('The total source referred RMS current noise $i_{ni}$ amounts: ' +
          '%s [A]'%(sp.N(IniRMS, ini.disp)))
```

Figure 5.42 shows the results of the SLiCAP *simulation. They are as expected.*

Noise analysis



A1 Include library

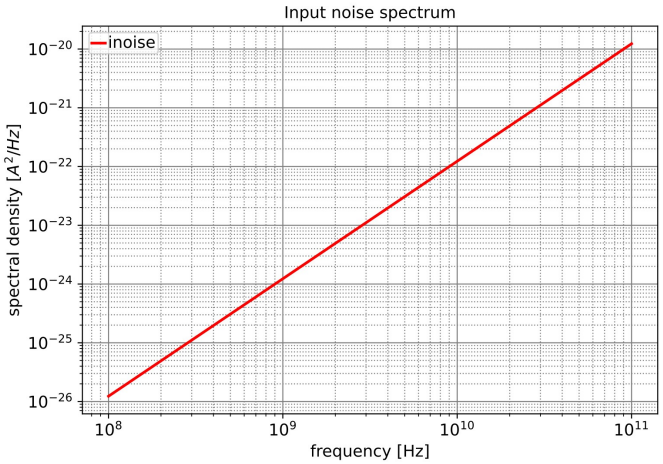
SLiCAP.lib

A2 Parameter definitions

.param C_s=0.1p W=57.4u L=180n ID=6.47m IG=0

Source referred noise

The figure below shows the source referred noise spectrum from 100MHz to 100GHz for $W = W_{opt}$ and at $IC = IC_{CRIT}$.



The total source referred RMS current noise i_{ni} amounts: 2.026e-8 [A]

Figure 5.42: SLiCAP simulation results.

Model inaccuracies

1. Model inaccuracies are identical as those mentioned with the resistive source. The right half plane zero in the current gain causes a change in the frequency spectrum of the source-referred voltage noise. In this case it adds a frequency independent part (noise floor) to the spectrum.
2. The contributions of the $\frac{1}{f}$ noise and the noise current associated with the gate leakage current I_G have been ignored.

Conclusion

The following conclusion can be drawn:

> A CS stage driven from a capacitive current source has the lowest noise figure if it operates in the saturation region and in strong inversion with g_m as large as possible and with its device width designed such that: $c_{iss} \approx C_s$. The spectral density of the total source-referred noise voltage can then be approximated by (5.95). The mean square value over a uniformly weighted frequency range from $f_{\min} \cdots f_{\max}$ is given by (5.83).

5.5 Conclusions

In the previous sections, we investigated the criteria to achieve optimum noise performance with a CS stage for resistive and capacitive sources. We have seen that in general a low noise contribution comes at the cost of area and current consumption. For the lowest noise contribution, the CS stage should operate in the saturation region in strong inversion and with its optimum device width. However, such operation is not always required nor desired. If the noise addition of a CS stage can be designed well above its lower limit, both the costs for area and for current consumption can be reduced.

In general, the noise design proceeds as follows:

1. Check if the requirements are feasible as this has been done in the examples. If the requirements are not feasible, the noise requirement specification imposes a show stopper on the design and either it needs to be relieved or another CMOS process with better performance parameters should be selected.

2. If the noise requirements can easily be met, reduce the current and the device width in such a way that the noise requirements are met at lower cost factors, while some reserve has been built in for device tolerances and model inaccuracies.

6 Balancing techniques

6.1 Introduction

In Chapter 5, we have seen that a CS stage operating in the saturation region can provide a large available power gain. We have also seen that the source-to-load transfer of a CS stage suffers from noise addition, nonlinearity and bandwidth limitation. In addition, we learned that a single CS stage shares one terminal of the input port with one terminal of the output port. Hence, such an amplifier stage cannot provide port isolation without using transformers. In this chapter, we will discuss the use of balancing techniques. These techniques can be applied for improvement of the port isolation and for reduction of reproducible errors due to offset and nonlinearity.

6.1.1 Additive compensation

Balancing is a form of additive compensation. We will speak of additive compensation if an undesired effect is compensated by adding an opposite effect. Additive compensation can be used to compensate for reproducible errors.

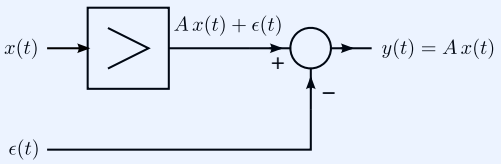


Figure 6.1: Principle of additive compensation.

The principle of additive compensation is elucidated in Figure 6.1. A known error signal $\epsilon(t)$ which equals the error caused by the nonideal behavior of an amplifier with small-signal gain A, is subtracted from the output signal of the amplifier. The resulting output signal is the amplified signal: $y(t) = Ax(t)$.

6.1.2 Balancing

Balancing is a technique in which anti-series, complementary series, anti parallel and complementary-parallel connections of amplifier stages are used to obtain such compensating effects. Parallel connections are used to add or subtract currents, while series connections are used to add or subtract voltages.

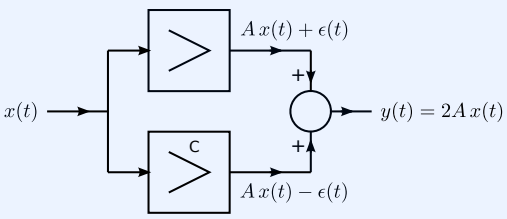


Figure 6.2: Balancing technique using amplifiers or amplifier stages with complementary characteristics.

Figure 6.2 illustrates this principle using complementary amplifiers or amplifier stages. The input signal is supplied to both amplifiers. If both amplifiers have complementary characteristics, their errors have opposite signs. After addition, the error is eliminated and the output signal is doubled.

Alternatively, one could use a setup as shown in Figure 6.3. There, both amplifiers are equal but they carry either the non-inverted or the inverted

signal at their input. Their output signals are subtracted which results in elimination of the error and doubling of the desired signal.

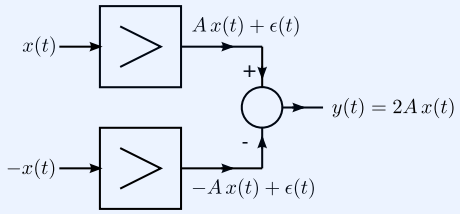


Figure 6.3: Balancing technique using identical amplifiers with inverted signals.

6.1.3 Multiplicative or cascaded compensation

Aside from additive compensation, there exists multiplicative or cascaded compensation. This type of compensation uses cascade connections of compensating systems. It can be used for linearization and for correction of the small-signal dynamic transfer of a system.

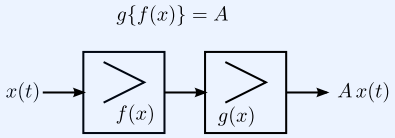


Figure 6.4: Multiplicative or cascaded compensation uses cascaded systems that together constitute a linear and instantaneous relation between the input and the output signal.

Figure 6.4 shows an arrangement in which two cascaded systems together constitute a linear and instantaneous relation between the input and the output signal.

6.1.4 Odd function synthesis

The $v-i$ characteristics of the input and the output port, as well as the transfer characteristics of an ideal amplifier, are first order odd characteristics. These functions are of the form

$$y(x) = a_1 x, \tag{6.1}$$

where y is the response signal and x is the excitation.

The instantaneous source-to-load transfer characteristics, as well as the $v-i$ characteristics of the input port and of the output port of amplifier stages such as the CS stage have both even and odd terms.

The series expansion of instantaneous nonlinear functions with odd and even nonlinearity can be written in the form

$$y(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots . \tag{6.2}$$

Odd and even functions can be synthesized from these functions. This can be seen if we write (6.2) as the sum of an odd function $y_{odd}(x)$ and an even function $y_{even}(x)$:

$$y(x) = y_{odd}(x) + y_{even}(x), \tag{6.3}$$

where:

$$y_{odd}(x) = a_1 x + a_3 x^3 + \dots , \tag{6.4}$$

and

$$y_{even}(x) = a_0 + a_2 x^2 + \dots . \tag{6.5}$$

The functions $y_{odd}(x)$ and $y_{even}(x)$ can be obtained from $y(x)$ as

$$y_{odd}(x) = \frac{1}{2}\left(y(x) - y(-x)\right), \tag{6.6}$$

$$y_{even}(x) = \frac{1}{2}\left(y(x) + y(-x)\right). \tag{6.7}$$

In order to obtain odd characteristics for amplifiers, even order terms can be compensated for by adding $-y(-x)$ to $y(x)$. This can be done with the aid of a complementary amplifier as shown in Figure 6.2, or by passing the $x(t)$ and $-x(t)$ through two equal amplifiers and subtracting their output signals as shown in Figure 6.3.

6.1.5 This chapter

In this chapter, we will discuss the application of balancing techniques with two-terminal elements and for two-ports. Balancing of two-terminal elements will be discussed in section 6.2. We will start with an introduction of the basic balancing techniques:

1. Anti-series connection of equal devices
2. Anti-parallel connection of equal devices
3. Series connection of complementary devices
4. Parallel connection of complementary devices.

We will see that balancing converts the biasing quantities into common-mode quantities and the signal quantities into differential-mode quantities. This makes balanced amplifiers less sensitive to changes in the operating conditions due to temperature variations. This mechanism is often referred to as *offset compensation* and compensation of the *offset drift*.

In section 6.3 we will discuss the balancing of two-ports. Practical implementations of balancing techniques with the basic amplifier stages will be discussed in section 6.4.

6.2 Balancing of two-terminal devices

In this section, we will discuss the balancing of two-terminal elements using complementary elements or reversely connected elements.

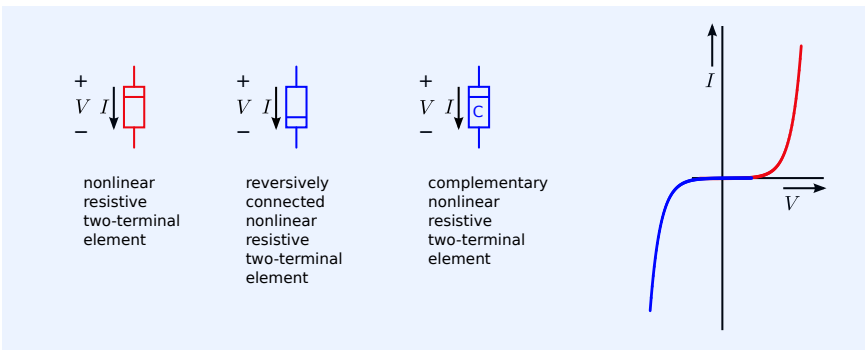


Figure 6.5: $v - i$ Characteristics of a two-terminal device (red), of its reversely connected and of its complementary device (both: blue).

Figure 6.5 shows the static $v - i$ relations of some nonlinear two-terminal device, its reversely connected device and its complementary device. The complementary device behaves similar as the reversely connected device.

This directly follows from the definition of a complementary device (see Chapter 3).

If a nonlinear two-terminal device is placed in series with a reversely connected identical device, we speak of an anti-series connection of equal devices or simply of an *anti-series connection*. If such a device is placed in series with its complementary device, we speak of a series connection of complementary devices or simply of a *complementary-series connection*. Anti-series and complementary-series connection can be used to obtain odd $v-i$ characteristics. This technique will be discussed in section 6.2.1.

If a nonlinear two-terminal device is placed in parallel with a reversely connected identical device we speak of an anti-parallel connection of equal devices or simply of a *anti-parallel connection*. If such a device is placed in parallel with its complementary device, we speak of a parallel connection of complementary devices or simply of a *complementary-parallel connection*. Complementary-parallel connections and anti-parallel connections can be used to obtain odd $v-i$ characteristics. This technique will be discussed in section 6.2.2.

6.2.1 Anti-series and complementary-series connection

Figure 6.6 shows the complementary-series connection and the anti-series connection of two nonlinear resistors. Let us assume that the $v-i$ relation of the nonlinear resistor is given by

$$v = v(i). \tag{6.8}$$

The $v-i$ relation of the anti-series or complementary-series connection can then be obtained as

$$v_s = v(i_s) - v(-i_s). \tag{6.9}$$

If we write $v(i)$ as a series expansion, we obtain

$$v(i) = a_o + a_1 i + a_2 i^2 + a_3 i^3 + \dots . \tag{6.10}$$

The series expansion of the $v-i$ relation of the anti-series or complementary-series connection becomes

$$\begin{aligned} v_s = v(i_s) - v(-i_s) = & \qquad (6.11)\\ & + a_0 + a_1 i_s + a_2 i_s^2 + a_3 i_s^3 + \dots \\ & - a_0 + a_1 i_s - a_2 i_s^2 + a_3 i_s^3 - \dots , \qquad (6.12)\\ & = 2\left(a_1 i_s + a_3 i_s^3 + \dots\right). \qquad (6.13) \end{aligned}$$

The resulting $v_s - i_s$ relation is an odd function. All even terms are canceled out, while the odd terms have twice the value of the corresponding terms of the $v-i$ relation of the single element.

Figure 6.7A shows a graph of $v(i)$ of a nonlinear resistor. Figure 6.7B shows the construction of the $v-i$ relation of the complementary-series or the anti-series connection from $v(i)$.

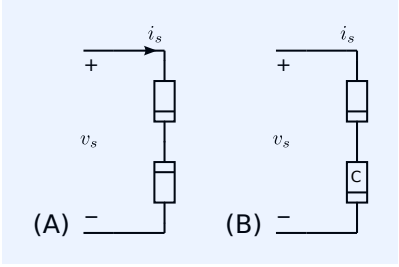


Figure 6.6:
A: anti-series connection of identical devices
B: series connection of complementary devices

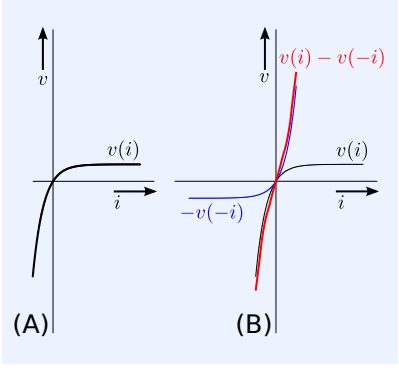


Figure 6.7:
A: $v-i$ relation of a non-linear resistor
B: construction of the $v-i$ relation of the anti-series or complementary series connection.

Differential-mode and common-mode quantities

Until now, we have studied the behavior of anti-series and complementary-series connections, isolated from their electrical environment. In practice there may exist a connection and/or parasitic current paths between the elements of the interconnected devices and other nodes of the circuits. An anti-series or a complementary-series connection is only *truly balanced* if there is no current flow to the reference node. This will always be the case if the

common-mode (signal) voltage equals zero and if the signal voltage at interconnection of the two devices equals zero. If the circuit is not truly balanced the $v(i)$ relation of the interconnected devices may deviate from the one described.

We will now study the common-mode behavior of two commonly used configurations of interconnected devices.

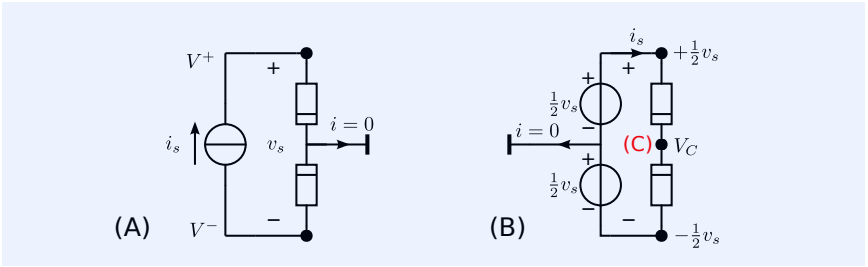


Figure 6.8: Common-mode and differential mode quantities in anti-series or complementary series connected two-terminal devices.
A: Anti-series connection with grounded interconnection point.
B: Anti-series connection with split input voltage source.

Figure 6.8A shows a situation in which the common node of the two interconnected resistors has been connected to the reference node. If this is the only connection between the interconnected resistors and the reference node, there will be no current flow to this reference. In this arrangement the common-mode voltage v_{cm} of the interconnected resistors is defined as:

$$v_{cm} = \frac{V^+ + V^-}{2}, \tag{6.14}$$

with $V^+ = v(i_s)$ and $V^- = v(-i_s)$ we may write:

$$\begin{aligned} v_{cm} = \frac{1}{2}\left(v(i_s) + v(-i_s)\right) = & \qquad (6.15)\\ & + a_0 + a_1 i_s + a_2 i_s^2 + a_3 i_s^3 + \ldots \\ & + a_0 - a_1 i_s + a_2 i_s^2 - a_3 i_s^3 + \ldots\,, \qquad (6.16)\\ & = a_0 + a_2 i_s^2 + \ldots\,. \qquad (6.17) \end{aligned}$$

Hence, the common-mode voltage comprises all the even terms while the differential-mode voltage v_s comprises all the odd terms of the function $v(i)$.

Figure 6.8B shows an arrangement in which the anti-series connection is driven from two voltage sources, carrying $\pm\frac{1}{2}v_s$ with respect to the reference node. The common-mode voltage now equals zero, while the voltage at node 'C' comprises all the even terms of $v(i)$:

$$V_C = -a_0 - a_2 i_s^2 - \ldots\,. \tag{6.18}$$

Any impedance between node 'C' and the reference node will change the $v(i)$ relation of the series connection. In fact, due to the rectifying properties of even functions, any parasitic capacitance or inductance between node 'C' and the reference node changes the operating point of the interconnected devices.

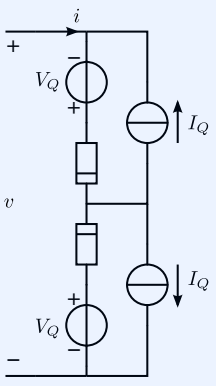


Figure 6.9: Anti-series connection of two biased two-terminal devices.

Anti-series connections of biased resistive elements

We will now study the behavior of anti-series and complementary series connection of biased devices. We will see that if such interconnections are used in a truly balanced way, the biasing can be done with common-mode current sources only.

The biasing of two-terminal resistive elements has been discussed in Chapter 3. Figure 6.9 shows the anti-series connection of two biased nonlinear resistors. The quiescent operating point Q of the resistors is fixed with a bias

voltage source V_Q and a bias current source I_Q. If the $v-i$ relation of the devices is described by the function $v(i)$ the bias voltage can be obtained from this function operating on I_Q:

$$V_Q = v(I_Q). \tag{6.19}$$

Figure 6.10 shows the construction of the $v-i$ characteristic of the anti-series connection from the single-device $v-i$ characteristic and the selection of the operating point.

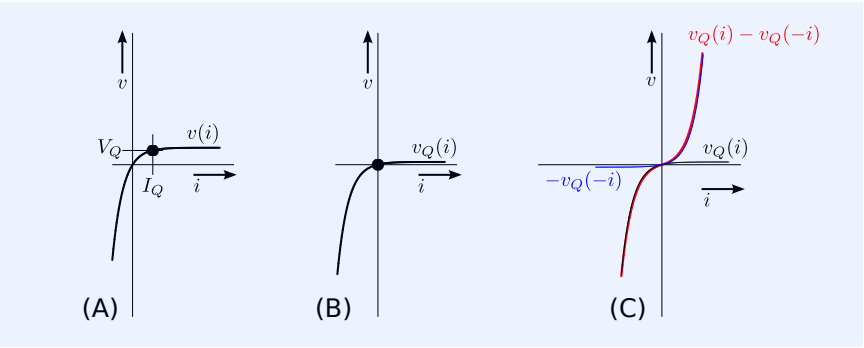


Figure 6.10:
A: $v-i$ relation of a non-linear resistor and selection of the quiescent operating point
B: $v-i$ relation of the biased device
C: construction of the $v-i$ relation of the anti-series connection of the biased devices

Common-mode biasing

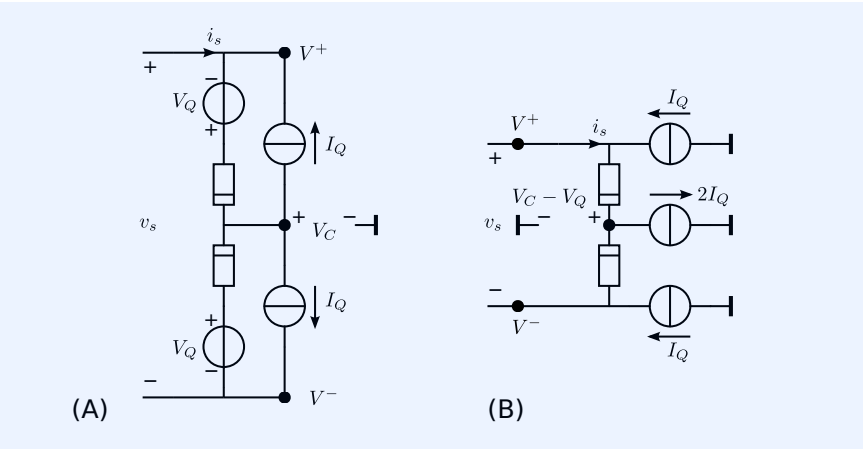


Figure 6.11: Biasing of an anti-series connection of two biased two-terminal devices.
A: Circuit from Figure 6.9 with nodal voltages V^+, V^- and V_C
B: Common-mode biasing of the circuit from Figure 6.9

Until now, the anti-series connection is considered to be isolated from the reference node. In Figure 6.11A the voltages at the outer nodes of the series connections are assumed V^+ and V^-; these values need to be defined by the electrical environment of this anti-series connection. The quiescent voltage at node 'C' equals the common-mode voltage $\frac{1}{2}(V^+ + V^-)$.

Figure 6.11B shows an alternative biasing scheme using exclusively common-mode bias sources. The current sources providing I_Q in Figure 6.9 have been redirected over the reference node using the current split theorem. The two bias voltage sources are transformed into one common-mode voltage source using the Blakesley[Blakesley1994][1] transformation. This voltage only changes the voltage across the current source that carries $2I_Q$. The quiescent voltage at the interconnection of the two resistors has changed to $\frac{1}{2}(V^+ + V^-) - V_Q$.

[1] T. A. Blakesley. A New Electrical Theorem. *Proc. Phys. Soc. London*, 13:65–67, 1994

Small-signal equivalent circuit

Until now, we have only studied the instantaneous behavior of anti-series or complementary series connections. An anti-series connection of two-terminal elements that also exhibits a nonlinear voltage-charge relation $q = q(v)$ and/or a nonlinear current-flux $\phi = \phi(i)$ relation, yields odd characteristics for these relations as well.

We will now study the small-signal behavior of anti-series connected nonlinear devices.

The small-signal resistance, capacitance and inductance of a biased nonlinear two-terminal element are defined as:

1. The small-signal resistance r_Q in an operating point $i = I_Q$ of a device modeled with a nonlinear $v - i$ relation $v(i)$ is defined as

$$r_Q = \left.\frac{dv(i)}{di}\right|_{i=I_Q}. \tag{6.20}$$

2. The small-signal capacitance c_Q in an operating point $v = V_Q$ of a device modeled with a nonlinear $q - v$ relation $q(v)$ is defined as

$$c_Q = \left.\frac{dq(v)}{dv}\right|_{v=V_Q}. \tag{6.21}$$

3. The small-signal inductance l_Q in an operating point $i = I_Q$ of a device modeled with a nonlinear $\phi - i$ relation $\phi(i)$ is defined as

$$l_Q = \left.\frac{d\phi(i)}{di}\right|_{i=I_Q}. \tag{6.22}$$

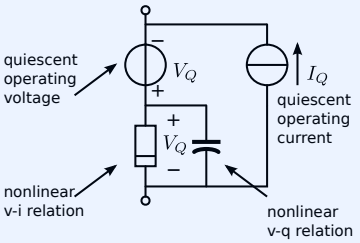


Figure 6.12: Biased two-terminal device with nonlinear $v - i$ and nonlinear $v - q$ relation.

Figure 6.12 shows the network model of a biased two-terminal device with a nonlinear $v - i$ relation and a nonlinear $q - v$ relation. Such behavior can be found in PN diodes.

Figure 6.13A, shows the anti-series connection of two devices. The current i_s represents a deviation from the quiescent operating point. It is not necessarily a small-signal current.

The impedance can be represented by the series connection of two parallel RC networks as depicted in Figure 6.13B. The small-signal impedance z_s of this circuit can then be written as

$$z_s = \frac{r_1}{1 + s\tau_1} + \frac{r_2}{1 + s\tau_2}, \tag{6.23}$$

where $\tau_1 = r_1 c_1$ and $\tau_2 = r_2 c_2$. The resistances and capacitances are found as the derivative of $v(i)$ and $q(v)$ in the device's operating point, respectively.

This impedance has two poles and one zero which is located between the two poles

$$z_s = (r_1 + r_2)\,\frac{1 + s\left(\tau_1 \frac{r_2}{r_1+r_2} + \tau_2 \frac{r_1}{r_1+r_2}\right)}{(1 + s\tau_1)\,(1 + s\tau_2)}. \tag{6.24}$$

Only if the two time constants are equal the impedance can be written as a single-pole function. If $\tau_1 = \tau_2 = \tau$ we may write

$$z_s = \frac{r_1 + r_2}{1 + s\tau}. \tag{6.25}$$

At zero excursion from the operating point ($i = 0$, and $v = 0$), this impedance can be simplified to a parallel connection of a resistance $2r_Q$ and a capacitance $\frac{1}{2}c_Q$, as shown in Figure 6.13C.

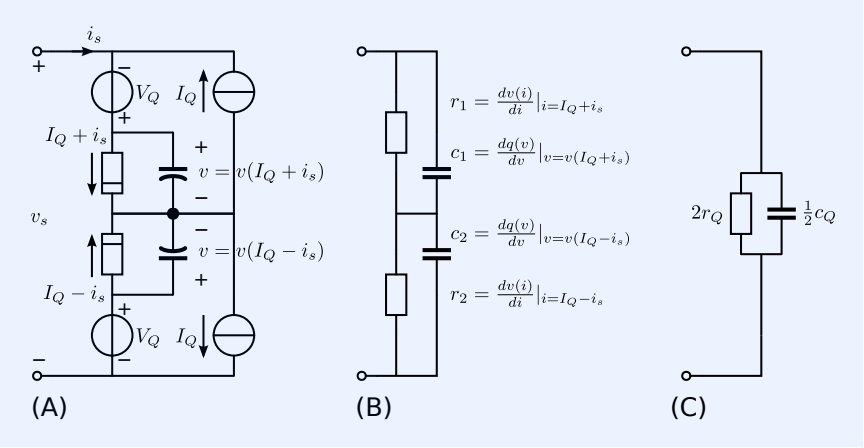


Figure 6.13:
A: Anti-series connection of nonlinear two-terminal devices.
B: Small-signal equivalent circuit
C: Small-signal equivalent circuit in the quiescent operating point ($i_s = 0$).

Stationary noise behavior

The spectral density S_{v_n} of the thermal noise voltage v_n associated with an impedance Z can be written as

$$S_{v_n} = 4kT\,\mathrm{Re}(Z)\ \mathrm{V^2Hz^{-1}}. \tag{6.26}$$

Hence, the voltage noise spectrum in [$\mathrm{V^2Hz^{-1}}$] of an anti-series connection, which is operating in its quiescent operating point, is twice that of its constituting biased device.

Since the small-signal impedance of nonlinear two-terminal devices depends on the operating conditions, the associated noise spectrum may vary with signal.

6.2.2 Anti-parallel and complementary-parallel connection

Anti-parallel connection or complementary-parallel connection is dual to anti-series connection or complementary-series connection. All conclusions from anti-series connection hold for anti-parallel connection if voltage is replaced with current, resistance with conductance, impedance with admittance, charge with flux, and series connection with parallel connection and vice versa.

Anti-parallel and complementary-parallel connection of resistive elements

Figure 6.14 shows the complementary-parallel connection and the anti-parallel connection of two nonlinear resistors. Let us assume that the $v - i$ relation of the nonlinear resistor is given by

$$i = i(v). \tag{6.27}$$

The $v - i$ relation of the anti-parallel or complementary-parallel connection can then be obtained as

$$i_p = i(v_p) - i(-v_p). \tag{6.28}$$

If we write $v(i)$ as a series expansion, we obtain

$$i(v) = a_o + a_1 v + a_2 v^2 + a_3 v^3 + \ldots\ . \tag{6.29}$$

The series expansion of the $v - i$ relation of the anti-parallel or complementary-

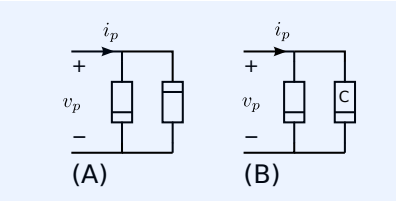


Figure 6.14:
A: anti-parallel connection of identical devices
B: parallel connection of complementary devices

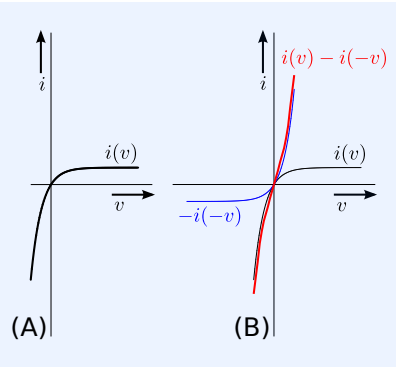


Figure 6.15:
A: $i - v$ relation of a non-linear resistor
B: construction of the $i - v$ relation of the anti-parallel or complementary parallel connection.

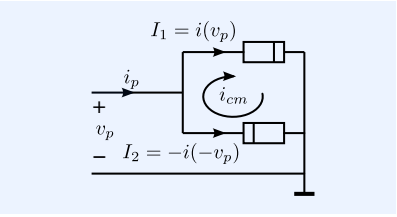


Figure 6.16: Common-mode current in anti-parallel or complementary parallel connected two-terminal devices.

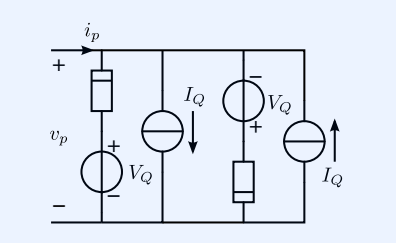


Figure 6.17: Anti-parallel connection of two biased two-terminal devices.

parallel connection becomes

$$i_p = i(v_p) - i(-v_p) = \tag{6.30}$$
$$+ a_o + a_1 v_p + a_2 v_p^2 + a_3 v_p^3 + ...$$
$$- a_o + a_1 v_p - a_2 v_p^2 + a_3 v_p^3 + ... , \tag{6.31}$$
$$= 2\left(a_1 v_p + a_3 v_p^3 + ...\right). \tag{6.32}$$

The resulting $v_p - i_p$ relation is an odd function. All even terms are canceled out and the odd terms have twice the value of the corresponding terms of the $v - i$ relation of the single element.

Figure 6.15A shows a graph of $i(v)$ of a nonlinear resistor. Figure 6.15B shows the construction of the $v - i$ relation of the complementary-parallel or the anti-parallel connection from $i(v)$.

Differential-mode and common-mode quantities

Until now, we have studied the behavior of anti-parallel and the complementary-parallel connection, isolated from their electrical environment.

Figure 6.16 shows a situation in which one side of the parallel connection has been connected to the ground.

In this arrangement, the signal current i_p is the sum of I_2 and I_2. The common-mode current i_{cm} in the interconnected resistors is now defined as

$$i_{cm} = \frac{I_1 - I_2}{2}, \tag{6.33}$$

with $I_1 = i(v_p)$ and $I_2 = -i(-v_p)$ we may write

$$i_{cm} = \frac{1}{2}\left(i(v_p) + i(-v_p)\right) = \tag{6.34}$$
$$+ a_o + a_1 v_p + a_2 v_p^2 + a_3 v_p^3 + ...$$
$$+ a_o - a_1 v_p + a_2 v_p^2 - a_3 v_p^3 + ... , \tag{6.35}$$
$$= 2\left(a_o v_p + a_2 v_p^2 + ...\right). \tag{6.36}$$

Hence, i_{cm} comprises all the even terms while the signal current i_p comprises all the odd terms of the function $i(v)$.

Anti-parallel connection of biased resistive elements

We will now study the behavior of anti-parallel and complementary-parallel connection of biased devices. We will see that the biasing can be done with voltage sources only.

The biasing of two-terminal resistive elements has been discussed in Chapter 3. Figure 6.17 shows the anti-parallel connection of two biased nonlinear resistors. The quiescent operating point Q of the resistors is fixed with a bias voltage source V_Q and a bias current source I_Q. If the $v - i$ relation of the devices is described by the function $i(v)$ we the bias voltage can be obtained from this function operating on V_Q:

$$I_Q = i(V_Q) \tag{6.37}$$

Since the sum of the two bias current sources equals zero, they can be omitted. In fact the two bias voltage sources deliver the bias current.

Figure 6.18 shows the construction of the $v - i$ characteristic of the anti-parallel connection from the single-device $v - i$ characteristic and the selection of the operating point.

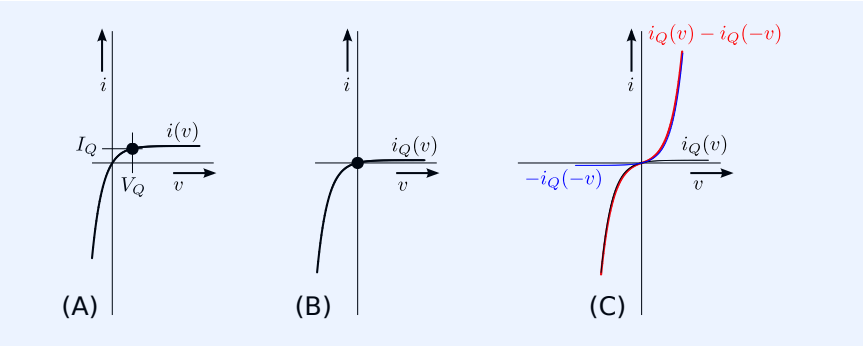


Figure 6.18:
A: $v - i$ relation of a non-linear resistor and selection of the quiescent operating point
B: $v - i$ relation of the biased device
C: construction of the $v - i$ relation of the anti-parallel connection of the biased devices

Common-mode biasing

Figure 6.19 shows a biasing scheme of a anti-parallel connection with one of its terminals connected to ground. The directions of I_1 and I_2 are according to those in Figure 6.16. For obvious reasons the positive direction of I_2 is usually taken opposite.

Although it appears as if both devices are in series for the biasing, the circuit is a complementary-parallel or anti-parallel connection from a signal processing point of view. The function $i_p\left(v_p\right)$ is an odd function, and i_p is constituted from the sum of the device currents.

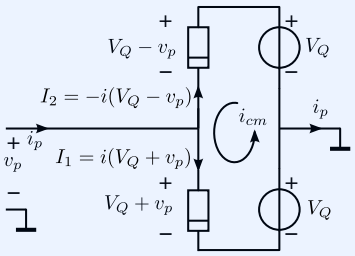


Figure 6.19: Biasing of an anti-parallel or complementary parallel connection with bias voltage sources only. The common mode-current is determined by the device characteristics and the bias voltage sources.

Small-signal equivalent circuit

In the case of the anti-series or complementary-series connection, we considered nonlinear dynamic elements that showed both a nonlinear $v - i$ relation and a nonlinear $q - v$ relation. The complementary situation is the anti-parallel or complementary-parallel connection of nonlinear devices that have both a nonlinear $v - i$ relation and a nonlinear $\phi - i$ relation. Because of its practical relevance, however, we will discuss the anti-parallel connection of devices with a nonlinear $v - i$ relation and a nonlinear $q - v$ relation. Figure 6.20A shows the anti-parallel connection two biased nonlinear devices from Figure 6.12. Figure 6.20B shows the small-signal equivalent circuit and Figure 6.20C shows the small-signal equivalent circuit that is valid for zero static excursion from the operating point.

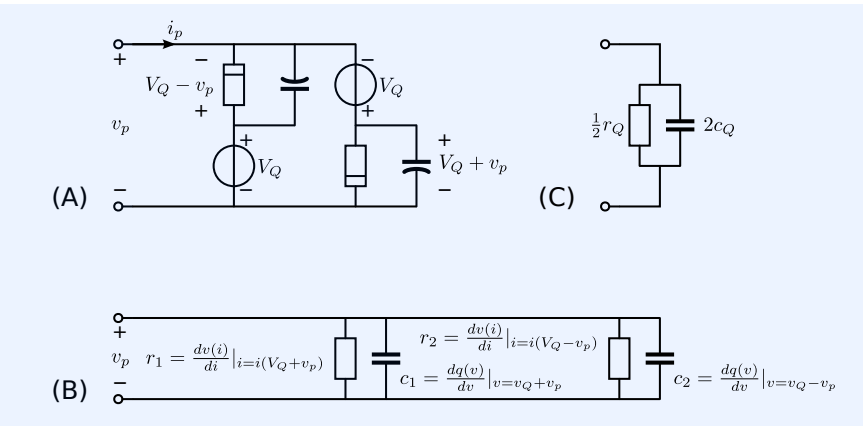


Figure 6.20: A: Anti-parallel connection of biased identical non-linear dynamic devices
B: Small-signal equivalent circuit for any static excursion from the operating point
C: Small-signal equivalent circuit for zero excursion from the operating point.

Stationary noise behavior

The spectral density S_{i_n} of the thermal noise current i_n associated with an admittance Y can be written as

$$S_{i_n} = 4kT\,\mathrm{Re}(Y)\ \mathrm{A^2Hz^{-1}}. \tag{6.38}$$

Hence, the current noise spectrum in [$\mathrm{A^2Hz^{-1}}$] of an anti-parallel connection, operating in its quiescent operating point is twice that of its constituting biased device.

Since the small-signal impedance of nonlinear two-terminal devices depends on the operating conditions, the associated noise spectrum may vary with signal.

6.3 Balancing of two-ports

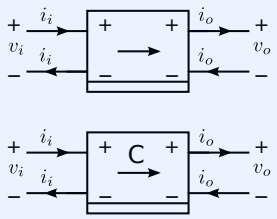


Figure 6.21: Upper: Symbol of a nonlinear two-port
Lower: Complementary version of the nonlinear two-port.

Before we will develop the balanced versions of the CS stage, we will apply the balancing techniques from the previous section to two-ports.

In a similar way as with two-terminal resistive elements, odd functions can be obtained through application of anti-series, complementary-series, anti-parallel and complementary-parallel connections at the ports.

Figure 6.21 shows the schematic symbols that we will use for a nonlinear two-port and for its complementary version. Their $v - i$ relations can be written as presented in section 3.3.

In section 6.3.1, we will give an overview of balanced two-ports that are constructed from anti-series, complementary-series, anti-parallel and complementary-parallel connections of two-ports. If not available, complementary devices can be constructed with the aid of transformers. Section 6.3.2 will be devoted to this topic.

6.3.1 Balanced two-port configurations

Below an inventory of balanced two-ports with odd transfer functions.

1. Anti-series or complementary-series connection of both input and output ports (Figure 6.22):

$$v_{ii} = v_i(i_i, i_o) - v_i(-i_i, -i_o), \tag{6.39}$$
$$v_{oo} = v_o(i_i, i_o) - v_o(-i_o, -i_o). \tag{6.40}$$

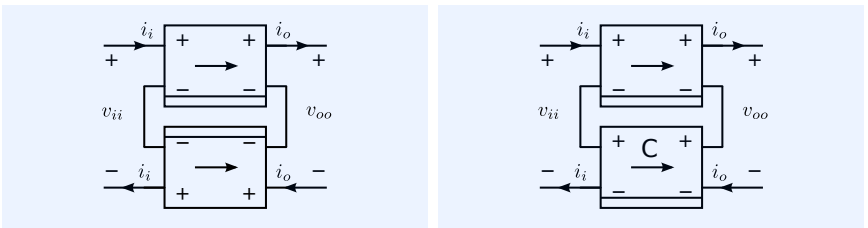


Figure 6.22: Left: anti-series connection of two-ports
Right: complementary-series connection of two-ports

2. Anti-series or complementary-series connection of the input ports and anti-parallel or complementary-parallel connection of the output ports (Figure 6.23):

$$v_{ii} = v_i(i_i, v_o) - v_i(-i_i, -v_o), \tag{6.41}$$
$$i_{oo} = i_o(i_i, v_o) - i_o(-i_i, -v_o). \tag{6.42}$$

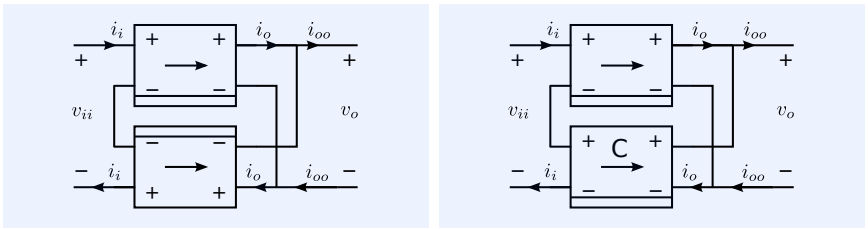


Figure 6.23: Left: input anti-series connection and output anti-parallel connection of two-ports
Right: input complementary-series connection and output complementarry-parallel connection of two-ports

3. Anti-parallel or complementary-parallel connection of the input ports and anti-series or complementary-series connection of the output ports (Figure 6.24):

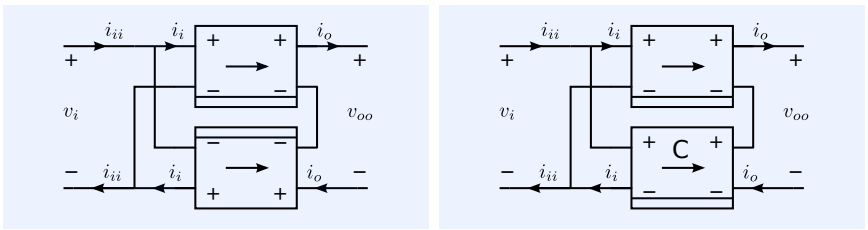


Figure 6.24: Left: input anti-parallel connection and output anti-series connection of two-ports
Right: input complementary-parallel connection and output complementary-series connection of two-ports

$$i_{ii} = i_i(v_i, i_o) - i_i(-v_i, -i_o), \tag{6.43}$$
$$v_{oo} = v_o(v_i, i_o) - v_o(-v_i, -i_o). \tag{6.44}$$

4. Anti-parallel or complementary-parallel connection of the input ports and anti-parallel or complementary-parallel connection of the output ports (Figure 6.25):

$$i_{ii} = i_i(v_i v_o) - i_i(-v_i, -v_o), \tag{6.45}$$
$$i_{oo} = i_o(v_i, v_o) - i_o(-v_i, -v_o). \tag{6.46}$$

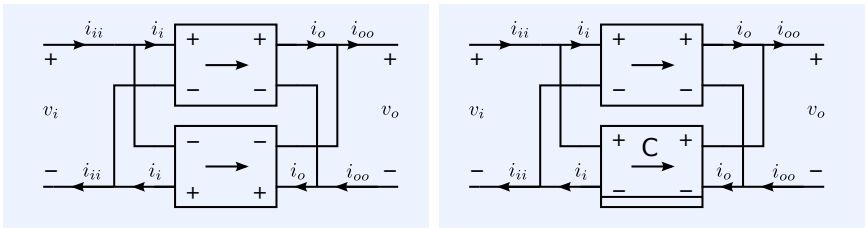


Figure 6.25: Left: input anti-parallel connection and output anti-parallel connection of two-ports
Right: input complementary-parallel connection and output complementary-parallel connection of two-ports

6.3.2 Design of complementary two-ports

Instead of using complementary elements, one can create complementary versions of a nonlinear two-ports with the aid of transformers. This is shown in Figure 6.26. This technique was used in vacuum tube amplifiers.

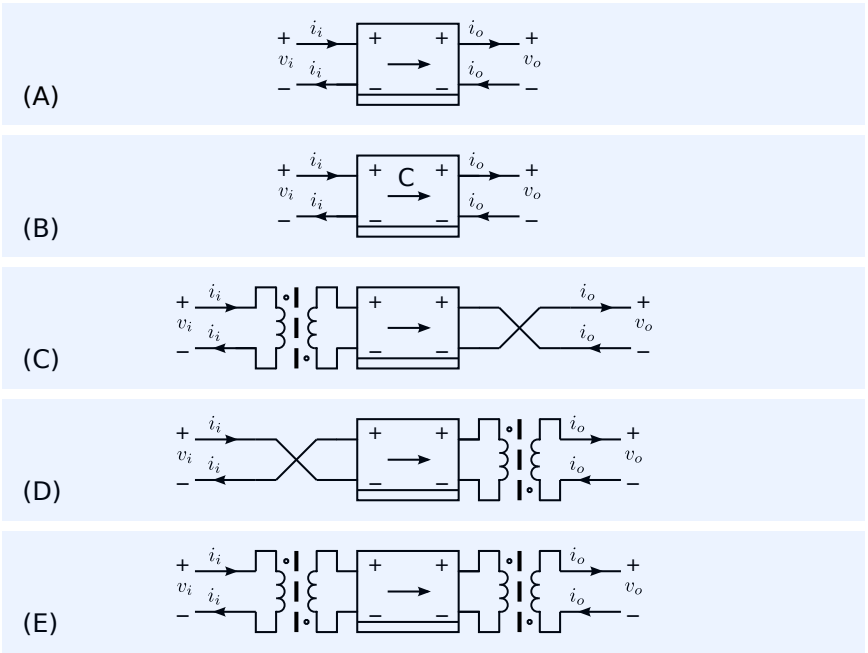


Figure 6.26: Nonlinear two-port and complementary versions. Complementary two-ports can be constructed with the aid of transformers:
A: Nonlinear two-port
B: Complementary two-port
C: Complementary two-port using isolated signal inversion at the input port and non isolated signal inversion at the output port
D: Complementary two-port using non isolated signal inversion at the input port and isolated signal inversion at the output port
E: Complementary two-port using isolated signal inversion at both ports.

6.4 Balanced CE and CS stages

Since active devices are three terminal elements, the input port and the output port of single-device amplifier stages share one common terminal. Due to this limitation, there are only two useful balanced configurations of elementary amplifier: the anti-series stage and the complementary-parallel stage.

The anti-series stage is known as the *differential pair* amplifier stage. It will be discussed in section 6.4.1.

The complementary-parallel stage is known as the *push-pull* stage. It will be discussed in section 6.4.3.

The properties of the anti-series stage and the complementary-parallel stage will be related to those of the single-device amplifier stage by considering behavioral modifications due to the application of balancing techniques.

6.4.1 Anti-series stages

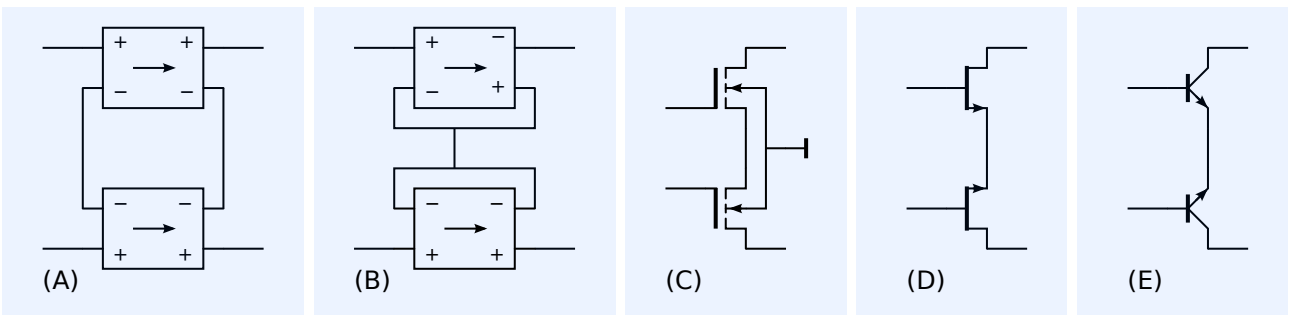


Figure 6.27: Anti-series connection of:
A: Identical two-ports
B: Identical three-terminal elements
C: IC NMOS devices
D: N channel JFETs
E: NPN transistors.

The anti-series connection of two identical devices is shown in Figure 6.27. The resulting balanced CE- and CS stages are usually called *differential pairs*.

We will first discuss the behavioral modifications in the small-signal dy-

namic behavior and the noise behavior as a result of anti-series connection of two identical two-ports. With the aid of these behavioral modifications we can easily predict the noise behavior and the small-signal behavior of all differential pair stages.

We will then discuss the behavior of the differential pair CS stage and the differential pair CE stage.

Small-signal dynamic behavior of anti-series connected stages

Let us consider the anti-series of two identical two ports as shown in Figure 6.28. The Transmission-1 matrix equation of each two port is given as

$$\begin{pmatrix} V_i \\ I_i \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_o \\ I_o \end{pmatrix}. \tag{6.47}$$

If V_{ii} is the voltage across the anti-series connection of the input ports and V_{oo} the voltage across the anti-series connection of the output ports, the Transmission-1 matrix equation of the anti-series connection can be written as

$$\begin{pmatrix} V_{ii} \\ I_i \end{pmatrix} = \begin{pmatrix} A & 2B \\ \frac{1}{2}C & D \end{pmatrix} \begin{pmatrix} V_{oo} \\ I_o \end{pmatrix}. \tag{6.48}$$

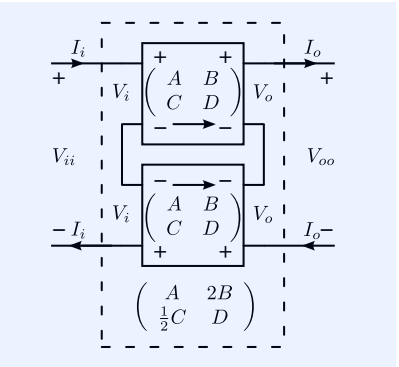


Figure 6.28: The effect of anti-series connection on the Transmission-1 matrix parameters.

Static noise behavior of anti-series stages

Figure 6.29 illustrates the way in which the equivalent input noise sources of the anti-series connection can be obtained from the equivalent input noise sources of the individual two-ports. Figure 6.29A shows the anti-series connection of the two two-ports with their individual equivalent input noise sources. This noise representation is not very convenient for evaluation of the equivalent input noise sources of the (anti) series connection. A representation with two voltage sources is more suited to this situation. With such a representation the total noise voltages at the input and at the output of the resulting two port are simply found by adding those of the constituting two-ports. An equivalent input representation can then be obtained by replacing the output voltage source with equivalent input sources, thereby using the transmission parameters of the (anti-) series connection.

Figure 6.29B shows the equivalent voltage noise representation while Figure 6.29C shows the final result.

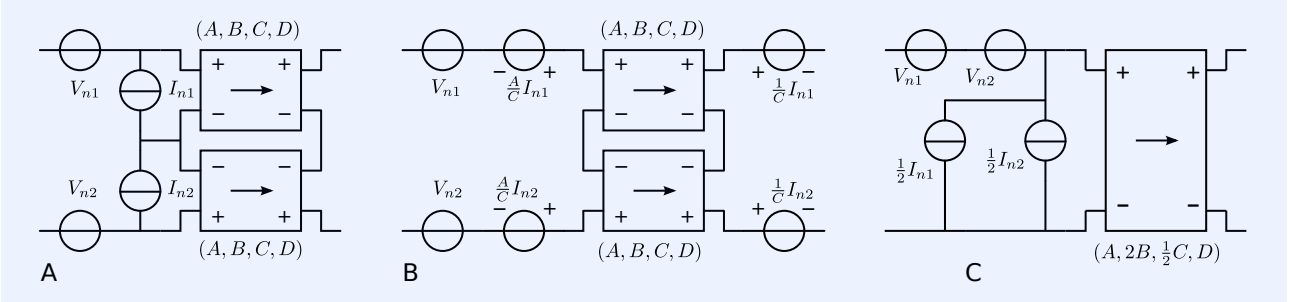


Figure 6.29: Equivalent input noise sources of (anti-) series connected two-ports.

The power spectral densities of the equivalent noise sources of the anti-series connection are thus obtained as

$$S_{vtot} = S_{vn1} + S_{vn2} \text{ [V}^2\text{/Hz]}, \tag{6.49}$$

$$S_{itot} = \frac{1}{4}S_{in1} + \frac{1}{4}S_{in2} \text{ [A}^2\text{/Hz]}. \tag{6.50}$$

If both constituting two-ports have identical noise behavior, we have: $S_{vn1} =$

$S_{vn2} = S_v$ [V²/Hz] and $S_{in1} = S_{1n2} = S_i$ [A²/Hz] and we obtain

$$S_{vtot} = 2S_v \text{ [V²/Hz]}, \tag{6.51}$$

$$S_{itot} = \frac{1}{2}S_i \text{ [A²/Hz]}. \tag{6.52}$$

6.4.2 Anti-series CS stage

In this section, we will apply the above theory about balancing to the biased anti-series CS stage. Figure 6.30 shows the anti-series connection of two biased CS stages. The DC common-mode voltages of the stage with respect to the ground have yet been undefined.

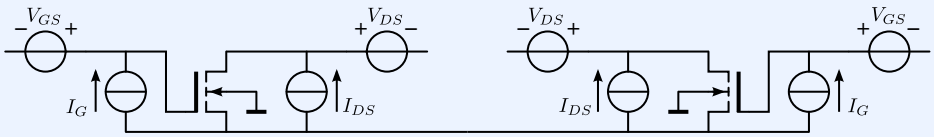


Figure 6.30: Anti-series connection of biased NMOS CS floating with respect to the substrate.

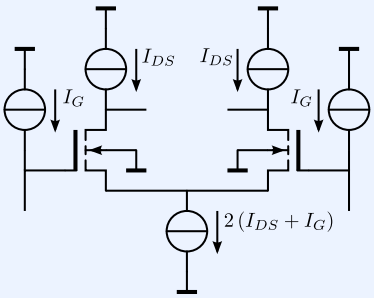


Figure 6.31: Anti-series connection of biased NMOS CS floating with respect to the substrate with its bias sources converted into common-mode bias sources.

Similar as with the anti-series connection of two-terminal elements, the bias voltages are converted into common-mode voltages and the bias current sources can be converted into common-mode current sources. The conversion of the bias current sources into common-mode sources is the result of redirecting these sources via the ground node. Figure 6.31 shows the result of this transformation.

In this figure the common-mode input voltage is defined by the voltage V_{ci}. The common-mode output voltage V_{co} equals

$$V_{co} = V_{ci} - V_{GS} + V_{DS}. \tag{6.53}$$

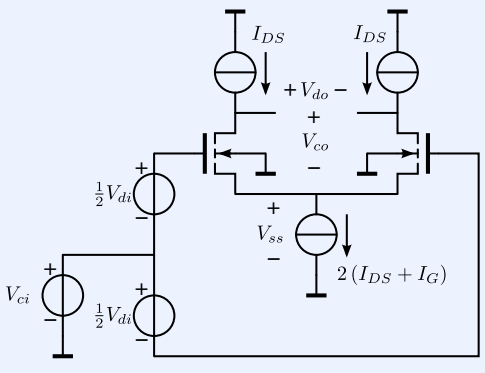
Due to the current source character of the output port of the CS stage, this common-mode output voltage will be extremely sensitive to temperature changes and device tolerances. In practice, additional measures will have to be taken to reduce this sensitivity. This will be discussed at a later stage.

If $V_{di} = 0$, V_{ss} equals

$$V_{ss} = V_{ci} - V_{GS}. \tag{6.54}$$

If $V_{di} \neq 0$, V_{ss} comprises even terms of V_d only, similar as with the anti-series connection of two-terminal elements.

Figure 6.32: Anti-series connection of biased NMOS CS stages with its input common-mode voltage defined by external circuitry.



Large-signal static behavior

The physical nature of the amplification mechanism of the active devices is a voltage-controlled current source. For this reason the voltage-to-current transfer is often considered the most important of the four transfer parameters. Figure 6.33 shows the simulation test bench for determination of the voltage-to-current transfer of the anti-series CS stage.

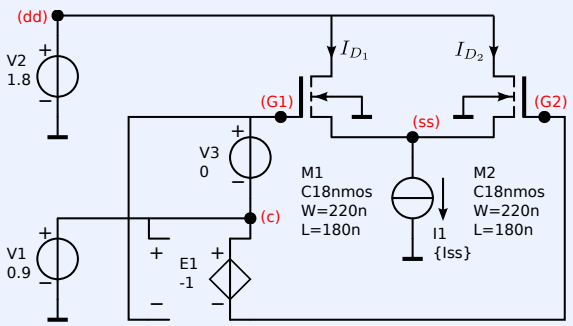


Figure 6.33: SPICE simulation test bench for plotting the DC voltage to current transfer of the anti-series CS stage.

The LTSPICE netlist of this circuit is shown below:

```
antiSeriesCS
* File:  antiSeriesCS.cir
* LTspice netlist file
.lib CMOS18TT.lib
M1 dd G1 ss 0 C18nmos W=220n L=180n
M2 dd G2 ss 0 C18nmos W=220n L=180n
V1 c 0 0.9
V2 dd 0 1.8
V3 G1 c 0
E1 c G2 c G1 -1
I1 ss 0 {Iss}
.param Iss=100n
* LTspice specific command section
.DC V3 -400m 400m 1m
.step param Iss list 100n 10u 20u 50u 100u
.end
```

The circuit from Figure 6.33 uses the biasing scheme from Figure 6.31. The common-mode input voltage V_{ci} is provided by `V1`. It is set to half the supply voltage. According to the definition of the transconductance, the output has been shorted. For proper biasing in the saturation region the drains are connected to the supply source `V2`. The stage is fully balanced because its input is driven symmetrically with respect to ground. Hence, there is no signal current flowing from the input to the output via the ground connection. The differential input voltage is twice the voltage of `V3`.

The bias current I_{ss} of this anti-series stage determines the inversion coefficient in the quiescent operating point of the devices. Since the *transconductance efficiency*[2] strongly depends on the inversion coefficient, the shape of the voltage to current transfer characteristic will strongly depend on the inversion coefficient.

[2] The transconductance efficiency is defined as the ratio of the transconductance and the drain current.

At weak inversion, the voltage-to-current transfer of a MOS transistor has an exponential character; see (4.147). In this operating region, the transconductance is proportional with the drain current. In strong inversion the voltage-to-current transfer has a quadratic relationship and the transconductance is proportional with the square root of the drain current; see (4.148). At very high inversion levels the transconductance saturates to a maximum value due to velocity saturation and vertical field mobility reduction.

The left plot from Figure 6.34 shows the simulated transfer characteristics for $I_{ss} = 100\text{nA}\cdots100\mu\text{A}$. At $I_{ss} = 100\text{nA}$, the transistors operate in weak

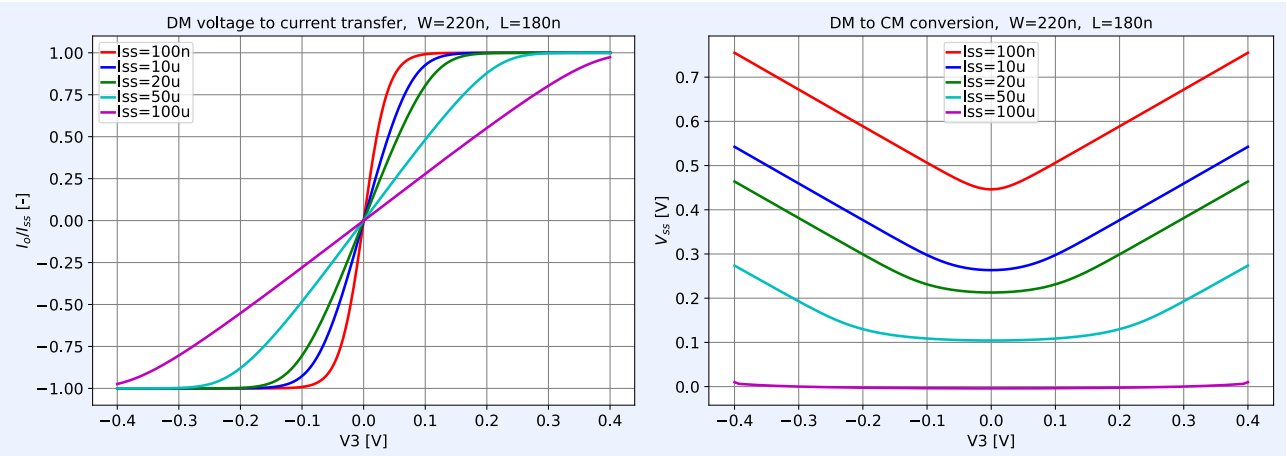


Figure 6.34: Characteristics of the anti-series CS stage, when operating from weak inversion to strong inversion with velocity saturation

Left: Normalized differential output current versus differential input voltage

Right: Voltage at source interconnection node versus differential input voltage,

inversion. The transfer characteristic can then be approximated by

$$I_{D_1} - I_{D_2} = I_{ss} \tanh \frac{V_{G_1} - V_{G_2}}{2nU_T}. \tag{6.55}$$

At about $\pm$180mV differential input voltage, the differential output current equals $\pm$99% of I_{ss}.

At $I_{ss} = 100\mu$A, the transistors operate at strong inversion with velocity saturation. In this operating region, the transconductance does not significantly depend on the drain current. The transfer is linear over a wide input voltage range, but the differential input voltage required for full differential output excursion increases to $\frac{I_{dss}}{g_m}$, where g_m is the transconductance of the unbalanced CS stage at full velocity saturation.

The right plot from Figure 6.34 shows the voltage V_{ss}. It clearly shows the rectifying properties of the even function. In the linear range of the differential-mode transfer, V_{ss} approximates $V_{ci} - V_{GS}$. Outside this range, V_{ss} increases with the absolute value of the differential-mode input voltage.

Small-signal dynamic behavior

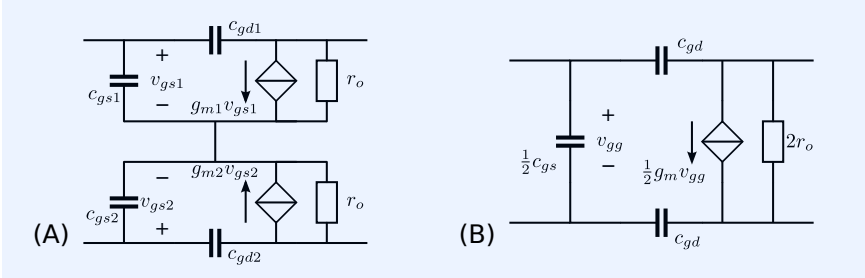


Figure 6.35: Simplified small-signal equivalent circuits of the anti-series CS stage. The bulk capacitances as well as the bulk transconductance g_b have been omited.

A: Small-signal equivalent circuit of the anti-series CS-stage

B: Small-signal equaivalent diagram of the anti-series CS stage in the quiescent operating point.

Figure 6.35A shows the simplified small-signal equivalent circuit of the anti-series CS stage. The bulk capacitances c_{gb}, c_{sb} and c_{db} have been omitted. The influence of the bulk transconductance g_b can usually be omitted. In a fully balanced environment and in the quiescent operating point it has no influence.

Figure 6.35B shows the simplified small-signal equivalent circuit for operation in the quiescent operating point. The element values are those of the

single CS stage, $g_m = g_{m1} = g_{m2}$, etc.

Differential-mode and common-mode equivalent circuits

Sometimes it is useful to model the differential-mode behavior and the common-mode behavior of balanced networks separately. In Chapter 7.7 a both practical and a theoretical modeling techniques will be presented.

Stationary noise behavior

The stationary noise behavior in fully balanced application, and in the quiescent operating point, has been described with equations (6.51) and (6.52). From these equations we can draw the following important design conclusion:

> For an anti-series CS stage to achieve the same noise performance at a given source impedance as a single CS stage, the quiescent operating current, as well as the width of each transistor of the anti-series connection, should be twice those of the single CS stage.

Hence, the cost factors for an equal noise performance with the differential pair, are four times area and four times current, when compared to the CS stage.

6.4.3 Complementary-parallel stages

The parallel connection of two complementary devices is shown in Figure 6.36. The resulting balanced CE- and CS stages are usually called *push-pull stages*.

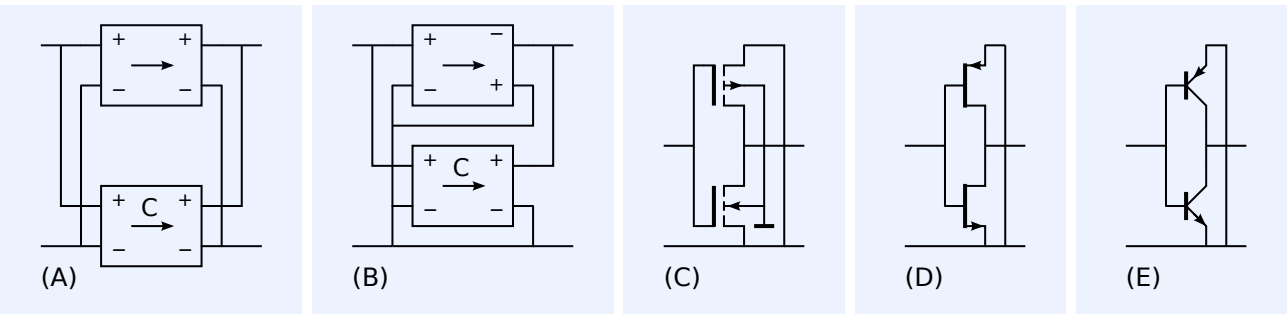


Figure 6.36: Parallel connection of:
A: Complementary two-ports
B: Complementary three-terminal elements
C: IC NMOS and PMOS devices
D: N channel and P channel JFETs
E:Bipolar NPN and PNP transistors.

We will first discuss the behavioral modifications in the small-signal dynamic behavior and in the noise behavior due to parallel connection. With the aid of these behavioral modifications we can easily predict the noise behavior and the small-signal behavior of all push-pull stages.

We will then discuss the behavior of the complementary-parallel CS stage and the complementary-parallel CE stage.

Small-signal dynamic behavior of complementary-parallel stages

Let us consider the anti-parallel of two identical two ports as shown in Figure 6.37. The Transmission-1 matrix equation of each two port is given as

$$\begin{pmatrix} V_i \\ I_i \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_o \\ I_o \end{pmatrix}. \tag{6.56}$$

If I_{ii} is the current that flows into the parallel connection of the input ports and I_{oo} the current that flows out of the parallel connection of the out-

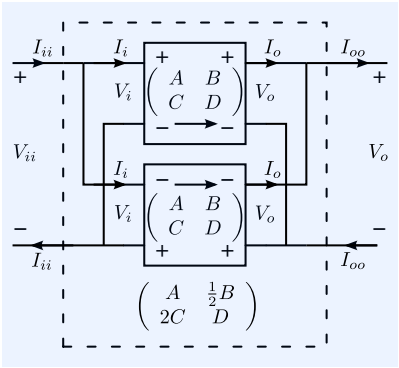


Figure 6.37: The effect of parallel connection of two-ports on the Transmission-1 matrix parameters.

put ports, the Transmission-1 matrix equation of the complementary-parallel connection can be written as

$$\begin{pmatrix} V_i \\ I_{ii} \end{pmatrix} = \begin{pmatrix} A & \frac{1}{2}B \\ 2C & D \end{pmatrix} \begin{pmatrix} V_o \\ I_{oo} \end{pmatrix}. \tag{6.57}$$

Static noise behavior of complementary-parallel stages

Figure 6.38 illustrates the way in which the equivalent input noise sources of the anti-series connection can be obtained from the equivalent input noise sources of the individual two-ports. Figure 6.38A shows the anti-parallel connection of the two two-ports with their individual equivalent input noise sources. This noise representation is not very convenient for evaluation of the equivalent input noise sources of the (anti) parallel connection. A representation with two current sources is more suited to this situation. With such a representation the total noise currents at the input and at the output of the resulting two port are simply found by adding those of the constituting two-ports. An equivalent input representation can then be obtained by replacing the output current source with equivalent input sources, thereby using the transmission parameters of the (anti) parallel connection.

Figure 6.38B shows the equivalent current noise representation, while Figure 6.38C shows the final result.

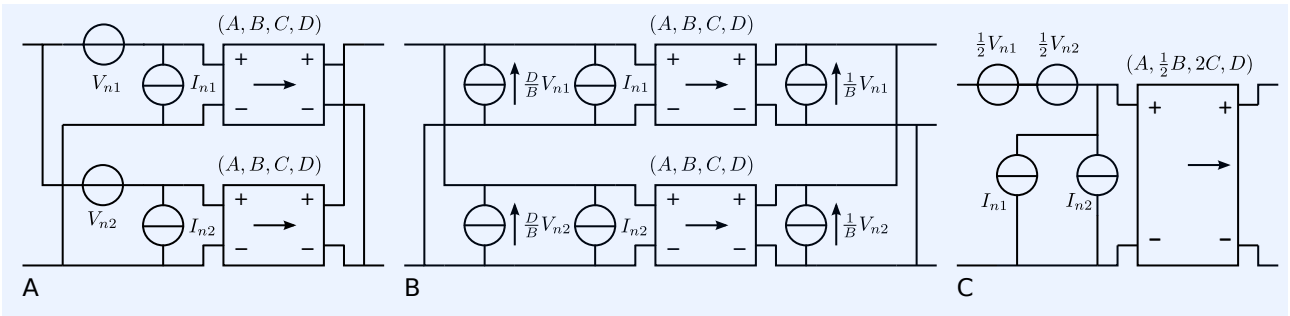


Figure 6.38: Equivalent input noise sources of (complementary-) parallel connected two-ports.

The power spectral densities of the equivalent noise sources of the anti-series connection are thus obtained as

$$S_{vtot} = \frac{1}{4}S_{vn1} + \frac{1}{4}S_{vn2}\ [\mathrm{V^2/Hz}], \tag{6.58}$$

$$S_{itot} = S_{in1} + S_{in2}\ [\mathrm{A^2/Hz}]. \tag{6.59}$$

If both constituting two-ports have identical noise behavior, we have: $S_{vn1} = S_{vn2} = S_v$ [V²/Hz] and $S_{in1} = S_{1n2} = S_i$ [A²/Hz] and we obtain

$$S_{vtot} = \frac{1}{2}S_v\ [\mathrm{V^2/Hz}], \tag{6.60}$$

$$S_{itot} = 2S_i\ [\mathrm{A^2/Hz}]. \tag{6.61}$$

6.4.4 Complementary-parallel CS stage

In this section, we will apply the above theory about balancing to the biased complementary-parallel CS stage. Figure 6.39 shows the parallel connection of two biased complementary CS stages.

Similar as with the complementary-parallel connection of two-terminal elements, the bias currents are converted into common-mode currents and the bias voltage sources can be converted into common-mode voltage sources.

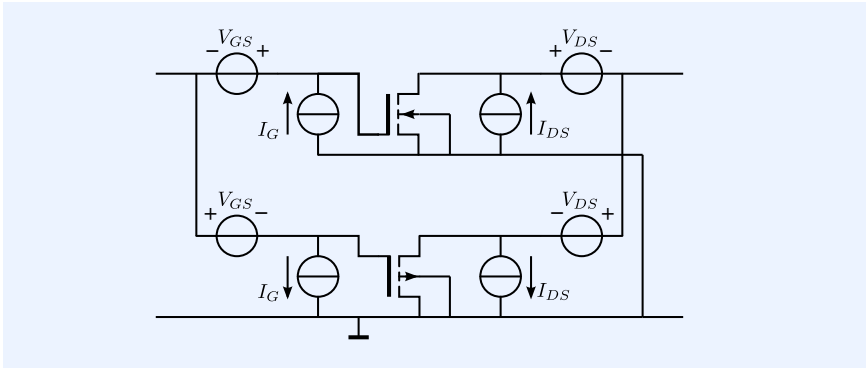


Figure 6.39: Parallel connection of biased complementary MOS CS stages.

The conversion of the bias voltage sources into common-mode sources is the result of shifting these sources through de MOS transistors. This effectively adds $-V_{DS}$ to the nodal voltages of the NMOS transistor and $+V_{DS}$ to those of the PMOS transistor. Figure 6.40 shows the result of this transformation. The output port bias voltages have been converted into power supply voltages. This circuit is known as the *CMOS inverter stage*.

The drain current and the gate current have become common-mode currents. They cannot be observed at the input and the output of the stage. They are determined by the voltage sources and the characteristics of the MOS transistors. This makes this stage very sensitive to variations in the power supply voltages and the temperature changes. In practice, measures will have to be taken to reduce these sensitivities.

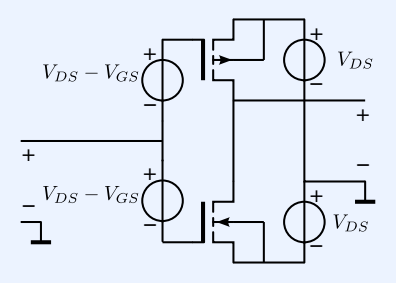


Figure 6.40: Parallel connection of biased CMOS CS stages. The output bias voltage V_{DS} has been shifted through the ground node.

Large-signal static behavior

Similar as with the anti-series stage, we will now study the large-signal dynamic behavior of the push-pull stage.

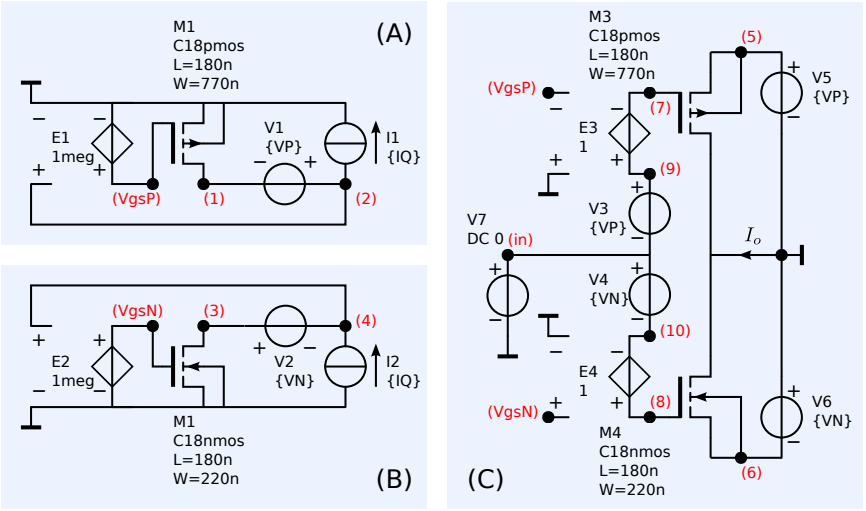


Figure 6.41: SPICE test bench for determination of the static voltage to current transfer of the complementary-parallel CS stage.
A: Circuit for determination of the bias voltage V_{GS} of the PMOS
B: Circuit for determination of the bias voltage V_{GS} of the NMOS
C: Biased complementary-parallel CS stage.

Figure 6.41 shows a simulation test bench for this purpose. Since in general we want to fix the operating voltage and current at the output port of an amplifier stage, we need to determine the input bias voltage V_{GS} for fixing I_{DS} to its desired value at a given value of V_{DS}. The width of the PMOS has been taken 3.5 times larger than that of the NMOS because the zero field mobility of electrons is larger than the mobility of holes. Relative scaling of the width of both devices can be used to minimize even-order distortion, or to

adjust for different values of the current source and current sink capabilities of the complementary-parallel stage.

Figure 6.41A shows the circuit for determination of V_{GS} of the NMOS and Figure 6.41B shows the circuit for determination of V_{GS} of the PMOS transistor. In practice, biasing techniques need to be applied that sufficiently accurate fix the quiescent operating current I_{DS} to its desired value. Such techniques will be discussed at a later stage.

Figure 6.41C shows the biased push-pull stage with its output shorted to ground. The output current is the sum of the drain currents and the common-mode current is half the difference between the two currents, similar as defined with the complementary-parallel connection of two-terminal devices.

The LTspice netlist of this circuit is shown below:

```
complParlCS
* File: complParlCS.cir
* LTspice netlist file
.lib CMOS18TT.lib
* Circuit for determination of VgsP
M1 1 VgsP 0 0 C18pmos L=180n W=770n
E1 VgsP 0 2 0 1meg
V1 2 1 {VP}
I1 2 0 {IQ}
* Circuit for determination of VgsN
M2 3 VgsN 0 0 C18nmos L=180n W=220n
V2 3 4 {VN}
I2 0 4 {IQ}
E2 VgsN 0 4 0 1meg
* Biased complementary parallel stage
V5 5  0  {VP}
V6 0  6  {VN}
V7 in 0  DC 0
V3 9  in {VP}
V4 in 10 {VN}
E3 9  7  0    VgsP 1
E4 8  10 VgsN 0    1
M3 0 7 5 5 C18pmos L=180n W=770n
M4 0 8 6 6 C18nmos L=180n W=220n
.param VP=0.9 VN=0.9 IQ=1u
*
* LTspice command section
*
.DC V7 -900m 900m 36m
.step param IQ list 10n 100n 200n 500n 1u 2u 5u 10u
.end
```

Figure 6.42 shows the differential-mode output current and the common-mode current both as a function of the input voltage. At zero input signal, the common-mode current equals the quiescent operating current. The complementary-parallel stage is capable of delivering a differential output current much larger than its quiescent current.

If the device is biased in weak inversion, the voltage to current transfer can be approximated by

$$\frac{I_o}{I_Q} = 2\sinh\frac{V_i}{nU_T}. \tag{6.62}$$

Hence, when biased in weak inversion, the gain around the operating point is very low and strongly varies with the input voltage. This phenomena is usually referred to as *cross-over distortion*. When biased in strong inversion, the gain around the operating point is larger than the gain at relatively large excursions because both transistors contribute to the output current, while at relatively large excursions either the PMOS or the NMOS delivers the output current.

Push-pull operating modes

Figure 6.42 illustrates that complementary-parallel stages can deliver currents that exceed their operating current. These stages are often referred to

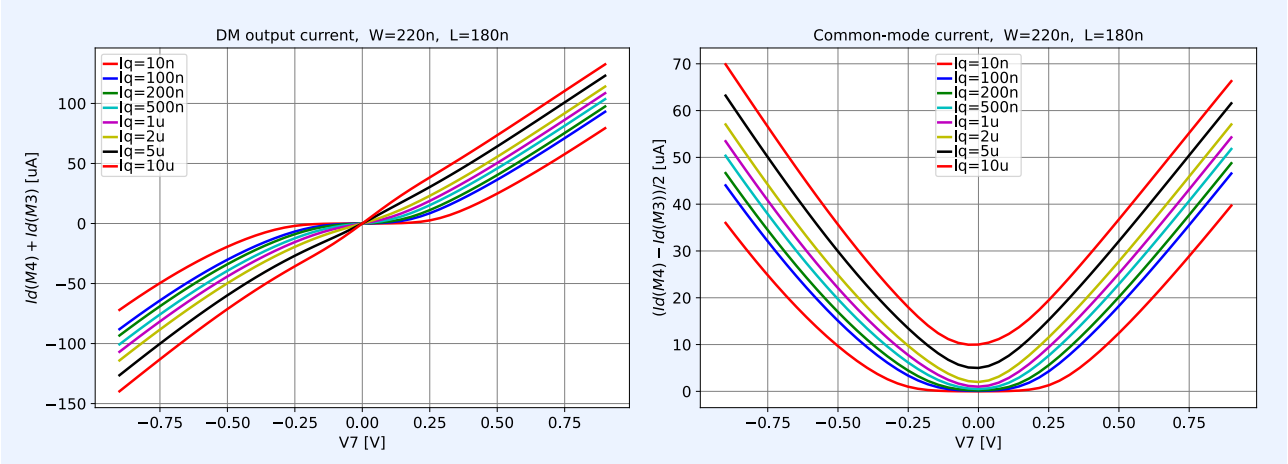


Figure 6.42: Characteristics of the complementary-parallel CS stage, when operating from weak inversion to strong inversion with velocity saturation

Left: Output current versus differential input voltage

Right: Common-mode current versus differential input voltage,

as push-pull stages, a name that indicates that one transistor is *pushing* or *sourcing* a current into a load, while the other is *pulling* or *sinking* the current from a load. Due to this property, these stages are often used as high-efficient output stages of amplifiers. Complementary-parallel stages can be biased in different operation modes or classes:

1. Class A operation: both transistors conduct for all positive and negative excursions of the output current. This is the case if the output current is smaller than the quiescent current.

2. Class B operation: for all values of the output current, either the push transistor or the pull transistor conducts. This is the case if the quiescent current equals zero.

3. Class AB operation: for small values of the output current both transistors conduct, but at high output currents either the push or the pull transistor conducts. This is the case if there is a non-zero quiescent current which is much smaller that the peak output current.

4. Class C operation: either the push or the pull transistor conducts, but there is a dead-zone in the transfer from input to output.

High-efficient, low-distortion operation, push-pull stages are often biased for class AB operation: a small bias current facilitates a smooth transition between the push and pull operation. Biasing of complementary-parallel stages will be discussed in Chapter 15.

Small-signal dynamic behavior

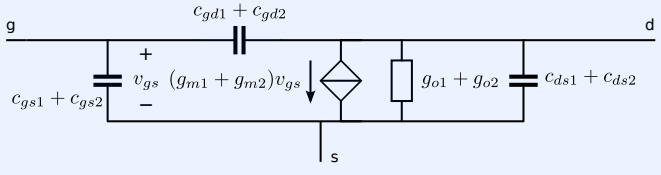


Figure 6.43: Simplified small-signal equivalent circuits of the complementary-parallel CS stage.

Figure 6.43 shows the simplified small-signal equivalent circuit of the complementary-parallel CS stage.

Stationary noise behavior

The stationary noise behavior in the quiescent operating point has been described with equations (6.60) and (6.61). From these equations we can draw the following important design conclusion:

For a complementary CS stage to achieve the same noise performance at a given source impedance as a single CS stage, the quiescent operating current as well as the width of each transistor of the anti-series connection should be half those of the single CS stage.

6.5 Conclusions

In this chapter we have studied the application of balancing techniques. We have introduced balancing as a form of additive compensation. This error reduction technique is capable of reducing reproducible errors due to even order nonlinearity, including its temperature dependency.

The two balancing techniques: anti-series connection and complementary-parallel connection are of special interest for amplifier design. Anti-series connection of the basic three-terminal amplifier stages yields a four terminal stage with improved port isolation. Complementary-parallel connection provides a push-pull stage that exhibits a high current drive capability at a low quiescent current. The noise performance of the balanced stage is similar to that of its unbalanced version. The same holds for the dynamic performance. Hence, their application for the improvement of the linearity, the reduction of the offset and temperature drift, the improvement of the port isolation and the extended current drive capability is only at the cost of circuit complexity.

7 Design of feedback amplifier configurations

7.1 Introduction

In Chapter 2, we have showed that all amplifiers can be characterized by their input impedance, their output impedance and their transfer characteristic. We determined nine different unilateral amplifier types and modeled their ideal behavior with the transmission-1 two port parameters. We found five different versions for each of those types if we also consider whether the ports are floating or if they have one their terminals connected to the ground.

7.1.1 Design tasks

In this chapter, we will discuss means to design the port impedances and the source-to-load transfer. We will do this by designing the transmission-1 two-port parameters of the amplifier.

We have seen that the port impedances should provide optimum sense and drive conditions for the source and the load, respectively. Sensing the open circuit voltage of the signal source, as well as driving the load with a current, requires an infinite port impedance. In order to sense the short-circuit current of the signal source, as well as drive the load with a voltage, the port impedance should be zero. For accurate termination of transmission lines, the port impedance should equal the characteristic impedance of the transmission line.

Let us, for example, consider a voltage amplifier that is driven from a voltage source with an open-circuit voltage V_s and a source impedance Z_s. The amplifier should provide a voltage V_ℓ across the load impedance Z_ℓ. This voltage should be an accurately amplified copy of the source voltage. The input impedance of the amplifier is Z_i and the output impedance of the amplifier equals Z_o, as shown in Figure 7.1.

The source-to-load voltage transfer for this configuration can be found as:

$$\frac{V_\ell}{V_s} = \frac{Z_i}{Z_i + Z_s} A_v \frac{Z_\ell}{Z_\ell + Z_o}, \tag{7.1}$$

where A_v is the voltage amplification factor of the amplifier: $A_v = \frac{V_o}{V_i}$.

Accurate information transfer from source-to-load is only possible if:

1. All impedances are accurately known
2. In cases in which Z_s and/or Z_i are not accurately known, the input impedance of the amplifier should be much larger than the source impedance.
3. In cases in which Z_ℓ and/or Z_o are inaccurately known, the output impedance of the amplifier should be much smaller than the load impedance.

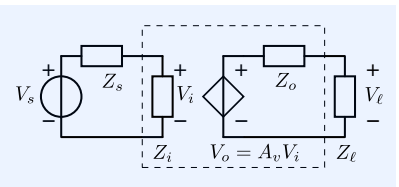


Figure 7.1: Simple model of a voltage amplifier with source and load.

7.1.2 Brute force port impedance design

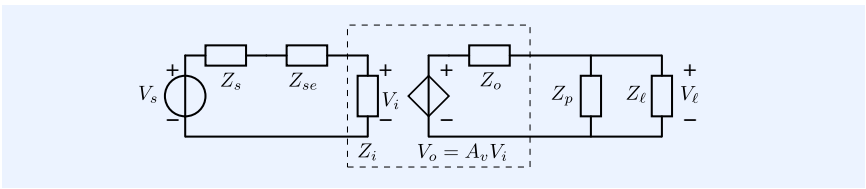


Figure 7.2: Voltage amplifier with brute force correction of its input and output impedance.

If the optimum drive or termination conditions are not met, accurate and linear impedances can be inserted into the signal path to improve the source sense and the load drive conditions. In the case of the voltage amplifier from Figure 7.1, this can be done by inserting a linear and accurately known impedance Z_{se} in series with Z_s and a linear and accurately known impedance Z_p in parallel with Z_ℓ, as shown in Figure 7.2.

Techniques by which accurate and linear port impedances are realized through their simple insertion in series or in parallel with the signal path, are called *brute force* techniques. In general, these techniques should be avoided, because they result in a deterioration of the signal-to-noise ratio and the power efficiency of the amplifier.

Example 7.1

Let us consider an amplifier with equivalent input noise sources V_{eq} and I_{eq}. The amplifier is driven from a voltage source with an open source voltage V_s and an impedance Z_s. The noise voltage associated with the source impedance is V_{ns}. The spectral density of this voltage equals $4kT\,\mathrm{Re}(Z_s)$ [V²/Hz]. An impedance Z_{se} has been placed in series with the signal path and an admittance Y_p has been placed in parallel with the signal path. The spectral density of the voltage noise V_{nY} associated with Z_{se} equals $4kT\,\mathrm{Re}(Z_{se})$ [V²/Hz]. The spectral density of the current noise I_{nY} associated with Y_p equals $4kT\,\mathrm{Re}(Y_p)$ [A²/Hz]. The complete configuration is shown in Figure 7.3.

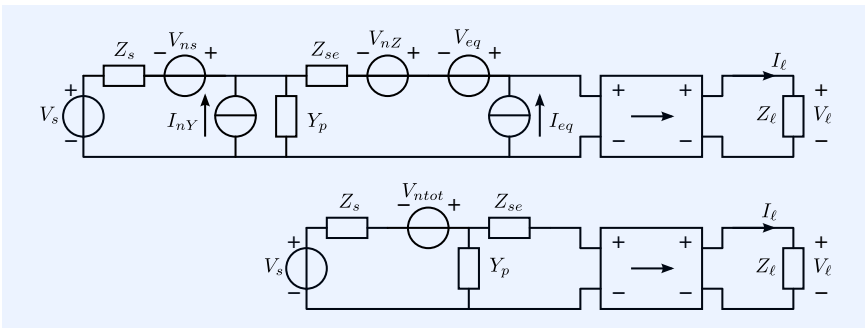


Figure 7.3: Impedances in parallel or in series with the signal path generally degrade the signal-to-noise ratio.

For the evaluation of the influence of the noise sources on the signal-to-noise ratio, the noise sources need to be transformed into one total equivalent noise voltage V_{ntot} in series with the signal voltage source. The spectral density $S_{V_{ntot}}$ of this total equivalent input noise voltage for the circuit from Figure 7.3, can be obtained as:

$$\begin{aligned} S_{V_{ntot}} = {} & 4kT\left(\mathrm{Re}\left(Z_s\right) + \mathrm{Re}\left(Z_{se}\right)\left|1 + Z_s Y_p\right|^2 + \mathrm{Re}\left(Y_p\right)\left|Z_s\right|^2\right) \\ & + S_{V_{eq}}\left|1 + Z_s Y_p\right|^2 \\ & + S_{I_{eq}}\left|Z_{se} + Z_s\left(1 + Z_{se} Y_p\right)\right|^2 \end{aligned} \tag{7.2}$$

This expression shows an enlarged contribution of both equivalent noise sources to the total equivalent input noise if Z_{se} and/or Y_p differ from zero. Only in narrow-band applications can the noise be improved by inserting impedances in series or in parallel with the signal path. This occurs if $Z_{se} = -\mathrm{Im}(Z_s)$. In cases in which current is the information-carrying quantity, this condition would be: $Y_p = -\mathrm{Im}(Y_s)$. Such improvement of the signal-to-noise ratio, however, only occurs in the vicinity of some (resonance) frequency at which these conditions are approximately met.

From the above analysis, we may draw the following conclusion:

Impedances in series or in parallel with the signal source will deteriorate the signal-to-noise ratio of the amplifier. This deterioration has two causes:

1. In general, the contribution of the amplifier's equivalent input noise sources to the total input noise will increase due to the insertion of these impedances as has been shown in the above example.

2. If the real part of these impedances differs from zero, the thermal noise associated with it, decreases the signal-to-noise ratio.

Impedances in series or in parallel with the load will deteriorate the power efficiency of the amplifier. This deterioration also has two causes:

1. Impedances in series or in parallel with the load require the amplifier's output current or voltage excursions to be larger for the same amount of load power. This results in an increase of the power losses in the amplifier. Only in narrow-band applications, can improvement be achieved by tuning out unwanted effects from reactive parts of impedances in series or in parallel with the load, with impedances that have opposite reactive parts.
2. If the impedances in series or in parallel with the load have a real part, extra power has to be delivered by the amplifier.

7.1.3 Negative feedback amplifiers

The first negative feedback amplifier was built in 1927 by H. Black.[Black1934][1] He applied negative feedback to obtain linear and stable-gain repeater amplifiers for long-distance telephone systems. Black's patent was awarded in 1937.[Black1932][2]

[1] Harold S. Black. Stabilized feed-back amplifiers. *Electrical Engineering*, 53(1):114–120, January 1934

[2] H.S. Black. Wave Translation System, March 1932

Negative feedback is a powerful error reduction technique that trades gain for quality improvement. With negative feedback, the source-to-load transfer of an amplifier, as well as the port impedances, can primarily be fixed with passive devices, while biased amplifying devices are used to provide power gain. Such an approach is useful because the characteristics of real world passive devices closely match those of their corresponding network abstraction. In other words, real world resistors, capacitors and inductors show approximate resistive, capacitive and inductive behavior over a wide operating range, respectively. This cannot be said of the biased amplifying devices that provide the power gain. Their characteristics are generally nonlinear, suffer from speed limitation, show relatively large fabrication tolerances and are rather sensitive to temperature variations.

We will show that the transmission-1 parameters of amplifiers can accurately be fixed with feedback elements around a high-gain amplifier or *controller*. Each nonzero parameter can be fixed to an accurate value with one feedback loop. In this way, we are able to give the port impedances and the source-to-load transfer their desired values, without significant degradation of other performance aspects, such as the noise performance, the accuracy, the linearity, the power efficiency and the dynamic response of the amplifier. As a matter of fact, with negative feedback, we are able to design those performance aspects almost independently.[3]

[3] This is usualy called *orthogonal design*.

7.1.4 This chapter

In section 7.2, we will present a design procedure for fixing the transmission-1 parameters through application of negative feedback. We will introduce two different feedback techniques: *direct feedback* and *indirect* or *model-based feedback*.

In section 7.3, we will discuss various implementation techniques. The best performance can be obtained with direct *nonenergic* feedback. Nonenergic feedback amplifiers use nonenergic feedback elements.[4] The practical use of nonenergic feedback is restricted due to the limited availability of those elements. Other techniques that will be discussed are: passive feedback, active feedback and balanced feedback.

The following sections will be devoted to a more elaborate treatment of all the techniques introduced. Section 7.4 discusses the design and the performance aspects of nonenergic negative feedback amplifiers. Conceptually, the

[4] A circuit element is nonenergic if the instantaneous power delivered to it is always zero. Hence, nonenergic network elements do not store energy and do not dissipate power. Nonenergic network elements are:

- Ideal transformers
- Ideal gyrators
- Short circuit
- Open circuit.

best performance can be achieved with this type of amplifiers. Section 7.5 is devoted to passive feedback. With passive feedback amplifiers, the feedback elements adversely affect the signal-to-noise ratio and the power efficiency. These effects, however, can be kept small, when compared to brute-force techniques. In amplifiers with passive feedback networks, the sign of the source-to-load transfer as well as the source-load isolation, is indissolubly connected to the type of transfer. Active feedback, balanced feedback and indirect feedback can be used to design the sign of the transfer, as well as the source-to-load isolation independent from the type of transfer. These techniques will be discussed in sections 7.6, 7.7 and 7.8, respectively.

7.2 Design of feedback configurations

The procedure for fixing a tranmission-1 parameter with the aid of negative feedback is as follows:

1. Sense the load quantity that needs to be related to a source quantity.

2. Multiply the sense result with the value of the transmission-1 matrix coefficient that needs to be fixed; this should yield a copy of the source quantity.

3. Nullify the error between the source quantity and the derived copy of it with the aid of a high-gain *error amplifier* or *controller*.

   This amplifier controls the output signal in such a way that the feedback signal is the best possible copy of the source signal.[5] At this stage of the design of negative feedback amplifiers, we will use nullors as error amplifiers. Nullors are two-ports that have a nullator as the input port and a norator as the output port. They have infinite available power gain. Nullors are introduced in Chapter 18.3.3.

4. Implement the nullor with a practical error amplifier, define its performance requirements and design or select it. Chapter 11, discusses in which way, and to what extend the amplifier's performance depends on properties of its controller.

[5] This is achieved with *corrective* or *negative feedback*.

7.2.1 Step-by-step design of the feedback configuration

The step-by-step design of a negative feedback voltage amplifier will be illustrated in the following example. The design procedure of other types of negative feedback amplifiers is similar to the one presented.

Example 7.2

Let us consider that we want to design a negative feedback voltage amplifier with a voltage gain A_v. Such an amplifier should have an infinite input impedance and zero output impedance. The values of the transmission-1 parameters of this amplifier should be:

$$A = \frac{1}{A_v},\ B = 0,\ C = 0,\ D = 0. \tag{7.3}$$

1. *Figure 7.4A shows the situation at the start: we have a signal voltage source with open-circuit voltage V_s and source impedance Z_s and a load impedance Z_ℓ. With the above values of the transmission-1 matrix parameters, the voltage across Z_ℓ will equal $A_v V_s$ and will not depend on Z_s and Z_ℓ.*

2. *We need to sense the load voltage V_ℓ (see Figure 7.4B).*

3. *Then, we create a copy of the source voltage by multiplying the sensed voltage by* $A = \frac{1}{A_v}$. *To this end, we may use a voltage-controlled voltage source as shown in Figure 7.4C. The gain of this so-called* feedback element *should equal the reciprocal value of the desired source-to-load transfer* A_v.

4. *We then determine the difference between the source voltage and this copy by connecting these two voltages anti-series: see Figure 7.4D.*

5. *Finally, we nullify this difference between the source voltage and the copy by placing a nullator between the anti-series connected voltage sources, and by controlling the load voltage through placement of a norator in parallel with the load (see Figure 7.4E).*

6. *The nullor is a network abstraction that cannot be realized. After we have designed the negative feedback configuration with the aid of a nullor, we need to design the error amplifier that replaces it. This error amplifier must have a large gain so that the error signal will be negligibly small and the signal at the load is a sufficiently accurate copy of the source signal.*

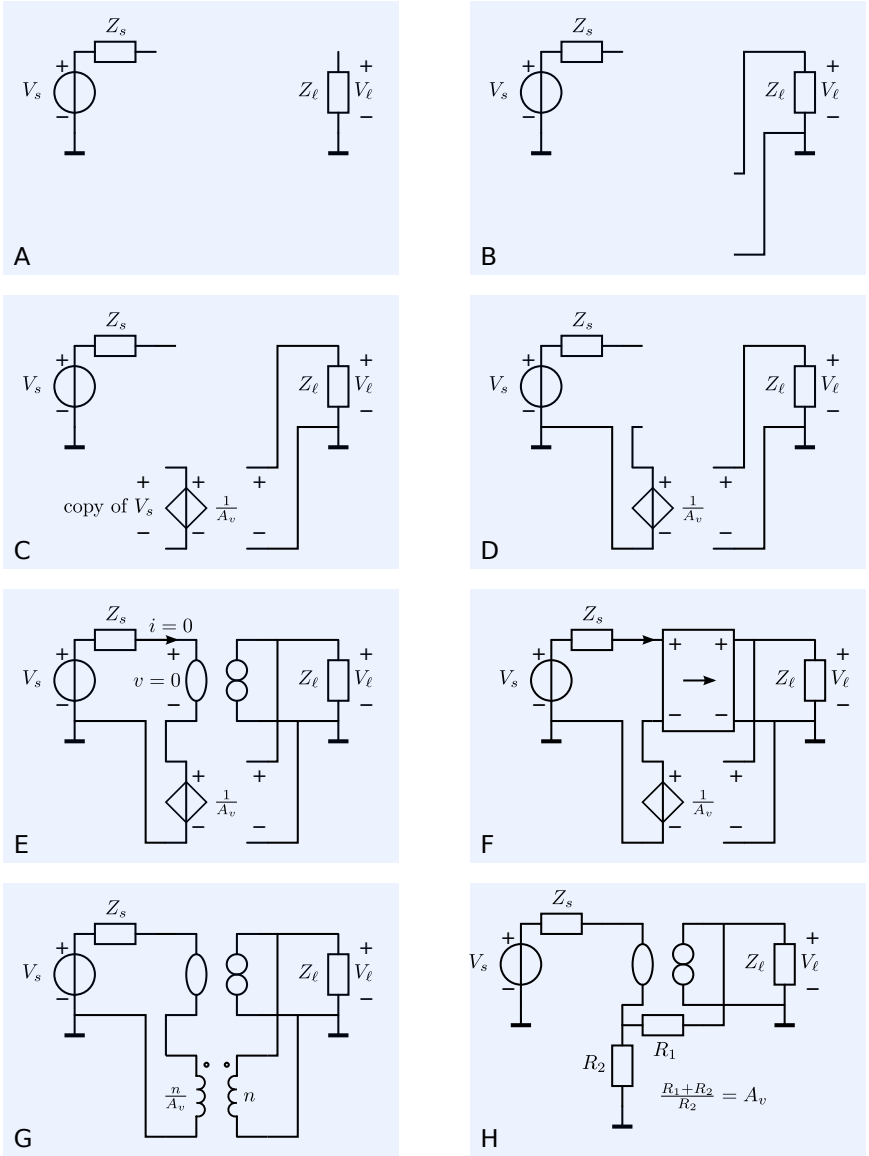


Figure 7.4: A until F: step-by-step development of the circuit concept for a negative feedback voltage amplifier with gain $\frac{V_\ell}{V_s} = A_v$.

A: The voltage V_s of a grounded source must evoke a voltage V_ℓ across the grounded load. Assume the voltage at the load equals its desired value V_ℓ.

B: The load voltage is sensed. Sensing of a voltage requires the sense element to be placed in parallel with the voltage to be sensed.

C: The sensed voltage is converted into a copy of the source voltage. Hereto, it is multiplied by the reciprocal value of the desired source-to-load transfer.

D: The difference between the source voltage and its copy is obtained through anti-series connection (series connection with opposite polartities).

E: This difference should be zero. The nullator sets the zero voltage and zero current condition for this *error signal*: the difference between the source signal and its intended copy. The norator drives the load and the feedback network in such a way that the condition set by the nullator will be satisfied. In this way, the load voltage V_ℓ is generated.

F: The nullor is replaced with a two-port with a large available power gain, in such a way that negative feedback is obtained.

G: The controlled source in the feedback path is replaced with a nonenergic voltage attenuator (a transformer).

H: The controlled source in the feedback path is replaced with a passive voltage attenuator.

Figure 7.4F shows the circuit after the nullor has been replaced with an error amplifier. The ports of the error amplifier are connected in such a way, that a

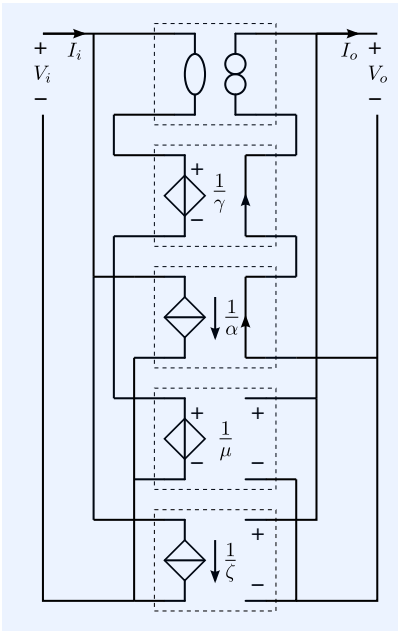


Figure 7.5: Quadruple loop negative feedback amplifier with controlled sources as feedback elements.

unintentional decrease of the load voltage results in an increase of the error signal that establishes a correction of the load signal. This technique is called corrective, degenerative or negative feedback. *If the opposite is the case, we have positive feedback, which may result in unstable behavior.*

In Figure 7.4C we used a voltage-controlled voltage source to generate a copy of the load voltage. This voltage-controlled voltage source is an active network element. We applied it to elucidate the synthesis of this negative feedback voltage amplifier. Amplifiers with active feedback elements are called *active feedback amplifiers*. Figure 7.4G shows the amplifier with a transformer as the feedback element. An ideal transformer is a nonenergic network element, and therefore this amplifier is called a *nonenergic feedback amplifier*. Figure 7.4H shows an arrangement in which the feedback network consists of a passive voltage divider. This technique is called *passive feedback*.

As already mentioned, the four transmission parameters of an amplifier can be independently fixed using negative feedback. Here is the procedure:

1. In order to fix the parameters A and C to a nonzero value, we need to sense the load voltage. Sensing of voltages means parallel connection of the sense network and the load impedance. This is called *load voltage sensing*, *parallel sensing*, *output voltage feedback*, *output parallel feedback* or *output shunt feedback*

2. In order to fix the parameters B and D to a nonzero value, we need to sense the load current. Sensing of a current means series connection of the sense network and the load impedance. This is called *load current sensing* or *series sensing*, *output current feedback* or *output series feedback*.

3. In order to fix the parameters A and B to a nonzero value, we need to compare the feedback voltage with the source voltage. Subtracting or comparison of voltages requires anti-series connection of the feedback voltage and the source voltage. This is called *source voltage comparison*, *series comparison*, *input voltage feedback* or *input series feedback*.

4. In order to fix the parameters C and D to a nonzero value, we need to compare the feedback current with the source current. Subtracting or comparing currents requires anti-parallel connection of the feedback current and the source current. This is called *source current comparison*, *parallel comparison*, *input current feedback*, *input parallel feedback* or *input shunt feedback*.

Figure 7.5 shows a feedback amplifier in which all four transmission-1 parameters have been independently fixed through application of negative or *corrective* feedback. The amplifier has its input port and its output port both floating with respect to the ground. Each feedback loop fixes one parameter only. This orthogonal design of the transmission-1 parameters is possible due to the ideal sense and drive conditions of the feedback elements:

- No voltage drop across current-sensing elements
- No current through voltage-sensing elements
- Zero output impedance of the feedback voltage sources
- Infinite output impedance of the feedback current sources

With passive feedback and nonenergic feedback amplifiers elements, this is generally not the case. This will be discussed later.

Table 7.1 gives an overview of the way the transmission parameters of the amplifiers can be fixed with the aid of negative feedback.

Table 7.2 gives an overview of the 16 amplifier types (including the nullor) that can be realized by leaving out one or more feedback elements from the circuit from Figure 7.5.

Param.	Feedback type at input port	Feedback type at output port	Transfer
A	V - comparison (series feedback)	V - sensing (parallel feedback)	[V/V]
B	V - comparison (series feedback)	I - sensing (series feedback)	[I/V]
C	I - comparison (parallel feedback)	V - sensing (parallel feedback)	[V/I]
D	I - comparison (parallel feedback)	I - sensing (series feedback)	[I/I]

Table 7.1: Feedback method for fixing the transmission-1 matrix parameters A, B, C or D.

Amplifier type	A	B	C	D
Nullor	0	0	0	0
Voltage amplifier (unilateral)	$\frac{1}{\mu}$	0	0	0
Transadmittance amplifier (unilateral)	0	$\frac{1}{\gamma}$	0	0
Transimpedance amplifier (unilateral)	0	0	$\frac{1}{\zeta}$	0
Current amplifier (unilateral)	0	0	0	$\frac{1}{\alpha}$
$Z_i = \frac{\alpha}{\gamma}$, $Z_o = \infty$ (unilateral)	0	$\frac{1}{\gamma}$	0	$\frac{1}{\alpha}$
$Z_i = \frac{\zeta}{\mu}$, $Z_o = 0$ (unilateral)	$\frac{1}{\mu}$	0	$\frac{1}{\zeta}$	0
$Z_i = 0$, $Z_o = \frac{\zeta}{\alpha}$ (unilateral)	0	0	$\frac{1}{\zeta}$	$\frac{1}{\alpha}$
$Z_i = \infty$, $Z_o = \frac{\mu}{\gamma}$ (unilateral)	$\frac{1}{\mu}$	$\frac{1}{\gamma}$	0	0
Transformer-like amplifier (non-unilateral)	$\frac{1}{\mu}$	0	0	$\frac{1}{\alpha}$
Gyrator-like amplifier (non-unilateral)	0	$\frac{1}{\gamma}$	$\frac{1}{\zeta}$	0
Triple loop 1 (non-unilateral)	$\frac{1}{\mu}$	$\frac{1}{\gamma}$	$\frac{1}{\zeta}$	0
Triple loop 2 (non-unilateral)	$\frac{1}{\mu}$	$\frac{1}{\gamma}$	0	$\frac{1}{\alpha}$
Triple loop 3 (non-unilateral)	$\frac{1}{\mu}$	0	$\frac{1}{\zeta}$	$\frac{1}{\alpha}$
Triple loop 4 (non-unilateral)	0	$\frac{1}{\gamma}$	$\frac{1}{\zeta}$	$\frac{1}{\alpha}$
Quadruple loop	$\frac{1}{\mu}$	$\frac{1}{\gamma}$	$\frac{1}{\zeta}$	$\frac{1}{\alpha}$

Table 7.2: Single-loop and multiple feedback configurations with controlled sources as feedback elements according to Figure 7.5. Removal of a feedback loop, while maintaining the drive and sense conditions for the remaining, makes the corresponding transmission coefficient zero.

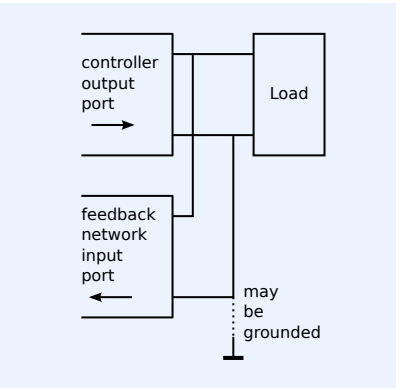


Figure 7.6: Direct voltage sensing: parallel sensing or load voltage sensing.

7.2.2 Direct sensing and comparison techniques

Until now we discussed the conceptual design of so-called *direct feedback amplifiers*. In direct feedback amplifiers, the load quantity is sensed and the source quantity is compared with the feedback quantity. With this technique, high-quality amplifiers can be designed. The opposite to direct feedback is indirect feedback. We will discuss it at a later stage.

We will now discuss direct sensing and comparison techniques in more detail.

Direct voltage sensing techniques

If we need to fix the transmission parameters A or C of an amplifier to nonzero values, we have to sense the voltage across the load. The input voltage of the feedback network will equal the load voltage if the input terminals of the feedback network are connected in parallel with the load. This is shown in Figure 7.6. Hence, if the load must be floating with respect to the ground, the input of the feedback network should also be floating with respect to ground. If one terminal of the load is connected to the ground, the input port of the feedback network will also be grounded.

Direct voltage comparison techniques

If we need to fix the transmission parameter A or B of an amplifier to a nonzero value, we have to subtract the output voltage of the feedback network from the source voltage. Methods for direct voltage comparison are shown in Figure 7.7.

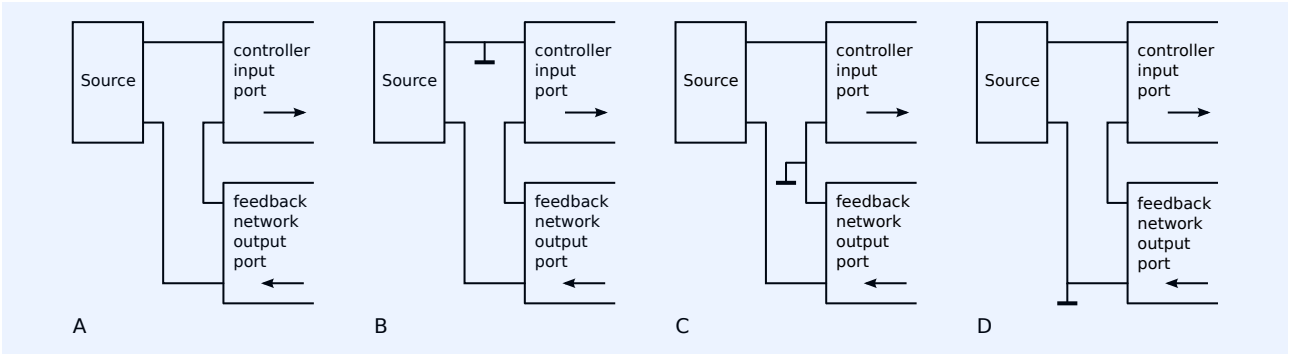


Figure 7.7: Direct voltage comparison, or input voltage feedback.

A: The controller input port, the source and the output port of the feedback network, are floating with respect to the ground.

B: Grounded input port of the controller and grounded load. The output port of the feedback network has to be floating with respect to the ground.

C: The input port of the controller and the output port of the feedback network are grounded. the source has to be floating with respect to the ground,

D: The source and the output port of the feedback network are grounded. The input port of the controller has to be floating with respect to the ground.

The principle of voltage comparison is to connect the source and the feedback network anti-series, with the controller input closing the loop. This is shown in Figure 7.7A.

Ideally, the differential-mode input voltage, the differential-mode input current and the common-mode input current of the controller are zero (controller is a natural two-port with A, B, C and D equal to zero).

Figure 7.7B though 7.7D show different grounding concepts for this loop. Similar as with direct current sensing, parasitic impedances in parallel with the source, or in parallel with the output of the feedback network, may affect the ideal gain of the amplifier (see section 7.3.2). This is of particular interest if the input port of the controller needs to be floating with respect to the ground (see Figure 7.7D).

Direct current sensing techniques

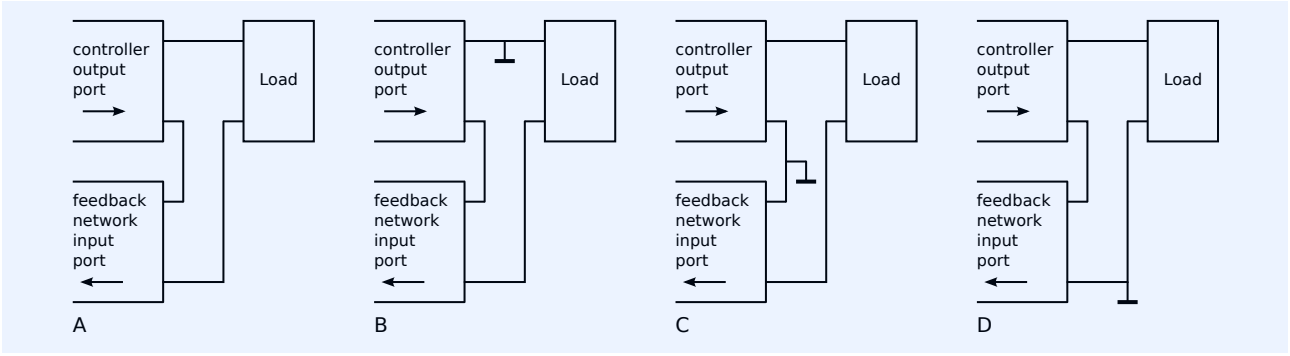


Figure 7.8: Direct current sensing, or output series series feedback.

A: The controller output port, the load and the input port of the feedback network, are floating with respect to the ground.

B: Grounded output port of the controller and grounded load. The input port of the feedback network has to be floating with respect to the ground.

C: The output port of the controller and the input port of the feedback network are grounded. the load has to be floating with respect to the ground,

D: The load and input port of the feedback network are grounded. The output port of the controller has to be floating with respect to the ground.

If we need to fix the transmission parameters B and D of an amplifier to nonzero values, we have to sense the current through the load. The input current of the feedback network equals the load current if the input port of the feedback network is placed in series with the load.

Methods for current sensing are shown in Figure 7.8. The principle of direct sensing of the load current is that the sensing element, the load and the output port of the controller are part of a current loop. This is shown in Figure 7.8A. Figure 7.8B though 7.8D show different grounding concepts for this loop.

If the load and the output port of the controller are grounded, the input port of the sensing element has to be floating with respect to the ground. This is shown in Figure 7.8B. Figure 7.8C illustrates a situation in which the

output port of the controller and the input port of the sensing element are connected to the ground. In such a situation, the load cannot share one of its nodes with the ground. Figure 7.8D shows a situation in which the load and the input port of the sensing element are connected to the ground. In this case, the output port of the controller has to be floating with respect to the ground.

In all cases, parasitic impedances in parallel with the load and in parallel with the input port of the sensing element have to be avoided.

Such impedances create parasitic current paths resulting in a difference between the load current and the input current of the sensing element. A parasitic impedance in parallel with the load results in a transfer that depends on the load impedance, while an impedance in parallel with the input port of the sensing element introduces an error in transfer of the feedback network, resulting in an error of the ideal gain of the feedback amplifier.[6] In practice, this complicates the design of a controller that requires a floating output port (see Figure 7.8D). Parasitic impedances between the output terminals of the controller and the ground should be large enough to keep those errors within acceptable limits.

[6] The definition of the ideal gain of a negative feedback amplifier is given in section 7.3.2.

Direct current comparison techniques

If we need to fix the transmission parameters C or D of an amplifier, we have to subtract the output current of the feedback network from the source current. To do so, the output port of the feedback network needs to be connected anti-parallel with the source. This is shown in Figure 7.9.

If the source must be floating with respect to the ground, the output port of the feedback network should also be floating with respect to the ground. If one terminal of the source is grounded, the output port of the feedback network should also be grounded.

Direct current comparison is dual to direct voltage sensing.

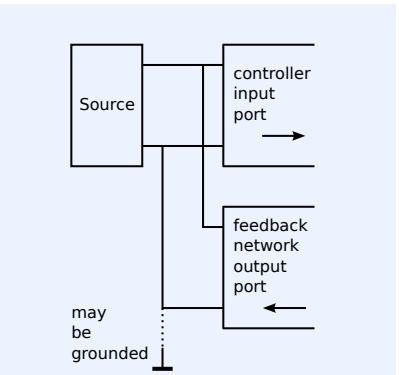


Figure 7.9: Direct current comparison: parallel comparison or source current comparison.

7.2.3 Indirect sensing and comparison techniques

Sometimes it is not possible to sense the load quantity or to compare the feedback quantity with the source quantity directly. In those cases, indirect feedback or *model-based feedback* can be applied. Basic indirect feedback techniques have been illustrated in Figure 7.10. Indirect feedback techniques are often used in IC design. A more detailed discussion of indirect feedback will be presented in section 7.8.

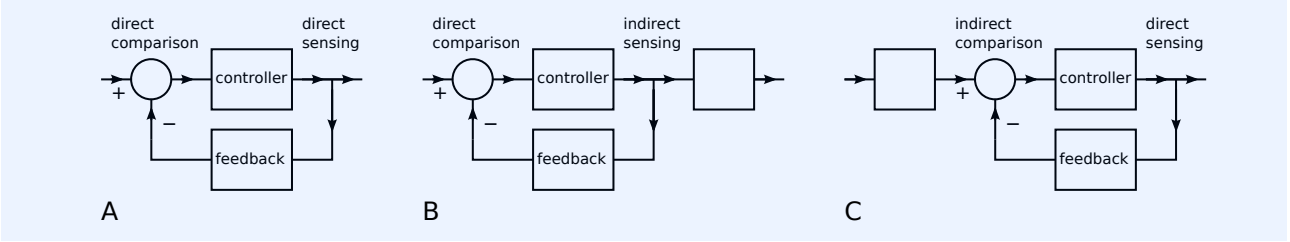


Figure 7.10: Indirect feedback amplifiers use indirect sensing of the load signal and/or indirect comparison with the source signal.

7.3 Implementation of negative feedback

Until now, we have discussed the principle of negative feedback: sensing of the load signal of interest, conversion of the sense result into a copy of the source signal and application of a controller that minimizes the difference between the source signal and its copy derived from the load. There are several ways to implement this. We have already mentioned nonenergic feedback,

passive feedback and active feedback. In this section, we will give a brief description of techniques and terms used. These techniques will be discussed in more detail in the following sections.

7.3.1 Feedback networks

Since the controller is to nullify the error between the input signal and the feedback signal, feedback networks predominantly determine the properties of the feedback amplifier. For example, their noise performance, signal transfer sign, and port isolation strongly depend on the type of feedback network applied. Considering this, it makes sense to classify feedback amplifiers based on their type of feedback network. Below is a list of feedback amplifier classes based on the type of feedback network.

1. Nonenergic feedback amplifiers

   Nonenergic negative feedback amplifiers use feedback elements that solely consist of nonenergic network elements. It will be shown that the noise performance and the power efficiency of these amplifiers is not affected by the feedback network. Theoretically, the best possible performance can be expected from this class of amplifiers. They will be discussed in section 7.4.

2. Passive feedback amplifiers

   Although nonenergic feedback theoretically gives the best possible performance, practical nonenergic feedback elements such as transformers are far from ideal, while gyrators can only be realized with the aid of active circuits. For this reason, the use of passive feedback networks is of more practical importance. Passive feedback amplifiers with either dissipative or non-dissipative feedback elements will be discussed in section 7.5.

3. Active feedback and balanced feedback amplifiers

   Not all amplifier types can be realized using passive feedback. In section 7.5, we will see that the sign of the transfer of a passive feedback amplifier does depend on the type of transfer. Moreover, with passive feedback networks, isolation between the input port and the output port of the amplifier cannot always be accomplished. In cases in which nonenergic feedback or passive feedback techniques are inadequate for the design of the desired amplifier type, other implementations of negative feedback can be used. Aside from taking one's refuge in indirect feedback techniques, one could also consider the application of:

   (a) Active feedback

   Active feedback amplifiers use one or more active networks as a feedback element. They will be discussed in section 7.6.

   (b) Balanced feedback

   Balanced feedback amplifiers can be used for the design of amplifier configurations with their input and/or output port floating with respect to the ground. These amplifier types will be discussed in section 7.7.

7.3.2 Ideal gain of a feedback amplifier

The *ideal gain* of a negative feedback amplifier is defined as its source-to-load transfer when all error amplifiers have been replaced with nullors. During the conceptual design of negative feedback amplifiers, we will always use nullors as controllers. As a matter of fact, we will use nullators and norators. The nullators determine which voltages and the currents need to be nullified and the norators provide the dependent currents to satisfy these conditions.

7.3.3 Negative and positive feedback

In feedback configurations with nullors as controllers, we cannot speak of negative or positive feedback. We just have feedback with an instantaneous error amplifier that provides an infinite available power gain, always yielding zero error signal at the input of a nullator. In fact, the nullor is the *ideal controller concept*, but it cannot be realized in practice.

The available power gain of a real-world error amplifier will not be infinite and will suffer from speed limitation and nonlinearity. Hence, in practice, the *error signal* will not be zero and there only exists a stable situation in case of corrective or negative feedback. Negative feedback requires a negative *loop gain*. A formal definition of the loop gain will be presented in Chapter 10. A negative loop gain requires a negative sign in the transfer from the positive output port quantity of the error amplifier to the positive input port quantity of the error amplifier. Or, in other words, the sign of the product of the gain of the controller and the transfer of the feedback network should be negative. Figure 7.11 shows the quadruple loop negative feedback amplifier of which the nullor-based concept was shown in Figure 7.5. The nullor is replaced with a controller with a finite available power gain and all feedback loops establish negative feedback.

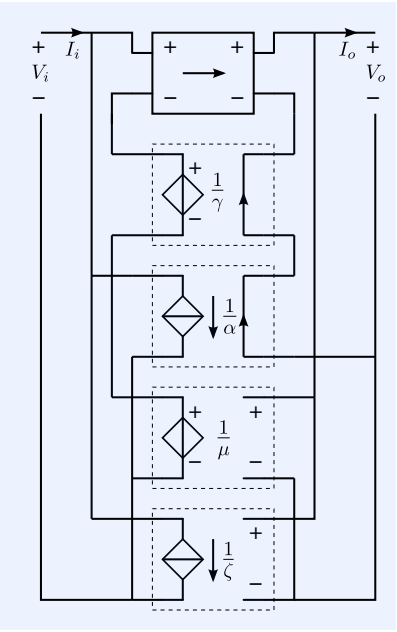


Figure 7.11: Quadruple loop negative feedback amplifier with controlled sources as feedback elements.

Although we do not have negative feedback or positive feedback using nullors, we will use the term 'negative feedback' throughout this chapter. This is because it is our aim to implement corrective feedback. We will speak of positive feedback, if it is our intention to implement non-corrective feedback.

If negative and positive feedback are combined, a stable situation is only obtained if the negative feedback dominates over the positive feedback.

Positive feedback with a loop gain less than unity does the opposite as negative feedback:

1. Parallel sensing of the load quantity or parallel comparison with the source quantity increases the corresponding port impedance.
2. Series sensing of the load quantity or series comparison with the source quantity decreases the corresponding port impedance.
3. Other negative effects of positive feedback will become clear in Chapter 10.

Positive feedback with a loop gain larger than unity results in bistable behavior or in oscillations.

7.4 Nonenergic feedback

Nonenergic negative feedback amplifiers have feedback networks that solely consist of nonenergic network elements. Nonenergic network elements have no energy storage and do not dissipate power; they are instantaneous and noise-free. It will be shown that the noise performance and the power efficiency of nonenergic feedback amplifiers are not affected by their feedback networks as it will be with passive feedback amplifiers.

The following network elements are nonenergic:

1. Short circuit (one port, $V = 0$, for all values of I)
2. Open circuit (one port, $I = 0$, for all values of V)
3. Ideal transformer (two port, $A = \frac{1}{n}$, $B = 0$, $C = 0$, $D = n$)
4. Ideal gyrator (two-port, $A = 0$, $B = \frac{1}{G}$, $C = G$, $D = 0$).

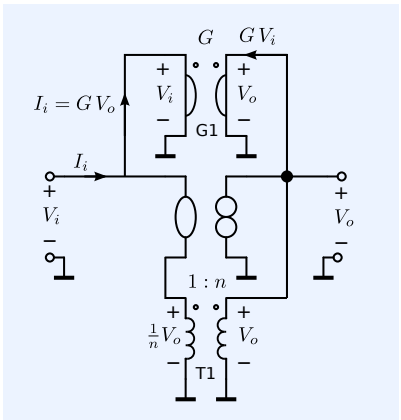


Figure 7.12: Two nonenergic feedback loops around a nullor that fix A and C independently.

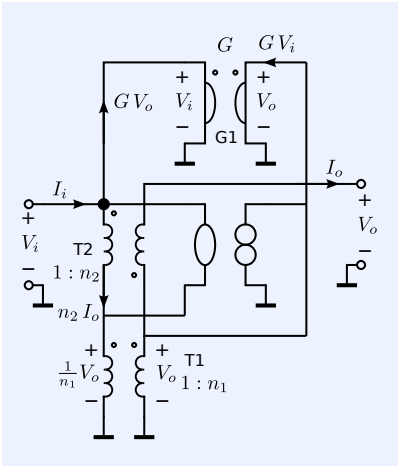


Figure 7.13: Three nonenergic feedback loops around a nullor that fix A , C and D independently.

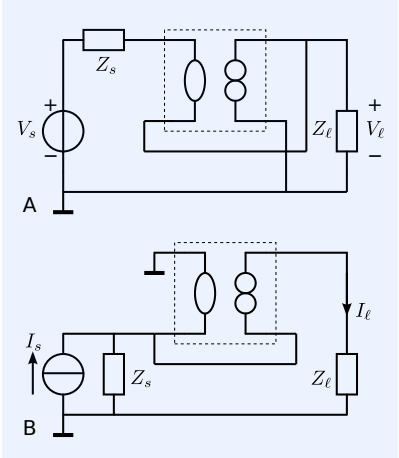


Figure 7.14: (A) Non-energic voltage follower and (B) nonenergic current follower.

[7] Ernst H. Nordholt. *Design of High-Performance Negative Feedback Amplifiers*. Delft Academic Press / VSSD, 1 edition, 1983-2006. ISBN: 9789040712470

7.4.1 Design of nonenergic amplifier configurations

All amplifier configurations with floating or grounded ports, having either inverting or noninverting transfer, can be designed as nonenergic feedback amplifiers, using ideal transformers and gyrators feedback elements. A nonenergic feedback voltage amplifier has already been shown in Figure 7.4G. In the following example, we will demonstrate the design of a dual-loop negative feedback amplifier.

Example 7.3

We will design a negative feedback amplifier that has zero output impedance and its input impedance accurately fixed to a nonzero value. Let us first study the expressions for the input and the output impedance of an amplifier (2.16) and (2.17), respectively. From these expression, we can conclude that we need to fix the transmission-1 parameters A and C to a nonzero value and have $B = D = 0$. The amplifier will then have $Z_i = A/C$ and a voltage gain factor $V_o/V_i = 1/A$. If we want to fix A with the aid of nonenergic negative feedback, we need to sense the output voltage V_o and compare it with the input voltage V_i, thereby using a nonenergic feedback element. A nonenergic voltage-to-voltage converter is a transformer. If we want to fix C with the aid of nonenergic negative feedback, we need to sense V_o, convert it into a current and subtract that current from I_i. A gyrator can be used for this purpose. If we combine both feedback loops, we obtain the circuit depicted in Figure 7.12. It has: $A = 1/n$, $B = 0$, $C = G$, and $D = 0$.

In the next example, we will add a third loop to fix D of this amplifier.

Example 7.4

If we also want to fix D, we need to sense the output current of the amplifier. If we do this without changing the sensing conditions for the voltage-sensing elements, the third loop will not change the values of the previously designed parameters A and C. Figure 7.13 shows a solution in which a second transformer is used to sense the output current and subtract an attenuated copy of it from the input current. The voltage drop across this transformer has been kept zero by placing its port for the input current comparison in parallel with the nullator. Now, the current through the voltage-sensing transformer T1 is not longer zero, but this does not affect the voltage sensing. Hence, the circuit from Figure 7.13 has: $A = 1/n_1$, $B = 0$, $C = G$ and $D = n_2$.

A fourth feedback loop can be added to fix the parameter B as well. This would require a second gyrator that converts the output current into a voltage and subtracts that voltage from the input voltage. If we were to place one port of the gyrator in series with the output of the amplifier and the other one in series with the input, while maintaining zero current through the voltage comparison port of this gyrator, it would fix the amplifier's transmission-1 parameter B without affecting the other parameters. Unfortunately, such an arrangement is not possible. Adding this gyrator would also affect the other transmission-1 parameters of the amplifier. This was demonstrated by Nordholt[Nordholt1983][7].

All 16 amplifier configurations listed in Table 7.2 can be obtained with the aid of negative feedback. One loop fixes one parameter only, if the voltage sensing and current sensing of the feedback elements can be realized independently, as demonstrated in the examples above.

Short circuits and open circuits are also nonenergic elements. The unity-gain voltage amplifier and the unity-gain current amplifier, as shown in Figure 7.14, use short circuits in series with the signal path and open circuits in parallel with the signal path as feedback elements. Hence, these amplifiers belong to the class of nonenergic feedback amplifiers.

7.4.2 Noise behavior of nonenergic feedback amplifiers

Figure 7.15 shows a nonenergic negative feedback voltage amplifier. The controller of the amplifier is modeled as a nullor with equivalent input noise sources V_n and I_n. Transformation of the noise sources from the input port of the nullor to the input port of the voltage amplifier shows that the equivalent input noise sources of the feedback amplifier equal those of the nullor. Hence, the feedback network does not affect the amplifier's noise performance.

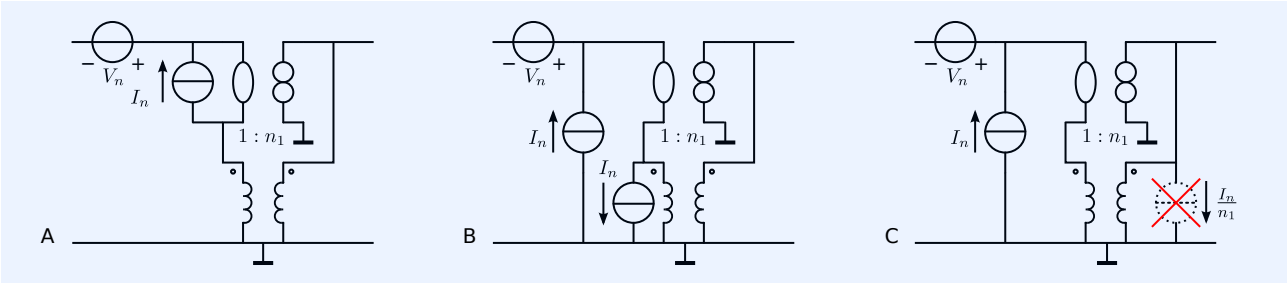


Figure 7.15: Noise behavior of a nonenergic negative feedback voltage amplifier.

This transformation proceeds as follows:

1. The voltage source V_n in series with the input of the nullor is already at the input of the amplifier (see Figure 7.15A).

2. The current source I_n in parallel with the input of the nullor can be redirected over the ground terminal of the amplifier. This is shown in Figure 7.15B. We then obtain a noise current source in parallel with the input of the amplifier and a correlated current source in parallel with the transformer.

3. The source in parallel with the transformer can be replaced with a current source I_n/n_1 in parallel with the other port of the transformer, as shown in Figure 7.15C. This source can be ignored, because it is in parallel with the norator. Replacing it with equivalent input noise sources of the nullor would yield zero values for those sources.

The remaining noise sources are a voltage source V_n in series with the input of the amplifier and a current source I_n in parallel with the input of the amplifier. Those two sources thus equal the original noise sources at the input of the nullor. Hence, the feedback network does not affect the noise performance of the amplifier.

7.4.3 Power efficiency of nonenergic feedback amplifiers

The nonenergic feedback element does not degrade the power efficiency of the amplifier. This can easily be seen for the voltage amplifier from Figure 7.15: its feedback element does not carry any current because one of the ports of the transformer is placed in series with the nullator.

7.5 Passive feedback

In this section, we will discuss the design of amplifier types that have one or more transmission parameters fixed to a nonzero value with the aid of a passive feedback network. We will discuss the design of single loop feedback configurations in section 7.5.1. We will study their influence on the power efficiency and on the noise performance in sections 7.5.2 and 7.5.3, respectively. The design of multiple-loop passive feedback configurations will be treated in section 7.5.4.

7.5.1 Single-loop passive feedback configurations

In the previous sections, we have seen that 15 types of negative feedback amplifiers can be realized using controlled sources as feedback elements. Application of nonenergic feedback elements resulted in the best possible noise performance and power efficiency of a feedback amplifier. In practice, nonenergic feedback cannot always be implemented. Gyrators can only be realized with the aid of active circuits, while transformers are expensive, show far from ideal behavior and cannot always be integrated in silicon technology. Practical application of nonenergic feedback is often restricted to followers in which short circuits and open circuits can be used as feedback elements.

In this section, we will focus on the design of the four basic single-loop configurations using passive feedback elements. These basic configurations are shown in Figure 7.16. They can be designed according to the procedure explained in Figure 7.4. Their ideal gain is the reciprocal transfer of the feedback network. It can easily be evaluated through network inspection, using the zero voltage and the zero current condition for the nullator.

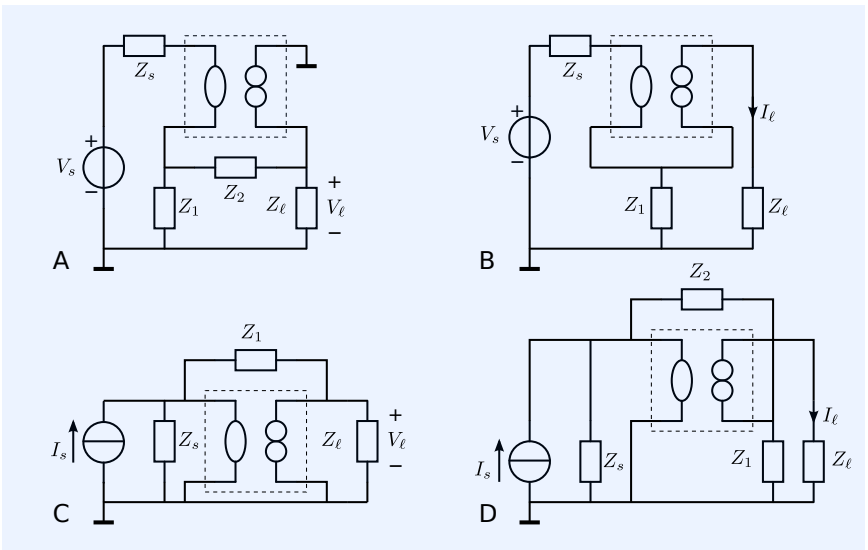


Figure 7.16: Single-loop passive feedback amplifiers for source and load referenced to ground.

A. Voltage amplifier: $\frac{V_\ell}{V_s} = \frac{Z_1+Z_2}{Z_1}$

B. Transadmittance amplifier: $\frac{I_\ell}{V_s} = -\frac{1}{Z_1}$

C. Transimpedance amplifier: $\frac{V_\ell}{I_s} = -Z_1$

D. Current amplifier: $\frac{I_\ell}{I_s} = \frac{Z_1+Z_2}{Z_1}$

Both the voltage and the current amplifier have noninverting transfers, and the transadmittance and transimpedance amplifiers have inverting transfers. This is a consequence of the noninverting character of the transfer of the passive feedback elements. With natural two-ports as feedback elements (such as transformers) we can realize both inverting and noninverting transfers of the feedback network. Using non natural two-ports, we have to find other ways to realize both inverting and noninverting amplifiers. These techniques are:

1. Active feedback
2. Balanced feedback

3. Indirect feedback.

These techniques will be discussed in sections 7.6, 7.7 and 7.8, respectively.

Another restriction due to the application of non-natural two-ports as feedback elements, is their lack of port isolation. This is because passive feedback networks establish interconnections between the input and the output port. With the aid of active feedback, balanced feedback and indirect feedback we will be able to design amplifier configurations with floating ports.

7.5.2 Noise behavior of passive feedback configurations

It is to be expected that the insertion of the passive feedback elements into the signal path causes a deterioration of the signal-to-noise ratio. It can be shown that such deterioration can be kept small when compared to the application of brute force techniques.

The influence of the feedback network on the noise performance of passive feedback amplifiers can be evaluated in several ways. Hand calculations can be done using source transformation techniques. Alternatively, one could use symbolic mathematical tools for this purpose. Network simulation programs (SPICE-like programs) can be used to obtain numeric results. We will demonstrate various methods in the following examples.

First, we will show the application of source transformation techniques. Then, we will show the combination of source transformation techniques and superposition. Finally, we will demonstrate the use of modified nodal analysis. The latter techniques are more suited for automated noise analysis as it has been implemented in SLiCAP.

Example 7.5

In this example, we will illustrate the use of source transformation technique.

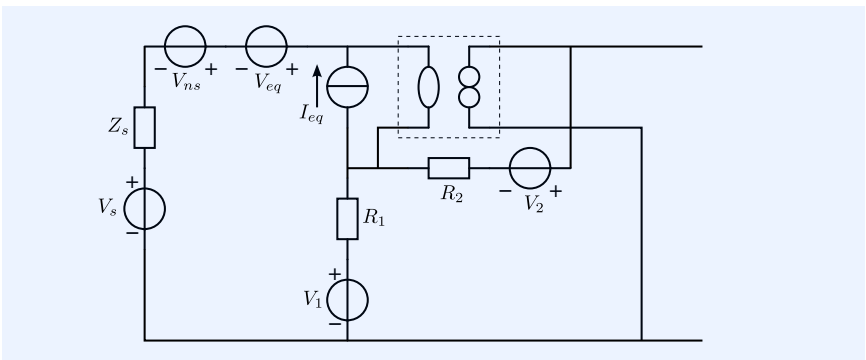


Figure 7.17: Passive-feedback voltage amplifier with noise sources.

Figure 7.17 shows a passive feedback voltage amplifier with its noise sources. In order to evaluate the deterioration of the signal-to-noise ratio due to passive feedback, these noise sources have to be transformed into one total equivalent input noise voltage source. The noise sources that have to be accounted for are:

- *The noise voltage V_{ns} associated with the signal source. The spectral density of this source is given by $4kT\,\mathrm{Re}\{Z_s\}$ $[V^2/Hz]$*
- *The equivalent input noise voltage V_{eq} of the nullor. The spectral density of this source is given by S_{veq} $[V^2/Hz]$*
- *The equivalent input noise current I_{eq} of the nullor. The spectral density of this source is given by S_{ieq} $[A^2/Hz]$*
- *The thermal noise V_1 associated with R_1. The spectral density of this source is given by $4kTR_1$ $[V^2/Hz]$*

- *The thermal noise V_2 associated with R_2. The power spectral density of this source is given by $4kTR_2$ [V^2/Hz].*

The starting situation is shown in Figure 7.17. The transformation of all the noise sources to one equivalent input voltage source will be illustrated in six steps.

1. *In the first step, as shown in Figure 7.18, the equivalent current source of the nullor is redirected via the ground. It is then represented by two equal sources: one in parallel with R_1 and one in parallel with the source.*

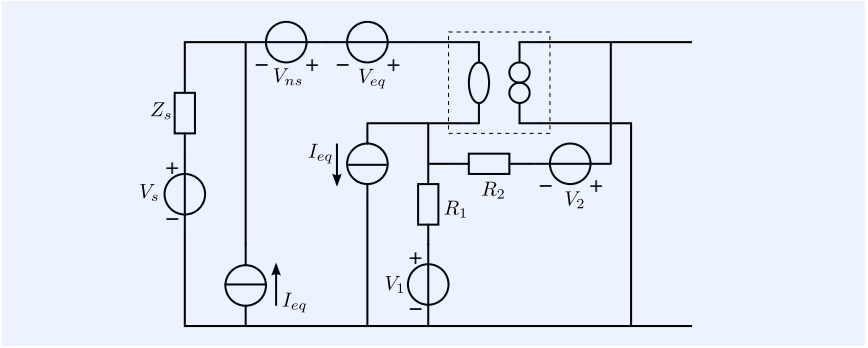


Figure 7.18: Application of the current-split theorem to I_{eq}.

2. *With the aid of the Thévenin transformation, both current sources can now be transformed into voltage sources, as shown in Figure 7.19. This results in two voltage sources, one in series with R_1 that has a value $I_{eq}R_1$, and one with a value of $I_{eq}Z_s$ in series with the source. Both sources are correlated, and therefore, we will keep track of the signs of these sources.*

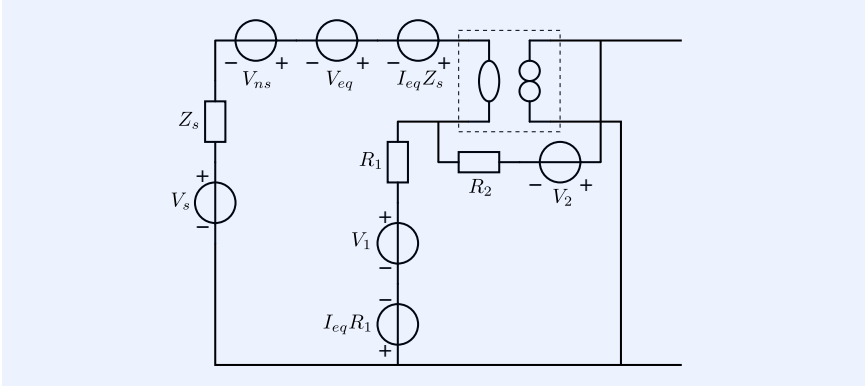


Figure 7.19: Conversion of noise current sources into voltage sources.

3. *The next step, the result of which is depicted in Figure 7.20, shows how the voltage sources in series with R_1 are shifted into the branch with R_2 and the branch with the input of the nullor.*

4. *The voltage sources in series with the input of the nullor can be shifted through the input port of the nullor. They now appear in series with the source. The voltage sources in series with R_2 can be shifted into the branch connected to the output of the voltage amplifier and the branch connected to the output of the nullor. This is shown in Figure 7.21.*

5. *The sources in series with the output port of the nullor can be transformed into equivalent input sources of the nullor by multiplying them by the corresponding transmission parameters. Since all of the transmission parameters of the nullor equal zero, these sources can be neglected.*

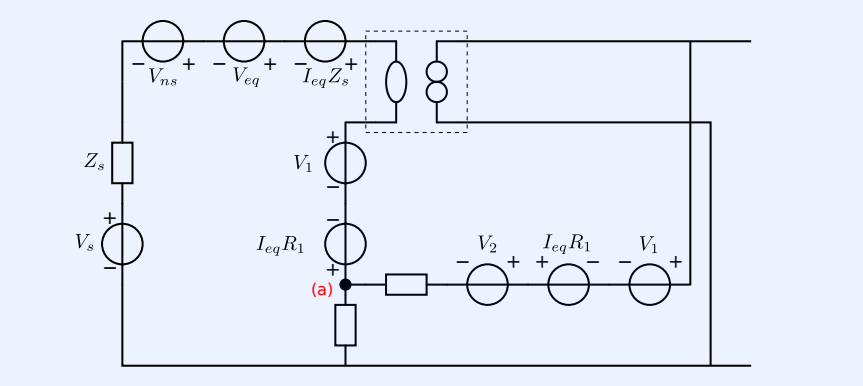


Figure 7.20: The voltage sources V_1 and $I_{eq}R_1$ are shifted through node (a).

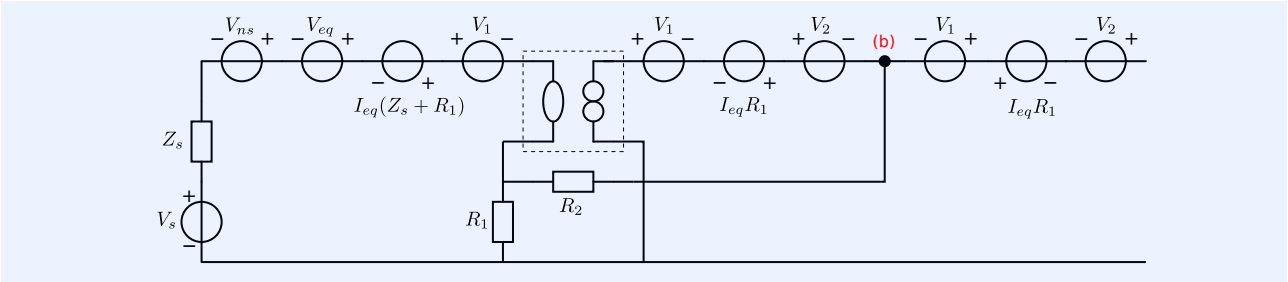


Figure 7.21: The noise voltage sources at the input are shifted through the input port of the nullor and the sources in series with R_2 are shifted through node (b).

The sources in series with the output port of the voltage amplifier can be transformed into equivalent input sources of the voltage amplifier by multiplying them by the corresponding transmission parameters of the voltage amplifier. Since, for the voltage amplifier, only the parameter A is fixed at a nonzero value, these output voltage sources can be represented by equivalent input voltage sources that have a value A times larger, with $A = R_1/(R_1 + R_2)$. These steps are shown in Figure 7.22.

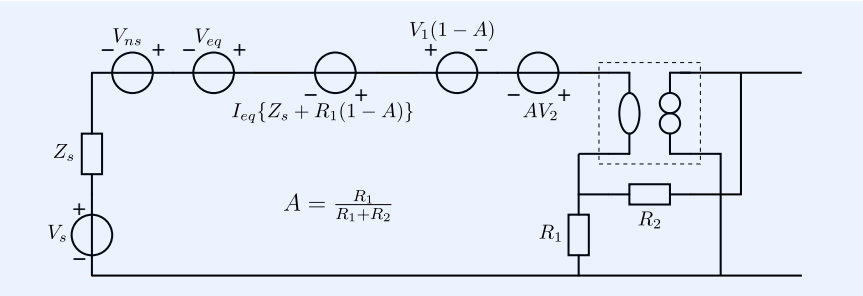


Figure 7.22: Transformation of all noise sources to the input.

6. *Figure 7.23 shows the final result of the noise transformations. The power spectral density of the total equivalent input noise voltage source can be expressed as:*

$$S_{V_{n,tot}} = 4kT\,\mathrm{Re}\{Z_s\} + S_{veq} + 4kT\frac{R_1R_2}{R_1+R_2} + S_{ieq}\left|Z_s + \frac{R_1R_2}{R_1+R_2}\right|^2. \quad (7.4)$$

7. *Thus, the influence of the feedback network on the noise performance of the passive feedback voltage amplifier can be calculated as if the parallel connection of R_1 and R_2 is in series with the signal source.*

   This confirms the equivalence of the noise models from Figure 7.17 and 7.24.

Conclusions with respect to the noise behavior of the other single loop passive feedback amplifier configurations can be derived in a similar way:

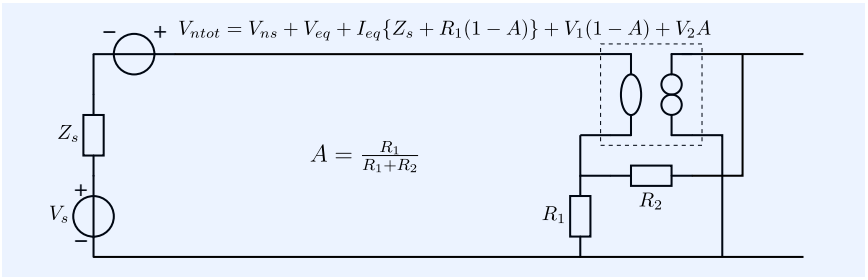


Figure 7.23: Final result of the noise transformations.

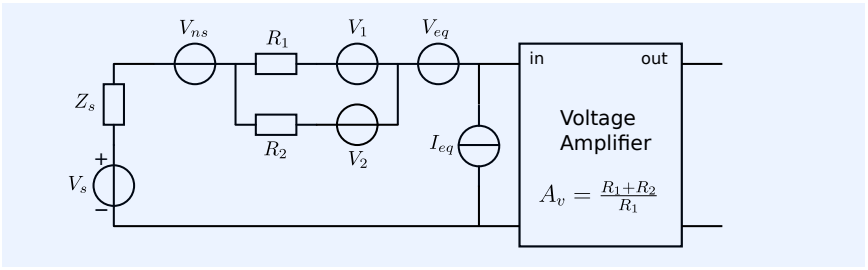


Figure 7.24: Noise model of the voltage amplifier, equal to that of Figure 7.17.

1. The total equivalent input noise of the voltage amplifier from Figure 7.16A can be calculated from the individual noise contributions as if the parallel connection of the feedback elements is in series with the source.
2. The total equivalent input noise of the transadmittance amplifier from Figure 7.16B can be calculated from the individual noise contributions as if the feedback element is in series with the source.
3. The total equivalent input noise of the transimpedance amplifier from Figure 7.16C can be calculated from the individual noise contributions as if the feedback element is in parallel with the source.
4. The total equivalent input noise of the current amplifier from Figure 7.16D can be calculated from the individual noise contributions as if the series connection of the feedback elements is in parallel with the source.

In the next example, we will demonstrate the use of source transformation techniques and the application of the superposition theorem. We start with the determination of the contribution of each individual noise source to the output noise of the amplifier. We then obtain the spectral density of the output noise by adding all of these contributions. The spectral density of the input noise is then obtained by dividing the output noise spectrum by the squared gain of the amplifier.

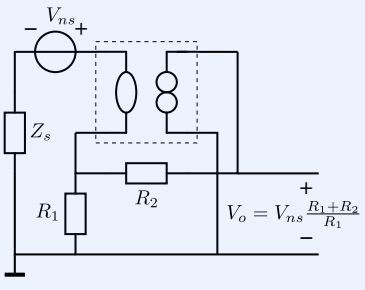


Figure 7.25: Contribution of V_{ns} to the output noise.

Example 7.6

We will evaluate the contributions of the individual noise sources to the total output noise spectrum S_{Vout}.

1. *The noise voltage V_{ns} associated with the signal source. The spectral density S_{vn} of this source equals: $4kT\,\mathrm{Re}\{Z_s\}$ [V^2/Hz].*

   The contribution S_1 of this noise source to S_{Vout} can be calculated by multiplying S_{vn} by the squared gain (see Figure 7.25):

$$S_1 = 4kT\,\mathrm{Re}\{Z_s\}\left(\frac{R_1+R_2}{R_1}\right)^2. \tag{7.5}$$

2. *The equivalent input noise voltage of the nullor V_{eq}. The spectral density of this source is given by: S_{veq} [V^2/Hz].*

   The contribution S_2 of this noise source to S_{Vout} can be calculated in a similar way:

$$S_2 = S_{veq}\left(\frac{R_1+R_2}{R_1}\right)^2. \tag{7.6}$$

3. *The equivalent input noise current I_{eq} of the nullor. The power spectral density of this source is given by: S_{ieq} [A^2/Hz].*

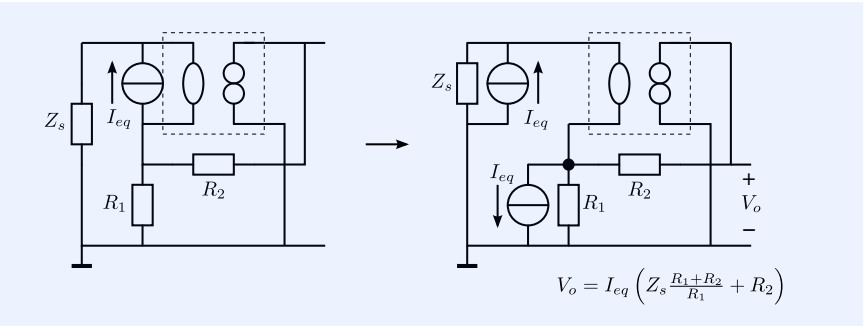


Figure 7.26: Evaluation of the contribution of I_{eq} to the output noise.

   We can split this source into two correlated sources as shown in Figure 7.26. The current source in parallel with the source impedance causes an output voltage V_1:

$$V_1 = I_{eq}Z_s\frac{R_1+R_2}{R_1}. \tag{7.7}$$

   The current source in parallel with R_1 causes an output voltage V_2:

$$V_2 = I_{eq}R_2. \tag{7.8}$$

   These are two correlated contributions, so we add these voltages and find their contribution S_3 to the spectrum of the output noise voltage by multiplying the spectral density of I_{eq} by the squared magnitude of the transfer from I_{eq} to the output voltage. In this way, we obtain

$$S_3 = S_{ieq}\left|Z_s\frac{R_1+R_2}{R_1}+R_2\right|^2. \tag{7.9}$$

   This can be written as

$$S_3 = S_{ieq}\left(\frac{R_1+R_2}{R_1}\right)^2\left|Z_s+\frac{R_1R_2}{R_1+R_2}\right|^2. \tag{7.10}$$

4. *The thermal noise V_1 associated with the feedback element R_1. The spectral density of this source is given by: $4kTR_1$ [V^2/Hz].*

   The gain of this noise source to the output voltage equals $-R_2/R_1$. The contribution S_4 to the spectral density of the out[put noise voltage is thus obtained by multiplying $4kTR_1$ with the squared magnitude of this transfer. We obtain:

$$S_4 = 4kT\frac{R_2^2}{R_1} \tag{7.11}$$

5. *The thermal noise V_2 associated with the feedback element R_2. The spectral density of this source is given by $4kTR_2$ [V^2/Hz].*

   This noise source can be converted into a current source I_2 with a spectral density $S_2 = 4kT/R_2$ [A^2/Hz]. It can be split into two correlated sources, as shown in

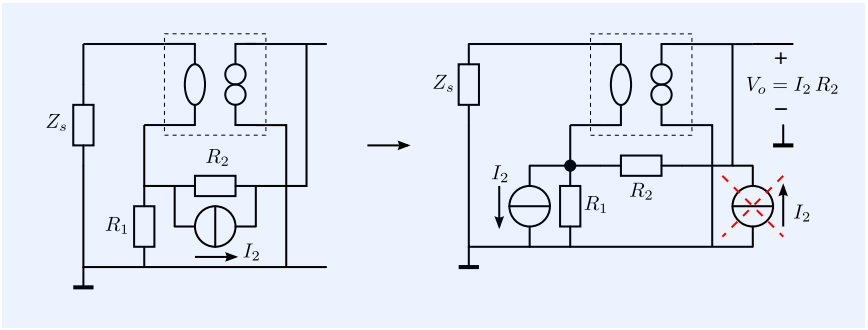


Figure 7.27: Evaluation of the contribution of the thermal noise of R_2 to the output noise.

Figure 7.27. The source in parallel with the norator has no effect. The remaining source contributes S_5 to the output spectrum:

$$S_5 = 4kTR_2. \tag{7.12}$$

6. *The spectral density of the output voltage noise is obtained by adding the contributions of S_1 through S_5:*

$$S_{Vout} = \left(\frac{R_1+R_2}{R_1}\right)^2 \left\{4kT\,\mathrm{Re}\{Z_s\} + S_{veq} + S_{ieq}\left|Z_s + \frac{R_1R_2}{R_1+R_2}\right|^2\right\} + \\ + 4kT\left(\frac{R_2^2}{R_1} + R_2\right), \tag{7.13}$$

which can be written as:

$$S_{Vout} = \left(\frac{R_1+R_2}{R_1}\right)^2 \times \\ \times \left\{4kT\,\mathrm{Re}\{Z_s\} + S_{veq} + S_{ieq}\left|Z_s + \frac{R_1R_2}{R_1+R_2}\right|^2 + 4kT\frac{R_2R_1}{R_2+R_1}\right\}. \tag{7.14}$$

7. *The source-referred noise spectrum is obtained after multiplication of this result by the reciprocal value of the squared magnitude of the voltage gain:*

$$S_{V_{n,tot}} = 4kT\,\mathrm{Re}\{Z_s\} + S_{veq} + S_{ieq}\left|Z_s + \frac{R_1R_2}{R_1+R_2}\right|^2 + 4kT\frac{R_1R_2}{R_1+R_2}. \tag{7.15}$$

Finally, we will evaluate the total equivalent input voltage noise of the passive-feedback voltage amplifier with the aid of the MNA method. This method has been implemented in SL1CAP.

Example 7.7

In order to keep the matrices as small as possible, we will replace the voltage noise sources V_1 and V_2 with current sources I_1 and I_2 and combine the noise sources V_{ns} and V_{eq} in one voltage source. In addition, we will omit the signal source V_s. The simplified equivalent noise model is shown in Figure 7.28. Using this model, we will calculate the total equivalent input noise from the total output noise. We will evaluate the total output noise and obtain the source-referred noise after multiplication of the output noise voltage by the reciprocal value of the voltage gain: $\frac{R_1}{R_1+R_2}$. The total output noise will be determined using MNA and Cramer's rule.

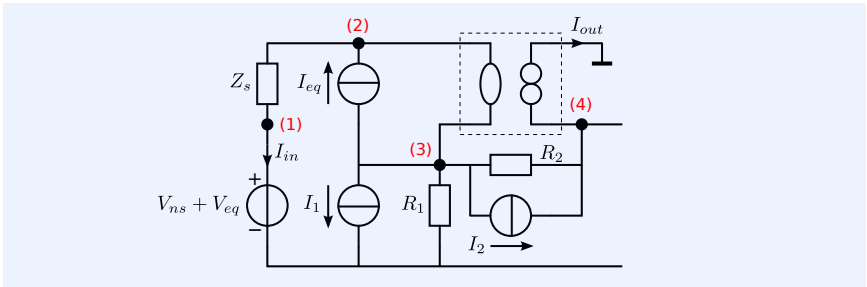


Figure 7.28: Equivalent noise model of the voltage amplifier for MNA.

The circuit can be described with the following matrix equation:

$$\begin{bmatrix} 0 \\ I_{eq} \\ -I_1 - I_2 - I_{eq} \\ I_2 \\ V_{ns} + V_{eq} \\ 0 \end{bmatrix} = \mathbf{M} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ I_{in} \\ I_{out} \end{bmatrix}, \tag{7.16}$$

where

$$\mathbf{M} = \begin{bmatrix} \frac{1}{Z_s} & -\frac{1}{Z_s} & 0 & 0 & 1 & 0 \\ -\frac{1}{Z_s} & \frac{1}{Z_s} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{R_1}+\frac{1}{R_2} & -\frac{1}{R_2} & 0 & 0 \\ 0 & 0 & -\frac{1}{R_2} & \frac{1}{R_2} & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix}. \tag{7.17}$$

The total equivalent input noise voltage V_{ntot} can be found from

$$V_{ntot} = \frac{R_1}{R_1 + R_2} V_{out} \tag{7.18}$$

The output voltage V_{out} equals V_4. It can be found with the aid of Cramer's rule:

$$V_{out} = \frac{\det(\mathbf{M}')}{\det(\mathbf{M})}, \tag{7.19}$$

where

$$\mathbf{M}' = \begin{bmatrix} \frac{1}{Z_s} & -\frac{1}{Z_s} & 0 & 0 & 1 & 0 \\ -\frac{1}{Z_s} & \frac{1}{Z_s} & 0 & I_{eq} & 0 & 0 \\ 0 & 0 & \frac{1}{R_1}+\frac{1}{R_2} & -I_1 - I_2 - I_{eq} & 0 & 0 \\ 0 & 0 & -\frac{1}{R_2} & I_2 & 0 & 1 \\ 1 & 0 & 0 & V_{ns} + V_{eq} & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix}. \tag{7.20}$$

With the aid of a symbolic math tool, this can quickly be simplified to

$$V_{ntot} = V_{eq} + V_{ns} + I_{eq}\left(\frac{R_1 R_2}{R_1 + R_2} + Z_s\right) + I_1 \frac{R_1 R_2}{R_1 + R_2} + I_2 \frac{R_1 R_2}{R_1 + R_2}. \tag{7.21}$$

Hence, for the spectral density $S_{V_{ntot}}$ of V_{ntot}, we write:

$$S_{V_{ntot}} = S_{V_{eq}} + S_{V_{ns}} + S_{I_{eq}} \left|\frac{R_1 R_2}{R_1 + R_2} + Z_s\right|^2 + \left(S_{I_1} + S_{I_2}\right)\left|\frac{R_1 R_2}{R_1 + R_2}\right|^2, \tag{7.22}$$

which can again be written as:

$$S_{V_{ntot}} = S_{V_{eq}} + 4kT\,\mathrm{Re}\,(Z_s) + S_{I_{eq}}\left|\frac{R_1R_2}{R_1+R_2} + Z_s\right|^2 + 4kT\frac{R_1R_2}{R_1+R_2}. \qquad (7.23)$$

In the following example, we will demonstrate the way in which SLiCAP can be used to perform such noise calculations.

Example 7.8

Figure 7.29 shows the SLiCAP *circuit diagram for the noise analysis of the passive feedback voltage amplifier. The spectral densities of all uncorrelated noise sources are given in [V²/Hz] or in [A²/Hz].*

The SLiCAP *netlist for this circuit is shown below. The controller of the amplifier is the 'noisy nullor'* `x1`*. It is a SLiCAP library element (sub circuit* `N_noise`*) that consists of a nullor with two equivalent-input noise sources. The spectral densities of these noise sources are parameters of this sub circuit.*

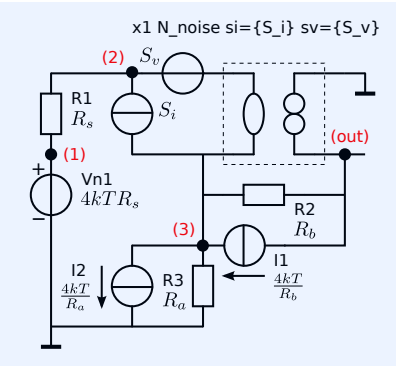


Figure 7.29: Equivalent noise model of the voltage amplifier for SLiCAP calculations.

