Theory of electronic circuits

Book · January 2013

CITATIONS
3

READS
17,514

2 authors:

Jitka Mohylová
VSB - Technical University of Ostrava
45 PUBLICATIONS **243** CITATIONS
SEE PROFILE

Josef Punčochář
VSB - Technical University of Ostrava
120 PUBLICATIONS **155** CITATIONS
SEE PROFILE

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Vysoká škola báňská – Technická univerzita Ostrava

THEORY OF ELECTRONIC CIRCUITS

Students book

Jitka Mohylová, Josef Punčochář

Ostrava 2013

Recenze:

Název: Theory of electronic circuits
Autor: Jitka Mohylová, Josef Punčochář
Vydání: první, 2013
Počet stran:
Náklad:

Studijní materiály pro studijní obor ..... fakulty FEI
Jazyková korektura: Mgr. Hana Vondrová

Určeno pro projekt:

Název:
Číslo:
Realizace: VŠB – Technická univerzita Ostrava
Projekt je spolufinancován z prostředků ESF a státního rozpočtu ČR

ISBN

CONTENTS

Instructions to stady

THEORY OF ELECTRONIC CIRCUITS

For the subject of Theory of Electronic circuits you received a studying package containing:

- integrated textbook for distance study with instructions for studying
- schedule for the course during semester and schedule of full-time studying
- dividing of students into groups to individual tutors and contacts to tutors
- contact to the study department

Prerequisites

For the study of this subject there is considered passing the following subjects: Electronic circuits I, Electronic circuits II and Electronic circuits III.

Course purpose and goals

After completing this course, the student should be able to analyze and construct electronic circuits (amplifiers, filters, oscillators, converters, etc.) with real active circuit elements: Operational amplifier; Transconductance amplifier; Transimpedance amplifier; Current conveyor; Analog multiplier.

For whom is the subject intended

The module is included into master studying of branch: Biomedical Engineering, Measurement and Control Engineering and Electronics, but it can be studied by anyone interested who is from any different branch provided he meets the needed prerequisites or he completed a course similar with content.

The textbook is divided into parts, chapters, which correspond to the logical division of substance, but they are not so extensive. Estimated time for studying of a chapter can vary greatly, therefore large chapters are divided into numbered subchapters and the structure described below corresponds to them.

REQUIREMENTS FOR ACCREDITATION

- Student must work out three tests. Student can receive up to 3 points for each of tests. Maximum number of points student can gain through tests is 3 * 3 = 9 points.
- Student can gain up to 12 points from the laboratory exercises (6 * 2 = 12 - 2 points each problem).
- Student can gain up to 9 points from the computer exercises (3 * 3 = 9 - 3 points each problem).
- Student can gain up to 14 points from the project.
- To pass the exercises part of course student has to gain at least 20 points.
- Student can gain up to 56 points from the final exam (circuit work - 32 points, oral part - 24 points).

- To pass the course student has to gain at least 51 points.

During the studying of each chapter we recommend the following procedure:

Time of study: xx hours

At the beginning of a chapter there is a given time for studying the matter. That time is approximate and can help like a guide to the layout of studying of the whole subject or of the chapter. It may seem too long for someone and vice versa. There are students who have never seen this problematic and those who have already experienced a lot in this field.

Goals: After studying this chapter you will be able to

- describe ...
- define ...
- solve ...

Right after that there are stated goals, which you should meet after studying the chapter – specific knowledge and skills.

EXPLANATION

Then it is followed by the explanation of the matter studied itself, introduction of new concepts and their explanation. Everything is followed by images, charts, solved examples and links to animations.

Solving problems

Summary

At the end of each chapter there are revised the main concepts which you should learn. If you do not understand any of these, go back to them again.

Questions

To verify that you well and fully understand the chapter material, there are several theoretical questions.

Problems

Because most of the theoretical concepts of this subject have immediate importance and use in practice, practical tasks to deal with are present to you at the end. The main importance of this subject is in them. The main goal of the whole subject is to apply the newly learnt knowledge in solving of real situations.

Example

Basic text

[1] Mohylová, J, Punčochář,J.: .....

Other text

[2]

[3]

[4]

Problems key 7

The results of the exercises given and theoretical questions are at the summary of the textbook in the Key. Use it only after solving the exercises by yourselves, only by this self-checking you will verify if you mastered the content of the chapter fully.

Authors of this educational material wish you successful and enjoyable time with this textbook!

Jitka Mohylová & Josef Punčochář

1. History of electronics, linearization

Time of study: 6 hours

Goals:

- basic electronics history
- basic models of amplifier structures
- quiescent points
- nonlinear and linear elements, basic parameters
- linearization in the quiescent point

EXPLANATION

1.1. History

The op amp has become the icon of analog electronics. This element has a long history, where the term op-amp was originally used to describe a chain of high performance dc amplifiers that was used as a basis for the construction of analog computers – to a realization of math operations.

Although the voltage OPA is more than sixty years old, it is still young – see figures and Table of electronics history below. The modern OPA series provides system, instrument, and equipment designers with an unprecedented combination of dc precision and wideband performance.

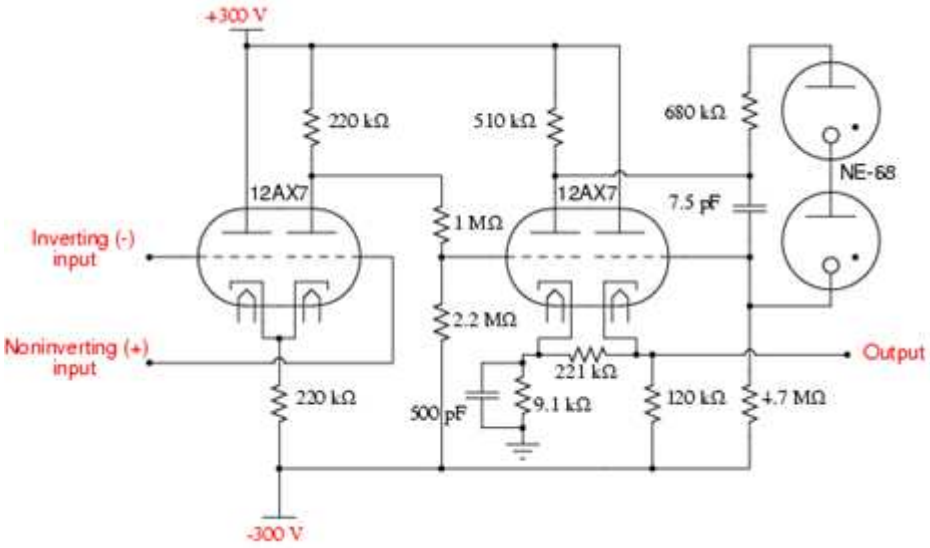


Fig. 1.1 Schematic diagram of the Philbrick Researches op-amp, model K2 - W

AD – anno Domini; AOPA – anno OP-AMP; BOPA – before OP-AMP

YAERS	OPA -years	text
1883 AD	**≈ 58 BOPA**	**Edison** discovers the Edison Effect in which a hot filament in a vacuum emits electrons to adjacent conductor.
1904 AD	**≈ 37 BOPA**	Prof. John Ambrose **Fleming** patents the two-element thermionic valve detector based upon the Edison Effect (**diode;** 1874 – F. Braun – point contact detector).
1906 AD	**≈ 37 BOPA**	Lee De **Forest** adds a grid to Fleming valve and produces the first **triode** vacuum tube (audion – later called triode) - **the beginning of electronics**.
1915 AD	**≈ 26 BOPA**	**Armstrong** - the first published explanation of how **positive feedback** could be used to increase greatly the voltage gain of amplifiers.
1925 AD	**≈ 16 BOPA**	Julius **Lilienfield** – the first patent for a **FET – like** transistor – but he never constructed a working device (1952 Shockley: unipolar field-effect transistor).
1927 AD	**≈ 14 BOPA**	Harold S. **Black -** the idea of the **negative feedback** amplifier.
1940 AD	???????	* Electronic **analog computer** developed by D. B. Parkinson and C. A. Lovell; * Russel **Ohl** – **PN junction**
≈1941 AD	**≈ 0 OPA**	In these days **operational amplifiers** (tube) **did exist**. They were extremely large and slow, not commercially available and **not named**.
≈1942 AD	**≈ 1 AOPA**	**Loebe** Julie – modular op amp using two twin triodes, **gain bandwidth product** about 100 kHz. Ragazzini published a paper after the war (1946 – 1947) explaining how HE designed the op amp. Mr Loebe was listed among the contributors only.
1947 -1952 AD	**≈ 6 - 11 AOPA**	John **Bardeen**, Walter **Brattain**, William **Shockley**: TRANSient resISTOR - **TRANSISTOR**. The transistor is an influential invention that changed the course of history for op amps, thus electronics.
1952 *	**≈ 11 AOPA**	The **first commercial** general purpose differential OP AMP was introduced to the market by G. A. **Philbrick** Researches: **K2 – W;** **DC gain A_u = 20 000**; gain- bandwidth product about 1 MHz; power consumption about 4,5 W; price 24 dollars.
1954 AD	**≈ 13 AOPA**	* **Teal, Buehler** (Bell Labs) – **single crystal silicon**; * **Texas Instruments** – introduces commercial **silicon junction transistor.**
1958 AD	**≈ 17 AOPA**	**Texas Instruments** and **Fairchild** announce development of the first **INTEGRATED CIRCUITS (silicon)**.
1962 AD	**≈ 21 AOPA**	The very first series of modular **solid state op amps** were introduced by **BURR-BROWN** Research Corporation and G. A. **Philbrick** Inc.
1963 AD	**≈ 22 AOPA**	The first op amp offered to the public was **μA702** made by **Fairchild Semiconductors.**
1965 AD	**≈ 24 AOPA**	The well known **μA709** (R.Widlar) was made by **Fairchild Semiconductors** (CZECH REPUBLIC – TESLA ROŽNOV – as MAA502 – about 1969)
------->		1967 - **LM101**; 1968 - **μA741;** 1975 - **OP07;** 1979 - **LF155/6/7; ..... etc.**
↓ 2008	**≈ 67 AOPA**	Today, the system designer can choose from numerous different kinds of what can be called traditional monolithic op amps, where the amplification factor A is usually reckoned to be extremely large. We will call this sort of amplifiers the voltage OPA. Today, the system designer can also choose from numerous other types of amplifier structures.

Tab. 1: Table of electronics history.

Fig. 1.2 OP AMP K2 - W

Controlled sources – are often used in models of electronic elements.

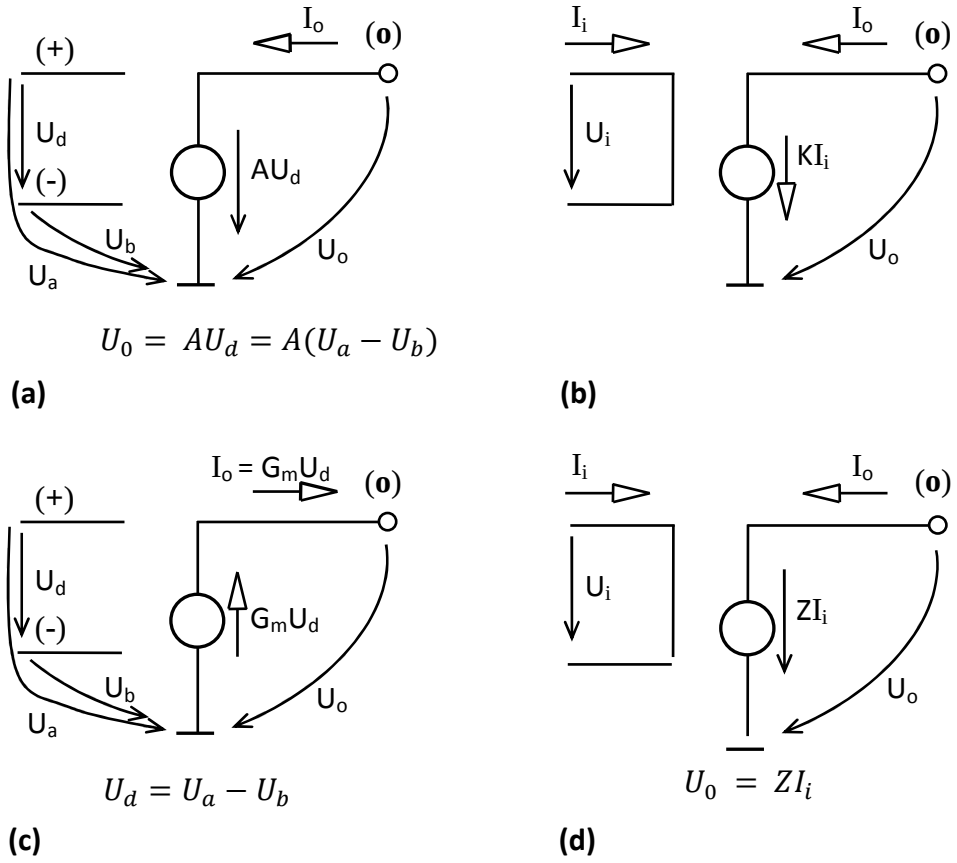


Fig.1.3. a) VCVS – $u_r = U_d$, zero input currents; b) CCCS – $i_r = I_i$, zero input voltage; c) VCCS – $u_r = U_d$, zero input currents; d) CCVS – $i_r = I_i$, zero input voltage; U – voltage; I – current; u_r, i_r - see Tab. 2

CONTROLLING VARIABLE	CONTROLLED SOURCE	
	VOLTAGE ***Internal (output) resistance is zero***	**CURRENT** ***internal (output) resistance is infinite***
VOLTAGE u_r ***input resistance is infinite***	**V**oltage **C**ontrolled **V**oltage **S**ource **VCVS**	**V**oltage **C**ontrolled **Current S**ource **VCCS**
CURRENT i_r ***input resistance is zero***	**C**urrent **C**ontrolled **V**oltage **S**ource **CCVS**	**C**urrent **C**ontrolled **C**urrent **S**ource **CCCS**

Tab. 2: Table of controlled sources.

- Controlling power $p_r = u_r \cdot i_r = 0$ (ideally).
- Actually controlled sources need some power to be delivered to the pairs of terminals with controlling voltages and currents, input resistances are no zero/infinite. Actual output resistances are not zero nor infinite.

  Simultaneously is true:

  Fig. 1.2a - ideal **voltage (differential) amplifier**; $A = U_o/U_d$ – voltage amplification factor. This is **ideal "voltage" OPA if $A \rightarrow \infty$.**

  Fig. 1.2b - ideal **current amplifier**; $K_i = I_0/I_i$ – current amplification factor

  Fig. 1.2c - ideal **transconductance (differential) amplifier**; $G_m = I_o/U_d$ – transconductance [S - Siemens]

  Fig. 1.2d - ideal **transimpedance amplifier**; $Z = U_o/I_i$ – transimpedance [Ω]

- Any linear electronic system, having two output terminals, can be fully modeled by a voltage source, U_{TH} (open - circuit output voltage), in series with an output impedance, Z_{OUT} – **Thévenin´s theorem** (step 1: Find the open-circuit output voltage; Step 2: In the actual circuit, replace any voltage sources with a short circuit and any current sources with an open circuit. Redraw the circuit and determine the total impedance. **OR another general way: Determine short circuit current I_{SC} with the actual circuit and calculate**; $\boldsymbol{Z_{OUT} = U_{TH}/I_{SC}}$).

Linear/nonlinear circuits

Electrical devices are built from nonlinear components. In the circuit world we classify a circuit as linear or non-linear by examining its $i - u$ graph (current-voltage graph). If the $i - u$ graph is a straight line, then the circuit is classified as liner (linear function). This

definition is valid for each circuit element – parameters of linear structures are no function of currents or voltages – Fig. 1.3.

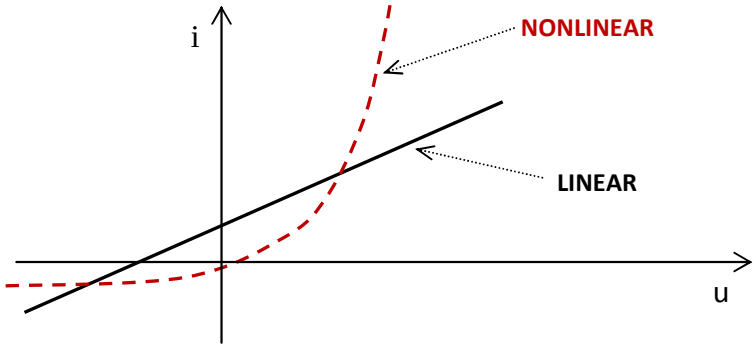


Fig. 1.3. Linear (———) and non-linear (– – –) $i-u$ characteristic

1.2 Linearization

Small-signal modeling is a common analysis method used in electrical engineering to describe nonlinear devices in the terms of linear equations (*parasite non-linearities;* amplifiers, filters). This ***linearization*** is done by first calculating the DC bias point (that is, the voltage/current levels present when no signal is applied), and then forming linear approximations about this point.

Electronic circuits generally involve small time-varying signals carried over a constant bias. This suggests using a method akin to approximation by differentials to analyze relatively small perturbations about the bias point – an example see Fig. 1.4. The $i_C - U_{BE}$ non-linear characteristic is approximated by a straight line with a slope given by

$$g_e = \Delta i_{CQ} / \Delta U_{BEQ} = d(i_{CQ})/d(U_{BEQ}) = I_{CQ}/U_{th} = 1/r_e$$

where U_{th} is the thermal voltage (derivation will be done later).

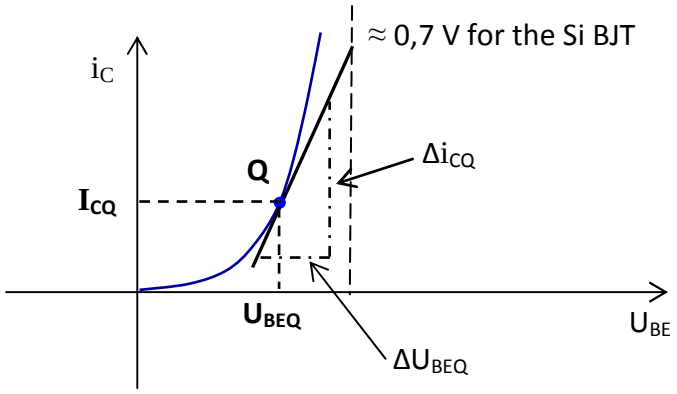


Fig. 1.4. Linear approximation for the Bipolar Junction Transistor's (BJT) input characteristics at the operating point Q.

On the other hand there is a wide spectrum of problems (rectifying, modulation, demodulation, multiplying etc.) which *can not be realized by linear circuits* – non-linearities of some circuit elements are necessary for correct function – *functional non-linearities*.

1.3 Three-terminal linearization

These equations are generally non- linear:

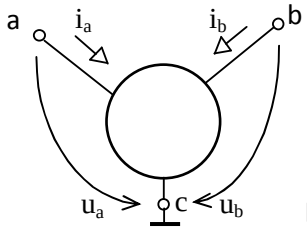


Fig. 1.5. Symbol for three-terminal and **voltage/current arrow convention** (small letters indicate variables)

$$i_a = f_1(u_a, u_b)$$
$$i_b = f_2(u_a, u_b) \qquad (1.1)$$

The quiescent point is described by: u_{a0}, u_{b0} .

If we are interested in small changes $\Delta u_{a,b}$ (signals) close this quiescent point, we can use linear approximation (linearization):

$$u_a = u_{a0} + \Delta u_a$$
$$u_b = u_{b0} + \Delta u_b$$

These signals produce currents:

$$i_a = f_1(u_{a0} + \Delta u_a, u_{b0} + \Delta u_b)$$
$$i_b = f_2(u_{a0} + \Delta u_a, u_{b0} + \Delta u_b) \qquad (1.2)$$

If we use Taylor expansion (the first two components only), we "do" a linear model of the three-terminal:

$$i_a \cong f_1(u_{a0}, u_{b0}) + \left(\frac{\partial i_a}{\partial u_a}\Delta u_a + \frac{\partial i_a}{\partial u_b}\Delta u_b\right)$$
$$i_b \cong f_2(u_{a0}, u_{b0}) + \left(\frac{\partial i_b}{\partial u_a}\Delta u_a + \frac{\partial i_b}{\partial u_b}\Delta u_b\right) \qquad (1.3)$$

Components $i_a \cong f_1(u_{a0}, u_{b0})$, $i_b \cong f_2(u_{a0}, u_{b0})$ describe quiescent point. They must be determined from the basic properties (non-linear) of the investigated three – terminal. The others components describe signal changes. From the eq. (1.3) is evident:

$$\Delta i_a \cong \frac{\partial i_a}{\partial u_a}\Delta u_a + \frac{\partial i_a}{\partial u_b}\Delta u_b$$
$$\Delta i_b \cong \frac{\partial i_b}{\partial u_a}\Delta u_a + \frac{\partial i_b}{\partial u_b}\Delta u_b \qquad (1.4)$$

Partial derivations define signal parameters (conductances) of the three - terminal – close by quiescent point:

$$g_{aa} = \frac{\partial i_a}{\partial u_a} \; ; \qquad g_{ab} = \frac{\partial i_a}{\partial u_b}$$

$$g_{ba} = \frac{\partial i_b}{\partial u_a} \; ; \qquad g_{bb} = \frac{\partial i_b}{\partial u_b} \tag{1.5}$$

This is **the linearized signal model of the three-terminal in the quiescent point** $\boldsymbol{u_{a0}}$, $\boldsymbol{u_{b0}}$, we suppose that conductances are independent of signals now.

We can substitute the time variables by phasors in the frequency domain (harmonic steady state) or generally by Laplace transforms of named variables ($g_{aa} \rightarrow Y_{aa}$, $g_{ab} \rightarrow Y_{ab}$, $g_{ba} \rightarrow Y_{ba}$ and $g_{bb} \rightarrow Y_{bb}$):

$$I_a = g_{aa}U_a + g_{ab}U_b$$
$$I_b = g_{ba}U_a + g_{bb}U_b \tag{1.6}$$

Thus the ***Y- matrix model (ordinary)*** is:

	a	b			
a	g_{aa}	g_{ab}	U_a	=	I_a
b	g_{ba}	g_{bb}	U_b		I_b

(1.7)

For a more-terminal elements we can extend the previous reflection analogously

$$I_a = g_{aa}U_a + g_{ab}U_b + g_{ac}U_c$$
$$I_b = g_{ba}U_a + g_{bb}U_b + g_{bc}U_c$$
$$I_c = g_{ca}U_a + g_{cb}U_b + g_{cc}U_c \tag{1.8}$$

where

$$g_{jk} = \frac{\partial i_j}{\partial u_k} \tag{1.9}$$

for $j = a, b, c$

$k = a, b, c$

The three-terminal is usually analyzed (measured) with one of the terminals referenced to ground. However, this selection is not always ideal for a circuit designer – for circuits analyze. If none of the terminals of a device is referenced to ground (external reference point), a three voltage network results – see Fig. 1.6:.

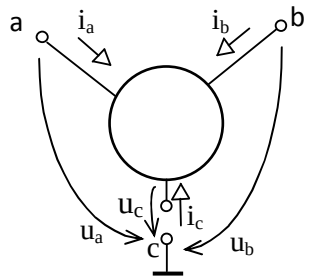


Fig. 1.6. The three-terminal with no ground point

1.4 The indefinite admittance matrix (extended, singular)

Consider the circuit Y – matrix (Fig. 1.6) presentation

$$\begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} = \begin{bmatrix} Y_{aa} & Y_{ab} & Y_{ac} \\ Y_{ba} & Y_{bb} & Y_{bc} \\ Y_{ca} & Y_{cb} & Y_{cc} \end{bmatrix} \cdot \begin{bmatrix} U_a \\ U_b \\ U_c \end{bmatrix} \tag{1.10}$$

Kirchhoff's current law says that the sum of currents coming to the circuit node is equal to zero, thus:

$$I_a + I_b + I_c = Y_{aa}U_a + Y_{ab}U_b + Y_{ac}U_c + Y_{ba}U_a + Y_{bb}U_b + Y_{bc}U_c + Y_{ca}U_a + Y_{cb}U_b + Y_{cc}U_c$$
$$= (Y_{aa} + Y_{ab} + Y_{ac}) \cdot U_a + (Y_{ba} + Y_{bb} + Y_{bc}) \cdot U_b + (Y_{ca} + Y_{cb} + Y_{cc}) \cdot U_c = 0$$

$\Rightarrow$ the sum of any Y – matrix column is zero.

Simultaneously is true:

if $U_a = U_b = U_c = U$

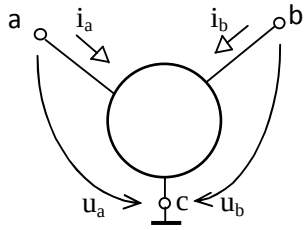
all the three currents must be zero. Thus

$$Y_{aa}U + Y_{ab}U + Y_{ac}U = Y_{ba}U + Y_{bb}U + Y_{bc}U = Y_{ca}U + Y_{cb}U + Y_{cc}U = 0$$

$\Rightarrow$ the sum of any Y -matrix row is zero.

This is the **definition** of the ***indefinite (extended) admittance matrix*** – ***<u>it makes possible to convert the measured data into other needed formats:</u>***

1) We know matrix (parameters) of the circuit in the Fig. 1.7, for example:



	a	**b**			
a	g_{aa}	g_{ab}	U_a	=	I_a
b	g_{ba}	g_{bb}	U_b		I_b

Fig. 1.7. Known parameters

2) We need matrix (parameters) of a circuit in the Fig. 1.8 (or in the other more complex circuitry; in *the same quiescent point* generally, of course):

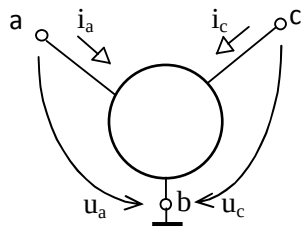


Fig. 1.8. Needed parameters

3) We extend the ordinary admittance matrix by derived rules for the extended matrix (we add one row – c– and one column – c):

	a	**b**	**c**				
a	g_{aa}	g_{ab}	$-g_{aa}-g_{ab}$		U_a		I_a
b	g_{ba}	g_{bb}	$-g_{ba}-g_{bb}$	=	U_b	=	I_b
c	$-g_{aa}-g_{ba}$	$-g_{ab}-g_{bb}$	$g_{aa}+g_{ab}+g_{ba}+g_{bb}$		U_c		I_c

(1.11)

This extended matrix is useful in the more complex circuitry if any node is connected to the reference point – Fig. 1.9:

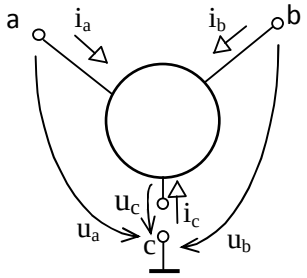


Fig. 1.9. The three-terminal from the Fig. 1.7 with no ground point

4) Now we connect the node b to the external reference point – so that its voltage is equal to zero. The ***respective equations are easily obtained by deleting the corresponding row and column*** (here b) in the eq. (1.11):

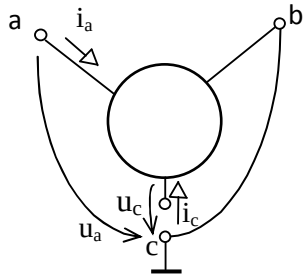


Fig. 1.9. Needed configuration

	a	**b**	**c**				
a	g_{aa}	g_{ab}	$-g_{aa}-g_{ab}$		U_a		I_a
b	g_{ba}	g_{bb}	$-g_{ba}-g_{bb}$	=	U_b	=	I_b
c	$-g_{aa}-g_{ba}$	$-g_{ab}-g_{bb}$	$g_{aa}+g_{ab}+g_{ba}+g_{bb}$		U_c		I_c

from (1.11)

Thus we get the needed matrix (ordinary) with reference node b:

	a	**c**			
a	g_{aa}	$-g_{aa}-g_{ab}$		U_a	I_a
c	$-g_{aa}-g_{ba}$	$g_{aa}+g_{ab}+g_{ba}+g_{bb}$		U_c	I_c

(1.12)

1.5 The quiescent point in general

Every electronics element must be good biased – we must adjust suitable quiescent point.

BJTs – in typical (linear) operation, the ***emitter – base junction is forward biased*** and the ***base – collector junction is reverse biased***. In order to bias the transistor correctly, external voltages are applied – transistor bias circuits: the function of a transistor bias circuit is to maintain a forward bias of the emitter – base junction and reverse bias on the base – collector junction. These bias circuits determine input and output resistances, too. We must always respect a base current of the BJT.

FETs – in typical (linear) operation, the ***gate – source voltage*** U_{GS} must create a current carrying (conductive) ***channel*** (control resistance of this channel) *between terminals* ***drain*** *and* ***source***. If drain-to-source voltage is great enough $[|U_{DS}| > |U_{PINC}H| - |U_{GS}|]$, the FET is said to be in saturation mode (*active mode*, for the better analogy with BJT operating regions). Because the gate is insulated (or reverse biased diode – JFET; there are no forward-biased junctions), the device operation requires very little DC currents (direct current) to flow in the input to the gate. Thus the device has an enormous input resistance (more than 10^{12} Ω). Nevertheless we must always ensure the DC "way" for this small current. That is an infinite resistance creates theoretically the infinite voltage even for the very small currents (in practice some supply voltage)!

Operational amplifiers (electronics structures in general) – we must always assure acceptable quiescent points:

- ***correct power-supply voltages*** (polarity, low impedance – right decoupling capacitors to ground plate - these capacitors are often placed at each power source as well as at each analog component in order to ensure that the supplies are as steady as possible);
- ***every input*** must have ***proper DC "way"*** for its input current – any infinite resistance creates theoretically the infinite voltage (in practice some supply voltage; inverting input – usually feedback resistor, noninverting input – be patient in case of capacitive coupling of noninverting structures!!! – input current without DC way will charge capacity – time variable voltage - ramp). For any "bipolar input" we may think currents 20 – 200 nA (proper resistors under 200 kΩ), for any "FET input" we may think currents 1 – 200 pA (proper resistors under 10 MΩ) – see simple the Ohm's law;
- ***negative DC feedback is usually used*** (amplifier and filter applications).

Summary

1) The operational amplifier (op – amp) is an amplifier consisting of dozens of transistor (≈1942 – op – amp using two twin triode; integrated circuits from ≈1962). An op – amp amplifies the voltage difference between input terminals, and produces a single – ended output voltage. An voltage gain is ideally infinite.

2) Independent sources are ideal circuit elements that passes a voltage or current value that is independent of the behaviour of the circuit to which they belong.
3) A linear circuit is one that contains linear elements, independent sources and linear dependent sources (controlled sources).
4) A linear amplifier is a device that increases the amplitude of a signal (a voltage or a current) whilst preserving waveform shape.
5) In analysis of the small signal´s contribution to the circuit, the non – linear components are modeled as linear components. In general, small signal transfer characteristic can be obtained via Taylor series expansion of large signal.
6) Properties of indefinite admittance matrix:
   - The sum of the elements of every column is zero.
   - The sum of the elements of every row is zero.

Questions 1

You can find the answers in this text.

1. Define the beginning of modern electronics.

2. Define a linear and a nonlinear circuit; linearization.

3. Define the matrix model of a three-terminal element, arrow convention.

4. Define basic components of electronic circuit theory.

5. What is the meaning of following abbreviations: VCVS, VCCS, CCVS, CCCS?

Problems 1

Example 1.1

A 10 V voltage source has the output (internal) resistance 10 kΩ.

a) What is the short current?
b) What load resistance (in series) decreases the short current for 1% (the current source with accuracy 1% for smaller load resistances)?
c) What is the open - circuit voltage?
d) What load resistance (in series) decreases the open - circuit voltage for 1% (the voltage source with accuracy 1% for greater load resistances)?

Example 1.2

You know the model of bipolar transistor – collector current as a function of base-emitor voltage – Ebers-Moll model (simplified):

$$I_c \cong I_S \cdot \left(e^{U_{BE}/U_{th}} - 1\right)$$

Treat I_S as a constant, for any transistor, except that I_S grows fast with temperature

Negligible if $U_{BE} \gg U_{th}$

The expression $U_{th} = k \cdot T/q$ is approximately 26 mV in the room-temperature value (thermal voltage). Thus is valid:

$$I_c \cong I_S \cdot e^{U_{BE}/\left(26 \cdot 10^{-3}\right)}$$

Derive $g_{BE} = 1/r_e = d(I_C)/d(U_{BE})$ as a function of dc current (r_e – intrinsic emitter resistance; it is not a resistor planted in the transistor; it only models the limited conductance of the device).

Example 1.3

Input bias current of an operational amplifier (bipolar) is 100 nA. Find the value of input resistance which creates the voltage just 1 mV.

Example 1.4

Input bias current of an operational amplifier (FET) is 100 pA. Find the value of input resistance which creates the voltage just 1 mV.

Basic texts

[1] Mikulec, M., – Havlíček, V.: Basic circuit theory. Vydavatelství ČVUT, Praha, 2005, ISBN 80-01-03172-1

[2] Mohylová, J.: Lineární obvody s elektronickými prvky-Sbírka příkladů, VŠB-TU Ostrava 2002, ISBN 80-248-0098-5

[3] Punčochář, J.: Lineární obvody s elektronickými prvky. Skriptum, VŠB-TU Ostrava 2002, ISBN 80-248-0040-3

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

2. Quiescent point of active tripoles; admittance models

Time of study: 6 hours

Goals:

- definition of the quiescent points set forth active tripoles
- definition of the BJT admittance model
- definition of the FET admittance model
- definition of the triode admittance model

EXPLANATION

Transistor is short for transfer resistor. The current flowing in the output circuit is determined by the current flowing in the input circuit. Since transistors are three-terminal devices, one electrode must remain shared (common) to both the input and the output.

Transistors fall into two main categories: bipolar junction transistors (BJT) and field effect transistors (FET) and they are also classified according to semiconductors material employed (silicon or germanium) and to their field of application (e.g. general-purpose, switching, high-frequency, etc.).

The transistor was not the first three terminal devices. The vacuum tube triode preceded the transistor by nearly 50 years. The triode allowed us to make an amplifier for audio signals, making AM radio possible.

2.1 THE QUIESCENT POINT OF BIPOLAR JUNCTION TRANSISTORS

The first step in any analysis of an alectronic circuit is to determine the operating point – quiescent point.

2.1.1 Basic description of a BJT principle

A transistor can be considered as two diodes with a shared region (base). In typical operation, the *emitter-base junction is forward biased* and the *base-collector junction is reverse biased*. In NPN transistor, for example, positive voltage applied to the base-emitter junction injects electrons into the base (type P) region. These electrons wander (diffuse) through the base towards the collector (type N). The collector- base junction is reverse biased → electrons that diffuse through the base are swept into the collector by the electric field in the depletion region of the collector-base junction. The ***base*** region of the transistor ***must be*** made ***thin***, so that carriers (electrons in the NPN structure) can diffuse across it in much less time than the semiconductor's minority carrier lifetime, to minimize recombination.

Figure 2.1 shows a schematic diagram. It is showing the flows of electrons within n^+-p-n bipolar junction transistor on the picture. The conventions for the electrical currents are shown. Since only a small portion of electrons will be recombined within base, we can use a small injection current (I_B) to control a much bigger current (I_C).

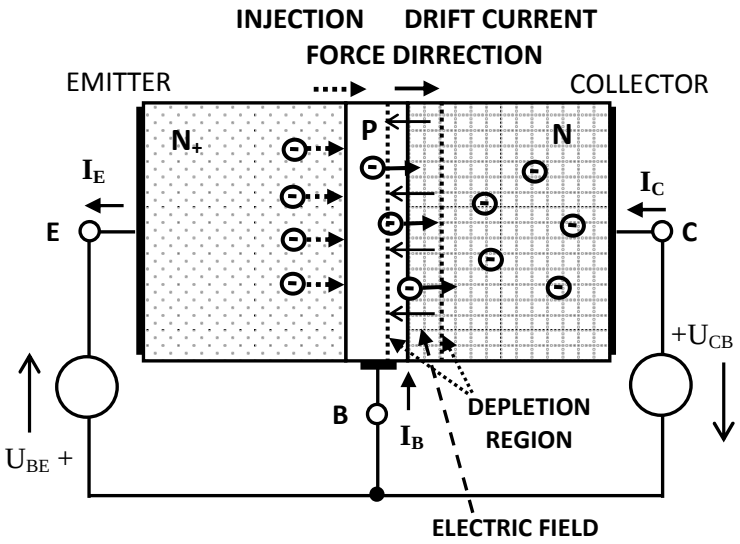


Fig. 2.1. Schematic diagram n^+-p-n bipolar junction transistor.

The equation that relates current flow in the collector, base and emitter currents is:

$$I_E = I_B + I_C$$

where I_E is the emitter current, I_B is the base current and I_C is the collector current (all expressed in the same units).

The current gain offered by a transistor is a measure of its effectiveness as an amplifying device. The most commonly quoted parameter is tha one which relates to common-emitter mode. In this mode, the input current is applied to the base and output current appears in the collector. The common-emitter current gain is given by

$$\beta = I_C / I_B$$

β is also known as h_{FE}.

The current gain in the common-base circuit is calculated in a method similar to that of the common emitter except that the input current is I_E not I_B. Alpha is the relationship of collector current (output current) to emitter current (input current). Alpha is calculated using the formula

$$\alpha = I_C / I_E$$

This is a current gain of less than 1. α is also known as h_{FB}.

By combining the expressions for both α and β, the mathematical relationship between these parameters and therefore the current gain of the transistor can be given as:

since

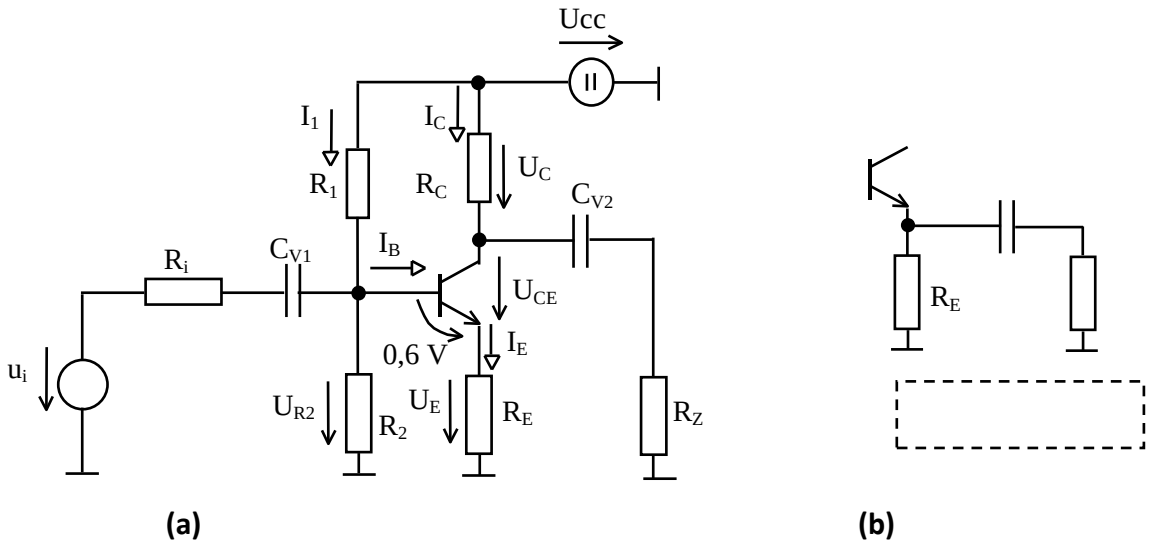


Fig. 2.2. A basic common emitter configuration (CE) – voltage divider biasing – a); a capacitor in fig. b) transfers AC signals through to the earth (reference point), too – parallel connection gives R_e for the AC signal

Example 2.1

Determine quiescent point of the circuit in the Fig. 2.2

☑ Solution:

1. It uses voltage negative feedback to stabilize the quiescent point – this is provided by .
2. Since there is a ± 50 mV spread in for defined value and about a -2 mV/°C temperature dependency on , it is best to allow U_E to be much larger than this variation. We therefore make between 1 V and 3 V (the larger the better from the point of view of quiescent point stability).
3. The fixed DC bias is guaranteed by the divider action of and . Since is drawn from this, it is usual to allow the voltage divider current to be → the BJT load (base) current does not pull the bias voltage down too much.
4. R_E introduces negative (DC) feedback. If the temperature increases, (circa 0, 6 V) decreases (influence of temperature transistor properties) and I_C increases and therefore increases, too. Since is fixed, this causes to fall (influence of feedback properties), which causes to fall back to its original value.

Biasing is determined by following steps (generally):

1. Choose the required (DC) value for the circuit – assume .
2. For a CE circuit, allow about 1 V ( ) to be dropped across – calculate it (you know its voltage and current – Ohm's law).
3. For a CE circuit (and CB), the rest of the power supply must be dropped approximately equally across BJT and – calculate (you know its voltage and current – Ohm's law). For a CC circuit, the rest of the power supply must be dropped across BJT only.
4. The quiescent base voltage is simply the emitter voltage plus (assume 0,6 V).
5. Find the lowest guaranteed for the BJT used and calculate .
6. The voltage divider circuit of and must provide the required .
   6.1. Assume that the lower resistor carries (5-10) times the value. Calculate (you know its voltage and current).
   6.2. The upper resistor therefore carries the current plus (Kirchhoff's current law). This means that carries

   The voltage across is found from the Kirchhoff's voltage law (KVL): – – calculate (you know its voltage and current).

Example 2.2

Design the voltage divider bias for a CE BJT circuit - Fig. 2.2, where , the power supply V and the load resistance kΩ (given, known).

☑ **Solution:** (steps are slightly modified because we know the load resistance now)

- It is known (thereinafter) that the resistance defines (approximately) an output resistance of the CE circuit. Thus we must choose - or else it will degrade voltage gain of the circuit (Thévenin's theorem).
- Thus we choose kΩ.
- We choose V.
- We choose V (V).
- Now we can calculate mA.
- We suppose mA.
- From this we can calculate kΩ.
- Now V.
- Suppose that the lowest value of is 150. Therefore — —— μA.
- The voltage divider must provide 1,2 V at 4 μA.
  Assume that carries 40 μA . Therefore kΩ

- Now R_1 carries $11 \cdot I_B = 44$ μA and voltage drop across R_1 is $\left(U_{CC} - U_{R_2}\right) = (12 - 1{,}2)$ V. Therefore $R_1 = 10{,}8\ \text{V}/44\ \mu\text{A} = 245{,}5$ kΩ.

Practically we can choose $R_2 = 27$ kΩ and R_1 to realize by means two resistors: the first part is "invariable" 220 kΩ *and the other part is "variable" 68* kΩ *- trimming resistor.*

2.1.2 The BJT simplest AC model – common emitter

The U_{BE} and I_C of the BJT are related by the equation

$$I_E \cong I_S \cdot \left(e^{U_{BE}/U_{th}} - 1\right) \approx I_S \cdot e^{U_{BE}/U_{th}}$$

where U_{th} is thermal voltage (of about 26 mV at room temperature) – Shockley equation. For β >>1 we can write

$$I_C \cong I_S \cdot e^{U_{BE}/U_{th}} \qquad (2.1)$$

We can easy derive that a signal conductance is

$$g_{BE} = 1/r_e = \frac{dI_C}{dU_{BE}} = I_S \cdot e^{U_{BE}/U_{th}} \cdot \frac{1}{U_{th}} = \frac{I_C}{U_{th}} \approx \frac{I_E}{U_{th}} \qquad (2.2)$$

We can model this property (in the given quiescent point I_C, of course) as shown in Fig. 2.3. E_i is internal emitter of an ideal transistor – zero voltage between base and E_i (signal model); the u_{BE} is external (signal) voltage; $r_e = 1/g_{BE} = U_{th}/I_C$ – intrinsic (internal) emitter resistance (signal); it is not a resistor planted in the transistor; it only models the real conductance of the device – the signal collector current changes as a function of the input signal voltages u_{BE} – according to the eq. (2.2); $dI_C \rightarrow i_C$; $dU_{BE} \rightarrow u_{BE}$ – signal changes.

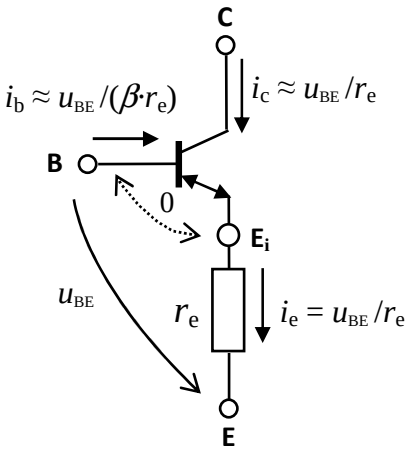


Fig. 2.3 The simplest signal model of BJT in the quiescent point I_C; β – current gain; all voltages referred to E

Now we can easily determine signal equations of the BJT – see Fig. 2.3 (E – common point)

$$i_b = i_c/\beta + 0 \cdot u_{CE} = \frac{u_{BE}/r_e}{\beta} + 0 \cdot u_{CE} \quad \Rightarrow \quad y_{BB} = 1/(\beta \cdot r_e)\,; \qquad y_{BC} = 0$$

$$i_c = u_{BE}/r_e + 0 \cdot u_{CE} \qquad \Rightarrow \quad y_{CB} = 1/r_e\,; \qquad y_{CC} = 0$$

So easily way we get the simplest *admittance model of the BJT*.

2.1.3 The BJT more complex AC model common emitter

Base-width modulation - BJT

As the applied collector-base voltage (U_{CB}) varies, the collector-base depletion region varies in size. An increase in the collector-base voltage causes a greatest reverse bias across the collector-base junction, increasing the collector-base depletion region width, and decreasing the width of the base – ***Early effect***. Narrowing of the base width has two consequences:

a) There is a lesser chance for recombination within the "smaller" base region.
b) The charge gradient is increased across the base, and consequently, the current of minority carriers injected across the emitter junction increases.

In the forward active region the Early effect modifies the collector current I_C and the forward common emitter current gain β as given by the following equations:

$$I_C \cong I_S \cdot e^{U_{BE}/U_{th}}(1 + U_{CB}/U_A) \tag{2.3}$$

$$\beta \cong \beta_0 \cdot (1 + U_{CB}/U_A) \tag{2.4}$$

where

U_{CB} … is the collector – base voltage (≈ U_{CE} always)
U_A … is the Early voltage (15 to 150 V)
β_0 … is forward common-emitter current gain when $U_{CB} = 0$

We can easily derive that a signal conductance is now

$$g_{BE} = 1/r_e = \frac{\delta I_C}{\delta U_{BE}} = I_S \cdot e^{U_{BE}/U_{th}} \cdot \frac{1}{U_{th}} \cdot \left(1 + \frac{U_{CB}}{U_A}\right) = \frac{I_C}{U_{th}} \approx \frac{I_E}{U_{th}} \tag{2.5}$$

and

$$g_{CE} = 1/r_{CE} = \frac{\delta I_C}{\delta U_{CE}} \approx \frac{\delta I_C}{\delta U_{CB}} = I_S \cdot e^{U_{BE}/U_{th}} \cdot \frac{1}{U_{th}} \cdot \left(\frac{1}{U_A}\right) \approx \frac{I_C}{U_A} \approx \frac{I_E}{U_A} \tag{2.6}$$

We can model this more complex property (in the given quiescent point I_C, of course) as shown in Fig. 2.4.

In Fig. 2.4 the u_{BE} is external (signal) voltage; the $r_{CE} = 1/g_{CE} = U_A/I_C$ – emitter - collector resistance (signal); it is not a resistor planted in the transistor; it only models the real conductance of the device – the signal collector current changes as a function of the signal voltages u_{CE}, too – in accordance with the eq. (2.6) – ideally it is infinite.