```
VampNoise.cir
X1 2 3 out 0 N_noise si={S_i} sv={S_v}
R1 out 3 {R_b}
R2 3 0 {R_a}
I1 out 3 I noise={4*k*T/R_b}
I2 3 0 I noise={4*k*T/R_a}
R3 2 1 {R_s}
V1 1 0 V noise={4*k*T*R_s}
.end
```

The SLiCAP *script for evaluation of the source-referred noise and the detector-referred noise is shown below. The signal source is the voltage source* `Vn1`*. The detector is a voltage detector between node (5) and the ground.*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Mon Mar 29 20:31:51 2021

@author: anton
"""
from SLiCAP import *

fileName = 'VampNoise';
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName + '.cir')
htmlPage('Circuit data')
head2html('Circuit diagram: ' + fileName)
img2html(fileName + '.svg', 600)
netlist2html(fileName + '.cir')
#
i1.setGainType('vi')
i1.setDataType('noise')
#
# Define the source and the detector
i1.setSource('V1');
i1.setDetector('V_out');
#
htmlPage('Symbolic noise analysis')
i1.setSimType('symbolic')
noiseResultSym = i1.execute()
noise2html(noiseResultSym)
```

The HTML output, generated by SLiCAP is shown in Figure 7.30.

Example 7.9

In thist example we will demonstrate the determination of show-stopper values for R_a, R_b, S_v and S_i for a given source resistance, noise figure and voltage gain. To do so, we assign values to R_s and the voltage gain A_v and establish a relation between R_a and R_b.

$$R_b = (A_v - 1)\,R_a, \qquad (7.24)$$

Symbolic noise analysis

Symbolic noise analysis results

Detector-referred noise spectrum

$$S_{out} = 4R_bTk + \frac{4R_b^2Tk}{R_a} + \frac{4R_sTk(R_a+R_b)^2}{R_a^2} + \frac{S_i(R_aR_b+R_aR_s+R_bR_s)^2}{R_a^2} + \frac{S_v(R_a+R_b)^2}{R_a^2}\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$$

Source-referred noise spectrum

$$S_{in} = \frac{4R_a^2R_bTk}{(R_a+R_b)^2} + \frac{4R_aR_b^2Tk}{(R_a+R_b)^2} + 4R_sTk + \frac{S_i(R_aR_b+R_aR_s+R_bR_s)^2}{(R_a+R_b)^2} + S_v\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$$

Contributions of individual noise sources

Noise source: I1

Spectral density:	$\frac{4Tk}{R_b}$	$\left[\frac{\mathrm{A}^2}{\mathrm{Hz}}\right]$
Detector-referred:	$4R_bTk$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$
Source-referred:	$\frac{4R_a^2R_bTk}{(R_a+R_b)^2}$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$

Noise source: I1_XU1

Spectral density:	S_i	$\left[\frac{\mathrm{A}^2}{\mathrm{Hz}}\right]$
Detector-referred:	$\frac{S_i(R_aR_b+R_aR_s+R_bR_s)^2}{R_a^2}$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$
Source-referred:	$\frac{S_i(R_aR_b+R_aR_s+R_bR_s)^2}{(R_a+R_b)^2}$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$

Noise source: I2

Spectral density:	$\frac{4Tk}{R_a}$	$\left[\frac{\mathrm{A}^2}{\mathrm{Hz}}\right]$
Detector-referred:	$\frac{4R_b^2Tk}{R_a}$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$
Source-referred:	$\frac{4R_aR_b^2Tk}{(R_a+R_b)^2}$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$

Noise source: V1

Spectral density:	$4R_sTk$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$
Detector-referred:	$\frac{4R_sTk(R_a+R_b)^2}{R_a^2}$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$
Source-referred:	$4R_sTk$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$

Noise source: V1_XU1

Spectral density:	S_v	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$
Detector-referred:	$\frac{S_v(R_a+R_b)^2}{R_a^2}$	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$
Source-referred:	S_v	$\left[\frac{\mathrm{V}^2}{\mathrm{Hz}}\right]$

Figure 7.30: SLiCAP symbolic noise results.

and calcuate the noise as a function of these parameters. This is done in lines 30-42 of the script:

```
30 #
31 # Let us find show-stopper values for R_a, S_v, and S_i for the case:
32 # - source resistance: R_s = 600 Ohm
33 # - voltage gain     : A_v = 20
34 # - noise figure     : NF  = 2 (3dB)
35 #
36 i1.defPar('R_s', 600)
37 # Define R_b = (A_v-1)*R_a
38 i1.defPar('R_b', '(A_v-1)*R_a')
39 i1.defPar('A_v', 20)
40 i1.setSimType('numeric')
41 # Calculate the noise with the given parameters.
42 noiseResultNum = i1.execute()
```

The remaining symbolic variables R_a, S_v and S_i as well as their show-stopper values $R_{a_{\max}}$, $S_{v_{\max}}$ and $S_{i_{\max}}$, and the noise factor NF are defined in the python environment in lines 43-47 of the script:

```
43 #
44 # Determine the noise figure NF: (the given procedure works with
45 # frequency-independent noise spectra only)
46 #
47 R_a, S_v, S_i, NF, R_a_max, S_i_max, S_v_max = sp.symbols('R_a, S_v, S_i,' +
48                                                 'NF, R_a_max, S_i_max, S_v_max')
```

The noise factor is calculated as the ratio of the total source-referred noise spectrum and the spectrum of the noise associated with the signal source. This is correct if the noise spectra do not depend on frequency (white noise); see lines 49-57 of the script:

```
49 htmlPage('Show-stopper values')
50 #
51 text2html('Let us find show-stopper values for $R_a$, $S_v$, and $S_i$ for ' +
52           'the case in which the noise factor $NF$ equals 2 (3dB).')
53 head2html('Noise factor NF')
54 text2html('The noise factor NF [-] is obtained as:')
55 NFact = sp.simplify(noiseResultNum.inoise/noiseResultNum.inoiseTerms['V1'])
56 eqn2html(NF, NFact);
57 #
```

The show stopper values of $R_{a_{\max}}$, $S_{v_{\max}}$ and $S_{i_{\max}}$ are found by solving the equation of NF for one variable, while assuming the other noise contributions zero. Since zero value for the feedback resistances is not meaningful, the show stopper values $S_{v_{\max}}$ and $S_{i_{\max}}$ are calculated as a function of R_a; see lines 49-57 of the script. Figure 7.31 shows the results.

```
58 # Show stopper (= maximum) value $R_{amax}$ for R_a with S_i=0 and S_v=0
59 Ra_max = sp.N(sp.solve(NFact.subs([(S_v, 0), (S_i, 0)])-2, R_a)[0], 3)
60 head2html('Show-stopper value $R_a$');
61 text2html('The show stopper value $R_{amax}$ for $R_a$ with $NF=2$, ' +
62           '$S_v=0$ and $S_i=0$ is obained as:');
63 eqn2html(R_a_max, Ra_max);
64 #
65 # Show stopper (= maximum) $S_{v,max}$ for S_v as a function of R_a and S_i=0
66 Sv_max = sp.N(sp.solve(NFact.subs(S_i, 0)-2, S_v)[0], 3);
67 head2html('Show-stopper value $S_v$');
68 text2html('The show stopper value for $S_v$ with $NF=2$ and $S_i=0$ can be ' +
69           'obained a function of $R_a$ (setting $R_a$ to zero would be ' +
70           'meaningless):')
71 eqn2html(S_v_max, Sv_max);
72 #
73 # Show stopper (= maximum) $S_{i,max}$ for S_i as a function of R_a and S_i=0
74 Si_max = sp.N(sp.solve(NFact.subs(S_v, 0)-2, S_i)[0], 3);
75 head2html('Show-stopper value $S_i$');
76 text2html('The show stopper value for $S_i$ with $NF=2$ and $S_v=0$ can be ' +
77           'obained a function of $R_a$: (setting $R_a$ to zero would be ' +
78           'meaningless):');
79 eqn2html(S_i_max, Si_max);
```

Show-stopper values

Let us find show-stopper values for R_a, S_v, and S_i for the case in which the noise factor NF equals 2 (3dB).

Noise factor NF

The noise factor NF [-] is obtained as:

$$NF = 0.001583R_a + 3.621 \cdot 10^{22} S_i (0.001583R_a + 1)^2 + 1.006 \cdot 10^{17} S_v + 1 \tag{1}$$

Show-stopper value R_a

The show stopper value R_{amax} for R_a with $NF = 2$, $S_v = 0$ and $S_i = 0$ is obained as:

$$R_{amax} = 632.0 \tag{2}$$

Show-stopper value S_v

The show stopper value for S_v with $NF = 2$ and $S_i = 0$ can be obained a function of R_a (setting R_a to zero would be meaningless):

$$S_{vmax} = 9.94 \cdot 10^{-18} - 1.57 \cdot 10^{-20} R_a \tag{3}$$

Show-stopper value S_i

The show stopper value for S_i with $NF = 2$ and $S_v = 0$ can be obained a function of R_a: (setting R_a to zero would be meaningless):

$$S_{imax} = \frac{2.76 \cdot 10^{-23} - 4.361 \cdot 10^{-26} R_a}{(0.00158R_a + 1)^2} \tag{4}$$

Figure 7.31: Derivation and solution of the component design equations for the noise performance.

7.5.3 Power efficiency of passive feedback configurations

In this section we will summarize the influence of the feedback networks of the basic passive feedback amplifier configurations on their power efficiency. These conclusions simply follow from network inspection:

1. The power efficiency of the voltage amplifier from Figure 7.16A is affected by the feedback elements as if the series connection of the feedback elements is in parallel with the load impedance.
2. The power efficiency of the transadmittance amplifier from Figure 7.16B is affected by the feedback elements as if the feedback element is in series with the load impedance.
3. The power efficiency of the transimpedance amplifier from Figure 7.16C is affected by the feedback elements as if the feedback element is in parallel with the load impedance.
4. The power efficiency of the current amplifier from Figure 7.16D is affected by the feedback elements as if the parallel connection of the feedback elements is in series with the load impedance.

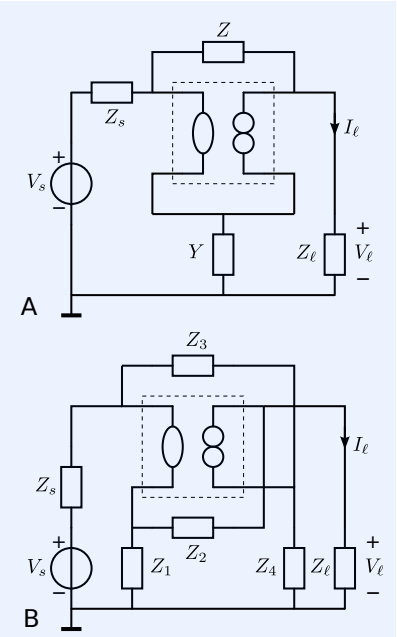


Figure 7.32: Dual-loop passive feedback configurations
A: Gyrator-like configuration
B: Transformer like configuration.

7.5.4 Dual-loop passive feedback configurations

Figure 7.32 shows the basic single-nullor dual-loop feedback configurations that can be realized with passive feedback alone. These are the gyrator-like configuration and the transformer-like configuration.

The gyrator-like configuration from Figure 7.32A has some interesting properties. The impedance Z establishes parallel feedback at both the input and the output port (load voltage sensing and source current comparison). Hence, it fixes transmission parameter C. The admittance Y establishes input and output series feedback (load current sensing and source voltage comparison), which fixes parameter B. Due to voltage drop across the current sensing element Y and current though the voltage-sensing element Z, all transmission parameters have been fixed by Z and Y. However, only two of them can be fixed independently.

The transmission parameters of this configuration are:

$$A = \frac{1}{1-YZ},\ B = \frac{Z}{1-YZ},\ C = \frac{Y}{1-YZ},\ D = \frac{1}{1-YZ}. \tag{7.25}$$

This configuration is often applied in characteristic impedance systems. Impedance matching at both ports is obtained if: $Z_s = Z_\ell = \sqrt{\frac{Z}{Y}}$.

In the transformer-like configuration from Figure 7.32B, a voltage attenuator (Z_1, Z_2) and a current attenuator (Z_3, Z_4) are used to fix transmission parameters A and D, respectively. Due to current flow through the voltage attenuator and voltage drop across the current attenuator, parameter C is also fixed. The transmission parameters are:

$$A = \frac{Z_1}{Z_1+Z_2},\ B = 0,\ C = \frac{Z_1+Z_4}{(Z_1+Z_2)(Z_3+Z_4)},\ D = \frac{Z_4}{Z_3+Z_4}. \tag{7.26}$$

An accurate input impedance can be obtained if the output port is left open. This is true in practice, if $Z_\ell \gg Z_1 + Z_2$.

The configuration has an accurate output impedance if it is driven from an ideal current source; in practice, if $Z_s \gg Z_3 + Z_4$.

Other passive single-nullor dual-loop configurations have opposite signs for their transmission parameters, which results in a negative input or output impedance. This will be demonstrated in the next example.

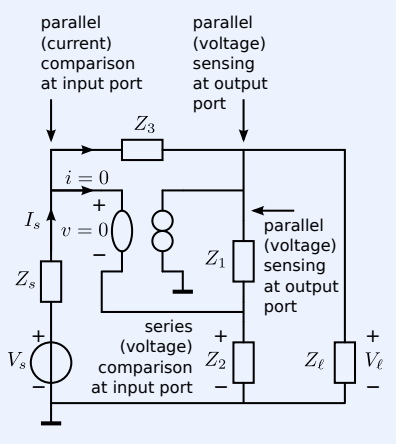


Figure 7.33: Dual-loop single-nullor passive feedback configuration with A and C fixed to a nonzero value and $B = D = 0$.

Example 7.10

Let us try to design a passive feedback amplifier that has a finite nonzero input impedance and zero output impedance. For such an amplifier, we need to fix the parameters A and C to a finite nonzero value, while having $B = D = 0$. Fixing A requires parallel sensing at the output port and series comparison at the input port. Fixing C requires parallel sensing at the output port and parallel comparison at the input port. With passive feedback, this would give the circuit shown in Figure 7.33.

Circuit analysis shows that

$$A = \frac{Z_2}{Z_1+Z_2},\ C = -\frac{Z_1}{Z_1+Z_2}\frac{1}{Z_3}, \tag{7.27}$$

which yields:

$$Z_i = \frac{A}{C} = -Z_3\frac{Z_2}{Z_1}. \tag{7.28}$$

Hence, the input impedance is negative! In order to obtain a positive input impedance, we need to create a signal inversion in one of the feedback loops. Means to achieve this will be discussed in the following sections.

7.6 Active feedback

We have seen that the sign of the transfer of the single-loop passive feedback amplifiers from Figure 7.16 is indissolubly connected to the type of their transfer. The voltage amplifier and the current amplifier have a noninverting transfer, while the transadmittance amplifier and the transimpedance amplifier have an inverting transfer. The sign of the transfer is the result of the feedback topology and the noninverting transfer of the passive feedback network. The inverting voltage amplifier, an inverting current amplifier, as well a noninverting transadmittance amplifier and a noninverting transimpedance amplifier, all require an inverting feedback network.

At a first glance, this may be a strange conclusion. Many experienced designers know the configuration depicted in Figure 7.34 as the inverting voltage amplifier. Looking at this circuit from a conceptual point of view, one might more reasonable call it an inverting transimpedance amplifier (the one from Figure 7.16C) with an input series impedance for (brute-force) voltage-to-current conversion. In cases in which the source impedance is very small and the noise performance is not critical, this solution may be good, but in other situations, brute force is not the way to go.

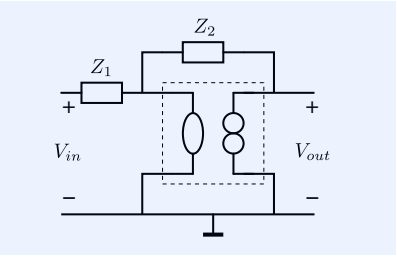


Figure 7.34: Circuit known as an inverting voltage amplifier.

Inverting feedback networks can be constructed with the aid of the transadmittance amplifier or the transimpedance amplifier. An inventory of inverting voltage networks and inverting current networks is shown in Figure 7.35. Since these feedback networks comprise an amplifier, they are active circuits. Amplifiers that use active feedback networks are called *active feedback amplifiers*.

We will discuss the design of single loop active feedback amplifiers in section 7.6.1 and that of multiple loop active feedback configurations in section 7.6.2.

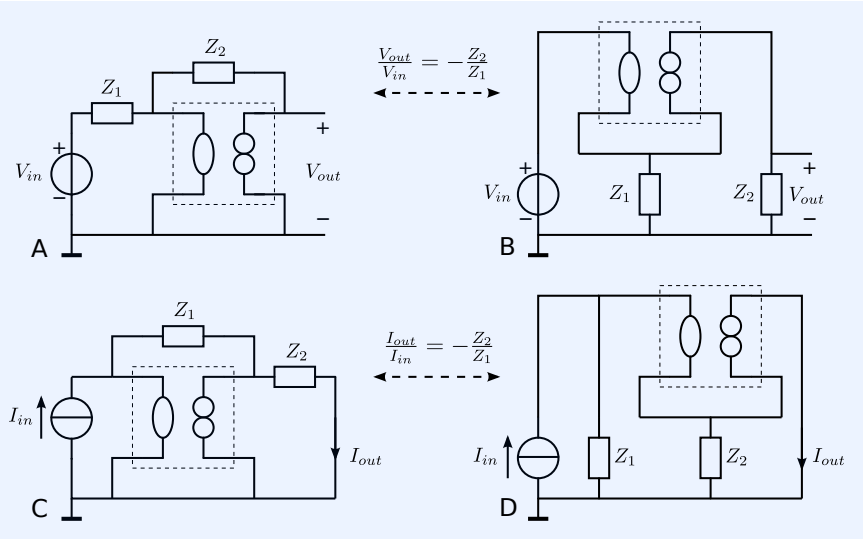


Figure 7.35: Inverting voltage attenuators (A,B) and inverting current attenuators (C, D). These attenuators are realized with the inverting passive-feedback transimpedance configuration (A, C) or the inverting passive-feedback transadmittance configuration (B, D).

7.6.1 Single-loop active feedback

Figure 7.36 shows an inverting voltage amplifier, using transimpedance feedback. An alternative solution using transadmittance feedback is left as an exercise to the reader. The inverting current amplifier, the noninverting transadmittance amplifier and the noninverting transimpedance amplifier can be similarly realized. These configurations are not drawn; they are also left as an exercise for the reader. The reader is also invited to evaluate the influence of the feedback elements on the noise behavior and the power efficiency for

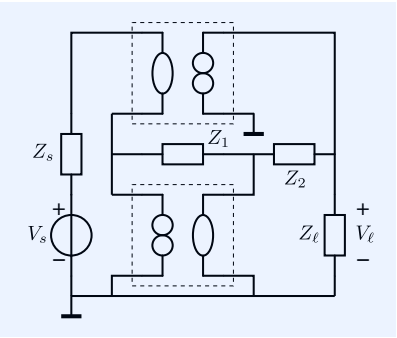


Figure 7.36: Active-feedback inverting voltage amplifier.

these single loop active feedback configurations.

Let us now consider the inverting voltage amplifier from Figure 7.36. We may generate a new circuit with similar properties by pairing the nullators and norators differently. This is because a network with nullors has a unique solution if the input conditions set by n nullators, can be satisfied with n norators. Theoretically, there exists no transfer from a nullator to a norator: a nullator sets a network condition by adding an equation and a norator adds the required extra dependent variable such that the condition can be satisfied.

After different pairings of the nullators and norators, a new circuit is found as the cascade connection of two single loop feedback amplifiers. This configuration is shown in Figure 7.37.

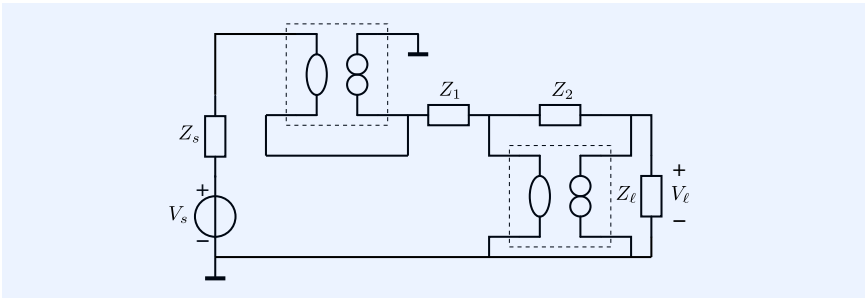


Figure 7.37: Alternative solutions for the active-feedback inverting voltage amplifiers from Figure 7.36.

7.6.2 Multiple-loop active feedback

Let us assume we need to design a negative feedback amplifier that has an accurately fixed finite nonzero input impedance and zero output impedance. We tried this already in section 7.5.4, but with a single nullor and passive feedback, this resulted in a negative input impedance because A and C had different signs. Equal signs for both transmission parameters requires a signal inversion in one of the feedback loops. Hence, there are two different realizations for each dual-loop feedback configuration.

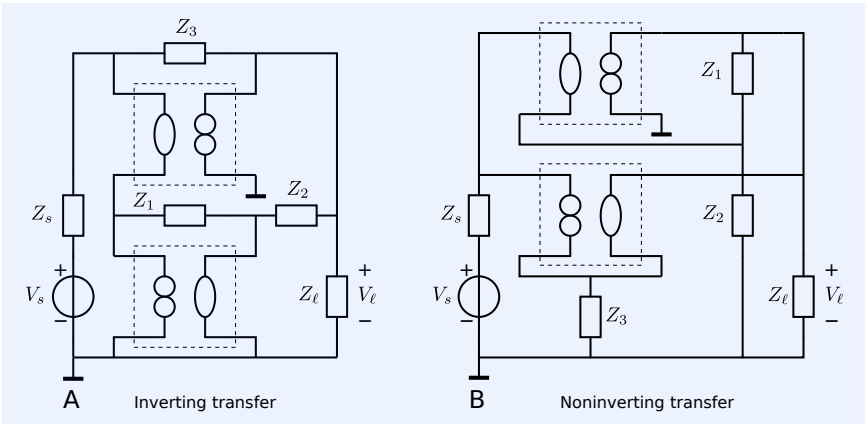


Figure 7.38: Examples of dual-loop active feedback configurations.
(A) A and C fixed, $B = D = 0$, inverting transfer.
(B) A and C fixed, $B = D = 0$, noninverting transfer.

Figure 7.38 shows two of those dual-loop configurations using active feedback. Both configurations have A and C fixed to a nonzero value while B and D are zero. The sign of their source-to-load transfer, however, is opposite.

7.7 Design of balanced amplifiers

All the passive and active negative feedback configurations that we have introduced until now had their source and their load connected to the ground. Configurations with one or two amplifier ports floating with respect to the ground have not yet been discussed.

In this section, we will focus on the design of amplifier configurations that have their input and/or output port floating with respect to the ground.

In section 7.7.1 we will introduce 'anti-series connection' as the basic technique for converting unbalanced amplifiers into a balanced ones and discuss the properties of anti-series connected amplifiers.

We will show that their transfer, their port impedances, their noise performance and their nonlinearity can easily be related to the those of the unbalanced amplifiers of which they are constructed.

Although anti-series connection of unbalanced amplifiers results in amplifiers that have their ports isolated from the ground, it does not always establish an isolation between the input port and the output port.

Modeling and analysis techniques for studying the common-mode behavior of balanced amplifiers will be discussed in section 7.7.3. In that section we will introduce techniques for converting a balanced circuit into equivalent circuits that model their common-mode behavior and their differential-mode behavior separately.

Section 7.7.4 will be devoted to the design of amplifiers that have natural two-port behavior.

7.7.1 Anti-series connected amplifiers

Amplifiers that have their ports floating with respect to the ground, can be obtained from anti-series connection of three-terminal amplifiers. Figure 7.39 shows the principle of anti-series connection of two identical two-ports.

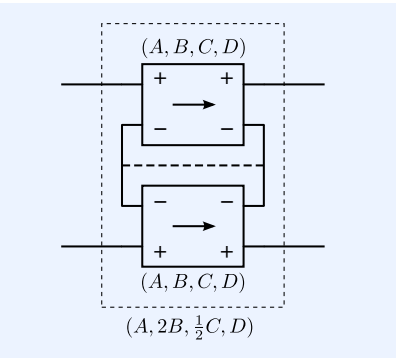


Figure 7.39: Anti-series connection of two equal two-ports. The dashed line shows the connection if both two-ports are three terminal networks. If both two-ports have transmission-1 parameters A, B, C, and D, the resulting two-port has A, $2B$, $\frac{1}{2}C$, and D, repectively.

Transmission-1 two-port parameters

It can be shown that the transmission parameters of a two-port that consists of a series or an anti-series connection of two identical two-ports with transmission parameters A, B, C, and D, equal A, $2B$, $\frac{1}{2}C$, and D, respectively.

Noise behavior of anti-series connected amplifiers

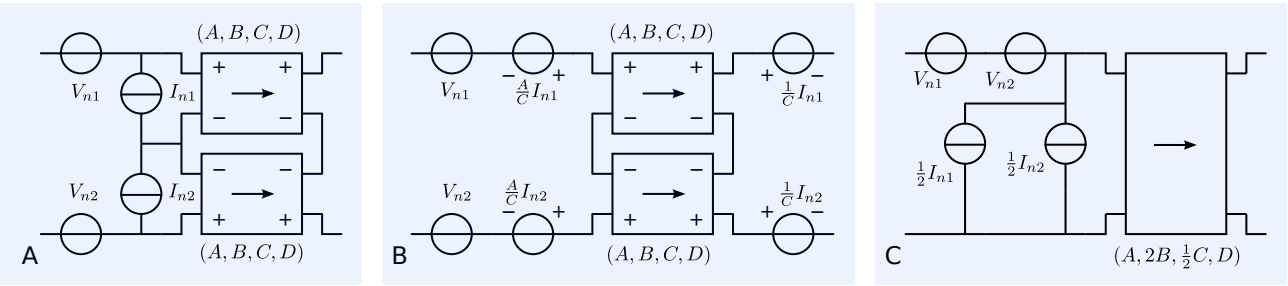


Figure 7.40: Equivalent input noise sources of (anti-) series connected two-ports.

The equivalent input noise sources of the (anti-) series connection can be obtained from the equivalent input noise sources of the individual two-ports. Figure 7.40A shows the anti-series connection of the two two-ports with their individual equivalent input noise sources. This noise representation, however, is not very convenient for evaluation of the equivalent input noise sources of the (anti-) series connection. A representation with two voltage sources is more suited to this situation, because with such a representation, the total noise voltages at the input and at the output of the resulting two-port

are simply found by adding those of the constituting two-ports. An equivalent input representation of the (anti-) series connection can then be obtained by replacing the output voltage source with equivalent input sources, thereby using the transmission parameters of the (anti-) series connection. Figure 7.40B shows the equivalent voltage noise representation (see Chapter 19.2). Figure 7.40C gives the equivalent sources of the (anti-) series connection. The power spectral densities of the equivalent noise sources of the (anti-) series connection are obtained as:

$$S_{vtot} = S_{vn1} + S_{vn2} \text{ [V}^2\text{/Hz]}, \tag{7.29}$$

$$S_{itot} = \frac{1}{4}S_{in1} + \frac{1}{4}S_{in2} \text{ [A}^2\text{/Hz]}. \tag{7.30}$$

If both constituting two-ports have identical noise behavior, we have: $S_{vn1} = S_{vn2} = S_v$ [V²/Hz] and $S_{in1} = S_{1n2} = S_i$ [A²/Hz]. We then obtain:

$$S_{vtot} = 2S_v \text{ [V}^2\text{/Hz]}, \tag{7.31}$$

$$S_{itot} = \frac{1}{2}S_i \text{ [A}^2\text{/Hz]}. \tag{7.32}$$

Consider a noise figure F obtained for a two-port driven from a source impedance Z_s. The anti-series connection of two of those two-ports then has the same noise figure when driven from an impedance $2Z_s$.

Nonlinearity of balanced amplifiers

Anti-series connection is an additive compensation technique. If an instantaneous nonlinear transfer of the individual (identical) two-ports is modeled by a Taylor series expansion, the corresponding transfer of the anti-series connection will show odd terms only; all even terms will be cancelled. As a consequence, with perfect balancing, the offset and all even order nonlinearity are cancelled.

7.7.2 Balanced single-loop configurations

Figure 7.41 shows the four single-loop balanced amplifiers that are obtained from anti-series connection of their unbalanced versions.

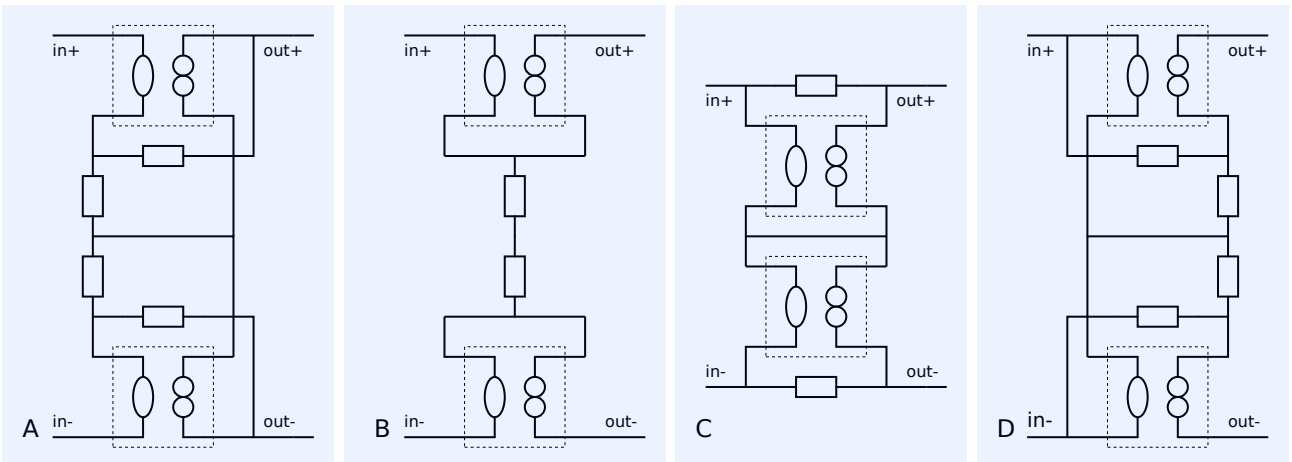


Figure 7.41: Single-loop balanced amplifiers:
A: Balanced voltage amplifier
B: Balanced transadmittance amplifier
C: Balanced transimpedance amplifier
D: Balanced current amplifier.

None of the terminals of these amplifiers has been connected to ground. Hence, their source and/or load can be floating with respect to the ground. However, if the amplifier itself does not behave as a natural two-port[8], par-

[8] See section 2.2.3.

asitic impedances from the source or the load to the ground may limit the *CMRR* of the amplifier. The analysis of the common-mode behavior as well as the design of amplifiers that exhibit natural two-port behavior, will be discussed in the following sections.

7.7.3 Common-mode behavior of balanced amplifiers

If an amplifier behaves as a natural two-port, it can be configured as a differential driver, a differential receiver, or a differential driver-receiver (see section 2.2.3). A natural two-port is a four terminal network to which the two-port conditions apply (see Chapter 18.6.1).

Figure 7.42 shows a four-terminal network with the definitions of the input and the output common-mode and differential-mode voltages and currents.

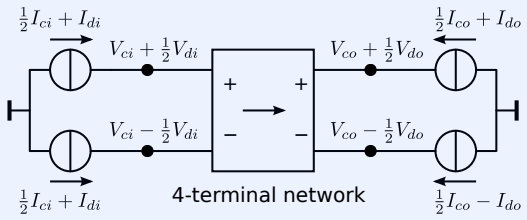


Figure 7.42: Definitions of the common-mode and differential-mode currents and voltages for a 4-terminal network.
V_{ci} common-mode input voltage
V_{di} differential-mode input voltage
I_{ci} common-mode input current
I_{di} differential-mode input current
V_{co} common-mode output voltage
V_{do} differential-mode output voltage
I_{co} common-mode output current
I_{do} differential-mode output current.

Natural two-ports have the following properties:

1. Under all conditions, the common-mode current into each port equals zero.

2. A common-mode voltage across the two ports does not change the differential-mode port quantities.

As stated before, an anti-series connection of two three-terminal amplifiers, does not automatically result in natural two-port behavior of the resulting balanced amplifier.

In order to study the differential-mode and the common-mode behavior of balanced amplifiers, we will introduce a method for decomposing balanced networks into common-mode and differential-mode equivalent networks, that describe their common-mode behavior and differential-mode behavior, respectively.

CM and DM equivalent circuits

There are two ways of decomposing a balanced circuit into common-mode and differential-mode equivalent circuits.

1. Decomposition based on changing the base of the MNA matrix.

2. Decomposition based on network inspection.

We will illustrate both techniques with the aid of an example.

In the following example we will decompose a balanced circuit by changing of the base of the MNA matrix.

Example 7.11

The technique of changing the base of a the MNA matrix has been presented in Chapter 18.6.1. We will apply this technique for decomposing the balanced voltage amplifier from Figure 7.41A into two circuits: one that describes the differential-mode behavior, and another that describes the common-mode behavior. Figure

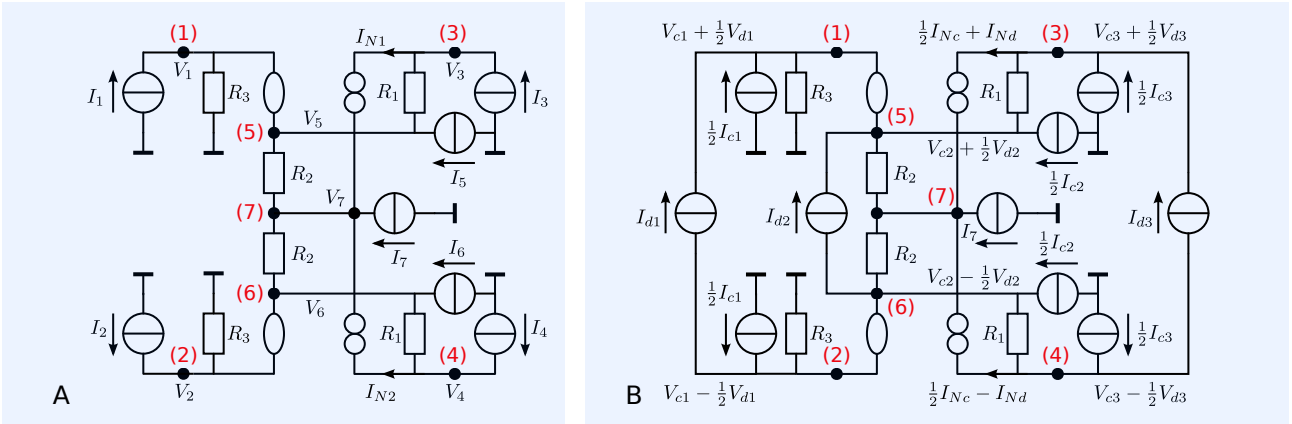


Figure 7.43: Balanced passive-feedback voltagre amplifier
A: Network for MNA analysis
B: Network with common-mode and differential-mode quantities.

7.43A shows the circuit for determination of the MNA matrix stamp of the complete amplifier. Figure 7.43B shows the definitions of the differential-mode and the common-mode currents and voltages.

The MNA matrix equation of the network from Figure 7.43A can be formulated as

$$\mathbf{I} = \mathbf{M}\mathbf{V}, \tag{7.33}$$

where $\mathbf{I}$ *is the vector with nodal currents and branch voltages:*

$$\mathbf{I} = \begin{pmatrix} I_1 & I_2 & I_3 & I_4 & I_5 & I_6 & I_7 & 0 & 0 \end{pmatrix}^T, \tag{7.34}$$

$\mathbf{M}$ *is the MNA matrix:*

$$\mathbf{M} = \begin{pmatrix} \frac{1}{R_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{R_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{R_1} & 0 & -\frac{1}{R_1} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{R_1} & 0 & -\frac{1}{R_1} & 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{R_1} & 0 & \frac{1}{R_1}+\frac{1}{R_2} & 0 & -\frac{1}{R_2} & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{R_1} & 0 & \frac{1}{R_1}+\frac{1}{R_2} & -\frac{1}{R_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{R_2} & -\frac{1}{R_2} & \frac{2}{R_2} & -1 & -1 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix}, \tag{7.35}$$

and $\mathbf{V}$ *is the vector with nodal voltages and branch currents:*

$$\mathbf{V} = \begin{pmatrix} V_1 & V_2 & V_3 & V_4 & V_5 & V_6 & V_7 & I_{N1} & I_{N2} \end{pmatrix}^T. \tag{7.36}$$

We would like to write this equation in the form

$$\mathbf{I}_{d,c} = \mathbf{M}' \cdot \mathbf{V}_{d,c}, \tag{7.37}$$

where $\mathbf{I}_{d,c}$ *is a vector with differential-mode and common-mode variables obtained from* $\mathbf{I}$ *,* $\mathbf{M}'$ *is the modified MNA matrix and* $\mathbf{V}_{d,c}$ *is the vector with differential-mode and common-mode variables obtained from* $\mathbf{V}$.[9]

These new vectors can be obtained by converting pairs of nodal voltages and branch currents into a common-mode, and differential-mode equivalents. Let us, for example convert the pair of nodal voltages (V_1, V_2) *into a differential-mode*

[9] $\mathbf{V}$ is the vector with nodal voltages and branch currents.

voltage V_d and a common-mode voltage V_c. We do this by writing

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 1 \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} V_d \\ V_c \end{pmatrix}. \tag{7.38}$$

Similarly, we can decompose a pair of nodal currents (I_1, I_2) into a differential-mode current I_d and a common-mode current I_c, by writing

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} I_d \\ I_c \end{pmatrix}. \tag{7.39}$$

We can now apply the above transformations to (7.36), and obtain

$$\mathbf{V} = \mathbf{A} \cdot \mathbf{V}_{d,c}. \tag{7.40}$$

The differential-mode quantities in $\mathbf{V}_{c,d}$ are grouped in the first rows, while the common-mode quantities are grouped in the last rows:

$$\mathbf{V}_{d,c} = \begin{pmatrix} V_{d1} & V_{d2} & V_{d3} & I_{Nd} & V_{c1} & V_{c2} & V_{c3} & I_{Nc} & V_7 \end{pmatrix}^T. \tag{7.41}$$

The base transformation matrix $\mathbf{A}$ then becomes

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}. \tag{7.42}$$

The vector $\mathbf{I}_{c,d}$ with the differential-mode and common-mode variables[10]*, can be obtained from $\mathbf{I}$. It can be shown that*

[10] The vector with nodal currents and branch voltages.

$$\mathbf{I}_{d,c} = \mathbf{A}^T \cdot \mathbf{I}, \tag{7.43}$$

where the differential-mode and the common-mode quantities in $\mathbf{I}_{d,c}$ are grouped as in $\mathbf{V}_{d,c}$:

$$\mathbf{I}_{d,c} = \begin{pmatrix} I_{d1} & I_{d2} & I_{d3} & 0 & I_{c1} & I_{c2} & I_{c3} & 0 & I_7 \end{pmatrix}^T. \tag{7.44}$$

The original MNA equation can now be rewritten in the form of (7.37):

$$\begin{aligned} \mathbf{I}_{d,c} &= \mathbf{A}^T \cdot \mathbf{I}, && (7.45)\\ &= \mathbf{A}^T \cdot (\mathbf{M} \cdot \mathbf{V}), && (7.46)\\ &= \mathbf{A}^T \cdot (\mathbf{M} \cdot (\mathbf{A} \cdot \mathbf{V}_{d,c})), && (7.47)\\ &= \left(\mathbf{A}^T \cdot \mathbf{M} \cdot \mathbf{A}\right) \cdot \mathbf{V}_{c,d}, && (7.48) \end{aligned}$$

from which we obtain $\mathbf{M}'$ as

$$\mathbf{M}' = \mathbf{A}^T \cdot \mathbf{M} \cdot \mathbf{A}. \tag{7.49}$$

Since $\det(\mathbf{A}) = 1$, we have $\det(\mathbf{M}') = \det(\mathbf{M})$. Hence, the base transformation does not change the characteristic equation of the network.

After substitution of $\mathbf{A}$ *and* $\mathbf{M}$ *in (7.49), we obtain*

$$\mathbf{M}' = \left(\begin{array}{cccc|ccccc} \frac{1}{2R_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2R_1}+\frac{1}{2R_2} & -\frac{1}{2R_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2R_1} & \frac{1}{2R_1} & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ - & - & - & - & + & - & - & - & - & - \\ 0 & 0 & 0 & 0 & \frac{2}{R_3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{2}{R_1}+\frac{2}{R_2} & -\frac{2}{R_1} & 0 & -\frac{2}{R_2} \\ 0 & 0 & 0 & 0 & 0 & -\frac{2}{R_1} & \frac{2}{R_1} & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{2}{R_2} & 0 & -1 & \frac{2}{R_2} \end{array}\right). \tag{7.50}$$

As a result of the grouping of the differential-mode quantities and the common-mode quantities in $\mathbf{I}_{c,d}$ *and in* $\mathbf{V}_{c,d}$*, the matrix* $\mathbf{M}'$ *can be split into four matrices:*

$$\mathbf{M}' = \begin{pmatrix} \mathbf{M}_{dd} & \mathbf{M}_{cd} \\ \mathbf{M}_{dc} & \mathbf{M}_{cc} \end{pmatrix}, \tag{7.51}$$

where

- $\mathbf{M}_{dd}$ *describes the differential-mode behavior*
- $\mathbf{M}_{cd}$ *describes the common-mode to differential-mode conversion*
- $\mathbf{M}_{dc}$ *describes the differential-mode to common-mode conversion*
- $\mathbf{M}_{cc}$ *describes the common-mode behavior.*

In this example, all of the coefficients of both $\mathbf{M}_{cd}$ *and* $\mathbf{M}_{dc}$ *are zero, because the circuit is perfectly balanced.*

The MNA matrix equation of the differential-mode equivalent network can be written as

$$\begin{pmatrix} I_{d1} \\ I_{d2} \\ I_{d3} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2R_3} & 0 & 0 & 0 \\ 0 & \frac{1}{2R_1}+\frac{1}{2R_2} & -\frac{1}{2R_1} & 0 \\ 0 & -\frac{1}{2R_1} & \frac{1}{2R_1} & 1 \\ 1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} V_{d1} \\ V_{d2} \\ V_{d3} \\ I_{Nd} \end{pmatrix}, \tag{7.52}$$

The equivalent differential-mode circuit is shown in Figure 7.44. It is of course similar to the circuit of one of the passive feedback voltage amplifiers of the anti-series connection.

The MNA matrix equation of the common-mode equivalent network can be written as

$$\begin{pmatrix} I_{c1} \\ I_{c2} \\ I_{c3} \\ 0 \\ I_7 \end{pmatrix} = \begin{pmatrix} \frac{2}{R_3} & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{R_1}+\frac{2}{R_2} & -\frac{2}{R_1} & 0 & -\frac{2}{R_2} \\ 0 & -\frac{2}{R_1} & \frac{2}{R_1} & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & -\frac{2}{R_2} & 0 & -1 & \frac{2}{R_2} \end{pmatrix} \begin{pmatrix} V_{c1} \\ V_{c2} \\ V_{c3} \\ I_{Nc} \\ V_7 \end{pmatrix}. \tag{7.53}$$

The common-mode equivalent circuit is shown in Figure 7.45.

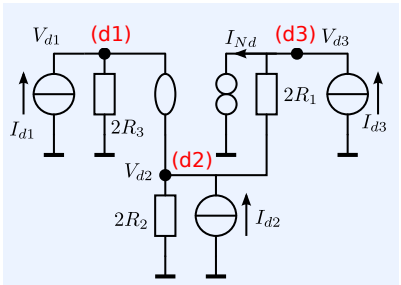


Figure 7.44: Differential-mode equivalent circuit of the balanced, passive-feedback voltage amplifier from Figure 7.43.

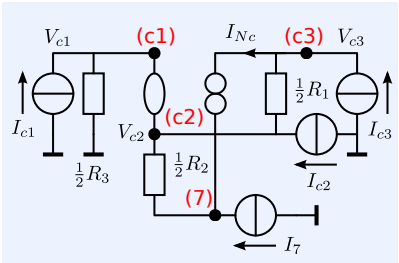


Figure 7.45: Common-mode equivalent circuit of the balanced, passive-feedback voltage amplifier from Figure 7.43.

In the following example, we will derive the differential-mode and the common-mode equivalent circuits through circuit inspection.

Example 7.12

Let us now consider the circuit from Figure 7.43B. We can derive a differential-mode equivalent circuit from it, by setting the common-mode currents to zero, and by redirecting the differential-mode currents via the node that connects both

amplifiers of the anti-series connection (node 7). If we also set $V_{c1} = 0$, the voltage at node 7 becomes zero and we may connect it to ground and take it as the reference node. In this way we obtain the circuit from Figure 7.44.

The common-mode equivalent circuit can be found by assuming all differential-mode currents and differential-mode voltages in the circuit from Figure 7.43B zero. In this way we obtain:

$$V_1 = V_2,\ V_3 = V_4,\ \textit{and}\ V_5 = V_6. \tag{7.54}$$

We then may connect the corresponding nodes of the two networks with each other, because there will be no current flow through these interconnections. By doing so, we obtain two identical networks of which all of their corresponding nodes have been interconnected. These two networks are the equivalent common-mode circuits. The circuit diagram of the common-mode equivalent circuit of the network from Figure 7.43 is shown in Figure 7.45.

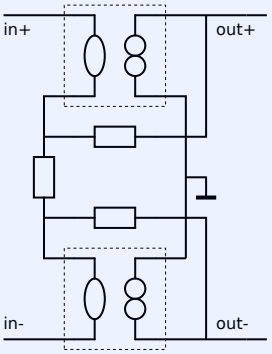


Figure 7.46: Balanced voltage amplifier with grounded norators.

CM behavior of single-loop balanced feedback configurations

Let us now study the common-mode behavior of the balanced, single-loop, passive-feedback configurations from Figure 7.41.

1. The balanced voltage amplifier in Figure 7.41A has unity common-mode voltage transfer, an infinite common-mode input impedance and, when driven from a finite common-mode source impedance, zero common-mode output impedance.

   The $CMRR$ of this amplifier concept is infinite, because the common-mode to differential-mode voltage transfer is zero and does not depend on the values of the feedback elements. The differential-mode to common-mode transfer depends on the values of the feedback elements. If the circuit is perfectly balanced, it equals zero.

   This circuit has none of the terminals of the nullors connected to the ground. This makes it impossible to replace them with operational amplifiers at a later stage of the design. However, the position of the norators can be changed in such a way that such replacement becomes possible. This is shown in Figure 7.46. Such a change, however, may adversely affect the $CMRR$.

2. The balanced transadmittance amplifier in Figure 7.41B has natural two-port behavior. It can be simplified by combining the two feedback elements as shown in Figure 7.47. This circuit has none of its norator terminals connected to ground. This makes it impossible to replace the nullors with operational amplifiers.

3. The balanced transimpedance amplifier in Figure 7.41C has non-unilateral common-mode behavior. The common-mode input impedance depends on the load impedance and the common-mode output impedance depends on the common-mode source impedance. The circuit can be simplified by replacing the two nullors with a single nullor, as shown in Figure 7.48.

4. The balanced current amplifier in Figure 7.41D exhibits natural two-port behavior. It has none of its norator terminals connected to ground. This makes it impossible to replace the nullors with operational amplifiers.

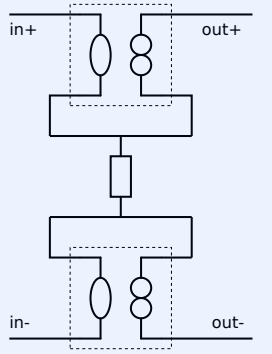


Figure 7.47: Balanced transadmitance amplifier.

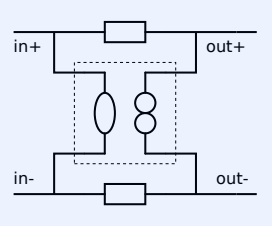


Figure 7.48: Simplified balanced transimpedance amplifier.

Common-mode feedback

The common-mode transfer parameters of an amplifier can be designed separately from its differential-mode transfer parameters. This can be done through application of both differential-mode feedback and common-mode

feedback. However, as a result of the lack of port isolation in passive feedback networks, the common-mode behavior of passive-feedback amplifiers cannot always be designed independently from their differential-mode behavior.

The procedure for designing the differential-mode behavior and the common-mode behavior independently is:

1. Design the differential-mode behavior in such a way that the amplifier behaves as a natural two-port.

2. Design the common-mode behavior with the aid of common-mode feedback loops.

In the following section we will give a brief discussion of the design of balanced amplifiers that show natural two-port behavior.

In section 7.7.6, we will discuss the application of negative feedback for design of common-mode impedances. In this section, we will also introduce techniques for the correction of the common-mode impedance of balanced amplifiers that do not exhibit natural two-port behavior.

7.7.4 Design of natural two-ports

In section 7.7.1, we introduced anti-series connection of unbalanced amplifiers as basic principle for the design of balanced amplifiers. We have seen that anti-series connection of two unbalanced passive feedback amplifiers does not necessarily result in amplifiers that behave as natural two-ports. In many cases such behavior is desired because it guarantees zero influence of common-mode disturbances on the quality of the (differential-mode) signal transfer, and it allows for significant flexibility in the drive and termination conditions of the input and output port, respectively. In this section, we will introduce design techniques for designing balanced amplifiers that behave as natural two-port behavior.

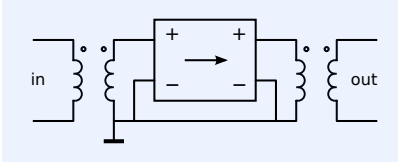


Figure 7.49: Creation of natural two-port behavior with the aid of transformers.

Application of transformers

One approach to the design of floating input ports is to use an cascade connection of a passive natural two-port element and an unbalanced amplifier. The only available practical component that approaches natural two-port behavior is the transformer. Figure 7.49 shows an arrangement in which input and output transformers have been used to convert a unbalanced amplifier into a natural two-port. This method is not possible for amplifiers that need to process signals with frequency components down to zero.

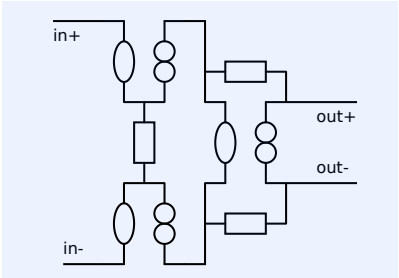


Figure 7.50: Concept of a balanced voltage amplifier that exhibits natural two-port behavior.

DC-coupled natural two-port amplifiers

Natural two-port behavior for amplifiers that have to process frequencies down to zero, can be designed using balanced amplifiers configurations that exhibit natural two-port behavior. Candidates are the balanced transadmittance amplifier and the balanced current amplifier from Figure 7.41B and 7.41D, respectively.

The natural two-port character of the balanced transadmittance amplifier makes it an ideal building block for designing amplifiers that behave as natural two-ports. It also owes its importance to the ease of implementation with the aid of three-terminal amplifying devices such as MOS transistors, BJTs and vacuum tubes.

Figure 7.50 shows the concept of a voltage amplifier with natural two-port behavior. It consists of a cascade connection of a balanced transadmittance amplifier and a balanced and transimpedance amplifier. The rejection factor of this amplifier can be much lower than that of the widespread used configuration from Figure 7.46. The CMRR can be very high.

Balanced active feedback

Figure 7.51 shows the concept of a voltage amplifier with natural two-port behavior that utilizes active feedback with a balanced transadmittance amplifier. The amplifier topology from Figure 7.51 is applied in the Analog Devices AMP01 instrumentation amplifier integrated circuit. The voltage gain of this circuit can be found as

$$\frac{V_\ell}{V_s} = \frac{Z_2}{Z_1}. \tag{7.55}$$

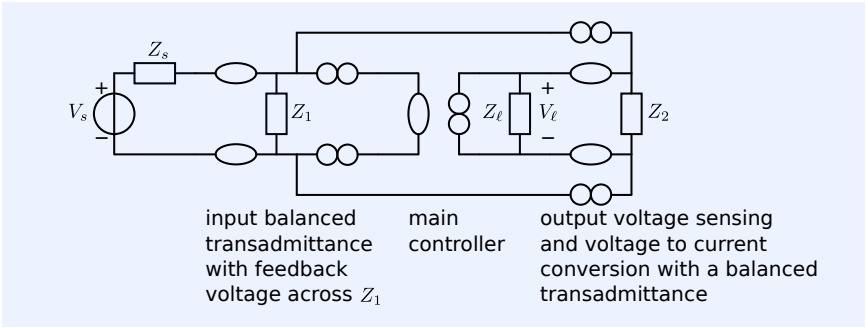


Figure 7.51: Floating port voltage amplifier with active feedback.

7.7.5 Dual-loop balanced passive feedback amplifiers

In balanced passive-feedback amplifiers, inversion in the passive feedback network can be realized with the aid of a cross-coupling in a feedback network. Some balanced dual-loop configurations that use this technique are shown in Figure 7.52. The reader is invited to evaluate their ideal gain and estimate the influence of the feedback network on the noise behavior of the amplifiers.

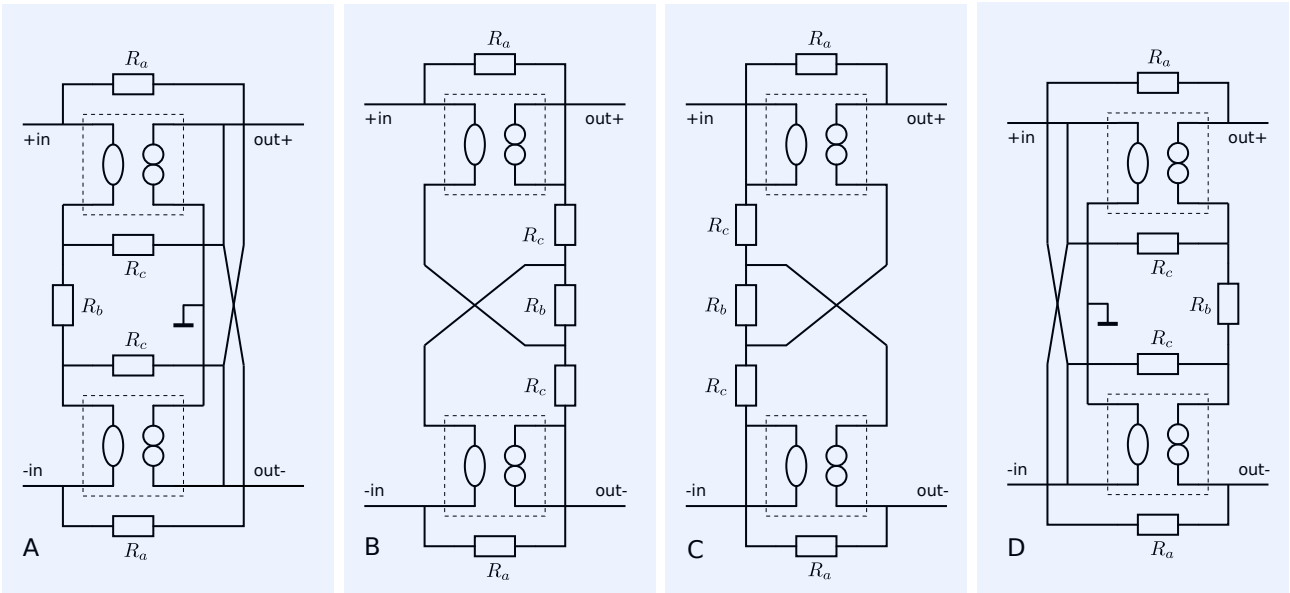


Figure 7.52: Balanced dual-loop feedback configurations with fixed input impedance (A, B) and fixed output impedance (C, D)
A: $Z_{in} = \frac{A}{C}$, $Z_{out} = 0$
B: $Z_{in} = \frac{B}{D}$, $Z_{out} = \infty$
C: $Z_{out} = \frac{B}{A}$, $Z_{in} = \infty$
D: $Z_{out} = \frac{D}{C}$, $Z_{in} = 0$.

CM behavior of dual-loop balanced amplifiers

Although cross-coupling in a feedback loop can be used to change the sign of a differential-mode transmission parameter, it does not affect the corresponding common-mode transmission parameter. In general, the common-

mode transmission-1 parameters of a balanced amplifier will differ from its differential-mode transmission-1 parameters. However, in many cases the common-mode impedance does not need to have an accurately defined value. Its value should be such that common-mode noise and interference will not cause unacceptable large values of the common-mode voltage at the port and the common-mode current through it.

The analysis of the common-mode port impedance can be done using a common-mode equivalent circuit, which can be obtained through circuit decompostion, as discussed in section 7.7.3. In the following section we will demonstrate such a decomposition for a balanced circuit that has cross-coupled feedback. The design of accurate common-mode port impedances using common-mode feedback will be discussed in section 7.7.6.

Decomposition of balanced amplifiers with cross-coupled feedback

In section 7.7.3, we have studied two ways for the decomposing balanced amplifiers into common-mode and differential-mode equivalent circuits. Transformation of the base of the MNA matrix, has been presented as a straightforward method that can easily be automated. In the following example we will demonstrate the method based upon circuit inspection, and illustrate the way in which a cross coupling in the feedback network can be dealt with.

Example 7.13

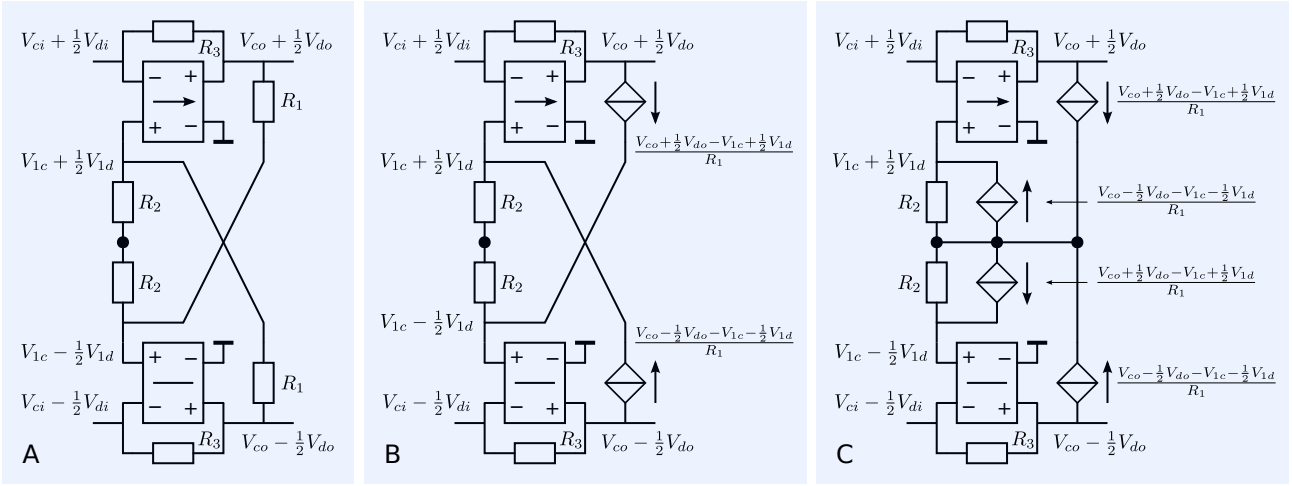


Figure 7.53: Decomposition of a balanced dual-loop feedback amplifier with cross-coupled feedback, into a common-mode and a differential-mode equivalent circuit.

A: Dual-loop passive feedback amplifier with A and C fixed by passive-feedback.

B: Application of substitution theorem: replace R_1 with a current source that carries the current of R_1.

C: Redirect the currents of the controlled sources over the common node of the anti-series connection.

Let us consider the balanced amplifier from Figure 7.53A. This amplifier has its differential-mode input impedance fixed to a positive value that is defined by the differential-mode transmission-1 parameters A and C. The common-mode input impedance, however, is infinite.

In order to investigate both the differential-mode and the common-mode behavior of this amplifier, we will decompose the circuit into a common-mode and a differential-mode equivalent circuit. In this example we will demonstrate the decomposition based upon circuit inspection.

In Figure 7.53B the substitution theorem has been applied to replace the cross-coupled resistor with a controlled current source. The controlled sources can then be redirected via the interconnection node of the anti-series connection (the black dot in the figure).

The result is shown in Figure 7.53C. The circuit can then be decomposed into a differential-mode equivalent circuit and a common-mode equivalent circuit as discussed in example 3.7.12. The result of this decomposition is shown in Figure 7.54.

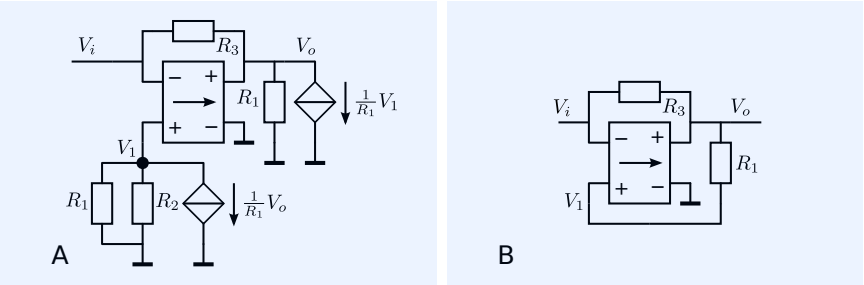


Figure 7.54: Result of the decomposition of the amplifier from Figure 7.53.
A: Differential-mode equivalent circuit
B: Common-mode equivalent circuit.

7.7.6 Design of common-mode port impedances

In the previous sections, we have seen that the common-mode behavior of balanced passive-feedback amplifiers cannot always be designed independently from their differential-mode behavior. This is because natural two-ports cannot be constructed using exclusively passive feedback elements. If elements that show natural two-port behavior can be used in combination with passive-feedback elements, independent design of common-mode and differential-mode behavior may become possible.

Figure 7.55 shows the concept of an amplifier that has its transmission parameters A and C fixed by means of negative feedback, both for common-mode and for differential-mode operation. The differential-mode input impedance and the common-mode input impedance of this amplifier both equal $\frac{A}{C} = \frac{R_f}{1+n}$. The analysis is left as an exercise to the reader.

Figure 7.56 shows a balanced passive-feedback amplifier of which the values of A and C have been fixed, both for common-mode and for differential-mode. To this end, common-mode feedback with two transadmittance amplifiers has been added to the balanced amplifier from Figure 7.52A.

The reader is invited to determine the values of the transmission-1 parameters, both for common-mode and for differential-mode operation.

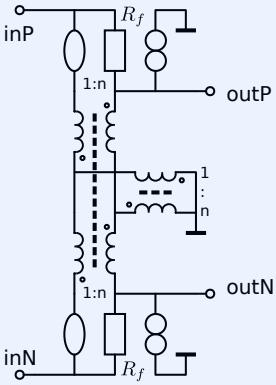


Figure 7.55: Concept of a low-noise amplifier of which the differential-mode input impedance and the common-mode input impedance equal $\frac{r_f}{1+n}$, and the differential-mode and the common-mode output impedance equal zero.

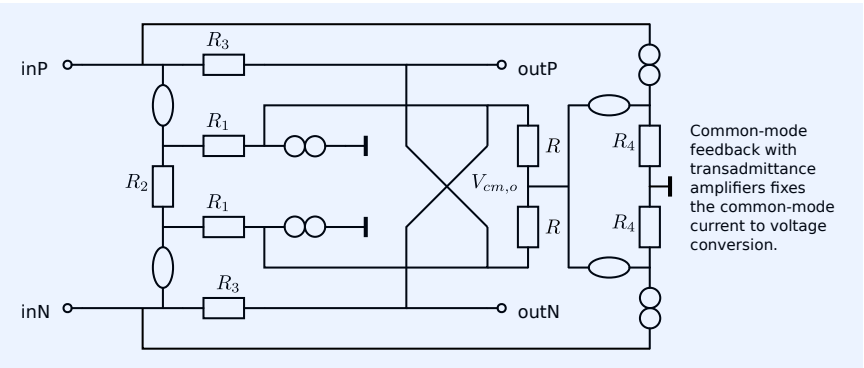


Figure 7.56: Balanced amplifier that has its common-mode and differential input impedance fixed by means of negative feedback.

The use of passive-feedback for establishing the common-mode impedance, may adversely affect the signal-to-noise ratio of the balanced amplifier. In the amplifier from Figure 7.56, the output current noise of the two transadmittance amplifiers is uncorrelated, and contributes to the differential-mode equivalent input current noise of the balanced amplifier.

7.8 Indirect feedback

Indirect feedback or model-based feedback uses indirect sensing of the load quantity or indirect comparison of the feedback quantity with the source

quantity. Compared with direct sensing techniques, indirect sensing and comparison techniques usually results in less accurate feedback amplifiers. This is because variations in the values of the source and/or the load impedance is not always accurately observed by indirect feedback.

7.8.1 Indirect sensing

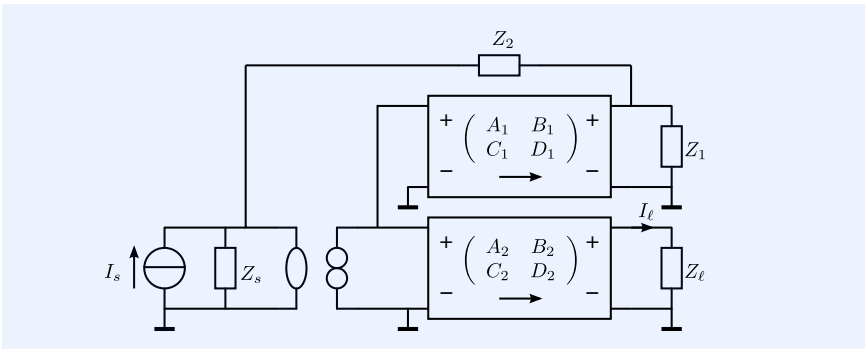


Figure 7.57: Inverting current amplifier using indirect feedback.

The circuit from Figure 7.57 shows an example of indirect current sensing in an indirect feedback current amplifier. The load and the feedback network are driven from two two-ports whose inputs have been connected in parallel. If those two-ports have $A_1 = A_2$ and $B_1 = B_2$, and if the parallel connection of Z_2 and Z_2 equals Z_ℓ, the transfer is solely determined by Z_1 and Z_2. This is expressed by (7.56):

$$\frac{I_\ell}{I_s} = \frac{Z_1 + Z_2}{Z_1}\left(\frac{A_1\frac{Z_1Z_2}{Z_1+Z_2} + B_1}{A_2Z_\ell + B_2}\right). \tag{7.56}$$

7.8.2 Indirect comparison

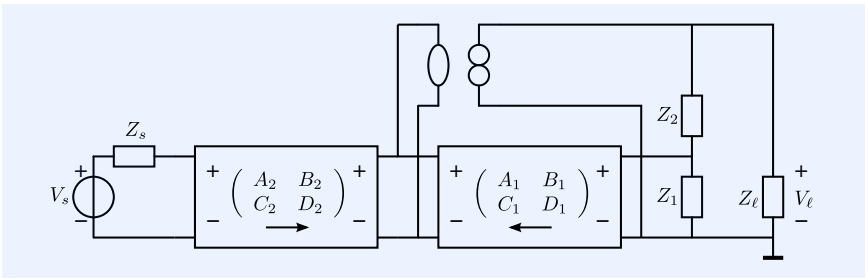


Figure 7.58: Voltage amplifier using indirect voltage comparison.

The circuit from Figure 7.58 shows an example of indirect voltage comparison in an indirect feedback voltage amplifier. The source voltage is converted into a current and then compared with the feedback current that is derived from the load voltage in a similar way. The difference between the two currents is nullified. As with indirect current sensing, the source-to-load voltage transfer is solely determined by Z_1 and Z_2 if their parallel connection equals Z_s and if $B_1 = B_2$ and $D_1 = D_2$. This follows from (7.57)

$$\frac{V_\ell}{V_s} = \frac{Z_1 + Z_2}{Z_1}\left(\frac{B_1 + D_1\frac{Z_1Z_2}{Z_1+Z_2}}{B_2 + D_2Z_s}\right). \tag{7.57}$$

8 Application and specification of operational amplifiers

Figure 8.1: K2-W commercial vacuum tube operational amplifier (1952).
Photo by courtesy of 'Studieverzameling Elektrotechniek', Delft University of Technology.

8.1 Introduction

Operational amplifiers (OpAmps) are high-gain voltage amplifiers that have a differential input and usually a single-ended output and that operate from DC. They were originally developed as controllers in feedback circuits that performed mathematical operations in analog computers.

Before integrated circuit operational amplifiers became available, there already existed operational amplifiers with vacuum tubes (Figure 8.1), and operational amplifier modules constructed with discrete semiconductors.

The first integrated-circuit operational amplifier, realized in a bipolar IC technology, was the uA702 (Bob Widlar, Fairchild 1964). Since then, JFET (1970s), MOSFET (1980s) and BiMOS operational amplifiers have become available, and their performance is continuously improved.

The operational amplifier is a general-purpose amplifier, intended to be used as *loop amplifier* or *controller* in circuits that exploit negative feedback. Its intended function thus differs from the one given in section 2.1.1. For its application as loop amplifier or controller , it needs to provide a huge gain over a wide frequency range, with less than 180 degrees of phase shift. This will become clear after studying the chapter on the modeling of negative feedback amplifiers (Chapter 10).

In modern embedded systems (systems that incorporate a microprocessor), operational amplifiers are mainly used in applications such as:

- Loop amplifiers in analog active filters
- Input amplifiers for amplification of low-level signals before A/D conversion can take place
- D/A converter output amplifiers for driving actuators and transmission lines
- Voltage buffers, summing amplifiers, etc.

8.1.1 Operational Amplifier types

Although the operational amplifier was originally intended to be a general-purpose building block, a huge number of versions have become available. Obviously, the operational amplifier is not particularly 'general-purpose' at all, and the immense diversity of its application justifies this apparently unlimited number of device types. Manufacturers of operational amplifiers provide the designers with selection tables to narrow the search for devices.

Another consequence of this trend is that data sheets of operational amplifiers often specify the devices for typical applications. If the application to be designed strongly deviates from this typical application, simulation and breadboarding are required to ensure performance to requirements.

Architecture

Integrated circuit operational amplifiers have been designed such that a positive current flow through the load (output source current) is delivered by the positive power supply and the negative current flow through the load (output sink current) is delivered by the negative power supply. This is shown in Figure 8.2. A small and almost signal-independent quiescent current flows through the operational amplifier from the positive to the negative power supply terminal (Class AB output stages).

There exist different architectures for operational amplifiers:

1. Voltage-feedback operational amplifiers

Voltage-feedback operational amplifiers have a balanced input stage that converts the differential input voltage into a current that drives an output stage with a transimpedance character. The output current of this type of input stage is limited. Parasitic capacitances, as well as the limited current-drive capability of the input stage, may seriously limit the rate of change of the output voltage (slew-rate). Since the input stage of voltage-feedback operational amplifiers is well-balanced, the input offset voltage and the offset current of these operational amplifiers can be very low. This makes these amplifiers very well suited for applications in which a high low-frequency (DC) accuracy is required.

Voltage-feedback operational amplifiers may have a rail-to-rail input stages. In fact, such amplifiers have two input stages in parallel. One operates at common-mode voltages up to the positive supply voltage and the other down to the negative supply voltage. Due to this architecture, the behavior of these devices may change with the common-mode input voltage.

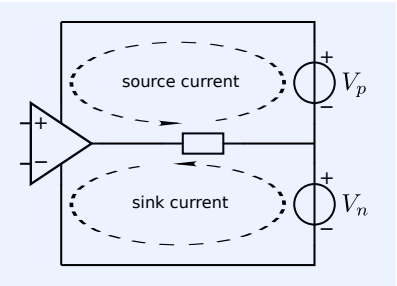


Figure 8.2: Current flow in the output port of the operational amplifier.

2. Current-feedback operational amplifiers

   Current feedback operational amplifiers have a push-pull class AB input stage with a high-current-drive capability. This input stage drives an output stage with a transimpedance character. Due to the high current-drive capability of the input stage, current-feedback operational amplifiers can have a very high rate of change of the output voltage. This makes them very well suited for high-speed applications. The input stage of the current-feedback operational amplifier, however, is strongly asymmetrical. This results in a relatively large offset voltage and offset current when compared to a voltage-feedback type.

3. Auto-zero operational amplifiers

   Auto-zero techniques are often applied for compensation of offset (zero) errors. Auto-zero operational amplifiers have very low offset and are mainly intended for low-frequency applications in which DC accuracy and temperature stability are of utmost importance.

4. Fully-differential operational amplifiers

   Relatively new to the family of operational amplifiers are the so-called fully-differential operational amplifiers. They are mainly intended for driving high-speed differential-input ADCs in digital radio and high-speed instrumentation systems with sampling frequencies in the GHz range. Fully differential operational amplifiers usually offer a limited design flexibility.[1]

[1] They are designed for a specific structure of the feedback network of which component values should be within a limited range.

8.1.2 Idealized models

As mentioned earlier, operational amplifiers are intended to be used as loop amplifiers or controllers in negative-feedback circuits. In Chapter 7 we introduced the nullor as the ideal controller. The operational amplifier is a practical implementation of the nullor. In fact, an operational amplifier can be regarded as a high-gain voltage-controlled voltage source, which is one of the possible controlled source approximations of the nullor.

Single-ended output operational amplifier

Figure 8.3 shows the nullor, the VCVS approximation of the nullor and the implementation with a single-ended output operational amplifier. One of the terminals of the norator is connected to the power supply terminals. In other words, the *return path* for the output current is the power supply.

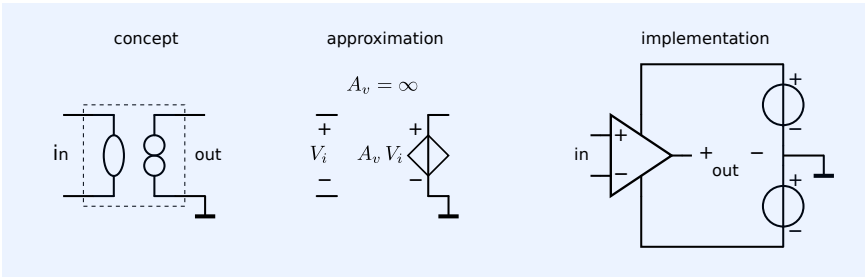


Figure 8.3: The operational amplifier can be regarded as an implementation of the nullor. One of the terminals of the norator is connected to the supply ground. Please notice that the nullator does not have inverting and non-inverting terminals. See Chapter 18.3 for the network equations of the nullor.

Fully differential operational amplifiers

Figure 8.4 shows an idealized model for a fully-differential operational amplifier. These amplifiers usually have a control input for the common-mode output voltage.

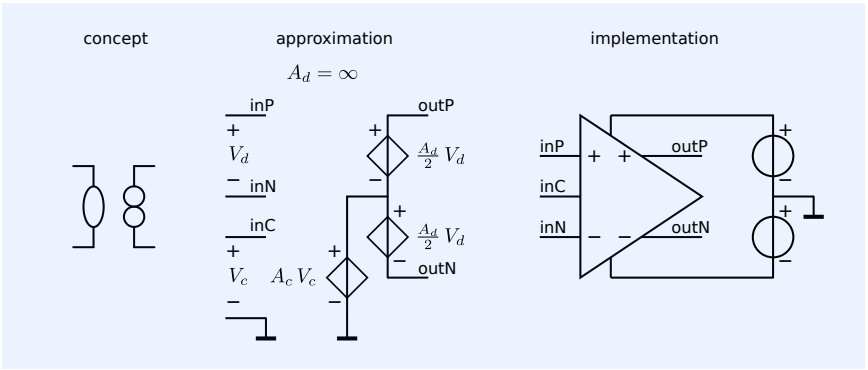


Figure 8.4: Implementation of the nullor with a fully-differential operational amplifier.

8.1.3 This chapter

In section 8.2, we will briefly discuss the characterization of operational amplifiers. Most of the parameters that describe the characteristics of these devices have already been introduced in Chapter 2.

In section 8.3, we will discuss the modeling of operational amplifiers. Although it seems attractive to use so-called macro models for numeric simulation with SPICE, we will not pay much attention to this. This is because manufacturers of operational amplifiers still encourage the use of data sheets for design, and prototyping for design verification. Moreover, not all aspects have been modeled correctly in these macro models. Verification of the macro models with test circuits that correspond to those given in the data sheets , is indispensable for reliable design.

Instead of working with one complete model for numeric simulation, we will pay attention to the modeling of individual performance aspects, both symbolically and numerically. This gives the designer much more control over the level of complexity of the models used at different stages of design.

8.2 Characterization of operational amplifiers

Operational amplifiers are often specified with the aid of a collection of parameters and plots. In this section, we will define the most commonly used parameters and list some frequently used plots. The reader is invited to

download some data sheets of operational amplifiers and study them to better understand the specification of these components.

8.2.1 Commonly used terms

As mentioned earlier, operational amplifiers are almost always used as controllers in feedback circuits. Hence, they are amplifiers and should be specified as such, as was discussed in Chapter 2. However, in common design practice , parameter names can be somewhat confusing. This is because the specification of operational amplifiers is very much connected to the design of negative-feedback (voltage) amplifiers. In data sheets of operational amplifiers , one often finds the terms *closed-loop* gain and *open-loop* gain. In this book, we try to avoid these terms. We speak rather of the *gain* of the operational amplifier, the *gain* of a negative feedback amplifier and the *loop gain* of a negative feedback amplifier. These terms have a well-defined meaning in the asymptotic gain model , that will be used for the design and analysis of negative feedback circuits (see Chapter 10).

The popular term *open-loop* gain in data sheets refers to the *gain* of the operational amplifier. It is a property of the operational amplifier itself and it does not need a feedback loop for its existence, nor for its definition. This also holds for the term *open-loop* output impedance. It is simply the output impedance of the operational amplifier itself.

The term *closed-loop* gain in data sheets refers to the *gain* of a *feedback amplifier* equipped with the operational amplifier. Hence, it is a property of the circuit with the operational amplifier, rather than a property of the operational amplifier itself. This is also the case for the term *closed-loop* output impedance.

In Chapter 10, we will discuss the analysis of circuits with feedback. We will then relate the gain of the negative feedback amplifier to its ideal gain as it is defined in Chapter 7, and to the gain of the operational amplifier.

8.2.2 Terminal voltages and currents

name	**symbol**	**definition**
common-mode input voltage	V_{cm}	$\frac{1}{2}(V_{i+} + V_{i-})$
differential-mode input voltage	V_{dm}	$V_{i+} - V_{i-}$
bias current	I_{BIAS}	$I_{i+},\ I_{i-}$
common-mode input current	I_{cm}	$I_{i+} + I_{i-}$
differential-mode input current	I_{dm}	$\frac{1}{2}(I_{i+} - I_{i-})$
output voltage	V_{out}	V_{out}
output current	I_{out}	I_{out}
supply current (no load)	I_{supply}	$\frac{1}{2}(I_P - I_N)$

Table 8.1: Definition of the operational amplifier's terminal voltages and currents

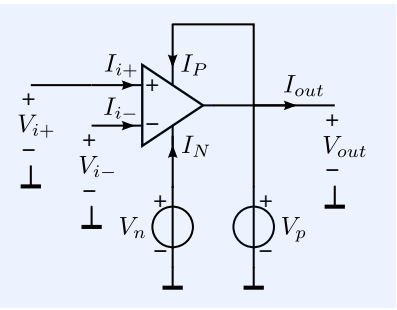


Figure 8.5: Definitions of the terminal voltages and currents of the operational amplifier.

Figure 8.5 shows the definitions of the terminal voltages and currents of the operational amplifier. Their general names and definitions are given in the Table 8.1. These names and definitions will be used throughout this chapter.

In the following sections, we will discuss parameters that describe the:

1. Static nonlinear behavior of operational amplifiers (section 8.2.3)

2. Small-signal dynamic behavior and noise behavior of operational amplifiers (section 8.2.4)

3. Large-signal dynamic behavior of operational amplifiers (section 8.2.5).

8.2.3 Static nonlinear behavior

The parameters that characterize the static nonlinear behavior (also large-signal instantaneous behavior) are given in Table 8.2.

name	symbol	definition
maximum positive V_{cm}	V_{cm+}	$V_{cm} < V_{cm+}$ for operation within specifications
maximum negative V_{cm}	V_{cm-}	$V_{cm} > V_{cm-}$ for operation within specifications
offset voltage	V_{off}	standard deviation of the differential-mode input voltage for $V_{out} = 0$
offset voltage drift	dV_{off}/dT	change of V_{off} with temperature
bias current drift	dI_{BIAS}/dT	change of I_{BIAS} with temperature
offset current	I_{off}	standard deviation of the differential-mode input current for $V_{out} = 0$
offset current drift	dI_{off}/dT	change of I_{off} with temperature
quiescent supply current (no load)	I_Q	$(I_+ - I_-)/2$
maximum positive V_{out} (no load)	V_{out+}	positive clipping value of V_{out}
maximum negative V_{out} (no load)	V_{out-}	negative clipping value of V_{out}
maximum output source current	I_{out+}	maximum positive value of I_{out}
maximum output sink current	I_{out-}	maximum negative value of I_{out}

Table 8.2: Parameters that describe the static nonlinear behavior of the operational amplifier

In many cases, manufacturers of operational amplifiers add graphs to these lists of parameters. These graphs show:

1. Temperature dependencies (i.e., bias current versus temperature or versus the common-mode input voltage)
2. Transfer characteristics (i.e., offset voltage versus the common-mode input voltage)
3. Statistical information (i.e., histogram of the offset voltage).

There are three effects that cause limitation of the low-frequency (static) signal handling capability of operational amplifiers: limitation of the common-mode input voltage range , limitation of the output current and limitation of the output voltage of the operational amplifier. The latter two are related to each other.

1. Limitation of the common-mode input voltage range of the operational amplifier:

   The common-mode input voltage range of operational amplifiers is usually smaller than the total supply voltage $V_P - V_N$. Only operational amplifiers with *rail-to-rail* inputs have a common-mode input voltage range equal to or (slightly) larger than the total power supply voltage. When driven beyond these limits, some operational amplifiers show *phase reversal* , an effect that may result in a response, as has been shown in Figure 2.48.

2. Limitation of the low-frequency output current and voltage handling capability:

   The maximum current that an operational amplifier can source or sink, depending on the load voltage. A typical plot of the voltage and current drive capability is shown in Figure 8.6.

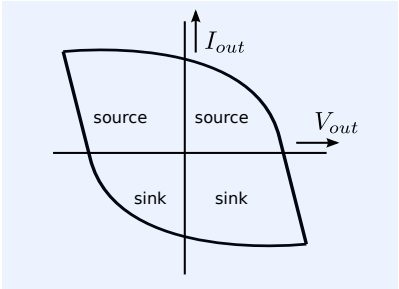


Figure 8.6: Voltage and current handling capability of an operational amplifier.

8.2.4 Noise and small-signal dynamic behavior

The small-signal voltage transfer of the operational amplifier is usually characterized with the aid of Bode plots of the gain and a small-signal step response of an application. Amplitude and phase characteristics of the operational amplifier's voltage transfer are indispensable for the design of stable negative feedback applications.

Modern rail-to-rail output operational amplifiers do not have a negligibly small output impedance. Bode plots of the output impedance of a feedback circuit with the operational amplifier are sometimes provided and may be used to estimate this impedance. In many applications, knowledge of this output impedance is indispensable for the design of the high-frequency stability of a circuit. Unfortunately, in many cases, the output impedance of an operational amplifier is not or incompletely specified.

name	**symbol**	**definition**
common-mode input impedance (complex)	Z_{cm}	dV_{cm}/dI_{cm}
differential-mode input impedance (complex)	Z_{dm}	dV_{dm}/dI_{dm}
voltage gain	A_{dm}	dV_{out}/dV_{dm}
output impedance	Z_{out}	$-dV_{out}/dI_{out}$
gain-bandwidth product	GB	unity-gain frequency
spectral density of input noise voltage	$S_{V_{inoise}}$	in series with input
spectral density of input noise current	$S_{I_{inoise}}$	parallel with input
common mode rejection ratio	$CMRR$	$dV_{icm}/dV_{idm}\big\vert_{V_{out}=0}$
positive power supply rejection ratio	$PSRR+$	$dV_p/dV_{idm}\big\vert_{V_{out}=0}$
negative power supply rejection ratio	$PSRR-$	$dV_n/dV_{idm}\big\vert_{V_{out}=0}$

Table 8.3: Parameters that describe the small-signal behavior and the stationary noise behavior of the operational amplifier.

Plots of the frequency-dependent spectral densities of the equivalent input voltage and current noise sources are almost always provided.

The parameters that describe the small-signal dynamic behavior and the stationary noise behavior of the operational amplifier are listed in Table 8.3.

8.2.5 Large-signal dynamic behavior

name	**symbol**	**definition**
positive slew rate	SR^+	maximum positive rate of change of V_{out}
negative slew rate	SR^-	maximum negative rate of change of V_{out}
full-power bandwidth	f_{fp}	maximum frequency for sine wave with amplitude $(V_{out+} - V_{out-})/2$
harmonic distortion	THD	see definition in Chapter 2
intermodulation distortion	IM	see definition in Chapter 2
differential gain		see definition in Chapter 2
differential phase		see definition in Chapter 2

Table 8.4: Parameters that describe the dynamic nonlinear behavior of the operational amplifier

Parameters that describe the large-signal dynamic behavior of operational amplifiers are listed in Table 8.4. Aside from these parameters, graphs of pulse responses in typical applications are often provided.

8.3 Modeling of the operational amplifier

The ongoing decrease of component dimensions and the development of high-density multi-layer printed circuit boards have lead to the development of complex printed circuit board assemblies. As a consequence, the design and production of a small number of prototypes of such PCAs have become relatively expensive. This forces designers to reduce the design risks with the aid of different design verification methods. In general, there are three design verification methods available to the designer:

1. Numerical simulation with accurate models that incorporate statistical in-

formation

Macro models for operational amplifiers are provided by IC manufacturers. Statistical information, however, is seldom incorporated to these models. Moreover, not all behavioral aspects that could be relevant to the design may have been modeled correctly. For this reason, the designer should investigate whether the measurement results obtained with test circuits described in the data sheet comply with the simulation results using a similar test setup. The use of macro models will be discussed in section 8.3.6

2. Circuit modeling and symbolic analysis

   During the design of circuits with operational amplifiers , the designer has to investigate the influences of various performance aspects of the operational amplifiers on the behavior of the complete circuit. If the type number of the operational amplifier is known, this can be done with the aid of numerical simulation using macro-models.

   However, such simulations can only be performed after a device has been selected. Selecting a device, however, is a design decision that needs to be motivated. Searching, selecting and evaluating devices without clear performance criteria easily turns into a time-consuming trial and error scenario.

   For the formulation of clear search criteria, we need to derive and solve the so-called design equations. These equations relate the performance parameters of a device to those of the circuit that comprises that device. Derivation of the design equation requires symbolic analysis. This required models that model the performance aspects of interest as simple as possible (but not too simple). Modeling techniques for various behavioral aspects will be discussed in the following sections.

   We will discuss the modeling of the following behavioral aspects:

   (a) Small-signal dynamic behavior (section 8.3.1)

   (b) Noise behavior (section 8.3.2)

   (c) PSRR and the CMRR (section 8.3.3)

   (d) Bias errors , offset and temperature effects (section 8.3.4)

   (e) Other aspects such as the current and the voltage drive capability and slew rate (section 8.3.5)

3. Conducting experiments

   In many cases, elimination of the most important design risks does not require the evaluation of the complete product. The manufacturing and evaluation of relatively small *proof of concepts* , designed to eliminate specific design risks at an early stage of the design project, strongly facilitates the first-time-right design of the complete product.

8.3.1 Small-signal dynamic behavior

Small-signal dynamic behavioral model for voltage-feedback opamps

Figure 8.7 shows the symbol and the small-signal dynamic equivalent model of a voltage-feedback operational amplifier. An overview of the model parameters is given in Table 8.5. This model suffices for evaluation of the small-signal frequency response, the small-signal step response, and the small-signal impulse response of circuits with operational amplifiers.

Name	Description
g_d	Differential-mode input conductance
C_d	Differential-mode input capacitance
g_c	Common-mode input conductance
C_c	Common-mode input capacitance
$A_v(s)$	Laplace transform of the voltage gain
$Z_o(s)$	Laplace transform of the output impedance

Table 8.5: Parameters of the small-signal model of the voltage-feedback operational amplifier.

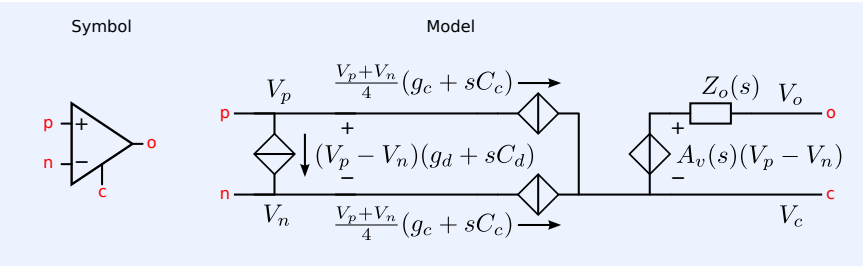


Figure 8.7: Simplified small-signal model of a voltage-feedback operational amplifier.

Common-mode rejection , power supply rejection , noise , bias current and offset current and voltage , as well as temperature effects , are not modeled. Errors due to these performance limitations will be modeled at a later stage.

Figure 8.8 shows the test benches for determination of the voltage gain and the output impedance of a voltage-feedback operational amplifier. The DC voltage of the signal source V_s equals zero. For very low frequencies (at DC), the circuit is a voltage follower , hence the DC output voltage equals zero. A DC error voltage at the output due to the DC current in the noninverting input of the operational amplifier can be kept low by taking R as low as possible. If necessary, the DC error voltage can be tuned to zero with the aid of V_s, or by adjusting the DC current source `I1`, which provides the current for the inverting input. The capacitor `C1` should eliminate the voltage feedback over the frequency range of interest. The conditions for proper measurement are:

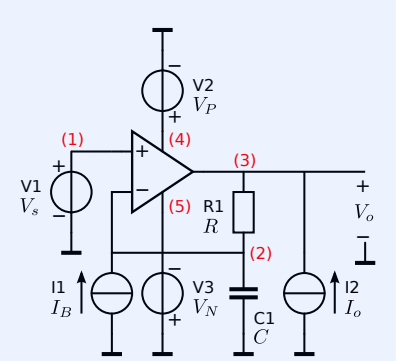


Figure 8.8: Test circuit for determination of the voltage gain and the output impedance of a voltage-feedback operational amplifier. $A_v = \frac{V_o}{V_s}$, $Z_o = \frac{V_o}{I_o}$

$$RC \gg \frac{A_v(0)}{2\pi f_{\min}}, \tag{8.1}$$

$$R \gg R_{out}, \tag{8.2}$$

where $A_v(0)$ is the expected maximum DC voltage gain of the operational amplifier, $f_{\min}$ is the lowest frequency of interest (it cannot be zero) and R_{out} is the expected maximum value of the output resistance of the operational amplifier. The voltage gain $A_v(s)$ is obtained as:

$$A_v(s) = \frac{V_o}{V_s}, \tag{8.3}$$

$$Z_o(s) = \frac{V_o}{I_o}. \tag{8.4}$$

Below is the syntax for the circuit from Figure 8.8, equipped with the LT1677 operational amplifier:

```
A_v_Z_o
* file: A_v_Z_o.cir
* LTSpice netlist file
* Test bench for A_v and Z_o of a
* voltage-feedback operational amplifier
* Default settings:
*** measurement of A_v
*** +/-5V supply voltage
```

```
*** LT1677 operational amplifier
.param vs=1          ;change to vs=0 for measurement of Z_o
.param io=0          ;change to io=1 for measurement of Z_o
V1 0 1 0 AC {vs} 0
I2 0 3 0 AC {io} 0
I1 0 2 0             ;Adjust this value for zero DC output voltage
VP 4 0 5             ;Positive supply voltage
VN 0 5 5             ;Negative supply voltage
C1 2 0 1meg          ;Adjust this value if necessary
R1 3 2 1meg          ;Adjust this value if necessary
X1 1 2 4 5 3 LT1677 ;Device Under Test (DUT)
.include LTC.lib     ;Library file with the subcircuit of DUT
.ac dec 20 1 10meg  ;AC sweep over frequency range of interest 20 points/
     decade
.save v(3)           ;Save the output voltage
.end
```

Poles and zeros of $A_v(s)$ and $Z_o(s)$

Figure 8.8 shows no difference between the test circuit for measurement of $A_v(s)$ of $Z_o(s)$. Only the values of two independent sources differ for the two different measurements. This implies that the voltage gain and the output impedance have the same poles. Only the zeros of both transfers (8.3) and (8.4) differ. However, not all of the poles of the voltage gain may be observable when measuring the output impedance and vice versa. If a pole in the output impedance cannot be observed in the voltage gain , then the voltage gain has a zero at the (complex) frequency of that pole in the output impedance; and vice versa. The concept of observability of poles is discussed in section 18.5.

In SPICE $A_v(s)$ can be modeled with a voltage-controlled voltage source of which the value is defined by a Laplace function. Below, you will find the syntax for an operational amplifier with a gain-bandwidth product GB, a DC gain A_0 and a first-order high-frequency roll-off. The circuit diagram is shown in Figure 8.9.

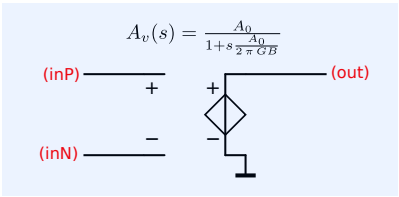


Figure 8.9: Model of a single-pole operational amplifier with infinite input impedance, zero output impedance, DC gain A_0 and gain-bandwidth product GB.

```
* file: simpleOpamp.cir
* LTspice subcircuit for simple a 3-terminal OpAmp
* A voltage controlled voltage source models A_v(s)
* of a single-pole operational amplifier with
* a DC gain A_0=1Meg and a gain-bandwidth product GB=10MHz
* noninverting input: inP
* inverting input:    inN
* output:             out
* reference node:     0
.subckt simpleopamp inP inN out params: A_0=1meg GB=10meg
E1 out 0 inP inN laplace = {A_0/(1+s*A_0/2/pi/GB)}
.ends  simpleopamp
```

In SPICE $Z_o(s)$ can be modeled with network elements like noise-free resistors, capacitors and inductors. Noise-free resistors can be alternatively be modeled with voltage-controlled current sources, as shown in Figure 8.10.

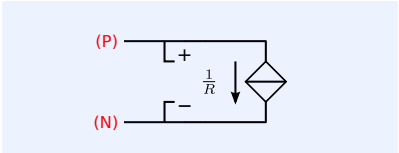


Figure 8.10: In SPICE, a noise-free resistor can be modeled with the aid of a voltage-controlled current source.

The dynamic behavior of $Z_o(s)$ can be modeled by replacing the DC value $\frac{1}{R}$ of the controlled source with a Laplace expression. Please notice that the poles of this expression will be the zeros of $Z_o(s)$ and vice versa. Since Laplace expressions in LTSPICE cannot have more zeros than poles, this technique cannot be used for modeling capacitive behavior at the highest frequencies. This limitation can be solved by adding capacitors, or by using a current-controlled voltage source for impedance modeling, as shown in Figure 8.11.

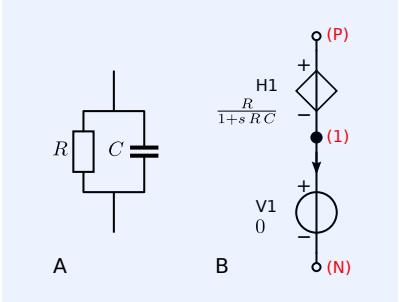


Figure 8.11: In SPICE, a noise-free impedance with a capacitive character for the highest frequencies, can be modeled with the aid of a current-controlled voltage source.
A: RC parallel connection
B: Network for a noise-free parallel RC connection in SPICE.

Below is the SPICE syntax for a parallel connection of a capacitor with capacitance C and a resistor with resistance R, according to the model from Figure 8.11:

```
* file: noiseFreeRC.cir
* LTspice subcircuit for a parallel connection of a noise-free resistor
* and a capacitor
```

```
* P: positive node
* N: negative node
* R: resistance
* C: capacitance
.subckt noisefreerc P N params: R=1 C=1
V1 1 N 0
H1 P 1 V1 laplace = {R/1+s*R*C}
.ends noisefreerc
```

Small-signal dynamic behavioral model for current-feedback opamps

Figure 8.12 shows the small-signal dynamic equivalent model of current-feedback operational amplifiers. The input impedances, output impedance and the differential-mode voltage transfer are modeled, either by using Laplace expressions or by using equivalent networks. This model suffices for evaluation of the small-signal frequency response, the small-signal impulse response and the small-signal step response of circuits with operational amplifiers.

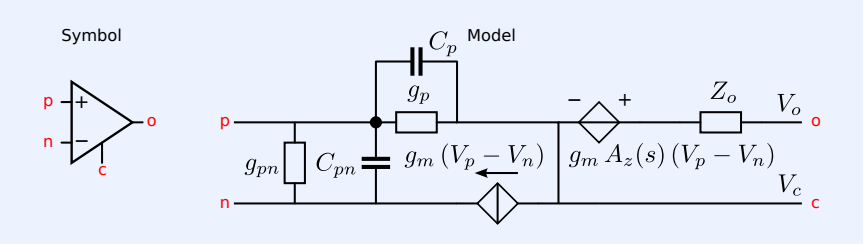


Figure 8.12: Simplified small-signal model of a current-feedback operational amplifier. The transimpedance gain A_z and the resistance R are usually specified in the data sheet.

An overview of the model parameters is given in Table 8.6. Common-mode rejection, power supply rejection, noise, bias current and offset current and voltage as well as temperature effects are not modeled. Errors resulting from these performance limitations will be modeled at a later stage.

Name	Description
g_p	Small-signal conductance between noninverting input and ground
C_p	Small-signal capacitance between noninverting input and ground
C_{pn}	Small-signal capacitance between noninverting input and inverting input
g_m	Transconductance of the input stage
$A_z(s)$	Laplace transform of the output stage's transimpedance
$Z_o(s)$	Laplace transform of the output impedance

Table 8.6: Parameters of the small-signal model of the current feedback operational amplifier.

Determination of the transconductance of the input stage and the transimpedance of the output stage can be done with the aid of the test circuit from Figure 8.13. The resistance R should be selected such that:

$$R \gg R_{out}, \tag{8.5}$$

$$R \gg A_z(0). \tag{8.6}$$

The transconductance g_m, the transimpedance gain $A_z(s)$ and the output impedance $Z_o(s)$ are obtained as:

$$g_m = \frac{I_i}{V_i}, \tag{8.7}$$

$$A_z(s) = \frac{V_o}{I_i}, \tag{8.8}$$

$$Z_o(s) = \frac{V_o}{I_o}. \tag{8.9}$$

The DC current I_B needs to be adjusted to bring the DC output voltage in its linear operating region.

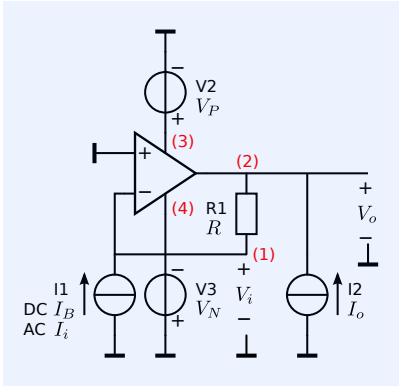


Figure 8.13: Test circuit for determination of the transconductance, the transimpedance and the output impedance of a current-feedback operational amplifier. $g_m = \frac{I_i}{V_i}, A_z = \frac{V_o}{I_i}, Z_o = \frac{V_o}{I_o}$

Rail-to-rail output operational amplifiers

Many modern operational amplifiers have so-called rail-to-rail outputs. These devices can drive the load with voltages almost equal to the power supply voltages. Older operational amplifiers were not capable of doing so. This all has to do with the topology of the output stage used in operational amplifiers. Rail-to-rail output stages usually exhibit a high low-frequency output impedance. This implies that the low-frequency voltage gain of the amplifier strongly depends on the low-frequency value of the load impedance. Although the output impedance of the operational amplifier often plays a significant role in the dynamic behavior of its application, it is not always specified in the data sheets. For reliable designs, only devices that should be used of which the relevant performance aspects for the application are specified.

8.3.2 Noise behavior

A simple noise model , with frequency-independent noise sources, can be obtained by adding input-equivalent noise sources to the small-signal dynamic model. Figure 8.14 shows the way in which two uncorrelated equivalent input noise sources can be added.[2] Both sources have uniform power density spectra (white noise). The spectra of the equivalent noise voltage and current source are given by S_v and S_i, respectively:

$$S_v = 4kTR_{nv}\ [\mathrm{V^2/Hz}], \tag{8.10}$$

$$S_i = \frac{4kT}{R_{ni}}\ [\mathrm{A^2/Hz}]. \tag{8.11}$$

The values of the resistors from Figure 8.14 are obtained from these spectra:

$$R_{nv} = \frac{S_v}{4kT}, \tag{8.12}$$

$$R_{ni} = \frac{4kT}{S_i}. \tag{8.13}$$

A SPICE model that includes $1/f$ noise can be constructed with the aid of a PN diode . The noise model of a PN diode is briefly discussed in Chapter 19.

Figure 8.15 shows the model for a voltage noise source including $1/f$ noise. The syntax for a SPICE subcircuit is listed below:

```
* file: vn.cir
* LTspice subcircuit of a voltage noise source for noise analysis
* sv: noise voltage density (noise-floor) in V/sqrt(Hz)
* fl: 1/f corner frequency
* Lowest frequency 100uHz
* For lower frequencies increase C1
.subckt vn 3 4 params: fl=1 sv=1n
I1 0 1 3.125u
D1 1 0 dnoise
C1 1 2 1
V1 2 0 0
H1 3 4 V1 {sv*1e12}
.model dnoise d kf={3.2e-19*fl} af=1
.ends vn
```

For frequencies at which the impedance of the capacitor `C1` is much smaller than the small signal impedance of the diode, the current noise of the diode flows through the voltage source `V1`. This current is converted into a voltage by the current-controlled voltage source `H1`.

[2] An equivalent input voltage noise source and an equivalent input current noise source.

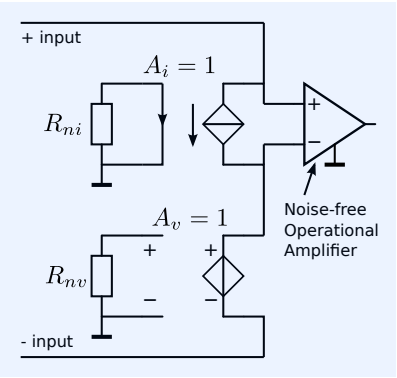


Figure 8.14: Modeling of the noise performance with the aid of two uncorrelated noise sources with uniform power density spectra.

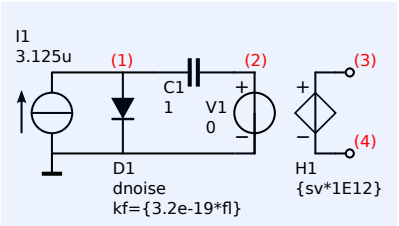


Figure 8.15: LTspice subcircuit for a voltage noise source with $1/f$ noise.

The spectral density S_i of the noise current through V1 can be written as:

$$S_i = 2qI_D\left(1 + \frac{\text{KF} I_D^{\text{AF}-1}}{2qf}\right), \tag{8.14}$$

where f is the frequency, KF and AF are model parameters, q the charge of the electron and I_D the DC current flowing through the diode. With $I_D = 3.125\mu\text{A}$, we have $2qI_D = 10^{-24}\text{A}^2/\text{Hz}$.

If $\text{AF} = 1$, equation 8.14 can be written as:

$$S_i = 2qI_D\left(1 + \frac{f_\ell}{f}\right), \tag{8.15}$$

where the $1/f$ corner frequency f_ℓ equals:

$$f_\ell = \frac{\text{KF}}{2q}. \tag{8.16}$$

Hence, if we want a corner frequency f_ℓ, we need $\text{KF} = 2qf_\ell$.

Figure 8.16 shows the model for a current noise source including $1/f$ noise. The syntax for a SPICE subcircuit is listed below:

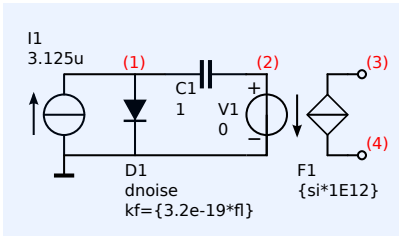


Figure 8.16: SPICE subcircuit for a current noise source with $1/f$ noise.

```
1  * file: in.cir
2  * LTspice subcircuit of a current noise source for noise analysis
3  * si: noise current density (noise-floor) in A/sqrt(Hz)
4  * fl: 1/f corner frequency
5  * Lowest frequency 100uHz
6  * For lower frequencies increase C1
7  .subckt in 3 4 params: fl=1 si=1p
8  I1 0 1 3.125u
9  D1 1 0 dnoise
10 C1 1 2 1
11 V1 2 0 0
12 F1 3 4 V1 {si*1e12}
13 .model dnoise d kf={3.2e-19*fl} af=1
14 .ends in
```

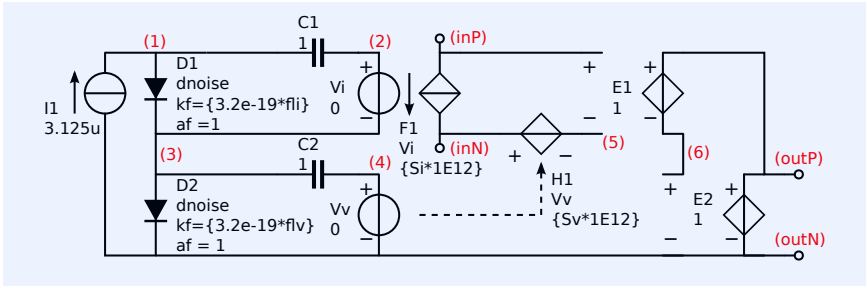


Figure 8.17: LTSpice model of a nullor aith added equivalent-input voltage and current noise sources.

Figure 8.17 shows an LTSpice sub circuit for a nullor with added equivalent-input voltage and current noise sources. Both noise sources are uncorrelated and exhibit $\frac{1}{f}$ noise. The LTSpice syntax for this circuit has been listed below.

```
1  * Sub circuit nNoise
2  * LTspice subcircuit for a noisy nullor for noise analysis
3  * Lowest frequency 100uHz
4  * For lower frequencies increase C1 and C2
5  * Nodes:       in+ in- out+ out-
6  .subckt nNoise inP inN outP outN params: Sv=1n Si=1p flv=1k fli=10k
7  * Parameters:
8  * Sv:  input noise voltage density (noise-floor) in V/sqrt(Hz)
9  * flv: 1/f corner frequency of voltage noise
10 * Si:  input noise current density (noise-floor) in A/sqrt(Hz)
11 * fli: 1/f corner frequency of current noise
12 E1 outP 6 inP 5 1
13 E2 outP outN 6 outN 1
14 D1 1 3 DnoiseV
15 D2 3 outN DnoiseI
16 I1 outN 1 3.125u
17 C1 1 2 1
```

```
C2 3 4 1
Vi 2 3 0
Vv 4 outN 0
F1 inP inN Vi {Si*1e12}
H1 inN 5 Vv {Sv*1e12}
.model DnoiseV D kf={3.2e-19*flv} af=1
.model DnoiseI D kf={3.2e-19*fli} af=1
.ends nNoise
```

8.3.3 PSRR and CMRR

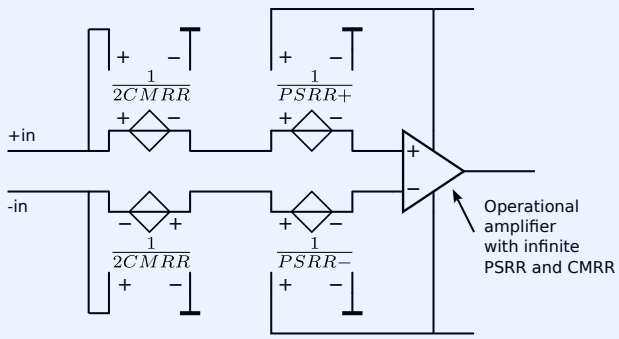


Figure 8.18: Modeling of the influence of the finite CMRR and PSRR. The PSRR for the positve power supply and for the negative power supply have been modeled with the transfers $PSRR+$ and $PSRR-$, respectively. Frequency dependencies can easily be implemented with the aid of passive transfer networks or Laplace blocks to the inputs of the controlled sources.

The $PSSR$ and the $CMRR$ are defined as the equivalent differential-mode input voltage sources that compensate for the output voltage change due to a change of the power supply voltage or of the common-mode input voltage, respectively. Modeling can be done according to Figure 8.18. Complex networks or Laplace expressions can be used for modeling dynamic effects.

The model from Figure 8.18 is not complete. This has already been discussed with the introduction of the two-port description for amplifiers. A complete model would also require current sources in parallel with the input, that are controlled by the power supply voltages and the common-mode voltage. Design data for these controlled sources, however, cannot be found in the data sheets, since the $CMRR$ and the $PSRR$ are only specified for voltage-driven inputs.

8.3.4 Bias and offset quantities

Modeling of the static nonlinear behavior (see Table 8.2 for parameters) can be done in various ways.

In many macro models , the bias current has been modeled with the aid of active devices. The offset current and the offset voltage have often been modeled with the aid of an independent current source and an independent voltage source, respectively. Unfortunately, this is not the correct way to do it. This is because the mean value of both of the offset quantities of a large number of devices tends to be zero.

The offset voltage source should be given statistical parameters for Monte Carlo analysis. Below the SPICE syntax for a voltage between node 1 and node 0 with a normal distribution with zero mean value and a standard deviation $\sigma = 10\text{mV}$.

```
MCexample

* file: MCexample.cir
* Example of a Monte Carlo simulation in LTspice
* vgauss :Voltage source with Gaussian distribution
* positive node = 1, negative node = 0
* vmean  :Mean value
```

```
* vsigma :Standard deviation

.params vmean=0 vsigma=10m
vgauss 1 0 {vmean + gauss(vsigma)}
.op
.step param run 1 200 1 ;200 Monte Carlo runs
.save V(1)
.end
```

Bias current sources can be realized in a similar way and placed between each of the operational amplifier's input terminals and the ground.

8.3.5 Modeling of other effects

Modeling of voltage and current limitations and slew-rate is not always necessary. Macro models that incorporate these effects can become very large and may cause numerical problems during simulations. An alternative approach for verification of the influence of these nonlinear effects is to use linear models exclusively:

1. Plot the derivative of the output voltage of the operational amplifier versus time and check if it exceeds the specification for the maximum slew rate
2. Plot the output voltage versus the output current of the operational amplifier and check if it stays within the specified region. Do this by using the graphs of the output voltage versus output current of the data sheet.
3. Plot the common-mode input voltage versus time and check if it stays within the specified region.

8.3.6 Macro models

Macro models are behavioral models that are used for numeric simulation with SPICE-like simulators. Macro models for operational amplifiers were introduced in 1974 by Boyle, Cohn, Pederson and Solomon [Boyle1974][3]. Macro models are comprised of fewer nonlinear elements than transistor-level models. This strongly facilitates fast computer simulations. The first macro models, however, only modeled a limited number of performance aspects. Simulation results with these models may strongly deviate from the real world behavior. Important improvements to these models have been suggested by Alexander and Derek [AlexanderDerek1990][4].

[3] Boyle, et al. Macromodeling of Integrated Circuit Operational Amplifiers. *IEEE Journal of solid-state circuits*, 9(6), December 1974

[4] M. Alexander and D.F. Bowers. SPICE-compatibe Op Amp Macro-Models. *EDN*, February, March 1990

Almost all operational amplifier manufacturers nowadays provide macro models for SPICE simulators. Unfortunately, the underlying equivalent circuit is not always given. Macro modeling with standardized models and with clearly defined model parameters, comparable to modeling of semiconductor devices, is not (yet) common practice. As a matter of fact, almost all operational amplifier manufacturers advice using the design information from the data sheets and making use of breadboards for testing specific applications.

It is strongly advised that designers who want to make use of macro models first evaluate whether the behavioral aspects of interest have been modeled correctly. This can be done by simulating test circuits identical to those given in the data sheets. Special attention has to be paid to the following aspects:

1. Operational amplifiers do not have a specific ground terminal. However, if the macro model of an operational amplifier is formed according to Boyle, the return path for the output current is the ground node rather than the power supply terminals. In such cases simulations fail if the operational amplifier is operating at a relatively high voltage with respect to the ground. This is corrected in the model presented by Alexander and Bowers.

An easy check as to whether the model has an internal connection to the ground is to sum all the currents flowing into the external nodes of the model when it is supplied from grounded voltage sources. This sum should, under all conditions, be zero; if not, some current is flowing to the ground.

2. Most macro models have the offset voltage modeled with a fixed independent voltage source. This may result in canceling of offset voltages in circuits with multiple operational amplifiers. In practice, this will not occur, because offset voltages and currents of different devices do not match.

3. The output impedance of the operational amplifier (often referred to as the "open-loop output impedance") is not always correctly modeled, neither is it fully specified in the data sheet. This makes it difficult to predict the dynamic behavior under various loads. Performance evaluation with the aid of breadboards may be necessary.

4. Rail-to-rail input operational amplifiers have two input stages. Related effects have not always been modeled in macro models.

5. A general rule for robust and first-time right design is not to depend on device parameters that have not fully been specified. If, for example, the output impedance is not specified, but it is critical in the application, it is wise to look for a device that has it specified, or to build a test circuit and verify its influence.

8.4 Design of feedback configurations with Op-Amps

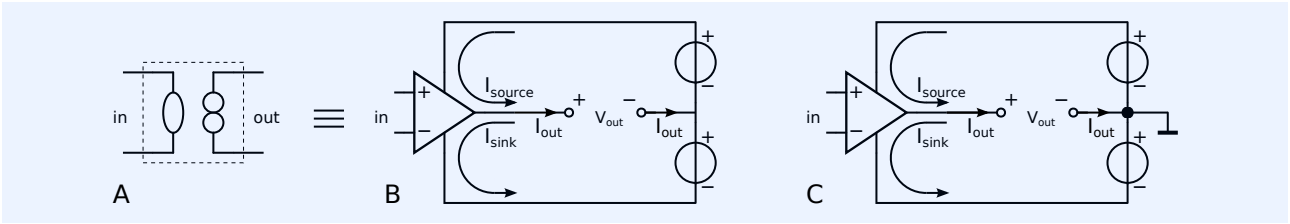


Figure 8.19: Implementation of the nullor with an operational amplifier hinders current sensing at the output port of the controller.

A. The nullor is the ideal controller, the nullator sets a network condition and the norator provides a current that satisfies this condition.

B. The operational amplifier can be regarded as an implementation of the nullor. However, for positive values of the output current the return path is the positive power supply, while for negative values of the output current the return path is the negative power supply.

C. The common power supply terminal is usually taken as ground. This makes it impossible to have both a grounded load and a grounded current sense element.

In this book we do not discuss the transistor-level design of controllers. We confine ourselves to the application of operational amplifiers as controllers. Implementation of controllers with operational amplifiers, puts serious constraints to the design of negative feedback amplifiers. This is because the return path for the output current differs for positive and negative values. This is illustrated in Figure 8.19. It shows that a grounded load cannot be combined with a sense element for the load current.

In this section we will discuss the design of negative feedback amplifier configurations of which the controllers can be implemented with operational amplifiers.

8.4.1 Single-loop passive feedback configurations

Figure 8.20 shows the basic single-loop negative feedback amplifiers in which the nullor has been replaced with an operational amplifier. Figure 8.20A shows the passive-feedback voltage amplifier. This configuration is often explicitly referred to as noninverting voltage amplifier configuration. Figure 8.20B shows the (inverting) transimpedance amplifier. Figure 8.20C shows

the transadmittance amplifier. This configuration has a grounded source and a floating load. This is the result of the fact that one of the output terminals of the controller is grounded (see section 7.2.2). This is also the case in the current amplifier from Figure 8.20D. In the following sections we will discuss these single-loop feedback configurations in more detail.

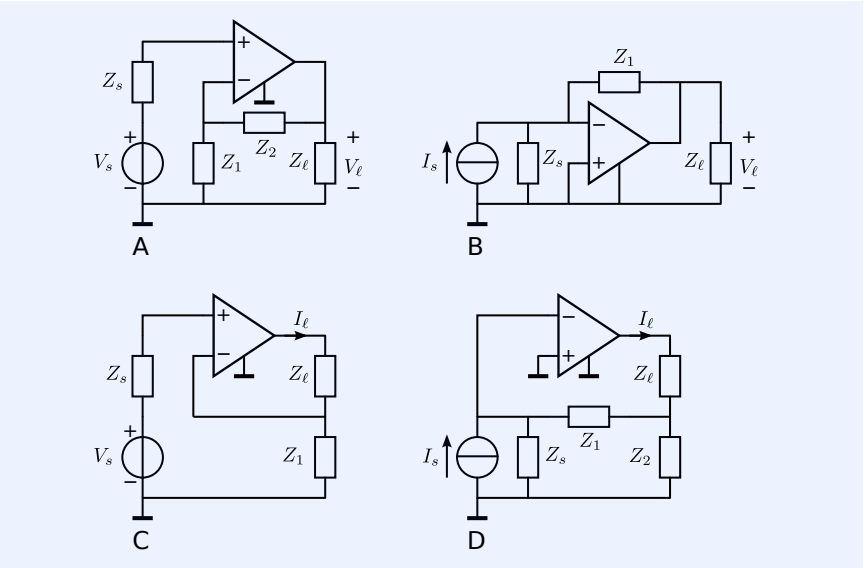


Figure 8.20: Basic single-loop negative feedback configurations with operational amplifiers.

Noninverting voltage amplifier

The ideal gain of the voltage amplifier from Figure 8.20A is

$$\frac{V_\ell}{V_s} = \frac{Z_1 + Z_2}{Z_1}. \tag{8.17}$$

The feedback network affects the noise behavior of the amplifier as if the parallel connection of Z_1 and Z_2 is in series with the source impedance. Low-noise design requires a low impedance for this parallel connection with respect to the source impedance.

The power efficiency of the amplifier is affected by the feedback network as if the series connection of Z_1 and Z_2 is in parallel with the load impedance. Low-power design requires a large impedance for this series connection with respect to the load impedance. The conditions for low-noise and low power design may conflict in low-gain situations.

Inverting transimpedance

The ideal gain of the transimpedance amplifier from Figure 8.20B is

$$\frac{V_\ell}{I_s} = -Z_1. \tag{8.18}$$

The feedback network affects the noise behavior of the amplifier as if Z_1 is in parallel with the source impedance. Low noise design requires a large impedance for Z_1. Independent design of the gain and the noise performance can be done by cascading the circuit with a low-noise voltage attenuator.

The power efficiency of the amplifier is affected by the feedback network like Z_1 is in parallel with the load impedance.

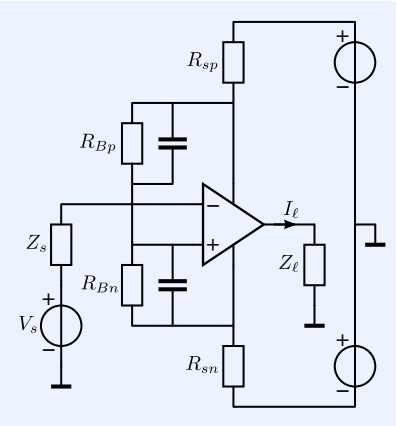


Figure 8.21: Transadmittance amplifier with grounded source and load.

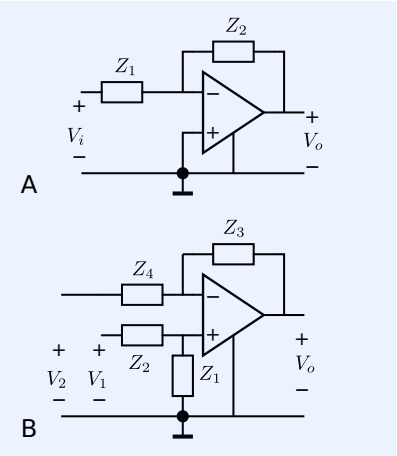


Figure 8.22: Basic configurations for constructing operational amplifier circuits with active feedback.

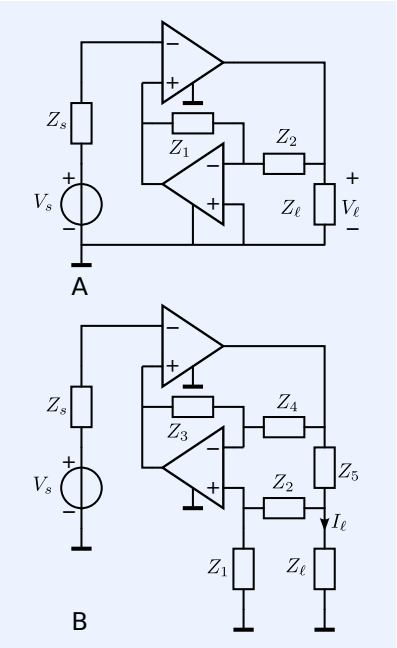


Figure 8.23: Two examples of active feedback amplifiers with opamps:
A: Low-noise inverting voltage amplifier
B: Inverting transadmittance amplifier with grounded load.

Transadmittance amplifier

Figure 8.20C shows a transadmittance amplifier with grounded source and floating load. The feedback impedance Z_1 affects the noise of the amplifier as if it is in series with the signal source. It affects the power efficiency as if it is in series with the load. A configuration for grounded source and load is drawn in Figure 8.21. The load current is now sensed in both power terminals of the amplifier.

The transfer is set by R_{sp} and R_{sn}.A difference between the values of R_{sp} and R_{sn} results in strong offset and even order distortion. This is a result of the class AB output stage of the operational amplifier. The resistor in the positive supply terminal carries the bias current plus the source current, while the resistor in the negative supply terminal carries the bias current plus the sink current. The resistors R_{Bp} and R_{Bn} provide the feedback voltage. In order to reduce their noise contributions, these resistors are shunted by capacitors. These capacitors also provide power-supply decoupling for the amplifier and reduce the even order distortion. The power supply noise adds up to the total input noise.

Although conceptually correct, this circuit is seldom used because of the drawbacks mentioned above.

Current amplifier

Figure 8.20D shows a current amplifier with grounded source and floating load. The principle of current sensing in the power supply leads, as shown in Figure 8.21, can also be applied to the current amplifier. In that case the noninverting input of the operational amplifier is connected to ground, while its inverting input is taken as the current input. The circuit has the same drawbacks as the transadmittance amplifier.

8.4.2 Active feedback amplifier configurations

As discussed before, active feedback can be exploited as a technique to construct amplifiers that cannot be obtained with passive feedback. Figure 8.22 shows active feedback elements with an operational amplifier as controller. The circuit from Figure 8.22A is known as the *inverting voltage amplifier*. However, a low-noise design of this amplifier would require a low value for Z_2, which conflicts with a high input impedance. It would be better to call this circuit a brute-force voltage to current converter cascaded with a transimpedance amplifier. If the circuit is used as a voltage attenuator in the feedback path of an active feedback amplifier, this conflict is resolved. Figure 8.23A shows its application as such in an active-feedback inverting voltage amplifier of which a low noise addition can be combined with a high input impedance.

A circuit that can be used to transfer a floating voltage into a grounded feedback voltage is depicted in Figure 8.22B. This circuit is known as the *differential voltage amplifier*, however, if it is applied as such, it has some serious drawbacks that are the result of the applied brute-force techniques:

1. Low-noise design conflicts with low-power design
2. Low-noise design conflicts with a high input impedance at both input terminals
3. Both input terminals have a different impedance to ground
4. The common-mode rejection is limited by the tolerances of the feedback impedances.

The output voltage V_o of the circuit can be written as a function of the two input voltages V_1 and V_2:

$$V_o = V_1 \frac{Z_1}{Z_1 + Z_2} \frac{Z_3 + Z_4}{Z_4} - V_2 \frac{Z_3}{Z_4}. \tag{8.19}$$

For differential to single-ended voltage conversion with a gain A_v and an infinite $CMRR$ we need

$$V_o = A_v (V_1 - V_2). \tag{8.20}$$

This requires

$$A_v = \frac{Z_1}{Z_1 + Z_2} \frac{Z_3 + Z_4}{Z_4} = \frac{Z_3}{Z_4}, \tag{8.21}$$

from which we obtain the design equation:

$$A_v = \frac{Z_1}{Z_2} = \frac{Z_3}{Z_4}. \tag{8.22}$$

Figure 8.23B shows application of this circuit in an inverting transadmittance amplifier with grounded load. In the following example we will evaluate ideal gain of this amplifier, as well as the contribution of the noise sources of the active feedback element to the source-referred and the load-referred noise.

Example 8.1

Figure 8.24 shows the concept of a voltage to current converter using active feedback. The load current is sensed by R_s and converted into a floating voltage $I_\ell R_s$. This voltage is converted into a voltage V_o, which is referenced to the ground. The difference between this voltage and the source voltage is nullified.

The ideal gain G_i of this transconductance amplifier equals

$$G_i = \frac{I_\ell}{V_s} = \frac{1}{AR_s} \frac{(1 + A)\, R}{(1 + A)\, R + R_\ell}. \tag{8.23}$$

If $(1 + A)\, R \gg R_\ell$, the sensed current approximates the load current, and the above expression can be simplified to

$$G_i = \frac{I_\ell}{V_s} = \frac{1}{AR_s}. \tag{8.24}$$

Under the above conditions, the spectrum S_{V_i} of the source-referred voltage noise can be obtained as

$$S_{V_i} = (S_v + 8kTAR) \left(1 + A^2\right) + (2AR)^2 S_i, \tag{8.25}$$

where S_v and S_i represent the spectrum of the equivalent input voltage noise and the equivalent input current noise of the controller in the active feedback element, respectively.

The contribution of the active feedback element to the total output current noise is obtained after multiplication of S_{V_i} with G_i^2. This yields

$$S_{I_o} = (S_v + 8kTAR) \frac{1 + A^2}{A^2 R_s^2} + \left(2 \frac{R}{R_s}\right)^2 S_i. \tag{8.26}$$

From this expression we may derive the following design conclusions:

1. *The noise contribution of the feedback network can be kept small by taking R_s as large as possible. This, however, conflicts with a high power efficiency.*
2. *For a low-noise performance we also require R as small as possible and $A > 1$. This conflicts with a high power efficiency and with a high accuracy. The latter one requires $(1 + A)\, R \gg R_\ell$.*

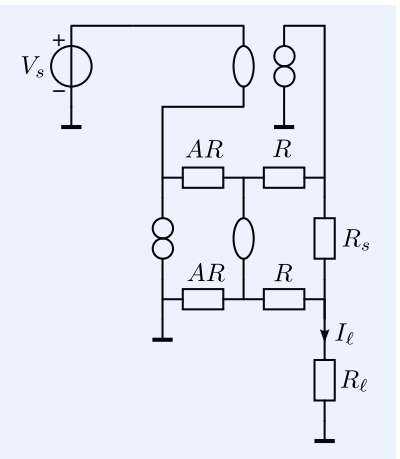


Figure 8.24: Transadmittance amplifier with active feedback.

9 Introduction to amplifier biasing

9.1 Introduction

In this chapter, we will give a brief introduction to the biasing of amplifiers or amplifier stages. We will start with a definition of the term *biasing:*

> Biasing is the application of a collection of techniques for fixing the electrical operating conditions of electronic devices, and deriving the required bias voltage and current sources from the power supply voltage(s).

In Chapter 2 we stated that the ideal behavior of an amplifier can be characterized by three curves:

1. The $v - i$ characteristic of the input port
2. The $v - i$ characteristic of the output port
3. The input-output characteristic.

Since amplifiers are intended to behave in a linear, stationary and instantaneous manner, these characteristics should all be straight lines that pass the origin (see Figure 2.10). These characteristics should not change over time. Conceptually, this is true, but in practice, we may need to add offset to these characteristics to shift them out of the origin.

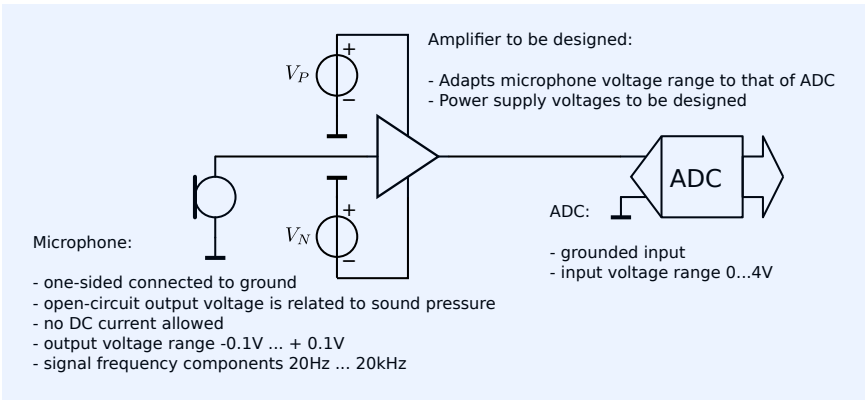


Figure 9.1: Simple audio digitizing system. Only aspects relevant to amplifier biasing are shown.

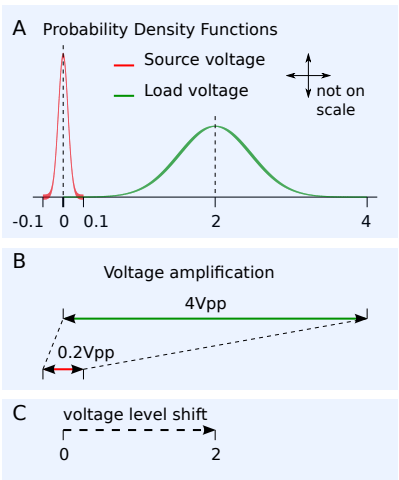


Figure 9.2: The amplifier in the system from Figure 9.1 must provide voltage amplification and level shift:
A. The probability density functions of the source voltage and the load voltage
B. The $0.2V_{pp}$ source voltage needs to be converted into a $4V_{pp}$ load voltage. This requires voltage amplification.
C. The mean value of the source voltage equals zero, while that of the load should be $2V$. This requires a level shift function.

Let us, for example, consider the audio digitizer shown in Figure 9.1. The amplifier must adapt the microphone output voltage range to the input voltage range of the analog to digital converter (ADC). Let us assume the microphone is a electrodynamic type. An electrodynamic microphone comprises a membrane that is connected to a voice coil. This voice coil is placed in a magnetic field, and any motion caused by a change in air pressure on the membrane induces a voltage at the output of the voice coil. Such a microphone can also act as a telephone.

When driven from a signal voltage, the membrane causes a variation of the air pressure, which can be experienced as sound. Similarly, a DC current through the voice coil brings the membrane out of its quiescent position. Such a *DC bias* for a dynamic microphone is undesirable.

Hence, the DC input bias current of the amplifier, as introduced in sections 2.4.7 and 8.2.3, is not allowed to flow through the microphone. This can be achieved by using *AC coupling* between the microphone and the amplifier. Alternatively, this bias current may be provided by a bias current source, as discussed in section 2.4.7.

Another important aspect of biasing is that the microphone produces a bipolar signal with an average value of zero, while the ADC can only accept input voltages between 0 to 4V. In the absence of a signal, the ADC input

voltage should equal its midrange value: 2V. Hence, aside from raising the signal level from $0.2V_{pp}$ to $4V_{pp}$, we need to change the zero-signal voltage level from 0V to 2V. Such an addition of an offset voltage is often referred to as application of a *voltage level-shift* , or simply a *level shift*.

Figure 9.2 illustrates the combination of the signal amplification function and the level shift function that both have to be performed by the amplifier.

9.1.1 This chapter

The concept of DC coupling , AC coupling and the application of voltage level shifts and bias current sources, as well as the selection of the power supply voltages, will be discussed in section 9.2.

Basic techniques, such as, AC coupling and application of level shifts do not always provide sufficiently accurate biasing. Changes in the power supply voltages, temperature variations and device tolerances may unacceptably affect the biasing of the amplifier and that of the source and the load. If the variations in the operating voltages and currents are too large, error reduction techniques such as compensation, negative-feedback biasing, auto-zero techniques and modulation techniques[1] need to be applied to improve the stability of the quiescent operating point. The application of error reduction techniques for improvement of the stability of the biasing will be discussed in section 9.4.

[1] Such techniques are applied in so-called *chopper stabilized amplifiers*.

A separate section will be devoted to common-mode biasing. In section 9.5, we will discuss techniques for fixing common-mode voltages and/or currents in balanced amplifiers to their desired values.

9.2 Basic techniques

In this section, we will discuss basic biasing techniques. First, we will discuss the concept of biasing with level shifts and bias current sources. Then, we will discuss implementations with AC coupling.

9.2.1 Basic biasing technique

The application of level shifts and bias currents has already been introduced in section 2.4.7. In Figure 2.42, we shifted the quiescent operating point of the amplifier from $(V_{pQ},\ I_{pQ})$ to $(0,\ 0)$ by placing a bias voltage source V_{pQ} in series with the amplifier port and a bias current source I_{pQ} in parallel with the amplifier port. This method can be applied at the input port of the amplifier and at the output port of the amplifier. In the following example, we will design the biasing concept of the amplifier from Figure 9.1 using bias current sources and voltage level shifts.

Example 9.1

Let us assume we have designed the amplifier from Figure 9.1 as a passive feedback voltage amplifier with a voltage gain of 20*. Figure 9.3 shows the concept of this amplifier. The source quantity is the one that has the best reproducing relation with the primary information (sound pressure), and the impedance of this source is the small-signal source impedance. Here, the voltage source V_s and its source impedance Z_s represent the microphone. The input impedance of the ADC is the load impedance Z_ℓ of the amplifier. During the conceptual design of the amplifier, we do not consider the biasing and the power supplies. The amplifier is simply assumed to behave as a linear two-port that has the $v-i$ characteristics of both ports passing through the origin. The ground node is the power supply reference node.*

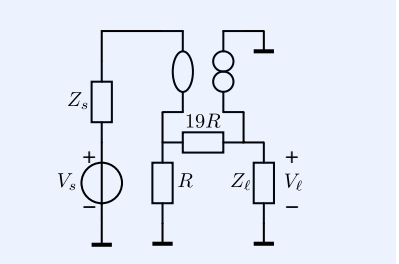
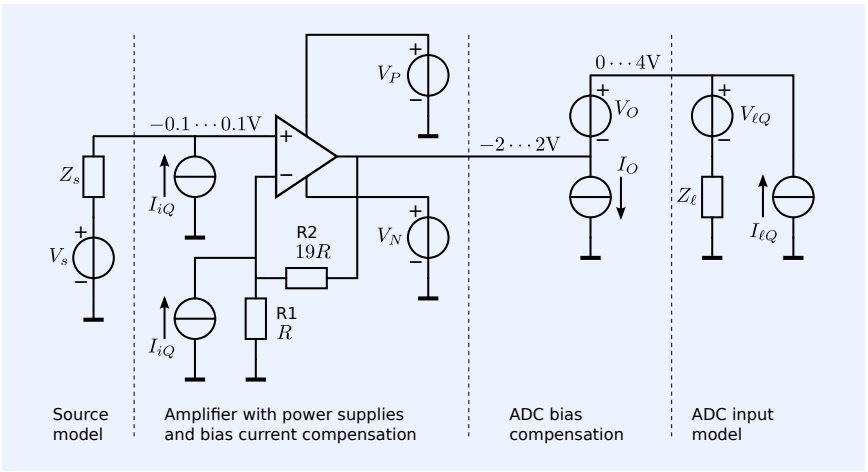


Figure 9.3: Concept design of the amplifier from Figure 9.1.

Figure 9.4 shows the amplifier in which the controller has been implemented with an operational amplifier and to which bias sources have been added.

Figure 9.4: Implementation of the controller with an operational amplifier and biasing concept of the amplifier from Figure 9.3.



We start the design of the biasing by choosing the power supply voltages of the operational amplifier such that it can drive the load with $\pm 2V$ signal excursion. Figure 9.4 shows these power supply voltages V_P and V_N for the positive power supply and the negative power supply, respectively.

Then, the operating point of the source and the load will be modeled. The $2V$ bias voltage of the ADC and the bias current of the ADC have been modeled with $V_{\ell Q}$ and $I_{\ell Q}$, respectively. The signal source should operate at zero bias, which requires no additional bias sources.

We then start with the design of the biasing of the amplifier. The output voltage range of the amplifier is adjusted to $0 \cdots 4V$ by adding an output offset voltage $V_O = 2V$ to the output voltage of the amplifier. This is done by inserting a $2V$ voltage source between the output of the amplifier and the input of the ADC. The bias current of the ADC is compensated for by an equally large current source in parallel with the output port of the amplifier. The input bias currents I_{iQ} of the operational amplifier have been compensated for by two current sources with a value I_{iQ}. In this way, both the input port and the output port of the amplifier operate at zero bias: the $v - i$ characteristics of both amplifier ports pass through the origin.[2] *Hence, at zero signal voltage, the source impedance, the load impedance, the amplifier input port, its feedback network and its output port all carry no current.*

[2] For the sake of simplicity, we assumed zero equivalent input offset voltage and zero equivalent offset current for the operational amplifier.