Now we can easily determine more complex signal equations of the BJT – see Fig. 2.4:

$$i_b = i_c/\beta + 0 \cdot u_{CE} = \frac{u_{BE}/r_e}{\beta} + 0 \cdot u_{CE} \qquad \Rightarrow \quad y_{BB} = 1/(\beta \cdot r_e)\,; \qquad y_{BC} = 0$$

$$i_c = u_{BE}/r_e + u_{CE}/r_{CE} \qquad \Rightarrow \quad y_{CB} = 1/r_e\,; \qquad y_{CC} = 1/r_{CE}$$

We don't reason about influence of equation (2.4) –we suppose β constant.

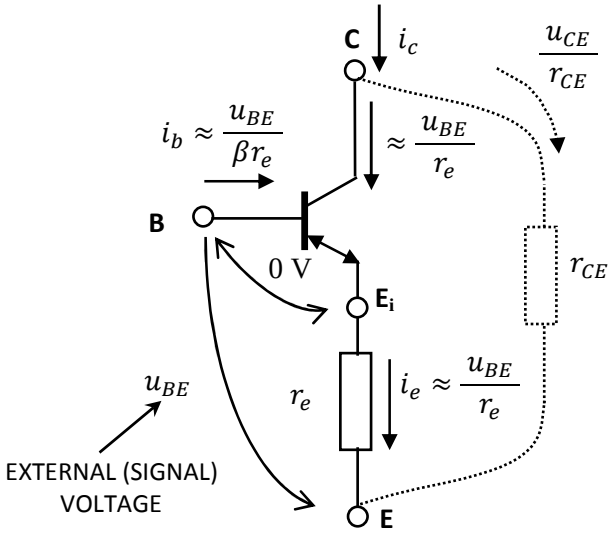


Fig. 2.4 The more complex signal model of BJT (extended) in a quiescent point I_C; it models the Early effect, too; all voltages referred to E

Thus the BJT matrixes are:

	B	C
B	Y_{BB}	Y_{BC}
C	Y_{CB}	Y_{CC}

(2.7)

Y- *matrix model (ordinary) of the BJT; common emitter connection*

	B	C	E
B	Y_{BB}	Y_{BC}	$-Y_{BB} - Y_{BC}$
C	Y_{CB}	Y_{CC}	$-Y_{CB} - Y_{CC}$
E	$-Y_{BB} - Y_{CB}$	$-Y_{BC} - Y_{CC}$	$+\sum Y$
	$\sum Y = Y_{BB} + Y_{BC} + Y_{CB} + Y_{CC}$		

(2.8)

extended Y- *matrix model of the BJT – derived from the common emitter connection*

2.1.4 The BJT AC model – collector capacitance

Most of the potential difference between base and collector is distributed over a thin region at the PN junction, within which the voltage gradient is very high. Consequently, in the vicinity of the junction there are two sets of electric charges (one on either side). The junction acts in much the same way as a capacitor in which charges are separated by a thin sheet of dielectric – for small-signal analysis the effect is closely equivalent to that of a capacitance shunted across the junction. The magnitude of this capacitance depends on the collector-base potential difference, being greatest at low voltages. This ***"depletion layer" capacitance*** is the dominant capacitance at the collector -we can model this capacitance C_{CB} (in the given quiescent point I_C, U_{CB} of course) as shown in Fig. 2.5.

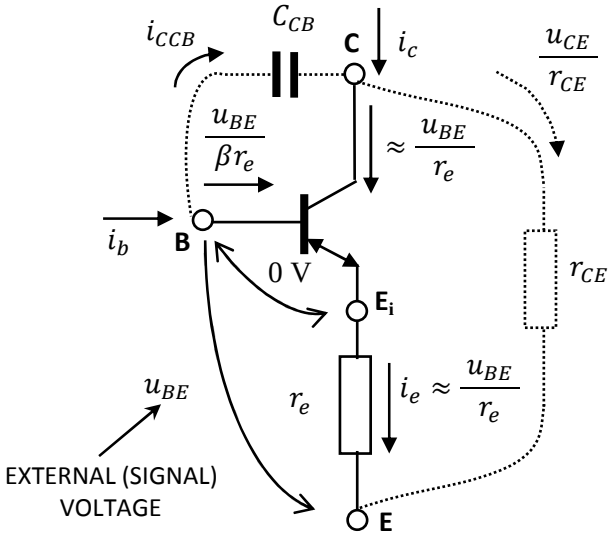


Fig. 2.5 A signal model of BJT (extended) with capacitance C_{CB} in a quiescent point I_C, U_{CB}; all voltages referred to E

Now we can easily derive (harmonic steady state: $p = j\omega$; or Laplace transform):

$$I_{CCB} = (U_{BE} - U_{CE}) \cdot p \cdot C_{CB};$$

$$I_B = \frac{U_{BE}/r_e}{\beta} + I_{CCB} = \frac{U_{BE}/r_e}{\beta} + (U_{BE} - U_{CE}) \cdot p \cdot C_{CB}$$

$$I_B = \left(\frac{1}{\beta \cdot r_e} + p \cdot C_{CB}\right) \cdot U_{BE} - p \cdot C_{CB} \cdot U_{CE}$$

$$\Rightarrow \qquad \boldsymbol{Y_{BB} = 1/(\beta \cdot r_e) + p \cdot C_{CB}}; \qquad \boldsymbol{Y_{BC} = -p \cdot C_{CB}} \tag{2.9}$$

$$I_C = \frac{U_{BE}}{r_e} + \frac{U_{CE}}{r_{CE}} - I_{CCB} = \frac{U_{BE}}{r_e} + \frac{U_{CE}}{r_{CE}} - (U_{BE} - U_{CE}) \cdot p \cdot C_{CB}$$

$$I_B = \left(\frac{1}{r_e} - p \cdot C_{CB}\right) \cdot U_{BE} + \left(\frac{1}{r_{CE}} + p \cdot C_{CB}\right) \cdot U_{CE}$$

$$\Rightarrow \qquad \boldsymbol{Y_{CB} = 1/r_e - p \cdot C_{CB}}; \qquad \boldsymbol{Y_{CC} = 1/r_{CE} + p \cdot C_{CB}} \tag{2.10}$$

see the equations (2.7) and (2.8).

We could add ***base-emitter capacitance*** C_{BE}, further – current that flows through C_{CB} is not amplified by the transistor. C_{BE} changes so rapidly with base current that it is not even specified on transistor datasheets; f_T (unity frequency is given instead).

2.2 The QUIESCENT POINT OF A FIELD EFFECT TRANSISTOR

The ***gate – source voltage*** U_{GS} must create a current carrying ***channel*** *between terminals* ***drain*** *and* ***source***. If drain-to-source voltage is great enough, the FET is said to be in saturation mode.

2.2.1 Basic description of Field Effect Transistor principle

Field Effect Transistor (FET) comprises a channel of P-type or N-type material surrounded by material of the opposite polarity. The ends of the channel (in which conduction take places) form electrodes known as the **source (S)** and **drain (D)**. The effective width of the channel (in which conduction take places) is controlled by a charge (voltage) placed on the third electrode – **gate (G)**. The effective resistance between the source and drain is thus determined by the voltage present at the gate.

Field effect transistors are available in two basic forms; **junction gate** and **insulated gate**. The gate source junction of a junction gate field effect transistor (JFET) is effectively a reverse-biased P-N junction. The gate connection of an insulated gate field effect transistor (MOSFET – Metal Oxide Semiconductor Field Effect Transistor or MISFET – Metal Insulator Semiconductor Field Effect Transistor), on the other hand, is insulated from the channel and charge is capacitively coupled to the channel.

FETs offer a very much higher input resistance when compared with bipolar transistors. For example, the input resistance of a bipolar transistor operating in common-emitter mode (see chapter 2.1) is usually around 2,5 kΩ whereas a JFET device operating in equivalent common-source mode (see chapter 2.3) would typically exhibit an input resistance of 100 MΩ. This feature makes JFET devices ideal for use in applications where very high input resistance is desirable.

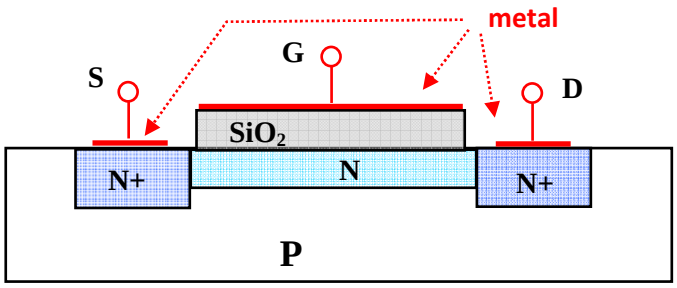


Fig. 2.6 A principle of depletion mode device – MOSFET; S – source; G – gait; D – drain; N – channel

2.2.1.1 Consider an n-channel “depletion-mode” device

Consider a n-channel "depletion-mode" device. (It exists conducting channel between S and D for $U_{GS} = 0$ V). A negative gate-to-source voltage causes a depletion region to expand in width and encroach on the channel from the sides, narrowing the channel. If the depletion region expands to completely close the channel, the resistance of the channel from source to drain becomes large, and the FET is effectively turned off like a switch. Likewise a positive gate-to-source voltage increases the channel size and allows electrons to flow easily – principle see Fig. 2.6 to Fig. 2.7.

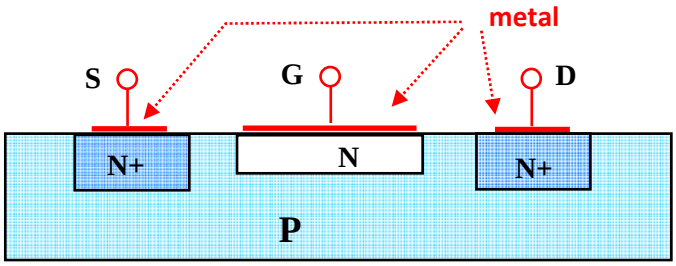


Fig. 2.7 A principle of depletion mode device; junction FET – JFET; $U_{GS} \leq 0$; N – channel; S–source; G–gait; D–drain; Do not forward bias the JFET gate. Forward gate current will burn out the JFET.

If drain-to-source voltage is increased, this creates a significant asymmetrical change in the shape of the channel due to a gradient of voltage potential from source to drain – see Fig. 2.8 to Fig. 2.12.

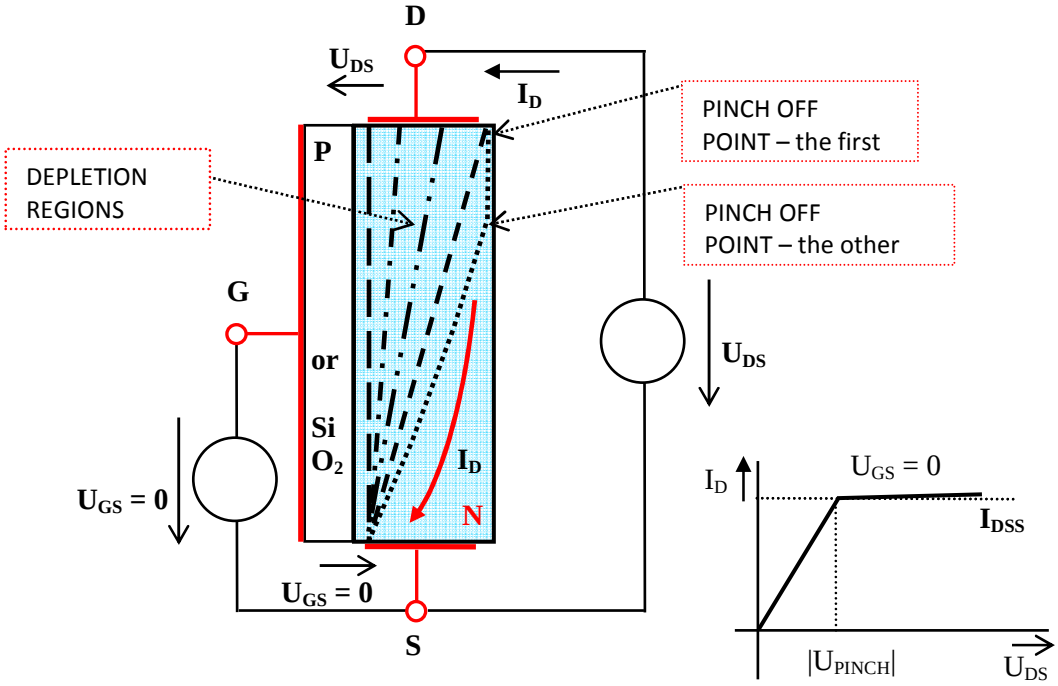


Fig. 2.8 An effect of U_{DS} on depletion region, $U_{GS} = 0$; N - channel
— — — $U_{DS} = 0$; – · – · – $U_{DS} = 0{,}5$ V; — · · — $U_{DS} = 1$ V;
– – – – – $U_{DS} = |U_{PINCH}| = 2$ V for example; $I_D = I_{DSS}$
····· $U_{DS} > |U_{PINCH}|$; $I_D \approx I_{DSS}$

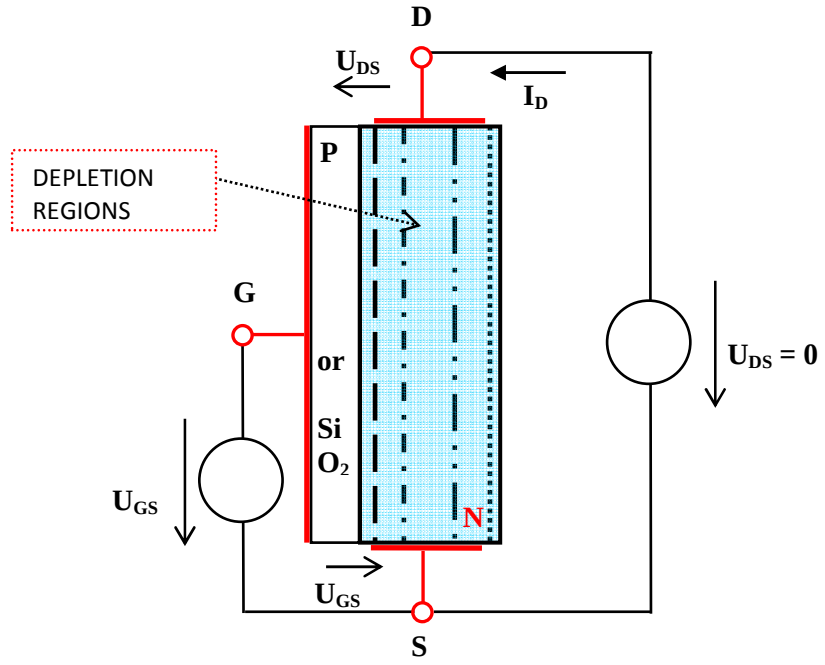


Fig. 2.9 An effect of U_{DS} on depletion region, $U_{GS} = 0$; N - channel
— — — $U_{DS} = 0$; - · - · - $U_{DS} = -0{,}5$ V; — · · — $U_{DS} = 1$ V;
····· $U_{DS} = |U_{PINCH}| = 2$ V for example; $I_D = 0$ for any U_{DS}
$U_{DS} < U_{PINCH}$; $I_D \approx 0$ for any U_{DS}

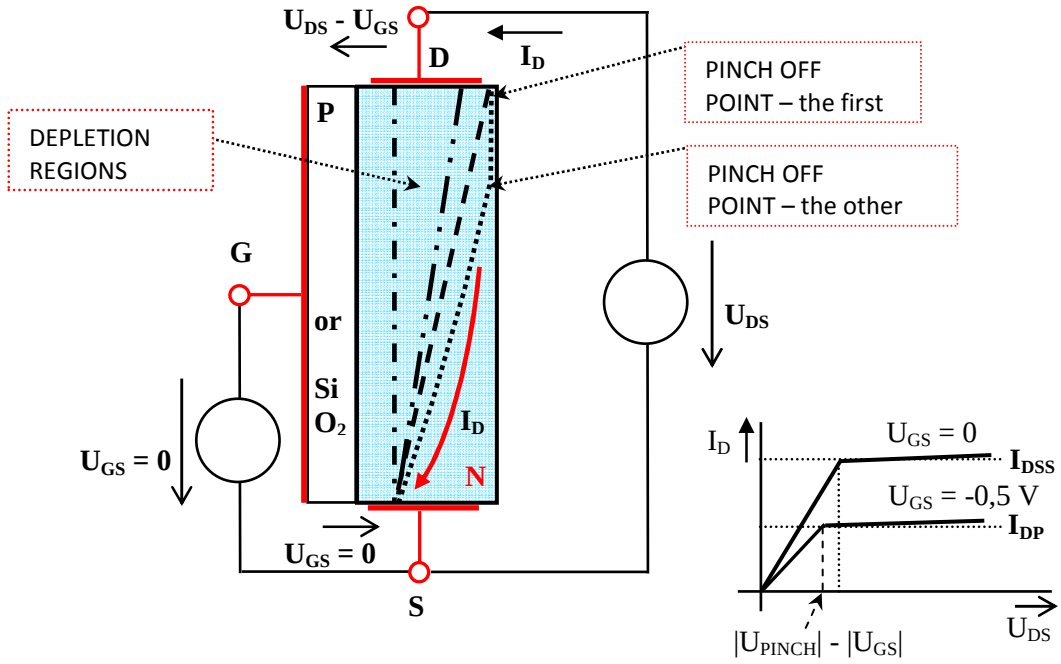


Fig. 2.10 An effect of U_{DS} on depletion region - superposition, U_{GS} = -0,5 V;
- · - · - · $U_{DS} = 0$; $U_{DS} - U_{GS} = 0{,}5$ V; — · · — $U_{DS} = 0{,}5$ V; $U_{DS} - U_{GS} = 1$ V;
- - - - - $U_{DSP} - U_{GS} = |U_{PINCH}| = 2$ V $\Rightarrow U_{DSP} = |U_{PINCH}| + U_{GS} = |U_{PINCH}| - |U_{GS}|$
$\Rightarrow U_{DSP} = -U_{PINCH} + U_{GS}$ (generally) $= 2 - 0{,}5 = 1{,}5$ V; $I_D < I_{DSS}$
····· $U_{DS} > U_{DSP}$; $I_D \approx I_D < I_{DSS}$

The shape of the conductive region (channel) is "pinched-off" near the drain end of the channel. If drain-to-source voltage is increased further, the pinch-off point of the channel begins to move away from the drain towards the source. The FET is said to be in ***saturation mode***

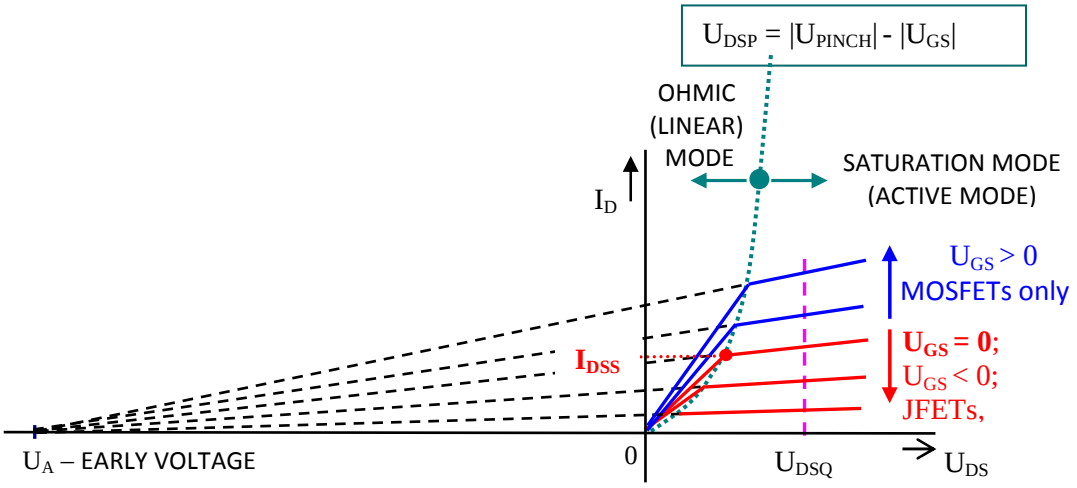


Fig. 2.11 There is $I - U$ characteristics of depletion FETs in this figure, U_{GS} – parameter; N – channel; U_{DSQ} – quiescent voltage; N - channel

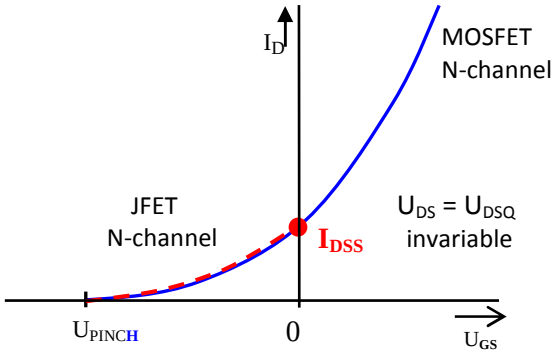


Fig. 2.12 There is $I - U$ characteristics of depletion FETs in this figure, U_{DSQ} – constant quiescent voltage; N - channel

2.2.1.2 Consider an n-channel "enhancement-mode" device

(***No conducting channel between S and D for $U_{GS} = 0$ V***). A positive gate-to-source voltage is necessary to create a conductive channel, since one does not exist naturally within the transistor. The ***positive voltage*** attracts free-floating electron within the body towards the gate, ***forming conductive channel***. But first enough electrons must be attracted near the gate to counter the dopant ions added to the body of the FET; this forms a region free of mobile carriers called a depletion region, and the phenomenon is referred to as the ***threshold voltage U_T*** of the FET. Further gate-to-source voltage will attract even more electrons towards the gate which are able to create a conductive channel from source S to drain D; this process is called ***inversion***; principle - see Fig. 2.13, 2.14

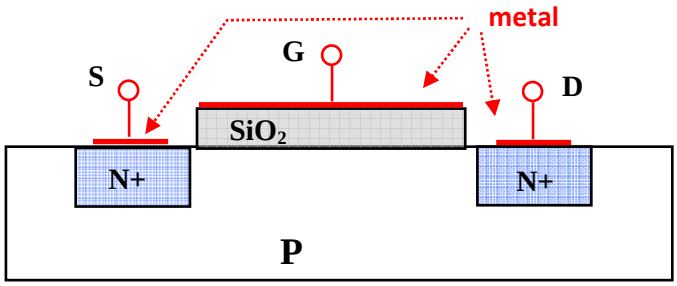


Fig. 2.13 Principle of enhancement mode device – MOSFET; S – source; G – gait; D – drain; N – channel – no induced now

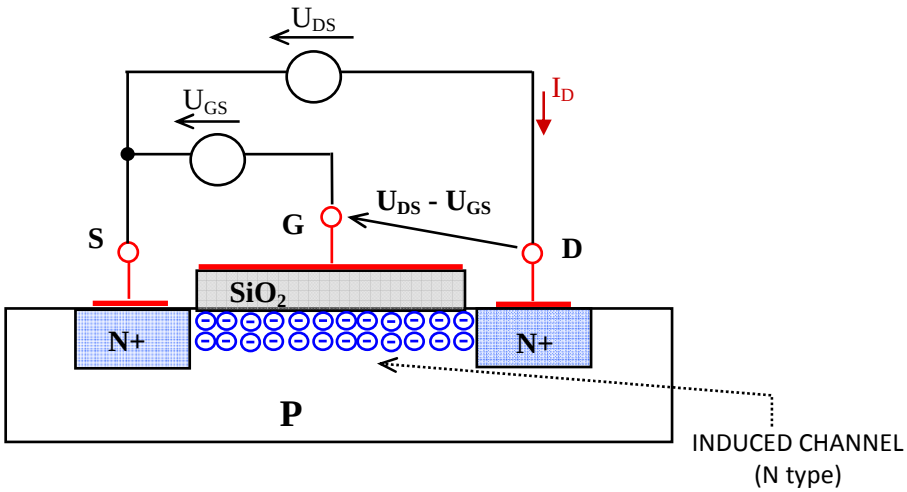


Fig. 2.14 Principle of enhancement mode device – MOSFET; $U_{GS} > U_T$; N – channel – induced now

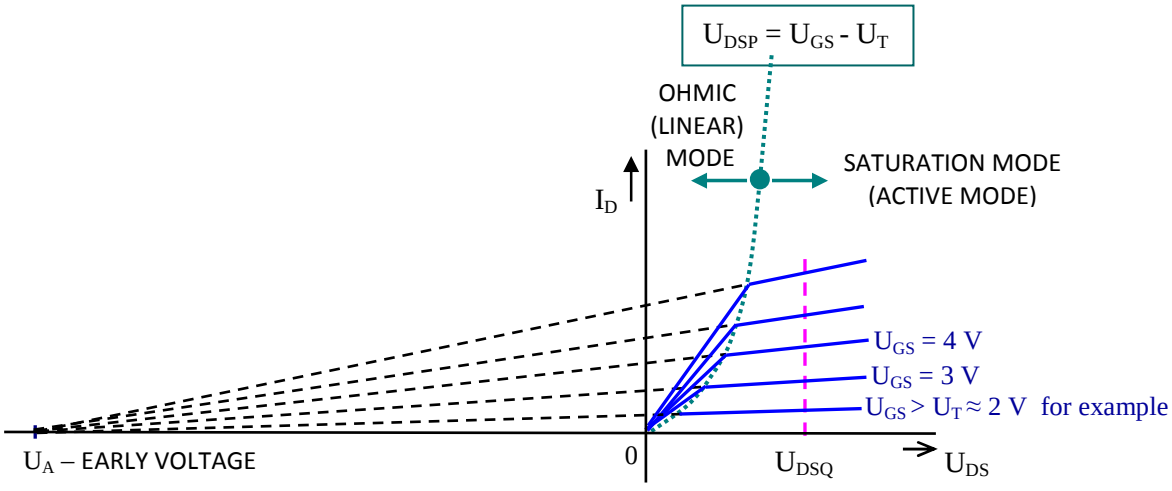


Fig. 2.15 There are $I-U$ characteristics of enhancement FETs in this figure, U_{GS} – parameter; N – channel induced if $U_{GS} > U_T$ U_{DSQ} – quiescent voltage; N – channel

Now there is valid the same thing for the induced channel as for depletion channel - see Fig. 2.8 to Fig. 2.12 – superpositon of influence of $\boldsymbol{U_{GS} > U_T}$ and $\boldsymbol{U_{DS}}$; positive voltage U_{DS} narrows the induced channel close the D; etc. The induced channel is "pinched off" if $\boldsymbol{U_{DS} - U_{GS} \to 0}$, too. We have no current if $\boldsymbol{U_{GS} = 0}$ V; thus it is no defined $I_{DSS,}$ see Fig. 2.15 to Fig. 2.16.

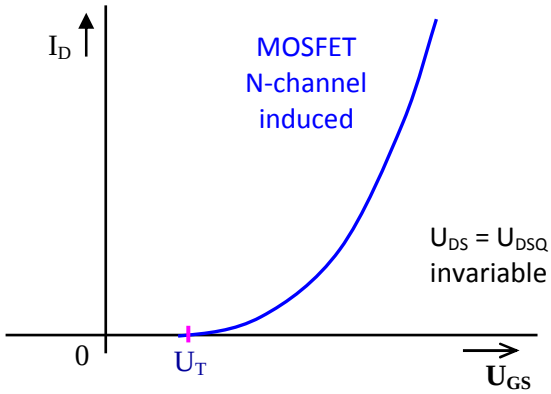


Fig. 2.16 There are $I-U$ characteristics of enhancement FETs in this figure, U_{DSQ} – constant quiescent voltage; N – channel – induced if $U_{GS} > U_T$

For either enhancement-or-depletion-mode devices, at drain-to-source voltages much less than gate-to-source voltages, changing the gate voltage will alter the channel resistance, and drain current will be proportional to drain voltage (referenced to source voltage). In this mode FET operates like a variable resistor and the FET is said to be operating in a ***linear mode*** (***ohmic*** mode).

For reference, here is the universal FET drain-current formula ($U_T \to U_{PINCH} \equiv U_P$ for the depletion type FETs):

$$I_D = 2k\left[(U_{GS} - U_T)\cdot U_{DS} - \frac{U_{DS}^2}{2}\right] \tag{2.11}$$

– it is valid **in linear region**.

If it is valid just $U_{DS} = U_{GS} - U_T$ than I_D "saturates" (pinch-off) and will be approximately constant further – ***saturation mode (region)***:

$$I_D = 2k\left[(U_{GS} - U_T)\cdot U_{DS} - \frac{U_{DS}^2}{2}\right] = 2k\left[(U_{GS} - U_T)\cdot (U_{GS} - U_T) - \frac{(U_{GS}-U_T)^2}{2}\right]$$

$$I_D = k\cdot (U_{GS} - U_T)^2 \tag{2.12}$$

– it is valid **in saturation mode (region)**.

The ***channel-length modulation effect*** (as a function of drain voltage U_{DS}) causes the characteristics to converge at a common intersection – U_A – ***Early voltage*** – see Fig. 2.11, 2.15 – models current dependence on drain voltage – equations (2.11) and (2.12) are modified:

$$I_D = 2k\left[(U_{GS} - U_T)\cdot U_{DS} - \frac{U_{DS}^2}{2}\right]\cdot\left(1 + \frac{U_{DS}}{U_A}\right) \tag{2.11a}$$

– it is **LINEAR REGION**

(2.12a)

– it is **SATURATION REGION**

For the depletion mode device we know that

⇒

Thus in the saturation mode we can rewrite equation (2.12a):

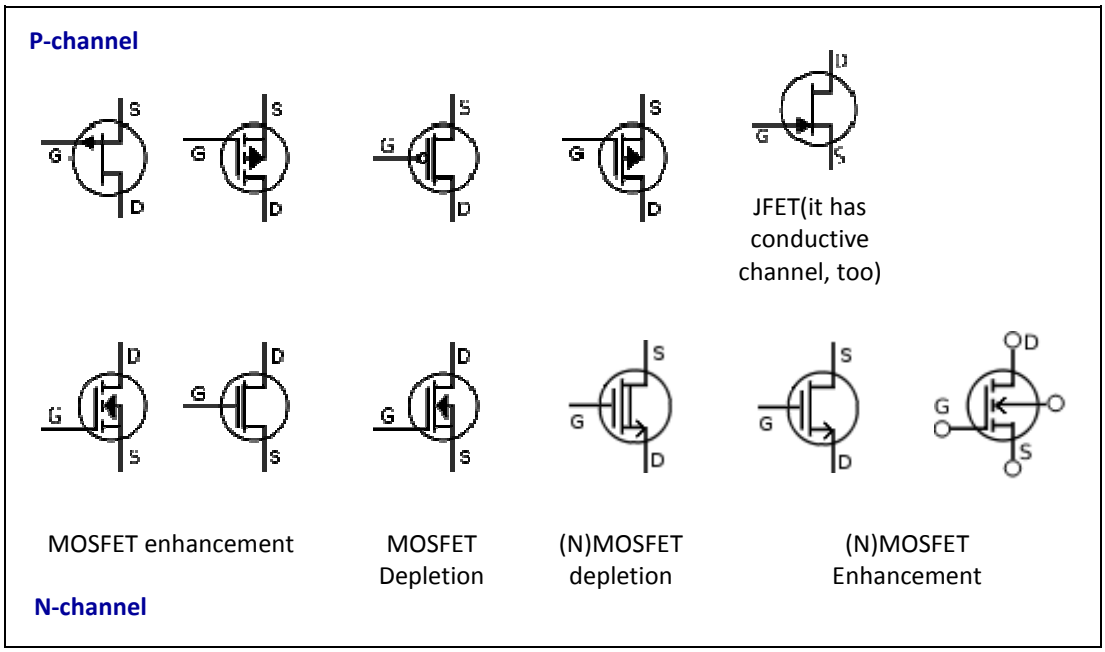
⇒

(2.12b)

– it is **SATURATION REGION –DEPLETION TYPES**

Some FET symbols

Comparison of different enhancement-mode and depletion-mode MOSFET symbols, along with JFET symbols:



Example 2.2

A depletion type MOSFET has a V, mA and V. Derive a quiescent point of the FET – see Fig. 2.17 (ignore Early effect influence, now – V; kΩ; kΩ.

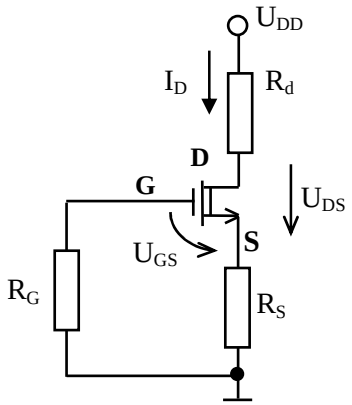


Fig. 2.17 Basic autobias circuit of depletion type FETs (JFETs); R_G usually (from 0,5 to 5) MΩ

☑ Solution:

1. We suppose that (Fig. 2.11)

$$U_{DS} > U_{DSP} = U_{GS} - U_P = U_{GS} - (-2) = U_{GS} + 2\ \text{V}$$

Thus the quiescent is in the saturation region and we can use eq. (2.12b), $U_{DS}/U_A \to 0$:

$$I_D = I_{DSS} \cdot \left(1 - \frac{U_{GS}}{U_P}\right)^2 \cdot (1 + 0) = I_{DSS} \cdot \left(1 - \frac{U_{GS}}{U_P}\right)^2 \qquad (2.12c)$$

2. The gate voltage U_{R_G} is 0 V because no significant current flows through R_G. Thus it is valid (resistor R_S generates the U_{GS})

$$U_{GS} = -R_s I_D$$

3. We establish $U_{GS} = -R_s I_D$ into (2.12c):

$$I_D = I_{DSS}\left(1 + \frac{I_D 1000}{U_P}\right)^2 = 5 \cdot 10^{-3}\left(1 + \frac{I_D 1000}{-2}\right)^2 = 5 \cdot 10^{-3} \cdot (1 - 500 I_D)^2 =$$

$$= 5 \cdot 10^{-3} \cdot (1 - 1000 I_D + 2{,}5 \cdot 10^5 I_D^2);$$

hence

$$I_D = 5 \cdot 10^{-3} \cdot (1 - 1000 I_D + 2{,}5 \cdot 10^5 I_D^2)$$

We easily get equation of the form

$$2{,}5 \cdot 10^5 I_D^2 - 1200 I_D + 1 = 0$$

This has a solution $I_D = 1{,}07$ mA (and $I_D = 3{,}73$ mA). We take the only physical solution $I_D = 1{,}07$ mA which gives

$$U_{GS} = -R_s I_D = -1000 \cdot 1{,}07 \cdot 10^{-3} = -1{,}07\ \text{V}.$$

No physical solution $3{,}73$ mA gives $U_{GS} = -3{,}73$ V – thus FET is would be fully "off".

4. We determine a voltage drain-source $U_{DS} = 10 - (5100 + 1000) \cdot 1{,}07 \cdot 10^{-3} = 3{,}47$ V (Kirchhoff's voltage law) and $U_{DSP} = -1{,}07 - (-2) = 0{,}93$ V. It is valid

$$U_{DS} = 3{,}47\ \text{V} > U_{DSP} = 0{,}93\ \text{V}.$$

Thus the quiescent point (1, 07 mA; 3, 47 V) is really in the saturation region as was supposed. Assumption in the design is right.

Example 2.3

A depletion (N) MOSFET (or JFET) in Fig. 2.17 has a V, mA; V. Calculate the component values if we need quiescent point (5 mA; 5 V).

☑ Solution:

1. We first suppose that (Fig. 2.11) – saturation (active) function region.Thus we can use the equation (2.12c) – we ignored Early effect:

2. For the given values we get from this equation

⇒

This has a solution V (and V). We take the only physical solution V No physical solution is V – thus FET is would be fully "off".

3. It is valid (no significant current flows through) ⇒

Ω.

4. It is valid

(Kirchhoff's voltage law).

Thus we can rearrange this formula:

kΩ

5. We determine

V

the quiescent point (5 mA; 5 V) is really in the saturation region.

6. We choose MΩ**.**

Component values are: Ω; kΩ; MΩ

Example 2.4

Suppose the component values gained in the example 2.3 (Fig. 2.17; V, kΩ, Ω). Now we use other FET – mA, 4 V (selected at random). Determine a new quiescent point.

☑ **Solution:**

1. It is valid now and (Early effect is ignored).

2. We establish into and easily get equation of the form

3. This has a solution I_D = 5, 869 mA (and I_D = 64, 89 mA). We take the only physical solution I_D = 5, 869 mA which gives U_{GS} = - $R_S I_D$ = - 205. 0, 005869 = - 1, 20 V. No physical solution 64, 89 mA gives U_{GS} = - 13, 3 V – thus FET would be fully "off".

Example 2.5

An enhancement (N)MOSFET shown in Fig. 2.18 has , and . Derive a quiescent point (ignore Early effect) when: , , , , and .

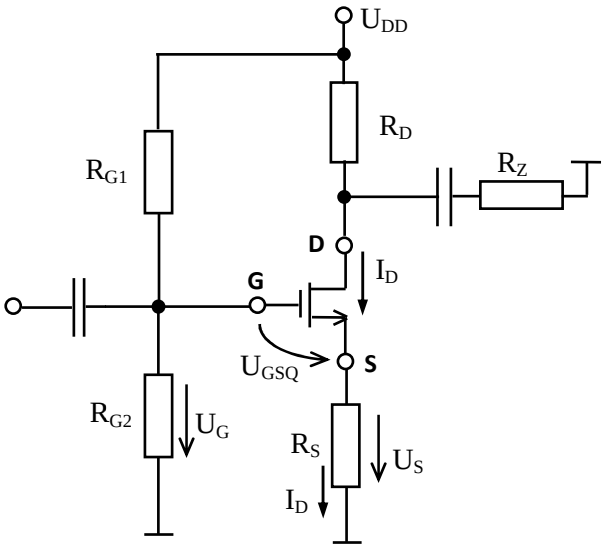


Fig. 2.18 Basic bias circuit of enhancement type FETs

☑ **Solution:**

1. We suppose that (Fig. 2.15)

Thus the quiescent is supposed in the saturation region and we can use eq. (2.12a), :

2. It is valid (voltage divider, supposed zero gait current):

3. It is valid (Ohm's law)

4. It is valid (Kirchhoff's voltage law)

We establish and into and easily get equation of the form

5. Rearranging this equation for given values gives us equation

This has solutions (and) – whereas the second value is not right value (why not?).

6. Now we can calculate

$\Rightarrow$

thus the FETs quiescent point is really (and correctly) in the saturation (active) region – as was supposed.

Example 2.6

An enhancement (N) MOSFET in Fig. 2.19 has , . Derive a quiescent point (ignore Early effect) when: , and

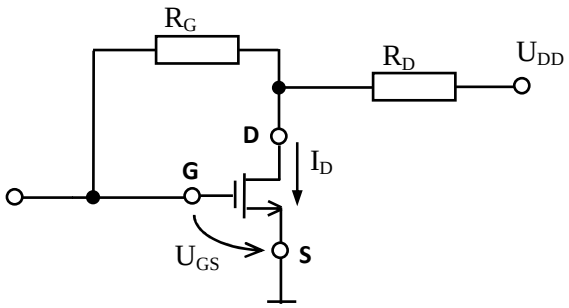


Fig. 2.19 Another bias circuit of enhancement type FETs

☑ Solution:

1. We can ignore gate current (in practice). Thus it is valid , and

2. Rearranging this formula we get easily

$$I_D^2 R_D^2 - \left[2R_D(U_{DD} - U_T) + \frac{1}{k}\right] I_D + (U_{DD} - U_T)^2 = 0$$

This gives for known values

$$2{,}25 \cdot 10^6 I_D^2 - 41500 I_D + 156{,}25 = 0$$

This has solutions $I_D = 5{,}27\ \text{mA}$ (and $13{,}17\ \text{mA}$) – whereas the second value is not right value.

If it is

$I_D = 5{,}27 \cdot 10^{-3} - U_{DS} = U_{DD} - R_D I_D = 15 - 7{,}905 = 7{,}095\ \text{V} = U_{GS}$ – suitable voltage.

If it is

$I_D = 13{,}17 \cdot 10^{-3} - U_{DS} = U_{DD} - R_D I_D = 15 - 19{,}76 = -4{,}76\ \text{V} = U_{GS}$ – bad voltage.

$U_{DSP} = U_{GS} - U_T = U_{GS} - (2{,}5) = 7{,}095 - 2{,}5 = 4{,}59\ \text{V}$ $\Rightarrow$ the FET is really in the active region.

2.3 The FET AC model – common source (in the active – saturation region)

The U_{GS} and I_D of the FET (in the saturation region) are related by the equations (2.12a) – general description – and (2.12b) – depletion types ($U_T \to U_P$ and $k = I_{DSS}/(U_P)^2$ for the depletion type FETs):

We can easily derive now that a "GS" signal conductance is now (equivalent rearranging are used only)

$$g_{GS} = \frac{1}{r_m} = \frac{\delta I_D}{\delta U_{GS}} = \frac{\delta}{\delta U_{GS}}\left[k(U_{GS} - U_T)^2 \cdot \left(1 + \frac{U_{DS}}{U_A}\right)\right] = 2k(U_{GS} - U_T) \cdot \left(1 + \frac{U_{DS}}{U_A}\right) =$$

$$= \frac{2k(U_{GS} - U_T)^2 \cdot \left(1 + \frac{U_{DS}}{U_A}\right)}{U_{GS} - U_T} = \frac{2I_D}{U_{GS} - U_T}$$

or another way

$$= 2 \cdot \sqrt{\left[k \cdot (U_{GS} - U_T) \cdot \left(1 + \frac{U_{DS}}{U_A}\right)\right]^2} = 2 \cdot \sqrt{k\left[k(U_{GS} - U_T)^2 \cdot \left(1 + \frac{U_{DS}}{U_A}\right)\right] \cdot \left(1 + \frac{U_{DS}}{U_A}\right)} =$$

$$= 2 \cdot \sqrt{kI_D \cdot \left(1 + \frac{U_{DS}}{U_A}\right)} = 2 \cdot \sqrt{\frac{I_{DSS} I_D \cdot \left(1 + \frac{U_{DS}}{U_A}\right)}{U_P^2}}$$

If we suppose (in practice rightly) that $\frac{U_{DS}}{U_A} \ll 1$, we get known formulas:

$$g_{GS} = \frac{1}{r_m} = 2k(U_{GS} - U_T) = \frac{2I_D}{(U_{GS} - U_T)} = 2\sqrt{kI_D} \qquad (2.13)$$

appropriate for enhancement types

$$g_{GS} = \frac{1}{r_m} = \frac{2I_D}{(U_{GS}-U_T)} = 2\sqrt{\frac{I_{DSS}I_D}{U_P^2}} \qquad (2.14)$$

appropriate for depletion types (and JFETs)

Further we derive a "DS" signal conductance (equivalent rearranging is used only):

$$g_{DS} = \frac{1}{r_d} = \frac{\delta I_D}{\delta U_{DS}} = \frac{\delta}{\delta U_{DS}}\left[k(U_{GS}-U_T)^2 \cdot \left(1+\frac{U_{DS}}{U_A}\right)\right] = \frac{k(U_{GS}-U_T)^2}{U_A} =$$

$$= \frac{k(U_{GS}-U_T)^2 \cdot \left(1+\frac{U_{DS}}{U_A}\right)}{\left(1+\frac{U_{DS}}{U_A}\right)\cdot U_A} = \frac{I_D}{\left(1+\frac{U_{DS}}{U_A}\right)\cdot U_A}$$

If we suppose (in practice rightly) that that $\frac{U_{DS}}{U_A} \ll 1$ again, we get known formulas:

$$g_{DS} = \frac{1}{r_d} = \frac{I_D}{U_A} \qquad (2.15)$$

We can model these known (derived) properties (in the given quiescent point I_D, U_{DS} of course) as shown in Fig. 2.20.

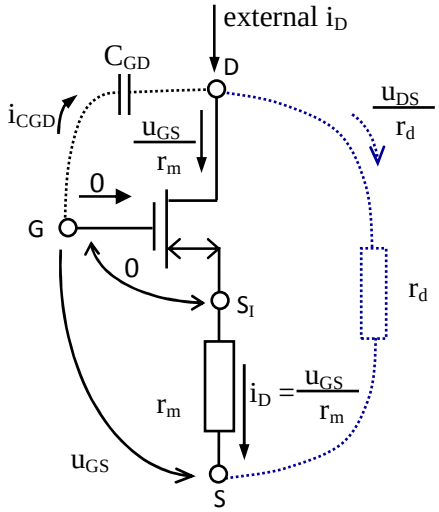


Fig. 2.20 A signal model of FET in a quiescent point I_D, U_{DS}; all voltages referred to S

Fig. 2.20 shows a signal model of a FET in a quiescent point I_D, U_{DS} in which are all voltages referred to S. S_I is an internal source of an ideal FE Tranzistor – zero voltage between gate and S_I (signal model); $r_d = 1/g_{DS} = U_A/I_D$ is drain –source resistance (signal); it is not a resistor planted in the transistor; it only models the real conductance of the device – the signal drain current changes as a function of the signal voltages u_{DS}; $r_m = 1/g_{GS}$ – intrinsic (internal) source resistance (signal); it is not a resistor planted in the transistor; it only models the real conductance of the device – the drain current changes as a function of the input signal voltages u_{GS}; $dI_D \to i_D$; $dU_{GS} \to u_{GS}$ – signal changes.

Now we can easily determine signal equations of the FET – see Fig. 2.20:

$$i_G = 0 \cdot u_{GS} + 0 \cdot u_{DS} \quad \Rightarrow \quad y_{GG} = 0;\ y_{GD} = 0$$

$$i_D = \frac{u_{GS}}{r_m} + \frac{u_{DS}}{r_d} \quad \Rightarrow \quad y_{DG} = \frac{1}{r_m} = g_m;\ y_{DD} = \frac{1}{r_d} = g_d$$

If we consider a ***gate-drain capacitance C_{DG}*** (see Fig. 2.20 – deshed line; harmonic steady state: $p = j\omega$; or Laplace transform), we can easily derive

$$I_{CGD} = (U_{GS} - U_{DS}) \cdot p \cdot C_{DG};$$
$$I_G = 0 + I_{CGD} = 0 + (U_{GS} - U_{DS}) \cdot p \cdot C_{DG}$$
$$I_G = p \cdot C_{DG} \cdot U_{GS} - p \cdot C_{DG} \cdot U_{DS}$$
$$\Rightarrow \quad y_{GG} = p \cdot C_{DG};\ \ y_{GD} = -p \cdot C_{DG} \tag{2.16}$$

$$i_D = \frac{U_{GS}}{r_m} + \frac{U_{DS}}{r_d} - I_{CGD} = \frac{U_{GS}}{r_m} + \frac{U_{DS}}{r_d} - (U_{GS} - U_{DS}) \cdot p \cdot C_{DG}$$
$$i_D = \left(\frac{1}{r_m} - p \cdot C_{DG}\right) \cdot U_{GS} + \left(\frac{1}{r_d} - p \cdot C_{DG}\right) \cdot U_{DS}$$
$$\Rightarrow \quad y_{DG} = \frac{1}{r_m} - p \cdot C_{DG};\ \ y_{DD} = \frac{1}{r_d} + p \cdot C_{DG} \tag{2.17}$$

Thus the BJT matrixes are:

	G	D
G	Y_{GG}	Y_{GD}
D	Y_{DG}	Y_{DD}

(2.18)

Y- *matrix model (ordinary) of the FET; common emitter connection*

	B	C	E
B	Y_{GG}	Y_{GD}	$-Y_{GG} - Y_{GD}$
C	Y_{DG}	Y_{DD}	$-Y_{DG} - Y_{DD}$
E	$-Y_{GG} - Y_{DG}$	$-Y_{GD} - Y_{DD}$	$+\sum Y$
	$\sum Y = Y_{GG} + Y_{GD} + Y_{DG} + Y_{DD}$		

(2.19)

extended Y- *matrix model of the FET – derived from the common emitter connection*

2.3.1 Note to the FET model

A simple FET model we can get another way. Suppose that instantaneous gate – to – source voltage is

$$u_{GS_I} = U_{GS_Q} + u_{GS};$$

where U_{GS_Q} is the DC component and u_{GS} is the AC component (signal voltage).

The instantaneous drain current is $i_{D_I} \cong k \cdot \left(u_{GS_I} - U_T\right)^2$.

Substituting u_{GS_I} into i_{D_I} produces

$$i_{D_I} \cong k \cdot \left(u_{GS_I} - U_T\right)^2 = k \cdot \left(U_{GS_Q} + u_{GS} - U_T\right)^2 = k \cdot \left(U_{GS_Q} - U_T + u_{GS}\right)^2 =$$

$$= k\left(U_{GS_Q} - U_T\right)^2 + 2k\left(U_{GS_Q} - U_T\right)^2 \cdot u_{GS} + k \cdot u_{GS}^2$$

The first term in equation is the DC (or quiescent) drain current I_D (or I_{D_Q}), the second term is the time – varying drain current component that is linearly related to the signal u_{GS}, and the third term is proportional to the square of the signal voltage u_{GS}. For a sinusoidal input signal u_{GS}, the squared term produces harmonics (nonlinear distortion) in the output current, thus output voltage, too. To minimize these harmonics (distortion), we require

$$u_{GS} \ll 2k\left(U_{GS_Q} - U_T\right) \cdot u_{GS}$$

This means that the third term will be much smaller than the second term. *The last equation represents the small – signal condition that must be satisfied for linear amplifier.* Neglecting the u_{GS}^2 term we can write

$$i_{D_I} \cong k\left(U_{GS_Q} - U_T\right)^2 + 2k\left(U_{GS_Q} - U_T\right)^2 \cdot u_{GS} = I_{D_Q} + i_D$$

The total current can be separated into a DC component (I_{D_Q}) and an AC component

$$i_{D_I} \cong 2k\left(U_{GS_Q} - U_T\right) \cdot u_{GS}$$

The small – signal drain current is related to the small – signal gate – to – source voltage by the transconductance

$$g_{GS} = g_m = \frac{1}{r_m} \cong 2k\left(U_{GS_Q} - U_T\right);$$

see equations (2.13) and (2.14), etc.

The term u_{GS}^2 we must not neglect if we solve problems of a distortion (parasite non – linearities) or modulation (functional non - linearities).

2.4 Triode – vacuum tubes

From a historical point of view it is appropriate to know the principle of vacuum tubes.

2.4.1 Basic description of triode principle

Underlying all tube operation is the fact that any hot metal is continuously emitting electrons (for an oxide-coated ***cathode*** under typical operating conditions, a 10% increase in temperature increases emission by about a factor of 3; current passing through the filament heats it).

Electrons, being negatively charged, are attracted to the positive plate – ***anode***.

The number of electrons depends on the current flow - the higher the current, the greater the number of electrons and therefore the greater the charge. Since the cathode feels the influence of the plate through the negatively-charged electrons between them, the increasing current reduces the attractive force of the plate until the two reach a balance. At this point of balance, the effective field at the surface of the cathode is reduced to zero. Moving away from the cathode, the electrons accelerate. If we have just one anode and one cathode (*vacuum diode*) there is valid Child – Langmuir law

$$I_a = k \cdot U_a^{3/2}$$

where:

U_a – an anode – cathode voltage;
I_a – an anode current;
k – constant – dependent on the tube construction.

A ***grid*** of wires between the cathode and the plate (fig. 2.21) is negative (*normally always*), which **decelerates** the electrons and hence controls the current to the plate – ***triode***. Making some simplifying assumptions, it showed that the electric field as seen at the cathode is equivalent to a plate voltage of:

$$U_{ef} = \frac{U_a}{\mu} + U_g$$

μ is a constant for a given electrode geometry. In other words, the actual plate voltage is divided by μ to get the effective voltage. For example, in a typical medium-μ triode under normal operating conditions, the effective voltage as seen at the cathode is only around 5 V, even though the plate is at 100 V or more. An idealized formula is

$$I_a = k \cdot \left(U_a + \mu U_g\right)^{3/2} = k' \cdot \left(\frac{U_a}{\mu} + U_g\right)^{3/2}$$

where:

U_g – a grid – cathode voltage;
μ is a constant for a given electrode geometry.

If we add the second grid, we get ***tetrode*** (double grid tube). If we add the third grid, we get ***pentode*** (triple grid tube).

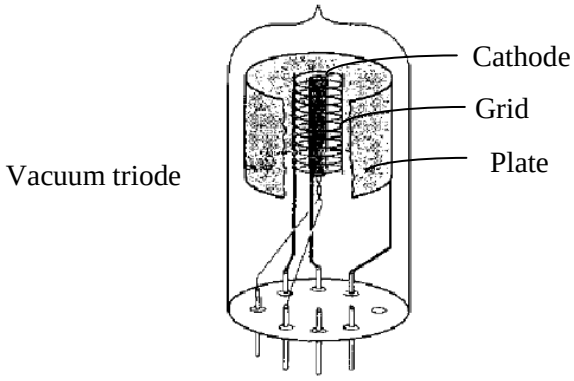


Fig. 2.21 A construction principle

A connection of the simple triode amplifier is shown in Fig. 2.22. The behavior of a triode is fully described by its plate curves, as shown in Fig. 2.22b. These show the anode (plate) current as a function of anode voltage (on the horizontal axis) and the grid voltage, becoming more negative as we move to the right of the family of curves. The used signal model is shown in Fig. 2.22c. Data sheets usually give a tube quiescent point as well as tube's parameters in Fig. 2.22c. The gate voltage U_{R_g} is 0 V because no significant current flows through R_g (if grid voltage is negative). Thus it is valid (resistor R_k generates the U_g – compare depletion FET – N channel): $U_g = -R_k I_a$. A capacitance C_k shorts out R_k – shorts AC signals.

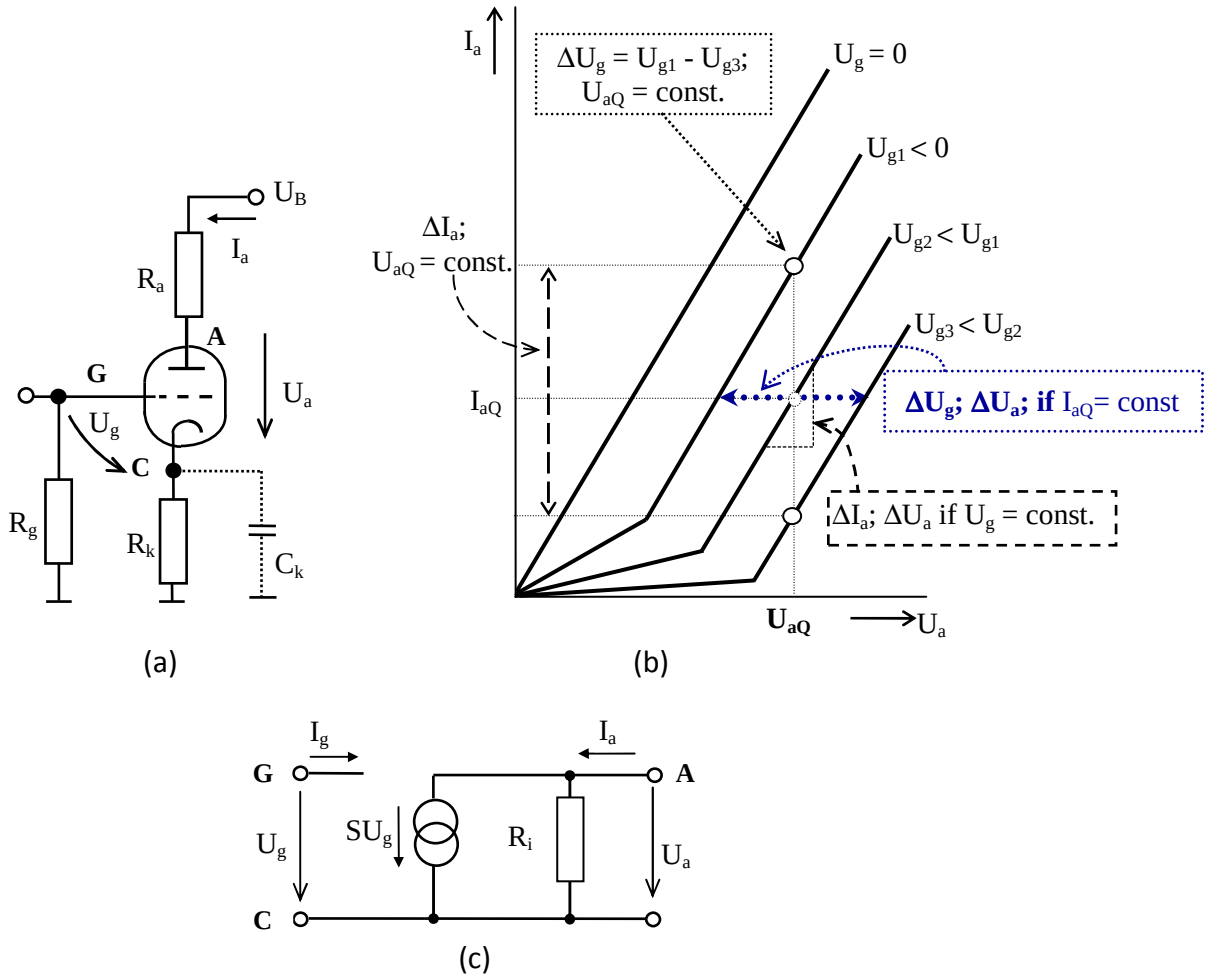


Fig. 2.22 a) An triode amplifier; b) A-V characteristic; c) small - signal model

2.4.2 Quiescent point

Triode ECC83 (double triode) – it is indicated: $U_A = 250$ V, $R_k = 1600\ \Omega$, $I_A = 1{,}2\ mA$, $\mu = 100$, $S = 1{,}6$ mA/V, $R_i = 62{,}5$ kΩ. If we require $R_a = 100$ kΩ for example and this quiescent point, we can calculate a supply voltage

$$U_B = U_A + (R_a + R_k) \cdot I_a = 372 \text{ V}.$$

2.4.3 Small – signal model (parameters) and small – signal triode model

Mutual transconductance is the slope of the transfer characteristic

$$g_m = \frac{1}{r_m} = S = \left.\frac{\Delta I_a}{\Delta U_g}\right|_{U_a=const}$$

It relates the mutual changes between anode current and voltage changes in grid circuit – see Fig. 2.22b, also – the transfer characteristic $I_a = f(U_g;\ U_a = U_{a_Q} = const.)$ we can derive, if we need it, from Fig. 2.22b.

Plate (anode, internal, output) resistance:

represents the change in voltage associated with a change in current in the anode circuit – see Fig. 2.22b, too.

$$R_i = \frac{1}{G_i} = \left.\frac{\Delta U_a}{\Delta I_a}\right|_{U_g=const}$$

Gain coefficient:

represents the change in voltage in the anode circuit associated with a change in voltage in the grid circuit – see Fig. 2.22b, too.

$$\mu = \left.\frac{\Delta U_a}{\Delta U_g}\right|_{I_a=const}$$

Barkhausen formula:

It is evident that for small – signal changes is valid

$$S \cdot R_i = \mu$$

"More physical model" is in the fig. 2.23.

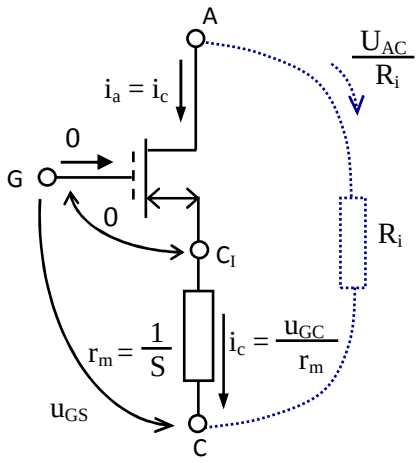


Fig. 2.23 Small – signal triode model; in the appropriate quiescent point

Small – signal triode model (in the appropriate quiescent point) is shown in Fig. 2.23. There are:

- internal cathode C_i of an ideal triode – zero voltage between gate and C_i (signal model);
- R_i - anode – cathode resistance (signal) is not resistor planted in triode, it only models the real conductance of the device – the signal anode current changes as a function of the signal voltages u_{AC};
- $r_m = 1/S$ – intrinsic (internal) cathode resistance (signal); it is not a resistor planted in the triode; it only models the real conductance of the device – the anode current changes as a function of the input signal voltages u_{GC}.

Triode admitance model:

If we consider just signal changes now, thus from Fig. 2.22c or Fig. 2.23 we can easily derive formulas (common cathode)

$$I_g = 0 \cdot U_g + 0 \cdot U_a$$

$$I_s = S \cdot U_g + G_i \cdot U_a$$

Thus we easily get (harmonic steady state) ordinary matrix model

	G	A
G	0	0
A	S	G_i

If we take into account a capacity anode – gate (C_{AG}), we easily determine more complex triode ordinary admittance matrix (compare FET matrix models)

	a	**c**		
a	pC_{AG}	$-pC_{AG}$	U_g	I_g
c	$S - pC_{AG}$	$G_i + pC_{AG}$	U_a	I_a

(2.20)

You can get ordinary extended matrix.

2.5 Admitance model of general active three – terminal device

(*We will suppose that input current is not function of a voltage* u_{23}*;* $i_1 \ll i$ *and the device quiescent point is correct*).

Everything described above can be summarized generally in Fig. 2.24 and two matrices and Table 1 – we need just one general model.

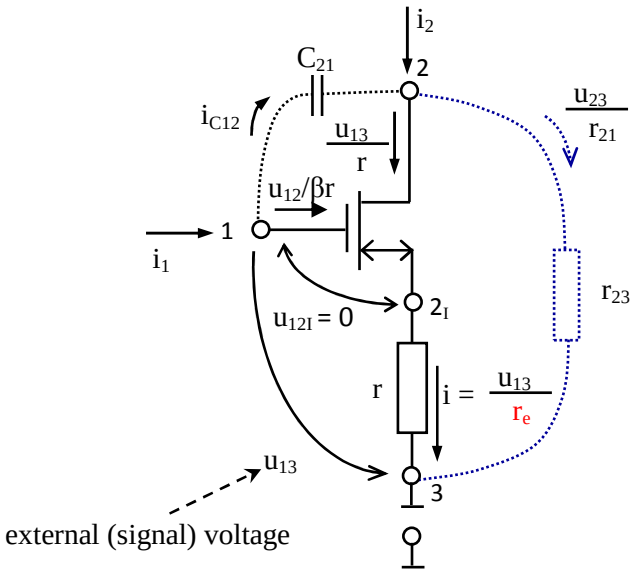


Fig. 2.24 A signal model of general active three-terminal device with capacitance C_{21}; all voltages referred to the terminal 3; input current is not function of a voltage; $i_1 \ll i$ and the device – see Table 1

In Fig. 2.24: 2_I – internal point of the ideal element (zero voltage u_{12_I})
i_1 – input current;
r – defines the i_2 changes as a function of the u_{13};
r_{23} – defines the i_2 changes as a function of the u_{23}

	1	2
1	$\frac{g}{\beta} + pC_{21}$	$-pC_{21}$
2	$G - pC_{21}$	$g_{23} + pC_{21}$

Ordinary admitance matrix – common terminal 3

	1	2	3
1	$\frac{g}{\beta} + pC_{21}$	$-pC_{21}$	$-\frac{g}{\beta}$
2	$g - pC_{21}$	$g_{23} + pC_{21}$	$-g - g_{23}$
3	$-\frac{g}{\beta} - g$	$-g_{23}$	$\frac{g}{\beta} + g + g_{23}$

Extended admitance matrix – external common terminal

general tripole device	**bipolar transistor; BJT**	**field effect transistor; FET**	**triode**
terminal 1	≡ Base	≡ Gate	≡ Grid
terminal 2	≡ Collector	≡ Drain	≡ Anode
terminal 3	≡ Emitter	≡ Source	≡ Cathode
input current i_1	$\frac{i}{\beta}$	$i_G \to 0$ it means that model $\beta \to \infty$	$i_G \to 0$ it means that model $\beta \to \infty$
$g = \frac{1}{r}$	$g_e = \frac{1}{r_e} = \frac{I_C}{U_{th}}$	$g_{GS} = g_m = \frac{1}{r_m} = 2k(U_{GS} - U_T) =$ $\frac{2I_D}{(U_{GS}-U_T)} = 2 \cdot \sqrt{kI_D} = 2 \cdot \sqrt{\frac{I_{DSS}I_D}{U_P^2}}$	$S = \frac{1}{r_m} = g_m$
$r_{23} = \frac{1}{g_{23}}$	$r_{CE} = \frac{U_A}{I_C}$	$r_D = \frac{U_A}{I_D}$	R_i
C_{21}	C_{CB}	C_{DG}	C_{AG}

Table 1. Sum of „all“ as mentioned above; $y_{12} = 0$ if input current is no function of u_{23} and parasitic capacitance is neglected. $U_T \to U_P$ – if depletion type

BJT versus FET

The FET transconductance is

$$g_m = \frac{1}{r_m} = 2 \cdot \sqrt{kI_D} = 2 \cdot \sqrt{\frac{I_{DSS}I_D}{U_P^2}}.$$

It increased only as the square root of I_D and is well below the transconductance

$$g_e = \frac{1}{r_e} = \frac{I_C}{U_{th}}$$

of a bipolar transistor at the same operating current. We will derive later that a low transconductance means a low voltage gain. The problem of the low voltage gain in FET amplifiers we can solve by means a current – source (active) load, but once again the bipolar transistor will be better in the same circuit. For this reason you seldom see FETs used as simple amplifiers, unless it is important to take advantage of their unique input properties – extremely high input resistance and low input current.

AC EQUIVALENT CIRCUIT (small-signal equivalent circuit)

Understanding and designing circuits can be made much easier if we can replace real devices by simplified mathematical models of the same thing. We know signal models (and their

matrix models) of BJT, FET and TRIODE, now ("signal" active components of a circuit). We must be able to determine a signal model of whole electronic circuit.

Capacitors

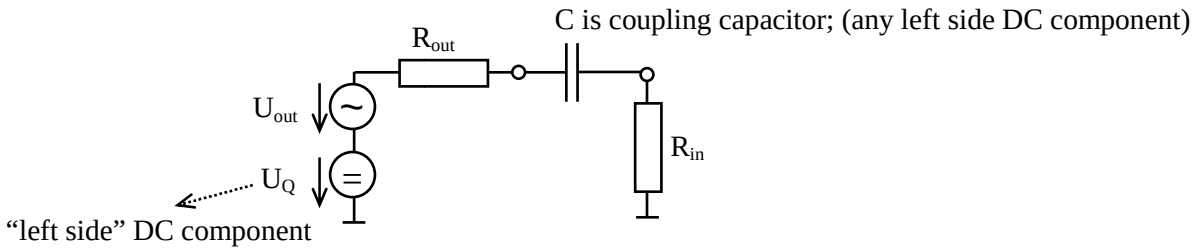


Fig. 2.25 A model circuit with a coupling capacitor

A capacitor has frequency-dependent impedance

— ; — .

We will ignore the phase-changing property as this is not relevant to our discussion here. As reduces to DC, then and we have . As increases towards infinity, then we have . This gives us two rules for capacitors in equivalent circuits:

1. Capacitors in DC equivalent circuits (determination of a quiescent point): replace capacitors by an *open circuit*.
2. Capacitors in AC equivalent circuits: replace capacitors by a *short circuit*.

In fact, the designer usually chooses values of which make negligible at all frequencies likely to be encountered for the amplifier. If we solve frequency properties of a circuit, the problem is more complex – capacitors must not be neglected (parasitic frequency properties of amplifiers, filters – functional capacitors).

Coupling capacitors (vazební)

Capacitors are used in this way to couple AC circuits together (without any steady DC bias conditions being affected) – must be small compared with other resistances in the circuit – at the minimum frequency specified for the circuit. This is an example of the *worst case design*.

Example 2.7

Find a suitable input coupling capacitor for an amplifier with a frequency response of to and , driven from a voltage source with (Fig. 2.25).

☑ **Solution:**

Referring to Fig. 2.25, we would choose

Therefore

For some designs this can generate enormous values, so designers opt for a more realistic value.

Bypass capacitors ("přemosťovací")

In Fig. 2.26 we can see that bypass capacitor transfers AC signal directly to earth, thus bypassing . For AC, is short circuited (and therefore its action is ignored), whereas for DC comes into play and is considered as an open circuit.

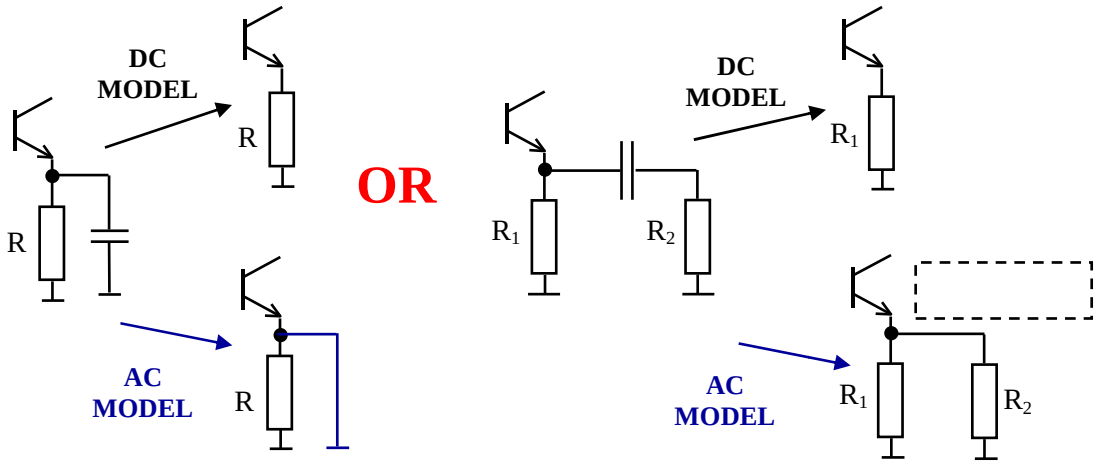


Fig. 2.26 A model circuits with a bypass capacitor

Example 2.8

Choose a capacitor to create an AC bypass across a 1 kΩ resistance at all frequencies in the range 25 Hz to 10 kHz.

☑ **Solution:**

Using

Equivalent circuits of (power) supplies

All active circuits require some form of energy sources. Usually we do not want a power supply the output of which varies with variation in load current drawn from it, thus we need an ideal DC voltage (power) source – its output resistance is zero. Therefore, the rules are:

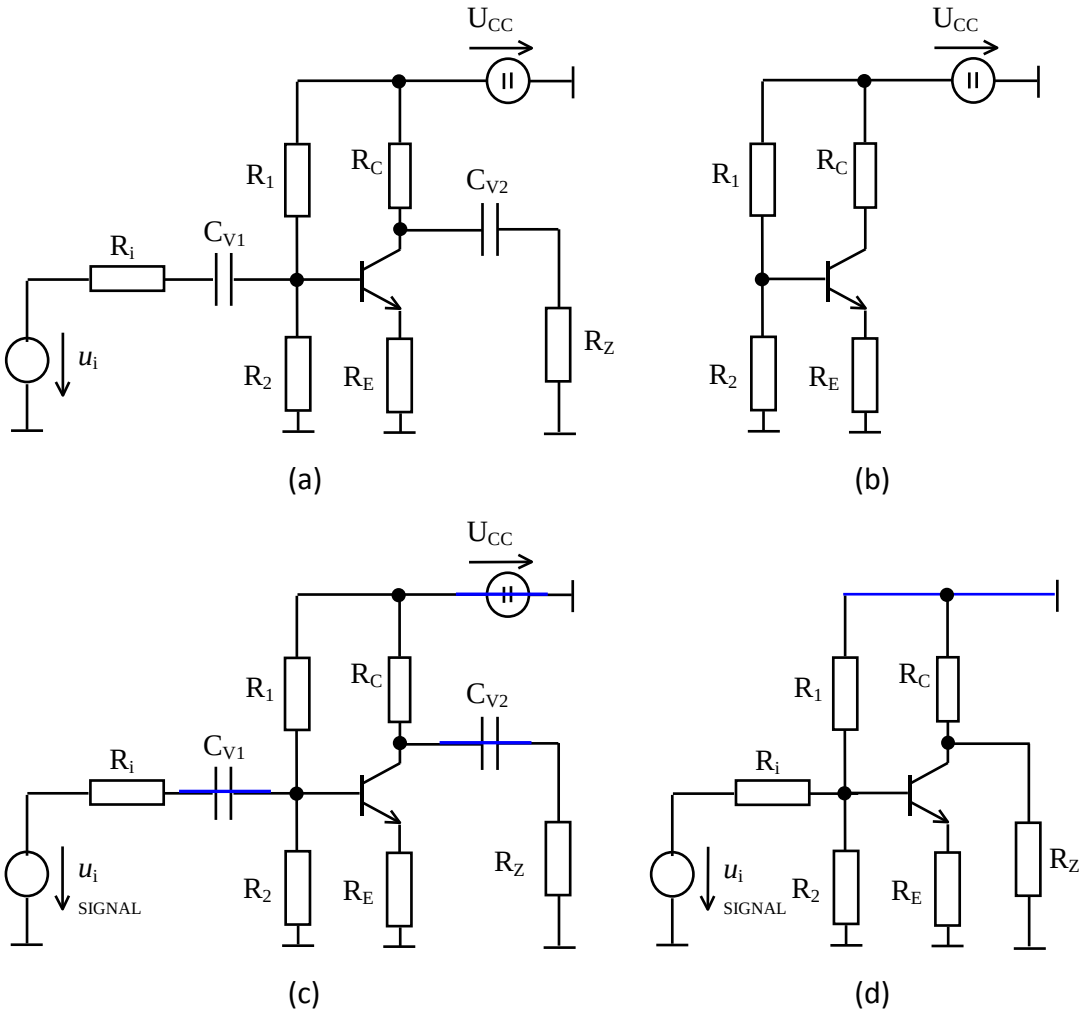


Fig. 2.27 a) A basic common emitter configuration (CE) – voltage divider biasing;
b) A DC equivalent circuit – all capacitors are opened;
c) AC - all capacitors and voltage power supply source are shorted;
d) an AC equivalent circuit – u_i - signal voltage we must not omit

RULE 1: For power supplies (sources) in DC equivalent circuits: show all power supply voltages and currents in full.

RULE 2: For power supplies (voltage sources) in AC equivalent circuits: replace all voltage sources with a short circuit – see Figs. 2.27, 2.29 and 2.30.

If we use an ideal current source (with infinite internal resistance), another rule is valid:

RULE 3: For current sources in AC equivalent circuits: replace all current sources with an open circuit – see Fig. 2.28.

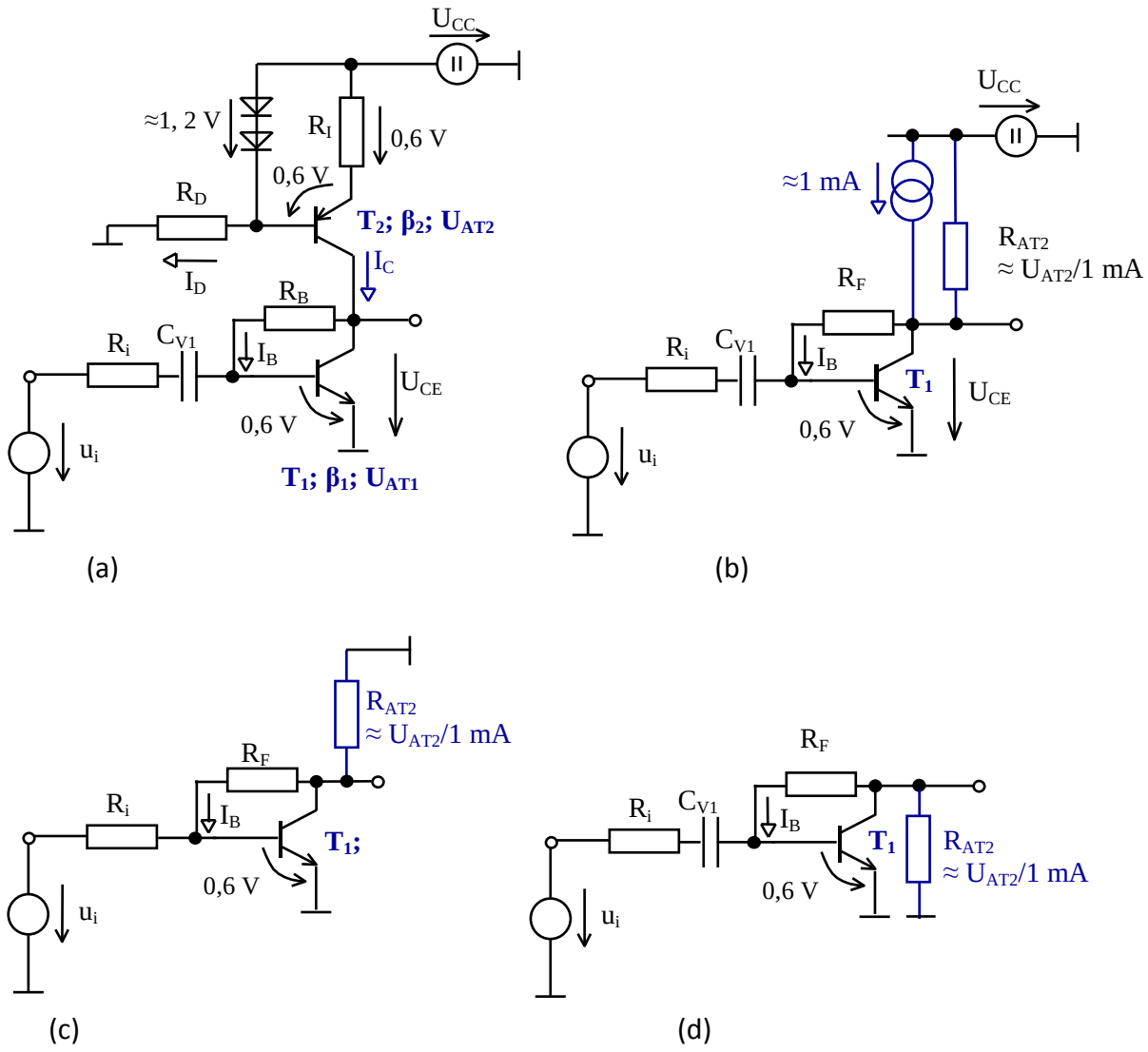


Fig. 2.28 a) A common emitter configuration (CE) – current source as a collector resistor (active): T2 (PNP), diodes, R_I and R_D create current source: it is evident that voltage across resistance R_I is ≈ 0,6 V; thus $I_C \approx 0{,}6/R_I = 1$ mA if $R_I = 600\ \Omega$; we must choose $I_D \gg I_C/\beta_2$;

b) A model of current source – $R_{AT_2} \approx U_{AT_2}/I_C$ – describes influence of T_2 Early voltage;

c) AC equivalent circuit and d) the rearranged AC equivalent circuit.

Resistance R_B defines a T_1 base current. It is valid $I_B \approx (U_{CE} - 0{,}6)/R_B = I_C/\beta_1$. We choose (properly) $U_{CE} = U_{CC}/2$ than than $R_B \approx \beta_1 \cdot (U_{CC}/2 - 0{,}6)/I_C$ – attention, it creates *negative feedback*!!

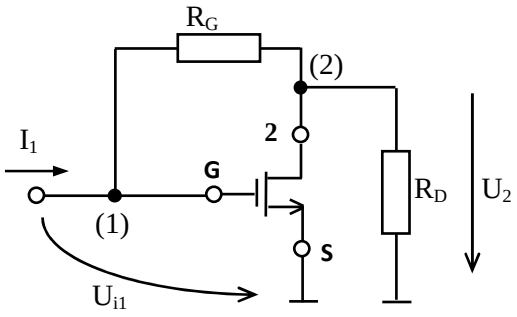


Fig. 2.29 AC equivalent circuit of the structure in Fig. 2.19

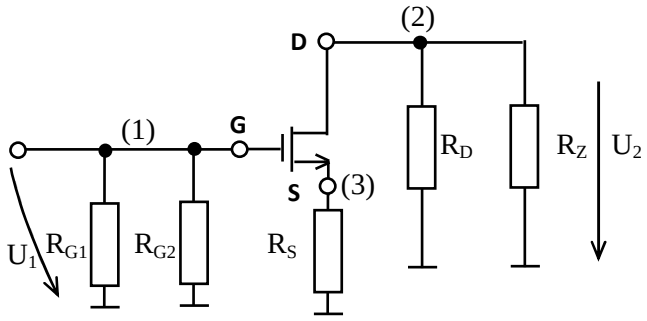


Fig. 2.30 AC equivalent circuit of the structure in Fig. 2.18

Summary

1) A BJT can be considered as two diodes with a shared region – base; the emitter – base junction is forward biased and the base – collector junction is reverse biased. The base regin must be made thin. BJT transistors are reffered to as current – controlled devices.
2) Analysis of biasing of a BJT:

- Typically, you start out assuming that the base – emitter junction is forward biased and $U_{BE} \approx 0{,}6\ V$.
- Check that the base – collector junction is reverse biased. If it is true and I_B is uniquely determined, then $I_C = \beta I_B \cong I_E$.
- In practice, we generally assume that a transistor is not saturated until $U_{CE} > 0{,}2\ V$.

3) Small – signal behavior of BJTs

   The basic trick is to linearize the very nonlinear exponentional relationship between U_{BE} and I_C by looking at a small enough region of the exponential; we get $r_e = U_{th}/I_E$, small signal equations and an admittance model of BJT (matrix – ordinary, extended).

4) A FET is three – terminal device that uses the voltage (field) of the two terminals to control the current flowing in the third terminal, conduction in a channel is controlled by an electric field.

5) Analysis of biasing of FETs:

- The nonlinear relationship between I_D and U_{GS} can complicate the mathematical approache to the DC analysis of FET configuration.
- The general relationship that can be applied to the DC analysis of all FET amplifiers are:

$$I_G = 0\ A \text{ and } I_D = I_S.$$

- The saturation region (constant currents) is the region typically employed in linear amplifiers.

6) Small signal behavior of FETs – the basic trick is to linearize the quadratic relationship between I_D and U_{GS}; we get $r_m = 1/g_m = (dI_D/dU_{GS})^{-1}$, small signal eqations and the admittance model of the FET.

7) From a historical point of view it is appropriate to know the principle of triode. After all, there are only three basic active elements: tubes, BJT and FET,

8) We must be able to determine a small signal models of "whole" electronic circuits.

Questions 2

You can find the answers in this text.

1. Explain basic principle of BJT.
2. Explain basic principle of FET.
3. Explain basic principle of triode.
4. Explain relationship between ordinary and extended matrices.
5. Explain a stabilization of transistor quiescent current (basic principle).

Problems 2

Example 2.1

A Determine an admittance matrix of a BJT in Fig. 2.1. Suppose that Early voltage is infinite (it is neglect influence of Early voltage), neglect capacity collector – base.

Example 2.2

Determine an admittance matrix of a FET in Fig. 2.18, neglect parasitic capacity.

Example 2.3

a) Determine a quiescent point of the NJFET in Fig. 2.31; $U_P = -5$ V, $I_{DSS} = 6$ mA, $U_A = 100$ V – we demand $U_{GSQ} = -2$ V. Determine U_{DD} and all resistances; we suppose load resistance 10 kΩ.

b) Determine an admittance matrix of a FET, neglect parasitic capacity.

c) Draw AC equivalent circuit.

Example 2.4

Suppose the component values gained in the example 3 (Fig. 2.17; $U_{DD} = 15$ V, $R_D = 1{,}795$ kΩ, $R_S = 205$ Ω, $R_G = 510$ kΩ). Now we use other FET - $I_{DSS} = 8$ mA, $U_P = -3$ V (selected at random). Determine a new quiescent current.

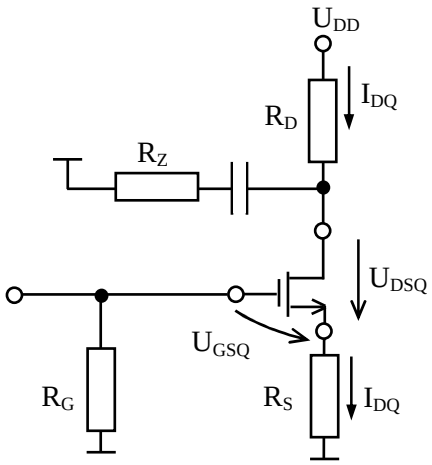


Fig. 2.31 Amplifier with R_S

PROBLEMS KEY 2

Ad example 2.1)

We know – solved example 2.1 – that $I_C = 0{,}6$ mA and β is 150. Thus

$$r_e = \frac{1}{y_{21}} \cong \frac{U_T}{I_C} = \frac{26 \cdot 10^{-3}}{0{,}6 \cdot 10^{-3}} = 43\ \Omega \qquad \frac{1}{y_{11}} \cong \beta \cdot r_e = 6\,500\ \Omega$$

If we neglect the Early voltage than $y_{22} = 0$; $y_{12} = 0$ – we suppose that input current is not a function of the voltage across collector and emitter.

Ad example 2.2)

We know – solved example 2.5 – that if $k = 2{,}96$ mA/V², $U_T = 2$ V, $U_A = 156$ V than $I_D = 5{,}2$ mA. We easily determine

$$g_m = 2 \cdot \sqrt{kI_D} = \sqrt{2{,}96 \cdot 10^{-3} \cdot 5{,}2 \cdot 10^{-3}} = 7{,}846 \text{ mS}$$

and also $r_d = 1/g_d = U_A/I_D = 156/5{,}2 \cdot 10^{-3} = 30 \text{ k}\Omega$.

Ad example 2.3 a)

It is known that in a quiescent point:

$$I_{DQ} = I_{DSS}(1 - U_{GS_Q}/U_P)^2 = 6 \cdot 10^{-3} \cdot (1 - (-2)/(-5))^2 = 2{,}16 \text{ mA}$$

and

$R_S I_{D_Q} = +2$ V – from this equation we get $R_S = 2/(2{,}16 \cdot 10^{-3}) = 926\ \Omega$.

Further

$$U_{DD} = R_D I_{D_Q} + U_{DS_Q} + R_S I_{D_Q}.$$

We suppose load resistance 10 kΩ. Thus we must choose $R_D \ll 10$ kΩ - or else it will degrade voltage gain of the circuit (Thévenin's theorem). Thus we choose $R_D = 2{,}2$ kΩ. We determine

$U_{DS_P} = U_{DS_Q} - U_P = -2 - (-5) = 3$ V – Fig. 2.11.

We choose

$U_{DD} = 12\ V$ – then $U_{DS_Q} = U_{DD} - R_D I_{D_Q} = 12 - 2{,}2 \cdot 10^3 \cdot 2{,}16 \cdot 10^{-3} = 5{,}25$ V

it is greater value than 3 V – the transistor is in the active (saturated) region – that is right quiescent point.

Ad example 2.3 b)

Now we can derive:

$$r_d = \frac{1}{g_d} = \frac{U_A}{I_{D_Q}} = \frac{100}{2{,}16 \cdot 10^{-3}} = 46{,}3 \text{ k}\Omega \qquad \Rightarrow \qquad g_d = 21{,}6 \cdot 10^{-6} \text{ S}$$

$$g_m = \frac{2I_{D_Q}}{U_{DS_Q} - U_P} = \frac{2 \cdot 2{,}16 \cdot 10^{-3}}{-2-(-5)} = 1{,}44 \text{ mS}$$

Ad example 2.3 c)

Equivalent AC you can see in fig. 2.32.

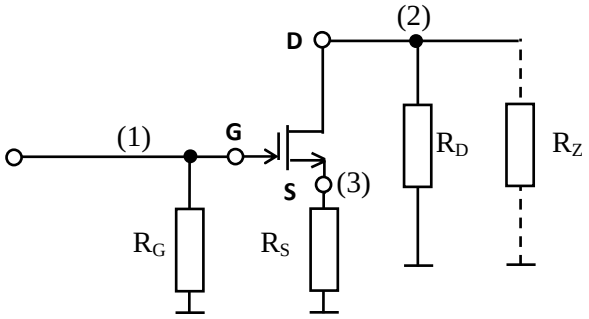


Fig. 2.32 Equivalent AC circuit of the amplifier in the Fig. 2.31

Ad example 2.4)

$I_D = 4{,}125$ mA – see solution in the solved example 2.4.

Basic texts

[1] Mikulec, M., – Havlíček, V.: Basic circuit theory. Vydavatelství ČVUT, Praha, 2005, ISBN 80-01-03172-1

[2] Mohylová, J.: Lineární obvody s elektronickými prvky-Sbírka příkladů, VŠB-TU Ostrava 2002, ISBN 80-248-0098-5

[3] Punčochář, J.: Lineární obvody s elektronickými prvky. Skriptum, VŠB-TU Ostrava 2002, ISBN 80-248-0040-3

[4] Mohylová, J. – Punčochář, J.: Elektrické obvody II. VŠB-TU Ostrava 2007, ISBN 978-80-248-1338-7

[5] Mohylová, J. – Punčochář, J.: Cvičení z Elektrických obvodů II. VŠB-TU Ostrava 2007, ISBN 978-80-248-1283-0

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

3. Modern amplifier structres

Time of study: 6 hours

Goals: The students will

- become familiar with modern amplifier structures
- be able to derive admittance matrixes of modern amplifier structures
- become familiar with VFA based on CFA structures
- be able to apply those matrixes to analyze linear electronics circuits

EXPLANATION

3.1 Basic principle of structures

There are only a few basic amplifying structures. Let us describe them (briefly).

3.1.1 Conventional circuit techniques – VFA (Voltage Feedback Amplifier)

An example (principle) of a conventional differential operational amplifier (OPA) you can see in Fig. 3.1. It can be seen in Fig. 3.1 that a capacitor C can be charged at the most with bias current I for rising and falling signals. We know that

$$i_C(t) = C \cdot \frac{du_C(t)}{dt} \quad \Rightarrow \quad \frac{du_C(t)}{dt} = \frac{i_C(t)}{dt} \quad \Rightarrow \quad \left.\frac{du_C(t)}{dt}\right|_{max} = \frac{i_{C_{max}}(t)}{dt}.$$

But the last equation defines ***slew rate*** (**SR**) of an OPA and it is evident that $i_{C_{max}} = \pm I$. Thus

$$SR = \frac{\pm I}{C} \tag{3.1}$$

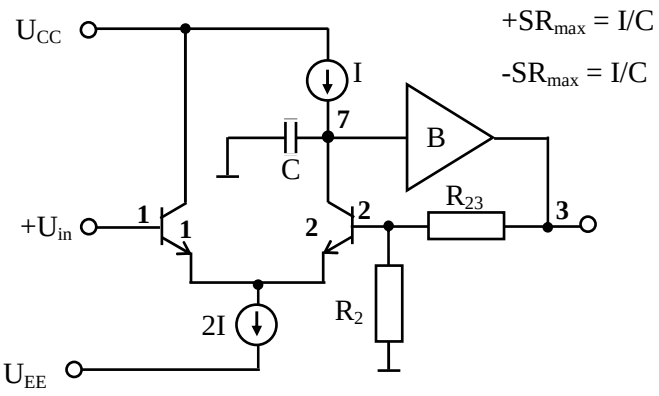


Fig. 3.1 Possible conventional structure of a differential amplifier; OPA or VFA

3.1.2 Current conveyor CCII (diamond transistor)

The basic current conveyor topology is shown in Fig. 3.2 below. Transistors T_1-T_4 buffer the input Y (B - diamond transistor convention, Y – current conveyor convention) – it is a bipolar complementary follower with a voltage gain coefficient higher than 0,98 (ideally 1). Notice that within the circuit, a unity gain buffer defines a voltage of an X - node (X – current conveyor convention, E - diamond transistor convention). In the ideal case, the output impedance of this buffer is zero (typically 10 to 100 Ω). D_1-T_5 and D_2-T_6 act as current mirrors that drive the high impedance node Z (Z – current conveyor convention, C - diamond transistor convention). Ideally it is valid

$$I_E = I_C \quad (I_X = I_Z)$$

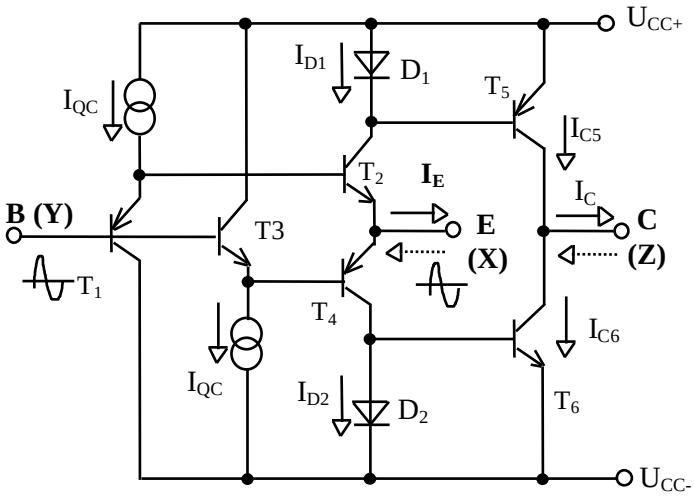


Fig. 3.2 Principle schematic of a CCII and diamond transistor

If we use E, B and C convention – we use *diamond transistor convention* (DT) (solid line current arrows). It is used to write

$$I_C = g_m \cdot U_{BE} \tag{3.2}$$

where

g_m – a transconductance of the diamond transistor (quiescent current I_{QC} defines a value of transconductance).

If we use X, Y and Z convention (it is the same structure) – we use current conveyor convention – it is CCII+ - *conventional positive current conveyor* (dotted line current arrow). We usually use other equations to describe it (ideal case):

$$U_X = U_Y \qquad I_Z = I_X \tag{3.3}$$

From this point of view we demand zero output resistance of X – node, which means that

$$g_m = \frac{1}{\text{output resistance of X mode}}$$

must tend theoretically to infinite!

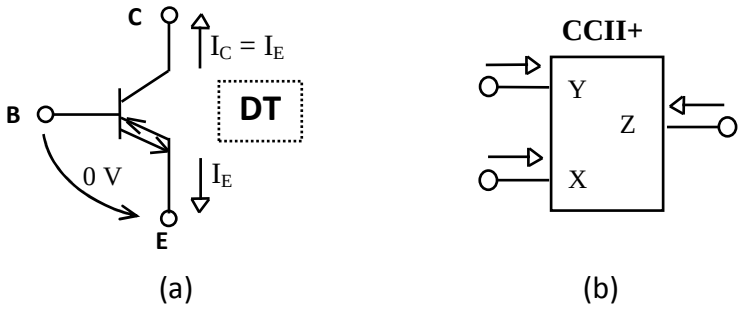


Fig.3 3 a) Symbol of DT; b) and CCII+

Used symbols of the diamond transistor (DT; early integrated circuit OPA660 for example) and CCII+ you can see in Fig. 3.3a and Fig. 3.3b.