Before we will find a practical implementation for the biasing concept from Figure 9.4, we will introduce the terms *DC coupling* and *AC coupling*.

9.2.2 DC coupling and AC coupling

DC coupling

We speak of DC coupling between a source and a load if there is a nonzero DC transfer from the source to that load. If there exists only a nonzero transfer for frequencies that differ from zero, we speak of AC coupling.

Figure 9.5A shows two disconnected networks. There is no transfer from the source current I_s to the load voltage V_ℓ. Figure 9.5B shows the two networks connected through a resistive branch. However, there is still no transfer from the source to the load. Figure 9.5C shows the two networks connected through two resistive branches. Now there exists a DC coupling from the source to the load because there is a nonzero DC transfer from I_s to V_ℓ. Figure 9.5D shows the two networks connected through one resistive and one

capacitive branch. In this case, there exists a nonzero transfer only for frequencies that differ from zero, hence we have an AC coupling between the source and the load.

AC coupling

Two networks are AC coupled if there only exists a transfer from one network to the other network for frequencies that differ from zero. Hence, a change in the DC operating point of a network that has an AC coupling with another network, does not cause a change the of the DC operating point of this other network. The following rules apply to AC coupling :

1. AC coupling *can* be applied if signal components with very low frequencies are not of interest to the observer.
2. AC coupling between a source or a detector and an amplifier *has to be* applied if bias currents or bias voltages of the amplifier ports are not allowed to appear at the source or at the load.
3. AC coupling between two networks can be established by creating a high-pass transfer between the two networks.
4. If a nonzero DC transfer has to be established, but errors due to bias quantities are too large, the frequency range of the information needs to be changed such that DC transfer is no longer required. This principle of modulation and demodulation is applied in so-called chopper amplifiers.

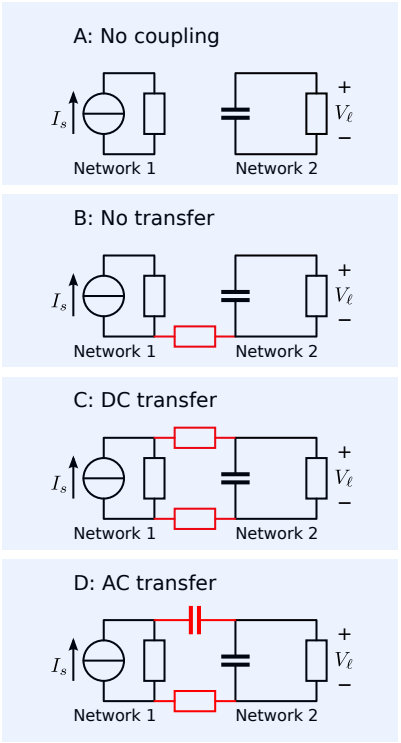


Figure 9.5: Coupling between networks:
A: Two disconnected networks
B: The two networks are connected, but there exists no transfer between the two networks
C: The two networks are connected and there exists a nonzero DC transfer from the left network to the one on the right
D: The two networks are connected, but there exists only a nonzero AC transfer between the two networks.

9.2.3 Deriving bias quantities from the power supply

There exist many different implementations of the biasing of the amplifier from Figure 9.1. In the following example, we will design a biasing scheme with DC coupling between the amplifier and the ADC and AC coupling between the signal source and the amplifier.

Example 9.2

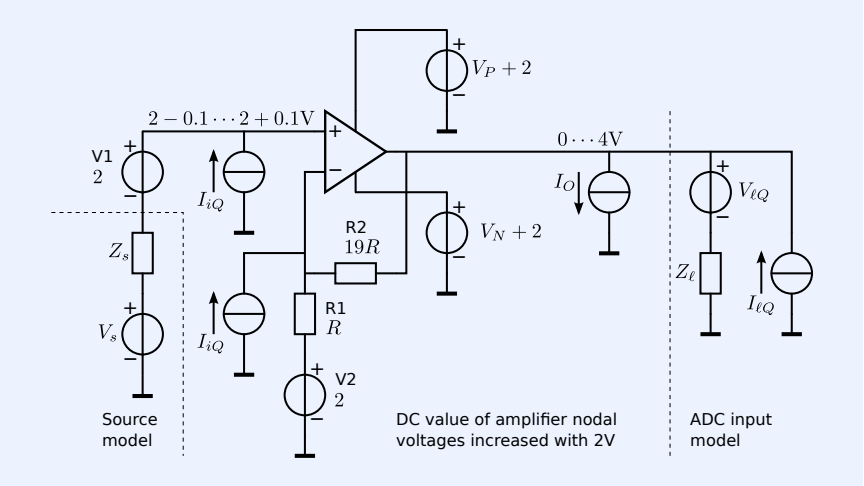


Figure 9.6: Amplifier from Figure 9.4 with 2V added to all nodal voltages.

In this example, we will implement the biasing of the amplifier from Figure 9.1 without the need for a level shift between the output of the amplifier and the ADC. We will start with the elimination of the 2V voltage source between the output of the amplifier and the input of the ADC.[3] *To this end, we increase the DC level of all the nodal voltages of the amplifier with 2V. Figure 9.6 shows the way in which this is done. Please notice that at the quiescent operating point, all voltage sources*

[3] This is because accurate low-noise voltage sources that are floating with respect to the ground, are hard to realize.

except the power supply voltage sources do not carry DC current! This is because all the required DC bias currents have been provided by DC current sources.

Since the frequency contents of the signal does not include DC, AC coupling may be applied. With AC coupling, level shifts that do not carry DC currents may be replaced with capacitors. These capacitors introduce impedances in series with the signal path. These impedances must be small enough not to deteriorate the signal-to-noise ratio and the power efficiency of the amplifier. This has to be the case at all frequencies of interest.

Figure 9.7 shows the result of this AC coupling. The capacitors C_1 and C_2 can be considered as small batteries implementing V_1 and V_2 from Figure 9.6, respectively. The DC voltage across these capacitors equals $2V$.

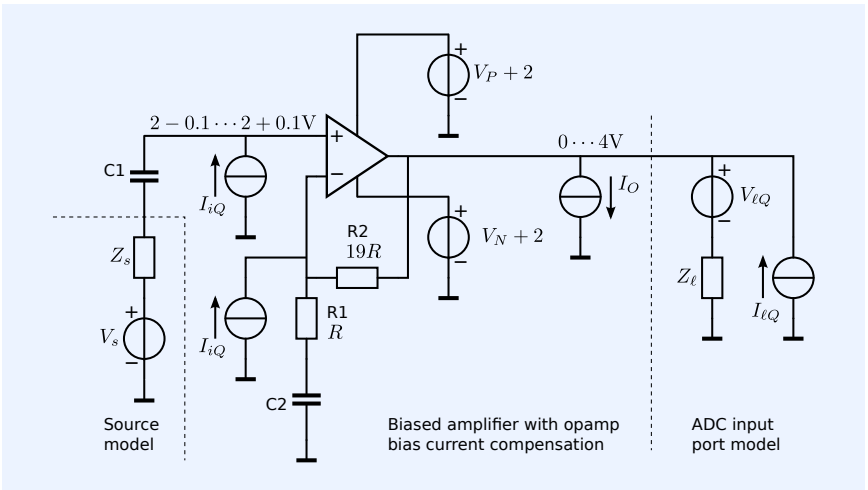


Figure 9.7: Amplifier from Figure 9.6 with the level shifts V1 and V2 replaced with capacitors.

In practice, the biasing scheme shown in Figure 9.7 will not provide a stable and well-defined biasing. This is because the bias currents of the operational amplifier are usually inaccurately known and strongly depend on temperature. In addition, the input impedance of the voltage amplifier is very high. These two properties result in a very badly defined DC voltage at the input of the amplifier, and measures have to be taken to convert this theoretically correct bias solution into a working one.[4]

For a complete design, we need to:

1. Determine the power supply voltages

   The minimum value of V_P equals the positive peak value of the output voltage plus the maximum value of the voltage drop in the output stage of the operational amplifier. This voltage drop depends on the internal structure of the output stage of the operational amplifier, on the temperature and on the current delivered by the amplifier. Figure 8.6 shows a plot of the output voltage and current drive capabilities of an operational amplifier.

   The *positive voltage headroom* is defined as the difference between this minimum required value of V_P and the actual value of V_P. The maximum value of V_N is determined in a similar way, now accounting for the *negative voltage headroom*. A large headroom is usually beneficial to the distortion, but it decreases the power efficiency of the amplifier.

2. Evaluate biasing errors

   Power supply tolerances, device tolerances and both the input offset voltage and the input offset current of the operational amplifier all introduce biasing errors. These errors may result in a reduced or even in a negative headroom.[5] Section 9.3 is devoted to this topic.

[4] This can be seen as follows: The common-mode input bias voltage error is the product of the common-mode input resistance and the error in the common-mode bias current. Operational amplifiers usually exhibit a very high input common-mode resistance, an inaccurately defined input bias current and a high relative input-bias current drift. The product of their common-mode input resistance and the unpredictable part of the bias current almost always exceeds the input common-mode voltage range.

[5] A negative headroom occurs if the voltage drive or the current drive capabilities are smaller than required.

3. If necessary, consider the application of error reduction techniques for improvement of the accuracy and the stability of the biasing

   If a certain biasing scheme results in unacceptably large biasing errors, error reduction techniques may be used to improve the biasing accuracy and stability. Application of such techniques will be discussed in section 9.4.

4. Verify the behavior with computer simulations and prototyping

In the following example, we will show a complete bias solution and discuss the above topics qualitatively.

Example 9.3

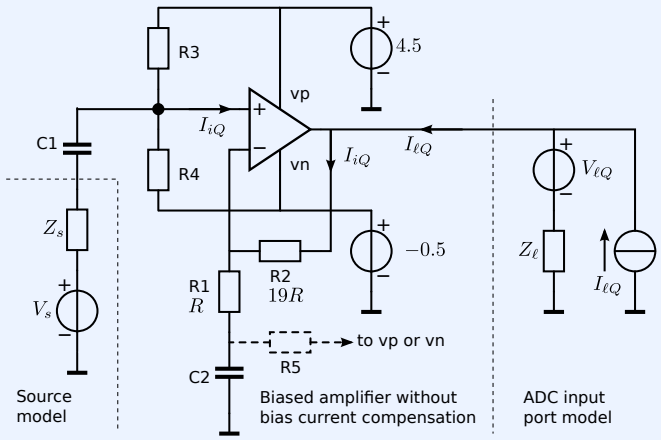


Figure 9.8: Amplifier from Figure 9.7 with brute-force input biasing and without input bias current and load bias current compensation.

Figure 9.8 shows a possible bias solution for the amplifier.

In this example, we use an operational amplifier with a rail-to-rail output stage. The positive and the negative voltage headroom are assumed to be sufficiently large with $V_P = 4.5V$ and $V_N = -0.5V$. In this solution, the common-mode input voltage of the amplifier should at least range from $1.9 \cdots 2.1V$, which is $2.4V$ below the positive supply and $2.4V$ above the negative supply voltage.

The resistors R_3 and R_4 fix the DC voltage at the noninverting input of the amplifier. This way of fixing the voltage at the noninverting input of the operational amplifier is basically a brute-force technique. Care should be taken as to the possible deterioration of the signal-to-noise ratio and the power efficiency of the amplifier. In order to minimize both the noise contribution and its effect on the gain, the linearity and the power supply rejection of the amplifier, the resistance of the parallel connection of R_3 and R_4 should be as large as possible. However, for a small biasing error, this resistance should be small enough. The biasing errors resulting from the input bias current and the offset current of the amplifier are proportional to $R_3 /\!/ R_4$.

If an operational amplifier with the required common-mode input voltage range cannot be found, there are two options:

1. *Increase the supply voltage such that the input common-mode voltage range of the operational amplifier satisfies the requirements.*

2. *Change the requirement for the common-mode input voltage range of the opamp.*

   This can be achieved by changing the values of V_1 and V_2 in the circuit from Figure 9.6. It can be implemented by changing the ratio $\frac{R_3}{R_4}$ and by inserting a DC current into the inverting input of the operational amplifier. This DC current is the Norton equivalent of the voltage change in V_2 in series with R_1. A low-noise implementation of this current source can be made by inserting a resistor R_5, as

shown in Figure 9.8. If the common-mode input voltage must be below the DC output voltage of the operational amplifier, R_5 should be connected to the negative supply, and otherwise to the positive supply voltage.

AC decoupling

AC decoupling means minimization of an AC transfer , while maintaining a DC transfer. It is the opposite of AC coupling ; it requires a low-pass filter action instead of a high-pass one. In example 9.3, the resistor R_5 connects the signal path of the amplifier to the power supply voltage. The capacitor C_2 performs two functions:

1. It acts as an AC coupling capacitor , because it establishes a high-pass character of the voltage amplifier

2. It acts as an AC decoupling capacitor , because it establishes a low-pass character in the voltage transfer from the power supply to the output.

In fact, the above can also be said for C_1 in combination with R_3 and R_4. However, the effectiveness of the AC decoupling by C_1 is limited by Z_s. At signal frequencies of interest, C_1 should act as a short with respect to Z_s, and at those frequencies, the transfer from the power supply voltages to the noninverting input of the operational amplifier will only be small if $|Z_s| \ll R_3 /\!/ R_4$.

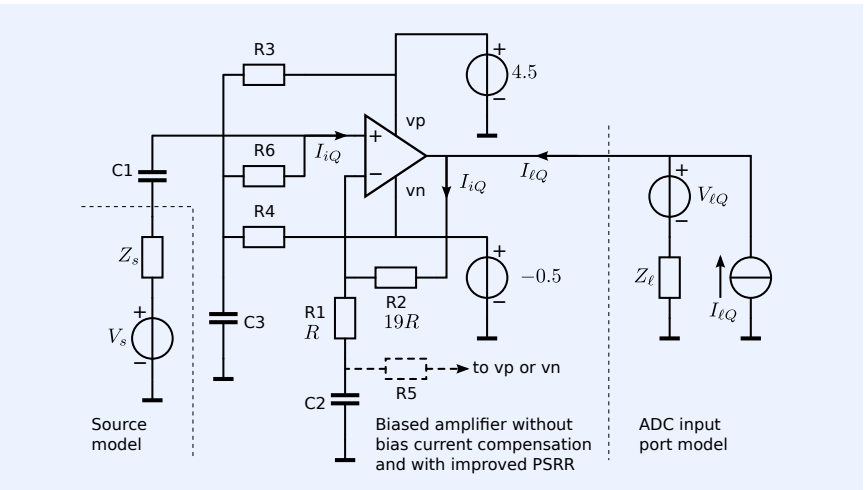


Figure 9.9: Amplifier from Figure 9.8 with a larger PSRR due to improvement of the power supply decoupling.

Figure 9.9 shows an arrangement for the amplifier with an improved PSRR in which extra power supply decoupling has been implemented with C_3. The power supplies, together with R_3, R_4 and C_3, constitute a low noise, low impedance voltage reference. The resistor R_6 connects the noninverting input of the operational amplifier to this reference, while maintaining a high input impedance for the voltage amplifier, just as R_3 and R_4 did in the circuit from Figure 9.8.

9.3 Evaluation of biasing errors

After we have designed a conceptual solution for the biasing, we need to evaluate possible error sources and find budgets for them. We will do this step by step:

1. First, we will evaluate the influence of power supply errors and resistor tolerances. To this end, we will replace the operational amplifier with a nullor, just as we did with the evaluation of the ideal gain. In order to find design budgets for tolerances and temperature coefficients, analysis should be done symbolically. We will demonstrate the use of SLiCAP for this purpose.

2. Then, we will study the influence of errors due to bias currents and offset currents and voltages of the controller. For this, we will add bias and offset sources to the nullor. Temperature effects, such as bias current drift and offset voltage drift, can be added as well. As mentioned above, design budgets will be derived with the aid of symbolic analysis.

3. Numeric verification can be done with the aid of SPICE using Monte-Carlo analysis at different temperature settings.

9.3.1 Power supply and resistor tolerances

For investigation of the influence of supply voltage changes and resistor tolerances, we will simplify the schematic of the amplifier as follows:

1. Capacitors will be left out; if this breaks a DC loop, all elements in that loop can be left out as well

2. All controllers will be replaced with nullors

Figure 9.10 shows the simplified diagram. In this case, the load has also been left out. This is done because, the controller having been replaced with a nullor, the load has no effect on the biasing of the amplifier.

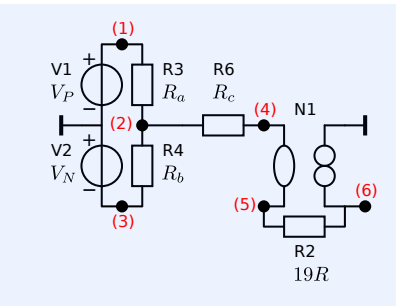


Figure 9.10: Simplified biasing scheme of the amplifier from Figure 9.9 for evaluation of the influence of power supply and resistor tolerances.

Determination of the numerical values

In order to find initial numerical values for the power supply voltages and the resistors, we need to derive and solve the design equations of the circuit. A symbolic expression for the voltage at the output of the amplifier can easily be obtained with SLiCAP symbolic DC analysis. The netlist of this circuit is shown below:

```
VampBiasNullor
* file: VampBiasNullor.cir
* SLiCAP netlist file
V1 1 0 V dc={V_P}
V2 0 3 V dc={V_N}
R3 1 2 {R_a}
R4 2 3 {R_b}
R6 2 4 {R_c}
R2 5 6 {19*R}
N1 6 0 4 5
.end
```

The SLiCAP script for evaluation of the DC voltage V_6 at node (6) is:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# file: VampBiasNullor.py

from SLiCAP import *

fileName = 'VampBiasNullor.cir'
i1 = instruction()          # Creates an instance of an instruction object
i1.setCircuit(fileName)     # Checks and defines the local circuit object, and
                            # sets the index page to the project index page.
i1.setDetector('V_6')
i1.setSimType('symbolic')
i1.setGainType('vi')
i1.setDataType('dc')
result = i1.execute()
```

```
DCvalue = result.dc
#
i1.setDataType('dcsolve')
result = i1.execute()
DCsolution = result.dcSolve
#
htmlPage('Netlist and circuit data')
netlist2html(fileName)
elementData2html(i1.circuit)
#
htmlPage('DC analysis')
text2html('The DC voltage $V_6$ is obtained as:')
eqn2html('V_6', DCvalue, units = 'V')
```

Running this script yields an expression for the detector voltage V_6:

$$V_6 = \frac{V_P R_b - V_N R_a}{R_a + R_b}. \tag{9.1}$$

This result could, of course, just as easily be obtained with hand calculations using the superposition theorem.

In order to assign numerical values to R_a and R_b, we need more design equations. Numerical values for V_P and V_N follow from the requirement for the load voltage range and budgets for the positive and the negative headroom:

$$V_P = V_{\max} + V_{H+}, \tag{9.2}$$

where $V_{\max}$ is the maximum positive voltage to be delivered to the load, and V_{H+} is the positive headroom required for proper operation of the operational amplifier.

An acceptable value for V_{H+} requires some knowledge about available components. Let us assume that, for example, a positive headroom $V_{H+} = 0.25 \cdots 0.5\text{V}$ leaves us a sufficiently large number of operational amplifiers from various manufacturers.

Similar things can be said about V_N. The minimum load voltage $V_{\min}$ should be zero, and with a negative headroom about equal to the positive headroom, we can work with $V_P = 4.5\text{V}$ and $V_N = 0.5\text{V}$.[6]

[6] Please notice that this results in a negative voltage at the power supply terminal of the operational amplifier.

This leaves us with two variables and one equation, so at this stage of the design, we can only determine the ratio of R_a and R_b:

$$R_b = \alpha R_a. \tag{9.3}$$

The design equation for α can be derived from:

$$V_6 = \frac{V_P \alpha R_a - V_N R_a}{R_a + \alpha R_a}. \tag{9.4}$$

The solution for α is found as:

$$\alpha = \frac{V_N + V_6}{V_P - V_6}. \tag{9.5}$$

With $V_N = 0.5$, $V_6 = 2$ and $V_P = 4.5$, we obtain $\alpha = 1$, and hence $R_a = R_b$

Influence of supply and resistor tolerances

For the investigation of the effects of device tolerances, designers have three techniques at their disposal:

1. Worst-case analysis

   This is a technique that accounts for maximum deviations in component values, in which the sign of the deviation is chosen so as to maximize its effect. Worst-case analysis can be performed symbolically, and the direction of the deviation can be obtained with the aid of sensitivity analysis.

2. Symbolic statistical analysis

   This technique uses the probability density functions of component values. It requires knowledge of algebra for random variables, which is a rather specialized discipline.

3. Monte-Carlo analysis

   Monte-Carlo analysis is a numerical technique running a simulation multiple times with different component values. At the beginning of each run, component values are randomly assigned according to pre-defined distribution functions and correlations.

Worst-case analysis may give clear a-priori insight into the effect of device tolerances, but it requires complex sensitivity analysis and may result in an overkill of robustness of the design. At first glance, this may appear attractive, but this overkill may seriously limit the feasibility of a design.

Monte-Carlo analysis can only be performed after devices have been selected, and it is therefore a verification technique, rather than a synthesis tool.

Symbolic statistical analysis would probably be the best way to go, but the algebra of random variables is complex and does not lend itself to hand calculations. Moreover, it is questionable whether such a complex technique will provide results that can be interpreted by the designer and yield clear design conclusions.

SLiCAP incorporates a simplified symbolic statistical analysis technique that, in many cases, gives clear design information with sufficient accuracy. The technique is only available for DC analysis and can be applied for setting up requirements for device tolerances , temperature drift and bias and offset voltages and currents of operational amplifiers. The method is based on the following principles:

1. In a linear network, the variance of a DC nodal voltage or a DC branch current can be obtained from linear superposition of the contributions of the variances of all of the independent sources in the network to that nodal voltage or branch current.

2. If tolerances of resistors are relatively small, errors due to these device tolerances can be converted into independent error current sources in parallel with the corresponding resistors. The contributions of these sources to the variance of a DC nodal voltage or a DC branch current can then be accounted for as described above.

3. This technique is allowed as long as the transfers of the independent sources to the output of the system are not noticeably affected by changes in the resistor values. This is often the case, and this condition can be easily be verified through symbolic analysis.

The conversion of the deviation of the resistance of a resistor from its mean value into an independent current source in parallel with the resistor proceeds as follows:

Let us assume we need a resistor with a resistance of R [Ω]. The DC voltage across this resistor in the application equals V [V], hence the DC current through it equals $I_R = \frac{V}{R}$[A]. Let us now replace this resistor with another one that has a resistance $R + \Delta R$ instead of R. If $\Delta R \ll R$, the voltage across it will not significantly change. The current then changes from I_R to $I_R + I_\varepsilon = \frac{V}{R+\Delta R}$, where the error current I_ε is defined as:

$$I_\varepsilon = \frac{V}{R + \Delta R} - I_R. \tag{9.6}$$

This can be written as:

$$I_\varepsilon = I_R \left(\frac{1}{1 + \frac{\Delta R}{R}} - 1 \right). \tag{9.7}$$

Since $\frac{\Delta R}{R} \ll 1$, we may write:

$$\frac{1}{1 + \frac{\Delta R}{R}} \approx 1 - \frac{\Delta R}{R}, \tag{9.8}$$

from which we obtain:

$$I_\varepsilon = I_R \left(\frac{\Delta R}{R} \right) \tag{9.9}$$

If $\frac{\Delta R}{R} \ll 1$ and the probability density function of R is symmetrical around its mean value, the probability density function of I will have the same shape as that of R. Hence, if R has a Gaussian distribution, the distribution function of the current I will also be Gaussian and the relative standard deviation of I will equal that of R.[7]

[7] The relative standard deviation of a random variable is the ratio of its standard deviation to the mean value.

Example 9.4

In this example, we will derive a symbolic expression for the standard deviation of the voltage at the output of a resistive voltage divider, given the standard deviation of the resistor values and of the driving voltage. To this end, consider the circuit shown in Figure 9.11. The standard deviation of the voltage V_s is $\sigma_1 V_s$, where σ_1 is the relative standard deviation of V_s, say, for example, 1%. The same holds for σ_2 and σ_3; these are the relative standard deviations of R_a and R_b, respectively. In this example, we assume no correlation between the resistor values, nor between the voltage V_s and the resistor values.

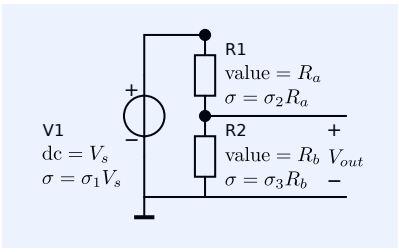


Figure 9.11: Resistive divider with driving voltage source. All element values have a Gaussian distribution function.

The resistor tolerances can be transformed into error currents as described above. According to expression (9.9), the standard deviation σ_{I_1} of the current of I_1 can be obtained as:

$$\sigma_{I_1} = \sigma_2 \frac{V_s}{R_a + R_b}. \tag{9.10}$$

The standard deviation of the current I_2 can be obtained in a similar way:

$$\sigma_{I_2} = \sigma_3 \frac{V_s}{R_a + R_b}. \tag{9.11}$$

The results are shown in Figure 9.12A.

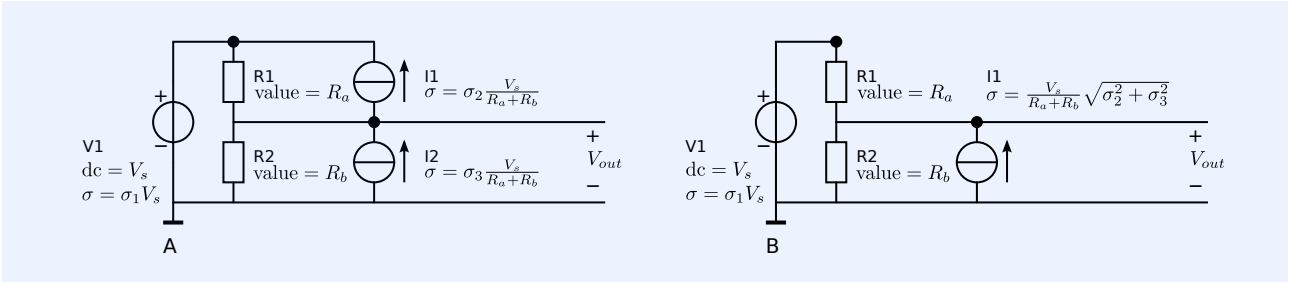


Figure 9.12: Resistor tolerances translated into error current sources.

If we assume a symmetrical distribution function for the resistor values, the signs of the error currents are not important. Since no correlation exists, we may calculate the variance[8] *$\sigma^2_{V_{out}}$ of the output voltage V_{out} as the sum of the contributions of the individual sources. Each of those contributions equals the product of the variance of the independent source and the square of its conversion gain to V_{out}. The conversion gain of the error currents I_1 and I_2 to the output voltage V_{out} equals the parallel connection $R_1 /\!/ R_2$.*

[8] The variance is the squared value of the standard deviation.

In this way, we may write:

$$\sigma^2_{V_{out}} = \left(\sigma_2^2 + \sigma_3^2\right)\left(\frac{V_s}{R_a + R_b}\right)^2\left(\frac{R_a R_b}{R_a + R_b}\right)^2 + \sigma_1^2 V_s^2\left(\frac{R_b}{R_a + R_b}\right)^2. \quad (9.12)$$

The standard deviation $\sigma_{V_{out}}$ of the output voltage V_{out} can thus be obtained as:

$$\sigma_{V_{out}} = V_s \frac{R_b}{R_a + R_b}\sqrt{\left(\frac{R_a}{R_a + R_b}\right)^2\left(\sigma_2^2 + \sigma_3^2\right) + \sigma_1^2}. \quad (9.13)$$

The relative standard deviation $\frac{\sigma_{V_{out}}}{V_{out}}$ of the output voltage is found to be

$$\frac{\sigma_{V_{out}}}{V_{out}} = \sqrt{\left(\frac{R_a}{R_a + R_b}\right)^2\left(\sigma_2^2 + \sigma_3^2\right) + \sigma_1^2}. \quad (9.14)$$

We may check this result by considering a few special cases:

1. *If $R_a = 0$, the relative tolerance equals that of the source. This is as expected, because in that case, the output voltage equals the source voltage!*
2. *With $R_b \ll R_a$ and equal value for all tolerances ($\sigma = \sigma_1 = \sigma_2 = \sigma_3$), the relative tolerance approaches a maximum value of $\sigma\sqrt{3}$. This is considerably less than we would have obtained from a worst-case scenario. A worst-case approximation with $+\delta$ relative error for V_s and R_b, and a relative error of $-\delta$ for R_a, yields a relative error $\delta_{V_{out}}$ for V_{out} of about 3δ.*

In the next example, we will demonstrate the application of SLiCAP for the symbolic determination of the standard deviation of the voltage at the output of the resistive divider.

Example 9.5

The SLiCAP *netlist file of the circuit from Figure 9.11 is listed below:*

```
"Voltage divider"
* File: vDivider.cir
* SLiCAP netlist file
V1 1 0 V dc={V_S} dcvar={(sigma_1*V_S)^2}
R1 1 out r value={R_a} dcvar={(sigma_2*R_a)^2}
R2 out 0 r value={R_b} dcvar={(sigma_3*R_b)^2}
.end
```

The variance of the supply voltage requires the use of the model 'V' for the voltage source. The mean value of a voltage source, or simply its DC voltage, can be defined with the model parameter 'dc', and the variance of the voltage of this source can be defined with the parameter 'dcvar'. The units of the variance are [V²]. Here, it has been defined as the square of the product of the DC voltage and the relative standard deviation σ_1.

Statistical properties for a resistor can be defined by using the model 'r' and the parameters 'value' for the mean resistance and 'dcvar' for the variance of the resistance in [Ω²]. The relative standard deviation of R_1 and R_2 are σ_2 and σ_3, respectively.

Below, the listing of the SLiCAP *script.*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: vDivider.py

from SLiCAP import *

fileName = 'vDivider.cir'
i1 = instruction()          # Creates an instance of an instruction object
i1.setCircuit(fileName)     # Checks and defines the local circuit object and
                            # sets the index page to the circuit index page
```

```
11 #
12 htmlPage('Netlist and circuit data');
13 netlist2html(fileName);
14 elementData2html(i1.circuit);
15 #
16 htmlPage('DC variance analysis');
17 i1.setDetector('V_out')
18 i1.setSimType('symbolic')
19 i1.setGainType('vi')
20 i1.setDataType('dcvar')
21 result = i1.execute()
22 dcVar2html(result)
```

Lines $15 \cdots 21$ *of the* SLiCAP *script define and execute the instruction for symbolic evaluation of the variance of the detector voltage and display the results on a separate html page.*

DC variance analysis

Symbolic dcvar analysis results

DC solution of the network

$$\begin{bmatrix} I_{V1} \\ I_{R2} \\ I_{R1} \\ V_1 \\ V_{out} \end{bmatrix} = \begin{bmatrix} -\frac{V_S}{R_a+R_b} \\ \frac{V_S}{R_a+R_b} \\ \frac{V_S}{R_a+R_b} \\ V_S \\ \frac{R_b V_S}{R_a+R_b} \end{bmatrix}$$

Detector-referred variance

$$\sigma^2_{out} = \frac{R_a^2 R_b^2 V_S^2 \sigma_2^2}{(R_a+R_b)^4} + \frac{R_a^2 R_b^2 V_S^2 \sigma_3^2}{(R_a+R_b)^4} + \frac{R_b^2 V_S^2 \sigma_1^2}{(R_a+R_b)^2} \left[\mathrm{V}^2\right]$$

Contributions of individual component variances

Variance of source: I_R1

Source variance: $\frac{V_S^2\sigma_2^2}{(R_a+R_b)^2}$ $\left[\mathrm{A}^2\right]$

Detector-referred: $\frac{R_a^2R_b^2V_S^2\sigma_2^2}{(R_a+R_b)^4}$ $\left[\mathrm{V}^2\right]$

Variance of source: I_R2

Source variance: $\frac{V_S^2\sigma_3^2}{(R_a+R_b)^2}$ $\left[\mathrm{A}^2\right]$

Detector-referred: $\frac{R_a^2R_b^2V_S^2\sigma_3^2}{(R_a+R_b)^4}$ $\left[\mathrm{V}^2\right]$

Variance of source: V1

Source variance: $V_S^2\sigma_1^2$ $\left[\mathrm{V}^2\right]$

Detector-referred: $\frac{R_b^2V_S^2\sigma_1^2}{(R_a+R_b)^2}$ $\left[\mathrm{V}^2\right]$

Figure 9.13: SLiCAP analysis results.

The output html page is shown in Figure 9.13. It also shows the DC solution of the network. With the aid of this solution, SLiCAP *translates the resistor tolerances into error current sources, just as discussed above. The current sources that*

model the errors of R1 and R2 have been named I_R1 and I_R2, respectively. Their contributions to the detector-referred variance are listed in the section "Contributions of individual variances" of the output page.

9.3.2 Controller bias imperfections

We will now investigate the influence of controller bias imperfections on the biasing solution of the network. In fact, we want to investigate the way in which and to what extent the offset voltages, the bias currents and the offset current of the controller affect the DC output voltage of the circuit from Figure 9.9.

If we know this, we will be able to define budgets for these offset and bias quantities and add them to the search criteria for operational amplifiers that may be used for implementation of the controller.

In order to keep things as simple as possible, we need a model of a controller that models these imperfections alone.

Figure 9.14 shows such a possible model. It consists of a nullor which provides the infinite available power gain and sources that model the imperfections of interest. The voltage source V1 models the equivalent-input offset voltage. The mean value of this source is taken as zero and the standard deviation σ_{vo} [V]. Zero mean value for this source is a good starting point, because manufacturers want to create zero offset error, but due to device tolerances, there will remain a small error that can be positive or negative.

The current source I2 models the offset current. For similar reasons as those related to the offset voltage source, the mean value of this source is also zero. Its standard deviation is taken as σ_{io}.

The two input bias currents that are required for the operation of the input stage have been taken as equal. In this model, they are correlated. This correlation is achieved by deriving them from one source: I1. This source has been given a mean value I_b and a standard deviation $\sigma_{ib} I_b$. Correlation between the two input bias currents usually occurs in operational amplifiers with a BJT input stage. The input bias currents of the JFET MOS input stage is usually very low. The bias currents of a MOS operational amplifier is usually dominated by leakage currents in the protection diodes. Correlation between the two input bias currents in those types of operational amplifiers may be considerably lower.

Unfortunately, correct modeling of the input bias currents and offset current requires manufacturing data that are not always provided to a level of detail that would include these correlations.

The SLiCAP netlist of the circuit from Figure 9.14 can be found in the library file: SLiCAP.lib. It is shown below:

```
.subckt ABCD 1 2 3 4 A={A} B={B} C={C} D={D}
N1 1 5 3 7
E1 7 8 1 2 {A}
H1 8 4 6 5 {B}
G1 3 4 1 2 {C}
F1 3 4 2 6 {D}
.ends
```

We will study the influence of the bias imperfections of the controller using the circuit from Figure 9.15. The biasing errors are a result of supply voltage errors and errors in the resistive divider, as shown in the previous results (Figure 9.13). This has been done to keep things simple through a step-by-step approach. At a later stage, we will evaluate the total biasing error, including these contributions.

The SLiCAP netlist of this circuit is given below:

```
VampBias
* file: VampBias.cir
```

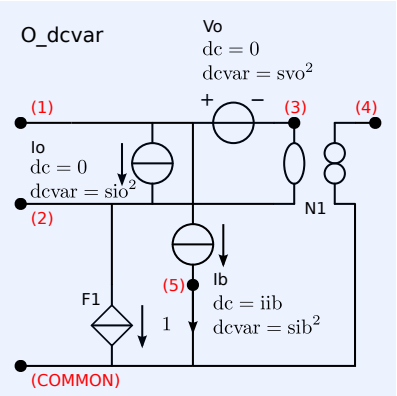


Figure 9.14: Operational amplifier modeled as a nullor with added bias and offset sources.

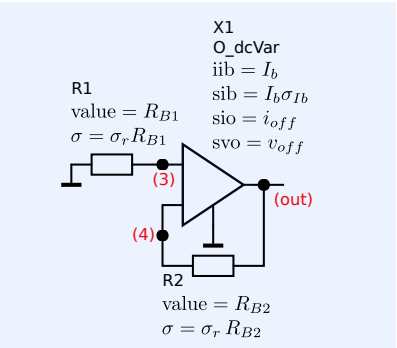


Figure 9.15: Circuit for studying the influence of the bias imperfections of the controller. It has been derived from the circuit from Figure 9.10. The power suppy voltage has been set to zero. The value of RB1 is: $R_a \,/\!/\, R_b + R_c$. The value of R2 is $19R$. The nullor has been replaced with the nullorWithBias sub circuit from Figure 9.14.

```
* SLiCAP netlist file
R1 0 3 r value={R_B1*sigma_r)^2}
R2 out 4 r value={R_B2} dcvar={(R_B2*sigma_r)^2}
X1 3 4 out 0 0 N_dcvar sib={I_b*sigma_Ib} sio={i_off} svo={v_off}
+ iib={I_b}
.end
```

The relative standard deviation of the resistor values have been set to σ_r [-]. The absolute standard deviations of the offset voltage and of the offset current of the controller have been set to v_{off} [V] and i_{off} [A], respectively. The bias current of the controller has been set to I_b [A] with a relative standard deviation of σ_{Ib} [-]. If desired, the temperature dependencies of these parameters can be added, but that has not been done at this stage.

The SLiCAP script for evaluation of the biasing errors of this circuit is:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: VampBias.py

from SLiCAP import *

fileName = 'VampBias.cir'
i1 = instruction()           # Creates an instance of an instruction object
i1.setCircuit(fileName)      # Checks and defines the local circuit object and
                             # sets the index page to the circuit index page
i1.setDetector('V_out')
i1.setSimType('symbolic')
i1.setGainType('vi')

# Obtain the DC detector voltage
i1.setDataType('dc')
detDC = i1.execute().dc

# Obtain the detector-referred variance
i1.setDataType('dcvar')
detVar = i1.execute().ovar

htmlPage('Biasing results')
head2html('DC detector voltage')
text2html("The DC detector voltage is:")
eqn2html("V_out", detDC, units="V")
head2html('Variance of the DC detector voltage')
text2html("The variance of the DC detector voltage is:")
eqn2html("(sigma_V_out)**2", detVar, units="V**2")
```

Biasing results

DC detector voltage

The DC detector voltage is:

$$V_{out} = -I_b R_{B1} + I_b R_{B2} \ [\mathrm{V}] \tag{1}$$

Variance of the DC detector voltage

The variance of the DC detector voltage is:

$$\sigma^2_{Vout} = I_b^2 R_{B1}^2 \sigma_r^2 + I_b^2 R_{B2}^2 \sigma_r^2 + I_b^2 \sigma_{Ib}^2 (-R_{B1} + R_{B2})^2 + i_{off}^2 (R_{B1} + R_{B2})^2 + v_{off}^2 \ [\mathrm{V}^2] \tag{2}$$

Figure 9.16: SLiCAP analysis results.

The contributions to the Detector-referred variance are shown in Figure 9.16.

The mean value of the DC output voltage V_{out} of the circuit is $I_b\,(R_{B2} - R_{B1})$. It shows the effect of the bias current I_b on V_{out}. It is zero if $R_{B2} = R_{B1}$ and the voltage drop across R_{B1} caused by the bias current equals the voltage drop across R_{B2}.

The contributions to the output variance are:

1. $I_b^2 \sigma_r {R_{B1}}^2$: this term represents an error voltage due to the combination of a current flow I_b through the resistor R1 and the variance of its resistance R_{B1}.

2. $I_b^2 \sigma_r {R_{B2}}^2$: this term represents an error voltage due to the combination of a current flow I_b through the resistor R2 and the variance of its resistance R_{B2}.

3. $I_b^2 \sigma_{Ib} (R_{B1} - R_{B2})^2$: this term shows the contribution of the variance of the bias current to the variance of V_{out}. Similar to the influence of the bias current on the mean DC voltage V_{out}, this influence can be made zero if $R_{B2} = R_{B1}$.

4. $i_{off}^2 (R_{B1} + R_{B2})^2$: the offset current of the controller flows through the series connection of R6 and R2. For many operational amplifiers, the order of magnitude of the offset current equals that of the bias current. This limits the effect of the biasing error compensation found in the DC output voltage V_{out}.

5. v_{off}^2: the variance of the equivalent-input offset voltage of the controller directly adds to that of V_{out}. This is because the circuit acts as a unity-gain voltage follower for any voltage in series with the input of the controller.

9.3.3 Total biasing error

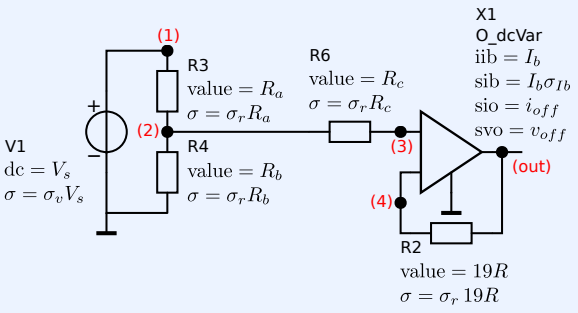


Figure 9.17: Circuit for evaluation of the total biasing error of the voltage amplifier from Figure 9.9. It is equal to the circuit from Figure 9.10 in which the nullor has been replaced with the nullor with bias sources from Figure 9.14.

In the previous sections, we have evaluated two contributions to the biasing error of the voltage amplifier. We investigated errors resulting from supply voltage tolerances and resistor tolerances, as well as errors resulting from biasing imperfections of the controller. We will now evaluate the total biasing error resulting from both contributions. Figure 9.17 shows the circuit for evaluation of the total biasing error. In this circuit, the relative standard deviation of the resistors is σ_r and that of the supply voltage equals σ_v.[9]

The netlist of this circuit is shown below.

[9] All resistors have the same relative standard deviation σ_r. But the resistance errors are uncorrelated.

```
VampBiasTotal
* file: VampBiasTotal.cir
* SLiCAP netlist file
V1 1 0 V dc={V_S} dcvar={(sigma_v*V_S)^2}
R3 1 2 r value={R_a} dcvar={(R_a*sigma_r)^2}
R4 2 0 r value={R_b} dcvar={(R_b*sigma_r)^2}
R6 2 3 r value={R_c} dcvar={(R_c*sigma_r)^2}
R2 out 4 r value={19*R} dcvar={(19*R*sigma_r)^2}
X 4 3 out 0 O_dcvar
+ sib={I_b*sigma_Ib}
+ sio={i_off}
+ svo={v_off}
+ iib={I_b}
.param R_b={R_a}
.end
```

9.3.4 Biasing design limits and budgets

We will now demonstrate the way in which show-stopper values for the different contributers for the output variance can be determined. We will do this for a simplified circuit.

1. The DC solution of the network from Figure 9.11 does not significantly change as a result of the offset and bias sources to the controller if $I_b \frac{R_a R_b}{R_a+R_b} \ll I_b R_c$, or $\frac{R_a R_b}{R_a+R_b} \ll R_c$.

2. From the DC analysis, we concluded that we required $R_a = R_b$

With the above assumtions the expression for the detector-referred variance becomes:

$$\sigma^2_{Vout} = \sigma^2_r \frac{V_s^2}{8} + \sigma^2_v \frac{V_s^2}{4} + v^2_{off} + i^2_{off} (38R)^2 + \sigma^2_r 2 (19R)^2 I^2_b. \tag{9.15}$$

The script for doing this is:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: VampBiasTotal.py

from SLiCAP import *

fileName = 'VampBiasTotal.cir'
i1 = instruction()          # Creates an instance of an instruction object
i1.setCircuit(fileName)     # Checks and defines the local circuit object and
                            # sets the index page to the circuit index page
i1.setDetector('V_out')
i1.setSimType('symbolic')
i1.setGainType('vi')

# Obtain the DC detector voltage
i1.setDataType('dc')
detDC = i1.execute().dc

# define the symbols in the python environment
R_a, R_b, R, R_c, alpha = sp.symbols('R_a, R_b, R, R_c, alpha')

detDC = detDC.subs({R_b:R_a, R_c:19*R-R_a/2})

# Assume R_a << R:
# Take R_a=alpha*R and evaluate lim (alpha --> 0)
detDC = detDC.subs(R_a, alpha*R)
detDC = sp.limit(detDC, alpha, 0)

i1.setDataType('dcvar')
detVar = i1.execute().ovar

# Use R_b=R_a and R_c = 19*R- R_a//R_b
detVar = detVar.subs({R_b:R_a, R_c:19*R-R_a/2})
# Assume R_a << R:
# Take R_a=alpha*R and evaluate lim (alpha --> 0)
detVar = detVar.subs(R_a, alpha*R)
detVar = sp.limit(detVar, alpha, 0)

# Make the report
htmlPage('Biasing results')
head2html("Simplifications")
text2html("We will simplify the expressions for the DC voltage and the " +
          " detector-referred variance using the following assumptions:" +
          "<ol><li>$R_c=19R-\\frac{R_aR_b}{R_a+R_b}$</li>" +
          "<li>$R_a=R_b$</li>" +
          "<li>$R_a\\ll R$</li></ol>")
head2html('DC detector voltage')
text2html("The DC detector voltage is:")
eqn2html("V_out", detDC, units="V")
head2html('Detector referred variance')
text2html("The variance of the DC detector voltage is:")
eqn2html("(sigma_V_out)**2", detVar, units="V**2")
```

The results are shown in Figure 9.18. They comply with those from 9.15.

Biasing results

Simplifications

We will simplify the expressions for the DC voltage and the detector-referred variance using the following assumptions:

1. $R_c = 19R - \frac{R_a R_b}{R_a + R_b}$
2. $R_a = R_b$
3. $R_a \ll R$

DC detector voltage

The DC detector voltage is:

$$V_{out} = 0.5 V_S \ [\mathrm{V}] \tag{1}$$

Detector referred variance

The variance of the DC detector voltage is:

$$\sigma^2_{Vout} = 722 I_b^2 R^2 \sigma_r^2 + 1444 R^2 i_{off}^2 + \frac{V_S^2 \sigma_r^2}{8} + \frac{V_S^2 \sigma_v^2}{4} + v_{off}^2 \ [\mathrm{V}^2] \tag{2}$$

Figure 9.18: SLiCAP analysis results.

Now that we have the means to find symbolic expressions for the contributions of resistor tolerances, power supply tolerances and controller biasing errors to the variance of the output voltage, we are able to set up design limits for such types of error sources.

The general way to find a design limit for one specific parameter that contributes to a certain error, is to assign the complete budget for this error for this contribution. For example, if the standard deviation of V_{out} is allowed to be 100mV, the design limit for the standard deviation σ_v of power supply voltage of the voltage amplifier is 200mV. This directly follows from the second term of (9.15). A design limit for the resistor tolerances can be found in a similar way. If we ignore the contribution of the bias current I_b, only the first term in (9.15) describes a contribution to σ_{Vout} due to the resistor tolerance σ_r. With a power supply voltage of 5V, we then obtain: $0.1^2 = \sigma_r^2 \frac{25}{8}$, which yields a maximum value for the resistor tolerances: $\sigma_r = 5.7\%$. A design limit for v_{off} can be found in a similar way. The definition of design limit for i_{off} depends on the selected value for R. This selection will primarily be based on the noise performance of the amplifier. The design limit for I_b strongly interacts with the selection of R and σ_r.

A good starting point for deriving budgets for multiple parameters that together affect one performance aspect, is to assume equal error contributions of all parameters. If, for some reason, such a budget appears to be too tight for one parameter, it can be relieved at the cost of a tighter budget for another one. Design limits, however, can never be relaxed. Non-compliance with a design limit means a *show-stopper*[10] for the design and requires another design concept.

[10] Action, condition, event, or problem that is serious enough to halt an activity, program, or process until it is resolved (http://www.businessdictionary.com).

9.4 Application of error reduction techniques

In the previous sections, we have studied basic amplifier biasing techniques. We found means to define the voltage levels and provide bias currents, and we evaluated errors resulting from device tolerances and imperfections. If

the biasing errors are too large, we have to apply so-called error reduction techniques to reduce them. This section is devoted to the application of such techniques. We will study them at a conceptual level only.

One error reduction technique that is often applied during biasing is compensation. As a matter of fact, adding level shifts and bias current sources to the design is a form of additive compensation. It has already been discussed in the previous sections and is considered as a basic biasing technique.

If remaining errors are too large, we need more powerful error reduction techniques to reduce them. In section 9.4.1, we will introduce negative-feedback biasing. We have already applied negative feedback for the design of amplifier concepts. In situations in which errors due to biasing can be distinguished from signal excursions, negative feedback biasing may offer further improvement.

9.4.1 Negative feedback and auto-zero biasing

Similar to the design of negative feedback amplifier configurations, the conceptual design of negative feedback biasing starts with measurement of the bias quantity to be fixed. In the amplifier from Figure 9.1 this is the DC operating voltage at the input of the ADC. In order to obtain maximum and equal positive and negative digital signal excursions at its output, the DC input voltage of the ADC should be set to 2V. During our study of biasing errors in section 9.3, we found all kinds of error sources that contribute to an error in this DC voltage. The idea behind negative-feedback biasing is that we measure this deviation from the desired operating point and apply a control signal to the amplifier to correct it. However, the output signal of the amplifier itself also causes such deviations, and those should not be compensated for! Hence, negative feedback biasing can only be applied if a distinction can be made between deviations caused by the signal and deviations caused by operating point changes. Such distinctions can be made in the frequency domain or in the time domain.

Negative-feedback biasing

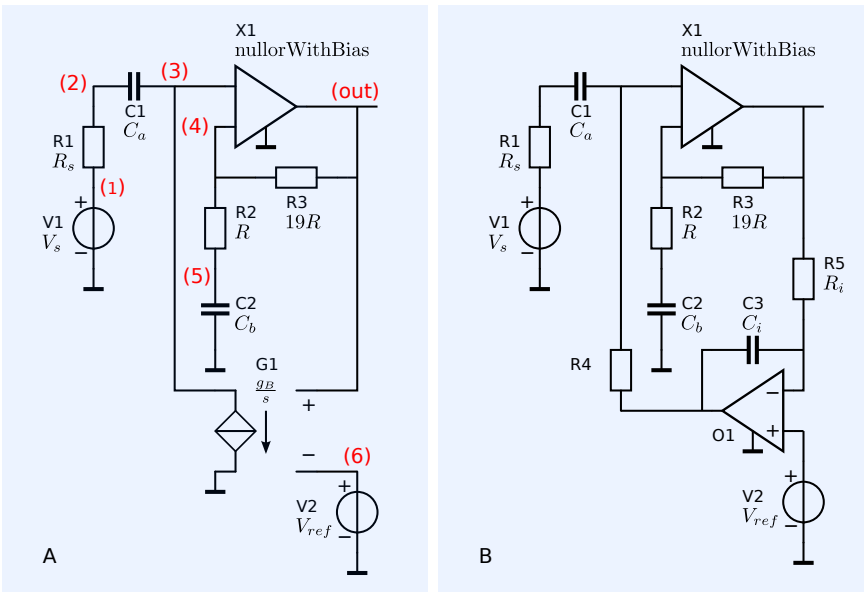


Figure 9.19:
A. Concept of negative feedback biasing applied to the amplifier from Figure 9.9
B. Implementation of the circuit with an integrating transimpedance amplifier as controller.

Changes in the DC operating point are caused by supply changes, device tolerances, biasing errors and temperature changes. These changes are usu-

ally slow and have only frequency components at very low frequencies. If those frequencies do not occur in the signal, a distinction between biasing and signal components can be made in the frequency domain.

Figure 9.19 shows a possible concept of negative feedback biasing for the voltage amplifier from Figure 9.19A.

The integrating transadmittance amplifier G1 is the *bias loop controller*. It acts as a first-order low-pass filter with an infinite DC gain. Hence, for DC, it can be replaced with a nullor. As a consequence, the DC output voltage of the amplifier will equal V_{ref}. Since both the input impedance and the output impedance are infinite, the source-to-load transfer, the noise and the power efficiency will not be affected by the bias control loop. This, of course, is under the assumption that the integrator gain g_B has been properly designed. In section 11.4.4, we will study the design of the high-pass cut-off frequency and in Chapter 12, we will discuss the design of the high-pass response and see the way in which the stability of negative feedback biasing can be assured. In the following example, we will derive symbolic expressions for the DC output voltage and the voltage transfer from the source to the output.

Example 9.6

The netlist of the amplifier from Figure 9.19A is shown below:

```
VampFeedbackBiasTotal
* file: VampFeedbackBiasTotal.cir
* SLiCAP netlist file
V1 1    0 V value = {V_s}
V2 6    0 V value = 0 dc={V_ref}
R1 1    2 {R_s}
R2 4    5 r value={R} dcvar={(R*sigma_r)^2}
R3 out  4 r value={19*R} dcvar={(19*R*sigma_r)^2}
C1 2    3 {C_a}
C2 5    0 {C_b}
X1 3 4 out  0 O_dcvar ; amplifier controller
+ sib={I_b*sigma_Ib}
+ sio={i_off}
+ svo={v_off}
+ iib={I_b}
G1 3 0 out 6 {g_B/s}  ; bias loop controller
.end
```

Below, the SLiCAP *script for evaluation of the DC output voltage, the source-to-load transfer and the high-frequency approximation of the source-to-load transfer.*

Feedback biasing

The DC output voltage V_{outDC} is:

$$V_{outDC} = V_{ref} \tag{1}$$

The voltage transfer A_v from source to load is:

$$A_v = \frac{C_a\left(20C_bRs^3 + s^2\right)}{g_B\left(\frac{C_aC_bRs^3}{g_B} + s\left(C_aR_s + 20C_bR\right) + 1 + \frac{s^2(20C_aC_bRR_sg_B + C_a)}{g_B}\right)} \tag{2}$$

For high frequencies this can be written as:

$$A_v = 20 \tag{3}$$

Figure 9.20: SLiCAP analysis results.

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File VampFeedbackBiasTotal.py

from SLiCAP import *

fileName = 'VampFeedbackBiasTotal'
i1 = instruction()              # Creates instance of instruction object
```

```
i1.setCircuit(fileName + '.cir') # Checks, defines the local circuit object,
                                 # and sets the index page to the circuit
                                 # index page
i1.setSimType('symbolic')
i1.setGainType('vi')
i1.setDataType('dc')
i1.setSource('V1')
i1.setDetector('V_out')
result = i1.execute()

htmlPage('Feedback biasing')
text2html('The DC output voltage $V_{outDC}$ is:')
eqn2html('V_outDC', result.dc)

# Laplace transfer function with feedback biasing

i1.setGainType('gain')
i1.setDataType('laplace')
result = i1.execute()
text2html('The voltage transfer $A_v$ from source to load is:')
eqn2html('A_v', normalizeRational(result.laplace))
hf = sp.limit(result.laplace, ini.Laplace, 'oo')
text2html('For high frequencies this can be written as:')
eqn2html('A_v', hf)
```

The html page with the results is shown in Figure 9.20.
The results comply with our expectations:

1. *The DC output voltage equals that of the reference voltage source.*
2. *The transfer has three poles and three zeros.*
3. *The high-frequency transfer equals the ideal gain of* 20.

Figure 9.19B shows a possible implementation of this concept. The transadmittance gain g_B is defined by the gain of the transimpedance integrator and the values of R4 and R5:

$$g_B = \frac{1}{R_B R_i C_i}. \tag{9.16}$$

Auto-zero biasing

If, during a short time interval, the load signal is not of interest, this time interval may be used to correct biasing errors with a feedback control signal. The corrected biasing should be maintained during the time of interest of the signal, which requires a memory element. If the load signal is not allowed to change due to this auto-zero process, a second memory element is required for storing the signal and passing it to the load during the auto-zero time interval. This principle is applied in so-called auto-zero operational amplifiers.

9.4.2 Modulation and demodulation techniques

Compensation, negative feedback and auto-zero techniques do not affect the way in which the information is embedded in the signal: they adapt the circuit to the signal. Another approach for reducing the amount of signal processing errors is to do the opposite: adapt the signal to the circuit. As a matter of fact, digitization can be seen as such a technique. By reducing the number of signal levels to two, the signal is made insensitive to nonlinearities that occur in the circuit. Modulation and demodulation techniques also belong to this class of error reduction techniques. With these techniques, the frequency range of a signal can be changed, and this may be an interesting option for the reduction of biasing errors. If frequency components of the signal are in the same range as those of temperature variations, changes in the

quiescent operating point of an amplifier due to temperature variations cannot be distinguished from signal excursions. Bringing the frequency range of the signal outside the frequency range of temperature changes and of other possible disturbances will then increase the signal-to-noise ratio. After the modulated signal has been amplified to a level at which temperature variations can be neglected, it can be demodulated to its original frequency band. This principle is applied in so-called chopper-stabilized amplifiers.

9.5 Common-mode biasing

If an amplifier shows natural two-port behavior, the common-mode port voltages need to be defined. This is because the floating ports of a natural two-port have an infinite common-mode impedance, and any common-mode current caused by external noise sources or temperature variations may drive the common-mode port voltage outside its desired operating range. In practice, this operating range will be limited by a power supply voltage or by physical breakdown voltages of circuit components in the amplifier.

9.5.1 AC coupling

Figure 9.21 shows a concept with a common-mode port impedance of zero for both ports. The input common-mode voltage has been set to ground level and the output common-mode level has been set to V_{cm} with respect to the ground.

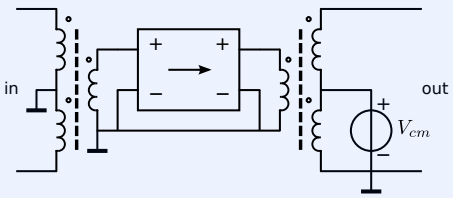


Figure 9.21: Balanced transformer coupled amplifier with zero input common-mode impedance and output common-mode impedance, zero input common-mode voltage and its output common-mode voltage set tos V_{cm}.

Alternatively, one may consider the use of brute-force techniques as they have been demonstrated for the biasing of the input port of the operational amplifier (see section 9.2.3).

9.5.2 DC-coupled floating port amplifiers

There are four techniques for defining DC common-mode voltages of floating ports:

1. Brute force methods

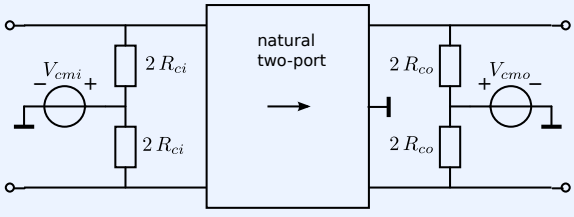


Figure 9.22: Brute-force common-mode biasing of the input port and the output port of a natural two-port.

We speak of brute force techniques if the common-mode voltages are fixed by simply connecting the port terminals to a voltage source by means of an impedance that allows for DC transfer. Figure 9.22 shows the principle with resistors. The common-mode input resistance of this circuit equals R_{ci}

and the common-mode output resistance R_{co}. The common-mode input voltage and the common-mode output voltage have been set to V_{cmi} and V_{cmo}, respectively. Brute force fixing of the common-mode input voltage of a floating port affects the differential-mode impedance of that port. In order to keep the possible adverse effects on the signal processing quality as small as possible, the sum of the two brute force impedances should be much larger than the differential driving point impedance.[11]

[11] The driving point impedance of a port is the in-circuit measured impedance at that port.

2. Feed forward techniques

   Feed forward techniques make use of the existing nonzero common-mode transfer of an amplifier or an amplifier stage.

3. Figure 9.23 combines brute force common-mode biasing of the input port of a two-port with feed forward biasing of its output port. An example of a feed forward biasing of the output port of a balanced transresistance amplifier has been shown in Figure 9.24. The common-mode output voltage of this circuit is fixed by both I_{cm} and V_{cmi}:

$$V_{ocm} = V_{cmi} + I_{cm} R_f. \tag{9.17}$$

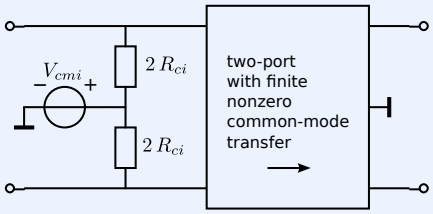


Figure 9.23: Feed forward common-mode biasing of the output port of a two-port with a finite, nonzero common-mode transfer.

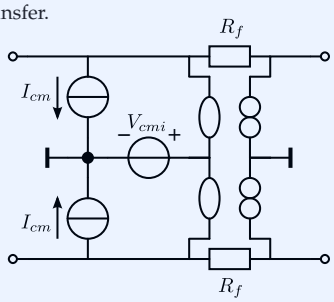


Figure 9.24: Feed forward common-mode biasing of the output port of a balanced transresistance amplifier.

4. Local common-mode feedback

   Local common-mode feedback is a technique in which the common-mode voltage or current of a port is measured, compared with a reference, and controlled by adding a common-mode current or voltage at the same port. This is shown in Figure 9.25.

   The common-mode input impedance of this circuit equals $1/G_{cm}$ while its differential-mode impedance is infinity. Figure 9.26 shows the concept with two transimpedance amplifiers.

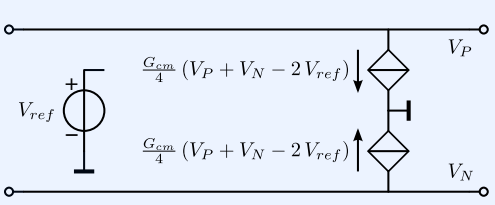


Figure 9.25: Principle of common-mode biasing of a floating port with the aid of local feedback.

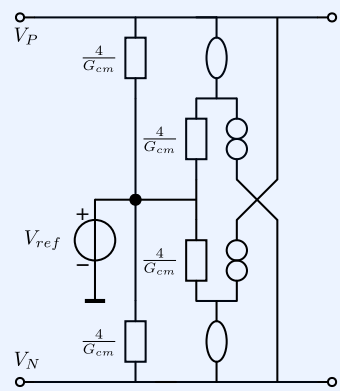


Figure 9.26: Concept with two nullors for implementation of the feed-forward common-mode biasing from Figure 9.25.

5. Over-all common-mode feedback

   With over-all common-mode feedback, a common-mode output quantity is measured and controlled by a common-mode input quantity. This is only possible if the two-port has a nonzero common-mode transfer from the controlling quantity to the controlled quantity.

   Figure 9.27 shows an arrangement in which the common-mode output voltage of a two-port is controlled by its common-mode input current. These techniques will be discussed in more detail in the volume about transistor-level design.

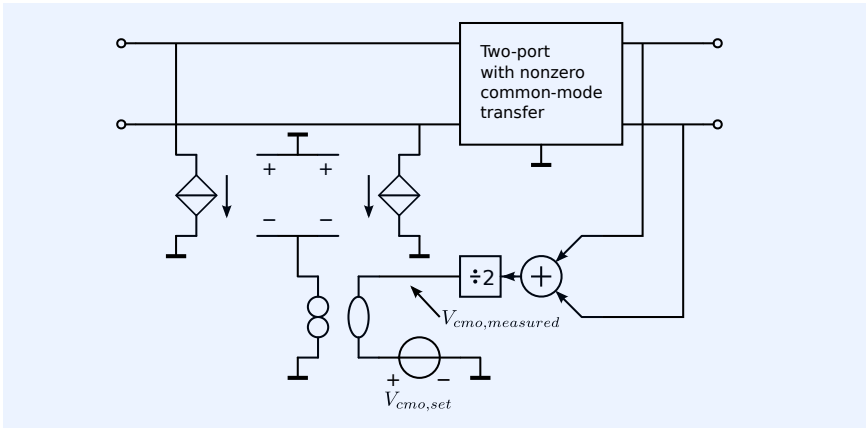


Figure 9.27: Common-mode biasing of a floating port with the aid of over-all common-mode feedback.

The common-mode output voltage is measured, compared with a set point, and contolled by inserting a common-mode input current. The two-port needs to have a nonzero common-mode transresistance.

10 Modeling of negative feedback circuits

10.1 Introduction

In Chapter 7, we showed that low-noise and power-efficient amplifiers can be realized with the aid of negative feedback. When applying negative feedback, the input impedance, the output impedance and the source-to-load transfer of an amplifier obtain their required value with the aid of feedback networks around high-gain loop amplifiers or *controllers*. If we use nullors as ideal controllers, the transfer characteristics of the amplifier are completely determined by the feedback networks. In fact, when designing this so-called feedback configuration, we perform the first design step in amplifier design. The result of this step is a conceptual design in which primary performance aspects such as the ideal gain, the port impedances and the port isolation configuration of the amplifier have been designed.

In the second design step, we would like to design the controllers without changing anything that has been designed during the first step.

10.1.1 Two-step design approach

If the design of a controller can be done without changing the amplifier concept, we have a straightforward two-step design approach with no iterative loops. Such a design approach requires decomposition of the allowed information processing errors in two independent contributions that can be designed during the two subsequent steps:

1. Error contributions due to imperfections of the feedback networks

   These error contributions have been discussed in Chapter 7. Since the feedback network determines the ideal gain of the negative feedback amplifier, tolerances of devices in the feedback network result in tolerances of the ideal gain. We have also seen that application of passive devices in the feedback network may result in an increase of noise and a decrease in power efficiency. It will also be clear that nonlinear and/or dynamic behavior of feedback elements results in a nonlinear and/or dynamic ideal transfer of the negative feedback amplifier. Sometimes, such effects are intended. In logarithmic amplifiers, for example, the output quantity intentionally changes logarithmically with the input quantity. Similarly, active filters exhibit an intended dynamic behavior.

2. Error contributions due to imperfect implementation of the controller(s)

   The transfer of a negative feedback amplifier only equals its ideal gain if all of the controllers are nullors. In practice, this will never be the case. As a matter of fact, it doesn't have to be the case. In general, small deviations from the ideal transfer of the negative feedback amplifier can be allowed. Hence, a part of the total error budget can be reserved for error contributions resulting from imperfect behavior of the controller. In order to assign such error budgets, or to judge whether given error budgets are realistic, we need to know the way in which and to what extent the performance aspects of the controller affect those of the negative feedback amplifier. We have already studied the influence of controller noise and are able to assign budgets for the equivalent-input noise voltage and noise current source of the controller. We have not yet studied the influence of gain and bandwidth limitations of the controller, nor do we have a clear understanding of the way in which the nonlinearity that occurs in the controller manifests itself in the transfer of the negative feedback amplifier.

At this stage, we would like to have a modeling technique at our disposal that clearly relates all kinds of behavioral aspects of the controller to relevant behavioral aspects of the negative feedback amplifier. However, not every model that provides this insight is useful. We need to have a model that

supports our two-step design approach. Hence, it should split error contributions into two parts: changes in the ideal gain due to imperfections of the feedback network and deviations from this (changed) ideal gain due to controller imperfections.

10.1.2 This chapter

In this chapter, we will discuss various techniques for modeling of feedback systems and circuits. Some techniques will only briefly be discussed, while a model that supports the two-step design method will be discussed in more detail.

A model that is widely used for analysis of negative feedback systems is the feedback model by Black.[1] Black's feedback model is very useful during system-level design with building blocks that have unilateral transfer and that show no interaction. These conditions are usually not satisfied in electronic circuits. As a matter of fact, satisfaction of these conditions would put impractical constraints on the design of application-specific negative feedback amplifiers: they would introduce requirements that adversely affects the feasibility of the design. Black's feedback model will be discussed in section 10.2.

[1] Harold Black (1898-1996).

One feedback model that supports the two-step design approach is the asymptotic gain model. It will be presented in section 10.3. This model is based upon the superposition model, which models the behavior of circuits that exploit negative feedback as accurately as network theory.

A feedback model that is very suited for measurements and simulations of complete feedback circuits has been described by Middlebrook's *generalized feedback theorem*. It is suited for analysis rather than for synthesis. It will not be discussed in this book.

In this book we will use the asymptotic gain model to relate the performance aspects of the controller to those of the negative-feedback amplifier. This will be done in Chapter 11.

10.2 Black's feedback model

In 1927, Black built the first negative-feedback amplifier [Black1934][2]. He applied negative feedback to obtain linear and stable-gain repeater amplifiers for long-distance telephone systems. Black's patent was awarded in 1937 [Black1932][3]. The theoretical understanding of the high-frequency stability of negative feedback amplifiers was developed by Black in cooperation with Nyquist [Nyquist1932][4] and Bode [Bode1945][5].

[2] Harold S. Black. Stabilized feed-back amplifiers. *Electrical Engineering*, 53(1):114–120, January 1934

[3] H.S. Black. Wave Translation System, March 1932

[4] H. Nyquist. Regeneration theory. *Bell System Technical Journal*, 11:126–147, 1932

[5] H.W. Bode. *Network Analysis and Feedback Amplifier Design*. Van Nostrand, New York, 1945

The use of Black's feedback model for negative feedback systems is widespread. The model is shown in Figure 10.1. The model uses three basic building blocks: a feedback network, a subtracter and a loop amplifier or controller.

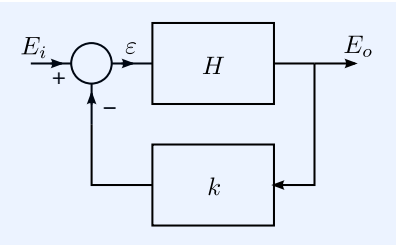


Figure 10.1: Black's feedback model.

10.2.1 Model description

We will now derive the input-output relation of a negative feedback system according to Black's feedback model.

Let E_i be the input signal and E_o be the output signal of a feedback system that consists of a feedback element with a transfer k, a controller with transfer H and a subtracter, as depicted in Figure 10.1. From Chapter 7, we know that the transfer k of the feedback element has been designed as the reciprocal gain of the desired input-to-output transfer of the feedback system. The output signal of the feedback element should be an accurate copy of the input signal of the system. Hence, the input signal of the controller, also

called the error signal ε, should approach zero. The error signal ε is the difference between the output signal of the feedback network and the input signal E_i. According to the block diagram from Figure 10.1 the error signal should satisfy two equations:

$$\varepsilon = E_i - kE_o, \tag{10.1}$$

$$\varepsilon = \frac{1}{H}E_o. \tag{10.2}$$

Hence, we may write:

$$\frac{1}{H}E_o = E_i - kE_o, \tag{10.3}$$

or:

$$E_o\left(\frac{1}{H} + k\right) = E_i, \tag{10.4}$$

from which we obtain the expression for the input-output relation of the feedback system:

$$\frac{E_o}{E_i} = \frac{H}{1 + Hk}. \tag{10.5}$$

If we consider H the gain of an amplifier, then we may say that negative feedback reduces the gain by a factor $(1 + Hk)$. The product of the gain of the controller and that of the feedback network is called the *loop gain* Hk. It is the gain enclosed in the loop that consists of the controller, the feedback element and the subtracter.

From a design point of view, we better consider H as the gain of a controller, that has to minimize the error between the E_i and the signal at the output of the feedback network. If the controller gain is infinite, the loop gain is infinite and the input-to-output transfer becomes the reciprocal transfer of the feedback network: the ideal gain of the negative feedback amplifier. This complies with the results from Chapter 7:

$$\lim_{H\to\infty}\left(\frac{E_o}{E_i}\right) = \frac{1}{k}. \tag{10.6}$$

Hence, a design-friendly notation that explicitly writes (10.5) as the product of the intended gain, and a factor that represents the effect of a limited controller gain, is:

$$\frac{E_o}{E_i} = \frac{1}{k}\left(\frac{Hk}{1 + Hk}\right). \tag{10.7}$$

10.2.2 Application of the model

Expression 10.5 describes the transfer from the input to the output of the system, rather than the transfer from the signal source to the load. This is a consequence of the modeling technique. Currents and voltages cannot easily be distinguished in these kinds of block diagrams.

The model also supposes ideal subtraction of the feedback signal from the input signal and no interaction between unilateral building blocks.

For a negative feedback voltage amplifier, for example, these conditions imply that the controller should have an infinite *CMRR*, infinite differential-mode and common-mode input impedances and zero output impedance. These, indeed, are prerequisites for voltage operational amplifiers, but they seriously complicate their design, because no electronic components with such properties exist.

In the following examples, we will show the application of Black's feed-

back model for electronic circuits.

Example 10.1 .

Figure 10.2 shows a passive-feedback voltage amplifier in which the controller has been modeled as a simple opamp: a voltage-controlled voltage source with voltage gain A_v. The SLiCAP *netlist of this circuit is:*

```
"Voltage amplifier with VCVS controller"
* file: vAmpBlack.cir
* SLiCAP circuit file
V1 1 0 {V_s}
R1 1 2 {R_s}
R2 3 0 {R_ell}
E1 3 0 2 4 {A_v}
R3 3 4 {R_a}
R4 4 0 {R_b}
.end
```

The following script can be used to obtain a symbolic expression for the transfer from `V1` *to the voltage across* `R2`*:*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: vAmpBlack.py

from SLiCAP import *

fileName = 'vAmpBlack'
prj = initProject(fileName)     # Creates the SLiCAP libraries and the
                                # project HTML index page
i1 = instruction()              # Creates an instance of an instruction object
i1.setCircuit(fileName+'.cir') # Checks and defines the local circuit object,
                                # and sets the index page to the project index
i1.setSource('V1')
i1.setDetector('V_3')
i1.setSimType('symbolic')
i1.setGainType('gain')
i1.setDataType('laplace')
result = i1.execute()

htmlPage('Voltage amplifier with VCVS controller')
text2html('The gain of the system is obtained as:')
V_ell, V_s = sp.symbols('V_ell, V_s')
eqn2html(V_ell/V_s, result.laplace)
```

The html page with the result of this run is shown in Figure 10.3.

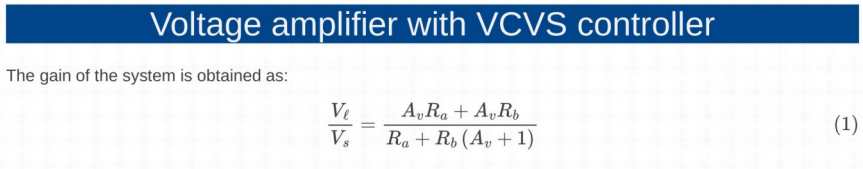


Figure 10.3: SLiCAP simulation results for the negative feedback voltage amplifier with VCVS controller.

The result can be written as

$$\frac{V_\ell}{V_s} = \frac{R_a + R_b}{R_b}\left(\frac{A_v \frac{R_b}{R_a+R_b}}{1 + A_v \frac{R_b}{R_a+R_b}}\right). \qquad (10.8)$$

In this notation, we clearly see that the loop gain L can be written as the product of the voltage gain of the controller and the voltage gain of the feedback network:

$$L = A_v \frac{R_b}{R_a + R_b}. \qquad (10.9)$$

The ideal gain is the gain of the amplifier in which the controller is replaced with a nullor. It equals the reciprocal value of the gain of the feedback network.

The above example shows the application of Black's feedback model for a simple situation. In this particular case, it appears as if the model supports

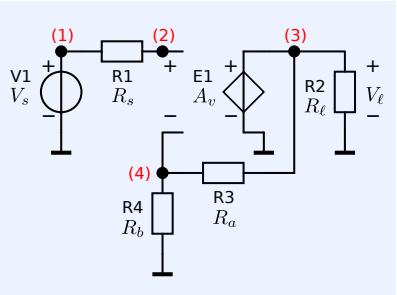


Figure 10.2: Passive feedback voltage amplifier with a voltage-controlled voltage source as controller.

a two-step design approach: the source-to-load transfer can be written as the product of the ideal gain $\frac{R_a+R_b}{R_b}$, and the *servo function* S, which is uniquely defined by the loop gain:

$$S = \frac{L}{1+L}. \tag{10.10}$$

This servo function is a measure for the error between the ideal gain, and the actual gain in which the controller a high-gain amplifier, rather than a nullor.[6]

[6] Middlebrook uses the term "discrepancy factor".

It appears as if the model of Black supports a two-step design approach:

The first step is the design of the ideal transfer using a nullor as controller, while the second step comprises the design of a controller that provides a sufficiently large loop gain.

However, the above is only true because the controller in this voltage amplifier is an ideal voltage-controlled voltage source: it has an infinite input impedance and zero output impedance. In this case, this has the following consequences:

1. The current through the source resistor R_1 equals zero, thus the input voltage of the amplifier equals the source voltage.
2. Current through the feedback network and current through load impedance do not affect the output voltage of the controller.
3. There is no direct transfer from the source to the load through the feedback network.

In fact, for negative-feedback voltage amplifiers, one would like to use an ideal voltage-controlled voltage source as controller. This would make the loop gain independent from the source and load impedance. However, from a design point of view it is not at all necessary. The design goal is to design servo function transfer close to unity.

This makes Black's feedback model not the best model for the design of negative feedback amplifiers. It does not support the two-step design approach.

We will elucidate this in the next example in which we select a current-controlled current source instead of a voltage-controlled voltage source as controller in the voltage amplifier.

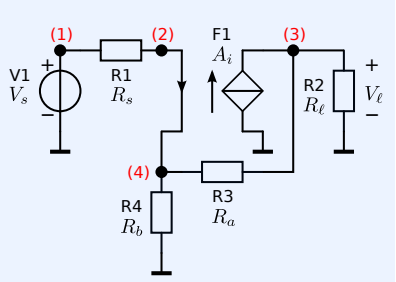


Figure 10.4: Passive feedback voltage amplifier with a current-controlled current source as controller.

Example 10.2

Figure 10.4 shows the passive feedback voltage amplifier from Figure 10.2 but with the controller replaced with a current-controlled current source. The SLiCAP netlist file for this model is:

```
"Voltage amplifier with CCCS controller"
* file: vAmpBlackF.cir
* SLiCAP circuit file
V1 1 0 {V_s}
R1 1 2 {R_s}
R2 3 0 {R_ell}
F1 0 3 2 4 {A_i}  ;The SLiCAP syntax for a CCCS differs from SPICE syntax
R3 3 4 {R_a}
R4 4 0 {R_b}
.end
```

The following script can be used to obtain a symbolic expression for the transfer from `V1` *to the voltage across* `R2`*:*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: vAmpBlackF.py

from SLiCAP import *

fileName = 'vAmpBlackF'
```

```
prj = initProject(fileName)     # Creates the SLiCAP libraries and the
                                # project HTML index page
i1 = instruction()              # Creates an instance of an instruction object
i1.setCircuit(fileName+'.cir') # Checks and defines the local circuit object,
                                # and sets the index page to the project index
i1.setSource('V1')
i1.setDetector('V_3')
i1.setSimType('symbolic')
i1.setGainType('gain')
i1.setDataType('laplace')
result = i1.execute()

htmlPage('Voltage amplifier with CCCS controller')
text2html('The gain of the system is obtained as:')
V_ell, V_s = sp.symbols('V_ell, V_s')
eqn2html(V_ell/V_s, result.laplace)
```

The result of the newly generated html page is shown in Figure 10.5.

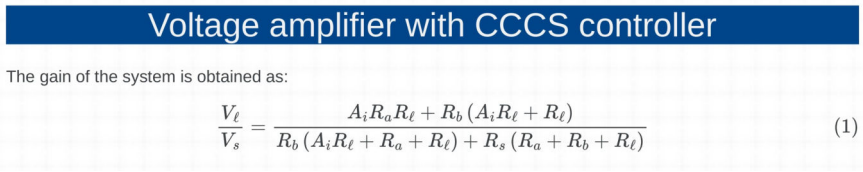


Figure 10.5: SLiCAP simulation results for the negative feedback voltage amplifier with CCCS controller.

It is clear that the transfer now depends on the source impedance, the load impedance and the loop gain, while the loop gain cannot simply be obtained as the product of the gain of the controller and the transfer of the feedback network.

The above example shows that changing the controller type does not only change the expression of the source-to-load transfer, but also complicates the interpretation of the results that are obtained with Black's feedback model.

In fact, there are three causes for the modification of the source-to-load transfer when we change the controller in the negative feedback voltage amplifier from a VCVS to a CCCS:

1. With the CCCS, the source current is no longer zero, hence the input voltage of the circuit does not equal the source voltage.
2. With the CCCS, the load voltage is not uniquely defined by the output quantity of the controller.[7]
3. With the CCCS, there exist two current paths from the source to the load with a nonzero transfer: one through the controller and one direct transfer from the source to the load through the feedback network. The latter one is not accounted for in Black's feedback model, which assumes unidirectional transfer only.

[7] The output quantity of the controller is current, while the load quantity is voltage.

In general, any practical controller circuit:

1. will have nonzero finite port impedances
2. will have bidirectional transfer
3. does not show natural two-port behavior (see Chapter 18.6.1 for two-port conditions and natural two-port behavior).

10.2.3 Conclusions

At this stage, we may conclude that Black's feedback model is very well suited to system level analysis with building blocks that show no interaction and have unilateral behavior, but it is not very suited to the design and the analysis of negative feedback amplifiers.

For the analysis of feedback circuits to match those obtained with network analysis, we need a model that does not pose additional requirements to

circuit parts that can be identified as controller or as feedback network. In the next section, we will show that the asymptotic gain model satisfies such requirements while also supporting the two-step design method in the best possible way.

10.3 Asymptotic gain model

In this section, we will discuss the asymptotic gain model. With this model, the source-to-load transfer of a feedback amplifier can be written in a way that supports the two-step design method, while it provides exactly the same results as network theory. The asymptotic gain model will be derived from the superposition model.

10.3.1 Superposition model

The superposition model gives a formal approach to the analysis of negative feedback amplifiers. The model is based upon the application of the superposition theorem that holds for linear networks. In order to apply the superposition theorem, the feedback amplifier is modeled as a network that has at least one controlled source E_c that is controlled by a controlling quantity E_i. The reference variable A, as shown in Figure 10.6, describes the relation between an arbitrarily selected controlled quantity E_c and its controlling quantity E_i : $E_c = AE_i$.[8]

[8] Please note that we will use the symbol A for one of the two-port's transmission parameters and for the reference variable in a negative feedback amplifier. The meaning will become clear from the context in which A will be used.

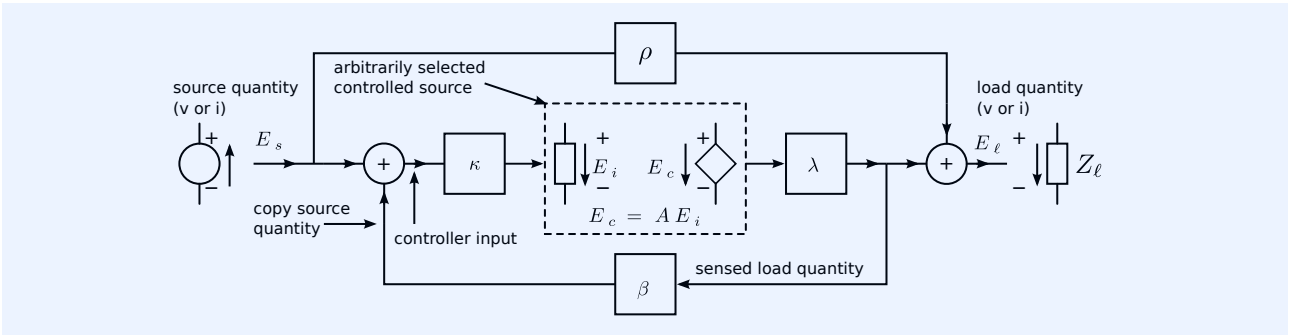


Figure 10.6: The superposition model is based upon the representation of a linear network with one arbitrarily selected controlled source.

Figure 10.6 shows a simplified schematic representation of an electrical network according to the superposition model. The source quantity E_s, the load quantity E_ℓ, the controlled quantity E_c and the controlling quantity E_i can be a current or a voltage.

In the superposition model, we will mathematically break the loop that exists from the input of the controlled source to its output by replacing the controlled source E_c with an independent source E_c, as shown in Figure 10.7.[9]

[9] This is a formal way to assign a value to E_c without knowing E_i.

With the aid of the superposition theorem, each current and voltage of the feedback amplifier can now be written as a linear combination of the two independent sources in the circuit: the current or voltage of the signal source (the source quantity E_s) and the current or voltage of the independent source E_c. In this way, we write both the load quantity E_ℓ and the controlling quantity E_i as a linear combination of the two sources E_s and E_c:

$$\begin{pmatrix} E_\ell \\ E_i \end{pmatrix} = \begin{pmatrix} \rho & \lambda \\ \kappa & \lambda\beta\kappa \end{pmatrix} \begin{pmatrix} E_s \\ E_c \end{pmatrix}. \tag{10.11}$$

This is the superposition model and its model parameters are defined as:

- $\rho = \left.\frac{E_\ell}{E_s}\right|_{E_c=0}$: Direct transfer from the source to the load;

- $\lambda = \left.\frac{E_\ell}{E_c}\right|_{E_s=0}$: Transfer from E_c to the load;
- $\kappa = \left.\frac{E_i}{E_s}\right|_{E_c=0}$: Transfer from the source to the driving quantity E_i;
- $\lambda\beta\kappa = \left.\frac{E_i}{E_c}\right|_{E_s=0}$: Transfer from E_c to E_i; a nonzero value of $\lambda\beta\kappa$ indicates the existence of feedback.

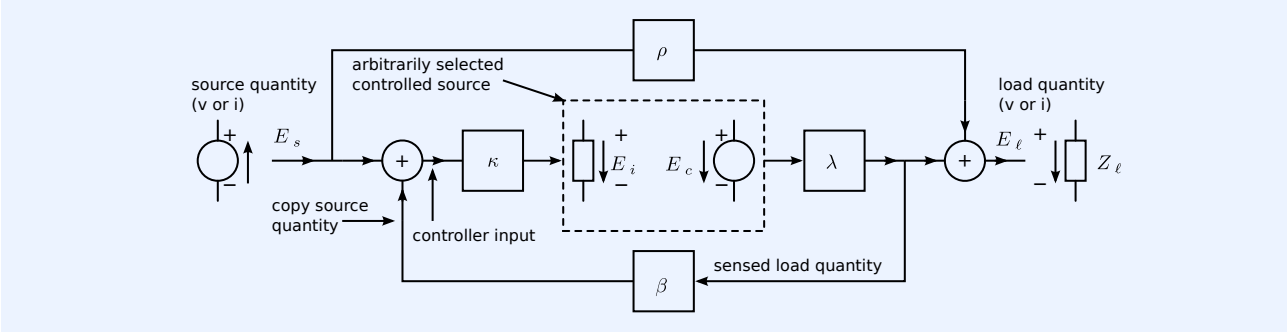


Figure 10.7: The superposition model. The gain of an arbitrarily selected controlled source is defined as 'reference variable' A, and the controlled source (diamond) is replaced by the independent source (circle) E_c.

The feedback factor β is thus defined as

$$\beta = \frac{1}{\lambda\kappa}\left(\left.\frac{E_i}{E_c}\right|_{E_s=0}\right). \tag{10.12}$$

Please note that all parameters depend on the selection of the loop gain reference variable A:

$$A = \frac{E_c}{E_i} \tag{10.13}$$

At a later stage we show that the loop gain reference A can be chosen in such a way that $-\beta$ equals the transfer of the feedback network. In such cases this feedback model provides useful design information.

Source-to-load transfer

We will now derive an expression for the transfer from source-to-load A_f of a feedback amplifier. The model equations are:

$$E_\ell = \rho E_s + \lambda E_c, \tag{10.14}$$
$$E_i = \kappa E_s + \lambda\beta\kappa E_c, \tag{10.15}$$
$$E_c = A E_i. \tag{10.16}$$

Substitution of (10.16) in (10.14) and in (10.15) yields:

$$E_\ell = \rho E_s + \lambda A E_i, \tag{10.17}$$
$$E_i = \kappa E_s + \lambda\beta\kappa A E_i. \tag{10.18}$$

We now solve E_i from (10.18) and substitute the result in (10.17). We obtain:

$$E_i = \frac{\kappa}{1-\lambda\beta\kappa A} E_s. \tag{10.19}$$

Substitution of this result in (10.17) yields:

$$E_\ell = E_s\left(\rho + \frac{\lambda A\kappa}{1-\lambda\beta\kappa A}\right). \tag{10.20}$$

We can now express the source-to-load transfer in terms of the parameters of the superposition model:

$$A_f = \frac{E_\ell}{E_s} = \rho - \frac{1}{\beta}\frac{-L}{1-L}, \tag{10.21}$$

where $L = \lambda\beta\kappa A$ is called the *loop gain* with respect to the reference variable A. Negative feedback gives a negative value for L. The factor $1 - L$ is called the *return difference* with respect to the reference variable A. Expression 10.21 gives an accurate description of the transfer from source-to-load for any selection of the reference variable. A different selection of the reference variable A, however, yields different values for the other model parameters ρ and L.

The superposition model does not support the two-step design approach as described in the introduction: if we have no circuit, we cannot select a reference variable and we do not know about ρ, κ and λ. The asymptotic gain model that will be described in the next section is suited for this purpose.

10.3.2 Asymptotic gain model

When the loop gain L approaches infinity, the source-to-load transfer obtains its so-called asymptotic gain value $A_{f\infty}$:

$$A_{f\infty} = \lim_{L\to-\infty} \frac{E_\ell}{E_s} = \rho - \frac{1}{\beta}. \tag{10.22}$$

With the aid of (10.22), we are able to express the source-to-load transfer $\frac{E_\ell}{E_s}$ from (10.21) in the asymptotic gain $A_{f\infty}$, the loop gain L and the direct transfer ρ:

$$A_f = \frac{E_\ell}{E_s} = \rho\frac{-L}{1-L} + \rho\frac{1}{1-L} - \frac{1}{\beta}\frac{-L}{1-L}, \tag{10.23}$$

$$A_f = A_{f\infty}\frac{-L}{1-L} + \frac{\rho}{1-L} \tag{10.24}$$

This expression resembles the expression obtained by Black's model. Before we draw any further conclusions, we will first study the influence of the direct transfer ρ.

Influence of the direct transfer

In most practical situations, the direct transfer can be neglected. This can be seen as follows. Assume the intended gain of the amplifier equals the ideal gain A_i. Also assume the loop gain reference variable A will be selected in such a way that the asymptotic gain equals the ideal gain.[10] We may then write (10.24) in the form

$$A_f = A_i\left(\frac{-L}{1-L}\right)\left(1 - \frac{1}{L}\frac{\rho}{A_i}\right). \tag{10.25}$$

The direct transfer ρ is usually much smaller than the ideal transfer A_i, so in most practical situations, we have[11]

$$\left|\frac{\rho}{A_i}\right| \ll 1. \tag{10.26}$$

In addition, for an accurate match between the actual source-to-load transfer and the ideal transfer of the amplifier, we need $|L| \gg 1$, which justifies

[10] This will be discussed later.

[11] If not, it is likely that no amplification was required!

the conclusion:

$$\left|\frac{1}{L}\frac{\rho}{A_i}\right| \ll 1. \tag{10.27}$$

So, with proper selection of the reference variable, we may write:

$$A_f = A_i\frac{-L}{1-L}. \tag{10.28}$$

Hence, if the asymptotic gain equals the ideal gain, the source-to-load transfer of a negative feedback amplifier can be written as the product of the ideal gain and the servo function S:

$$S = \frac{-L}{1-L}. \tag{10.29}$$

This servo function is uniquely defined by the loop gain.

Now, we have our two-step design approach: first, we design the ideal gain, and second, we design the controller such that the errors with respect to this ideal gain are sufficiently small.

Comparison with Black's model

Expression (10.28) has the same appearance as (10.5), but it is more arrurate and it does not require the assumptions of Black's feedback model.

However, only if the reference variable has been selected such that the ideal gain A_i equals the asymptotic gain $A_{f\infty}$, the loop gain L is the only measure for the correspondence between the ideal gain and the gain. If so, the asymptotic gain model supports the desired two-step design approach.

Please notice that the servo function according to the asymptotic gain model (10.29) has $-L$ in the numerator and in the denominator, while the servo function according to Black's feedback model (10.10) has $+L$ in the numerator and the denominator. This is because the asymptotic gain model does not include a subtracter. In the asymptotic gain model, a negative value of the loop gain indicates negative feedback, while in Black's feedback model, a positive value of the loop gain indicated negative feedback.

10.3.3 Selection of the loop gain reference

The two-step approach requires proper selection of the loop gain reference variable A.

First of all, this reference variable needs to be selected such that the asymptotic gain $A_{f\infty}$ equals the ideal gain A_i. Hence, the condition $L \to \infty$ for the selected reference variable should be equal by replacing the *controller* with a nullor. In other words, when the loop gain approaches infinity, the voltage across the input terminals of the controller *and* the current through the input port of the controller should both become zero:

> The asymptotic gain $A_{f\infty}$ of a direct feedback amplifier equals its ideal gain if the reference variable A is selected such that both the input voltage and the input current of the controller approach zero if $A \to \infty$.

Secondly, the reference variable should be selected such that the contribution of the direct transfer ρ to the gain A_f should be minimized. This can usually easily be achieved over the frequency range of interest. Outside this range, at relatively high frequencies, the contribution of the direct transfer may become dominant.[12]

If the above criteria have been met, expression (10.28) describes the source-to-load transfer and the design of the amplifier can be performed in two subsequent steps:

[12] In balanced amplifiers, or in amplifiers with a split signal path, selection of the reference variable in one of the signal paths results in an ideal gain that equals the asymptotic gain, but also in a large direct transfer. As a result the servo function will not be a measure for the difference between the gain and the ideal gain.

1. The design of the ideal gain, as discussed in Chapter 7
2. The design of a controller that provides a sufficiently large loop gain over the operating range of interest.

In the following examples, we will show that the source-to-load relation of a feedback amplifier is always accurately described by (10.24), while it is only accurately described by (10.28) if the loop gain reference variable A has been properly selected.

Example 10.3

Figure 10.8A shows a negative feedback voltage amplifier built with a current feedback operational amplifier as the controller. A simplified small-signal model of this operational amplifier is shown in Figure 10.8B.

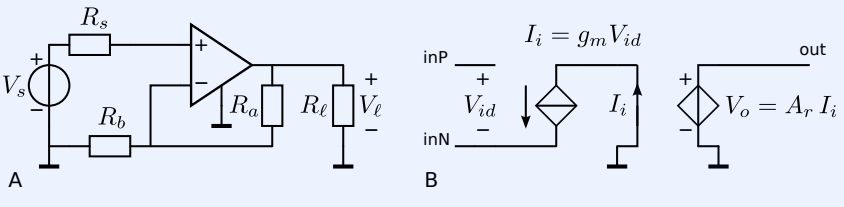


Figure 10.8:
Left: Voltage amplifier with passive direct negative feedback.
Right: Strongly simplified small-signal model of the current feedback operational amplifier that is used as active part for the voltage amplifier.

The ideal source-to-load voltage transfer A_v of this amplifier is obtained by replacing the controller with a nullor, as shown in Figure 10.9; we obtain:

$$A_v = \frac{V_\ell}{V_s} = \frac{R_a + R_b}{R_a} \tag{10.30}$$

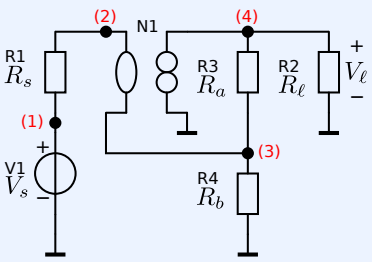


Figure 10.9: Circuit for evaluation of the ideal gain of the voltage amplifier from Figure 10.8A.

Figure 10.10 shows the small-signal equivalent model of the complete amplifier. We will evaluate the asymptotic gain, the direct transfer, the loop gain and the source-to-load transfer of this amplifier with the aid of the asymptotic gain feedback model. We will do this for two different selections of the loop gain reference variable A. First, we will select the gain g_m of the voltage-controlled current source as the loop gain reference, and then we will select the gain A_r of the current-controlled voltage source.

The SLiCAP *netlist for the circuit from Figure 10.10 is:*

```
cfbVamp
* file: cfbVamp.cir
* SLiCAP circuit file
V1 1 0 {V_s}
R1 1 2 {R_s}
R2 4 0 {R_ell}
R3 4 5 {R_a}
R4 5 0 {R_b}
G1 3 5 2 5 {g_m}
H1 4 0 0 3 H value={A_r}
.end
```

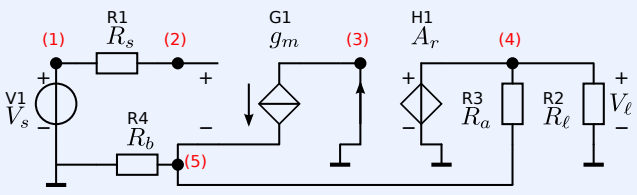


Figure 10.10: Small signal equivalent model of the voltage amplifier from Figure 10.8.

The first part of the script calculates the gain:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: cfbVamp.py

from SLiCAP import *

fileName = 'cfbVamp'
prj = initProject(fileName)    # Creates the SLiCAP libraries and the
                               # project HTML index page
i1 = instruction()             # Creates an instance of an instruction object
i1.setCircuit(fileName+'.cir') # Checks and defines the local circuit object,
                               # and sets the index page to the project index
i1.setSource('V1')
i1.setDetector('V_4')
i1.setSimType('symbolic')
i1.setGainType('gain')
i1.setDataType('laplace')
result = i1.execute()

gain = result.laplace
```

The second part of the script calculates the parameters of the asymptotic gain model with the voltage-controlled current source `G1` selected as the loop gain reference and generates an html page displaying the results:

```
V_ell, V_s, A_f                = sp.symbols('V_ell, V_s, A_f')
L_G1, rho_G1, S_G1, A_infty_G1 = sp.symbols('L_G1, rho_G1, S_G1, A_oo_G1')
L_H1, rho_H1, S_H1, A_infty_H1 = sp.symbols('L_H1, rho_H1, S_H1, A_oo_H1')

# Calculations with H1 as loop gain reference

i1.setLGref('G1')
i1.setGainType('asymptotic')
result = i1.execute()
AG1 = result.laplace

i1.setGainType('loopgain')
result = i1.execute()
LG1 = result.laplace

i1.setGainType('servo')
result = i1.execute()
SG1 = result.laplace

i1.setGainType('direct')
result = i1.execute()
DG1 = result.laplace

htmlPage('Asymptotic-gain model G1 ref')
text2html('The gain of the circuit is obtained as:')
eqn2html(V_ell/V_s, gain)

text2html('The asymptotic-gain $A_{\\infty_G1}$ is found as:')
eqn2html(A_infty_G1, AG1)

text2html('The loop gain $L_{G1}$ is found as:')
eqn2html(L_G1, LG1)

text2html('The servo function $S_{G1}$ is found as:')
eqn2html(S_G1, SG1)

text2html('The direct transfer $\\rho_{G1}$ is found as:')
eqn2html(rho_G1, DG1)

text2html('The gain $A_f$ calculated from $A_{\\infty_{G1}}$, $S_{G1}$ ' +
          'and  $\\rho_{G1}$ is obtained as:')
eqn2html(A_f, sp.simplify(AG1*SG1 + DG1/(1-LG1)))
```

Figure 10.11 shows the results of the gain calculation directly from the MNA matrix and the results according to the asymptotic gain model with the voltage-controlled current source `G1` selected as the loop gain reference.

The results clearly show that the asymptotic gain does not equal the ideal gain, but the gain calculated according to the asymptotic gain model is the same as the one calculated directly from network analysis. This can be understood if we study

the circuit from Figure 10.10 in more detail. The asymptotic gain is calculated for the case in which g_m approaches infinity. If this is the case, the controlled source `G1` can deliver any current while its controlling voltage equals zero. Hence, in this case, the input voltage of the controller equals zero. However, since the controlled source `H1` has a finite gain, it requires a nonzero input current for any nonzero output voltage. This input current flows in the inverting input of the controller, thus the controller does not act as a nullor, and as a consequence, the asymptotic gain differs from the ideal gain.

Asymptotic-gain model G1 ref

The gain of the circuit is obtained as:

$$\frac{V_\ell}{V_s} = \frac{g_m\left(A_r R_a + A_r R_b\right)}{R_a + R_b g_m\left(A_r + R_a\right) + R_b} \tag{1}$$

The asymptotic-gain $A_{\infty_{G1}}$ is found as:

$$A_{ooG1} = \frac{A_r R_a + A_r R_b}{R_b\left(A_r + R_a\right)} \tag{2}$$

The loop gain L_{G1} is found as:

$$L_{G1} = -\frac{R_b g_m\left(A_r + R_a\right)}{R_a + R_b} \tag{3}$$

The servo function S_{G1} is found as:

$$S_{G1} = \frac{R_b g_m\left(A_r + R_a\right)}{R_a + R_b g_m\left(A_r + R_a\right) + R_b} \tag{4}$$

The direct transfer ρ_{G1} is found as:

$$\rho_{G1} = 0 \tag{5}$$

The gain A_f calculated from $A_{\infty_{G1}}$, S_{G1} and ρ_{G1} is obtained as:

$$A_f = \frac{A_r g_m\left(R_a + R_b\right)}{R_a + R_b g_m\left(A_r + R_a\right) + R_b} \tag{6}$$

Figure 10.11: SLiCAP simulation results of the voltage amplifier with current-feedback operational amplifier. The voltage-controlled current source `G1` has been selected as loop gain reference variable.

The third part of the script calculates the parameters of the asymptotic gain model with the current-controlled voltage source `H1` selected as the loop gain reference and generates an html page displaying the results:

```
65 # Calculations with H1 as loop gain reference
66
67 i1.setLGref('H1')
68 i1.setGainType('asymptotic')
69 result = i1.execute()
70 AH1 = result.laplace
71
72 i1.setGainType('loopgain')
73 result = i1.execute()
74 LH1 = result.laplace
75
76 i1.setGainType('servo')
77 result = i1.execute()
78 SH1 = result.laplace
79
80 i1.setGainType('direct')
81 result = i1.execute()
82 DH1 = result.laplace
83
84 htmlPage('Asymptotic-gain model H1 ref')
85 text2html('The gain of the circuit is obtained as:')
86 eqn2html(V_ell/V_s, gain)
```

```
87
88  text2html('The asymptotic-gain $A_{\\infty_{H1}}$ is found as:')
89  eqn2html(A_infty_H1, AH1)
90
91  text2html('The loop gain $L_{H1}$ is found as:')
92  eqn2html(L_H1, LH1)
93
94  text2html('The servo function $S_{H1}$ is found as:')
95  eqn2html(S_H1, SH1)
96
97  text2html('The direct transfer $\\rho_{H1}$ is found as:')
98  eqn2html(rho_H1, DH1)
99
100 text2html('The gain $A_f$ calculated from $A_{\\infty_G1}$, $S_{G1}$ and ' +
101           '$\\rho_{G1}$ is obtained as:')
102 eqn2html(A_f, sp.simplify(AH1*SH1 + DH1/(1-LH1)))
```

Asymptotic-gain model H1 ref

The gain of the circuit is obtained as:

$$\frac{V_\ell}{V_s} = \frac{g_m\left(A_r R_a + A_r R_b\right)}{R_a + R_b g_m\left(A_r + R_a\right) + R_b} \tag{1}$$

The asymptotic-gain $A_{\infty_{H1}}$ is found as:

$$A_{ooH1} = \frac{R_a + R_b}{R_b} \tag{2}$$

The loop gain L_{H1} is found as:

$$L_{H1} = -\frac{A_r R_b g_m}{R_a R_b g_m + R_a + R_b} \tag{3}$$

The servo function S_{H1} is found as:

$$S_{H1} = \frac{A_r R_b g_m}{A_r R_b g_m + R_a R_b g_m + R_a + R_b} \tag{4}$$

The direct transfer ρ_{H1} is found as:

$$\rho_{H1} = 0 \tag{5}$$

The gain A_f calculated from A_{∞_G1}, S_{G1} and ρ_{G1} is obtained as:

$$A_f = \frac{A_r g_m\left(R_a + R_b\right)}{A_r R_b g_m + R_a R_b g_m + R_a + R_b} \tag{6}$$

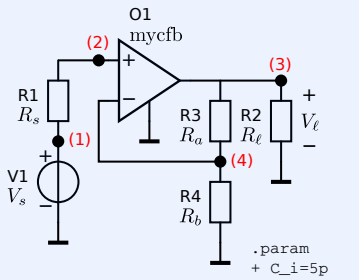
Figure 10.12: SLiCAP simulation results of the voltage amplifier with a current-feedback operational amplifier. The voltage-controlled current source `H1` has been selected as the loop gain reference variable.

Figure 10.12 shows the results of the gain calculation directly from the MNA matrix and the results according to the asymptotic gain model with the current-controlled voltage source `H1` *selected as the loop gain reference. In this case, the reference variable has been selected properly: the asymptotic gain equals the ideal gain. Since we also have no direct transfer ($\rho = 0$), the servo function completely describes the non ideal behavior of the controller. Hence, the design of this amplifier can be performed in two subsequent steps: the design of the ideal gain and the design of the controller.*

In the following example, we will use a more elaborate model for the operational amplifier and illustrate the use of the asymptotic gain model in more detail.

Example 10.4

Figure 10.13 shows the circuit of the passive feedback voltage amplifier with a current feedback operational amplifier as the controller.




```
.param
+ C_i=5p
+ R_i=1M
+ C_d=2p
+ R_d=10k
+ g_m=20m
+ R_t=1M
+ R_o=2k
+ tau=50u
+ R_a=2k
+ R_b=500
+ R_s=50
+ R_ell=500
```

```
.model mycfb OC
+ cp={C_i}
+ gp={1/R_i}
+ cpn={C_d}
+ gpn={1/R_d}
+ gm={g_m}
+ zt={R_t/(1+s*tau)}
+ zo={R_o/(1+s*tau)}
```

Figure 10.13: Voltage amplifier with a current feedback operational amplifier as controller.

The SLiCAP *netlist of this circuit has been listed below.* SLiCAP *has a built-in model for a current feedback operational amplifier. Please see the* SLiCAP *help file for detailed information on this model.*

```
cfbVampExtended
* file: cfbVampExtended.cir
* SLiCAP circuit file
V1 1 0 {V_s}
R1 1 2 {R_s}
O1 2 4 3 0 mycfb
R2 3 0 {R_ell}
R3 3 4 {R_a}
R4 4 0 {R_b}
* Model definition for the operational amplifier 'mycfb'
.model mycfb OC cp={C_i} gp={1/R_i} cpn={C_d} gpn={1/R_d} gm={g_m}
+               zt={R_t/(1+s*tau)} zo={R_o/(1+s*tau)}
* parameter values for numeric simulation
.param C_i=5p R_i=1M g_m=20m R_t=1M R_o=2k tau=50u R_a=2k R_b=500 R_s=100
+            R_ell=500 V_s=1 C_d=2p R_d=10k
.end
```

SLiCAP *expands this netlist to the circuit depicted in Figure 10.13. The element* `H_01` *is a current-controlled voltage source with series impedance. It has two associated values: the first one for the transimpedance factor and the second one for the output impedance. The data of all circuit elements after expansion of the netlist can be displayed by* SLiCAP*. The result is shown in Figure 10.14. The script below shows how this should be done:*

```
#!/usr/bin/env python2
# -*- coding: utf-8 -*-
"""
Created on Thu Jul  2 14:59:59 2020

@author: anton
"""
from SLiCAP import *

fileName = 'cfbVampExtended'
prj = initProject(fileName)    # Creates the SLiCAP libraries and the
                               # project HTML index page
i1 = instruction()             # Creates an instance of an instruction object
i1.setCircuit(fileName+'.cir') # Checks and defines the local circuit object,
                               # and sets the index page to the project index
htmlPage('Circuit data')
elementData2html(i1.circuit)
```

Figure 10.14: Netlist of the amplifier from Figure 10.13, expanded by SLiCAP.

Circuit data

Table: Element data of expanded netlist 'cfbVampExtended'

RefDes	Nodes	Refs	Model	Param	Symbolic	Numeric
Cp_01	2 0		C	value	C_i	$5.0 \cdot 10^{-12}$
Cpn_01	2 4		C	value	C_d	$2.0 \cdot 10^{-12}$
Gm_01	1_01 4 2 4		g	value	g_m	0.02
Gp_01	2 0 2 0		g	value	$\frac{1}{R_i}$	$1.0 \cdot 10^{-6}$
Gpn_01	2 4 2 2		g	value	$\frac{1}{R_d}$	0.0001
H_01	3 0 0 1_01		HZ	value	$\frac{R_t}{s\tau+1}$	$\frac{1000000}{5.0 \cdot 10^{-5} s+1}$
				zo	$\frac{R_o}{s\tau+1}$	$\frac{2000}{5.0 \cdot 10^{-5} s+1}$
R1	1 2		R	value	R_s	100
R2	3 0		R	value	R_ℓ	500
R3	3 4		R	value	R_a	2000
R4	4 0		R	value	R_b	500
V1	1 0		V	value	V_s	1

The second part of the script shows how to combine the magnitude characteristics and the phase of the transfers of the asymptotic gain model in one magnitude and one phase plot. Figure 10.15 shows the Bode plots generated by this script. The part of the script for generating the Bode plots is shown below:

```
i1.setSource('V1')
i1.setDetector('V_3')
i1.setLGref('H_O1')
i1.setSimType('numeric')
i1.setDataType('laplace')
i1.setGainType('gain')
G = i1.execute()
i1.setGainType('asymptotic')
A = i1.execute()
i1.setGainType('loopgain')
L = i1.execute()
i1.setGainType('servo')
S = i1.execute()
i1.setGainType('direct')
D = i1.execute()
figdBmag = plotSweep('cfbVampdBmag.svg', 'dB magnitude plots asymptotic-' +
                     'gain model', [G,A,L,S,D], 1e4, 1e9, 200,
                     funcType = 'dBmag', show=True)
figPhase = plotSweep('cfbVampPhase.svg', 'Phase plots asymptotic-gain model',
                     [G,A,L,S,D], 1e4, 1e9, 200, funcType = 'phase', show=True
                       )
htmlPage('Bode plots')
fig2html(figdBmag, 800)
fig2html(figPhase, 800)
```

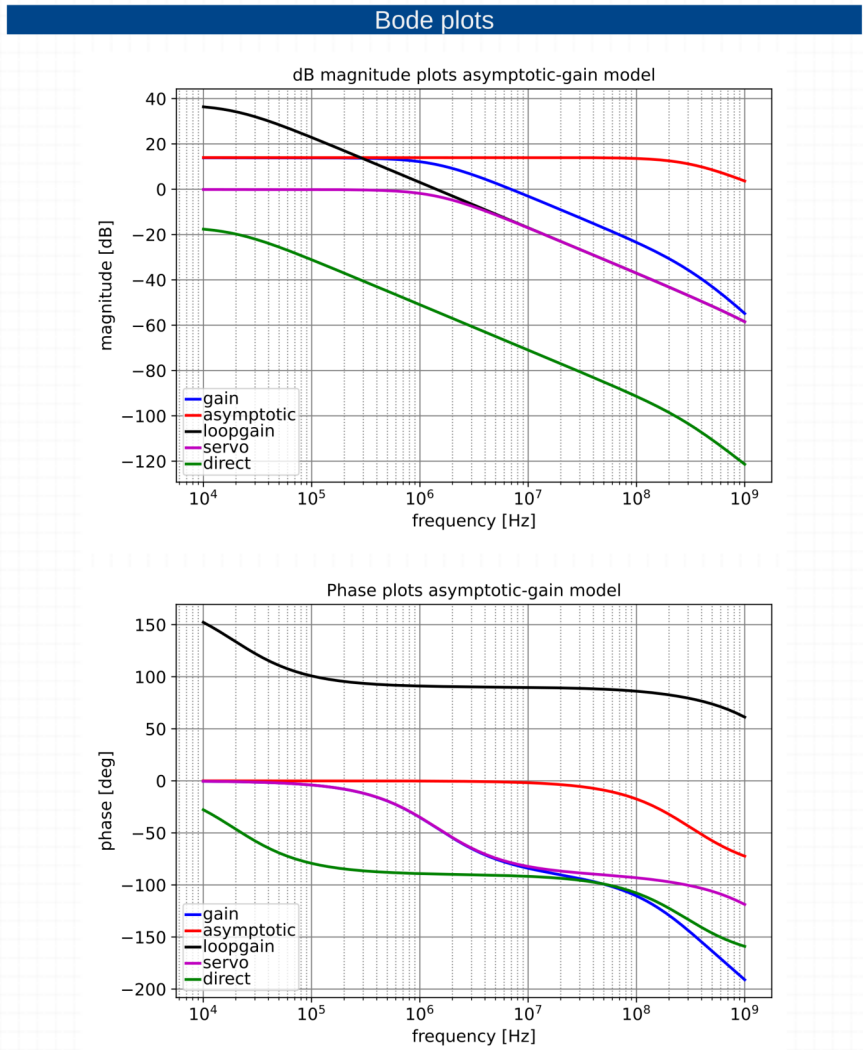


Figure 10.15: Bode plots of the transfers according to the asymptotic gain model of the voltage amplifier from Figure 10.13.

Mismatch between ideal gain and asymptotic gain

In the above example, the asymptotic gain approximates the ideal gain over a wide frequency range. At low frequencies, the nonzero resistance of the source and the nonzero conductance between the noninverting controller input and the ground cause a small error: together, they constitute a voltage divider. At high frequencies, the asymptotic gain drops due to the nonzero values of the source resistance and the input capacitance of the controller. In general, a finite impedance between an input terminal of the controller and the ground will cause inaccuracy and nonlinearity in the transfer of amplifiers that have their controller inputs floating with respect to the ground. Such impedances can almost never be avoided and should be kept as large as possible because their influence cannot be reduced by increasing the loop gain (see direct voltage comparison in Figure 7.7).

Below the part of the script for evaluation of the asymptotic gain.

```
43 htmlPage('Symbolic asymptotic-gain')
44 i1.setSimType('symbolic')
45 i1.setGainType('asymptotic')
46 result = i1.execute()
47 A = result.laplace
48 text2html('The asymptotic-gain is found as:')
49 eqn2html('A_f_oo', A)
```

The html page generated by this script has been shown in Figure 10.16.

Figure 10.16: Results of the symbolic evaluation of the asymptotic gain of the circuit from Figure 10.13

Symbolic asymptotic-gain

The asymptotic-gain is found as:

$$A_{foo} = \frac{R_i\,(R_a + R_b)}{C_i R_b R_i R_s s + R_b R_i + R_b R_s} \tag{1}$$

This result can be written as:

$$A_{f\infty} = \left(\frac{R_a + R_b}{R_b}\right)\left(\frac{R_i}{R_s + R_i}\right)\left(\frac{1}{1 + sC_i\frac{R_i R_s}{R_s + R_i}}\right), \tag{10.31}$$

which clearly shows the influence of R_i and C_i on the asymptotic gain.

Two-step design approach

The servo function approximates unity at frequencies at which the magnitude of the loop gain is much larger than unity, while it approximates the loop gain at frequencies for which the magnitude of the loop gain is much smaller than unity. In the above example, the magnitude of the direct transfer is much smaller than that of the asymptotic gain. Hence, a two-step design approach steps, seems to be possible.

10.3.4 Hand calculations of the loop gain

Although symbolic analysis programs such as SLiCAP provide an easy way to derive expressions for the different transfer functions of the asymptotic gain model, they do not always present these expressions in a design-friendly way. In other words, it is not always easy to derive design conclusions from them. In the following examples, we will perform hand calculations based upon the calculation of current division factors. This method is known as design-oriented analysis, resulting in so-called *low-entropy expressions*, as described by R.D. Middlebrook.[Middlebrook1991][13]

[13] R.D. Middlebrook. Low-Entropy Expressions: The Key to Design-Oriented Analysis. *Frontiers in Education Conference*, pages 399–403, 1991

Ideal gain and asymptotic gain

In the following example, we will evaluate the ideal gain and the asymptotic gain of a passive feedback voltage amplifier.

Example 10.5

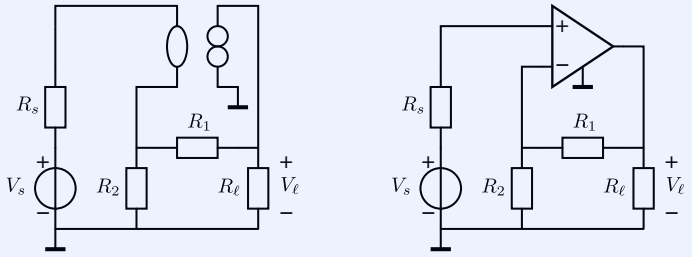


Figure 10.17:
Left: Voltage amplifier with a nullor as the controller for evaluation of the ideal gain.
Right: Voltage amplifier with an operational amplifier for evaluation of the asymptotic-gain, the loop gain and the direct transfer.

Figure 10.17 shows two amplifier circuits: one with a nullor as the controller and another with an operational amplifier.

We will evaluate the ideal transfer from the configuration with the nullor. Since both the input voltage and the input current of the nullor equal zero, we find:

$$\frac{V_\ell}{V_s} = \frac{R_1 + R_2}{R_2}. \tag{10.32}$$

The asymptotic gain will be evaluated from the configuration with the operational amplifier. The small-signal model of the operational amplifier that is used for the controller is shown in Figure 10.18. It is a single-pole operational amplifier with a gain-bandwidth product of GB [Hz], a DC voltage gain A_0, an input resistance R_i and an output resistance R_o.

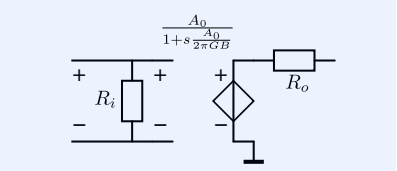


Figure 10.18: Simple small-signal model of the single-pole operational amplifier.

The small-signal model of the voltage amplifier with the operational amplifier is shown in Figure 10.19. We will determine the asymptotic gain, the loop gain and the direct transfer. To this end, we need to select the loop gain reference variable A. A proper selection is one in which the controller behaves as a nullor if $A \to \infty$.

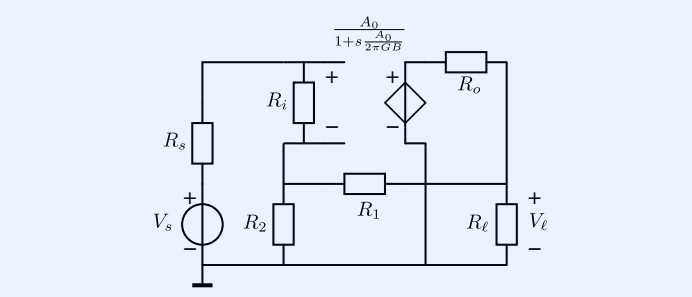


Figure 10.19: Small-signal equivalent circuit of the passive feedback voltage amplifier.

This is the case if we select the voltage transfer of the operational amplifier[14] *as the loop gain reference variable:*

[14] The voltage-controlled voltage source.

$$A = \frac{A_0}{1 + s\frac{A_0}{2\pi GB}} \tag{10.33}$$

If the voltage gain of this voltage-controlled voltage source approaches infinity, the voltage across the input port of the controller approaches zero for any finite output voltage. Under this condition the current through the input port also approaches zero.[15] *Hence, the asymptotic gain equals the ideal gain of the amplifier:*

[15] The current through the input port of the active part flows through R_i only.

$$A_{f\infty} = \left.\frac{V_\ell}{V_s}\right|_{A\to\infty} = \frac{R_1 + R_2}{R_2}. \tag{10.34}$$

Loop gain

In the following example, we will determine the loop gain of the amplifier.

Example 10.6

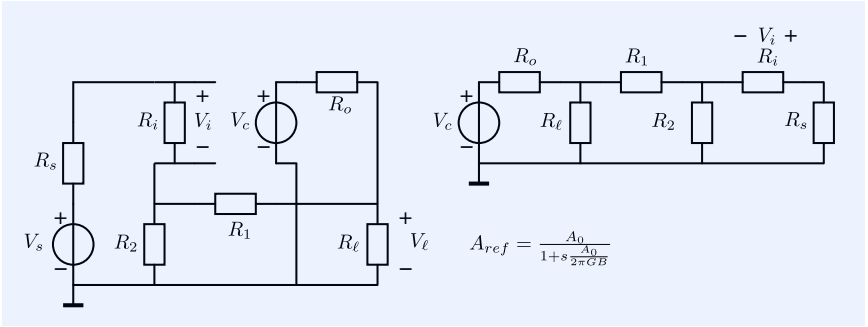


Figure 10.20: Small-signal equivalent circuits for calculation of the loop gain. The voltage gain of the operational amplifier has been selected as the loop gain reference.

Left: Amplifier circuit in which the controlled source has been replaced with an independent source.

Right: The circuit redrawn to easy hand calculations of the loop gain.

For the evaluation of the loop gain, we need to replace the controlled source of the reference variable by an independent source V_c. The controlling voltage is the dependent variable V_i. According to its definition, the loop gain $L = A\beta$ should be obtained as:

$$L = A\left.\frac{V_i}{V_c}\right|_{V_s=0}. \tag{10.35}$$

Figure 10.20 shows the equivalent circuits for the above steps. The circuit on the left shows the small-signal model with the controlled source replaced with an independent source and the controlling voltage V_i. To simplify the hand calculations, the circuit has been redrawn on the right side.

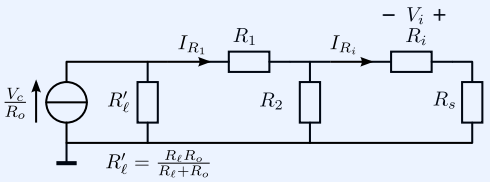


Figure 10.21: Simplified small-signal equivalent circuit for calculation of the loop gain.

A method, that is very well suited for the analysis of so-called ladder networks *is based upon current division analysis: the voltage across R_i is the product of the current through R_i and the resistance R_i. The current I_{R_i} that flows through R_i can easily be obtained from network inspection. Figure 10.21 shows the modified network in which the series connection of V_c and R_o has been replaced with its Norton equivalent. With the aid of this schematic, we write:*

$$V_i = -I_{R_1}\frac{R_2}{R_2 + R_i + R_s}R_i, \tag{10.36}$$

$$I_{R_1} = \frac{V_c}{R_o}\frac{R'_\ell}{R'_\ell + R_1 + (R_2 /\!\!/ (R_1 + R_s))}, \tag{10.37}$$

$$R'_\ell = (R_o /\!\!/ R_\ell), \tag{10.38}$$

where the short notation $(R_a /\!\!/ R_b)$ represents the equivalent resistance of the parallel connection of R_a and R_b:

$$(R_a /\!\!/ R_b) = \frac{R_a R_b}{R_a + R_b}, \tag{10.39}$$

from which we obtain:

$$V_i = \frac{V_c}{R_o}\frac{(R_o /\!\!/ R_\ell)}{(R_o /\!\!/ R_\ell) + R_1 + (R_2 /\!\!/ (R_i + R_s))}\frac{-R_2 R_i}{R_2 + R_1 + R_s}. \tag{10.40}$$

For design purposes, it is often convenient to leave the short notations for parallel connections. From this, we find

$$\lambda\beta\kappa = \left.\frac{V_i}{V_c}\right|_{V_s=0} = \frac{R_\ell}{R_o /\!\!/ R_\ell}\frac{R_2}{(R_o /\!\!/ R_\ell) + R_1 + (R_2 /\!\!/ (R_i + R_s))}\frac{-R_i}{R_2 + R_i + R_s}. \tag{10.41}$$

The loop gain $L = \lambda\beta\kappa A$ is now found as:

$$L = -A_0\frac{1}{1 + s\frac{A_0}{2\pi GB}}\frac{R_\ell}{R_o + R_\ell}\frac{R_2}{(R_o /\!\!/ R_\ell) + R_1 + (R_2 /\!\!/ (R_i + R_s))}\frac{R_i}{R_2 + R_i + R_s}. \tag{10.42}$$

The five terms of the loop gain expression (10.42) give the following design information:

1. *The first term $-A_0$ shows the contribution of the DC voltage gain of the operational amplifier to the loop gain.*
2. *The second term $\frac{1}{1+s\frac{A_0}{2\pi GB}}$ represents a unity gain low-pass transfer due to the pole of the operational amplifier.*
3. *The third term $\frac{R_\ell}{R_o+R_\ell}$ represents the DC attenuation due to the finite value of the load resistance R_ℓ and the nonzero value of the output resistance R_o of the operational amplifier.*
4. *The fourth term $\frac{R_2}{(R_o /\!\!/ R_\ell)+R_1+(R_2 /\!\!/ (R_i+R_s))}$ represents the DC attenuation in the loop gain caused by the feedback network. Note that this attenuation also depends on R_o, R_i, R_s and R_ℓ. For ideal drive conditions of the feedback network: $R_o = 0$, $R_i = \infty$, $R_s = 0$ and $R_\ell = \infty$, this term simplifies to $\frac{R_2}{R_1+R_2}$, which is the reciprocal value of the ideal gain of the voltage amplifier.*
5. *The fifth term $\frac{R_i}{R_2+R_i+R_s}$ shows the DC attenuation in the loop gain caused by the nonzero input resistance R_i of the operational amplifier.*

Writing the loop gain as a product of static and dynamic transfers as above is very helpful for finding ways to implement frequency compensation. We will show this at a later stage.

Direct transfer

In the following example, we will derive the direct transfer ρ of this passive feedback voltage amplifier. It has been defined as the source-to-load transfer with the reference variable set to zero.

Example 10.7

Figure 10.22 shows the equivalent circuits for evaluation of the direct transfer. The circuit on the left side is the small-signal equivalent circuit with the reference variable set to zero. The circuit on the right has the components rearranged to facilitate hand calculations. It clearly shows zero direct transfer if $R_i = \infty$ or if $R_o = 0$. With the aid of the current division method, we obtain:

$$\rho = \frac{R_2}{R_s + R_i + R_2}\frac{(R_o /\!\!/ R_\ell)}{R_2 /\!\!/ (R_s + R_i) + R_1 + (R_o /\!\!/ R_\ell)}. \tag{10.43}$$

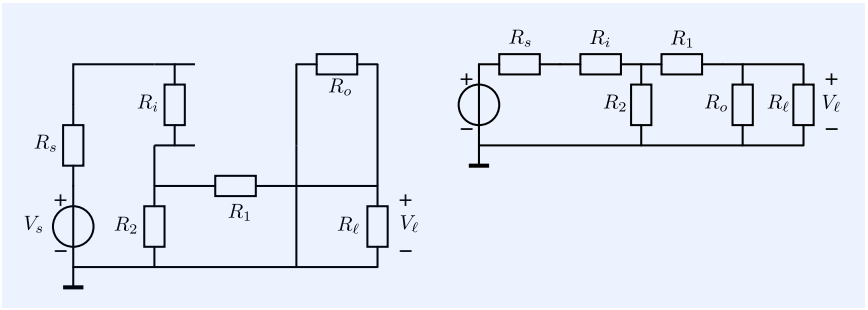


Figure 10.22: Small-signal equivalent circuit for calculation of the direct transfer, with the voltage gain of the operational amplifier selected as loop gain reference variable.

10.3.5 Impedance model

With the asymptotic gain model, we have a negative feedback model that relates one design parameter (the loop gain L) to the error of the source-to-load transfer of the feedback amplifier with respect to its ideal gain. This error includes:

1. Imperfect sensing of the load source quantity due to a non-ideal input impedance of the amplifier

2. Imperfect driving of the load due to a non-ideal output impedance of the amplifier

3. Imperfect input-to-output transfer due to a lack of controller gain.

In many cases, we are only interested in the total error of the source-to-load transfer and it is not necessary to resolve this error in the above contributions. However, if accurate termination of the source and/or the load is required, we need to know the values obtained for the input and/or output impedance of the amplifier. We will show that the asymptotic gain model can also be used for this purpose.

[16] Immittance: impedance or admittance.

Figure 10.23 shows the setup for the determination of the port *immittance*[16] of a feedback amplifier. We will show that the immittance of the amplifier's input or output port can be expressed in its asymptotic value and two functions that express the influence of the finite loop gain for two different termination conditions:

1. The port is shorted, $L = L_{sc}$

   This is the case if we determine the input admittance Y_{xf} of the port by driving it from a voltage source V_x while measuring the port current I_x, as shown in Figure 10.23A for the input port of the amplifier and Figure 10.23C for its output port

   $$Y_{xf} = \frac{I_x}{V_x}. \tag{10.44}$$

   Any shunt feedback to or from this port will become maximally ineffective while series feedback is maximally effective and the loop gain obtained is L_{sc}.[17]

[17] Under the ideal sense and comparison conditions shown in Figure 10.23 A and B, the shunt feedback is non-existent.

2. The amplifier port is left open, $L = L_o$

   This is the case if we determine the input impedance Z_{xf} of the port by driving it from a current source I_x while measuring the port voltage V_x, as shown in Figure 10.23C for the input port of the amplifier and Figure 10.23D for its output port

   $$Z_{xf} = \frac{V_x}{I_x}. \tag{10.45}$$

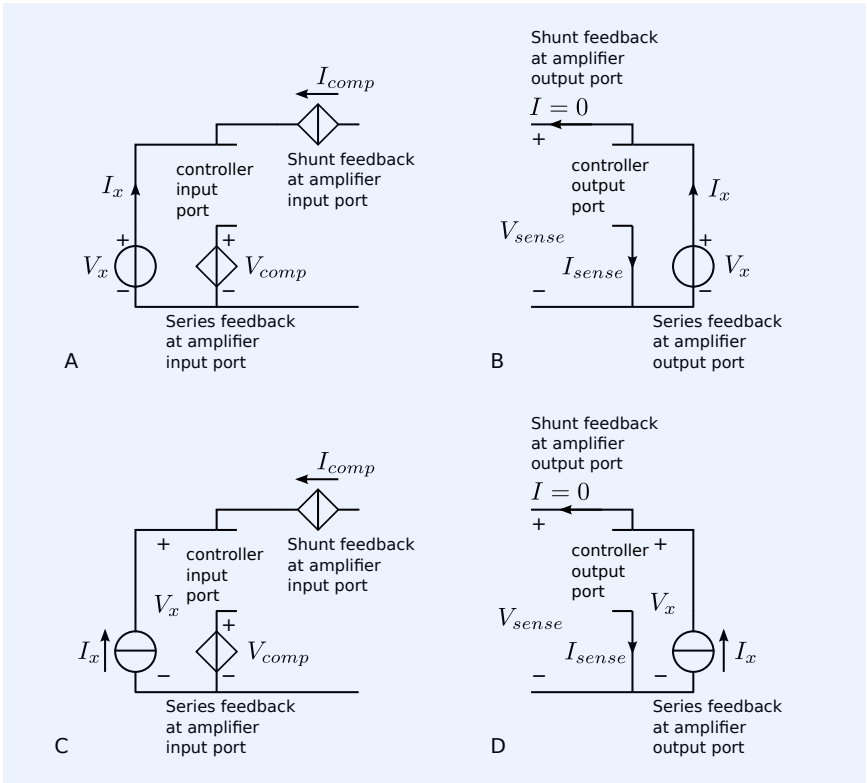


Figure 10.23: Measurement of the port immitances of a negative feedback amplifier with ideal output sensing and input comparison conditions:

A: Measurement of the input admittance $Y_x = \frac{I_x}{V_x}$. The input shunt feedback is ineffective. The feedback current I_{comp} cannot change the input voltage or current of the controller, because it flows only in V_x.

B: Measurement of the output admittance $Y_x = \frac{I_x}{V_x}$. The output shunt feedback is ineffective. The feedback voltage V_{sense} cannot be changed by the controller, because it equals V_x.

C: Measurement of the input impedance $Z_x = \frac{V_x}{I_x}$. The input series feedback is ineffective. The feedback voltage V_{comp} cannot change the input voltage or current of the controller, because the driving impedance at the opposite terminal of the controller input is infinity.

D: Measurement of the output impedance $Z_x = \frac{V_x}{I_x}$. The output series feedback is ineffective. The feedback current I_{sense} cannot be changed by the controller, because it equals I_x.

Any series feedback to or from this port will become maximally ineffective while shunt feedback is maximally effective and we obtain a loop gain L_o.[18]

[18] Under the ideal sense and comparison conditions shown in Figure 10.23 C and D, the series feedback is non existent.

We will now demonstrate the use of the superposition model and asymptotic gain model for evaluation of the port impedance of port x (input or output). For measurement of the port impedance, we drive the port with a current I_x and measure the response voltage V_x across the port.

For the current drive condition from Figure 10.23B, we apply the superposition model with the following substitutions:

$$E_s = I_x, \qquad \text{current flow in the port} \tag{10.46}$$

$$E_\ell = V_x, \qquad \text{voltage across the port} \tag{10.47}$$

$$\frac{E_\ell}{E_s} = Z_{xf}, \qquad \text{port impedance} \tag{10.48}$$

The model equations of the superposition model now become

$$\begin{pmatrix} V_x \\ E_i \end{pmatrix} = \begin{pmatrix} \rho & \lambda \\ \kappa & \lambda\beta_o\kappa \end{pmatrix} \begin{pmatrix} I_x \\ E_c \end{pmatrix}. \tag{10.49}$$

The direct transfer ρ is now defined as the port impedance when E_c has been set to zero:

$$\rho = \left.\frac{V_x}{I_x}\right|_{E_c=0}.$$

The parameter $\lambda\beta_o\kappa$ is defined as the transfer from the controlled quantity E_c to the controlling quantity E_i at $I_x = 0$; in words, when the port has been

left open

$$\lambda\beta_o\kappa = \left.\frac{E_i}{E_c}\right|_{I_x=0}. \tag{10.50}$$

We now have the following set of equations:

$$V_x = \rho I_x + \lambda E_c, \tag{10.51}$$
$$E_i = \kappa I_x + \lambda\beta_o\kappa E_c, \tag{10.52}$$
$$E_c = AE_i. \tag{10.53}$$

Substitution of (10.53) in (10.51) and in (10.52) yields

$$V_x = \rho I_x + \lambda AE_i, \tag{10.54}$$
$$E_i = \kappa I_x + \lambda\beta_o\kappa AE_i. \tag{10.55}$$

We now solve E_i from (10.55) and obtain

$$E_i = \frac{\kappa I_x}{1 - \lambda\beta_o\kappa A}. \tag{10.56}$$

We substitute this result in (10.54), which yields:

$$V_x = I_x\left(\rho + \frac{\lambda\kappa A}{1 - \lambda\beta_o\kappa A}\right). \tag{10.57}$$

We now obtain an expression for Z_{xf} in terms of the parameters of the superposition model:

$$Z_{xf} = \rho\left(1 + \frac{\lambda\kappa A}{\rho}\frac{1}{1 - \lambda\beta_o\kappa A}\right). \tag{10.58}$$

We will now define a term $\lambda\beta_{sc}\kappa$ as

$$\lambda\beta_{sc}\kappa = \left.\frac{E_i}{E_c}\right|_{V_x=0} \tag{10.59}$$

We can obtain this factor $\lambda\beta_{sc}\kappa$ from (10.54) and (10.58) for $V_x = 0$:

$$0 = \rho I_x + \lambda E_c, \tag{10.60}$$
$$E_i = \kappa I_x + \lambda\beta_o\kappa E_c. \tag{10.61}$$

We then eliminate I_x:

$$I_x = -\frac{\lambda E_c}{\rho}, \tag{10.62}$$
$$E_i = E_c\left(\lambda\beta_o\kappa - \frac{\lambda\kappa}{\rho}\right). \tag{10.63}$$

From (10.63) we obtain:

$$\left.\frac{E_i}{E_c}\right|_{V_x=0} = \lambda\beta_{sc}\kappa = \lambda\beta_o\kappa - \frac{\lambda\kappa}{\rho}. \tag{10.64}$$

From which we obtain

$$\beta_{sc} = \beta_o - \frac{1}{\rho}. \tag{10.65}$$

The direct transfer ρ is the port impedance with the loop gain reference set to zero. It is found as:

$$\rho = \frac{1}{\beta_o - \beta_{sc}}. \tag{10.66}$$

We may then rewrite (10.58) as:

$$Z_{xf} = \rho \left(\frac{1 - \lambda \beta_{sc} \kappa A}{1 - \lambda \beta_o \kappa A} \right). \tag{10.67}$$

If we substitute $\lambda A \beta_o \kappa = L_o$ and $\lambda A \beta_{sc} \kappa = L_{sc}$ we obtain:

$$Z_{xf} = \rho \frac{1 - L_{sc}}{1 - L_o}. \tag{10.68}$$

In the following sections, we will discuss this result for both single-loop feedback amplifiers and multiple-loop feedback amplifiers.

10.3.6 Port impedance of single-loop feedback amplifiers

In single-loop feedback amplifiers, we have either parallel feedback or series feedback at a port. In a case of solely parallel feedback, the ideal value of the port impedance is zero. In case of solely series feedback at a port, the ideal value of the port impedance is infinite. We have already concluded this in Chapter 7. These conclusions can also be derived from (10.68).

Parallel feedback at a port

In a case of parallel feedback, the loop gain will be zero if we short the port. Hence, we have $L_{sc} = 0$, and the port impedance can be written as

$$Z_{xf} = \rho \frac{1}{1 - L_o}. \tag{10.69}$$

The asymptotic value of the port impedance then equals zero:

$$Z_{f\infty} = \lim_{A \to \infty} Z_{xf}, \tag{10.70}$$

from which we obtain

$$Z_{f\infty} = \lim_{-L_o \to \infty} \rho \frac{1}{1 - L_o} = 0. \tag{10.71}$$

This is, of course, as expected: both forms of parallel feedback, voltage sensing at the output port and current comparison at the input port, were used to create zero port impedance.

Series feedback at a port

In a case of series feedback, the loop gain will be zero if we leave the port open. Hence, we have $L_o = 0$, and the port impedance can be written as

$$Z_{xf} = \rho \left(1 - L_{sc} \right). \tag{10.72}$$

The asymptotic value of the port impedance then equals infinity:

$$Z_{f\infty} = \lim_{-L_{sc} \to \infty} \rho \left(1 - L_{sc} \right) = \infty. \tag{10.73}$$

This is also as expected: both forms of series feedback, current sensing at the output port and voltage comparison at the input port, were used to create an infinite port impedance.

Determination of the output impedance of an operational amplifier

In the following example, we will evaluate the output impedance of a negative feedback voltage amplifier. As stated earlier, the asymptotic gain model

accurately describes the source-to-load transfer of negative feedback amplifiers, including error contributions resulting from the non-ideal input and output impedance of the amplifier. For this reason, there is no need to evaluate the port impedances of single-loop negative feedback amplifiers. The ideal port impedances of these amplifiers equal either zero or infinite. However, many data sheets of operational amplifiers specify the so-called 'closed loop' output impedance. This 'closed loop' output impedance is a property of a negative feedback amplifier equipped with the operational amplifier, rather than a property of the operational amplifier itself. The following example shows the way in which the output impedance of the operational amplifier itself[19] relates to the output impedance of the negative feedback amplifier.

[19] The output impedance of the operational amplifier is often referred to as the *open loop output impedance*.

Example 10.8

Let us consider the unity gain voltage amplifier from Figure 10.24A. It comprises an operational amplifier the model of which is shown in Figure 10.24B. For the sake of simplicity, the impedance of the signal source has been taken as zero. In practice, it would, of course, be useless to cascade an ideal voltage source with a voltage follower!

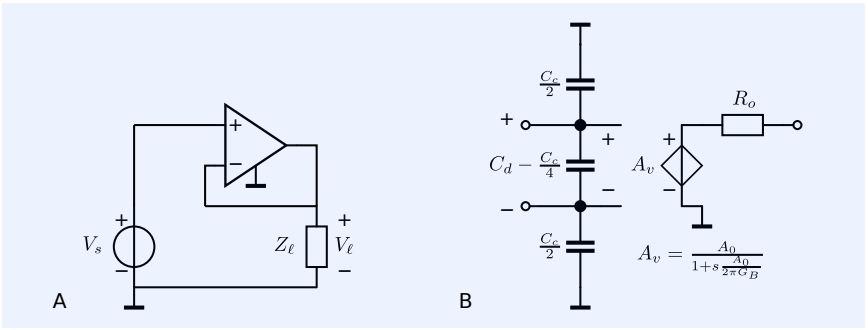


Figure 10.24:
A: Unity gain negative feedback voltage amplifier
B: Small-signal equivalent model of the voltage follower from (A).

The circuit for the determination of the output impedance is shown in Figure 10.25A. A current is driven into the output port and the voltage across the output port is measured. The output impedance Z_f is defined as

$$Z_f = \frac{V_o}{I_o}. \tag{10.74}$$

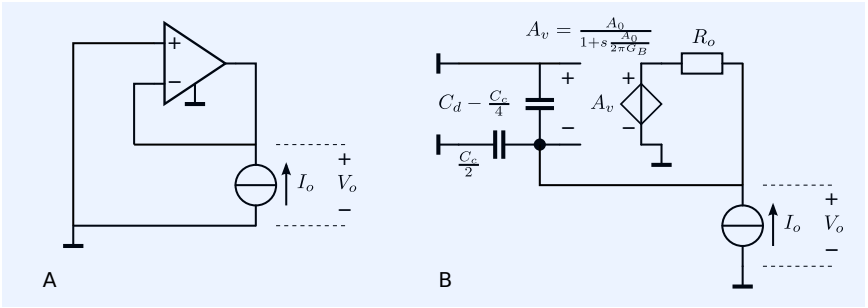


Figure 10.25:
A: Circuit for determination of the output impedance of the voltage follower.
B: Small-signal equivalent circuit of (A).

Figure 10.25B shows the model for evaluation of the output impedance with the aid of the negative feedback model. The voltage gain A_v of the operational amplifier will be selected as the loop gain reference:

$$A = \frac{A_0}{1 + s\frac{A_0}{2\pi G_B}}. \tag{10.75}$$

The controlled source will be replaced with an independent source V_c.

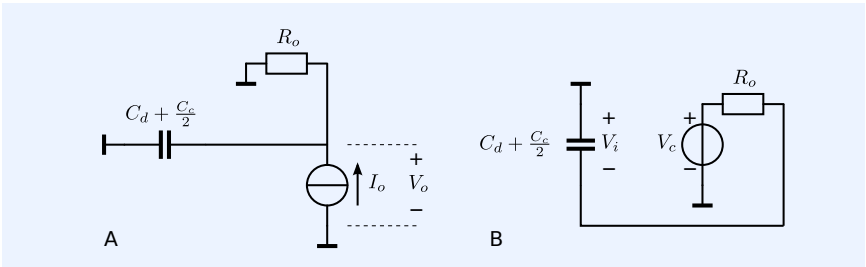


Figure 10.26:
A: Circuit for determination of ρ.
B: Circuit for determination of $\lambda\beta_o\kappa$.

Figure 10.26A shows the circuit for evaluation of ρ and Figure 10.26B shows the circuit for evaluation of $\lambda\beta_o\kappa$:

$$\rho = \left.\frac{V_o}{I_o}\right|_{V_c=0}, \quad \lambda\beta_o\kappa = \left.\frac{V_i}{V_c}\right|_{I_o=0}. \tag{10.76}$$

Please notice that the output impedance of the amplifier is defined with no load connected; the output port is left open.

The direct transfer ρ can easily be found from network inspection of the circuit from Figure 10.26A:

$$\rho = \frac{R_o}{1 + sR_o(C_d + \frac{C_c}{2})}. \tag{10.77}$$

The loop gain L_o is found from the circuit from Figure 10.26B:

$$L_o = A\lambda\beta_o\kappa = \frac{A_0}{1 + s\frac{A_0}{2\pi G_B}} \frac{-1}{1 + sR_o(C_d + \frac{C_c}{2})}. \tag{10.78}$$

The output impedance Z_f can now be obtained with the aid of (10.69) as

$$Z_f = \frac{\frac{R_o}{1+sR_o(C_d+\frac{C_c}{2})}}{1 + \frac{1}{1+sR_o(C_d+\frac{C_c}{2})} \frac{A_0}{1+s\frac{A_0}{2\pi G_B}}}. \tag{10.79}$$

With $A_0 \gg 1$, this expression can be simplified to

$$Z_f = \frac{R_o}{A_0} \frac{1 + s\frac{A_0}{2\pi G_B}}{1 + s\left(\frac{R_o(\frac{1}{2}C_c+C_d)}{A_0} + \frac{1}{2\pi G_B}\right) + s^2\frac{R_o(\frac{1}{2}C_c+C_d)}{2\pi G_B}}. \tag{10.80}$$

From this expression, we see that the low frequency value of the output impedance of the voltage follower is A_0 times smaller than the output resistance of the operational amplifier. The output impedance has a zero at the pole of the voltage gain of the amplifier. In most cases, this zero will be dominant. At higher frequencies, we find two poles. Hence, at frequencies above the frequency of the zero, the output impedance of the voltage follower will be inductive, while at the highest frequencies, it will be capacitive. This can be seen by letting s approach infinity:

$$\left(Z_f\right)_{s\to\infty} = \frac{1}{s\left(\frac{1}{2}C_c + C_d\right)}. \tag{10.81}$$

10.3.7 Port impedance of multi-loop feedback amplifiers

In this section, we will show that the port impedance of a multi-loop feedback amplifier can be expressed in its asymptotic-value and the product of two servo functions. These two servo functions are defined by the loop gain in

the situation in which the port is shorted, and in the situation in which the port is left open. If the loop gain reference has been properly selected, the asymptotic value of the port impedance is equal to its ideal value, just as we found for the source-to-load transfer.

The asymptotic value of the port impedance is defined as the value of the port impedance (10.67) for $A \to \infty$:

$$Z_{f\infty} = \lim_{A\to\infty} Z_{xf} = \rho \frac{\beta_{sc}}{\beta_o}. \tag{10.82}$$

From this expression, we obtain:

$$\rho = Z_{f\infty} \frac{\beta_o}{\beta_{sc}}. \tag{10.83}$$

After we substitute (10.83) in (10.67) we obtain an expression for the port impedance in terms of its asymptotic value $Z_{f\infty}$ and two servo functions. These servo functions describe the influence of the finite loop gains L_o and L_{sc}:

$$Z_{xf} = Z_{f\infty} \frac{L_o}{1 - L_o} \frac{1 - L_{sc}}{L_{sc}}. \tag{10.84}$$

Just as in the case of the source-to-load transfer, if the reference variable is selected properly, the asymptotic value of the port impedance is equal to the ideal value of the port impedance. The ideal value is obtained assuming nullor properties for the controller.

Expression 10.84 shows that idealized single loop negative feedback amplifiers have either zero or infinite port impedances. The realization of accurate finite nonzero port impedances requires the application of two feedback loops, one establishing series feedback and another establishing shunt feedback at that port.

10.3.8 Application of asymptotic gain model in balanced amplifiers

In section 7.7, we have discussed the design of balanced amplifiers. We obtained balanced amplifiers through anti-series connection of unbalanced amplifiers. We have seen that the differential-mode transmission-1 parameters of balanced amplifiers show a simple relationship with the transmission-1 matrix parameters of their unbalanced version. The design and analysis of the differential mode behavior of a balanced amplifier is therefore usually performed by designing their unbalanced version and then connecting two of those unbalanced amplifiers anti-series. Cross-coupling of feedback networks can be used to change the sign of the a (differential-mode) transmission parameter of a balanced amplifier with multiple feedback loops (see section 7.7.5).

We have seen that in truly balanced amplifiers, the common-mode to differential-mode conversion and the differential-mode to common-mode conversion are both zero (see section 7.7.3 for decomposition of balanced circuits into common-mode and differential-mode equivalent circuits.). In such cases we speak of orthogonality between common-mode and differential-mode quantities.

Although the information processing is governed by the differential-mode behavior, the common-mode behavior of balanced negative feedback amplifiers also needs to have our interest. Even in the absence of common-mode to differential-mode conversion, the common-mode signal excursions should be within the linear operating range of the devices and common-mode stability must be ensured. If not, the quality of the differential-mode transfer will be adversely affected by common-mode signals.

The asymptotic gain model can be for the design of both the common-mode behavior and the differential-mode behavior of balanced amplifiers. To this end, we need to decompose the balanced amplifier into a differential-mode and a common-mode equivalent circuit, as described in section 7.7.5. The differential-mode behavior can then be design using the differential-mode equivalent circuit and the common-mode behavior can be designed with the aid of the common-mode equivalent circuit, both under application of the asymptotic gain model. In this way we find differential-mode and common-mode values for the asymptotic gain, the loop gain, the servo function and the direct transfer. It should be clear that the differential-mode ideal gain and the common-mode ideal gain are obtained after replacing the controller in their respective equivalent circuits with a nullor.

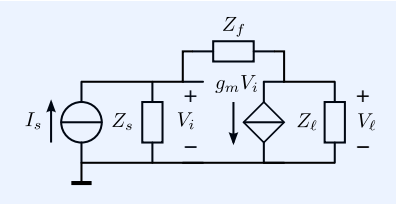


Figure 10.27: Passive-feedback transimpedance amplifier with a voltage-controlled current source as controller.

10.3.9 Asymptotic gain model and network analysis

In this section we will discuss the way in which the ideal gain, the asymptotic gain, the loop gain, the servo function, the direct transfer and the gain can be calculated from the network equations, that are obtained from modified nodal analysis. We will illustrate this using an example of a transimpedance amplifier that uses a voltage-controlled current source as a controller. The small-signal equivalent circuit of this amplifier is shown in Figure 10.27.