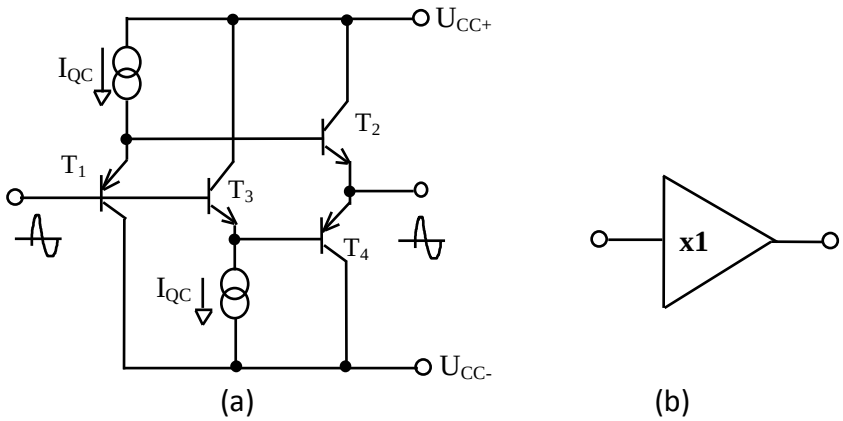


Fig.3.4 Principle schematic of buffer and its symbol

If we add a simple buffer now – Fig.3.4 – we have a *complete set* of needed functional *blocks to construct* (and understand) *modern amplifying structures*. It is evident that the structure in Fig. 3.4 is a bipolar complementary follower, too. Just current mirrors are omitted

3.1.3 Current feedback amplifier (CFA, transimpedance amplifier)

The basic CFA topology (principle - AD846) is shown in Fig. 3.5 below. The current I_X (error signal) is mirrored into high impedance (*transimpedance*) $\boldsymbol{Z}$, and the voltage developed across $\boldsymbol{Z}$ is equal to product $\boldsymbol{Z} \cdot I_X$. The current mirrors supply *current-on-demand* from the power supplies (theoretically no signal current limitations). There is no slew-rate limitation in an ideal CFA. The negative feedback forces the output voltage to a value that reduces input error current (I_X) to zero.

A denomination "current feedback amplifier - CFA" by itself is not a very good name (But the CFA is such as commonly used name today that we can hardly change it). From the point of view of "feedback" - the feedback in inverting and noninverting amplifiers, for example, is always "voltage" - dependent on the output voltage U_o (and not on the output current!).

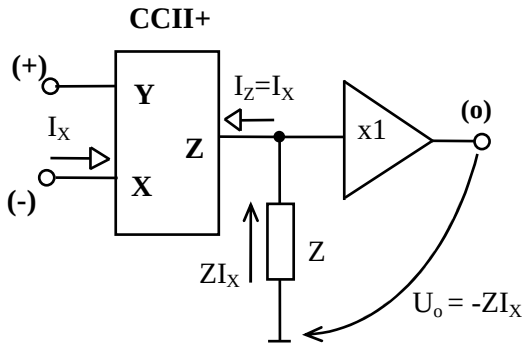


Fig. 3.5 Principle schematic of CFA; Z is usually the input resistance (generally impedance) of the output buffer (x1) – a load impedance may change this parameter.

But there are different properties of inverting inputs, there:

- VFA - the inverting input impedance is infinite;
- CFA - the inverting input impedance is zero.

The inverting input voltage is equal to noninverting input voltage - for CFA and VFA, but:

- VFA - the inverting input voltage is created by feedback circuit - an error signal is difference of voltages;
- CFA - the inverting input voltage is created by input follower - sometimes named OTA - an error signal is inverting input current I_X - converted on the output voltage by means of transimpedance).

If a CFA transimpedance is large enough (ideally infinite), a resulting inverting input current is the same for CFA and VFA - ideally zero.

The low inverting input impedance (and simple CFA construction) assures (for the same technology) better frequency and dynamic (high slew rate - 1000 V/µs) response of CFA compared to VFA (but worse DC properties of CFA).

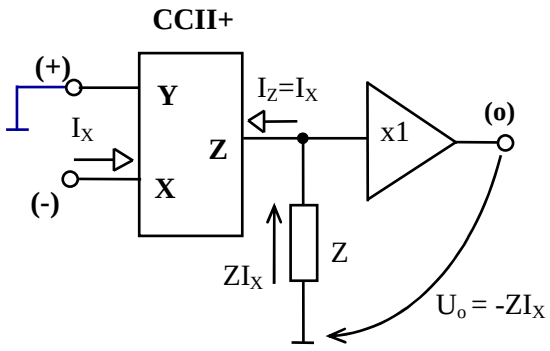


Fig. 3.6 Inverting transimpedance amplifier; node X voltage is ideally zero; output voltage $U_o = -ZI_X$; thus X is inverting input

We can investigate two possibilities without feedback – see Fig.3.6 and Fig. 3.7. An *open loop voltage gain* – Fig. 3.7 – is

$$A_0 = \frac{U_0}{U_+} = \frac{Z}{R_X} \tag{3.4}$$

In reality we must reason about *output resistance* R_{OX} *of the node* X (it is output resistance of the "input follower"), then we substitute simply $R_X \rightarrow R_X + R_{OX}$ and get

$$A_0 = \frac{U_0}{U_+} = \frac{Z}{R_X + R_{OX}} \tag{3.4a}$$

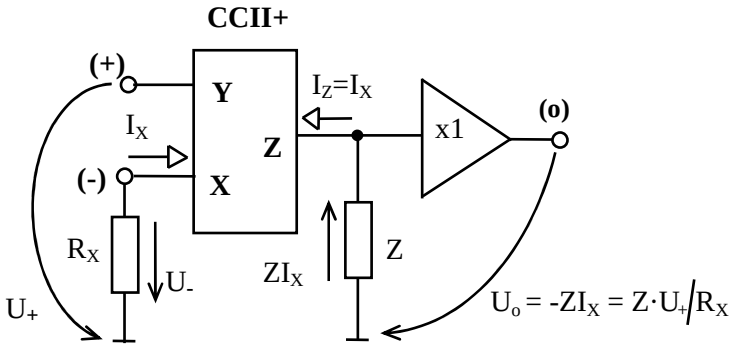


Fig. 3.7 Determination of an open loop voltage gain; ideally $U_- = U_+$; thus $I_X = -U_+/R_X$ (current arrow and voltage arrow have opposite direction, thus minus in Ohm's law) $\rightarrow$ $U_o = -ZI_X = U_+ \cdot Z/R_X$

The following results from this:

- ***Maximum possible open voltage gain*** is (if $R_X = 0$)

$$A_0 = \frac{U_0}{U_+} = \frac{Z}{R_{OX}} \tag{3.5}$$

- ***Open voltage gain depends on feedback resistor circuitry*** – R_X is equivalent resistance of the node X (Thévenin's theorem). No current I_X flows if R_X is infinite, and thus $A_0 \rightarrow 0$. This result is very important if we investigate system stability. CFA producers recommend appropriate value of feedback resistors for needed gains.

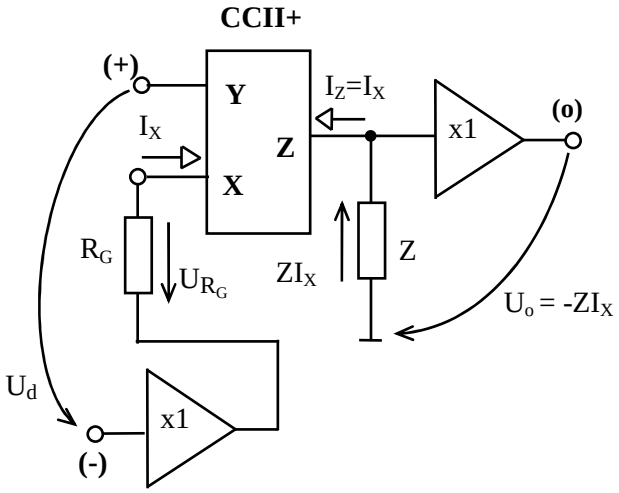


Fig.3.8 Principle VFA based on CFA

Remember, as feedback resistor (thus R_X) approaches zero ohms, the stability of a structure decreases (while the open loop bandwidth increases); thus placing diodes or capacitors across the feedback resistor (their signal resistance goes to zero) will cause oscillations in CFA structures.

3.1.4 VFA based on CFA structure

The VFA based on CFA (structure) are for example **EL5102/5103/5202/5203/5302)**

If we simply add one buffer – Fig. 3.8 – we get „VFA based on CFA“. Ideally it is

$$U_{R_G} = U_d$$

Thus

$$I_X = -U_d/R_G$$

and

$$U_0 = -Z \cdot I_X = U_d \cdot Z/R_G$$

The open loop voltage gain is

$$U_0 = Z/R_G \tag{3.6}$$

In reality we must reason about *output resistances of "input followers"* (R_O represents sum of output resistances of the two "input followers"), then we substitute simply $R_G \rightarrow R_G + R_O$ and get

$$A_0 = \frac{U_0}{U_+} = \frac{Z}{R_G + R_O} \tag{3.6b}$$

Now both inputs have high impedance (ideally infinite) and dynamic properties are the same with CFA structure – this structure gives current-on-demand, too – see Fig. 3.9, as well.

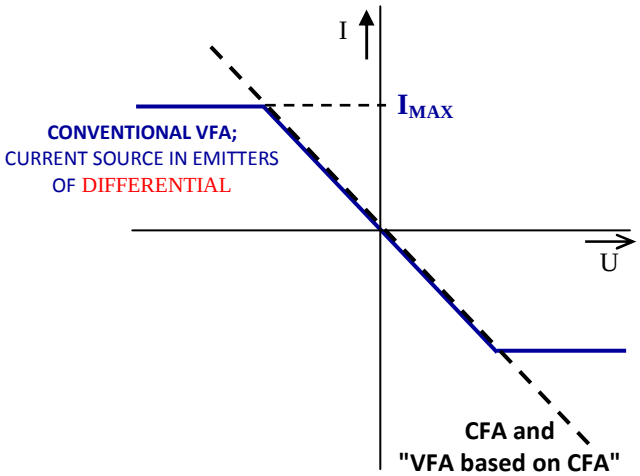


Fig. 3.9 Signal currents in amplifier structures as a function of differential voltage

We can place diodes or capacitors across the feedback resisto, now.

3.2 Basic signal (AC) and matrix models

If we know AC (linear signal) model of an amplifier structure, we can easily determine its admittance matrix model that way we describe all node (terminal) currents as a linear combination (linear function) of all node voltages.

3.2.1 VFA and matrix model

The assignment and sign convention for input (U_+, U_-) and output (U_o) voltages and currents (I_+, I_-, I_o) of differential amplifiers is shown in Fig. 3.10. The simplified model (idealized input properties; $I_+ = I_- = 0$) of differential amplifiers is shown in Fig. 3.10b. The equation below defines open output voltage.

$$U_o(I_o = 0) = U_{o_{open}} = AU_d = A(U_+ - U_-)$$

We can easily determine equations

$$I_+ = 0 \cdot U_+ + 0 \cdot U_- + 0 \cdot U_O$$

$$I_- = 0 \cdot U_+ + 0 \cdot U_- + 0 \cdot U_O$$

$$I_O = (U_O - A \cdot U_d)/R_O = -AG_O U_+ + AG_O U_- + G_O U_O$$

Input impedances (if I_+ and $I_- \neq 0$) of the VFA we can „put in the passive part“ of the analysed circuit.

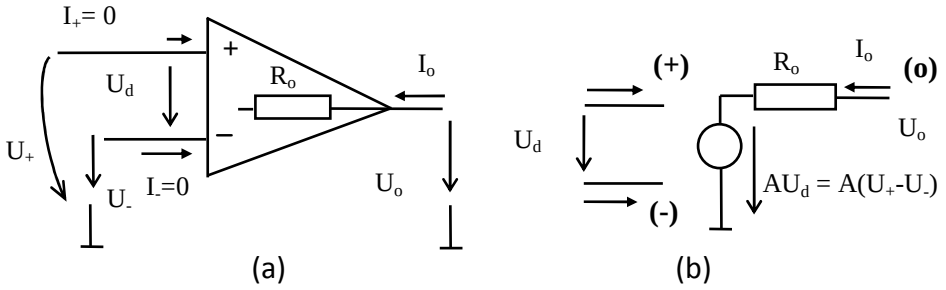


Fig. 3.10 a) A differential amplifier; b) a signal (AC) model

A matrix form (model) of equations is

	(+)	(-)	(o)				
(+)	0	0	0	U_+		I_+	
(-)	0	0	0	U_-	=	I_-	(3.7)
(o)	$-AG_o$	AG_o	G_o	U_o		I_o	

The eq. (3.7) describes the *matrix admittance model* of the differential amplifier with an output resistance $R_o = 1/G_o$ and a voltage gain A. If the amplifier is ideal (ideal VFA, OPA) than $A \to \infty$ and $U_+ = U_-$.

If only one input is needed (+ or -) we can non request input connect to the reference point and the corresponding raw and column of the matrix must be omitted.

3.2.2 One input voltage amplifier model

The assignment and sign convention for input (U_a) and output (U_b) voltages and currents (I_a, I_b) of an amplifier is shown in Fig. 3.11. The simplified model (idealized input properties; $I_a = 0$) of the amplifier is shown in Fig. 3.11b. The equation below defines open output voltage.

$$U_b(I_b = 0) = K \cdot U_a$$

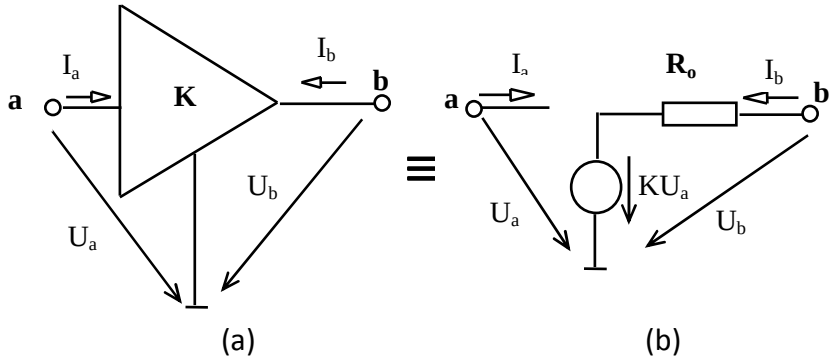


Fig. 3.11 a) A one input amplifier; b) a signal (AC) model

We can easily determine equations

$$I_a = 0 \cdot U_a + 0 \cdot U_b$$
$$I_b = \frac{U_b - KU_a}{R_o} = -KG_oU_a + G_oU_b$$

Thus corresponding admittance model is

	a	b			
a	0	0	U_a	=	I_a
b	$-KG_o$	G_o	U_b		I_b

(3.8)

3.2.3 Current conveyor signal and matrix model

The assignment and sign convention for input and output voltages and currents is shown in Fig. 3.12.

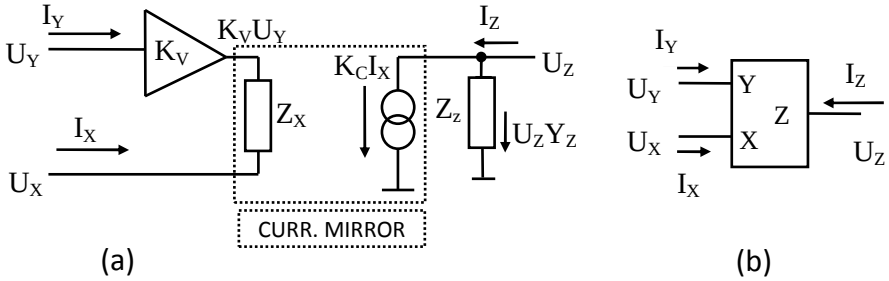


Fig. 3.12 a) Signal model of the three-port second-generation current conveyor; b) symbol of current conveyor

The ground (reference) terminal provides a reference point for the three others. The simplified signal model is shown in Fig.3.12a. The terminal Y is the input of a voltage amplifier (voltage gain K_V is 1 or -1 for the ideal case), input current $I_Y = 0$ for the ideal case. The terminal X is the output of the voltage amplifier; I_X is the current through output impedance Z_X of the voltage amplifier. Z_X is zero for the ideal case, consequently $U_X = K_V U_Y$ for the ideal case.

The output current I_Z (Z - output terminal) will mirror the current I_X (by means of current mirrors; K_C - current gain):

$$I_Z = K_C I_X$$

An admittance $Y_z = 1/Z_z$ (zero for the ideal case) represents real properties of the current source (Fig.3.12a).

All possibilities are summarised in a table 3.1.

Table 3.1. Different types of three-port second -generation current conveyors

K_V	K_C	NAMED	SYMBOL
1	1	*conventional positive current conveyor*	CCII+
1	-1	*conventional negative current conveyor*	CCII-
-1	1	*inverting positive current conveyor*	ICCII+
-1	-1	*inverting negative current conveyor*	ICCII-

From the Fig. 3.12a we can easily determine equations

$$I_X = (U_X - K_V U_Y)/Z_X = Y_X U_X - K_V Y_X U_Y + 0 \cdot U_Z$$
$$I_Y = 0 \cdot U_X + 0 \cdot U_Y + 0 \cdot U_Z$$
$$I_Z = K_C I_X + U_Z/Z_Z = K_C Y_X U_X - K_C K_V Y_X U_Y + Y_Z U_Z$$

The matrix form of these equations is

$$\begin{array}{c|ccc} & X & Y & Z \\ \hline X & Y_X & -K_V Y_X & 0 \\ Y & 0 & 0 & 0 \\ Z & K_C Y_X & -K_C K_V Y_X & Y_Z \end{array} \begin{bmatrix} U_X \\ U_Y \\ U_Z \end{bmatrix} = \begin{bmatrix} I_X \\ I_Y \\ I_Z \end{bmatrix} \qquad (3.9)$$

Equation (3.9) thus describes all *four different types* of three-port current conveyors from the Table 3.1.

3.2.4 CFA signal and matrix model

The assignment and sign convention for input and output voltages and currents is shown in Fig. 3.13.

A terminal (+) is an input of a voltage follower (ideal case $K_V = 1$, in general case $K_V \neq 1$). An input current $I_+ = 0$ for the ideal CFA. A terminal (-) is an output of the follower; I_- is the

current through an output resistance R_{O1} of the input follower. R_{O1} is zero for the ideal CFA consequently $U_+ = U_-$ for the ideal case.

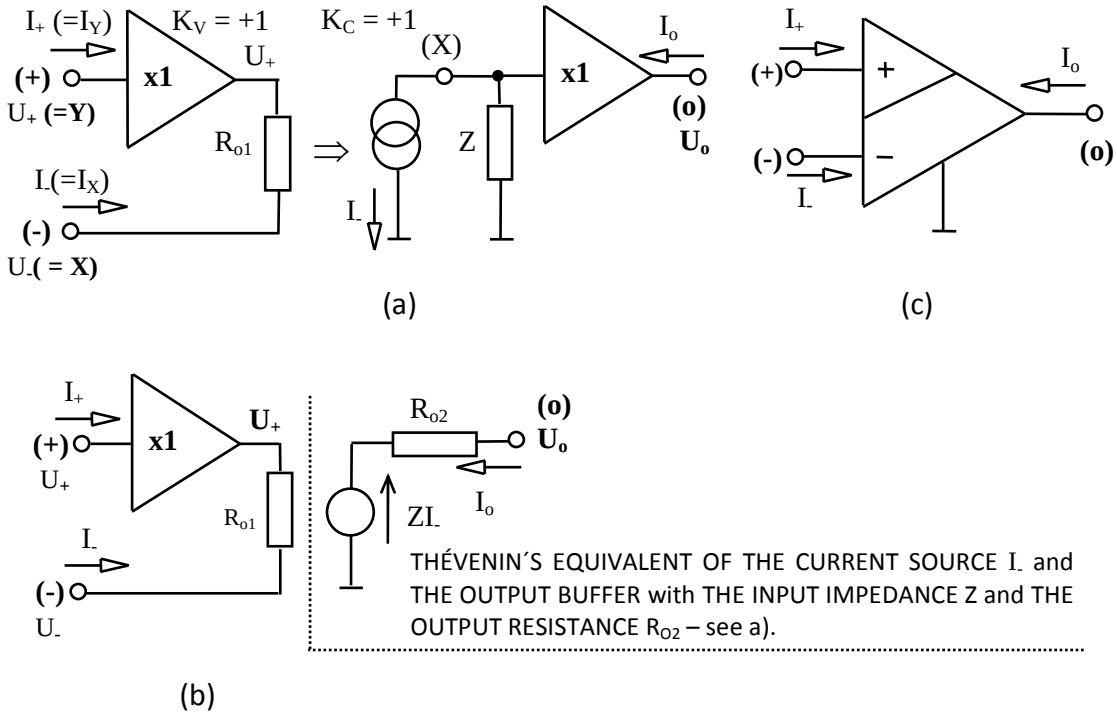


Fig.3. 13 a) A basic signal model of CFA; b) rearranging by means Thévenin's theorem; c) possible symbol of a CFA

The output voltage U_o will follow the current I_- thanks to a *transimpedance* Z (and current mirrors, of course). An output resistance R_{O2} represents real properties of the voltage source controlled by I- (zero for the ideal CFA).

From the Fig. 3.13b we can easily determine equations – we suppose:

$I_+ = 0$; $G_{O1} = 1/R_{O1}$; $G_{O2} = 1/R_{O2}$; K_V is voltage gain of the follower.

$$I_+ = 0 \cdot U_+ + 0 \cdot U_- + 0 \cdot U_O$$
$$I_- = (U_- - KU_+)/R_{O1} = |_{\text{for } K_V=1}| = -G_{O1}U_+ + G_{O1}U_- + 0 \cdot U_O$$
$$I_O = \big(U_O - (-ZI_-)\big)/R_{O2} = -G_{O1}G_{O2}ZU_+ + G_{O1}G_{O2}ZU_- + G_{O2}U_O$$

A matrix form (model) of the equations is

	(+)	(-)	(o)			
(+)	0	0	0	U_+		I_+
(-)	$-G_{O1}$	G_{O1}	0	U_-	=	I_-
(o)	$-G_{O1}G_{O2}Z$	$G_{O1}G_{O2}Z$	G_{O2}	U_o		I_o

(3.10)

3.2.5 VFA based on CFA signal and matrix model

Formally is valid the same signal model as for conventional VFA, thus see equation (3.7).

Only frequency dependence of voltage gain will be different (DC gain usually lower, dynamic properties better – high frequency, high slew rate).

3.2.6 Transadmitance amplifier (TAA, OTA) signal and matrix model

An idealised signal model of TAA (zero input currents I_+ and I_-; thus infinite input impedance) is in Fig. 3.14 - voltage-controlled (U_d) current-source (I_O); thus infinite output impedance. G_m is mutual conductance - commonly transadmitance $Y_m(j\omega)$.

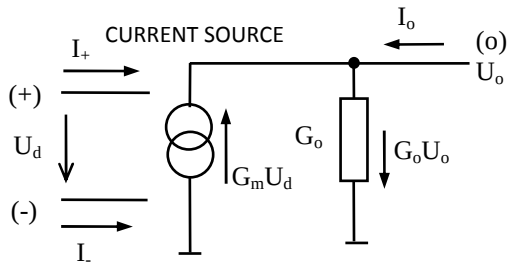


Fig. 3.14 The idealized signal model of the TAA

From the Fig.14 it can be easy determined that

$$I_+ = 0 \cdot U_+ + 0 \cdot U_- + 0 \cdot U_O$$
$$I_- = 0 \cdot U_+ + 0 \cdot U_- + 0 \cdot U_O$$
$$I_O = -G_m \cdot U_d + G_O U_O = -G_m(U_+ + U_-) + G_O U_O = -G_m \cdot U_+ + G_m \cdot U_- + G_O U_O$$

Hence a model is

	(+)	(-)	(o)				
(+)	0	0	0		U_+	=	I_+
(-)	0	0	0		U_-		I_-
(o)	$-G_m$	G_m	G_o		U_o		I_o

(3.11)

3.3 Conclusion

In the matrices models we can easily define frequency responses, too. We can do substitutions: $A \rightarrow A(j\omega)$; $G_m \rightarrow Y_m(j\omega)$; $Z \rightarrow Z(j\omega)$; $K \rightarrow K(j\omega)$; etc. Analysis will be more difficult, of course, but more useful, as well. Generally

$$X = \frac{X_o \omega_1}{p + \omega_1} \tag{3.12}$$

where X_o is DC "quality" and ω_1 defines first pole of this parameter X – se fig. 3.15, too. If it describes an OPA gain thus $X = A$ and $X_o \omega_1 = \omega_T$ is gain bandwidth product.

It is really evident that there are no outstanding difficulties involved in incorporating the described devices (models) into the framework of conventional linear network theory. Usefulness of the signal models with parameters that are specified on manufacturers` data sheets is evident, too.

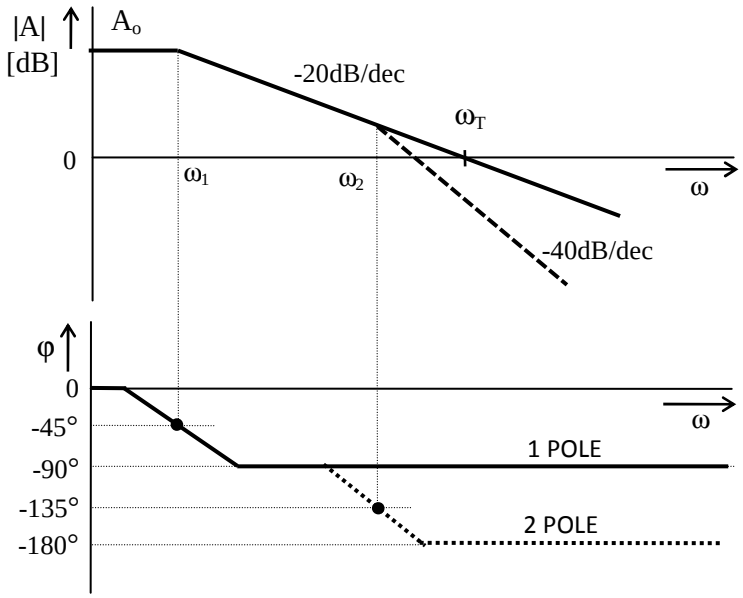


Fig. 3.15 One and two-pole models of the real OPA gain.

Summary

1) VFA or "VFA based on CFA" – with negative feedback (as is generally used with op – amps except as oscillators or comparators), the two input terminals are forced to the same potential by the feedback (virtual ground).
2) VFA or "VFA based on CFA" are voltage controlled voltage sources (ideally: input impedances are infinite, the output impedance is zero; the open loop gain is infinite, the bandwidth is infinite).
3) "Real" versus "ideal" VFAs:

- There are unwanted currents at the inputs – the input resistance is not infinite.
- There offset currents and voltages.
- Signals applied to both inputs (which should not be amplified) are amplified to some extent.
- The output resistance is not zero.
- Op – amps do not have infinite frequency response.
- The slew rate specifies the maximum rate at which the op – amp can swing its output.

4) Current conveyor (diamond transistor) is a current controlled current source (ideally: the current input impedance is zero, the voltage input impedance is infinite, the output impedance is infinite; voltage of the "current input" is equal to voltage of the "voltage input" – even without feedback – always; the bandwidth is infinite).
5) CFA is a current controlled voltage source [ideally: the current input (inverting) impedance is zero, the voltage input (noninverting) impedance is infinite, the output impedance is zero; inverting input voltage is equal to noninverting input voltage – even without feedback; the bandwidth is infinite; there is no slew – rate limitation – current on demand].
6) VFA based on CFA – we simply propend an follower and resistance to the CFA inverting input and so we get required current – so we get inverting voltage input.
7) Dynamic and frequency characteristics of the CFA and "VFA based on CFA" are much better than for conventional VFA. On the contrary DC properties are worse.
8) If we know linear signal models of amplifying structures, we can determine their admittance matrices that way we describe all node currents as a linear combination of all node voltages.

Questions 3

You can find the answers in this text.

1. Define basic properties of a conventional VFA.
2. Define basic properties of a CFA.
3. Define basic properties of a CCII.
4. Define basic properties of an OTA.
5. Define basic properties of VFA based on CFA.

Problems 3

1. Identify ω_T of an OPA if it is $A_o = 10^6$ and $f_1 = 5$ Hz.
2. Derive an admittance model (matrix) of an idealized operational amplifier if you know that: its $A(j\omega) = A_o\omega_1/(p + \omega_1)$; A_o is DC gain; ω_1 is a dominant pole of the gain; $A_o = 10^6$; $f_1 = 5$ Hz; a output resistance $R_o = 100$ Ω.
3. Make out a model of CFA transimpedance if $Z_o = 5 \cdot 10^5$ Ω and its dominant pole is 50 kHz.

4. Derive an admittance model (matrix) of an idealized CFA if you know that: output resistances are $R_{o1} = R_{o2} = 50\ \Omega$ and transimpedance is defined as $Z_o = 5 \cdot 10^5\ \Omega$; its dominant pole is 50 kHz.
5. Entitle correctly current conveyor if is valid: $K_V = 1$ and $K_C = -1$. Derive its admittance matrix.

PROBLEMS KEY 3

Ad 1) Use the equation (3.12) and Fig. 3.15.

Ad 2) Use the equation (3.7).

Ad 3) Use the equation (3.12).

Ad 4) Use the equation (3.10).

Ad 5) See the Table 3.1 and the equation (3.9).

Basic texts

[1] Punčochář, J.: Admittance models of modern linear amplifying structures. Transactions of the VŠB – TU Ostrava, 1, 2003, vol. VI, pp 151 – 161, ISBN 80-248-0223-6

[2] Mohylová, J.: Lineární obvody s elektronickými prvky-Sbírka příkladů, VŠB-TU Ostrava 2002, ISBN 80-248-0098-5

[3] Punčochář, J.: Lineární obvody s elektronickými prvky. Skriptum, VŠB-TU Ostrava 2002, ISBN 80-248-0040-3

[4] Punčochář, J.: Operační zesilovače v elektronice. BEN – Praha, 2002, ISBN 80-7300-059-8

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

4. Basic feedback theory, stability

Time of study: 6 hours

Goals: the student should be able to

- define positive and negative feedback
- be able to derive admittance matrixes of modern amplifier structures
- describe an influence of the feedback on frequency properties of amplifiers
- describe an influence of the feedback on impedance properties of amplifiers
- define stability of electronics structures

EXPLANATION

4.1 Feedback – principle

A principle of feedback you can see in Fig. 4.1. A part of *output signal* (X_2; excited quantity) is carried back as a *feedback signal* (X_Z) via *feedback network* with a transfer function $\hat{P}_Z = \hat{X}_Z/\hat{X}_2$ into the input port of a *general two-port – forward path*, with the transfer function $\hat{P}_a = \hat{X}_2/\hat{X}_1$; $\hat{X}_1$ - the input quantity of the forward path (controlling quantity). Here the signal X can be both voltages and currents. Transfer ratios are generally functions of frequency.

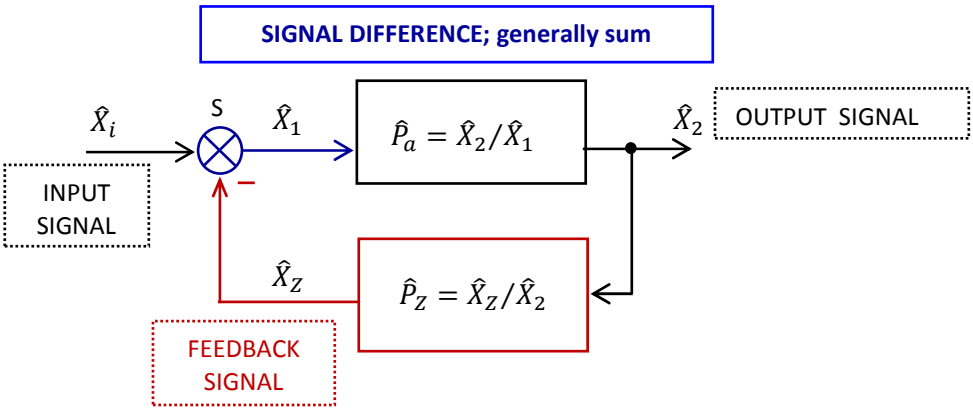


Fig. 4.1 Feedback system block diagram (a negative feedback is supposed as basic state)

A summation element (S) produces

$$\hat{X}_1 = \hat{X}_i - \hat{X}_Z \tag{4.1}$$

if we suppose a negative feedback as a basic analyzed state (generally we can write $\hat{X}_1 = \hat{X}_i + \hat{X}_Z$; it gives other formal models, but the same conclusions); $\hat{X}_i$ - the input signal (exciting quantity).

We can easily derive

$$\hat{X}_2 = \hat{P}_a\hat{X}_1 = \hat{P}_a(\hat{X}_i - \hat{X}_Z) = \hat{P}_a(\hat{X}_i - \hat{P}_Z\hat{X}_2)$$

thus

$$\hat{X}_2(1 + \hat{P}_a\hat{P}_Z) = \hat{P}_a\hat{X}_i$$

After rearranging we get *Black relation* (closed loop gain)

$$\hat{P} = \frac{\hat{X}_2}{\hat{X}_i} = \frac{\hat{P}_a}{1+\hat{P}_a\hat{P}_Z} \tag{4.2}$$

and it is a basis for the investigation of all the transfer properties of feedback circuits. Its denominator appears in many relations describing these properties and it is usually called the *feedback factor (činitel zpětné vazby).*

The feedback is *negative* (degenerative) if

$$|\hat{P}| = \left|\frac{\hat{P}_a}{1+\hat{P}_a\hat{P}_Z}\right| = \frac{|\hat{P}_a|}{|1+\hat{P}_a\hat{P}_Z|} < |\hat{P}_a|$$

thus

$$|1 + \hat{P}_a\hat{P}_Z| > 1 \tag{4.3}$$

The feedback is *positive* (regenerative) if

$$|\hat{P}| = \left|\frac{\hat{P}_a}{1+\hat{P}_a\hat{P}_Z}\right| = \frac{|\hat{P}_a|}{|1+\hat{P}_a\hat{P}_Z|} > |\hat{P}_a|$$

thus

$$|1 + \hat{P}_a\hat{P}_Z| < 1 \tag{4.4}$$

In actual devices there are always some storing elements (L, C), at least parasitic ones. It is evident, that owing to the different phase shifts at different frequencies, the feedback maybe negative in some frequency region and positive in another one (sinusoidal steady state).

The limiting case of positive feedback resulting from the equation

$$1 + \hat{P}_a\hat{P}_Z \to 0$$

is called *critical* and it can be attained after fulfilling the conditions:

$$1 + Re[\hat{P}_a\hat{P}_Z] + jIm[\hat{P}_a\hat{P}_Z] = 0 \quad \Rightarrow$$

$$1 + Re[\hat{P}_a\hat{P}_Z] = 0 \tag{4.5}$$

$$Im[\hat{P}_a\hat{P}_Z] = 0 \tag{4.6}$$

The circuit behaves as a *generator* – no exciting quantity is needed (it is undesirable effect if we project amplifiers or filters, but it is desirable effect if we project generators: narrow-band feedback = sinus generator; wide-band feedback = multivibrator, comparator). Closed loop gain goes to infinity. An oscillation amplitude results from non-linear properties of investigated (analyzed) electronic structure.

4.2 Transfer function sensitivity

The amplification of actual amplifiers is temperature and time dependent (because of aging). To investigate this influence let us consider now $P_a \equiv K$ (real transfer function) and $P_Z \equiv \beta < 1$ (the feedback path is formed by a divider; generally there exist structures with "active" feedback path – for example logarithmic amplifier – its $P_Z \equiv \beta > 1$ - it is more complicated situation). Then

$$P = \frac{K}{1+\beta K} \tag{4.7}$$

where

$1 + \beta K$ - is feedback factor.

Supposing that the transfer function of the forward path (of the amplifier) is increased by ΔK, the increment of the resultant transfer is

$$\Delta P = P(K + \Delta K) - P(K) = \frac{K+\Delta K}{1+\beta(K+\Delta K)} - \frac{K}{1+\beta K} = \frac{(K+\Delta K)\cdot(1+\beta K)-K[1+\beta(K+\Delta K)]}{[1+\beta(K+\Delta K)]\cdot(1+\beta K)} \Rightarrow$$

$$\Delta P = \frac{\Delta K}{[1+\beta(K+\Delta K)]\cdot(1+\beta K)}$$

and the relative (fractional) change

$$\frac{\Delta P}{P(K)} = \frac{\Delta K/K}{1+\beta(K+\Delta K)} \tag{4.8}$$

We can easily get normed form:

$$\frac{\Delta P/P(K)}{\Delta K/K} = \frac{1}{1+\beta(K+\Delta K)}; \quad \text{if} \quad \Delta \to 0 \quad \text{than} \quad \frac{\delta P/P(K)}{\delta K/K} \to \frac{1}{1+\beta K} \tag{4.9}$$

This means a change of P is $(1 + \beta K)$ – times less than relative change of original (forward path; amplifier) transfer K.

The higher is the feedback factor, the smaller is the influence of the properties of the forward path (of the amplifier). Particularly for $\beta K \gg 1$ it is valid

$$P = \frac{K}{1+\beta K} \to \frac{1}{\beta} \tag{4.10}$$

The resultant transfer is given by only passive elements in the feedback path, regardless of the properties of the forward path.

4.3 Frequency characteristic

Let us consider an amplifier with the complex transfer function

$$\hat{P}_a = K \cdot \frac{\omega_0}{j\omega+\omega_0} = K \cdot \frac{1}{1+j\omega/\omega_0} \tag{4.11}$$

the modulus characteristic of which is introduced in Fig. 4.2.

Supposing real negative feedback $\hat{P}_Z \equiv \beta$ we get from the equations (4.2) and (4.11)

$$\hat{P} = \frac{\frac{K}{1+j\omega/\omega_0}}{1+\beta\cdot\frac{K}{1+j\omega/\omega_0}} = \frac{K}{1+j\omega/\omega_0+\beta\cdot K} = \frac{K}{1+\beta\cdot K}\cdot\frac{1}{1+j\frac{\omega}{(1+\beta\cdot K)\cdot\omega_0}} \tag{4.12}$$

$$P_{ID} = \frac{K}{1+\beta\cdot K} \tag{4.13}$$

is an *ideal transfer function* (frequency independent now).

$$\hat{E} = \frac{1}{1+j\frac{\omega}{(1+\beta\cdot K)\cdot\omega_0}} = \frac{1}{1+j\frac{\omega}{\omega_3}}; \qquad \omega_3 = (1+\beta\cdot K)\cdot\omega_0 \tag{4.14}$$

is an *error function* (frequency dependent).

As seen, the characteristic angular frequency

$$\omega_3 = \omega_0(1+\beta K)$$

of the amplifier with the feedback is $(1+\beta K)$ – times increased – see the new modulus characteristic (dotted line) in fig. 4.2.

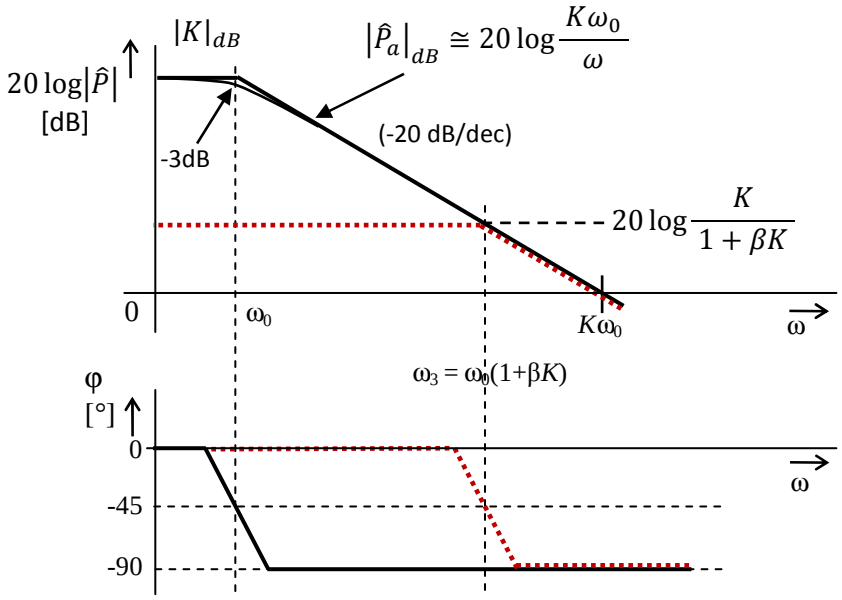


Fig.4.2 Modulus and phase characteristics; solid lines – no feedback; dotted line – with feedback

4.4 Input impedance

According to the type of feedback the original input impedance of the forward path ($\hat{Z}_{v_1}$) can be radically changed. Let us investigate the relations among the circuit variables in the input circuit of a *series connection*, see Fig. 4.3. Derivation of feedback information (signal) is not investigated now. It is steadily valid

$$\hat{U}_Z = \hat{U}_1\hat{P}_a\hat{P}_Z \quad \hat{I}_Z = \hat{I}_1\hat{P}_a\hat{P}_Z$$

Product $\hat{P}_a\hat{P}_Z$ *has not unit (physical).*

The "feedback" input impedance $\hat{Z}_{v_s}$ (loading impedance of signal source) of the structure in Fig. 4.3a (series connection) is

$$\hat{Z}_{v_s} = \frac{\hat{U}_i}{\hat{I}_i} = \frac{\hat{U}_1 + \hat{U}_z}{\hat{I}_i} = \frac{\hat{U}_1 + \hat{U}_1 \hat{P}_a \hat{P}_Z}{\hat{U}_1 / \hat{Z}_{v_1}} = \hat{Z}_{v_1} \cdot \left(1 + \hat{P}_a \hat{P}_Z\right) \tag{4.15}$$

If a *negative feedback* and the *series connection* are used then $\left|1 + \hat{P}_a \hat{P}_Z\right| > 1$ and thus the *module of input impedance increases*:

$$\left|\hat{Z}_{v_s}\right| = \left|\hat{Z}_{v_1}\right| \cdot \left|1 + \hat{P}_a \hat{P}_Z\right| \geq \hat{Z}_{v_1} \tag{4.16}$$

Similarly for the ***parallel connection*** – Fig. 4.3b – the input impedance $\hat{Z}_{v_p}$ can be expressed as follows:

$$\hat{Z}_{v_p} = \frac{\hat{U}_i}{\hat{I}_i} = \frac{\hat{U}_1}{\hat{I}_1 + \hat{I}_Z} = \frac{\hat{Z}_{v_1} \hat{I}_1}{\hat{I}_1 + \hat{I}_1 \hat{P}_a \hat{P}_Z} = \frac{\hat{Z}_{v_1}}{1 + \hat{P}_a \hat{P}_Z} \tag{4.17}$$

If a *negative feedback* and the *parallel connection* are used then $\left|1 + \hat{P}_a \hat{P}_Z\right| > 1$ and thus the *module of input impedance decreases*, or a module of input admittance

$$\hat{Y}_{v_p} = \frac{1}{\hat{Z}_{v_p}} = \frac{1 + \hat{P}_a \hat{P}_Z}{\hat{Z}_{v_1}} = \hat{Y}_{v_1} \cdot \left(1 + \hat{P}_a \hat{P}_Z\right) \tag{4.18}$$

increases (it is identical statement)

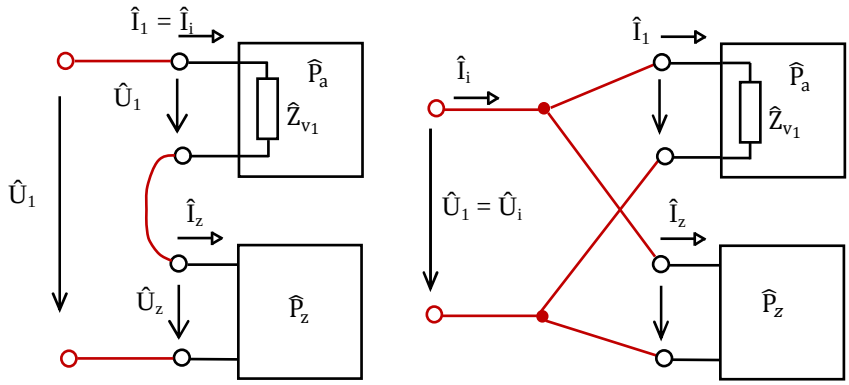


Fig. 4.3 a) A *series connection* (of the amplifier input and feedback output $\hat{I}_1 = \hat{I}_i = \hat{I}_z$ - *"series requirement"*) – ideally we suppose that the output impedance of the feedback circuit is zero (an ideal voltage source; feedback voltage $\hat{U}_z$ is no function of the input current $\hat{I}_1 = \hat{I}_i$;

b) A *parallel connection* (of the amplifier input and feedback two-port output; $\hat{U}_1 = \hat{U}_i = \hat{U}_z$ - *"parallel requirement"*;) – ideally we suppose that the output impedance is infinite (an ideal current source; feedback current $\hat{I}_z$ is no function of the input voltage $\hat{U}_1 = \hat{U}_i$)

4.5 Output impedance

Owing to the Thévenin's theorem the terminal output impedance (of the linear circuit) can be determined as a ratio of the open-circuit voltage $\hat{U}_{2_p}$ and the short-circuit current $\hat{I}'_{2_K}$.

In the case of **voltage feedback** is suitable to use the model in Fig. 4.4a. The *open voltage* $\hat{U}_{2_p}$ is (Fig. 4.4a; $R_Z \to \infty$; the voltage feedback is **„ON"**; we suppose voltage input $\hat{U}_i$):

$$\hat{P} = \frac{\hat{U}_{2p}}{\hat{U}_i} = \frac{\hat{P}_a}{1+\hat{P}_a\hat{P}_Z} \quad \Rightarrow \quad \hat{U}_{2_p} = \frac{\hat{U}_i \cdot \hat{P}_a}{1+\hat{P}_a\hat{P}_Z} \tag{4.19}$$

The ***short-circuit current*** $\hat{I}'_{2_K}$ in this case is (Fig. 4.4a; $R_Z \to 0$):

$$\hat{I}'_{2_K} = \frac{\hat{U}_i \cdot \hat{P}_a}{\hat{Z}_{v_2}} \tag{4.20}$$

since, owing to the short-circuited output, voltage feedback is set out of function (the voltage feedback is **„OFF"** now). Therefore the output impedance (with voltage feedback) will be

$$\hat{Z}_{v_{2n}} = \frac{\hat{U}_{2p}}{\hat{I}'_{2_K}} = \frac{\hat{Z}_{v_2}}{1+\hat{P}_a\hat{P}_Z} \tag{4.21}$$

If a *negative feedback* and the *voltage feedback* are used then $\left|1 + \hat{P}_a\hat{P}_Z\right| > 1$ and thus the *module of output impedance* $\left|\hat{Z}_{v_{2n}}\right|$ *falls* below the output impedance module $\left|\hat{Z}_{v_2}\right|$ of the amplifier.

If $\left(1 + \hat{P}_a\hat{P}_Z \to \infty\right)$ (an ideal case) then $\left|\hat{Z}_{v_2}\right| \to 0$, the system behaves as an *ideal voltage source.*

In the case of *current feedback* it is suitable to use the model in Fig. 4.4b – Norton's representation of the amplifier output. The *open voltage* $\hat{U}_{2_p}$ is (Fig. 4.4b; $R_Z \to \infty$; we suppose current input $\hat{I}_i$):

$$\hat{U}_{2_p} = \frac{-\hat{P}_a\hat{I}_i}{\hat{Y}_{v_p}} = \frac{1+\hat{P}_a\hat{P}_Z}{\hat{Z}_{v_1}} = -\hat{P}_a\hat{I}_i\hat{Z}_{v_2} \tag{4.22}$$

since at the open-circuit output feedback is set out of function; $\hat{I}_1 = \hat{I}_i$ and $\hat{I}_2 = 0$; the current feedback is **„OFF".**

The *short-circuit current* $\hat{I}'_{2_K}$ in this case is (fig. 4.4b; $R_Z \to 0$)

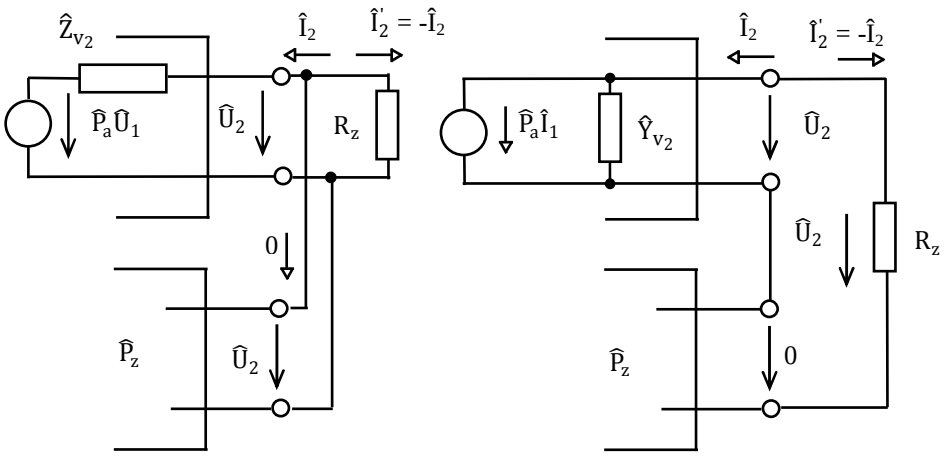
$$\hat{I}'_{2_K} = -\hat{I}'_{2_K} = -\frac{\hat{P}_a\hat{I}_i}{1+\hat{P}_a\hat{P}_z} \tag{4.23}$$

because of that current feedback is **"ON",** now. Therefore the output impedance (with current feedback) will be

$$\hat{Z}_{v_{2I}} = \frac{\hat{U}_{2p}}{\hat{I}'_{2_K}} = \frac{-\hat{P}_a\hat{I}_i\hat{Z}_{v_2}}{\frac{-\hat{P}_a\hat{I}_i}{1+\hat{P}_a\hat{P}_Z}} = \hat{Z}_{v_2} \cdot \left(1 + \hat{P}_a\hat{P}_z\right) \tag{4.24}$$

If a *negative feedback* and the *current feedback* are used then $\left|1 + \hat{P}_a\hat{P}_Z\right| > 1$ and thus the *module of output impedance* $\left|\hat{Z}_{v_{2I}}\right|$ *increases* over the output impedance module $\left|\hat{Z}_{v_2}\right| =$

$1/\hat{Y}_{v_2}$ of the amplifier. If $\left(1+\hat{P}_a\hat{P}_Z\right)\to\infty$ (an ideal case) then $\left|\hat{Z}_{v_2 I}\right|\to\infty$, the system behaves as an *ideal current source.*



(a) (b)

Fig. 4.4 a) A *voltage feedback* (the amplifier output and feedback two-port input in parallel – identical voltages) – ideally we suppose that the input impedance of the feedback circuit is infinite – no feedback current flows;

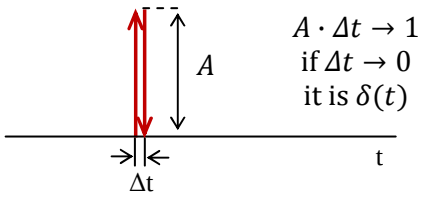
b) A current feedback (the amplifier output and feedback two-port input in series – identical currents) – ideally we suppose that the input impedance of the feedback circuit is zero – no feedback circuit input voltage

4.6 Feedback system stability

We know, due to the presence of storing elements, that the feedback can be positive in some frequency range and negative in another one. This also means that in circuits projected with a negative feedback, in a certain region, positive feedback may arise, thus, *unwanted oscillations* may appear (oscillators – wanted oscillations) – the *circuit is unstable*.

The basic idea of stability investigation results from the properties of the closed – loop transfer function (2). This is always a *rational function* and actually it is the Laplace transform of the *unit impulse response* – see BOX below (in harmonic steady state it is simply $p = j\omega$).

BOX



$A\cdot\Delta t\to 1$ if $\Delta t\to 0$ it is $\delta(t)$

$$X_i(t)=\delta(t);\quad \mathcal{L}[\delta(t)]=1$$

$$P(p)=\frac{X_2(p)}{X_i(p)}=\frac{P_a(p)}{1+P_a(p)\cdot P_z(p)}$$

For $X_i(p)=\mathcal{L}[\delta(t)]=1$ is

$$X_2(p)=\frac{P_a(p)}{1+P_a(p)\cdot P_z(p)}\cdot X_i(p)=\frac{P_a(p)}{1+P_a(p)\cdot P_z(p)}=P(p)$$

The nature of the unit impulse response is given by the types of its *poles* [poles are roots of equation $1 + P_a(p) \cdot P_z(p) = 0 \Rightarrow (p - p_{p_1}) \cdot (p - p_{p_2}) \cdots (p - p_{p_n}) = 0$]. If they are *negative* or complex with *negative real part*, the unit impulse response is composed of decreasing exponential $exp(p_{p_n} t)$ functions or sinusoidal functions with decreasing amplitude. The steady value of the output quantity is equal to zero. Thus, the time axis is an asymptote for the unit impulse response and we say the circuit is *asymptotically stable*.

Any positive pole or a pole with the positive real part indicates the presence of an increasing exponential function or a sinusoidal function with increasing amplitude.

The previous considerations can be reduced to the following statement:

The transfer function of an asymptotically stable circuit has poles only in the left half of the complex plain.

To obtain the poles, an algebraic (characteristic) equation must be solved which might be tedious in higher order circuits. Therefore other methods needing only knowledge of the coefficients of the transfer function denominator, or the frequency characteristic of the open-loop transfer function, were elaborated – *stability criterions*.

4.6.1 Nyquist stability criterion

$P_a(p) \cdot P_z(p)$ - ***open loop transfer function*** (characteristic, ***loop gain***) – feedback loop is open at some suitable point. See equations (4.5) and (4.6), too.

The transfer function has only poles in the left-half on the complex plain (with exception of a single or double pole at $p = 0$), if the critical point $(-1; j0)$ is on the left hand-side of the locus $\hat{P}_a \hat{P}_z$ in the direction of increasing value of ω (from 0 to ∞) - in this case the closed loop is stable – see Fig. 4.5, Fig. 4.6.

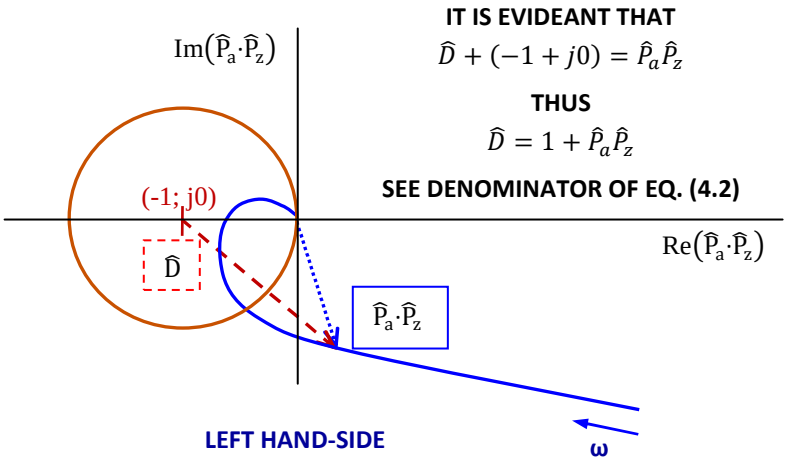


Fig. 4.5 Basic Nyquist criterion depiction

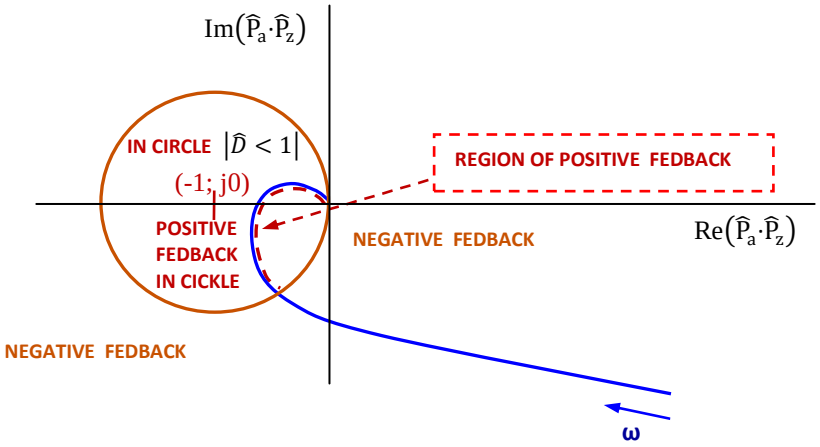


Fig. 4.6 "Type of feedback" depiction

In order to make a feedback system stable, the loop gain $\hat{P}_a\hat{P}_z$ must be less than unity when the phase shift is of 180° (we suppose a negative feedback as a basic analyzed state; a total of 360°), and the phase shift must be less than 180° (total 360°) when the gain around loop is unity. Thus two important notions can be derived: *phase and gain stability margins* – fig. 4.7.

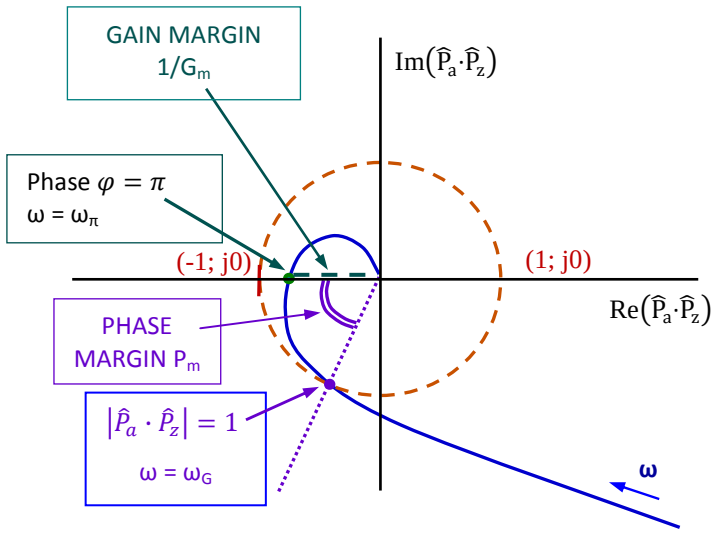


Fig. 4.7 Phase and gain stability margins

Their formal definitions are given by

$$P_m = 180° + arg\left[\hat{P}_a(j\omega_G)\cdot\hat{P}_z(j\omega_G)\right] \tag{4.25}$$

$$G_m[dB] = 20\log\frac{1}{|\hat{P}_a(j\omega_G)\cdot\hat{P}_z(j\omega_G)|} \quad [dB] \tag{4.26}$$

where

ω_G and ω_π stand for, respectively, the *gain and phase crossover frequencies*, which from Fig. 4.7 are obtained as

$$\left|\hat{P}_a(j\omega_G)\cdot\hat{P}_z(j\omega_G)\right| = 1 \quad \Rightarrow \quad \omega_G$$

$$arg\left[\hat{P}_a(j\omega_G)\cdot\hat{P}_z(j\omega_G)\right] = 180° \equiv \pi \quad \Rightarrow \quad \omega_\pi$$

You can see some typical examples in Fig. 4.8.

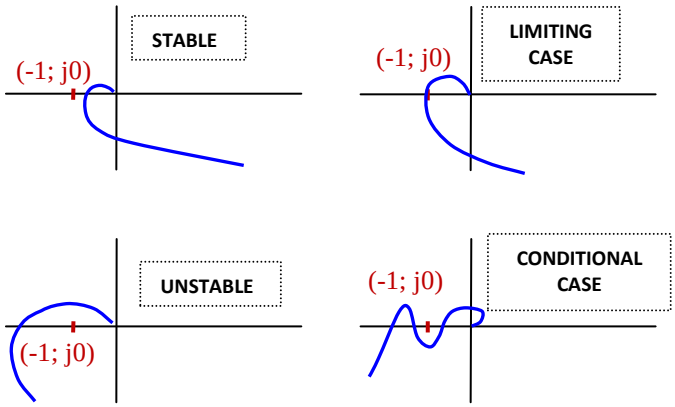


Fig. 4.8 Examples of different systems

4.6.2 Nyquist stability criterion – Bode plots

Notice the scale problems we have with these plots. Clearly, the plot at the right doesn't show the entire plot, but the detail near the unit circle is still not good – see figures above. Here's a Bode's plot for the same system – Fig. 4.9.

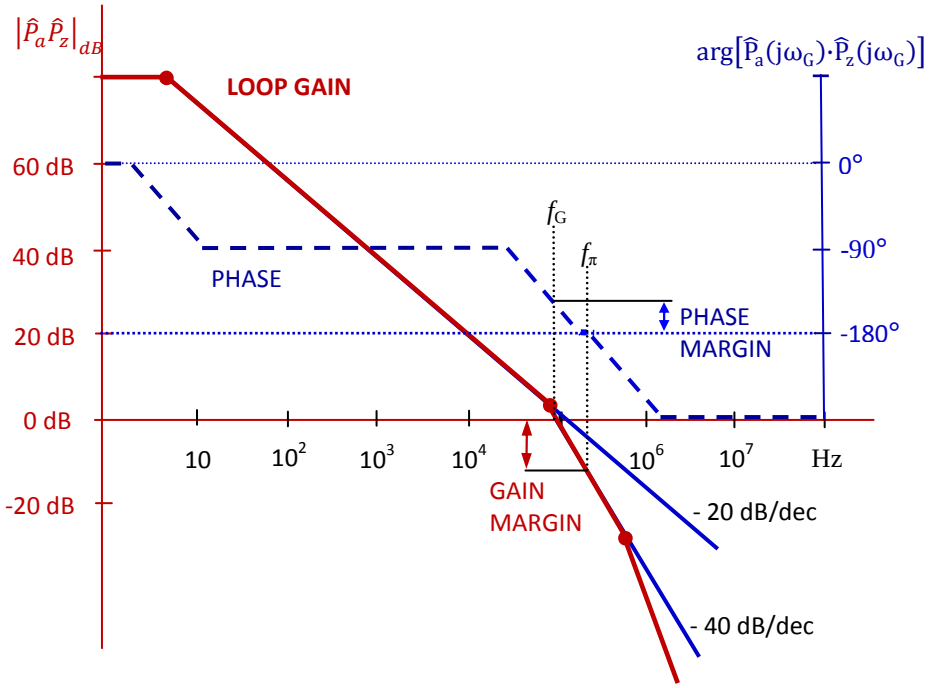


Fig. 4.9 Phase and gain margins – on a Bode plots system

Notice that the scale problem we had with the Nyquist plots is better here. You can see the entire plot, and with a reasonable size plot, you can see all the details you need.

We can summarize how to evaluate the Nyquist stability criterion when the frequency response data is plotted on a Bode's plot.

- Plot the amplitude and phase plots - in Bode's plot form.
- The closed loop system is stable if the zero dB crossing f_G occurs at a lower frequency than the -180° crossing f_π.

Another useful depiction can be used. We use just amplifier characteristics and reciprocal feedback modulus – see Fig. 4.10 – there is frequency independent feedback there.

If the feedback is frequency dependent, the situation will be more complicated – Fig. 4.11. If there is used feedback $\hat{P}_{z_2}$ the phase margin changes sign – feedback structure is unstable.

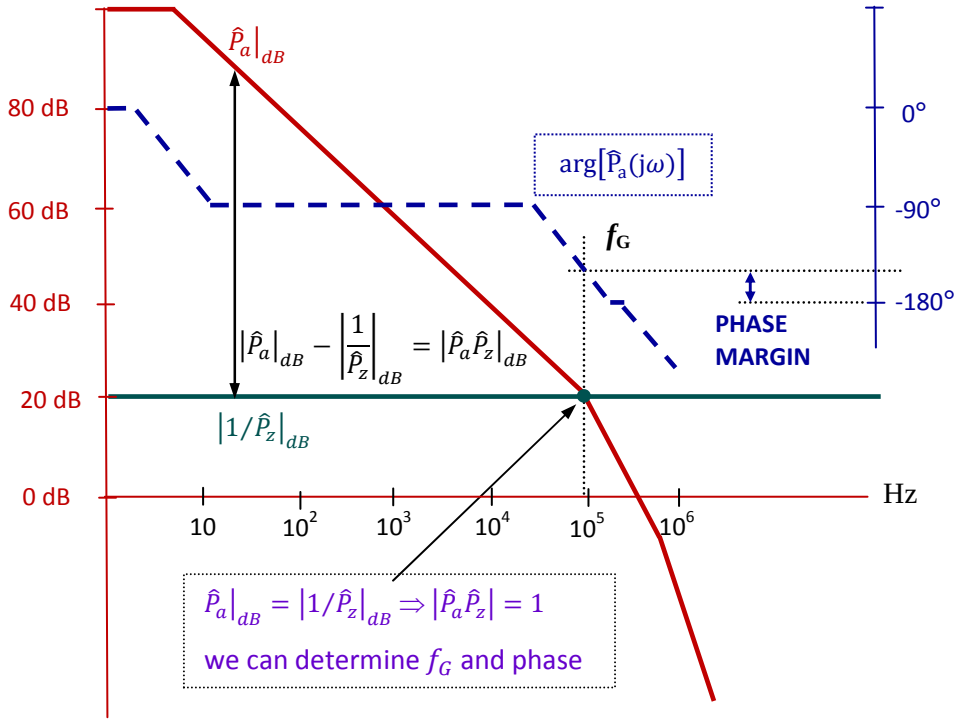


Fig. 4.10 Another depiction of phase margin – on a Bode plots system; stable circuit; $\arg\{\hat{P}_z\} = 0$ here – frequency independent feedback

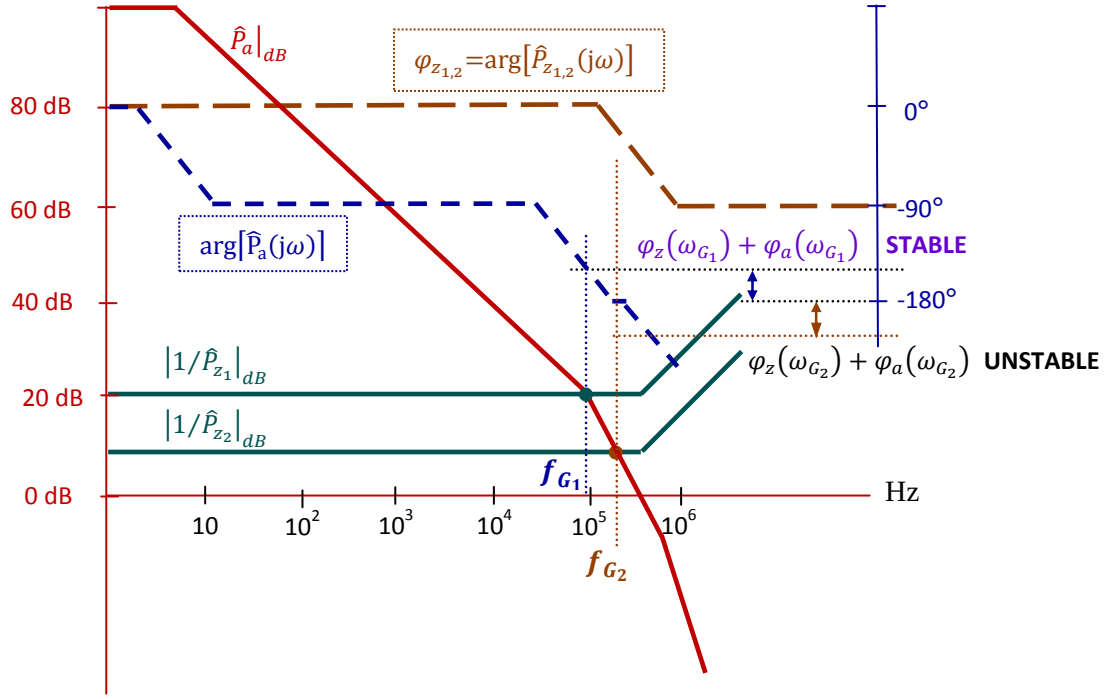


Fig. 4.11 Another depiction of phase margin – on a Bode plots system; stable $(\hat{\boldsymbol{P}}_{\boldsymbol{z_1}})$ and unstable $(\hat{\boldsymbol{P}}_{\boldsymbol{z_2}})$ circuit; if $\boldsymbol{arg}\{\hat{\boldsymbol{P}}_{\boldsymbol{z_1}}\}=\boldsymbol{arg}\{\hat{\boldsymbol{P}}_{\boldsymbol{z_2}}\}$

Changes of only AC feedback properties are depicted in Fig. 4.12. If it is used feedback $\hat{P}_{z_2}$ the phase margin changes sign – feedback structure is unstable again.

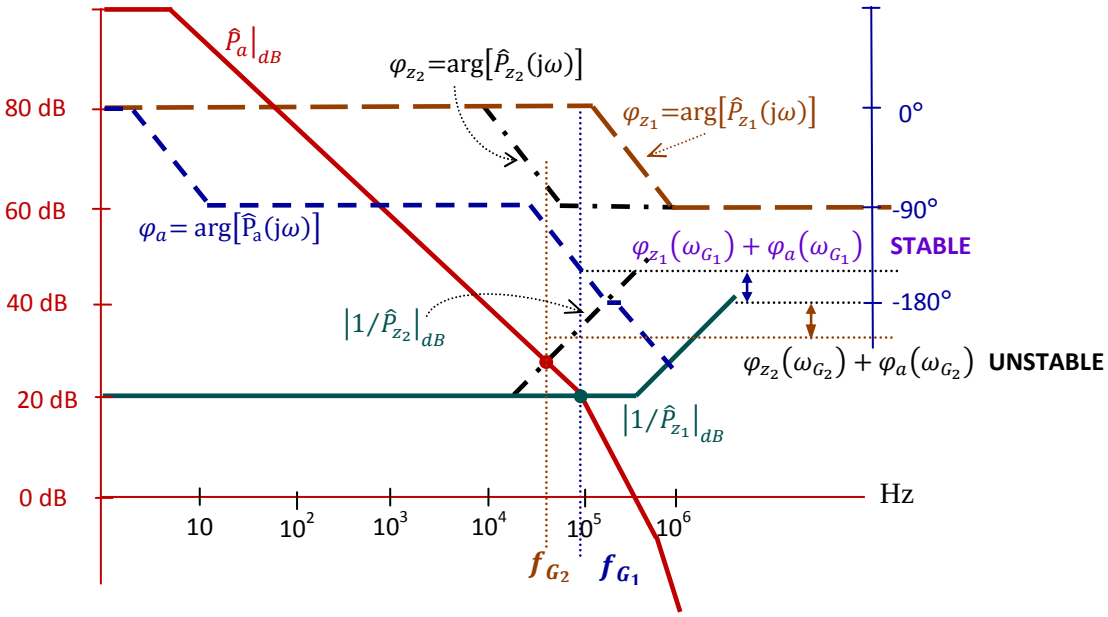


Fig. 4.12 Another depiction of phase margin – on a Bode plots system; stable $(\hat{\boldsymbol{P}}_{\boldsymbol{z_1}})$ and unstable $(\hat{\boldsymbol{P}}_{\boldsymbol{z_2}})$ circuit; different feedback arguments

Summary

A principle of feedback is – a part of output signal is carried back via feedback network into the input port.

- The feedback is negative (degenerative) if $|\hat{P}| < |\hat{P}_a|$; $\hat{P}_a$ is an open loop gain and $\hat{P}$ is a closed loop gain.
- The feedback is positive (regenerative) if $|\hat{P}| > |\hat{P}_a|$.
- The circuit behaves as a generator if

  $1 + \hat{P}_a\hat{P}_Z \to 0$,

  $\hat{P}_Z$ is a feedback network transfer function.
- According to the type of feedback original input and output impedances can be radically changed.
- The transfer function of an asymptotic stable circuit has poles only in the left half of the complex plain – Nyquist stability criterion.

Questions 4

You can find the answers in this text.

1. Explain the basic principle of a feedback

2. Explain a series feedback connection and a parallel feedback connection.

3. Explain a voltage feedback and a current feedback.

4. Explain frequency properties of a feedback system

5. What kind of feedback we must use if we need an oscillator?

Problems 4

Example 4.1

Identify properties of the amplifier in Fig. 4.13. You know $R_2 = 99$ kΩ, $R_1 = 1$ kΩ and real OPA properties: $P_{a_o} \equiv A_o = 2 \cdot 10^5$; $\omega_0 = \omega_1 = 2\pi \cdot 5$**;** a differential input resistance is $\hat{Z}_{v_1} = R_d = 1$ MΩ ; an output resistance is $\hat{Z}_{v_2} = R_o = 100$ Ω.

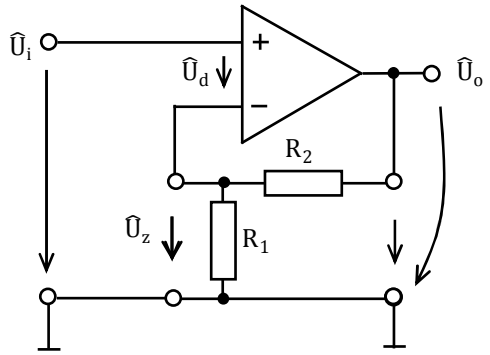


Fig. 4.13 A noninverting amplifier with real OPA

Example 4.2

Identify an input impedance of a current-voltage converter in Fig. 4.14 if you know its $\hat{A}$ and its input differential resistance R_d; suppose zero output resistance. (Note: $\hat{A} = A_0 \cdot \frac{\omega_1}{j\omega+\omega_1}$)

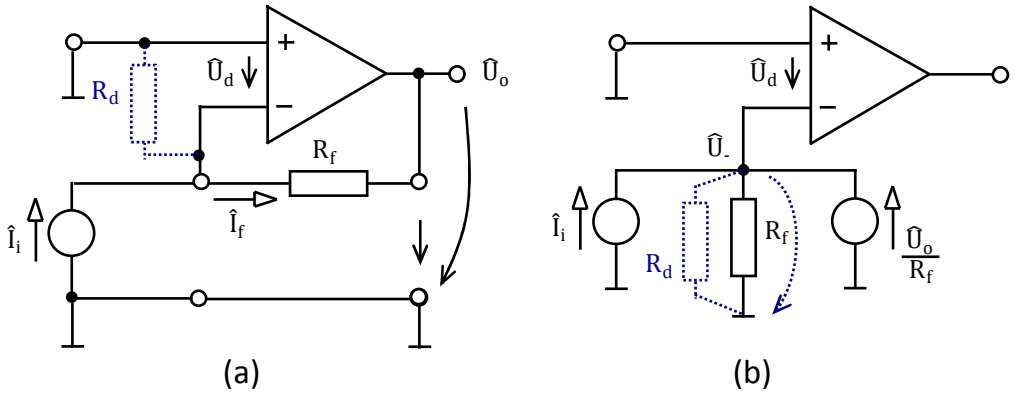


Fig. 4.14 a) A *current-voltage converter*. It is always valid $\hat{U}_0 = \hat{A}\hat{U}_d$;
b) Norton equivalent of feedback circuit (R_f)

Example 4.3

Identify a transfer function of a structure in Fig. 4.15 if you know an OPA gain. Employ experience and findings from the problem 4.2.

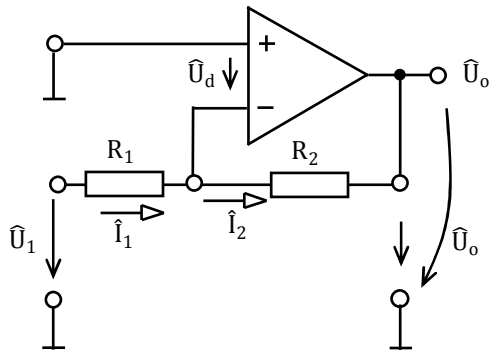


Fig. 4. 15 An inverting amplifier with real OPA

PROBLEMS KEY 4

Ad example 4.1)

It is a negative feedback, voltage, series. Thus $\left(P_{a_o} \equiv K \equiv A_o;\ \omega_0 = \omega_1\right)$

$$\hat{P}_a = P_{a_0} \cdot \frac{\omega_0}{j\omega+\omega_0} = A_0 \cdot \frac{\omega_1}{j\omega+\omega_1}$$

A fraction of the output voltage, determined by the voltage divider at the output of the OPA, is applied to the inverting input; feedback transfer function is not practically affected by the OPA properties. Thus

$$\hat{U}_z = \frac{\hat{U}_0 R_1}{R_1+R_2}$$

and (feedback transfer function we often mark as β)

$$\hat{P}_z = \frac{\hat{U}_z}{\hat{U}_0} = \frac{R_1}{R_1+R_2} = \frac{10^3}{10^3+99\cdot10^3} = 10^{-2} = \beta$$

From the eq. (4.13):

$$\hat{P}_o = \frac{P_{a_0}}{1+P_z P_{a_0}} = \frac{A_0}{1+\beta A_0} = \frac{2\cdot10^5}{1+10^{-2}\cdot2\cdot10^5} = 99{,}95$$

From the eq. (4.13):

$$\omega_3 = \omega_0\left(1 + P_z P_{a_0}\right) = \omega_1(1 + \beta A_0) = 2\pi \cdot 5 \cdot (1 + 2 \cdot 10^3)$$

and

$$f_3 = \frac{\omega_3}{2\pi} = 5 \cdot (1 + 2 \cdot 10^3) = 10{,}005\ \text{kHz}$$

Thus

$$\hat{P} = \frac{P_{a_0}}{1+P_z P_{a_0}} \cdot \frac{1}{1+\frac{j\omega}{\omega_H(1+P_z P_{a_0})}} = \frac{P_{a_0}}{1+P_z P_{a_0}} \cdot \frac{1}{1+\frac{j\omega}{\omega_{H_Z}}} = 99{,}95 \cdot \frac{1}{1+\frac{j\omega}{2\pi 10005}}$$

From the eq. (4.15), $\omega = 0$ we get:

$$\hat{Z}_{v_s} = \hat{Z}_{v_1} \cdot \left(1 + P_z P_{a_0}\right) = R_d(1 + \beta A_0) = 10^6 \cdot (1 + 2 \cdot 10^3) \cong 2 \cdot 10^9 = 2\ \text{G}\Omega$$

From the eq. (4.21), $\omega = 0$ we get:

$$\hat{Z}_{v_{2N}} = \frac{\hat{Z}_{v_2}}{1+\hat{P}_a \hat{P}_z} = R_o(1 + \beta A_0) = 100 \cdot (1 + 2 \cdot 10^3) \cong 50 \cdot 10^{-3} = 50\ \text{m}\Omega$$

Ad example 4.2)

It is a negative feedback, voltage, parallel. But feedback circuit is not an ideal current source. We must use Norton theorem for precise analyze – see Fig. 4.14b – so include real feedback properties. Then

$$\hat{U}_- = \left(R_d \| R_f\right) \cdot \left(\hat{I}_i + \frac{\hat{U}_0}{R_f}\right) = \left(R_d \| R_f\right) \cdot \left(\hat{I}_i + \frac{\hat{U}_d \hat{A}}{R_f}\right) \Rightarrow \Big|_{\hat{U}_d = -\hat{U}_- = -\hat{U}_i}$$

$$\Rightarrow\ \hat{U}_- + \left(R_d \| R_f\right) \cdot \frac{(\hat{U}_-)\hat{A}}{R_f} = \left(R_d \| R_f\right) \cdot \hat{I}_i \qquad \Rightarrow$$

$$\hat{Z}_{v_p} = \frac{\hat{U}_i}{\hat{I}_i} = \frac{(R_d \| R_f)}{1+(R_d \| R_f)\cdot\frac{\hat{A}}{R_f}} = \left| \begin{matrix} R_d \| R_f \equiv \frac{R_d R_f}{R_d+R_f} \\ \text{parallel conection} \end{matrix} \right| = \frac{(R_d \| R_f)}{1+\frac{\hat{A} R_d}{R_d+R_f}}$$

Generally it is $R_d \gg R_f$ a thus

$$\hat{Z}_{v_p} = \frac{R_f}{1+\hat{A}}$$

Now we easily determine

$$\hat{U}_o = \hat{U}_d \cdot \hat{A} = -\hat{U}_- \cdot \hat{A} = -\hat{I}_i \cdot \hat{Z}_{v_p} \cdot \hat{A} = \frac{-\hat{I}_i \cdot R_f \cdot \hat{A}}{1+\hat{A}}$$

Ad example 4.3)

From the problem 2) we know that input impedance of "OPA + R_2" is $\hat{Z}_{v_p} = R_2/(1+\hat{A})$. Now we can easily determine $\hat{I}_1 = \hat{U}_1/\left(R_1 + \hat{Z}_{v_p}\right)$ and thus

$$\hat{I}_1 = \hat{U}_1/\left(R_1 + \hat{Z}_{v_p}\right) = \frac{\hat{U}_1}{R_1+\frac{R_2}{1+\hat{A}}}.$$

If we suppose zero OPA input current it is valid

$$\hat{U}_o = -\hat{U}_d - R_2\hat{I}_1 = \frac{-\hat{U}_o}{\hat{A}} - \frac{R_2\hat{U}_1}{R_1+\frac{R_2}{1+\hat{A}}}$$

and

$$\frac{\hat{U}_o}{\hat{U}_1} = -\frac{R_2}{R_1}\cdot\frac{1}{1+\frac{1+R_2/R_1}{\hat{A}}}$$

The same result we get (we suppose zero OPA input current) this way:

$$\hat{U}_o = -\hat{U}_d - R_2\hat{I}_1 = \frac{-\hat{U}_o}{\hat{A}} - R_2\frac{\hat{U}_{R_1}}{R_1} = \frac{-\hat{U}_o}{\hat{A}} - R_2\frac{\hat{U}_1-\hat{U}_d}{R_1} = \frac{-\hat{U}_o}{\hat{A}} - R_2\frac{\hat{U}_1-\frac{\hat{U}_o}{\hat{A}}}{R_1} = \cdots$$

$$\cdots = -\hat{U}_o\frac{1+\frac{R_2}{R_1}}{\hat{A}} - \hat{U}_1\cdot\frac{R_2}{R_1}$$

and so

$$\frac{\hat{U}_o}{\hat{U}_1} = -\frac{R_2}{R_1}\cdot\frac{1}{1+\frac{1+R_2/R_1}{\hat{A}}}$$

Basic texts

[1] Mikulec, M., – Havlíček, V.: Basic circuit theory. Vydavatelství ČVUT, Praha, 2005, ISBN 80-01-03172-1

[2] Punčochář, J.: Operační zesilovače v elektronice. BEN – Praha, 2002, ISBN 80-7300-059-8

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

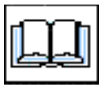
[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

5. Generalized nodal voltage analysis

Time of study: 6 hours

Goals: the student should be able to

- construct an admittance model of an electronics circuit
- derive impedance properties of an electronic structure
- derive transfer functions of an electronic structure
- judge stability of an electronics structure



EXPLANATION

5.1 An admittance model of a $n+1$ – terminal

Let us consider a linear n+1 – terminal that is supplied by the external current sources (signal currents; generally phasors or Laplace transforms) $I_1, I_2, \cdots, I_n, I_{n+1}$ with an external reference point (common node) – see Fig. 5.1 (common node being not incident with any of n+1-terminal nodes). Then there exist just (n+1) node voltages $U_1, U_2, \cdots, U_n, U_{n+1}$ (phasors or Laplace transforms) – they are a linear function of external currents. It may be written in the following matrix form:

$$\begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1r} & Y_{1s} & Y_{1t} & \cdots & Y_{1n} & Y_{1z} \\ Y_{21} & Y_{22} & \cdots & Y_{2r} & Y_{2s} & Y_{2t} & \cdots & Y_{2n} & Y_{2z} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ Y_{r1} & \cdots & \cdots & Y_{rr} & Y_{rs} & Y_{rt} & \cdots & Y_{rn} & Y_{rz} \\ Y_{s1} & \cdots & \cdots & Y_{sr} & Y_{ss} & Y_{st} & \cdots & Y_{sn} & Y_{sz} \\ Y_{t1} & \cdots & \cdots & Y_{tr} & Y_{ts} & Y_{tt} & \cdots & Y_{tn} & Y_{tz} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ Y_{n1} & \cdots & \cdots & Y_{nr} & Y_{ns} & Y_{nt} & \cdots & Y_{nn} & Y_{nz} \\ Y_{z1} & \cdots & \cdots & Y_{zr} & Y_{zs} & Y_{zt} & \cdots & Y_{zn} & Y_{zz} \end{bmatrix} \cdot \begin{bmatrix} U_1 \\ U_2 \\ \cdots \\ U_r \\ U_s \\ U_t \\ \cdots \\ U_n \\ U_z \end{bmatrix} = \begin{bmatrix} I_1 \\ I_2 \\ \cdots \\ I_r \\ I_s \\ I_t \\ \cdots \\ I_n \\ I_z \end{bmatrix} \tag{5.1}$$

Elements Y_{sk} of the square matrix have dimension of admittance. The meaning of these elements results from the relation

$$Y_{sk} = I_s / U_k \tag{5.2}$$

where

Y_{sk} is the element in the $s-th$ row and $k-th$ column, I_s the current of the $s-th$ node when all voltages except of the $k-th$ node are set to zero, and U_k the voltage of this $k-th$ node.

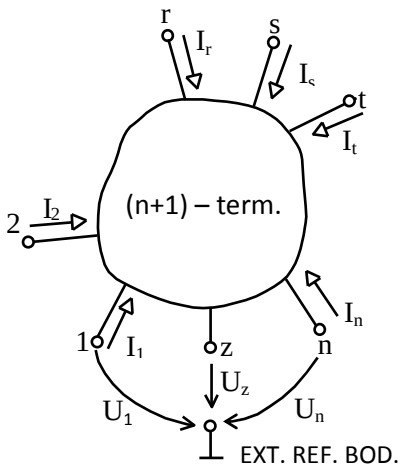


Fig. 5.1 $n+1\ (=z)$ – terminal circuit

Let us have now n – terminal containing just passive two – terminals. We may supply only node s – see Fig. 5.2a – others nodes are connected to the reference point. Now we can easily determine (Ohm's and KCL) that

$$I_s = U_s Y_1 + \cdots + U_s Y_k + \cdots + U_s Y_n$$

It is evident that diagonal elements are sums of admittances connected to the $s-th$ node – they are always *positive*:

$$Y_{ss} = I_s/U_s = Y_1 + \cdots + Y_k + \cdots + Y_n \tag{5.3}$$

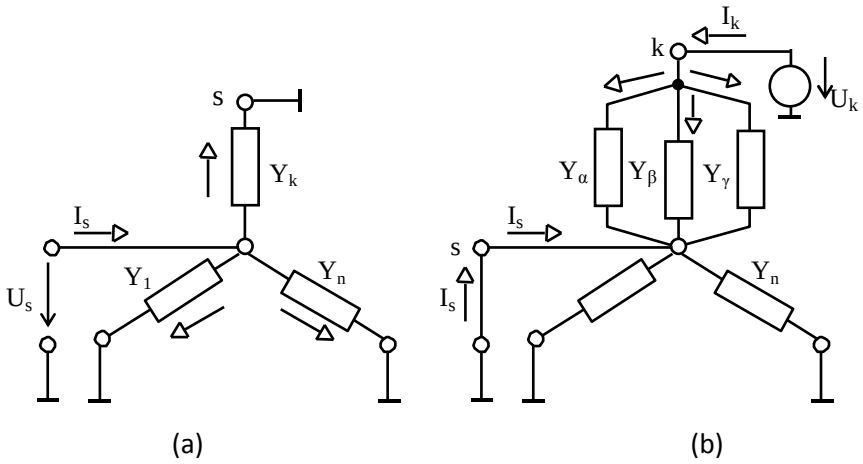


Fig. 5.2 Ilustration of determining admittance matrix parameters

Now we can supply only $k-th$ node – see Fig. 5.2b – others nodes are connected to the reference point and we "measure" the current I_s. All voltages across the admittances are zero except the voltage across $s-th$ node and $k-th$ node. It is evident that

$$I_s = -I_k = -U_k\left(Y_\alpha + Y_\beta + Y_\gamma\right)$$

Then

$$Y_{sk} = I_s/U_k = -I_k/U_k = -\left(Y_\alpha + Y_\beta + Y_\gamma\right) \tag{5.4}$$

Thus other matrix elements Y_{sk} (no diagonal) are sums of admittances of elements connected between $s-th$ node and $k-th$ node – all these elements are *negative*.

The matrix defined by eq. (5.1) – Fig. 5.1 – we call extended matrix (indefinite, singular) because *common node is external*. This means that the sum of all elements in any column (row) is equal to zero – see Chapter 1, too.

If the $r-th$ node is *connected to the external reference* point (now $r-th$ node *becomes internal reference point*) – its voltage is equal to zero. The *respective equations are easily obtained by deleting the corresponding row and column* (here r) in the eq. (5.1). We get an ordinary matrix of the $n+1$ – terminal (with $r-th$ common node). The same is true if we will connect other nodes to the common node (whether extended or ordinary).

The arrow convention in Fig.1 will always be used.

5.2 The parallel connection of n – terminals

Let us connect two n–terminals, „large one" and "small one" – Fig. 5.3

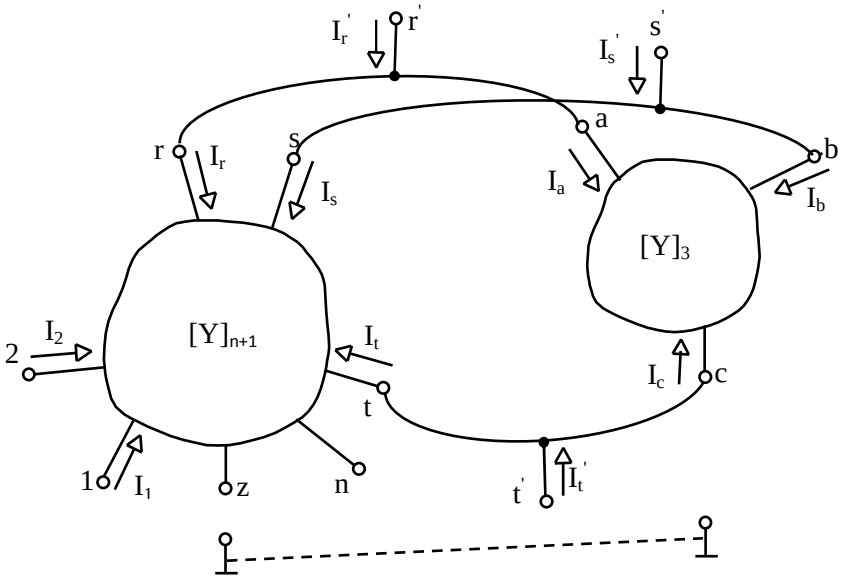


Fig. 5.3 The parallel n – terminal connection

The "large" n - terminal is described by eq. (1), the "small" by equation:

$$\begin{array}{c|ccc} & a & b & c \\ a & Y_{aa} & Y_{ab} & Y_{ac} \\ b & Y_{ba} & Y_{bb} & Y_{bc} \\ c & Y_{ca} & Y_{bc} & Y_{cc} \end{array} \cdot \begin{array}{|c|}\hline U_a \\ \hline U_b \\ \hline U_c \\ \hline \end{array} = \begin{array}{|c|}\hline I_a \\ \hline I_b \\ \hline I_c \\ \hline \end{array} \qquad (5.5)$$

It is fact that (*sign of the parallel connection*)

$$U_r = U_a; \qquad U_s = U_b; \qquad U_t = U_c$$

We use KCL, then for "new" nodes r', s', t' we get:

$$I_r' = I_r + I_a =$$
$$= \left(Y_{r1}U_1 + Y_{r2}U_2 + \cdots + Y_{rr}U_r + Y_{rs}U_s + Y_{rt}U_t + \cdots + Y_{rn}U_n + Y_{r,n+1}U_{n+1}\right) + \cdots$$
$$\cdots + (Y_{aa}U_r + Y_{ab}U_s + Y_{ac}U_t)$$

We rearrange it and get:

$$I_r' = I_r + I_a =$$
$$= (Y_{r1}U_1 + Y_{r2}U_2 + \cdots + (Y_{rr} + Y_{aa})U_r + (Y_{rs} + Y_{ab})U_s + (Y_{rt} + Y_{ac})U_t + \cdots$$
$$\cdots + Y_{rn}U_n + Y_{r,n+1}U_{n+1}\big)$$

In the same way we get:

$$I_s' = I_s + I_b =$$
$$= Y_{s1}U_1 + Y_{s2}U_2 + \cdots + (Y_{sr} + Y_{ba})U_r + (Y_{ss} + Y_{bb})U_s + (Y_{st} + Y_{bc})U_t + \cdots$$
$$\cdots + Y_{sn}U_n + Y_{s,n+1}U_{n+1}$$
$$I_t' = I_t + I_c =$$
$$= Y_{t1}U_1 + Y_{t2}U_2 + \cdots + (Y_{tr} + Y_{ca})U_r + (Y_{ts} + Y_{cb})U_s + (Y_{tt} + Y_{cc})U_t + \cdots$$
$$\cdots + Y_{tn}U_n + Y_{t,n+1}U_{n+1}$$

From set forth above result

1) First we determine the admittance matrix of the large n-terminal.
2) Now we rewrite matrix of the small n-terminal in the same system of nodes (we assign the nodes of small matrices to the large matrix nodes; we do not need mark I_r', … in the future).
3) In „places of coincidences" we add the respective matrix elements from the matrix of small n-terminal, see the matrix below – see the table below.
4) The resultant admittance matrix, thus, the circuit equations, describe the linear network and we can solve the problem - the analysis of circuit with two n-terminals by means of Cramer's rule.

	1	2	⋯	r(a)	s(b)	t(c)	⋯	n	z
1	Y_{11}	Y_{12}	⋯	Y_{1r}	Y_{1s}	Y_{1t}	⋯	Y_{1n}	Y_{1z}
2	Y_{21}	Y_{22}	⋯	⋯	⋯	⋯	⋯	Y_{2n}	Y_{2z}
⋮	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
r(a)	Y_{r1}	⋯	⋯	$Y_{rr} + (Y_{aa})$	$Y_{rs} + (Y_{ab})$	$Y_{rt} + (Y_{ac})$	⋯	Y_{rn}	Y_{rz}
s(b)	Y_{s1}	⋯	⋯	$Y_{sr} + (Y_{ba})$	$Y_{ss} + (Y_{bb})$	$Y_{st} + (Y_{bc})$	⋯	Y_{sn}	Y_{sz}
t(c)	Y_{t1}	⋯	⋯	$Y_{tr} + (Y_{ca})$	$Y_{ts} + (Y_{cb})$	$Y_{tt} + (Y_{cc})$	⋯	Y_{tn}	Y_{tz}
⋮	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
n	Y_{n1}	⋯	⋯	Y_{nr}	Y_{ns}	Y_{nt}	⋯	Y_{nn}	Y_{nz}
z	Y_{z1}	⋯	⋯	Y_{zr}	Y_{zs}	Y_{zt}	⋯	Y_{zn}	Y_{zz}

If we can determine the admittance matrices of modern electronic devices: *operational amplifier, current conveyors, current feedback (transimpedance) amplifier, transadmitance amplifier, Norton amplifier and active three – terminals,* we can modify the algorithm above.

Practical algorithm

1) First we determine the admittance matrix of the circuit „without“ active -terminal [„passive part“ of the circuit - are exciting currents, are nodal voltages, are elements of the admittance matrix; are sums of admittances of elements connected to the node - they are always positive; are sums of admittances of elements connected between the and nodes - all these elements are negative]. For this purposes we must consider that also single terminals are nodes.
2) We rewrite matrices of the active n-terminals (linearized models) in the same system of nodes (into the admittance matrix „without“ active - terminals).
3) In „places of coincidences“ we add the respective matrix elements from the matrices of active -terminals.
4) The resultant admittance matrix, thus, the circuit equations, describe the linear active network and we can solve the problem - the analysis of circuit with active -terminals by means of Cramer's rule.