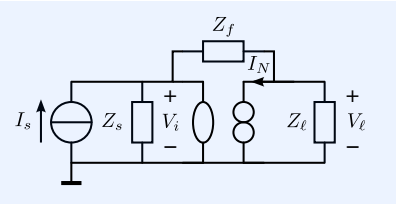


Figure 10.28: Passive-feedback transimpedance amplifier from Figure 10.27 in which the controller has been replaced with a nullor.

Ideal gain

The ideal gain of a feedback amplifier is obtained as the source-to-load transfer of the circuit in which the controller has been replaced with a nullor. This will be demonstrated in the following example.

Example 10.9

Figure 10.28 shows the transimpedance amplifier from Figure 10.27, in which the controller has been replaced with a nullor.

The MNA matrix equation of the circuit is:

$$\begin{pmatrix} I_s & 0 & 0 \end{pmatrix}^T = \mathbf{M} \begin{pmatrix} V_i & V_\ell & I_N \end{pmatrix}^T, \tag{10.85}$$

where

$$\mathbf{M} = \begin{pmatrix} \frac{1}{Z_s} + \frac{1}{Z_f} & -\frac{1}{Z_f} & 0 \\ g_m - \frac{1}{Z_\ell} - \frac{1}{Z_f} & \frac{1}{Z_\ell} + \frac{1}{Z_f} & 1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{10.86}$$

The source-to-load transfer $\frac{V_\ell}{I_s}$ can be obtained as

$$\frac{V_\ell}{I_s} = \frac{\mathcal{C}_{1,2}}{\det \mathbf{M}}, \tag{10.87}$$

where $\mathcal{C}_{2,1}$ is an element of the cofactor matrix of $\mathbf{M}$ (see Chapter 18).

In this way we obtain

$$\frac{V_\ell}{I_s} = \frac{(-1)^{1+2} \det \begin{pmatrix} g_m - \frac{1}{Z_\ell} - \frac{1}{Z_f} & 1 \\ 1 & 0 \end{pmatrix}}{\det \mathbf{M}}, \tag{10.88}$$

$$\frac{V_\ell}{I_s} = -Z_f. \tag{10.89}$$

Asymptotic-gain

The asymptotic gain of a feedback amplifier is obtained as the source-to-load transfer of the circuit in which the loop gain reference has been replaced with a nullor.

Example 10.10

For calculating the asymptotic gain of the circuit from Figure 10.27, we need to replace the selected reference variable with a nullor. In this simple circuit, the voltage-controlled current source is the controller and the loop gain reference variable. Hence, replacement of the reference variable with a nullor yields the circuit from Figure 10.28 and the asymptotic gain $A_{f\infty}$ equals the ideal gain:

$$A_{f\infty} = -Z_f. \tag{10.90}$$

Loop gain

[20] H.W. Bode. *Network Analysis and Feedback Amplifier Design*. Van Nostrand, New York, 1945

In 1945, Bode [Bode1945][20] published 'Network Analysis and Feedback Amplifier Design'. In this fundamental work, Bode stated that a useful distinction between a forward transfer and a feedback transfer as it was made in Black's feedback model, cannot be made in real physical circuits. This is because in such circuits, parasitic feedback paths almost always exist. Therefore, in practice, the loop gain as the product of the forward gain H, and the feedback factor k, is not always a measure for the quality improvement of the signal transfer.[21]

[21] The term $(1 + Hk)$ in 10.5.

For this reason Bode introduced the terms *return difference* and *return ratio* that have a meaning for a specific selection of a controlled source. The relation between the return difference F and the return ratio T is

$$F = (1 + T)\,. \tag{10.91}$$

The asymptotic gain model is based upon the work of Bode. The loop gain as defined in the asymptotic gain model can be obtained from the return ratio as

$$L = -T. \tag{10.92}$$

Bode showed, that for a given selection of the reference variable, the return difference can be calculated as

$$F = \frac{\det \mathbf{M}}{\det \mathbf{M}^0}\,, \tag{10.93}$$

where $\mathbf{M}$ is the MNA matrix of the network and $\mathbf{M}^0$ is obtained from $\mathbf{M}$, after replacing the gain of the selected reference variable with zero.

Example 10.11

Let us now calculate the loop gain for the circuit from Figure 10.27. We will first calculate the return difference F with g_m as reference. The MNA matrix equation of the network can be formulated as

$$\begin{pmatrix} I_s \\ 0 \end{pmatrix} = \mathbf{M} \begin{pmatrix} V_i \\ V_\ell \end{pmatrix}, \tag{10.94}$$

where

$$\mathbf{M} = \begin{pmatrix} \frac{1}{Z_s} + \frac{1}{Z_f} & -\frac{1}{Z_f} \\ g_m - \frac{1}{Z_f} & \frac{1}{Z_\ell} + \frac{1}{Z_f} \end{pmatrix}. \tag{10.95}$$

The matrix $\mathbf{M}^0$ *is obtained after substitution of* $g_m = 0$ *in* $\mathbf{M}$*:*

$$\mathbf{M}^0 = \begin{pmatrix} \frac{1}{Z_s} + \frac{1}{Z_f} & -\frac{1}{Z_f} \\ -\frac{1}{Z_f} & \frac{1}{Z_\ell} + \frac{1}{Z_f} \end{pmatrix}. \tag{10.96}$$

With the aid of (10.93) the return difference can be obtained as

$$F = 1 + \frac{g_m Z_s Z_\ell}{Z_f + Z_s + Z_\ell}. \tag{10.97}$$

The loop gain L, as it has been defined in the asymptotic gain model, can be obtained from the return difference as

$$L = 1 - F. \tag{10.98}$$

Example 10.12

Let us now calculate the loop gain with g_m *as reference variable for the circuit from Figure 10.27. After substitution of (10.97) in (10.98), we obtain*

$$L = -\frac{g_m Z_s Z_\ell}{Z_f + Z_s + Z_\ell}. \tag{10.99}$$

Servo function

The servo function can be obtained from the return difference as

$$S = \frac{1 - F}{F}. \tag{10.100}$$

Example 10.13

We will evaluate the servo function for the circuit from Figure 10.27 using g_m *as loop gain reference. After substitution of (10.93) in (10.100), we obtain*

$$S = \frac{\det \mathbf{M}^0}{\det \mathbf{M}} - 1, \tag{10.101}$$

which yields

$$S = \frac{-Z_s g_m Z_\ell}{Z_f + Z_s + Z_\ell + Z_s g_m Z_\ell}. \tag{10.102}$$

Direct transfer

The direct transfer is defined as the source-to-load transfer with the gain of the loop gain reference set to zero. Hence, it can be calculated from the matrix $\mathbf{M}^0$.

Example 10.14

The direct source-to-load transfer for the circuit from Figure 10.27, with g_m *selected as loop gain reference variable can be obtained as*

$$\rho = \left.\frac{V_\ell}{I_s}\right|_{g_m = 0} = \frac{(-1)^3 \left(-\frac{1}{Z_f}\right)}{\det \mathbf{M}^0}, \tag{10.103}$$

which yields

$$\rho = \frac{Z_s Z_\ell}{Z_f + Z_s + Z_\ell}. \tag{10.104}$$

Gain

An important property of the asymptotic-gain model is that the source-to-load transfer can be calculated from the asymptotic gain, the loop gain and the direct transfer, or directly from the matrix **M**. Both calculation methods yield the same result.

Example 10.15

$$\frac{V_\ell}{I_s} = \frac{(-1)^3 \left(g_m - \frac{1}{Z_f}\right)}{\det \begin{pmatrix} \frac{1}{Z_s} + \frac{1}{Z_f} & -\frac{1}{Z_f} \\ g_m - \frac{1}{Z_f} & \frac{1}{Z_\ell} + \frac{1}{Z_f} \end{pmatrix}}, \tag{10.105}$$

which yields

$$\frac{V_\ell}{I_s} = \frac{-Z_s Z_\ell \left(Z_f g_m - 1\right)}{Z_f + Z_s + Z_\ell + Z_s g_m Z_\ell}. \tag{10.106}$$

This expression could also have been obtained through application of the asymptotic gain model:

$$\frac{V_\ell}{I_s} = A_{f\infty} S + \frac{\rho}{1 - L}. \tag{10.107}$$

after substitution of (10.90), (10.102), (10.104) and (10.99) into (10.107), we also obtain (10.106).

10.3.10 Conclusions

We have seen that the asymptotic gain model is very well suited for a two-step design of negative feedback amplifiers. The only important condition is that the loop gain reference variable is chosen properly.

The ideal gain as it has been fixed with the feedback elements, matches the asymptotic gain, if the controller behaves as a nullor when the controlled source that has been selected as the loop gain reference variable is replaced with a nullor.

This will not be the case if:

1. The reference variable is selected in a local feedback loop in the controller.

   With operational amplifiers as controllers, this situation can almost always be avoided.

2. Parasitic impedances at one or more port terminals:

   - Introduce an attenuation in the coupling between the source and the amplifier, and/or the amplifier and the load
   - Introduce an attenuation in the feedback network(s).

   In such cases, the asymptotic gain model can still be applied. Those impedances can easily be extracted from the controller and assume to be part of the feedback network, the source or the load. In doing so, the ideal gain is modified and the asymptotic gain will equal this modified ideal gain.

11 Amplifier performance and controller requirements

11.1 Introduction

Negative feedback is an error reduction technique that uses the available power gain of a controller for quality improvement of the transfer of the feedback amplifier. With a high-gain controller, the bandwidth, the linearity and the accuracy of a negative feedback amplifier are primarily defined by the feedback network(s). In the following sections, we will discuss to what extent performance limitations of the controller affect those of the negative feedback amplifier. This knowledge will help us in setting up the performance requirements for the controller.

We have already seen that the controller adds noise to the signal. For given noise sources of the controller, the smallest noise contribution can be achieved with nonenergic feedback. With other implementations of feedback, the feedback network enlarges this contribution and possibly adds noise itself. This deterioration of the noise behavior can be kept small by placing only relatively small impedances in series with the signal path and/or small admittances in parallel with the signal path. These small impedances and admittances can have a large effect on the port impedances of the feedback amplifier (see section 10.3.5). Hence, when compared with the use of brute-force techniques, we are able to design the port impedances, the transfer and the signal-to-noise ratio of a feedback amplifier almost independently. We will see that this is also true for other performance aspects, such as the static accuracy, the nonlinearity and the bandwidth of feedback amplifiers.

With the aid of the asymptotic gain model, the source-to-load transfer of a feedback amplifier can be approximated by the product of its ideal transfer and the servo function.[1] With a nullor as controller the servo function equals unity and the source-to-load transfer equals the ideal transfer. A practical controller cannot behave as a nullor; it will have a finite available power gain and suffer from speed and power limitations. Therefore, in practice, the servo function will only approximate unity over a limited operating range of the controller. Since the servo function is completely determined by the loop gain, designing the servo function means designing the loop gain. The elements that contribute to the loop gain are:

[1] If the loop gain reference variable has been selected properly and if the direct transfer can be neglected.

1. The controller
2. The feedback network
3. The source and the load impedance

The source and the load impedances are part of the application specification of the amplifier. The feedback network has been designed on grounds of the desired transfer, its noise contribution and its effect on the power efficiency. Design or selection of the controller is what remains. This starts with setting up its performance requirements. The specification of the controller requirements in relation to the performance of the amplifier is the main topic of this chapter.

11.1.1 This chapter

In the following sections, we will study the influence of the following performance limitations:

1. Finite static loop gain

   A finite static (DC) loop gain can be caused by:

   (a) A finite static gain of the controller
   (b) One or more zeros in the loop gain at $s = 0$

(c) One or more zeros at $s = 0$ in the impedance or in the admittance of the source and/or the load

The effect on the static accuracy of the feedback amplifier will be discussed in section 11.2.

2. Nonlinear loop gain

Nonlinearity of the loop gain can be caused by:

(a) Nonlinear transfer of the controller

(b) Nonlinearity in the transfer of the feedback network

(c) Nonlinear behavior of the source or the load

The relation between nonlinear behavior of the loop gain and that of the feedback amplifier will be discussed in section 11.3.

3. Bandwidth limitation of the loop gain

Bandwidth limitation of the loop gain can be caused by:

(a) Bandwidth limitation of the controller gain

(b) Bandwidth limitation in the transfer of the feedback network

(c) Dynamic character of the source and/or the load

The relation between the dynamic behavior of the loop gain and that of the feedback amplifier will be discussed in section 11.4.

11.2 Accuracy design considerations

The static inaccuracy of a system is defined in section 17.5.2. The static inaccuracy of a negative feedback amplifier is determined by the static inaccuracy of the ideal gain and by that of the servo function. The static inaccuracy of the ideal gain is determined by the inaccuracy of the static transfer of the feedback network. The static inaccuracy of the servo function is determined by the value of the *DC* loop gain.

11.2.1 Static inaccuracy of the servo function

The static inaccuracy δ_{DC} introduced by the servo function is its difference from unity at DC:

$$\delta_{DC} = S_{DC} - 1, \tag{11.1}$$

where S_{DC} is the value of the servo function at zero frequency. With the aid of the DC loop gain L_{DC}, we may write

$$\delta_{DC} = \frac{-L_{DC}}{1 - L_{DC}} - 1 = -\frac{1}{1 - L_{DC}}. \tag{11.2}$$

For $|L_{DC}| \gg 1$, this can be approximated by

$$\delta_{DC} \approx \frac{1}{L_{DC}}. \tag{11.3}$$

Hence, for a static inaccuracy of -1%, we need a DC loop gain of about -100. Negative feedback will result in a one-sided static inaccuracy, in other words, the servo function will always be less than unity. The error can be made symmetrical by compensating 50% of the error budget in the ideal gain.

11.2.2 Design conclusion

In general, we may say that a high accuracy requires a large loop gain:

> In order to obtain the static inaccuracy of a feedback amplifier within specifications, the controller must provide a sufficiently large DC loop gain.

11.3 Nonlinearity design consideration

Negative feedback is capable of reducing errors due to the nonlinear behavior of the controller. The nonlinearity of the controller causes nonlinearity of the loop gain and of the servo function. Let us assume the asymptotic gain is linear and equals the ideal gain. The nonlinearity of the gain of the feedback amplifier is then fully determined by the nonlinearity of the loop gain.

11.3.1 Loop gain differential gain error

Let $\varepsilon_{y,L}$ be the differential gain error of the loop gain L_y at a load signal excursion y from the quiescent operating point. The loop gain in the quiescent operating point equals L_Q. The differential gain of the loop gain is defined as

$$\varepsilon_{y,L} = \frac{L_y - L_Q}{L_Q}. \tag{11.4}$$

The differential gain ε_{y,A_f} of the source-to-load transfer is defined as

$$\varepsilon_{y,A_f} = \frac{A_{f,y} - A_{f,Q}}{A_{f,Q}} \tag{11.5}$$

in which $A_{f,y}$ is the gain of the source-to-load transfer at a load signal excursion from the quiescent operating point of y:

$$A_{f,y} = A_{f\infty}\frac{-L_y}{1 - L_y}, \tag{11.6}$$

and $A_{f,Q}$ is the source-to-load transfer in the quiescent operating point Q:

$$A_{f,Q} = A_{f\infty}\frac{-L_Q}{1 - L_Q}. \tag{11.7}$$

Substitution of (11.6) and (11.7) in (11.5) yields

$$\varepsilon_{y,A_f} = \frac{A_{f\infty}\frac{-L_y}{1-L_y} - A_{f\infty}\frac{-L_Q}{1-L_Q}}{A_{f\infty}\frac{-L_Q}{1-L_Q}}. \tag{11.8}$$

This can be written as

$$\varepsilon_{y,A_f} = \frac{L_y - L_Q}{L_Q}\frac{1}{1 - L_y}. \tag{11.9}$$

With the aid of (11.4), we may write this as

$$\varepsilon_{y,A_f} = \frac{\varepsilon_{y,L}}{1 - L_y}. \tag{11.10}$$

Hence, the differential gain of a negative feedback amplifier at a load signal excursion y from the quiescent operating point equals that of the controller at the load operating point y, divided by the return difference at that operating point.

If the differential gain of the loop gain at a load signal excursion y from the quiescent operating point is small, we may also write this as:

$$\varepsilon_{y,A_f} \cong \frac{\varepsilon_{y,L}}{1 - L_Q}. \tag{11.11}$$

Hence, the differential gain of a negative feedback amplifier can be reduced by increasing the loop gain in such a way that the differential gain error of the loop gain is not increased.

11.3.2 Design conclusion

The contribution of the controller to the nonlinearity of the amplifier can be minimized by minimizing its contribution to the differential error-to-gain ratio of the loop gain.

For this reason, we will use the contribution of the controller to the *differential error-to-gain ratio* of the loop gain, as a figure of merit for a controller and for any amplification stage in it.

11.4 Bandwidth design considerations

In this section, we will study the dynamic behavior of negative feedback amplifiers and see the way in which the dynamic behavior of the controller contributes to that of the feedback amplifier. To do so, the negative feedback amplifier will be modeled as a linear time-invariant dynamic system.

With the aid of the asymptotic gain model, the source-to-load transfer of a negative feedback amplifier can be written as

$$A_f(s) = A_{f\infty}(s)S(s), \tag{11.12}$$

in which s is the Laplace variable.

With proper selection of the loop gain reference variable, the asymptotic gain $A_{f\infty}(s)$ equals the ideal gain of the amplifier. During the conceptual design of the amplifier, any desired filter characteristic can be given to the ideal gain. This is done by designing feedback networks that establish the required dynamic behavior. Hence, at first glance, it seems difficult to give a definition of the bandwidth of a negative feedback amplifier, because it can have many intended filter characteristics. However, with the aid of the servo function $S(s)$, we are able to decouple the desired filter characteristics from bandwidth limitations caused by controller imperfections. These limitations are described by the servo function:

1. A high loop gain over the frequency range of interest results in a servo function whose magnitude approaches unity at these frequencies.

2. A lack of loop gain at high frequencies and/or at low frequencies manifests itself in a high-frequency and/or low-frequency roll-off of the servo function.

Hence, the bandwidth of an approximate unity-gain servo function can be used as a figure of merit for the dynamic performance of a negative feedback amplifier. A servo function with a large bandwidth implies that the transfer of the negative feedback amplifier approximates its ideal transfer over a large frequency range.

11.4.1 Midband frequency range and bandwidth

If we want to achieve a small error over the frequency range of interest, we need a large loop gain at these frequencies. The frequency range over which the loop gain is much larger than unity will be referred to as the *midband frequency range*. The loop gain at these frequencies will be referred to as the *midband loop gain* L_{MB}. In this frequency range, the midband inaccuracy equals $\delta_{MB} \cong \frac{1}{L_{MB}}$; this follows from (11.3). Outside the midband frequency range, the error with respect to the asymptotic gain increases due to a lack of loop gain: $|L(j\omega)| < 1$, and the controller is no longer capable of reducing errors. We will define the bandwidth of a negative feedback amplifier as follows:

> The bandwidth of the negative feedback amplifier is defined as the distance between the lowest frequency ω_ℓ and the highest frequency ω_h at which the servo function drops -3dB below its midband frequency value. Otherwise, the bandwidth of a negative feedback amplifier is the -3dB bandwidth of the servo function.

With this definition, we have decoupled the intended filter characteristic of the amplifier from the bandwidth limitation caused by controller limitations. If $|L(j\omega)| \gg 1$, the bandwidth of a negative feedback amplifier defines the frequency range over which the gain of the feedback amplifier approaches its ideal gain.

In cases in which $L(s)$ has no zeros, the midband frequency loop gain L_{MB} equals the DC value of the loop gain L_{DC}, and the servo function $F_s(s)$ has a low-pass filter characteristic. We will first study the design considerations for this case and later study situations in which there are zeros in the loop gain.

11.4.2 All-pole loop gain functions

For all-pole loop gain functions that have no poles at $s = 0$, $L(s)$ can be expressed in its DC value L_{DC} and its poles. A loop gain function with n poles p_i can then be written as

$$L(s) = \frac{L_{DC}}{\prod_{i=1}^{n}\left(1 - \frac{s}{p_i}\right)}. \tag{11.13}$$

With the aid of (10.29), the servo function can then be obtained as

$$S(s) = \frac{-L_{DC}}{\prod_{i=1}^{n}\left(1 - \frac{s}{p_i}\right) - L_{DC}}. \tag{11.14}$$

We will write this expression as the product of a DC term and a unity-gain dynamic transfer:

$$S(s) = \frac{-L_{DC}}{1 - L_{DC}} \frac{1}{1 + ... + (-1)^n \frac{s^n}{(1-L_{DC})\prod_{i=1}^{n} p_i}}. \tag{11.15}$$

The DC term describes the static inaccuracy with respect to the asymptotic gain. The static inaccuracy has already been discussed in section 11.2. The unity-gain dynamic transfer function has our interest and particularly the

highest coefficient of s of the denominator. This term describes the roll-off at the highest frequencies. From (11.15), we see that this coefficient equals the reciprocal value of the product of the poles and the DC return difference $(1 - L_{DC})$. With $-L_{DC} \gg 1$, it approximates the product of the DC loop gain and the poles. In the following section, we will derive a relation between this so-called *loop gain-poles product* and the bandwidth of the servo function.

11.4.3 Low-pass cut-off and loop gain-poles product

The aim of the design of the high-frequency behavior is to let the servo function $S(s)$ approximate unity over as wide a frequency range as possible and have a smooth transition to roll-off at high frequencies, while all poles of $S(s)$ have a negative real part. Hence, we will assume that a Maximum Flat Magnitude (MFM) or Butterworth filter characteristic has to be established for the servo function.

MFM response

The bandwidth ω_n of a system with an $n-th$ order MFM characteristic shows a very simple relation to the n system poles $p_1 \cdots p_n$:

$$\omega_n = \sqrt[n]{\prod_{i=1}^{n} |p_i|}. \tag{11.16}$$

The poles of MFM systems can be found by taking the n poles with a negative real part from the $2n$ poles spaced equally over a circle centered at the origin of the complex plane. The pole positions of first, second and third order Butterworth filters are shown in Figure 11.1.

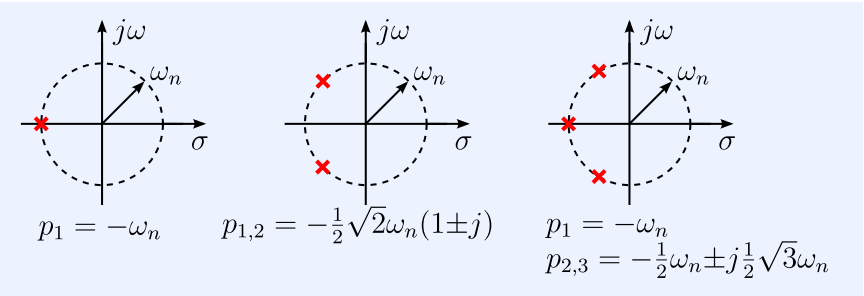


Figure 11.1: Pole positions of first, second and third order Butterworth filters.

The unity-gain MFM transfers of the first-order, second-order and third-order are:

$$H_1(s) = \frac{1}{1 + s/\omega_n}, \tag{11.17}$$

$$H_2(s) = \frac{1}{1 + s\sqrt{2}/\omega_n + s^2/\omega_n^2}, \tag{11.18}$$

$$H_3(s) = \frac{1}{1 + 2s/\omega_n + 2s^2/\omega_n^2 + s^3/\omega_n^3}. \tag{11.19}$$

In order to achieve an MFM characteristic for the servo function, the poles of $S(s)$ have to be manipulated into Butterworth positions. We will assume that the poles of $S(s)$ can be divided into two groups:

1. Dominant poles

   The dominant poles are the poles that can be manoeuvred into Butterworth positions; they will determine the bandwidth of the servo function.

2. Non-dominant poles

   The non-dominant poles are assumed to have no significant influence on the high-frequency behavior, since their frequency is far beyond the -3dB cut-off frequency of the servo function.

Hence, according to equations (11.15) and (11.16), the bandwidth ω_n of a negative feedback amplifier with an $n-th$ order MFM characteristic can be obtained as:

$$\omega_n = \sqrt[n]{\left|(1 - L_{DC}) \prod_{i=1}^{n} p_i\right|}. \tag{11.20}$$

With $|L_{DC}| \gg 1$, we see that the bandwidth of a negative feedback amplifier is determined by the product of the DC loop gain and the *dominant poles* of the loop gain. Since this is larger than the bandwidth of the loop gain itself, we see that negative feedback is capable of reducing the errors as a result of the speed limitations introduced by the controller.

In order to establish an MFM characteristic, the servo function must be an all-pole function, and the dominant poles of the servo function, have to be manipulated into the desired filter positions. Before discussing *frequency compensation techniques* for manipulating these poles into Butterworth positions, we will introduce techniques for separating the high-frequency poles of $L(s)$ into dominant and non-dominant poles.

Other filter types

The design method for high frequency behavior is also valid for other filter characteristics, e.g. Gaussian or Chebyshev filters. For filters with a Gaussian response, such as Bessel Filters, ω_n is smaller, and for filters with an MFM response. For a Chebyshev response, ω_n is larger than that obtained from 11.16. In all cases, however, the loop gain-poles product can be taken as a design parameter for the bandwidth of a negative feedback amplifier.

Dominant poles

We will now present a way to determine whether poles are dominant or non-dominant. Figure 11.2 shows an example of the asymptotes of the magnitude characteristics of the asymptotic gain $A_{f\infty}(\omega)$, the source-to-load transfer $A_f(\omega)$, the loop gain $L(\omega)$ and the servo function $S(\omega)$ of a negative feedback amplifier. The intended frequency characteristic of the amplifier's source-to-load transfer is that of $A_{f\infty}(\omega)$. In this figure, the loop gain $L(\omega)$ has been given three poles, p_1, p_2 and p_3, of which the corresponding frequencies are denoted as ω_{p1}, ω_{p2} and ω_{p3}, respectively.[2] The high-frequency MFM -3dB frequency ω_h of the negative feedback amplifier is the frequency at which the asymptote of the magnitude characteristic of the servo function intersects with its DC value. This frequency equals ω_2. This figure clearly shows the second order low-pass character of the magnitude characteristic. Obviously, the third pole does not belong to the dominant group.

[2] The frequency ω_p of a pole p is defined as its magnitude: $\omega_p = \sqrt{\mathrm{Re}(p)^2 + \mathrm{Im}(p)^2}$.

This can be seen as follows. If only p_1 was assumed dominant, the servo function would have had a bandwidth of ω_1. This assumption would have been wrong, because $\omega_{p2} < \omega_1$. If all three poles were to have been assumed dominant, the servo bandwidth would have been ω_3. This assumption would also have been wrong, because at ω_3, the servo function has already dropped below its DC value.

The procedure for finding the order of the system (the number of dominant poles) is thus as follows:

1. Rank the high-frequency poles of $L(s)$ with increasing frequency: start with the most dominant pole[3] and end with the pole with the highest

[3] The pole with the lowest frequency.

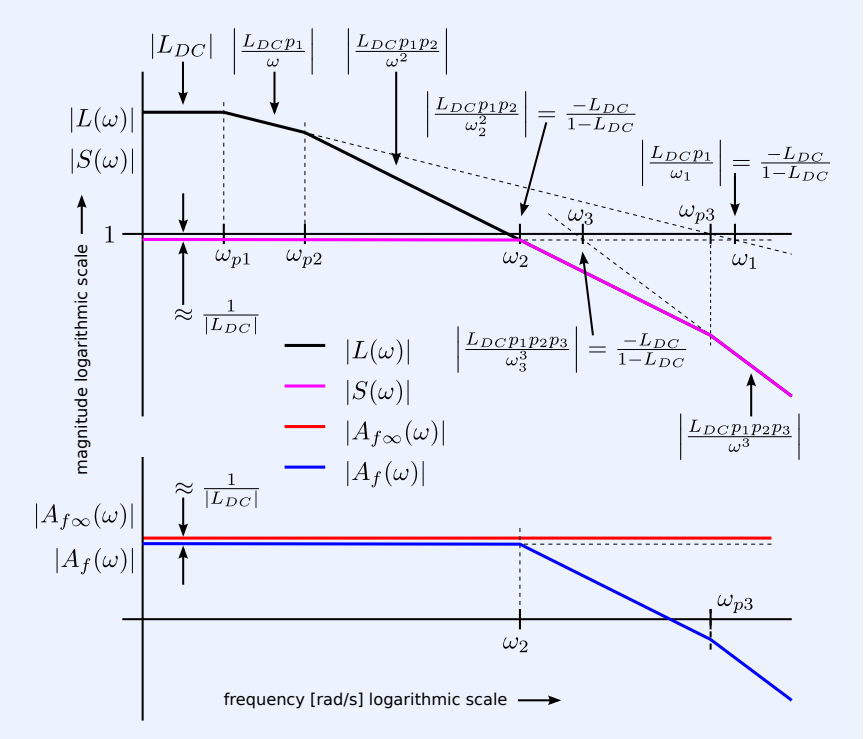


Figure 11.2: Example of the frequency characteristics of the asymptotic gain, the loop gain and servo function of a negative feedback amplifier.

frequency; here, $|p_1| < |p_2| < |p_3|$.

2. Calculate the -3 dB frequency for increasing order of the system: for the amplifier from this example, we find

$$\omega_1 = |(1 - L_{DC})\, p_1|\,, \tag{11.21}$$

$$\omega_2 = \sqrt{|(1 - L_{DC})\, p_1 p_2|}, \tag{11.22}$$

$$\omega_3 = \sqrt[3]{|(1 - L_{DC})\, p_1 p_2 p_3|}. \tag{11.23}$$

3. The order n is the number for which ω_n has the smallest value.

Alternatively, we may say that a pole is dominant if its frequency is below the frequency of intersection of its corresponding asymptote of the magnitude characteristic of the loop gain, and unity (see Figure 11.2).

Design conclusion

We are now able to relate the bandwidth of a negative feedback amplifier to properties of the controller and formulate design requirements for the controller.

> In order to obtain the desired low-pass cut-off frequency of a negative feedback amplifier, the controller should raise the product of the loop gain and the dominant poles to a sufficiently large value.

Although the asymptotic description provides a method for finding the requirements for the controller's contribution to the loop gain-poles product, the magnitude characteristic of the servo function can deviate considerably from the asymptotic values. We have already seen that an MFM characteristic requires the poles of the servo function to be in Butterworth positions. Usually, the dominant poles of the servo function will not have their desired

positions and *frequency compensation techniques* have to be applied to bring these poles into their desired positions. Independent design of the bandwidth and the frequency response requires that application of frequency compensation techniques does not affect the product of the loop gain and the poles. Such techniques will be discussed in Chapter 12. Before we discuss these techniques, we need to study the influence of zeros in the loop gain on the bandwidth of the servo function. This will be done in the following sections.

We will conclude this section with two examples in which we will derive requirements for the properties of an operational amplifier that determine the bandwidth of a transimpedance amplifier in which it has been used as a controller. In the first example, we will derive the design equations manually. In the second example, we will use numeric values and demonstrate the use of SLiCAP.

Example 11.1

In this example, we will derive requirements for the gain-bandwidth product of an operational amplifier that will be used as a controller in a transimpedance amplifier. The transimpedance amplifier is shown in Figure 11.3A. It converts the current of a capacitive source into a voltage across a resistive load. The small-signal equivalent model of the single-pole operational amplifier is shown in Figure 11.3B. The complete small-signal model of the circuit is depicted in Figure.11.4.

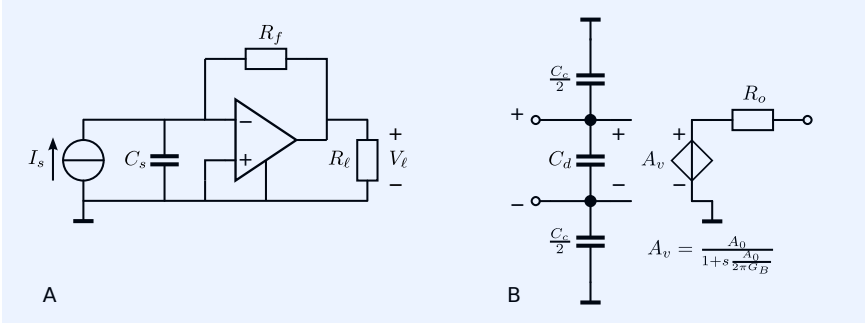


Figure 11.3:
A: Transimpedance amplifier that converts the current from a capacitive source into a voltage across a resistive load.
B: Model of the operational amplifier used in (A).

With the aid of the asymptotic gain model, we are able to find a relation between the bandwidth of the transimpedance amplifier and properties of the operational amplifier. To do so, we select the voltage gain A_v of the operational amplifier as the loop gain reference. With this selection, the asymptotic gain equals the ideal gain of the amplifier:

$$A_{f\infty} = -R_f. \tag{11.24}$$

The error with respect to this ideal transfer is now completely determined by the servo function, while, in its turn, the servo function is completely defined by the loop gain.

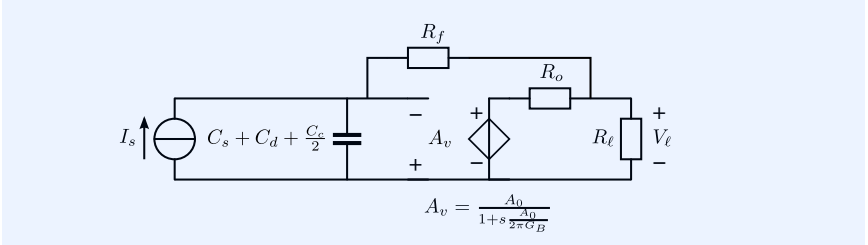


Figure 11.4: Small-signal model of the transimpedance amplifier from Figure 11.3 with the operational amplifier model from Figure 11.3B.

An expression for the loop gain can be found from network inspection:

$$L = -\frac{A_0}{1+s\frac{A_0}{2\pi G_B}}\frac{R_\ell}{R_\ell+R_o}\frac{1}{1+s\left((R_o/\!/R_\ell)+R_f\right)\left(C_s+C_d+\frac{C_c}{2}\right)}. \quad (11.25)$$

The DC loop gain L_{DC} equals

$$L_{DC} = -A_0\frac{R_\ell}{R_\ell+R_o}. \quad (11.26)$$

The two poles of the loop gain are found to be

$$p_1 = -\frac{2\pi G_B}{A_0}\ [rad/s] \quad (11.27)$$

$$p_2 = -\frac{1}{\left((R_o/\!/R_\ell)+R_f\right)\left(C_s+C_d+\frac{C_c}{2}\right)}\ [rad/s]. \quad (11.28)$$

If both poles are dominant, the amplifier can be realized with a second order MFM response with bandwidth B_ω that can be obtained from the product of the DC loop gain and the poles:

$$B_\omega = \sqrt{\frac{2\pi G_B R_\ell}{R_\ell+R_o}\frac{1}{\left((R_o/\!/R_\ell)+R_f\right)\left(C_s+C_d+\frac{C_c}{2}\right)}}\ [rad/s]. \quad (11.29)$$

From (11.29), we can derive a relation between the circuit parameters and the achievable bandwidth B_f (in [Hz]) of the transimpedance amplifier:

$$G_B\frac{R_\ell}{R_\ell+R_o}\frac{1}{\left((R_o/\!/R_\ell)+R_f\right)\left(C_s+C_d+\frac{C_c}{2}\right)} \geq 2\pi B_f^2\ [Hz]. \quad (11.30)$$

The unknown parameters in this equation are G_B, R_o, C_d and C_c. The other parameters C_s, R_f and R_ℓ, as well as the required bandwidth B_f, follow from the specification of the amplifier.

The maximum contribution that the operational amplifier can deliver to the loop gain-poles product of the transimpedance amplifier is G_B. This occurs if

$$C_d + \frac{C_c}{2} \ll C_s, \quad (11.31)$$

$$R_o \ll R_\ell, \quad (11.32)$$

$$R_o \ll R_f. \quad (11.33)$$

If these conditions are met, the relation between the achievable bandwidth B_f and the gain-bandwidth product G_B of the operational amplifier simplifies to

$$B_f = \sqrt{\frac{G_B}{2\pi R_f C_s}}\ [Hz], \quad (11.34)$$

and we find a design limit for the gain-bandwidth product of the amplifier:

$$G_B \geq 2\pi R_f C_s B_f^2\ [Hz]. \quad (11.35)$$

$G_B = 2\pi R_f C_s B_f^2$ is a lower limit for the gain-bandwidth product. Any device with a smaller value would be a show-stopper.[4]

Until now, we have assumed that both poles are dominant. This is not necessarily true. If we require a small bandwidth, e.g., smaller than the frequency of p_2, this pole may be kept away from the dominant pole. The expression for the

[4] Action, condition, event, or problem that is serious enough to halt an activity, program, or process until it is resolved (http://www.businessdictionary.com).

loop gain then changes to:

$$L = -\frac{A_0}{1 + s\frac{A_0}{2\pi G_B}} \frac{R_\ell}{R_\ell + R_o}. \tag{11.36}$$

The lower limit for G_B then follows from

$$G_B > \frac{R_\ell + R_o}{R_\ell} B_f. \tag{11.37}$$

However, this first order approximation is only valid if the second pole is not dominant thus,

$$B_f \ll \frac{1}{2\pi \left((R_o /\!/ R_\ell) + R_f \right) \left(C_s + C_d + \frac{C_c}{2} \right)}, \tag{11.38}$$

and if the gain-bandwidth product is sufficiently small

$$G_B \ll \frac{R_\ell + R_o}{R_\ell} \frac{1}{2\pi \left((R_o /\!/ R_\ell) + R_f \right) \left(C_s + C_d + \frac{C_c}{2} \right)}. \tag{11.39}$$

In the next example, we will use numeric values and demonstrate the use of SLiCAP for determination of controller properties that contribute to the low-pass cut-off of the amplifier.

Example 11.2

We will start with the definition of the target specifications of the transimpedance amplifier from the previous example. We will consider its dynamic behavior only. The target specifications that are relevant for the dynamic behavior are:

1. *Source capacitance:* $5pF$
2. *Load resistance:* $2k\Omega$
3. *Transimpedance gain:* $100k\Omega$
4. *Signal frequency range:* $500kHz$

The circuit for the generation of the netlist is depicted in Figure 11.5.

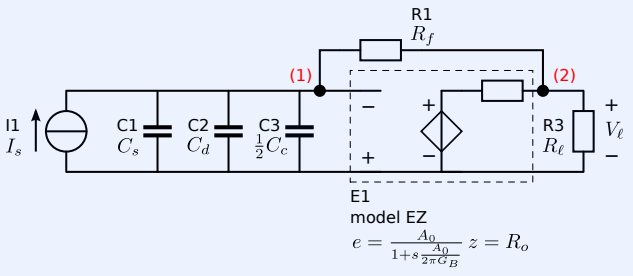


Figure 11.5: Circuit diagram of the transimpedance amplifier for evaluation with SLiCAP.

The SLiCAP *netlist:*

```
transimpedance
* file: transimpedance.cir
* SLiCAP circuit file
I1 0 1 {I_s}
C1 1 0 {C_s}
C2 1 0 {C_d}
C3 1 0 {C_c/2}
R1 1 2 {R_f}
R2 2 0 {R_ell}
E1 2 0 0 1 EZ value={A_0/(1+s*A_0/2/pi/G_B)} zs={R_o}
```

```
.param C_s=5p R_f=100k R_ell=2k
.end
```

Below is the part of the listing of the script file that performs the checking of the circuit and the definition of the simulation variables, such as the simulation type, the gain type, the data type, the source, the load and the loop gain reference. The simulation type has been set to numeric, so all parameter values defined with the `.param` *statement in the circuit file will be substituted into the element expressions.*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# file: transimpedance.py

from SLiCAP import *

fileName = 'transimpedance'
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName + '.cir')
htmlPage('Example controller GB product requirements')
i1.setSimType('numeric')
i1.setDataType('laplace')
i1.setSource('I1')
i1.setDetector('V_2')
```

Lines 16-19 of the script evaluate the loop gain. The function `coeffsTransfer` *in line 21 returns the coefficients of the numerator and the denominator of its argument: a Laplace rational function. From these coefficients we will calculate the loop gain-poles product. This script is intended to work with an all-pole loop gain that has a finite nonzero DC value. Hence, the coefficient of* s^0 *in the denominator should not be zero and the numerator should only have a nonzero coefficient for* s^0*. This is checked in line 21-27. Line 29-31 print DC loop gain on the html page. The loop gain-poles product is calculated in line 35, and displayed on the html page (line 37 -40):*

```

i1.setLGref('E1')
i1.setGainType('loopgain')
result = i1.execute()
#
gain, numerCoeffs, denomCoeffs = coeffsTransfer(result.laplace)
if len(numerCoeffs) > 1:
  text2html('Zeros found, GB determination method not valid.')
else:
  text2html('Found the nonzero DC loop gain.')
if denomCoeffs[0] == 0:
  text2html('Found poles in the origin, GB determination method not valid.')
#
LP = sp.symbols('LP')
text2html('The DC loop gain equals:')
eqn2html('L_DC', gain)

if len(denomCoeffs) == 1:
  text2html('No poles found, GB determination method not valid.')
LPproduct = -gain/denomCoeffs[-1]

text2html('The loop gain-poles product is found as:')
eqn2html(LP, LPproduct)
order = len(denomCoeffs) - 1
text2html('The order of the LP product is: ' + str(order))
```

This LP product is evaluated as the ratio of the coefficient of s^0 *of the numerator and the highest order coefficient of s of the denominator. The order of the loop gain is found as the number of coefficients of the denominator minus one.*[5]

[5] All poles are assumed dominant.

With the expression for the LP product and the requirement for the bandwidth, we are able to derive requirements for the unknown variables.

We will now calculate a show-stopper value for the gain-bandwidth product GB of the operational amplifier. The lowest possible gain-bandwidth product is found if all other contributions to the bandwidth limitation of the amplifier are set

to zero. This is the case if $R_o = 0$, $C_d = 0$, and $C_c = 0$.

```
42 B_f = 500e3
43 R_o, C_d, C_c, G_B, GB_min, A_0 = sp.symbols('R_o, C_d, C_c, G_B, GB_min, A_0'
       )
44
45 GB_minAll = sp.solve(LPproduct -(B_f*2*sp.pi)**order, G_B)[0]
46 text2html('The required bandwidth  = ' + str(B_f/1000) + 'kHz')
47 text2html('With this value, the show stopper value of the gain-bandwidth
       product $G_B$ is:')
48 GB_minNum = GB_minAll.subs([(R_o, 0), (C_d, 0), (C_c, 0)])
49 eqn2html('GB_min', GB_minNum/1e6, units="MHz")
```

This minimum requirement for the GB product is calculated in lines 42-49 of the script. It consists of two parts. In line 45 we find a symbolic expression for GB as a function of the other parameters of the operational amplifier and the required bandwidth from solving:

$$LP_n - \left(2\pi B_f\right)^n = 0.$$

In line 48 we subsitute $R_o = 0$, $C_d = 0$, and $C_c = 0$ in this expression. This yields the lower limit $GB_{\min}$ of GB.

Figure 11.6 shows the html page generated by this script.

Example controller GB product requirements

Found the nonzero DC loop gain.

The DC loop gain equals:

$$L_{DC} = -\frac{4.0 \cdot 10^{11} A_0}{2.0 \cdot 10^8 R_o + 4.0 \cdot 10^{11}} \tag{1}$$

The loop gain-poles product is found as:

$$LP = \frac{4.0 \cdot 10^{11} A_0 \left(2.0 \cdot 10^8 \pi G_B R_o + 4.0 \cdot 10^{11} \pi G_B\right)}{\left(2.0 \cdot 10^8 R_o + 4.0 \cdot 10^{11}\right)\left(5.1 \cdot 10^{12} A_0 C_c R_o + 1.0 \cdot 10^{16} A_0 C_c + 1.02 \cdot 10^{13} A_0 C_d R_o + 2.0 \cdot 10^{16} A_0 C_d + 51.0 A_0 R_o + 1.0 \cdot 10^5 A_0\right)} \tag{2}$$

The order of the LP product is: 2

The required bandwidth = 500.0kHz

With this value, the show stopper value of the gain-bandwidth product G_B is:

$$GB_{min} = 0.7854 \text{ [MHz]} \tag{3}$$

Figure 11.6: Result of the execution of the SLiCAP script.

We will now select an operational amplifier based on the results of the previous example. Operational amplifiers with zero input capacitance are not available, so we need to select one with a larger GB product than the show-stopper value. As far as its dynamic performance is concerned, the `OPA627` looks promising. It has a differential-mode input capacitance of 8pF, a common-mode input capacitance of 7pF, an output resistance of about 55Ω and a gain-bandwidth product of 16MHz. Its magnitude and phase plot show a somewhat more complicated behavior than a simple first-order roll-off, but we will ignore this because the extra phase shift above 10MHz may not be at all relevant if the servo bandwidth turns out to be much less.

In the next example, we will verify the design with the `OPA627` device with SLiCAP.

Example 11.3

The script for the device selection and the verification are listed below. Lines 51-63 perform the substitution of the parameters of the `OPA627`, and the calculation of the bandwidth that can be achieved with this device. The results are displayed on the html page; see Figure 11.7.

```
51 htmlPage('Device selection and verification')
52 i1.defPar('A_0', '1M')
53 i1.defPar('C_d', '8p')
54 i1.defPar('C_c', '7p')
55 i1.defPar('R_o', 55)
56 i1.defPar('G_B', '16M')
57 i1.defPar('I_s', 1)
58 params2html(i1.circuit)
59
60 Bf = 1/(2*sp.pi)*LPproduct**(1/order)
61 BfOPA627 = sp.N(Bf.subs([(R_o, 55), (C_d, 8e-12), (C_c, 7e-12), (G_B,16e6),  (
       A_0, 1e6)]), 4)
62 text2html('The achievable low-pass cut-off frequency $f_h$ with the OPA627 in
       [MHz] is:')
63 eqn2html('f_h', BfOPA627*1e-6)
```

Lines 64-77 create the Bode plots for all transfer functions of the asymptotic gain model. These plots are shown in Figure 11.8.

Figure 11.7: Result of the execution of the SLiCAP script.

Device selection and verification

Table: Parameter definitions in 'transimpedance'.

Name	Symbolic	Numeric
A_0	$1.0 \cdot 10^{6}$	$1.0 \cdot 10^{6}$
C_c	$7.0 \cdot 10^{-12}$	$7.0 \cdot 10^{-12}$
C_d	$8.0 \cdot 10^{-12}$	$8.0 \cdot 10^{-12}$
C_s	$5.0 \cdot 10^{-12}$	$5.0 \cdot 10^{-12}$
G_B	$1.6 \cdot 10^{7}$	$1.6 \cdot 10^{7}$
I_s	1	1
R_ℓ	2000.0	2000.0
R_f	$1.0 \cdot 10^{5}$	$1.0 \cdot 10^{5}$
R_o	55	55.0

The achievable low-pass cut-off frequency f_h with the OPA627 in [MHz] is:

$$f_h = 1.225 \tag{1}$$

```
60 htmlPage('Bode plots')
61 L = i1.execute()
62 i1.setGainType('asymptotic')
63 A = i1.execute()
64 i1.setGainType('servo')
65 S = i1.execute()
66 i1.setGainType('direct')
67 D = i1.execute()
68 i1.setGainType('gain')
69 G = i1.execute()
70 figMag = plotSweep('TrimpMag', 'Magnitude characteristics', [L, A, S, D, G],
       10e3, 10e6, 200, funcType = 'mag', show = True)
71 fig2html(figMag, 800)
72 figPhase = plotSweep('TrimpPhase', 'Phase characteristics', [L, A, S, D, G],
       10e3, 10e6, 200, funcType = 'phase', show = True)
73 fig2html(figPhase, 800)
```

The plots clearly show the low-pass cut-off at the predicted frequency. However, the magnitude characteristic of the gain shows significant peaking and is far from maximally flat. *This was not yet our aim. Before we corcern ourselves about the frequency stability and the filter type of the response, we need to establish a sufficiently large bandwidth!*

In the previous examples, we have demonstrated the way in which the gain bandwidth requirements for the controller can be derived from the specification of the transimpedance amplifier. Initial target values can be derived with SLiCAP, and after a device has been selected, the performance can be verified. In this case, this has been done with the aid of Bode plots.

The next step in the design of the amplifier is the design of the frequency stability and the filter response type. These topics will be discussed in Chapter 12. However, before dealing with these topics, we need to study the way

in which the above results can be applied if we do not have all-pole loop gain functions, or if we have poles at $s = 0$, which would yield an infinite DC loop gain.

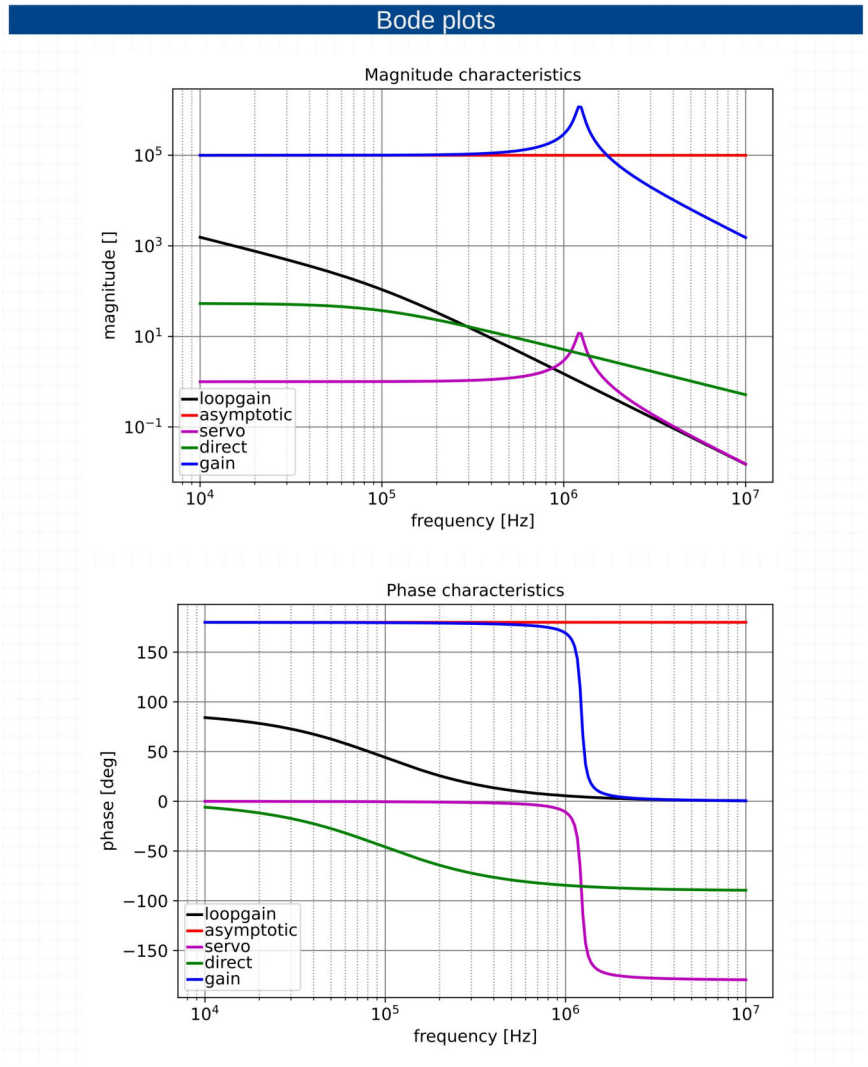


Figure 11.8: Result of the execution of the SLiCAP script.

In section 11.4.4, we will study the high-pass cut-off that arises if we have DC blocking elements in the loop of a feedback amplifier. In section 11.4.5, we will formulate the general case for relating the low-pass cut-off of the servo function to the loop gain with poles and zeros. In section 11.4.6, we will do this for the high-pass cut-off. In section 11.4.7, we combine the methods found in sections 11.4.5 and 11.4.6 and derive a clear procedure for relating the low-pass and the high-pass cut-off of the servo function to the properties of the loop gain function. The algorithm developed in this section has been implemented in SLiCAP.

11.4.4 High-pass cut-off and DC loop gain

In the previous section, we studied the relation between the low-pass cut-off frequency of a negative feedback amplifier and the DC value and the poles of the loop gain. We have seen that, due to a lack of loop gain at high frequencies, the servo function drops below unity. This causes a low-pass behavior with respect to the desired ideal gain of the amplifier. In this section, we will

study the influence of a lack of loop gain below certain frequencies. Such effects may be caused by capacitive branches in series with the signal path or inductive branches in parallel with the signal path in the loop equivalent circuit. If the loop gain drops below unity for decreasing frequencies, the servo function obtains a high-pass character.

DC blocking elements in the loop

DC blocking elements in the loop cause a loss of loop gain at low frequencies. DC blocking elements are capacitors in series with the signal path or inductors in parallel with the signal path, they introduce zeros at zero frequency in the loop gain. In order to study their effect, we assume a loop gain function with q DC blocking elements in the loop and no other dynamic elements. In such cases the loop transfer can be written as:

$$L(s) = \frac{b_q s^q}{\prod_{i=1}^{q}\left(1 - \frac{s}{p_i}\right)} \tag{11.40}$$

The factor b_q can be written as the product of a DC term which is the product of the DC gain A_{DC} provided by the controller, a DC attenuation factor α_{DC} due to resistive networks in the loop transfer, and a product of q time constants introduced by the q DC blocking elements in the loop:

$$b_q = A_{DC}\alpha_{DC}\prod_{i=1}^{q}\tau_i \tag{11.41}$$

As stated earlier, all other elements that introduce low-pass transfer (speed limitations) have been ignored.

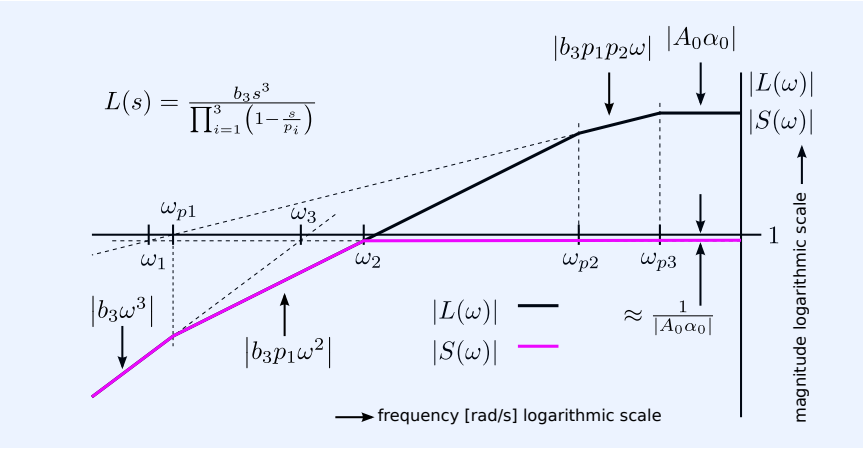


Figure 11.9: High-pass character of the servo function due to 3 DC blocking elements in the loop transfer circuit.

The loop gain now has q zeros in the origin and q poles. The order of the high-pass cut-off equals the number of zeros minus the number of poles below the frequency of intersection of the asymptote of the magnitude characteristic of the loop gain and unity. If the frequency of the pole is in a frequency range where the loop gain is smaller than unity, it reduces this order. We will call such a pole dominant because it affects the high-pass cut-off. This is illustrated in Figure 11.9.

With r dominant poles, the loop gain at the frequency of intersection appoximates:

$$L(s) = \frac{b_q s^q}{\prod_{i=1}^{r}\left(1 - \frac{s}{p_i}\right)}. \tag{11.42}$$

In the vicinity of the frequency of intersection, this can be approximated by its asymptote:

$$L(s) = b_q s^{q-r} \prod_{i=1}^{r} -p_i. \tag{11.43}$$

The servo function is then obtained as

$$S(s) = \frac{-b_q s^{q-r} \prod_{i=1}^{r} -p_i}{1 - (-1)^r b_q s^{q-r} \prod_{i=1}^{r} p_i}. \tag{11.44}$$

This function has the structure of a unity-gain high-pass filter of order $q-r$. If all poles are in MFM positions, it has a -3dB cut-off frequency $\omega_{\min}$ that can be found as:

$$\omega_{\min} = \sqrt[q-r]{\frac{1}{\left| b_q \prod_{i=1}^{r} p_i \right|}}. \tag{11.45}$$

If we substitute (11.41) in this equation, we see that the low-frequency cut-off can be brought to an arbitrarily low value by increasing the DC gain of the controller:

$$\omega_{\min} = \sqrt[q-r]{\frac{1}{\left| A_{DC}\alpha_{DC} \left(\prod_{i=1}^{q} \tau_i \right) \left(\prod_{i=1}^{r} p_i \right) \right|}}. \tag{11.46}$$

Design conclusion

From (11.46), we may conclude:

> A high-pass cut-off frequency $\omega_{\min}$ of the servo function, caused by DC blocking elements in the loop transfer, can be kept below its specified value if the controller provides a sufficiently large DC gain.

The design of the high-pass cut-off frequency, will be elucidated in the example below.

Example 11.4

In this example, we will design the high-pass cut-off frequency of a current integrator. Integration of electrical current is often required in charge detection or charge control amplifiers. In this example, we will only focus on one performance aspect: high-pass cut-off due to a lack of loop gain at low frequencies.

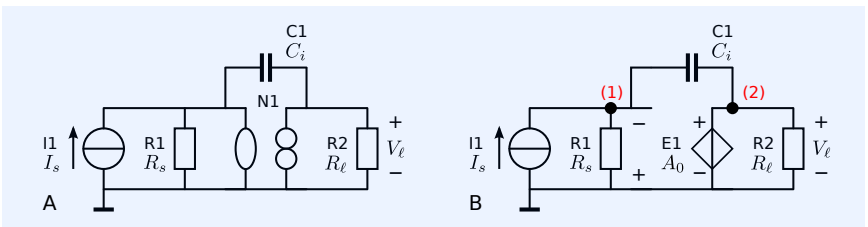


Figure 11.10: Charge amplifier:
A: The amplifier concept for a grounded resistive current source and a grounded resistive load
B: The amplifier in which the ideal controller has been replaced with a voltage-controlled voltage source with DC gain A_0.

High-pass behavior results in droop or tilt in the step response (see section 2.4.6). Such droop or tilt may pose requirements to the high-pass cut-off frequency of charge amplifiers.

Figure 11.10A shows the concept of the charge amplifier under study. It converts the charge of a grounded source into a voltage for a grounded load. The

source is resistive and has a source resistance R_s. The load is also resistive and has a resistance R_ℓ. Figure 11.10B shows the circuit in which the ideal controller (a nullor) has been replaced with a voltage-controlled voltage source with a DC gain A_0. Since we are only studying the high-pass cut-off, low-pass behavior of the controller has not been modeled.

We will assume the following requirements for the charge amplifier:

1. *The peak value of the charge delivered by the source is $10pC$. This peak charge should evoke a peak load voltage of $2V$.*

   From this, we can calculate the value of the integration capacitance: $C_i = 5pF$.

2. *The source resistance is $50k\Omega$ and the load resistance is $2k\Omega$.*

3. *The high-pass cut-off frequency f_ℓ of this circuit should be below $1kHz$.*

Below the SLiCAP *netlist of the circuit from Figure 11.10B.*

```
simpleQamp
* file: simpleQamp.cir
* SLiCAP circuit file
I1 0 1 {I_s}
R1 1 0 {R_s}
C1 1 2 {C_i}
R2 2 0 {R_ell}
E1 2 0 0 1 {A_0}
.param C_i=5p R_s=50k R_ell=2k
.end
```

The design task that we will try to fulfill in this example is: find a minimum value for the DC gain of the controller to meet the requirement for f_ℓ.

A simple and straightforward way to do this is to calculate the servo function. This function describes the behavior of the transfer with respect to the ideal behavior. The script below (lines 1 to 20) prints the expression for the servo function with A_0 as variable to the output html page.

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# file: simpleQamp.py

from SLiCAP import *

fileName = 'simpleQamp'
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName + '.cir')
htmlPage('High-pass cut-off design')
i1.setSource('I1')
i1.setDetector('V_2')
i1.setLGref('E1')
i1.setSimType('numeric')
i1.setGainType('servo')
i1.setDataType('laplace')
result = i1.execute()
exprGain = result.laplace
eqn2html('S', exprGain);
```

The next part of the script (lines 22-30), calculates the minimum requirement for A_0. This script only works if there are no low-pass effects modeled! It calculates the high-pass cut-off frequency ω_{min} from the first coefficient a_0 and the last coefficient a_q of the denominator polynomial of the servo function:

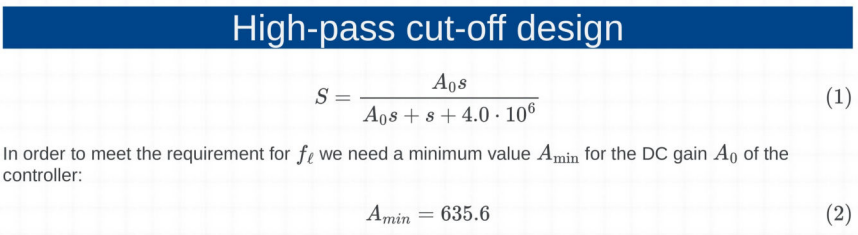
$$\omega_{\text{min}}^q = \frac{a_0}{a_q}. \tag{11.47}$$

It then solves

$$\omega_{\text{min}}^q - (2\pi f_\ell)^q = 0. \tag{11.48}$$

The output of this script is shown in Figure 11.11.

Figure 11.11: Result of the execution of the SLiCAP script.




```
21 gain, numerCoeffs, denomCoeffs = coeffsTransfer(exprGain)
22 order          = len(denomCoeffs) - 1
23 omegaLow       = denomCoeffs[0]/denomCoeffs[-1]
24 f_l            = 1e3
25 A0min          = sp.solve(omegaLow - (2*sp.pi*f_l)**(1/order))[0]
26 A_min = sp.Symbol('A_min')
27 text2html('In order to meet the requirement for $f_{\\ell}$ we need a minimum
       value $A_{\\min}$ for the DC gain $A_0$ of the controller:');
28 eqn2html(A_min, A0min);
29 i1.defPar('A_0', A0min);
```

The Bode plots are shown in Figure 11.12. They are generated in lines 32-46.

```
30 L = i1.execute()
31 i1.setGainType('asymptotic')
32 A = i1.execute()
33 i1.setGainType('direct')
34 D = i1.execute()
35 i1.setGainType('gain')
36 G = i1.execute()
37 figMag = plotSweep('magQamp', 'Magnitude characteristics', [A, L, S, D, G],
       10, 10e6, 200, funcType='mag', show=True)
38 fig2html(figMag, 800)
39 figPhase = plotSweep('phaseQamp', 'Magnitude characteristics', [A, L, S, D, G
       ], 10, 10e6, 200, funcType='phase', show=True)
40 fig2html(figPhase, 800);
```

The plots of the asymptotic gain clearly show the integrating character of the charge amplifier. The plots of the loop gain show a lack of loop gain at low frequencies. This is due to the high-pass filter in the loop formed by the integration capacitor and the source resistance. The gain of the integrator closely matches the asymptotic gain at frequencies above 1kHz. The loop gain reaches its maximum of 635.6 above the frequency of the pole caused by the integration capacitor and the source resistance.

The analysis performed in the example above, requires all low-pass effects not to have been modeled. This can be done by using different circuit models during the design of the high-pass transfer and of the low-pass transfer. In amplifier design, such an approach is usually possible because the frequency regions of high-pass and low-pass transfer are well separated. Only when designing narrow-band active filters would this way of working not be possible.

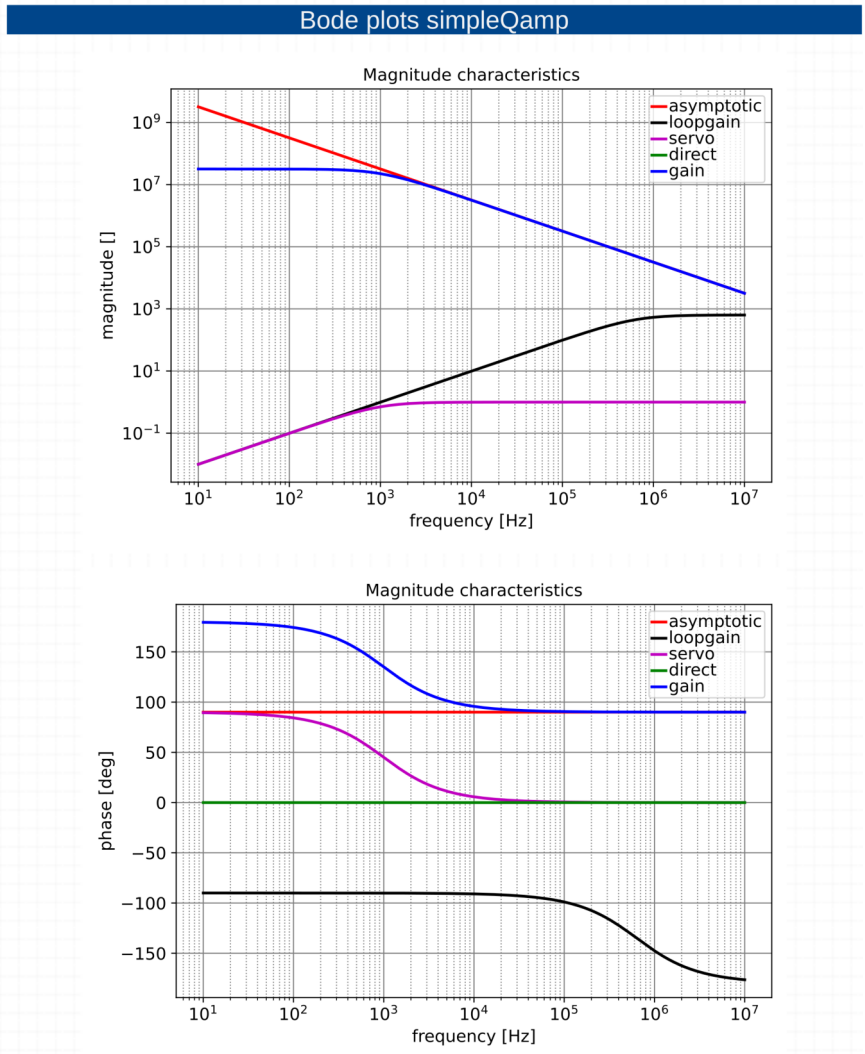


Figure 11.12: Result of the execution of the SLiCAP script.

11.4.5 Low-pass cut-off with poles and zeros

In section 11.4.3, we have studied the low-pass cut-off frequency of the servo function for all-pole loop gain transfers with a finite DC loop gain. In this, we will derive a general expression for the low-pass cut-off frequency ω_h of the servo function. To this end, we assume a loop gain with k poles at $s = 0$ and ℓ zeros at $s = 0$, and a total number of poles and zeros of n and m, respectively. The loop gain function can then be written as

$$L(s) = \frac{b_\ell s^\ell}{a_k s^k} \frac{\prod_{j=\ell+1}^{m} \left(1 - \frac{s}{z_j}\right)}{\prod_{i=k+1}^{n} \left(1 - \frac{s}{p_i}\right)}, \tag{11.49}$$

in which b_ℓ and a_k are the trailing coefficients of s of the numerator polynomial and of the denominator polynomial of the loop gain, respectively.[6]

We obtain a low-pass character of the loop gain if the number of poles exceeds the number of zeros, thus if $n > m$ in expression 11.49. If so, the magnitude characteristic has a falling slope at the highest frequencies and

[6] The coefficients of the lowest order of the Laplace variable s.

the loop gain may be approximated by its asymptote as:

$$L = \frac{C_{n-m}}{\omega^{n-m}}, \tag{11.50}$$

where the coefficient C_{n-m} can be written as

$$C_{n-m} = \left| \frac{b_\ell}{a_k} \frac{\prod_{i=k+1}^{n} p_i}{\prod_{j=\ell+1}^{m} z_j} \right|. \tag{11.51}$$

The low-pass cut-off of the servo function occurs at approximately unity loop gain. Hence, with the aid of (11.50), we obtain:

$$\omega^{n-m} \approx C_{n-m}. \tag{11.52}$$

With the aid of (11.51), we obtain the value of the low-pass cut-off frequency ω_h for a feedback circuit of which the loop gain has m dominant zeros and n dominant poles, with $n > m$. Dominant poles and zeros are those that have a frequency smaller than ω_h, as discussed in section 11.4.3

$$\omega_h \approx \sqrt[n-m]{\left| \frac{b_\ell}{a_k} \frac{\prod_{i=k+1}^{n} p_i}{\prod_{j=\ell+1}^{m} z_j} \right|}. \tag{11.53}$$

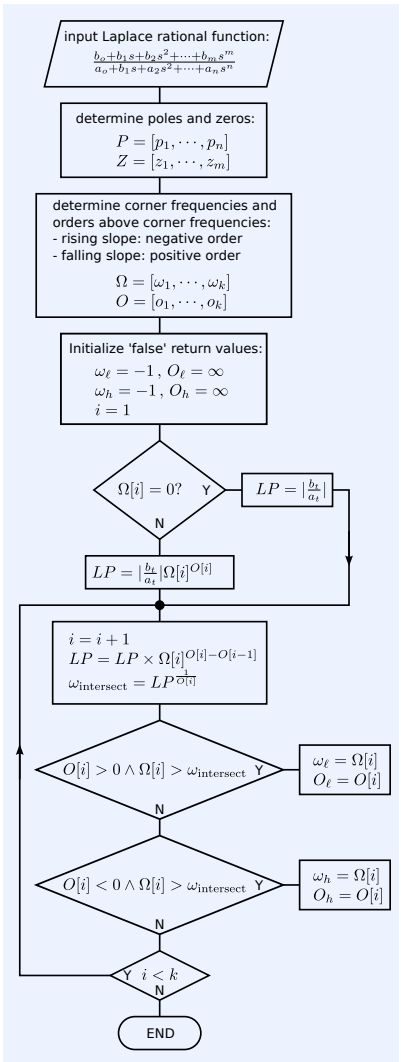


Figure 11.13: Simplified flow chart for the determination of the servo bandwidth from the asymptotic magnitude characteristic of the loop gain.

11.4.6 High-pass cut-off with poles and zeros

At low frequencies, zeros may cause the loop gain to drop below its midband value. In such cases the servo function obtains a high-pass character with a high-pass cut-off at ω_ℓ. This cut-off frequency can be found in a similar way as the low-pass cut-off frequency ω_h. We now only account for the p zeros and the q poles with frequencies smaller than ω_ℓ and use the asymptotic approximation according to (11.50) with $p > q$. In this way we obtain:

$$\omega_\ell \approx \sqrt[p-q]{\left| \frac{b_\ell}{a_k} \frac{\prod_{i=k+1}^{p} p_i}{\prod_{j=\ell+1}^{q} z_j} \right|}. \tag{11.54}$$

11.4.7 Procedure for determination of the servo bandwidth

With the results of the theory from sections 11.4.5 and 11.4.6, we are able to design a procedure for the determination of the bandwidth of the servo function with an arbitrary number of poles and zeros. The outline of the method is:

1. From the rational expression of the loop gain, determine the ratio of the trailing coefficients of the numerator polynomial and the denominator polynomial: $\frac{b_\ell}{a_k}$.

2. Determine the m zeros (the solutions of the numerator polynomial) and the n poles (the solutions of the denominator polynomial).

3. Find the corner frequencies and the order of the loop gain function for frequencies above each corner. This can be done as follows:

   (a) Calculate the $m + n$ corner frequencies; these are the absolute values of the poles and the zeros.

   (b) Assign order $= +1$ to a pole frequency and order $= -1$ to a zero frequency

   (c) Rank the frequencies and their corresponding orders in ascending order of frequencies.

   (d) Make a new list with unique frequencies ω_i and corresponding orders o_i. These orders are the sum of the orders of the original corner frequencies that had the same value. Hence, two poles with the same frequency result in one corner frequency with order $+2$. A pole-zero pair results in one corner frequency with order 0, etc.

   (e) Replace the order of the corner frequencies with the order of the loop gain function. The order of the loop gain function above a corner frequency is the sum of the orders of all corner frequencies up to that corner frequency. Hence, if we first have a pole, then another pole and then a zero, we have the orders: $1, 2, 1$.

4. Determine the coefficients C_i (see (11.51)) of the asymptotes for each corner frequency. This is done as follows:

   (a) At the first corner frequency, we have $C_1 = \frac{b_\ell}{a_k}$. If the first corner frequency is not zero, we have to multiply this by the first corner frequency ω_1.

   (b) At following corner frequencies ω_i, we have $C_i = C_{i-1} \times \omega_i^{(o_i - o_{i-1})}$.

5. Determine the cut-off frequencies ω_ℓ and ω_h.

   To this end, we evaluate the intersection points Ω of the asymptotes of the loop gain function and unity for each corner frequency, starting at the lowest frequency and proceeding with increasing frequencies. These intersection points are obtained as

$$\Omega_i = C_i^{\frac{1}{o_i}}, \tag{11.55}$$

   in which o_i is the order at the corner frequency ω_i. This intersection point can be a valid cut-off frequency if its value is larger than that of the corresponding corner frequency. It is assigned to the high-pass cut-off ω_ℓ if the order at the corresponding corner frequency is negative and to the low-pass cut-off frequency ω_h if it is positive:

$$\omega_\ell = \Omega_i \text{ if } o_i > 0 \wedge \Omega_i > \omega_i; \tag{11.56}$$

$$\omega_h = \Omega_i \text{ if } o_i < 0 \wedge \Omega_i > \omega_i. \tag{11.57}$$

   Each time a new valid cut-off frequency is found, the old value will be overwritten, resulting in correct values of ω_ℓ and ω_h of the servo function.

   This function has been implemented in SLiCAP.
   The function `findServoBandwidth` returns:

- The high-pass cut-off frequency ω_ℓ and the order at this cut-off frequency
- The low-pass cut-off frequency ω_h of the servo function and the order at this frequency
- The maximum *asymptotic* value of the loop gain and the first frequency of occurrence.

11.4.8 Design conclusions

In the preceding sections, we have studied the relationship between the achievable bandwidth of the feedback amplifier and the gain and pole-zero pattern of the loop gain function. We have seen that the gain and the pole-zero pattern of the loop gain function is governed by the dynamic behavior of the source, the load, the feedback network and the controller. At this stage, the design of the controller has our interest, and we found that the controller must contribute to the loop gain-poles product such that it provides a sufficiently high low-pass cut-off. If the loop gain drops below unity at low frequencies, the DC loop gain can be used to achieve a sufficiently low high-pass cut-off.

In many cases, we are interested in a smooth step response and a smooth magnitude and phase response of the amplifier. Such responses can be established with all-pole transfers. Since zeros in the loop gain transfer appear in the servo function, zeros should be avoided in the frequency range of interest. Hence, in many situations, the controller should contribute an all-pole characteristic to the loop gain. Zeros in the loop gain arise from inevitable influences of the source, the load and the feedback networks.

11.5 Stability of negative feedback amplifiers

In the previous sections, we studied the design of the low-pass and high-pass cut-off frequencies of the servo function. We have not yet investigated whether the feedback amplifier is stable. A system is called stable if all system poles have a negative real part. If there are single poles with a real part equal to zero, the system is called marginally stable. Since it is our intention is to build stable amplifiers (not oscillators), the ideal transfer of an amplifier will have all its poles in the left half plane. With the loop gain reference selected properly, the stability with respect to the ideal gain is thus governed by the servo function. Hence, in order to investigate the stability of a feedback amplifier, we need to investigate whether the servo function has all its poles in the left half plane.

There are several ways to do this, and they will be discussed briefly in the following sections. For further study on these topics, the reader is referred to control theory, for which references will be given.

1. Routh-Hurwitz criterion

   The quote[7] taken from Truxal [Truxal1972][8] introduces the Routh test, which will be presented in section 11.5.1. The Routh test states that the number of solutions of a polynomial with a positive real part equals the number of sign changes in the first column of the Routh array. A brief and clear description of the method is also given by Dorf [Dorf1995][9].

2. Nyquist criterion

   The Nyquist [Nyquist1932][10] criterion states that the number of poles in the right half plane of a feedback system is equal to the number of clockwise encirclements around the point -1 in the polar plot of the loop gain. [11] This criterion can also be used for systems that incorporate delay lines. The Nyquist theorem will be discussed in section 11.5.2. For a detailed treatment of this criterion, the reader is referred to literature on control theory.[12]

3. Root locus analysis

   The root locus technique was introduced by Evans [Evans1948][13]. It is a graphical method for finding the poles of the servo functions based on the poles and the zeros of the loop transfer function. It will briefly be discussed in section 11.5.3.

[7] In the late 1870s Maxwell remarked (in a published paper) that determining stability from the coefficients of the characteristic polynomial (without factoring) was, unfortunately, apparently an insoluble problem. After reading this, Routh worked diligently for three years to develop the Routh test. He presented it with the opening remarks, "It has recently come to my attention that my good friend James Clerck Maxwell has difficulty with a rather trivial problem. ..."

[8] John G. Truxal. *Introductory System Engineering*. McGraw-Hill, Tokyo, 1972

[9] Richard C. Dorf and Robert H. Bishop. *Modern Control Systems*. Addison-Wesley Publishing Company, Inc., USA, 1995. ISBN: 0-201-84559-8

[10] H. Nyquist. Regeneration theory. *Bell System Technical Journal*, 11:126–147, 1932

[11] Assuming that the loop gain itself has no poles in the right half plane.

[12] The encirclement of the point -1 is based upon the application of Black's feedback model. With the asympotic-gain model, it is +1 in the polar plot of the loop gain.

[13] Walter R." "Evans. Graphical Analysis of Control Systems. *Trans. AIEE*, 67(1):547–551, January 1948

11.5.1 Routh–Hurwitz criterion

The Routh–Hurwitz stability criterion is a mathematical test that is a necessary and sufficient condition for the stability of a linear time-invariant system. The Routh test was proposed in 1876 by the English mathematician Edward John Routh. The German mathematician Adolf Hurwitz independently proposed an equivalent procedure in 1895, but the Routh test is more efficient.

The Routh test takes the coefficients of the characteristic polynomial of a system as input to determine the Routh array. The number of sign changes in the first column of the Routh array is equal to the number of solutions of the characteristic polynomial that have a positive real part. The procedure is extensively discussed in the references to control theory. In this section, we will only briefly describe the procedure.

The Routh test has been implemented in SLiCAP.

The procedure

The Routh array is determined as follows:

1. The polynomial $D(s)$ under test is

$$a_n s^n + a_{n-1} s^{n-1} + a_{n-2} s^{n-2} + \cdots + a_1 s + a_0. \tag{11.58}$$

2. The coefficients of this polynomial are arranged in the first two rows of the Routh array. The number of rows of this array equals $n+1$, and the number of columns is m, with $m = \frac{n}{2}$ if n is odd and $m = \frac{n+1}{2}$ if n is even, n being the order of the polynomial

$$\begin{array}{c} s^n \\ s^{n-1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ s^0 \end{array} \begin{bmatrix} a_n & a_{n-2} & a_{n-4} & \cdot & \cdot \\ a_{n-1} & a_{n-3} & a_{n-5} & \cdot & \cdot \\ c_{3,1} & c_{2,2} & \cdot & \cdot & c_{3,m} \\ c_{4,1} & c_{5,1} & \cdot & \cdot & c_{4,m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{n,1} & c_{n,1} & \cdot & \cdot & c_{n,m} \end{bmatrix}. \tag{11.59}$$

The remaining coefficients $c_{i,j}$ are obtained as follows:

$$c_{i,j} = -\frac{1}{c_{i-1,j}} \det \begin{bmatrix} c_{i-2,j} & c_{i-2,j+1} \\ c_{i-1,j} & c_{i-1,j+1} \end{bmatrix}. \tag{11.60}$$

3. The number of solutions of the polynomial that have a positive real part equals the number of sign changes in the first column of the Routh array. There are two special cases; see Dorf [Dorf1995][14] for details:

   (a) If the entry in the first column of a row becomes 0 while other entries are nonzero, the first entry will be replaced with ε. This allows the array to be completed. After completion of the array, we let ε approach zero to check for stability

   (b) If a complete row of zeros is detected, the *auxiliary polynomial* of the above row is differentiated and its coefficients are used as entries. The auxiliary polynomial P_i of the $i-th$ row is

$$P_i = \sum_{j=1}^{m} R_{i,j} s^{n+3-i-2j}. \tag{11.61}$$

   A row of zeros indicates instability or marginal instability.

[14] Richard C. Dorf and Robert H. Bishop. *Modern Control Systems*. Addison-Wesley Publishing Company, Inc., USA, 1995. ISBN: 0-201-84559-8

In the following example, we will investigate the stability of the transimpedance amplifier from example 11.3 with the aid of the Routh array.

Example 11.5

```
79 htmlPage('Routh array')
80 i1.setGainType('gain')
81 i1.setDataType('denom')
82 denomLaplace = i1.execute().denom
83 text2html('The characteristic equation of the gain is:')
84 eqn2html('charPoly', denomLaplace)
85 text2html('The Routh array of this poly is:')
86 eqn2html('RA', routh(denomLaplace))
```

The html page with the Routh array RA is shown in Figure 11.14. It shows that the transimpedance amplifier is stable: all numbers in the first column of the array are positive.

Figure 11.14: Result of the execution of the SLiCAP-MATLAB script.

Routh array

The characteristic equation of the gain is:

$$charPoly = -1.687 \cdot 10^{-13} s^2 - 1.022 \cdot 10^{-7} s - 10 \tag{1}$$

The Routh array of this poly is:

$$RA = \begin{bmatrix} -1.687 \cdot 10^{-13} & -10 \\ -1.022 \cdot 10^{-7} & 0 \\ -10 & 0 \end{bmatrix} \tag{2}$$

Design application of the Routh array

Because it can be evaluated symbolically, the Routh array can be used as a design tool.

11.5.2 Nyquist criterion

The Routh array provides a clear method for the determination of the stability of a feedback system, but it requires the transfer of a system to be written as a rational function of the Laplace variable s. This is not possible for systems that include delay elements. The Laplace transfer of a delay τ is $\exp(-s\tau)$, which cannot be written as a rational function. The Nyquist stability criterion (see: [Nyquist1932][15]) is applicable for all types of loop gain functions. The Nyquist stability criterion states that

$$n = p - q, \tag{11.62}$$

if the loop gain $-L(s)$ is plotted as a contour in the complex L plane: $\mathrm{Re}\{-L(s)\}$ versus $\mathrm{Im}\{-L(s)\}$, with $s = j\omega$ varying along the contour from 0 to $j\infty$ and from $-j\infty$ to 0, where:

1. n is the number of clockwise encirclements of the point $(-1, 0)$ in the complex L plane by the contour
2. p is the number of poles of the servo function $S(s) = \frac{-L(s)}{1-L(s)}$ inside the right half of the s plane
3. q is the number of poles of the loop gain $L(s)$ inside the right half of the s plane

[15] H. Nyquist. Regeneration theory. *Bell System Technical Journal*, 11:126–147, 1932

Please notice that we use $-L$ throughout this section about the Nyquist criterion. This is because the Nyquist criterion has been derived with Black's feedback model. That model includes a -1 transfer in the loop that is not included in the asymptotic gain model.

In almost all cases, the loop gain itself will be a stable transfer and the Nyquist stability criterion simplifies to $n = p$. Nyquist stability analysis is a graphical technique. The contour plot is usually only constructed for positive values of $j\omega$. For negative values, the contour can be completed with a complex conjugated curve. Respective examples of so-called *Nyquist plots* for an unstable and a stable system are shown in Figure 11.15.

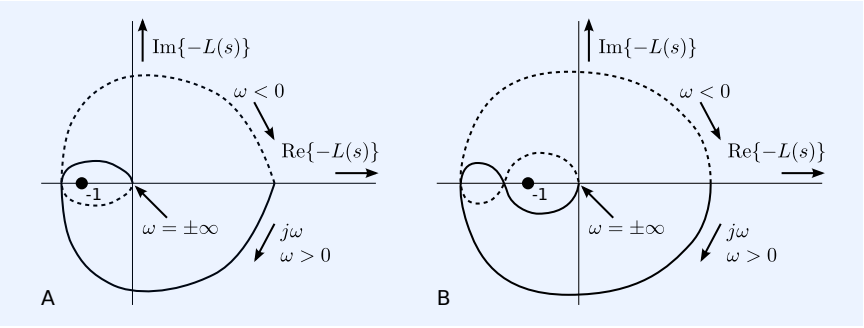


Figure 11.15: Examples of Nyquist plots:
A: Instable feedback system: clockwise encirclement of (-1,0).
B: Stable feedback system: counterclockwise encirclement of (-1,0).

Gain and phase margin

It is common practice to use the *gain margin* and the *phase margin* as measures for the margin in the loop gain function before instability of the servo function occurs. Figure 11.16 illustrates the definition of these terms:

- The gain margin is the magnitude of $1 + L$ if the loop gain $(-L)$ has a phase shift of 180 degrees.
- The phase margin is the difference between 180 degrees and the phase shift of the loop gain at a frequency where the magnitude of the loop gain equals unity.

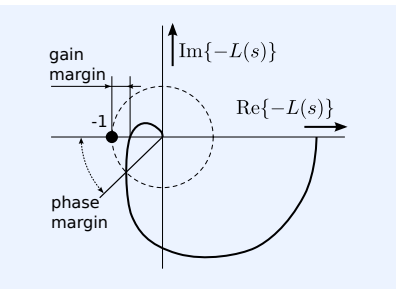


Figure 11.16: Definition of the gain and the phase margin in the Nyquist plot.

However, single-frequency gain and phase margin descriptions do not completely characterize the dynamic behavior of feedback systems. First of all, such a description is incomplete for higher order systems. Secondly, we are interested in the stability of the gain (source-to-load transfer) and the loop gain only has a well defined relationship with the gain if the asymptotic gain equals the ideal gain.[16]

In the following example, we will demonstrate the use of SLiCAP for making Nyquist plots.

Example 11.6

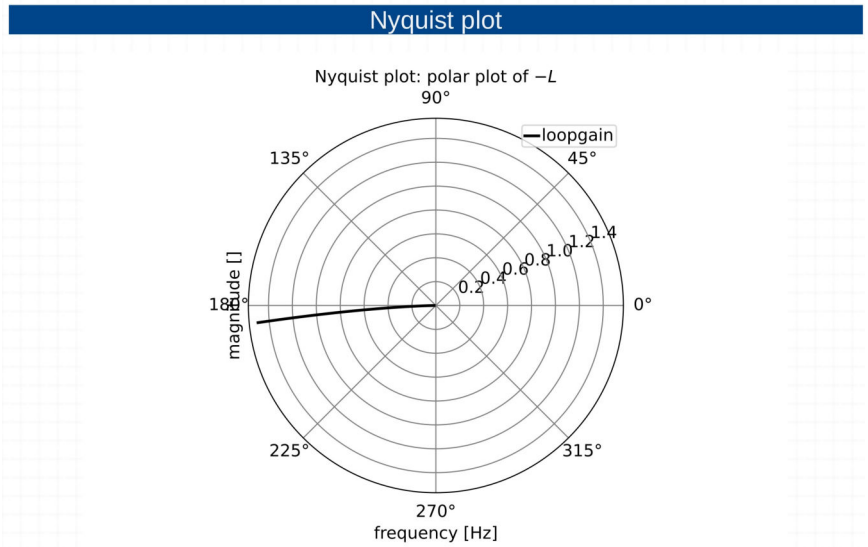
Lines 113 to 119 of the `transimpedance.m` *script generate the Nyquist plot of the transimpedance amplifier and display it on the html page. This page is shown in Figure 11.17. The Nyquist plot in* SLiCAP *only works with data type* `LAPLACE` *and gain type* `LOOPGAIN`*. It shows a polar plot of* $-L$*.*

```
88 htmlPage('Nyquist plot')
89 i1.setGainType('loopgain')
90 i1.setDataType('laplace')
91 result = i1.execute()
92 result.laplace = -result.laplace
93 figNyquist = plotSweep('Nyquist', 'Nyquist plot: polar plot of $-L$', result,
       1e6, 10e6, 100, axisType='polar', funcType='mag', show=True)
94 fig2html(figNyquist, 800)
```

[16] At this point, it should be clear that a system is stable if and only if all roots of the charactersitic equation have a negative real part, and that the filter characteristic is defined by the complete pole-zero pattern of the source-to-load transfer.

Figure 11.17 clearly shows the very small phase margin of the uncompensated transimpedance amplifier. This corresponds with the peaking in the magnitude characteristic of the servo function, as shown in Figure 11.8.

Figure 11.17: SLiCAP simulation results.



11.5.3 Root locus analysis

The root locus method is a graphical method for determining the poles of the servo function. It was introduced by Evans in 1948 [Evans1948][17] [Evans1950][18] and is also often used in control engineering practice. The root locus shows how the poles of a system move around in the complex plane as a parameter is varied. In negative feedback systems, the DC loop gain is used as a variable for the root locus. For electronic amplifiers, varying the DC loop gain might be interesting for analytical purposes, but we will see that it is not the best method for maneuvering poles into their desired positions. However, the root locus technique is very useful for finding the poles of the servo function from the poles and zeros of the loop gain and the DC value of the loop gain, and it can be helpful in understanding frequency compensation, which will be discussed in Chapter 12.

In this section, we will briefly discuss the rules for drawing the root locus plots. For complete proofs of the rules, the reader is referred to literature on control theory.

A general expression of a loop transfer function with poles and zero has been given in (11.49). The general expression for the servo function $S(s)$ can thus be obtained as

$$S(s) = \frac{-b_\ell s^\ell \prod_{j=\ell+1}^{m}\left(1-\frac{s}{z_j}\right)}{a_k s^k \prod_{i=k+1}^{n}\left(1-\frac{s}{p_i}\right) - b_\ell s^\ell \prod_{j=\ell+1}^{m}\left(1-\frac{s}{z_j}\right)}. \tag{11.63}$$

The poles of the servo function are the solutions of the denominator:

$$a_k s^k \prod_{i=k+1}^{n}\left(1-\frac{s}{p_i}\right) - b_\ell s^\ell \prod_{j=\ell+1}^{m}\left(1-\frac{s}{z_j}\right), \tag{11.64}$$

[17] Walter R." "Evans. Graphical Analysis of Control Systems. *Trans. AIEE*, 67(1):547–551, January 1948

[18] Walter R." "Evans. Control Systems Synthesis by Root Locus Method. *Trans. AIEE*, 69(1):66–69, January 1950

which can be written as

$$s^k \prod_{i=k+1}^{n} \left(1 - \frac{s}{p_i}\right) - \frac{b_\ell}{a_k} s^\ell \prod_{j=\ell+1}^{m} \left(1 - \frac{s}{z_j}\right). \tag{11.65}$$

If $\ell = 0$ and $k = 0$, the gain factor $\frac{b_\ell}{a_k}$ equals the DC loop gain L_{DC}. From (11.65), we can easily see that the poles of the servo function depend on:

1. The n poles p_i of the loop gain
2. The m zeros z_j of the loop gain
3. The gain factor $\frac{b_\ell}{a_k}$, or the DC loop gain L_{DC} if both ℓ and k are zero.

The root locus is the path in the complex plane as it is traced by the poles when the gain factor $\frac{b_\ell}{a_k}$ is varied. The construction rules for the root loci are given without proof:

1. The number of branches of the root locus is equal to the number of poles.
2. Poles are either real or complex conjugates. As a consequence, the root locus is symmetrical with respect to the real axis.
3. The root locus starts ($\frac{b_\ell}{a_k} = 0$) in the poles of $L(s)$.
4. The root locus ends ($\frac{b_\ell}{a_k} = \infty$) on the zeros of $L(s)$; if the number of zeros m is smaller than the number of poles n, $n - m$ zeros are supposed to be at infinity.
5. Parts of the real axis left to an odd number of poles and zeros are part of the root locus.
6. If there are n poles and m zeros, $n - m$ branches go to infinity, the angle with the positive real axis of the asymptotes of these branches is given by

$$\theta_i = \frac{2i+1}{n-m}\pi, \qquad i = 0, 1, 2, \ldots \tag{11.66}$$

7. These asymptotes intersect the real axis ($\sigma - axis$) at: $\sigma = \frac{\sum_{k=1}^{n} p_k - \sum_{i=1}^{m} z_i}{n-m}$.
8. If a branch of the real axis lies between two poles or two zeros, the point where the poles leave or arrive at the real axis are called the break away points. The value of s at these break away points follows from

$$\frac{d}{ds} L(s) = 0. \tag{11.67}$$

9. The tangents to the loci at the break away points are equally spaced over 2π. Two poles break away at $\pm\frac{\pi}{2}$, three poles at π and $\pm\frac{\pi}{3}$, four poles break away at $\pm\frac{\pi}{4}$, and $\pm\frac{3\pi}{4}$, etc.
10. Each point of the root locus satisfies the condition:
    $|L(s)| = 1$ and $\arg(L(s)) = k2\pi$; $k = 0, \pm1, \pm2, \ldots$

Below, we will show how to generate root-locus plots with SLiCAP.

The circuit that we will use for drawing the root locus plots is shown in Figure 11.18. It is a unity-gain voltage follower with a voltage-controlled voltage source as the controller. This controlled source will be selected as the loop gain reference, hence the loop gain equals the gain of the controller. With each root locus plot, we will plot the poles and the zeros of the loop

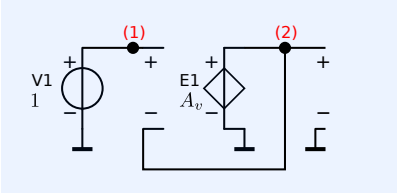


Figure 11.18: Circuit for drawing root locus plots with SLiCAP.

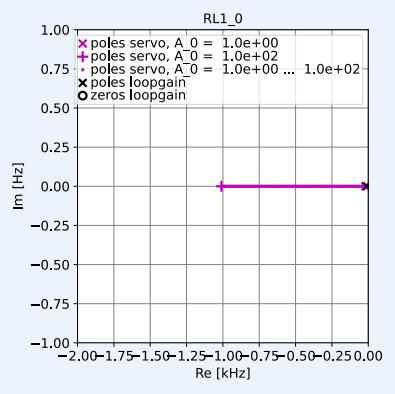


Figure 11.19: SLiCAP root-locus plot.

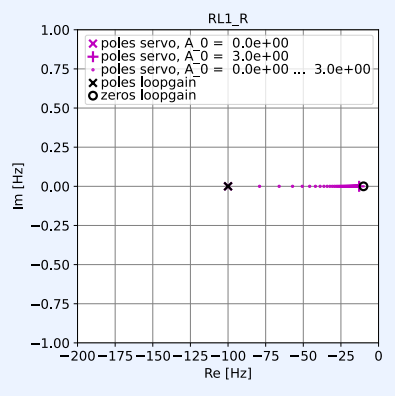


Figure 11.20: SLiCAP root-locus plot.

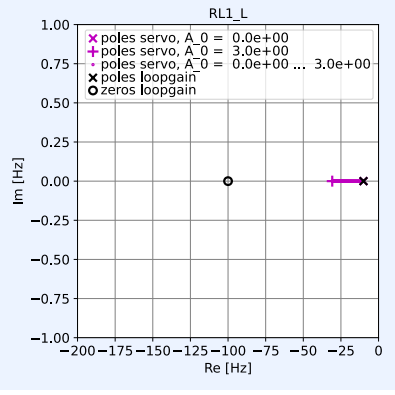


Figure 11.21: SLiCAP root-locus plot.

gain and the poles of the servo function while stepping the DC loop gain or, if there are any poles or zeros in the origin, the gain factor $\frac{b_\ell}{a_k}$.

The SLiCAP netlist of the circuit from Figure 11.18 with a single-pole transfer is listed below. The transfer of the voltage-controlled voltage source `E1` will be changed for the other root locus plots depicted in Figure 11.19 to 11.26

The script `RL1_0.py` plots the root locus of the single-pole circuit. Similar scripts have been used to generate the root locus plots from Figure 11.20 to 11.26. Lines 1-19 of the scripts check the circuit and define the source, the load and the loop gain reference.

```
1  RL1_0
2  * file: RL1_0.cir
3  V1 1 0 0
4  E1 2 0 1 2 {A_0/(1-s/2/pi/p_1)}
5  .param A_0=100 p_1=-10
6  .end
```

```
1  #!/usr/bin/env python3
2  # -*- coding: utf-8 -*-
3  """
4  Created on Thu Mar 18 12:11:37 2021
5
6  @author: anton
7  """
8  from SLiCAP import *
9
10 fileName = 'RL1_0'
11 #prj = initProject(fileName)
12 i1 = instruction()
13 i1.setCircuit(fileName + '.cir')
14 htmlPage('Root locus plot: ', fileName)
15 i1.setSource('V1')
16 i1.setDetector('V_2')
17 i1.setLGref('E1')
18 i1.setSimType('numeric')
19 i1.setGainType('loopgain')
```

Lines 20 and 21 evaluate the poles and the zeros of the loop gain.

```
20 i1.setDataType('pz')
21 pzL = i1.execute()
```

Lines 22 to 30 evaluate the poles of the servo function.

```
22 i1.setGainType('servo')
23 i1.setDataType('poles')
24 i1.stepOn()
25 i1.setStepVar('A_0')
26 i1.setStepStart(1)
27 i1.setStepStop(100)
28 i1.setStepMethod('lin')
29 i1.setStepNum(100)
30 plsS = i1.execute()
```

Lines 31 and 32 generate the root locus plot and place it on the html page.

```
31 figPZ = plotPZ(fileName, fileName, [plsS, pzL], xscale='k', xmin=-2, xmax=0,
       ymin=-1, ymax=1, show=True)
32 fig2html(figPZ, 500)
```

The root-locus plot generated by this script is shown in Figure 11.19. The loop gain of the circuit is:

$$L(s) = \frac{A_0}{1 - \frac{s}{2\pi p_1}}, \tag{11.68}$$

with $p_1 = -10$ Hz. The root-locus has one branch (rule 1), it starts at p_1, (rule 3) and it ends at $-\infty$ (rule 4). The part of the real axis, left from p_1, belongs

to the root locus (rule 5), and there is one asymptote with an angle of 180 degrees (rule 6). The root-locus in Figure 11.19 is drawn for $A_0 = 0 \cdots 100$.

Figure 11.20 shows the root locus if a zero is located right from the pole: $z_1 = -10\text{Hz}$, $p_1 = -100\text{Hz}$. The root locus starts on the pole and ends on the zero, while the part of the real axis between the pole and the zero belongs to the root locus (rule 5). The root-locus in Figure 11.20 is drawn for $A_0 = 0 \cdots 3$.

Figure 11.21 shows the root locus if a zero is located left from the pole: $p_1 = -10\text{Hz}$, $z_1 = -100\text{Hz}$.

The root locus starts on the pole and ends on the zero, while the part of the real axis between the pole and the zero belongs to the root locus (rule 5). The root-locus in Figure 11.21 is drawn for $A_0 = 0 \cdots 100$.

Figure 11.22 shows the root locus for a loop gain with two real negative poles: $p_1 = -10\text{Hz}$, $p_2 = -40\text{Hz}$. The poles of the servo function move along two asymptotes with an angle of $\frac{\pi}{2}$ and $\frac{3\pi}{2}$ towards infinity(rule 6). The asymptotes intersect with the real axis at -25Hz (rule 7) and the poles leave the real axis at -25Hz (rule 8). The expression for the loop gain $L(s)$ with two poles is

$$L(s) = \frac{A_0}{\left(1 - \frac{s}{2\pi p_1}\right)\left(1 - \frac{s}{2\pi p_2}\right)}. \tag{11.69}$$

The derivative $\frac{d}{ds}L(s)$ is obtained as

$$\frac{d}{ds}L(s) = \frac{-A_0\left(2s - 2\pi p_1 - 2\pi p_2\right)}{4\pi^2 p_1 p_2 \left(1 - \frac{s}{2\pi p_1}\right)^2 \left(1 - \frac{s}{2\pi p_2}\right)^2}. \tag{11.70}$$

The solution of $\frac{d}{ds}L(s) = 0$ is

$$\frac{s}{2\pi} = \frac{p_1 + p_2}{2} = -25\text{Hz}. \tag{11.71}$$

Figure 11.22 shows the root-locus from $A_0 = 0 \cdots 100$.

Figure 11.23 shows the root locus for a loop gain with two real negative poles: $p_1 = -10\text{Hz}$, $p_2 = -40\text{Hz}$ and one zero, left from the poles: $z_1 = -100$. One of the poles moves to the zero and the other to infinity along an asymptote that has an angle of π with respect to the positive real axis. The part of the real axis between the poles and the part of the real axis left from the zero belongs to the root locus. The frequency at which the poles leave and arrive at the real axis can be found according to rule 8. The loop gain $L(s)$ with two poles and one zero can be written as

$$L(s) = \frac{A_0\left(1 - \frac{s}{2\pi z_1}\right)}{\left(1 - \frac{s}{2\pi p_1}\right)\left(1 - \frac{s}{2\pi p_2}\right)}. \tag{11.72}$$

The derivative $\frac{d}{ds}L(s)$ is obtained as

$$\frac{d}{ds}L(s) = \frac{2\pi A_0 p_1 p_2\left(s^2 - 4\pi z_1 s - 4\pi^2 p_1 p_2 + 4\pi^2 p_1 z_1 + 4\pi^2 p_2 z_1\right)}{z_1\left(s - 2\pi p_1\right)^2\left(s - 2\pi p_2\right)^2}. \tag{11.73}$$

The arrival points are found as the solutions of

$$\left(\frac{s}{2\pi}\right)^2 - \frac{z_1}{\pi}s - p_1 p_2 + p_1 z_1 + p_2 z_1 = 0, \tag{11.74}$$

from which we obtain

$$\frac{s_{1,2}}{2\pi} = z_1 \pm \sqrt{\left(p_2 - z_1\right)\left(p_1 - z_1\right)}, \tag{11.75}$$

$$= -100 \pm 10\sqrt{55}\ \text{Hz}. \tag{11.76}$$

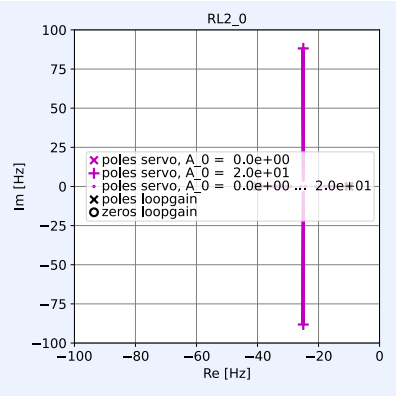


Figure 11.22: SLiCAP root-locus plot.

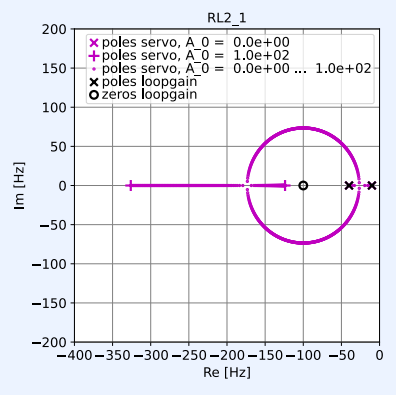


Figure 11.23: SLiCAP root-locus plot.

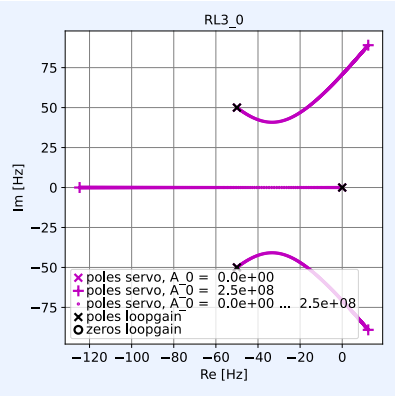


Figure 11.24: SLiCAP root-locus plot.

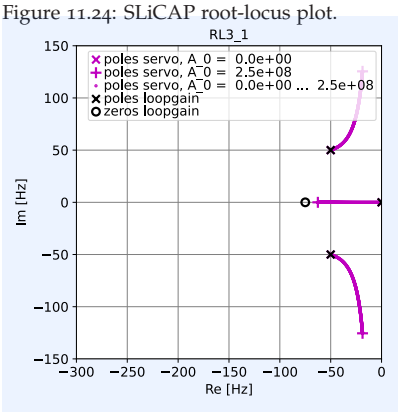


Figure 11.25: SLiCAP root-locus plot.

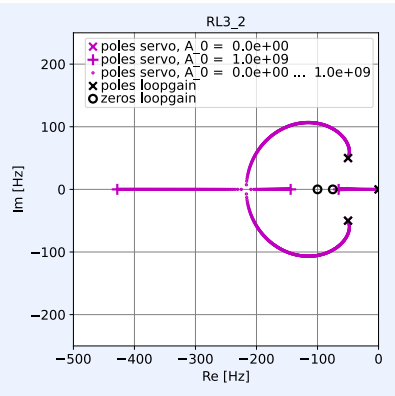


Figure 11.26: SLiCAP root-locus plot.

[19] The characteristic equation of a network is the determinant of the MNA matrix.
[20] Independent capacitor voltages and independent inductor currents are the state variables of electrical networks.

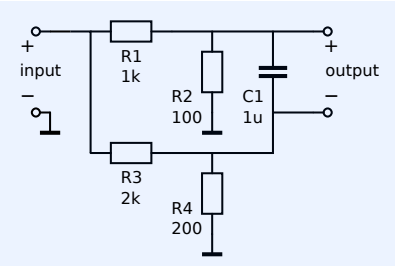


Figure 11.27: Passive two-port network with a non controllable state.

Hence, the poles leave the axis at -25.84Hz and arrive at -174.2Hz. It can be shown that when they leave the real axis, the poles move along a circular path around the zero. The center of that circle is z_1 and the radius $\sqrt{(p_2 - z_1)(p_1 - z_1)}$.

Figure 11.24 shows an example of a root locus of a third order system. This loop gain function $L(s)$ of this system has one pole in the origin and two complex conjugated poles:

$$L(s) = \frac{A_0}{s\left(20 \times 10^3 \pi^2 + 200\pi s + s^2\right)}. \tag{11.77}$$

Figure 11.24 shows the root locus with A_0 stepping from zero to 2.5×10^8. According to rule 4, the poles move to infinity along three asymptotes. The angles of the asymptotes with the positive x axis are $\frac{\pi}{3}$, π and $\frac{2\pi}{3}$ (rule 6). The pole in the origin moves along the x axis (rule 5). The asymptotes intersect the real axis at -33.33Hz (rule 7).

Figure 11.25 shows the root locus of this system after one zero at -75Hz has been added to $L(s)$:

$$L(s) = \frac{A_0\left(1 + \frac{s}{150\pi}\right)}{s\left(20 \times 10^3 \pi^2 + 200\pi s + s^2\right)}. \tag{11.78}$$

The number of asymptotes now reduces to two, and the intersection point of the asymptotes and the real axis is found at -12.5Hz. The pole in the origin moves towards the zero.

Figure 11.26 shows the root locus of this system after another real zero at -100Hz has been added to $L(s)$:

$$L(s) = \frac{A_0\left(1 + \frac{s}{150\pi}\right)\left(1 + \frac{s}{200\pi}\right)}{s\left(20 \times 10^3 \pi^2 + 200\pi s + s^2\right)}. \tag{11.79}$$

The number of asymptotes now reduces to one: two poles will end on the zero and another will move towards infinity. The part of the real axis between the two zeros does not belong to the root locus.

11.5.4 Non-observable and non-controllable states

Unstable behavior cannot always be detected though determination of a time domain response. The only way to ensure detection of instability is to solve the characteristic equation of a network.[19] The solutions of this equation are the solutions of the homogeneous differential equation of the network and yield all the poles of the network, including those associated with non-observable or non-controllable state variables.[20]

Determination of instability with the aid of a time domain simulation does not disclose instability if, for the given excitations and responses, the state variables associated with poles in the right half plane are not controllable or not observable.

We will elucidate the terms 'observable' and 'controllable' below.

Non-controllable state variable

> A non-controllable state variable of a network is one that cannot be changed by excitation(s) applied to the input(s) of a network.

Consider, for example, the network from Figure 11.27. It shows a passive network with an input port and an output port. The output voltage of the network is the capacitor voltage, which is the only state variable of the network. Although this voltage can be observed at the output port, it cannot be

controlled by a current or by a voltage applied to the input port. In other words, the transfer from a current or voltage applied to the input port to the voltage across the capacitor is zero. Hence, in this network, the state variable is not controllable.

If we determine the poles and the zeros of the transfer parameters of this network, we would find one pole and one zero with equal frequencies. The pole is associated with the independent capacitor voltage. Because it cannot be controlled, it is canceled by a zero.

Non-observable state variable

A non-observable state variable of a network is one that cannot be observed at the output(s) of the network.

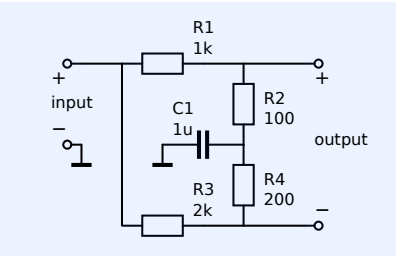


Figure 11.28: Passive two-port network with a non observable state.

Let us now consider the network from Figure 11.28. It shows a passive network with an input port and an output port. There exists a nonzero transfer from a voltage or from a current applied to the input port to the voltage across the capacitor. Hence, the voltage across the capacitor is controllable. However, the transfer from the capacitor voltage to the voltage or current in the output is zero. So, in this network, the state variable is not observable and called a non-observable state.

If we were to determine the poles and the zeros of the transfer parameters of this network, we would find one pole and one zero with equal frequencies. The pole is associated with the independent capacitor voltage. Because it cannot be observed, it is canceled by a zero.

11.5.5 Design conclusions

At this stage, the design of adequate high-pass and low-pass cut-off frequencies of negative feedback amplifiers had our interest. The DC gain of the controller can be used as a design parameter for high-pass cut-off frequency, while the gain-poles product can be used as a design parameter for the low-pass cut-off frequency.

The Nyquist stability analysis and the root locus technique show that negative feedback may yield unstable behavior. The root locus technique in particular indicates that feedback circuits with a loop gain with more than two poles may become unstable.[21] It also shows that the number of asymptotes can be decreased by inserting zeros into the loop. This changes the coefficients of the denominator of the servo function. The Routh array method could be used to investigate the conditions under which stability is achieved. In Chapter 12, we will discuss frequency compensation techniques that can be applied to achieve the desired frequency response of the amplifier without affecting the high-pass and low-pass cut-off frequencies that have been designed at this stage.

[21] Rule 6 shows that asymptotes move into the right half plane if the loop gain has more than two poles.

12 Frequency compensation

12.1 Introduction

In the previous chapter, we discussed the design of the bandwidth of a negative feedback amplifier. We have studied effects that contribute to high-pass cut-off and effects that contribute to low-pass cut-off. We have seen that speed limitations cause low-pass cut-off. Low-pass cut-off should occur above the frequency range of interest, and we found design criteria to achieve this. We have also seen that DC blocking elements in the loop or, more generally, the use of AC coupling techniques may result in high-pass cut-off. High-pass cut-off should occur below the frequency range of interest, and we have also found ways to achieve this.

However, until now, the actual character of the roll-off behavior has not yet been studied. We have found ways to determine the stability of the amplifier, but we have not yet defined the means to achieve stable behavior, nor have we studied ways to design a specific roll-off character. This will be the subject of this chapter. In order to deal with this topic in a structured way, we will assume that the high-frequency cut-off and the low-frequency cut-off can be designed independently. In other words, the pole-zero pattern that determines the high-pass behavior can be designed independently from the pole-zero pattern that determines the low-pass behavior, and vice versa. Since in this book, we confine ourselves to the design of wide-band amplifiers, this will be the case.

In practice, this means that capacitors and inductors that play a role in the low-frequency cut-off can be considered as short circuits and open circuits at high-frequency cut-off, respectively. Similarly, capacitors and inductors that determine the high-frequency roll-off can be considered as open circuits and short circuits during the low-frequency roll-off, respectively.

The design of the desired cut-off characteristic is usually referred to as *frequency compensation*. The use of this term implies that after we have designed the bandwidth, the pole zero pattern of the transfer is usually not as desired and should somehow be corrected. This is usually the case, and frequency compensation is a collective term for application of such corrections:

> Frequency compensation comprises a collection of techniques that can be used to correct pole-zero patterns in such a way as to obtain a desired frequency response or time response.

Frequency compensation techniques can be applied to obtain different types of responses. Examples of such responses are:

- Maximally flat magnitude (MFM) response
- No overshoot in step response
- Maximally flat group delay

In this chapter, we will focus on establishing an MFM response, but the techniques discussed can also be applied for obtaining other types of responses. Only the extent to which such techniques will be or can be applied, as well as the related design equations, will differ for other types of responses.

We will first study frequency compensation techniques that do not affect the designed bandwidth. This ensures that the design of the frequency response of a negative feedback amplifier can be performed in two subsequent steps:

1. The design of an adequate bandwidth
2. The design of a desired pole-zero patterns

Only in cases in which the amplifier's bandwidth exceeds the required bandwidth may frequency compensation result in bandwidth reduction. However, in order to obtain the largest possible performance-to-cost ratio, we will show that a limitation of the bandwidth of the ideal gain and maximization of the bandwidth of the servo function is the best strategy in such cases.

Frequency compensation should also not result in an unacceptable degradation of other performance aspects, such as, among others

- Signal-to-noise ratio
- Distortion and overdrive recovery
- Accuracy

12.1.1 Filter design approach

Before we will discuss frequency compensation techniques, we will formulate the goal of frequency compensation in a mathematical way. This facilitates the development of strategies and procedures for frequency compensation.

In the previous chapters, we have seen that the source-load transfer $A_f(s)$ of a negative feedback amplifier can be approximated by the product of its ideal transfer $A_i(s)$ and the servo function $\frac{-L(s)}{1-L(s)}$:

$$A_f(s) = A_i(s)\frac{-L(s)}{1-L(s)}. \tag{12.1}$$

This is the case if the loop gain reference variable has been selected such that the controller becomes a nullor for $|L| \to \infty$ and the influence of the direct transfer ρ is negligible.[1] From now on, we will assume this to be the case.

[1] In that case, the asymptotic gain equals the ideal gain and the loop gain is a measure for the difference between the ideal gain and the gain.

Low-pass cut-off

We will assume that the desired filter characteristic of the amplifier has been designed in the ideal transfer $A_i(s)$.[2] The bandwidth over which the source-load transfer $A_f(s)$ approaches the ideal transfer is determined by the bandwidth of the servo function. During the design of the low-pass cut-off we will assume that there exists a midband frequency range where the loop gain equals L_{MB}, with $|L_{MB}| \gg 1$, at higher frequencies, but below the low-pass cut-off frequency ω_h, the loop gain has n dominant poles and no zeros.[3] For studying the low-pass cut-off the loop gain may then be written as an $n-th$ order all-pole low-pass filter :

[2] See example 7.11.4: the design of a current integrator.

[3] At a later stage, we will see that poles with a frequency above ω_h can also be dominant. A proper definition of a dominant pole is "a pole that contributes to the bandwidth".

$$L(s) = \frac{L_{MB}}{\prod_{i=1}^{n}\left(1-\frac{s}{p_i}\right)}. \tag{12.2}$$

Figure 12.1 shows two asymptotes of the magnitude characteristic of the loop gain, as well as the asymptotes of the magnitude characteristic of the servo function for this situation. The two asymptotes of the loop gain are:

1. The asymptote at midband frequencies
2. The asymptote at the low-pass cut-off frequency ω_h

With the aid of (11.15), $A_f(s)$ can be written as:

$$A_f(s) = A_i(s)\frac{-L_{MB}}{1-L_{MB}}\frac{1}{1+a_1 s+a_2 s^2+\cdots a_n s^n}, \tag{12.3}$$

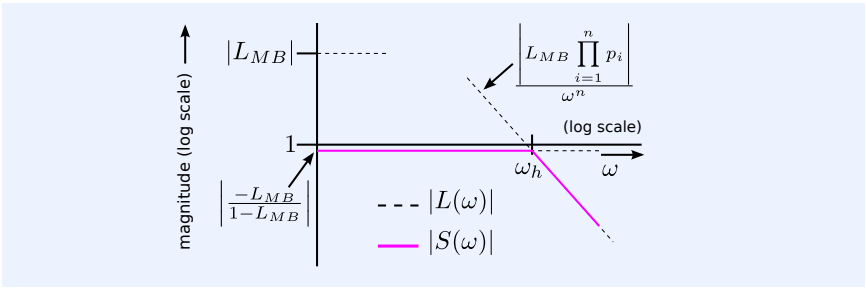


Figure 12.1: Asymptotes of the magnitude characteristics of the loop gain and the servo function for a loop gain with a midband value of L_{MB} and n dominant poles above midband frequencies.

where

$$a_n = \frac{1}{(1 - L_{MB}) \prod_{i=1}^{n} p_i}. \tag{12.4}$$

In case of an MFM characteristic , the low-pass cut-off frequency ω_h can be obtained as:

$$\omega_h = |a_n|^{-\frac{1}{n}}. \tag{12.5}$$

The goal of high-frequency compensation is to give the coefficients $a_1 \cdots a_{n-1}$ the values that correspond to those of an $n - th$ order low-pass filter with the desired characteristic and with a cut-off frequency close to ω_h.

High-pass cut-off

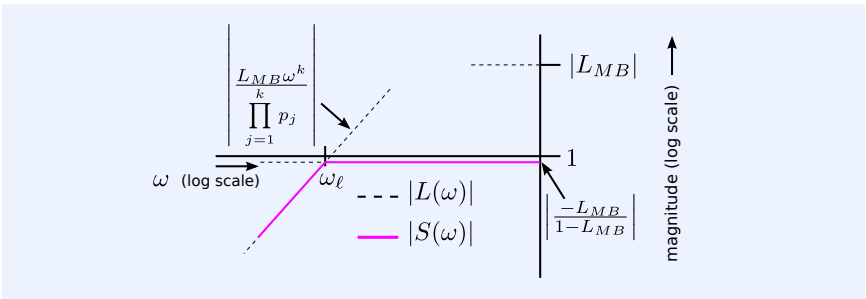


Figure 12.2: Asymptotes of the magnitude characteristics of the loop gain and the servo function for a loop gain with a midband value of L_{MB} and n dominant poles below midband frequencies.

Figure 12.2 illustrates a high-pass cut-off. For this type of transfer, we will assume that a midband frequency range can be defined where the value of the loop gain equals L_{MB}, with $|L_{MB}| \gg 1$, at lower frequencies, but above the high-pass cut-off frequency ω_ℓ, the loop gain has k dominant poles p_j, where k equals the number of zeros below ω_ℓ.[4] For studying high-pass cut-off the loop gain may then be approximated by a $k - th$ order high-pass filter and may be written as

$$L(s) = L_{MB} \frac{\prod_{j=1}^{k} \left(-\frac{s}{p_j}\right)}{\prod_{j=1}^{k} \left(1 - \frac{s}{p_j}\right)}. \tag{12.6}$$

The source-load transfer $A_f(s)$ can then be written in the form

$$A_f(s) = A_i(s) \frac{-L_{MB}}{1 - L_{MB}} \frac{b_k s^k}{1 + b_1 s + b_2 s^2 + \cdots b_k s^k}, \tag{12.7}$$

where

$$b_k = \frac{1 - L_{MB}}{\prod_{j=1}^{k} p_j}. \tag{12.8}$$

[4] At a later stage, we will see that poles with a frequency below ω_ℓ can also be dominant. Again, a proper definition of a dominant pole is "a pole that contributes to the bandwidth".

In case of an MFM characteristic , the high-pass cut-off frequency ω_ℓ can be obtained as

$$\omega_\ell = |b_k|^{-\frac{1}{k}}. \tag{12.9}$$

The goal of low-frequency compensation is to give the coefficients $b_1 \cdots b_{k-1}$ the values that correspond to those of an $k-th$ order high-pass filter with the desired characteristic and with a cut-off frequency close to ω_ℓ.

Design conclusion

We may draw the following conclusions:

- If, below certain frequencies, the loop gain drops below unity, the servo function will obtain a high-pass character.
- If, above certain frequencies, the loop gain drops below unity, the servo function will obtain a low-pass character.
- If both situations occur, the servo function will have a band-pass character. If the low-pass cut-off and high-pass cut-off are well separated, the source-load transfer can be written as a product of four terms:

$$A_f(s) = A_i(s)\ \frac{-L_{MB}}{1-L_{MB}}\ \frac{b_k s^k}{1+b_1 s+b_2 s^2+\cdots b_k s^k}\ \frac{1}{1+a_1 s+a_2 s^2+\cdots a_n s^n}, \tag{12.10}$$

  in which:

  - $A_i(s)$ is the ideal transfer, which has been designed using nullors as controllers.
  - The term $\frac{-L_{MB}}{1-L_{MB}}$ is a measure for the accuracy at midband frequencies.
  - The term $\frac{b_k s^n}{1+b_1 s+b_2 s^2+\cdots b_k s^n}$ describes the high-pass roll-off with respect to the ideal transfer. The MFM high-pass cut-off frequency ω_ℓ is found as $|b_k|^{-\frac{1}{k}}$, in which $b_k = \frac{1-L_{MB}}{\prod_{j=1}^{k} p_j}$ with p_j being a pole that contributes to high-pass cut-off , and L_{MB} the midband frequency loop gain.
  - The term $\frac{1}{1+a_1 s+a_2 s^2+\cdots a_n s^n}$ describes the low-pass roll-off with respect to the ideal transfer. The MFM low-pass cut-off frequency ω_h is found as $|a_n|^{-\frac{1}{n}}$, in which $a_n = \frac{1}{(1-L_{MB})\prod_{i=1}^{n} p_j}$ with p_i being a pole that contributes to the low-pass cut-off and L_{MB} the midband frequency loop gain.

12.1.2 Compensation techniques

There are different ways to adjust the coefficients a_i, $(i < n)$and b_j, $(j < k)$ to their desired values. Separate sections will be devoted to each of the techniques listed below.

1. The most powerful frequency compensation technique is the insertion of so-called *phantom zeros*.
2. Another technique is to change the positions of two (dominant) poles of the loop gain in such a way that their product remains the same, while their sum changes. In this way, the bandwidth of the servo function is preserved, but lower order coefficients of s in the denominator of the servo function can be modified. This technique is often referred to as *pole-splitting*.
3. Instead of splitting two poles by changing their interaction, poles can also be split with the aid of pole-zero canceling techniques.

4. An alternative to changing the sum of two poles, is to trade midband loop gain with the frequency of a dominant pole such that the product of the midband loop gain and the dominant poles is not affected. This technique is known as *resistive broadbanding*.

12.1.3 Compensation strategies

We have seen that the product of the midband loop gain and the dominant poles determine the bandwidth of the servo function. We have also seen that a large value of the loop gain is also beneficial to a high accuracy and a high linearity of the source-load transfer. Hence, there may exist situations in which severe requirements for the linearity or the accuracy result in a servo bandwidth that is far more than that required. In such cases, we may reduce the bandwidth while performing frequency compensation.

In general, there are three different strategies for frequency compensation:

1. Maintain the designed bandwidth of the source-load transfer. This is a useful strategy if the requirement for the midband loop gain that follows from the bandwidth design prevails over the one that follows from linearity or accuracy design considerations and the bandwidth has been designed to its desired value.

2. Exchange the bandwidth of the ideal transfer with the bandwidth of the servo function. This is a useful strategy if the requirement for the midband loop gain that follows from linearity and/or accuracy requirements dominates over the one obtained from bandwidth design considerations and if the bandwidth obtained in this way is more than required. This is the most powerful method of bandwidth reduction. However, we will see that a complicating side effect is that non-dominant poles of the loop gain before compensation may have become dominant after compensation.

3. Reduce the bandwidth of the servo function through reduction of the loop gain poles product. This may be a useful strategy if the bandwidth is larger than required while there exist too many non-dominant poles in the loop gain such that the previous method could not be implemented.

12.1.4 This chapter

Phantom zero compensation will be discussed in section 12.2. We will introduce the concept of phantom zeros, calculate their values for second and third order systems and present implementation methods. We will also present a technique for bandwidth reduction with the aid of phantom zeros and discuss the influence of non-dominant poles.

Pole-splitting by means of capacitive negative feedback across a gain stage will be discussed in section 12.3. This technique is often referred to as *Miller compensation* [Miller1920][5].

Pole-splitting by means of pole-zero canceling will be presented in section 12.4.

Resistive broadbanding is a brute-force technique for exchanging a pole frequency with the midband loop gain, while maintaining the product of the midband loop gain and the dominant poles. It will be discussed in section 12.5.

Phase margin and amplitude margin are properties of the loop gain that are often used as a measure for the stability of a negative feedback amplifier. Although the method is not advocated here, frequency compensation driven by phase margin improvement will be illustrated in section 12.6.

Reduction of the servo bandwidth can be useful for reducing the number of dominant poles. It can be achieved either by reducing the midband loop

[5] John M. Miller. Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit. *Scientific Papers of the Bureau of Standards*, 15(351):367–385, 1920

gain, lowering the frequency of one or more poles or adding one dominant pole. It will be discussed in section 12.7.

In section 12.9, we will discuss the design of the controller using cascaded feedback amplifiers. In control theory this is known as *nested loop control*. This technique should be applied if the controller should have a well-defined dynamic behavior, such as in analog PID controllers.

In section 12.8, we will discuss the frequency compensation of circuits that use negative feedback biasing. This is often referred to as low-frequency compensation.

Large variations in the drive and termination impedance of an amplifier may require special measures to ensure stability. Such variations may be part of the character of the signal source or the load, but they may also be a result of a shorted or a disconnected amplifier port. In section 12.10 we will discuss techniques for dealing with these impedance variations.

Non-dominant poles are poles that do not contribute to the bandwidth of the servo function. The frequency of these poles exceeds the unity-gain frequency of the loop gain. Although they cannot be included in the design of the bandwidth of the servo function, they may cause severe deviations from the MFM characteristic designed considering dominant poles alone. Their influence and measures for its reduction will be discussed in section 12.11.

12.2 Phantom zero compensation

The root locus method , presented in section 11.5.3, showed us that while increasing the loop gain, the poles of the servo function may obtain a very large imaginary part or even move into the right half plane. The method also showed that poles tend to move towards zeros. Hence, insertion of left half-plane zeros in the loop may be beneficial to the stability. However, zeros inserted into the loop transfer function are also zeros in the source load transfer. This directly follows from the expression for the servo function in which the loop transfer function appears in the numerator. Since an MFM roll-off requires the realization of an all-pole function, it appears as if frequency compensation by insertion of zeros in the loop is not an option. However, there is a way out of this using phantom zeros.

12.2.1 The phantom zero concept

A *phantom zero* owes its name to the fact that it is not visible in the source-load transfer. A phantom zero can have a beneficial effect on the stability, while it does not appear in the source-load transfer. The concept of the phantom zero can be understood from the expression for the gain $A_f(s)$ of a negative feedback amplifier:

$$A_f(s) = A_{f\infty}(s)\frac{-L(s)}{1-L(s)}. \tag{12.11}$$

A phantom zero is a zero in the loop gain $L(s)$ that coincides with a pole in the asymptotic gain $A_{f\infty}(s)$.

Such a zero does not appear as a zero in the source-load transfer, but it changes the numerator $1-L(s)$ of the servo function. Let us, for example, insert a phantom zero z_1 into $L(s)$ and modify expression (12.11) accordingly. Hence, we need to add a zero z_1 into the loop gain and a pole z_1 into the asymptotic gain $A_{f\infty}(s)$. Expression (12.11) then changes to

$$A_f(s) = \frac{A_{f\infty}(s)}{(1-s/z_1)}\frac{-L(s)\,(1-s/z_1)}{1-L(s)\,(1-s/z_1)}, \tag{12.12}$$

which can be simplified to:

$$A_f(s) = A_{f\infty}(s)\frac{-L(s)}{1-L(s)\,(1-s/z_1)}. \tag{12.13}$$

This clearly shows that the phantom zero only changes the denominator (and thus the poles) of the servo function. It does not change the numerator of $A_f(s)$.

In the following sections, we will study the conditions for phantom zero compensation of feedback amplifiers with a loop gain with two and three poles. SLiCAP examples will be given with simplified circuits of voltage amplifiers in which both the controller and the feedback network have been modeled with the aid of voltage-controlled voltage sources. At a later stage, we will discuss the compensation of amplifiers with passive feedback networks with the aid of phantom zeros.

12.2.2 Second order compensation

The dynamic part $F_2(s)$ of the servo function with a second-order MFM characteristic will have the form

$$F_2(s) = \frac{1}{1+s\sqrt{2}/\omega_h + s^2/\omega_h^2}, \tag{12.14}$$

in which ω_h is the -3 dB low-pass cut-off frequency.

Loop gain with two poles

Let us consider a loop gain $L(s)$ with two dominant poles p_1 and p_2 above midband frequencies, and a midband loop gain L_{MB}. The loop gain $L(s)$ can then be written as:

$$L(s) = \frac{L_{MB}}{(1-s/p_1)\,(1-s/p_2)}. \tag{12.15}$$

With the aid of (10.29) and (12.15), the second order servo function $S_2(s)$ can be obtained as

$$S_2(s) = \frac{-L_{MB}}{1-L_{MB}}\frac{1}{1-s\frac{p_1+p_2}{(1-L_{MB})p_1p_2}+s^2\frac{1}{(1-L_{MB})p_1p_2}}. \tag{12.16}$$

After equating the coefficients of (12.14) and (12.16), we find that a second order MFM characteristic is obtained if

$$p_1 + p_2 = -\sqrt{2}\omega_h, \tag{12.17}$$

in which the low-pass cut-off frequency ω_h equals

$$\omega_h^2 = (1-L_{MB})\,p_1p_2. \tag{12.18}$$

In the following example, we use SLiCAP to study the root locus of a second-order voltage follower that has an MFM response without compensation.

Example 12.1

Figure 12.3 shows a second order feedback system with poles p_1 and p_2 and DC loop gain L_{DC}.

The SLiCAP *netlist of the circuit, including the definitions of the values of the poles and the DC controller gain, is shown below:*

```
1 " Root Locus 2nd order voltage follower"
2 * File: RLvFollower-2.cir
```

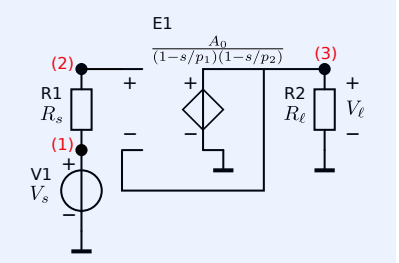


Figure 12.3: Voltage follower with MFM response.

```
* SLiCAP netlist file
V1 1 0 {V_s}
R1 1 2 {R_s}
E1 3 0 2 3 {A_0/(1-s/p_1)/(1-s/p_2)}
R2 3 0 {R_ell}
* For pole-zero analysis all parameters must have a numeric value
.param A_0=1M
+ p_1={-2*pi/sqrt(2)}
+ p_2={-2*pi*1M*sqrt(2)}
+ R_ell=1 R_s=1 V_s=1
.end
```

Because of the unity-gain feedback, the DC loop gain equals the controller gain: $L_{DC} = -10^6$. The loop gain has a dominant pole at $-\frac{1}{\sqrt{2}}$Hz and a second pole at $-10^6\sqrt{2}$Hz.

The low-pass cut-off frequency can be calculated with the aid of (12.18). We obtain: $f_h = 1$MHz. The sum of the poles is determined by p_2, and condition (12.17) has been satisfied. The voltage follower should have an MFM characteristic without compensation.

The script for plotting the root locus is listed below.

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: RLvFollower_2.py

from SLiCAP import *

fileName = 'RLvFollower_2'
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName + '.cir')
i1.setSource('V1')              # Signal source is V1
i1.setDetector('V_3')           # Voltage detector at node (3)
i1.setLGref('E1')               # Loop gain reference variable = E1
i1.setGainType('servo')         # Source to load transfer
i1.setSimType('numeric')        # Numeric simulation
i1.setDataType('poles')         # Calculate the poles
i1.setStepVar('A_0')            # Step the DC controller gain for RL plot
i1.setStepMethod('lin')
i1.setStepStart(0)
i1.setStepStop('1M')
i1.setStepNum(100)
i1.stepOn()
RL = i1.execute()
i1.stepOff()
i1.setGainType('gain')
polesGain = i1.execute()
i1.setGainType('loopgain')
polesLoopGain = i1.execute()
plots = [RL, polesLoopGain, polesGain]
htmlPage('Root locus')
fig_PZ = plotPZ('RL_' + fileName, fileName, plots, xmin=-2, xmax=0, ymin=-1,
     ymax=1, xscale='M', yscale='M', show=True)
fig2html(fig_PZ, 600)
```

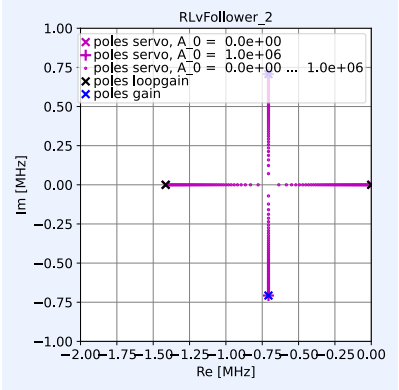


Figure 12.4: Root locus plot of the voltage follower with MFM response.

The root locus plot is shown in Figure 12.4. The root loci start at the poles of the loop gain. The endpoints for $A_0 = 10^6$ are the pole positions of the servo function. Since there are no poles or zeros present in the asymptotic gain, the poles of the gain equal those of the servo function. The poles of the loop gain, as well as the DC value of the loop gain, have been designed such that the system has an MFM response. The poles of the gain are in Butterworth positions (see Figure 11.1).

Lines 33-36 of the script displays the DC value and the poles and zeros of the gain in the Python *command window. Lines 37-47 print the phase margin and the frequency at which the magnitude of loop gain equals unity.*

```
# Calculate the phase margin
i1.setDataType('laplace')
i1.setGainType('loopgain')
L = i1.execute()
loopGain = L.laplace
pmResults = phaseMargin(loopGain)
```

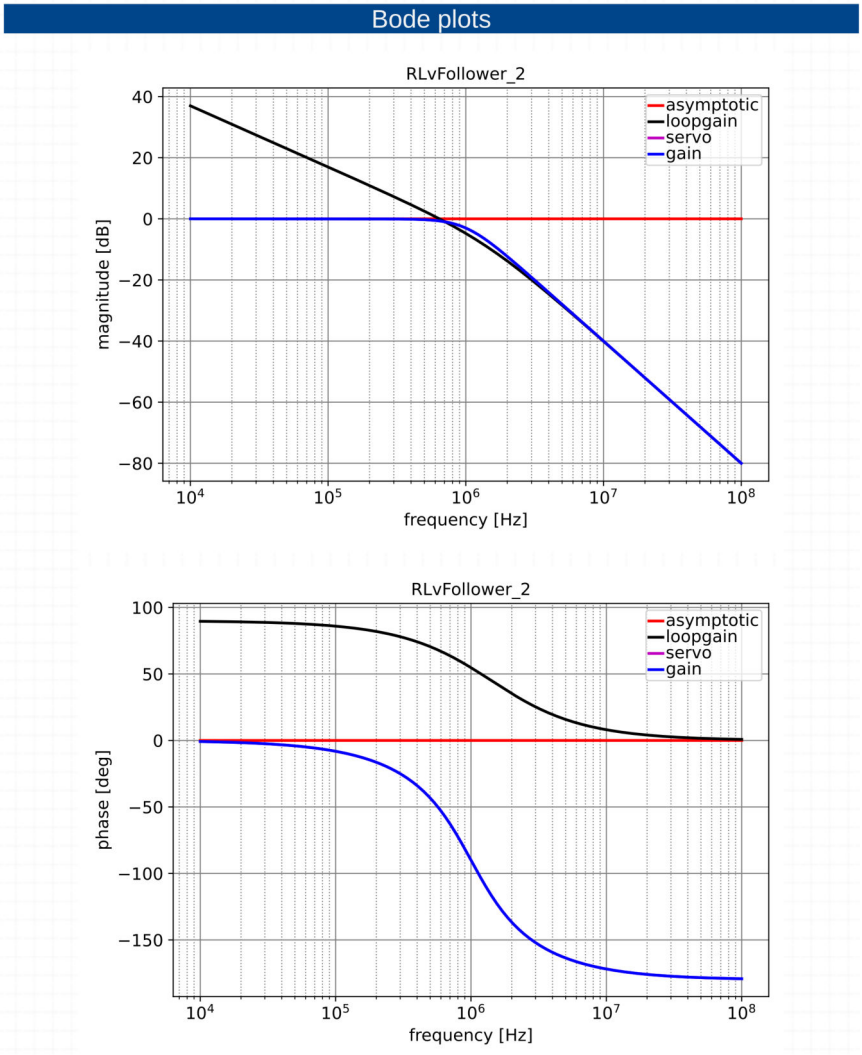
```
43
44 uF = pmResults[1]
45 pM = pmResults[0]
46
47 print('Loop gain: phase margin = {:3.2f}deg at f = {:8.2e}Hz'.format(pM, uF))
```

The results are shown below.

```
1  DC value of gain:  1.00e+0
2
3  Poles of gain:
4
5   n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
6  --  --------------  --------------  --------------  --------
7   0        -7.07e+5        -7.07e+5         1.00e+6   7.07e-1
8   1        -7.07e+5         7.07e+5         1.00e+6   7.07e-1
9
10 Found no zeros.
11
12 Loop gain: phase margin = 65.53deg at f = 6.44e+05Hz
```

Lines 49-61 generate the Bode plots of the asymptotic gain, the loop gain, the servo function and the gain of this voltage follower. The dB magnitude plots of these transfer functions are shown in Figure 12.5.

Figure 12.5: SLiCAP simulation result.




```
49 # Generate Bode plots
50 htmlPage('Bode plots');
51 i1.setGainType('servo');
```

```
S = i1.execute();
i1.setGainType('gain');
G = i1.execute();
i1.setGainType('asymptotic');
A = i1.execute();
plots = [A, L, S, G];
BodeMag  = plotSweep('dBmag_' + fileName, fileName, plots, 1e4, 1e8, 200,
     funcType='dBmag',  show=True)
BodePhas = plotSweep('Phase_' + fileName, fileName, plots, 1e4, 1e8, 200,
     funcType='phase', show=True)
fig2html(BodeMag, 800)
fig2html(BodePhas,800)
```

Since the asymptotic gain equals unity, and the direct transfer equals zero, the gain equals the servo function. The curve for the servo function has become invisible, because it has been overwritten by the gain curve.

The magnitude characteristic of the gain shows a second order MFM low-pass behavior with a $-3dB$ frequency at $1MHz$. The loop gain equals $0dB$ at $640kHz$. Please notice that negative feedback requires a negative midband frequency value of the loop gain. For this reason the phase plot of the loop gain starts at 180 degrees! The phase margin is 66 degrees.

Loop gain with two poles and one phantom zero

The sum of the poles of the servo function can be changed by adding one zero to the loop gain. In doing so, the expression for the loop gain (12.15) changes to

$$L(s) = L_{MB}\frac{1 - s/z}{(1 - s/p_1)(1 - s/p_2)}. \tag{12.19}$$

Let p_a and p_b be the poles of the servo function, thus allowing their sum to be found from the coefficient of s in the denominator of the servo function. Substitution of (12.19) in (10.29) shows that the sum of the poles of the servo function has been changed by the zero:

$$p_a + p_b = p_1 + p_2 - \frac{p_1 p_2}{z}L_{MB}. \tag{12.20}$$

With the aid of (12.14), it follows that an MFM characteristic is obtained if

$$-(p_1 + p_2) + \frac{p_1 p_2}{z}L_{MB} = \sqrt{2}\omega_h. \tag{12.21}$$

With $-L_{MB} > 1$, a second order MFM characteristic can be established if

$$z = -\frac{\omega_h^2}{\sqrt{2}\omega_h + p_1 + p_2}. \tag{12.22}$$

Hence, a second order feedback system can be given an MFM characteristic with one negative phantom zero if

$$-(p_1 + p_2) < \sqrt{2}\omega_h, \tag{12.23}$$

where $\omega_h = \sqrt{|(1 - L_{MB})\, p_1 p_2|}$.

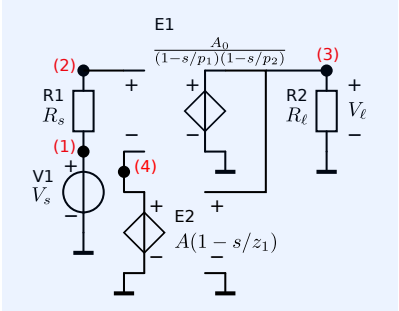


Figure 12.6: Voltage follower with MFM response and phantom-zero compensation.

Example 12.2

Figure 12.6 shows a voltage amplifier with a controller with two poles and one zero in the transfer of the feedback network. Since the ideal gain of this circuit is the reciprocal value of the transfer of the feedback network, this zero establishes a pole in the ideal gain. The transfer of the feedback network is also part of the loop gain. Thus, the zero in the feedback circuit is also a zero of the loop gain. Hence, it is a phantom zero, and it is not observable in the gain of the amplifier. With `E1` *the as loop gain reference variable, we are able to evaluate the asymptotic gain and the loop gain.*

The asymptotic gain $A_{f\infty}$ equals

$$A_{f\infty} = \frac{1}{A(1 - s/z_1)}. \tag{12.24}$$

The loop gain $L(s)$ equals

$$L(s) = \frac{-A_0 A(1 - s/z_1)}{(1 - s/p_1)(1 - s/p_2)}. \tag{12.25}$$

Let us assume the following numerical values:

$$\begin{aligned} p_1 &= -1\ Hz, \\ p_2 &= -100\ Hz, \\ A_0 &= 10^6\ [\text{-}], \\ A &= \frac{1}{100}\ [\text{-}]. \end{aligned} \tag{12.26}$$

The achievable second order bandwidth B_2 equals

$$\sqrt{1 \cdot 100 \cdot \frac{1}{100} \cdot 10^6} = 1000\ [Hz]. \tag{12.27}$$

The sum of the poles satisfies (12.23), so according to (12.22), MFM compensation can be achieved with a phantom zero which needs to be located at

$$z = -\frac{10^6}{1000\sqrt{2} - 101} = -761.49\ [Hz]. \tag{12.28}$$

The netlist of this circuit is:

```
"Root Locus 2nd order voltage amplifier"
* File: RLvAmp-2.cir
* SLiCAP netlist file
V1 1 0 {V_s}
R1 1 2 {R_s}
E1 3 0 2 4 {A_0/(1-s/p_1)/(1-s/p_2)}
E2 4 0 3 0 {A*(1-s/z_1)}
R2 3 0 {R_ell}
* For pole-zero analysis all parameters must have a numeric value
.param A_0 = 1M A = 10m
+ p_1 = {-2*pi}
+ p_2 = {-2*pi*100}
+ z_1 = {-2*pi*761.49}
+ R_ell = 1 R_s=1 V_s=1
.end
```

The script for plotting the root locus is:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: RLvAmp_2.py

from SLiCAP import *

fileName = 'RLvAmp_2'
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName + '.cir')
i1.setSource('V1')              # Signal source is V1
i1.setDetector('V_3')           # Voltage detector at node (3)
i1.setLGref('E1')               # Loop gain reference variable = E1
i1.setGainType('servo')         # Source to load transfer
i1.setSimType('numeric')        # Numeric simulation
i1.setDataType('poles')         # Calculate the poles
i1.setStepVar('A_0')            # Step the DC controller gain for RL plot
i1.setStepMethod('lin')
i1.setStepStart(0)
i1.setStepStop('1M')
i1.setStepNum(100)
i1.stepOn()
```

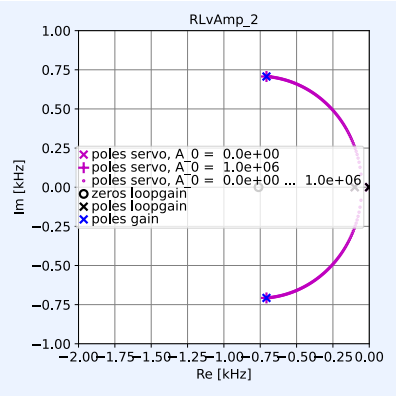


Figure 12.7: Root locus plot of the voltage amplifier with MFM response.

```
23 RL = i1.execute()
24 i1.stepOff()
25 i1.setGainType('gain')
26 polesGain = i1.execute()
27 i1.setGainType('loopgain')
28 polesLoopGain = i1.execute()
29 i1.setDataType('zeros')
30 zerosLoopGain = i1.execute()
31 plots = [RL, zerosLoopGain, polesLoopGain, polesGain]
32 htmlPage('Root locus')
33 fig_PZ = plotPZ('RL_' + fileName, fileName, plots, xmin=-2, xmax=0, ymin=-1,
       ymax=1, xscale='k', yscale='k', show=True)
34 fig2html(fig_PZ, 600)
```

The root locus plot is shown in Figure 12.7. The root loci start at the poles of the loop gain. The endpoints for $A_0 = 10^6$ are the pole positions of the servo function. The zero in the loop gain coincides with the pole in the asymptotic gain. Its effect on the root locus can clearly be observed.

Lines 35-38 of the above script display the DC value, the poles, and the zeros of the gain in the PYTHON *command window. Lines 39-49 print the phase margin and the frequency at which the magnitude of the loop gain equals unity.*

```
39 # Calculate the phase margin
40 i1.setDataType('laplace')
41 i1.setGainType('loopgain')
42 L = i1.execute()
43 loopGain = L.laplace
44 pmResults = phaseMargin(loopGain)
45
46 uF = pmResults[1]
47 pM = pmResults[0]
48
49 print('Loop gain: phase margin = {:3.2f}deg at f = {:8.2e}Hz'.format(pM, uF))
```

The results are shown in the PYTHON *command window. According to our expectations, the two poles are in MFM positions and the low-pass cut-off frequency is found at* 1kHz.

```
1  DC value of gain:  1.00e+2
2
3  Poles of gain:
4
5   n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
6  --  --------------  --------------  --------------  --------
7   0        -7.07e+2        -7.07e+2         1.00e+3   7.07e-1
8   1        -7.07e+2         7.07e+2         1.00e+3   7.07e-1
9
10 Found no zeros.
11
12 Loop gain: phase margin = 66.61deg at f = 1.47e+03Hz
```

Lines 51-63 create the Bode plots of the asymptotic gain, the loop gain, the servo function, and the gain of this voltage follower:

```
51 # Generate Bode plots
52 htmlPage('Bode plots');
53 i1.setGainType('servo');
54 S = i1.execute();
55 i1.setGainType('gain');
56 G = i1.execute();
57 i1.setGainType('asymptotic');
58 A = i1.execute();
59 plots = [A, L, S, G];
60 BodeMag  = plotSweep('dBmag_' + fileName, fileName, plots, 1, 1e4, 200,
       funcType='dBmag',  show=True)
61 BodePhas = plotSweep('Phase_' + fileName, fileName, plots, 1, 1e4, 200,
       funcType='phase', show=True)
62 fig2html(BodeMag, 800)
63 fig2html(BodePhas,800)
```

The dB magnitude plots and phase of these transfer functions is shown in Figure 12.8. The phantom zero can clearly be observed as a zero in the loop gain and a pole in the asymptotic gain. Its influence in the servo function can

also be observed by an increase of the magnitude around 1kHz. The gain has an all-pole MFM characteristic. Please notice that negative feedback requires a negative midband frequency value of the loop gain. For this reason, the phase plot of the loop gain starts at 180 degrees! The plot clearly shows the influence of the phantom zero on the phase margin. The phantom zero shifts the unity-gain frequency of the loop gain upwards and increases the phase margin. The MFM-compensated amplifier has a phase margin of 67 degrees.

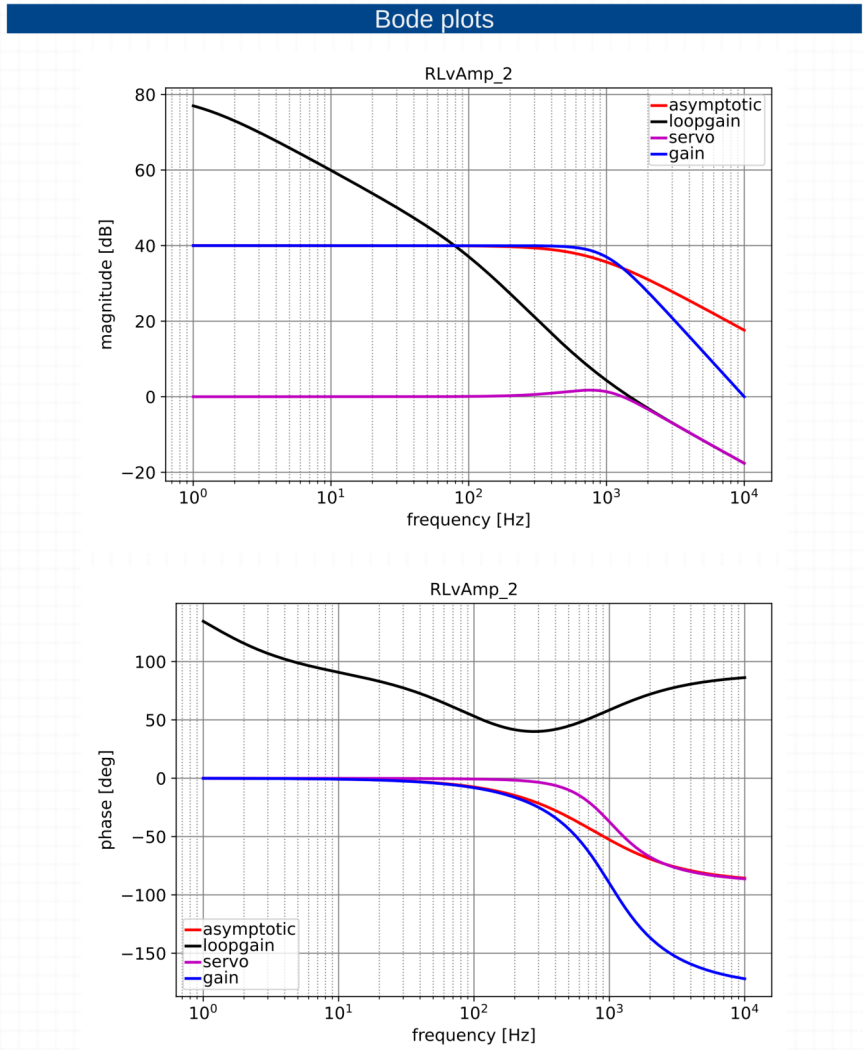


Figure 12.8: SLiCAP simulation result.