Example 5.1

Derive an admittance model of the signal model in Fig. 5.4.

a) neglect capacity base - collector
b) with capacity base - collector

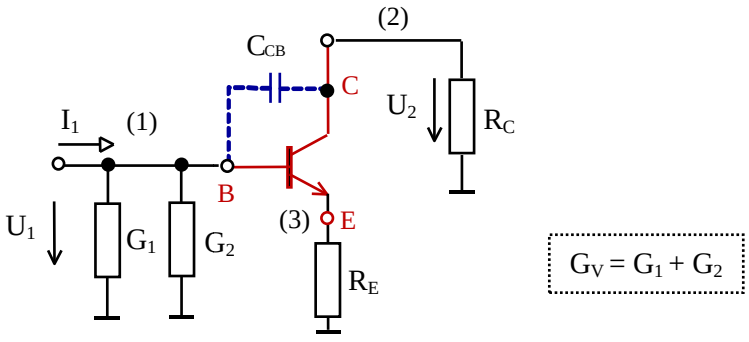


Fig.5.4 Small – signal equivalent circuit - Figs.1 and 26; Chapter 2; (we can add an admittance of the no ideal current source , too).

☑ **Solution:**

Solution of the problem a)

There are three nodes in the circuit and one reference node – we can use ordinary "passive" matrix. The transistor must be defined by its extended matrix – we need all its three nodes. We have just one signal current source .

1. We disconnect the transistor and specify node numbers. We create matrix 3x3.

- Input current of the node 1 is I_1 – it goes "into" node thus + sign (the first row).
- To the node 1 there is connected the admitance $G_V = 1/R_V$, thus $Y_{11} = +G_V$.
- To the node 2 there is connected the admitance $G_C = 1/R_C$, thus $Y_{22} = +G_C$.
- To the node 3 there is connected the admittance $G_E = 1/R_E$, thus $Y_{33} = +G_E$.

There are no other admittances there. The "passive" (ordinary) matrix is:

	1	2	3				
1	G_V				U_1	=	I_1
2		G_C		·	U_2		0
3			G_E		U_3		0

2. We mark the parallel connection of the nodes (rewrite the small matrix nodes) into the „passive matrix.

	1 (B)	2 (C)	3 (E)				
1 (B)	G_V				U_1	=	I_1
2 (C)		G_C		·	U_2		0
3 (E)			G_E		U_3		0

and use the extended transistor matrix (we neglect Y_{CC} and capacitor)

	B	C	E
B	Y_{bb}	0	$-Y_{bb}$
C	Y_{cb}	0	$-Y_{cb}$
E	$-Y_{bb} - Y_{cb}$	0	$Y_{bb} + Y_{cb}$

3. Now we add the respective elements of the transistor matrix:

$B-B \rightarrow$ put Y_{bb} into „area 11"of passive matrix

$B-C \rightarrow$ put 0 into „area 12"of passive matrix¨

……

$E-E \rightarrow$ put $Y_{bb} + Y_{cb}$ into „area 33"of passive matrix

The resulting admittance model is then

	1 (B)	2 (C)	3 (E)				
1 (B)	$G_V + (Y_{bb})$	0	$(-Y_{bb})$		U_1	=	I_1
2 (C)	(Y_{cb})	$G_C + 0$	$(-Y_{cb})$	·	U_2		0
3 (E)	$(-Y_{bb}-Y_{cb})$	0	$G_E + (Y_{bb} + Y_{cb})$		U_3		0

Solution of the problem b)

There are three nodes in the circuit and one reference node, again – we can use ordinary "passive" matrix. The transistor must be defined by its extended matrix – we need all its three nodes. We have just one signal current source I_1.

1. We disconnect the transistor and specify node numbers. But we inscribe the capacity C_{CB} into passive circuit, and we create matrix 3x3, again. There are still only three nodes there.
   - Input current of the node 1 is I_1 – it goes "into" node thus + sign (the first row).
   - To the node 1 there is connected the admittance $G_V = 1/R_V$, and the capacity C_{CB} → thus $Y_{11} = +G_V + pC_{CB}$ ($p = j\omega$ - harmonic steady state).
   - To the node 2 there is connected the admittance $G_C = 1/R_C$, and the capacity C_{CB} thus $Y_{22} = +G_C + pC_{CB}$.
   - To the node 3 there is connected the admittance $G_E = 1/R_E$, thus $Y_{33} = +G_E$.
   - Between the nodes 1 and 2 there is connected just the capacity C_{CB} thus $Y_{12} = Y_{21} = -pC_{CB}$

	1	2	3				
1	$G_V + pC_{CB}$	$-pC_{CB}$			U_1	=	I_1
2	$-pC_{CB}$	$G_C + pC_{CB}$		·	U_2	=	0
3			G_E		U_3	=	0

The "passive" (ordinary) matrix is now:

2. We mark the parallel connection of the nodes (rewrite the small matrix nodes) into the „passive matrix.

	1 (B)	2 (C)	3 (E)				
1 (B)	$G_V + pC_{CB}$	$-pC_{CB}$			U_1	=	I_1
2 (C)	$-pC_{CB}$	$G_C + pC_{CB}$		·	U_2	=	0
3 (E)			G_E		U_3	=	0

3. Now we add the respective elements of the transistor matrix (we neglect Y_{CC} still):

$B-B$ → put Y_{bb} into „area 11"of passive matrix

$B-C$ → put 0 into „area 12"of passive matrix

……..

$E-E$ → put $Y_{bb} + Y_{cb}$ into „area 33"of passive matrix

The resulting admittance model is then

	1 (B)	2 (C)	3 (E)				
1 (B)	$G_V + pC_{CB} + (Y_{bb})$	$-pC_{CB} + 0$	$(-Y_{bb})$		U_1	=	I_1
2 (C)	$-pC_{CB} + (Y_{cb})$	$G_C + pC_{CB} + 0$	$(-Y_{cb})$	·	U_2	=	0
3 (E)	$(-Y_{bb} + Y_{cb})$	0	$G_E + (Y_{bb} + Y_{cb})$		U_3	=	0

Example 5.2

Let us have an electronic structure with an equivalent signal model as it is in Fig. 5.5 There are just two impedances (and) there, and one active three – terminal defined by its extended matrix

	a	b	c		
a					
b				=	
c					

(5.6)

There is just one signal current source (ideal).

Determine an admittance model of this circuit.

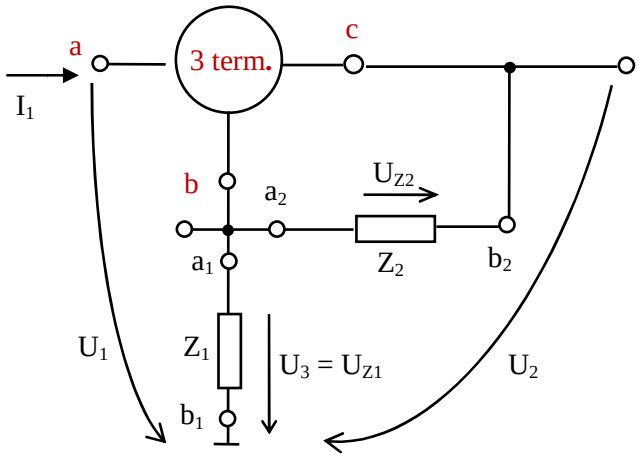


Fig. 5.5 An equivalent signal model of electronic structure with general three-terminal defined by matrix (5.6)

☑ Solution:

1. We disconnect the three-terminal and specify node numbers – Fig. 5.6. We create matrix 3x3.

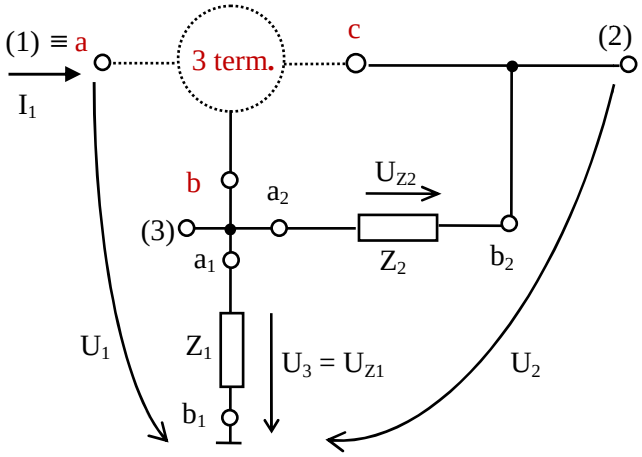


Fig. 5.6 An equivalent signal model after disconnecting the three-terminal

- Input current of the node 1 is I_1 – it goes "into" node thus + sign (the first row).

Thus the "passive" matrix of the structure is

	1	2	3				
1					U_1		I_1
2		Y_2	$-Y_2$	$\cdot$	U_2	$=$	0
3		$-Y_2$	Y_1+Y_2		U_3		0

2. We mark the parallel connection of the nodes (rewrite the small matrix nodes) into the „passive matrix, and add respective admittances of the "active" (three-terminal) matrix. So we are getting the *admittance model of the electronics circuit* in Fig. 5.5.

	1,(a)	2,(c)	3,(b)				
1,	$+(Y_{aa})$	$+(Y_{ac})$	$+(Y_{ab})$		U_1		I_1
2,	$+(Y_{ca})$	$Y_2+(Y_{cc})$	$-Y_2+(Y_{cb})$	$\cdot$	U_2	$=$	0
3,	$+(Y_{ba})$	$-Y_2+(Y_{bc})$	$Y_1+Y_2+(Y_{bb})$		U_3		0

(5.7)

5.3 Determination of nodal voltages, transfer functions and circuit currents

We use as an example the matrix model from the example 2. We must determine unknown nodal voltages. We suppose that the determinant D

$$D=\begin{vmatrix} Y_{aa} & Y_{ac} & Y_{ab} \\ Y_{ca} & Y_2+Y_{cc} & -Y_2+Y_{bc} \\ Y_{ba} & -Y_2+Y_{bc} & Y_1+Y_2+Y_{bb} \end{vmatrix}$$

of the admittance matrix (5.7) is not zero – that investigated electronic system is stable.

Now we can determine (Cramer's rule) that (unknown voltages)

$$U_i = D_i/D \tag{5.8}$$

where D_i is the determinant formed by replacing the $i-th$ column of D - eq. (5.7) - by the column "vector I" – the vector of signal currents (known "right side"); index i – the nodal (voltage) number.

Thus we can derive the all *nodal voltages*:

$$U_1=\frac{D_1}{D}=\frac{\begin{vmatrix} I_1 & Y_{ac} & Y_{ab} \\ 0 & Y_2+Y_{cc} & -Y_2+Y_{bc} \\ 0 & -Y_2+Y_{bc} & Y_1+Y_2+Y_{bb} \end{vmatrix}}{D} \tag{5.9}$$

$$U_2=\frac{D_2}{D}=\frac{\begin{vmatrix} Y_{aa} & I_1 & Y_{ab} \\ Y_{ca} & 0 & -Y_2+Y_{bc} \\ Y_{ba} & 0 & Y_1+Y_2+Y_{bb} \end{vmatrix}}{D} \tag{5.10}$$

$$U_3 = \frac{D_3}{D} = \frac{\begin{vmatrix} Y_{aa} & Y_{ac} & I_1 \\ Y_{ca} & Y_2+Y_{cc} & 0 \\ Y_{ba} & -Y_2+Y_{bc} & 0 \end{vmatrix}}{D} \qquad (5.11)$$

Determinant of 3x3 matrixes we can calculate by means of Sarrus' rule or Sarrus' scheme.

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - [a_{13}a_{22}a_{31} + a_{23}a_{32}a_{11} + a_{33}a_{12}a_{21}]$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix}$$

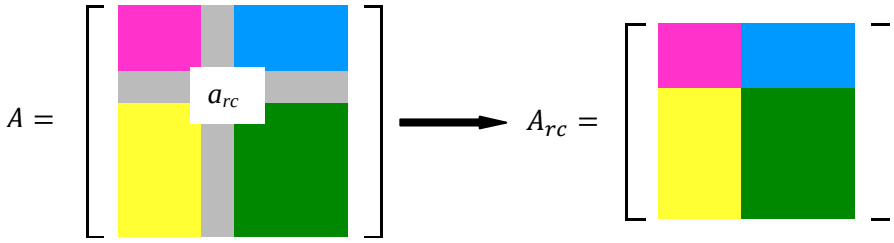
$$D = \begin{vmatrix} 2 & 1 & 0 \\ -3 & 0 & 7 \\ -5 & 4 & -1 \end{vmatrix} = 2 \cdot 0 \cdot (-1) - 3 \cdot 4 \cdot 0 - 5 \cdot 1 \cdot 7 - [0 \cdot 0 \cdot (-5) + 7 \cdot 4 \cdot 2 - 1 \cdot 1(-3)] = -94$$

$$\begin{matrix} 2 & 1 & 0 \\ -3 & 0 & 7 \end{matrix}$$

In general case we can use *cofactor expansion (we prefer column).*

Cofactor expansion (column)

Let $\boldsymbol{A} = (a_{rc})$ be an n by n matrix. By deleting the $r-$ th row and the $c-$ th column, we get an $n-1$ by $n-1$ matrix $\boldsymbol{A}_{rc}$.



The *cofactor* of the rc-entry a_{rc} is defined as the number $A_{rc} = (-1)^{r+c} det\boldsymbol{A}_{rc}$.

$r-raw; c-column; c-th$ column then

$$detA = a_{1c}A_{1c} + a_{2c}A_{2c} + \cdots + a_{nc}A_{nc}$$
$$detA = a_{1c}(-1)^{1+c}detA_{1c} + a_{2c}(-1)^{2+c}detA_{2c} + \cdots + a_{nc}(-1)^{n+c}detA_{nc}$$

If we choose $c = 2$ then (an example)

$$D = \begin{vmatrix} 2 & 1 & 0 \\ -3 & 0 & 7 \\ -5 & 4 & -1 \end{vmatrix} \begin{matrix} \leftarrow r=1 \\ \leftarrow r=2 \\ \leftarrow r=3 \end{matrix} = 1 \cdot (-1)^{1+2} \begin{vmatrix} -3 & 7 \\ -5 & -1 \end{vmatrix} + 0 \cdot (-1)^{2+2} \begin{vmatrix} 2 & 0 \\ -5 & -1 \end{vmatrix} + 4 \cdot (-1)^{3+2} \begin{vmatrix} 2 & 0 \\ -3 & 7 \end{vmatrix}$$

c = 2

r = 1, a_{12} =1 | r = 2, a_{22} =0 | r = 3, a_{32} =4

D = -(3 + 35) + 0 - 4.(14 - 0) = - 94

If it is the node 1 the input node we can derive easily *transfer functions (voltage)*:

$$\frac{U_2}{U_1} = \frac{\frac{D_2}{D}}{\frac{D_1}{D}} = \frac{D_2}{D_1} \tag{5.12}$$

$$\frac{U_3}{U_1} = \frac{\frac{D_3}{D}}{\frac{D_1}{D}} = \frac{D_3}{D_1} \tag{5.13}$$

In this case we need not to determine the determinant D.

The *circuit currents* we can easy derive as follows (current arrows see Fig. 5.3):

$$U_{ab_2} = U_{Z_2} = U_3 - U_2$$
$$I_{Z_2} = U_{Z_2}/Z_2 = U_{Z_2}Y_2 = (U_3 - U_2)Y_2 = Y_2(D_3 - D_2)/D$$
$$U_{ab_1} = U_{Z_1} = U_3$$
$$I_{Z_1} = U_{Z_1}/Z_1 = U_{Z_1}Y_1 = Y_1 D_3/D$$

The currents I_a, I_b and I_c we get by means of the KCL.

5.4 Determination of input impedance

We simply use the Ohm's law

$$Z_{IN} = \frac{U_1}{I_1} = \frac{\frac{D_1}{D}}{I_1} \tag{5.14}$$

For our example thus is valid

$$U_1 = \frac{D_1}{D} = \frac{\begin{vmatrix} I_1 & Y_{ac} & Y_{ab} \\ 0 & Y_2 + Y_{cc} & -Y_2 + Y_{bc} \\ 0 & -Y_2 + Y_{bc} & Y_1 + Y_2 + Y_{bb} \end{vmatrix}}{D} = \frac{(-1)^{1+1} I_1 \begin{vmatrix} Y_2 + Y_{cc} & -Y_2 + Y_{bc} \\ -Y_2 + Y_{bc} & Y_1 + Y_2 + Y_{bb} \end{vmatrix}}{D}$$

$$Z_{in} = \frac{(-1)^{1+1} I_1 \begin{vmatrix} Y_2 + Y_{cc} & -Y_2 + Y_{bc} \\ -Y_2 + Y_{bc} & Y_1 + Y_2 + Y_{bb} \end{vmatrix}}{I_1 D} = \frac{\begin{vmatrix} Y_2 + Y_{cc} & -Y_2 + Y_{bc} \\ -Y_2 + Y_{bc} & Y_1 + Y_2 + Y_{bb} \end{vmatrix}}{D} \tag{5.15}$$

5.5 Voltage signal source (ideal)

There is an ideal voltage source as the signal source in Fig. 5.7.

It is evident that some input current I_1 exists – but we don't know this current. But we know $U_i = U_1$ – we know the first node voltage, now.

Formally must be valid the same system (5.7) of circuits equations as before.

$$Y_{aa}U_1 + Y_{ac}U_2 + Y_{ab}U_3 = I_1$$
$$Y_{ca}U_1 + (Y_{cc} + Y_2)U_2 + (Y_{cb} - Y_2)U_3 = 0$$

$$Y_{ba}U_1 + (Y_{bc} - Y_2)U_2 + (Y_{bb} + Y_1 + Y_2)U_3 = 0$$

But we must rearrange equations now. Known variables go "to the right" – here U_1, unknown ones go "to the left" – here I_1 I_1, thus

$$-I_1 + Y_{ac}U_2 + Y_{ab}U_3 = -Y_{aa}U_1$$
$$0 \cdot I_1 + (Y_{cc} + Y_2)U_2 + (Y_{cb} - Y_2)U_3 = -Y_{ca}U_1$$
$$0 \cdot I_1 + (Y_{bc} - Y_2)U_2 + (Y_{bb} + Y_1 + Y_2)U_3 = -Y_{ba}U_1$$

Then new matrix model (with the known U_1) is

$$\begin{bmatrix} -1 & Y_{ac} & Y_{ab} \\ 0 & Y_{cc} + Y_2 & Y_{cb} - Y_2 \\ 0 & Y_{bc} - Y_2 & Y_{bb} + Y_1 + Y_2 \end{bmatrix} \begin{bmatrix} I_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} -Y_{aa}U_1 \\ -Y_{ca}U_1 \\ -Y_{ba}U_1 \end{bmatrix} \qquad (5.16)$$

After we get experience in the described procedure, we are able to get the result (5.16) – from the (5.7) – directly.

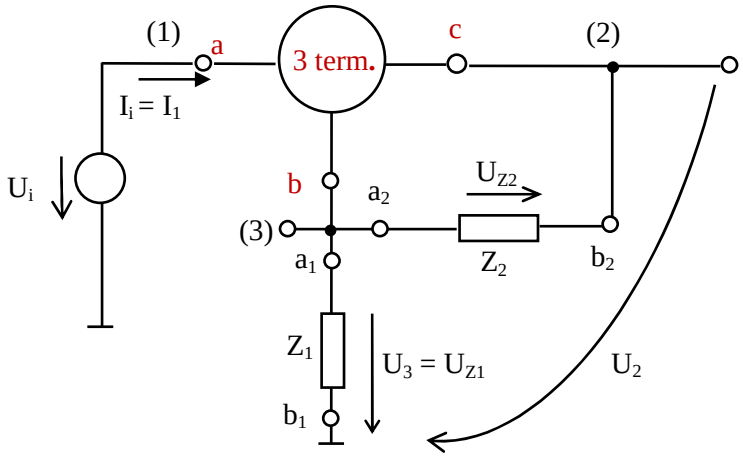


Fig. 5.7 An ideal voltage signal source U_i is used

It is evident that we don't need raw 1 if we solve U_2 and U_3 only, thus it is enough to solve equations (matrix model)

$$\begin{bmatrix} Y_{cc} + Y_2 & Y_{cb} - Y_2 \\ Y_{bc} - Y_2 & Y_{bb} + Y_1 + Y_2 \end{bmatrix} \begin{bmatrix} U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} -Y_{ca}U_1 \\ -Y_{ba}U_1 \end{bmatrix} \qquad (5.17)$$

If we have two unknown variables we need only two equations.

The ideal voltage source (here U_1) is not a function of any current (here I_1) → that's why we scratch the first raw. But its source properties are valid in rows 2 and 3, still → that's why we transfer $Y_{ca}U_1$ and $Y_{ba}U_1$ to the right – attention, sign changes (remember – the internal resistance of an ideal voltage source is zero).

If we know U_2 and U_3 we can determine all circuit variables – so I_1. Or we can solve directly equations (5.16). If we know the I_1, we determine the input impedance $Z_{IN} = U_1/I_1$.

Example 5.3

Determine an admittance model of the circuit in Fig. 5.8. There are just two impedances (and) there, and one active three – terminal defined by its extended matrix

	a	b	c			
a						
b					=	
c						

There is a non ideal voltage source as the signal source in Fig. 5.8 – its input resistance is described by a resistor .

☑ Solution:

1. We determine Norton´s equivalent of the real voltage source and specify node numbers.
2. We determine a "passive" matrix model of the circuit – now the signal current is

	1	2	3			
1				U_1		
2				U_2	=	
3				U_3		

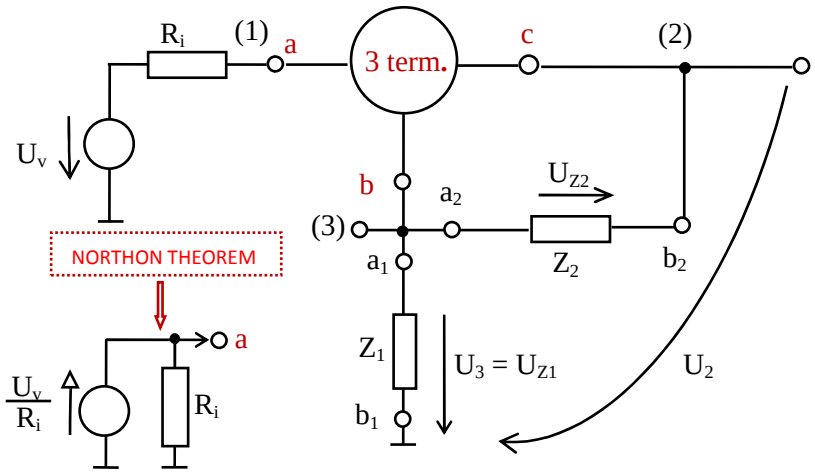


Fig. 5.8 A real voltage signal source is used

3. We rewrite the small matrix nodes into the „passive" matrix, and add respective admittances of the "active" (three-terminal) matrix

	1,(a)	2,(c)	3,(b)			
1,(a)						
2,(c)					=	0
3,(b)						0

(5.18)

This is the admittance model of the circuit in Fig. 5.8 if real voltage source as a signal source is used.

5.6 Determination of output impedance – generally node impedance

Since we investigate linear circuits, we can use a „measure approache a problem“. The signal sources we substitute (replace) by their internal impedances. It means that $I_1 = 0$ (and infinite impedance) or $U_v = 0$ (and zero impedance). But the wiring is still the same, hence the admittance model too – see Fig. 5.9. It is advantageous to use a real signal source – we can directly investigate the influence of R_i.

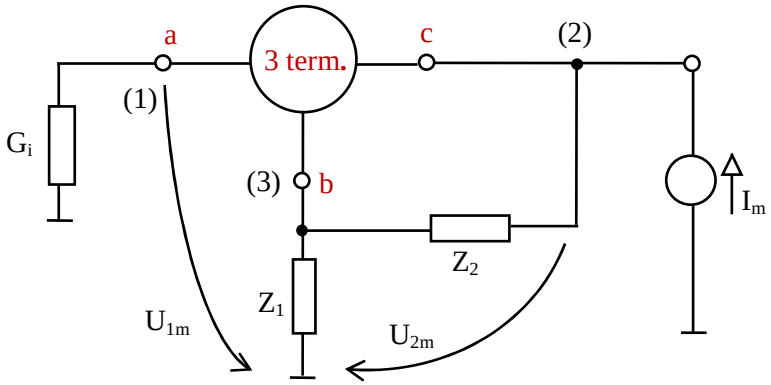


Fig. 5.9 Measuring of the node impedance – here node 2

Let us derive the node 2 impedance – see Fig. 5.9. We force on the measuring current I_m into the node 2. Derivation of the output "node 2" impedance is that way transferred onto derivation of the input "node 2" impedance. We easily get (see the known algorithm above):

	1, a	2, c	3, b		
1, a	$Y_{aa} + G_i$	Y_{ac}	Y_{ab}	U_{1m}	0
2, c	Y_{ca}	$Y_{cc} + Y_2$	$Y_{cb} - Y_2$	U_{2m} =	I_m
3, b	Y_{ba}	$Y_{bc} - Y_2$	$Y_{bb} + Y_1 + Y_2$	U_{3m}	0

(5.19)

It is evident that admittance matrix is the same as before – see eq. (5.18). Only signal currents vector changes because measuring current I_m gets in the node 2. The helping nodal voltages suffix "m" indicates the above described situation, only. Now we can determine that impedance of the node 2 is (output or input)

$$Z_{out_2} = \frac{\frac{U_{m_2}}{I_m}}{\frac{D_2}{I_m D}} = \frac{\begin{vmatrix} Y_{aa} + G_i & 0 & Y_{ab} \\ Y_{ca} & I_m & Y_{cb} - Y_2 \\ Y_{ba} & 0 & Y_{bb} + Y_1 + Y_2 \end{vmatrix}}{I_m D} \tag{5.20}$$

D is determinant of the admittance matrix (5.19).

5.7 Interconnection of two nodes

Suppose we know an admittance model of some electric circuit.

If we connect any node $\boldsymbol{s}$ with the reference point (node), we must scrape correspondent raw and column (the $s-th$ raw and $s-th$ column; signal source I_s falls into the $s-th$ raw) because corresponding nodal voltage is zero.

If we connect each two nodes ("no reference nodes") together, the $s-th$ one and the $k-th$, for example, it is valid that

$$U_k = U_s = U_{k,s} \tag{5.21}$$

$$I_k + I_s = I_{k,s} \tag{5.22}$$

The suffix "k, s" indicates a "connection history" – an origin one new node (coupling).

The equation (5.21) defines a summation of the $k-th$ column and the $s-th$ column in the admittance matrix.

The equation (5.22) defines a summation of the $k-th$ raw and the $s-th$ raw in the admittance matrix.

Summary

1) Nodal voltage analysis can be applied to any linear (linearized) circuit.
2) It is always necessary to follow the arrow convention of Fig. 5.1. Then we can use the algorithm in section 5.2 to obtain an admittance model of each linear (linearized) electronic circuit. Now the "circuit" problem is converted to a mathematical one.
3) We are able to derive all nodal voltages and therefore also transfer functions.
4) To determine the node impedance we force on the "measuring" current into them.

Questions 5

You can find the answers in this text.

1. Define an admitance matrix of a passive n – terminal (why does it have to be a linear system?).

2. Explain a parallel connection of two linear n – terminals – resulting matrix model (algorithm).
3. What is an admittance model of a linear (linearized) electronic circuit?
4. Define generally a voltage transfer function between nodes.
5. How can we generally derive node impedance?

Problems 5

Example 5.1

a) Derive an admittance model of the circuit in Fig. 5.10 (see Chapter 2, example 2.6; $k = 0{,}25$ mA/V^2, $U_T = 2{,}5$ V, $U_{DD} = 15$ V, $R_G = 510$ kΩ and $R_D = 1{,}5$ kΩ), suppose an ideal signal voltage source U_{i_1}
b) Derive a transfer function U_2/U_{i_1}
c) Derive input impedance (impedance of the node 1).

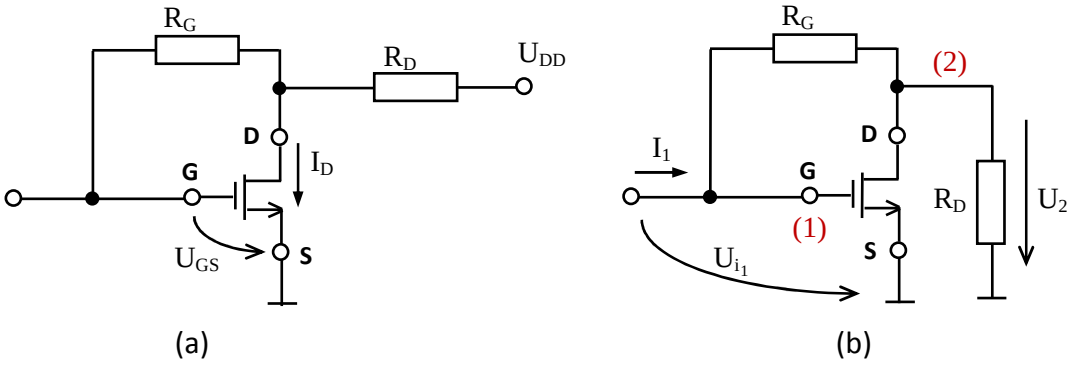


Fig. 5.10 a) An quiescent point adjustment (N channel MOSFET); b) An equivalent (AC) model of the circuit

Example 5.2

a) Is the quiescent point of the circuit in Fig. 5.11 the same as the quiescent point of the circuit in fig. 5.10, if there are used the same transistor, R_G and R_D and $U_{DD} = 15$ V? Why yes/no?
b) Draw a signal equivalent of the circuit in Fig. 5.11 and specify node numbers.
c) Derive an admittance model of the circuit – U_{i_1} is an ideal voltage source, $U_{i_1} + R_I$ models a real voltage source.

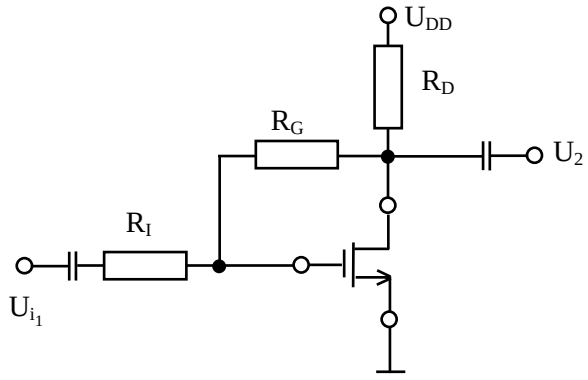


Fig. 5.11 a) A structure with a real input voltage source

PROBLEMS KEY 5

Ad example 5.1)

a)

There are only two nodes in the circuit (and reference one, of course) – see Fig. 5.10. We can easily construct needed admittance matrix. The transistor "S" node is connected directly into reference point, thus "raw S" and "column S" we can "scrape" – we use ordinary transistor matrix only.

	G	D	S
G	0	0	0
D	g_m	g_d	$-g_m - g_d$
S	$-g_m$	$-g_d$	$g_m + g_d$

	1,G	2,D		
1,G	G_G	$-G_G$	U_{i_1}	I_1
2,D	$-G_G + g_m$	$G_D + G_G + g_d$	U_2	0

If we know nodal voltage U_{i_1}, we must rearrange equations – "I_1 to the left; U_{i_1} to the right":

1,G	2,D		
-1	$-G_G$	I_1	$-G_G U_{i_1}$
0	$G_D + G_G + g_d$	U_2	$-(-G_G + g_m)U_{i_1}$

This matrix is not an admittance matrix, now.

b)

It is evident that (very simple case here)

$$U_2(G_D + G_G + g_d) = -(-G_G + g_m)\, U_{i_1}$$

Thus

$$\frac{U_2}{U_{i_1}} = \frac{G_G - g_m}{G_D + G_G + g_d} = \frac{G_G}{G_D + G_G + g_d} - \frac{g_m}{G_D + G_G + g_d}$$

and then

$$\frac{U_2}{U_{i_1}} = \frac{-g_m R_D}{1 + \frac{R_D}{R_G} + \frac{R_D}{r_d}} + \frac{1}{1 + \frac{R_D}{R_G} + \frac{R_D}{r_d}}$$

The first term defines a *gain* of the structure, the second one a *feed through* of the input signal – via R_G to the amplifier output.

If we use results gained in Chapter 2:

$$I_D = 5{,}27 \text{ mA};\ U_{GS} = 7{,}095 \text{ V};\ g_m = \frac{2I_D}{U_{GS} - U_T} = \frac{2 \cdot 5{,}27 \cdot 10^{-3}}{7{,}095 - 2{,}5} = 2{,}294 \text{ mA/V};$$

$$g_d = \frac{1}{r_d}\Big|_{\text{suppose } U_A = 200 \text{ V}} \Big| \frac{5{,}27 \cdot 10^{-3}}{200} = 26{,}35 \text{ µS}$$

then *gain* is -3,301 and *feed through* is 0,003; thus a total value is -3,298.

c)

From the first raw of the circuit model we get

$$-I_1 - G_G U_2 = -G_G U_{i_1}$$

Then

$$\frac{U_{i_1}}{I_1} = \frac{R_G}{1 - U_2/U_{i_1}}$$

This is a very good example of Miller effect – apply to the resistor R_G which is connected as feedback resistor. We might solve this problem generally and establish rate U_2/U_{i_1} as mentioned above. But it is a ceremonious matter. For given conditions we get the input resistance

$$R_{in} = \frac{U_{i_1}}{I_1} = \frac{510 \cdot 10^3}{1 - (-3{,}298)} = 118{,}66 \text{ kΩ}$$

The same result we must get in this ordinary way: The voltage across R_G is

$$U_{i_1} - U_2 = U_{i_1} - A_U U_{i_1}.$$

Then the resistor current is

$$I_1 = \frac{U_{i_1} - U_2}{R_G} = \frac{U_{i_1}(1 - A_U)}{R_G}.$$

Then equivalent input resistance is

$$R_{in} = \frac{U_{i_1}}{I_1} = \frac{R_G}{1 - A_U} = \frac{R_G}{1 - U_2/U_{i_1}}.$$

The feedback is not critical if we use an ideal signal voltage source (its signal resistor is zero).

Ad example 5.2)

a)

The quiescent point is the same. The resistor R_I can no influence DC currents – due to an input coupling capacity.

b)

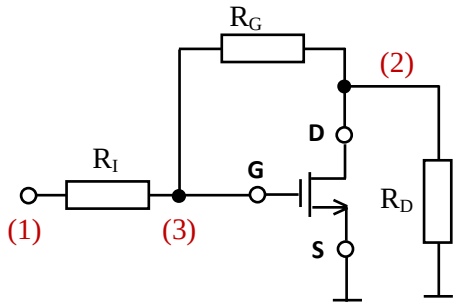


Fig. 5.12 An equivalent AC circuit (signal)

c)

Now we can ordinarily derive an admittance model

	1	2, D	3, G			
1	G_I	0	$-G_I$	U_1	I_1
2, D	0	$G_D+G_G+g_d$	$-G_G+g_m$	$\cdot U_2 =$	0
3, G	$-G_I$	$-G_G$	G_I+G_G	U_3	0

If we again suppose known input voltage (ideal signal voltage source) we must rearrange above equations

-1	0	$-G_I$	I_1	$-G_I U_1$
0	$G_D+G_G+g_d$	$-G_G+g_m$	$\cdot U_2 =$	0
0	$-G_G$	G_I+G_G	U_3	$G_I U_1$

It is enough to solve just two last equations, thus

$G_D+G_G+g_d$	$-G_G+g_m$	U_2	$=$	0
$-G_G$	G_I+G_G	$\cdot U_3$		$G_I U_1$

From these equations we get

$$U_2 = \frac{-(-G_G+g_m)G_I U_1}{(G_D+G_G+g_d)(G_I+G_G)+G_G(-G_G+g_m)}$$

Further – after rearranging – we get

$$\frac{U_2}{U_1} = \frac{-g_m+G_G}{(G_D+G_G+g_d)} \cdot \frac{G_I}{G_I+G_G\left(1-\frac{G_G-g_m}{G_D+G_G+g_d}\right)}$$

The first factor evidently defines the structure gain of the circuit in Fig. 5.10 – then with the exclusion of the resistor R_I:

$$A_{U_0} = -\frac{-G_G+g_m}{G_D+G_G+g_d}$$

The second factor includes a term

$$G_{IN_0} = G_G\left(1-\frac{-G_G+g_m}{G_D+G_G+g_d}\right) = G_G \cdot \left(1-A_{U_0}\right) = \frac{1-A_{U_0}}{R_G} = \frac{1}{R_{IN_0}}$$

This term describes the input resistance of the structure in Fig. 5.10.

We can define the structure gain (in Fig. 5.12) now as

$$\frac{U_2}{U_1} = \frac{A_{U_0} G_I}{G_I + G_{IN_0}} = \frac{A_{U_0} R_{IN_0}}{R_I + R_{IN_0}}$$

Physical sense is evident – resistor R_I and the input resistance R_{IN_0} create a voltage divider and so

$$\frac{U_3}{U_1} = \frac{R_{IN_0}}{R_I + R_{IN_0}}$$

Further we can easily determine that it is really true (from the known equations above)

$$\frac{U_2}{U_3} = -\frac{-G_G + g_m}{G_D + G_G + g_d}$$

Furtherit is evident that

$$\frac{U_2}{U_1} = \frac{\frac{U_2}{U_3}}{\frac{U_3}{U_1}} = -\frac{-G_G + g_m}{G_D + G_G + g_d} \cdot \frac{R_{IN_0}}{R_I + R_{IN_0}} = A_{U_0} \cdot \frac{R_{IN_0}}{R_I + R_{IN_0}}$$

A total input resistance is now

$$R_{IN} = R_I + R_{IN_0}$$

The feedback resistor R_G decreases the input circuit resistance. It means, if we suppose real signal source resistance R_I, the gain decreases, too.

Basic texts

[1] Punčochář, J.: Admittance models of modern linear amplifying structures. Transactions of the VŠB – TU Ostrava, 1, 2003, vol. VI, pp 151 – 161, ISBN 80-248-0223-6

[2] Mohylová, J.: Lineární obvody s elektronickými prvky-Sbírka příkladů, VŠB-TU Ostrava 2002, ISBN 80-248-0098-5

[3] Punčochář, J.: Lineární obvody s elektronickými prvky. Skriptum, VŠB-TU Ostrava 2002, ISBN 80-248-0040-3

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

6. Analysis of amplifying structures (ideally no frequency dependent) and oscillators

Time of study: 6 hours

Goals: the student should be able to

- define ideal gain (ideal transfer function) of basic amplifier structures (inverting, noninverting, differential)
- define an error function of amplifier structures (influence of real active elements properties)
- define characteristic amplifier frequency (gain drop = 3 dB)
- define oscillation conditions

EXPLANATION

6.1 An linear circuits – tranzistors and triode

There are analysed basic transistor and triode circuits in this text.

6.1.1 Common emitter connection (CE)

A basic common emitter connection of a bipolar transistor you can see in Fig. 6.1. Stricly speaking, the resistance R_E (its AC equivalent) should be zero. It is known (thereinafter) that the resistance R_C defines (approximately) an output resistance of the CE circuit. Thus we must always choose $R_C \ll R_z$ (load resistor) - or else it will degrade voltage gain of the circuit (Thévenin's theorem). The quiescent point was determined in the Chapter 2 (Fig. 2.1, too) so as the admitance model of the bipolar transistor (Fig. 2.23, table 2.1, Problem 2.1).

We neglect R_i (zero), a capacity C_{CB} and "Early voltage" - $Y_{22} = 0$, first. Then we easily get

	1 **(B)**	2 **(C)**	3 **(E)**			
1 **(B)**	$G_V + (\boldsymbol{Y_{11}})$	0	$(-\boldsymbol{Y_{11}})$	U_1	=	I_1
2 **(C)**	$(\boldsymbol{Y_{21}})$	$G_C + 0$	$(-\boldsymbol{Y_{21}})$	U_2		0
3 **(E)**	$(-\boldsymbol{Y_{11}} - \boldsymbol{Y_{21}})$	0	$G_e + (\boldsymbol{Y_{11}} + \boldsymbol{Y_{21}})$	U_3		0

Then we get

$$D = G_C[G_V(G_e + Y_{11} + Y_{21}) + G_e Y_{11}]$$

The nodal voltage of the node 1 is

$$U_1 = \frac{D_1}{D} = \frac{I_1 G_C (G_e + Y_{11} + Y_{21})}{D}$$

The nodal voltage of the node 2 is

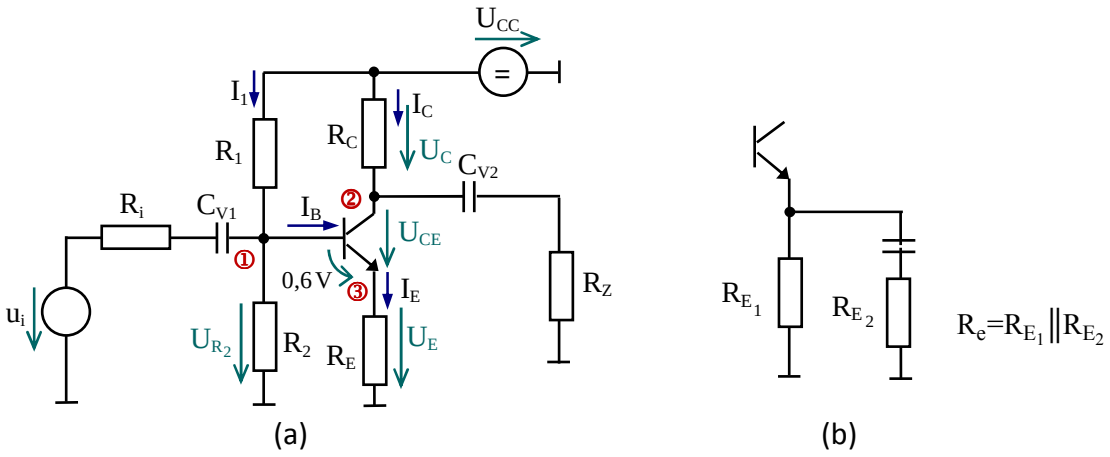


Fig. 6.1 a) A basic common emitter configuration (CE) – voltage divider biasing; b) a capacitor in Fig. transfers AC signals through R_{E_2} to the earth (reference point), too – parallel connection $R_{E1}\|R_{E_2}$ gives R_e for the AC signal

$$U_2 = \frac{D_2}{D} = \frac{-I_1 G_e Y_{21}}{D}$$

The nodal voltage of the node 3 is

$$U_3 = \frac{D_3}{D} = \frac{I_1 G_C (Y_{11}+Y_{21})}{D}$$

Now we can derive voltage gain

$$A_{U_{CE}} = \frac{U_2}{U_1} = \frac{-G_e Y_{21}}{G_C(G_e+Y_{11}+Y_{21})} = -\frac{R_C}{R_e}\cdot\frac{1}{1+\frac{1}{\beta}+\frac{r_e}{R_e}} \qquad (6.1)$$

If $\beta \gg 1$ and $R_e \gg r_e \cong U_T/I_E$ we get familiarly known relation $A_{U_{CE}} = -R_C/R_e$. The "rest" of the eq. (6.1) defines an error function. If we ensure $R_e = 0$, then

$$A_{U_{CE}} = \lim_{R_e \to 0}\left(-\frac{R_C}{R_e}\cdot\frac{1}{1+\frac{1}{\beta}+\frac{r_e}{R_e}}\right) = -\frac{R_C}{R_e}$$

Further we determine an input resistance (generally impedance)

$$R_{in} = \frac{U_1}{I_1} = \frac{G_C(G_e+Y_{11}+Y_{21})}{G_C[(G_e+Y_{11}+Y_{21})+G_e Y_{11}]} = \frac{1}{G_V+\frac{1}{\beta(r_e+R_e)+R_e}} \qquad (6.2)$$

It is evident that the input resistance is created by parallel connection of $G_V = G_1 + G_2$ and „transistor base input resistor"

$$R_{V_B} = \beta(r_e + R_e) + R_e \qquad (6.3)$$

The eq. (6.3) we can get from the eq. (6.2) if we put $G_V = 0$ – thus we are neglecting input "DC" divider influence. It is evident, too, that there always is $R_{in} < R_{V_B}$. The input divider decreases the input resistance.

The output resistance (resistance of the node 2) we get by means the measuring current I_{2m} into node 2, I_1 we equate zero. Then we get

	1 **(B)**	2 **(C)**	3 **(E)**			
1 **(B)**	$G_V + (\boldsymbol{Y_{11}})$	0	$(-\boldsymbol{Y_{11}})$	U_{1m}	$=$	0
2 **(C)**	$(\boldsymbol{Y_{21}})$	$G_C + 0$	$(-\boldsymbol{Y_{21}})$	U_{2m}		I_{2m}
3 **(E)**	$(-\boldsymbol{Y_{11}} - \boldsymbol{Y_{21}})$	0	$G_e + (\boldsymbol{Y_{11}} + \boldsymbol{Y_{21}})$	U_{3m}		0

For given supposing we get

$$R_{out} = \frac{U_{2m}}{I_{2m}} = R_C \tag{6.4}$$

The same result we can get by means of Thévenin's theorem. The open output voltage is $U_{2_n} \cong - U_1 R_C / R_e$. The short output current I_{ZK} defines only resistor R_e: $I_{ZK} = -I_E \cong -U_1/R_e$. Thus the output resistance is $R_{out} \cong U_{2_n}/I_{ZK} = R_C$. If we use more complex transistor model the result will by slightly change – but still it will be R_C dominant.

6.1.2 Common collector connection (CC)

Common collector conection is shown in Fig. 6.2a, Fig. 6.2b –an AC equivalent model (signal) of CC; $G_V = G_1 + G_2 + G_i$; $I_1 = U_i G_i$ – a Norton equivalent; neglected R_Z; G_i – dash line.

In the circuit „above" we let out R_C and choose $R_1 = R_2$, so $U_E \approx U_{CE} \approx U_{CC}/2$. Then we can easy determine quiescent current and an admittance model of transistor.

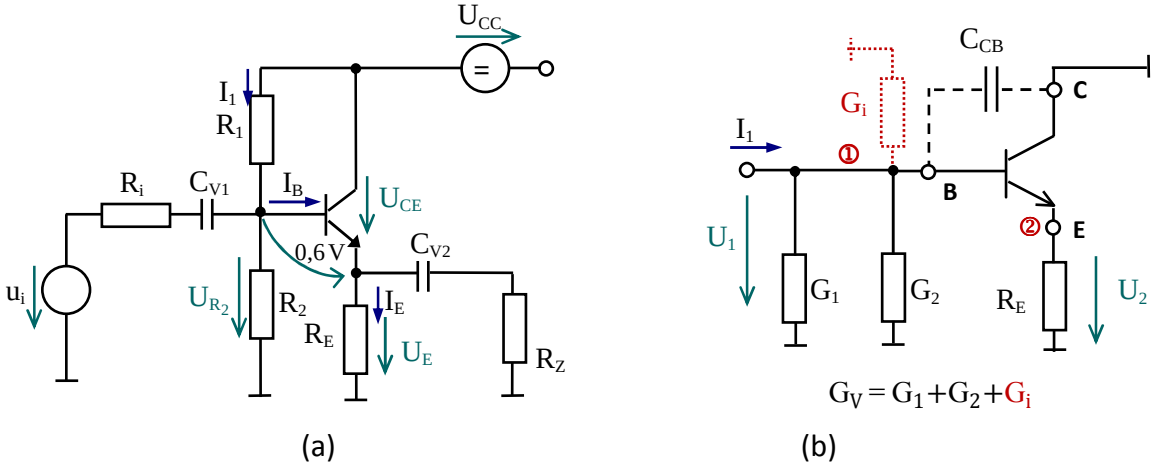


Fig. 6.2 Common collector connection (follower)

We easily determine an admittance circuit model (C_{CB} neglected now):

	1,B	2,E			
1,B	$G_V + \boldsymbol{Y_{11}}$	$-\boldsymbol{Y_{11}}$	U_1	$=$	I_1
2,E	$-\boldsymbol{Y_{11}} - \boldsymbol{Y_{21}}$	$G_E + \boldsymbol{Y_{11}} + \boldsymbol{Y_{21}}$	U_2		0

Then

$$D = G_V(G_E + \boldsymbol{Y_{11}} + \boldsymbol{Y_{21}}) + \boldsymbol{Y_{11}} G_E$$
$$D_1 = I_1(G_E + \boldsymbol{Y_{11}} + \boldsymbol{Y_{21}})$$
$$D_2 = I_1(\boldsymbol{Y_{11}} + \boldsymbol{Y_{21}})$$

Voltage gain (transfer function) is

$$A_{U_{CC}} = \frac{D_2}{D_1} = \frac{U_2}{U_1} = \frac{(Y_{11}+Y_{21})}{G_e+Y_{11}+Y_{21}} = \frac{1}{1+\frac{\beta}{\beta+1}\cdot\frac{r_e}{R_e}}$$

If $\beta \gg 1$, we get simple formula $A_{U_{CC}} \cong 1/(1 + r_e/R_E)$.

Input resistance derive from equation

$$\frac{U_1}{I_1} = \frac{G_e+Y_{11}+Y_{21}}{G_v(G_e+Y_{11}+Y_{21})+G_eY_{11}}$$

After rearranging we get

$$R_{in} = \frac{U_1}{I_1} = \frac{G_e+Y_{11}+Y_{21}}{G_v(G_e+Y_{11}+Y_{21})+G_eY_{11}} = \left| Y_{11}=\frac{1}{\beta r_e}; Y_{21}=\frac{1}{r_e} \right| = \frac{1}{G_V+\frac{1}{\beta(r_e+R_e)+R_e}}$$

Compare this formula with eq. (6.2) – it is formally the same input resistance.

For derivation of an output resistance we use signal model in fig. 6.3 (I_{2m} into node 2, I_1 we equate zero)

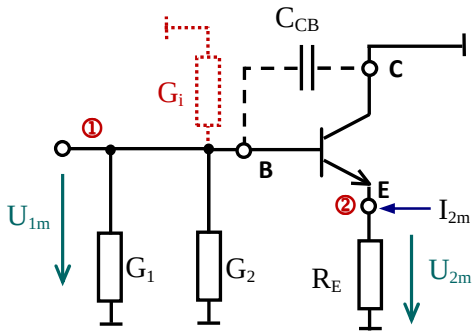


Fig. 6.3 Determination of an output resistance – signal model

The admittance matrix is still the same, "current" column changes, only.

	1,B	2,E		
1,B	$G_V + \boldsymbol{Y_{11}}$	$-\boldsymbol{Y_{11}}$	U_{1m}	0
2,E	$-\boldsymbol{Y_{11}} - \boldsymbol{Y_{21}}$	$G_E + \boldsymbol{Y_{11}} + \boldsymbol{Y_{21}}$	U_{2m}	I_{2m}

Now we can get

$$D = G_V(G_E + \boldsymbol{Y_{11}} + \boldsymbol{Y_{21}}) + \boldsymbol{Y_{11}}G_E$$

$$D_{2m} = \begin{vmatrix} G_V + Y_{11} & 0 \\ -Y_{11} - Y_{22} & I_{2m} \end{vmatrix} = I_{2m}(G_V + Y_{11}).$$

The node voltage of the node 2 is now

$$U_{2m} = \frac{I_{2m}(G_V+Y_{11})}{G_V(G_E+Y_{11}+Y_{21})+Y_{11}G_E};$$

and thus the output resistance

$$R_{out} = \frac{G_V+Y_{11}}{G_V(G_E+Y_{11}+Y_{21})+Y_{11}G_E}$$

If the *signal source is an ideal voltage* source, then $R_i = 0 \rightarrow G_V = G_1 + G_2 + G_i \rightarrow \infty$ and thus

$$R_{out}(R_i = 0) = \lim_{G_V \to \infty} \left(\frac{G_V + Y_{11}}{G_V(G_E + Y_{11} + Y_{21}) + Y_{11}G_E} \right) = \frac{1}{G_E + Y_{11} + Y_{21}}$$

$$= \frac{\beta \cdot r_e R_E}{\beta \cdot r_e + R_E + \beta \cdot R_E} = \left| \text{large value } \beta \right| \cong \frac{r_e R_E}{r_e + R_E} = \frac{r_e}{1 + \frac{r_e}{R_E}}$$

6.1.2.1 Notice – buffer large signal properties

The basic circuit in Fig. 6.2 is suitable for small signal changes only. Transistor supplies (sources) current into the load and capacity C_{V2}, when the input is positive – for both small and large inputs. The circuit time constant is small. When the input goes negative and small, the transistor decreases its current – it is still in the "active" region. The time constant is still small. If the input signal is negative and large the transistor is closed – the output capacity discharges only via R_E – the time constant is "long" – it is very bad situation. That's why there is used the circuit shoved below (principle only).

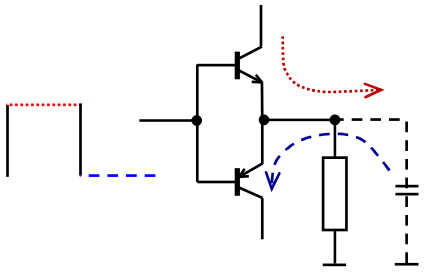


Fig. 6.4 Common collector connection (follower) – complementary

NPN transistor sources (supplies) current into the load when the input goes positive and transistor PNP sinks load current when the input goes negative.

6.1.3 Common base configuration (CB)

The common base configuration (Fig. 6.5a) of BJT we easily get from the circuit in Fig. 6.1 – the same quiescent point. The capacitor C_B connects the BJT base to the common node (for AC signals) - input current enters into the BJT emitter. An equivalent AC model is in fig. 6.5b.

We determine an admittance circuit model (C_{CB} neglected now and $Y_{22} = 0$, too):

	1, E	2, C			
1,E	$G_E + \boldsymbol{Y_{11}} + \boldsymbol{Y_{21}}$	0	U_1	=	I_1
2,C	$-\boldsymbol{Y_{21}}$	G_C	U_2		0

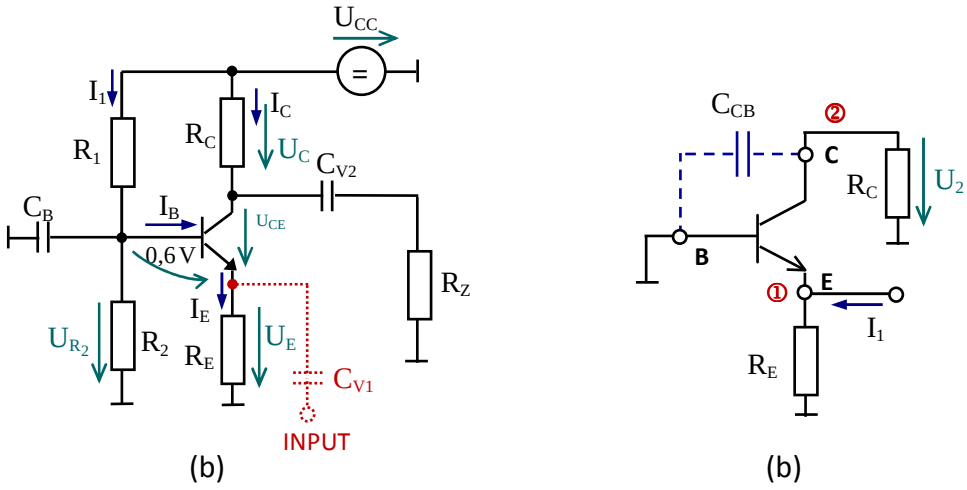


Fig. 6.5 a) Common base connection; b) Equivalent signal model of CB

It is evident (see the second raw) that

$$-Y_{21}U_1 + G_C U_2 = 0$$

Thus we can determine that the CB gain is

$$\frac{U_2}{U_1} = \frac{Y_{21}}{G_C} = \frac{R_C}{r_e}$$

It is evident, too, (see the first matrix raw) that

$$U_1(G_E + Y_{11} + Y_{21}) = I_1$$

The CB input resistance is

$$R_{in} = \frac{U_1}{I_1} = \frac{1}{G_e + Y_{11} + Y_{21}} \cong \frac{r_e R_E}{r_e + R_E}$$

We can easily determine that the output resistance equals R_C.

6.1.4 Effect of the C_{CB}

Now we use the BJT admittance matrix "with the C_{CB}"; we suppose simplified formulas (if $\beta \gg 1$)

$$Y_{11} + Y_{21} = \frac{1}{\beta r_e} + \frac{1}{r_e} \cong Y_{21}.$$

We will solve common emitter connection. Then we get (see Chapter 5, Example 5.1)

$$\begin{bmatrix} G_V + Y_{11} + \boldsymbol{pC_{CB}} & -\boldsymbol{pC_{CB}} & -Y_{11} \\ Y_{21} - \boldsymbol{pC_{CB}} & G_C + \boldsymbol{pC_{CB}} & -Y_{21} \\ -Y_{21} & & G_E + Y_{21} \end{bmatrix} \cdot \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} I_1 \\ 0 \\ 0 \end{bmatrix}$$

From this admittance circuit model we easily get

$$A_{U_{CCE}} = \frac{Y_{21}G_E - pC_{CB}(G_E + Y_{21})}{G_C(G_E + Y_{21}) + pC_{CB}(G_E + Y_{21})} = \frac{p - \frac{Y_{21}G_E}{C_{CB}(G_E + Y_{21})}}{p + \frac{G_C}{C_{CB}}} = \frac{p - \omega_n}{p + \omega_p}$$

where

$\omega_p = \frac{1}{R_C C_{CB}}$ is a pole of the transfer function and

$\omega_n = \frac{Y_{21}G_E}{C_{CB}(G_E+Y_{21})} = \frac{1}{C_{CB}(R_E+r_e)} = \frac{R_C}{R_E+r_e} \cdot \frac{1}{C_{CB}R_C} = \frac{R_C}{R_E} \cdot \frac{1}{1+\frac{r_e}{R_E}} \cdot \omega_p$

is a zero of the transfer function.

If $\omega \to 0$ then $p \to 0$, too, and

$$A_{U_{CCE}}(\omega < \omega_p) = \frac{-\omega_n}{\omega_p} = -\frac{R_C}{R_E} \cdot \frac{1}{1+\frac{r_e}{R_E}}$$

If $\omega \to \infty$ then $A_{U_{CCE}}(\omega \to \infty) = 1$; it is a feed through via C_{CB} (into the transistor collector) – see Fig. 6.6.

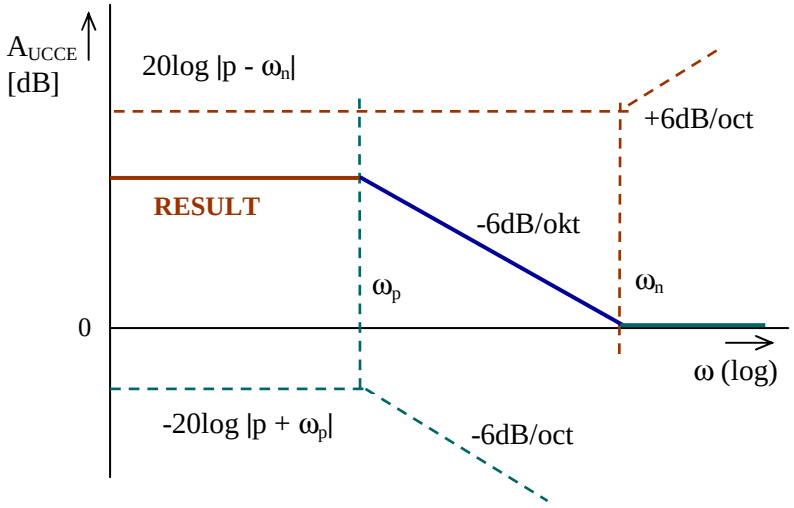


Fig. 6.6 A modulus characteristic of the CE connection – the influence of the C_{CB}

We do not investigate an influence of coupling capacitors.

Determination of input and output impedances from the known admittance model is easy.

Another explanation - Miller effect (more physical):

Consider an ideal voltage amplifier of gain A_v with an impedance Z connected between its input and output nodes. The output voltage is therefore $U_o = A_v U_i$ and the input current is

$$I_i = \frac{U_i - U_o}{Z} = \frac{U_i(1-A_v)}{Z}$$

The equivalent input impedance of the circuit is

$$Z_{in} = \frac{U_i}{I_i} = \frac{Z}{1-A_v}$$

If Z represents a capacitor C_{CB}, then $Z = 1/(j\omega C_{CB})$ and the resulting input impedance is

$$Z_{in} = \frac{Z}{1-A_v} = \frac{1}{j\omega C_{CB}(1-A_v)} = \frac{1}{j\omega C_M}$$

where

$$C_M = C_{CB}(1 - A_v) = \left| A_v = \frac{-R_C}{R_e}; common\ emitter\ connection \right| = C_{CB}\left(1 + \frac{R_C}{R_e}\right)$$

This "Miller capacitance" creates a large input time constant – and it is not good. The frequency properties are poorer.

The "collector" time constant is $\approx R_C C_{CB}$, always. The voltage across C_{CB} is

$$U_2 - U_1 = U_2 - U_2/A_{UCCE} \approx U_2;$$

the component U_2/A_{UCCE} has no cardinal meaning – it changes not "output" frequency ratios.

If you investigate signal models, you can see that in the common collector connection (Fig. 6.2; follower) the C_{CB} is connected directly to the common node (reference) – it creates "simple" time constant in the base. In the common base connection the C_{CB} creates "simple" time constant in the collector (Fig. 6.5). From this point of view it is the worst situation in the common emitter configuration where "Miller capacitance exists".

connection	voltage gain	current gain	input resistance	output resistance	frequency properties
CE	$-R_C/(R_e + r_e)$	$\approx \beta$	middle	R_C	worst
CC	$R_E/(R_E + r_e)$	$\approx \beta$	large	r_e	good
CB	R_C/r_e	≈ 1	r_e- little	R_C	good

6.1.5 Cascode amplifier with two BJTs

The cascode connection of two BJTs you can see in Fig. 6.7a. We suppose the same transistor properties. A quiescent point definition is evident – see Fig. 6.7a, too. An AC model is depicted in Fig. 6.7b; the C_B capacity – its X_C is negligible at all frequencies likely to be encountered for the amplifier.

We will suppose that $Y_{21} \gg Y_{11}$ ($\beta \gg 1$), then $Y_{21} + Y_{11} \cong Y_{21} = g_e = 1/r_e$ and the transistor admittance matrix is:

	B	C	E
B	Y_{11}	0	$-Y_{11}$
C	Y_{21}	0	$-Y_{21}$
E	$-Y_{21}$	0	Y_{21}

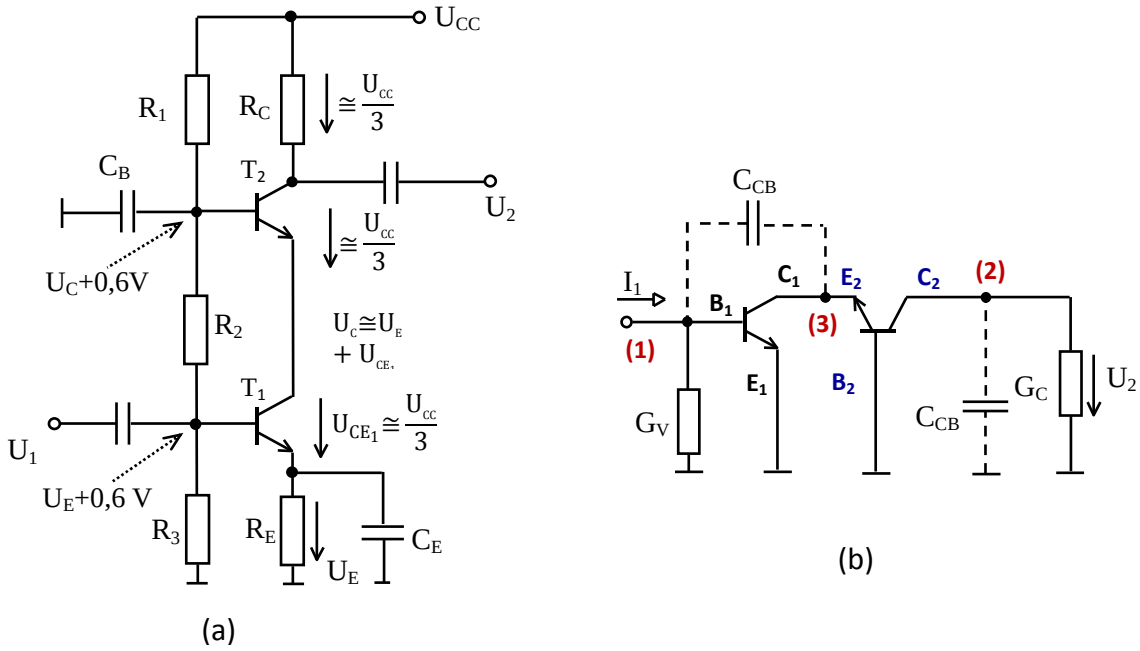


Fig. 6.7 BJTs cascode connection– a) and its AC model – b); $G_V = G_i + G_2 + G_3$; G_i is equivalent signal source conductance

We make up a matrix model of the circuit:

	1(**B1**)	2 [**C2**]	3 (**C1**) [**E2**]			
1(**B1**)	$G_V + pC_{CB} + (\boldsymbol{Y_{11}})$		$-pC_{CB} + (\mathbf{0})$	U_1	=	I_1
2[**C2**]		$G_C + pC_{CB} + [\mathbf{0}]$	$[-\boldsymbol{Y_{21}}]$	U_2		0
3(**C1**)[**E2**]	$-pC_{CB} + (\boldsymbol{Y_{21}})$	$[\mathbf{0}]$	$pC_{CB} + [\boldsymbol{Y_{21}}] + (\mathbf{0})$	U_3		0

$B_1 - B_1 \rightarrow$ put Y_{11} into „area 11“of passive matrix

$B_1 - C_2 \rightarrow$ „area 12“of passive matrix - it is not a coincidence – ***nodes of different transistors (in the investigated circuit; generally electronic elements) do not create coincidences***

$B_1 - C_1 \rightarrow$ put 0 into „area 13“of passive matrix

$B_1 - E_2 \rightarrow$ „area 13“of passive matrix - it is not coincidence

$C_2 - B_1 \rightarrow$ „area 21“of passive matrix - it is not coincidence; .....

etc.

Now we can determine the input impedance, if $\omega < 1/(r_e C_{CB})$,

$$Z_{in} = \frac{1}{G_v + Y_{11} + 2pC_{CB}}$$

It is evident that the C_{CB} has double effect only.

T_1 is a common emitter amplifier – it drives a common base stage (T_2) with input resistance about r_e. It means that a T_1 voltage gain is just $-r_e/r_e = -1$. Thus

The “creates” only current gain. “creates” only voltage gain about . We can determine that

We rearrange this formula

;

A influence is suppressed now – but we must use two transistors. So we get good frequency properties.

6.1.6 Common source configuration (CS; NMOSFET enhancement mode)

Let us solve an example 6.1.

Example 6.1

This basic connection you can see in Fig. 6.8 – , , and . We suppose that parameters of a FET are: , , . Determine basic properties of the circuit.

☑ Solution:

We must determine a quiescent point first. We suppose saturation transistor mode. It is evident that DC voltage is

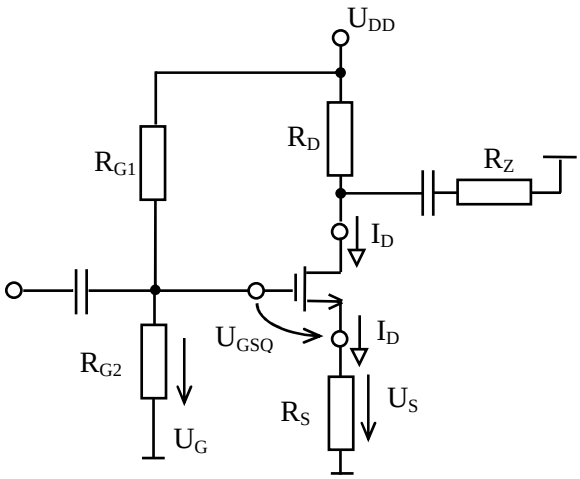


Fig. 6.8 Basic common source connection

Simultaneously

$$U_S = R_S I_D$$

so that

$$U_{GS} = U_G - R_S I_D.$$

We can determine now that

$$I_D = K[(U_G - R_S I_D) - U_T]^2 = K[(U_G - U_T) - R_S I_D]^2$$

Rearranging this formula we get

$$10^4 I_D^2 - 707 I_D + 3{,}41 = 0$$

We take the only physically solution $I_D = 5{,}2$ mA which gives

$$U_{DS} = U_{DD} - I_D(R_D + R_S) = 10 - 5{,}2 \cdot 10^{-3} \cdot 1{,}1 \cdot 10^3 = 4{,}28 \text{ V}$$

In the saturation mode must be

$$U_{DS} > U_{GS} - U_T = U_G - R_S I_D - U_T = 3{,}85 - 0{,}52 - 2 = 1{,}33 \text{ V}$$

This condition is true, our presumption was right. Thus we can determine

$$g_m = 2 \cdot \sqrt{K \cdot I_D} = 2 \cdot \sqrt{2{,}96 \cdot 10^{-3} \cdot 5{,}2 \cdot 10^{-3}} = 7{,}845 \text{ mS}$$

and

$$r_d = \frac{1}{g_d} = \frac{U_A}{I_D} = \frac{156}{5{,}2 \cdot 10^{-3}} = 30 \text{ k}\Omega$$

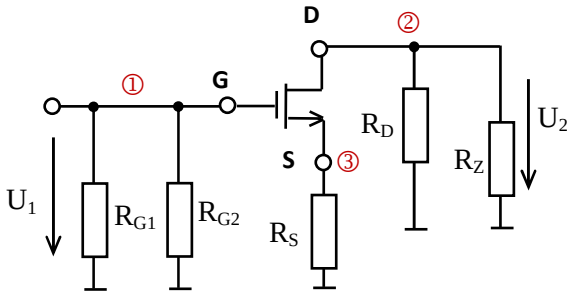


Fig. 6.9 An equivalent AC model

The AC equivalent you can see in the Fig. 6.9 and a corresponding matrix model is:

	1, G	2, D	3, S			
1, G	$G_{G_1} + G_{G_2}$	0	0	U_1	I_1
2, D	g_m	$G_D + G_Z + g_d$	$-g_m - g_d$	$\cdot U_2$	$= 0$
3, S	$-g_m$	$-g_d$	$G_S + g_m + g_d$	U_3	0

If we suppose voltage signal source U_1 (known nodal voltage) we can solve just equations

$G_D + G_Z + g_d$	$-g_m - g_d$	U_2	$=$	$-g_m U_1$
$-g_d$	$G_S + g_m + g_d$	U_3		$g_m U_1$

We can add the second raw to the first row (Gaussian elimination), we get

		=	

and so

We substitute known parameters and get

An input resistance is defined by parallel connection of and , thus 92,3 kΩ. An output resistance we determine by means of "measuring current " into the node 2. We suppose that voltage source is disconnected now. We will see that resistors and have no influence on the output resistance (for MOSFETs). Corresponding matrix model is now

If we do not reason about load influence ( ), we simply give . Then we ordinarily get

where

For the given parameters we get

Ω

and so

Ω

6.1.7 Coscode amplifier with FET and BJT

Let us solve an example 6.2.

Example 6.2

Determine a voltage gain of the circuit in Fig. 6.10 if: , , , , , and , (FET): , , ; (BJT): ,

☑ Solution:

We must determine quiescent points of transistors, first. Well defined is voltage – see resistor divider :

$$U_B = \frac{U_{CC}R_{B_2}}{R_{B_1}+R_{B_2}} = \frac{15\cdot 3{,}9\cdot 10^3}{(4{,}7+3{,}9)\cdot 10^3} = 6{,}8\ \mathrm{V}$$

We suppose that base current of T_2 is insignificant now. The DC voltage of the emitter E (T_2) is

$$U_E \cong U_B - 0{,}6 = 6{,}2\ \mathrm{V}.$$

A current of T_1 is defined by formula

$$I_D = I_{DSS}\left(1 - \frac{U_{GS}}{U_P}\right)^2.$$

At the same time there is

$$U_{GS} = -R_S I_D.$$

The DC current I_D creates DC gate bias voltage.

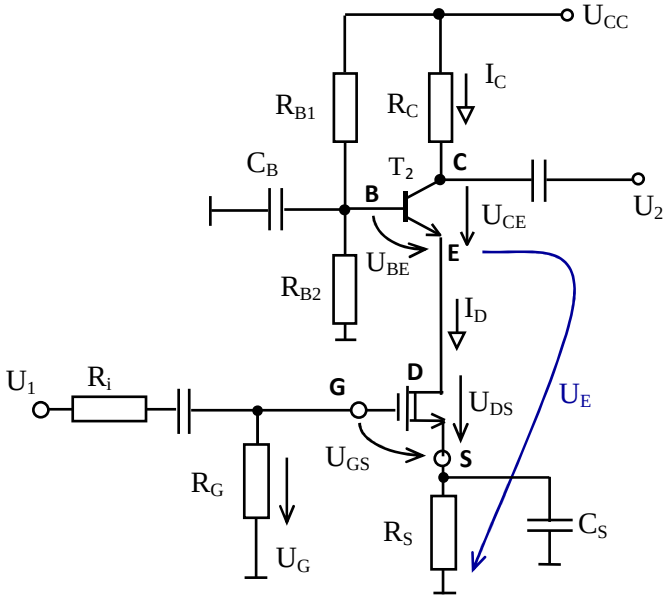


Fig. 6.10 FET and BJT cascode connection

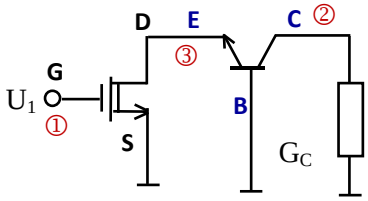


Fig. 6.11 An AC equivalent

If we substitute U_{GS} into I_D, we get (after basic rearranging)

$$I_D^2\left(\frac{R_S}{U_P}\right)^2 + I_D\left(\frac{2R_S}{U_P} - \frac{1}{I_{DSS}}\right) + 1 = 0$$

and for the given parameters

$$302500 \cdot I_D^2 - 1300 \cdot I_D + 1 = 0$$

Solving this, we get two different currents 1,004 mA and 3,294 mA. Right solution is value 1,004 mA because it creates right DC gate voltage $U_{GS} = -R_S \cdot 1{,}004 \cdot 10^{-3} = -2{,}209$ V.

Now we can determine voltage

$$U_{DS} = U_E - R_S I_D = 6{,}2 - 2{,}209 = 4 \text{ V } (T_1).$$

This voltage must be greater than

$$U_{DS_P} = U_{GS} - U_P = -2{,}209 - (-4) = 1{,}791 \text{ V}$$

in order to work the T_1 in the active mode (fulfilled).

Thus it is $I_D \cong 1\text{mA}$, $U_{DS} = 4$ V. Further $I_C \cong I_D$ (if we have large β of T_2) and

$$U_{CE} \cong U_{CC} - U_E - R_C I_D = 15 - 6{,}2 - 3{,}9 = 4{,}9 \text{ V}.$$

It means that the BJT works right – in the active mode, too.

Now we can determine needed parameters of the T_1:

$$g_{m_1} = \frac{2I_D}{U_{GS} - U_P} = \frac{2 \cdot 10^{-3}}{-2{,}209 - (-4)} = 1{,}117 \text{ mS}$$

$$r_{d_1} = \frac{1}{g_{d_1}} = \left|\frac{U_A}{I_D}\right| = \frac{200}{10^{-3}} = 200 \text{ kΩ}$$

	G	D	S
G	0	0	0
D	g_{m_1}	g_{d_1}	$-g_{m_1} - g_{d_1}$
S	$-g_{m_1}$	$-g_{d_1}$	$g_{m_1} + g_{d_1}$

$\cong$

	G	D	S
G	0	0	0
D	g_{m_1}	g_{d_1}	$-g_{m_1}$
S	$-g_{m_1}$	$-g_{d_1}$	g_{m_1}

And we can determine parameters of the T_2, too:

$$g_e = \frac{1}{r_e} = \frac{I_C}{26 \cdot 10^{-3}} = \frac{10^{-3}}{26 \cdot 10^{-3}} = 38{,}46 \text{ mS}$$

$$r_{d_2} = \frac{1}{g_{d_2}} == \frac{250}{10^{-3}} = 250 \text{ kΩ}$$

	B	C	E
B	g_e/β	0	$-g_e/\beta$
C	g_e	g_{d_2}	$-g_e - g_{d_2}$
E	$-g_e - g_e/\beta$	$-g_{d_2}$	$g_e + g_e/\beta + g_{d_2}$

$= |\beta \gg 1;\ g_e \gg g_{d_2}| \cong$

	B	C	E
B	g_e/β	0	$-g_e/\beta$
C	g_e	g_{d_2}	$-g_e$
E	$-g_e$	$-g_{d_2}$	g_e

Notice that g_e is much greater than g_{m_1} if the quiescent current is the same!!!!

The AC equivalent is in Fig. 6.11. We neglect resistance 600 Ω and R_G. We easily get an admittance model of the circuit (in the given quiescent point):

1, G 2, C 3, D,E

1, G	0	0	0	$U_1 \equiv U_i$	=	$I_1 = 0$
2, C	0	$G_C + g_{d_2}$	$-g_e$	U_2		0
3, D, E	g_{m_1}	$-g_{d_2}$	$g_{d_1} + g_e$	U_3		0

We suppose voltage input signal $U_1 \equiv U_i$. In this case is $I_1 = 0$. Further there is

$$G_C \gg g_{d_2} \text{ a } g_e \gg g_{d_1}$$

so that

G_C	$-g_e$	U_2	=	0
$-g_{d_2}$	g_e	U_3		$-g_{m_1}U_i$

and we easily get

$$U_2 = \frac{-g_e g_{m_1} U_i}{G_C g_e - g_e g_{d_2}} = \frac{-g_e g_{m_1} U_i}{(G_C - g_{d_2}) g_e} \cong \frac{g_{m_1} U_i}{G_C}$$

The voltage gain is

$$\frac{U_2}{U_i} \cong -g_{m_1} R_C = -1{,}117 \cdot 10^{-3} \cdot 3{,}9 \cdot 10^3 = -4{,}356$$

If we do not neglect some components (see text above), we get $U_2/U_i = -4{,}335$. You can see that this deviation is very small.

6.1.8 Triode amplifier

There is a basic connection of the triode amplifier in the Fig. 6.12. A grid – cathode voltage is always negative. Data sheets usually give a tube quiescent point as well as tube's parameters. The resistor R_g voltage U_{R_g} is 0 V because no significant current flows through R_g (if grid voltage is negative). Thus it is valid (resistor R_C generates the U_g – compare depletion FET – N channel): $U_g = -R_C I_a$. A capacitance C_C shorts out R_C - shorts AC signals – see an AC model in Fig. 6.13.

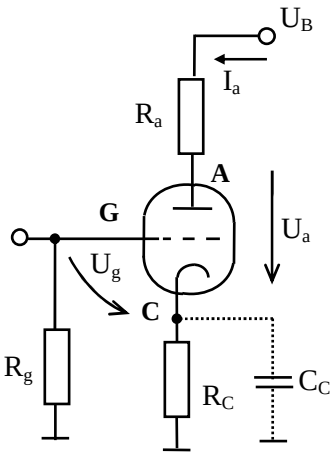


Fig. 6.12 Basic triode amplifier

The known admittance triode model is (extended)

	G	A	C			
G	0	0	0	U_g	=	I_g
A	S	G_i	$-S-G_i$	U_a		I_a
C	$-S$	$-G_i$	$S+G_i$	U_C		I_C

Formally accordance with FETs depiction is evident – just give $S \rightarrow g_m$ and $G_i \rightarrow g_d$.

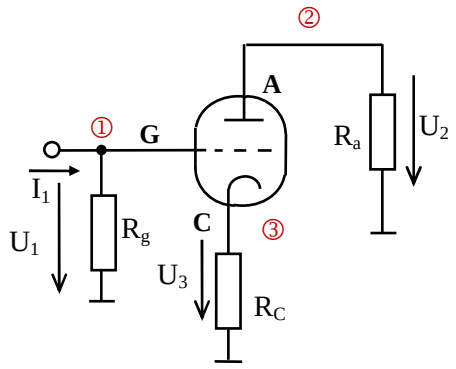


Fig. 6.13 AC model of the amplifier (no C_C now)

If we know an admittance triode model, the next procedure is evident. The AC amplifier model is in Fig. 6.13. There is just one signal source (I_1) in the figure (we suppose "no C_C now"):

	1, G	2, A	3, C			
1, G	G_g	0	0	U_1	=	I_1
2, A	S	G_a+G_i	$-S-G_i$	U_2		0
3, C	$-S$	$-G_i$	G_C+S+G_i	U_3		0

We can add the second raw to the third row (Gaussian elimination), we get

G_g	0	0	U_1	=	I_1
S	G_a+G_i	$-S-G_i$	U_2		0
0	G_a	G_C	U_3		0

From this admittance circuit model we derive

$$\frac{U_2}{U_1} = \frac{-SG_C}{(G_a+G_i)G_C+(S+G_i)G_a}$$

Another rearranging gives us

$$\frac{U_2}{U_1} = \frac{-SR_iR_a}{R_i+R_a+(1+SR_i)R_C} = \frac{-\mu R_a}{R_i+R_a+(1+\mu)R_C}$$

Similarly we get

$$\frac{U_3}{U_1} = \frac{SG_C}{(G_a+G_i)G_C+(S+G_i)G_a} = \frac{SR_iR_C}{R_i+R_a+(1+SR_i)R_C} = \frac{\mu R_C}{R_i+R_a+(1+\mu)R_C}$$

Example 6.3

Determine a voltage gain of the circuit in Fig. 6.12 if it is given: , , , , , (double triode ECC83) and we require .

☑ Solution:

For this quiescent point and we can easily calculate a supply voltage from a simple equation .

Now we can calculate

— —————————

— —————————

If a bypass capacitor is used - it transfers AC signal directly to the ground, thus bypassing . For AC, is short circuited (and therefore its action is ignored), whereas for DC it comes still into play and is considered as an open circuit. Thus in formulas above we can simply substitute

— ——— ———————

If we suppose now an infinite value of (so that infinite , too), we get

— ———

If the is omitted (and too) we must reduce the supply voltage for – we get voltage follower (cathode) – but not very good this time:

— ——————— ———————

Hereafter compare a solution of a similar circuit with FE.

6.1.9 FE amplifier with resistors

Let us solve an example 6.4.

Example 6.4

An analyzed circuit is in Fig. 6.14 (; kΩ; Ω). The quiescent point as so as an admittance JFET model we know: , (see Chapter 2; Fig. 2.30). Determine properties of the circuit.

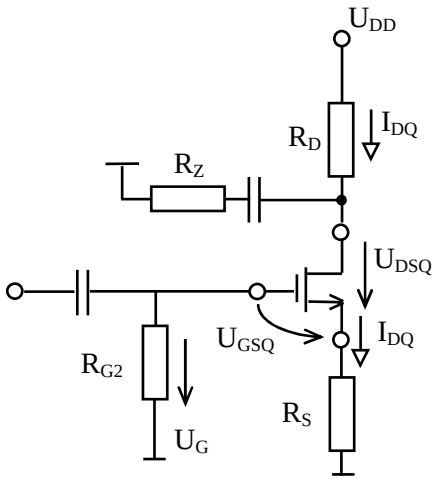


Fig.6.14a) An amplifier with R_S

	G	D	S
G	0	0	0
D	g_m	g_d	$-g_m-g_d$
S	$-g_m$	$-g_d$	g_m+g_d

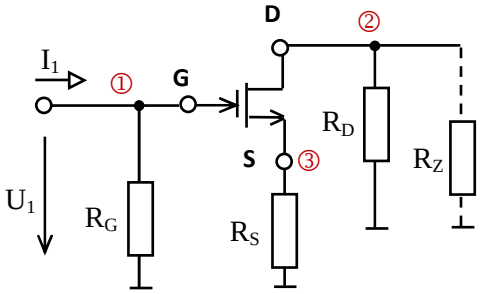


Fig. 6.14b) An AC model

☑ Solution:

We easily get an admitance model of the circuit – see Fig. 6.14b (in the given quiescent point, we neglect load resistance R_Z now):

	1(G)	2(D)	3(S)			
1(G)	$G_G+(0)$	$0+(0)$	$0+(0)$		U_1	I_1
2(D)	$0+(g_m)$	$G_D+(g_d)$	$0-(g_m+g_d)$	$\cdot$	U_2	$=$ 0
3(S)	$0+(-g_m)$	$0+(-g_d)$	$G_S+(g_m+g_d)$		U_3	0

We can determine

$$\frac{U_2}{U_1}=\frac{-g_m G_S}{(G_D+g_d)G_S+g_m G_D+g_d G_D}=\frac{-g_m r_d R_D}{r_d+R_D+(1+g_m r_D)R_S}$$

(See the „triode amplifier"– result is formally the same if we substitute $S \to g_m$, $G_D \to G_a$, $G_i \to g_d$, $G_C \to G_S$).

If we substitute known circuit parameters, we get:

$$\frac{U_2}{U_1}=\frac{1{,}44\cdot10^{-3}\cdot46{,}3\cdot10^3 2{,}2\cdot10^3}{46{,}3\cdot10^3+2{,}2\cdot10^3+926\cdot(1+1{,}44\cdot10^{-3}\cdot46{,}3\cdot10^3)}=-1{,}319$$

If we need to analyze an influence of R_Z we simply do a substitution $R_D \to R_D \| R_Z$ (1,8 kΩ for given values – it is just a parallel connection of R_Z and R_D):

$$\frac{U_2}{U_1}=\frac{-g_m r_d(R_D\|R_Z)}{r_d+(R_D\|R_Z)+(1+g_m r_D)R_S}=\frac{1{,}44\cdot10^{-3}\cdot46{,}3\cdot10^3 1{,}8\cdot10^3}{46{,}3\cdot10^3+1{,}8\cdot10^3+926\cdot(1+1{,}44\cdot10^{-3}\cdot46{,}3\cdot10^3)}=-1{,}083$$

If a bypass capacitor C_S is used - it transfers AC signal directly to the ground, thus bypassing R_S. For AC, R_S is short circuited (and therefore its action is ignored), whereas for DC it comes still into play and C_S is considered as an open circuit. Thus in formulas above we can simply substitute

$$\frac{U_2}{U_1}(R_S\to 0)=\frac{-g_m r_d(R_D\|R_Z)}{r_d+(R_D\|R_Z)}=\frac{1{,}44\cdot10^{-3}\cdot46{,}3\cdot10^3 1{,}8\cdot10^3}{46{,}3\cdot10^3+1{,}8\cdot10^3}=-2{,}49$$

We can determine a transfer function into node 3 – follower. We leave in the resistor R_D so the quiescent point is the same. We easily determine ("no R_Z" – it is neglected now):

$$\frac{U_3}{U_1} = \frac{g_m G_D}{G_D(G_S+g_m+g_d)+g_d G_S} = \frac{g_m r_d R_S}{r_d+R_D+(1+g_m r_d)R_S}$$

For the given „circuit values"it gives $U_3/U_1 = 0{,}555$. Only if $g_m r_d R_S \gg r_d+R_D$ we can get U_3/U_1 near the value 1 – we must use a transistor with large g_m (or use a current source instead of R_S). It is evident, too, that R_D has not large influence if $R_D \ll r_d$.

Let us derive an output resistance of the circuit (of the node 2). We "use" a current source I_2 in to node 2, now. The admittance model is the same – see matrix model below:

	1(G)	2(D)	3(S)			
1(G)	$G_G + (0)$	$0 + (0)$	$0 + (0)$	U_1'	=	0
2(D)	$0 + (g_m)$	$G_D + (g_d)$	$0 - (g_m + g_d)$	U_2'	=	I_2
3(S)	$0 + (-g_m)$	$0 + (-g_d)$	$G_S + (g_m + g_d)$	U_3'	=	0

Now is valid

$$R_{out_2} = \frac{U_2'}{I_2} = \frac{G_S+g_m+g_d}{G_D(G_S+g_m+g_d)+g_d G_S} = \frac{1}{\frac{1}{R_D}+\frac{1}{R_{OD}}}$$

The output resistance we get as a parallel connection of the R_D and a transistor output resistance R_{OD} (without influence of the R_D now):

$$R_{OD} = \frac{G_S+g_m+g_d}{g_d G_S} = r_d + R_S(1 + g_m r_d)$$

Thus

$$R_{out_2} = R_D \| R_{OD}$$

and so

$$R_{OD} = 46{,}3 \cdot 10^3 + 926 \cdot (1 + 1{,}44 \cdot 10^{-3} \cdot 46{,}3 \cdot 10^3) = 108{,}9\ \text{k}\Omega$$

$$R_{out_2} = 108{,}9 \cdot 10^3 \| 2{,}2 \cdot 10^3 = 2{,}16\ \text{k}\Omega$$

In the same way we can determine an output resistance of the node 3. We "use" a current source I_3 in to node 3, now. The admittance model is the same – see matrix model below:

	1(G)	2(D)	3(S)			
1(G)	$G_G + (0)$	$0 + (0)$	$0 + (0)$	U_1''	=	0
2(D)	$0 + (g_m)$	$G_D + (g_d)$	$0 - (g_m + g_d)$	U_2''	=	0
3(S)	$0 + (-g_m)$	$0 + (-g_d)$	$G_S + (g_m + g_d)$	U_3''	=	I_3

Now is valid

$$R_{out_3} = \frac{U_3'''}{I_3} = \frac{1}{G_S+\frac{G_D(g_m+g_d)}{G_D+g_d}} = \frac{1}{\frac{1}{R_S}+\frac{1}{R_{OS}}}$$

The output resistance we get as a parallel connection of the R_S and a transistor output resistance R_{OS} (without influence of the R_S now):

$$R_{OS} = \frac{G_D+g_d}{(g_m+g_d)G_D} = \frac{r_d+R_D}{1+g_m r_d}$$

$$R_{OUT_3} = R_S \| R_{OS}$$

Thus

$$R_{OS} = \frac{(46{,}3+2.2)\cdot 10^3}{1+1{,}44\cdot 10^{-3}\cdot 46{,}3\cdot 10^3} = 717\ \text{k}\Omega$$

$$R_{OUT_3} = 926 \| 717 = 404\ \Omega$$

It is evident that a principal solution of linearized circuits with tri - poles is always the same. If we use a general description "with Y_{11}, Y_{21}, etc." results will be formally identical. We will just substitute different values of Y_{11}, Y_{21}, etc. for different tri – poles.

We can see that followers with FETs are not very good – a value of g_m is small (considerably smaller than g_e of BJTs for the same quiescent current). The small value of g_m gives a small value of voltage gain for CS connection, too. This problem we can solve by means of a current source instead of R_D – it defines the proper quiescent current but its signal resistance is ideally infinite – really about 100 kΩ at 1 mA (it is valid for small voltages, too). Notice that passive resistance 100 kΩ at 1 mA needs a voltage 100 V!!! The same solution can be used for BJTs – we get gains 1000 (60 dB) in one amplifier stage.

6.2 An linear circuits with operational amplifiers (OPA)

An idealized OPA model is described in Chapter 3:

	(+)	(-)	(o)			
(+)	0	0	0	U_+		I_+
(-)	0	0	0	U_-	=	I_-
(o)	$-A\ G_o$	$A\ G_o$	G_o	U_o		I_o

An OPA gain is frequency dependent – always. The simplest model of gain is

$$A(p) = \frac{A_o\omega_1}{p+\omega_1} = |\omega_T = A_o\omega_1| = \frac{\omega_T}{p+\omega_1} \qquad (6.5)$$

$p = j\omega$ (harmonic steady state)
A_o is DC OPA gain
ω_1 is the first pole (frequency) of an OPA (dominant)

Correspondent module and phase characteristic you can see in Fig. 6.15.

If $\omega \gg \omega_1$, it is valid (1 pole model)

$$A(p) \cong \frac{A_o\omega_1}{p} = \frac{\omega_T}{p} = \frac{-j\omega_T}{\omega} \qquad (6.6)$$

$\omega_T \cong A_o\omega_1$ is gain-bandwidth product of the OPA (GBW).

Just from eq. (6.6) it is evident that a slope for frequencies $\omega > \omega_1$ is -20dB/dec. If $\omega = \omega_T$ than$|A(\omega_T)| = |-j| = 1$. A phase for $\omega > \omega_1$ (1 pole model) is still -90°.

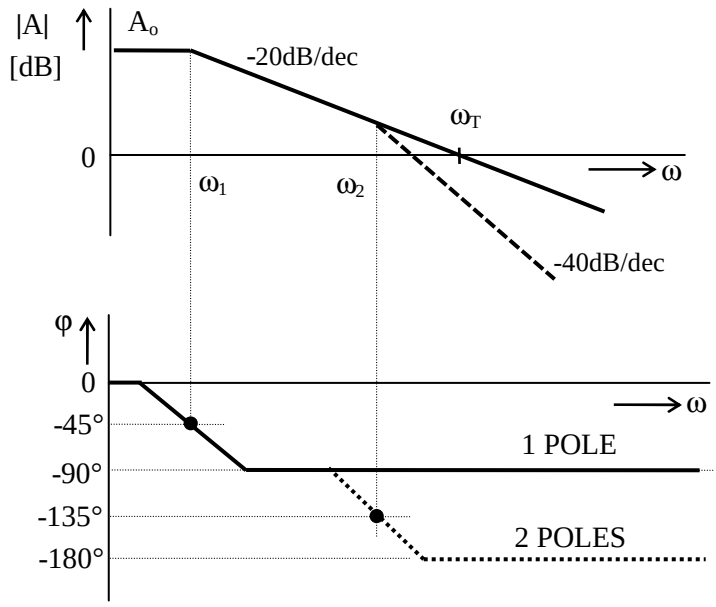


Fig. 6.15 Module and phase characteristic of an OPA; 1 POLE model – solid line; 2 POLE model – dashed line, ω_2 – the second OPA pole

If we investigate frequency stability of a structure, it is useful to apply "more-pole models" of the OPA gain A.

Frequency degradation of A is the main cause of errors in structures with OPAs because ideally we suppose that $A \to \infty$.

OPA output impedance is usually „ohmic" only – just R_o. Only for very high frequencies we must respect inductive character of the OPA output impedance. Ideally there is $G_o = 1/R_o \to \infty$.

An idealization of input OPA impedances we can easily correct – this input impedances we draw directly in "passive" part of the circuit.

6.2.1 Current – to – voltage converter

Basic circuit diagram you can see in Fig. 6.16.

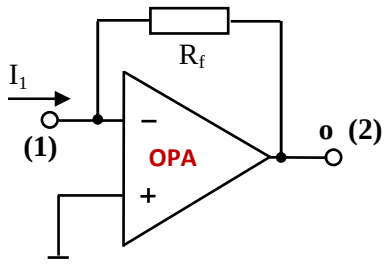


Fig. 6.16 Current – to voltage - converter

There are three nodes in the circuit only - thus matrix model (ordinary) is “2x2” only. We determine “passive matrix”:

	1	2			
1	G_f	$-G_f$	U_1	$=$	I_1
2	$-G_f$	G_f	U_2		0

Then we „add“ used active element matrix (OPA now; OPA “node +” is connected to a reference node – it has no effect now, correspondent row and column we can delete anyway):

	1, **(-)**	2, **(o)**			
1, **(-)**	G_f	$-G_f$	U_1	=	I_1
2, **(o)**	$-G_f + \boldsymbol{AG_o}$	$G_f + \boldsymbol{G_o}$	U_2		0

Now we determine

$$U_2 = \frac{D_2}{D} = \frac{\begin{vmatrix} G_f & I_1 \\ -G_f + AG_o & 0 \end{vmatrix}}{\begin{vmatrix} G_f & -G_f \\ -G_f + AG_o & G_f + G_o \end{vmatrix}} = \frac{-I_1(AG_o - G_f)}{G_f G_o (1 + A)} = -R_f I_1 \frac{1 - \frac{G_f}{AG_o}}{1 + \frac{1}{A}}$$

The first component

$$U_{2_{ID}} = -R_f I_1$$

defines an ideal transfer function (with $A \to \infty$), the second component defines an error function which is ideally 1 if $A \to \infty$.