12.2.3 Third order compensation

The dynamic part of a third order MFM servo function $S_3(s)$ has the form:

$$S_3(s) = \frac{1}{1 + 2s/\omega_h + 2s^2/\omega_h^2 + s^3/\omega_h^3}, \tag{12.29}$$

in which ω_h is the -3 dB low-pass cut-off frequency.

Loop gain with three poles

We will consider an all-pole loop gain with three poles p_1, p_2, and p_3:

$$L(s) = L_{MB}\frac{1}{(1-s/p_1)(1-s/p_2)(1-s/p_3)}. \tag{12.30}$$

After substitution of (12.30) in (10.29), we obtain the expression for the servo function as

$$S_3(s) = \frac{-L_{MB}}{1-L_{MB}}\frac{1}{1-s\frac{p_1p_2+p_1p_3+p_2p_3}{(1-L_{MB})p_1p_2p_3}+s^2\frac{p_1+p_2+p_3}{(1-L_{MB})p_1p_2p_3}-s^3\frac{1}{(1-L_{MB})p_1p_2p_3}} \tag{12.31}$$

With the aid of (12.29), we find that all the poles are in MFM positions if:

$$p_1p_2+p_1p_3+p_2p_3 = 2\omega_h^2 \tag{12.32}$$

and

$$p_1+p_2+p_3 = -2\omega_h. \tag{12.33}$$

in which $\omega_h = \sqrt[3]{|(1-L_{MB})\,p_1p_2p_3|}$.

Example 12.3

A third order feedback system with poles p_1, p_2 and p_3 obtains an MFM response without compensation if the conditions from (12.32) and (12.33) are satisfied. If we add the condition that the poles of the loop gain should not have a positive real part, we find only one valid solution. It has one pole in the origin $p_1 = 0$ and two complex poles $p_{2,3} = -\omega_n \pm j\omega_n$. The loop gain can be written as

$$L(s) = \frac{-\omega_h^3}{s\left(s^2+2s\omega_h+2\omega_h^2\right)}. \tag{12.34}$$

This will be demonstrated for the voltage follower from Figure 12.9. The voltage follower described in the SLiCAP *netlist below has been given an MFM bandwidth of* 1MHz*:*

```
" Root Locus 3nd order voltage follower"
* File: RLvFollower-3.cir
* SLiCAP netlist file
V1 1 0 {V_s}
R1 1 2 {R_s}
E1 3 0 2 3 {8*pi^3*f_h^2*A_0/s/(s^2+s*4*pi*f_h+8*pi^2*f_h^2)}
R2 3 0 {R_ell}
* For pole-zero analysis all parameters must have a numeric value
.param f_h = 1M R_ell=1 R_s=1 V_s=1 A_0=1M
.end
```

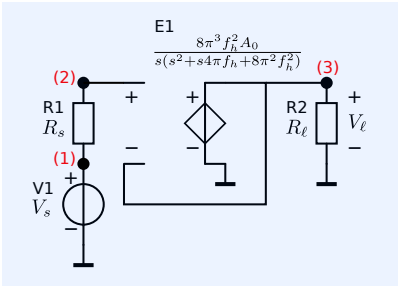


Figure 12.9: Voltage follower with third-order MFM response.

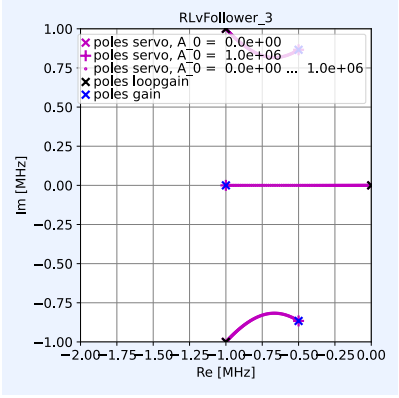


Figure 12.10: Root locus plot of the voltage follower with MFM response.

The script `RLvFollower_3.py` *for plotting the root locus is similar to* `RLvFollower_2.py`*. It only uses the circuit from Figure 12.9 instead of that from Figure 12.5.*

The root locus plot is shown in Figure 12.10. The root loci start at the poles of the loop gain. The endpoints for $A_0 = 10^6$ are the pole positions of the servo function. Since there are no poles or zeros present in the asymptotic gain, the poles of the gain equal those of the servo function. The poles of the loop gain, as well as the loop gain poles product have been designed such that the system has a 3-rd order MFM response. The poles of the gain are in Butterworth positions (see Figure 11.1).

Below is the output generated by `RLvFollower_3.py` *as it is displayed in the* Python *command window. This confirms our expectations.*

```
DC value of gain:  1.00e+0

Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
```

```
--  --------------  --------------  --------------  --------
 0        -5.00e+5        -8.66e+5         1.00e+6   1.00e+0
 1        -5.00e+5         8.66e+5         1.00e+6   1.00e+0
 2        -1.00e+6        0.00e+00         1.00e+6

Found no zeros.

Loop gain: phase margin = 60.49deg at f = 4.96e+05Hz
```

The dB magnitude plots and the phase plots of the transfer functions of the asymptotic gain model are shown in Figure 12.11. Since the asymptotic gain equals unity, and the direct transfer equals zero, the gain equals the servo function. The curve for the servo function has become invisible, because it has been overwritten by the gain curve. Please notice that negative feedback requires a negative midband frequency value of the loop gain. For this reason the phase plot of the loop gain with a pole in the origin starts at 90 degrees! The phase margin of the loop gain is 60deg and the unity-gain frequency of the loop gain is 500kHz.

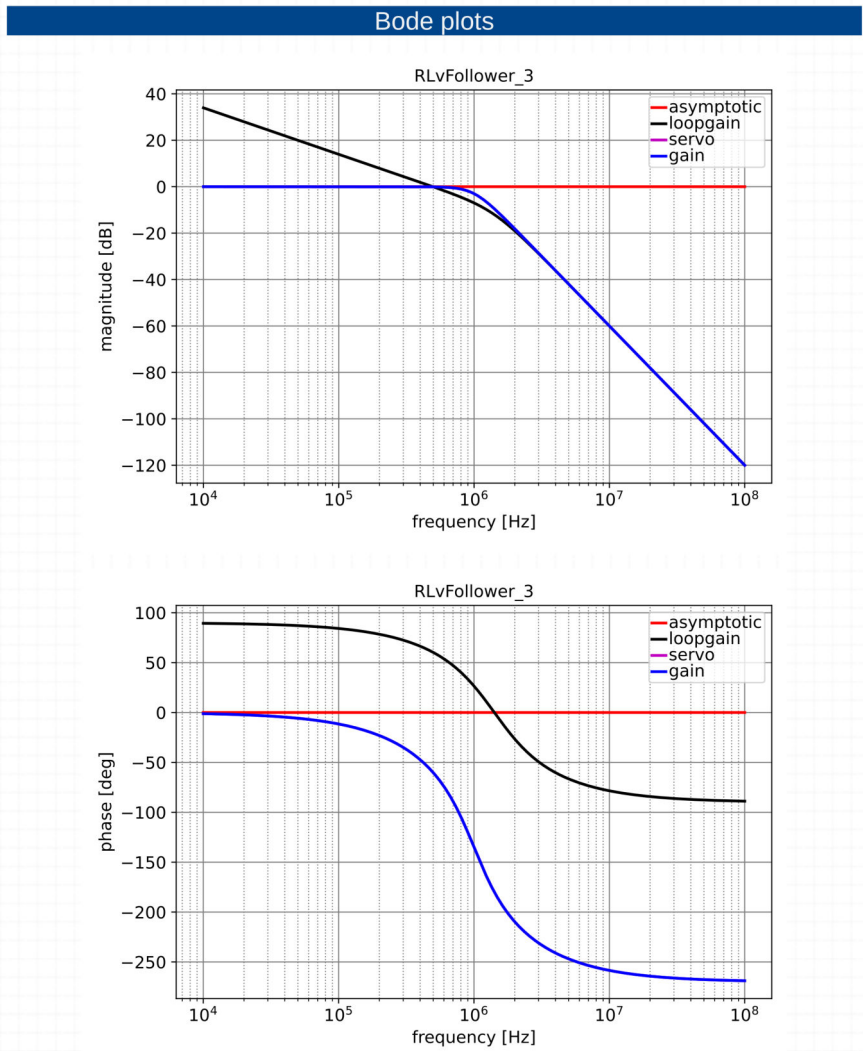


Figure 12.11: SLiCAP simulation result.

Loop gain with three poles and one phantom zero

We will now investigate the conditions under which a third-order system can be compensated with one phantom zero. A loop transfer function with one

zero z_1 and three poles p_1, p_2, and p_3 can be written as

$$L(s) = L_{MB}\frac{1 - s/z_1}{(1 - s/p_1)(1 - s/p_2)(1 - s/p_3)}. \tag{12.35}$$

After substitution of (12.35) in (10.29), it can be seen that the coefficient of s^2 in the denominator depends on z_1. A third-order MFM response for the servo function now requires the following two conditions to be satisfied:

$$p_1 + p_2 + p_3 = -2\omega_h, \tag{12.36}$$

$$p_1p_2 + p_1p_3 + p_2p_3 - \frac{L_{MB}p_1p_2p_3}{z_1} = 2\omega_h^2. \tag{12.37}$$

Assuming $-L_{MB} > 1$, a third-order MFM characteristic can be accomplished if:

$$z_1 = \frac{\omega_h^3}{p_1p_2 + p_1p_3 + p_2p_3 - 2\omega_h^2}. \tag{12.38}$$

From this, we see that a third-order system can be given an MFM characteristic with one negative real phantom zero if

$$p_1p_2 + p_1p_3 + p_2p_3 < 2\omega_h^2 \tag{12.39}$$

and if

$$p_1 + p_2 + p_3 = -2\omega_h \tag{12.40}$$

in which $\omega_h = \sqrt[3]{|(1 - L_{MB})\, p_1p_2p_3|}$.

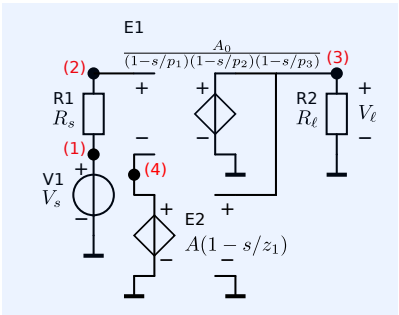


Figure 12.12: Circuit of a voltage amplifier with three poles and one phantom zero.

Example 12.4

Figure 12.12 shows the circuit of a third-order voltage amplifier that has been compensated with one phantom zero. The circuit is equivalent to the one from Figure 12.6, only the expressions for the voltage-controlled voltage sources have been changed. The pole positions have been chosen to comply with (12.39) and (12.40):

$$p_1 = -10\ Hz,\ p_2 = -590\ Hz,\ p_3 = -1400\ Hz. \tag{12.41}$$

The DC loop gain L_{DC} equals

$$L_{DC} = -A_0A = -120, \tag{12.42}$$

from which we obtain an MFM bandwidth f_h:

$$f_h = \sqrt[3]{121 \cdot 10 \cdot 590 \cdot 1400} = 1000. \tag{12.43}$$

The sum of the poles equals -2000, which is $-2f_h$. This satisfies (12.40). The sum of the products of two poles is less than two times the squared bandwidth. This satisfies condition (12.39). The amplifier can thus obtain an MFM characteristic by inserting one negative real phantom zero. The position of the zero follows from (12.38):

$$z_1 = \frac{10^9}{10 \cdot 590 + 10 \cdot 1400 + 590 \cdot 1400 - 2 \cdot 10^6} = -866.48Hz. \tag{12.44}$$

The netlist of this circuit is shown below:

```
"Root Locus 3nd order voltage amplifier, one zero"
* File: RLvAmp-3_1.cir
* SLiCAP netlist file
V1 1 0 {V_s}
R1 1 2 {R_s}
E1 3 0 2 4 {A_0/(1-s/p_1)/(1-s/p_2)/(1-s/p_3)}
E2 4 0 3 0 {A*(1-s/z_1)}
R2 3 0 {R_ell}
```

```
* For pole-zero analysis all parameters must have a numeric value
.param A_0 = 12k A = 10m
+ p_1 = {-2*pi*10}
+ p_2 = {-2*pi*590}
+ p_3 = {-2*pi*1400}
+ z_1 = {-2*pi*866.48}
+ R_ell = 1 R_s=1 V_s=1
.end
```

The script `RLvAmp_3_1.py` *is similar to* `RLvAmp_2.py`*. It plots the root locus, but uses the circuit from Figure 12.12 instead of that from Figure 12.6.*

The root locus plot generated by this script is shown in Figure 12.13. It clearly shows a third-order response with the poles in MFM positions. The phantom zero in the loop gain is not observable in the gain, because it coincides with a pole in the asymptotic gain.

Below is the output generated by the script `RLvAmp_3_1.py` *as it is displayed in the* PYTHON *command window. It confirms our expectations.*

```
Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0       -4.95e+2        -8.63e+2         9.95e+2   1.00e+0
 1       -4.95e+2         8.63e+2         9.95e+2   1.00e+0
 2       -1.01e+3        0.00e+00         1.01e+3

Found no zeros.

Loop gain: phase margin = 49.24deg at f = 8.29e+02Hz
```

The phase margin of this MFM-compensated third-order amplifier is 49 *degrees.*

The dB magnitude plots and the phase plots of the transfer functions of the asymptotic gain model are shown in Figure 12.14. Please notice that negative feedback requires a negative midband frequency value of the loop gain. For this reason, the phase plot of the loop gain starts at 180 *degrees!*

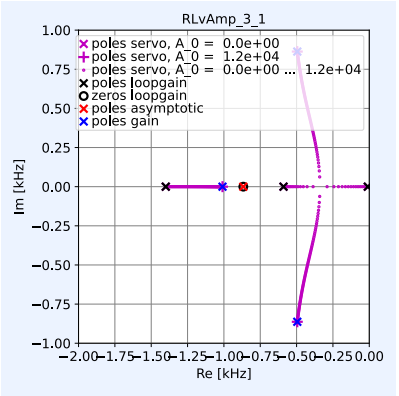


Figure 12.13: Root locus plot of the voltage amplifier with third order MFM response, compensated with one phantom-zero.

Loop gain with three poles and two phantom zeros

A loop transfer function with three poles p_1, p_2, and p_3 and two zeros z_1 and z_2 can be written as

$$L(s) = L_{MB}\frac{(1-s/z_1)\,(1-s/z_2)}{(1-s/p_1)\,(1-s/p_2)\,(1-s/p_3)}. \tag{12.45}$$

From this expression, we can calculate the frequency-dependent part of the servo function. We then obtain expressions for the first order and the second order coefficients of s of the denominator of the servo function. We may then conclude that the poles will be in MFM positions if

$$p_1p_2 + p_1p_3 + p_2p_3 - L_{MB}p_1p_2p_3\left(\frac{1}{z_1}+\frac{1}{z_2}\right) = 2\omega_h^2, \tag{12.46}$$

$$p_1 + p_2 + p_3 - \frac{p_1p_2p_3L_{MB}}{z_1z_2} = -2\omega_h. \tag{12.47}$$

From (12.46) and (12.46), we obtain values for the product and the sum of the phantom zeros:

$$z_1z_2 = \frac{\omega_h^3}{p_1+p_2+p_3+2\omega_h}, \tag{12.48}$$

$$z_1 + z_2 = \frac{p_1p_2 + p_1p_3 + p_2p_3 - 2\omega_h^2}{p_1+p_2+p_3+2\omega_h}. \tag{12.49}$$

With two phantom zeros that have a negative real part, a third order MFM

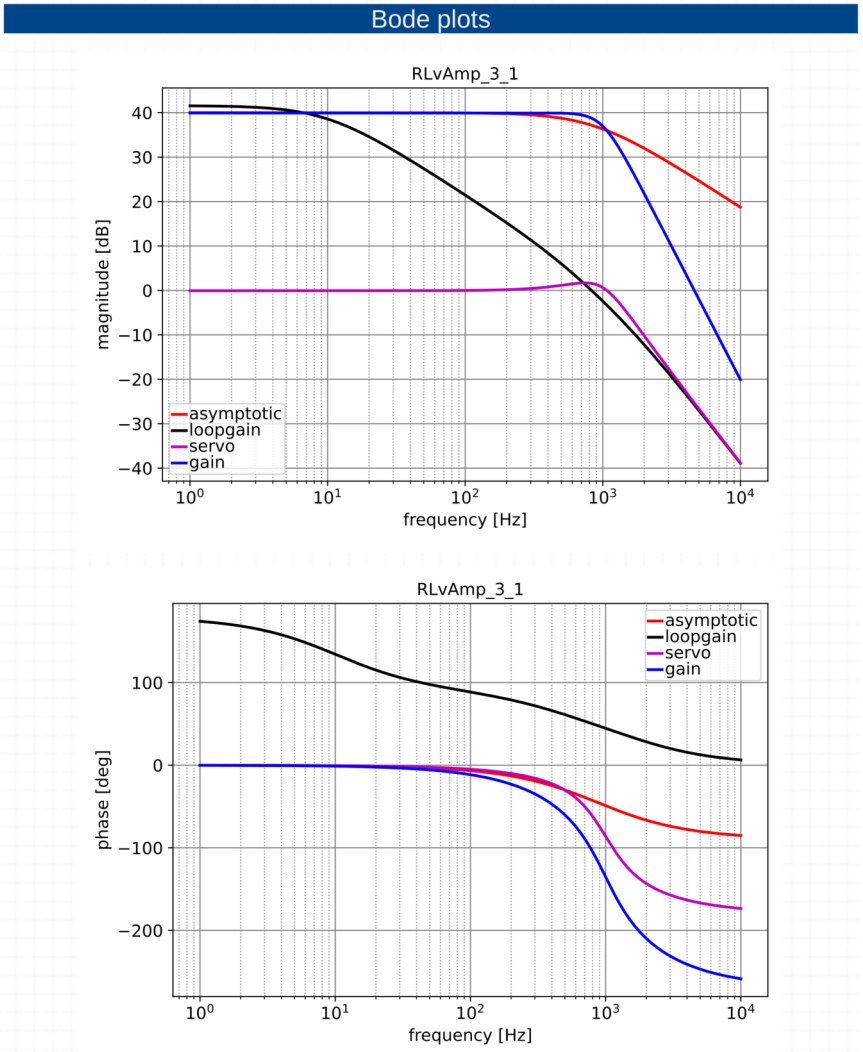
transfer can be realized if

$$p_1 p_2 + p_1 p_3 + p_2 p_3 < 2\omega_h^2, \tag{12.50}$$

$$-(p_1 + p_2 + p_3) < 2\omega_h, \tag{12.51}$$

in which $\omega_h = \sqrt[3]{|(1 - L_{MB})\, p_1 p_2 p_3|}$.

Figure 12.14: SLiCAP simulation result.



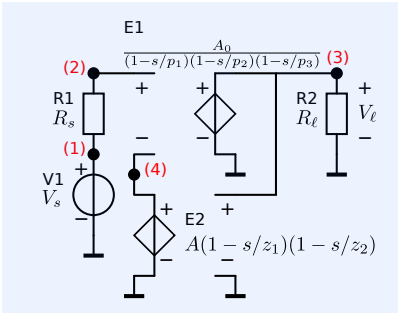


Figure 12.15: Circuit of a voltage amplifier with three poles and two phantom zeros.

The phantom zeros can then be calculated from 12.46 and 12.47 using

$$z_{1,2} = \frac{1}{2}(z_1 + z_2) \pm \frac{1}{2}\sqrt{(z_1 + z_2)^2 - 4z_1 z_2}. \tag{12.52}$$

In the following example, we will demonstrate the compensation of a third order system with two phantom zeros.

Example 12.5

Figure 12.15 shows the circuit of a third order voltage amplifier that has been compensated with two phantom zeros. The circuit is equivalent to the one from Figure 12.6, only the expressions for the voltage-controlled voltage sources have been changed. The pole positions have been chosen to comply with (12.50) and (12.51):

$$p_1 = -20\ Hz,\ p_2 = -50\ Hz,\ p_3 = -1250\ Hz. \tag{12.53}$$

The DC loop gain L_{DC} equals:

$$L_{DC} = -A_0 A = -799, \tag{12.54}$$

from which we obtain an MFM bandwidth f_h:

$$f_h = \sqrt[3]{800 \cdot 20 \cdot 50 \cdot 1250} = 1000\ Hz. \tag{12.55}$$

The sum of the poles equals $-1320Hz$, which satisfies (12.40). The sum of the products of two poles equals 88.5×10^3, which satisfies condition (12.39). The amplifier can thus be given an MFM characteristic by inserting two phantom zeros that have a negative real part. The product of the zeros follows from (12.48):

$$z_1 z_2 = \frac{1000^3}{-1320 + 2 \cdot 1000} = 1.47 \times 10^6\ Hz^2. \tag{12.56}$$

The sum of the zeros follows from (12.49):

$$z_1 + z_2 = \frac{88.5 \times 10^3 - 2 \cdot 1000^2}{-1320 + 2 \cdot 1000} = -2811\ Hz. \tag{12.57}$$

With the aid of (12.52), we obtain

$$z_1 = -695\ Hz, \tag{12.58}$$

$$z_2 = -2116\ Hz. \tag{12.59}$$

The netlist of this circuit is shown below.

```
"Root Locus 3nd order voltage amplifier, two zeros"
* File: RLvAmp-3_2.cir
* SLiCAP netlist file
V1 1 0 {V_s}
R1 1 2 {R_s}
E1 3 0 2 4 {A_0/(1-s/p_1)/(1-s/p_2)/(1-s/p_3)}
E2 4 0 3 0 {A*(1-s/z_1)*(1-s/z_2)}
R2 3 0 {R_ell}
* For pole-zero analysis all parameters must have a numeric value
.param A_0 = 1M A = 799u
+ p_1 = {-2*pi*20}
+ p_2 = {-2*pi*50}
+ p_3 = {-2*pi*1250}
+ z_1 = {-2*pi*(695)}
+ z_2 = {-2*pi*(2116)}
+ R_ell = 1 R_s=1 V_s=1
.end
```

The script `RLvAmp_3_2.py` *for plotting the root locus is similar to* `RLvAmp_2.py`. *It only uses the circuit from Figure 12.15 rather than that from Figure 12.12.*

The root locus plot generated by this script is shown in Figure 12.16. It clearly shows a third order response with the poles in MFM positions. The phantom zeros in the loop gain are not observable in the gain, because they coincide with poles in the asymptotic gain.

Below is the output generated by the scrip as it is displayed in the PYTHON *command window.*

```
DC value of gain:  1.25e+3

Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -4.99e+2        -8.66e+2         9.99e+2   1.00e+0
 1        -4.99e+2         8.66e+2         9.99e+2   1.00e+0
 2        -1.00e+3        0.00e+00         1.00e+3

Found no zeros.

Loop gain: phase margin = 47.98deg at f = 1.13e+03Hz
```

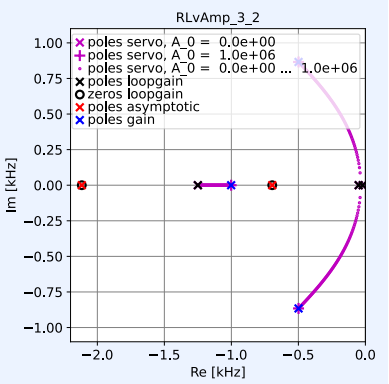
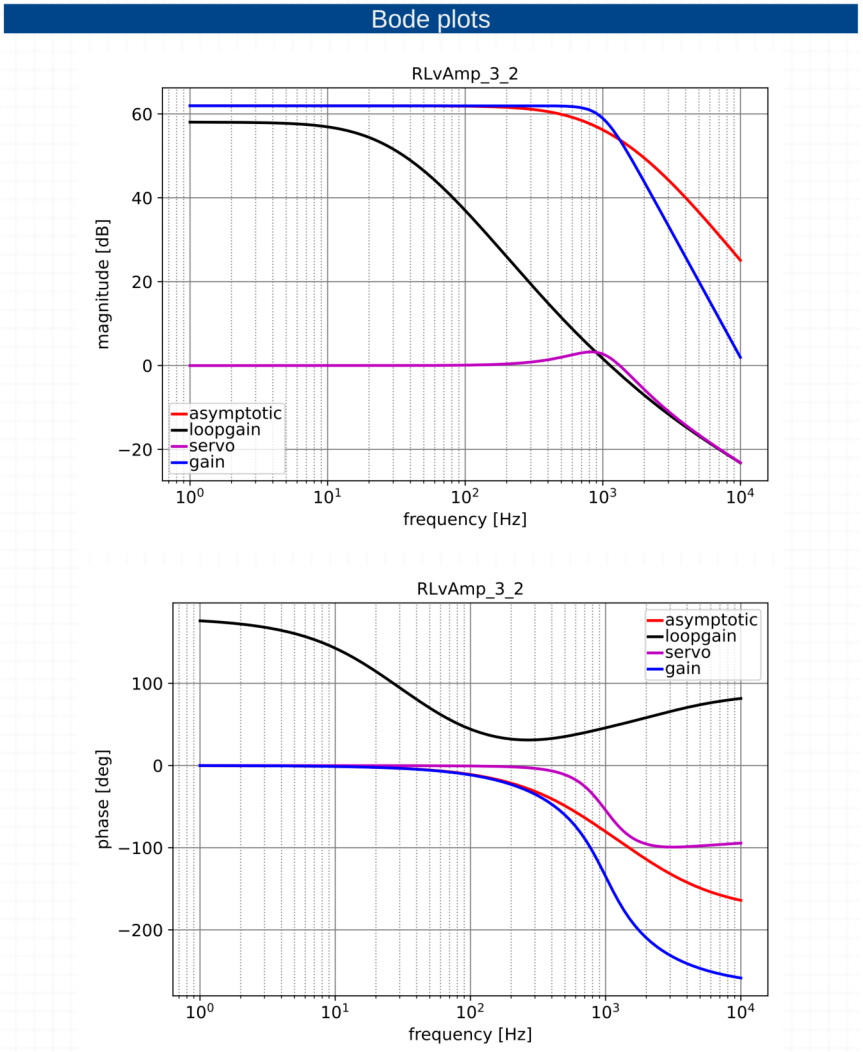


Figure 12.16: Root locus plot of the voltage amplifier with third order MFM response, compensated with two complex phantom-zeros.

The bandwidth, as well as the response type, are as expected.

The dB magnitude plots of the transfer functions of the asymptotic gain model are shown in Figure 12.17. Please notice that negative feedback requires a negative midband frequency value of the loop gain. For this reason, the phase plot of the loop gain starts at 180 *degrees!*

Figure 12.17: SLiCAP simulation result.



In the following example, we will demonstrate the frequency compensation in the case in which the loop gain has three poles in the origin. This may be of practical interest for systems that have the poles of the loop gain relatively close to the origin. The example shows that MFM compensation of such systems requires two complex conjugated phantom zeros with a quality factor of $\frac{1}{2}\sqrt{2}$ and a frequency of $\frac{1}{2}\sqrt{2}f_h$, where f_h is the -3dB MFM cut-off frequency in Hz.

Example 12.6

Let us now consider a loop gain $L(s)$ with three poles in the origin. This loop transfer function can be written as

$$L(s) = \frac{1}{a_3 s^3}, \tag{12.60}$$

in which $a_3 < 0$ (negative feedback). The expression for the servo function is then

obtained as

$$\frac{-L(s)}{1-L(s)} = \frac{1}{1-a_3 s^3}. \tag{12.61}$$

The MFM low-pass cut-off frequency ω_h is then found as

$$\omega_h = \sqrt[3]{\left|\frac{1}{a_3}\right|}. \tag{12.62}$$

Both the sum of the poles and the sum of the products of two poles are zero. Hence, conditions (12.50) and (12.51) have been met and an MFM low-pass cut-off of the servo function can be established with two phantom zeros. With the aid of (12.50), (12.51) and (12.52), we find:

$$z_1, z_2 = \frac{1}{2}\omega_h \pm \frac{1}{2} j\omega_h. \tag{12.63}$$

12.2.4 Implementation of phantom zeros

In the previous sections, we have studied the conditions under which second and third order systems can be compensated with the aid of phantom zeros. We also found means to determine their values. Although for higher order systems, the procedure is similar, we will confine ourselves to second order and third order systems. We will discuss the design of the frequency response of higher order amplifiers at a later stage.

The examples of amplifiers compensated with phantom zeros, that we discussed in the previous sections, were primarily intended to illustrate the concept of phantom zero compensation. The presented implementations of phantom zeros in the transfer of the controlled sources in the feedback path were conceptually correct, but far from practical. These zeros would result in an infinite gain of those controlled sources for $j\omega \to \infty$, which is non-physical.

In this section, we will discuss the implementation of phantom zeros in practical circuits. We will study the way in which phantom zeros can be realized and investigate possible side effects that may limit their effectiveness.

Phantom zero locations

By definition, phantom zeros are zeros in the loop gain that coincide with poles in the asymptotic gain. If we have selected the loop gain reference variable properly, the asymptotic gain equals the ideal gain. The ideal gain has been defined as the source-load transfer of the feedback amplifier in which the controllers have been replaced with nullors. From this, we obtain an important design conclusion:

> Poles in the asymptotic gain can only be established in circuitry around the nullor(s). Hence, the internal circuitry of the controller is not a place to search for implementation of phantom zeros!

The circuitry around the nullor comprises:

- The feedback network

  A pole in the asymptotic gain can be realized by inserting a zero into the transfer of the feedback network. This has already been illustrated in the examples in the previous section.

- The signal source

Insertion of a pole into the signal transfer from the signal source to the input of the amplifier causes a pole in the asymptotic gain. By doing so, we insert a low-pass filter between the signal source and the input of the amplifier.

- The load

  Similarly as above, insertion of a pole into the transfer from the output of the amplifier to the load also causes a pole in the asymptotic gain.

Phantom zero effectiveness

After we have realized the poles in the asymptotic gain, we need to investigate whether they also appear as *effective zeros* in the loop gain. We will call a phantom zero *effective* if it does what it should do: bring the dominant poles of the servo function into their desired positions. The effectiveness of phantom zeros may strongly be reduced if:

- The insertion of the phantom zero changes the original pole positions such that the bandwidth is reduced.
- The insertion of a phantom zero introduces a new dominant pole.

We will see that the insertion of phantom zeros is never free of such side effects, but in many cases, these effects may be kept within acceptable limits.

12.2.5 Phantom zeros in the feedback network

Zeros in the feedback network can be established as described in Chapter 18.5.3. The basic techniques are:

1. Change the feedback network in such a way that its transfer becomes zero at the (complex) frequency of the zero. In passive feedback networks, such as voltage attenuators, current attenuators, current-to-voltage conversion impedances and voltage-to-current conversion admittances, this works as follows:

   (a) Create zero admittance (an open circuit) at the (complex) frequency of the zero in a branch of the feedback network that appears in series with the signal path. This is done by placing an admittance in parallel with this branch whose value at the (complex) frequency of the zero is opposite to the value of the admittance of the original branch.[6]

   (b) Create zero impedance (a short) at the (complex) frequency of the zero in a branch of the feedback network that appears in parallel with the signal path. This is done by placing an impedance in series with that branch whose value at the (complex) frequency of the zero is opposite to the value of the impedance of the original branch.[7]

2. Add a second feedback path that cancels the transfer of the original one at the (complex) frequency of the zero.

In the following example, we will demonstrate the first method.

Example 12.7

Figure 12.18A shows a passive voltage attenuator that has been used as a feedback network in the voltage amplifier from Figure 12.18B. The resistor R1 *appears in series with the signal path, while* R2 *appears in parallel with the signal path.*[8] *Now, let us assume that a phantom zero at the complex frequency* $s = z_1$ *has to be implemented. As described above, we then need to create a zero in the transfer of the feedback network at this frequency. We have two possibilities:*

[6] Such a branch can be considered as one that performs voltage-to-current conversion.

[7] Such a branch can be considered as one that performs current-to-voltage conversion.

[8] A network branch or a sub network is in series with the signal path if, when replaced with an open circuit, the signal transfer becomes zero.
A network or a sub network is in parallel with the signal path if, when replaced with a short circuit, the signal transfer becomes zero.

1. *Place an admittance $Y_p(s)$ in parallel with R1 that cancels the admittance of R1 at $s = z_1$. After doing so, the admittance $Y_{tot}(s)$ of the modified branch is obtained as*

$$Y_{tot}(s) = Y_p(s) + \frac{1}{R_a}. \tag{12.64}$$

This function has a unique negative real solution if $Y_p(s)$ is capacitive. With a capacitance C_{phz}, we have $Y_p(s) = sC_{phz}$ and obtain

$$Y_{tot}(s) = sC_{phz} + \frac{1}{R_a}. \tag{12.65}$$

Hence, we have created a negative real zero z_1, that is the solution of $Y_{tot}(s) = 0$:

$$z_1 = -\frac{1}{R_a C_{phz}}. \tag{12.66}$$

Figure 12.19 shows the modified voltage attenuator.

2. *Place an impedance $Z_{se}(s)$ in series with R2 that cancels the resistance of R2 at $s = z_1$. After doing so, the impedance $Z_{tot}(s)$ of the modified branch is obtained as*

$$Z_{tot}(s) = Z_{se}(s) + R_b. \tag{12.67}$$

This function has a unique negative real solution if $Z_{se}(s)$ is inductive. With an inductance L_{phz}, we have $Z_{se}(s) = sL_{phz}$ and obtain

$$Z_{tot}(s) = sL_{phz} + R_b. \tag{12.68}$$

Hence, we have created a negative real zero z_1 that is the solution of $Z_{tot}(s) = 0$:

$$z_1 = -\frac{R_b}{L_{phz}}. \tag{12.69}$$

Figure 12.20 shows the modified voltage attenuator.

The example above gives a formal approach to the design of zeros in the transfer of the feedback network. We have not yet studied any side effects of such implementations, nor have we discussed the effectiveness of the phantom zeros. This will be done at a later stage. At this stage, we may draw the following conclusions for implementation of zeros in the transfer of feedback networks:

1. Branches in series with the signal path should be changed in such a way that the order of s in the admittance is increased and zero admittance is achieved at the (complex) frequency of the phantom zero:

   (a) If a series branch behaves resistively at the frequency of the phantom zero, a capacitive branch placed in parallel will cause a negative real zero in the transfer of the feedback network (see Figure 12.21A).

   (b) If a series branch behaves inductively at the frequency of the phantom zero, a resistive branch placed in parallel will cause a zero and a capacitive branch placed in parallel will cause two (complex) zeros in the transfer of the feedback network (see Figure 12.21B).

   (c) If a series branch behaves capacitively at the frequency of the phantom zero, then there is no means for creating a zero in this branch (see Figure 12.21C).

2. Branches in parallel with the signal path should be changed in such a way that the order of s in the impedance is increased and zero impedance is achieved at the (complex) frequency of the phantom zero:

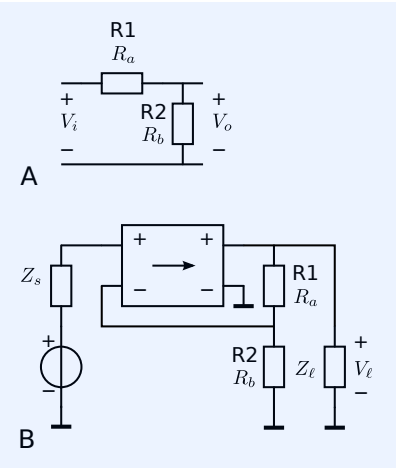


Figure 12.18: A. Resistive voltage attenuator
B. Application of the resistive voltage attenuator in a passive feedback voltage amplifier.

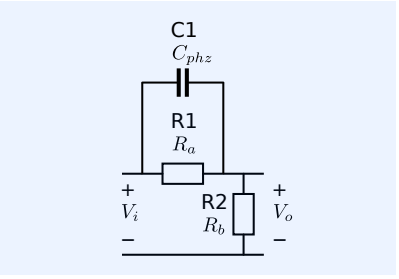


Figure 12.19: Resistive voltage attenuator with added zero in the series branch.

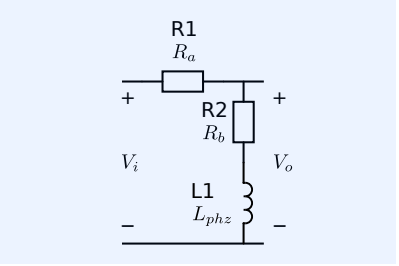


Figure 12.20: Resistive voltage attenuator with added zero in the parallel branch.

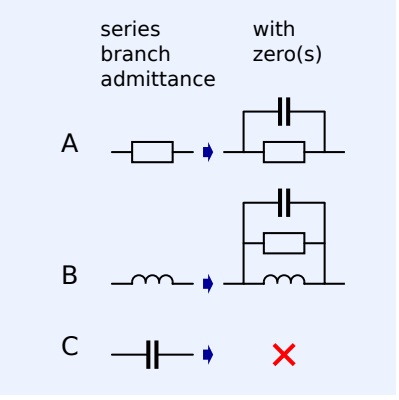


Figure 12.21: Implementation of zeros in series branch admittances.

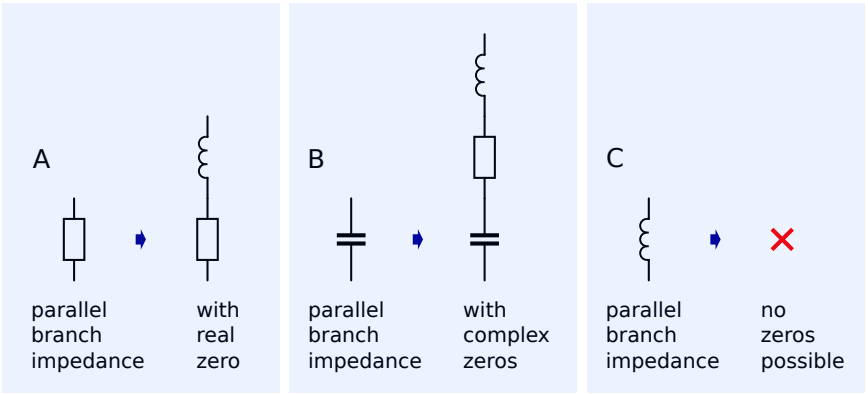


Figure 12.22: Implementation of zeros in parallel branch impedances.

(a) If a parallel branch behaves resistively at the frequency of the phantom zero, an inductive branch placed in series will cause a negative real zero in the transfer of the feedback network (see Figure 12.22A).

(b) If a parallel branch behaves capacitively at the frequency of the phantom zero, a resistive branch placed in series will cause a negative real zero and an inductive branch placed in series will cause two (complex) zeros in the transfer of the feedback network (see Figure 12.22B).

(c) If a parallel branch behaves inductively at the frequency of the phantom zero, then there is no means for creating a zero in this branch (see Figure 12.22C).

In Figure 12.23, the above method has been applied for implementing phantom zeros in the elementary single-loop resistive feedback amplifiers. The zeros in the transfers of the feedback networks cause poles in the asymptotic gain and zeros in the loop gain.

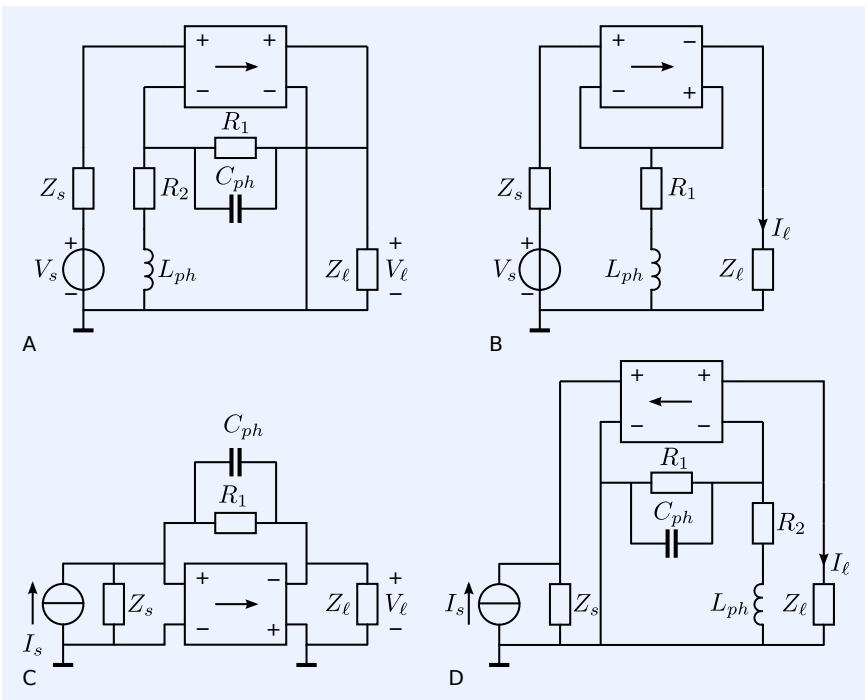


Figure 12.23: Realization of phantom zeros in the feedback network of the basic single-loop resistive feedback amplifier configurations.
A. Two phantom zeros in the voltage amplifier
B. One phantom zero in the transconductance amplifier
C. One phantom zero in the transresistance amplifier
D. Two phantom zeros in the current amplifier

Effectiveness of phantom zeros in the feedback network

A phantom zero in the feedback network is effective , if the branch of the feedback network in which the zero will be implemented causes a significant attenuation in the loop gain at the frequency of the zero. The zero will then increase the loop gain by reducing this attenuation for frequencies above the frequency of the zero, while it causes loop gain to become zero at the (complex) frequency of the zero. This will be elucidated with the aid of Figure 12.24.

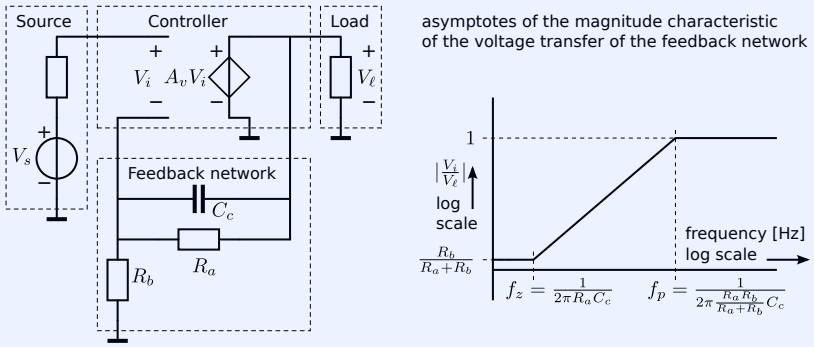


Figure 12.24: Left: small-signal diagram of a passive-feedback voltage amplifier showing the signal source, the load, the feedback network and the controller. A phantom zero has been implemented in the series branch of the feedback network.

Right: the magnitude characteristic of the transfer of the feedback network. The phantom zero establishes an increase of the loop gain for frequencies above f_z and below f_p. In this frequency range, it changes the order of the transfer from 0 to -1 (first order differentiating) and is called *effective*.

Figure 12.24 shows the small-signal diagram of a passive-feedback voltage amplifier as well as the magnitude characteristic of the transfer of the feedback network. If we select the gain A_v of the voltage-controlled voltage source as the loop gain reference variable, the loop gain is proportional to the voltage transfer of the feedback network. This transfer has one zero $z = -\frac{1}{R_a C_b}$, and one pole $p = -\frac{R_a+R_b}{R_a R_b C_c}$. This figure shows that a nonzero value of C_c reduces the loop gain attenuation caused by the resistive voltage divider. The frequency range over which the phantom zero is effective increases with the attenuation of the voltage divider: $\frac{f_z}{f_p} = \frac{R_b}{R_a+R_b}$.

In the following example, we will apply the above methods for frequency compensation of the transimpedance amplifier from example 11.3.

Example 12.8

Figure 12.25 shows the small-signal equivalent diagram of the transimpedance amplifier from example 11.3 with phantom zero compensation. The implementation of the phantom zero is as discussed in Figure 12.21A. Other implementations will be discussed later.

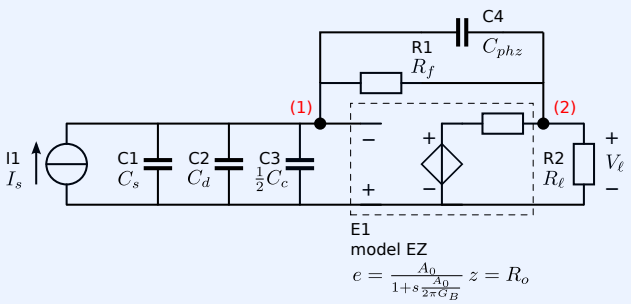


Figure 12.25: Circuit diagram of the transimpedance amplifier from example 7.11.3 with phantom zero compensation.

Below is the SLiCAP *netlist of the circuit in which the initial value of the compensation capacitance C_{phz} has been set to zero:*

```
1 transimpedanceCompensated
2 * file: transimpedanceCompensated.cir
3 * SLiCAP circuit file
```

```
I1 0 1 0
C1 1 0 {C_s}
C2 1 0 {C_d}
C3 1 0 {C_c/2}
C4 1 2 {C_phz}
C5 2 0 {C_ell}
R1 1 2 {R_f}
R2 2 0 {R_ell}
E1 2 0 0 1 EZ value={A_0/(1+s/2/pi/16)} zo={R_o}
.param C_d=8p C_c=7p A_0=1M R_o=55 G_B=16M
.param C_s=5p R_f=100k R_ell=2k C_ell=0 C_phz=0
.end
```

Lines 1 to 20 of the SLiCAP *script calculate the DC value of the loop gain, as well as the poles and zeros:*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: transimpedanceCompensated.py

from SLiCAP import *

fileName = 'transimpedanceCompensated'
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName + '.cir')
htmlPage('Phantom zero compensation')
i1.setSource('I1')
i1.setDetector('V_2')
i1.setLGref('E1')
i1.setSimType('numeric')
i1.setGainType('loopgain')
i1.setDataType('pz')
result = i1.execute()
# Display the DC value and poles and zeros of the loop gain
listPZ(result)
```

The results below will be shown in the Python *command window (instruction in line 20):*

```
DC value of loopgain: -9.73e+5

Poles of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.60e+1        0.00e+00         1.60e+1
 1        -9.64e+4        0.00e+00         9.64e+4

Found no zeros.
```

This data will be used to calculate the value of the compensation capacitance. Expression (12.22) is used to calculate C_{phz}. Lines 21 to 34 perform this calculation and display the value of C_{phz} in the Python *command window:*

```
polesLG = result.poles
# Extract data for compensation
p_1 = polesLG[0]/2/sp.pi
p_2 = polesLG[1]/2/sp.pi
L_0 = result.DCvalue
# Calculate achievable bandwidth
Bw = sp.sqrt(sp.Abs((1-L_0)*p_1*p_2))
R_f = i1.getParValue('R_f')
# Calculate compensation capacitance
C_phz = sp.N((sp.sqrt(2)*Bw+p_1+p_2)/R_f/Bw**2/2/sp.pi)
# Pass the value to the circuit
i1.defPar('C_phz', C_phz)
# Print the value of C_phz
print('C_phz = {:9.2e}F\n'.format(C_phz))
```

The output shows:

```
C_phz =  1.73e-12F
```

Lines 35-37 of the script display the data of the loop gain after compensation:

```
35 # Display the DC value and poles and zeros of the loop gain
36 LGpz = i1.execute()
37 listPZ(LGpz)
```

The output shows a third pole and one zero added to the loop gain. The zero is the intended phantom zero and the third pole arises because the compensation capacitor adds one independent capacitor voltage to the network.

```
14 DC value of loopgain: -9.73e+5
15
16 Poles of loopgain:
17
18  n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
19 --  --------------  --------------  --------------  --------
20  0        -1.60e+1        0.00e+00         1.60e+1
21  1        -8.72e+4        0.00e+00         8.72e+4
22  2        -1.89e+9        0.00e+00         1.89e+9
23
24 Zeros of loopgain:
25
26  n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
27 --  --------------  --------------  --------------  --------
28  0        -9.17e+5        0.00e+00         9.17e+5
```

The location of the third pole can be found from network inspection.[9] *It can be calculated from the time constant constituted by R_o in parallel with the series connection of C_{phz} and the parallel connection of C_s, C_d and $C_c/2$:*

[9] Please study section 18.5.3 if you are not familiar with the estimation of poles in networks.

$$p_3 \approx -\frac{C_{phz}+C_s+C_d+\frac{1}{2}C_c}{2\pi R_o C_{phz}\left(C_s+C_d+\frac{1}{2}C_c\right)} = -1.85GHz. \tag{12.70}$$

Lines 38-41 of the script display the properties of the gain *in the* PYTHON *command window:*

```
44 # Display the DC value and poles and zeros of the gain
45 i1.setGainType('gain')
46 Gpz = i1.execute()
47 listPZ(Gpz)
```

The result shows three poles and two zeros:

```
30 DC value of gain: -1.00e+5
31
32 Poles of gain:
33
34  n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
35 --  --------------  --------------  --------------  --------
36  0        -7.84e+5        -8.63e+5         1.17e+6   7.43e-1
37  1        -7.84e+5         8.63e+5         1.17e+6   7.43e-1
38  2        -1.89e+9        0.00e+00         1.89e+9
39
40 Zeros of gain:
41
42  n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
43 --  --------------  --------------  --------------  --------
44  0         1.63e+8        0.00e+00         1.63e+8
45  1        -1.64e+8        0.00e+00         1.64e+8
```

The three poles are as expected: two dominant poles in approximate MFM positions and the third non-dominant pole. The two zeros are caused by the nonzero direct transfer. This can be seen from the magnitude plots of the transfer functions of the asymptotic gain model (see Figure 12.26). At high frequencies the gain curve becomes equal to that of the direct transfer. This transition causes the two zeros in the gain: the order decreases from 2 *to* 0.

In the above example, we demonstrated the compensation of the transimpedance amplifier with the aid of a phantom zero in the feedback network. This phantom zero is very effective: it introduces a new pole which is not at all dominant. This is because before compensation, the feedback

resistor introduced a large attenuation in the loop gain at the frequency of the phantom zero.

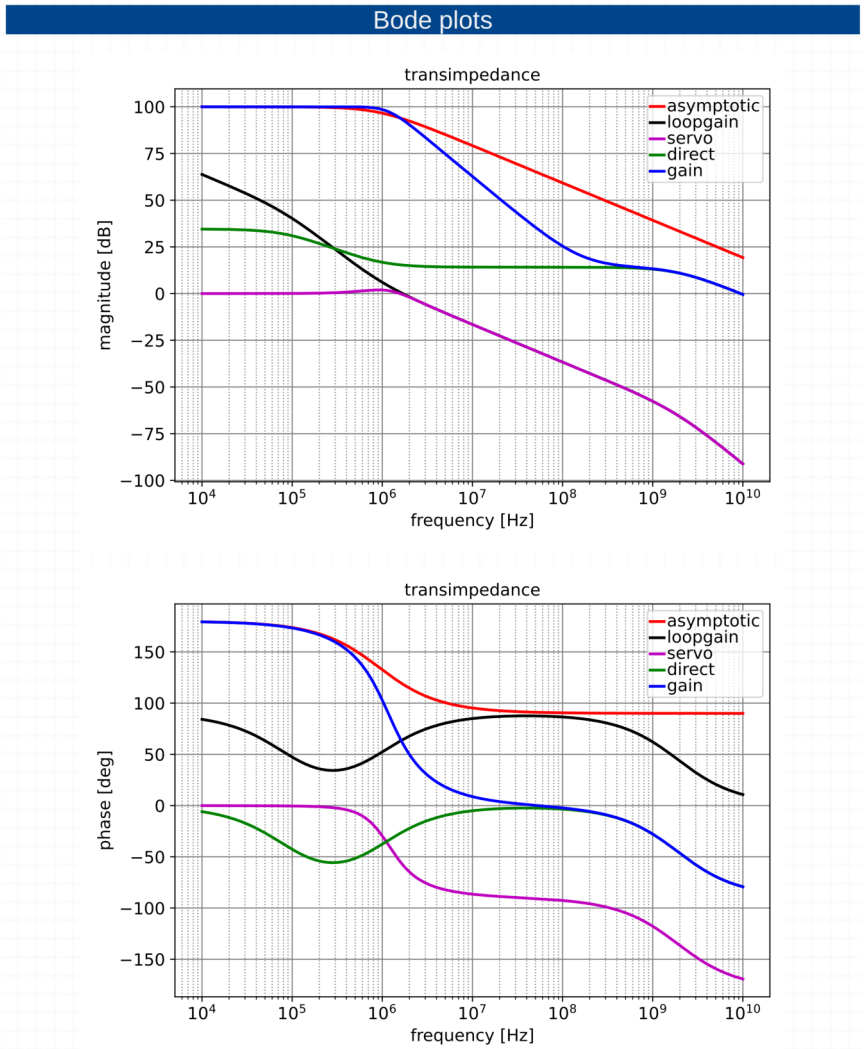


Figure 12.26: Magnitude and phase plots of the transfer functions of the asymptotic gain model for the transimpedance amplifier with phantom zero compensation.

12.2.6 Phantom zeros at source and load

A pole in the asymptotic gain can also be established by inserting a low-pass transfer between the source and the input of the amplifier or between the output of the amplifier and the load. The source impedance or the load impedance can be included in this low-pass filter. An effective phantom zero is then established if this filter also establishes a zero in the loop gain without causing a significant change of the loop gain poles product.

Phantom zeros at source or load with shunt feedback

In the case of parallel feedback at an amplifier port, hence with voltage sensing at the output port or with current comparison at the input port (see Figure 12.27), a phantom zero can be implemented by inserting an impedance Z_{se} in series with the source or load.

The character of Z_{se} depends on the behavior of the source or load impedance at the frequency of the phantom zero:

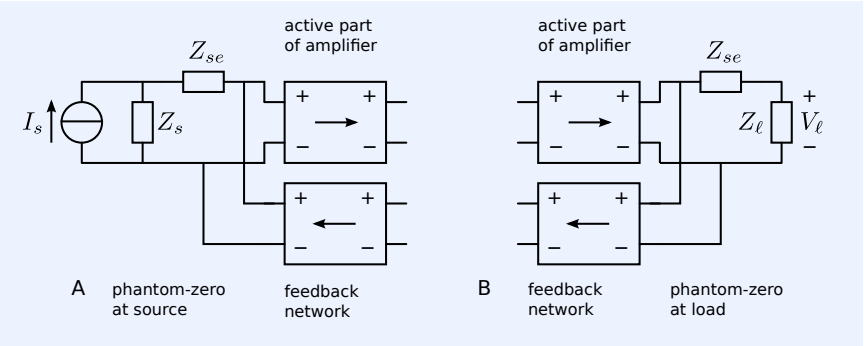


Figure 12.27: Implementation of phantom zeros for shunt feedback

1. If, at that frequency, the source or load impedance character is inductive, there is no possible implementation for Z_{se}
2. If, at that frequency, the source or load impedance character is resistive, Z_{se} must be inductive
3. If, at that frequency, the source or load impedance character is capacitive, Z_{se} must be resistive. Two real or complex conjugated phantom zeros can be implemented by having both an inductive and a resistive part for Z_{se}.

The pole in the asymptotic gain causes an effective phantom zero if, at the frequency of this pole, the source impedance or load impedance causes a significant attenuation in the loop gain. In such cases, the insertion of Z_{se} reduces this attenuation for frequencies above the frequency of this pole. The phantom zero frequency is the complex frequency at which $Z_{se} + Z_s = 0$ or $Z_{se} + Z_\ell = 0$. In fact, at this frequency, the series connection of Z_{se} and Z_s or the series connection of Z_{se} and Z_ℓ introduces a short in parallel with the signal path of the loop gain.

Phantom zeros at source or load with series feedback

In the case of series feedback at the amplifier port, hence with current sensing at the output port or voltage comparison at the input port (as shown in Figure 12.28), the phantom zero can be implemented with an impedance Y_p in parallel with the source or load.

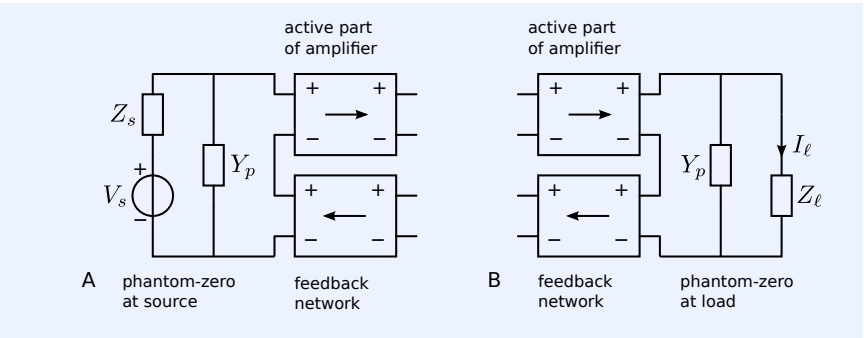


Figure 12.28: Phantom zero implementation at series feedback.

The character of Y_p depends on the behavior of the source or load impedance at the frequency where the phantom zero must be effective:

1. If, at that frequency, the source or load impedance character is capacitive, there is no possible implementation for Y_p
2. If, at that frequency, the source or load impedance character is resistive, Y_p must be capacitive

3. If, at that frequency, the source or load impedance character is inductive, Y_p must be resistive. Two real or complex conjugated phantom zeros can be implemented by having both an inductive and a capacitive part for Y_p.

The pole in the asymptotic gain causes an effective phantom zero if, at the frequency of this pole, the source impedance or load impedance causes a significant attenuation in the loop gain. In such cases, the insertion of Y_p reduces this attenuation for frequencies above the frequency of this pole. The phantom zero frequency is the complex frequency at which $Y_p + 1/Z_s = 0$ or $Y_p + 1/Z_\ell = 0$. In fact, at this frequency, the parallel connection of Y_p and Z_s or the series connection of Y_p and Z_ℓ introduces an open circuit in series with the signal path of the loop gain.

Effectiveness of phantom zeros at source or load

Creation of a zero in the loop with the aid of passive networks implies that an attenuation in the loop gain is reduced for frequencies above that of the zero. This has already been elucidated in Figure 12.24.

Hence, if the zero is realized by inserting a resistive element, it will shift an existing pole of the loop gain towards a higher frequency at which the reduction of the attenuation, as it was caused by the zero, stops.

If the zero is realized through insertion of an inductive or capacitive element, a pole may be added to the loop transfer function at a frequency at which the reduction of the attenuation, as it was caused by the zero, stops.

If the new pole is too close to the phantom zero, the phantom zero is called *ineffective*.

In the following example, we will try to compensate the transimpedance amplifier from example 11.3 with the aid of a phantom zero at the source. We will show that this phantom zero implementation is less effective than the one from the previous example.

Example 12.9

Figure 12.29 shows the small-signal equivalent diagram of the transimpedance amplifier from example 11.3 with phantom zero compensation at the source. The phantom zero has been implemented according to the concept from Figure 12.27A. Since the source is capacitive, a resistor has been inserted between the source and the input of the transimpedance amplifier.

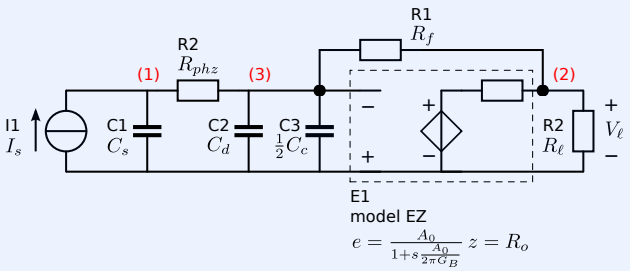


Figure 12.29: Circuit diagram of the transimpedance amplifier from example 7.11.3 with phantom zero compensation at the source.

Below is the SLiCAP *netlist of the circuit in which the initial value of the compensation capacitance C_{phz} has been set to zero:*

```
transimpedanceCompensatedSource
* file: transimpedanceCompensatedSource.cir
* SLiCAP circuit file
I1 0 1 0
C1 1 0 {C_s}
C2 3 0 {C_d}
C3 3 0 {C_c/2}
R1 3 2 {R_f}
```

```
R2 2 0 {R_ell}
R3 1 3 r value={R_phz}
E1 2 0 0 3 EZ value={A_0/(1+s*A_0/2/pi/G_B)} zo={R_o}
.param C_s=5p R_f=100k R_ell=2k C_d=8p C_c=7p A_0=1M R_o=55 G_B=16M R_phz=0
.end
```

Lines 1-20 of the script below list the poles and zeros of the transfer in the PYTHON *command window.*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: transimpedanceCompensatedSource.py

from SLiCAP import *

fileName = 'transimpedanceCompensatedSource'
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName + '.cir')
htmlPage('Phantom zero compensation')
i1.setSource('I1')
i1.setDetector('V_2')
i1.setLGref('E1')
i1.setSimType('numeric')
i1.setGainType('loopgain')
i1.setDataType('pz')
result = i1.execute()
# Display the DC value and poles and zeros of the loop gain
listPZ(result)
```

Expression (12.22) is used to calculate R_{phz}. Lines 21 to 34 perform this calculation and display the value of R_{phz} in the PYTHON *command window:*

```
polesLG = result.poles
# Extract data for compensation
p_1 = polesLG[0]/2/sp.pi
p_2 = polesLG[1]/2/sp.pi
L_0 = result.DCvalue
# Calculate achievable bandwidth
Bw = sp.sqrt(sp.Abs((1-L_0)*p_1*p_2))
C_s = i1.getParValue('C_s')
# Calculate compensation resistance
R_phz = sp.N((sp.sqrt(2)*Bw+p_1+p_2)/C_s/Bw**2/2/sp.pi)
# Pass the value to the circuit
i1.defPar('R_phz', R_phz)
# Print the value of R_phz
print('R_phz = {:9.2e} Ohm\n'.format(R_phz))
```

The output shows:

```
DC value of loopgain: -9.73e+5

Poles of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.60e+1        0.00e+00         1.60e+1
 1        -9.64e+4        0.00e+00         9.64e+4

Found no zeros.

R_phz =   3.47e+4 Ohm
```

Lines 35-37 of the script display the data of the loop gain after compensation:

```
# Display the DC value and poles and zeros of the loop gain
LGpz = i1.execute()
listPZ(LGpz)
```

The output shows a third pole and one zero added to the loop gain. The zero is the intended phantom zero and the third pole arises because the compensation resistor adds one independent capacitor voltage to the network. It removes the loop of the two capacitors from the circuit.

```
DC value of loopgain: -9.73e+5
```

```
Poles of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.60e+1        0.00e+00        1.60e+1
 1        -9.32e+4        0.00e+00        9.32e+4
 2        -1.36e+6        0.00e+00        1.36e+6

Zeros of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -9.17e+5        0.00e+00        9.17e+5
```

The location of the third pole can be found from network inspection.[10] *It can be calculated from the time constant constituted by R_{phz} in parallel with the series connection of C_s and the parallel connection of C_d and $C_c/2$:*

$$p_3 \approx -\frac{C_s + C_d + \frac{1}{2}C_c}{2\pi R_{phz} C_s \left(C_d + \frac{1}{2}C_c\right)} = -1.32MHz. \tag{12.71}$$

This is a dominant pole that limits the effectiveness of the phantom zero.

[10] Please study section 18.5.3 if you are not familiar with the estimation of poles in networks.

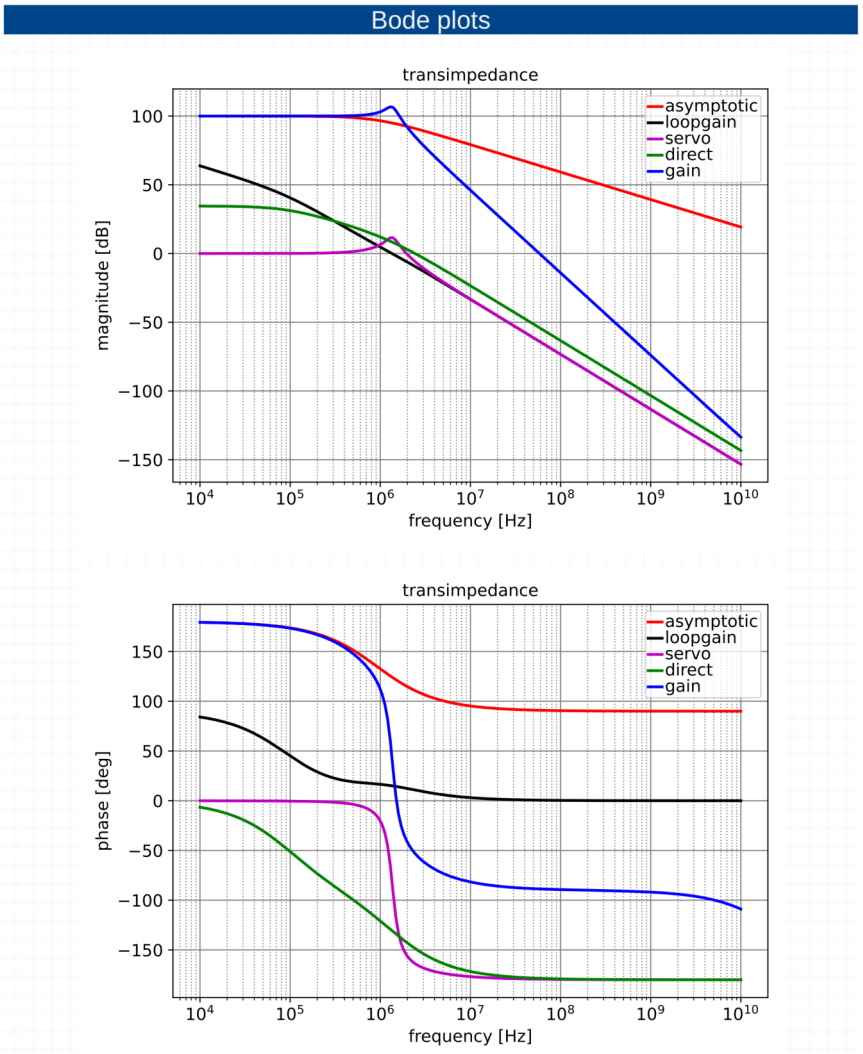


Figure 12.30: Magnitude plots of the transfer functions of the asymptotic gain model for the transimpedance amplifier with phantom zero compensation at the source.

Lines 38-41 of the script display the properties of the gain in the PYTHON

command window:

```
# Display the DC value and poles and zeros of the gain
i1.setGainType('gain')
Gpz = i1.execute()
listPZ(Gpz)
```

The result shows three poles and one zero:

```
DC value of gain: -1.00e+5

Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.06e+6        0.00e+00         1.06e+6
 1        -1.97e+5        -1.35e+6         1.36e+6   3.46e+0
 2        -1.97e+5         1.35e+6         1.36e+6   3.46e+0

Zeros of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        2.91e+10        0.00e+00        2.91e+10
```

The three poles are as expected: three dominant poles that are not in MFM positions due to the ineffective compensation. The zero is caused by the nonzero direct transfer. This can be seen from the Bode plots of the transfer functions of the asymptotic gain model (see Figure 12.30). At very high frequencies (not visible on the plot), the gain curve becomes equal to that of the direct transfer. This transition causes the zero in the gain: the order decreases from 3 to 2.

In the previous example, we demonstrated the compensation of the transimpedance amplifier with the aid of a phantom zero at the source. In this case, this compensation is not very effective: the implementation of the phantom zero introduces a new dominant pole. MFM compensation with C_{phz} in parallel with R_f as discussed in the previous example turns out to be far more effective.

In the following example, we will demonstrate the frequency compensation of a low-noise voltage reference that can be used as reference for an AD converter.

Many AD converters with an unbuffered reference input draw a high-frequency switching current from their reference. In order to maintain a low-ripple reference voltage under these conditions, a decoupling capacitor needs to be placed between the reference voltage input of the ADC and ground.

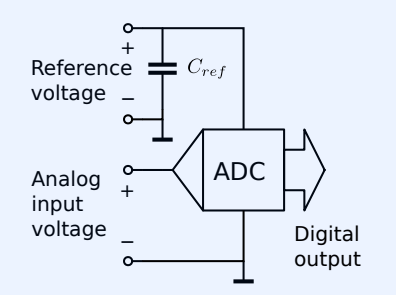


Figure 12.31: ADC with single-ended signal input and single-ended external reference input.

Example 12.10

In this example, we will study the frequency compensation of a voltage driver of a capacitive load. This driver is intended to drive the reference input of an AD converter with unbuffered reference input that has been decoupled with a large capacitance C_{ref}. This situation is illustrated in Figure 12.31.

We will assume that a DC reference for the ADC with the correct DC voltage is available, but the associated noise voltage of this reference is too large. The noise performance can then be improved by placing a low-pass filter cascaded with a unity-gain voltage buffer between the voltage reference and the ADC reference input. This is shown in Figure 12.32. In this example, we will only focus on the frequency compensation of the voltage buffer, and particularly on phantom zero compensation. We will simply assume $C_{ref} = 10\mu F$, $C_f = 1\mu F$, $R_f = 15k\Omega$ and a voltage buffer that has been designed with an `AD8610` operational amplifier. This operational amplifier has a resistive output impedance.

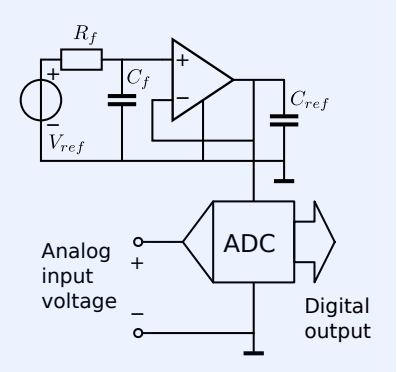


Figure 12.32: Reference voltage with noise filter and unity-gain buffer.

SLiCAP *has two models for this operational amplifier in the* `SLiCAP.lib` *file:*

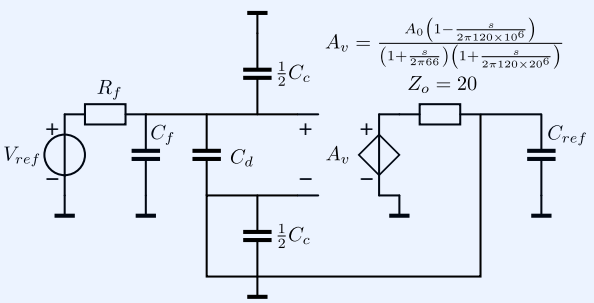
```
.model AD8610 OV
+ cd = 15p  ; differential-mode input capacitance
+ cc = 8p   ; common-mode input capacitance
```

```
+ av = {300k*(1-s/2/pi/120M)/(1+s*300k/2/pi/20M)/(1+s/2/pi/120M)} ; voltage
    gain
+ zo = 20   ; output impedance

.model AD8610_A0 OV
+ cd = 15p  ; differential-mode input capacitance
+ cc = 8p   ; common-mode input capacitance
+ av = {A_0*(1-s/2/pi/120M)/(1+s*300k/2/pi/20M)/(1+s/2/pi/120M)} ; voltage
    gain
+ zo = 20   ; output impedance
```

The model `AD8610_A0` *can be used for sweeping the DC gain* A_0 *of this opamp to generate root-locus plots. The model* `AD8610` *does not contain symbolic parameters. Figure 12.33 shows the complete small-signal circuit of the buffer.*

Figure 12.33: Small-signal equivalent circuit of the voltage reference circuit.



The gain of the `AD8610` *has been modeled with the aid of an expression for* A_v*. It comprises a DC gain* $A_0 = 300 \times 10^3$*, two poles and one zero. The output impedance* Z_o *equals* 20Ω*.*

The poles and zeros of the loop gain are listed below.[11]

[11] The voltage gain A_v of the operational amplifier has been selected as loop gain reference variable.

```
Poles of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.06e+1        0.00e+00        1.06e+1
 1        -6.67e+1        0.00e+00        6.67e+1
 2        -7.96e+2        0.00e+00        7.96e+2
 3        -1.20e+8        0.00e+00        1.20e+8
```

The loop gain has four poles: two in A_v *and two from independent capacitor voltages.*[12] *The pole positions can be estimated with the aid of the time constant matrix, as explained in Chapter 18.5.3.*

[12] The poles are obtained with the data type set to `'poles'`.

#	time constant / transfer	value [Hz]
p_1	$\frac{-1}{2\pi\tau_1}$, $\tau_1 \approx C_f R_f$	-10.61
p_2	pole in A_v	-66
p_3	$\frac{-1}{2\pi\tau_3}$, $\tau_3 \approx C_{ref} Z_o$	-796
p_4	pole in A_v	$-120 \cdot 10^6$

The estimated values match those obtained from the SLiCAP *calculations. The loop gain has two zeros:*[13]

[13] The zeros are obtained with the data type set to `'zeros'`.

```
Zeros of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.06e+1        0.00e+00        1.06e+1
 1         1.20e+8        0.00e+00        1.20e+8
```

The zeros can be estimated as follows:

#	time constant network / transfer	value [Hz]
z_1	zero in A_v	$120 \cdot 10^6$
z_2	$\frac{-1}{2\pi\tau_2}$, $\tau_2 \approx C_f R_f$	-10.61

The estimated values comply with those obtained by SLiCAP.

Within the displayed accuracy, z_2 cancels p_1. With the SLiCAP *data type set to* `'pz'`, *p_1 and z_2 are no longer shown, but the DC loop gain will be shown:*

```
DC value of loopgain: -3.00e+5

Poles of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -6.67e+1        0.00e+00        6.67e+1
 1        -7.96e+2        0.00e+00        7.96e+2
 2        -1.20e+8        0.00e+00        1.20e+8

Zeros of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0         1.20e+8        0.00e+00        1.20e+8
```

The servo bandwidth can be calculated from p_1 and p_2 and the DC value of the loop gain. It can also be obtained from SLiCAP *using the* `findServoBandwidth()` *function. Both methods show that a second order MFM low-pass cut-off of the servo function can be obtained at $f_\ell = 125.5kHz$.*

The poles of the gain are in the left half plane, but the quality factor is very high and frequency compensation is required.

```
Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.06e+1        0.00e+00        1.06e+1
 1        -2.99e+2       -1.26e+5        1.26e+5   2.11e+2
 2        -2.99e+2        1.26e+5        1.26e+5   2.11e+2
 3        -1.20e+8        0.00e+00        1.20e+8
```

Since the two dominant poles are relatively close to the origin, the frequency ω_z of the phantom zero can be approximated as:

$$\omega_z = -2\pi f_\ell \frac{1}{2}\sqrt{2}. \tag{12.72}$$

A negative real phantom zero can be implemented at the load. To this end, we need to insert a resistor between the output of the amplifier and the load. Its resistance R_{phz} should equal

$$R_{phz} = \frac{1}{-\omega_z C_{ref}} \approx \frac{1}{\sqrt{2}\pi f_\ell C_{ref}} = 0.18\Omega. \tag{12.73}$$

Figure 12.34A shows the compensated driver.

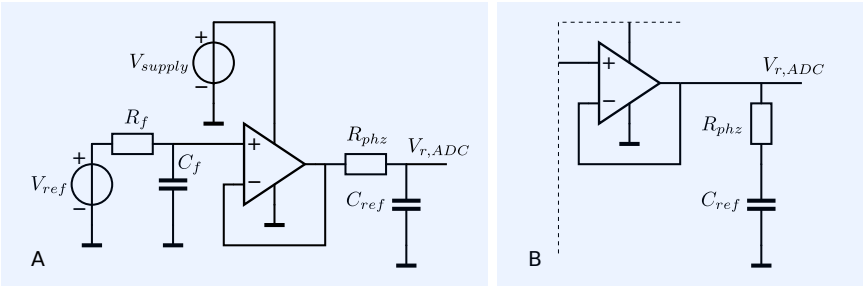


Figure 12.34:
A: compensated voltage reference circuit
B: modified compensation; the ESR of C_{ref} is used for compensation. The zero is not longer a phantom zero: it also appears in the gain.

The dominant poles of the compensated amplifier are in MFM positions while the non-dominant poles all have negative real values. Please notice that the insertion of the phantom zero resistor created a new independent capacitor voltage, and thus a new pole:

```
DC value of gain:  1.00e+0
```

```
Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.06e+1        0.00e+00        1.06e+1
 1        -8.98e+4        -8.81e+4        1.26e+5   7.01e-1
 2        -8.98e+4         8.81e+4        1.26e+5   7.01e-1
 3        -1.20e+8        0.00e+00        1.20e+8
 4       -4.70e+10        0.00e+00       4.70e+10

Zeros of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -8.84e+4        0.00e+00        8.84e+4
 1         4.16e+7        -6.74e+7        7.92e+7   9.53e-1
 2         4.16e+7         6.74e+7        7.92e+7   9.53e-1
 3        -2.03e+8        0.00e+00        2.03e+8
```

If the ADC reference takes a DC current from this buffer, the phantom zero resistor shown in Figure 12.34A may introduce a DC inaccuracy.

This DC error will not occur if the reference output is taken directly at the output of the operational amplifier, as it is shown in Figure 12.34B. In this circuit, the resistance R_{phz} introduces a zero in the loop gain that is no longer a phantom zero. Hence, it also appears in the transfer from V_{ref} to the output voltage. In many cases, this minor penalty on the noise transfer is justified by the improvement of the DC transfer. As a matter of fact, in this way, R_{phz} can often be implemented using the ESR of the reference capacitor.[14] *Figure 12.35 shows the pole positions of the servo function as a function of R_{phz}. Please notice that the pole at $-10.61Hz$ is canceled by a zero. The zeros have not been shown in this plot.*

[14] ESR = Equivalent Series Resistance.

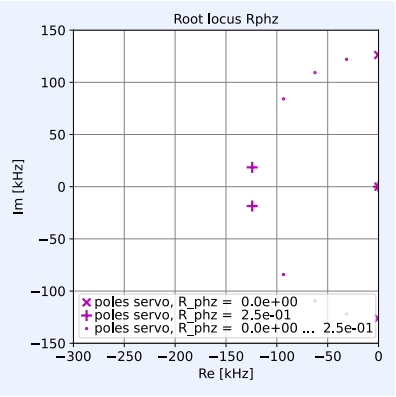


Figure 12.35: Pole positions of the servo function as a function of R_{phz}.

12.2.7 Active phantom zeros

Until now, we have studied implementation methods for phantom zeros based on modifications of existing passive attenuators in the loop transfer. However, the most straightforward implementation of phantom zeros, based on the concepts discussed in sections 12.2.1 through 12.2.3, uses differentiators in the feedback networks. Unfortunately, due to the fundamental limitation of speed, practical implementations of differentiators also add poles, thereby reducing the effectiveness of the frequency compensation. Only in cases in which these poles can be kept away from the dominant group might active phantom zero compensation be feasible. In this section, we will study frequency compensation with the aid of active phantom zeros. We will confine ourselves solely to compensation with a first order active differentiator.

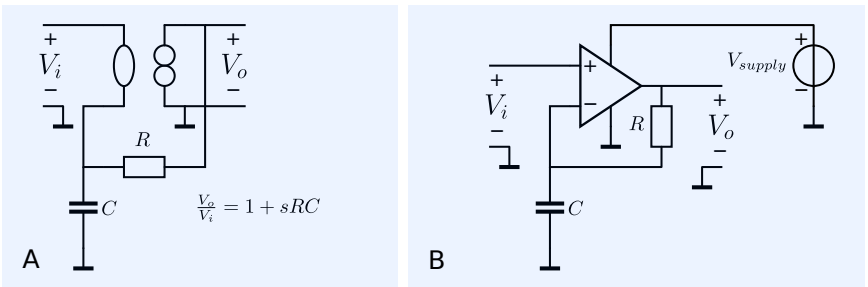


Figure 12.36:
A: Negative feedback first order differentiator concept.
B: implementation with an operational amplifier.

Figure 12.36A shows the concept of a first order voltage differentiator circuit. The circuit has unity DC gain and a zero: $z = -\frac{1}{RC}$. An implementation with the aid of an operational amplifier is shown in Figure 12.36B.

In the following example, we will apply this circuit for frequency compen-

sation of the ADC reference circuit from example 12.10.

Example 12.11

The circuit from Figure 12.34A has a phantom zero realized with a resistor R_{phz} between the output of the amplifier and the load. A disadvantage of this implementation is the increase of the low-frequency output resistance of the reference buffer to R_{phz}. This reduces the DC accuracy of the reference under nonzero DC load conditions. One possible solution to this is to use the arrangement shown in Figure 12.34B. This solution, however, has its high-frequency output impedance increased to R_{phz}. A solution that has both a low DC output resistance and a low high-frequency output impedance can be obtained with active phantom zero compensation.

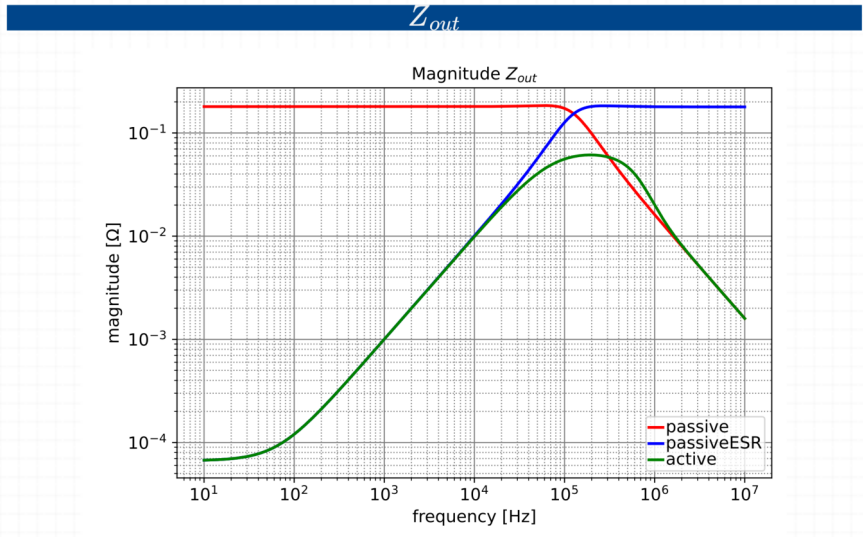


Figure 12.37: Magnitude plots of the output impedance for three different compensation methods:
Red: Passive compensation according to Figure 12.34A
Blue: Passive phantom zero compensation according to Figure 12.34B
Green: Active phantom zero compensation according to Figure 12.38.

Figure 12.38 shows the circuit of the reference buffer with active phantom zero compensation. The feedback network of this circuit consists of the differentiator with unity DC gain from Figure 12.36B. R_{phz} in Figure 12.38 implements a phantom zero in the differentiator. Figure 12.37 shows the magnitude characteristic of the output impedances of the three reference buffers. The following values have been used for the circuit from Figure 12.38: $R_d = 1k\Omega$, $C_d = 2.7nF$, $R_{phz} = 82\Omega$.

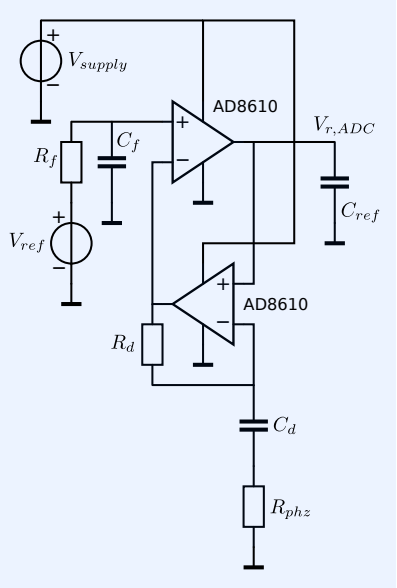


Figure 12.38: ADC reference buffer with active phantom zero compensation.

12.2.8 Interaction with other performance aspects

The realization of passive phantom zeros requires insertion of impedances in series or in parallel with the signal path. The general design rule was to avoid this, which may lead to the conclusion that passive phantom zeros may seriously deteriorate other performance aspects, such as the noise performance, the power efficiency and the linearity. Moreover, insertion of impedances into the signal path may also result in a reduction of the product of the midband loop gain and the dominant poles. Fortunately, implementation of phantom zeros seldom causes significant deterioration of performance aspects. This can be understood as follows:

1. Phantom zeros are usually located at the edge of the transmission band of the amplifier. We have seen that second order systems and third order systems that have all poles in the origin require the magnitude of the frequency of their phantom zeros at $\frac{1}{2}\sqrt{2}\omega_n$. Below this frequency, the impedance inserted into the signal path for the realization of the phantom zeros plays no role compared to existing impedances.

2. If phantom zeros are implemented at the source, they constitute a low-pass filter between the signal source and the amplifier. Such a low-pass filter is generally beneficial to the IMFDR since it suppresses out-of-band interference signals.[15]
3. Phantom zeros increase the loop gain above the frequency of the zero. This reduces the nonlinearity of the amplifier at those frequencies.

[15] IMFDR = InterModulation-Free Dynamic Range.

12.2.9 Bandwidth limitation with phantom zeros

In the previous sections, we discussed the frequency compensation of amplifiers with the aid of phantom zeros. With phantom zero compensation, we are able to obtain a well-defined frequency response of a feedback amplifier without paying a significant penalty on other performance aspects.

We will now discuss the application of phantom zeros for limitation of the bandwidth of the source-load transfer. Such bandwidth limitation may be desired if the bandwidth of the servo function is much larger than required. This can be the case if a low distortion drives the specification for a high value of the midband loop gain.

Bandwidth limitation with phantom zeros occurs in negative feedback amplifiers with phantom zeros if the loop gain at the frequency of the phantom zeros is much larger than unity.

The method

A phantom zero is a pole in the asymptotic gain that coincides with a zero in the loop gain. As a consequence, a phantom zero cannot be observed in the source-load transfer. However, the root locus technique shows us that if, at the frequency of the zero, the magnitude of the loop gain is much larger than unity, one of the poles of the servo function will move towards it. This pole is visible in the source-load transfer and limits the bandwidth of the source-load transfer to a frequency that approximates that of the zero.

We will elucidate this with the aid of an example in which we will limit the bandwidth of the transimpedance amplifier from example 12.8 with the aid of a phantom zero to 200kHz.

Example 12.12

In example 12.8, we demonstrated the compensation of a transimpedance amplifier with the aid of a phantom zero. The phantom zero has been implemented with the aid of a capacitance C_{phz} in parallel with the feedback resistor (see Figure 12.25). We will now use C_{phz} to limit the bandwidth B of the transimpedance amplifier to 200kHz. To do so, we need to increase its value to $\frac{1}{2\pi B R_f} = 7.96pF$.

The results of the pole-zero analysis of the transfer functions of the asymptotic gain model are shown below:

```
=== Bandwidth limitation to 200kHz ===
DC value of asymptotic: -1.00e+5

Poles of asymptotic:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -2.00e+5        0.00e+00         2.00e+5

Found no zeros.

DC value of loopgain: -9.73e+5

Poles of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.60e+1        0.00e+00         1.60e+1
```

```
 1        -6.51e+4        0.00e+00         6.51e+4
 2        -5.54e+8        0.00e+00         5.54e+8

Zeros of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -2.00e+5        0.00e+00         2.00e+5

DC value of servo:  1.00e+0

Poles of servo:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -2.06e+5        0.00e+00         2.06e+5
 1        -4.97e+6        0.00e+00         4.97e+6
 2        -5.49e+8        0.00e+00         5.49e+8

Zeros of servo:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -2.00e+5        0.00e+00         2.00e+5

DC value of gain: -1.00e+5

Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -2.06e+5        0.00e+00         2.06e+5
 1        -4.97e+6        0.00e+00         4.97e+6
 2        -5.49e+8        0.00e+00         5.49e+8

Zeros of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0         7.62e+7        0.00e+00         7.62e+7
 1        -7.64e+7        0.00e+00         7.64e+7
```

Lines 48-56 show the data of the asymptotic gain. As expected, the asymptotic gain has a DC value of 10^5*, a pole at* $-200kHz$*, and no zeros.*

The results for the loop gain are listed in lines 58-72. The loop gain has a zero at $-200kHz$*; this is the phantom zero, and three poles, of which* p_1 *and* p_2 *are dominant.*

Lines 74-88 show the results for the servo function. The servo function has the zero of the loop gain and three poles that can be obtained from the poles of the loop gain using the root locus method. The frequency of p_1 *approximately equals the frequency of the zero in the loop gain. The root locus (see Figure 12.39) shows it approaches the zero from the left side. Since the frequency of* p_1 *approximately equals that of* z_1*, the servo function can be approximated as a first order low-pass function with a* $-3dB$ *frequency of* $4.9706MHz$*. This is about five times larger than that of the transimpedance amplifier from example 12.8.*

Lines 90 to 105 list the data of the gain. The phantom zero at $-200kHz$ *is not visible and the dominant pole equals* p_1 *of the servo function. This pole limits the bandwidth of the transimpedance amplifier.*

The magnitude characteristics of the transfer functions of the asymptotic gain model of the compensated amplifier are shown in Figure 12.40. Comparing them with those from Figure 12.26, clearly shows the reduced bandwidth of the gain is reduced, and the increased bandwidth of the servo. This is the result of the insertion of a zero in the loop gain at a frequency at which the magnitude of the loop gain is much larger than unity.

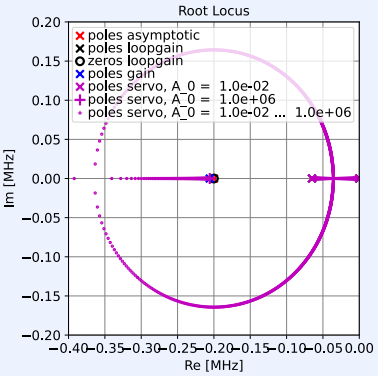


Figure 12.39: Root locus of the transimpedance amplifier with phantom zero bandwidth limitation.

This example shows that bandwidth limitation with a phantom zero is a powerful method. Above the frequency of the phantom zero the loop gain increases, which is beneficial to the linearity. If the reduction of the weak nonlinearity (intermodulation distortion) is of utmost importance, one might

consider bandwidth limitation with a phantom zero at the source. It reduces the influence of out-of-band interference on the intermodulation distortion in two ways:

1. The pole in the asymptotic gain reduces the amplitude of high-frequency interference

2. The zero in the loop gain reduces the high-frequency intermodulation distortion.

A disadvantage of limiting the bandwidth with a phantom zero is the increased influence of non-dominant poles. This is an inevitable result of the increase of the loop gain above the frequency of the phantom zero. If this results in unacceptable peaking or instability, frequency compensation with additional phantom zeros may be required.

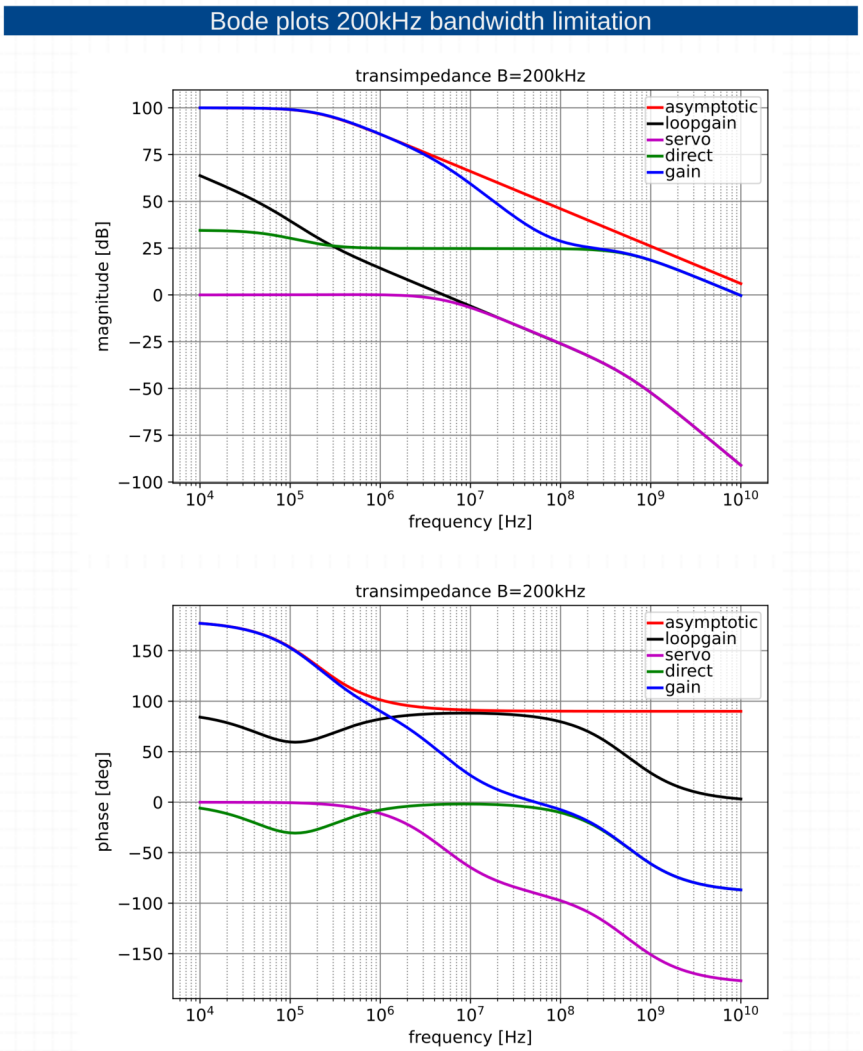


Figure 12.40: Bode plots of the transfer functions of the asymptotic gain model of the transimpedance amplifier with phantom zero bandwidth limitation.

12.3 Pole-splitting

If phantom zero compensation is not possible, frequency compensation may also be accomplished by changing the initial positions of the poles of the

loop gain. A compensation technique based on this principle is the application of local capacitive feedback around one or more gain stages in the controller. With this technique, the distance between two poles can be increased while their product remains virtually unaffected. This technique is called *pole-splitting*. Pole-splitting reduces the controller gain in the frequency range between the two poles. In fact, it trades controller gain with local loop gain in the gain stages enclosed in the capacitive feedback loop. Reduction of the controller gain may adversely affect the nonlinearity of the feedback amplifier. Since phantom zero compensation increases the controller gain, it is preferred over pole-splitting. Pole-splitting should only be applied if phantom zero compensation is not feasible.

12.3.1 Pole-splitting in operational amplifier circuits

Figure 12.41A shows a gain stage that is part of a controller. This gain stage consists of a single-pole operational amplifier driven from an *RC* network.

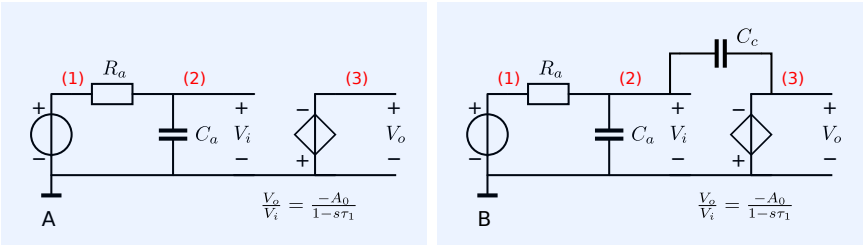


Figure 12.41:
A: Controller gain stage with a single-pole operational amplifier
B: Gain stage with pole-splitting.

The two poles of this circuit can be estimated from network inspection: $p_1 = -\frac{1}{R_a C_a}$ and $p_2 = -\frac{1}{\tau}$. These poles are the solutions of the characteristic equation.[16] The characteristic equation can be calculated symbolically by SLiCAP. The netlist file of the circuit from Figure 12.41A has is below:

[16] The determinant of the MNA matrix.

```
"Pole splitting with opamps"
* file: poleSplitOpamp.cir
* SLiCAP netlist file
V1 1 0 0
R1 1 2 {R_a}
C1 2 0 {C_a}
E1 0 3 2 0 {A_0/(1+s*tau)}
C2 2 3 {C_c}
.param C_c=0
.end
```

The script for calculating the poles (with $C_c = 0$) is:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: poleSplitOpAmp.py

from SLiCAP import *

fileName = 'poleSplitOpamp'
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName + '.cir')
i1.setSimType('numeric') # This substitutes C_c=0 in the element expressions
i1.setDataType('poles')
i1.setGainType('gain')
i1.setDetector('V_3')
poles = i1.execute().poles
print("Poles:",poles)
```

The result is shown in the PYTHON console window:

```
Poles: [-1/tau, -1/(C_a*R_a)]
```

This corresponds with the estimation above.

After adding C_c to the circuit as shown in Figure 12.41B, the characteristic equation of the network changes to:

$$1 + s\left(R_a\left(C_a + C_c\left(1 + A_0\right)\right) + \tau\right) + R_a\left(C_a + C_c\right)\tau s^2 = 0. \tag{12.74}$$

Insertion of C_c did not change the order of the characteristic equation. This is because C2 (with capacitance C_c) is part of a loop of capacitors and voltage sources. Hence, it does not introduce a new independent capacitor voltage.

Expression (12.74) shows if $C_c \ll C_a$ the product of the poles $\frac{1}{(C_a+C_c)\tau}$ does not change with C_a. However, if $C_c\left(1 + A_0\right) \gg C_a + \tau$, the magnitude of the sum of the poles increases considerably with C_c. This means that one of the poles must shift to a higher frequency while the other one moves to a lower frequency. This will be illustrated with a numeric example.

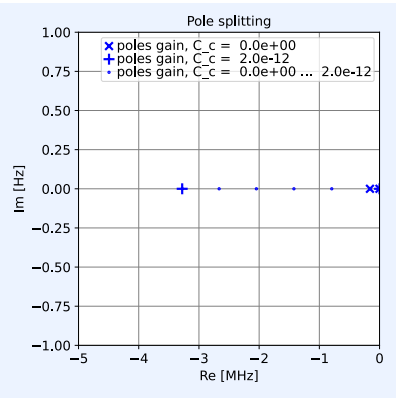


Figure 12.42: Pole-splitting in the gain stage from Figure 12.41. The pole with the highest frequency moves away from the origin while increasing C_c

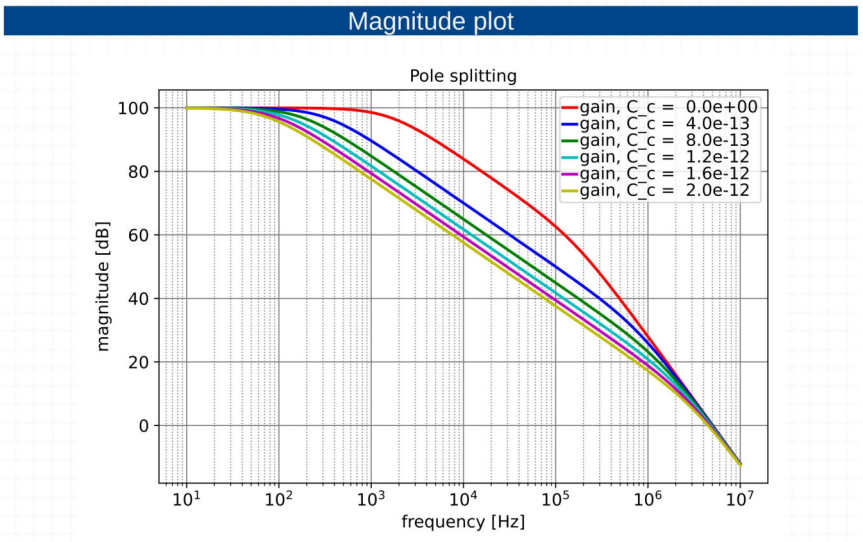
Example 12.13

We will plot the pole positions of the circuit from Figure 12.41B as a function of C_c. In this numerical example, we use: $C_a = 100pF$, $R_a = 10k\Omega$, $\tau = 1ms$ and $A_0 = 10^5$. Figure 12.42 shows the pole positions as a function of C_c. It can clearly be seen that the pole with the highest frequency moves to a higher frequency.

Figure 12.44 shows the low-frequency pole moving towards the origin for increasing C_c.

Pole-splitting can also clearly be observed in the voltage transfer of the network. Figure 12.43 shows the magnitude plots of the voltage transfer with C_c stepping from $0 \cdots 2pF$. At very high frequencies, all curves approach the same asymptote. This indicates that the product of the poles is not affected by C_c.

Figure 12.43: Pole splitting in the circuit from Figure 12.41B.



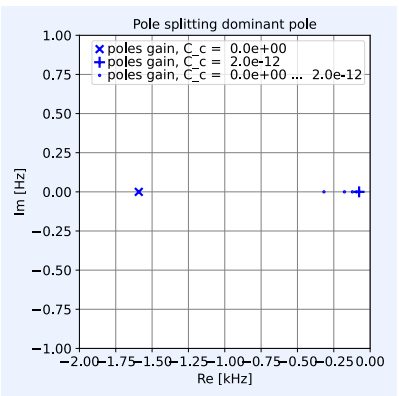


Figure 12.44: Pole splitting in the gain stage from Figure 12.41. The pole with the lowest frequency moves towards the origin while increasing C_c.

Undesired pole-splitting may occur due to parasitic capacitances that establish negative feedback around gain stages in a feedback amplifier. As a consequence, poles that move towards higher frequencies may no longer belong to the dominant group. This results in reduction of the achievable MFM bandwidth of the amplifier.

12.3.2 Miller effect

The earliest description of pole-splitting in amplifier stages dates from 1920.[Miller1920][17] In the honour of its discoverer, this phenomenon is often referred to as the *Miller effect*.

[17] John M. Miller. Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit. *Scientific Papers of the Bureau of Standards*, 15(351):367–385, 1920

Figure 12.45 shows the small-signal equivalent circuit of a current-driven basic amplification stage driving an RC load ($R_\ell \,/\!/\, C_\ell$). In a case of vacuum tubes and MOS transistors, the resistance r_i is extremely large and is usually left out of the model. If the reverse capacitance c_r equals zero, the poles of the transimpedance gain can easily be found from network inspection: $p_1 = -\frac{1}{r_i c_i}$ and $p_2 = -\frac{1}{R_\ell C_\ell}$.

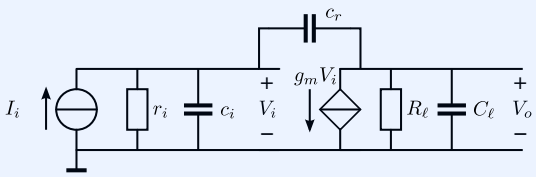


Figure 12.45: Pole-splitting in a basic amplifier stage caused by the reverse transfer capacitance c_r.

If $c_r \ll c_i$ and $c_r \ll C_\ell$, the reverse capacitance does not affect the product of the poles. If, in addition, $g_m R_\ell \gg 1$ and $g_m R_\ell c_r \gg c_i$, the sum of the poles increases considerably with c_r and pole-splitting occurs. This can be seen by solving the characteristic equation of the circuit.[18]

[18] A similar circuit is analysed in example 18.3.

12.3.3 Interaction with other performance aspects

We have seen that pole-splitting reduces the controller gain in the frequency range between the split poles. This may increase the nonlinearity of the feedback amplifier. This may be of particular interest if an amplifier stage with a current-limiting character drives a stage in which pole-splitting has been implemented. This is because an amplifier stage in which pole splitting has been implemented by means of parallel capacitive feedback obtains an integrating transimpedance character. If such a stage is driven from a current-limiting stage, the rate of change of the output voltage is limited to the quotient of the drive current and the feedback capacitance. In fact, pole-splitting by means of local capacitive feedback is the common cause of the voltage slew rate limitation as it occurs in operational amplifiers.

The deterioration of the signal-to-noise ratio and of the power efficiency that is associated with this technique can usually be neglected. This is because the feedback capacitance that provides the pole splitting is usually small with respect to existing capacitances in the amplifier stage.

12.4 Pole-zero canceling

Another pole-splitting technique that can be used to change initial pole positions of the loop gain is *pole-zero canceling*. With this technique, a zero is inserted into the loop gain that coincides with a pole in the loop gain. However, we have seen that passive zeros can only be inserted into existing attenuators in the loop. In well-designed amplifiers, the source impedance, the load impedance and the feedback network are the only places where impedances in series or in parallel with the signal path cause attenuation of the loop gain. This is because in well-designed amplifiers, we have not inserted impedances in parallel or in series with the signal path, unless this was inevitable. The source impedance, the load impedance and the impedances of passive feedback elements are such inevitable impedances, and in sections 12.2.5 and 12.2.6 we have seen that we can use these impedances for the creation of phantom zeros.

12.4.1 Insertion of impedances into the signal path

According to the above, the insertion of zeros that are not phantom zeros requires the insertion of an extra attenuator into the loop. This should be accomplished through insertion of an impedance in series or in parallel with the signal path. For this reason, the penalty on other performance aspects, such as, noise performance, power efficiency and linearity is much larger than with pole-splitting by means of negative feedback or with phantom zero compensation.

In the following example, we will demonstrate this technique.

Example 12.14

Let us consider the network depicted in Figure 12.46A. It shows two cascaded single-pole gain stages that are assumed to be part of a controller. The first stage has a pole $p_1 = -1Mrad/s$ and the second stage has a pole $p_2 = -100krad/s$. The aim of pole-splitting is to move the most dominant pole to a lower frequency while moving the pole with the highest frequency to a higher frequency. In this example, we will change the sum of the poles from $-1100krad/s$ to $14Mrad/s$, while maintaining their product at $100Grad^2/s^2$.

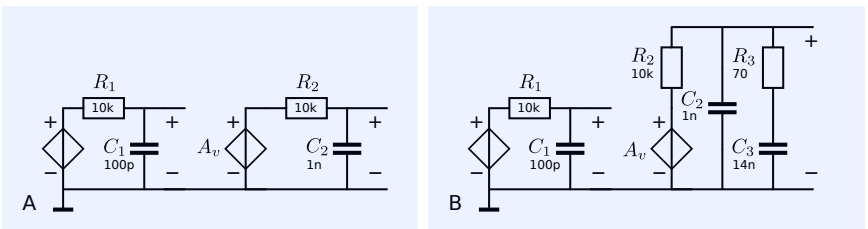


Figure 12.46: Pole-splitting by means of pole-zero canceling
A: Two cascaded gain stages of a controller B: With added pole-zero canceling network.

We will move the most dominant pole to a lower frequency by placing a capacitance C_3 in parallel with C_2. We can then place a zero on p_1 by inserting a resistor R_3 in series with C_3. The zero will be located at $-\frac{1}{R_3C_3}$. Figure 12.46B shows the modified circuit. The second gain stage now has two poles and one zero. If the poles are well separated ($C_3 \gg C_2$) and ($R_2 \gg R_3$), their frequencies can be estimated as discussed in section 18.5.3: $p_2 \approx -\frac{1}{(R_2+R_3)C_3}$ and $p_3 \approx -\frac{1}{R_3C_2}$. The zero in the second stage cancels p_1 of the first stage, so the sum of the new poles of the two cascaded stages equals $p_2 + p_3$. Hence, we obtain the following design equations:

$$\frac{1}{(R_2+R_3)\,C_3} + \frac{1}{R_3C_2} = 14 \times 10^6, \tag{12.75}$$

$$R_3C_3 = 10^{-6}, \tag{12.76}$$

where $C_2 = 1nF$ and $R_2 = 10k\Omega$. Numeric solution yields: $C_3 \approx 14nF$, $R_3 \approx 70\Omega$. With these values, the product of the poles is slightly degraded to $95.2Grad^2/s^2$. This is because the symbolic expressions for p_2 and p_3 are not accurate.

In the following example, we will demonstrate this technique for the compensation of the transimpedance amplifier from example 11.3. We will show that this technique is inferior with respect to the phantom zero compensation discussed in example 11.3.

Example 12.15

Figure 12.47 shows the small-signal equivalent circuit of the uncompensated transimpedance amplifier from example 11.3. In that example, we already determined the values of the dominant poles and the DC loop gain. We found two dominant poles: $p_1 = -16Hz$ and $p_2 = -96.4kHz$. The DC loop gain was obtained as $-973 \cdot 10^3$.

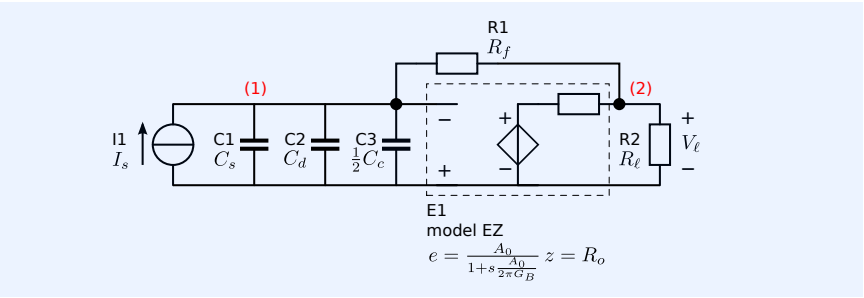


Figure 12.47: Small-signal diagram of the uncompensated transimpedance amplifier from example 7.11.3.

With these values, the achievable MFM bandwidth amounts to 1.225*MHz. This largely exceeds the absolute value of the sum of the poles, thus MFM compensation is required. Unfortunately, the poles cannot be split by moving the dominant pole closer to the origin and the other one further away from the origin. The dominant pole is a pole of the voltage gain of the operational amplifier and its value cannot be altered. We can move* p_2 *to a higher frequency by placing a resistor in parallel with the input of the operational amplifier. This technique is known as resistive broadbanding and will be discussed in section 12.5. In order to shift* p_2 *to* $-\sqrt{2} \times 1.225MHz$*, we need to place a resistance* $R_z = 5.896k\Omega$ *in parallel with the input of the operational amplifier. If we want to apply pole-splitting and maintain an MFM characteristic, we need to create a zero at* $-16Hz$*, which can be achieved by placing a capacitance* $C_z = 1.69\mu F$ *in series with this resistor. In fact, we now created a new dominant pole at* $-0.89Hz$*. Figure 12.48 shows the modified circuit.*

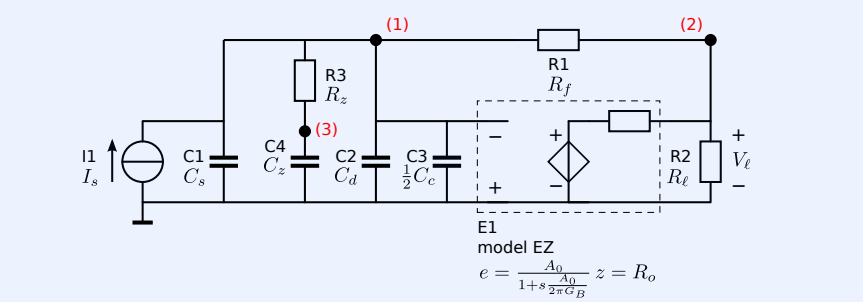


Figure 12.48: Small-signal diagram of the transimpedance amplifier from example 7.11.3 with pole-splitting by means of pole-zero canceling.

Below is the SLiCAP *netlist of the circuit:*

```
transimpedancePZcancel
* file: transimpedancePZcancel.cir
* SLiCAP circuit file
I1 0 1 0
C1 1 0 {C_s}
C2 1 0 {C_d}
C3 1 0 {C_c/2}
C4 3 0 {C_z}
R1 1 2 {R_f}
R2 2 0 {R_ell}
R3 1 3 {R_z}
E1 2 0 0 1 EZ value={A_0/(1+s/2/pi/16)} zo={R_o}
.param C_s=5p R_f=100k R_ell=2k C_d=8p C_c=7p A_0=1M
+ R_o=55 C_z={1/R_z/2/pi/16} R_z=5.896k
.end
```

The value of the capacitance C_z *has been defined symbolically. During numeric analysis, it will be calculated with a high accuracy such that the zero exactly cancels the pole at* $-16Hz$*.*

Below is the SLiCAP *script for analysis of the locations of the poles and the zeros. For determination of the poles, the data type has been set to* `'poles'`*, and*

for determination of the zeros the data type has been set to `'zeros'`. Combined pole-zero analysis is performed with the data type set to `'PZ'`. With this analysis type, poles and zeros that coincide will not be displayed.

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: transimpedancePZcancel.py

from SLiCAP import *

fileName = 'transimpedancePZcancel'
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName + '.cir')
i1.setSource('I1');
i1.setDetector('V_2')
i1.setLGref('E1')
i1.setSimType('numeric')
i1.setGainType('loopgain')
i1.stepOff()
# Display poles
i1.setDataType('poles')
result = i1.execute()
listPZ(result)

# Display zeros
i1.setDataType('zeros')
listPZ(i1.execute())

# Display, poles, zeros and DC gain
i1.setDataType('pz')
listPZ(i1.execute())

# Calculate the phase margin
i1.setDataType('laplace')
i1.setGainType('loopgain')

L = i1.execute()
loopGain = L.laplace
pmResults = phaseMargin(loopGain)

uF = pmResults[1]
pM = pmResults[0]

print('Loop gain: phase margin = {:3.2f}deg at f = {:8.2e}Hz\n'.format(pM, uF)
    )

# Show the poles, zeros and DC value of the gain
i1.setDataType('pz')
i1.setGainType('gain')
listPZ(i1.execute())
```

The results of this script will be displayed in the PYTHON *console window:*

```
Poles of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -8.90e-1        0.00e+00        8.90e-1
 1        -1.60e+1        0.00e+00        1.60e+1
 2        -1.73e+6        0.00e+00        1.73e+6

Zeros of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.60e+1        0.00e+00        1.60e+1

DC value of loopgain: -9.73e+5

Poles of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -8.90e-1        0.00e+00        8.90e-1
 1        -1.73e+6        0.00e+00        1.73e+6

Found no zeros.
```

```
Loop gain: phase margin = 65.52deg at f = 7.89e+05Hz

DC value of gain: -1.00e+5

Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -8.66e+5        -8.67e+5         1.23e+6   7.07e-1
 1        -8.66e+5         8.67e+5         1.23e+6   7.07e-1

Zeros of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        2.91e+10        0.00e+00        2.91e+10
```

The results show exact canceling of the pole of the operational amplifier, while the product of the poles is maintained. This yields an MFM characteristic of the gain.

12.4.2 Interaction with other performance aspects

Although pole-splitting with the aid of pole-zero canceling can successfully be applied for high-frequency compensation, the penalty on other performance aspects can be relatively high. This can easily be seen from example 12.15. There, we achieved the MFM characteristic by placing an RC series network in parallel with the input of the operational amplifier. This network causes an attenuation of the loop gain for frequencies between 0.89Hz and 1.22MHz. This frequency range covers almost the whole operating frequency range of the amplifier, which has the following consequences:

1. The resistor `R3` adds noise and increases the contribution of the equivalent input voltage noise source of the operational amplifier for frequencies between 16Hz and 1.22MHz. This causes a serious deterioration of the signal-to-noise ratio.

2. The capacitor `C3` increases the contribution of the equivalent input voltage noise source of the operational amplifier for frequencies between 0.89Hz and 16Hz. This causes a serious deterioration of the signal-to-noise ratio.

3. The reduction of the loop gain over almost the whole frequency range of the amplifier increases the nonlinearity of the amplifier.

4. The overdrive recovery time of the amplifier is enlarged due to charge storage on `C3`.

   This can be seen as follows. A (too) large source current may cause voltage clipping at the output of the amplifier. In such an event, the input voltage of the controller will not longer be zero and charge will be stored on `C3`. The amplifier can only return to its normal state after this charge has been removed. In the absence of an input signal, this time is determined by the pole frequency set by `C3`. With phantom zero compensation, this pole frequency was about 96kHz. With pole-zero canceling this pole frequency is about 0.89Hz. Hence, in the transimpedance amplifier discussed above, overdrive recovery will take about 76 times longer with pole-zero canceling when compared to phantom zero compensation.

If this technique is applied in a part of the circuit that carries a large signal level, there may be a severe penalty on the energy storage and the power efficiency of the feedback amplifier. In general, reduction of the loop gain in a certain frequency range will result in a degraded performance of the feedback amplifier in that frequency range.

possible improvements

The negative effects described above can be reduced if we accept zeros in the transmission band of the amplifier. In other words if we accept deviation from the all-pole MFM characteristic. If our design goal is to guarantee stability and accept a certain amount of overshoot and tilt in the step response, we may follow a different compensation strategy. This is often done when using the phase margin as a measure for stability. In section 12.6 we will pay attention to this design approach.

12.5 Resistive broadbanding

Resistive broadbanding is a technique that can be applied to increase the frequency of a pole associated with an independent capacitor voltage or an independent inductor current, while maintaining the contribution of this subnetwork to the gain-pole product.

12.5.1 Insertion of resistors into the signal path

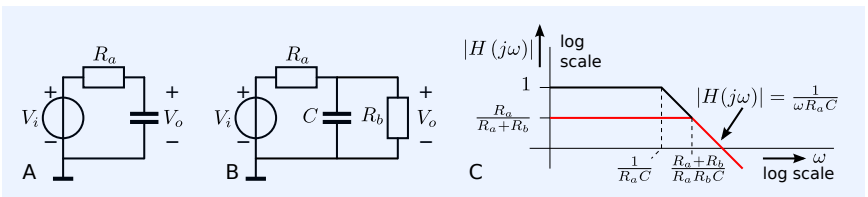


Figure 12.49: Resistive broadbanding
A. Circuit with single-pole voltage transfer: $H(j\omega) = \frac{V_o}{V_i}$
B. R_b in parallel with C implements resistive broadbanding
C. Magnitude characteristics of the voltage transfers from circuits A and B.

Resistive broadbanding can be done by inserting a resistor in parallel with a capacitor or in series with an inductor. Figure 12.49 illustrates the first method. We have already illustrated the method for compensation of the transimpedance amplifier in example 12.15. There, resistive broadbanding was implemented by placing a resistor in parallel with the input of the operational amplifier.

12.5.2 Interaction with other performance aspects

Resistive broadbanding can be regarded as a brute force method for exchanging controller gain with controller bandwidth. If it is done at the input of the controller, the noise penalty, as well as the penalty on the DC error and the temperature drift thereof, can be large. If it is done at stages that carry a large signal level, e.g., at the output of the controller, the penalty on the power efficiency may be large. In general, reduction of the midband loop gain by means of resistive broadbanding will result in a degraded performance of the feedback amplifier.

12.6 Phase margin design

Phase margin and gain margin are parameters of the loop gain that are often used as a measure for the stability of negative feedback amplifiers (see section 11.5.2). In this book, the design of the dynamic response of a negative feedback amplifier has been based on manipulation of the solutions of the characteristic equation of the source-load transfer of an amplifier. This is the only mathematically proper approach. The solutions of the characteristic equation are all the poles, including those associated with non-observable or

not-controllable states. Despite its popularity, the design of the frequency response of a source-load transfer driven by the gain margin and phase margin of the loop gain is not advocated. In general, there does not exist a unique correspondence between the dynamic properties of the source-load transfer and the gain and phase margin of the loop gain. However, for the sake of completeness, we will demonstrate the method.

12.6.1 Lag and lead compensators

The phase margin compensation method makes use of *lag compensators* and of *lead compensators*. Both types of compensator add a pole and a zero to the loop gain. In a lag compensator, the frequency of the pole is below that of the zero. The dominant pole causes phase lag and the zero reduces this phase lag. A lead compensator has the frequency of the pole above that of the zero. The dominant zero introduces phase lead and the pole reduces it. Figure 12.50 shows an example of a lag compensation network, and Figure 12.51 shows an example of a lead compensation network. Phantom zero implementations use lead networks, while pole-zero canceling is realized with lag networks.

The compensation strategy is to provide enough phase margin to ensure stability of the circuit. In the following example, we will demonstrate the design of a lag compensator for the transimpedance amplifier, thereby using the phase margin design approach.

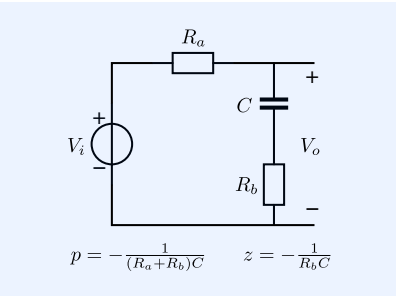


Figure 12.50: Example of a lag compensation network. The frequency of the zero in the voltage transfer is above that of the pole.

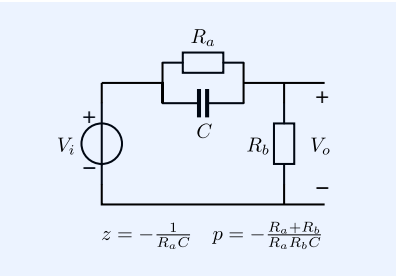


Figure 12.51: Example of a lead compensation network. The frequency of the zero in the voltage transfer is below that of the pole.

Example 12.16

In example 12.15, we discussed pole-zero canceling, and our objective was to achieve an MFM characteristic. In Figure 12.48, the feedback resistance R_f and the pole-zero canceling network that consists of the series connection of R_z and C_z can be regarded as a lag compensator (see Figure 12.50). After compensation with R_z and C_z, we found a rather large degradation of the signal-to-noise ratio, an increased overdrive recovery and possibly an increased nonlinearity.

If we look at this compensation problem from the perspective of improvement of the phase margin, not caring much about the MFM requirement, we may obtain different values for C_z and R_z and the degradation of other performance aspects may be less. The resistance R_z cannot be changed much. This is because at high frequencies where C_z acts as a short with respect to R_z, this resistance, together with the total capacitance at the input of the operational amplifier, determines the frequency of the pole with the highest frequency. In order to approximate an MFM response, the frequency of this pole should be at about $\sqrt{2}$ times the bandwidth of the amplifier. However, if we do no longer aim accurate pole-zero canceling, the capacitance C_z can be taken much smaller. The frequency of the new pole will then be above that of the pole of the operational amplifier, while the frequency of the zero no longer coincides with the pole of the operational amplifier. Its frequency will be above that of the new pole. The loop gain will then have three poles and one zero:

$$p_1 = -16\ Hz, \tag{12.77}$$

$$p_2 \approx -\frac{1}{2\pi R_f\left(C_s + C_z + C_d + \frac{1}{2}C_c\right)}\ Hz, \tag{12.78}$$

$$p_3 \approx -\frac{1}{2\pi R_z\left(C_s + C_d + \frac{1}{2}C_c\right)}\ Hz, \tag{12.79}$$

$$z_1 = -\frac{1}{2\pi R_{fz}C_z}\ Hz. \tag{12.80}$$

Below the script for determination of the DC value, the poles, and the zeros of the gain and the loop gain with $C_z = 200pF$, as well as the phase margin of the loop gain.

```
print('=== Modified compensation C_z = 200p. ===\n')
i1.defPar('C_z', '200p')

# Show the poles, zeros and DC value of the gain
listPZ(i1.execute())

i1.setGainType('loopgain')
# Show the poles, zeros and DC value of the gain
listPZ(i1.execute())
# Show the phase margin of the loop gain
i1.setDataType('laplace')
L = i1.execute()
loopGain = L.laplace
pmResults = phaseMargin(loopGain)

uF = pmResults[1]
pM = pmResults[0]

print('Loop gain: phase margin = {:3.2f}deg at f = {:8.2e}Hz\n'.format(pM, uF)
    )
```

The output is:

```
=== Modified compensation C_z = 200p. ===

DC value of gain: -1.00e+5

Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.64e+5        0.00e+00         1.64e+5
 1        -8.52e+5        -7.14e+5         1.11e+6   6.52e-1
 2        -8.52e+5         7.14e+5         1.11e+6   6.52e-1

Zeros of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.35e+5        0.00e+00         1.35e+5
 1        2.91e+10        0.00e+00        2.91e+10

DC value of loopgain: -9.73e+5

Poles of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.60e+1        0.00e+00         1.60e+1
 1        -6.99e+3        0.00e+00         6.99e+3
 2        -1.86e+6        0.00e+00         1.86e+6

Zeros of loopgain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0        -1.35e+5        0.00e+00         1.35e+5

Loop gain: phase margin = 58.25deg at f = 7.59e+05Hz
```

The gain has three poles and two zeros. The positive zero is a result of the direct transfer. The zero at $-135kHz$ *is the zero of the loop gain. At the frequency of this zero, the magnitude of the loop gain exceeds unity. Hence, in the root locus, one of the poles of the loop gain will move towards this zero. Because the loop gain is finite this pole does not reach the zero, and a pole-zero pair remains in the gain. This is illustrated in the root locus plots from Figure 12.53 and 12.55. At the available loop gain, the pole that moves towards the zero reaches* $-100.69kHz$*. This can be seen in the listing of the poles and the zeros of the gain with* $C_z = 200pF$*.*

The phase margin after compensation is 49 *degrees, which almost equals the phase margin of the circuit with pole-zero canceling* (52 *degrees). Figure 12.52 shows the phase margin as a function of the compensation capacitance* C_z*. It shows that increasing* C_z *above* $200pF$ *does not result in a significant improvement of*

the phase margin. The script below shows how to display the phase margin and the poles and zeros of the loop gain and how to plot this phase margin against C_z.

```
70 # Plot the phase margin versus C_z
71 i1.setStepVar('C_z')
72 i1.setStepStart(0)
73 i1.setStepStop('200p')
74 i1.setStepNum(100)
75 i1.setStepMethod('lin')
76 i1.stepOn()
77
78 result = i1.execute()
79
80 PM = [i1.stepList, phaseMargin(result.laplace)[0]] # x, y trace data
81 htmlPage('Phase margin')
82 plotData = {'PhaseMargin vs C_z': PM}
83 figPM = plot('PM', 'Phase margin versus $C_z$', 'lin', plotData, xName = '$' +
       sp.latex(i1.stepVar) + '$', xScale = 'p', xUnits = 'F', yName = 'Phase
      margin', yUnits = 'deg', show = True)
84 fig2html(figPM, 800)
```

Phase margin

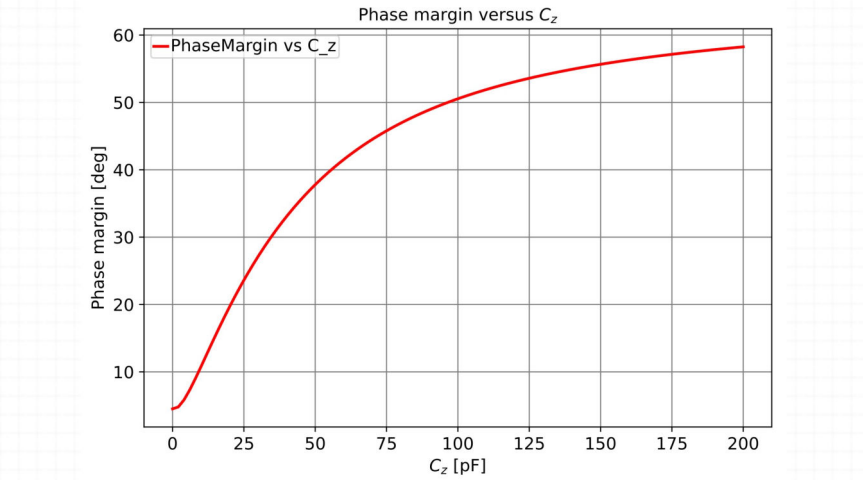


Figure 12.52: Phase margin as a function of the compensation capacitance C_z.

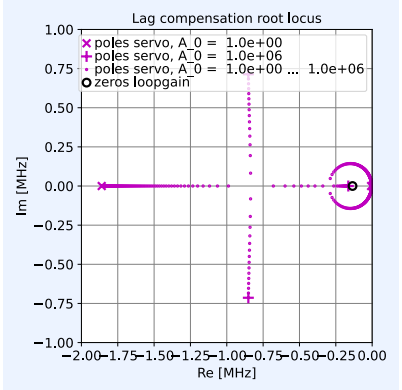


Figure 12.53: Root locus plot of the transimpedance amplifier with lag compensation.

Step response PZ canceling

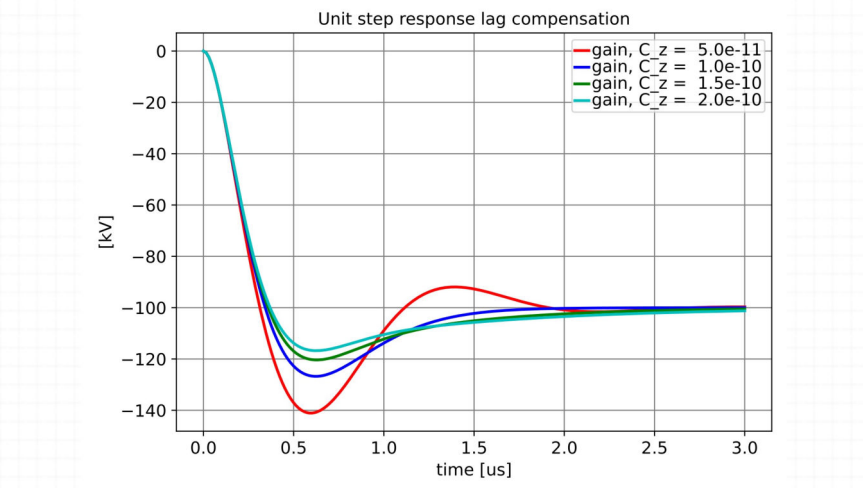


Figure 12.54: Step response of the transimpedance amplifier with lag compensation as a function of the compensation capacitance C_z.

Please note that this figure shows the unit step response which has a step size of 1A. The output voltage of this linear circuit with a tranfer of about 10^5 will thus be 10^5V!

The effect of such a pole-zero pair in the transfer of the amplifier can clearly be observed in the step response of the amplifier and in the magnitude and phase characteristics of its transfer. The unit step response shown in Figure 12.54 shows

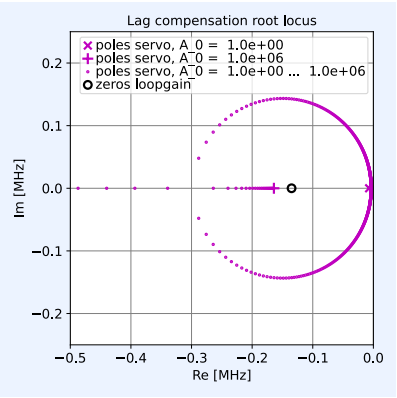


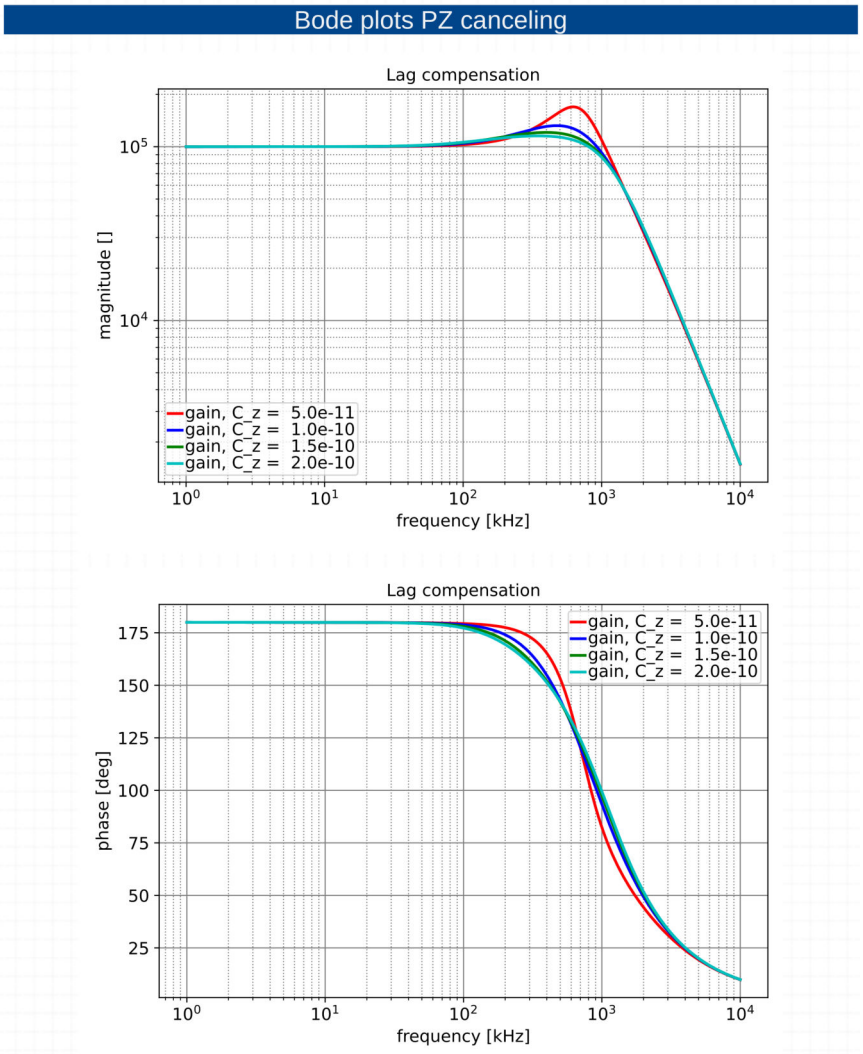
Figure 12.55: Detail of the root locus plot of the transimpedance amplifier with lag compensation.

sequential settling to two different levels. First, the output voltage appears to settle at about −103kV, and after some time final settling to −100kV takes place.

If the first settling is much faster than the second, the ratio between these two levels approximates the ratio of the frequency of the pole and that of the zero. The settling time to the first level is determined by the poles with the highest frequency. The settling time to the final value is determined by the pole at −164kHz. Generally, it may be concluded that a pole-zero pair in the transmission band of the amplifier causes droop or tilt in its step response.

A similar effect can be observed in the magnitude characteristic. The magnitude characteristic in Figure 12.56 shows the effect of the pole-zero pair. Since the frequency of the zero is below that of the pole, the gain increases in the region between the frequency of the zero and that of the pole. Again, the relative increase equals the ratio of the frequency of the pole and the frequency of the zero. Aside from this effect, the curve deviates from an MFM response due to insufficient compensation. If, after optimization, the zero is not canceled by a pole, the response will deviate from an MFM response.

Figure 12.56: Magnitude characteristic and phase of the transfer of the transimpedance amplifier with lag compensation as a function of the compensation capacitance C_z.



12.6.2 Interaction with other performance aspects

Although the phase margin design approach with lag or lead networks may result in stable behavior, the most powerful method for frequency compensation is the application of phantom zeros.

12.7 Reduction of the servo bandwidth

We have seen that a large loop gain is beneficial to almost all performance aspects of a feedback amplifier. The only performance degradations that cannot be cured by a large loop gain are:

1. The addition of noise at the input of the controller and in the feedback network.
2. The addition of DC offset at the input of the controller and associated temperature effects.
3. Energy storage and power losses that occur in (parasitic) impedances in series and/or in parallel with the load, the source or elements of the feedback network.

There may be situations in which the loop gain needs to be very large, yielding servo bandwidth that is much larger than required. If the order of the servo function exceeds two or three, frequency compensation while maintaining the bandwidth may become difficult or even impossible.

Consider, for example, an audio power amplifier that needs to drive electrodynamic loudspeakers. Such an amplifier needs to have a large power gain, a very low distortion at low frequencies, and a very low output impedance, but a rather modest frequency range. The requirements for the low distortion combined with the large power gain and the low output impedance, often result in a feedback amplifier with a very large loop gain. As a result, the servo bandwidth may be much larger than required and the number of dominant poles may be too large to deal with during frequency compensation.

In the above situations, it may be a useful approach to limit the number of dominant poles without limiting the low-frequency loop gain. This can be achieved with the aid of pole-splitting techniques or by using brute force methods.

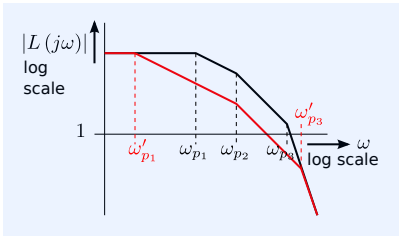


Figure 12.57: Bandwidth reduction by means of excessive pole-splitting. The poles p_1 and p_3 have been split. After splitting, p_3 not longer belongs to the dominant group and the servo bandwidth is reduced.

12.7.1 Excessive pole-splitting

In sections 12.3 and 12.4, we have introduced two pole splitting techniques that can be applied for increasing the sum of two dominant poles of the loop gain, without changing their product. If two dominant poles after splitting are still dominant poles of the loop gain, the bandwidth of the servo function will be preserved. The splitting of the poles can also be performed in such a way that one of the poles is moved out of the dominant group. If so, the order of the servo function is decreased by one and the bandwidth of the servo function will be reduced. Such a reduction of the bandwidth may be desired in the situation described above. Figure 12.57 shows the intended effect on the magnitude characteristics of the loop gain and the servo function.

12.7.2 Pole frequency reduction

Instead of splitting one of the poles out of the dominant group, the bandwidth can also be reduced by moving a pole of the dominant group closer to the origin or by adding a dominant pole. This can be achieved by inserting a capacitance in parallel with the signal path. Figure 12.58 illustrates the

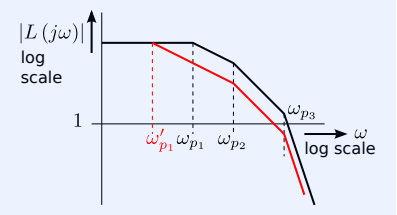


Figure 12.58: Bandwidth reduction by moving a dominant pole closer to the origin. The poles p_1 is moved closer to the origin. As a result p_3 does not longer belong to the dominant group.

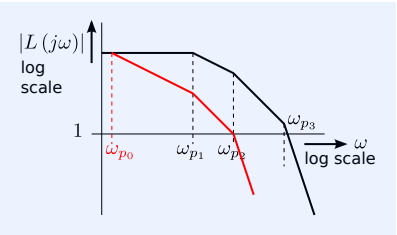


Figure 12.59: Bandwidth reduction by addition of a dominant pole.

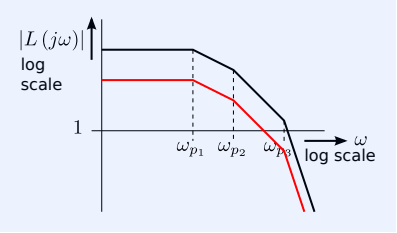


Figure 12.60: Bandwidth reduction by means of DC loop gain reduction .

reduction of the bandwidth of the servo function by moving the most dominant pole closer to the origin. In cases in which the most dominant pole is associated with an independent capacitor voltage, this can be achieved by placing a capacitor in parallel with the existing capacitor.

Figure 12.59 illustrates the effect of the addition of a new dominant pole. This can be achieved by insertion of a capacitor in parallel with the signal path in a resistive network with a nonzero resistance.

12.7.3 DC loop gain reduction

In section 11.5.4, we studied a balanced voltage amplifier that exhibited a common-mode loop gain much larger than the differential-mode loop gain. Since the common-mode DC performance is usually of less interest than the differential-mode DC performance the common-mode frequency compensation may be accomplished by reduction of the common-mode DC loop gain. The effect of this compensation method is illustrated in Figure 12.60.

12.7.4 Interaction with other performance aspects

Reduction of the loop gain in some frequency range generally causes deterioration of accuracy and linearity in this frequency range. If the reduction of the loop gain is achieved by brute force insertion of impedances in series or in parallel with the signal path, it may result in degradation of the signal-to-noise ratio and/or an increase of power losses and/or energy storage. The latter may result in an increase of the overdrive recovery error.

12.8 Feedback biasing frequency compensation

Until now, we have only paid attention to the design of the low-pass roll-off characteristic of the transfer of negative feedback amplifiers. Any transfer of a physical system will show a low-pass behavior due to the fundamental limitation of speed.

Amplifiers that exhibit AC coupling, may also show a high-pass character at low frequencies. AC coupling may be the result of the design of the desired transfer function or of the biasing. In the negative-feedback integrator, discussed in section 11.4.4, AC coupling and the high-pass character of the servo function were a result of the design of the integration operation. In the voltage amplifier discussed in section 9.3, the high-pass character of the transfer was implemented during biasing.

In this section, we will discuss the design of the high-pass characteristic when using negative feedback biasing.

12.8.1 Negative feedback biasing concepts

Negative feedback biasing, as it was introduced in section 9.3, can only be applied if DC signal components are not of interest. In that case DC bias quantities can be separated from the signal. A DC control loop that keeps the bias quantities in their desired range can then be added to the amplifier without affecting the transmission coefficients of the amplifier in the frequency range of interest.

Figure 12.61 shows two arrangements in which a bias loop controls the DC operating voltage of the output port of a negative feedback amplifier. Figure 12.61A shows a circuit concept in which the bias controller is modeled as a voltage-controlled current source. Figure 12.61B shows a circuit concept in

which the bias controller is modeled as a voltage-controlled voltage source. Similar arrangements can be designed to control the DC operating current of the output port.

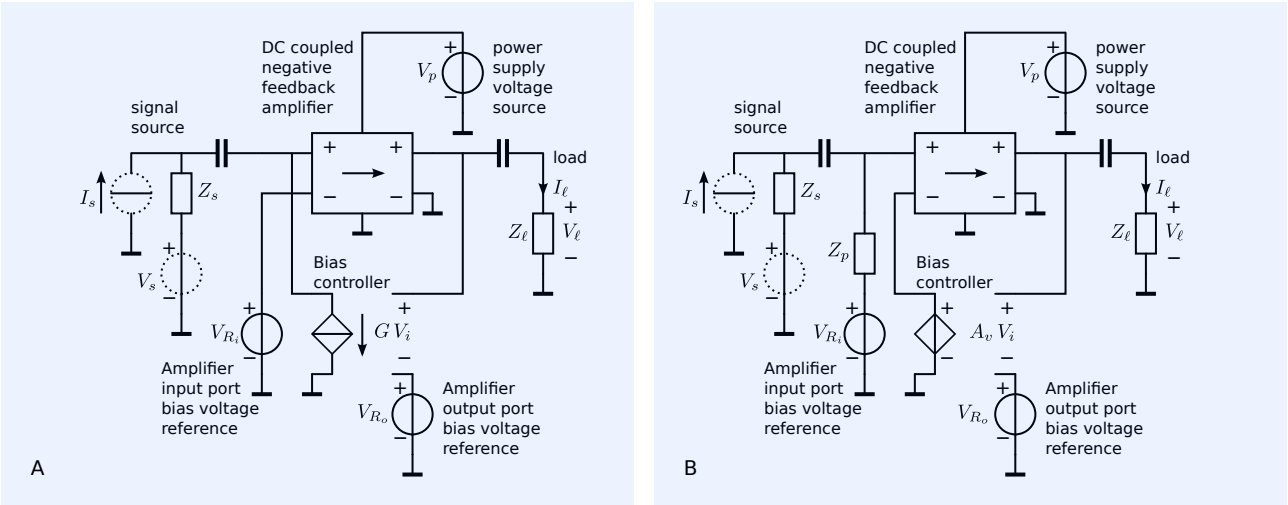


Figure 12.61: Feedback biasing concepts
A. Negative feedback biasing concept with VCCS bias controller
B. Negative feedback biasing concept with VCVS bias controller

In the circuit from Figure 12.61A, the bias circuit controls the DC operating voltage of the output port by adding a DC current to the input of the negative feedback amplifier. This feedback biasing scheme can be applied if the negative feedback amplifier has a nonzero DC transimpedance factor. AC coupling at the source or the load is not always necessary. It has to be applied if no DC current is allowed to flow through the source or the load and if the DC source impedance or the DC load impedance equals zero.

In the circuit from Figure 12.61B, the bias circuit controls the DC operating voltage of the output port by adding a DC voltage to its input. This feedback biasing scheme can be applied if the negative feedback amplifier has a nonzero DC voltage gain factor. AC coupling at the source or the load is not always necessary. However, it has to be applied if no DC current is allowed to flow through the source or the load and if the DC load impedance equals zero. If the source has been AC coupled to the amplifier, a nonzero DC bypass has to be created for setting the DC voltage at the noninverting input of the feedback amplifier. In the circuit from Figure 12.61B, Z_p performs this task.

12.8.2 Dynamic behavior with feedback biasing

In the frequency range of interest, the values of the transmission parameters of the amplifier should not be affected by the biasing. Hence, at those frequencies, the coupling capacitors should behave as short circuits and the gain in the bias control loop should approximate zero. In the concept from Figure 12.61A, the transmission parameter C can be affected by the bias loop, while in the concept from Figure 12.61A, the transmission parameter A could be changed. Aside from affecting the transmission parameters of the amplifier, too small values of the coupling capacitors also deteriorate the signal-to-noise ratio and the power efficiency.

The poles introduced by the coupling capacitors, as well as the dynamic behavior of the controller, should be designed such that a stable MFM high-pass character will be obtained. The general approach to the design of feedback biasing circuitry is as follows:

1. Use the idealized model for the negative feedback amplifier.

   This is allowed if the low-pass cut-off and the high-pass cut-off of the

amplifier are well separated.

2. If necessary, apply AC coupling of the source and the load and define the lower limits of the coupling capacitors on grounds of:
   (a) Their influence in the noise behavior
   (b) The high-pass cut-off frequency
   (c) The deterioration of the power efficiency of the amplifier

[19] Including the frequency compensation of the biasing.

3. Design the desired transfer characteristic of the bias controller.[19]
   This is done by evaluating the source-load transfer of the amplifier, including its bias loop, and equating the coefficients of s of this transfer with those of the desired high-pass characteristic.

We will demonstrate this for the design of the biasing of the charge integrator from example 11.4.

Example 12.17

In example 11.4, we discussed the design of the high-pass behavior of a current integrator. We derived a requirement for the DC gain of the controller such that the high-pass cut-off of the integrator was below 1kHz. However, we did not finalize the design and did not focus on the biasing. The design of the biasing will be discussed in this example. We will use the following additional requirements for the biasing of the integrator:

1. *A bias current of maximally $\pm 1\mu A$ is allowed to flow through the source.*
2. *A DC current of maximally $+2.5mA$ is allowed to flow through the load.*
3. *The DC voltage at the output of the amplifier should be $2.5 \pm 0.01V$.*
4. *A power supply of $\pm 5V$ is available.*

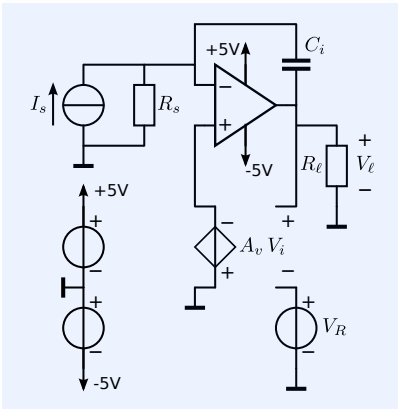


Figure 12.62: Current integrator from example 11.4 with negative feedback biasing according to Figure 12.61B.

Figure 12.62 shows the current integrator in which an operational amplifier has been used for the controller and feedback biasing according to Figure 12.61B has been applied. The aim of the biasing circuit is to keep the quiescent DC output voltage within specifications.

During the design of the biasing, the transfer of the reference voltage V_R to the DC output voltage is of interest. The ideal value of this transfer should be unity. We will use the asymptotic gain model to evaluate the dynamic behavior of this biasing circuit and determine the requirements for the controller. To this end, we select A_v of the controller as the reference variable and evaluate the asymptotic gain, the loop gain and the servo function. With A_v as the loop gain reference variable, the asymptotic gain equals unity. Hence, it equals the ideal gain and the servo function will describe the deviation from this ideal transfer.

In order to find an expression for the servo function, we will determine the loop gain $L_B(s)$ of the biasing loop. For the evaluation of the loop gain, we will approximate the dynamic behavior of the integrator by its low-frequency behavior.[20] *By doing so, we obtain*

[20] Low-frequency behavior is the behavior below midband frequencies as described in section 12.1.1.

$$L_B(s) = -A_v(s)\frac{A_0}{1 + sA_0R_sC_i}. \tag{12.81}$$

We can now evaluate the servo function of the bias loop:

$$\frac{-L_B(s)}{1 - L_B(s)} = \frac{A_v(s)A_0}{1 + sA_0R_sC_i + A_v(s)A_0}. \tag{12.82}$$

The biasing accuracy will be completely determined by the DC properties of the bias controller if it approximates nullor behavior. For this reason, we will design the bias controller as an ideal integrator:

$$A_v(s) = \frac{1}{s\tau}. \tag{12.83}$$

With this controller, the servo function becomes:

$$\frac{-L_B(s)}{1-L_B(s)} = \frac{1}{1+s\frac{\tau}{A_0}+s^2\tau R_s C_i}. \tag{12.84}$$

For a second order cut-off at 1kHz, we need

$$\tau = \frac{1}{4\pi^2 10^6 R_s C_i} = 0.1s. \tag{12.85}$$

The loop gain has one pole in the origin ($s = 0$) and one pole very close to the origin ($s = -\frac{1}{2\pi A_0 R_s C_i} Hz$). A smooth MFM high-pass cut-off can be achieved with phantom zero compensation. With the two poles of the loop gain in the origin, this zero should be located at: $z = -\frac{1000}{\sqrt{2}} Hz$.

In the next example, we will verify the results with SLiCAP.

Example 12.18

Figure 12.63 shows the circuit of the biased charge amplifier. The phantom zero for frequency compensation of the biasing has been implemented with C_c. The OPA627 has been selected as the operational amplifier for implementation of both controllers. The SLiCAP *netlist of the small-signal equivalent circuit is listed below:*

```
QampBias
* file: QampBias.cir
* SLiCAP circuit file
I1 0 1 0
R1 1 0 {R_s}
R2 2 0 {R_ell}
R3 2 4 {R_B}
C1 1 2 {C_i}
C2 2 4 {C_c}
C3 3 4 {C_B}
O1 3 1 2 0 OPA627
O2 0 4 3 0 OPA627
.param R_s=50k R_ell=2k C_i=5p R_B=1M C_B=100n C_c=0
.end
```

With $R_B = 1M\Omega$ and $C_B = 100nF$, we have $\tau = 0.1s$, and with $C_c = 220pF$, we have the value of the phantom zero as defined in the previous example. The SLiCAP *script for plotting the magnitude characteristics of the uncompensated and compensated circuit, as well as for plotting the pole positions while stepping C_c, is listed below:*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: QampBias.py

from SLiCAP import *

fileName='QampBias'
prj = initProject(fileName)
i1 = instruction()
i1.setCircuit(fileName+ '.cir')
htmlPage('Circuit data')
netlist2html(fileName+ '.cir')
i1.setSource('I1')
i1.setDetector('V_2')
i1.setSimType('numeric')
i1.setGainType('gain')
i1.setDataType('laplace')
GainUncomp = i1.execute()
GainUncomp.label = 'uncomp.' # Assign a plot label to this result
i1.defPar('C_c','220p')
GainComp = i1.execute()
GainComp.label = 'phz comp.' # Assign a plot label to this result
htmlPage("Bode plots charge amplifier with feedback biasing")
dBmag = plotSweep('dBmagQamp', 'Charge amplifier with feedback biasing', [
     GainUncomp, GainComp], 100, 10e6, 500, funcType='dBmag', show=True)
```

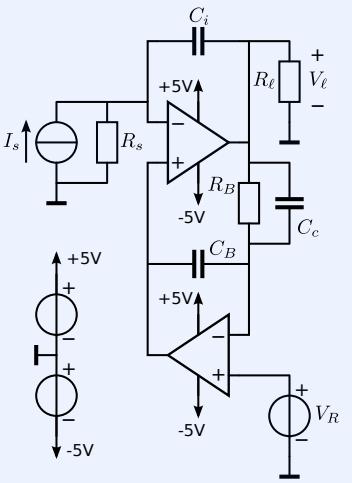
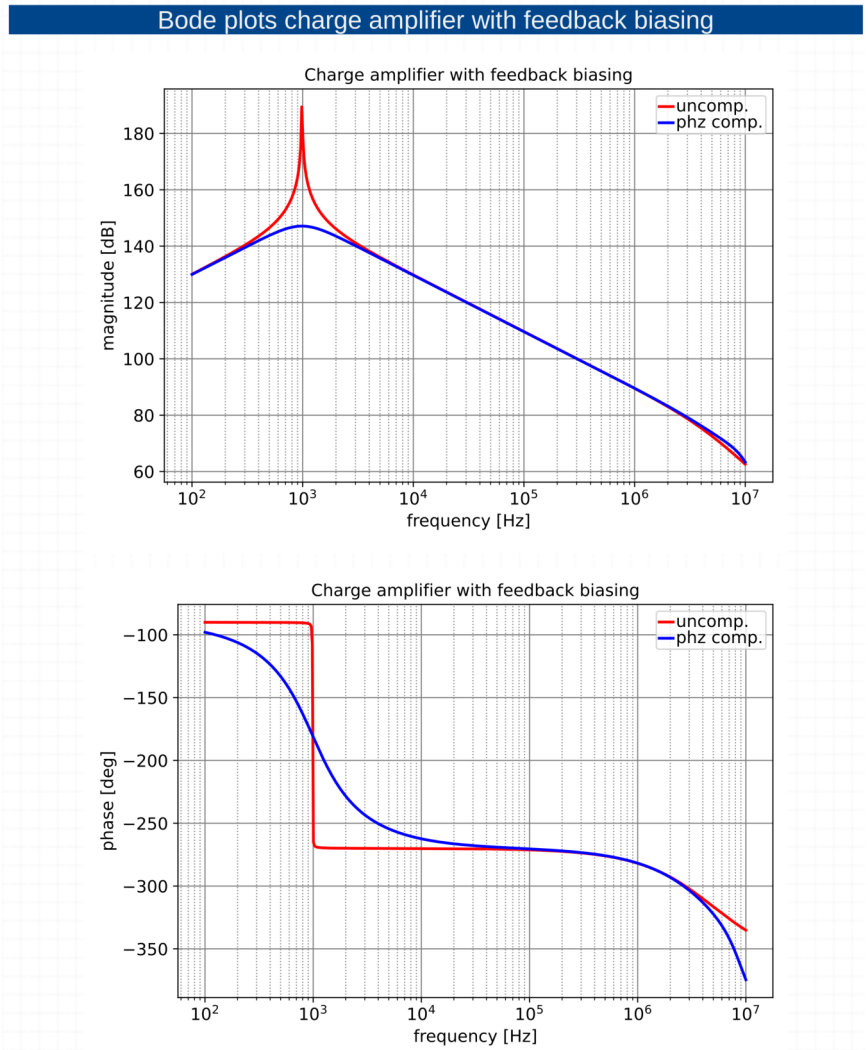


Figure 12.63: Current integrator from Figure 12.62 with negative feedback biasing with phantom zero compensation

```
25 phase = plotSweep('phaseQamp', 'Charge amplifier with feedback biasing', [
       GainUncomp, GainComp], 100, 10e6, 500, funcType='phase', show=True)
26 fig2html(dBmag, 800)
27 fig2html(phase, 800)
28 i1.setDataType('poles')
29 i1.stepOn()
30 i1.setStepVar('C_c')
31 i1.setStepStart(0)
32 i1.setStepStop('500p')
33 i1.setStepNum(20)
34 i1.setStepMethod('lin')
35 pzPlot = plotPZ('RLqAmp', 'Biased charge amplifier', i1.execute(), xmin=-3,
       xmax=0, ymin=-1.5, ymax=1.5, xscale='k', yscale='k', show=True)
```

The Bode plots are shown in Figure 12.64. They clearly show a second order high-pass character of the integrator with a high-pass cut-off frequency of 1kHz. This is exactly as designed. The uncompensated amplifier shows a large peaking in the vicinity of the high-pass cut-off frequency.

Figure 12.64: Magnitude plots of the uncompensated and the compensated charge amplifier with negative feedback biasing.



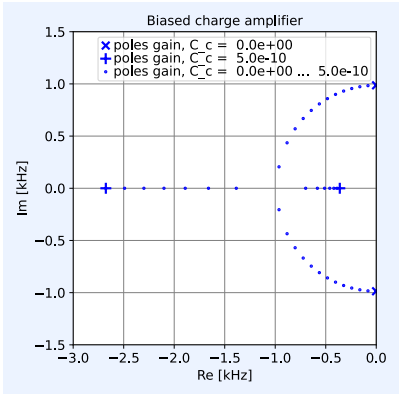


Figure 12.65: Poles of the gain of the charge amplifier as a function of C_c (25pF per step).

Figure 12.65 shows that the two poles of the uncompensated amplifier, that determine the high-pass behavior are almost located on the imaginary axis. This confirms the peaking in the magnitude characteristic and is according to the calculations from the previous example. Figure 12.65 also shows that phantom zero compensation with 225pF brings the poles into MFM positions. The magnitude characteristic of the compensated amplifier shows a smooth transition from a dif-

ferentiator below 1kHz to an integrator above 1kHz.

12.9 Nested control

Until now, we have used the following strategy for the design of negative feedback amplifiers:

1. Design the ideal gain of the amplifier using a nullor as controller
2. Relate the controller requirements to those of the amplifier
3. Select an operational amplifier that meets the controller requirements
4. Perform frequency compensation of the low-pass cut-off
5. Apply biasing
6. If negative feedback biasing is applied, perform frequency compensation of the high-pass roll-off

This is a useful strategy for amplifiers comprising only one controller.[21] However, there may be situations in which it is preferable to build a controller with cascaded local feedback amplifier stages. Each of these amplifier stages has its transfer accurately fixed by means of negative feedback, which gives the controller an accurately fixed gain and pole-zero pattern. In control theory, this is often referred to as *nested control loops*.

[21] These amplifiers only exhibit *over-all feedback*.

12.9.1 PID controllers with local feedback amplifiers

We will illustrate the application and design of controllers with local feedback amplifier stages with the design of a motor driver. Figure 12.66 shows the application of the motor driver. The motor driver has to drive a voice coil motor with a current that is accurately related to the output voltage of a digital-to-analog converter (DAC). One of the motor terminals is connected to ground. The following requirements apply:

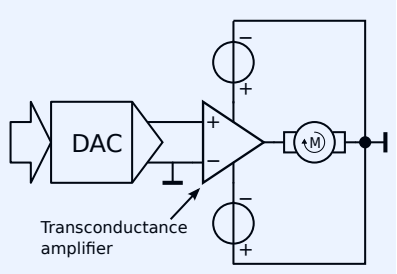


Figure 12.66: Half-bridge motor current driver, dirven from a voltage output DAC.

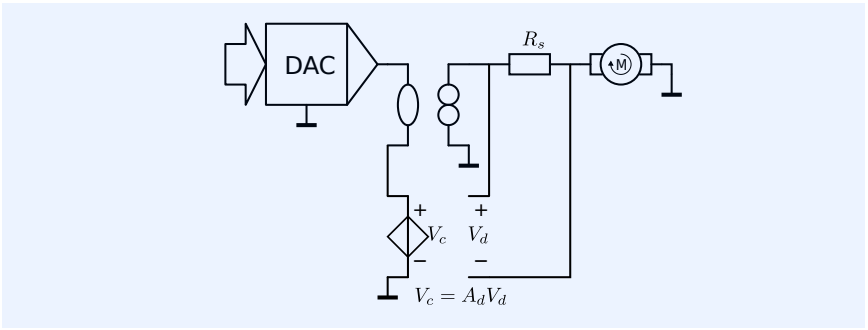
1. The motor driver should be capable of driving multiple types of motors. The dynamic response of the motor driver should not depend on the properties of the motor. The small-signal impedance of the voice-coil motors can be represented by a series LR connection. The inductance L_m and the resistance R_m of the voice coil differ for each motor.
2. Digital settings may be applied to adjust the gain and the dynamic response for each motor type.
3. A single-pole high-efficiency power voltage amplifier with a voltage gain $A_v(s) = \frac{A_0}{1+s\tau}$ is available and should be used for driving the motor.

Over-all feedback architecture

According to the design method discussed in Chapter 7, we could design a single-controller active feedback current driver as shown in Figure 12.67.

In this amplifier structure, a sense resistor with a value R_s has been used to convert the motor current into a (floating) voltage. A differential voltage converts this voltage into a single-ended one that can be compared with the single-ended DAC output voltage. The gain of this differential amplifier equals A_d. A nullator sets the zero condition for this comparison, while a norator provides the dependent motor current for satisfying this condition.

Figure 12.67: Conceptual design of the motor driver amplifier.

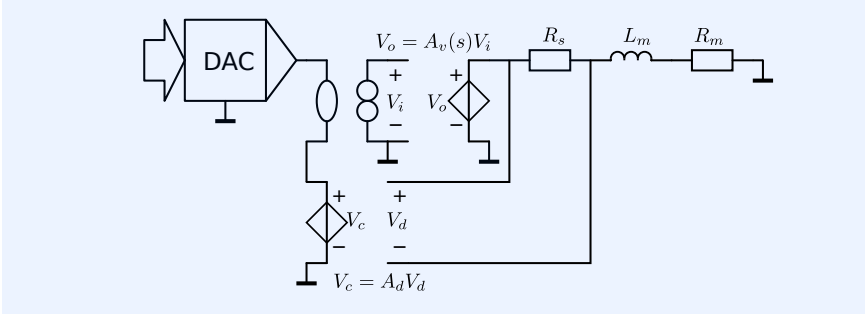


At a first glance, this seems to be a perfect solution. The ideal transconductance gain of the motor driver equals A_d/R_s. The bandwidth can be designed by inserting a zero into the gain of the differential amplifier. This establishes bandwidth limitation with the aid of an active phantom zero. These techniques have been discussed in sections 12.2.9 and 12.2.7, respectively.

The above design approach can also be used if the single-pole high-efficiency power amplifier is applied as the final stage in the controller. Figure 12.68 shows the concept in which the impedance $R_m + sL_m$ represents the small-signal motor impedance. If the nullor were to be implemented with a single-pole operational amplifier and an active phantom zero in the differential voltage amplifier, the gain would show three poles while the servo function would be second order.[22] One extra phantom zero will be required to give the servo function an MFM characteristic.

[22] The gain of a controlled source that implements the nullor has been taken as the loop gain reference.

Figure 12.68: Conceptual design of the motor driver amplifier.



Nested control

The design approach sketched above cannot always be applied. This is because active phantom zero compensation and bandwidth limitation with the aid of phantom zeros are two techniques that increase the effect of non-dominant poles. Particularly if the load impedance varies over a wide range, it may become difficult to establish robust control in this way.

A more robust control can be obtained by designing a controller with an accurate and programmable dynamic performance. This can be done as follows:

1. The static error of the transfer will be zero if the controller has an integrating character at low frequencies.

2. The bandwidth of the system will not depend on the load impedance if the product of the loop gain and the dominant poles does not depend on the motor impedance.

3. The pole due to the motor inductance can be compensated for in the controller. This can be done by creating a programmable zero in the transfer of the controller.

4. The low-frequency loop gain variation due to R_m can be compensated for by making the controller gain programmable.

When designed in this way, the controller gain $A_c(s)$ becomes

$$A_c(s) = \frac{A}{s}\frac{R_s + R_m}{R_s}\left(s\frac{L_m}{R_s + R_m} + 1\right). \tag{12.86}$$

With this controller gain, we obtain a loop gain $L(s)$:[23]

$$L(s) = \frac{A}{s}\frac{A_0}{1 + s\tau}A_d. \tag{12.87}$$

[23] The gain of a controlled source that implements the nullor has been taken as loop gain reference.

If the pole of the voltage amplifier is dominant, the bandwidth B of the transfer can be obtained as

$$B = \frac{1}{2\pi}\sqrt{\frac{AA_0A_d}{\tau}}\ [\text{Hz}]. \tag{12.88}$$

This yields a design equation for the integrator gain. The poles of the servo function can be brought into MFM positions with the aid of a phantom zero in A_d.

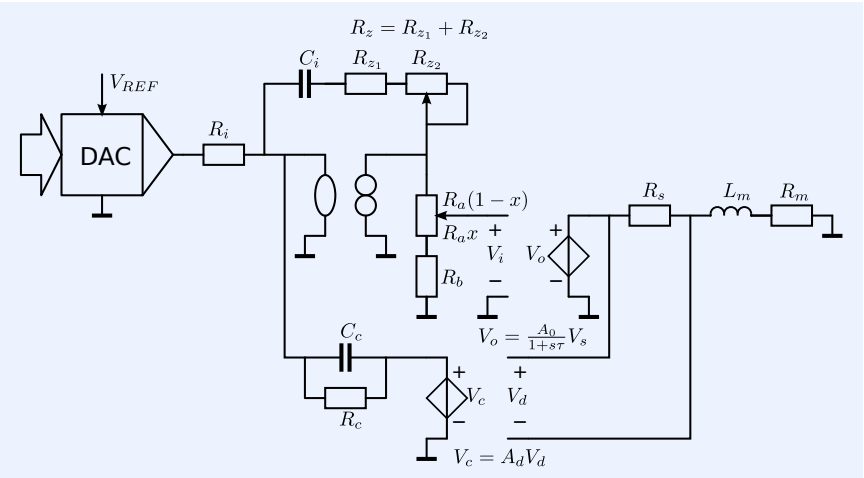


Figure 12.69: Conceptual design of the motor driver amplifier.

Figure 12.69 shows a more detailed concept of an inverting motor driver with nested control. The integrator gain is fixed with the aid of C_i. The zero at $s = -\frac{1}{R_z C_i}$ should coincide with the pole caused by the motor impedance. For this reason, R_z has been made adjustable. The controller gain can be adjusted with the aid of the relative potentiometer setting x. A phantom zero at $s = -\frac{1}{R_c C_c}$ brings the two poles of the servo function into MFM positions.

In terms of classical control theory, the part of the circuit that consists of the differential voltage amplifier with gain A_d and the nullor with R_c, C_c, C_i and R_z can be regarded as a *PID-controller*. The *Proportional* action has been implemented with R_c and R_z, the *Integration* action with C_i, and the *Differentiation* action with C_c. The controller gain $A_c = \frac{V_o}{V_d}$ can be obtained as

$$A_c = \frac{V_o}{V_d} = -A_d\frac{(1 + sC_iR_z)\,(1 + sC_cR_c)}{sC_iR_c}\frac{R_b + xR_a}{R_b + R_a}\frac{A_0}{1 + s\tau}. \tag{12.89}$$

The gain of the amplifier can be changed by changing the value of R_i or of the DAC reference voltage V_{ref}.

The transfer of the controller gain has been fixed accurately with the aid of local feedback amplifiers. The frequency of the zero that should cancel the pole of the load can be adjusted by making R_z programmable using a digitally controlled potentiometer. The controller gain can be adjusted with another digitally controlled potentiometer. Aside from the design equations that follow from the required transfer, the controller gain and the motor impedance, budgets should be defined for the influence of non-dominant poles. Non-dominant poles may be introduced by the nonzero output impedances the nonzero input capacitances of the amplifiers. The circuit should be dimensioned such that the influence of non-dominant poles is kept within acceptable limits.

12.9.2 Increasing bandwidth without adding dominant poles

Sometimes, it is not possible to find operational amplifiers that combine a sufficiently large gain bandwidth product with other performance aspects that are relevant for the controller. There are two approaches to solving this problem:

1. Split the over-all feedback amplifier into several cascaded over-all feedback amplifiers. This generally results in more design flexibility.

2. Use several cascaded amplifiers within the over-all feedback loop. In order to limit the number of dominant poles in this loop, apply local feedback or pole-splitting techniques in one or more of these amplifier stages.

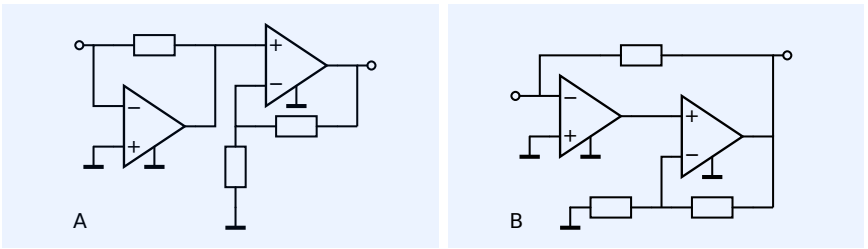


Figure 12.70: Dual stage transimpedance amplifiers:
A: A feedback transimpedance amplifier cascaded with a feedback voltage amplifier
B: An over-all feedback transimpedance amplifier with a two-stage controller.

Figure 12.70 illustrates the two design approaches. If both circuits exhibit equal over-all transimpedance gain, equal second stage voltage amplifiers and equal first stage operational amplifiers, the noise performance of the circuit from Figure 12.70B will be better than that of the circuit from Figure 12.70A.

If, in the circuit from Figure 12.70B, the bandwidth of the voltage amplifier is much larger than that of the servo function of the over-all loop, the voltage gain of this amplifier will contribute to the loop gain poles product of the over-all loop without adding dominant poles.

12.9.3 Interaction with other performance aspects

The application of over-all feedback with a single controller and phantom zero frequency compensation usually results in the best possible performance. However, if there are many dominant poles, this design approach may result in a too complex dynamic behavior. In those situations, nested loop control may be a useful design approach. With nested loop control, the impedances inserted into the signal path of the local feedback stages usually have a larger impact on other design aspects than the feedback impedances in the over-all design would have. Special care should be taken inserting integrators within an over-all feedback loop. Clipping of one of the other stages in the loop

may result in so-called *integrator windup*, which may seriously degrade the overdrive recovery behavior.[24]

[24] Integrator windup is a term used in control theory to indicate the charge accumulated in the integrator during overdrive.

12.10 Compensation for open and shorted ports

The source impedance and/or the load impedance of an amplifier may be subjected to large changes. Those changes may be the result of varying operating conditions or due to defects. Shorted or interrupted connections at the amplifier ports are common defects that have to be considered during design. Such defects may result in a considerable change in the loop gain and cause instability. Unstable behavior and oscillations may cause permanent damage to the amplifier due to increased power losses. In practice, a shorted port will never be terminated with zero impedance and a disconnected port will never be terminated with an infinite impedance. During these fault situations unpredictable parasitic impedances will determine the port termination impedance.

According to the definition of stability, the amplifier will be stable for all possible port terminations if, for all these terminations, the poles are located in the left half of the complex plane. Frequency compensation techniques that have been discussed in this chapter can be applied to achieve this. However, performing root-locus analysis and frequency compensation for an unlimited number of port terminations is not doable. It is far better to perform these analyses and apply frequency compensation techniques only for normal operating conditions. Additional corrective measures can be taken to ensure stability under fault conditions. In this section, we will discuss such measures.

12.10.1 Compensation of shorted ports

In a case of a shorted port or a port terminated with a relatively low impedance, it is useful to study the real part of the port impedance of the amplifier:[25]

[25] A relatively low termination impedance has a magnitude smaller than that of the typical termination impedance.

An amplifier is stable for all port termination impedances if the real part of the port impedance is positive for all frequencies.

Hence, in order to check the stability for relatively low termination impedances, we need to study the real part of the output impedance. The test benches for determination of the input and output impedances have been defined in section 2.3.3.

If, in some frequency range, the real part of the port impedance is negative, stable short circuit behavior can be ensured by placing a correction impedance in series with the port. This series impedance should establish a positive real part of the corrected port impedance at all frequencies.

If no positive feedback has been applied at DC, the DC output resistance will usually be positive. Correction of a possible negative real part of the port impedance at higher frequencies can then be achieved by inserting a parallel LR network in series with the amplifier port. This is shown in Figure 12.71. The real part of the admittance of this network is positive and has a high-pass character:

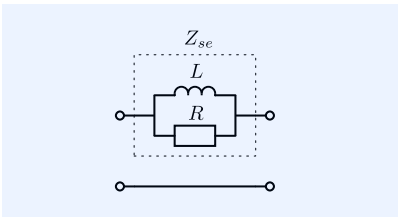


Figure 12.71: Network that can be be inserted between the output port and the load or between the input port and the source to protect for short circuit instability.

$$\mathrm{Re}\,(Z_{se}) = R\frac{\frac{L^2}{R^2}\omega^2}{1+\frac{L^2}{R^2}\omega^2}. \tag{12.90}$$

Hence, it can provide effective compensation for frequencies above $\omega = \frac{R}{L}$.

12.10.2 Compensation of open ports

In a case we leave the port of the amplifier open or terminate it with a relatively high impedance,[26] it is easier to consider the real part of the port admittance of the amplifier:

[26] A relatively high termination impedance has a magnitude larger than that of the typical termination impedance.

An amplifier is stable for all port termination admittances if the real part of the port admittance is positive for all frequencies.

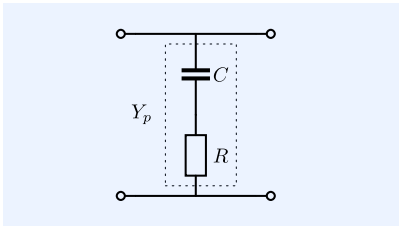


Figure 12.72: Network that can be be inserted between the output port and the load or between the input port and the source to protect for open circuit instability.

If, in some frequency range, this real part is negative, we can ensure open circuit stability by placing a correction admittance in parallel with the port. This parallel admittance should establish a positive real part of the corrected port admittance at all frequencies.

If no positive feedback has been applied at DC, the DC output conductance will usually be positive. Correction of a possible negative real part of the port admittance at higher frequencies can then be achieved by placing a series RC network in parallel with the amplifier port. This is shown in Figure 12.72. The real part of the admittance of this network is positive and has a high-pass character:

$$\mathrm{Re}\left(Y_p\right) = \frac{1}{R}\frac{\omega^2 R^2 C^2}{\omega^2 R^2 C^2 + 1}. \tag{12.91}$$

Hence, it can provide effective compensation for frequencies above $\omega = \frac{1}{RC}$.

12.11 Influence of non dominant poles

Non-dominant poles are poles that do not contribute to the bandwidth of the servo function. They do not significantly change the dynamic response of the amplifier if their frequency is much larger than the bandwidth of the servo function. If not, extra measures have to be taken to reduce their influence. All high-frequency compensation techniques discussed in this chapter can be applied for this purpose.

12.11.1 Bandwidth limitation with phantom zeros

Although bandwidth limitation with the aid of phantom zeros is a powerful technique, care has to be taken while applying it. The increase of the loop gain and of the servo bandwidth established by the phantom zero may turn non-dominant poles (before compensation) into dominant ones. This may pose limits to the application of this technique.

13
Local feedback stages

13.1 Introduction

In Chapter 5, we have learned that a properly biased CS stage can provide a large available power gain. We studied the noise behavior, the dynamic behavior and the nonlinear behavior of this basic amplifier stage.

We have seen that the noise addition of the stage can be minimized through optimization of the device geometry and the operating current.

After studying the small-signal dynamic behavior of the stage driven from and loaded with an $R \parallel C$ network, we found that if the parasitic feedback capacitance c_{dg} is much smaller than the total input capacitance and the total output capacitance, the product of the poles equals the product of the eigenfrequencies of both RC networks, while, if the stage has a high low-frequency voltage gain, the parasitic feedback capacitance largely influences the sum of the poles.

The nonlinear behavior of the CS stage is governed by the voltage dependency of the active capacitance, the current dependency of the transconductance factor and the voltage dependency of the output impedance.

Aside from these performance limitations, the CS stage does not show natural two-port behavior because the input port and the output port share one terminal.

In Chapter 6, we introduced balancing techniques such as complementary-parallel connection and anti-series connection of CS stages. Application of these techniques yields odd transfer functions with a reduced sensitivity for even order effects, such as, threshold voltage mismatch and drift. In addition, anti-series connected CS stages provide improved port isolation, while complementary-parallel connections provide a current drive capability that exceeds its quiescent operating current. We found that application of balancing techniques does not result in a significant change of the small-signal dynamic behavior, the gain accuracy, the odd nonlinearity and the noise behavior.

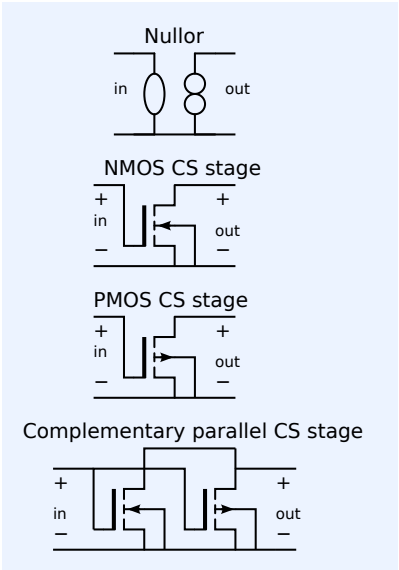


Figure 13.1: Three-terminal nullor and CS stage approximations.

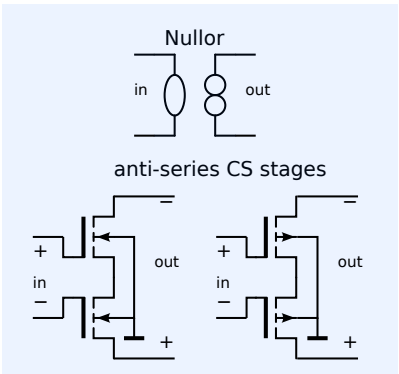


Figure 13.2: Nullor and approximations with anti-series CS stages.

13.1.1 Local feedback amplifier stages

Negative feedback is considered a much more powerful error reduction technique than compensation. Negative feedback uses a controller to minimize the error between actual transfer and the desired transfer of an amplifiers. Through application of negative feedback, the dependency of the source-to-load transfer from the controller properties can theoretically be reduced to zero. This will be the case if the controller is a natural two-port with an infinite available power gain. The nullor is the network concept of the ideal controller.

In this chapter, we will study the design of so-called local feedback amplifier stages. These stages are negative feedback amplifiers of which the controller is implemented with a single transistor CS stage, or its balanced version. Figure 13.1 shows the implementation of a nullor with NMOS or PMOS CS stages or with a complementary-parallel CS stage. Figure 13.2 shows the implementation of the nullor with anti-series connected NMOS or PMOS CS stages. In both figures, the transistors are assumed to be biased; their bias sources have been omitted.

13.1.2 This chapter

In this chapter, we will discuss design and the behavior of local feedback amplifier stages. Local feedback amplifier stages can be applied as single-stage amplifiers or as amplifier stages in multiple-stage feedback amplifiers.

Prerequisite knowledge

The design of negative feedback amplifiers using operational amplifiers as controllers has been described in Chapter 7 to Chapter 12.

1. The synthesis of feedback amplifier configurations
2. The modeling of feedback circuits
3. The relation between the controller performance and the amplifier performance
4. Frequency compensation techniques

The reader is assumed to have a clear understanding of these topics.

This chapter

In section 13.2, we will discuss direct feedback stages. In section 13.4, we will discuss feedback stages that use indirect feedback techniques. Section 13.3 will be devoted to the design of amplifier stages using both balancing and feedback.

13.2 Direct feedback stages

In feedback amplifiers, the values of the nonzero transmission parameters are established with feedback networks. Direct feedback is a technique in which the input quantity of the feedback network is sensed at the load, and the output quantity of the feedback network is compared with the source quantity. This in contrary with indirect feedback or model-based feedback, in which comparison or sensing is performed with a copy of the source or the load signal, respectively.

Feedback networks may comprise nonenergic network elements, passive network elements or amplifiers (active elements).

13.2.1 Nonenergic feedback stages

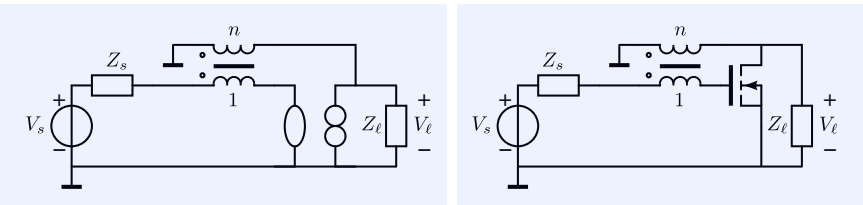


Figure 13.3: Inverting nonenergic feedback voltage amplifier:
Left: Concept with ideal controller
Right: Controller implemented with (biased) CS stage.

Nonenergic feedback stages use feedback network elements that have no energy storage and no dissipation. Such feedback elements do not adversely affect the noise performance and the power efficiency of the feedback amplifier. Ideal transformers, gyrators, open circuits and short circuits are nonenergic feedback elements.

Figure 13.3 shows an example of a nonenergic feedback inverting voltage amplifier. The bias sources in this figure have been omitted. In practice, transformer feedback can only be applied over a limited frequency band.

Figure 13.4 shows the concept of the nonenergic feedback voltage follower and its implementation with a CS stage controller. This stage is known as the *common-drain stage* or shortly the *CD stage*, because the drain is the common terminal for the input port and the output port.[1] Hence, the CD stage can be considered a feedback version of the CS stage.

[1] In this figure, the bias sources have been omitted.

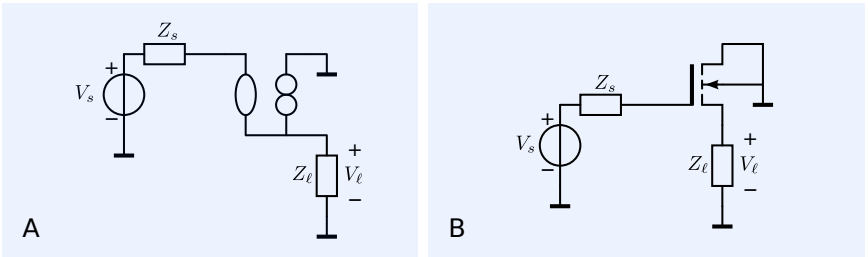


Figure 13.4: Nonenergic feedback voltage follower:
A. Concept with ideal controller
B. Controller implemented with (biased) CS stage.

In section 13.2.2, we will study the noise behavior and the small-signal dynamic behavior of the CD stage. We will see that the behavior of this stage can be related to that of the CS stage, simply by considering the error reduction capabilities of negative feedback.

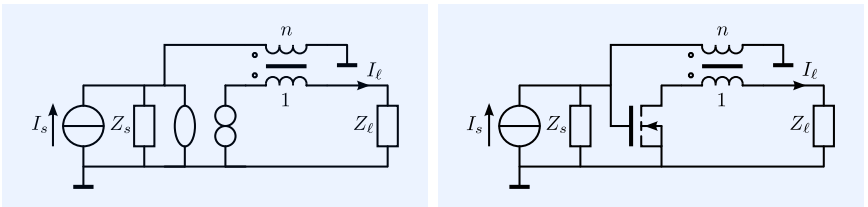


Figure 13.5: Noninverting nonenergic feedback current amplifier:
Left: Concept with ideal controller
Right: Controller implemented with (biased) CS stage.

Figure 13.5 shows an example of an inverting, nonenergic feedback current amplifier, using transformer feedback. The bias sources in this figure have been omitted.

Figure 13.6 shows the concept of the nonenergic feedback current follower and its implementation with a CS stage controller. This stage is known as the *common-gate stage* or shortly the *CG stage*, because the gate is the common terminal for the input port and the output port.[2] Hence, the CG stage can be considered a feedback version of the CS stage.

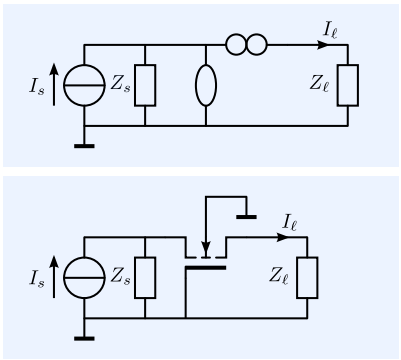


Figure 13.6: Nonenergic feedback current follower:
Upper: Concept with ideal controller
Lower: Controller implemented with CS stage.

[2] The bias sources in this figure have been omitted.

In section 13.2.3 we will study the noise behavior and the small-signal dynamic behavior of the CG stage. We will see that the behavior of this stage can be related to that of the CS stage, simply by considering the error reduction capabilities of negative feedback.

13.2.2 Common Drain Stage

The common drain stage, CD stage, or source follower is a nonenergic feedback voltage follower. It exhibits parallel (voltage) sensing at the load and series (voltage) comparison at the source. The feedback network consists of a short circuit between the load and the inverting input of the controller and an open circuit between the inverting input of the controller and the ground. This feedback structure establishes $A = 1$ without affecting the other transmission parameters. The ideal voltage follower thus has zero output impedance and infinite input impedance.

In this section, we will discuss the noise behavior and the small-signal dynamic behavior of the CD stage.

Noise behavior

Because the CD stage is a nonenergic feedback stage, its equivalent input noise sources equal those of its controller. The controller in the CD stage is a CS stage and the optimization of its noise behavior for capacitive and resistive source has been discussed in section 5.4. The conclusions for the

design of the noise performance of the CS stage are also valid for the CD stage.

Small-signal dynamic behavior

The small-signal transfer from the source to the load can be analyzed in different ways. A straightforward method is to model the stage with its small-signal equivalent network and use network analysis techniques to find the source-to-load transfer. Since the network is rather simple, one would be able to derive design conclusions from this expression.

However, if we consider the CD stage a feedback version of the CS stage, the application of the feedback model provides much more design information. Evaluation of the asymptotic gain, the loop gain and the direct transfer will provide clear design information for obtaining the desired transfer.

We will study the small-signal dynamic behavior of a CD stage which is driven from a resistive source and loaded with $R_\ell /\!\!/ C_\ell$, with the aid of the asymptotic gain feedback model. The analysis for different source and load types proceeds in a similar way.

Figure 13.7 shows the signal diagram of the CD stage. The biasing circuitry has been omitted in the figure. The NMOS transistor is assumed to be biased in the forward saturated operating region.

Figure 13.8 shows the small signal equivalent circuit of the CD stage from Figure 13.7.

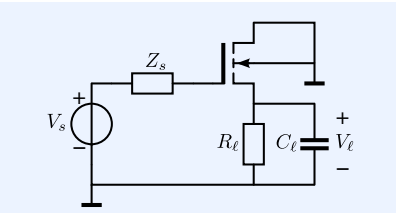


Figure 13.7: Signal diagram (bias circuitry not drawn) of a CD stage driven from a resistive source and loaded with an $R /\!\!/ C$ network.

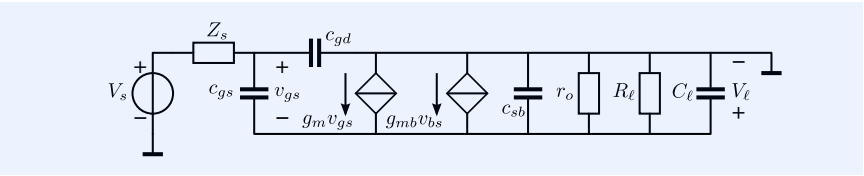


Figure 13.8: Small signal diagram of the CD stage from Figure 13.7.

Selection of the loop gain reference variable

In order to obtain meaningful results from the feedback analysis, the loop gain reference variable should be selected in such a way that the controller obtains nullor properties if the reference variable approaches infinity. If we select g_m as the loop gain reference variable, this will be the case at zero frequency. At frequencies that differ from zero, the current through the parasitic feedback capacitance c_{gd} does not approach zero for any value of V_ℓ with $g_m \to \infty$. This implies that the asymptotic gain will not equal the ideal voltage gain and, as a consequence, the loop gain will not completely describe the deviation of the source-to-load transfer from the ideal transfer. This hinders a two-step design approach in which we first design the ideal gain of an amplifier, and then design the controller in such a way that brings the accuracy and the linearity to their desired values over the frequency range of interest.

Asymptotic gain and ideal gain

The asymptotic gain can easily be found from network inspection if the circuit is redrawn as shown in Figure 13.9.

The parasitic feedback capacitance is now placed outside the controller. Together with the source resistance it establishes a first order low-pass filter in front of the modified CD stage. If we now select g_m as loop gain reference, the modified controller behaves as a nullor. This can be seen as follows: if g_m approaches infinity, the voltage across c_{gs} approaches zero. If the voltage across c_{gs} approaches zero, the current through it also approaches zero. Since the voltage across c_{gs} is the input voltage of the controller and the current

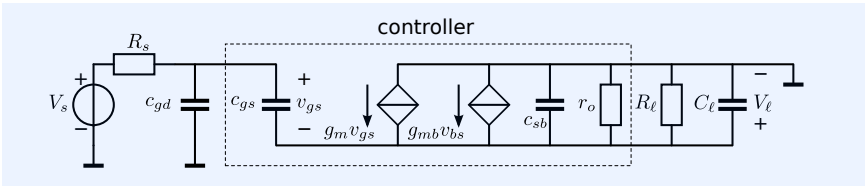


Figure 13.9: Small signal diagram of the CD stage from Figure 13.8 redrawn with the feedback capacitance c_{gd} outside the controller. This easies the calculation of the asymptotic gain from network inspection.

through c_{gs} is the input current of the controller, the controller behaves as a nullor for $g_m \to \infty$. Hence, the asymptotic gain of the modified CD stage equals its ideal gain. The asymptotic gain of the complete CD stage is thus found as

$$A_{f\infty} = \left.\frac{V_\ell}{V_s}\right|_{A\to\infty} = \frac{1}{1 + sR_s c_{gd}} \text{ [-]}. \tag{13.1}$$

From this expression we may conclude that the influence of any signal dependancy of R_s and/or c_{gd} on the source-to-load transfer will not be reduced by negative feedback.

Evaluation of the loop gain

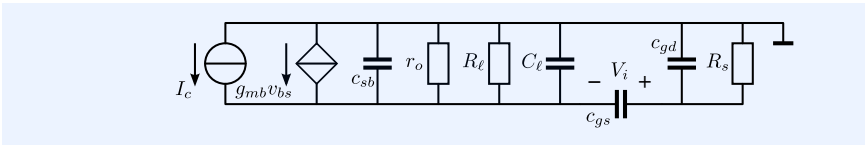


Figure 13.10: Small signal diagram of the CD stage from Figure 13.8 redrawn for evaluation of the loop gain.

Figure 13.10 shows the circuit for evaluation of the loop gain $L = A\lambda\beta\kappa$. According to the asymptotic-gain model, where, in this case, $\lambda\beta\kappa$ is defined as

$$\lambda\beta\kappa = \left.\frac{V_i}{I_c}\right|_{V_s=0} \text{ [}\Omega\text{]}. \tag{13.2}$$

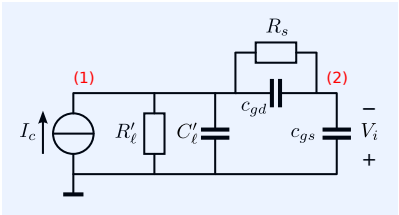


Figure 13.11: Alternative way of drawing the circuit from Figure 13.10.

Figure 13.11 shows the equivalent circuit for evaluation of the loop gain, drawn in a more convenient way. The total load resistance R'_ℓ is defined as

$$\frac{1}{R'_\ell} = g_{mb} + \frac{1}{r_o} + \frac{1}{R_\ell} \text{ [A/V]}. \tag{13.3}$$

The total load capacitance C'_ℓ is defined as

$$C'_\ell = C_\ell + c_{sb} \text{ [F]}. \tag{13.4}$$

DC loop gain

The DC value of the loop gain L_{DC} can be found from network inspection

$$L_{DC} = -g_m R'_\ell \text{ [-]}. \tag{13.5}$$

It has its maximum value $L_{DC\,\max}$ if $R_\ell = \infty$:

$$L_{DC\,\max} = -\frac{g_m}{g_{mb} + g_o} = -\frac{1}{n - 1 + \frac{1}{\mu}} \text{ [-]}, \tag{13.6}$$

where n is the substrate factor and μ the low-frequency voltage gain factor. The substrate factor varies with the source-to-bulk voltage. For a 180nm NMOS devices the substrate factor varies between 1.15 and 1.35 (see Binkley[Binkley2008][3]). The intrinsic voltage gain μ strongly depends on the

[3] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0

channel length and on the inversion coefficient.

A high substrate factor, or *body effect* strongly limits the loop gain in the CD stage. With $n = 1.35$ the maximum achievable DC loop gain is -2.86. This limits the DC voltage gain of the follower to 0.74.

Poles and zeros

The loop gain also comprises two poles and one zero. The zero z is caused by the parallel branch in series with the signal path formed by $R_s /\!\!/ C_{gd}$, as shown in Figure 13.11. We find

$$z = -\frac{1}{R_s c_{gd}} \text{ [rad/s]}. \tag{13.7}$$

This zero coincides with the pole in the asymptotic gain, which makes it a phantom zero.

The values of the two poles of the loop gain depend on the specific element values. Their product and their sum can be found from symbolic analysis of the loop gain.

The MNA equation of the circuit from Figure 13.11 can be written as

$$\begin{pmatrix} I_c \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{R'_\ell} + \frac{1}{R_s} + s\left(C'_\ell + c_{gd}\right) & -\frac{1}{R_s} - sc_{gd} \\ -\frac{1}{R_s} - sc_{gd} & \frac{1}{R_s} + s\left(c_{gs} + c_{gd}\right) \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}. \tag{13.8}$$

The loop gain L with reference variable g_m is defined as

$$L = g_m \frac{-V_2}{I_c} \text{ [-]}. \tag{13.9}$$

After solving the equations (application of Cramer's rule) we obtain

$$L = -\frac{R'_\ell g_m \left(1 + sR_s c_{gd}\right)}{1 + s\left(R_s\left(c_{gd} + c_{gs}\right) + R'_\ell\left(C'_\ell + c_{gs}\right)\right) + s^2 R'_\ell R_s\left(c_{gd}c_{gs} + C'_\ell c_{gd} + C'_\ell c_{gs}\right)} \text{ [-]}. \tag{13.10}$$

Let us now study the small-signal dynamic behavior, for cases in which the feedback capacitance is negligibly small:

$$c_{gd} \ll c_{gs} \text{ and } c_{gd} \ll C'_\ell. \tag{13.11}$$

In that case, the product of the two poles p_1 and p_2 can be approximated by

$$p_1 p_2 = \frac{1}{R'_\ell R_s C'_\ell c_{gs}} \text{ [rad}^2\text{/s}^2\text{]}. \tag{13.12}$$

The sum of the poles can then be approximated by

$$p_1 + p_2 = -\frac{1}{R'_\ell C'_\ell} - \frac{C'_\ell + c_{gs}}{R_s C'_\ell c_{gs}} \text{ [rad/s]}. \tag{13.13}$$

Achievable bandwidth

If both poles are dominant, the bandwidth B of the voltage transfer of the CD stage can be estimated from

$$B = \sqrt{(1 - L_{DC})\, p_1 p_2} \text{ [rad/s]}. \tag{13.14}$$

If the body effect causes the dominant limitation of the loop gain, we have $L_{DC} \approx -\frac{1}{n-1}$ and $R'_\ell \approx \frac{1}{g_m(n-1)}$. If we also use $f_T \approx \frac{g_m}{2\pi c_{gs}}$, the maximum achievable bandwidth of an RC loaded CD stage, driven from a resistive

source, is approximately:

$$B \approx \sqrt{\frac{2\pi n f_T}{R_s C'_\ell}} \text{ [Hz]}. \tag{13.15}$$

Hence, if we need to drive a capacitance C_ℓ from a source with an internal resistance R_s, a CD stage placed between source and load will increase the bandwidth of the voltage transfer if $2\pi n f_T R_s C_\ell > 1$.

Although the DC loop gain in the CD stage is relatively low, the poles of the voltage transfer may become complex, and frequency compensation for MFM behavior may be required. In such cases, phantom zero compensation can be achieved through insertion of a capacitance in parallel with the input of the CD stage.

In the following example we will demonstrate phantom zero compensation by placing a capacitance in parallel with the input of the CD stage.

Example 13.1

Let us study the behavior of a CD stage with an NMOS transistor fabricated in a standard CMOS18 process. The bulk of the NMOS is connected to the ground (see Figure 13.7). The transistor has a length of $1\mu m$*, a width of* $4\mu m$ *and operates in the forward saturation region at a drain current of* $250\mu A$*. The CD stage is driven from a resistive source with a source resistance of* $15k\Omega$*, and drives a capacitive load of* $250fF$*. The netlist of this circuit is shown below.*

```
CDcompM18
XU1 0 2 out 0 CMOS18N W={W} L={L} ID={ID}
V1 1 0 V value=0 dc=0 dcvar=0 noise=0
R1 2 1 {R_s}
C1 out 0 {C_ell}
C2 2 0 {C_phz}
.lib C18.lib
.end
```

We will evaluate the DC value, the poles, and the zeros of the source-to-load transfer. The SLiCAP *script that performs this analysis has been shown below:*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File CDstageCompensation.py

from SLiCAP import *
fileName = 'CDcompM18'
prj = initProject(fileName)

# Create the netlist if you didn't do it before
#makeNetlist(fileName + '.asc', fileName)

# Create an instruction object
i1 = instruction();

# Define the circuit
i1.setCircuit(fileName + '.cir')
#
i1.defPar('W', '4u')
i1.defPar('L', '1u')
i1.defPar('ID', '250u')
i1.defPar('R_s', '15k')
i1.defPar('C_ell', '250f')
i1.defPar('C_phz', 0)

# Display the circuit information on an HTML page
htmlPage('Ciruit data')
head2html('Schematic diagram')
img2html(fileName + '.svg', 500)
netlist2html(fileName + '.cir')
params2html(i1.circuit)

i1.setSimType('numeric')
i1.setSource('V1')
i1.setDetector('V_out')
```

```
# List the DC value, the poles, and the zeros of the gain
i1.setDataType('pz')
i1.setGainType('gain')
gainData = i1.execute()
listPZ(gainData)
```

The results of the pole-zero analysis are shown below. As a result of the substrate effect, the loop gain is very low and the low-frequency gain of the CD stage is only 0.69. *The poles are complex, but close to MFM positions:*

```
DC value of gain:  6.88e-1

Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0       -3.66e+08       -3.89e+08        5.34e+08   7.29e-1
 1       -3.66e+08        3.89e+08        5.34e+08   7.29e-1

Zeros of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0       -5.75e+09        0.00e+00        5.75e+09
```

The poles can be brought into MFM positions by adding a phantom zero capacitance C_{phz} in parallel with the input of the CD stage. It effectively increases c_{dg} of the transistor.

```
i1.defPar('C_phz', '1.7f')
# List the DC value, the poles, and the zeros of the gain after compensation
i1.setDataType('pz')
i1.setGainType('gain')
gainData = i1.execute()
listPZ(gainData)
```

With $C_{phz} = 1.7fF$ the poles are in MFM positions:

```
DC value of gain:  6.88e-1

Poles of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0       -3.67e+08       -3.67e+08        5.19e+08   7.07e-1
 1       -3.67e+08        3.67e+08        5.19e+08   7.07e-1

Zeros of gain:

 n  Real part [Hz]  Imag part [Hz]  Frequency [Hz]     Q [-]
--  --------------  --------------  --------------  --------
 0       -5.75e+09        0.00e+00        5.75e+09
```

The zero in the voltage transfer is due to the direct transfer ρ. This transfer has a first order low-pass character. This can be seen from network inspection by taking $g_m = 0$ in Figure 13.9. At very high frequencies, the transfer equals the direct transfer and the order of the high-frequency roll-off changes from two to one. This causes the zero at this transition frequency.

Figure 13.12 shows the Bode plots of the transfers of the asymptotic-gain model after compensation. It clearly shows the low value of the loop gain, which is the result of the substrate effect, and the influence of the direct transfer as described above.

The substrate effect of back-gate effect limits the loop gain in the CD stage. Much larges values of the loop gain can be obtained, if the bulk of the MOS transistor can be isolated from the substrate, and connected to the source. This, however, may also increase the load capacitance of the stage with the relatively large bulk-substrate capacitance. In such situations phantom zero compensation as described above, may be required.

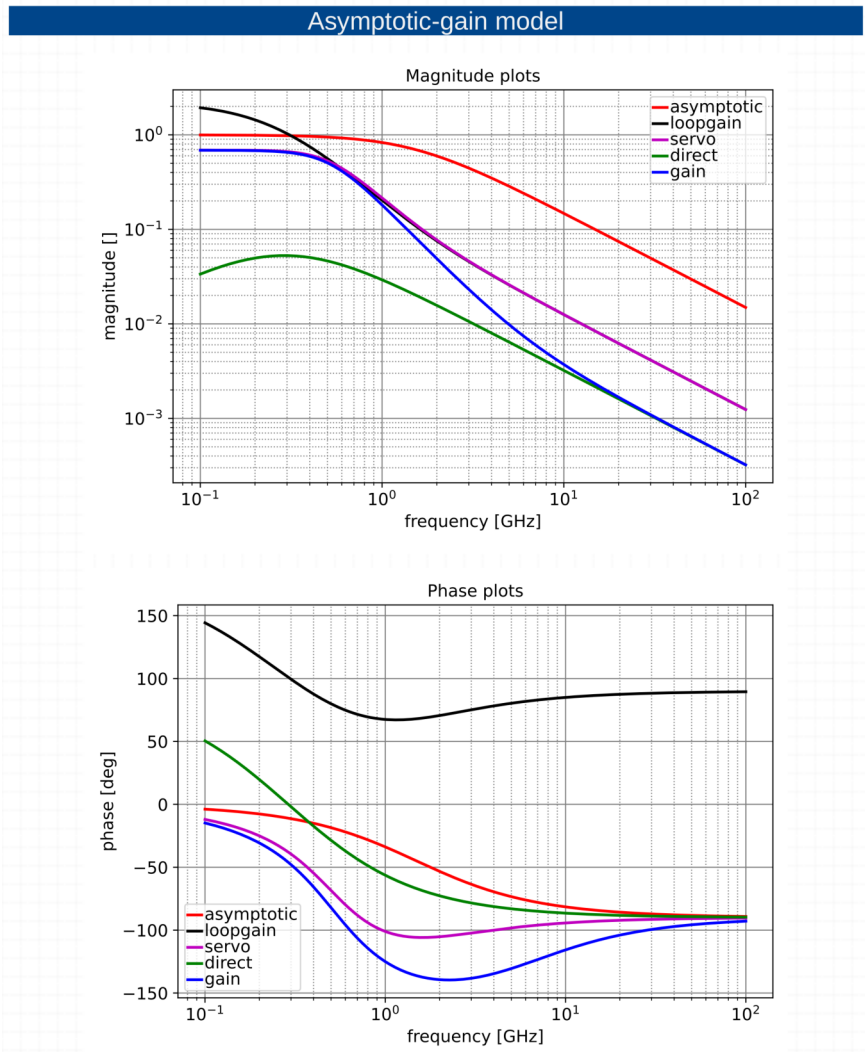


Figure 13.12: Bode plots of the transfers of the asymptotic-gain model, of the CD stage from Figure 13.7

13.2.3 Common Gate Stage

The common gate stage (CG-stage) is a nonenergic feedback stage with series (current) sensing at the load and parallel (current) comparison at the source. The feedback network consists of a short circuit between the inverting output of the controller and the source, and an open circuit between the inverting output of the controller and the ground. This feedback structure establishes $D = 1$, without affecting the other transmission parameters. The ideal current follower thus has an infinite output impedance and zero input impedance.

In this section we will discuss the noise behavior and the small-signal dynamic behavior of the CG stage.

Noise behavior

The CG stage is a nonenergic feedback stage. Hence, its equivalent input noise sources equal those of the controller. The controller in the CG stage is a CS stage and the optimization of its noise behavior for capacitive and resistive source has been discussed in section 5.4. The conclusions for the design of the noise performance of the CS stage are also valid for the CG stage.

Small-signal dynamic behavior

Similar as with the CD stage, the small-signal transfer from the source to the load can be analyzed using network analysis techniques. However, if we consider the CG stage a feedback version of the CS stage, and we apply the asymptotic gain model for performance analysis, we will obtain much more design information.

We will study the small-signal dynamic behavior of a CG stage driven from a current source with an internal impedance that consists of $R_s /\!\!/ C_s$, and loaded with $R_\ell /\!\!/ C_\ell$. The analysis for different source and load types proceeds in a similar way.

Figure 13.13 shows the signal diagram of the CG stage. The biasing circuitry has been omitted in the figure. The NMOS transistor is assumed to be biased in the forward operating region.

Figure 13.14 shows the small-signal equivalent circuit of the CD stage from Figure 13.13.

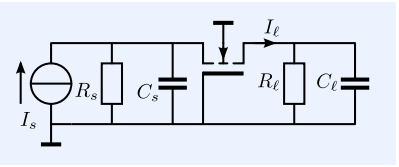


Figure 13.13: Signal diagram (bias circuitry not drawn) of a CG stage driven from an $R /\!\!/ C$ source and loaded with an $R /\!\!/ C$ network.

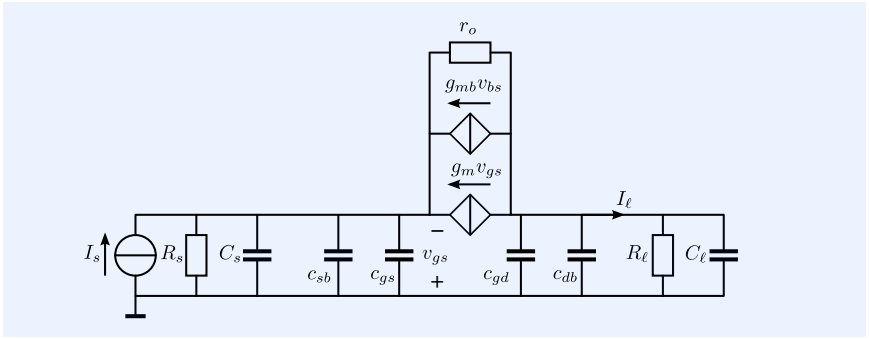


Figure 13.14: Small signal equivalent circuit of the CG stage from Figure 13.13.

Selection of the loop gain reference variable

In order to obtain meaningful results from the feedback analysis, the loop gain reference variable should be selected in such a way, that the controller obtains nullor properties if the reference variable approaches infinity. If we select g_m as loop gain reference, this will be the case at zero frequency. However, the capacitance $c_{gd} + c_{db}$ appears in parallel with the load. Its influence cannot be reduced by increasing the loop gain to infinity. This can be seen by redrawing the circuit as depicted in Figure 13.15. This figure shows that the capacitance $c_{gd} + c_{db}$ establishes an attenuator between the output of the controller and the load.

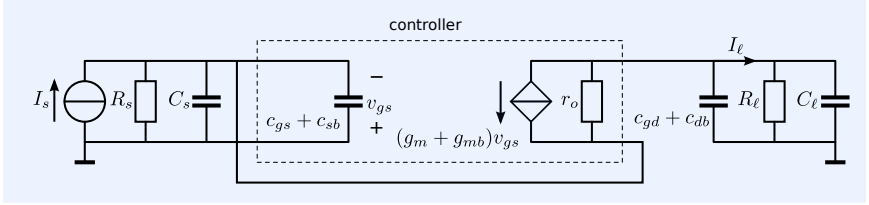


Figure 13.15: Alternative way of drawing the circuit from Figure 13.14. It clearly shows the series feedback at the load (load current sensing) and the parallel feedback at the soure (source current comparison).

Asymptotic gain and ideal gain

The asymptotic gain can be found from network inspection using the circuit representation from Figure 13.15. The sum of the two transconductances is selected as loop gain reference:

$$A = g_m + g_{mb} \text{ [-]}. \tag{13.16}$$

If $A \to \infty$ the controller behaves as a nullor and the asymptotic gain $A_{i\infty}$

is found as

$$A_{i\infty} = \left.\frac{I_\ell}{I_s}\right|_{A\to\infty} = \frac{1+sR_\ell C_\ell}{1+sR_\ell\left(C_\ell + c_{gd} + c_{db}\right)}\ [\text{-}]. \tag{13.17}$$

Evaluation of the loop gain

Figure 13.16 shows the equivalent circuit for evaluation of the loop gain. According to the asymptotic gain model, $\lambda\beta\kappa$ is defined as

$$\lambda\beta\kappa = \left.\frac{V_i}{I_c}\right|_{I_s=0}\ [\text{A/V}]. \tag{13.18}$$

In order to obtain compact expressions we use the representation from Figure 13.17

The MNA equations in matrix form can be formulated as

$$\begin{pmatrix} -I_c \\ I_c \end{pmatrix} = \begin{pmatrix} \frac{1}{R_\ell}+\frac{1}{r_o}+sC'_\ell & -\frac{1}{r_o} \\ -\frac{1}{r_o} & \frac{1}{R_s}+\frac{1}{r_o}+sC'_s \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}. \tag{13.19}$$

The loop gain L with reference variable $g_m + g_{mb}$ is defined as

$$L = (g_m + g_{mb})\,\frac{-V_2}{I_c}\ [\text{-}]. \tag{13.20}$$

After solving the equations (application of Cramer's rule) we obtain

$$L = -\frac{R_s r_o\,(g_m+g_{mb})}{R_s+r_o+R_\ell}\,\frac{1+sC'_\ell R_\ell}{1+s\left(\frac{C'_s R_s(r_o+R_\ell)}{R_s+r_o+R_\ell}+\frac{C'_\ell R_\ell(R_s+r_o)}{R_s+r_o+R_\ell}\right)+s^2\frac{C'_s C'_\ell R_s r_o R_\ell}{R_s+r_o+R_\ell}}\ [\text{-}]. \tag{13.21}$$

DC loop gain

The DC loop gain is found as

$$L_{DC} = -\frac{R_s r_o\,(g_m+g_{mb})}{R_s+r_o+R_\ell}\ [\text{-}]. \tag{13.22}$$

Under ideal drive and load conditions: $R_s \to \infty$ and $R_\ell \to 0$, the DC loop gain obtains its maximum value $L_{\max}$,which equals

$$L_{\max} = -r_o\,(g_m+g_{mb})\ [\text{-}]. \tag{13.23}$$

This can be written as the product of the substrate factor and the voltage gain factor

$$L_{\max} = -n\mu\ [\text{-}]. \tag{13.24}$$

Under the given conditions, the current gain A_i of the CG stage equals unity, even though the loop gain is not infinite. This is because under ideal drive and load conditions, the direct transfer ρ equals unity. According to the asymptotic gain model we obtain the source-to-load transfer as

$$A_i = A_{i\infty}\frac{-L}{1-L} + \frac{\rho}{1-L}\ [\text{-}]. \tag{13.25}$$

Under ideal drive and load conditions we have $\rho = 1$, which yields $A_i = A_{i\infty} = 1$.

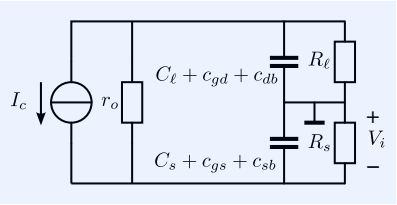


Figure 13.16: Circuit for evaluation of the loop gain of the CG stage from Figure 13.15.

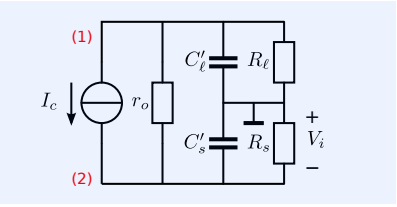


Figure 13.17: Compact representation of the circuit for evaluation of the loop gain of the CG stage from Figure 13.15.

Poles and zeros

The poles and the zeros of the loop gain can be found from (13.21). This expression shows that the loop gain has one zero and two poles. If $R_\ell \ll r_o + R_s$, one of the poles coincides with the zero and the loop gain has a first order low-pass character with a pole at $s = -\frac{r_o+R_s}{r_o R_s C_s'}$. Hence, it is determined by the parallel connection of r_o, R_s and C_s'. The bandwidth B of the current gain is then found from the first order loop gain poles product:

$$B = \left(1 + \frac{r_o R_s \left(g_m + g_{mb}\right)}{R_s + r_o}\right) \frac{r_o + R_s}{r_o R_s \left(C_s + c_{gs} + c_{gb} + c_{bs}\right)} \text{ [rad/s]}. \quad (13.26)$$

If the DC loop gain is much larger than unity, this may be simplified to

$$B = \frac{g_m + g_{mb}}{C_s + c_{gs} + c_{gb} + c_{bs}} \text{ [rad/s]}. \quad (13.27)$$

Hence, if the stage is driven from a high impedance ($C_s + c_{bs} \ll c_{gs}$), and it is terminated with a low impedance ($R_\ell \ll r_o + R_s$), the bandwidth equals about ω_T of the transistor.

This makes this stage very well suited as a high-frequency buffer between two CS stages. This application will be discussed at a later stage.

13.2.4 Passive feedback stages

The CD stage and CG stage are two nonenergic negative feedback stages that have the CS stage as controller. We will now discuss single-transistor passive feedback stages. The synthesis of amplifiers exploiting passive feedback has been discussed in Chapter 7. Passive feedback stages use feedback networks that consist of passive elements. Due to energy storage and/or power dissipation, and due to noise generation or signal attenuation, the power efficiency and the noise performance of passive feedback amplifiers are not as good as those of nonenergic feedback amplifiers.

Figure 13.18 gives an overview of passive feedback amplifier stages that have a CS stage as controller. The following stages are shown:

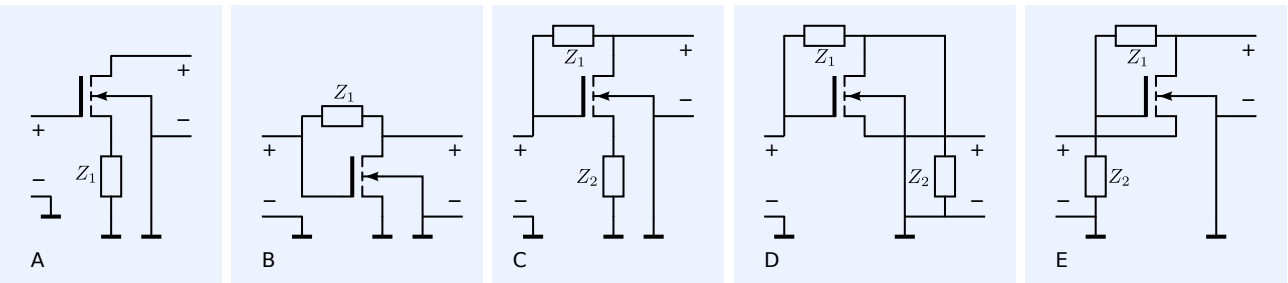


Figure 13.18: Passive feedback stages with a CS stage as controller.

A The series stage is a single-stage voltage to current converter. The impedance Z_1 senses the output current and converts it into a voltage that is subtracted from the input voltage. It establishes parameter B of the stage. It degrades the noise performance of the CS stage as if it is placed in series with the source. The power efficiency and the energy storage are affected as if this impedance is placed in series with the load. Due to the series feedback, the output impedance and the input impedance are both larger than those of the CS stage.

B The shunt stage is a single-stage current to voltage converter. The impedance Z_1 senses the output voltage and converts it into a current that is subtracted from the input voltage. It establishes parameter C of the stage. It

degrades the noise performance of the CS stage as if it is placed in parallel with the source. The power efficiency and the energy storage are affected as if this impedance is placed in parallel with the load. Due to the parallel feedback, the output impedance and the input impedance are both smaller than those of the CS stage.

C The gyrator-like stage exhibits series feedback at the input and at the output as well as parallel feedback at the input and at the output. The feedback networks intentionally establish nonzero parameters for B and C of the stage. However, due to interaction of the sensing and comparison networks, all four transmission parameters of the stage are determined by the feedback networks, although not independently.

D The transformer-like configuration has the parameters A and D intentionally fixed to a nonzero value by means of negative feedback. In this configuration we have $A = 1$.

E This version of the transformer-like configuration has $D = 1$.

13.3 Application of balancing

Application of balancing techniques can be twofold:

1. The controller of the negative feedback amplifier consists of a balanced stage (anti-series or complementary-parallel CS-stage).

   Some examples of these configurations will be discussed in section 13.3.1.

2. A negative feedback amplifier consists of two anti-series or complementary-parallel connected local-feedback stages.

   Some examples of these configurations will be discussed in section 13.3.2.

13.3.1 Local feedback with balanced CS stage

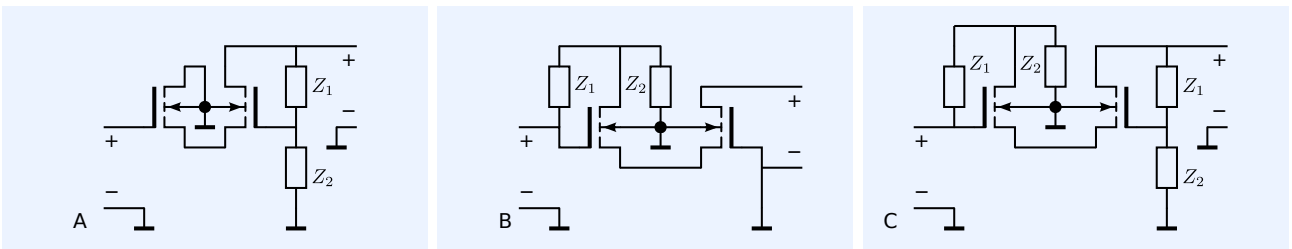


Figure 13.19: Passive feedback stages with an anti-series CS stage as controller.

Balanced amplifier stages can be applied as controller in negative feedback amplifiers. Anti-series connected CS stages can be used as four-terminal controllers and complementary-parallel stages can be used as push-pull stages with a high current drive capability. Figure 13.19 shows some examples of local feedback stages with an anti-series connected CS stage as controller. The application of complementary-parallel connected stages as controller does not give new configurations than those with a single CS stage as controller.

13.3.2 Balanced local feedback amplifier stages

Anti-series connection of three-terminal local feedback stages can be applied to obtain amplifiers with floating ports. Figure 13.20A shows two anti-series connected series stages.

Figure 13.20B shows two complementary-parallel connected CD stages. This combination results in a voltage follower with high current source and sink capability.

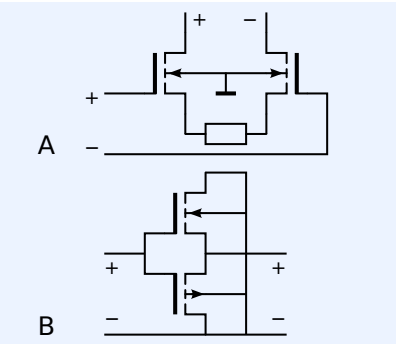


Figure 13.20: Balanced local feedback stages
A: Anti-series connection of two series stages
B: Complementary parallel connection of two CD stages.

13.4 Indirect feedback stages

Indirect feedback stages use indirect sensing and/or indirect comparison techniques. In case of indirect sensing, a copy of the load voltage or the load current is sensed and converted by the feedback network. This generally is less accurate than direct sensing.

In case of indirect comparison, a copy of the source voltage or current is compared with the output quantity of the feedback network. In this section we will give two examples of indirect feedback stages: the current mirror and the voltage mirror.

13.4.1 Current mirror

Figure 13.21A shows the concept of an inverting current amplifier using indirect feedback. A copy of the load current is sensed, attenuated and compared with the source current. The source-to-load transfer can be found as

$$\frac{I_\ell}{I_s} = -\frac{Z_i + Z_2}{Z_2}\left(\frac{A\frac{Z_1 Z_2}{Z_1+Z_z} + B}{AZ_\ell + B}\right). \tag{13.28}$$

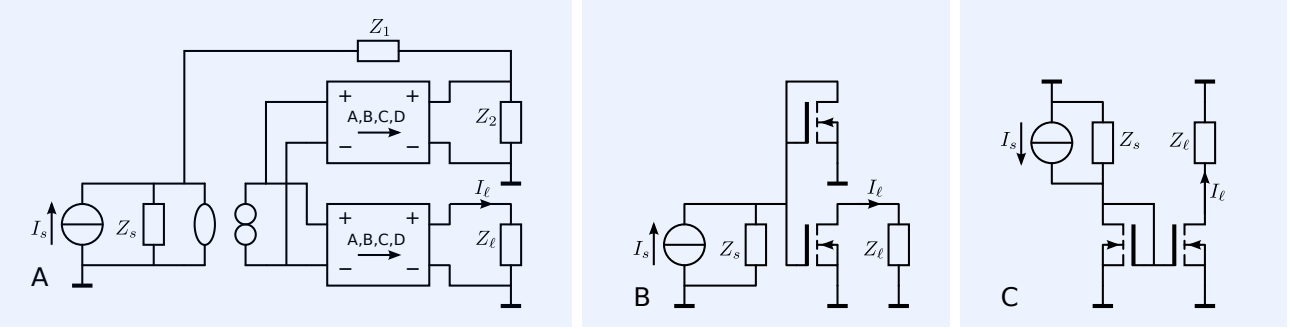


Figure 13.21: Inverting indirect feedback current amplifier.
A: Concept of an inverting current amplifier using indirect current sensing.
B: Implementation with unity gain feedback, while using the two port with feedback as controller.
C: The circuit redrawn known as the current mirror.

The above expression clearly shows that the ideal current transfer (nullor as controller) depends on the load impedance if $Z_\ell \neq \frac{Z_1 Z_2}{Z_1+Z_z}$.

Figure 13.21B shows an implementation of this concept with two CS stages. This circuit is known as the current mirror and usually drawn as shown in Figure 13.21C.

13.4.2 Voltage mirror

Figure 13.22A shows the concept of an inverting voltage amplifier that uses indirect voltage comparison. The load voltage is directly sensed and attenuated. The attenuated version of the load voltage is indirectly compared with the source voltage.

Figure 13.22B shows the implementation with two CS stages known as the voltage mirror.

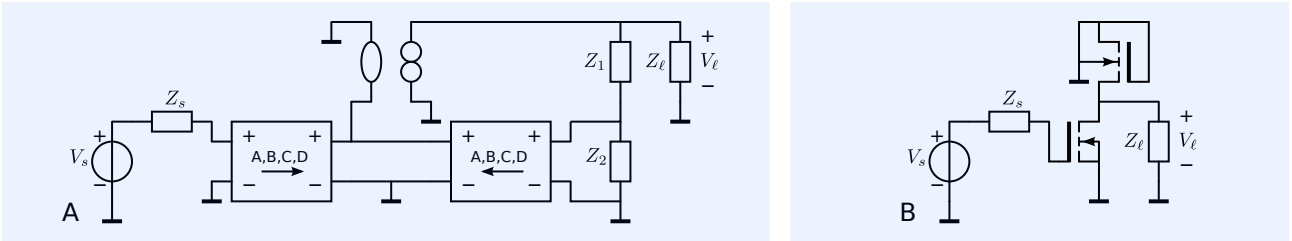


Figure 13.22: Inverting indirect feedback current amplifier.

A: Concept of an inverting voltage amplifier using indirect voltage comparison.

B: Voltage mirror: an implementation with unity gain feedback, while using the two port with feedback as controller.

14
Multi-stage Feedback Amplifiers

14.1 Introduction

In this chapter, we will discuss the design of multi-stage negative feedback amplifiers. We will formulate considerations for the design of the various performance aspects of such an amplifier, and present a step-by-step approach for the design of high-performance negative feedback amplifiers.

14.1.1 Summary of previous chapters

The design approach presented in this chapter is based upon the theory discussed in the preceding chapters. We will briefly summarize the conclusions of these chapters below.

Modeling and characterization of amplifiers

In Chapter 2, we started with an introduction to amplifiers. We have seen that an amplifier has to provide its load with an amplified copy of the source signal. The word 'amplified' in this context means that the available signal power at the amplifier's output exceeds that of the source.

We continued with the classification and characterization of amplifiers. We have seen that amplifiers can be classified according to their source and load characteristics. The appropriate amplifier type for a specific source and load can be found by answering the following questions:

1. What would be the ideal value of the amplifier's input impedance, or alternatively, which electrical quantity of the source (current or voltage) should be taken as input quantity for the amplifier?
2. What would be the ideal value of the amplifier's output impedance, or, alternatively, which electrical quantity should be used to drive the load?
3. What is the desired source-to-load transfer that needs to be established?
4. Which ports (input, output and power port) should be isolated with respect to each other?
5. Is unilateral behavior required?

The answers to these questions together define the functional behavior of the amplifier. They give us the ideal values of the required port impedances[1] and of the transfer characteristics. We have seen that the functional behavior of an amplifier can be described with the aid of two-ports. The Transmission-1 matrix parameters A, B, C and D, in conjunction with the source and load impedances, provide high-level descriptions for the source-to-load transfer and the port impedances of the idealized amplifier.

[1] Diferential-mode and common-mode port impedances.

Practical amplifiers suffer from non-idealities. Like all information processing systems, imperfections in their behavior arise as a result of the physical limitation of speed and power, the addition of noise and imperfections in the physical operating principle used for their implementation. These imperfections limit the amount of information that can be processed by the amplifier. We have seen that application-specific error description methods should be used to reflect the effects of the information processing errors, as they will be observed in the application.

Another error source in amplifiers is the imperfect port isolation. It introduces susceptibility to power supply noise and common-mode noise. We have seen that the regularly used description methods with the (voltage) CMRR, the (voltage) common-mode rejection factor and the (voltage) PSRR are far from complete. Application-specific test benches are required to complete such specifications.

Amplification mechanism

In Chapter 3, we studied the amplification mechanism. We have seen that the application of specific passive devices together with (active) bias sources, provide the means for building electronic amplifiers. Because of their amplifying capabilities, such passive devices are usually called *active devices*.

Device modeling

In order to understand the amplification mechanism and its physical limitations, we briefly discussed the construction, the operation and the modeling of active devices in Chapter 4. We have seen that properly biased bipolar junction transistors (BJTs), junction field-effect transistors (JFET) and MOSFETs can provide a large available power gain.

In that chapter, we also presented simplified device models for symbolic analysis. Such models can be used for finding methods to affect specific behavioral aspects of interest by design and setting up design equations.

CS stage

The common-source (CS) stage can be considered as the basic MOS amplifier stage; its performance is discussed in Chapter 5. When properly biased, this stage provides a large available power gain. The name of this stage refers to the fact that the source is the common terminal for the input port and the output port. Hence, port isolation with this stage is not possible.

The drain-gate capacitance causes non-unilateral behavior and may seriously limit the bandwidth of the stage or the stage's contribution to the loop gain-poles product in a negative feedback amplifier. Such effects can be kept small by shorting the stage, while taking the short circuit output current as information carrying quantity.

The noise addition by a CS stage can be minimized by optimizing both its device geometry and its drain current. In high-frequency applications, the device should operate in strong inversion and at a high cut-off frequency. At low frequencies, the influence of flicker-noise may become dominant and a longer channel and a lower operating current may be optimal.

The nonlinearity of the CS stage strongly depends on the frequency, the drive and termination conditions and the operating conditions.

The accuracy of the source-to-load transfer is limited due to fabrication tolerances. All performance aspects of the CS stage depend on temperature.

CE stage

The common-emitter or CE stage can be considered as the basic three-terminal BJT amplifier stage. The name of this stage refers to the fact that the emitter is the common terminal for the input port and the output port. Hence, port isolation with this stage is not possible.

The collector-base capacitance causes non-unilateral behavior and may seriously limit the bandwidth of the stage or the stage's contribution to the loop gain-poles product in a negative feedback amplifier. Such effects can be kept small by shorting the stage, while taking the short circuit output current as information carrying quantity.

The noise addition by a CE stage can be minimized by optimizing its device geometry and its collector current. Low-noise amplification with a CE stage requires a transistor with a large DC current gain factor, a high cut-off frequency and a small base resistance.

The nonlinearity of the CE stage strongly depends on the frequency and the drive and termination conditions of the stage. At low frequencies, the parameters A and D can have a relatively low differential-gain, while the pa-

rameter B is inversely proportional to the collector current and the parameter C is proportional to the collector current. At high frequencies, the nonlinearity is strongly dominated by the nonlinear $Q - V$ relations of the junctions and the voltage and current dependency of the transit time.

Even when accurately biased, the accuracy of the small-signal source-to-load transfer of a CE stage is limited due to fabrication tolerances. Since most device parameters depend on temperature, all performance aspects of the CE stage will be temperature dependent. Hence, accurate amplification with a single CE stage is then only possible with properly selected devices over a small temperature range.

Single-stage amplifiers

In many RF applications, the accuracy and the temperature stability of the gain are not the dominant requirements. In such applications, automatic gain control is often used to stabilize the gain. Usually, low-noise and low-distortion operation over a limited frequency range, is then of primary interest. In such applications, and at low signal levels, a single CE stage or CS stage amplifier may perform well enough.

Straightforward design of the different performance aspects of the CE stage or the CS stage is only possible if design parameters can be found that more or less independently fix the different performance aspects.[2] Usually, we only have a limited number of design parameters at our disposal to fix all kinds of different performance aspects. The most important design parameters are:

[2] Orthogonal design.

1. The device type and geometry
2. The operating current
3. The operating voltage

With these three parameters we are not able to design all the performance aspects independently. An extra degree of freedom can be obtained through application of impedance transformation techniques at the source and/or at the load. Impedance transformation can be implemented with wide-band transformers, resonant networks or transmission lines.

In many situations the demand for high-quality information processing requires improvement of performance of the CE or CS stage. Such improvements can be obtained through application of error reduction techniques.

Balanced amplifier stages

In Chapter 6, we discussed the application of balancing techniques and introduced the anti-series stage or *differential pair* and the complementary-parallel stage or *push-pull stage*.

Balancing is a compensation technique will limited error reduction capabilities. With balancing techniques, we can improve the following performance aspects:

1. Linearity

   Ideally, the anti-series and the complementary-parallel connection of CE or CS stages have odd transfer characteristics. Even terms such as the offset and its associated temperature drift are compensated. Differential-mode bias sources are replaced with common-mode bias sources.

2. Port isolation

With the aid of anti-series connection of CE or CS stages, we obtain so-called differential pairs that can be considered as basic four-terminal amplifier stages. They have improved isolation between the input port and the output port.

3. Power efficiency

   With the aid of complementary-parallel connected CE or CS stages, we obtain push-pull stages. These stages can source and sink currents that exceed the quiescent operating current of its constituting CE or CS stages.

The odd order nonlinearity, the dynamic behavior, the noise behavior, as well as the accuracy and the temperature dependency of the transfer, however, cannot be improved with balancing.

Negative feedback amplifiers

A significant improvement of the quality of the amplifier's signal transfer can be obtained through application of negative feedback. Negative feedback is a powerful error reduction technique in which available gain of amplifying devices is used for quality improvement of the source-to-load transfer. Each transmission parameter of a negative feedback amplifier can be given an accurate value with the aid of a feedback loop around a high-gain controller. The task of this controller is to nullify the error between the actual source-to-load transfer and the one determined by the feedback networks.

Negative feedback has stronger error-reduction capabilities than balancing. The noise performance and the influence of input offset and bias sources, and the load drive capability cannot be improved by negative feedback, they at best equal those of the controller. The accuracy, the linearity and the bandwidth are at best determined by the feedback networks. Such networks can be realized with accurate passive components.

The synthesis of negative feedback amplifier configurations is discussed in Chapter 7. All unilateral and non-unilateral, isolated and non-isolated amplifier types that have been defined in Chapter 2, can be realized with the aid of negative feedback. The basic synthesis techniques for negative feedback amplifiers are:

1. Sensing of the load quantity
2. Design of a feedback network that converts the sensing result into an accurate copy of the source quantity
3. Comparison of this copy with the source quantity
4. Nullification of the comparison result with the aid of a high-gain controller.

If nonenergic feedback elements are used, the noise performance and the power efficiency of the negative feedback amplifier equal those of its controller. Practical use of nonenergic feedback is often limited to unity-gain feedback. Amplifiers with a gain that differs from unity are often realized with passive feedback.

Passive feedback networks increase the contribution of the equivalent input noise sources of the controller to the total noise. If those networks comprise resistive elements, they also contribute noise themselves.

Passive feedback networks may also contribute to the energy storage or dissipation of the amplifier. As a result, the power efficiency of the amplifier will be less than that of its controller.

Not all amplifier configurations can be realized with the aid of passive feedback around a single controller. This is because feedback elements that

behave as natural two-ports cannot be constructed using exclusively passive elements. Configurations that cannot be realized with passive feedback around a single controller, can be realized with balanced feedback, active feedback or indirect feedback techniques.

The design of the feedback configuration is the first step in the design of negative feedback amplifiers. In this step the controller is assumed to have nullor properties, which means that it has an infinite available power gain and that it behaves as a natural two-port. The gain of a negative feedback amplifier with a nullor as controller is called the *ideal gain*.

Feedback modeling

The second step in the design of negative feedback amplifiers, is the design of the controller. This step requires knowledge about the relation between the performance aspects of the controller and those of the negative feedback amplifier which it is part of. The asymptotic gain model provides this information; it is presented in Chapter 10. If the loop gain reference variable is chosen such that an infinite loop gain turns the controller into a nullor, the design of negative feedback amplifiers can be performed in two steps:

1. Design of the feedback network with a nullor as controller
2. Design of the controller

Controller design

In Chapter 11, we related the performance aspects of the controller to those of the negative feedback amplifier. We obtained the following design conclusions:

1. If the loop gain reference variable has been selected properly, the servo function is a measure for the deviation of the source-to-load transfer from the ideal gain. The servo function is completely defined by the loop gain, which comprises contributions of the controller, the source impedance, the load impedance and the feedback networks. The ideal value of the servo function is unity.
2. The relative inaccuracy of the servo function at midband frequencies is determined by the mid-band loop gain.
3. The low-pass cut-off frequency of the servo function is determined by the product of the dominant poles of the loop gain and the midband value of the loop gain.
4. The differential gain of the servo function at midband frequencies, equals the ratio of the differential gain of the loop gain and the value of the loop gain at mid-band frequencies.

Frequency compensation

With a sufficiently large loop gain-poles product, the negative feedback amplifier can obtain its required bandwidth. However, this does not mean that the error with respect to the ideal behavior of the amplifier is small. For this to be the case, the servo function should have an all-pole transfer with its poles in, or close to MFM positions. In most cases we will have to apply frequency compensation techniques to manipulate the poles into these positions. These techniques have been discussed in Chapter 12. We have seen that compensation with phantom zeros minimally affects other performance aspects, such as noise and over drive recovery, while maintaining the bandwidth of the servo function close to its designed value.

Local feedback amplifier stages

The application of negative feedback around a single CE or CS stage results in local feedback stages. These stages have been studied in Chapter 13. Commonly used nonenergic local feedback stages are the emitter follower, the source follower, the CB stage and the CG stage. Single-loop, single-transistor, passive feedback stages are the series stage and the shunt stage. Dual-loop, single transistor feedback stages, as well as stages that exploit indirect feedback, such as the voltage mirror and the current mirror, have also been introduced in Chapter 13. Balanced, local feedback stages are obtained after applying balancing techniques to the basic feedback stages.

The properties of feedback stages can easily be predicted by considering the behavioral modifications that are a result of negative feedback:

1. Each feedback loop fixes one transmission parameter.
2. Parallel feedback at a port reduces the port impedance (ideally to zero).
3. Series feedback at a port increases the port impedance (ideally to infinity).
4. Both series and shunt feedback at a port establishes a finite non-zero port impedance primarily defined by the feedback networks.

The properties of balanced feedback stages can easily be predicted by considering the behavioral modifications that are the result of balancing:

1. Balancing provides odd characteristics.
2. Differential-mode bias sources convert into common-mode bias sources.
3. Anti-series connection provides a four-terminal stage with a current limiting charachter.
4. Complementary-parallel connection provides push-pull stages with a voltage limiting character. Those stages can source and sink currents that exceed the stage's quiescent operating current.

14.1.2 This chapter

In this chapter, we will put all the knowledge obtained from the previous chapters together, and define an approach for the design of multi-stage negative feedback amplifiers.

In section 14.2, we will formulate basic considerations for the design of the controller. The following topics that be discussed:

1. Design of the input stage
2. Design of the output stage
3. Design of the number of stages
4. Interconnection of stages
5. Interconnection of the feedback network and the controller
6. The use of cascode stages
7. Application of local feedback stages

Chapter 15 is devoted to biasing of the stages in the controller.

14.2 Controller design considerations

In this section, we will formulate the answers to the following design questions.

1. Which type of stage should be applied as input stage of the controller?

   We have seen that the equivalent input noise sources of a negative feedback amplifier, at best equal those of its controller. It will become clear that if we select the proper type of input stage, the noise performance of the controller will predominantly be defined by that of the first stage. Hence, at an early stage of the design, the noise contributions of other stages can be ignored. The design of the first stage of the controller will be discussed in section 14.2.1.

2. Which type of stage should be applied as output stage of the controller?

   We have seen that the voltage and current drive capability of a negative feedback amplifier, at best equal those of its controller. If we select the proper type of output stage, it will also be the dominant contributor to the nonlinearity. Hence, at an early stage of the design, contributions to the nonlinearity of other stages can be ignored. The design of the output stage of the controller will be discussed in section 14.2.2.

3. How many stages are required in the controller?

   A rough estimation of the required number of stages at an early stage of the design is important. In practice, the use of more than two or three stages in a multi-stage negative feedback amplifier complicates the design of the frequency response and the stability. If more that two or three stages are required, one may consider the use of cascaded negative feedback amplifiers for the amplifier or for the controller, as discussed in section 12.9. The minimum number of stages that is required for the controller will be discussed in section 14.2.3.

4. In which way should the stages be interconnected?

   The non-unilateral behavior of the basic amplifier stages due to the collector-base capacitance in BJTs or the drain-gate capacitance in field effect devices, causes interactions between stages that complicate the dynamic behavior of interconnected stages. In addition, improper interconnection of amplifier stages causes the gain to deviate from the ideal gain. Such a deviation cannot be reduced by increasing the loop gain, and should be avoided in high-accuracy amplifiers. The aspects of the interconnection of the stages will be discussed in section 14.2.3.

5. In which way should the biasing of the controller be arranged?

   During the design of the signal path, the stages are assumed to be biased with four ideal sources. This has been discussed in Chapter 3. Although this biasing method is very unpractical, it allows us to separate the design of the signal path from the design of the biasing. As soon as we have a feasible solution for the signal path of the amplifier, we need to convert this theoretical biasing scheme into a practical one. This means that all bias sources have to be derived from the available power supply sources and that measures have to be taken to ensure proper bias conditions over the required temperature range and for all device tolerances. In Chapter 15 we will pay attention to the biasing of the amplifier stages with sources that can be derived from the power supplies. However, we will not discuss the implementation of the bias sources themselves. This will be discussed in a separate chapter in a future edition of this book.

14.2.1 Design of the input stage

The equivalent input noise sources of a negative feedback amplifier, at best equal those of its controller. Hence, a low-noise amplifier requires a low-noise controller. A low-noise controller requires a minimum number of amplifier stages that contribute as little as possible to the source-referred noise.

Type of input stage

We have seen that if the first stage of the amplifier has nullor properties, noise of the following stages does not contribute to total source-referred noise. Hence, the preferred type of input stage for the controller is a stage that does not have intentional feedback. Such stages have their transmission parameters (A, B, C and D) as small as possible. Hence, a CE or CS stage, or one of their balanced versions is preferred as input stage.

Local feedback stages have at least one of these transmission parameters increased with respect to that of the CE or CS stage. Let us, for example consider a CD stage as input stage of the controller. Such a stage has $A = 1$, so the noise voltage of the second stage fully contributes to the input voltage noise of the controller.

Application of local feedback stages

If, for some reason, CE stages or CS stages or one of their balanced versions cannot be applied as input stage, one has to apply a local feedback stage. In such situations, nonenergic local feedback stages are preferred over passive feedback stages. This is because the feedback elements in passive feedback stages increase the contribution of the equivalent input noise of sources of their controller and, in cases those feedback elements have a real part, they contribute noise themselves.

Conclusions

The order of preference for the input stage of the controller of a negative feedback amplifier is:

1. CE or CS stage or one of their balanced versions
2. CB, CG, CC or CD stage or one of their balanced versions
3. Passive feedback stages or one of their balanced versions.

The selection of the device type, its geometry and its operating conditions is governed by the source impedance, the frequency range of interest and the noise requirements. This has been discussed in Chapter 5 for the CS stage. We have seen that these conclusions also hold for their balanced versions.

14.2.2 Design of the output stage

The voltage and current drive capability of a negative feedback amplifier, at best equal those of its controller. Hence, the drive capability of the output stage is the most important performance aspect to be dealt with during its design.

In Chapter 11, we have seen that the differential gain error of the servo function equals the quotient of the differential gain error of the loop gain and the loop gain. A small differential gain error is thus obtained if the smallest number of stages contribute as little as possible to the differential gain error, and as much as possible to the loop gain.[3] Since only stages that have non-zero signal excursions contribute to the differential gain, their

[3] A stage has to contribute maximally to the differential error to gain ratio.

number has to be minimized. This can be done by using an output stage that has its Transmission-1 matrix parameters as small as possible.

Type of output stage

If the output stage has nullor properties, signal excursions in preceding stages are zero and the output stage is the only contributor to the differential gain. From this we see that the preferred type of output stage is a CE or CS stage or one of their balanced versions.

Application of local feedback stages

If, for some reason, CE or CS stages or one of their balanced versions cannot be applied as output stage, one has to apply a local feedback stage. In such situations, nonenergic local feedback stages are preferred over the passive ones. This is because impedances in series or in parallel with the signal path of the passive feedback stages usually increase the differential error to gain ratio.

Conclusions

The order of preference for the output stage of the active part of a negative feedback amplifier is:

1. CE or CS stage or one of their balanced versions
2. CB, CG, CC or CD stage or one of their balanced versions
3. Passive feedback stages or one of their balanced versions.

The selection of the device type, its geometry and its operating conditions is governed by the current and voltage drive requirements of the output stage, and by the power efficiency of the amplifier. A high-efficiency output stage uses a complementary-parallel stage (push-pull stage).

14.2.3 Design of the number of stages

If the type of stage, the requirements for its geometry and operation conditions for the input stage overlap with the corresponding properties of the output stage, a single-stage controller might be an option. The requirements for the midband accuracy, the bandwidth and the differential gain error of the servo function then determine the number of stages to be used. If a single-stage solution satisfies all requirements, there is no need for a multi-stage controller.

If the requirements of the first and the second stage do not overlap, at least two stages are required: an input stage and an output stage. A two-stage solution is then feasible if:

1. The midband accuracy of the servo function is large enough.
2. The bandwidth of the servo function is large enough.
3. The differential gain error of the servo function is small enough.

   This includes the requirements that the current drive capability and the voltage drive capability of the first stage are large enough to drive the output stage.

If at least one of the above requirements[4] is not met, more stages are required.

[4] Over the temperature range of interest and for all possible device tolerances.

Type of intermediate stages

In order not to complicate the design of the dynamic performance of the negative feedback amplifier, the number of dominant poles of the loop gain should be kept as small as possible. Hence, we would like to realize a sufficiently small inaccuracy, large bandwidth and low distortion, with a minimum number of stages. This can be achieved by letting the stages contribute as much as possible to the product of the loop gain and the poles, and as little as possible to the differential error-to-gain ratio of the loop gain.

The contribution of CE and CS stages and their balanced versions to the mid-band value of the loop gain, depends very much on the drive and load conditions of the stages. Generally this contribution is maximized if no impedances are inserted in series or in parallel with the signal path.

CE and CS stages and their balanced versions maximally contribute ω_T to the loop gain-poles product. This is the case if the stage is driven from a current source and shorted at its output, while the short circuit output current is passed to the next stage. If it is not shorted, and if the voltage gain of a capacitively loaded stage is large, pole splitting may drive one of the poles of the transfer out of the dominant group, which reduces the bandwidth of the servo function.

CE and CS stages and their balanced versions also contribute as little as possible to the differential error to gain ratio of the loop gain. Balanced stages mainly contribute to odd order distortion, while unbalanced stages contribute to both even and odd distortion. Application of local feedback stages causes a larger contribution to the differential error-to-gain ratio of preceding stages, while passive feedback stages usually have a larger differential error-to-gain ratio than nonenergic feedback stages, or stages without feedback.

Cascode stages

From the above, we may conclude that all amplifier stages preferably are CE or CS stages, or their balanced versions. The high-frequency behavior of cascaded CE or CS stages, however, is very complicated. The parasitic capacitive feedback in these stages causes a strong interaction between them. The interaction between stages can considerably be reduced by using unity gain current amplifiers as buffer between cascaded CE or CS stages. These stages effectively short the first stage, while passing the short circuit to the next stage. The CB or CG stages can be used for this purpose. These unity-gain current amplifier stages have a low input impedance and a high output impedance for a wide range of drive and termination impedances. The cascade connection of a CE or CS stage with a CB or CG stage is called a *cascode stage*. The use of cascode stages strongly facilitates the design of the dynamic behavior of a negative feedback amplifier.

Another advantage of the cascode stage is its high output impedance. This is a result of output the series feedback in the CG or CB stage. As a result, the Transmission-1 matrix parameters A and C of the cascode stage are much smaller than those of its CS stage.[5] This makes a cascode stage a better approximation of the nullor than its CS or CE stage both at low frequencies and at high frequencies.

[5] The transimpedance factor and the voltage gain factor of the cascode stage exceed those of its CS or CE stage.

Replacement of a CE or CS stage in a controller by a cascode stage usually results in an increase of loop gain. This results in a larger bandwidth of the servo function, a better (gain) accuracy and a lower distortion of the feedback amplifier. Their application will be discussed in more detail in section 14.2.6.

Application of local feedback stages

If the midband accuracy or the differential gain of the servo function is out of specifications, while the bandwidth of the servo function is large enough, local feedback stages may be used to increase the loop gain without adding another dominant pole.

As a result of local feedback, the bandwidth of a local feedback stage may exceed that of the servo function. Hence, it do not introduce a dominant pole in the loop gain of the outer loop of the feedback amplifier.

Conclusions

The order of preference for intermediate stages in the controller is:

1. CE or CS stages or their balanced versions; use cascode stages if pole splitting brings one of the poles out of the dominant group, or if the low-frequency inaccuracy or distortion needs to be reduced.
2. Nonenergic feedback stages or their balanced versions.
3. Passive feedback stages or their balanced versions.

The device type, its geometry and its operating conditions should be optimized for the performance aspect of the servo function that needs to be improved through the application of this extra stage.

14.2.4 Interconnection of stages

The interconnection of the stages in the controller should be done in such a way that the controller behaves as a nullor if one of the stages behaves as such. This ensures that the performance of the feedback amplifier is predominantly fixed by its feedback network(s).

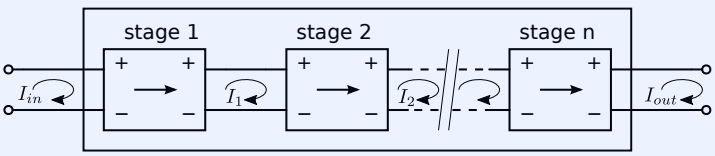


Figure 14.1: Preferred method for interconnecting stages in the controller: Both terminals of the input port of each following stage are connected to the output port of its preceding stage.

Figure 14.1 shows the way to achieve this: the input port of each following stage should be connected to the output port of its preceding stage. If one stage of the controller behaves as a natural two-port, the common-mode port current through the controller is zero and common-mode port voltages do not affect the differential-mode port quantities.

Generally, stages will not behave as natural two-ports and imperfect port isolation will cause errors in the source-to-load transfer and limit the CMRR and the PSRR of the amplifier. Such limitations can be kept small as possible by using balanced stages in the controller.

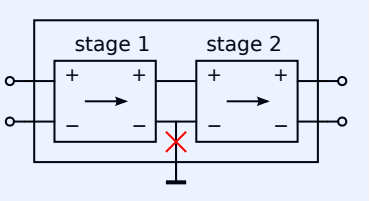


Figure 14.2: In cases where the input port and/or the output port of the controller is / are floating with respect to the ground, current paths from inside the controller to the ground should be avoided. Such paths may increase the gain inaccuracy as a result of an imperfect port isolation in the stages.

Figure 14.2 shows an example in which the controller has an internal connection to the ground, while both ports are floating with respect to the ground. In integrated circuit amplifiers, such paths exist as a result of the coupling between the devices and the substrate. In general, intentional creation of such paths should be avoided because such paths increase the effects of the imperfect port isolation of the individual stages of the controller on the CMRR and the PSSR of the amplifier.

Figure 14.3 shows five different options for a dual-stage four-terminal controller. All devices are drawn as generalized biased active three-terminal devices, as introduced in Chapter 3.5.1. We will discuss these configurations below.

A The circuit shown in Figure 14.3A represents a simple dual-device, dual-stage, four-terminal controller. The input port of the second stage is connected to the output port of the first stage. However, since both stages are three-terminal devices, the input voltage of the second stage equals the common-mode voltage across the two-port. Hence, the stage has a huge common-mode to differential mode conversion and does not at all approximate a natural two-port. Such a stage can only be applied in conjunction with an element that provides port isolation, such as a transformer.

B The circuit depicted in Figure 14.3B shows another simple dual-device, dual-stage, four-terminal controller. Because of its relative simplicity, it is a popular structure in simple dual-transistor feedback amplifiers. Figure 14.4 shows the signal path of a negative feedback voltage amplifier that uses this controller.

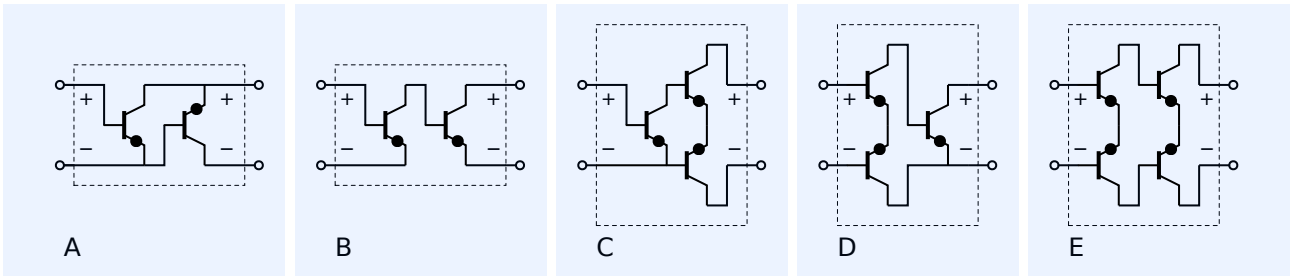


Figure 14.3: Signal path of all NPN two-stage controllers.

A: Simple two-transistor, two-stage controller. This controller has no port isolation because the input voltage of the second stage equals the common-mode voltage across the two-port.

B: Simple, two-transistor, two-stage controller in which the return path for the current between the two stages has to flow through the external network. This introduces deviations from the ideal gain that cannot be reduced by increasing the loop gain.

C: Two-stage controller with properly interconnected stages that uses a single-transistor input stage and an anti-series output stage.

D: Two-stage controller with properly interconnected stages that uses an anti-series input stage and a single transistor, or complementary-parallel output stage.

E: Two-stage controller with properly interconnected stages that consists of a cascade connection of two anti-series stages.

The circuit in Figure 14.3B also forms the basic structure for so-called *current feedback operational amplifiers*. A complementary-parallel connection of two controllers from Figure 14.6 is the basic architecture of such an operational amplifier.

The main disadvantage of this controller is that it does not obtain nullor properties if only one of the stages is replaced with a nullor. This implies that, if the loop gain approaches infinity, the gain of an amplifier equipped with such a controller does not equal its ideal gain. This makes the amplifier more sensitive to device parameter tolerances than amplifiers with controllers with properly connected stages, as shown in Figure 14.3C to E. Please study Example 14.1 and Example 14.2 for a more in-depth treatment of this topic.

C The circuit from Figure 14.3C, shows a dual-stage controller with properly interconnected stages. It uses a single-device input stage and an anti-series output stage. The latter one provides the port isolation.

D The circuit from Figure 14.3D also shows a two-stage controller with properly interconnected stages. It uses an anti-series input stage that provides the port isolation, cascaded with a single-device output stage. For high-efficiency operation, a complementary-parallel output stage can be used. The structure with a differential pair input stage and a complementary-parallel output stage is the basic structure used in *voltage-feedback operational amplifiers*.

E The controller circuit from Figure 14.3E consists of two cascaded anti-series stages.

In the following two examples we will demonstrate the effect of improper interconnection of stages in the controller as done in Figure 14.3B.

Example 14.1

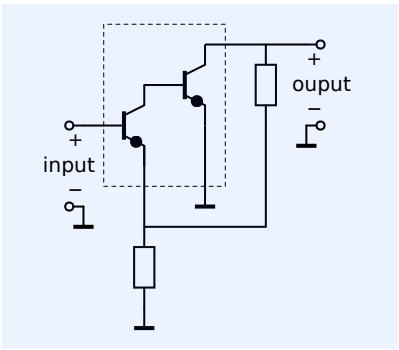


Figure 14.4: Voltage amplifier that uses passive feedback around the two-transistor, dual-stage controller from Figure 14.3B.

Let us consider the voltage amplifier from Figure 14.4, constructed with two MOS transistors as shown in Figure 14.5A. Figure 14.5B shows its small-signal equivalent circuit for evaluation of the asymptotic gain with the transconductance of the second stage selected as loop gain reference. In this circuit all dynamic elements have been omitted.

According to the asymptotic gain feedback model, the asymptotic gain $A_{v\infty}$ should now be calculated as

$$A_{v\infty} = \frac{V_\ell}{V_s}. \tag{14.1}$$

The network equations for the circuit can be found from network inspection:

$$V_s = V_i + V_1, \tag{14.2}$$

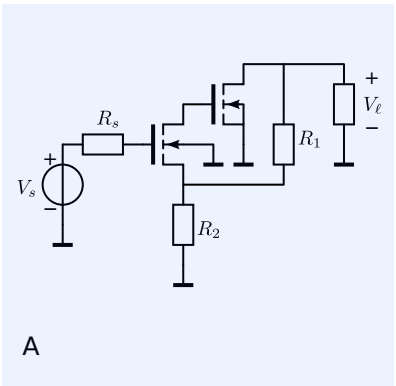
$$g_m V_i = V_1 \left(g_o + g_{mb}\right), \tag{14.3}$$

$$\frac{V_1}{R_2} + \frac{V_1 - A_{v\infty} V_s}{R_1} = 0, \tag{14.4}$$

where (14.2) follows from the loop of the voltages V_s, V_i and V_1. Equation (14.3) states that the sum of the currents that flow into node (2) equals zero, and (14.4) states that the sum of the currents in node (1) equals zero.

The asymptotic gain $A_{v\infty}$ is found as

$$A_{v\infty} = \frac{R_2 + R_1}{R_2} \frac{g_m}{g_m + g_o + g_{mb}}. \tag{14.5}$$



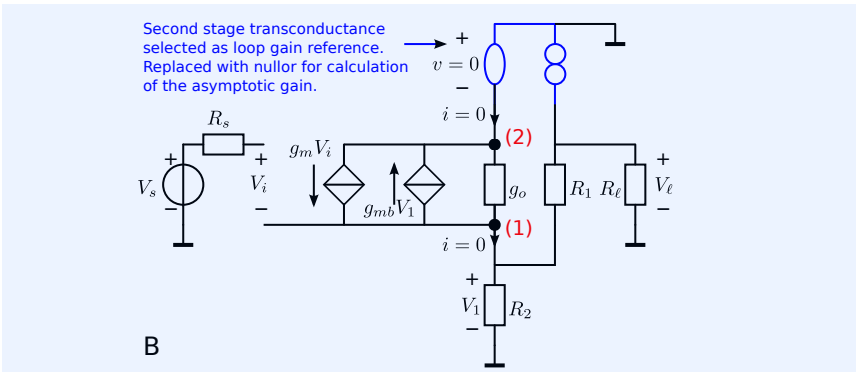


Figure 14.5: NMOS passive feedback voltage amplifier according to Figure 14.4.
A: Signal path diagram.
B: Small-signal equivalent circuit for the analysis of the static (DC) asymptotic gain in which the transconductance of the second stage has been selected as the loop gain reference.

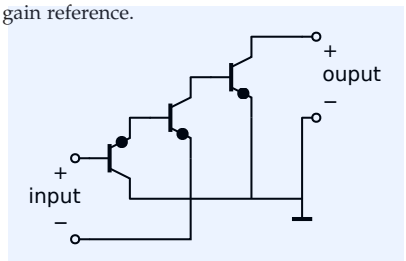


Figure 14.6: The so-called *current-feedback operational amplifier* is contructed from a complementary parallel connection of two of these circuits.

Hence, although one stage of the controller behaves as a nullor, the complete controller does not behave as a nullor. This is a consequence of the non-ideal interconnection of the two stages of the controller. The influence of the improper connection on the static inaccuracy of the source-to-load transfer can be kept small by using an input stage that has $\frac{g_m}{g_o+g_{mb}}$ as large as possible. This can be achieved by using a cascode stage for the first stage of the controller. The cascode stage will be described later.

Figure 14.6 shows the structure of the signal path of the so-called current feedback amplifier, which is derived from the structure from Figure 14.3B.

In the following example we will elucidate a similar effect in a two-stage, two-transistor current amplifier with bipolar transistors.

Example 14.2 *Let us now consider the current amplifier, constructed with two bipolar transistors as shown in Figure 14.7A. Figure 14.7B shows the small-signal equivalent circuit for evaluation of the asymptotic gain when the transconductance of the first stage has been selected as loop gain reference. In this circuit, all dynamic elements have been omitted. For the sake of simplicity, the output resistance of the second stage transistor has been omitted as well.*

According to the asymptotic gain feedback model, the asymptotic gain $A_{i\infty}$ is obtained as

$$A_{i\infty} = \frac{I_\ell}{I_s}. \tag{14.6}$$

The network equations for the circuit can be found from network inspection:

$$V_1 = -I_s R_1, \tag{14.7}$$

$$(1 + \beta_{AC})\, I_i + I_s = \frac{V_1}{R_2}, \tag{14.8}$$

$$\beta_{AC} I_i = -A_{i\infty} I_s, \tag{14.9}$$

where (14.7) follows from the fact that the input voltage and the input current of the nullor equal zero. As a result, the source current flows through R_1. Equation (14.8) states that the sum of the currents that flow into node (1) equals zero, and (14.9) states that the sum of the currents in node (2) equals zero.

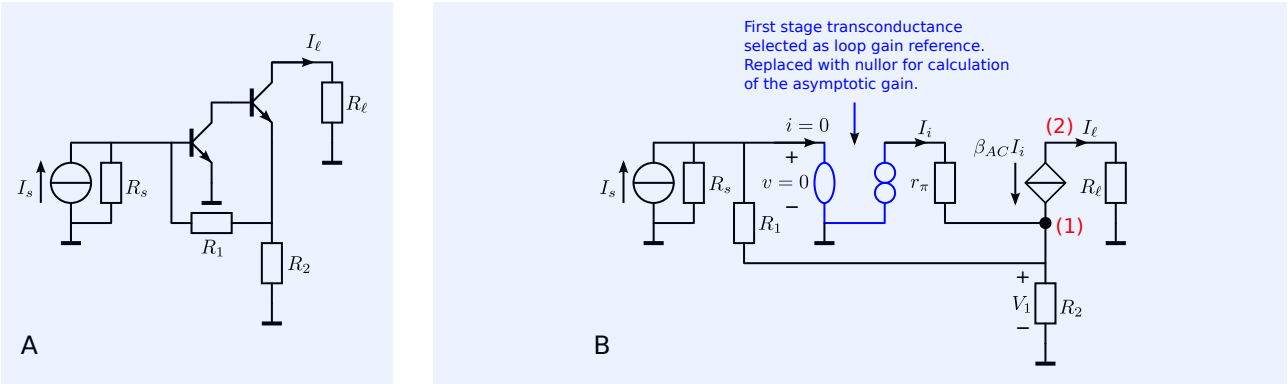


Figure 14.7: Bipolar passive feedback current amplifier with the controller according to Figure 14.3B.
A: Signal path diagram.
B: Small-signal equivalent circuit for the analysis of the static (DC) asymptotic gain in which the transconductance of the first stagehas been selected as the loop gain reference.

The asymptotic gain $A_{i\infty}$ is found as

$$A_{i\infty} = \frac{R_1 + R_2}{R_2} \frac{\beta_{AC}}{1 + \beta_{AC}}. \tag{14.10}$$

Hence, similar as in the previous example, the complete controller does not behave as a nullor if one stages behaves as such. This is a consequence of the non-ideal interconnection of the two stages of the controller. The influence of the improper connection on the static inaccuracy of the source-to-load transfer can be kept small by using an output stage with a high DC current gain, such as a field effect device.

Interconnection of unbalanced stages and balanced stages

Let us now consider cases in which a balanced stage is cascaded with an unbalanced one. Such a structure exists, for example, in the controller from Figure 14.3D. In such a situation, the unbalanced stage will convert common-mode output quantities of the differential stage into differential-mode quantities. In practice, common-mode currents in the balanced stage may be a result of the finite impedances of common-mode bias sources, and of parasitic coupling of the balanced stage with the power supply or the ground.

Figure 14.8 illustrates the conversion of a common-mode current flowing out of a balanced stage into a differential-mode current. For the sake of simplicity, the unbalanced stage has been connected to the ground. However, this is not necessary for common-mode to differential-mode conversion.

Figure 14.8 shows that the input current of the unbalanced stage equals the differential-mode output current of the balanced stage, plus half of its common-mode output current. This common-mode to differential-mode con-

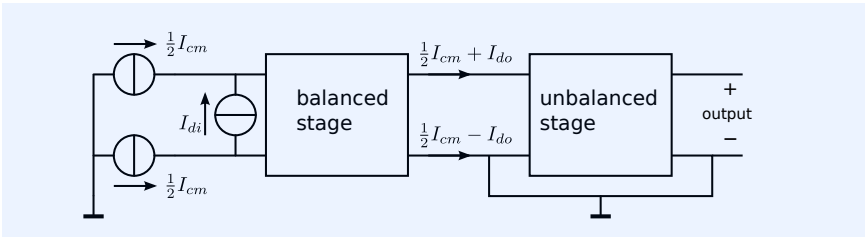


Figure 14.8: Common-mode to differential-mode conversion in a cascade connection of a balanced stage and an unbalanced stage.

version causes a reduction of the CMRR. If the common-mode current is a result of an undesired coupling between the signal path and the power supply, it also decreases the PSRR.

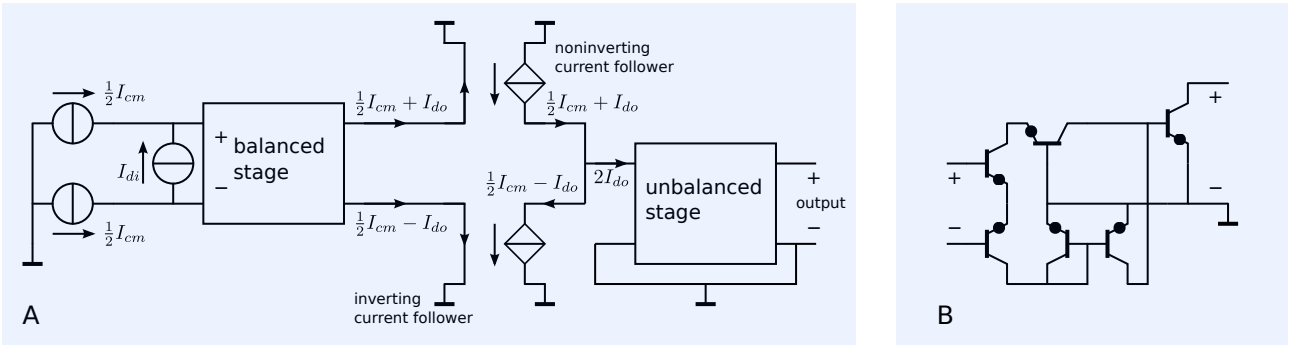


Figure 14.9: Reduction of common-mode to differential-mode conversion in a cascade connection of a balanced stage and an unbalanced stage with the aid of noninverting and inverting current followers.

The common-mode to differential-mode conversion as described above, can be prevented by subtracting the output currents of the differential stage before feeding them to the unbalanced stage. This has been illustrated in Figure 14.9.

Figure 14.9A shows the concept of this common-mode compensation technique. The current out of one output terminal of the balanced stage is directed to a noninverting unity-gain current amplifier, while the current out of the other output terminal is directed to an inverting unity-gain current amplifier. The output currents of both current amplifiers are added, and their sum is directed to the input of the unbalanced stage. In this way, the common-mode component has been removed from the input current of the unbalanced stage, which now equals twice the differential-mode output current of the preceding stage.

The use of both an inverting and a noninverting current amplifier, each connected to one output of the differential stage, generally improves the balanced operating conditions for the differential stage.

Figure 14.9B shows an implementation example in which a current mirror is used for the implementation of the inverting unity-gain current amplifier, while a CG stage or a CB stage implements the noninverting current amplifier.

14.2.5 Interconnection of controller and feedback networks

Let us now consider four-terminal controllers with an asymmetrical structure. Examples of such controllers have been shown in Figure 14.3B, Figure 14.3C and Figure 14.3D. If such a controller is connected to its external electrical environment, an alternative amplifier circuit can be obtained by swapping

the port terminals, both at the input port and at the output port. Figure 14.10 illustrates this for the voltage amplifier from Figure 14.4.

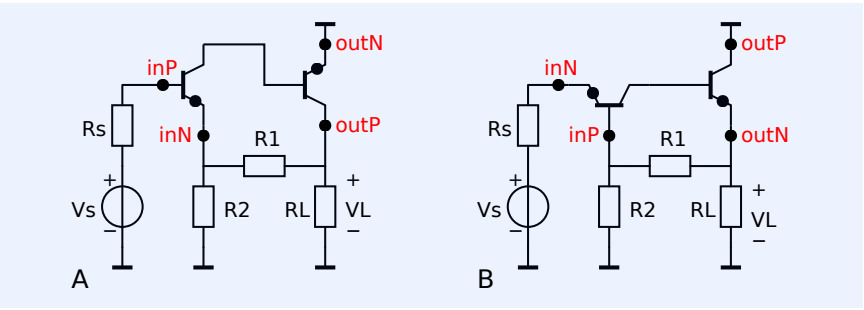


Figure 14.10: Swapping of the pins at the input port and at the output port of an asymmetrical, four-terminal controller, yields an alternative amplifier circuit.

A. Voltage amplifier from Figure 14.4 with pinnames at the input port and at the output port of the controller.

B. Passive feedback voltage amplifier, obtained from swapping the pins at the input port and at the output port of the controller.

Although both amplifiers are two-stage noninverting voltage amplifiers, their performance as well as their biasing will differ. Such differences may result in a strong preference for one of the two circuits. It it common practice to connect the controlling node[6] of the input device of the controller to the signal source, as it is done in Figure 14.10A. However, performance evaluation and comparison of both solutions may, in specific situations, yield another preference.

[6] Controlling node of an active device: gate (FET), base (BJT) or grid (vacuum tube).

14.2.6 Cascode stages

We have seen that the CS or CE stage or one of its balanced versions are the preferred controller stages. However, as a result of the parasitic feedback capacitance in those stages, cascaded stages strongly interact, which results in a rather complex dynamic behavior. Moreover, pole splitting in these stages may drive one or more poles of the loop gain out of the dominant group. If so, more than the minimum number of stages might be needed to achieve a sufficiently large bandwidth. Pole-splitting can be prevented by shorting the stages and driving the following stage with the short-circuit current. By doing so, each shorted stage has a simple $R/\!/C$ input impedance that introduces a pole at $s = -\frac{1}{RC}$. In this way, the controller needs the smallest number of stages for the required servo bandwidth. All poles are as close as possible to the origin!

Shorting of a CS or CE stage, while delivering the short-circuit current to the following stage can be accomplished by cascading a stage with a unity-gain current amplifier. Unity-gain current amplifiers can be implemented with local feedback current followers, such as the CG stage or the CB stage. The cascade connection of a CS or CE stage with a CG or CB stage is called a *cascode* stage.[7]

Figure 14.11 shows the cascode stage with generalized biased three-terminal active devices. Figure 14.12A shows the small-signal equivalent circuit of the stage from Figure 14.11. If BJTs are used we have $g_i = g_\pi = g_m/\beta_{AC}$, for field effect devices we have $g_i = 0$. Figure 14.12B shows a simplified model that, in most cases, models the behavior sufficiently accurate.

In case of MOS transistors, the output resistance r_o can be obtained as

$$r_o = r_{o_1}\left(1 + g_{m_2} r_{o_2}\right). \tag{14.11}$$

[7] Also: the cascade connection of a common cathode stage and a common grid stage in vacuum tube technology.

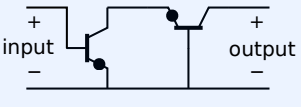


Figure 14.11: Cascode stage with generalized 3-terminal biased active devices.

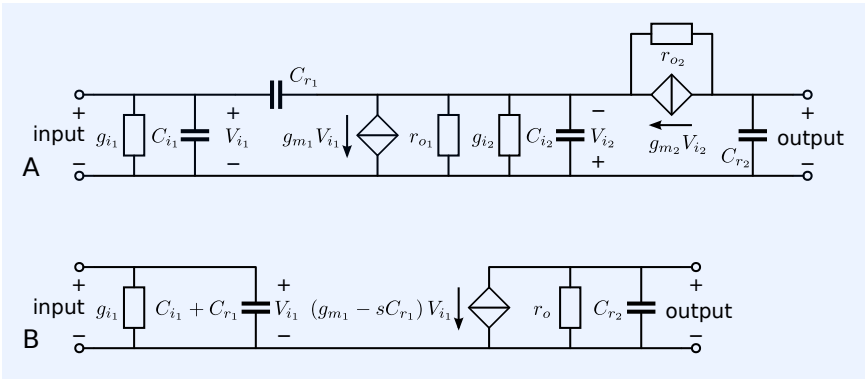


Figure 14.12: Small-signal models of the cascode stage.
A. Small-signal model contructed with the generalized small-signal models from Figure 4.52.
B. Simplified model of the cascode stage. The interaction between the input and output circuits has been eliminated by the current follower.

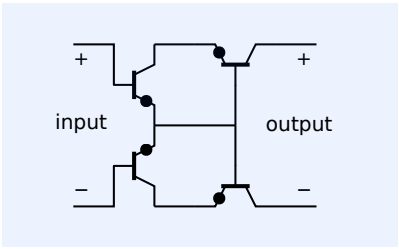


Figure 14.13: Anti-series connection of two cascode stages from Figure 14.11.

Figure 14.14: Small signal model of the anti-series cascode stage.

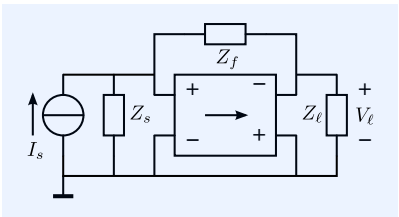
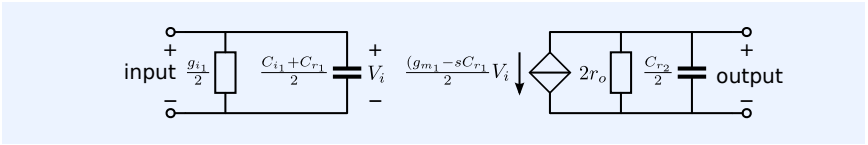


Figure 14.15: A passive-feedback transimpedance amplifier of which the source and the load are both connected to the ground, can be equipped with a three-terminal controller with an inverting transfer.

Figure 14.13 shows the anti-series connection of two cascode stages. The small-signal equivalent circuit is shown in Figure 14.14. When compared with the anti-series stage, the anti-series cascode stage exhibits:

- A unilateral transfer
- An improved port isolation at high frequencies
- No pole-splitting (Miller effect).



14.2.7 Application of local feedback stages

Until now, we have studied the application of the CS stage and the CE stage and their balanced versions as amplifier stages in the controller. We have also applied local feedback current followers as high-frequency buffers between cascaded CE or CS stages. From this we have seen that a CE or CS stage, cascaded with a CB or CG stage can be considered as a single amplifier stage (the cascode stage) with improved unilateral behavior.

In this section, we will study the application of other local feedback stages in the controller.

Dual-device three-terminal controllers

Let us consider a passive-feedback transimpedance amplifier in which both the source and the load share one terminal with the ground. Such an amplifier can be equipped with a three-terminal controller that has an inverting transfer (see Figure 14.15). Since two cascaded CE or CS stages provide an noninverting transfer, it is not possible to construct a dual-device controller with only CE or CS stages. So, if we need a two-stage controller we need at least one anti-series stage, hence, three devices.

Let us investigate other options for dual-device controllers. A three-terminal dual device controller with an inverting transfer can be built with one inverting amplifier stage and one noninverting amplifier stage. The CE or CS stage is an inverting amplifier stage and the CB or CG stage and the CC or CD

stage can be applied as noninverting amplifier stages. Hence, we can construct four different dual-device controllers. Figure 14.16 shows the available options, they will be discussed below.

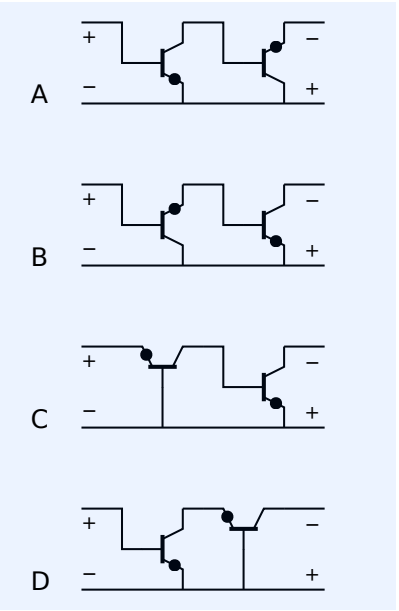


Figure 14.16: Options for three-terminal dual-device controllers.
A. A CE or CS stage cascaded with a CC or CD stage.
B. A CC or CD stage cascaded with a CE or CS stage.
C. A CB or CG stage cascaded with a CE or CS stage.
D. A CE or CS stage cascaded with a CG or CB stage.

A Figure 14.16A shows a CE or CS type input stage and a CC or CD type output stage. If the output impedance of the first stage is larger than the input impedance of the CC or CD stage, the latter one can be considered to be current-driven. The loop gain in the output stage is then very low and it tends to behave as a noninverting CE or CS stage.

Since the voltage gain factor of the output stage is unity, the voltage swing at the output of the first stage equals the output voltage. This makes it contribute to the distortion.

B Figure 14.16B shows a CC or CD type input stage and a CE or CS type output stage. In this solution, the first stage of the controller has a voltage gain factor of unity. This implies that the contribution of the voltage noise source of the second stage to the total equivalent input noise voltage cannot be ignored.

C Figure 14.16C shows a CB or CG stage type input stage and a CE or CS type output stage. In this configuration the contribution of the equivalent input current noise source of the second stage to the equivalent input current noise of the controller cannot be ignored.

D Figure 14.16D shows a CE or CS type input stage and a CB or CG type output stage. As discussed before, such a stage is often considered as a single stage. This is because the CB or CG stage does not increase the product of the dominant poles and the loop gain: their current gain is unity with a pole at ω_T, which is almost always outside the dominant group.

CC-CB and CD-CG cascade connection

Let us now investigate the application of the circuit from Figure 14.3C as two-stage controller in the transimpedance amplifier from Figure 14.15. Figure 14.17 shows the complete signal path of a transimpedance amplifier with this controller. This figure clearly shows that the anti-series input stage is not perfectly balanced. One device of the differential stage has its input grounded, while the other one has its output connected to the ground. Such an arrangement can better be considered as the cascade connection of a CC or CD stage and a CB or CG stage, or alternatively, as a cascade connection of a voltage follower and a current follower.

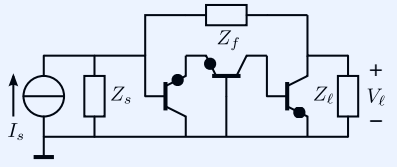


Figure 14.17: Transimpedance amplifier from Figure 14.15 with the controller from Figure 14.3C.

This combination of local feedback stages behaves as a non-inverting CE or CS stage. The output resistance r_o strongly depends on the driving resistance of the stage. If the stage is driven from a current source, it equals the output resistance of the cascode stage. If the stage is driven from a voltage source, it equals the output resistance of the basic CE or CS stage. The small-signal equivalent circuit of the stage is shown in Figure 14.18.

The noise behavior of this stage differs from that of a truly balanced anti-series stage. Since the first stage in the controller is a voltage follower, its

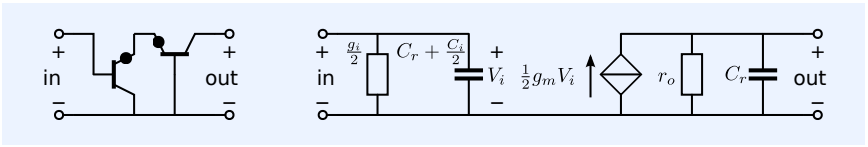


Figure 14.18: Small-signal model of cascade connection of the voltage follower and the current follower. This stage can be regarded as a basic noninverting amplifier stage. The element values refer to those of the single device. The output resistance r_o strongly depends on the driving resistance of the stage. If the stage is driven from a current source, it equals the output resistance of the cascode stage. If the stage is driven from a voltage source, it equals the output resistance of the basic CE or CS stage.

transmission parameter A equals unity. This implies that the noise voltage source of the current follower contributes as much to the total equivalent input voltage as the voltage noise of the voltage follower. Other contributions of the noise of the current follower to the equivalent input noise sources can be ignored because the parameters B, C, and D of the first stage are not enlarged as a result of negative feedback. Figure 14.19 shows the noise transformation for the cascade connection of the voltage follower and the current follower.

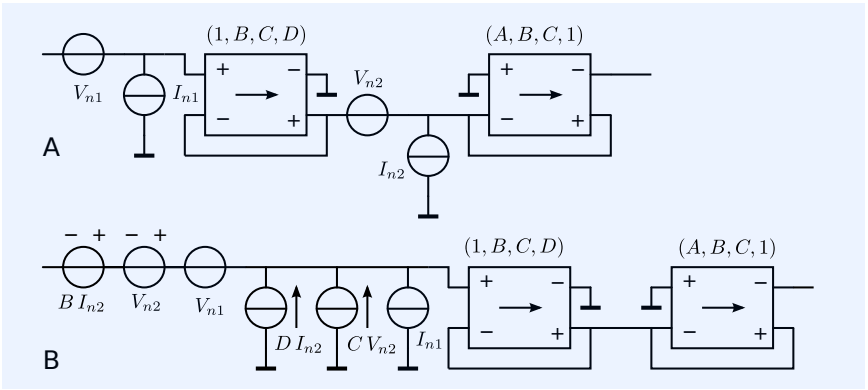


Figure 14.19: Noise transformations in a cascade connection of a voltage follower and a current follower.

A. Since the voltage follower and the current follower are nonenergic feedback stages, their equivalent input noise sources equal those of their controller (a CE or CS stage). B. Negative feedback in the voltage follower establishes $A = 1$, while the other transmission parameters equal those of the controller (CE or CS stage).

The contributions of BI_{n_2}, DI_{n_2} and CV_{n_2} to the equivalent input noise sources can usually be ignored. If we assume equal spectral densities of the corresponding equivalent input noise sources of both stages, the spectral density $S_{v,tot}$ of the total equivalent input voltage noise can be approximated by

$$S_{v,tot} = 2S_v, \tag{14.12}$$

where $S_v = S_{V_1} = S_{V_2}$ is the spectrum of the equivalent input voltage noise of both stages.

The spectral density $S_{i,tot}$ of the total equivalent input current noise can be approximated as

$$S_{i,tot} = S_i, \tag{14.13}$$

where $S_i = S_{I_1} = S_{I_2}$ is the spectrum of the equivalent input current noise of both stages.

If we compare this result with the one shown in Figure 6.29, we see that the spectrum of the total equivalent input voltage noise source equals that of a truly balanced stage, but the spectrum of the total equivalent input current noise source is twice that of a truly balanced stage.[8] This implies that the unbalanced application of the anti-series stage has a slightly larger noise contribution.

[8] Assuming equal devices and operating points for the truly balanced stage and the cascade connection of the voltage follower and the current follower.

15 Amplifier Biasing

15.1 Introduction

Until now, we have discussed the design of the amplifier's signal-path, thereby assuming that the desired operating conditions of the active devices were fixed with ideal current and voltage sources, as described in Chapter 3. In this chapter, we will discuss methods to minimize the number of bias sources, as well as the way in which the remaining bias sources can be derived from the power supplies. We will not yet discuss the design of the bias sources themselves.

15.1.1 A structured approach to biasing

A definition of biasing has been given in Chapter 9.

> Biasing is the process of fixing the electrical operating conditions of electronic devices, and deriving the required bias voltage and current sources from the power supply voltage(s).

The biasing design problem can thus be formulated as follows:
Derive the input and output bias quantities of the amplifier stages from the power supply voltages in such a way that:

1. *The amplifier stage is accurately biased in the desired operating point over the whole temperature and power supply range.*
2. *The amplifier stage is sufficiently isolated from the power supply.*
3. *The degradation of the amplifiers's performance caused by the biasing, is acceptable.*

The biasing of amplifier stages and amplifiers is known to be difficult. First of all, this is because the design and the implementation of a feasible biasing concept, show a strong interaction with the design of all kinds of performance aspects of the amplifier. In particular, the number of feasible circuit topologies shows a strong interaction with the supply voltage. Secondly, biasing is a complex process because the bias elements themselves are passive elements each of which needs to be biased in a proper operating point as well.[1]

For these reasons, biasing circuits are often designed concurrently with the signal path. Possible design conflicts that result from interactions between the bias circuits and the signal path are usually resolved in design loops. Such a design approach may be suited to experienced designers, but it may be very confusing for novice designers and it does not lend itself to automation.

In this book, we will introduce a structured approach to biasing that clearly shows the interaction between the design of the signal path, the design of the biasing scheme and the implementation of bias sources. We will use a step-by-step approach in which we will explicitly motivate and describe the subsequent design steps. By doing so, this approach is believed to be suited to non-experienced designers and may also be used as a basis for algorithm-based design automation.

We will introduce techniques that can be used to minimize interactions between the design of the biasing circuits and the signal path, as well as techniques for optimization of the design of the signal path with respect to biasing.

15.1.2 Basic passive biasing elements

Since the power supply is the only active element in a circuit, biasing elements have to be realized with passive network elements that behave as

[1] A network element is called passive if it dissipates power. This is the case if the real part of the sum of the products of all its branch currents and branch voltages is positive. A network element is called active if it delivers power to its external circuit. This is the case if the real part of the sum of products of all its branch currents and branch voltages is negative.

voltage or current sources. Within the context of biasing, we will use the following definitions for elements that behave as voltage or current sources:

1. A two-terminal network element has a voltage source character if, at frequencies of interest, the magnitude of its small-signal impedance is smaller than the ratio of the DC voltage across the element and the DC current through it.

2. A two-terminal network element has a current source character if, at frequencies of interest, the magnitude of its small-signal impedance is larger than the ratio of the DC voltage across the element and the DC current through it.

With the above definitions, it will be clear that at nonzero frequencies, inductors have a current source character and capacitors a voltage source character. Hence, for biasing purposes capacitors can be inserted in series with the signal path and inductors can be placed in parallel with the signal path. This is called *AC coupling*.

Nonlinear resistors

Biasing elements that exhibit a voltage or current source character at all frequencies, can be realized with nonlinear resistive elements, or shortly: nonlinear resistors. However, since the $V-I$ characteristic of a *passive* nonlinear resistive element passes through the origin, such an element can only exhibit a voltage or a current source character over a limited operating range. In the vicinity of $(V,I)=(0,0)$ any passive nonlinear resistor behaves as a linear resistor:[2]

$$\frac{V}{I}=\frac{dV}{dI}, \tag{15.1}$$

where V and I are the voltage across the device and the current through the device, respectively.

[2] Assume no discontinuity in the $V-I$ characteristic and its derivative at $(V,I)=(0,0)$.

Figure 15.1A shows a $V-I$ characteristic of a nonlinear resistor with a current source character over a wide operating range. The lower limit of the voltage across a nonlinear resistor, at which it exhibits a current source character, is called the saturation voltage. The upper limit of this voltage is called the breakdown voltage.

Figure 15.1B shows a $V-I$ characteristic of a nonlinear resistor with a voltage source character over a wide operating range. The lower limit of the current through a nonlinear resistor, at which it exhibits a voltage source character, is called the cut-off current. The upper limit of this current is called the saturation current.

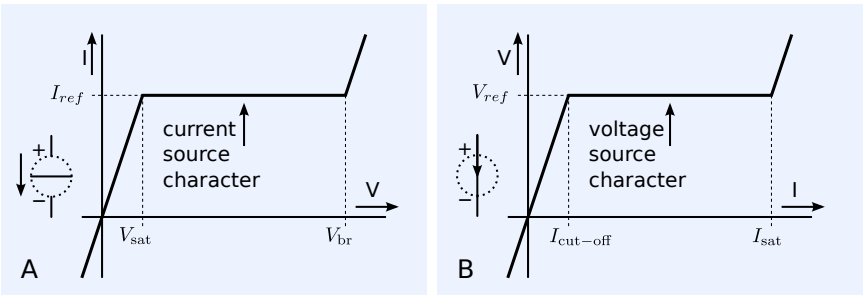


Figure 15.1: Nonlinear resistors that exhibit a current source or a voltage source character over a limited operating range:
A: Symbol and characteristic of a nonlinear resistor with a current source character when $V_{\rm sat}<V<V_{\rm br}$.
B: Symbol and characteristic of a nonlinear resistor with a voltaget source character when $I_{\text{cut-off}}<I<I_{\rm sat}$.

One of the difficult aspects of biasing is that the biasing elements themselves need to be biased in an operating point at which they deliver the proper value of the bias quantity, while they maintain their current or voltage source character for all signal values and power supply variations.

15.1.3 Outline of the biasing approach

The outline of the structured approach to the biasing of amplifier stages and amplifiers is sketched below:

1. The signal path design with four bias sources per transistor, as discussed in Chapter 3, is the starting point for the design of the biasing. Theoretically, all of these bias sources are capable of delivering power.

   In practice, only the power supply sources deliver power to the circuit, all other elements are passive elements that either dissipate power or store electrical energy. The latter ones are capable of delivering power over a limited time interval, but the time average of the power, delivered by those elements, equals zero.

2. The power supply voltage sources will be added to the signal path diagram. One of the terminals of each power supply will be connected to the reference node (the ground) of the signal path.

3. The bias currents need to be delivered by the power supply voltage(s). This is achieved by redirecting the bias current sources via the power supplies and the ground. In order to minimize the number of biasing elements, parallel connections of bias current sources will be replaced with a single current source and series connections of voltage sources will be replaced with a single voltage source.

4. Since the power supply source is the only active device in the circuit, all remaining bias sources need to be replaced with passive nonlinear resistors that exhibit a voltage or current source character. This requires that the branch current of each biasing element flows from the terminal with the most positive voltage to the one with the most negative voltage.[3] In this design step, bias sources will be added to ensure such behavior.

5. At a later stage, we will see that floating voltage sources are difficult to realize. For this reason, their number needs to be minimized. We will show that this can be done either by changing the operating voltages of the active devices, by replacing them with grounded voltage sources, or by generating alternative amplifier topologies that have similar signal performance, but different biasing requirements.

6. Before judging the feasibility of biasing elements, we need to define their requirements.

7. If one or more feasible bias solutions can be found, we will select the most promising solution and continue with step 8, else we will evaluate the application of error reduction techniques and go back to step 3.

   Error reduction techniques that can be applied are:

   - Compensation (temperature compensation, model-based biasing)
   - Error feed forward
   - Negative feedback (application of local, over-all, differential-mode and common-mode feedback techniques)
   - Sampling/switching (*auto-zero* biasing)
   - Modulation/demodulation (*chopper stabilized* biasing)

8. Start the detailed design of the biasing elements.

[3] The power dissipation in the biasing element is positive.

15.1.4 Drawing conventions

In order to illustrate the process of biasing in a comprehensible way, we will use some circuit drawing conventions. Figure 15.2 shows the drawing conventions for some devices, as they will be used in this chapter:

- The positive direction of the current flow in a schematic is from top to bottom. Hence, p-type devices are drawn with their emitter or source up, while n-type devices are drawn with their collector or drain down. Devices that operate at a higher voltage level are drawn at a higher position in the circuit diagram than devices that operate at a lower voltage level.

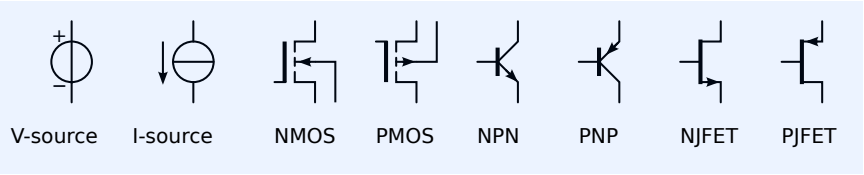


Figure 15.2: Drawing conventions.

- Information flow in a schematic is from left to right. In feedback networks it may be from right to left.
- For the sake of simplicity, the substrate connection of bipolar devices is not drawn. When properly biased, the substrate voltage has only a minor influence on the performance aspects of the device.
- In IC technology the common bulk of equal NMOS devices is connected to the supply with their largest negative voltage in the circuit. The common bulk of equal PMOS devices is connected to the one with their largest positive voltage.

15.1.5 This chapter

The motivation for separating the design of the signal path and the biasing has already been given in section 4.5.1. In this chapter, we will introduce a step-by-step approach to the biasing of amplifier stages and amplifiers.

In section 15.2, we will discuss the way in which the bias quantities of amplifier stages can be derived from one or more power supply voltage sources, thereby using only passive nonlinear resistors (step 1 through 4 of the list in section 15.1.3). We will demonstrate this technique for CE or CS stages, local feedback stages and balanced stages.

Other sections, such as the minimization of floating voltage sources, setting up the requirements for biasing elements and improvement of the biasing using error-reduction techniques still need to be added.

15.2 Setting up the initial biasing scheme

In this section, we will discuss the design of an initial biasing scheme of an amplifier. The associated design steps are:

1. Set up the signal path design with four bias sources per transistor, as discussed in Chapter 3.
2. Add the power supply voltage sources to the signal path diagram.
3. Redirect the current sources via the supplies and the ground and minimize the number of bias sources.
4. If necessary, add bias currents through voltage sources in such a way that those voltage sources can be replaced with passive nonlinear resistors with a voltage source character.
5. If necessary, add (or change) supply voltages in such a way that the bias current sources can be replaced with passive nonlinear resistors with a current source character.

We will demonstrate this for a single-transistor CE and CS stage amplifier in section 15.2.1, for local feedback stages in section 15.2.2, for cascode stages in section 15.2.3 for anti-series stages in section 15.2.4, for complementary parallel stages in section 15.2.5.

15.2.1 CE and CS stage biasing

In this section, we will demonstrate the biasing of a transimpedance amplifier with a single CE or CS stage.

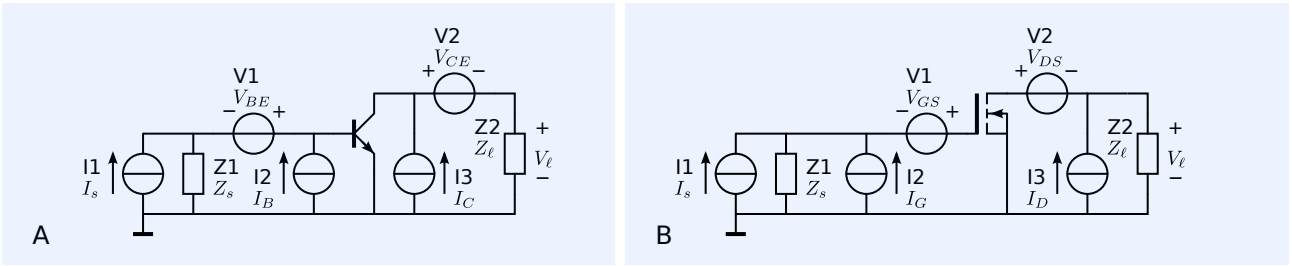


Figure 15.3: Starting points for the biasing of a transimpedance amplifier with a CE stage or CS stage. Please note that I_s and V_ℓ represent the source signal and the load signal, respectively. Their values are zero in the quiescent operating point.

A. Biased CE stage with zero quiescent current through V1 and V1 and nonzero quiescent voltages across the current sources.

B. Biased CS stage with zero quiescent voltage across I2 and I3 and nonzero quiescent currents through V1 and V2.

Figure 15.3A shows the signal diagram of the transimpedance amplifier equipped with a single CE stage and its four bias sources. In this arrangement, the bias voltage sources carry no current, while the voltages across these sources equal the operating point voltages. Figure 15.3B shows the signal diagram of the transimpedance amplifier equipped with a single CS stage with the four bias sources. In this arrangement, the bias voltage sources carry the quiescent operating currents, while the voltage across the bias current sources equals zero. Both arrangements of bias sources can be converted into each other through application of the Blakesley[Blakesley1994][4] voltage shift theorem.

Figure 15.4 shows the initial biasing schemes of both circuits from Figure 15.3. Please note that the voltage sources `V1` and `V2` in both circuits cannot be replaced with passive nonlinear resistors. In Figure 15.3A the current through these sources is zero, while the voltage across them is nonzero. Hence, their $V - I$ characteristic cannot pass through the origin. In Figure 15.3B the product of the branch voltage and the branch current is negative. Hence, by definition, these are active elements.

[4] T. A. Blakesley. A New Electrical Theorem. *Proc. Phys. Soc. London*, 13:65–67, 1994

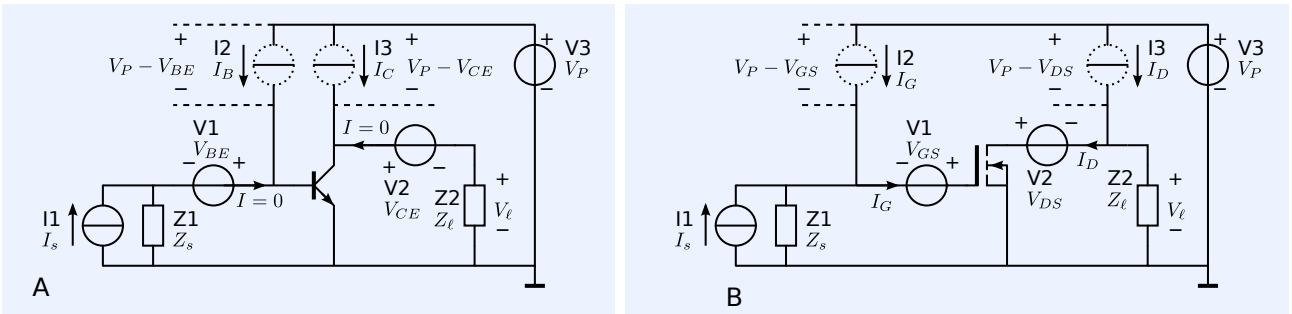


Figure 15.4: Initial biasing schemes of the transimpedance amplifiers from Figure 15.3:

A. Initial biasing scheme of the CE stage transimpedance amplifier.

B. Initial biasing scheme of the CS stage transimpedance amplifier.

In order to replace these voltage sources with passive elements, we need to add current through them in such a way that the power dissipation in the element is positive. If we do this for `V2`, we need a passive element with a current source character (`I5`), as shown in Figure 15.5. Since this element itself needs to be passive, it requires a nonzero voltage across it. This voltage needs to be provided by the negative power supply source `V4`.

The biasing of the transimpedance amplifier with the CS stage from Fig-

ure 15.4B proceeds in a similar way. In order to replace the bias voltage sources with passive devices, the current direction in these sources needs to be changed. This can be accomplished by changing the initial biasing arrangement from Figure 15.3B, to the one used in Figure 15.3A.

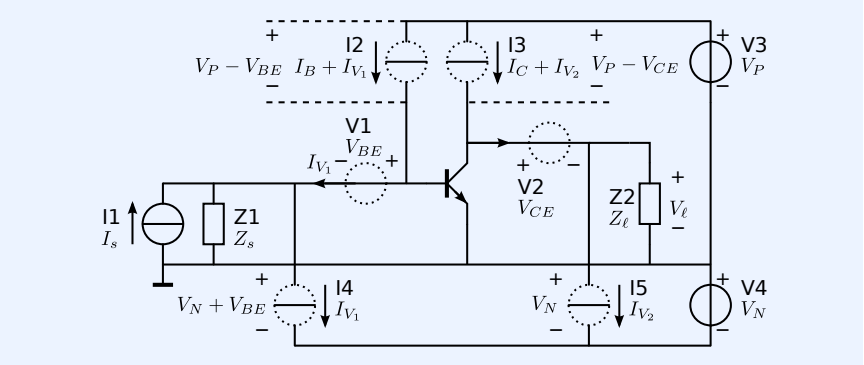


Figure 15.5: Initial biasing schemes of the transimpedance amplifiers from Figure 15.3A. All bias sources have now been derived from the power supply voltages with the aid of passive nonlinear resistive elements.

The result for the CMOS version of the transimpedance amplifier is then similar to that of the bipolar version; it is shown in Figure 15.6. Please note that the bulk of the NMOS has now been connected to the most negative voltage in the circuit. Whether this is necessary or not, depends on the applied integrated circuit technology.

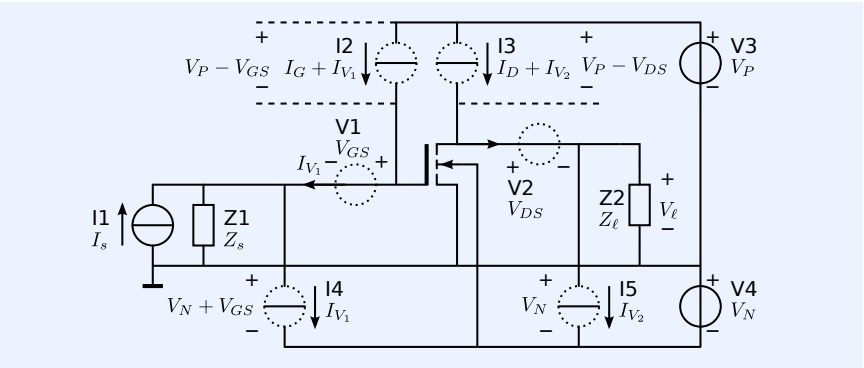


Figure 15.6: Initial biasing schemes of the transimpedance amplifiers from Figure 15.3B. All bias sources have now been derived from the power supply voltages with the aid of passive nonlinear resistive elements.

At a later stage of the biasing, we will discuss means to reduce the number of floating voltage sources. Now, we will continue with designing initial biasing schemes for other kinds of stages.

15.2.2 Biasing of local-feedback stages

The biasing of local-feedback stages is performed similarly as the biasing of CE or CS stages. The starting point is the biased N or P device with its four bias sources. The feedback networks are added around the biased devices. In this section, we will discuss the design of an initial biasing scheme for local feedback stages with n-type devices. The biasing of stages with p-type devices proceeds in a similar way, but the biasing elements have opposite signs. We will use examples with BJT or MOS devices. The biasing procedure is independent of the device type and technology. The implementation of the biasing scheme, however, strongly depends on the technology and the device type.

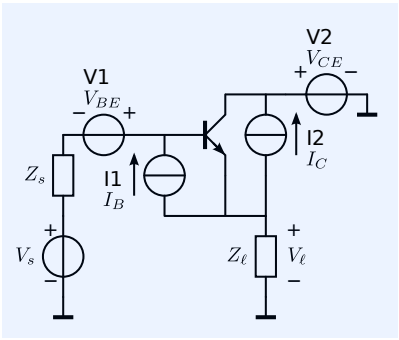


Figure 15.7: CC stage with source, load and bias sources.

Setting up the initial biasing scheme of CD and CC stages

Figure 15.7 shows an NPN CC stage with its source, load and bias sources. A similar circuit can be designed with a MOS transistor.

Similar as with the biasing of the CE and CS stage, the bias quantities have to be derived from the power supply voltages with the aid of the nonlinear resistive elements. This process is illustrated in Figure 15.8.

Figure 15.8A shows the result after adding the power supply voltage V_P, redirecting the current sources via the power supply and ground, and replacing them with nonlinear resistors. In order to have a nonzero voltage drop across I2, a negative supply voltage V_N has to be added. After doing so, I2 can be replaced with a nonlinear resistor with a current source character. Figure 15.8B shows the addition of a current I_{V_1} through V1, so that this voltage source can be replaced with a nonlinear resistor.

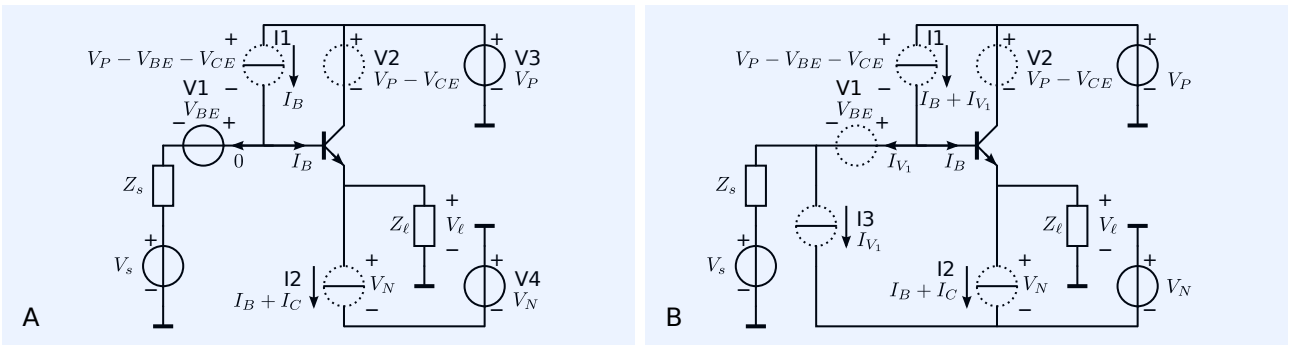


Figure 15.8: Biasing of the CC stage. Please note that V_s and V_ℓ represent the signal voltages. These voltages are zero in the quiescent operating point.

A. Addition of the power supply voltage to the signal path from Figure 15.7, redirecting of the current sources via the power supplies and the ground and replacing them with passive nonlinear resistors.

B. Replacing V1 with a passive nonlinear resistor, requires the addition of a bias current I_{V_1} through it.

The biasing of an NMOS and a PMOS CD stage proceeds in a similar way; it is left as an exercise to the reader.

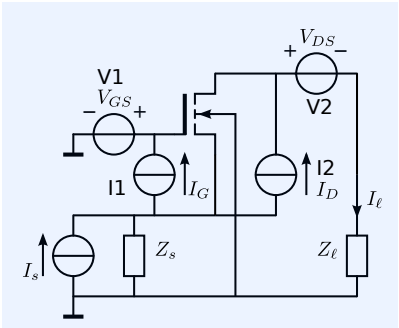


Figure 15.9: CG stage with source, load and bias sources.

Setting up the initial biasing scheme of CG and CB stages

Figure 15.9 shows an NMOS CG stage with its source, load and bias sources. A similar circuit can be designed with bipolar transistor. The biasing of an NPN and PNP CB stage is left as an exercise to the reader.

Similar as with the biasing of the CE and CS stage, the bias quantities have to be derived from the power supply voltages with the aid of the nonlinear resistive elements. This process is illustrated in Figure 15.10.

Figure 15.10A shows the result after adding the power supply voltages, redirecting the current sources via the power supply voltage sources and replacing them with nonlinear resistors. In order to have a nonzero voltage drop across I2, a negative supply voltage V_N has been added. Figure 15.8B shows the addition of a current I_{V_1} through V1, so that this voltage source can be replaced with a nonlinear resistor. The biasing of an NMOS CD stage proceeds similarly. This is left as an exercise to the reader.

Initial biasing of other three-terminal configurations

Until now, we have discussed the design of initial biasing schemes for some basic amplifier stages of which the source and the load share one terminal. By taking the initial bias scheme from Chapter 3 as starting point, both the source and the load operate at $(V, I) = (0, 0)$. Hence, the DC voltage across the source and the load and the DC current through the source and the load all equal zero. These conditions are not always required nor desired. In fact, some signal sources like PIN diodes in optical receivers require a nonzero DC bias voltage for high-speed operation, while a low-noise amplifiers for

radio applications may need to drive a mixer that requires a nonzero DC bias current. A signal source or a load that needs, or is allowed to operate under nonzero bias conditions can be modeled with the aid of a bias-free source or load and bias sources, similar to the bias sources to two-terminal elements in Chapter 3. A biasing example with added output offset can be found in Chapter 9.

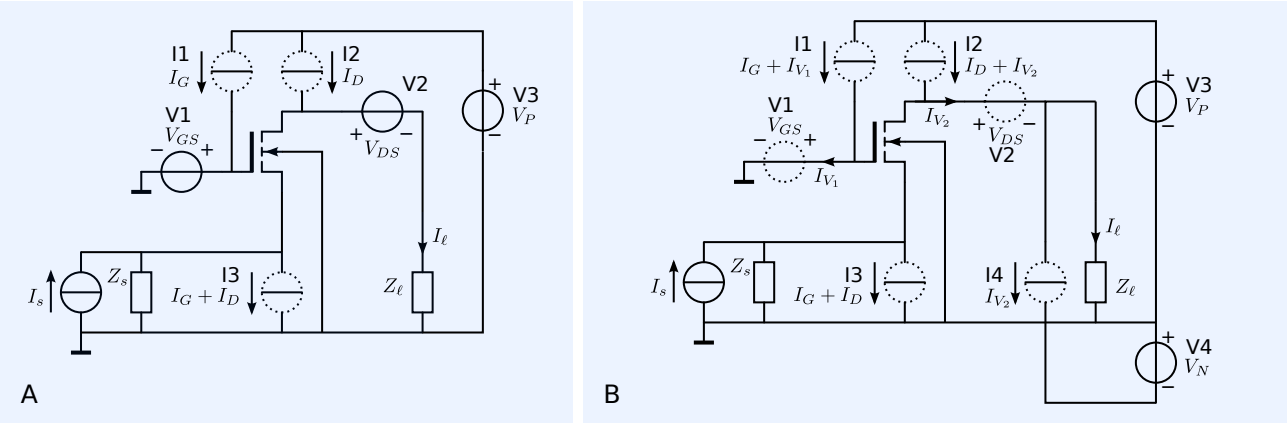


Figure 15.10: Biasing of the CG stage. Please note that I_s and I_ℓ represent the signal currents. These currents are zero in the quiescent operating point.

A. Addition of the power supply voltage to the signal path from Figure 15.9, redirecting the current sources via the power supplies and the ground and replacingthem with passive nonlinear resistors.

B. Replacing the bias voltage sources from Figure 15.10A with passive nonlinear resistors requires the addition of bias currents through them.

Although the return path for the source signal and the load signal, in three terminal amplifiers, are the same, they do not always need to operate at the same DC voltage with respect to the ground. If the source and the load share the signal return paths, but allow a DC offset between their common terminal, a DC bias voltage can be added to the biasing scheme. At a later stage of the design, the number of voltage sources can minimized.

Example 15.1

In this example we will discuss the design of a series stage with an NPN transistor, of which the load:

1. *is a nonlinear resistor,*
2. *is electrically isolated from the source and the ground,*
3. *is allowed to operate at a nonzero common-mode voltage,*
4. *must be biased with a current I_Q,*
5. *carries a voltage V_Q at this bias current.*

Figure 15.11 shows a model of this load.

The series stage will be driven from a voltage source with a source impedance Z_s, of which one terminal is connected to the ground. The source should operate at zero bias current and zero bias voltage.

The series stage should provide a voltage-to-current transfer for signals of which the frequency components of interest include zero (DC).

Figure 15.12 shows the circuit diagram of the biased series stage connected to its source and load.

Figure 15.13 shows the biasing scheme after the power supply sources have been added, the bias current sources have been redirected via the power supplies and replaced with passive nonlinear resistors with a current source character. The negative power supply voltage source `V3` is required for biasing `I3`. The bias voltage across `I4` amounts V_{CE}. It can be increased by connecting `I4` to the negative supply as well.

The bias current through V1 equals zero. In order to replace this voltage source with a passive nonlinear resistor, a bias current needs to be added. The biasing

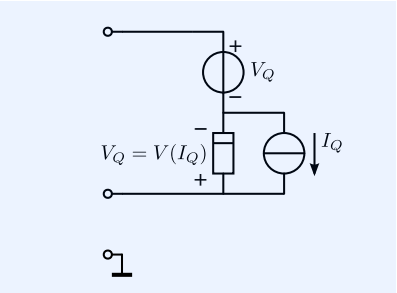


Figure 15.11: Model of a nonlinear resistor biased at $(V, I) = (V_Q, I_Q)$, and isolated from the ground.

scheme can be simplified by taking the operating current of the BJT equal to that of the load: $I_C = I_Q$. Changing the operating current or voltage of an amplifier stage, a biasing element, the source, or the load is one of the techniques to simplify the biasing scheme of an amplifier. We will give more examples of the application of this technique at a later stage. The result of the above actions is shown in Figure 15.14.

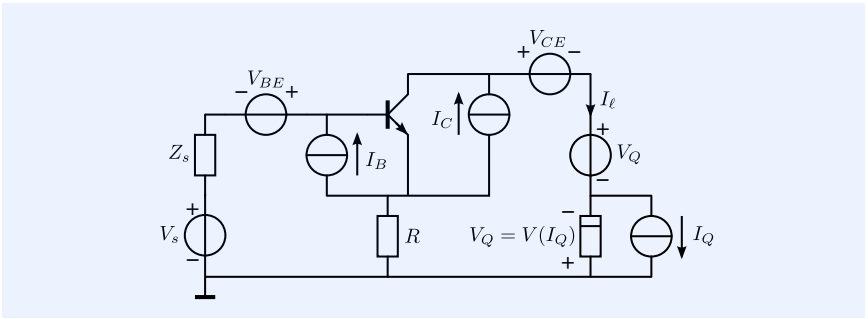


Figure 15.12: Biased series stage, driven from its source and loaded with the biased load from Figure 15.11.

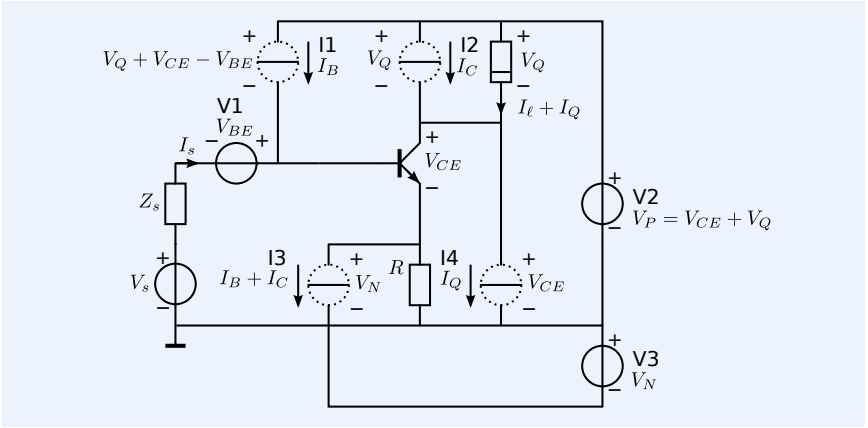


Figure 15.13: Simplified biasing scheme for the circuit from Figure 15.12.

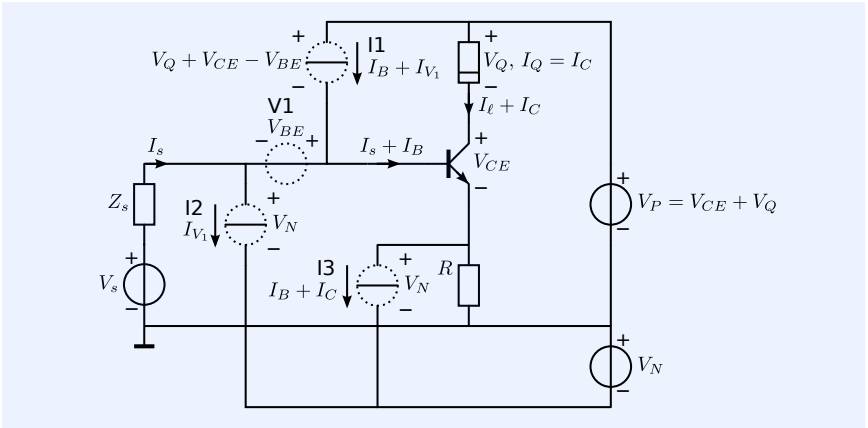


Figure 15.14: Biasing scheme for the circuit from Figure 15.13 for the case in which $I_Q = I_C$.

Although we confine ourselves in this chapter to the design of the initial biasing scheme of amplifiers, it may be helpful to show a possible final implementation of the biasing of the series stage from the preceding example. By doing so, it will become clear that separation of the design of the signal path and the biasing helps us with clearly defining requirements and design limits for biasing circuitry before designing them. This makes it possible to

estimate the feasibility of the biasing of an amplifier (stage) at an early stage of the design.

In the following we will show relatively simple implementation of the biasing in which the biasing elements with a current source character will be replaced with resistors and the biasing element with a voltage source character will be replaced with a diode.

Example 15.2

Figure 15.15 shows a simple implementation of the biasing scheme from Figure 15.14. In this implementation the nonlinear resistors with a current source character have been replaced with linear resistors and the nonlinear resistor with a voltage source character (V1) has been replaced with a diode. By doing so, the signal performance of the series stage from Figure 15.14 differs from that of the initial series stage from Figure 15.12:

1. *The input impedance of the series stage is decreased, this reduces the transadmittance of the stage.*

2. *The spectral densities of the equivalent input noise sources of the stages are increased.*

3. *The nonzero conductance of R3 increases the transadmittance of the stage.*

Another disadvantages of this implementation is that the PSSR for both positive and negative supply voltages is not as good as it would be using nonlinear resistors with a current source character.

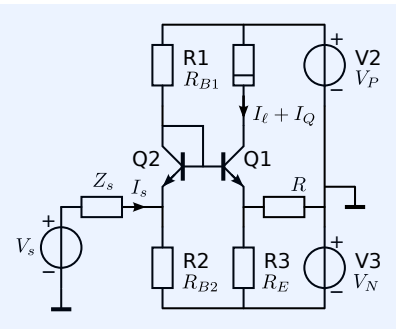


Figure 15.15: Possible final implementation of the series stage from Figure 15.12.

Although the biasing solution in presented in the preceding example seems attractive because of its low complexity, the penalties on the signal behavior may be too large. Hence, the design of the biasing elements should start with setting up their performance specification. This will be discussed at a later stage.

15.2.3 Biasing of cascode stages

In Chapter 14, we have introduced the cascode stage as a basic amplifier stage that consists of two transistors. The reason for this is that, when applied in negative feedback amplifiers, the CG or CB stage does not add a dominant pole to the loop gain poles product. In this section we will study the biasing of this cascode stage.

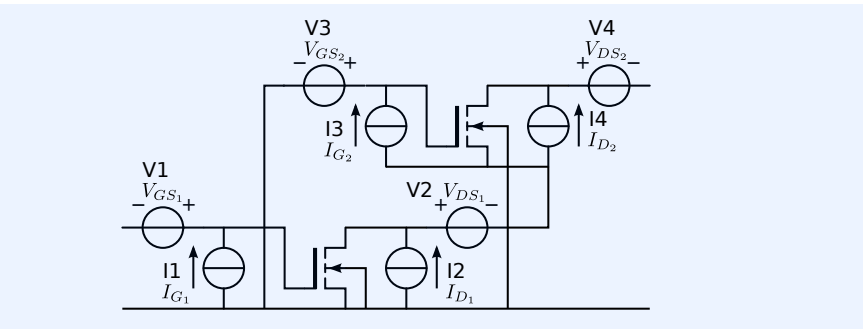


Figure 15.16: Initial schematic of a biased NMOS cascode stage.

Figure 15.16 shows an NMOS cascode stage with its initial bias sources. Each of the two transistors of the cascode stage is biased with the aid of four bias sources. The voltages and currents provided by these sources should be derived from the power supply voltage. Figure 15.17 shows two steps of this biasing process.

Figure 15.17A shows that the biasing scheme can be simplified by raising the voltage level of Q2 with respect to the ground. This is done by shifting the voltage source `V2` through the device. It then appears in series with `V3` and `V4`. The bias current sources have been redirected via the power supply and replaced with nonlinear resistors with a current source character.

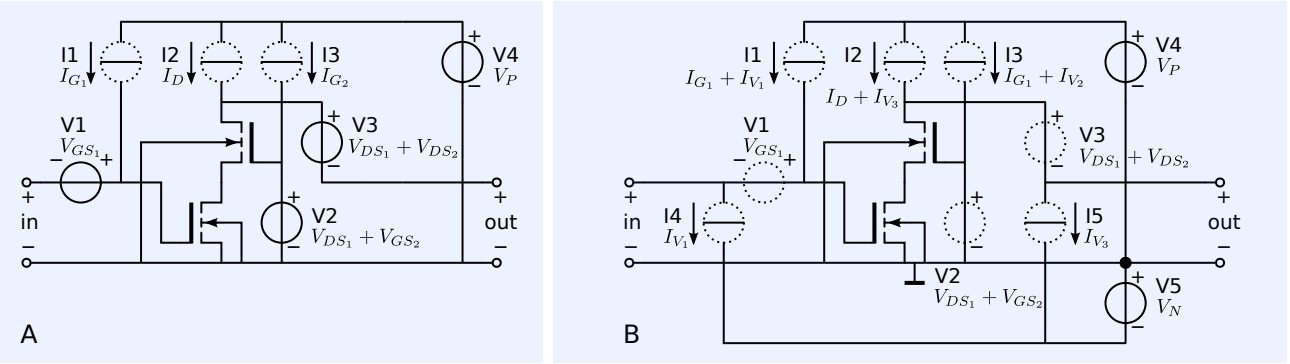


Figure 15.17: Biasing of the NMOS cascode stage:
A. Current sources replaced with biased passive elements
B. Voltage sources replaced with biased passive elements

Figure 15.17B shows the result after bias currents have been added through the bias voltage sources. By doing so, these bias sources can be replaced with passive nonlinear resistors. The negative supply source `V5` is added to obtain a positive voltage drop across the current source elements `I4` and `I5`. As a matter of fact, the positive supply voltage V_P, or the drain-to-source voltages of the transistors may need to be adjusted to obtain sufficiently large bias voltages for `I1`, `I2` and `I3`.

Changing the voltage level and the bias voltages of devices are techniques that can be used for simplification of the biasing scheme. These techniques will be discussed at a later stage.

Another technique that can be applied for simplification of the biasing scheme is the use of complementary transistors. Figure 15.18 shows the so-called *folded cascode* stage in which the CG stage transistor is the complementary type of the CS stage transistor. All the previously mentioned techniques have been used to simplify the biasing scheme that now requires current source elements only. The application of complementary devices will also be discussed at a later stage.

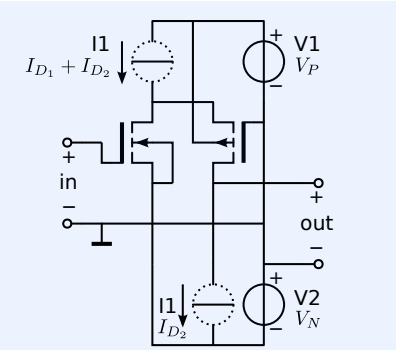


Figure 15.18: Biased folded cascode..

15.2.4 Biasing of anti-series stages

In this section, we will discuss the biasing of the balanced anti-series stages. In Chapter 6, we have seen that anti-series stages can be biased with common-mode current sources only (see Figure 6.31). Figure 15.19 shows the way in which this is done for an NMOS anti-series stage. The bias current sources have been replaced with using nonlinear resistors with a current source character. However, by doing so, the node voltages with respect to ground are undefined. In fact, in this biasing concept, the differential pair is floating with respect to the ground.

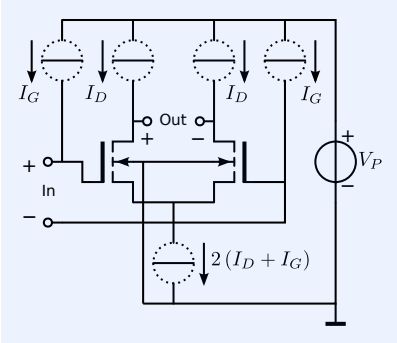


Figure 15.19: NMOS anti-series stage biased with with nonlinear resistors with a current source character. The common-mode voltages are yet undefined.

In practice, we need to fix the common-mode voltages of the input port and the output port with respect to ground in such a way, that under all operating conditions, the biasing elements and the transistors are properly biased. This can be accomplished in various ways. In Chapter 9, several methods for common-mode biasing have been introduced. In order to grasp the idea that these are the concepts behind the circuit implementations given below, the reader is invited to study this chapter.

Source or load common-mode biasing

Source or load common-mode biasing is a technique in which the common-mode voltage at the input port is defined by that of the source and/or the common-mode voltage at the output port is defined by that of the load. Figure 15.20 shows an application of this technique for the case in which both the source and the load have a common-mode connection, and the common-mode voltages are fixed with the aid of nonlinear resistive elements.

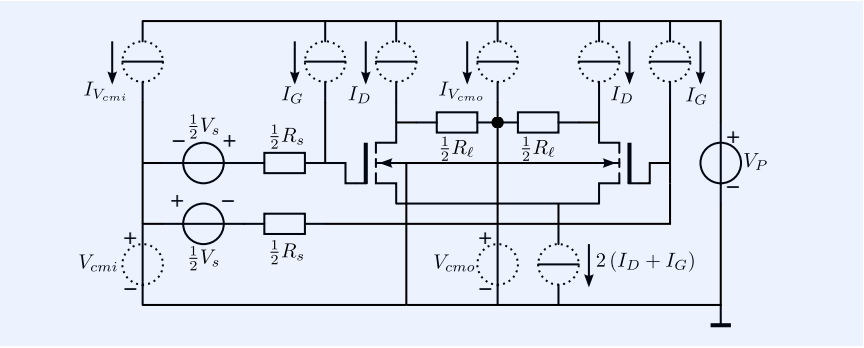


Figure 15.20: Differential-pair NMOS stage of which the common-mode voltages at the input port and the output port have been fixed by those of the source and the load, respectively.

This way of biasing requires a signal source or load that has a finite DC resistance to a common-mode voltage. Such source and load configurations are not always available. In those cases the common-mode voltages have to be defined by using different biasing techniques.

Brute force common-mode biasing

Brute force common-mode biasing is a technique in which the common-mode level at a port is defined by connecting the terminals of the port with resistors to a common-mode reference voltage. Figure 15.21 shows the application of this technique for fixing the common-mode voltage and the common-mode resistance of a signal voltage source that is floating with respect to the ground. The common-mode resistance of the source is reduced from infinity to R_{cm}. This way of biasing may deteriorate the signal performance of the stage:

1. The bias resistors generate noise and increase the noise contribution of the equivalent input noise sources of the differential pair
2. The bias resistors attenuate the source signal and increase the influence of the source impedance on the signal transfer of the differential pair.

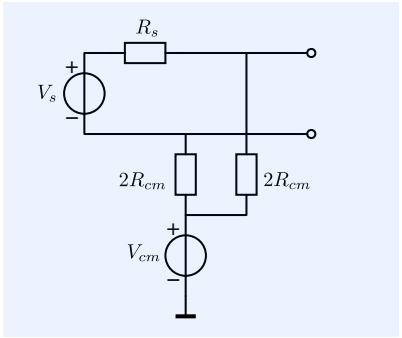


Figure 15.21: Method for fixing the common-mode voltage and the common-mode resistance of a floating source using a brute force resistive connection to a common-mode voltage source.

Local common-mode feedback

Similar to the design of feedback amplifiers (see Chapter 7), the common-mode impedances, voltages or currents can be fixed to an accurate value with the aid of negative feedback. Figure 15.22 shows the concept for fixing the common-mode impedance and bias voltage of a floating port.

The open-circuit common-mode voltage of this circuit equals V_{cm} and the common-mode impedance equals $\frac{1}{g_{cm}}$. The differential-mode impedance is not affected by the common-mode feedback.

Figure 15.23 gives an implementation of this principle with four equal NMOS transistors. In the quiescent operating point the gate-source voltages of these devices all equal V_{GS}, which is the common-mode voltage of the port. As a result of the finite output resistance of the MOSFETs, the differential input resistance of the port is limited to r_{ds}, which is the DC output resistance of one transistor in its operating point. For a large differential input voltage

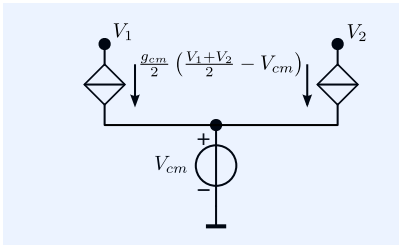


Figure 15.22: Method for creating a common-mode resistance $R_{cm} = \frac{1}{g_m}$ tied to a common-mode voltage V_{cm}, without affecting the differential-mode impedance.

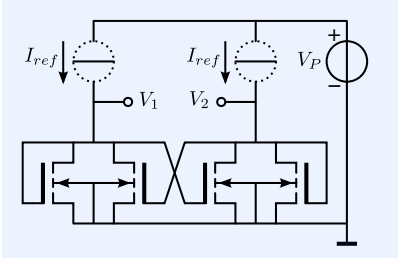


Figure 15.23: Implementation of the concept from Figure 15.22 with NMOS transistors and nonlinear resistors with a current source character.

range and a large differential-mode input resistance, the devices should be as long as possible and operate in strong inversion. In the following example we will plot the DC characteristics of the common-mode biasing circuit from Figure 15.23.

Example 15.3

In this example we will evaluate the DC port characteristics of the circuit from Figure 15.23. Figure 15.24 shows the test circuit and the DC characteristics.

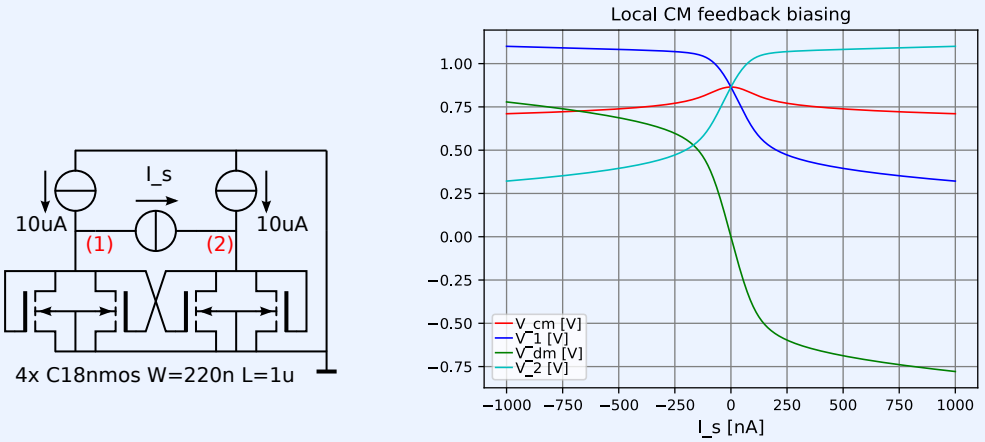


Figure 15.24:
Left: SPICE circuit for determination of the DC characteristics from Figure 15.23.
Right: Node voltages, common-mode and differential mode voltage as a function of the differential-mode input current.

The listing of the SPICE input file is shown below:

```
RcmVcmNMOS
* File: RcmVcmNMOS.cir
* LTspice circuit file
.include CMOS18TT.lib
M1 1 1 0 0 C18nmos W=220n L=1u
M2 1 2 0 0 C18nmos W=220n L=1u
M3 2 2 0 0 C18nmos W=220n L=1u
M4 2 1 0 0 C18nmos W=220n L=1u
I1 0 1 10u
I2 0 2 10u
I3 1 2 0
.dc I3 -1u 1u 1n
.end
```

Over a differential input voltage range of $\pm 0.25V$ the differential input resistance is about $10M\Omega$. Outside this range the input resistance drops and the differential-mode to common-mode conversion becomes significant.

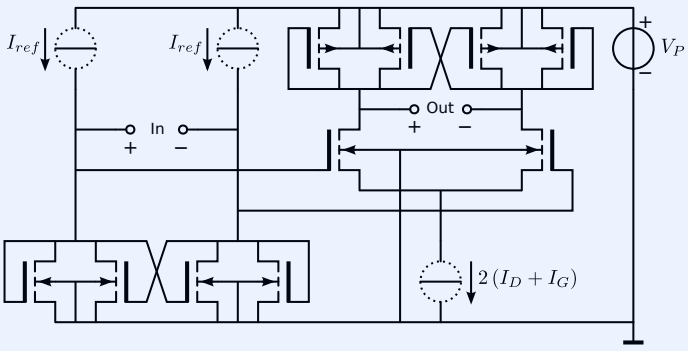


Figure 15.25: Anti-series CS stage of which the common-mode input voltage, the common-mode input resistance, the common-mode output voltage and the common-mode output resistance are fixed according to the method from Figure 15.22.

Figure 15.25 shows an arrangement in which this type of biasing has been

applied at both the input port and the output port of an anti-series CS stage. At the output port, the common-mode voltage is referred to the power supply voltage. It is controlled by four PMOS devices.

It should be noted that this way of common-mode biasing may result in a deterioration of the signal-to-noise ratio. Since the channel noise of the four MOS devices is uncorrelated, it contributes to differential current noise. This noise contribution can be kept low by making the bias currents of these devices as low as possible. In fact, this bias current should be large enough to maintain proper biasing for all values of the common-mode input current.

Another disadvantage of the circuit is its contribution to the differential-mode input capacitance of the biased port.

Feed forward common-mode biasing

Feed forward common-mode biasing is a technique in which the common-mode voltage or current at the output port is fixed by means of fixing the common-mode voltage or current at the input port, together with the common-mode transfer of a balanced stage. This technique can be used to establish a well-defined common-mode output voltage in anti-series connected stages with DC parallel feedback at the output port:

1. The balanced common drain stage or the balanced common collector stage
2. The balanced shunt stage
3. Anti-series connected multiple-loop passive feedback stages with parallel feedback at the output port.

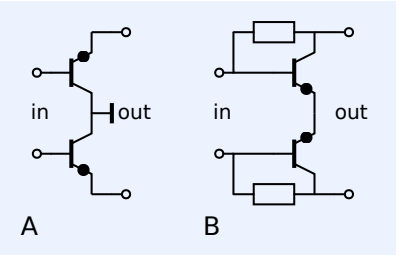


Figure 15.26: Anti-series connected stages with a well-defined common-mode transfer:
A. The anti-series connection of grounded voltage follower stages has a common-mode voltage gain of unity.
B. The floating anti-series connection of shunt stages has a common-mode voltage gain of unity.

Figure 15.26A shows the signal diagram of the balanced voltage follower, of which the common terminals of the stages (collectors or drains) have been connected to the ground. Please note that this does not affect the balanced operation if both the source and the load are truly balanced with respect to the ground. Since this stage has a common-mode voltage gain of unity, the common-mode voltage at the output port will be defined by that at the input port.

Figure 15.26B shows the signal diagram of a balanced shunt stage with resistive feedback. In this circuit, the common terminal of the stages (the emitters of sources) have been left floating with respect to the ground. In this case, the common-mode voltage gain of the stage equals unity and the common-mode voltage at the output port can be defined by that at the input port. If the common terminal of the two stages is grounded the common-mode transresistance equals the DC feedback resistance. In that case, the common-mode output voltage can be fixed with the aid of a common-mode input current.

Over-all common-mode feedback.

Over-all common-mode feedback uses the balanced stage to be biased as part of the circuit that controls the common-mode voltage at one of its ports. Figure 15.27 shows the application of overall-feedback for fixing the common-mode voltage at the input port of an anti-series CS stage. The common-mode voltage at the sources of the differential pair (`M1`, `M2`) is used to control the input common-mode current generated by `M3` and `M4`. The input common-mode voltage is set to the sum of the gate-to-source voltages of (`M1`, `M2`) and (`M3`, `M4`). These voltages are defined by the technology the geometry and the drain currents of the MOS transistor.

Figure 15.28 gives an arrangement in which the common-mode voltage at the output of the differential pair is also fixed by using over-all common-mode feedback. This voltage equals the sum of the gate-to-source voltages

of (M5, M6) and M7. The anti-series CD stage constructed with M5 and M6, detects the common-mode voltage at the output of the anti-series stage (M1, M2) and adds a voltage level shift to it. In order to keep the differential-mode to common-mode conversion low, the pair (M5, M6) should operate in strong inversion.

It should be noted that common-mode feedback may result in common-mode instability and frequency compensation techniques may be required to ensure frequency stability of the common-mode loop(s). In general, the bandwidth of the common-mode loop should be kept as large as possible. Bandwidth reduction of the common-mode loop, combined with differential to common-mode conversion may result in operating point shift and reduction of overdrive recovery. The design of the frequency behavior of the common-mode behavior proceeds similar to that of the differential-mode behavior.

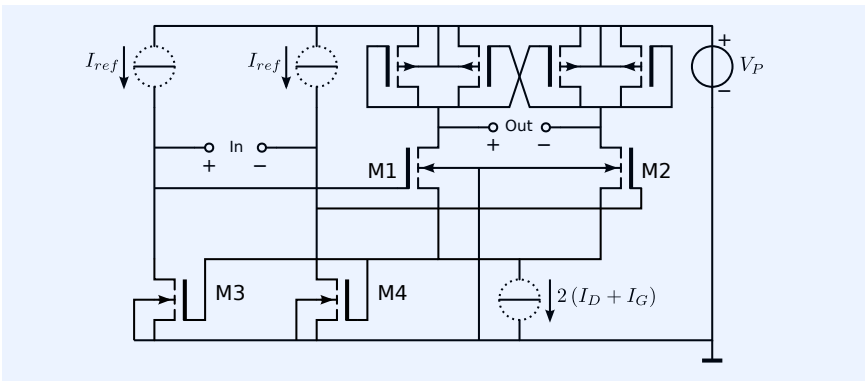


Figure 15.27: Anti-series CS stage of which the common-mode input voltage is fixed using over-all common-mode feedback, while the common-mode voltage of the output port is fixed with the aid of local common-mode feedback.

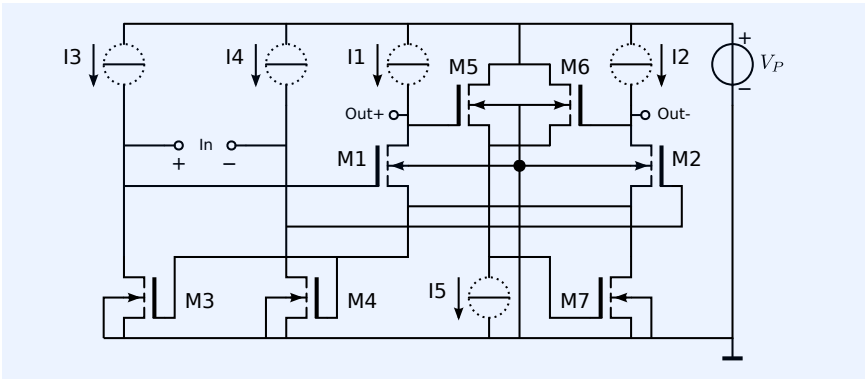


Figure 15.28: Anti-series CS stage of which the common-mode input and output voltages have been fixed using over-all common-mode feedback.

15.2.5 Biasing of complementary-parallel stages

In this section, we will discuss the biasing of the balanced complementary parallel stages. In Chapter 6, we have seen that these stages can be biased with common-mode voltage sources only (see Figure 6.40).

Figure 15.29 shows the biasing of the complementary parallel stage. Figure 15.29A shows the initial biasing scheme from Figure 6.40. Figure 15.29B shows the biasing with two supply sources and nonlinear resistors with a voltage source or current source character. Because of its importance as class AB output stage in operational amplifiers, the biasing of the complementary parallel stage has been a subject of research and many solutions have been proposed. The most important aspects related to the design of the biasing of the complementary parallel stage are:

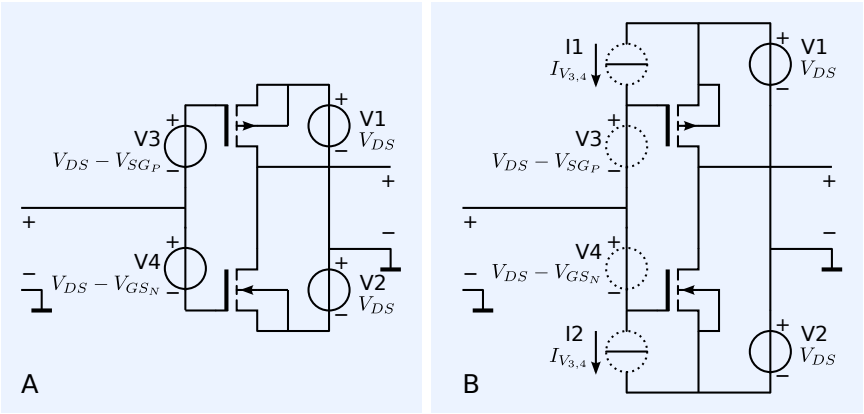


Figure 15.29: Biasing of the complemantary parallel stage:
A. Biasing scheme according to Figure 6.40
B. Biasing using nonlinear resistors with a current source and a voltage source character.

1. Accurately fixing of the quiescent current over a wide temperature range and supply voltage range.

   In order to make the quiescent (common-mode) current of the stage constant over a wide supply range and temperature range, the voltages of `V3` and `V4` should accurately track with the supply voltages and with the gate-to-source voltages of the MOS transistors.

2. Minimization of the cross-over distortion of the stage.

   If the complementary parallel stage operates in class AB mode, the transfer during the zone in which the two transistors contribute equally to the output current may strongly differ from the transfer in the regions in which either the P or the N device contributes to the output current. This behavior has already been illustrated in Figure 6.42. This change in gain during the take-over of the current causes so-called *cross-over distortion*. Minimization of this distortion requires optimization of the quiescent current of the complementary-parallel stage.

In this section we will present three solutions for the biasing of the complementary parallel stage. All solutions use model-based biasing techniques.

First we will introduce a low-voltage solution in which the power supply defines the biasing current. Secondly, we will present a solution for a complementary parallel stage that can be biased from a higher voltage and driven from a single input signal. Subsequently we will present a solution for the biasing of a class AB output stage in which the P device and the N device are driven from separate signals.[5] Examples of implementations will be given for MOS technology only.

[5] This will be referred to as a *split-signal* output stage.

As stated earlier, the design of class AB output stages has been a topic of intensive research. The reader is invited to study literature on this subject, investigate the concepts behind circuit implementations and find alternative implementations of those concepts.

Model-based biasing of the CMOS inverter

Figure 15.30 shows a CMOS complementary-parallel CS stage, of which the transistors are biased at zero drain-gate voltage. The quiescent current of the stage is determined by the power supply voltages. Hence, for accurate biasing at an operating point with minimized cross-over distortion, the supply voltages should track with the quiescent gate-source voltages. This can be done with the aid of a low-drop voltage regulator as shown in Figure 15.31. In the circuit shown in this figure, the total supply voltage $V_{SG_P} + V_{GS_N}$ is

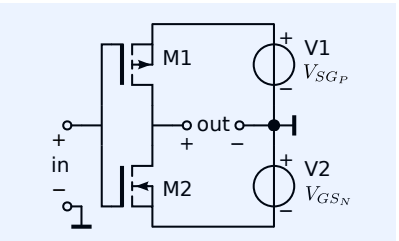


Figure 15.30: The CMOS inverter biased with the supply sources.

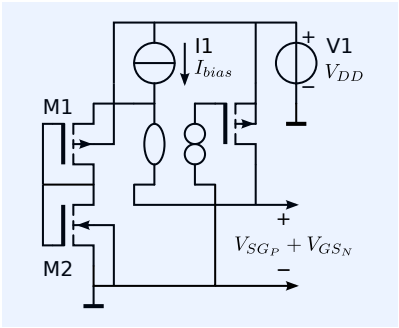


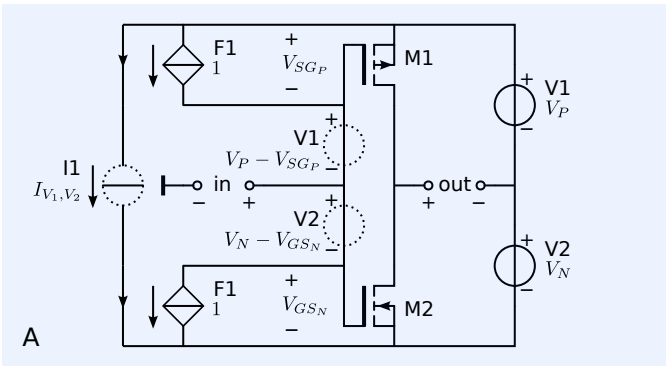
Figure 15.31: Supply voltage generation for accurate biasing of the CMOS inverter.

generated by letting the required bias current I_{bias} flow through the series connection of an NMOS and a PMOS with shorted gate-drain.

Model-based biasing with single input signal

Figure 15.32A shows the concept of the common-mode biasing of a complementary parallel stage CMOS stage using biasing with two floating voltage sources (`V1, V2`). In order to make the quiescent current of this stage independent of the temperature and insensitive for device tolerances, the voltages of these sources should track with the quiescent gate-to-source voltages of the MOS devices. This can be achieved using model-based biasing.

In the CMOS implementation of this biasing concept, from Figure 15.32B this is done by generating the reference current with two complementary MOS devices (`M3, M5`), biased at zero drain-to-gate voltage. Their quiescent operating conditions are copied to (`M4, M6`). Hence, the voltages across the two resistors (`R2, R3`) equal $V_{P-}V_{SG_P}$ and $V_N - V_{GS_N}$, respectively. If all N devices are equal, and all P devices are equal, the quiescent current of all complementary pair are equal. They do, however, depend on the power supply voltage.



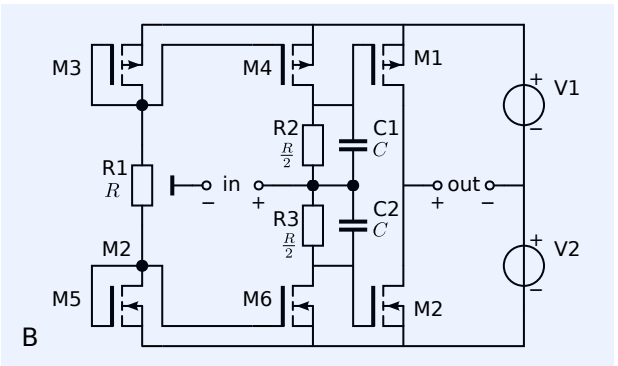


Figure 15.32: Model-based biasing of the complemantary parallel stage:
A. Biasing concept
B. Model-based CMOS implementation.

[6] A. Torralba, R. G. Carvajal, J. Ramirez-Angulo, J. Tombs, and T. Galan. Class ab output stages for low voltage cmos opamps with accurate quiescent current control by means of dynamic biasing. In *ICECS 2001. 8th IEEE International Conference on Electronics, Circuits and Systems (Cat. No.01EX483)*, volume 2, pages 967–970, 2001

Low-voltage push-pull stage biasing

A similar biasing method can be used for biasing a complementary-parallel stage at low supply voltages. We speak of low-voltage design if the power supply voltage can be as low as the sum of one gate-to-source or base-to-emitter voltage and one drain-to-source or collector-to-emitter saturation voltage.

A. Torralba, R. G. Carvajal, J. Ramirez-Angulo, J. Tombs and T. Galan [Torralba2001][6], present a solution for low-voltage biasing of a complementary-parallel stage of which the concept is shown in Figure 15.33.

In the right part of the circuit, the transistors `M3` and `M4`, together with the resistor `R2` that implements a floating voltage source and the power supply source `V1`, constitute the biased complementary parallel stage. The voltage across `R2` is generated by two controlled current sources of which the current is a copy of the output current of the controller in the bias control circuit. This bias control circuit is shown left from the dashed line. This circuit generates the current I_N that will cause a voltage drop of $V_{SG_P} + V_{GS_N} - V_P$ across a resistor with a resistance R, where V_{SG_P} and V_{GS_N} are the diving voltages for letting the PMOS and the NMOS operate at a quiescent current I_{bias}. This is achieved as follows.

The required bias current I_{bias} is generated by the reference current source I1. This current flows through the gate-drain shorted PMOS, while a copy of

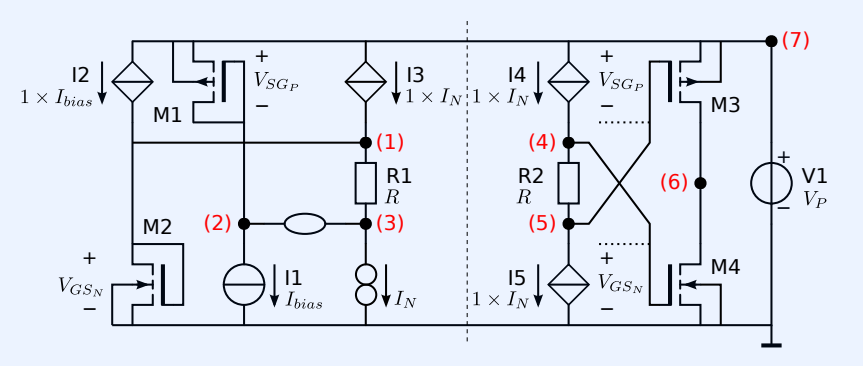


Figure 15.33: Low-voltage model-based biasing of the complemantary parallel stage.

it flows through the gate-drain shorted NMOS. In this way the two driving voltages for the MOS transistors V_{SG_P} and V_{GS_N} are obtained.

In order to create a voltage drop $V_{SG_P} + V_{GS_N} - V_P$ across a resistor, `R1` is tied to node (1) and a nullator is placed between node (2) and (3). The norator draws a current I_N through `R1`, which causes the required voltages drop across it. In order not to affect the current trough `M1`, a copy of the norator current is delivered by `I3`. In this way, the current I_N causes the required voltage drop $V_{SG_P} + V_{GS_N} - V_P$ across `R1`. The controlled current sources `I4` and `I5` provide a copy of this current for biasing `M3` and `M4` at a quiescent current I_{bias}.

It should be noted that the location of the norator and the current-controlled current source `I3` can be interchanged. This also holds for `I1` and `I2`.

[7] D. M. Monticelli. A quad CMOS single-supply opamp with rail-to-rail output swing. *IEEE Journal of Solid-State Circuits*, 21:1026–1034, December 1986

Model-based biasing with split input signal

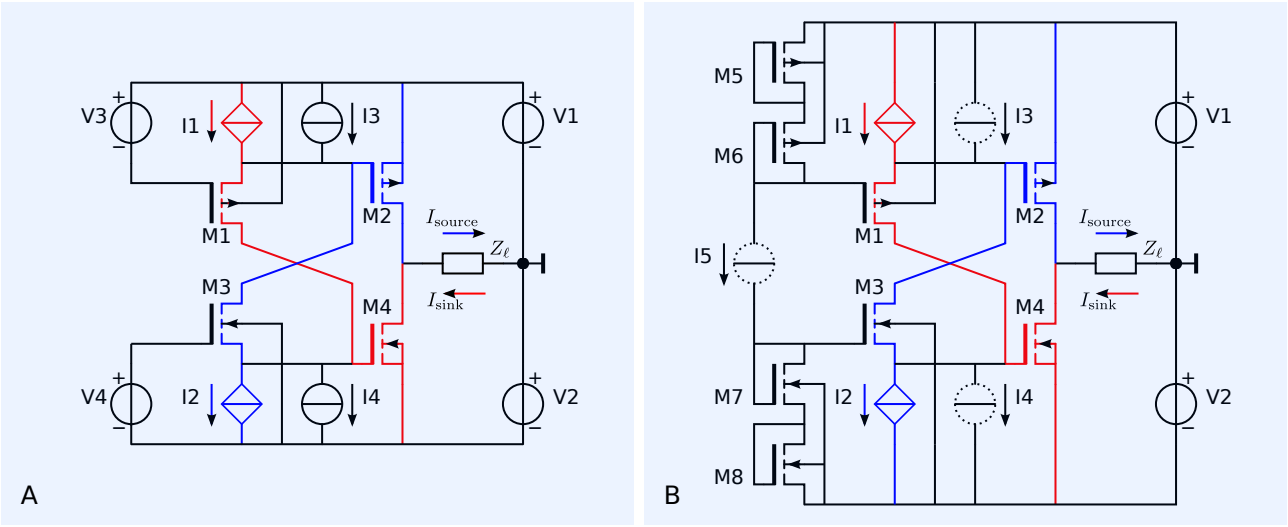


Figure 15.34: Output stage with split signal path for source and sink output current, presented by Monticelli.

Figure 15.34 shows a solution presented by Monticelli[Monticelli1986][7] that uses a so-called *split signal path*. The circuit consists of two amplifier halves, one providing the source output current and the other providing the sink output current. The source and the sink current are added by connecting the outputs of the source and the sink halves in parallel.

Figure 15.34A shows the basic concept. The controlled current sources (`I1, I2`) provide the signal current, while the bias voltages provided by `V1` and `V2` and the bias currents provided by `I3` and `I4` set the quiescent current of the output complementary CS stages (`M2, M4`). A positive signal current

provided by I1 causes a sink output (signal) current, while a positive signal current provided by I2 causes a source output (signal) current. The signal currents from (I1, I2) are passed through the common-gate stages (M1, M3) to the output transistors (M2, M4), respectively.

Figure 15.34B shows the implementation in which the current source element I5, together with M5 through M8 provide the model-based biasing voltages for the transistors in the signal path.

As a result of the inclusion of the common-gate stages in the signal path, circuit is not a true complementary-parallel connection of CS stages. Figure 15.35A shows the simplified signal diagram with generic biased devices and Figure 15.35B shows the small-signal equivalent circuit.

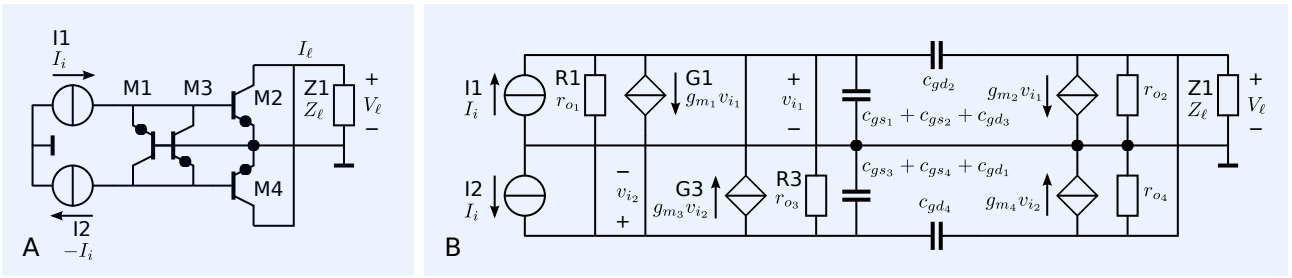


Figure 15.35: Signal diagram and small-signal equivalent circuit of the push-pull stage from Figure 15.34A.

Figure 15.35A shows that if the two inputs are connected together, the stage acts as complementary-parallel CS stage. For non-zero frequencies this can be achieved by placing a capacitor across the two input terminals.

In the quiescent operating point both the push and the pull stage contribute to the output current. If both stages have equal parameters, the controlled current sources G1 and G3 generate equal but opposite currents and the resistors R1 and R3 carry no voltage. The stage then acts as a parallel connection of two complementary CS stages.

In cases in which the output source or sink current exceeds the quiescent bias current, only the source or sink output transistor contributes to the output current.

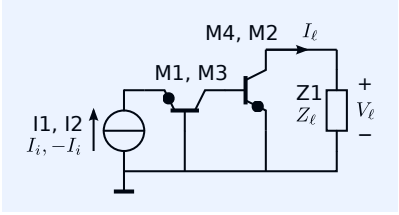


Figure 15.36: Signal diagram of the push-pull stage from Figure 15.34A for the sink/source phase.

Figure 15.36 shows the signal diagram and the small-signal equivalent circuit for the sink or the source phase. During the sink phase M1 and M4 are active, while during the source phase M3 and M2 are active. The stage then acts as a cascade connection of a common-gate and a common-source stage.

II

Background knowledge

16
Signal Modeling (selected topics)

16.1 Introduction

Signal modeling provides descriptions of signals that can be used for the evaluation of the performance of information processing systems. Signal modeling can be done using statistical or deterministic description methods. A deterministic model describes the signal as a time function, while random signal modeling uses statistical description methods such as amplitude distribution functions, average values, etc. It should be noted that signals are neither deterministic nor stochastic. It is up to the designer to decide whether to use deterministic or stochastic description methods.

Deterministic signal modeling will be discussed in section 16.2 and random signal modeling will be discussed in section 16.3.

16.2 Deterministic signal modeling

Deterministic signal modeling is based upon the resolution of signals into elementary or basic signals. This way of modeling is convenient because many information processing systems are intended to be linear and decomposition of arbitrary signals into elementary or basic signals facilitates the performance analysis of those systems using superposition. If an arbitrary signal can be written as a finite or infinite sum of basic signals that all have the same shape, the system response to an arbitrary signal can then easily be derived from the response to the basic signal. Basic signals that are often used for this purpose are:

1. Unit impulse and unit step functions

   In section 16.2.2, we will discuss the resolution of signals into a continuum of time-shifted impulse functions or of step functions. These signals have special interest for studying the performance of linear stationary dynamic systems. If the response to a single impulse or step is known,[1] the system response to an arbitrary signal can easily be related to it. Step functions are often used as test signals for those systems. The popular square wave test signal can be constituted from two step functions. System analysis using impulses or step functions as test signals is also referred to as *time domain analysis*.

[1] For example, by measurement or analysis.

2. Imaginary exponentials and sinusoidal signals

   In section 16.2.3, we will discuss the resolution of signals into a discrete series, or into a continuum of imaginary exponentials. Periodic signals can be resolved into a discrete series of imaginary exponentials. Such a resolution is described by the Fourier series, and by the sine or cosine transform. Non-periodic signals can, under certain conditions, be resolved into a continuum of imaginary exponential functions. The Fourier transform describes such a resolution. The Fourier transform only applies for signals that are absolute integrable. For a signal $x(t)$, this means that the integral of the absolute value of the signal over all time is finite:

$$\int_{-\infty}^{\infty} |x(t)|\, dt < \infty. \tag{16.1}$$

   Practically, this means that the signal values only differ from zero over a limited time interval. Signals that meet this condition are called *energy signals* or *pulse signals*.[2]

[2] Energy signals have a finite nonzero energy. Their average power over all time is zero.

   Since exponential signals and sinusoidal signals retain their shape under differentiation and integration, they are often used for characterization of linear dynamic systems. Any deviation of the shape of the response from a sinusoid, indicates system nonlinearity. The amount of nonlinearity of a system can thus be evaluated by measuring the distortion in

the system's response to a sinusoidal signal. Sinusoidal test signals can be constructed from two complex conjugated imaginary exponentials. System analysis with the aid of imaginary exponentials is referred to as *frequency domain analysis*.

3. Complex exponentials
   In section 16.2.6, we will discuss the resolution of signals into a continuum of complex exponentials. This resolution is described by the Laplace transform. Application of the Laplace transform is not limited to energy signals; it can also be applied for signals that are not absolutely integrable. Theoretically, these signals may occur in unstable systems: systems that have an unbounded response to a bounded excitation.

Above, we have introduced the concepts of *power signals* and *energy signals*. Before we continue with signal modeling, we will give the definitions of power signals and energy signals.

16.2.1 Power signals and energy signals

The average power that can be extracted from a signal $x(t)$ is proportional to the mean square value $\overline{x(t)^2}$ of that signal. The mean square value is defined as

$$\overline{x(t)^2} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \{x(t)\}^2 dt. \tag{16.2}$$

A signal with a nonzero mean square value is called a *power signal*. If the mean square value of a signal equals zero, it is called an *energy signal*. The energy $W\{x(t)\}$ of a signal $x(t)$ is defined as

$$W\{x(t)\} = \lim_{T\to\infty} \int_{-T/2}^{T/2} \{x(t)\}^2 dt. \tag{16.3}$$

The energy of a power signal is (theoretically) unlimited. In the real world, all signals are energy signals, but signal modeling uses abstractions for which it is useful to make a distinction between power signals and energy signals.

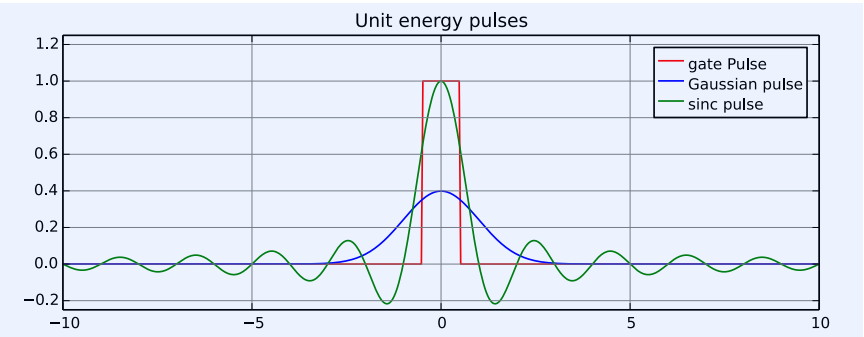


Figure 16.1: Three examples of pulse signals with unit energy and unit width ($\tau = 1$).

Figure 16.1 shows some examples of energy signals with their energy normalized to unity.

The unit gate pulse

The unit gate pulse signal $x(t, \tau)$ is defined as:

$$x(t, \tau) = 1, \quad |t| < \frac{\tau}{2}, \tag{16.4}$$

$$x(t, \tau) = 0, \quad |t| \geq \frac{\tau}{2}. \tag{16.5}$$

The unit Gaussian pulse

A unit Gaussian pulse signal $x(t, \tau)$ is defined as

$$x(t, \tau) = \frac{1}{\tau\sqrt{2\pi}} \exp\left(\frac{-t^2}{2\tau^2}\right). \tag{16.6}$$

The unit sinc pulse

A unit sinc pulse signal $x(t, \tau)$ is defined as

$$x(t, \tau) = \frac{\sin(\pi t/\tau)}{\pi t}. \tag{16.7}$$

The unit impulse $\delta(t)$

The unit impulse function can be obtained as $\lim_{\tau \to 0}$ of all three above pulse signals. The area (strength) of the unit impulse function is unity

$$\int_{-\infty}^{\infty} \delta(t) = 1, \tag{16.8}$$

but the value at time instants other than $t = 0$ equals zero:

$$\delta(t) = 0 \text{ for } t \neq 0. \tag{16.9}$$

16.2.2 Time-domain modeling of signals

In this section, we will discuss the resolution of arbitrary time functions $x(t)$ into unit impulse and unit step functions. We will see that an arbitrary time function can be considered to be constituted from a linear superposition of time delayed unit impulse- or unit step functions. As a consequence, the behavior of linear systems can be written as a linear superposition of the responses to these unit impulses or unit steps. This superposition, which is known as the convolution integral, is often used for graphical determination of the response of a linear system to an arbitrary input signal. Linear dynamic systems can thus be characterized by their so-called unit impulse response or unit step response. Step functions are often used as test signals for the verification of the behavior of linear dynamic systems.

Resolution in unit impulse functions

A signal $x(t)$ can be resolved into a continuum of time-shifted unit impulses. We will do this for a signal that has nonzero values between $t = 0$ and $t = T$. The signal value at time instant $0 < \tau < T$ is then obtained as the strength of a unit impulse at $t = \tau$ with the scalar $x(\tau)$:

$$x(t) = \int_0^T x(\tau)\delta(t - \tau)d\tau. \tag{16.10}$$

Resolution in unit step functions

The resolution of a signal in unit step functions $\mu(t)$ can be found in a similar way:

$$x(t) = \int_0^T \dot{x}(\tau)\mu(t - \tau)d\tau, \tag{16.11}$$

where $\dot{x}(t) = \frac{d}{dt}\{x(t)\}$.

16.2.3 Frequency-domain modeling of signals

Periodic signals can be modeled as a linear superposition of undamped sinusoidal signals. These signals play an important role in the verification of dynamic systems. This is due to the property that linear dynamic systems with sinusoidal excitations give sinusoidal responses.

The Fourier series description is based upon the resolution of periodic signals into a series of imaginary exponentials. It can be shown that an arbitrary signal $x(t)$, within a limited interval $-\frac{T}{2} < t < \frac{T}{2}$, under certain conditions may be resolved into a series of imaginary exponentials. It will be clear that only if $x(t)$ itself is periodic with period T does this decomposition hold for *all* values of t. The set of complex amplitudes X_i of the imaginary exponentials constituting this periodic signal $x(t)$ is called the Fourier series for $x(t)$:

$$x(t) = \sum_{n=-\infty}^{n=\infty} X_n \exp(jn\omega_0 t). \tag{16.12}$$

The complex amplitudes X_n are obtained as

$$X_n = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\exp(-jn\omega_0 t)dt. \tag{16.13}$$

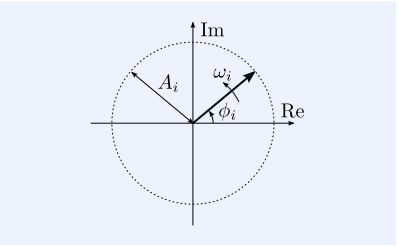


Figure 16.2: An imaginary exponential $A_i \exp(\omega_i t + \phi_i)$ represented by a rotating phasor.

Phasor representation

A single exponential component can be depicted as a phasor (rotating arrow) by plotting the imaginary part along the y-axis and the real part along the x-axis. This is shown in Figure 16.2. Any signal that can be resolved into complex exponentials can be written as a superposition of phasors. Phasor representations of signals are often used for modulated signals. Figure 16.3 shows a real sinusoidal signal constructed from two complex conjugated phasors.

The complex amplitude X_n of an element of the Fourier series can be represented by a phasor with amplitude $|X_n|$ and initial phase φ_n rotating with an angular speed $n\omega_0$.

$|X_n|$ and φ_n are known as the discrete magnitude and phase spectrum of the signal $x(t)$, respectively.

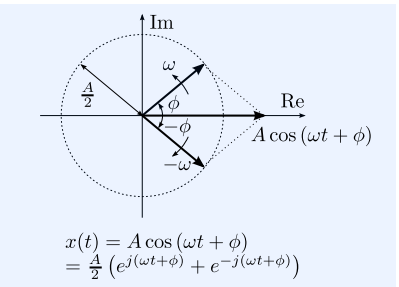


Figure 16.3: Two phasors constitute a real sinusoidal signal.

Power spectrum, Parseval's theorem

The mean square value of a periodic signal $x(t)$ can be obtained from the complex amplitudes of the elementary sinusoids from which it is constituted. Parseval's theorem asserts that the mean square value of a periodic signal equals the sum of the squares of the amplitudes of the harmonics that compose that signal:

$$\overline{\{x(t)\}^2} = \sum_{n=-\infty}^{n=\infty} X_n.X_n^* = \sum_{n=-\infty}^{n=\infty} |X_n|^2. \tag{16.14}$$

The graphical representation of $|X_n|^2$ is called the power spectrum of the periodic signal $x(t)$.

AC and DC signals

The mean value, the time-average value, or the DC value of a signal $x(t)$ over a time interval T, is defined as

$$x_{DC} = \overline{\{x(t)\}} = \frac{1}{T}\int_{-T/2}^{T/2} x(t)dt. \tag{16.15}$$

The AC signal is the signal that remains after subtraction of the DC value from the original signal:

$$x_{AC}(t) = x(t) - x_{DC}. \tag{16.16}$$

The power of a signal can be written in terms of its DC value and the power of the AC signal:

$$\overline{\{x(t)\}^2} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \{x_{AC}(t) + x_{DC}\}^2 dt \tag{16.17}$$

Since the average value of $x_{AC}(t)$ equals zero, we obtain

$$\overline{\{x(t)\}^2} = \overline{\{x_{AC}(t)\}^2} + x_{DC}^2. \tag{16.18}$$

The RMS (root mean square) value x_{rms} of a signal $x(t)$ is defined as

$$x_{\text{rms}} = \sqrt{\overline{\{x(t)\}^2}} = \sqrt{x_{AC_{rms}}^2 + x_{DC}^2}. \tag{16.19}$$

A signal with an RMS value x_{rms} delivers the same power to a load as a signal having a constant value of $x_{DC} = x_{\text{rms}}$.

16.2.4 Cosine transformation

A periodic signal can also be resolved in a series of cosine functions. This resolution can be written as

$$x(t) = x_{DC} + \sum_{n=1}^{\infty} C_n \cos(n\omega_0 t - \varphi_n). \tag{16.20}$$

in which $C_n = \sqrt{A_n^2 + B_n^2}$, and $\varphi_n = \arctan \frac{B_n}{A_n}$ with

$$A_n = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \cos(n\omega_0 t)\ dt, \tag{16.21}$$

$$B_n = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \sin(n\omega_0 t)\ dt. \tag{16.22}$$

This follows directly from the Fourier series description of a periodic signal. With the aid of Parseval's theorem for periodic signals, the mean square value of a periodic signal $x(t)$ can thus be written as

$$\overline{\{x(t)\}^2} = x_{DC}^2 + \frac{1}{2} \sum_{n=0}^{n=\infty} C_n^2. \tag{16.23}$$

16.2.5 Fourier transform

Non-periodic signals cannot be modeled by a discrete Fourier series. Under certain conditions, however, these signals can be resolved into a continuum of imaginary exponentials. An important condition is that the signals are absolute integrable:

$$\int_{-\infty}^{\infty} |x(t)| dt < \infty. \tag{16.24}$$

For real-world pulse signals, the conditions are always satisfied. We can then define an energy spectrum by letting the period T approaching infinity. It is apparent that such a resolution will then have frequency components at all frequencies, but the amplitudes of these components will all approach zero. For this reason, we speak of a frequency density spectrum. The resolution of a signal into a continuum of imaginary exponentials is known as the

inverse Fourier transform:

$$x(t) = \mathcal{F}^{-1}\{X(j\omega)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\omega)\exp(j\omega t)d\omega. \tag{16.25}$$

The complex amplitude function $X(j\omega)$ is known as the Fourier transform of $x(t)$:

$$X(j\omega) = \int_{-\infty}^{\infty} x(t)\exp(-j\omega t)dt. \tag{16.26}$$

The functions $x(t)$ and $X(j\omega)$ form a so-called Fourier pair for which we will use the short hand notation:

$$X(j\omega) = \mathcal{F}\{x(t)\}. \tag{16.27}$$

Some frequently used Fourier pairs are listed in Table 16.1.

time description	**time function**	**frequency function**	**frequency description**
impulse at $t = 0$	$\delta(t)$	1	flat spectrum
impulse at $t = t_0$	$\delta(t - t_0)$	$\exp(-j\omega t_0)$	imaginary exponential
impulse sequence	$\sum_{-\infty}^{\infty}\delta(t - nT)$	$\frac{2\pi}{T}\sum_{-\infty}^{\infty}\delta\left(\omega - \frac{2\pi n}{T}\right)$	impulse sequence
real sinusoid	$\cos(\omega_0 t)$	$\pi\delta(\omega \pm \omega_0)$	impulse pair
Gaussian pulse	$\exp(-at^2)$	$\sqrt{\frac{\pi}{a}}\exp\left(\frac{-\omega^2}{4a}\right)$	Gaussian pulse
unit gate pulse	$G_{\Delta\tau}(t)$	$\tau\frac{\sin(\omega\tau/2)}{\omega\tau/2}$	sinc function

Table 16.1: Fourier pairs

Important properties of the Fourier transformation are shown in table 16.2.

property	**time function**	**frequency function**
linearity	$a\,x_1(t) + b\,x_2(t)$	$a\,X_1(j\omega) + b\,X_2(j\omega)$
scaling	$x(at)$	$\frac{1}{\lvert a\rvert}X\left(\frac{j\omega}{a}\right)$
time shift	$x(t - \tau)$	$X(j\omega)\exp(-j\omega\tau)$
frequency shift	$x(t)\exp(j\omega_0 t)$	$X(j\omega - j\omega\tau)$
duality	$X(-t)$	$2\pi x(j\omega)$
time derivative	$\frac{d^n}{dt^n}\{x(t)\}$	$(j\omega)^n X(j\omega)$
frequency derivative	$(-jt)^n x(t)$	$\frac{d^n}{d\omega^n}X(j\omega)$
time integration	$\int_{-\infty}^{t} x(\tau)\,d\tau$	$\frac{1}{j\omega}X(j\omega) + \pi X(0)\,\delta(\omega)$
average power	$\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}\{x(t)\}^2\,dt$	$\lim_{T\to\infty}\frac{1}{2\pi T}\int_{-\infty}^{\infty}\lvert X(j\omega)\rvert^2\,d\omega$
total energy	$\lim_{T\to\infty}\int_{-T/2}^{T/2}\{x(t)\}^2\,dt$	$\frac{1}{2\pi}\int_{-\infty}^{\infty}\lvert X(j\omega)\rvert^2\,d\omega$
time-domain convolution	$x(t) * h(t)$	$X(j\omega)\,H(j\omega)$
time domain multiplication	$x(t)\,g(t)$	$\frac{1}{2\pi}X(j\omega) * G(j\omega)$
time domain correlation	$\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}x(t)\,x(t+\tau)\,dt$	$\lim_{T\to\infty}\frac{1}{2\pi T}\lvert X(j\omega)\rvert^2$

Table 16.2: Properties of the Fourier Transform

Energy density spectrum and Parseval's theorem

The energy of a pulse signal $x(t)$ may as well be obtained from its frequency domain description $X(j\omega)$. This relation is given by Parseval's theorem, given in 16.28:

$$W\{x(t)\} = \lim_{T\to\infty}\int_{-T/2}^{T/2}\{x(t)\}^2 dt = \frac{1}{2\pi}\int_{-\infty}^{\infty}|X(j\omega)|^2 d\omega, \tag{16.28}$$

where $|X(j\omega)|^2$ is called the energy density spectrum of $x(t)$.

Fourier transform of periodic signals

The Fourier transform gives the spectral density of a signal. Since periodic signals have a discrete amplitude spectrum, the density must be infinite at multiples of the fundamental frequency. From this, we expect the Fourier

transform of a periodic signal to be written as a series of frequency impulses occurring at $\omega = n\omega_0$. Moreover, we expect the strength of the impulses to be $2\pi X_n$. Without exact mathematical derivation, it follows that the Fourier transform of a periodic signal can be written as

$$X(j\omega) = \mathcal{F}\left\{\sum_{n=-\infty}^{n=\infty} X_n \exp(jn\omega_0 t)\right\} = \sum_{n=-\infty}^{n=\infty} 2\pi X_n \delta(\omega - n\omega_0). \qquad (16.29)$$

16.2.6 Complex frequency domain modeling

As discussed earlier, periodic signals and absolute integrable pulse signals can be resolved into undamped sinusoids. Many theoretical signals have no power limitation (unbounded signals) and are not absolutely integrable. These signals can be decomposed into complex exponentials. The inverse Laplace transform of $X(s)$, denoted as $\mathcal{L}^{-1}\{X(s)\}$, writes a time function $x(t)$ as a continuum of complex exponentials:

$$x(t) = \mathcal{L}^{-1}\{X(s)\} = \frac{1}{2\pi j}\oint_{\sigma-j\omega}^{\sigma+j\omega} X(s)\exp(st)ds. \qquad (16.30)$$

The continuum of the complex coefficients $X(s)$ is called the Laplace transform of $x(t)$:

$$X(s) = \mathcal{L}\{x(t)\} = \int_{-\infty}^{\infty} x(t)\exp(-st)dt. \qquad (16.31)$$

The time function $x(t)$ and its Laplace transform $X(s)$ form a so-called Laplace pair for which we will use the short hand notation:

$$X(s) = \mathcal{L}\{x(t)\}. \qquad (16.32)$$

Some often used Laplace pairs are given in Table 16.3.

Table 16.3: Laplace pairs

time description	time function	frequency function
unit impulse at $t = 0$	$\delta(t)$	1
unit step at $t = t_0$	$\mu(t)$	$\frac{1}{s}$
power function	$\frac{t^{n-1}}{(n-1)!}$	$\frac{1}{s^n}$
real exponential	$\exp(-at)$	$\frac{1}{s+a}$
cosine function	$\cos(\omega t)$	$\frac{s}{s^2+\omega^2}$
sine function	$\sin(\omega t)$	$\frac{\omega}{s^2+\omega^2}$

The function $X(s)$ from expression 16.31 is called the two-sided Laplace transform of $x(t)$. Most problems can be formulated in a manner that permits all signals to be zero for $t < 0$. The Laplace transform then reduces to the one-sided Laplace transform:

$$\mathcal{L}\{x(t)\} = \int_{0}^{\infty} x(t)\exp(-st)dt. \qquad (16.33)$$

For analysis of linear dynamic systems, the state of the system before $t = 0$ can be described by means of so-called initial conditions.

Some important properties of the Laplace transform are listed in Table 16.4.

property	time function	Laplace transform
linearity	$a\{x_1(t)\}+b\{x_2(t)\}$	$a\{X_1(s)\}+b\{X_2(s)\}$
scaling	$x\left(\frac{t}{a}\right)$	$aX(as)$
time-domain shift	$x(t-\tau)\,\mu(t-\tau)$	$X(s)\exp(-s\tau)$
s-domain shift	$x(t)\exp(-at)$	$X(s+a)$
time derivative	$\frac{d^{n-1}}{dt^{n-1}}\{x(t)\}\|_{t=0}$	$s^nX(s)-s^{n-1}x(0)-...$ $...-s^0\frac{d^{n-1}}{dt^{n-1}}\{x(t)\}\|_{t=0}$
s-domain derivative	$tx(t)$	$\frac{d}{ds}X(s)$
time-domain integration	$\int_{-\infty}^{t}x(\tau)\,d\tau$	$\frac{1}{s}X(s)+\frac{1}{s}\int_{-\infty}^{t}x(\tau)d\tau\|_{t=0}$
time-domain convolution	$x(t)*h(t)$	$X(s)\,H(s)$
time-domain multiplication	$x(t)\,g(t)$	$X(s)*G(s)$
initial value theorem	$x(0^+)=\lim_{s\to\infty}sX(s)$	
final value theorem	$\lim_{t\to\infty}x(t)=\lim_{s\to 0}sX(s)$	

Table 16.4: Properties of the Laplace transform

16.3 Random signal modeling

Random processes are characterized by an ensemble of sample functions generated by the process. As an example of a random process consider the noise generation in a resistor. Across the terminals of the resistor, a noise voltage can be measured, recorded, and plotted versus time. We would like to have expressions for the time average, the average power, the RMS value and the power spectral density for this signal. This, however, is not possible, since we do not have a time-domain description of a random signal. In many situations, we can obtain these quantities from a statistical description of the random process using its so-called probability density function.

To do so, we make different recordings for a large number of equal resistors. These recordings are made over a limited time interval T. This so-called ensemble of truncated sample signals $x_{T_i}(t)$, corresponding to the sample signals from the different resistors $x_i(t)$, can be defined as

$$x_{T_i}(t)=\begin{cases} x_i(t) & 0\le t\le T \\ 0 & t<0,\ t>T \end{cases}. \tag{16.34}$$

Now, we consider a specific time instant t_1. The amplitudes of the truncated sample functions at that time instant constitute a random variable $\underline{x}_1(t_1)$. The possible values of this random variable are denoted as x_1. We are now able to set up an histogram by counting the number of samples of $\underline{x}_1(t_1)$ that fall within a range between x_1 and $x_1+\Delta x_1$, for all successive values of x_1. This histogram is a discrete approximation of the probability density function $P(\underline{x},t_1)$ of the random variable $\underline{x}$ at a time instant $t=t_1$. The probability that a random variable $\underline{x}(t)$ has a value between a and b, at time instant t is then obtained as

$$\Pr(a\le\underline{x}\le b,t)=\int_a^b P(\underline{x},t)dx. \tag{16.35}$$

By definition:

$$\int_{-\infty}^{\infty}P(\underline{x},t)dx=1. \tag{16.36}$$

Some examples of probability density functions are uniform and Gaussian probability functions. Uniform probability functions have a probability $P(\underline{x},t)$ that does not depend on x. A Gaussian probability density function is characterized by its mean value μ and its root mean square value σ for $\mu=0$:

$$P(\underline{x},t)=\frac{1}{\sigma\sqrt{2\pi}}\exp-\frac{\{x(t)-\mu\}^2}{2\sigma^2}. \tag{16.37}$$

16.3.1 Stationary and ergodic processes

A random process is called stationary if its statistical properties do not change with time. For stationary processes, the averages, the correlation functions and the power spectral densities do not depend on time. It is important to notice that, from the observer's point of view, a process can be called stationary as long as its properties do not change appreciably over the time of interest.

A process is called ergodic if it can fully be described by one sample function only. For ergodic processes, we may obtain the values for the random variable $\underline{x}(t)$ only from one truncated sample function $x_T(t)$. Ergodic processes are always stationary. A stationary process, however, is not necessarily ergodic. Knowledge about the random process that generates the information is indispensable for setting up methods for extracting characteristic properties of the information from the process. This extraction is needed to find the characteristic properties of the information-carrying signals and to formulate the requirements for the information processing system. For the verification of the behavior of an information processing system, relevant properties of the deterministic test signals must be related to these characteristic properties of the information source. For this purpose, we will introduce some characteristic properties of signals, generated by random processes, and relate them to properties of deterministic signals that are often used as test signals.

16.3.2 Time average and ensemble average

The ensemble average, or the *expectation* $E\{\underline{x}(t)\}$ of a random variable $\underline{x}(t)$ at time instant t, is defined as the weighted sum of all values of the sample functions at that time instant t, each weighted according to its probability:

$$E\{\underline{x}(t)\} = \int_{-\infty}^{\infty} xP(\underline{x}, t)dx. \tag{16.38}$$

If we have n recordings or truncated sample functions $x_{T_i}(t)$, we can estimate the expectation at $t = t_1$ from the average value of the n samples $\overline{\underline{x}_1(t_1)}$:

$$\overline{\underline{x}_1(t_1)} = \frac{1}{n}\sum_{i=1}^{i=n} x_i(t_1). \tag{16.39}$$

The time average $\overline{x_i(t)}$ of the sample signal $x_i(t)$ is the average value of the truncated recording $x_{T_i}(t)$. It is obtained as

$$\overline{x_i(t)} = \frac{1}{T}\int_0^T x_{T_i}(t)dt. \tag{16.40}$$

If a process is stationary, the ensemble average does not depend on time. If the random process is also ergodic, the ensemble average $E\{\underline{x}_1(t_1)\}$ of a random variable $\underline{x}(t)$ equals the time average $\overline{x_i(t)}$ of one recording $x_i(t)$.

16.3.3 Correlation function

We have seen that a random variable at time instant $t = t_1$ can be described by its probability density function $P(x, t_1)$. Now, let us assume that, in another random process, a random variable $\underline{y}(t)$ at $t = t_2$ is described by its probability density function $P(y, t_2)$. The so-called joint probability density function $P(x, t_1; y, t_2)$ gives the probability that $\underline{x}(t_1)$ has a value between x and $x + dx_1$, while $\underline{y}(t_2)$ has a value between y and $y + dy$, which can be written as $P(x, t_1; y, t_2)dxdy$. The correlation function tells us something about the similarity between these two processes. It is defined as the expectation of the

product of the two random variables $\underline{x}(t_1)$ and $\underline{y}(t_2)$ at time instants t_1 and t_2, respectively:

$$E\{\underline{x}(t_1)\underline{y}(t_2)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyP(x,t_1;y,t_2)dxdy. \tag{16.41}$$

If both processes are stationary, the correlation function only depends on the time difference $t_2 - t_1$. If, in addition, the process is ergodic, we may approximate the correlation using two truncated recordings $x_T(t)$ and $y_T(t)$. For a time difference $t_2 - t_1 = \tau$, we may then estimate the correlation function $r_{xy}(\tau)$ of two sample signals $x_T(t)$ and $y_T(t)$ generated by an ergodic process as

$$r_{xy}(\tau) = \frac{1}{T}\int_0^T x_T(t)y_T(t+\tau)dt. \tag{16.42}$$

From this expression, we intuitively expect that the correlation function can be interpreted as a measure for the joint power of $x(t)$ and $y(t)$. This can be expressed by saying that two signals are correlated when they have a nonzero joint power. As a consequence, the similarity between two signals $x(t)$ and $y(t)$ can also be found from the average power of the sum of both signals and the sum of the average powers of the two signals individually. The average power of their sum is proportional to

$$\overline{(x+y)^2} = \overline{x^2} + \overline{y^2} + 2\overline{xy}. \tag{16.43}$$

The third term in the expression can be written as $2r_{xy}(0)$, and if it is zero, the sum of the powers of $x(t)$ and $y(t)$ equals the power of the sum of both signals. Two signals $x(t)$ and $y(t)$ are said to be uncorrelated or orthogonal when $r_{xy}(0) = 0$.

16.3.4 Autocorrelation function

Of particular importance is the so-called autocorrelation function. This function tells us something about the correspondence between the values of one random variable, at two time instants. If the rate of change of a random variable is large, the correspondence will rapidly drop with the difference between the two time instants. The autocorrelation function is defined as the expectation of the product of a random variable at one time instant and the same random variable at a second time instant. It is written as:

$$E\{\underline{x}(t_1)\underline{x}(t_2)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2P(x,t_1;x,t_2)dx^2. \tag{16.44}$$

For an ergodic process, the autocorrelation function $r_x(\tau)$ may be approximated using one sample function of the process with its replica, shifted τ in time. It is an important characteristic for the correspondence between signal values at time distance τ and tells us something about the rate of change of a signal:

$$r_x(\tau) = \frac{1}{T}\int_0^T x_T(t)x_T(t+\tau)dt. \tag{16.45}$$

As an example, consider random noise $n(t)$, which has been passed through a low-pass filter. Since the output noise of the filter cannot change rapidly, the similarity between values will be large for small values of τ. For increasing values of τ, the similarity will become smaller, and finally, $r_x(\tau)$ will drop to zero.

16.3.5 Mean square value

The mean square value of a random variable $\underline{x}(t)$ is defined as the expectation of the squared value of the random variable that equals the value of the autocorrelation function for $t_1 = t_2$:

$$E\{\underline{x}(t)\}^2 = \int_{-\infty}^{\infty} x^2 P(\underline{x};t)dx. \tag{16.46}$$

If we are dealing with ergodic processes, the mean square value of a truncated sample function random $x_T(t)$ can be estimated from $r_x(0)$ as

$$\overline{\{x(t)\}^2} = r_x(0) = \frac{1}{T}\int_0^T \{x_T(t)\}^2 dt. \tag{16.47}$$

16.3.6 Wiener-Khintchine theorem

We have seen that the autocorrelation function $r_x(\tau)$ tells us something about the correspondence between signal values at two time instants. When $r_x(\tau)$ drops to zero for a very small value of τ, the signal $x(t)$ changes rapidly and its bandwidth must be large. The relation between the frequency contents of a signal and its autocorrelation function is given by the Wiener-Khintchine theorem. This theorem asserts that for a stationary process, the power spectral density $S_x(\omega)$ is the Fourier transform of the autocorrelation function:

$$S_x(\omega) = \mathcal{F}\{r_x(\tau)\}, \tag{16.48}$$

$$r_x(\tau) = \mathcal{F}^{-1}\{S_x(\omega)\}. \tag{16.49}$$

where $\mathcal{F}[\{x(t)\}$ is the Fourier Transform of $x(t)$.

16.3.7 Power spectral density

The power spectral density $S(f)$ of a signal $x(t)$ specifies the power or the mean square value of that signal per unit of bandwidth [Hz]. For electrical signals, $S(f)$ will have the dimension of W/Hz. The mean square value of a time-limited recording of a signal $x(t)$, of which the frequency contents is limited between f_1 and f_2, and with a recording time T over which $x(t)$ can be considered to be stationary, can be obtained both from the time-domain and the frequency-domain descriptions:

$$\overline{\{x(t)\}^2} = \frac{1}{T}\int_0^T \{x(t)\}^2 dt = \int_{f_1}^{f_2} S(f)df \tag{16.50}$$

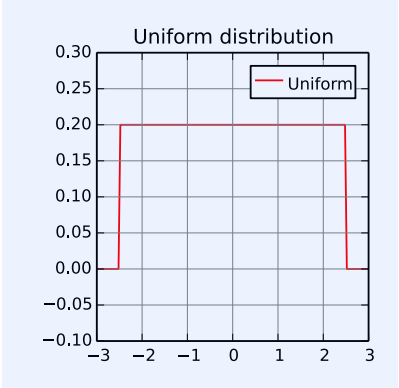


Figure 16.4: Example of a uniform distribution function.

16.4 Signals, data and information

In this section, we will summarize description methods for the amount of data and information, present in signals.

16.4.1 Amount of data

The amount of data D that is received on the retrieval on a value between a and b of a random variable $\underline{x}$ at time instant t is defined as

$$D = -\log_2\{\text{Prob}(a \leq \underline{x} \leq b, t)\} \text{ [bit]}. \tag{16.51}$$

The probability of the occurrence of a signal value between a and b can be calculated from the probability density function $P(\underline{x},t)$ of the random

variable $\underline{x}$ at time instant t:

$$\text{Prob}(a \leq \underline{x} \leq b, t) = \int_a^b P(\underline{x}, t)dx \text{ [-]}. \tag{16.52}$$

Since the probability of obtaining a signal value is unity, we have, by definition:

$$\int_{-\infty}^{\infty} P(\underline{x}, t)dx = 1 \text{ [-]}. \tag{16.53}$$

An example of a uniform probability density function is shown in Figure 16.4. A uniform probability density function is characterized by its mean value μ and its width $w = \underline{x}_{\max} - \underline{x}_{\min}$:

$$P(\underline{x}, t) = \frac{1}{w} \text{ if } |\underline{x} - \mu| < \frac{w}{2}, \tag{16.54}$$

$$P(\underline{x}, t) = 0 \text{ if } |\underline{x} - \mu| \geq \frac{w}{2}. \tag{16.55}$$

The mean value or the expectation of a random variable is defined in section 16.3.2.

Figure 16.5 shows two examples of a Gaussian probability density function. A Gaussian probability density function is characterized by its mean value μ and its standard deviation σ :

$$P(\underline{x}, t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\underline{x} - \mu)^2}{2\sigma^2}\right\}. \tag{16.56}$$

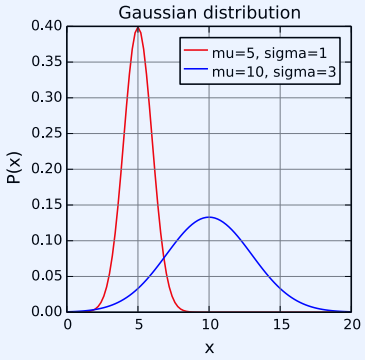


Figure 16.5: Examples of Gaussian distribution functions.

[3] H. Nyquist. Certain topics in telegraph transmission theory. *Trans. Am. Inst. Elec. Eng.*, 47:617–644, 1928

16.4.2 Bandwidth and minimum sample rate

The power spectral density of a signal, is a measure for the frequency contents of a signal. It is defined in section 16.3.7. The bandwidth B of a signal is the width of the frequency range in which signal components can be found. The minimum required sample rate S, also the Nyquist[Nyquist1928-2][3] rate, is two times the bandwidth of the signal:

$$S \geq 2B \text{ [1/s]}. \tag{16.57}$$

16.4.3 Crest factor

The crest factor C is defined as the ratio of the maximum absolute value and the RMS value of a signal:

$$C = \frac{|x|_{\max}}{x_{RMS}}. \tag{16.58}$$

16.4.4 Data rate

The data rate is the amount of data contained in a signal over a time span of one second:

$$R = 2B \log_2 \frac{x_{pp}}{n_{RMS}} \text{ [bit/s]}, \tag{16.59}$$

where n_{RMS} is the RMS value of the noise associated with the signal over the bandwidth of B [Hz].

In the following example, we will evaluate the data rate of an analog signal perturbed by noise.

Example 16.1

Consider and analog signal perturbed by noise. The signal is contained in a frequency band with a bandwidth $B = 100kHz$. The RMS value of the signal x_{pp}

in this frequency band equals $2V$. *Random noise in the same frequency band is added to the signal. The RMS value* n_{RMS} *of this noise is* $0.3mV$. *The data rate* R *of this signal is found as*

$$R = 2B\log_2\frac{x_{pp}}{n_{RMS}} = 2\times 10^5\log_2\frac{2}{0.3\times 10^{-3}} = 2.54\times 10^6 bit/s. \quad (16.60)$$

In the next example, we will evaluate the data rate of a stereo digital audio signal.

Example 16.2

The signal comprises the audio data of two channels. The update rate is $44ksps$ *and the number of bits per sample is* 16. *The data rate* R *equals:*

$$R = 2\times 16\times 44\times 10^3 = 1.41\times 10^6 bit/s. \quad (16.61)$$

16.4.5 Information rate

The information rate I [bits/s] is often much less than the data rate. Nowadays, when using high quality audio compression techniques, the data rate of a stereo audio signal of 1.41Mbit/s can be reduced to 256kbit/s without noticeable errors. The data rate for speech can be much less. In general, the data rate can be reduced if:

1. Not all data is relevant
2. Data is redundant

16.4.6 Relevant signal properties

From the above, it seems important to know which signal properties (data) should be preserved for retrieval of the information. A few examples of signals with specific signal properties are:

- FM and PM signals:
  - Information is embedded in the momentary frequency or in the momentary phase of the signal
- Analog composite video signal:
  - Intensity information is embedded in the signal level
  - Color information is embedded in the phase of a carrier with respect to reference burst during the black level.

16.4.7 Channel capacity

According to Shannon,[ShannonWeaver1963][4] the maximum amount of information that can be transported per second with an arbitrarily low number of errors over a linear channel with:

- Channel bandwidth B
- Added white Gaussian noise with power N
- Maximum signal power level S.

  is limited by the so-called channel capacity C:

$$C = B\log_2\left(1+\frac{S}{N}\right)\ [\text{bit/s}]. \quad (16.62)$$

[4] Claude E. Shannon and Warren Weaver. *The Mathematical Theory of Communication*. The University of Illinois Press, Ubana, 1 edition, 1963. ISBN: 0-88179-205-5

This expression clearly shows the three fundamental physical limitations to the amount of information that can be processed and that have been introduces above:

- Noise limitation: any physical system adds noise
- Power limitation: the power of any physical signal is limited
- Speed limitation: the rate of change of any physical signal is limited

16.4.8 Spectral efficiency

The spectral efficiency E is a measure for the information rate I per unit of bandwidth. It is defined as

$$E = \frac{I}{B} \text{ [(bit/s)/Hz]}. \tag{16.63}$$

17
System Modeling (selected topics)

17.1 Introduction

The aim of system modeling is to provide description methods for the evaluation of the performance of a system. These description methods should be as simple as possible, but adequate and sufficiently accurate for the investigation of a performance aspect of interest. Hence, different models may be used for the evaluation of different performance aspects.

Most real world systems can be considered to be time-variant, nonlinear dynamic and distributed, and such behavior can only be analyzed with complex models and computer-based numerical analysis. Because of their complexity, such models do not provide much design information. In fact, numerical analysis can only be performed on systems that have already been designed. Symbolic analysis techniques, however, can help us find design parameters that determine specific behavioral aspects of the system.

Most symbolic analysis techniques have a limited scope, and proper analysis techniques have to be selected for the investigation of specific performance aspects. Analysis of the influence of the dynamic behavior of a system requires a modeling technique other than the analysis of its nonlinear behavior. For this reason, we will start this overview of system modeling with a classification of systems and assign appropriate modeling techniques to each class of systems. This will be done in section 17.2. In the sections following this classification, we will give a short summary of the most commonly used modeling and analysis techniques. We will discuss the modeling of lumped continuous-time analog systems only.

Many idealized systems, e.g. ideal amplifiers, are considered linear, instantaneous and stationary for signals of interest.[1] Their ideal behavior will be described in section 17.3. We also need techniques for modeling their non-ideal behavior: their nonlinear behavior, their dynamic behavior and time-variant behavior. Modeling of dynamic stationary linear systems is discussed in section 17.4. Modeling of nonlinear stationary and instantaneous systems will be discussed in section 17.5, and modeling of time-variant linear instantaneous systems in section 17.6. Some notes will be made on the modeling of combinations of non-idealities, that occur in nonlinear dynamic systems. This will be done in section 17.7.

[1] Stationary is also called *static*, *fixed* or *time-invariant*.

17.2 Classification of systems

Continuous time and discrete-time systems

In a continuous time system, the inputs and the outputs are capable of changing at any time instant. Otherwise, the input and output signal values are of interest at any time instant. In discrete time systems or sampled systems, the signal values are of interest at discrete time instants only. Between these instants, the signals need not be defined.

Analog and digital systems

In an analog system, both the input and output signals are capable of having any value in a limited interval. Otherwise any signal value in a limited interval is of interest. In quantized systems, or digital systems, only a countable number of signal values is used.

Linear and nonlinear systems

In a strictly mathematical way, a system is linear, if and only if both properties of homogeneity and additivity hold. Mathematically, this can be written as

$$\mathcal{H}\{\alpha x_1(t) + \beta x_2(t)\} = \alpha\mathcal{H}\{x_1(t)\} + \beta\mathcal{H}\{x_2(t)\}, \qquad (17.1)$$

where $\mathcal{H}$ is the system operator, α and β are two scalars, and $x_1(t)$ and $x_2(t)$ are two excitations.

Instantaneous and dynamic systems

A system is called instantaneous if its responses, at any instant, only depend on the excitations at the same instant (not on past and future values). A system is called dynamic if a response at any time instant not only depends on the present input, but also on at least one of the past values of an excitation. A dynamic system is said to have a memory of length T if the output at time t is completely determined by the input values in the interval $(t - T, t)$.

Fixed and time-variant systems

A system is called fixed, time-invariant or stationary, if the system response does not depend on the time of the application of the excitation, when the system is assumed to be at rest prior to the application of the excitation. This implies that $\mathcal{H}$ is not a function of time. This can be expressed as

$$\text{if } \mathbf{Y}(t) = \mathcal{H}\{\mathbf{X}(t)\}, \text{ then } \mathbf{Y}(t-\tau) = \mathcal{H}\{\mathbf{X}(t-\tau)\}, \text{ for all } \tau. \tag{17.2}$$

Otherwise, a system is stationary if the *system properties* do not change with time.

Stable and unstable systems

A system is said to be stable if and only if any bounded input signal, results in a bounded output signal. A bounded time function $x(t)$ is one that never becomes infinite.

Causal systems

A causal system (physical or non-anticipatory system) is one whose responses to any input does not depend on any future value of the excitations.

Lumped and distributed systems

A dynamic system is said to be lumped if its behavior is governed by a set of ordinary differential (or difference) equations. Systems that require the use of partial differential equations to describe their behavior are said to be distributed. In this book, we will only deal with causal, analog, stationary, continuous time, lumped systems.

17.3 Linear stationary instantaneous systems

The response $y(t)$ of a linear stationary instantaneous system to an excitation $x(t)$ can be obtained through multiplication of $x(t)$ by a scalar:

$$y(t) = g\, x(t), \tag{17.3}$$

where the scalar g is called the gain of the system.

17.4 Linear stationary dynamic systems

Linear, dynamic, time-invariant, lumped, continuous-time systems, can be modeled with (a set of) ordinary linear differential equations with constant

coefficients. If such a system has only one input and one output, the input-output relation is given by a linear differential equation of the order n with fixed coefficients. The general form of such a differential equation will be:

$$\sum_{i=0}^{i=n} a_i \frac{d^i y(t)}{dt^i} = \sum_{k=0}^{k=m} b_k \frac{d^k x(t)}{dt^k}, \tag{17.4}$$

where $x(t)$ is the input signal or excitation, $y(t)$ the output signal or response, and a_i and b_k the time-independent (fixed) real coefficients. Due to the physical limitation of speed, we have $n > m$.

The response of a linear dynamic system to an arbitrary input signal can be found by solving the above differential equation. Direct solution of a differential equation requires an exact time domain description of the excitation $x(t)$, as well as the definition of n initial conditions. Such descriptions are not always available. However, since we are dealing with linear systems, the response of the system to an arbitrary signal, may also be obtained as a linear superposition of responses to elementary signals in which the arbitrary signal can be resolved. To this end, signals are decomposed into series of unit impulses $\delta(t)$, imaginary exponentials $\exp j\omega t$, or complex exponentials $\exp st$, where $s = \sigma + j\omega$. In this section, we will discuss modeling techniques, based on such decompositions.

17.4.1 Time domain analysis

1. Solution with the aid of the unit impulse response $h(t) = \mathcal{H}\{\delta(t)\}$, or the unit step response $a(t) = \int_{-\infty}^{t} h(\tau)d\tau$.

   Since all time functions can be resolved into unit impulse or unit step functions, the response $y(t)$ of a linear system to an arbitrary signal $x(t)$ can be found as the sum of the responses to the unit impulse, or the unit step functions into which the signal is resolved. The operation for finding the time-domain response with the aid of resolution in unit impulses is called convolution:

$$y(t) = x(t) * h(t) = \int_0^t x(\tau)h(t-\tau)d\tau, \tag{17.5}$$

$$y(t) = \int_0^t \dot{x}(\tau)a(t-\tau)d\tau = \dot{x}(t) * a(t). \tag{17.6}$$

2. Analysis of the behavior using signals resolved in elementary exponentials.

   Since exponential functions retain their shape under the operation of differentiation and integration, differential equations can be transformed into algebraic equations if the excitation and response are written as exponential functions. The algebraic equation obtained in this way, can be solved analytically. For this reason, we often resolve signals in imaginary or complex exponentials, using the Fourier and Laplace transformation techniques, respectively. These techniques will be discussed below.

17.4.2 Frequency domain analysis

The resolution of an arbitrary time signal into imaginary exponentials is convenient both from the mathematical and practical point of view. Resolution of signals into elementary imaginary exponentials, allows us to use frequency domain descriptions for both information-carrying signals and test signals. Linear amplifiers can be characterized with the aid of sinusoidal test signals, and deviations from linear behavior can be observed as signal distortions at the amplifier's load.

We will briefly demonstrate this technique for the differential equation given in (17.4). To this end, we resolve $x(t)$ and $y(t)$ into elementary imaginary exponentials. One element of the resolved excitation $x(t)$ is then of the form $X(j\omega)\exp j\omega t$ and the corresponding response is of the same shape and can be written as $Y(j\omega)\exp j\omega t$, in which $X(j\omega)$ and $Y(j\omega)$ are the complex amplitudes of the imaginary exponentials $\exp j\omega t$ at the input and the output, respectively. The $k-th$ derivative of an element of the resolved input signal is obtained as $(j\omega)^k X(j\omega)\exp j\omega t$ and the $i-th$ derivative of its corresponding response as $(j\omega)^i Y(j\omega)\exp j\omega t$. All initial conditions are assumed to be zero. Substituting these exponentials into the differential equation yields

$$\sum_{i=0}^{i=n} a_i (j\omega)^i Y(j\omega)\exp j\omega t = \sum_{k=0}^{k=m} b_k (j\omega)^k X(j\omega)\exp j\omega t. \tag{17.7}$$

We now obtain the *transfer function* $H(j\omega)$ of the amplifier, which relates, for a given frequency ω, the complex amplitude of the response $Y(j\omega)$ to that of the excitation $X(j\omega)$:

$$H(j\omega) = \frac{Y(j\omega)}{X(j\omega)} = \frac{\sum_{k=0}^{k=m} b_k (j\omega)^k}{\sum_{i=0}^{i=n} a_i (j\omega)^i}. \tag{17.8}$$

Transfer function

The transfer function $H(j\omega)$ of a linear fixed dynamic system is the Fourier transform of the system's unit impulse response:

$$H(j\omega) = \mathcal{F}\{h(t)\}. \tag{17.9}$$

Information-carrying signals are often characterized by their power spectral density $S(\omega)$, which is a statistical quantity defined as the average signal power in [W/Hz] at a certain frequency, over a bandwidth of one Hz. If the input signal $x(t)$ of a system with a transfer function $H(j\omega)$, has a power spectral density of $S_x(\omega)$, the power spectral density $S_y(\omega)$ of the response signal $y(t)$ can be obtained as:

$$S_y(\omega) = S_x(\omega)\,|H(j\omega)|^2\,. \tag{17.10}$$

Bode plots

A graphical representation of the transfer function can be given by plotting the magnitude $|H(j\omega)|$ and the argument $\arg\{H(j\omega)\}$ of the transfer function $H(j\omega)$ as a function of ω. These so-called Bode plots show the magnitude $20\log_{10}|H(j\omega)|$ in dB and the argument $\arg\{H(j\omega)\}$ in degrees or radians, both on a linear scale. The angular frequency ω or the frequency f is usually plotted on a logarithmic scale.

17.4.3 Complex frequency domain analysis

The system function $H(s)$ is obtained in a similar way as the transfer function $H(j\omega)$; it relates the complex amplitudes $Y(s)$ of the complex exponentials of the resolved output signal to those of the resolved input signal. $H(s)$ is the Laplace transform of the amplifier's unit impulse response:

$$H(s) = \frac{Y(s)}{X(s)} = \frac{\sum_{k=0}^{k=m} b_k s^k}{\sum_{i=0}^{i=n} a_i s^i} = \mathcal{L}\{h(t)\}. \tag{17.11}$$

The Laplace transform is also defined for signals that have no power limitation. These (theoretical) signals arise in linear(ized) unstable dynamic sys-

tems. We will show that the stability of linear(ized) dynamic systems can be investigated through evaluation of the system function.

The numerator of the system function (see expression 17.11) is a polynomial of degree m. The m roots of the numerator are called the zeros of the system functions. The n roots of the denominator are called the poles of the system function. Any system function of a system that does not incorporate delay lines can be written in terms of these poles and zeros:

$$H(s) = \frac{b_m}{a_n} \frac{\prod_{k=1}^{m} (s - z_k)}{\prod_{i=1}^{n} (s - p_i)}, \tag{17.12}$$

where z_k is a zero of $H(s)$ and p_i is a pole of $H(s)$.

If there are no poles or zeros at $s = 0$, this may also be written as

$$H(s) = \frac{b_0}{a_0} \frac{\prod_{k=1}^{m} (1 - \frac{s}{z_k})}{\prod_{i=1}^{n} (1 - \frac{s}{p_i})}. \tag{17.13}$$

The system function is completely described by means of its poles and zeros and the coefficients a_0 and b_0 or a_n and b_m. The factor $\frac{b_0}{a_0}$ is the zero frequency transfer or the DC transfer of the system.

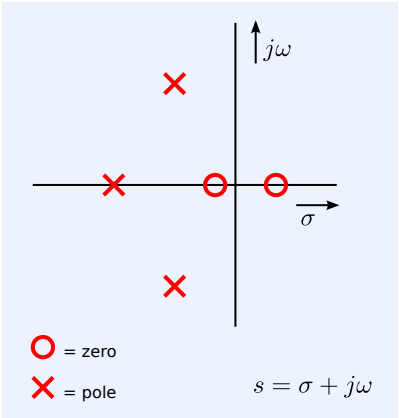


Figure 17.1: Example of a pole-zero pattern.

Pole-zero pattern

The graphical representation of the poles and zeros in the complex plane is referred to as the pole-zero pattern of $H(s)$. A typical pole-zero pattern of a system function is depicted in Figure 17.1.

Poles and zeros always appear as single real or pairs of complex conjugates; this is because the coefficients of the differential equations are real.

Not all system functions can be represented by a pole-zero pattern. A system with a time delay τ has a system function like: $H(s) = \exp(-s\tau)$, which cannot be expressed in terms of poles and zeros.

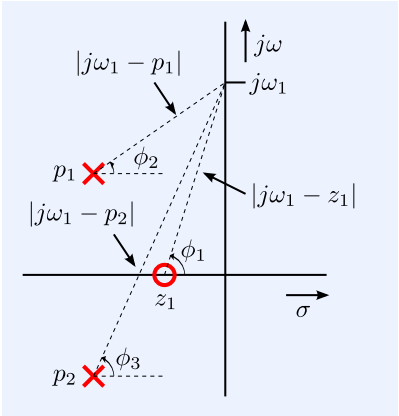


Figure 17.2: Relation between pole-zero pattern and Bode plots.

Pole-zero pattern and Bode plots

The relation between the Bode plots and the pole-.zero pattern can be found by substituting $s = j\omega$ in the expression for the system function $H(s)$. The magnitude of the transfer function $|H(j\omega)|$ can then be evaluated from

$$|H(j\omega)| = \frac{b_m}{a_n} \frac{\prod_{k=0}^{k=m} |j\omega - z_k|}{\prod_{i=0}^{i=n} |j\omega - p_i|}. \tag{17.14}$$

The magnitude of the transfer function is thus proportional to the quotient of the product of the magnitudes of the terms of the numerator and the product of the magnitudes of the denominator. The magnitude of a single factor $|j\omega - z_k|$ at an angular frequency ω_1 equals the distance between the location of the zero z_k and the position $j\omega = j\omega_1$ on the imaginary axis. The magnitude $|j\omega - p_i|$ can be found in a similar way. This is shown in Figure 17.2.

The argument of the transfer function can be evaluated as:

$$\arg\{H(j\omega)\} = \arg b_m - \arg a_n + \sum_{k=0}^{k=m} \arg(j\omega - z_k) - \sum_{i=0}^{i=n} \arg(j\omega - p_i). \tag{17.15}$$

The arguments of a_n and b_m are either 0 or π rad, depending on their sign. The argument of a single term $(j\omega - z_k)$ at an angular frequency ω_1 is found as the angle formed by the vector that connects the zero z_k with the point $j\omega = j\omega_1$ on the imaginary axis and the positive real axis. This relation is shown in Figure 17.2.

Minimum phase system

A system is called a minimum phase system if the inverse of $H(s)$ is stable and causal. This implies that $H(s)$ has no zeros in the right half plane. The magnitude and the phase characteristic of a minimum phase system are related by the *Hilbert Transform*:

$$\arg\{H(j\omega)\} = -\mathcal{H}_{\mathcal{I}}\{\ln|H(j\omega)|\}, \tag{17.16}$$

in which the Hilbert transform operator $\mathcal{H}_{\mathcal{I}}$ is defined as:

$$\mathcal{H}_{\mathcal{I}}\{x(t)\} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}d\tau. \tag{17.17}$$

In the frequency domain, this is equivalent to

$$\mathcal{H}_{\mathcal{I}}\{X(j\omega)\} = -j\,\mathrm{sgn}\,\omega\; X(j\omega). \tag{17.18}$$

A minimum phase system has the smallest possible delay of all linear time-invariant dynamic systems that have equal magnitude characteristics.

17.4.4 Time domain analysis using the Laplace transform

Time domain analysis, using signals resolved in elementary exponentials, is performed with the aid of the Laplace transform. The procedure for the determination of the time domain response to a known time function can be presented as

$$y(t) = \mathcal{L}^{-1}\{Y(s)\} = \mathcal{L}^{-1}\{H(s)X(s)\} = \mathcal{L}^{-1}[H(s)\mathcal{L}\{x(t)\}]. \tag{17.19}$$

For determination of the Laplace transform of a time function, we use tables of the Laplace pairs and the properties of the Laplace transform.

Stability

The system's unit impulse response $h(t)$ can be found from the inverse Laplace transform of its system function $H(s)$. The unit impulse response of a system with more poles than zeros can be written as a sum of exponentials with $p_k t$ as its argument, in which p_k is a pole of the system function. A general expression for the impulse response of a system with more poles than zeros is:

$$h(t) = \sum_{i=1}^{n}\sum_{j=0}^{\ell-1} A_{i,j}t^j \exp p_i t \tag{17.20}$$

in which ℓ is the number of occurrences for the pole p_i. The coefficients $A_{i,j}$ depend on the poles and the zeros. Expression 17.20 clearly shows that the impulse response is bounded if all poles of the system function have a negative real part. As a consequence, if all poles have a negative real part, the response to any bounded signal is bounded and the system is *stable*.

Definition: A system is said to be stable, if the poles of its system function $H(s)$ are all located in the left half of the complex s-plane.

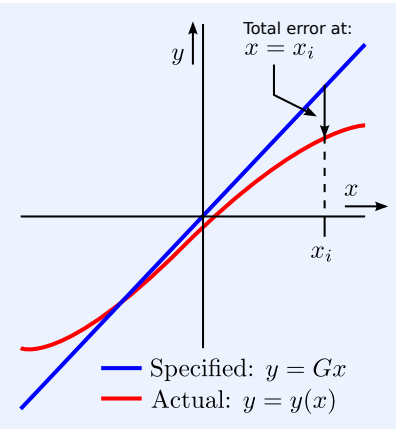


Figure 17.3: Blue: ideal static instantaneous linear input-output relation
Red: realized static instantaneous nonlinear input output relation.

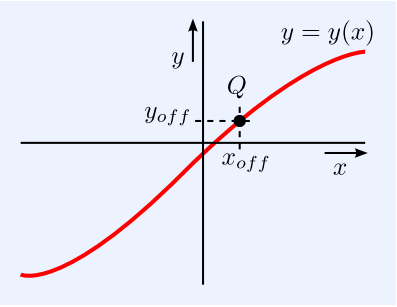


Figure 17.4: Definition of the desired operating point for linearization.

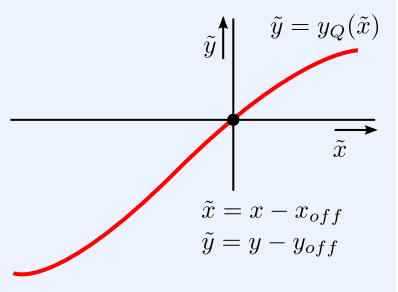


Figure 17.5: Modified source-load relation.

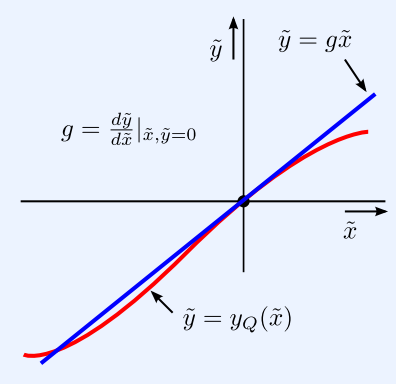


Figure 17.6: Definition of small-signal gain and inaccuracy.

17.5 Fixed instantaneous nonlinear systems

In this section, we will introduce the basic concepts for characterization of nonlinear time-invariant instantaneous systems. Information processing systems are often intended to behave linearly. Deviations from this linear behavior will be modeled in this section. There are many different description methods for nonlinear behavior. Harmonic distortion , intermodulation , differential-gain , and gain compression are terms that are often used to express the perception of nonlinear effects. In general, description methods that show the best correspondence to the observer's error perception must be used. This perception strongly depends on the way in which the information is embedded in the signal and how this information is interpreted by the observer.

A single-input single-output nonlinear instantaneous fixed system can be characterized by a curve in its so-called input-output plane. Figure 17.3 shows an example of a static nonlinear instantaneous input-output relation, with x and y representing the input and the output signal values, respectively. One could measure such a curve through application of an input signal that varies with time and plot the output signal values as a function of the input signal values. Figure 17.3 also shows the intended ideal input-output relation of this system.

The *total error* shown in the figure is the difference between the intended system output and the actual output. It depends on the value of the input signal.

This total error can be decomposed in a number of contributions. These contributions are defined in the following sections.

17.5.1 Operating point, input and output offset

The static nonlinear curve from Figure 17.3 can mathematically be described as

$$y = y(x). \tag{17.21}$$

In order to specify the deviations from the ideal behavior, we have to select our intended origin, also called the *quiescent operating point* Q. This is shown in Figure 17.4.

The curve is shifted to this new origin through application of the so-called input and output offset quantities x_{off} and y_{off}, respectively. For a voltage amplifier, x_{off} and y_{off} are the voltages of a voltage source in series with the signal source and a voltage source in series with the load, respectively. The quescent operating point Q can be selected on various grounds. At this stage, we will not discuss this in detail, but just introduce the concept of an operating point.

After application of x_{off} and y_{off}, the operating point Q is the origin of the modified source-to-load relation that is written as:

$$\tilde{y} = y_Q(\tilde{x}) \tag{17.22}$$

In which $\tilde{x}$ and $\tilde{y}$ are the deviations of the input and output quantities from the operating point Q. This is illustrated in Figure 17.5.

The new function $\tilde{y} = y_Q(\tilde{x})$ is free of offset. It can be approximated by a Taylor series expansion as

$$\tilde{y} = \sum_{n=1}^{n=\infty} g_n^n \tilde{x}^n, \tag{17.23}$$

with

$$g_n = \frac{1}{n!} \frac{d^n\{y_Q(\tilde{x})\}}{d\tilde{x}^n}\bigg|_{\tilde{x}=0}. \tag{17.24}$$

17.5.2 Small-signal gain and inaccuracy

For small signals, the system can be linearized by approximating the input-output relation by a straight line through the (new) origin. This is shown in Figure 17.6. The tangent of the angle between this line and the x-axis is the small-signal gain of the system. It is found as g_1 from expression 17.24:

$$\tilde{y} = g_1 \tilde{x}, \text{ with } g_1 = \left. \frac{d\{y_Q(\tilde{x})\}}{d\tilde{x}} \right|_{\tilde{x}=0}. \tag{17.25}$$

The inaccuracy δ of the linearized system is defined as the difference between the actual gain g_1 and the required gain of the system, while the relative inaccuracy δ_{rel} equals the quotient of the inaccuracy and the desired gain.

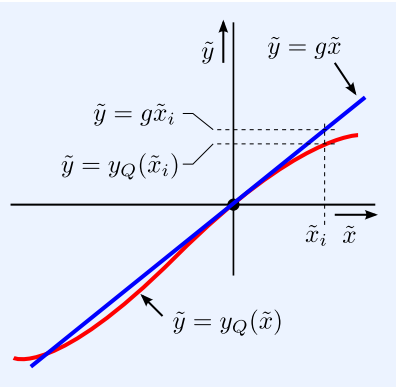


Figure 17.7: Definition of the nonlinearity

17.5.3 Nonlinearity

For large signal excursions, the approximation by a straight line is not accurate. The nonlinearity Δ is defined as the difference between the actual output signal and the output signal that would be obtained from the linearized system. Figure 17.7 shows Δ as a function of the relative input excursion:

$$\Delta = y_Q(\tilde{x}) - g_1 \tilde{x}. \tag{17.26}$$

The relative nonlinearity Δ_{rel} equals the ratio of the nonlinearity and the ideal output value gx_1 :

$$\Delta_{rel} = \frac{y_Q(\tilde{x}) - g_1 \tilde{x}}{g_1 \tilde{x}}. \tag{17.27}$$

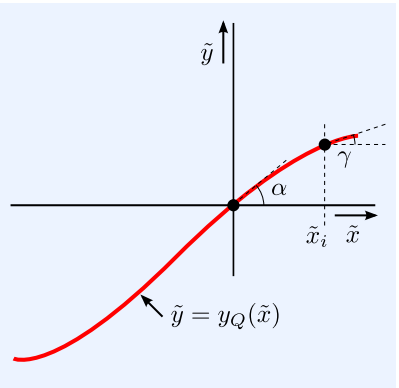


Figure 17.8: Definition of the differential gain.

17.5.4 Differential gain

Due to nonlinearity, the gain changes with the operating point. At an arbitrary operating point, for example, at $\tilde{x} = x_1$, the differential gain error ϵ is defined as the difference between the gain at $\tilde{x} = x_1$ and the gain at $\tilde{x} = 0$, divided by the gain at $\tilde{x} = 0$:

$$\epsilon(\tilde{x}) = \frac{\left. \frac{d\{y_Q(\tilde{x})\}}{d\tilde{x}} \right|_{\tilde{x}=x_1} - \left. \frac{d\{y_Q(\tilde{x})\}}{d\tilde{x}} \right|_{\tilde{x}=0}}{\left. \frac{d\{y_Q(\tilde{x})\}}{d\tilde{x}} \right|_{\tilde{x}=0}}. \tag{17.28}$$

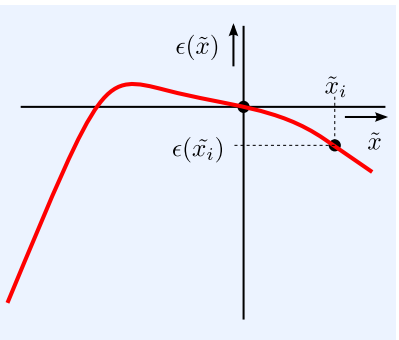


Figure 17.9: Differential gain as a function of the source signal deviation from the operating point.

17.5.5 Harmonic distortion

We have seen that in linear fixed dynamic systems, sinusoidal signals retain their shape. The nonlinearity of a system is therefore often characterized by the amount of distortion of the response to a sinusoidal input signal. The response of a fixed nonlinear instantaneous system to a sinusoidal input signal is a periodic signal whose period equals the period of the sine wave. The nonlinearity of the system generally introduces a change in the output offset and in the harmonic contents. The total harmonic distortion THD is defined as the relative *RMS*-value of all the harmonics in the output signal that have a frequency larger than the fundamental frequency:

$$THD = \frac{1}{d_1} \sqrt{\sum_{n=2}^{n=\infty} d_n^2}, \tag{17.29}$$

where d_n equals the amplitude of the n-th harmonic in the output signal.

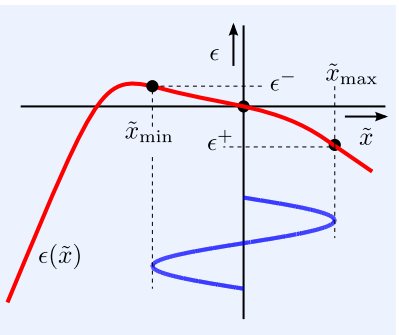


Figure 17.10: Determination of ϵ^+ and ϵ^- for harmonic and intermodulation distortion.

For weak nonlinear, instantaneous systems, the THD is related to the differential gain. To illustrate this, we consider a nonlinear system with an input-output relation as given by expression 17.23, and apply an input signal $\tilde{x}(t)$ given by

$$\tilde{x}(t) = X\cos\omega t. \tag{17.30}$$

The output signal $\tilde{y}(t)$ is then obtained as

$$\tilde{y}(t) = g_1 X\cos\omega t + g_2 X^2\cos^2\omega t + g_3 X^3\cos^3\omega t + ... , \tag{17.31}$$

which can be approximated by

$$\tilde{y}(t) = g_1 X\cos\omega t + \frac{g_2}{2}X^2\cos 2\omega t + \frac{g_3}{4}X^3\cos 3\omega t + ... \tag{17.32}$$

The relative amplitude d_2 of the second harmonic is thus found as $g_2X/2g_1$. The relative amplitude of the third harmonic d_3 is found as $g_3X^2/4g_1$.

In order to find the relation between the THD and the differential gain , we will write the differential gain as a function of $g_1.....g_n$. Since the differential gain depends on the large signal excursion, we use approximations for the differential gain at maximum and minimum signal excursion. The differential gain at the largest positive signal excursion ($\tilde{x}(t) = X$) will be denoted as ϵ^+ and is obtained from the definition as

$$\epsilon^+ = \frac{2g_2X + 3g_3X^2}{g_1}. \tag{17.33}$$

In a similar way, the differential gain at the largest negative signal excursion ($x(t) = -X$), denoted by ϵ^-, is found to be

$$\epsilon^- = \frac{-2g_2X + 3g_3X^2}{g_1}. \tag{17.34}$$

With the aid of 17.33 and 17.34, we are able to calculate the second and order third harmonic distortion from the differential gain values. We obtain

$$d_2 = \frac{\epsilon^+ - \epsilon^-}{8}, \text{ and } d_3 = \frac{\epsilon^+ + \epsilon^-}{24}, \tag{17.35}$$

in which ϵ^+ and ϵ^- are the differential gain errors at $\tilde{x}_{\max}$ and $\tilde{x}_{\min}$, respectively. This is shown in Figure 17.10.

17.5.6 Intermodulation distortion

Let us consider a situation in which the excitation of a fixed nonlinear instantaneous system consists of two sinusoidal components with different frequencies ω_1 and ω_2, but equal amplitudes. Due to the nonlinearity of the system the output signal will have components at multiples of ω_1 and ω_2 (known as harmonic distortion) and at frequencies $m\omega_1 \pm n\omega_2$ (m and n are integers). The latter effect is called intermodulation distortion. The amplitudes of the components at these frequencies will be denoted as $A_{m\omega_1 \pm n\omega_2}$. The second order intermodulation distortion IM_2 is defined as the *relative* RMS-value of the component of the output signal with $m = 1$ and $n = 1$, when the components at ω_1 and ω_2 have equal amplitudes:

$$IM_2 = \frac{|A_{\omega_1 \pm \omega_2}|}{|A_{\omega_{1,2}}|}. \tag{17.36}$$

The third order intermodulation distortion IM_3 is defined as the relative RMS-value of the component with $m = 2$ and $n = 1$ or $m = 1$ and $n = 2$,

when both the components at ω_1 and ω_2 have equal amplitudes:

$$IM_3 = \frac{|A_{\omega_1 \pm 2\omega_2}|}{|A_{\omega_{1,2}}|} = \frac{|A_{2\omega_1 \pm \omega_2}|}{|A_{\omega_{1,2}}|}. \tag{17.37}$$

For weak nonlinear, instantaneous systems, the second and third order intermodulation distortion are related to the differential gain. In order to find this relation, we apply

$$x(t) = X\{\cos\omega_1 t \ + \cos\omega_2 t\} \tag{17.38}$$

to the input of the system. The second order term from the Taylor approximation of the output signal is then obtained as

$$\begin{aligned} Y_2 = g_2 X^2\{\tfrac{1}{2} + \tfrac{1}{2}\cos(2\omega_1 t\ ) \ + \cos(\omega_1 t \ + \omega_2 t\ ) + \\ \cos(\omega_1 t \ - \omega_2 t\ ) \ + \tfrac{1}{2} + \tfrac{1}{2}\cos(2\omega_2 t\ )\}, \end{aligned} \tag{17.39}$$

which yields

$$IM_2 = \frac{g_2}{g_1} X = \frac{\epsilon^+ - \epsilon^-}{4}. \tag{17.40}$$

In a similar way, we find

$$IM_3 = \frac{3g_3}{4g_1} X^2 = \frac{\epsilon^+ + \epsilon^-}{8}, \tag{17.41}$$

where ϵ^+ and ϵ^- are the differential gain errors at $\tilde{x}_{\max}$ and $\tilde{x}_{\min}$, respectively. This is shown in Figure 17.10.

17.6 Linear time-variant instantaneous systems

The behavior of a linear time-variant instantaneous system can be modeled with the aid of a time variant gain $g(t)$:

$$y(t) = g(t)x(t). \tag{17.42}$$

Sometimes it can be convenient to model a system differently for different components of the input signal. Suppose a system shows strong instantaneous nonlinear behavior (for example a limiter). The system's response for the large sinusoidal signal can then be calculated from its nonlinear instantaneous behavior. The large sinusoidal signal causes periodic small-signal gain variations. The system response to a small signal or to noise may then be analyzed using a linear time-variant approach: the small signal and the noise are then multiplied by a gain function that is a function of the shape of the large signal and the nonlinearity of the system.

17.7 Modeling of nonlinear dynamic systems

Dynamic nonlinear systems can be modeled with the aid of nonlinear differential equations. The solution of these equations can almost always only be found through numerical methods. In this section, we will give some parameters that describe dynamic nonlinear effects. Aside from these parameters, the following parameters from the previous paragraph also apply to dynamic nonlinear systems:

1. Differential gain

   In dynamic nonlinear systems, the differential gain depends on frequency. Since the small-signal gain in dynamic systems is complex, the differential

gain is also complex with both magnitude and phase (differential gain and differential phase).

2. THD

   In dynamic nonlinear systems, the harmonic distortion becomes a function of the frequency and the relations between the differential gain and the second and third order harmonic distortions are no longer valid.

3. Intermodulation

   In dynamic nonlinear systems, the intermodulation distortion becomes a function of the frequency and the relations between the differential gain and the second and third order intermodulation distortions are no longer valid.

4. Gain compression

   In dynamic nonlinear systems the 1dB compression point becomes a function of the frequency and the relation between the IP3 and the 1dB compression point is no longer valid.

5. Slew rate

   The limitation of the maximum rate of change of the output signal is called the slew rate limitation. Let $y(t)$ be the system response to an input signal. The positive slew rate SR^+ is defined as

$$SR^+ = \left.\frac{dy(t)}{dt}\right|_{\max} ; \text{where } \frac{dy(t)}{dt} \geq 0. \tag{17.43}$$

   The negative slew rate SR^- is defined as

$$SR^- = -\left|\left.\frac{dy(t)}{dt}\right|_{\max}\right| ; \text{ where } \frac{dy(t)}{dt} < 0. \tag{17.44}$$

18
Network Theory (selected topics)

18.1 Introduction

Designers of analog electronic circuits often need to investigate or modify the small-signal static or dynamic behavior of circuits. Numerical circuit analysis is often performed with the aid of CAD programs. Although the obtained numeric results can be accurate, they do not provide design information. This is because the underlying mechanisms are not shown. Symbolic analysis of simplified circuits often does give a lot more design information, but it also often requires unpopular and cumbersome hand calculations.

Nowadays, many symbolic analysis tools are available to help designers with symbolic calculations. Adequate use of these programs requires proper formulation of the problem, which is often in the form of a matrix equation. One software tool for symbolic and numeric analysis of linear circuits is SLiCAP. It provides symbolic expressions for the circuit's signal transfer. Aside from its symbolic analysis capabilities, it can perform parametric numeric analysis and generate frequency domain, time domain and complex frequency domain plots.[1]

[1] Parametric pole-zero plots.

The theory behind the operation of SLiCAP is summarized in this chapter. It starts with a summary of Nodal Analysis in section 18.2. Nodal Analysis (NA) can be used to solve network equations for networks with only voltage-controlled elements. Networks that also include elements whose behavior can only be described in a current-controlled form, such as voltage sources, can be transformed into networks with only voltage-controlled elements, by using the Norton transformation and Blakesley's voltage shift theorem. Alternatively, the so-called Modified Nodal Analysis (MNA) can be used. MNA is implemented in SLiCAP and in most SPICE-like simulators. It will be discussed in section 18.3. An overview of MNA matrix stamps for commonly used network elements, as well as for Laplace transfer functions, will be given in section 18.3.3.

For more information on these topics, the reader is referred to Desoer Chua and Kuh [Chua1987][2].

[2] Charles A. Desour Leo O. Chua and Ernest S. Kuh. *Linear and Nonlinear Circuits*. McGraw-Hill, Inc., Singapore, 1987. ISBN: 0-07-010898-6

Circuit designers often need to determine and/or manipulate the complex frequencies of poles and zeros. Techniques for numeric and symbolic determination of pole and zero frequencies will be presented in section 18.5. Symbolic approximation of pole and zero frequencies will be added to future versions of SLiCAP.

In many engineering cases, it is convenient to model an electrical network as a two-port. In those cases, the port quantities (voltage and current) of one port are related to those of the other port and the electrical behavior of the two-port is described with the aid of a 2×2 matrix. The conditions under which four-terminal networks can be modeled as two-ports, as well as two-port representation methods and properties, will be discussed in section 18.6. Understanding of these conditions is important for correct application of the asymptotic gain feedback model discussed in Chapter 10.

18.2 Nodal Analysis

An electric network consists of interconnected network elements. The connections are called nodes and the connecting elements between nodes form the branches of a network. The graph of a network shows the branches as lines and the nodes as dots. Below are a few definitions that we will use throughout this text:

1. A connected graph is a graph that has at least one path among the branches that connects all the nodes.

2. A sub-graph is a subset of branches with their corresponding nodes.

3. A closed path of branches is called a loop.

4. A collection of branches that isolates a sub-graph when removed is called a cut set.

5. A tree is a collection of branches that connects all the nodes but has no loops.

Figure 18.1 illustrates the definitions of loops, cut sets and of the tree.

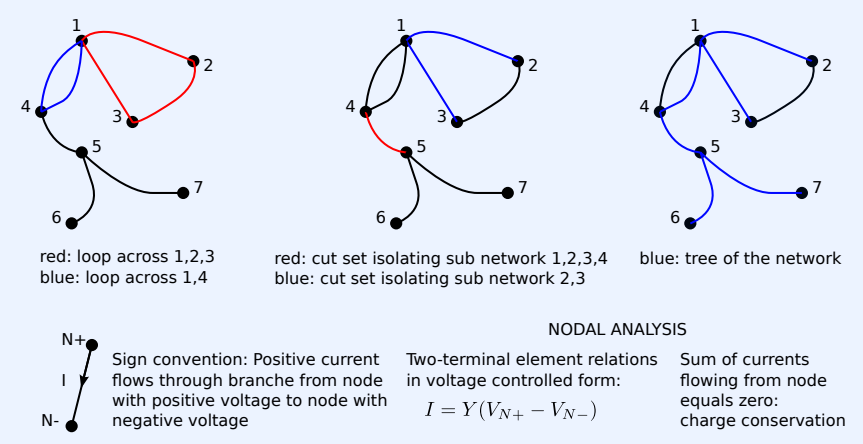


Figure 18.1: Graph of a network, showing loops, cut sets and a tree.

Nodal analysis provides the network's nodal voltages from the independent currents flowing into the nodes. It is based upon the application of Kirchhoff's current law (see Figure 18.2), which states that the sum of the electric currents that flow into a node equals zero.

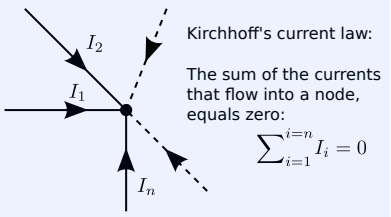


Figure 18.2: Kirchhoff's current law.

Networks with only voltage-controlled elements ($I = f(V)$) can directly be solved with the nodal analysis method. A network element is said to be voltage-controlled if its branch currents are uniquely defined by its branch voltages. This is not the case for current-controlled elements such as voltage sources. These elements have their voltage unambiguously defined by their current ($V = f(I)$). As we will see later, networks with both voltage-controlled and current-controlled elements can be solved using modified nodal analysis (MNA).

In the next section, we will demonstrate the nodal analysis method for a simple network comprising voltage-controlled elements only.

18.2.1 The procedure

The procedure for nodal analysis is as follows:

1. Set up the circuit diagram, select a reference node and number all remaining nodes.

2. Set up the nodal equations for all nodes except the reference node.

   This results in the following matrix equation:

$$\mathbf{I} = \mathbf{Y} \cdot \mathbf{V}, \tag{18.1}$$

   where $\mathbf{I}$ is the vector of independent currents that flow into a node, $\mathbf{Y}$ is the admittance matrix that depends both on the graph and the element relations of the network elements, and $\mathbf{V}$ is the vector of the nodal voltages with respect to the voltage at the reference node. A network having n nodes requires $n-1$ nodal equations. The voltage of the reference node is usually assigned zero (ground potential).

3. Find the network solution (all nodal voltages and branch currents).

   The nodal voltages are found from

$$\mathbf{V} = \mathbf{Y}^{-1} \cdot \mathbf{I}. \tag{18.2}$$

   The branch currents $I(j,k)$ $(j \neq k)$ are found from

$$I(j,k) = \left(V_j - V_k\right) Y_{j,k}. \tag{18.3}$$

   The transimpedance from a current I_k flowing *into* node k to a voltage V_j at node j is found as:

$$\frac{\mathbf{V}_j}{\mathbf{I}_k} = \mathbf{Y}^{-1}_{j,k}, \tag{18.4}$$

   in which $\mathbf{Y}^{-1}_{j,k}$ is the coefficient j,k of the inverse of the admittance matrix $\mathbf{Y}^{-1}$. It can be obtained as

$$\left(\mathbf{Y}^{-1}\right)_{j,k} = \frac{\mathcal{C}_{k,j}}{\det(\mathbf{Y})}, \tag{18.5}$$

   in which the *cofactor* $\mathcal{C}_{k,j}$ is a coefficient of the cofactor matrix $\mathcal{C}$ of $\mathbf{Y}$. It is defined as the determinant of the minor matrix $\mathcal{Y}_{j,k}$, multiplied by $(-1)^{j+k}$. The minor matrix $\mathcal{Y}_{j,k}$ is the matrix $\mathbf{Y}$ with the $j-th$ row and the $k-th$ column left out. Hence, $\mathcal{C}_{k,j}$ is defined as

$$\mathcal{C}_{kj} = (-1)^{j+k} \det(\mathcal{Y}_{j,k}). \tag{18.6}$$

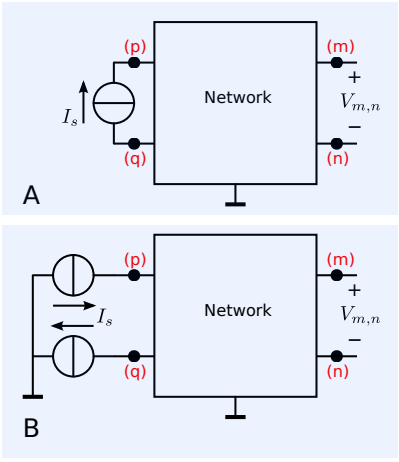


Figure 18.3: Current to voltage transfer in a network.

4. If an independent current source is supplying a current I_s into node p and drawing a current from node q, the transfer Z_t from that current source to the voltage between node m and node n is found as

$$Z_t = \frac{V_m - V_n}{I_s} = \frac{\mathcal{C}_{p,m} - \mathcal{C}_{q,m} - \mathcal{C}_{p,n} + \mathcal{C}_{q,n}}{\det(\mathbf{Y})}. \tag{18.7}$$

   This is illustrated in Figure 18.3. Figure 18.3A shows the circuit with I_s flowing from node p to node q. This branch current can be redirected via the ground node as shown on Figure 18.3B. We now have contributions to the differential voltage $V_m - V_n$. Each transfer is modeled by one term in (18.7):

   (a) The transfer from I_s flowing into node p, to the nodal voltage V_m
   (b) The transfer from I_s flowing into node p, to the nodal voltage V_n
   (c) The transfer from I_s flowing from node q, to the nodal voltage V_m
   (d) The transfer from I_s flowing from node q, to the nodal voltage V_n.

5. The poles of the system described by $\mathbf{Y}$, are obtained by solving the characteristic equation:

$$\text{poles} : \det \mathbf{Y}(s) = 0. \tag{18.8}$$

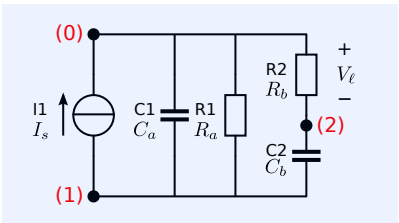


Figure 18.4: Circuit for demonstration of the nodal analysis method.

Example 18.1

In this example, we will derive expressions for the transfer of the current source `I1` *to the voltage* V_ℓ *of the circuit drawn in Figure 18.4. Please notice that the device identifiers (also called* `REFDES`*: reference designators) have been typeset with a fixed width font and the device values with a math font: (*`R1`*,* R_a*).*

The node at the top of the current source `I1` *has been selected as the reference node. The value of* `I1` *is* I_s *and the load voltage* V_ℓ *is the voltage between node (0) and node (2), which equals* $-V_2$.

The nodal equations for nodes (1) and (2) are:

$$\textit{node 1:} \qquad 0 = I_s + V_1 sC_a + V_1 \frac{1}{R_a} + (V_1 - V_2)\, sC_b; \tag{18.9}$$

$$\textit{node 2:} \qquad 0 = V_2 \frac{1}{R_b} + (V_2 - V_1)\, sC_b. \tag{18.10}$$

These equations can be written using matrix notation as:

$$\begin{pmatrix} -I_s \\ 0 \end{pmatrix} = \begin{pmatrix} s\,(C_a + C_b) + \frac{1}{R_a} & -sC_b \\ -sC_b & sC_b + \frac{1}{R_b} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}, \tag{18.11}$$

or, alternatively

$$\begin{pmatrix} -I_s & 0 \end{pmatrix}^T = \mathbf{Y} \cdot \begin{pmatrix} V_1 & V_2 \end{pmatrix}^T, \tag{18.12}$$

where

$$\mathbf{Y} = \begin{pmatrix} s\,(C_a + C_b) + \frac{1}{R_a} & -sC_b \\ -sC_b & sC_b + \frac{1}{R_b} \end{pmatrix}. \tag{18.13}$$

The current-to-voltage transfer can be obtained with the aid of (18.7) using $p = 0$, $q = 1$, $m = 0$ and $n = 2$, which yields[3]

$$Z_t = \frac{\mathcal{C}_{1,2}}{\det(\mathbf{Y})}, \tag{18.14}$$

[3] Row 0 and column 0 do not exist, because node (0) is selected as the reference node.

where

$$\mathcal{C}_{1,2} = (-1)^3 \det\left(-sC_b\right), \tag{18.15}$$

$$\mathcal{C}_{1,2} = sC_b, \tag{18.16}$$

and

$$\det(\mathbf{Y}) = \det \begin{pmatrix} s\,(C_a + C_b) + \frac{1}{R_a} & -sC_b \\ -sC_b & sC_b + \frac{1}{R_b} \end{pmatrix}, \tag{18.17}$$

$$= \frac{1 + s\left(\left(C_a + C_b\right) R_a + C_b R_b\right) + s^2 C_a C_b R_a R_b}{R_a R_b}. \tag{18.18}$$

After substitution of (18.16) and (18.18) in (18.14), we obtain

$$Z_t = \frac{sC_b R_a R_b}{1 + s\left(\left(C_a + C_b\right) R_a + C_b R_b\right) + s^2 C_a C_b R_a R_b}. \tag{18.19}$$

The transfer Z_t has one zero at $s = 0$ (no DC transfer) and two poles that are the solutions for s of (18.18).

In the following example, we will demonstrate the use of SLiCAP for setting-up the matrix equations for the circuit from Figure 18.4 and for evaluation of the transfer.

Example 18.2

The SLiCAP *netlist for the circuit from Figure 18.4 is listed below:*

```
"NA-2"
* file: NA-2.cir
* SLiCAP netlist for nodal analysis
I1 1 0 {I_s}
C1 0 1 {C_a}
R1 0 1 {R_a}
R2 0 2 {R_b}
C2 2 1 {C_b}
.param R_a=100k R_b=1k C_a=100p C_b=10n I_s=1
.end
```

The SLiCAP *script that displays the matrix equation on a html page is:*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: NA-2.py

from SLiCAP import *
prj = initProject('NA-2')
instr = instruction()
instr.setCircuit('NA-2.cir')
instr.setSimType('symbolic')
instr.setGainType('vi')
instr.setDataType('matrix')
result = instr.execute()
htmlPage("Example Nodal Analysis")
head2html('MNA equation')
matrices2html(result)
```

The current-to-voltage transfer from expression (18.19) can also be found with SLiCAP. *To this end, we need to define the source and the detector, set the gain type to* `GAIN` *and set the required data type to* `LAPLACE`. *The lines 21 through 33 of the script show the way to do this:*

```
instr.setSource('I1')
instr.setDetector(['V_0', 'V_2'])
instr.setGainType('gain')
instr.setDataType('laplace')
result = instr.execute()
head2html('Transimpedance')
eqn2html('Z_t', result.laplace)
```

The html page generated by this script is shown in Figure 18.5. The matrix equations match those of (18.11) and the expression for the gain equals (18.19).

Figure 18.5: SLiCAP simulation results.

Example Nodal Analysis

MNA equation

Matrix equation:

$$\begin{bmatrix} -I_s \\ 0 \end{bmatrix} = \begin{bmatrix} C_a s + C_b s + \frac{1}{R_a} & -C_b s \\ -C_b s & C_b s + \frac{1}{R_b} \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{1}$$

Transimpedance

$$Z_t = \frac{C_b R_a R_b s}{C_a C_b R_a R_b s^2 + s\left(C_a R_a + C_b\left(R_a + R_b\right)\right) + 1} \tag{2}$$

18.2.2 General form of the admittance matrix

The general form of the nodal equation (Kirchhoff's current law) for node k is:

$$\sum i_k = -\sum \mathbf{Y}_{k,1} v_1 - \sum \mathbf{Y}_{k,2} v_2 \quad \ldots + \sum \mathbf{Y}_{k,k} v_k \quad \ldots - \sum \mathbf{Y}_{k,n-1} v_{n-1}, \tag{18.20}$$

where:

$$\begin{aligned} \sum \mathbf{Y}_{k,j} &= \text{sum of the admittances connected between node } k \text{ and node } j; \\ \sum \mathbf{Y}_{k,k} &= \text{sum of the admittances connected to node } k; \\ \sum i_k &= \text{sum of the independent currents flowing into node } k. \end{aligned} \tag{18.21}$$

In other words: a diagonal element $\mathbf{Y}_{k,k}$ of the admittance matrix $\mathbf{Y}$ equals

the sum of the admittances of each network element connected to node k. So, the first diagonal element is the sum of admittances connected to node 1, the second diagonal element is the sum of admittances connected to node 2, and so on.

The off-diagonal elements $\mathbf{Y}_{k,j}$ equal the sum of the negative admittances of the network elements connected between node k and node j. Hence, an admittance between nodes 1 and 2 appears in the admittance matrix $\mathbf{Y}$ at locations $(1,1)$ and $(2,2)$ with a positive sign and at locations $(1,2)$ and $(2,1)$ with a negative sign. This is shown in Figure 18.6.

A network with passive elements alone, has a symmetrical admittance matrix $\mathbf{Y}_{j,k} = \mathbf{Y}_{k,j}$.

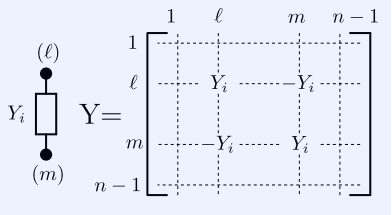


Figure 18.6: A passive admittance and its contributions to the admittance matrix for a network having n nodes including the reference node.

18.2.3 Voltage-controlled current sources

Voltage-controlled current sources can easily be handled in nodal analysis. A voltage-controlled current source G_x, with its current flowing from node ℓ into node m and which is controlled by the voltage between node p (positive) and node q (negative) and a gain of g [A/V], adds, in the ℓ-th row, $+g$ to column p and $-g$ to column q, and in the m-th row $-g$ to column p and $+g$ to column q. This is illustrated in Figure 18.7.

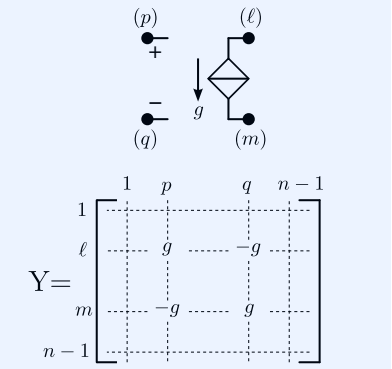


Figure 18.7: Voltage-controlled current source and the coefficients in the Y matrix for a network having n nodes, including the reference node.

Example 18.3

In this example, we will determine the poles and the zeros of the transimpedance of a current-driven and RC-loaded CE-stage. Figure 18.8 shows the small-signal equivalent circuit of a current-driven and RC-loaded CE-stage. The emitter terminal is taken as the reference node.

Nodal analysis yields the following matrix equation

$$\begin{pmatrix} I_s \\ 0 \end{pmatrix} = \mathbf{Y} \cdot \begin{pmatrix} V_i \\ V_\ell \end{pmatrix}, \tag{18.22}$$

in which

$$\mathbf{Y} = \begin{pmatrix} s\left(C_\pi + C_\mu\right) + \frac{1}{r_\pi} & -sC_\mu \\ g_m - sC_\mu & \frac{1}{R'_\ell} + s\left(C_\mu + C_\ell\right) \end{pmatrix}, \tag{18.23}$$

and $R'_\ell = \frac{R_\ell r_o}{R_\ell + r_o}$.

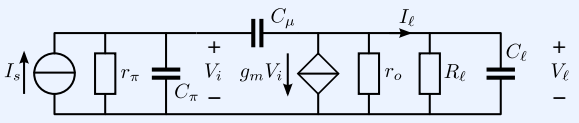


Figure 18.8: Circuit for the determination of the input impedance, the transimpedance and the current gain of the current-driven RC-loaded intrinsic CE-stage

The transimpedance of this stage is defined as

$$Z_t = \frac{V_\ell}{I_s} = \left(\mathbf{Y}^{-1}\right)_{2,1} \tag{18.24}$$

$$= \frac{-R'_\ell r_\pi \left(g_m - sC_\mu\right)}{s^2 r_\pi R'_\ell \left(C_\pi C_\mu + C_\mu C_\ell + C_\pi C_\ell\right) + s\left\{r_\pi g_m R'_\ell C_\mu + R'_\ell \left(C_\mu + C_\ell\right) + r_\pi \left(C_\pi + C_\mu\right)\right\} + 1}.$$

The zeros are the solutions for s from

$$g_m - sC_\mu = 0. \tag{18.25}$$

The poles of this circuit are the solutions for s from

$$s^2 r_\pi R'_\ell \left(C_\pi C_\mu + C_\mu C_\ell + C_\pi C_\ell\right) + \tag{18.26}$$

$$+ s\left\{r_\pi g_m R'_\ell C_\mu + R'_\ell \left(C_\mu + C_\ell\right) + r_\pi \left(C_\pi + C_\mu\right)\right\} + 1. \tag{18.27}$$

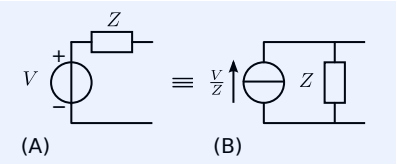


Figure 18.9: Norton equivalent circuit
A) Voltage source in series with an impedance
B) Current source representation using Norton equivalent circuit.

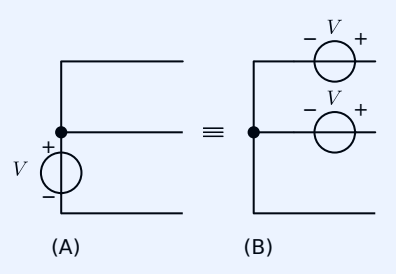


Figure 18.10: Norton equivalent circuit
A) A voltage source connected to a number of branches
B) Equivalent representation using the Blakesly Voltage Shift Theorem.

18.2.4 Network transformations

Nodal Analysis can be applied to analyze networks comprising only elements of which their branch currents can be written as a function of their branch voltages. This is not the case for, e.g., voltage sources; their current is not defined by their voltage. With the aid of network transformations, we can replace voltage sources with current sources and find the network solution with Nodal Analysis.

Any voltage source in series with an impedance can be replaced with a current source in parallel with that impedance (Norton equivalent circuit), thereby reducing the number of nodes by one. This is shown in Figure 18.9.

If a voltage source is connected to a multiple of branches, it must be "shifted through the node" before it can be replaced with a current source. This shifting is known as the Blakesley Voltage Shift; it is shown in Figure 18.10.

In the following example, we will evaluate the transfer of a negative feedback voltage follower with an operational amplifier and use the above transformation techniques.

Example 18.4

Evaluation of the small-signal voltage transfer of a voltage follower realized with an operational amplifier.

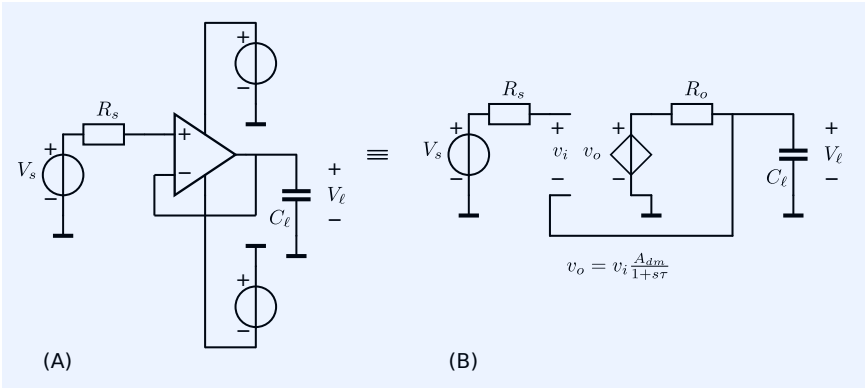


Figure 18.11: Voltage follower and its small-signal model
(A) Voltage follower with operational amplifier and power supplies
(B) Simplified small-signal model of (A).

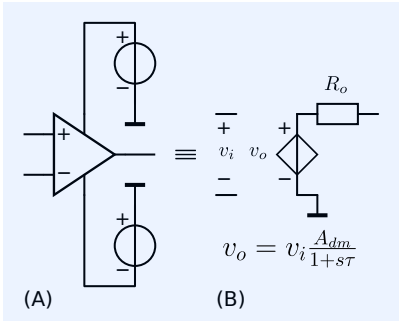


Figure 18.12: Simple OpAmp model
A) Operational Amplifier with power supplies
B) Simplified small-signal model of (A).

In this example, we use a simple operational amplifier model as shown in Figure 18.12. This model includes the DC voltage gain A_{dm}, the DC output resistance R_o and a first order low-pass cut-off. The complete circuit of the voltage follower with the operational amplifier and its small-signal equivalent circuit are shown in Figure 18.11.

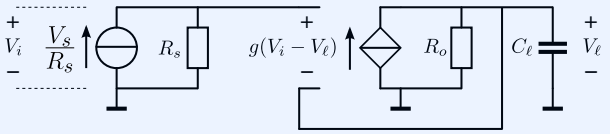
With the aid of Norton equivalent representation, the small-signal model from Figure 18.11B can be further simplified to that of Figure 18.13.

This circuit now has three nodes and its 2×2 admittance matrix can easily be found using Nodal Analysis:

$$\begin{pmatrix} \frac{V_s}{R_s} \\ 0 \end{pmatrix} = \mathbf{Y} \cdot \begin{pmatrix} V_i \\ V_\ell \end{pmatrix}, \tag{18.28}$$

where:

$$\mathbf{Y} = \begin{pmatrix} \frac{1}{R_s} & 0 \\ -\frac{A_{dm}}{R_o(1+s\tau)} & \frac{1}{R_o} + sC_\ell + \frac{A_{dm}}{R_o(1+s\tau)} \end{pmatrix}. \tag{18.29}$$



$$g = \frac{A_{dm}}{R_o(1+s\tau)}$$

Figure 18.13: Norton equivalent circuit for the small-signal model from Figure 18.11B.

The source-to-load transfer can be obtained as:

$$\frac{V_\ell}{V_s} = \frac{\mathcal{C}_{1,2}}{\det(\mathbf{Y})}\frac{1}{R_s} \tag{18.30}$$

$$= \frac{A_{dm}}{s\tau + A_{dm} + sR_oC_\ell + s^2\tau R_oC_\ell + 1}. \tag{18.31}$$

Alternatively, this can be written as

$$\frac{V_\ell}{V_s} = \frac{A_{dm}}{1+A_{dm}}\frac{1}{1 + s\frac{\tau+R_oC_\ell}{1+A_{dm}} + s^2\frac{\tau R_oC_\ell}{1+A_{dm}}}. \tag{18.32}$$

18.3 Modified Nodal Analysis

The nodal analysis technique can only be applied for networks with elements of which the $v-i$ relation can be described in voltage-controlled notation: $i = f(v)$. We have seen that network transformation methods, such as, Blakesley's voltage shift and Thévenin-Norton conversion, can be applied to convert current-controlled elements like voltage sources into voltage-controlled equivalents. In this section, we will introduce the Modified Nodal Analysis method. This technique allows both voltage-controlled and current-controlled notations for network elements.

18.3.1 The procedure

Like Nodal Analysis, Modified Nodal Analysis is based upon the application of Kirchhoff's current law (KCL). Current-controlled elements are included by adding the unknown currents through these elements to the vector with nodal voltages. Additional equations are found from the relations between the nodal voltages to which these elements are connected. In this way the voltages of independent voltage sources are added to the vector with nodal currents. The procedure is as follows:

1. Set up the circuit diagram, select a reference node and number all remaining nodes as in Nodal Analysis.

2. Define the m unknown currents through the m branches with current-controlled notation.

3. Set-up the $n-1$ nodal equations for a network having n nodes

4. Relate the m branch voltages of the elements with current-controlled notation to the nodal voltages, in this way we obtain a square matrix with

$n-1+m$ rows and columns, and a matrix equation of the form

$$\begin{pmatrix} \mathbf{I} \\ \mathbf{V} \end{pmatrix} = \mathbf{M} \cdot \begin{pmatrix} \mathbf{V}_n \\ \mathbf{I}_v \end{pmatrix}, \tag{18.33}$$

in which:

$\begin{pmatrix} \mathbf{I} \\ \mathbf{V} \end{pmatrix}$ = vector with independent current and voltage sources;

$\begin{pmatrix} \mathbf{V}_n \\ \mathbf{I}_v \end{pmatrix}$ = vector with $n-1$ unknown nodal voltages and m unknown currents through the m current-controlled branches;

$\mathbf{M} = \begin{pmatrix} \mathbf{Y} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{pmatrix}$ = matrix that consists of 4 sub matrices:

$\mathbf{Y} = (n-1) \times (n-1)$ admittance matrix as with NA;

$\mathbf{B} = m \times (n-1)$ incidence matrix with topology information of the current-controlled branches;

$\mathbf{C} = (n-1) \times m$ incidence matrix with topology information, $= \mathbf{B}^T$ for networks that have only *independent* voltage sources;

$\mathbf{D} = m \times m$ matrix containing zeros only for *independent* sources.

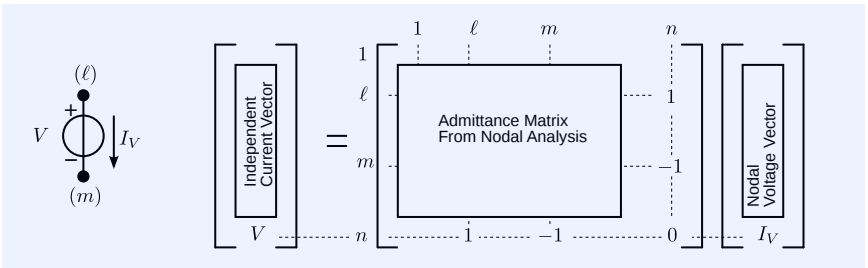


Figure 18.14: Contributions of an independent voltage source to the MNA matrix and the vectors for a network having n nodes, including the reference node. All not shown positions outside the admittance matrix have zeros.

5. The $\mathbf{B}$ matrix is an $m \times (n-1)$ incidence matrix with only 0, 1 and -1 elements. Each location in the matrix corresponds to a branch (row) or node (column). If the positive terminal of the $i-th$ voltage source is connected to node ℓ, then $\mathbf{B}_{i.\ell} = 1$. If the negative terminal of the $i-th$ voltage source is connected to node m, then $\mathbf{B}_{i,m} = -1$. If the network has only *independent* voltage sources, all other elements of the $\mathbf{B}$ matrix are zero and the $\mathbf{C}$ matrix is the transposed version of the $\mathbf{B}$ matrix. This is illustrated in Figure 18.14. The network equation added by the voltage source is:

$$V = V_\ell - V_m \tag{18.34}$$

Since $+I_v$ flows from node ℓ, it is added to the $\ell-th$ nodal equation. Similarly, -1 is added to the $m-th$ equation.

6. Find the network solution (all nodal voltage and all branch currents). All the nodal voltages and unknown branch currents are obtained from

$$\begin{pmatrix} \mathbf{V}_n \\ \mathbf{I}_v \end{pmatrix} = \begin{pmatrix} \mathbf{Y} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{I} \\ \mathbf{V} \end{pmatrix}. \tag{18.35}$$

The branch currents through voltage-controlled elements $I_{j,k}$ ($j \neq k$, $j < n$ and $k < n$) are found from

$$I_{j,k} = (\mathbf{V}_j - \mathbf{V}_k)\, \mathbf{Y}_{j,k}. \tag{18.36}$$

Before we will discuss the coefficients of the **B** and **C** matrix for all kinds of network elements, we will apply the above in an example of MNA for a network with some voltage sources.

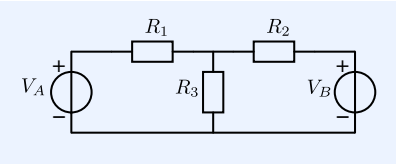


Figure 18.15: Circuit for the demonstration of modified nodal analysis.

Example 18.5

Figure 18.15 shows a resistive network with two voltage sources. We will derive an expression for the voltage across R_3 with the aid of modified nodal analysis. We will start with the selection of the reference node and numbering of the remaining nodes. Figure 18.16 shows the network for netlist generation. The device identifiers (also called `REFDES`*: reference designators) have been typeset with a fixed width font and the device values with a math font: (*`R1`*: R_1). The common terminal of the two voltage sources has been selected as the reference node. The unknown currents through the voltage sources* `V1` *and* `V2` *are I_A and I_B, respectively. This is shown in Figure 18.16. We are now able to set up the MNA matrix equation:*

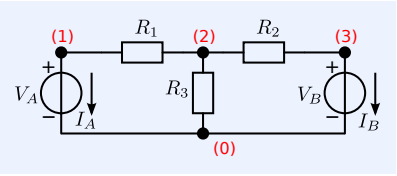


Figure 18.16: Circuit from Figure18.15 with numbered nodes.

$$\begin{pmatrix} 0 & 0 & 0 & V_A & V_B \end{pmatrix}^T = \mathbf{M} \cdot \begin{pmatrix} V_1 & V_2 & V_3 & I_A & I_B \end{pmatrix}^T, \tag{18.37}$$

in which

$$\mathbf{M} = \begin{pmatrix} \frac{1}{R_1} & -\frac{1}{R_1} & 0 & 1 & 0 \\ -\frac{1}{R_1} & \frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3} & -\frac{1}{R_2} & 0 & 0 \\ 0 & -\frac{1}{R_2} & \frac{1}{R_2} & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \tag{18.38}$$

The nodal voltage V_2 can be found as

$$V_2 = V_A \mathcal{C}_{4,2} + V_B \mathcal{C}_{5,2} = \frac{R_2 R_3 V_A + R_1 R_3 V_B}{R_1 R_2 + R_1 R_3 + R_2 R_3}. \tag{18.39}$$

It will be clear that the results of the above example could just as well have been obtained in other ways. With the aid of Norton equivalent circuits, we can eliminate two nodes and obtain quick results.

In the above example, we found V_2 by linear superposition of the voltages caused by the two independent voltage sources V_A and V_B, and by using the transfers from both independent voltage sources to the output voltage. Alternatively, V_2 could be found with the aid of Cramer's rule.

Let us consider a linear system described by the following matrix equation:

$$\mathbf{I} = \mathbf{M} \cdot \mathbf{V}, \tag{18.40}$$

where **I** is the vector of nodal currents and branch voltages and **V** is the vector with nodal voltages and branch currents, Cramer's rule states that the solution of V_i can be found from

$$V_i = \frac{\det \mathbf{M}'}{\det \mathbf{M}}, \tag{18.41}$$

in which $\mathbf{M}'$ is the matrix **M** in which the $i-th$ column has been replaced with the vector **I**.

Example 18.6

We will now find V_2 from the previous example with the aid of Cramer's rule. To do so we obtain the matrix $\mathbf{M}'$ by substituting the 2^{nd} column of **M** *with the*

vector $\begin{pmatrix} 0 & 0 & 0 & V_A & V_B \end{pmatrix}^T$. *This yields*

$$\mathbf{M}' = \begin{pmatrix} \frac{1}{R_1} & 0 & 0 & 1 & 0 \\ -\frac{1}{R_1} & 0 & -\frac{1}{R_2} & 0 & 0 \\ 0 & 0 & \frac{1}{R_2} & 0 & 1 \\ 1 & V_A & 0 & 0 & 0 \\ 0 & V_B & 1 & 0 & 0 \end{pmatrix}. \tag{18.42}$$

Application of Cramer's rule yields

$$V_2 = \frac{\det \mathbf{M}'}{\det \mathbf{M}} = \frac{R_2 R_3 V_A + R_1 R_3 V_B}{R_1 R_2 + R_1 R_3 + R_2 R_3}. \tag{18.43}$$

In the next example, we will determine the MNA matrix equation, as well as the voltage V_2 at node (2) with the aid of SLiCAP.

Example 18.7

Below is the listing of the SLiCAP *netlist file for the circuit from Figure 18.16:*

```
NA_7
* file: NA_7.cir
* SLiCAP netlist for nodal analysis
V1 1 0 {V_A}
R1 1 2 {R_1}
R2 2 3 {R_2}
R3 2 0 {R_3}
V2 3 0 {V_B}
.end
```

The html output page with the SLiCAP *matrix equations and the voltage* V_2 *is shown in Figure 18.17.*

Figure 18.17: SLiCAP simulation results.

Example Modified Nodal Analysis

MNA equation

Matrix equation:

$$\begin{bmatrix} V_A \\ V_B \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & \frac{1}{R_1} & -\frac{1}{R_1} & 0 \\ 0 & 0 & -\frac{1}{R_1} & \frac{1}{R_3}+\frac{1}{R_2}+\frac{1}{R_1} & -\frac{1}{R_2} \\ 0 & 1 & 0 & -\frac{1}{R_2} & \frac{1}{R_2} \end{bmatrix} \cdot \begin{bmatrix} I_{V1} \\ I_{V2} \\ V_1 \\ V_2 \\ V_3 \end{bmatrix} \tag{1}$$

Output voltage

The voltage V_2 at node (2) is obtained as:

$$V_2 = \frac{R_3 (R_1 V_B + R_2 V_A)}{R_1 R_2 + R_1 R_3 + R_2 R_3} \tag{2}$$

The SLiCAP *script that displays the matrix equation on this html page is:*

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
# File: NA-2.py

from SLiCAP import *
prj = initProject('NA-7')
instr = instruction()
instr.setCircuit('NA-7.cir')
instr.setSimType('symbolic')
instr.setGainType('vi')
```

```
instr.setDataType('matrix')
result = instr.execute()
htmlPage("Example Modified Nodal Analysis")
head2html('MNA equation')
matrices2html(result)
```

Evaluation of the voltage V_2 at node (2) requires the definition of the source and the detector. The data type needs to be set to `LAPLACE`*. The script below shows the way in which this has been done:*

```
instr.setDetector('V_2')
instr.setDataType('laplace')
result = instr.execute()
head2html('Output voltage')
text2html("The voltage $V_2$ at node (2) is obtained as:")
eqn2html('V_2', sp.simplify(result.laplace))
```

The matrix equations in Figure 18.17 match those of (18.38) and the expression for the voltage V_2 equals (18.43).

18.3.2 DC and AC network solutions

The DC solution of a network is defined as the solution of a network for $\mathbf{s} = 0$, thus the solution of the simplified equation:

$$\mathbf{I} = \mathbf{G} \cdot \mathbf{V}. \tag{18.44}$$

The DC network described by (18.44) can be obtained from the original network by:

1. Removing all capacitors (replacing them with open circuits)
2. Replacing all inductors with short circuits
3. Substitution of $s = 0$ in the Laplace expressions of the transfer of controlled sources.

A network has a unique DC solution if the matrix $\mathbf{G}$ can be inverted, which requires:

$$\det(\mathbf{G}) \neq 0 \tag{18.45}$$

It is common practice to use the term '*AC solution*' of a network instead of '*small-signal solution*'. In linear networks, the AC solution of a network equals the solution of the complete network, including all capacitors, inductors and complex transfers.

18.3.3 MNA stamps

A netlist of a circuit can be converted into an MNA matrix equation with the aid of '*MNA stamps*'. The MNA stamp of an admittance and a voltage-controlled current source have been shown in Figure 18.6 and in Figure 18.7, respectively. In this section, we will discuss some MNA stamps of two frequently used network elements: the voltage-controlled voltage source and the nullor.

Voltage-controlled voltage sources

Figure 18.18 shows the MNA stamp for a voltage-controlled voltage source E_x with voltage gain A_v, connected between node (1) and (2) and controlled by the voltage between node (3) and (4). The unknown current I_{Ex} through the voltage source is added to the vector with the nodal voltages. The last row of the matrix stamp shows the relation between the branch voltages and the nodal voltages: $V_1 - V_2 = A_v\,(V_3 - V_4)$.

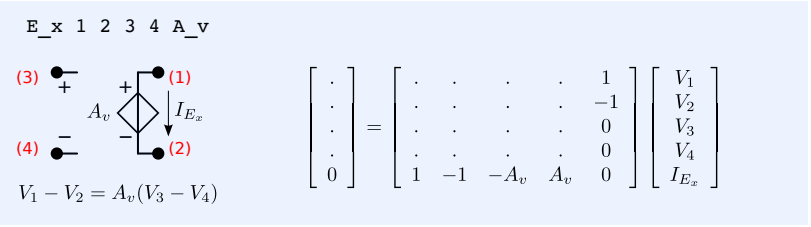


Figure 18.18: Voltage-controlled voltage source with voltage gain A_v, spice syntax, symbol, device equation and MNA stamp.

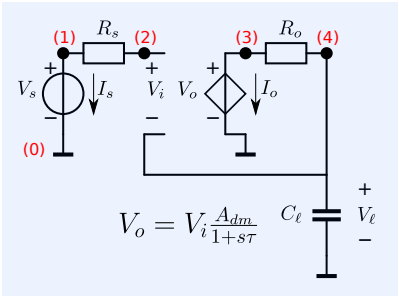


Figure 18.19: Small-signal equivalent circuit from example18.4, prepared for MNA.

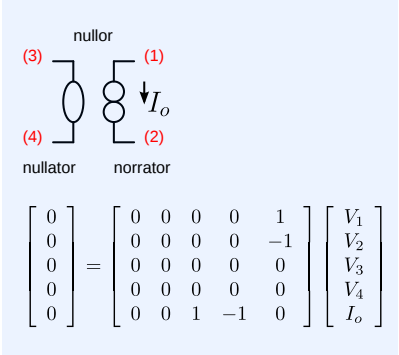


Figure 18.20: The nullor, its symbol, device equations and MNA stamp.

Example 18.8

Let us now solve the problem from example C.18.4 with the aid of MNA.

Figure 18.19 shows the circuit with its nodes and the currents through the voltage sources. The MNA equation for this circuit is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & V_s & 0 \end{pmatrix}^T = \mathbf{M} \cdot \begin{pmatrix} V_1 & V_2 & V_3 & V_4 & I_s & I_o \end{pmatrix}^T, \qquad (18.46)$$

in which:

$$\mathbf{M} = \begin{pmatrix} \frac{1}{R_s} & -\frac{1}{R_s} & 0 & 0 & 1 & 0 \\ -\frac{1}{R_s} & \frac{1}{R_s} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{R_o} & -\frac{1}{R_o} & 0 & 1 \\ 0 & 0 & -\frac{1}{R_o} & \frac{1}{R_o} + sC_\ell & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{A_{dm}}{1+s\tau} & 1 & \frac{A_{dm}}{1+s\tau} & 0 & 0 \end{pmatrix}$$

The transfer from the source voltage V_s to the load voltage $V_\ell = V_4$ is found from

$$\frac{V_\ell}{V_s} = \frac{\mathcal{C}_{5,4}}{\det(\mathbf{M})}, \qquad (18.47)$$

where $\mathcal{C}_{5,4}$ is the element on the fifth row and the fourth column of the cofactor matrix $\mathcal{C}$ of $\mathbf{M}$. We obtain

$$\frac{V_\ell}{V_s} = \frac{A_{dm}}{s\tau + A_{dm} + sR_oC_\ell + s^2\tau R_oC_\ell + 1}, \qquad (18.48)$$

which, of course, is equal to the result from example 18.4, expression 18.30.

The Nullor

Evaluation of the performance of idealized negative-feedback circuits is an almost daily practice for designers of analog circuits. Such circuits comprise a high-gain amplifying device: the *controller* and one or more feedback loops. The network abstraction of an ideal controller is a *nullor*. A nullor consists of two network elements: a *nullator* and a *norator*. The nullator sets a network condition, while the norator adds an independent variable such that this condition can be satisfied. Figure 18.20 shows the symbol and the MNA stamp of a nullor.

The nullator and norator of the nullor only communicate through the external circuit: there exists no relationship between the input port quantities and the output port quantities. A network comprising n nullators only has a valid solution if it also has n norators of which the port variables together create the zero drive conditions for the n nullators. If more nullors appear in one circuit, one can arbitrarily combine nullators and norators in pairs to form nullors without changing the operation of the circuit.

Many simulators do not have the nullor implemented as a standard com-

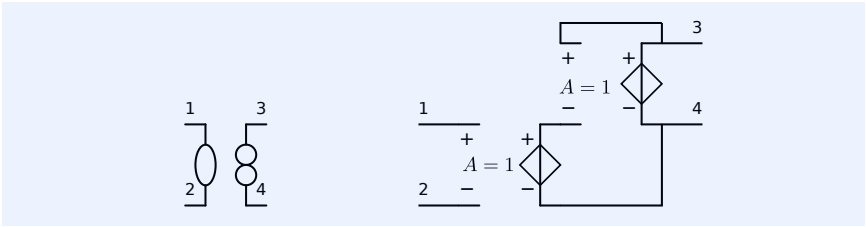


Figure 18.21: The nullor and an implementations with unity-gain voltage-controlled voltage sources.

ponent. Figure 18.21 shows an implementation of a nullor with two unity-gain voltage-controlled voltage sources.

Below is the definition of a nullor sub-circuit for LTSPICE according to the model from Figure 18.21.

```
* file: nullor.cir
* LTspice nullor subcircuit
.subckt nullor 3 4 1 2
E1 3 4 3 5 1
E2 5 4 1 2 1
.ends nullor
```

In the following example, we will evaluate the gain of a negative feedback voltage amplifier that has a nullor as controller.

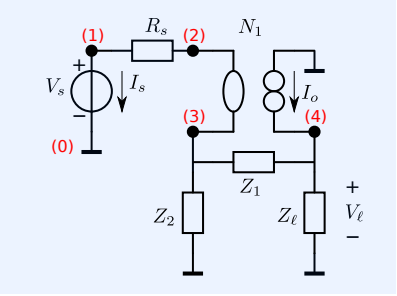


Figure 18.22: Negative-feedback voltage amplifier with a nullor.

Example 18.9

The circuit from Figure 18.22 shows the concept of a passive feedback voltage amplifier. We will evaluate the voltage gain from source to load using MNA. The MNA matrix equation can be found as

$$\begin{pmatrix} 0 & 0 & 0 & 0 & V_s & 0 \end{pmatrix}^T = \mathbf{M} \cdot \begin{pmatrix} V_1 & V_2 & V_3 & V_4 & I_s & I_o \end{pmatrix}^T, \tag{18.49}$$

in which the matrix $\mathbf{M}$ *is defined as:*

$$\mathbf{M} = \begin{pmatrix} \frac{1}{Z_s} & -\frac{1}{Z_s} & 0 & 0 & 1 & 0 \\ -\frac{1}{Z_s} & \frac{1}{Z_s} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{Z_1}+\frac{1}{Z_2} & -\frac{1}{Z_1} & 0 & 0 \\ 0 & 0 & -\frac{1}{Z_1} & \frac{1}{Z_\ell}+\frac{1}{Z_1} & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \end{pmatrix}, \tag{18.50}$$

from which the source-to-load transfer can be found as

$$\frac{V_\ell}{V_s} = \frac{\mathcal{C}_{5,4}}{\det(\mathbf{M})} = \frac{Z_1+Z_2}{Z_2}. \tag{18.51}$$

In the following example, we will evaluate the gain of a unity-gain second order active low-pass filter.

Example 18.10

Figure 18.23 shows the circuit of the two-pole unity-gain active low-pass filter with a nullor as controller. We will derive an expression for the transfer function $T(s) = \frac{V_{out}}{V_s}$.

To this end, we set up the MNA matrix equation of the circuit and obtain:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & V_s & 0 \end{pmatrix}^T = \mathbf{M} \cdot \begin{pmatrix} V_1 & V_2 & V_3 & V_{out} & I_s & I_o \end{pmatrix}^T, \tag{18.52}$$

in which I_s *is the unknown current through the signal voltage source and* I_o *the*

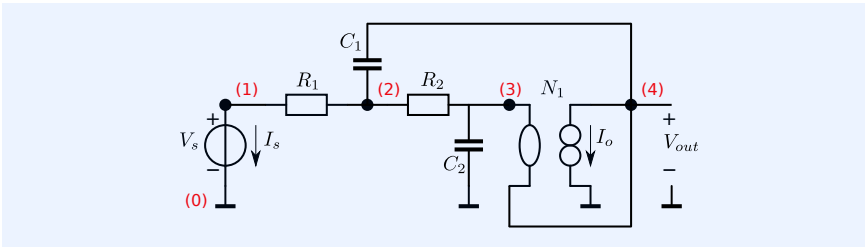


Figure 18.23: Second order active low-pass filter with voltage follower.

current delivered by the nullor. The matrix $\mathbf{M}$ *is obtained as:*

$$\mathbf{M} = \begin{pmatrix} \frac{1}{R_1} & -\frac{1}{R_1} & 0 & 0 & 1 & 0 \\ -\frac{1}{R_1} & \frac{1}{R_1}+\frac{1}{R_2}+sC_1 & -\frac{1}{R_2} & -sC_1 & 0 & 0 \\ 0 & -\frac{1}{R_2} & \frac{1}{R_2}+sC_2 & 0 & 0 & 0 \\ 0 & -sC_1 & 0 & sC_1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix}. \tag{18.53}$$

The transfer function $T(s)$ *that describes the transfer from* V_s *to* V_{out} *can be found as*

$$T(s) = \frac{V_{out}}{V_s} = \frac{\mathcal{C}_{5,4}}{\det(\mathbf{M})} = \frac{1}{1+sC_2\left(R_1+R_2\right)+s^2C_1C_2R_1R_2}. \tag{18.54}$$

Overview of MNA stamps

Figure 18.24 gives an overview of the matrix stamps of two-terminal network elements. All elements are considered to be floating with respect to the reference node. Hence, the dimension of their MNA stamp is the sum of the number terminals and the number of dependent currents of the current-controlled elements. Figure 18.24G shows a compact matrix stamp for a voltage source in series with an impedance. It assumes that the impedance Z can be written as a Laplace rational function: $Z = \frac{N_z(s)}{D_z(s)}$. Both the numerator $N_z(s)$ and the denominator $D_z(s)$ are polynomials of the Laplace variable s. This way of implementing Laplace rational expressions will be discussed in section 18.4.1. This also applies to the matrix stamps of controlled sources. These will be shown in section 18.4.1.

Figure 18.25 shows the matrix stamps of the nullor (A), the gyrator (B), two coupled inductors (C) and the ideal transformer (D).

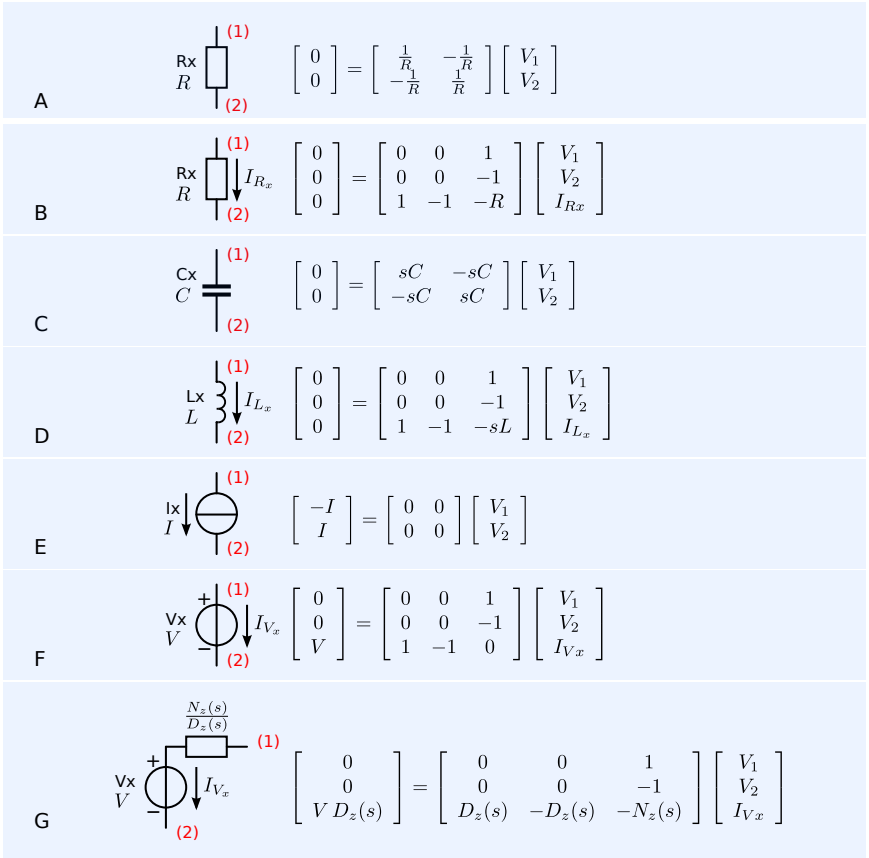


Figure 18.24: MNA stamps of two-terminal network elements:
A: Resistor; cannot have zero resistance
B: Resistor; can have zero resistance
C: Capacitor
D: Inductor
E: Independent current source
F: Independent voltage source
G: Independent voltage source with series impedance; compact matrix stamp.

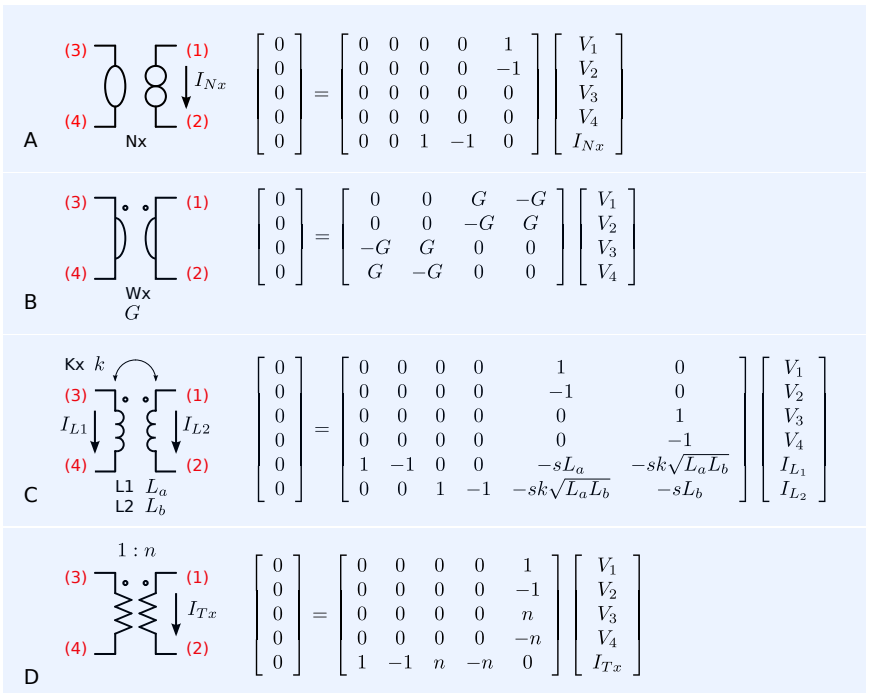


Figure 18.25: MNA stamps of:
A: Nullor
B: Gyrator
C: Coupled inductors
D: Ideal transformer

18.4 Implementation of transfer functions

During the design of electronic circuits, it is sometimes convenient to have high-level models of dynamic systems at one's disposal. Many SPICE-like simulators have the possibility of using Laplace building blocks for this purpose. The general form of a transfer function $H(s)$ that can be modeled with these blocks is that of an univariate rational function of the Laplace variable s:

$$H(s) = \frac{N(s)}{D(s)} = \frac{\sum_{j=0}^{m} b_j s_j}{\sum_{i=0}^{n} a_i s_i}. \tag{18.55}$$

Like SPICE, SLiCAP accepts transfer functions that can be written as a quotient of two univariate polynomials of the Laplace variable s with the order of the denominator $D(s)$ equal to or larger than the order of the numerator $N(s)$ ($n \geq m$). In order to allow such descriptions, they need to be incorporated into the circuit's MNA matrix. In this section, we will discuss two methods for doing this:

1. Direct insertion of $s-$polynomials for numerator and denominator in the MNA matrix. This will be discussed in section 18.4.1.

2. Conversion of $H(s)$ into an equivalent network. This will be discussed in section 18.4.2.

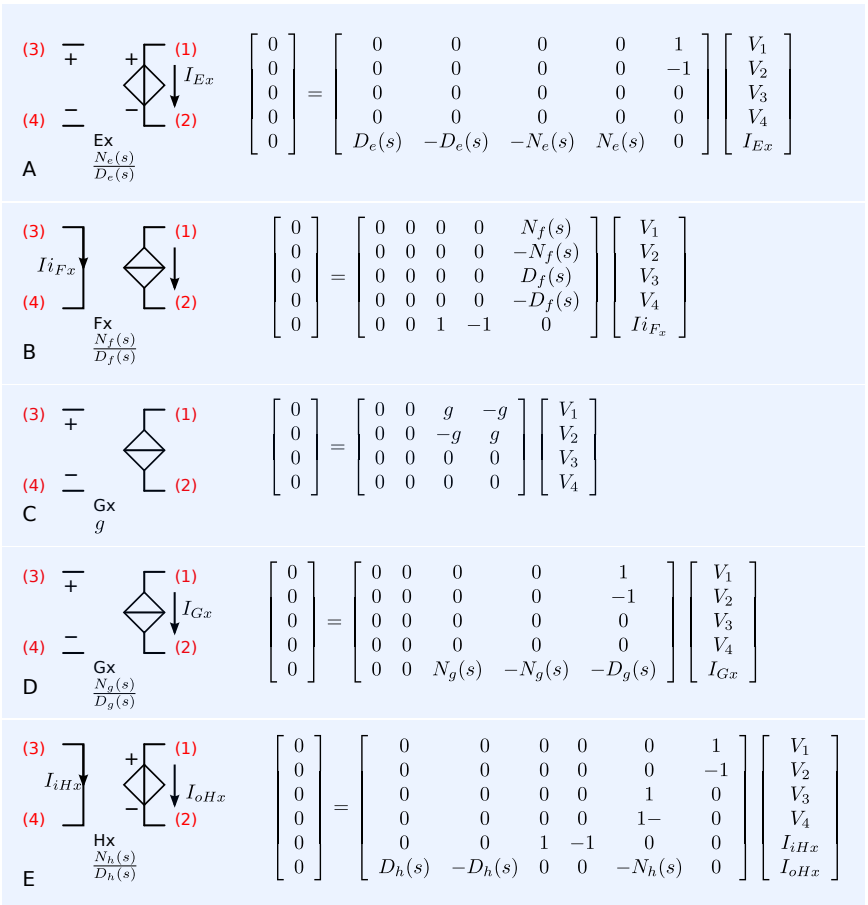


Figure 18.26: MNA stamps of controlled sources of which the transfer can be entered as a Laplace rational function.
A: Voltage-controlled voltage source; VCVS
B: Current-controlled current source; CCCS
C: Voltage-controlled current source; VCCS, compact matrix stamp. The gain must be a constant and not a Laplace rational.
D: Voltage-controlled current source; VCCS.
E: Current-controlled voltage source; CCVS.

18.4.1 Numerator and denominator substitution

In order to determine the poles and the zeros of a network, both the cofactors and the determinant of the MNA matrix need to be obtained as polynomials of the Laplace variable s. For accurate results, all the determinants should then be evaluated through symbolic[4] expansion of minors.[5]

Figure 18.26A, B, D and E show the MNA matrix stamps for the controlled sources of which the transfer can be written as a Laplace rational function. Figure 18.26C shows a compact stamp for the voltage-controlled current source with instantaneous transfer.

[4] By keeping $s^n : n > 0$ symbolically in the expression.

[5] Determination of the cofactors and the determinant of a matrix through expansion of minors, is known as being accurate but slow. For an $n \times n$ matrix with nonzero numeric coefficients, it requires $n!$ multiplications!

18.4.2 Network expansion method

With the aid of the network expansion method, we will be able to write the MNA matrix $\mathbf{M}$ in the form of:

$$\mathbf{M} = \mathbf{G} + s\mathbf{C}. \tag{18.56}$$

The poles (and zeros) of a transfer can now be found by solving the so-called generalized eigenvalue problem; for the poles: $\det(\mathbf{G} + s\mathbf{C}) = \mathbf{0}$, which gives rank($\mathbf{M}$) nonzero solutions. Each pole (or zero) $s_i = \frac{-1}{\tau_i}$ corresponds to an eigenvalue τ_i of the *time-constant matrix* $\mathbf{T} = \mathbf{G}^{-1}\mathbf{C}$. Since the time-constant matrix $\mathbf{T}$ can only be defined if $\det(\mathbf{G}) \neq 0$, the network described by $\mathbf{M}$ must have a unique DC solution.

The method for finding poles and zeros with the aid of the time-constant matrix is presented in section 18.5. In this section, we will focus solely on the network expansion method.

The basis of this technique is to write higher order differential equations as a set of first order differential equations. To illustrate this, let us assume a system of order k:

$$H(s) = \frac{Y(s)}{X(s)} = \frac{b_0 + b_1 s + b_2 s^2 + \cdots + b_k s^k}{a_0 + a_1 s + a_2 s^2 + \cdots + a_k s^k}. \tag{18.57}$$

We may describe this transfer with k first order differential equations. To this end, we introduce k state variables: $Q_1 \cdots Q_k$ with $Q_i = sQ_{i-1}$, with $i = 1 \cdots k$, and write

$$Y(s) = b_0 + b_1 Q_1 + b_2 Q_2 + \cdots + b_k Q_k, \tag{18.58}$$
$$X(s) = a_o + a_1 Q_1 + a_2 Q_2 + \cdots + a_k Q_k, \tag{18.59}$$
$$Q_1 = s, \tag{18.60}$$
$$Q_2 = sQ_1, \tag{18.61}$$
$$\vdots, \tag{18.62}$$
$$Q_k = sQ_{k-1}. \tag{18.63}$$

In matrix form, this set of equations appears as

$$\begin{pmatrix} X(s) \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & a_0 & a_1 & a_2 & \cdots & a_{k-1} & a_k \\ 1 & -b_0 & -b_1 & -b_2 & \cdots & -b_{k-1} & -b_k \\ 0 & s & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & s & -1 \end{pmatrix} \begin{pmatrix} Y(s) \\ 1 \\ Q_1 \\ Q_2 \\ \vdots \\ Q_{k-1} \\ Q_k \end{pmatrix}. \tag{18.64}$$

Such a matrix stamp can easily be added to the network's MNA matrix. If $X(s)$ and $Y(s)$ are available as input and output variables, we can assign either voltage or currents to them.

In some cases, it is convenient to represent the system with an equivalent network and assign the state variables to capacitor voltages or inductor currents.

As we have seen with nodal analysis, the most compact network description is obtained with voltage-controlled elements. Such a matrix description has the form

$$\mathbf{I} = \mathbf{Y} \cdot \mathbf{V}, \tag{18.65}$$

in which $\mathbf{I}$ is the vector with nodal currents that flow and branch voltages, $\mathbf{Y}$ the admittance matrix and $\mathbf{V}$ the vector with nodal voltages and branch currents. Hence, a network that is compatible with Nodal Analysis, requires the state variables to be voltages across k capacitors:

$$Q_0 \cdots Q_k = V_0 \cdots V_k. \tag{18.66}$$

Such a network is shown in Figure 18.27. The k capacitors are driven from voltage-controlled current sources, whose gain is set by the coefficients $a_0 \cdots a_k$ and $b_0 \cdots b_k$. The current-driven capacitors act as integrator elements. In order to find the circuit and its matrix stamp, we need to convert the expression for $H(s)$ (18.55) into the integral form. This is achieved by dividing both the numerator and denominator of $H(s)$ (18.57) by s^k. This yields:

$$H(s) = \frac{\frac{b_0}{s^k} V_k + \frac{b_1}{s^{k-1}} V_{k-1} + \frac{b_2}{s^{k-2}} V_{k-2} + \cdots + b_k V_0}{\frac{a_o}{s^k} V_k + \frac{a_1}{s^{k-1}} V_{k-1} + \frac{a_2}{s^{k-2}} V_{k-2} + \cdots + a_k V_0}. \tag{18.67}$$

For a system order of k, the circuit has $k+2$ nodes: one node for the output, one node for the dummy variable V_o, and k nodes for the state variables. The circuit shown in Figure 18.27 has been designed from the equations

$$I_{out} = \frac{b_0}{s^k} V_k + \frac{b_1}{s^{k-1}} V_{k-1} + \frac{b_2}{s^{k-2}} V_{k-2} + \cdots + b_k V_0, \tag{18.68}$$

$$I_{in} = \frac{a_0}{s^k} V_k + \frac{a_1}{s^{k-1}} V_{k-1} + \frac{a_2}{s^{k-2}} V_{k-2} + \cdots + a_k V_0, \tag{18.69}$$

$$V_i = \frac{1}{s} V_{i-1}. \tag{18.70}$$

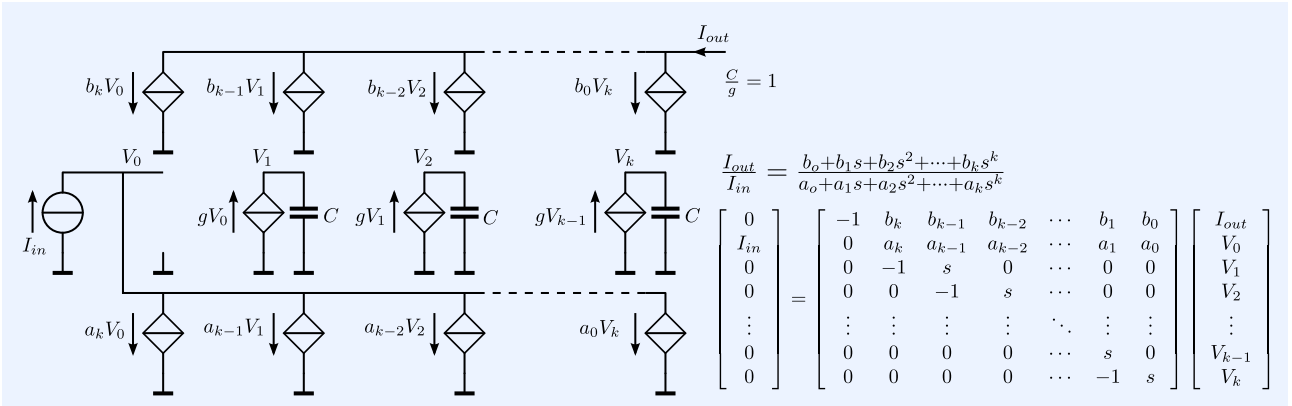


Figure 18.27: Integrator-based circuit for implementation of transfer functions that can be written as a quotient of polynomials, using current-driven capacitors as integrators and its MNA matrix stamp.

18.4.3 Matrix stamps of expanded transfer functions

We will give the matrix stamps for controlled sources whose transfer can be described by a quotient of two polynomials in s. The descriptions are based upon equivalent networks with integrators, in which voltage-controlled current sources drive the capacitors that carry the state variables.[6]

[6] The capacitor voltages.

VCVS

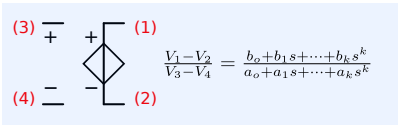


Figure 18.28: Voltage-controlled voltage source with Laplace transfer function.

[7] Modified Nodal Analysis
[8] With $C = 1$ and $g = 1$

Below is the SLiCAP syntax of the voltage-controlled voltage source from Figure 18.28:

```
E_x 1 2 3 4 (b_0+b_1*s+...+b_k*s^k)/(a_0+a_1*s+...+a_k*s^k)
```

Its MNA[7] matrix stamp[8] and the vector $\mathbf{V}$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & b_k & b_{k-1} & \cdots & b_1 & b_0 \\ 0 & 0 & -1 & 1 & 0 & a_k & a_{k-1} & \cdots & a_1 & a_0 \\ 0 & 0 & 0 & 0 & 0 & -1 & s & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & s & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -1 & s \end{pmatrix}, \begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ I_{E_x} \\ V_{0_{E_x}} \\ V_{1_{E_x}} \\ \vdots \\ V_{k-1_{E_x}} \\ V_{k_{E_x}} \end{pmatrix}. \quad (18.71)$$

CCCS

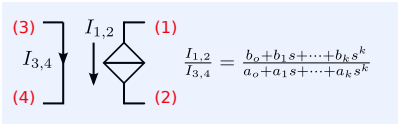


Figure 18.29: Current-controlled current source with Laplace transfer function.

[9] With $C = 1$ and $g = 1$

Below is the SLiCAP syntax of the current-controlled current source from Figure 18.29:

```
F_x 1 2 3 4 (b_0+b_1*s+...+b_k*s^k)/(a_0+a_1*s+...+a_k*s^k)
```

Its MNA matrix stamp[9] and the vector $\mathbf{V}$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & b_k & b_{k-1} & \cdots & b_1 & b_0 \\ 0 & 0 & 0 & 0 & 0 & -b_k & -b_{k-1} & \cdots & -b_1 & -b_0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & a_k & a_{k-1} & \cdots & a_1 & a_0 \\ 0 & 0 & 0 & 0 & 0 & -1 & s & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & s & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -1 & s \end{pmatrix}, \begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ I_{F_{xi}} \\ V_{0_{F_x}} \\ V_{1_{F_x}} \\ \vdots \\ V_{k-1_{F_x}} \\ V_{k_{F_x}} \end{pmatrix}. \quad (18.72)$$

VCCS

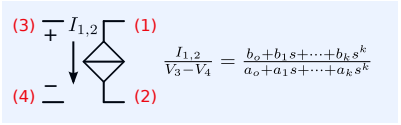


Figure 18.30: Voltage-controlled current source with Laplace transfer function.

[10] With $C = 1$ and $g = 1$

Below is the SLiCAP syntax of the voltage-controlled current source from Figure 18.30:

```
G_x 1 2 3 4 (b_0+b_1*s+...+b_k*s^k)/(a_0+a_1*s+...+a_k*s^k)
```

Its MNA matrix stamp[10] and the vector $\mathbf{V}$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & b_k & b_{k-1} & \cdots & b_1 & b_0 \\ 0 & 0 & 0 & 0 & -b_k & -b_{k-1} & \cdots & -b_1 & -b_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -1 & a_k & a_{k-1} & \cdots & a_{k-1} & a_k \\ 0 & 0 & 0 & 0 & -1 & s & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & s & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -1 & s \end{pmatrix}, \begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_{0_{G_x}} \\ V_{1_{G_x}} \\ \vdots \\ V_{k-1_{G_x}} \\ V_{k_{G_x}} \end{pmatrix}. \quad (18.73)$$

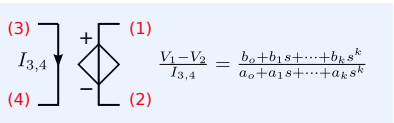


Figure 18.31: Current-controlled voltage source with Laplace transfer function.

[11] With $C = 1$ and $g = 1$

CCVS

Below is the SLiCAP syntax of the current-controlled voltage source from Figure 18.31:

```
H_x 1 2 3 4 (b_0+b_1*s+...+b_k*s^k)/(a_0+a_1*s+...+a_k*s^k)
```

Its matrix stamp[11] and the vector **V** is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & b_k & b_{k-1} & \cdots & b_1 & b_0 \\ 0 & 0 & 0 & 0 & -1 & 0 & a_k & a_{k-1} & \cdots & a_1 & a_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & s & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & s & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -1 & s \end{pmatrix}, \begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ I_{H_{xi}} \\ I_{H_{x_o}} \\ V_{0_{H_x}} \\ V_{1_{H_x}} \\ \vdots \\ V_{k-1_{H_x}} \\ V_{k_{H_x}} \end{pmatrix}. \quad (18.74)$$

18.5 Determination of poles and zeros

Knowledge of poles and zeros is indispensable for the analysis and design of the stability of amplifiers. A network of which the solutions of s of its characteristic equation are located in the left half of the complex plane is stable. Instability cannot always be observed at the circuits's output. In a case of non-observable instability, zeros with the same frequency as the right half plane poles cover up instability. Hence, observation of a stable time-domain response of a circuit does not guarantee the circuit's stability.

In order to get a clear picture of both observable and non-observable instability, only the solutions of the characteristic equations should be evaluated. These solutions are the poles of the network.

We will discuss three ways for the determination of the poles:

1. Direct solution of the characteristic equation; this will be discussed in section 18.5.1.
2. Determination of the eigenvalues of the time-constant matrix; this will be discussed in section 18.5.2.
3. Direct estimation from a network; this will be discussed in section 18.5.3.

18.5.1 Solving the characteristic equation

Consider a network of which the MNA equations are given as

$$\mathbf{I} = \mathbf{M} \cdot \mathbf{V} \tag{18.75}$$

in which $\mathbf{I}$ is the vector with nodal currents and branch voltages, $\mathbf{M}$ the MNA matrix, and $\mathbf{V}$ the vector with nodal voltages and branch currents. The poles of the network can be found by solving the equation

$$\det(\mathbf{M}) = 0. \tag{18.76}$$

The zeros are found by solving

$$\det(\mathbf{M}') = 0, \tag{18.77}$$

in which $\mathbf{M}'$is obtained from $\mathbf{M}$ and $\mathbf{I}$, after application of Cramer's rule.[12]

[12] See section 18.3 for theory on this topic.

In practice, calculation of the determinant of a matrix using Gaussian elimination may generate extra poles due to rounding effects. For accurate determination of the poles and the zeros, powers of the Laplace variable s^n $(n > 0)$ should be kept symbolically in the matrix elements and the determinant should be evaluated through expansion of minors. In order to keep $\mathbf{M}$ or $\mathbf{M}'$ as small as possible, transfer functions can be implemented as discussed in section 18.4.1. Symbolic expansion of minors, however, is a time-consuming process and can be used for small circuits only. This method has been implemented in the current version of SLiCAP.

Another way is to calculate the poles and zeros from a state equation representation of the network. Formulation of the state equations for a network with passive elements and controlled sources, however, is complex, and the procedure for it differs considerably from setting up the MNA matrix. The reader is referred to Chua [Chua1975][13] for more information on setting up the state equations of a network.

[13] Leo O. Chua and Pen-Min Lin. *Computer aided analysis of electronic circuits*. Prentice-Hall, Inc., USA, 1975. ISBN: 0-13-165415-2

Alternatively, poles can be found as the eigenvalues of the so-called *time-constant matrix*, which can be derived from the MNA matrix. This will be discussed in section 18.5.2.

18.5.2 Eigenvalues of the time-constant matrix

Consider the network equations written as a set of first order differential equations.[14] The matrix $\mathbf{M}$ then has the form:

[14] In differential form.

$$\mathbf{M} = \mathbf{G} + s\mathbf{C}. \tag{18.78}$$

The poles are obtained by solving the equation:

$$\det(\mathbf{G} + s\mathbf{C}) = 0. \tag{18.79}$$

This method for determination of the poles, solves the so-called generalized eigenvalue problem as stated in (18.79). This equation can be written in the form

$$\det(\mathbf{I} + s\mathbf{T}) = 0. \tag{18.80}$$

in which $\mathbf{T} = \mathbf{G}^{-1}\mathbf{C}$. Notice that $\det(\mathbf{I} + s\mathbf{T}) = 0$ can be written in the form of the generalized eigenvalue problem: $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, with $\lambda = -\frac{1}{s}$ Hence, if τ_i is an eigenvalue of $\mathbf{T}$, the complex frequency of the corresponding pole p_i is $s = -\frac{1}{\tau_i}$.

However, in most cases the number of eigenvalues of $\mathbf{T}$ exceeds the number of independent state variables.[15] Thus, evaluated in this way, the n poles found from the eigenvalues of $\mathbf{T}$ are the k system poles plus $n - k$ poles at infinity, or after numeric solution, at very high frequencies.

[15] The number of solutions that are found in this way, is equal to the dimension of $\mathbf{M}$. However, the number of nonzero eigenvalues equals the number of independent states, which is the sum of the number of independent capacitor voltages and independent inductor currents.

For systems with only finite nonzero eigenvalues[16], the modified nodal

[16] A circuit has only finite non zero eigenvalues if the number of poles equals the sum of the number of capacitors plus the number of inductors and there are no cut sets of capacitors (and current sources) or loops of inductors (and voltage sources).

equations can be transformed into the state equations, using the transformation technique described by Haley [Haley1988][17]. The time constant matrix can then be written as

$$\mathcal{T} = \mathcal{R} \cdot \mathcal{C}, \tag{18.81}$$

in which $\mathcal{C}$ is a diagonal matrix, having $\mathcal{C}_{ii} = C_i \ (i = 1 \cdots p)$ for a circuit with p capacitors, and $\mathcal{C}_{p+j,p+j} = -L_j \ (j = 1 \cdots q)$ for a circuit having q inductors:

$$\mathcal{C} = \begin{pmatrix} C_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & C_p & 0 & \cdots & 0 \\ 0 & \cdots & 0 & -L_1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & -L_q \end{pmatrix}. \tag{18.82}$$

The resistance matrix $\mathcal{R}$ is obtained as

$$\mathcal{I}^T \cdot \mathbf{G}^{-1} \cdot \mathcal{I}, \tag{18.83}$$

where $\mathcal{I}$ is the incidence matrix. This matrix holds the positions of the elements of $\mathcal{C}$ in $\mathbf{C}$. Its number of rows i equals that of $\mathbf{G}$ and its number of columns j equals the sum of the number of capacitors and the number of inductors. The elements of $\mathcal{I}$ can be obtained as follows:

- $\mathcal{I}_{ij} = 1$
  - if the positive terminal of a capacitor at $\mathcal{C}_{jj}$ is connected to node i.
  - or if $\mathbf{V}_i$ is the current through an inductor at $\mathcal{C}_{jj}$.[18]
- $\mathcal{I}_{ij} = -1$ if the negative terminal of a capacitor at $\mathcal{C}_{jj}$ is connected to node i.
- Otherwise, $\mathcal{I}_{ij} = 0$

Example 18.11 shows the application of the incidence matrix.

The time-constant matrix can only be defined if, $\mathbf{G}^{-1}$ exists. This means that the circuit needs to have a unique DC solution.

[17] Stephen B. Haley. The Generalized Eigenproblem: Pole-Zero Computation. *Proceedings of the IEEE*, 76(2):103–120, February 1988

[18] $\mathbf{V}$ is the vector with dependent variables

Structure and meaning of the resistance matrix $\mathcal{R}$

If we have a network with p capacitors and q inductors, the network can be drawn as a $p+q$ port network in which a capacitor $C_i : i = 1 \cdots p$ is connected to its port i, and an inductor $L_j : j = 1 \cdots q$ is connected to its port $p+j$, as shown in Figure 18.32A.

The resistance matrix can be found from the $(p+j)^2$ DC transfers that exist in this $p+q$ port. To find these transfers, we replace all capacitors C_i with independent current sources I_i, and all inductors L_j with independent voltage sources V_{p+j}, as shown in Figure 18.32B. Then, the voltage V_i across the independent current source I_i at port i, and the current I_{p+j} through the independent voltage source V_{p+j} at port $p+j$, can be found from the following matrix equation:

$$\begin{pmatrix} V_1 & \cdots & V_p & I_{p+1} & \cdots & I_{p+q} \end{pmatrix}^T = \mathcal{R} \cdot \begin{pmatrix} I_1 & \cdots & I_p & V_{p+1} & \cdots & V_{p+q} \end{pmatrix}^T, \tag{18.84}$$

in which $\mathcal{R}$ is the resistance matrix.

The resistance matrix $\mathcal{R}$ can be decomposed into four sub-matrices:

$$\mathcal{R} = \begin{pmatrix} r & \mu \\ \alpha & g \end{pmatrix}. \tag{18.85}$$

Expression (18.85) shows this decomposition in:

1. A $p \times p$ transresistance matrix r
2. A $q \times q$ transconductance matrix g
3. A $p \times q$ voltage transfer matrix μ
4. A $p \times q$ current transfer matrix α

The resistance matrix $\mathcal{R}$ thus has the form:

$$\mathcal{R} = \begin{pmatrix} r_{11} & \cdots & r_{1p} & \mu_{11} & \cdots & \mu_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ r_{p1} & \cdots & r_{pp} & \mu_{p1} & \cdots & \mu_{pq} \\ \alpha_{11} & \cdots & \alpha_{1p} & g_{11} & \cdots & g_{q1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{q1} & \cdots & \alpha_{qp} & g_{1q} & \cdots & g_{qq} \end{pmatrix} \tag{18.86}$$

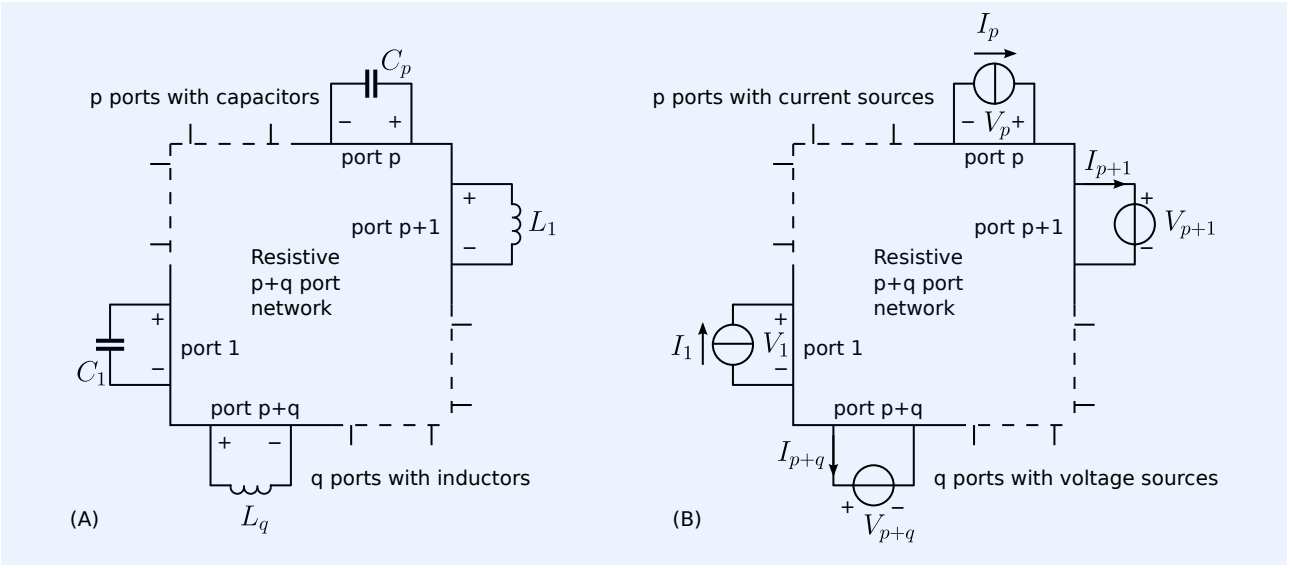


Figure 18.32: Interpretation of the resistance matrix

(A) The complete network is drawn as a resistive multi-port network with capacitors and inductors connected to the ports.

(B) The capacitors are replaced with independent current sources, and the inductors with independent voltage sources.

The resistance matrix describes the relation between the independent sources and the dependent port variables.

The physical meaning of the elements of $\mathcal{R}$ is as follows:

1. A diagonal element r_{ii}, with $i = 1 \cdots p$, represents the port resistance at port i.

   If an independent current source I_i is replaced with the capacitor C_i, and all other independent sources have been set to zero, r_{ii} is the resistance that will dissipate the energy stored in C_i.

2. A diagonal element g_{jj}, with $j = 1 \cdots q$, represents the port conductance of port $p + j$.

   If an independent voltage source V_{p+j} is replaced with the inductor L_j, and all other independent sources have been set to zero, g_{jj} is the conductance that will dissipate the energy stored in L_j.

3. An off-diagonal element $r_{i,k}$, with $i = 1 \cdots p, k = 1 \cdots p$ and $i \neq k$, represents the transresistance from port k to port i.

   A nonzero coefficient $r_{i,k}$ implies charge exchange between C_i and C_k in the original network.

4. An off-diagonal element $g_{j,\ell}$, with $j = 1 \cdots q, \ell = 1 \cdots q$ and $j \neq \ell$, represents the transconductance from port $p + \ell$ to port $p + j$.

   A nonzero coefficient $g_{j,\ell}$ implies flux exchange between inductor L_j and L_ℓ in the original network.

5. An off-diagonal element $\mu_{i,j}$, with $i = 1 \cdots p$ and $j = 1 \cdots q$, represents the voltage transfer from port $p + j$ to port i.

   A nonzero coefficient $\mu_{i,j}$, indicates the exchange of energy storage between L_j and C_i in the original network.

6. An off-diagonal element $\alpha_{i,j}$, with $i = 1 \cdots q$ and $j = 1 \cdots p$, represents the current transfer from port i to port $p + j$.

   A nonzero coefficient $\alpha_{i,j}$, indicates the exchange of energy storage between C_i and L_j in the original network.

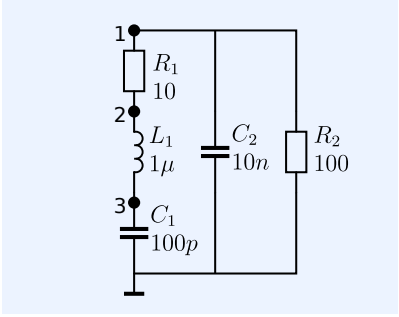


Figure 18.33: Example of a passive circuit for pole-zero calculation.

Example 18.11

We will evaluate the poles of the circuit from Figure 18.33, with the aid of the time-constant matrix. The MNA equations for this circuit are:

$$\begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}^T = (\mathbf{G} + s\mathbf{C}) \cdot \begin{pmatrix} V_1 & V_2 & V_3 & I_{L_1} \end{pmatrix}^T, \tag{18.87}$$

in which

$$\mathbf{G} = \begin{pmatrix} \frac{1}{R_1} + \frac{1}{R_2} & -\frac{1}{R_1} & 0 & 0 \\ -\frac{1}{R_1} & \frac{1}{R_1} & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \tag{18.88}$$

and

$$\mathbf{C} = \begin{pmatrix} C_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & C_1 & 0 \\ 0 & 0 & 0 & -L_1 \end{pmatrix}. \tag{18.89}$$

We find rank $\mathbf{C}$ *is* 3*, which is equal to the number of reactive elements. Hence, there are no zero eigenvalues, and three finite eigenvalues. All capacitor voltages and inductor currents can be taken as state variables.*

We will define the diagonal matrix $\mathcal{C}$ *as:*

$$\mathcal{C} = \begin{pmatrix} C_1 & 0 & 0 \\ 0 & C_2 & 0 \\ 0 & 0 & -L_1 \end{pmatrix}. \tag{18.90}$$

Please notice the minus sign for L; it corresponds with the definition of the positive directions of the port variables as shown in Figure 18.32B.

The incidence matrix $\mathcal{I}$ *can be obtained from the netlist as illustrated in Figure 18.34. For the given circuit we obtain*

$$\mathcal{I} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{18.91}$$

Note that $\mathbf{C}$ *can be obtained from* $\mathcal{I}$ *and* $\mathcal{C}$ *as:*

$$\mathbf{C} = \mathcal{I}\mathcal{C}\mathcal{I}^T. \tag{18.92}$$

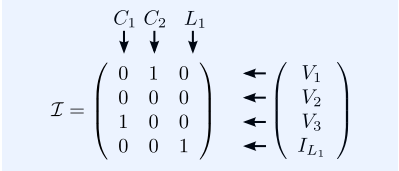


Figure 18.34: Construction of the incidence matrix:
C_1 is connected between node 3 and node o
C_2 is connected between node 1 and node o
The current through L_3 is the dependent variable at the fourth row of the MNA equation.

With the aid of (18.83), we find the resistance matrix $\mathcal{R}$ as

$$\mathcal{R} = \mathcal{I}^T \mathbf{G}^{-1} \mathcal{I} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}^T \cdot \begin{pmatrix} \frac{1}{R_1} + \frac{1}{R_2} & -\frac{1}{R_1} & 0 & 0 \\ -\frac{1}{R_1} & \frac{1}{R_1} & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{18.93}$$

which yields

$$\mathcal{R} = \begin{pmatrix} R_1 + R_2 & R_2 & -1 \\ R_2 & R_2 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{18.94}$$

The time-constant matrix of the circuit from Figure 18.33 is according to its definition from 18.81, obtained as

$$\mathcal{T} = \mathcal{R}\mathcal{C} = \begin{pmatrix} R_1 + R_2 & R_2 & -1 \\ R_2 & R_2 & 0 \\ -1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} C_1 & 0 & 0 \\ 0 & C_2 & 0 \\ 0 & 0 & -L_1 \end{pmatrix}, \tag{18.95}$$

which yields:

$$\mathcal{T} = \begin{pmatrix} C_1 (R_1 + R_2) & C_2 R_2 & L_1 \\ C_1 R_2 & C_2 R_2 & 0 \\ -C_1 & 0 & 0 \end{pmatrix}, \tag{18.96}$$

with τ_i being an eigenvalue of $\mathbf{T}$. The poles p_i are found from these eigenvalue:

$$p_i = -\frac{1}{\tau_i}. \tag{18.97}$$

Let us now substitute the numeric element values for the elements from Figure 18.33 and evaluate the pole frequencies:

$$R_1 = 10,\ R_2 = 100,\ L_1 = 10^{-6},\ C_1 = 10^{-10},\ C_2 = 10^{-8}. \tag{18.98}$$

We then have

$$\mathcal{T} = \begin{pmatrix} 1.1 \times 10^{-8} & 1.0 \times 10^{-6} & 1.0 \times 10^{-6} \\ 1.0 \times 10^{-8} & 1.0 \times 10^{-6} & 0 \\ -1.0 \times 10^{-10} & 0 & 0 \end{pmatrix}, \tag{18.99}$$

and the eigenvalues of $\mathbf{T}$ are obtained as

$$\tau_1 = 1.01 \times 10^{-6}, \tag{18.100}$$

$$\tau_2 = 4.955\,4 \times 10^{-10} + 9.938\,0 \times 10^{-9} j, \tag{18.101}$$

$$\tau_3 = 4.955\,4 \times 10^{-10} - 9.938\,0 \times 10^{-9} j. \tag{18.102}$$

Hence, the poles of the circuit are

$$p_1 = -\frac{1}{\tau_1} = -9.901 \times 10^5, \tag{18.103}$$

$$p_2 = -\frac{1}{\tau_2} = -5.005 \times 10^6 + 1.004 \times 10^8 j, \tag{18.104}$$

$$p_3 = -\frac{1}{\tau_3} = -5.005 \times 10^6 - 1.004 \times 10^8 j. \tag{18.105}$$

From this example, we see that the poles can easily be found from the eigenvalues of the time-constant matrix $\mathbf{T}$. A similar procedure can be followed for zeros. The time-constant matrix is then derived from $\mathcal{R}'$ and $\mathcal{C}'$ which are obtained from $\mathcal{R}$ and $\mathcal{C}$ after application of Cramer's rule.

In the next example, we will demonstrate the determination of the resis-

tance matrix $\mathcal{R}$ by network inspection.

Example 18.12

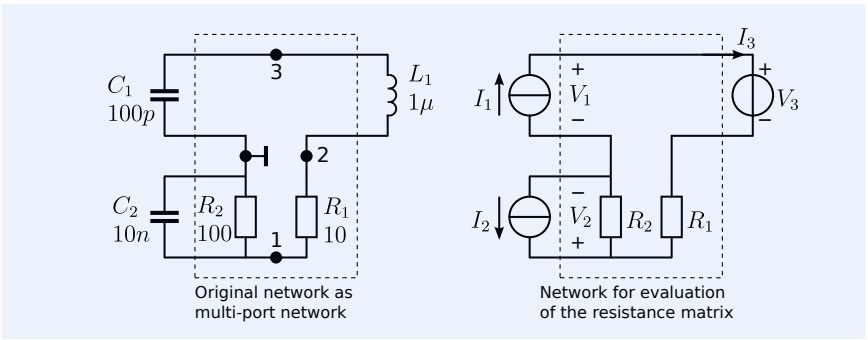


Figure 18.35: Setting up the R-matrix by network inspection.

Figure 18.35 shows the circuit from Figure 18.33 redrawn as a multi-port network. The left side of the figure shows the original circuit with the capacitors and the inductor connected to a resistive multi-port. Capacitor C_1 is connected to port 1, capacitor C_2 to port 2 and inductor L_1 to port 3.

The right side of the figure shows the same network with the capacitors replaced with current sources and the inductor with a voltage source. The resistor matrix $\mathcal{R}$ holds the relations between the dependent and independent port variables:

$$\begin{pmatrix} V_1 \\ V_2 \\ I_3 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \cdot \begin{pmatrix} I_1 \\ I_2 \\ V_3 \end{pmatrix}. \tag{18.106}$$

The coefficients R_{ij} are found from network inspection:

$$R_{11} = \left.\frac{V_1}{I_1}\right|_{I_2=0,V_3=0} = R_1 + R_2, \tag{18.107}$$

$$R_{12} = \left.\frac{V_2}{I_1}\right|_{I_1=0,V_3=0} = R_2, \tag{18.108}$$

$$R_{13} = \left.\frac{V_1}{V_3}\right|_{I_1=0,I_2=0} = 1, \tag{18.109}$$

$$R_{21} = \left.\frac{V_1}{I_2}\right|_{I_1=0,V_3=0} = R_2, \tag{18.110}$$

$$R_{22} = \left.\frac{V_2}{I_2}\right|_{I_1=0,V_3=0} = R_2, \tag{18.111}$$

$$R_{23} = \left.\frac{V_2}{V_3}\right|_{I_1=0,I_2=0} = 0, \tag{18.112}$$

$$R_{31} = \left.\frac{I_3}{I_1}\right|_{I_2=0,V_3=0} = 1, \tag{18.113}$$

$$R_{32} = \left.\frac{I_3}{I_2}\right|_{I_1=0,V_3=0} = 0, \tag{18.114}$$

$$R_{33} = \left.\frac{I_3}{V_3}\right|_{I_1=0,I_2=0} = 0. \tag{18.115}$$

Hence, we obtain:

$$\mathcal{R} = \begin{pmatrix} R_1 + R_2 & R_2 & 1 \\ R_2 & R_2 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{18.116}$$

This result corresponds to (18.94).

As stated earlier, loops of capacitors and/or capacitors with voltage sources, and cut sets of inductors and/or inductors with current sources, result in eigenvalues with zero value.[19] Due to numerical rounding effects, these poles will appear at finite frequencies rather than at infinity. However, in most cases, these frequencies will be outside the frequency range of interest, and can easily be separated from the desired ones, and ignored. According to Haley [Haley1988][20], the method is robust if the number of capacitor loops and inductor cut sets is relatively low with respect to the number of capacitors and inductors.

[19] Poles at infinity.

[20] Stephen B. Haley. The Generalized Eigenproblem: Pole-Zero Computation. *Proceedings of the IEEE*, 76(2):103–120, February 1988

18.5.3 Symbolic estimation of poles and zeros

During the design of the frequency response of electronic circuits, designers often need to manipulate poles and zeros into their desired positions. To do so, the designer needs to know which network elements determine the positions of specific poles and zeros. One method for this is to perform a numeric sensitivity analysis. If the value of a specific pole or zero depends on a component value, then the value of that pole or zero can be changed with that component. This requires analyses of the influence of the values of all the components that constitute the elements of the time-constant matrix. This approach may be used for computer analysis, but it is not suited for quick hand estimations.

In this section, we will demonstrate the way in which poles and zeros of a network can be estimated. Limitations of the presented technique will also be shown.

Number of poles

The number of poles of a network equals the sum of the number of independent capacitor voltages and the number of independent inductor currents.

1. The number of independent capacitor voltages equals the number of capacitors minus the number of loops of capacitors or loops of capacitors and voltage sources.

2. The number of independent inductor currents equals the number of inductors minus the number of cut sets of inductors or cut sets of inductors and current sources.

3. In networks with controlled sources, there is no straightforward way to determine the number of poles:

   The presence of feedback across controlled sources may effectively add or remove loops of controlled voltage sources and capacitors, or cut sets of controlled current sources and inductors. If there are no feedback loops involving controlled sources, then these sources can be treated as independent sources.

4. In networks with ideally coupled inductors[21]:
   The number of independent inductor currents is reduced by the number of unity coupling factors;
   The number of independent capacitor voltages is reduced by one for each capacitor driven from voltage-driven, ideally coupled inductors.

[21] Coupling factor $k = 1$.

Some situations sketched above will be elucidated in the next example.

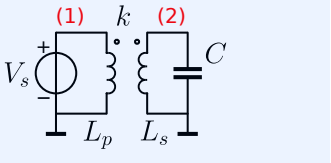


Figure 18.36: Two coupled inductors, driven from a voltage source and terminated with a capacitor. If $k < 1$, the circuit has three poles. If $k = 1$, the number of poles reduces to one.

Example 18.13

Figure 18.36A shows a circuit with two coupled inductors.

One of the coupled inductors is driven by an ideal voltage source, while the other is connected to a capacitor C.

If $k = 1$, the number of independent inductor currents is 1, and there exists a loop of a controlled voltage source and a capacitor. This can be seen from the equivalent circuit shown in Figure 18.36B. Having $k = 1$, the inductance $\left(1 - k^2\right) L_p$ in series with the voltage source V_s becomes zero. As a consequence,. the current feedback α cannot affect the driving impedance for C, which makes C driven from a voltage source.

This can also be seen by evaluating the determinant Δ of the MNA matrix. The MNA matrix equations for this circuit is

$$\begin{pmatrix} 0 \\ 0 \\ V_s \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & sC & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -sL_p & -sk\sqrt{L_pL_s} \\ 0 & 1 & 0 & -sk\sqrt{L_pL_s} & -sL_s \end{pmatrix} \cdot \begin{pmatrix} V_1 \\ V_2 \\ I_{V_s} \\ I_{L_p} \\ I_{L_s} \end{pmatrix}. \quad (18.117)$$

The determinant Δ can be found as:

$$\Delta = s^3 C L_p L_s \left(k^2 - 1\right) - L_p s. \quad (18.118)$$

Hence, the order of the circuit is 3 for $k < 1$ and 1 for $k = 1$.

This example also shows the effect of a controlled source on the number of independent capacitor voltages. Although Figure 18.36B shows a loop of a controlled voltage source and the capacitor C, this loop does not reduce the number of independent capacitor voltages if $k < 1$. This is due to the current feedback with the current-controlled current source. This feedback creates a nonzero drive impedance for C. Hence, due to this feedback, the voltage across C is an independent capacitor voltage.

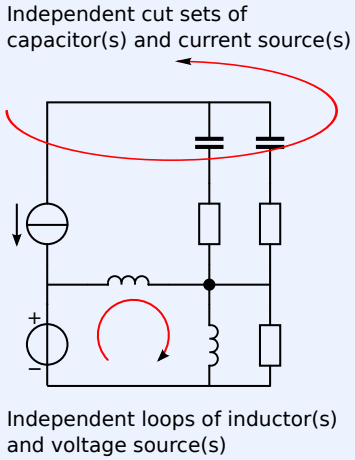


Figure 18.37: Circuit that has two independent capacitor voltages and two independent inductor currents. Hence it has four poles. Since it has one loop of inductors and voltage sources and one cut set of capacitors and current sources, two of the four poles have zero frequency.

Number of zero-frequency poles

The number of poles at zero frequency equals the number of independent cut sets of capacitors and independent cut sets of capacitors and current sources, plus the number of independent loops of inductors and independent loops of inductors and voltage sources.

The number of independent loops of certain types of elements is equal to the minimum number of branches with these elements, which needs to be removed to break all the loops.

The number of independent cut sets of certain types of elements is equal to the minimum number of branches that have to be added to connect all the sub-networks, that were disconnected from each other, after all branches that were part of a cut set were removed.

Example 18.14

The circuit depicted in Figure 18.37 has two capacitors and two inductors. All capacitor voltages and inductor currents are independent., so the circuit has four poles. There exists one cut set of capacitors and current sources and one loop of inductors and voltage sources. Hence, two of the four poles have zero frequency.

Note that feedback loops that include controlled sources may also affect the number of zero-frequency poles.

Estimation of pole positions for circuits without feedback

Now that we know the total number of poles and the number of zero-frequency poles, we will discuss a method for finding symbolic expressions for the finite nonzero poles of a network. In order to provide design information, these expressions must be easy to interpret, in other words, they must be relatively simple. Simple expressions are obtained if there is no charge exchange between capacitors, flux exchange between inductors and charge and flux exchange between capacitors and inductors. For circuits with controlled sources, an extra requirement is the absence of negative feedback. In these cases, the time-constant matrix comprises only diagonal elements that are the eigenvalues of the matrix.

Example 18.15

We will estimate the poles of the circuit shown in Figure 18.38.

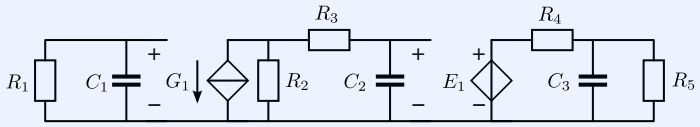


Figure 18.38: Example of a circuit with no charge exchange between capacitors.

The unilateral transfer of the controlled sources prevents charge exchange between the capacitors. As a consequence, the time-constant matrix $\mathcal{T}$ of this circuit comprises nonzero diagonal elements only:

$$\mathcal{T} = \begin{pmatrix} R_1 C_1 & 0 & 0 \\ 0 & (R_2 + R_3) C_2 & 0 \\ 0 & 0 & \frac{R_4 R_5}{R_4 + R_5} C_3 \end{pmatrix}. \tag{18.119}$$

This time-constant matrix is the product of the resistance matrix $\mathcal{R}$ and the capacitance matrix $\mathcal{C}$. The resistance matrix can easily be found from network inspection, as discussed in section 18.5.2.

The poles of the circuit are the negative reciprocal values of the diagonal elements of $\mathcal{T}$:

$$p_1 = -\frac{1}{R_1 C_1} \text{ [rad/s]}, \tag{18.120}$$

$$p_2 = -\frac{1}{(R_2 + R_3) C_2} \text{ [rad/s]}, \tag{18.121}$$

$$p_3 = -\frac{R_4 + R_5}{R_4 R_5 C_3} \text{ [rad/s]}. \tag{18.122}$$

If the exchange of charge, flux or charge and flux is limited between pairs of components, the situation becomes more difficult, but it many cases, useful symbolic approximations for the poles can still be found. This will be illustrated with the aid of the circuit from Figure 18.39.

The circuit from Figure 18.39A has two capacitors whose voltages are independent. Hence, the circuit has two poles. Since there exists one cut set of capacitors, one of the poles has zero frequency.

The second pole can be found as follows:

1. Replace the largest capacitor with a short and find the second pole from the time constant $R_1 C_2$ as shown in Figure 18.39B1. In this case, this is sufficiently accurate if $C_1 \gg C_2$.

2. Replace the two capacitors by their series connection and find the second pole from the time constant $R_1 \frac{C_1 C_2}{C_1 + C_2}$ as shown in Figure 18.39B2.

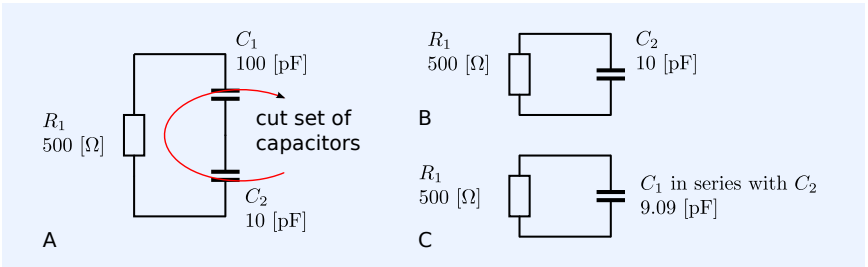


Figure 18.39: Circuit with two poles. Since there exists a cut set of capacitors, it has one zero-frequency pole.

Since the circuit has no DC solution, the time-constant matrix is not defined and cannot be of any help for finding the poles.

First order approximation for finite nonzero poles

For symbolic analysis of the poles of a network that has a DC solution, we use a procedure based on the evaluation and interpretation of the time-constant matrix, as illustrated in Figure 18.32. Consider hereto the network of Figure 18.40A.

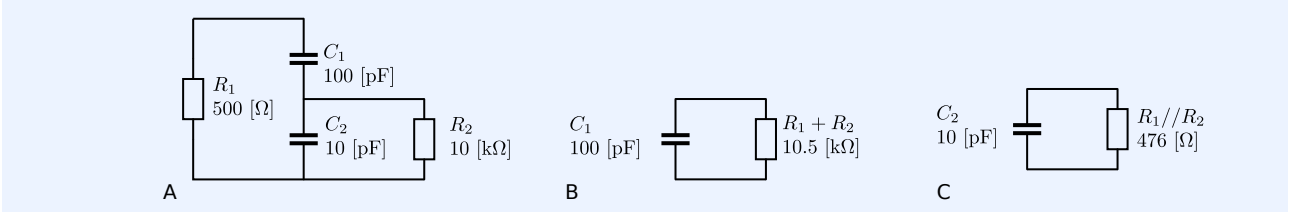


Figure 18.40: Circuit with two poles and interaction between capacitor voltages.

The procedure applied to this network is as follows:

1. Find the largest time constant in the circuit, by evaluating the DC port resistances for the capacitors and the DC port conductances for the inductors. The dominant pole is determined by this time constant.

   For the given circuit, the DC port resistance for C_1 is $R_1 + R_2$. The DC port resistance for C_2 is R_2. The largest time constant of the circuit is $C_1 (R_1 + R_2)$; the dominant pole p_1 is thus found from the equivalent circuit shown Figure 18.40A:

$$p_1 = -\frac{1}{C_1 (R_1 + R_2)} \text{ [rad/s]}. \tag{18.123}$$

2. Now, short the port with the capacitor that caused the largest time constant.

   This is justified because, at frequencies above the frequency of the dominant pole, the capacitor impedance is lower than the DC port impedance. Thus, above the frequency of the dominant pole, the capacitor that determines the dominant pole acts as a short for its driving resistance.

3. Similarly, if an inductor was connected to this port, leave the port open.

4. Now find the largest time constant of the modified circuit.

   With C_1 shorted, we obtain the circuit according to Figure 18.40C, and thus obtain the second pole p_2 as

$$p_2 = -\frac{R_1 + R_2}{C_2 R_1 R_2}. \tag{18.124}$$

The above presented method only gives exact results if there is no exchange of charge and/or flux between capacitors and/or inductors.

Second order approximation for finite nonzero poles

For second order systems, a more accurate solution is obtained by using all elements of the circuit's time-constant matrix. According to Haley and Hurst [Haley1989][22], the sum of the eigenvalues of the time-constant matrix equals its first order trace T_1, and the sum of the mutual products of the eigenvalues equals its second order trace T_2.[23,24] For a second order system this yields

$$\frac{1}{p_1} + \frac{1}{p_2} = -T_1, \tag{18.125}$$

$$\frac{1}{p_1 p_2} = T_2. \tag{18.126}$$

We will demonstrate this for the circuit from Figure 18.40A. For this circuit, we have

$$\mathbf{G} = \begin{pmatrix} \frac{1}{R_1} & 0 \\ 0 & \frac{1}{R_2} \end{pmatrix}, \; \mathcal{C} = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}, \; \mathcal{I} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}. \tag{18.127}$$

Hence, according to its definition, the time-constant matrix is found as[25]

$$\mathcal{T} = \mathcal{I}^T \cdot \mathbf{G}^{-1} \cdot \mathcal{I} \cdot \mathcal{C} = \begin{pmatrix} C_1 (R_1 + R_2) & -C_2 R_2 \\ -C_1 R_2 & C_2 R_2 \end{pmatrix}. \tag{18.128}$$

From the first order trace, we find the sum of the eigenvalues:

$$\tau_1 + \tau_2 = C_1 (R_1 + R_2) + C_2 R_2. \tag{18.129}$$

From the second order trace, we find the product of the eigenvalues:

$$\tau_1 \tau_2 = C_1 C_2 (R_1 + R_2) R_2 - C_1 C_2 R_2^2, \tag{18.130}$$

$$= C_1 C_2 R_1 R_2. \tag{18.131}$$

If the two poles are well separated: $p_1 \ll p_2$, the sum of the eigenvalues is dominated by the largest. We then obtain the dominant pole p_1 as

$$p_1 = -\frac{1}{\tau_1} = -\frac{1}{C_1 (R_1 + R_2) + C_2 R_2}. \tag{18.132}$$

With $R_1 \ll R_2$, this can be simplified to

$$p_1 = -\frac{1}{(C_1 + C_2) R_2}. \tag{18.133}$$

The non-dominant pole is then found from the product $p_1 p_2$ and the dominant pole p_1:

$$p_2 = \frac{p_1 p_2}{p_1} = -\frac{C_1 (R_1 + R_2) + C_2 R_2}{C_1 C_2 R_1 R_2}. \tag{18.134}$$

An overview of the numeric and symbolic results is given in Table 18.1. The most accurate numeric values are those found from the eigenvalues of the time-constant matrix.

The first order approximation for this circuit has an inaccuracy of about 10%. However, it provides symbolic results that can be interpreted clearly:

- The dominant pole is predominantly caused by[26]: $C_1 \| (R_1 + R_2)$.
- The non-dominant pole is predominantly caused by: $C_2 \| R_1 \| R_2$.

Such simple expressions are useful during design.

[22] Stephen B. Haley and Paul J. Hurst. Pole and Zero Estimation in Linear Circuits. *IEEE Transactions on Circuits and Systems*, 36(6):838–845, June 1989

[23] The trace of a matrix is the sum of the diagonal elements of a matrix.

[24] The second-order trace of a matrix is the sum of the determinants of all 2×2 submatrices, that can be formed by intersecting any two rows of a matrix with the same two columns.

[25] Alternatively, the time-constant matrix can be found from network inspection.

[26] $\|$: in parallel with.

The second order approach still gives relatively simple expressions, but, compared to the first order approximation, the poles cannot be calculated from parallel connections of a an equivalent resistor and capacitor. The accuracy of the second order approximation, however, is much better.

Exact symbolic expressions for the eigenvalues of the time-constant matrix can be found for first, second and third order networks. However, the complexity of those expressions often decreases their usefulness for deriving design conclusions. For this reason, exact symbolic expressions for the poles are not listed in Table 18.1.

	Eigenvalues(T)	**First order approximation**		**Second order approximation**	
	numeric	symbolic	numeric	symbolic	numeric
p_1	-873×10^3	$-\frac{1}{C_1(R_1+R_2)}$	-954×10^3	$-\frac{1}{C_1(R_1+R_2)+C_2R_2}$	-870×10^3
p_2	-230×10^6	$-\frac{R_1+R_2}{C_2R_1R_2}$	-210×10^6	$-\frac{C_1(R_1+R_2)+C_2R_2}{C_1C_2R_1R_2}$	-230×10^6

Table 18.1: Poles of the circuit from Figure 18.40A in [rad/s], according to different calculation methods.

LCR resonators

In passive circuits and in circuits with controlled sources that are not part of a feedback loop, only interaction between capacitors and inductors may result in complex pole pairs. If so, one speaks of resonance. Figure 18.41 shows four passive second order *LCR* networks in which resonance may occur. The general representation of the characteristic equation Δ of these networks is

$$\Delta = s^2 + s\frac{\omega_0}{Q} + \omega_0^2, \tag{18.135}$$

in which ω_0 is the resonance frequency and Q the quality factor. This factor is a measure for the losses that occur during exchange of energy between the capacitor and the inductor:

$$Q = 2\pi \frac{\text{Energy stored in the resonator}}{\text{Energy losses per cycle}}. \tag{18.136}$$

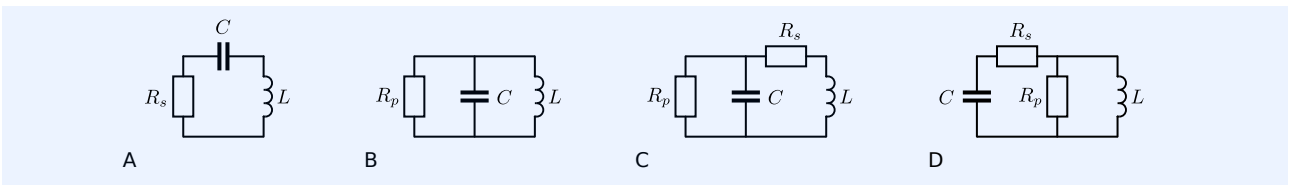


Figure 18.41: Second-order LCR networks.

Expression (18.135) can alternatively be written as

$$\Delta = s^2 - s\,(p_1 + p_2) - p_1p_2.$$

The poles p_1 and p_2 of these second order *LCR* resonators can thus be found from their product and their sum:

$$p_1p_2 = \omega_0^2, \tag{18.137}$$

$$p_1 + p_2 = -\frac{\omega_0}{Q}. \tag{18.138}$$

Table 18.2 gives the product and the sum of the poles, as well as the resonance frequency and the quality factor for the networks from Figure.18.41.

	Figure 18.41A	Figure 18.41B	Figure 18.41C	Figure 18.41D
$p_1 p_2$	$\frac{1}{LC}$	$\frac{1}{LC}$	$\frac{1}{LC}\frac{R_p+R_s}{R_p}$	$\frac{1}{LC}\frac{R_p}{R_p+R_s}$
$p_1 + p_2$	$-\frac{R_s}{L}$	$-\frac{1}{R_p C}$	$-\frac{1}{R_p C} - \frac{R_s}{L}$	$-\frac{1}{C(R_p+R_s)} - \frac{R_p R_s}{L(R_p+R_s)}$
ω_0^2	$\frac{1}{LC}$	$\frac{1}{LC}$	$\approx \frac{1}{LC};\ R_p \gg R_s$	$\approx \frac{1}{LC};\ R_p \gg R_s$
Q	$\frac{1}{R_s}\sqrt{\frac{L}{C}} = Q_s$	$R_p\sqrt{\frac{C}{L}} = Q_p$	$\approx \frac{1}{\frac{1}{Q_s}+\frac{1}{Q_p}};\ R_p \gg R_s$	$\approx \frac{1}{\frac{1}{Q_s}+\frac{1}{Q_p}};\ R_p \gg R_s$

Table 18.2: Product of the poles, sum of the poles, resonance frequency and quality factor for the circuits from Figure 18.41

Estimation of pole positions for circuits with feedback

Feedback in circuits that comprise controlled sources can have a significant influence on the position of the poles. In section 11.4 we will discuss a graphical method for determination of the poles in a feedback circuit.

Estimation of zeros

We have seen that determination of the poles of a network, does not require the definition of a source-to-load transfer. As a matter of fact, the poles of the network are found from the solution of the set of homogeneous differential equations that describe the network. For determination of the poles, all independent sources have been set to zero: all current sources become open circuits and voltage sources short circuits.

This differs for the determination of the zeros: this requires the definition of a source and a detector. Zeros appear at the complex frequencies for which the transfer from the source to the detector equals zero. The physical causes for zero transfer can be:

1. At some complex frequency, there exists an open circuit in series with the signal path.

Example 18.16

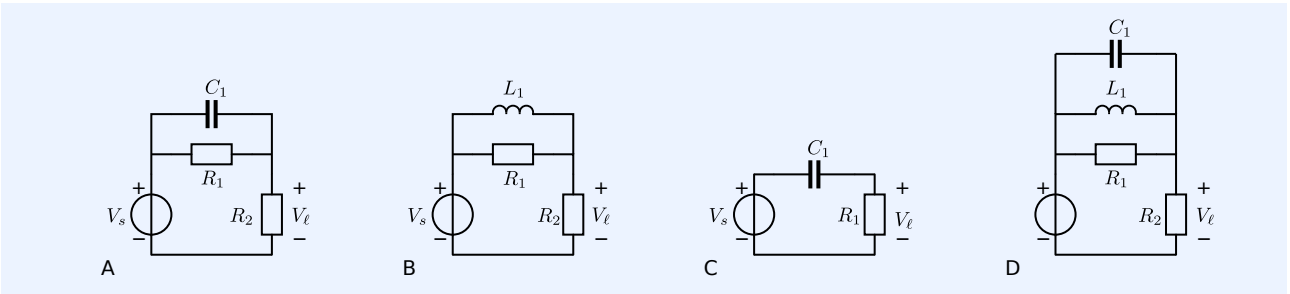


Figure 18.42: At some complex frequency, an open circuit in series with the signal path causes a zero.

(a) Let us consider the transfer from V_s to V_ℓ for the circuit from Figure 18.42A. If, at some complex frequency, the current through R_1 is cancelled by an opposite current through C_1, no signal transfer is possible through the parallel connection of R_1 and C_1. This occurs if both branches have opposite conductance: $sC_1 = -\frac{1}{R_1}$, hence at a complex frequency $s = -\frac{1}{R_1 C_1}$. This is the frequency of a zero.

(b) For the circuit from Figure 18.42B, this situation occurs if $\frac{1}{sL_1} = -\frac{1}{R_1}$, or if $s = -\frac{R_1}{L_1}$.

(c) For the circuit from Figure 18.42C, no transfer is possible for $s = 0$. A capacitor in series with the signal path causes a zero-frequency zero.

(d) The circuit from Figure 18.42D has a parallel resonance circuit in series with the signal path. Here, signal transfer becomes zero if $\frac{1}{R_1} + \frac{1}{sL_1} + sC_1 = 0$. This results in two zeros, z_1 and z_2. They can be found from their product, $z_1 z_2 = \frac{1}{L_1 C_1}$ and their sum, $z_1 + z_2 = -R_1 C_1$. The two zeros are complex conjugated if the quality factor of the parallel resonator exceeds 0.5.

2. At some complex frequency, there exists a short in parallel with the signal path.

Example 18.17

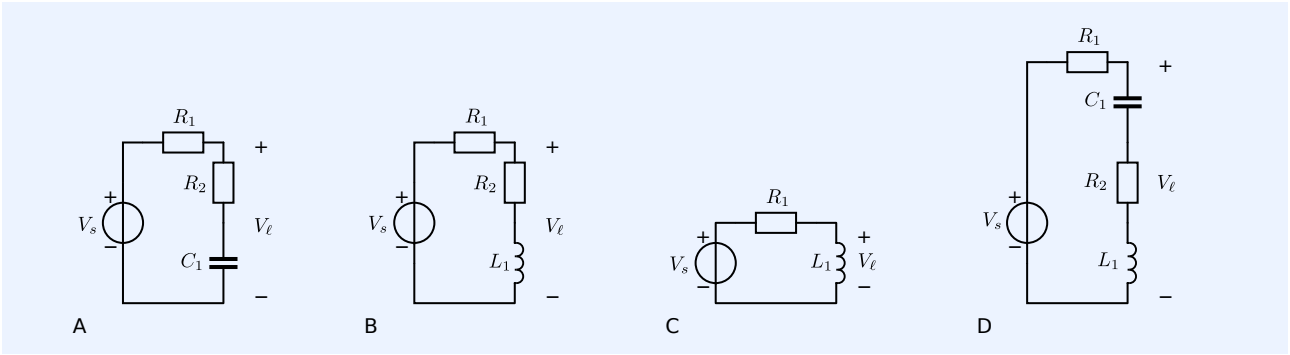


Figure 18.43: At some complex frequency, a short circuit in parallel with the signal path causes a zero.

(a) Let us now consider the transfer from V_s to V_ℓ for the circuit from Figure 18.43A. If, at some complex frequency, the series connection of R_2 and C_1 becomes a short; V_ℓ will then be zero. This occurs if $\frac{1}{sC_1} = -R_2$. Hence, a zero is thus found at $s = -\frac{1}{R_2 C_1}$

(b) Similarly, the circuit from Figure 18.43B has a zero at $s = -\frac{R_2}{L_1}$.

(c) The circuit from Figure 18.43C has a zero-frequency zero, caused by the inductor in parallel with the signal path.

(d) The circuit from Figure 18.43D has a series resonance circuit in parallel with the signal path. Here, signal transfer becomes zero if $R_2 + \frac{1}{sC_1} + sL_1 = 0$. This results in two zeros, z_1 and z_2. They can be found from their product, $z_1 z_2 = \frac{1}{L_1 C_1}$ and their sum, $z_1 + z_2 = -\frac{L_1}{R_2}$. The two zeros are complex conjugated if the quality factor of the series resonator exceeds 0.5.

3. At some complex frequency, multi-path source-load transfers cancel each other out.

Example 18.18

Figure 18.44 shows a dual-path input-output transfer. Both paths have their frequency-dependent transfer modeled by a ratio of two polynomials of the Laplace variable s. The two transfers cancel each other out if $\frac{N_1(s)}{D_1(s)} = -\frac{N_2(s)}{D_2(s)}$. The zeros are the solutions for s of this equation.

4. If a pole of a network cannot be observed in a specific transfer, then a zero cancels out that pole.

Example 18.19

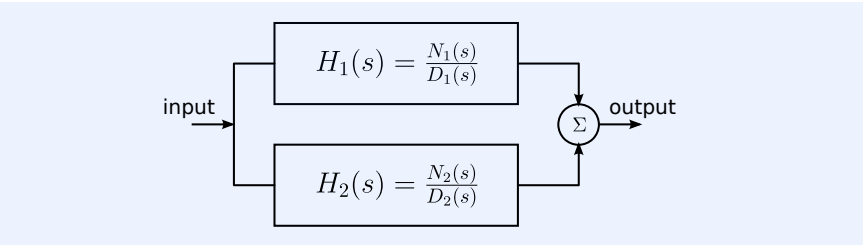


Figure 18.44: Zeros due to multiple transmission paths.

Consider the circuit from Figure 18.45. Let us estimate the poles and zeros of the transfer from I_1 to V_1. The poles of the circuit can easily be found from network inspection. The circuit has one independent inductor current and one independent capacitor voltage. Hence, it has two poles. There exists no loop of inductors (or of inductors and voltage sources) and no cut sets of capacitors (or capacitors and current sources). So, we have no poles in the origin.

With the aid of the time-constant method we find two time-constants $\tau_1 = R_1C_1$ and $\tau_2 = \frac{L}{R_2}$, and since there is no exchange of energy between the inductor and the capacitor, we have two real poles: $p_1 = -\frac{1}{\tau_1}$ and $p_2 = -\frac{1}{\tau_2}$.

At zero frequency, L_1 shorts the output so we have a zero $z_1 = 0$.

Let us now consider the transfer for I_1 to V_1. The current I_1 flows through the parallel connection of L and R_2 and circuit analysis yields

$$\frac{V_1}{I_1} = \frac{sL}{1 + s\frac{L}{R_2}}. \tag{18.139}$$

From this expression, we clearly see one pole and the zero in the origin. The pole caused by C is not observable in this transfer. So, the transfer from I_1 to V_1 has two poles and two zeros. One of the zeros cancels out one of the poles and makes it non-observable in the transfer function (18.139). A proper expression for the transfer would be

$$\frac{V_1}{I_1} = \frac{sL(1 + sR_1C)}{\left(1 + s\frac{L}{R_2}\right)(1 + sR_1C)}. \tag{18.140}$$

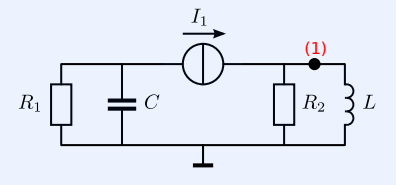


Figure 18.45: Circuit for estimating poles and zeros of the transfer $\frac{V_1}{I_1}$.

18.6 Two-ports

In many engineering situations, it is convenient to model an electrical network as a two-port. In those cases, the port quantities (voltage and current) of one port are related to the port quantities of the other port, and the electrical behavior of the two-port is described with the aid of a 2×2 matrix.

As is shown in the nodal analysis, a four-terminal network requires at least a 3×3 matrix description, thus a two-port description can only be complete under additional constraints. These constraints are called the two-port constraints. They require that the current that flows into one port terminal equals the current flowing out of the corresponding port terminal, and a voltage placed between one of the input port terminals and one of the output port terminals does not cause a change in the port voltages and currents.

The two-port constraints are always valid if both ports are terminated with one-ports, or if the four-terminal network is a natural two-port. Ideal transformers and gyrators, as well as the controlled sources are examples of nat-

ural two-ports. A formal derivation of the two-port conditions will be given in section 18.6.1.

With input and output port voltages and currents as port variables, there are six different combinations of dependent and independent port variables, resulting in six different two-port models. The choice of an appropriate two-port model strongly depends on the effect one wishes to study. In section 18.6.2, we will present the different representation methods. Finally, important classes of two-ports and their characteristic properties are listed in section 18.6.3.

18.6.1 Two-port conditions

[27] The two-port conditions are the conditions that should apply to a four-terminal network, such that its two-port equations correctly describe its behavior, for any external network.

We will now formally derive the two-port conditions.[27] To this end, we will set up the Y parameter representation of the four terminal network from Figure 18.46A. We will then define the differential-mode and common-mode port quantities and relate them to the nodal voltages and currents. This is illustrated in Figure 18.46B. Since the two-port model only describes relations between the differential-mode quantities, we then find the conditions under which the two-port model correctly describes the behavior of the four-terminal network.

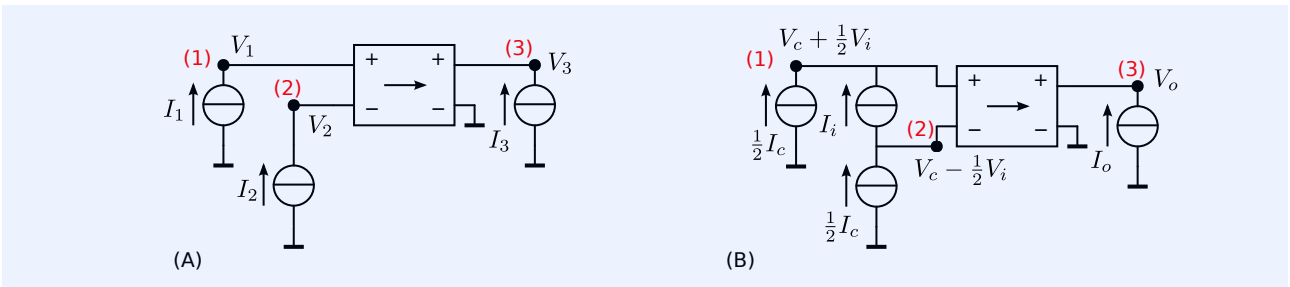


Figure 18.46: Two four-terminal network representations:
(A) Four-terminal network with nodal voltages and currents.
(B) Nodal voltages and currents decomposed into common-mode and differential-mode quantities. The differential-mode quantities are the two-port variables.

The four terminal network can be characterized by its Y parameters:

$$\begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix} = \begin{pmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{21} & Y_{22} & Y_{23} \\ Y_{31} & Y_{32} & Y_{33} \end{pmatrix} \cdot \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix}, \tag{18.141}$$

denoted in short as

$$\mathbf{I} = \mathbf{Y} \cdot \mathbf{V}. \tag{18.142}$$

[28] This is not essential: two-port conditions for a floating output port are similar to those for a floating input port.

In order to describe it as a two-port, we need to select the port variables. Since here the output port is connected to the ground[28], we select $V_o = V_3$ and $I_o = I_3$. Nodes 1 and 2 constitute the input port. The input port variables become the differential-mode port voltage and the differential-mode port current, $V_i = V_1 - V_2$ and $I_i = \frac{1}{2}(I_1 - I_2)$, respectively. The common-mode input port voltage and the common-mode input port current are defined as: $V_c = \frac{1}{2}(V_1 + V_2)$ and $I_c = I_1 + I_2$, respectively. These definitions are illustrated in Figure 18.46B.

The differential-mode port and common-mode currents can be expressed in the nodal currents with the aid of a base transformation matrix $\mathbf{A}$:

$$\begin{pmatrix} I_i \\ I_c \\ I_o \end{pmatrix} = \mathbf{A} \cdot \begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix}, \tag{18.143}$$

where:

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{18.144}$$

Equation (18.142) can now be written in the form:

$$\begin{pmatrix} I_i \\ I_c \\ I_o \end{pmatrix} = \mathbf{A} \cdot \mathbf{Y} \cdot \mathbf{V}. \tag{18.145}$$

Similarly, we may obtain the nodal voltages from the differential-mode port voltages from the nodal voltages as

$$\mathbf{V} = \mathbf{A}^T \cdot \begin{pmatrix} V_i \\ V_c \\ V_o \end{pmatrix}. \tag{18.146}$$

Substitution of this result in (18.145) yields the modified network equation:

$$\begin{pmatrix} I_i \\ I_c \\ I_o \end{pmatrix} = \mathbf{A} \cdot \mathbf{Y} \cdot \mathbf{A}^T \begin{pmatrix} V_i \\ V_c \\ V_o \end{pmatrix}. \tag{18.147}$$

Since matrix multiplication is associative, we may write

$$\begin{pmatrix} I_i \\ I_c \\ I_o \end{pmatrix} = \mathbf{Y}' \cdot \begin{pmatrix} V_i \\ V_{ic} \\ V_o \end{pmatrix}, \tag{18.148}$$

where:

$$\mathbf{Y}' = \mathbf{A} \cdot \mathbf{Y} \cdot \mathbf{A}^T. \tag{18.149}$$

In this way we obtain $\mathbf{Y}'$ as

$$\mathbf{Y}' = \begin{pmatrix} \frac{1}{4}Y_{11} - \frac{1}{4}Y_{12} - \frac{1}{4}Y_{21} + \frac{1}{4}Y_{22} & \frac{1}{2}Y_{11} + \frac{1}{2}Y_{12} - \frac{1}{2}Y_{21} - \frac{1}{2}Y_{22} & \frac{1}{2}Y_{13} - \frac{1}{2}Y_{23} \\ \frac{1}{2}Y_{11} - \frac{1}{2}Y_{12} + \frac{1}{2}Y_{21} - \frac{1}{2}Y_{22} & Y_{11} + Y_{12} + Y_{21} + Y_{22} & Y_{13} + Y_{23} \\ \frac{1}{2}Y_{31} - \frac{1}{2}Y_{32} & Y_{31} + Y_{32} & Y_{33} \end{pmatrix}. \tag{18.150}$$

The coefficients of $\mathbf{Y}'$ have the following meaning:

1. $\mathbf{Y}'_{1,1}$: Differential-mode input admittance.
2. $\mathbf{Y}'_{1,2}$: Common-mode input voltage into differential-mode input current conversion.
3. $\mathbf{Y}'_{1,3}$: Differential-mode reverse transadmittance.
4. $\mathbf{Y}'_{2,1}$: Differential-mode input voltage into common-mode input current conversion.
5. $\mathbf{Y}'_{2,2}$: Common-mode input admittance.
6. $\mathbf{Y}'_{2,3}$: Differential-mode output voltage into common-mode input current conversion.
7. $\mathbf{Y}'_{3,1}$: Differential-mode transadmittance.
8. $\mathbf{Y}'_{3,2}$: Differential-mode output voltage into common-mode input current conversion.
9. $\mathbf{Y}'_{3,3}$: Differential-mode output admittance.

Equation (18.150) can be reduced to a two-port equation, if the port voltages V_i, V_c and V_o do not cause any common-mode current I_c, and if the common-mode voltage V_c across the two-port does not affect the port currents I_i and I_o. These conditions are satisfied if $\mathbf{Y}'_{1,2}$, $\mathbf{Y}'_{2,1}$, $\mathbf{Y}'_{2,2}$, $\mathbf{Y}'_{2,3}$ and $\mathbf{Y}'_{3,2}$ are zero.

This yields the following set of equations:

$$\frac{1}{2}Y_{11} + \frac{1}{2}Y_{12} - \frac{1}{2}Y_{21} - \frac{1}{2}Y_{22} = 0 \tag{18.151}$$

$$\frac{1}{2}Y_{11} - \frac{1}{2}Y_{12} + \frac{1}{2}Y_{21} - \frac{1}{2}Y_{22} = 0 \tag{18.152}$$

$$Y_{11} + Y_{12} + Y_{21} + Y_{22} = 0 \tag{18.153}$$

$$Y_{13} + Y_{23} = 0 \tag{18.154}$$

$$Y_{31} + Y_{32} = 0 \tag{18.155}$$

The solution of this set of equations is:

$$Y_{11} = Y_{22},\ Y_{12} = Y_{21},\ Y_{21} = -Y_{22},\ Y_{13} = -Y_{23} \text{ and } Y_{31} = -Y_{32}. \tag{18.156}$$

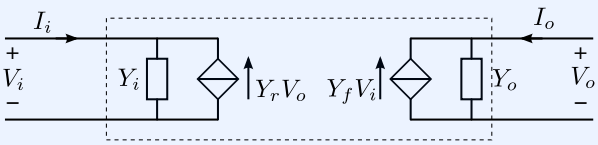


Figure 18.47: Four terminal element that satisfies the two-port conditions for arbitrary port connections. Such a circuit is referred to as a natural two-port.

Four-terminal networks that satisfy these conditions are called *natural two-ports*. Figure 18.47 shows such a network. When it is modeled as a four-terminal network according to Figure 18.46A, it satisfies the above conditions:

$$Y_i = Y_{11} = Y_{22} = -Y_{12} = -Y_{21}, \tag{18.157}$$

$$Y_o = Y_{33}, \tag{18.158}$$

$$Y_r = -Y_{13} = Y_{23}, \tag{18.159}$$

$$Y_f = Y_{31} = -Y_{32}. \tag{18.160}$$

18.6.2 Two-port representations

Linear four-terminal elements are often modeled as two-ports, one pair of terminals is considered as input port, while the other pair is considered as output port. Two-ports are then represented by matrices having only four coefficients. To do so, we select two independent variables from the four port quantities. In this way we obtain six different representation methods:

1. Z-matrix, or current-controlled representation:

$$\begin{pmatrix} V_i \\ V_o \end{pmatrix} = \mathbf{Z} \cdot \begin{pmatrix} I_i \\ I_o \end{pmatrix}. \tag{18.161}$$

2. Y-matrix, or voltage-controlled representation:

$$\begin{pmatrix} I_i \\ I_o \end{pmatrix} = \mathbf{Y} \cdot \begin{pmatrix} V_i \\ V_o \end{pmatrix}. \tag{18.162}$$

3. Hybrid-1 matrix, or current-controlled input and voltage-controlled out-

put representation:

$$\begin{pmatrix} V_i \\ I_o \end{pmatrix} = \mathbf{H} \cdot \begin{pmatrix} I_i \\ V_o \end{pmatrix}. \tag{18.163}$$

4. Hybrid-2 matrix, or voltage-controlled input and current-controlled output representation:

$$\begin{pmatrix} I_i \\ V_o \end{pmatrix} = \mathbf{H}' \cdot \begin{pmatrix} V_i \\ I_o \end{pmatrix}. \tag{18.164}$$

5. Transmission-1 matrix, or anti causal representation (the positive direction of the output current is opposite):

$$\begin{pmatrix} V_i \\ I_i \end{pmatrix} = \mathbf{T} \cdot \begin{pmatrix} V_o \\ -I_o \end{pmatrix}. \tag{18.165}$$

6. Transmission-2 matrix representation (the positive direction of the output current is opposite):

$$\begin{pmatrix} V_o \\ I_o \end{pmatrix} = \mathbf{T}' \cdot \begin{pmatrix} V_i \\ -I_i \end{pmatrix}. \tag{18.166}$$

Although these six representation methods are fully equivalent, in a particular situation, one specific representation can give more insight and reduce calculations considerably. If, for example, two-ports are connected in parallel, the Y parameters of the combination are easily found by adding the individual Y matrices. If, in another situation, the inputs are connected in series and the outputs are connected in parallel, the hybrid 1 representation is convenient. In situations in which two-ports are cascaded, the transmission-1 matrix representation is convenient: the transmission-1 matrix of the cascaded two-ports is the product of the transmission-1 matrices of the individual two-ports.

For design purposes, we will often use the anti-causal transmission-1 matrix representation (**T** representation). It will be shown that this representation is very convenient for deriving design strategies for amplifiers. As we have seen, nodal analysis uses the admittance **Y** representation.

Table 18.3 gives the relations between the two-port parameters of the six representations.

18.6.3 Two-port properties

Linear two-ports

A two-port is linear if both the properties of superposition and homogeneity hold:

1. The response to two excitations is equal to the sum of the responses to the individual excitations.
2. The response to an enlarged excitation is equal to the equally enlarged response to the excitation.

This means that a two-port is linear if it consists of linear elements only and apart from the excitations, it has no independent sources.

Reciprocal two-ports

A reciprocal two-port is a two-port from which the input port and the output port can be interchanged without affecting the network solution. This is the case if

$$z_{12} = z_{21},\ y_{12} = y_{21},\ \det(\mathbf{T}) = 1,\ h_{12} = -h_{21}. \tag{18.167}$$

	Z	**Y**	**T**	**T′**	**H**	**H′**
Z	$\begin{matrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{matrix}$	$\begin{matrix} \frac{y_{22}}{Y_\Delta} & -\frac{y_{12}}{Y_\Delta} \\ -\frac{y_{21}}{Y_\Delta} & \frac{y_{11}}{Y_\Delta} \end{matrix}$	$\begin{matrix} \frac{t_{11}}{t_{21}} & \frac{T_\Delta}{t_{21}} \\ \frac{1}{t_{21}} & \frac{t_{22}}{t_{21}} \end{matrix}$	$\begin{matrix} \frac{t'_{22}}{t'_{21}} & \frac{1}{t'_{21}} \\ \frac{T'_\Delta}{t'_{21}} & \frac{t'_{11}}{t'_{21}} \end{matrix}$	$\begin{matrix} \frac{H_\Delta}{h_{22}} & \frac{h_{21}}{h_{22}} \\ -\frac{h_{21}}{h_{22}} & \frac{1}{h_{22}} \end{matrix}$	$\begin{matrix} \frac{1}{h'_{11}} & -\frac{h'_{12}}{h'_{11}} \\ \frac{h'_{21}}{h'_{11}} & \frac{H'_\Delta}{h'_{11}} \end{matrix}$
Y	$\begin{matrix} \frac{z_{22}}{Z_\Delta} & -\frac{z_{12}}{Z_\Delta} \\ -\frac{z_{21}}{Z_\Delta} & \frac{z_{11}}{Z_\Delta} \end{matrix}$	$\begin{matrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{matrix}$	$\begin{matrix} \frac{t_{22}}{t_{12}} & \frac{T_\Delta}{t_{12}} \\ \frac{1}{t_{12}} & \frac{t_{11}}{t_{12}} \end{matrix}$	$\begin{matrix} \frac{t'_{11}}{t'_{12}} & -\frac{1}{t'_{12}} \\ -\frac{T'_\Delta}{t'_{12}} & \frac{t'_{22}}{t'_{12}} \end{matrix}$	$\begin{matrix} \frac{1}{h_{11}} & -\frac{h_{12}}{h_{11}} \\ \frac{h_{21}}{h_{11}} & \frac{H_\Delta}{h_{11}} \end{matrix}$	$\begin{matrix} \frac{H'_\Delta}{h'_{22}} & \frac{h'_{12}}{h'_{22}} \\ -\frac{h'_{12}}{h'_{22}} & \frac{1}{h'_{22}} \end{matrix}$
T	$\begin{matrix} \frac{z_{11}}{z_{21}} & \frac{Z_\Delta}{z_{21}} \\ \frac{1}{z_{21}} & \frac{z_{22}}{z_{21}} \end{matrix}$	$\begin{matrix} -\frac{y_{22}}{y_{21}} & -\frac{1}{y_{21}} \\ -\frac{Y_\Delta}{y_{21}} & -\frac{y_{11}}{y_{21}} \end{matrix}$	$\begin{matrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{matrix}$	$\begin{matrix} \frac{t'_{22}}{T'_\Delta} & \frac{t'_{12}}{T'_\Delta} \\ \frac{t'_{21}}{T'_\Delta} & \frac{t'_{11}}{T'_\Delta} \end{matrix}$	$\begin{matrix} -\frac{H_\Delta}{h_{21}} & -\frac{h_{11}}{h_{21}} \\ -\frac{h_{22}}{h_{21}} & -\frac{1}{h_{21}} \end{matrix}$	$\begin{matrix} h'_{11} & h'_{12} \\ h'_{21} & h'_{22} \end{matrix}$
T′	$\begin{matrix} \frac{z_{22}}{z_{12}} & \frac{Z_\Delta}{z_{12}} \\ \frac{1}{z_{12}} & \frac{z_{11}}{z_{12}} \end{matrix}$	$\begin{matrix} -\frac{y_{11}}{y_{12}} & -\frac{1}{y_{12}} \\ -\frac{Y_\Delta}{y_{12}} & -\frac{y_{22}}{y_{12}} \end{matrix}$	$\begin{matrix} \frac{t_{22}}{T_\Delta} & \frac{t_{12}}{T_\Delta} \\ \frac{t_{21}}{T_\Delta} & \frac{t_{11}}{T_\Delta} \end{matrix}$	$\begin{matrix} t'_{11} & t'_{12} \\ t'_{21} & t'_{22} \end{matrix}$	$\begin{matrix} \frac{1}{h_{12}} & \frac{h_{11}}{h_{12}} \\ -\frac{h_{22}}{h_{12}} & -\frac{H_\Delta}{h_{12}} \end{matrix}$	$\begin{matrix} -\frac{H'_\Delta}{h'_{12}} & -\frac{h'_{22}}{h'_{12}} \\ -\frac{h'_{11}}{h'_{12}} & \frac{1}{h'_{12}} \end{matrix}$
H	$\begin{matrix} \frac{Z_\Delta}{z_{22}} & \frac{z_{12}}{z_{22}} \\ -\frac{z_{21}}{z_{22}} & \frac{1}{z_{22}} \end{matrix}$	$\begin{matrix} \frac{1}{y_{11}} & -\frac{y_{12}}{y_{11}} \\ \frac{y_{21}}{y_{11}} & \frac{Y_\Delta}{y_{11}} \end{matrix}$	$\begin{matrix} \frac{t_{12}}{t_{22}} & \frac{T_\Delta}{t_{22}} \\ -\frac{1}{t_{22}} & \frac{t_{21}}{t_{22}} \end{matrix}$	$\begin{matrix} \frac{t'_{12}}{t'_{11}} & \frac{1}{t'_{11}} \\ -\frac{T'_\Delta}{t'_{11}} & \frac{t'_{21}}{t'_{11}} \end{matrix}$	$\begin{matrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{matrix}$	$\begin{matrix} \frac{h'_{22}}{H'_\Delta} & \frac{h'_{12}}{H'_\Delta} \\ \frac{h'_{21}}{H'_\Delta} & \frac{h'_{11}}{H'_\Delta} \end{matrix}$
H′	$\begin{matrix} \frac{1}{z_{11}} & -\frac{z_{12}}{z_{11}} \\ \frac{z_{21}}{z_{11}} & \frac{Z_\Delta}{z_{11}} \end{matrix}$	$\begin{matrix} \frac{Y_\Delta}{y_{22}} & \frac{y_{12}}{y_{22}} \\ -\frac{y_{21}}{y_{22}} & \frac{1}{y_{22}} \end{matrix}$	$\begin{matrix} \frac{t_{21}}{t_{11}} & -\frac{T_\Delta}{t_{11}} \\ \frac{1}{t_{11}} & \frac{t_{12}}{t_{11}} \end{matrix}$	$\begin{matrix} \frac{t'_{21}}{t'_{22}} & -\frac{1}{t'_{22}} \\ -\frac{T'_\Delta}{t'_{12}} & \frac{t'_{12}}{t'_{22}} \end{matrix}$	$\begin{matrix} \frac{h_{22}}{H_\Delta} & -\frac{h_{12}}{H_\Delta} \\ -\frac{h_{21}}{H_\Delta} & \frac{h_{11}}{H_\Delta} \end{matrix}$	$\begin{matrix} h'_{11} & h'_{12} \\ h'_{21} & h'_{22} \end{matrix}$

Table 18.3: Two-port transformations: all matrices in the same row are equal.
$X_\Delta = x_{11}x_{22} - x_{12}x_{21}$

Any linear two-port without controlled sources is reciprocal.

An amplifier generally needs to have a large gain from input to output and a negligible gain from output to input. Hence, in amplifier design, one tries to minimize the reciprocity.

Unilateral two-ports

Unilateral behavior means that any change of the port termination at one port cannot be noticed at the other port. The input impedance of a unilateral two-port does not depend on the load impedance, and vice versa. This is often a desirable property of amplifiers: unilateral amplifiers have no reverse transfer. A unilateral two-port has

$$\det(\mathbf{T}) = 0. \tag{18.168}$$

Nonenergic two-ports

A two-port is called nonenergic if it has no energy storage and no power dissipation. A nonenergic two-port has:

$$|\det(\mathbf{T})| = 1. \tag{18.169}$$

Examples of nonenergic two-ports are ideal transformers and gyrators.

19
Noise in electronic systems

19.1 Introduction

In any physical information-processing system, noise is added to the signals. Noise is a collective noun for all undesired current and voltage fluctuations in a circuit. Hence, DC offset voltages and currents including their temperature fluctuations, as well as uncertainty of device parameters including their temperature dependency and ageing effects, can be regarded as noise. All of these effects have in common that they increase the uncertainty that the circuit's current and voltage variations arise from information-carrying signals only. In this way, all these effects limit the amount of information that can be processed by the circuit.

In this section, we will introduce the main physical mechanisms for noise generation and introduce noise-modeling techniques for amplifiers.

19.1.1 Thermal noise

The noise caused by the thermal movement of electrons in conductive elements is called thermal noise or Johnson noise. In 1928, J.B. Johnson [Johnson1928][1] experimentally found that thermal noise variance was proportional to the absolute temperature. His colleague at the AT&T Bell labs, H. Nyquist [Nyquist1928-1][2] explained this. At constant temperature, the generation of thermal noise is a stationary and ergodic process.

The thermal noise in resistors can be modeled by a noise voltage source V_n in series with a noise-free resistor or a noise current source I_n in parallel with a noise-free resistor (see Figure 19.1).

[1] J.B. Johnson. Thermal agitation of electricity in conductors. *Phys. Rev.*, 32:97–109, 1928

[2] H. Nyquist. Thermal agitation of electrical charge in conductors. *Phys. Rev.*, 32:110–113, 1928

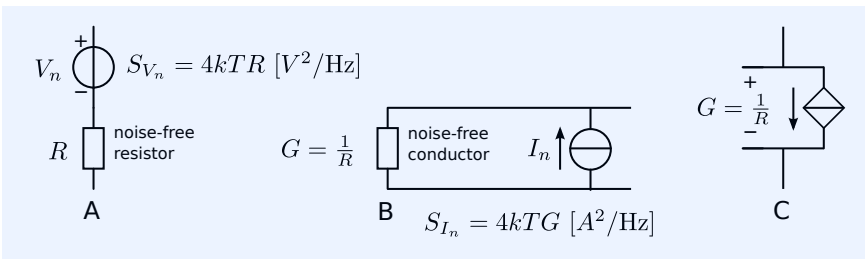


Figure 19.1: Models for noisy resistors (A and B) and for a noise-free resistor (C).

These thermal noise sources have a Gaussian amplitude distribution density function and a uniform or *white* power spectral density up to very high frequencies. The power spectral densities of V_n and I_n of a resistor with value R are given by:

$$S_{V_n} = 4kTR \qquad [\text{V}^2/\text{Hz}], \tag{19.1}$$

$$S_{I_n} = 4kTG \qquad [\text{A}^2/\text{Hz}], \tag{19.2}$$

where T represents the absolute temperature in K, and k the Boltzmann constant: $k = 1.38 \cdot 10^{-23}$ in J/K.

19.1.2 Shot noise

Variations in the transport of charge carriers across a potential barrier is the cause of so-called *shot noise*. Currents through PN junctions have an associated shot noise current. Shot noise current sources have a Gaussian distribution density function and a uniform spectral density, as long as the junction transit time is small with respect to the reciprocal value of the frequency. The spectral density of a shot noise current associated with a DC junction current I_J is given by

$$S_{I_n} = 2qI_J \qquad [\text{A}^2/\text{Hz}], \tag{19.3}$$

where q equals the electrical charge of the charge carrier: $q = 1.6 \cdot 10^{-19}$ C.

19.1.3 Excess noise

Fluctuations in conduction mechanisms give rise to so-called excess noise. Excess noise is found in resistors, in electrolytic capacitors and in semiconductor devices. The amplitude distribution function of excess noise sources is Gaussian, and the spectral density is inversely proportional to the frequency.

In resistors, the spectral density is usually modeled as proportional to the squared voltage V_R across the resistor:

$$S_{V_n} = K\frac{V_R^2}{f} \qquad [\mathrm{V}^2/\mathrm{Hz}], \tag{19.4}$$

where K is a material constant.

Excess noise in resistors is usually specified in μV/V/decade: the ratio of the RMS noise voltage v_n in μV and the DC voltage V_R, over a frequency range of one decade. The constant K can then be found by equating v_n^2 with the integral of the voltage noise spectral density over one decade:

$$\int_f^{10f} \frac{KV_R^2}{\nu} d\nu = v_n^2 \times 10^{-12} \qquad [\mathrm{V}^2], \tag{19.5}$$

from which we obtain

$$K = \frac{v_n^2}{V_R^2 \ln 10} \times 10^{-12} \qquad [\text{-}]. \tag{19.6}$$

The excess noise figure N is often specified in dB. In such cases, 1 μV/V/decade is referred to as 0 dB. The material constant K can be derived from N as

$$K = \frac{10^{\frac{N}{10}} \times 10^{-12}}{\ln 10} \qquad [\text{-}]. \tag{19.7}$$

The $1/f$ corner frequency f_ℓ is the frequency at which the spectral density of the excess noise equals that of the thermal noise:

$$\frac{KV_R^2}{f_\ell} = 4kTR. \tag{19.8}$$

It follows that

$$f_\ell = \frac{KV_R^2}{4kTR} = K\frac{P_R}{4kT} \qquad [\mathrm{Hz}], \tag{19.9}$$

in which P_R is the power dissipated in the resistor.[3]

[3] Excess noise in resistors is sometimes called *current noise*.

Example 19.1

We will evaluate K and f_ℓ for a low-noise metal film SMD 0402 resistor of 100Ω *with a specified excess noise of* $0.03\mu V/V$*, operating at a DC voltage of* $1.5V$ *and at a temperature* $T = 300K$.

With the aid of expression 19.6, we obtain $K = \frac{0.03^2 \times 10^{-12}}{\ln 10} = 3.91 \times 10^{-16}$.

The power dissipation in the resistor P_R *equals* $1.5^2/100 = 22.5mW$. *With the aid of expression 19.9, we obtain* $f_\ell = \frac{0.0225 \times 3.91 \times 10^{-16}}{4 \times 1.38 \times 10^{-23} \times 300} = 531Hz$.

In semiconductors, excess noise is caused by energy traps in the depletion layer of PN junctions or in the oxide of MOS structures. For PN junctions, it is modeled with a noise current source I_n in parallel with the junction. The

spectral density S_{I_n} of I_n is

$$S_{I_n} = \frac{K I_J^{\alpha}}{f} \qquad [\mathrm{A}^2/\mathrm{Hz}], \tag{19.10}$$

in which K and α are material constants and I_J is the DC current through the junction.

19.1.4 Noise temperature

The noise power of a source is sometimes defined with the aid its noise temperature. The noise temperature T_n is defined as

$$T_n = \frac{P}{kB}, \tag{19.11}$$

[4] The noise power dissipated in a load of which the impedance is the complex conjugated of that of the noise source.

in which P [W] is the available noise power[4], B [Hz] is the total bandwidth over which the noise power is measured, and k [J/K] is the Boltzmann constant.

19.2 Noise-modeling in two-ports

From the four port variables (see Chapter 18.6), two can be selected as independent variables. This results in six equivalent noise models for two-ports. They are shown in Figure 19.2.

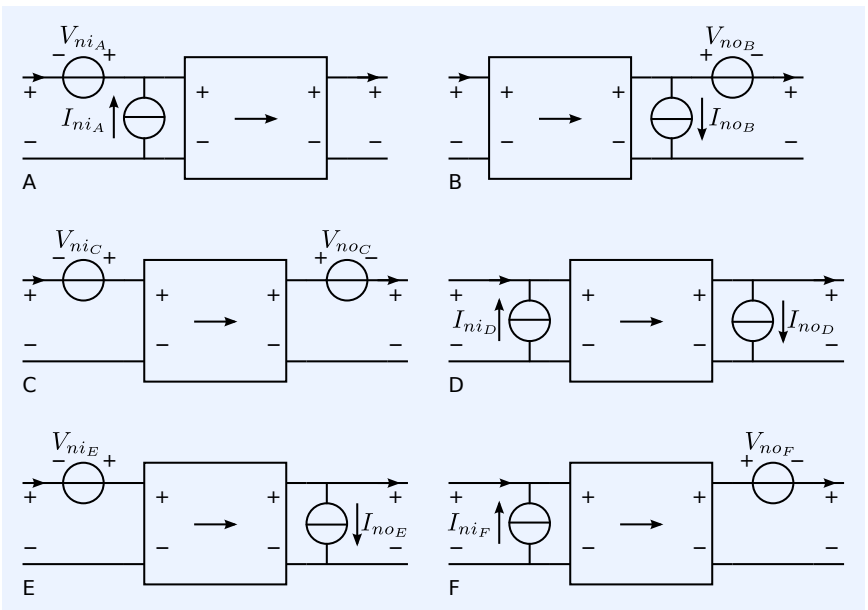


Figure 19.2: Six ways to model a noisy two-port. The positive directions for the signs of the port voltages (V_i, V_0) and of the port currents (I_i, I_o) have been indicated by means of pus and minus signs and arrows at the port terminals, respectively.

In the example below, we will demonstrate the transformation of the noise representation according to Figure 19.2A into the representation from Figure 19.2C.

Example 19.2

The two-port equations for the model according to Figure 19.2A are

$$\begin{pmatrix} V_i + V_{ni_A} \\ I_i + I_{ni_A} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_o \\ I_o \end{pmatrix}. \tag{19.12}$$

We eliminate I_{ni_A} from the input current vector by subtracting it from the current

equation (row 2):

$$I_i = CV_o - I_{ni_A} + DI_o. \tag{19.13}$$

We then define the new output voltage of the noisy two-port as $V_o - I_{ni_A}\frac{1}{C}$*, and write expression 19.13 as*

$$I_i = C\left(V_o - I_{ni_A}\frac{1}{C}\right) + DI_o. \tag{19.14}$$

We then substitute this new output voltage into the voltage equation (row 1). This changes the voltage equation to

$$V_i + V_{ni_A} = A\left(V_o - I_{ni_A}\frac{1}{C}\right) + \frac{A}{C}I_{ni_A} + BI_o. \tag{19.15}$$

After bringing the voltage $\frac{A}{C}I_{ni}$ *from the right side of (19.15) to the left side of this equation, we obtain the corrected input voltage:*

$$V_i + V_{ni_A} - \frac{A}{C}I_{ni_A} = A\left(V_o - I_{ni_A}\frac{1}{C}\right) + BI_o. \tag{19.16}$$

Equations (19.14) and (19.16) are the new two-port equations:

$$\begin{pmatrix} V_i + V_{ni_A} - \frac{A}{C}I_{ni_A} \\ I_i \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} V_o - I_{ni_A}\frac{1}{C} \\ I_o \end{pmatrix}, \tag{19.17}$$

from which we find the equivalent noise voltage sources according to the representation in Figure 19.2C:

$$V_{ni_C} = V_{ni_A} - \frac{A}{C}I_{ni_A}, \tag{19.18}$$

$$V_{no_C} = -I_{ni_A}\frac{1}{C}. \tag{19.19}$$

Note: if V_{ni_A} *and* I_{ni_A} *are uncorrelated, then* V_{ni_C} *and* V_{no_C} *are partially correlated.*

19.3 Noise performance characterization

In this section, we will describe some quantities that are used for the description of the noise behavior of information processing systems.

19.3.1 Signal-to-noise ratio

The signal-to-noise ratio SNR, of a signal perturbed by noise is defined as the ratio of the signal power and the noise power , and is often interpreted as a quality measure for the signal:

$$SNR = \frac{P_{signal}}{P_{noise}}\ [\text{-}], \tag{19.20}$$

$$(SNR)_{\text{dB}} = 10\log_{10}\left(\frac{P_{signal}}{P_{noise}}\right)\ [\text{dB}]. \tag{19.21}$$

19.3.2 Dynamic range

A quality measure for amplifiers that is closely related to the channel capacity is the dynamic range D of an amplifier. It is defined as the ratio of the

maximum signal power $P_{signal,\max}$ at which the retrieval of the information is still possible, and the noise power $P_{noise,\min}$ in absence of a signal:

$$D = 10\log_{10}\left(\frac{P_{signal,\max}}{P_{noise,\min}}\right) \text{ [dB]}. \tag{19.22}$$

The maximum power $P_{signal,\max}$ that is used in this definition differs for each application. The Intermodulation-Free Dynamic Range $IMFDR$ of a low-noise RF amplifier is based on the output power level where the power of the intermodulation components equals that of the noise. In audio amplifiers, the dynamic range is measured at an output power for a given percentage of total harmonic distortion.

19.3.3 Noise figure

The noise figure NF [dB] or F [-] of an amplifier tells us something about the deterioration of the signal-to-noise ratio by the amplifier. NF and F are defined as:

$$NF = 10\log F, \tag{19.23}$$

$$F = \frac{S/N \text{ at the input of the amplifier}}{S/N \text{ at the output of the amplifier}}. \tag{19.24}$$

The noise figure is usually defined at a temperature of 290K.

19.3.4 Equivalent noise bandwidth

The equivalent noise bandwidth B_n of a system with a transfer function $H(j\omega)$ is defined as the bandwidth of a brickwall filter with a pass band gain equal to the maximum magnitude of $H(j\omega)$, that would produce the same output noise power as $H(j\omega)$:

$$B_n = \frac{1}{2\pi}\int_0^\infty \left|\frac{H(j\omega)}{H_{\max}}\right|^2 d\omega \text{ [Hz]}. \tag{19.25}$$

Example 19.3

We will evaluate the noise bandwidth of a first order low-pass filter of which the transfer $H(j\omega)$ can be written as

$$H(j\omega) = \frac{H_0}{1+j\omega\tau}. \tag{19.26}$$

The maximum value of $|H(j\omega)|$ is H_0, from which we obtain

$$B_n = \frac{1}{2\pi}\int_0^\infty \left|\frac{1}{1+j\omega\tau}\right|^2 d\omega \tag{19.27}$$

$$= \frac{1}{2\pi\tau}\int_0^\infty \frac{1}{1+\omega^2\tau^2} d\omega\tau$$

$$= \frac{1}{4\tau} \tag{19.28}$$

$$= \frac{\pi}{2} B_{-3dB}. \tag{19.29}$$

Hence, the noise bandwidth of a first order low-pass filter is $\pi/2$ times larger than its $-3dB$ bandwidth in Hz.

Bibliography

[AlexanderDerek1990] M. Alexander and D.F. Bowers. SPICE-compatibe Op Amp Macro-Models. *EDN*, February, March 1990.

[Altshuller1999] Altshuller. *The Innovation Algorithm, TRIZ, systematic innovation and technical creativity, Translated by Lev Shulyak and Steven Rodman*. Technical Innovation Center, Inc., USA, 1999. ISBN: 0-9640740-4-4.

[BSIM3-1995] Cheng, Y. et al. *BSIM3 Version 3.0 Manual*. University of California/Berkeley, Electronics Research Laboratory, 1995.

[BSIM4-2017] Chetan Kumar Dabhi, Shivendra Singh Parihar, Harshit Agrawal, Navid Paydavosi, Tanvir Hasan Morshed, Darsen D. Lu, Wenwei(Morgan)Yang, Mohan V. Dunga, Xuemei (Jane) Xi, Jin He, Weidong Liu, Kanyu, M. Cao, Xiaodong Jin, Jeff J. Ou, Mansun Chan, Yogesh Singh Chauhan, Ali M. Niknejad, Chenming Hu. *BSIM4 4.8.1 MOSFET Model*. University of California/Berkeley, 2017.

[Binkley2008] Binkley, David M. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons Inc., 1997. ISBN: 978-0-470-03136-0.

[Black1932] H.S. Black. Wave Translation System, March 1932.

[Black1934] Harold S. Black. Stabilized feed-back amplifiers. *Electrical Engineering*, 53(1):114–120, January 1934.

[Blakesley1994] T. A. Blakesley. A New Electrical Theorem. *Proc. Phys. Soc. London*, 13:65–67, 1994.

[BlanchardEA1998] Benjamin S. Blanchard and Wolter J. Fabrycky. *Systems engineering and analysis*. Prentice-Hall, Inc, New York, 1981. ISBN: 013-135047-1.

[Bode1945] H.W. Bode. *Network Analysis and Feedback Amplifier Design*. Van Nostrand, New York, 1945.

[Boyle1974] Boyle, et al. Macromodeling of Integrated Circuit Operational Amplifiers. *IEEE Journal of solid-state circuits*, 9(6), December 1974.

[Chen1991] W.K. Chen. *Active Network Analysis*. World Scientific Publishing Co. Pte. Ltd., Singapore, 1991. ISBN: 9971-50-912-1.

[Chua1975] Leo O. Chua and Pen-Min Lin. *Computer aided analysis of electronic circuits*. Prentice-Hall, Inc., USA, 1975. ISBN: 0-13-165415-2.

[Chua1987] Charles A. Desour Leo O. Chua and Ernest S. Kuh. *Linear and Nonlinear Circuits*. McGraw-Hill, Inc., Singapore, 1987. ISBN: 0-07-010898-6.

[Cross1989] Nigel Cross. *Engineering Design Methods, Strategies for Product Design*. John Wiley & Sons Ltd., England, 1989. ISBN: 0-471-94228-6.

[DaceyRoss1952] G.C. "Dacey and I.M." Ross. Unipolar "Field-Effect" Transistor. *Proc. of the I.R.E*, pages 970–979, August 1953.

[Dorf1995] Richard C. Dorf and Robert H. Bishop. *Modern Control Systems*. Addison-Wesley Publishing Company, Inc., USA, 1995. ISBN: 0-201-84559-8.

[Early1952] J.M. Early. Effects of space-charge layer widening in junction transistors. *Proceedings IRE*, 40:1401–1406, 1952.

[EbersMoll1954] J.J. Ebers and J.L Moll. Large-signal behavior of junction transistors. *Proceedings of the IRE*, 42(12):1761–1772, 1954.

[EnzVittoz2006] Christian C. Enz, and Eric A. Vittoz. *Charge-based MOS Transistor Modeling*. John Wiley & Sons Inc., 2006. ISBN: 978-0-470-85541-6.

[Evans1948] Walter R." "Evans. Graphical Analysis of Control Systems. *Trans. AIEE*, 67(1):547–551, January 1948.

[Evans1950] Walter R." "Evans. Control Systems Synthesis by Root Locus Method. *Trans. AIEE*, 69(1):66–69, January 1950.

[Forest1907] L. De Forest. Wireless Telegraphy, Januari 1907.

[Foty1997] Daniel P. Foty. *MOSFET modeling with SPICE: principles and practice*. Prentice-Hall, Inc., USA, 1997. ISBN: 0-13-227935-5.

[Friis1944] H.T. Friis. Noise Figures of Radio Receivers. *Proceedings of the IRE*, 32:419–422, February 1944.

[GummelPoon1970] H. K. Gummel and H. C. Poon. An integral charge control model of bipolar transistors. *Bell Syst. Tech. J.*, 49(5):827–852, May-June 1970.

[HSPICE1993] *HSPICE User's Manual*. Campbell, California, 1993.

[Haley1988] Stephen B. Haley. The Generalized Eigenproblem: Pole-Zero Computation. *Proceedings of the IEEE*, 76(2):103–120, February 1988.

[Haley1989] Stephen B. Haley and Paul J. Hurst. Pole and Zero Estimation in Linear Circuits. *IEEE Transactions on Circuits and Systems*, 36(6):838–845, June 1989.

[Johnson1928] J.B. Johnson. Thermal agitation of electricity in conductors. *Phys. Rev.*, 32:97–109, 1928.

[KahngDawon1960] Kahng and Dawon. Electric Field Controlled Semiconductor Device, May 1960.

[Kirk1962] C. T. Kirk. A theory of transistor cut-off frequency (f_t) falloff at high current densities. *IEEE Trans. Electron Devices*, ED-9(2):164–174, February 1962.

[KossiakoffEA2003] Alexander Kossiakoff and William N. Sweet. *Systems engineering principles and practice*. John Wiley and Sons, Inc., New Jersey, 2003. ISBN: 0-471-23443-5.

[Lilienfeld1925] Julius Edgar Lilienfeld. Method and apparatus for controlling electric currents, October 1925.

[MOS9] D. Velghe, R. Klaassen and F. Klaassen. MOS Model 9, 1994. Unclassified Report NL-UR 003/94.

[McDermottEA1996] Robin E. McDermott, Raymond J. Mikulak, Michael R. Beauregard. *The basics of FMEA*. Productivity Inc., Portland, USA, 1996. ISBN: 0-527-76320-9.

[Meyer1971] J. Meyer. MOS models and circuit simulation. *RCA Review*, 32:42–63, 1971.

[Middlebrook1975] R.D. Middlebrook. Measurement of loop gain in feedback systems. *Int. J. Electronics*, 38:485–512, April 2001.

[Middlebrook1991] R.D. Middlebrook. Low-Entropy Expressions: The Key to Design-Oriented Analysis. *Frontiers in Education Conference*, pages 399–403, 1991.

[Miller1920] John M. Miller. Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit. *Scientific Papers of the Bureau of Standards*, 15(351):367–385, 1920.

[Monticelli1986] D. M. Monticelli. A quad CMOS single-supply opamp with rail-to-rail output swing. *IEEE Journal of Solid-State Circuits*, 21:1026–1034, December 1986.

[Nagel-M382] Laurence W. Nagel and D.O. Pederson. Technical Report UCB/ERL M382, EECS Department, University of California, Berkeley, April 1973.

[Nagel-M520] Laurence W. Nagel. *SPICE2: A Computer Program to Simulate Semiconductor Circuits*. PhD thesis, EECS Department, University of California, Berkeley, 1975.

[Nordholt1983] Ernst H. Nordholt. *Design of High-Performance Negative Feedback Amplifiers*. Delft Academic Press / VSSD, 1 edition, 1983-2006. ISBN: 9789040712470.

[Nyquist1928-1] H. Nyquist. Thermal agitation of electrical charge in conductors. *Phys. Rev.*, 32:110–113, 1928.

[Nyquist1928-2] H. Nyquist. Certain topics in telegraph transmission theory. *Trans. Am. Inst. Elec. Eng.*, 47:617–644, 1928.

[Nyquist1932] H. Nyquist. Regeneration theory. *Bell System Technical Journal*, 11:126–147, 1932.

[Papoulis1965] Athanasios Papoulis. *Probability Random Variables and Stochastic Processes*. McGraw-Hill, New York, 1965.

[Quarles-M89/42] Thomas L. Quarles. *Analysis of Performance and Convergence Issues for Circuit Simulation*. PhD thesis, EECS Department, University of California, Berkeley, 1989.

[Rosenstark1974] Solomon Rosenstark. A Simplified Method of Feedback Amplifier Analysis. *Transactions on education*, E-17(4):192–198, November 1974.

[SPICE1988] P. Antognetti and G. Massobrio. *Semiconductor Device Modeling with SPICE*. McGraw-Hill, 1988.

[ShannonWeaver1963] Claude E. Shannon and Warren Weaver. *The Mathematical Theory of Communication*. The University of Illinois Press, Ubana, 1 edition, 1963. ISBN: 0-88179-205-5.

[SheuEtAll1987] B. Sheu, D. Scharfetter, P. Ko, and M. Jeng. BSIM: Berkely Short Channel IGFET Model for MOS Transistors. *IEEE Journal of Solid State Circuits*, SC-22:558–566, 1987.

[ShichmanHodges1968] H. Shichman and D. Hodges. Modeling and simulation of insulated-gate field-effect transistor switching circuits. *IEEE J. Solid-State Circuits*, 3(3):285–289, September 1968.

[Shockley1952] William Shockley. Circuit element utilizing semiconductive material, June 1948.

[ShockleyFET1952] W. Shockley. A Unipolar "Field-Effect" Transistor. *Proc. of the I.R.E*, pages 1365–1376, November 1952.

[Terninko1998] Terninko et al. *Systematic Innovation, An introduction to TRIZ*. CRC Press LLC, USA, 1998.

[Torralba2001] A. Torralba, R. G. Carvajal, J. Ramirez-Angulo, J. Tombs, and T. Galan. Class ab output stages for low voltage cmos opamps with accurate quiescent current control by means of dynamic biasing. In *ICECS 2001. 8th IEEE International Conference on Electronics, Circuits and Systems (Cat. No.01EX483)*, volume 2, pages 967–970, 2001.

[Truxal1972] John G. Truxal. *Introductory System Engineering*. McGraw-Hill, Tokyo, 1972.

[VBIC1996] Colin C. McAndrew, Jerold A. Seitchik, Derek F. Bowers, Mark Dunn, Mark Foisy, Ian Getreu, Marc McSwain, Shahrir Moinian, James Parjer, David J. Roulston, Michael Schroeter, Paul van Wijnen, Lawrence F. Wagner. VBIC95, The Vertical Bipolar Inter-Company Model. *IEEE Journal of solid-state circuits*, 31(10):1476–1483, October 1996.

[Verhoeven2003] C.J.M. Verhoeven, A. van Staveren, G.L.E. Monna, M.H.L. Kouwenhoven and E. Yildiz. *Structured Electronic Design*. Kluwer Academic Publishers, Boston - Dordrecht - London, 2003. ISBN: 1-4020-7590-1.

[Wanlass-Sah1963] F. "Wanlass and C." Sah. Nanowatt logic using field-effect metal-oxide semiconductor triodes. *1963 IEEE International Solid-State Circuits Conference. Digest of Technical Papers*, VI:32–33, 1963.

[Ward1998] D. Ward and R. Dutton. A Charge-Oriented Model for MOS Transistor Capacitances. *IEEE Solid-state Circuits*, sc-13(5):703–708, October 1978.

[spiceEKV2.6] Matthias Bucher, Christoff Lallement, Christian Enz, Fabien Théodoloz and François Krummenacher. The EPFL-EKV MOSFET Model Equations for Simulation. Technical report, Electronics Laboratory, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, July 1998.

Index

Structured Electronics Design -

A conceptual approach to amplifier design - 3rd ed.

Anton J.M. Montagne

Many people consider analog electronic circuit design complex. This is because theoretical concepts, circuit topologies, electronic devices, their operating conditions, and the system's physical construction constitute an enormous design space in which it is easy to get lost. For this reason, analog electronics often is regarded as an art rather than a solid discipline.

Structured Electronics Design:

- Defines a step-by-step hierarchically organized design process.
- Is based on solid principles from systems engineering, physics, signal processing, control theory, and network theory.
- Provides a solid foundation for circuit design education and automation.
- Has been developed at the TU Delft since the 1980s.

Anton J.M. Montagne TU Delft | Fac. EEMCS

Anton Montagne (Leiden, The Netherlands, 1953) received his master's degree in 1984 at the Delft University of Technology. In 1983, he joined Philips Semiconductors in Nijmegen where he designed analog integrated circuits for audio and video applications. At Philips, he also developed training courses on analog electronics. In 1989, together with Catena Microelectronics, the Delft University of Technology and the Institute of Microelectronics in Stuttgart, he cooperated in the development of an intensive training course, covering many topics of analog circuit design. Since 1997, he works as an independent consultant, trainer and designer in the field of analog circuit design. Over the past 38 years, he developed analog electronics for instrumentation and communication systems and carried out many training courses on analog electronics. Since 2016 Anton Montagne is coaching students and giving lectures and masterclasses "Structured Electronics Design" at the Delft University of Technology.

TU Delft

ISBN 978-94-6366-712-8
DOI: https://doi.org/10.59490/t.2023.005

textbooks.open.tudelft.nl