Input impedance we determine from the known node voltage U_1:

$$U_1 = \frac{D_1}{D} = \frac{\begin{vmatrix} I_1 & -G_f \\ 0 & G_f + G_o \end{vmatrix}}{G_f G_o (1 + A)} = \frac{I_1(G_o + G_f)}{G_f G_o (1 + A)}$$

$$Z_{in} = \frac{U_1}{I_1} = \frac{R_o}{1 + A} + \frac{R_f}{1 + A}$$

Output impedance we determine by means of current I_2 into node 2 of the circuit:

	1, **(-)**	2, **(o)**			
1, **(-)**	G_f	$-G_f$	U_1'	$=$	0
2, **(o)**	$-G_f + \boldsymbol{AG_o}$	$G_f + \boldsymbol{G_o}$	U_2'		I_2

We determine voltage U_2' and so we will determine output impedance of the structure (node 2 impedance):

$$U_2' = \frac{D_2'}{D} = \frac{\begin{vmatrix} G_f & 0 \\ -G_f + AG_o & I_2 \end{vmatrix}}{\begin{vmatrix} G_f & -G_f \\ -G_f + AG_o & G_f + G_o \end{vmatrix}} = \frac{-I_2 G_f}{G_f G_o (1 + A)}$$

$$Z_{out} = \frac{U_2'}{I_2} = \frac{R_o}{1 + A}$$

Influence of OPA differential resistance

We simply add in R_d – see Fig. 6.17 – and solve this problem by routine method.

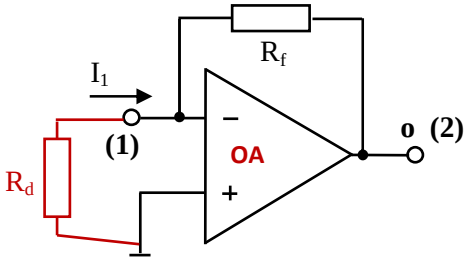


Fig. 6.17 Influence of R_d

Thus

	1, (-)	2, (o)			
1, (-)	$G_f + G_d$	$-G_f$	U_1	=	I_1
2, (o)	$-G_f + AG_o$	$G_f + G_o$	U_2		0

We can do all needed calculations from this simply model "with incorporated R_d".

6.2.2 Inverting amplifier with OPA

An inverting amplifier with OPA is in Fig. 6.18.

Noninverting input has no effect, again. We determine matrix model (ordinary) with one current signal source:

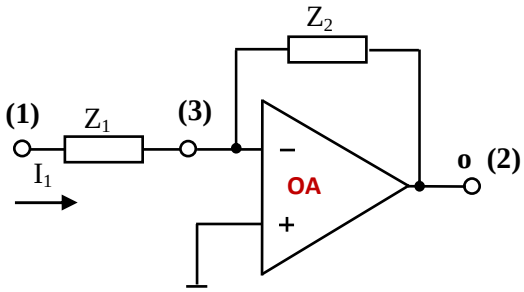


Fig. 6.18 An inverting amplifier

	1	2(o)	3(-)				
1	Y_1	0	$-Y_1$		U_1		I_1
2(o)	0	$Y_2+(G_o)$	$-Y_2+(AG_o)$	$\cdot$	U_2	$=$	0
3(-)	$-Y_1$	$-Y_2$	Y_1+Y_2		U_3		0

Let us suppose known voltage signal source U_1, now. The above model is still formally valid. We must give known "terms with U_1" to the "right" and unknown current I_1 to the "left":

	1	2(o)	3(-)				
1	-1	0	$-Y_1$		I_1		$-Y_1U_1$
2(o)	0	$Y_2+(G_o)$	$-Y_2+(AG_o)$	$\cdot$	U_2	$=$	0
3(-)	0	$-Y_2$	Y_1+Y_2		U_3		Y_1U_1

It is enough to solve the below given equation system only

$Y_2+(G_o)$	$-Y_2+(AG_o)$	$\cdot$	U_2	$=$ 0
$-Y_2$	Y_1+Y_2		U_3	Y_1U_1

From this we determine

$$U_2=\frac{\begin{vmatrix}0 & -Y_2+AG_o\\ Y_1U_1 & Y_1+Y_2\end{vmatrix}}{\begin{vmatrix}Y_2+G_o & -Y_2+AG_o\\ -Y_2 & Y_1+Y_2\end{vmatrix}}=\frac{-Y_1U_1(AG_o-Y_2)}{Y_1Y_2+G_o(Y_1+Y_2)+Y_2AG_o}$$

Searched gain of the invertor is

$$\frac{U_2}{U_1}=\frac{-Y_1(AG_o-Y_2)}{Y_1Y_2+G_o(Y_1+Y_2)+Y_2AG_o}=-\frac{Y_1}{Y_2}\cdot\frac{AG_o-Y_2}{Y_1+G_o\left(1+\frac{Y_1}{Y_2}\right)+AG_o}$$

The first term defines an ideal structure gain (for $A\to\infty$):

$$\left.\frac{U_2}{U_1}\right|_{ID}=-\frac{Y_1}{Y_2}=-\frac{Z_2}{Z_1}$$

The second term defines an error function of the inverting structure

$$E_{IN}=\frac{AG_o-Y_2}{Y_1+G_o\left(1+\frac{Y_1}{Y_2}\right)+AG_o}=\frac{AG_o}{Y_1+G_o\left(1+\frac{Y_1}{Y_2}\right)+AG_o}-\frac{Y_2}{Y_1+G_o\left(1+\frac{Y_1}{Y_2}\right)+AG_o}$$

This is ideally 1 (0 dB) if we suppose the ideal OPA with $A\to\infty$.

It is evident that a term

$$\frac{Y_2}{Y_1+G_o\left(1+\frac{Y_2}{Y_1}\right)+AG_o}=0$$

if $G_o\to\infty$ ($R_o=0$) – in this case quantity of A has no influence.

This term defines a **feedthrough** – *a transit of the input signal via Z_1 and Z_2 into output (nonzero) resistance R_o of the OPA. As the "A" decreases (with increasing frequency) the feedthrough increases.* In practice we use the OPA only for frequencies where A is large enough and so we usually can neglect the feedthrough - thus it is valid:

$$E_{IN} = \frac{AG_o}{Y_1 + G_o\left(1 + \frac{Y_1}{Y_2}\right) + AG_o} = \cdots = \frac{1}{1 + \frac{Z_2 + R_o}{Z_1}} = \left|\text{if } R_o \ll |Z_2|\right|$$

$$= \frac{1}{1 + \frac{Z_2}{Z_1}}$$

If we need to know input impedance we must determine current I_1 and node voltage U_3:

$$U_3 = \frac{\begin{vmatrix} Y_2 + G_o & 0 \\ -Y_1 & Y_1U_1 \end{vmatrix}}{\begin{vmatrix} Y_2 + G_o & -Y_2 + AG_o \\ -Y_2 & Y_1 + Y_2 \end{vmatrix}} = \frac{Y_1U_1(G_o + Y_2)}{Y_1Y_2 + G_o(Y_1 + Y_2) + Y_2AG_o}$$

By means of Ohm's law we determine

$$I_1 = \frac{U_1 - U_3}{Z_1} = \cdots = Y_1U_1 \cdot \frac{Y_2G_o + Y_2AG_o}{Y_1Y_2 + G_o(Y_1 + Y_2) + Y_2AG_o}$$

therefore

$$Z_{in} = \frac{U_1}{I_1} = \frac{1}{Y_1} \cdot \frac{Y_1Y_2 + G_o(Y_1 + Y_2) + Y_2AG_o}{Y_2G_o + Y_2AG_o} = \cdots = \frac{1}{Y_1} + \frac{1}{G_o + AG_o} + \frac{1}{Y_2 + Y_2A}$$

Thus

$$Z_{in} = Z_1 + \frac{R_o}{1 + A} + \frac{Z_2}{1 + A}$$

Compare this product with the input impedance of the current – to – voltage converter. It is evident that we must add Z_1 only.

If we want to determine output impedance (= node 2 impedance) of the inverting structure, we replace the signal voltage source U_1 by its internal resistance – it means ideally zero – and we feed signal current I_2 into node 2. It means that node 1 is short-circuited (to reference node) and we must omit the first row ant the first column in the admittance model (as marked below):

	1	2(o)	3(-)			
1	Y_1	0	$-Y_1$	U_1	=	0
2(o)	0	$Y_2 + (G_o)$	$-Y_2 + (AG_o)$	U_2'		I_2
3(-)	$-Y_1$	$-Y_2$	$Y_1 + Y_2$	U_3'		0

An admittance model for determination of output impedance is:

$$\begin{array}{|c|c|}\hline Y_2+(G_o) & -Y_2+(AG_o)\\ \hline -Y_2 & Y_1+Y_2\\ \hline\end{array}\cdot\begin{array}{|c|}\hline U_2'\\ \hline U_3'\\ \hline\end{array}=\begin{array}{|c|}\hline I_2\\ \hline 0\\ \hline\end{array}$$

Now we determine that

$$U_2'=\frac{\begin{vmatrix} I_2 & -Y_2+AG_o\\ 0 & Y_1+Y_2\end{vmatrix}}{\begin{vmatrix} Y_2+G_o & -Y_2+AG_o\\ -Y_2 & Y_1+Y_2\end{vmatrix}}=\frac{I_2(Y_1+Y_2)}{Y_1Y_2+G_o(Y_1+Y_2)+Y_2AG_o}$$

From this equation we get the output impedance of the inverting amplifier

$$Z_{out}=\frac{U_2'}{I_2}=\frac{Y_1+Y_2}{Y_1Y_2+G_o(Y_1+Y_2)+Y_2AG_o}=\cdots=\frac{R_o}{1+A\frac{R_1}{R_1+R_2}+\frac{R_o}{R_1+R_2}}$$

6.2.3 Noninverting amplifier with OPA

A noninverting amplifier with OPA is in Fig. 6.19.

Resistor R_P has functional meaning – it makes possible to pass a DC input OPA current if input signal is fed into noninverting input via coupling capacity. If this DC path is not available – the OPA will be saturate.

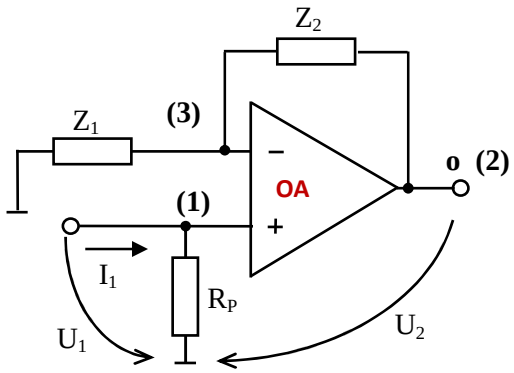


Fig. 6.19 A noninverting amplifier

We determine admittance model of the noninverting structure:

$$\begin{array}{c|c|c|c|}\hline & 1 & 2(o) & 3(-)\\ \hline 1 & G_P & 0 & -Y_1\\ \hline 2(o) & -(AG_o) & Y_2+(G_o) & -Y_2+(AG_o)\\ \hline 3(-) & 0 & -Y_2 & Y_1+Y_2\\ \hline\end{array}\cdot\begin{array}{|c|}\hline U_1\\ \hline U_2\\ \hline U_3\\ \hline\end{array}=\begin{array}{|c|}\hline I_1\\ \hline 0\\ \hline 0\\ \hline\end{array}$$

Let us suppose known voltage signal source U_1, again. The above model is still formally valid. We must give known "terms with U_1" to the "right" and unknown current I_1 to the "left":

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & Y_2+(G_o) & -Y_2+(AG_o) \\ 0 & -Y_2 & Y_1+Y_2 \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} -G_P U_1 \\ AG_o U_1 \\ 0 \end{bmatrix}$$

We solve this equation system „without I_1 “:

$$\begin{bmatrix} Y_2+(G_o) & -Y_2+(AG_o) \\ -Y_2 & Y_1+Y_2 \end{bmatrix} \cdot \begin{bmatrix} U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} AG_o U_1 \\ 0 \end{bmatrix}$$

Now we easily get node 2 voltage

$$U_2 = \frac{\begin{vmatrix} AG_o U_1 & -Y_2+AG_o \\ 0 & Y_1+Y_2 \end{vmatrix}}{\begin{vmatrix} Y_2+G_o & -Y_2+AG_o \\ -Y_2 & Y_1+Y_2 \end{vmatrix}} = \frac{AG_o U_1 (Y_1+Y_2)}{Y_1Y_2+G_o(Y_1+Y_2)+Y_2AG_o}$$

thus noninverting amplifier gain is

$$\frac{U_2}{U_1} = \frac{AG_o(Y_1+Y_2)}{Y_1Y_2+G_o(Y_1+Y_2)+Y_2AG_o} = \frac{Y_1+Y_2}{Y_2} \cdot \frac{AG_o}{Y_1+G_o\left(1+\frac{Y_1}{Y_2}\right)+AG_o}$$

We rearrange this formula and

$$\frac{U_2}{U_1} = \left(1+\frac{Z_2}{Z_1}\right) \cdot \frac{1}{1+\frac{1+\frac{Z_2+R_o}{Z_1}}{A}}$$

The first term defines an ideal structure gain (for $A \to \infty$):

$$\left.\frac{U_2}{U_1}\right|_{ID} = 1+\frac{Y_1}{Y_2} = 1+\frac{Z_2}{Z_1}$$

The second term defines an error function of the noninverting structure

$$E_N = \frac{1}{1+\frac{1+\frac{Z_2+R_o}{Z_1}}{A}}$$

Notice that this formula is the same as the error function of the inverting amplifier neglecting $R_o(=0)$ – thus feedthrough of the inverter. This notice is acceptable if OPA gain (A) is large enough – and this is an area of normal OPA function.

We can easily determine that input resistance is defined just by the R_P value and the output resistance is the same as for the inverting structure.

6.2.4 Differential amplifier with OPA

A differential amplifier with OPA is in Fig. 6.20.

We could determine an admittance model, again. It is easy. But we can also use superposition principle (superposition property) which is valid for all linear systems:

The net response at a given place and time caused by two or more stimuli is the sum of the responses which would have been caused by each stimulus individually.

Let us determine an influence of the voltage source U_A, first. The second source (U_B) we replace by its internal resistance – ideally zero – see Fig. 6.21.

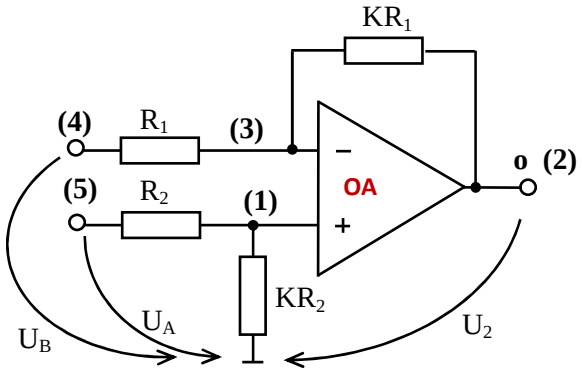


Fig. 6.20 A differential amplifier

It is evident that it is a *noninverting amplifier* with input voltage divider R_2 and KR_2. We easily determine an output voltage (caused by U_A). Thus

$$U_{2_A} = U_A \cdot \frac{KR_2}{R_2+KR_2} \cdot \left(1 + \frac{KR_1}{R_1}\right) \cdot E_N = KU_A E_N$$

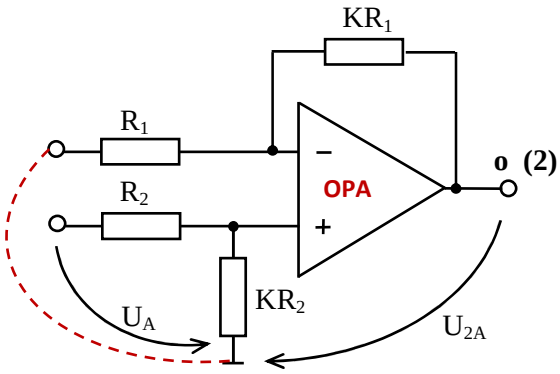


Fig. 6.21 An influence of the voltage U_A

Next we determine an influence of the voltage source U_B – the voltage source U_A we replace by its internal resistance – ideally zero – see Fig. 6.22.

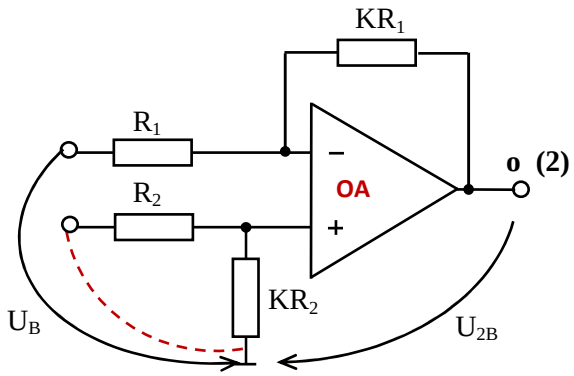


Fig. 6.22 An influence of the voltage U_B

A parallel connection of R_2 and KR_2 has no effect now (if OPA is ideal – zero input currents) – in Fig. 6.22 is *inverting amplifier*. Thus we easy determine an output voltage (caused by U_B):

$$U_{2_B} = -U_B \cdot \frac{KR_1}{R_1} \cdot E_{IN} = -KU_B E_{IN}$$

Resulting voltage is (superposition)

$$U_2 = U_{2_A} + U_{2_B} = KU_A E_N - KU_B E_{IN}$$

If feedthrough is neglectable (see inverting and noninverting amplifier error functions), it is valid

$$E_{IN} \cong E_N = \frac{1}{1 + \frac{1 + \frac{Z_2 + R_o}{Z_1}}{A}}$$

and output voltage of the differential amplifier is

$$U_2 = K(U_A - U_B) \cdot \frac{1}{1 + \frac{1 + \frac{Z_2 + R_o}{Z_1}}{A}}$$

$$Z_2 = KR_1\,; \quad Z_1 = R_1$$

The output voltage is a function of input voltages difference only.

If we have different rates of resistors ($R_1 - K_1R_1$; $R_2 - K_2R_2$) we get none zero output voltage for $U_A = U_B$ – bad common voltage rejection.

6.2.5 3 – dB frequency – f_3

Let us suppose that feedthrough is neglectable and we have the same error function for inverting and noninverting amplifiers:

$$E = \frac{1}{1 + \frac{1 + \frac{R_2}{R_1}}{A}}$$

Voltage gain off all analyzed structures we can define as below (U_{out} – output voltage; U_{in} – input voltage)

$$\frac{U_{out}}{U_{in}} = \frac{U_{out}}{U_{in}}\bigg|_{ID} \cdot E,$$

where

$\frac{U_{out}}{U_{in}}\Big|_{ID} = -\frac{R_2}{R_1}$ - inverting amplifier ideal gain

$\frac{U_{out}}{U_{in}}\Big|_{ID} = 1 + \frac{R_2}{R_1}$ - noninverting amplifier ideal gain

$\frac{U_{out}}{U_{in}}\Big|_{ID} = \frac{U_{out}}{U_A - U_B}\Big|_{ID} = K$ - differential amplifier ideal gain

We use absolute values and determine gain „in dB“:

$$\frac{U_{out}}{U_{in}}\bigg|_{ID} = 20\log\left|\frac{U_{out}}{U_{in}}\right|_{ID} + 20\log|E|$$

3 – dB decrease compared with the ideal value occurs if $|E| = \frac{1}{\sqrt{2}}$.

Thus it is enough to investigate the same error function for all structures above. We suppose simplified formula of the OPA gain $A \cong -j\omega_T/\omega$, where $\omega_T \cong A_o\omega_1$ is gain – bandwidth product. We easily get

$$E = \frac{1}{\frac{1 + \frac{R_2}{R_1}}{A}} = \frac{1}{1 + j\frac{1 + \frac{R_2}{R_1}}{\frac{\omega_T}{\omega}}} = \frac{1}{1 + j\frac{1 + \omega}{\frac{\omega_T}{\left(1 + \frac{R_2}{R_1}\right)}}} = \frac{1}{1 + j\frac{\omega}{\omega_3}}$$

Just if $\omega = \omega_3$, it is valid

$$\frac{\omega_3}{\frac{\omega_T}{\left(1 + \frac{R_2}{R_1}\right)}} = 1$$

and

$$|E| = \left|\frac{1}{1 + j\frac{\omega}{\omega_3}}\right| = \left|\frac{1}{1 + j}\right| = \frac{1}{\sqrt{2}} \qquad \Rightarrow \; 20\log\sqrt{2} = -3\ \text{dB}$$

Thus $\omega_3 = \omega_T/(1 + R_2/R_1) \;\rightarrow\; f_3 = f_T/(1 + R_2/R_1) \;\Rightarrow\; (1 + R_2/R_1) = \omega_T/\omega_3$

From these equations it is evident that if $\omega = \omega_3$ *then the term* $1 + R_2/R_1$ *equates to a modulus of the OPA gain – for all three described structures – see Fig. 6.23.*

Real behavior of the gain is marked by a dotted line. It is evident that *shift in phase* of the amplifier is -45° (compared with the ideal value) if $\omega = \omega_3$.

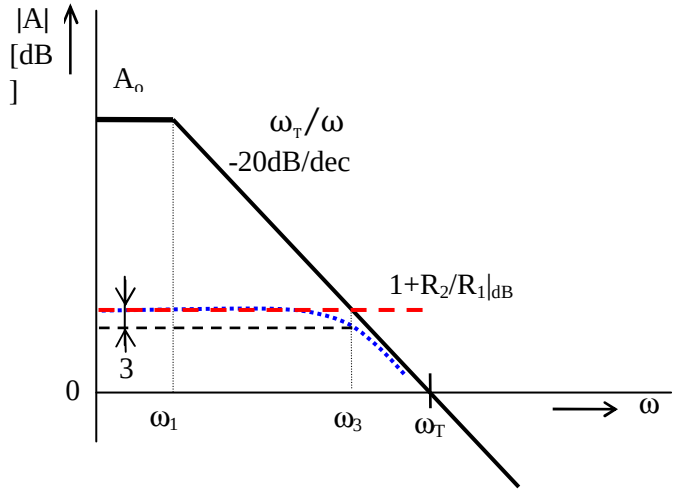


Fig. 6.23 A depiction of the modulus OPA gain and its coincidence with $1 + R_2/R_1$

6.2.6 Wien oscillator

A basic Wien oscillator structure we can see in Fig. 6.24.

An oscillator needs no signal sources – it generates signals by itself. An admittance model does not change if we feed a current source (signal) into any node – for example I_1 into node 2 in Fig. 6.24. Ideal current source internal resistance is infinite so its admittance is zero and does not change node admittance. In the same way the admittance model does not change if we put an ideal voltage source on series with any circuit element – U_i on series with R_1 in Fig. 6.24. Ideal voltage source internal resistance is zero and does not change impedance of the circuit element (on series) – so it does not change the admittance model. (A voltage source between any node and reference node shorts out this node – signal short circuit – thus changes the admittance model of the circuit; a current source on series with R_1 disconnects this resistor – signal open circuit – thus changes the admittance model of the circuit).

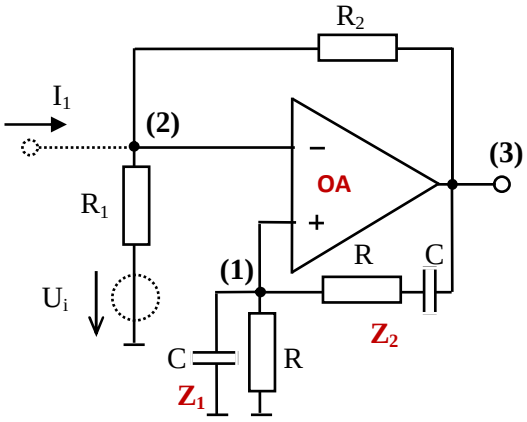


Fig. 6.24: A basic Wien oscillator structure

It enables us to determine an admittance model of the circuit in Fig. 6.24 – we suppose signal source I_1 now. Let us classify

$$Z_1 = R/(1+pCR);\quad Y_1 = 1/Z_1 = G + pC$$
$$Z_2 = R + 1/(pC);\quad Y_2 = 1/Z_2 = pCG/(G+pC)$$

Routine method gives us

	1 (+)	2 (-)	3(o)			
1(+)	$Y_1 + Y_2$	0	$-Y_2$	U_1	=	I_1
2(-)	0	$G_1 + G_2$	$-G_2$	U_2	=	0
3(o)	$-AG_o - Y_2$	$-G_2 + AG_o$	$G_2 + Y_2 + G_o$	U_3	=	0

We get matrix determinant

$$D = G_o[A(G_2Y_1 - G_1Y_2) + (Y_1 + Y_2)(G_1 + G_2)] + G_1G_2(Y_1 + Y_2) + Y_1Y_2(G_1 + G_2)$$

If we suppose infinite OPA gain, it is enough to use only

$$D(A \to \infty) = G_oA(G_2Y_1 - G_1Y_2)$$

(other terms are insignificant). We substitute Y_1 and Y_2 and rearrange equation:

$$D(A \to \infty) = \frac{G_oAG_2C^2\left[p^2 + pG\dfrac{\left(2 - \dfrac{G_1}{G_2}\right)}{C} + \dfrac{G^2}{C^2}\right]}{G + pC}$$

We need just zero real part of equation root (*Laplace domain*)

$$p^2 + pG\frac{\left(2 - \frac{G_1}{G_2}\right)}{C} + \frac{G^2}{C^2} = 0$$

- it is critical stability**.** We must ensure

$$2 - \frac{G_1}{G_2} = 0 \quad \Rightarrow \quad \frac{R_2}{R_1} = 2$$

Then are two complex conjugate roots

$$p^2 + \frac{G^2}{C^2} = 0 \quad \Rightarrow \quad p_{1,2} = \frac{\pm j}{RC} = \pm j\omega_o$$

Then the circuit *will be oscillate* on frequency

$$\omega_o = \frac{1}{RC}$$

Let us investigate our problem in the *frequency domain* (harmonic steady state). Then is valid $p = j\omega$ and we get

$$D(A \to \infty) = \frac{G_oAG_2C^2\left[(j\omega)^2 + j\omega\dfrac{2 - R_2/R_1}{RC} + \left(\dfrac{1}{RC}\right)^2\right]}{G + j\omega C}$$

The nodal voltages must be infinite – this means the circuit will oscillate – if $D(\omega_o) \to 0$, thus

$$\left(\frac{1}{RC}\right)^2 - (\omega_o)^2 + j\omega_o \frac{2 - \frac{R_2}{R_1}}{RC} = 0$$

then

$$(\omega_o)^2 = \left(\frac{1}{RC}\right)^2 \qquad \text{and} \qquad R_2/R_1 = 2$$

We get the same solution as above.

Previous solution comes out from the admittance description of the linear electronic circuit. It does not use a feedback theory construction. But ***right determined model always includes "feedback properties" of the investigated circuit – "automatically"****. The same problem we can solve by means of feedback theory. We must get the same solution, of course:*

Let us separate the circuit in Fig. 6.24 into two parts. Node 1 is an input of a noninverting amplifier, node 3 is an output of this amplifier. An ideally gain of the noninverting amplifier is $U_3/U_1 = 1 + R_2/R_1$ – see Fig. 6.25.

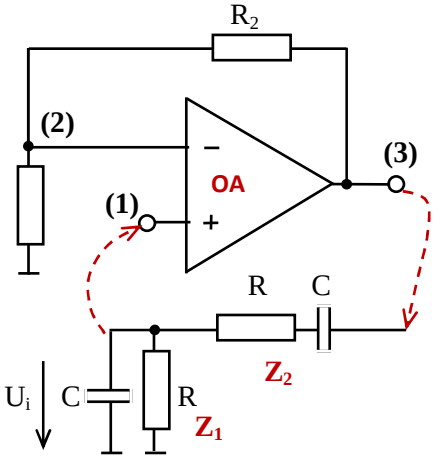


Fig. 6.25 A basic Wien oscillator structure – feedback point of view

A part of the output voltage is fed back to the noninverting amplifier input via the RC network. It can be determined that a RC network transfer is (impedance divider)

$$\frac{U_1}{U_3} = \frac{pRC}{p^2 + p \cdot \frac{3}{RC} + \left(\frac{1}{RC}\right)^2} = \Big|_{p = j\omega} = \frac{j\omega RC}{\left(\frac{1}{RC}\right)^2 - \omega^2 + j\omega \frac{3}{RC}}$$

It is evident that it is a noninverting band pass filter with characteristic (here resonant) frequency $\omega_0 = 1/RC$ and $U_1/U_3\ (\omega = \omega_0) = 1/3$ (quality factor $Q = 1/3$).

Loop gain on $\omega = \omega_0$ is

$$\frac{U_3}{U_1} \cdot \frac{U_1}{U_3} = \left(1 + \frac{R_2}{R_1}\right) \cdot \frac{1}{3}$$

From the feedback theory it must be just one (or larger) if circuit is supposed to oscillate (stability criterion, critical stability). Thus it must be

$$\left(1+\frac{R_2}{R_1}\right)\cdot\frac{1}{3}=1$$

Then we must guarantee $R_2/R_1 > 2$. The voltage gain of the noninverting amplifier must be 3 (or larger) - set by the resistor network R_2 and R_1.

6.2.7 Phase shift oscillator

One version of such oscillator we can see in Fig. 6.26. We determine an admittance model of the oscillator – we do admittance matrix only:

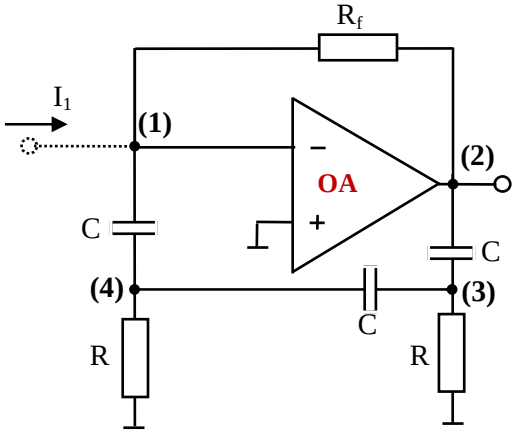


Fig. 6.26 Possible phase shift oscillator

	1 (-)	2 (o)	3(o)	4
1(-)	$G_f + pC$	$-G_f$	0	$-pC$
2(o)	$-G_f + AG_{of}$	$G_f + pC$+G$_o$	$-pC$	0
3	0	$-pC$	$2pC + G$	$-pC$
4	$-pC$	0	$-pC$	$2pC + G$

If we suppose infinite OPA gain ($A \to \infty$) it is enough to use only hereinafter significant terms of determinant (*Cofactor expansion – 1. column;* $p = j\omega$; frequency domain)

$$-G_f(G^2 - 3\omega^2C^2) - j\omega C\left(4G_fG - \omega^2C^2\right)$$

Others terms are insignificant. The circuit will oscillate if

$$-G_f(G^2 - 3\omega^2C^2) - j\omega C\left(4G_fG - \omega^2C^2\right) = 0$$

Hence we get

$$-G_f(G^2 - 3\omega^2C^2) = 0 \quad \to \quad (\omega_o)^2 = \frac{G^2}{3C^2} = \frac{1}{3R^2C^2}$$

And (if $\omega = \omega_o$)

$$4G_fG - \omega^2C^2 = 0 \quad \to \quad R_f = \frac{1}{G_f} = \frac{4}{\omega_o^2RC^2} = 12R$$

If we want to investigate real OPA properties we simply substitute real A and G_o of OPAs. The problem (solution) will be more complex but more difficult, too.

Let us solve the same problem ***by means of feedback theory***. OPA + R_f creates simply current to voltage converter – inverting – see fig. 6.27 – it is valid (ideally) $U_2 = -R_f I_f$.

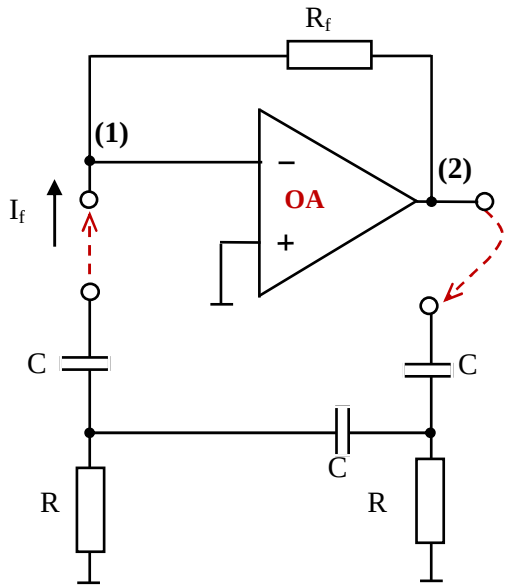


Fig. 6.26 Phase shift oscillator with current to voltage converter – feedback point of view

We must determine a transfer function of a feedback circuit RC; now U_2 is the input voltage and I_f is the output current – see Fig. 6.27 below.

By any known method (impedance divider; nodal voltage analysis ...) we determine that

$$I_f = \frac{U_2 p^3 C^3}{3p^2C^2 + 4pCG + G^2}$$

Then *loop gain* is

$$\left.\frac{U_2}{I_f}\right|_{\text{CONVERTER}} \cdot \left.\frac{I_f}{U_2}\right|_{\text{RC-FEEDBACK}} = -R_f \cdot \frac{p^3C^3}{3p^2C^2 + 4pCG + G^2} = \left|_{p\,=\,j\omega}\right| = \cdots$$

$$= R_f \cdot \frac{j\omega^3 \frac{C}{3}}{-\omega^2 + \frac{1}{3R^2C^2} + j\omega\frac{4G}{3C}}$$

If we need oscillations it must be

$$R_f \cdot \frac{j\omega^3 \frac{C}{3}}{-\omega^2 + \frac{1}{3R^2C^2} + j\omega\frac{4G}{3C}} = 1$$

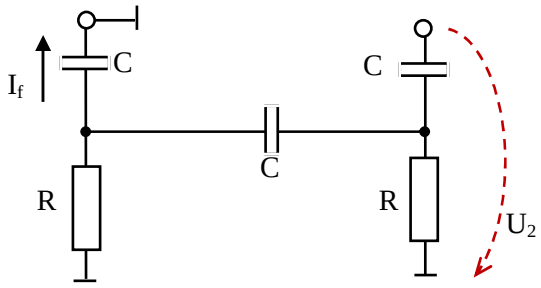


Fig. 6.27 A feedback circuit RC from the fig. 26

Phase condition (requirement) (phase shift 180° creates feedback circuit RC, another shift 180° creates converter) will be right if

$$-\omega_o^2 + \frac{1}{3R^2C^2} = 0$$

Thus we can determine that

$$\omega_o^2 = \frac{1}{3R^2C^2}$$

This we substitute to the formula above and get

$$R_f \cdot \frac{j\omega_o^3 \frac{C}{3}}{0 + j\omega_o \frac{4G}{3C}} = \frac{\omega_o^2 RC^2}{4} = 1$$

Then to fulfill ***magnitude condition*** *(requirement)* it must be

$$R_f = \frac{4}{\omega_o^2 RC^2} = \frac{4(3R^2C^2)}{RC^2} = 12R$$

We get the same solution as above – from the matrix model.

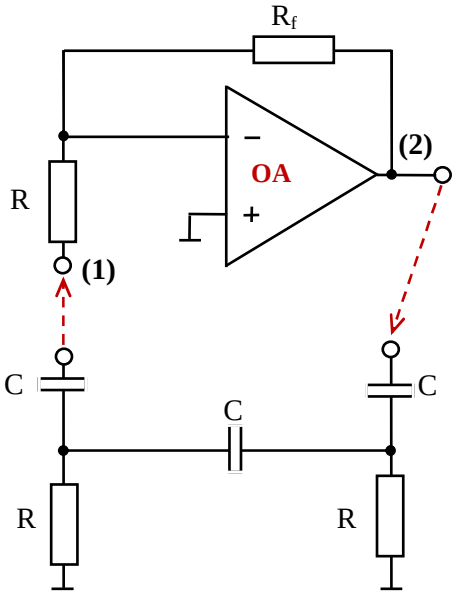


Fig. 6.28 A more known phase shift oscillator

More known circuit with phase shift you can see in fig. 6.28. The OPA creates an inverting voltage amplifier with input resistance R and gain (voltage) – R_f/R. Thus we must determine voltage transfer function of a feedback circuit RC – from the node 2 to the node 1– equivalent signal circuit see fig. 6.29.

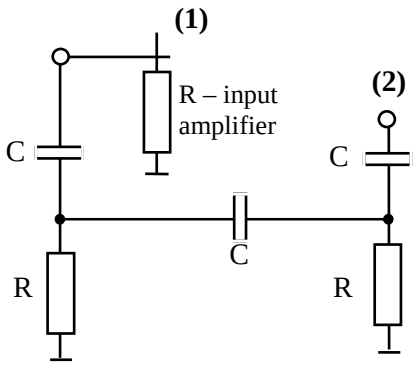


Fig. 6.29 A feedback circuit RC from the fig. 28

By any known method (impedance divider; nodal voltage analysis ...) we determine that

$$\frac{U_1}{U_2} = \frac{p^3C^3}{p^3C^3+p^2C^26G+pC5G^2+G^3} = \cdots = \frac{-j\omega^3}{j\omega\left(\frac{5}{R^2C^2-\omega^2}\right)+\frac{1}{RC}\left(\frac{1}{R^2C^2}-6\omega^2\right)}$$

If

$$\frac{1}{R^2C^2} - 6\omega^2 = 0,$$

thus for

$$\omega_o^2 = \frac{1}{6R^2C^2}.,$$

is a feedback transfer function just

$$\frac{U_1}{U_2}(\omega_o) = \frac{-j\omega_o^3}{j\omega_o\left(\frac{5}{R^2C^2}-\omega_o^2\right)+\frac{1}{RC}(0)} = -\frac{1}{29}$$

This way is phase condition realized.

To realize magnitude condition we must ensure

$$\frac{-R_f}{R}\cdot\left(-\frac{1}{29}\right) = 1.$$

Thus we must realize

$$R_f = 29R$$

Summary

1) There are analysed basic transistor, triode and OPA circuits in this chapter – by means of the generalized nodal voltage analysis.
2) Interesting (important) problems:

a. Large signal properties of the buffer

b. An effect of C_{CB} – Miller effect

c. 3 – dB frequency f_3

d. Oscillation – a phase condition,a magnitude condition

3) A value of g_m (FET) is considerably smaller than g_e of BJTs for the same quiescent current – the small value of g_m gives a small value of voltage gain.

4) Right determined admittance models of electronic circuits always include “feedback informations” - automatically.

Questions 6

You can find the answers in the text of this chapter.

1. Define a voltage gain of inverting amplifier with an ideal OPA.
2. Define a voltage gain of a noninverting amplifier with an ideal OPA.
3. May we use the superposition principle to determine a voltage gain of a differential amplifier with an ideal OPA if we know voltage gains from points 1. and 2.?
4. What is the most severe degradation of real amplifier structures?
5. Describe phase and magnitude conditions of oscillations – couple it with admittance model.

Problems 6

Example 6.1

Analyze properties of circuit in Fig. 6.1 (A quiescent point is determined in Chapter 2: $U_{CC} = 12$ V; $R_E = 1$ kΩ; $R_C = 10$ kΩ; $R_1 = 270$ kΩ; $R_2 = 30$ kΩ; $\beta = 100$; $r_e = 43$ Ω). Determine

a) voltage gain
b) input resistance
c) output resistance

Example 6.2

Let us suppose the same conditions as above. Determine

a) a pole of the transfer function
b) a zero of the transfer function
c) an equivalent input capacitance if you know that $C_{CB} = 3$ pF

Example 6.3

Solve tasks from points 1. and 2. if (Fig. 6.1b) you know that a resistor $R'_E = 100\ \Omega$ is used. Compare results and give reasons for differences.

PROBLEMS KEY 6

Ad example 6.1)
Use findings in the point 6.1.1. *Common emitter connection* (CE); $R_e = R_E \| R'_E = R_E$

Ad example 6.2)
Use findings in the point 6.1.4. *Effect of the* C_{CB}; $R_e = R_E \| R'_E = R_E$

Ad example 6.3)
Use findings in the point 6.1.1. *Common emitter connection* (CE) and .1.4. *Effect of the* C_{CB}; $R_e = R_E \| R'_E$

Basic texts

[1] Mikulec, M., – Havlíček, V.: Basic circuit theory. Vydavatelství ČVUT, Praha, 2005, ISBN 80-01-03172-1

[2] Punčochář, J.: Admittance models of modern linear amplifying structures. Transactions of the VŠB – TU Ostrava, 1, 2003, vol. VI, pp 151 – 161, ISBN 80-248-0223-6

[3] Mohylová, J.: Lineární obvody s elektronickými prvky-Sbírka příkladů, VŠB-TU Ostrava 2002, ISBN 80-248-0098-5

[4] Punčochář, J.: Lineární obvody s elektronickými prvky. Skriptum, VŠB-TU Ostrava 2002, ISBN 80-248-0040-3

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

7. Analysis of frequency dependent structures 2nd order filters

Time of study: 6 hours

Goals: the student should be able to

- define ideal transfer functions of 2nd order filters: Low Pass (LP); High Pass (HP); Band Pass (BP) and Band Reject (BR; band stop, notch)
- define basic properties of common approximation functions (Butterworth, Chebyshev, Bessel)
- analyze 2nd order filter properties – with an ideal amplifiers
- judge an influence of real amplifier properties on LP filters

EXPLANATION

Filters are classified according to the functions they are to perform, in terms of ranges of frequencies, as *pass bands* and *stop bands*.

A LOWPASS (LP) – filter characteristic is one in which the pass band extends from ω = 0 to ω = ω_0, where ω_0 is known as the ***cutoff frequency*** – Fig. 7.1

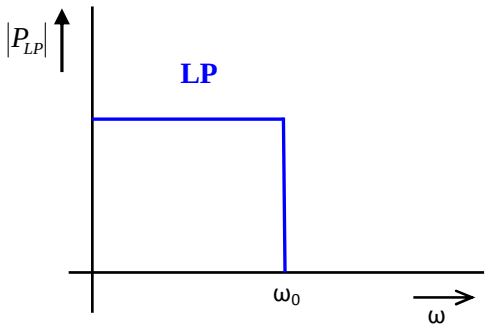


Fig. 7.1 Magnitude response $|P_{LP}|$ of an ideal lowpass-brick wall

A HIGH PASS (HP) filter is a complement to the lowpass filter in that the frequency range from 0 to ω_0 is a stop band, while from ω_0 to infinity is a pass band – Fig. 7.2.

A BAND PASS (BP) filter is one which frequencies extending from ω_1 to ω_2 are passed, while other frequencies are stopped – Fig. 7.3.

A BAND STOP (BS) filter is a complement of the bandpass filter where the frequencies from ω_1 to ω_2 are stopped and others are passed – Fig. 7.4. These filters are sometimes known as ***notch*** filters or as ***reject*** filters.

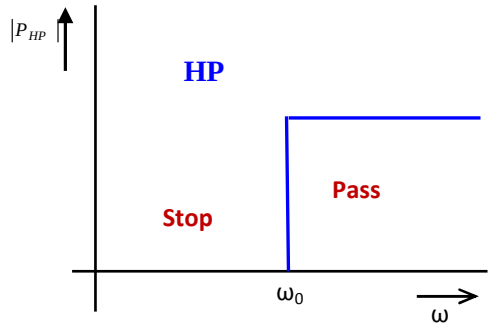


Fig. 7.2 Magnitude response $|P_{HP}|$ of an ideal lowpass-brick wall

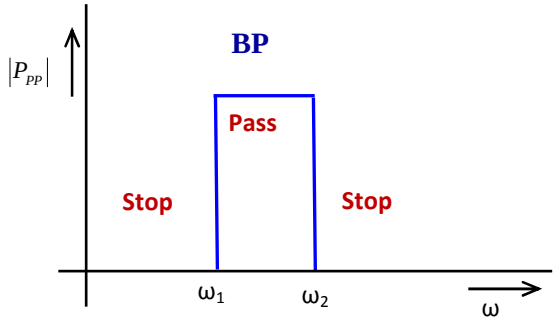


Fig. 7.3 Magnitude response $|P_{BP}|$ of an ideal lowpass-brick wall

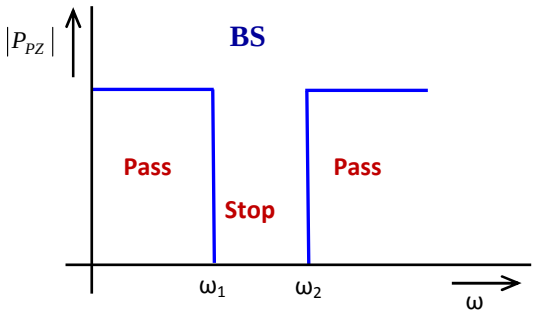


Fig. 7.4 Magnitude response $|P_{BS}|$ of an ideal lowpass-brick wall

It is not possible to realize the ideal characteristic above with a finite number of elements (it needs infinite number of elements). Realistic LP characteristic is shown in fig.7.5. The sharpness of the transition from a stop band to a pass band can be controlled to some extent in the design of the filters.

A PASS BAND – the magnitude of the transfer function $|P_{LP}|$ is always greater than a value designated A_1 [the ***attenuation*** $\alpha_1 = 20\log(A/A_1)$ is less than a value designated as α_{max}].

A STOP BAND – the magnitude of the transfer function $|P_{LP}|$ is always less than a value designated A_2 [the attenuation $\alpha_2 = 20\log(A/A_2)$ is greater than a value designated as α_{min}].

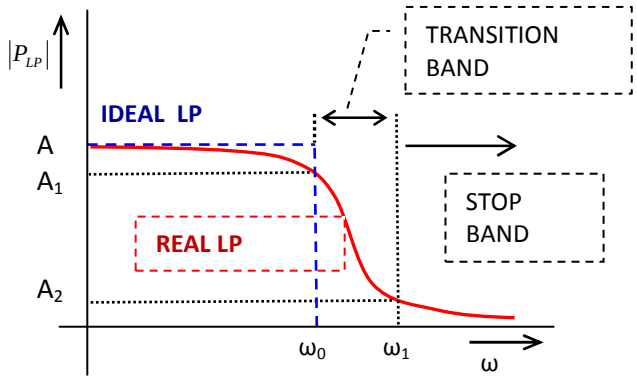


Fig. 7.5 Magnitude response $|P_{LP}|$ of a real lowpass

A TRANSITION BAND – a band of frequencies between the stop band and the pass band.

The same definitions are analogously valid for the other above described filter types.

7.1 Approximation LP normalized *functions – general – for a cascade realization*

Fig. 7.6, Tab. 7.1, 2, 3.

Even n (order $n = 2, 4, 6, \cdots$) – cascade realization by means $n/2$ second order filters.

$$P_{LP}(s) = \prod_{k=1}^{n/2} A_k \frac{b_k}{s^2 + a_k s + b_k}$$

Odd n (order $n = 1, 3, 5, \cdots$) – cascade realization by means one first order filter and $(n-1)/2$ second order filters.

$$P_{LP}(s) = \frac{A_0}{s + b_0} \cdot \prod_{k=1}^{(n-1)/2} A_k \frac{b_k}{s^2 + a_k s + b_k}$$

If we need denormalized LP, we substitute $s = p/\omega_0$.
If we need denormalized HP, we substitute $s = \omega_0/p$.

Here ω_0 is generally the "all filter frequency".

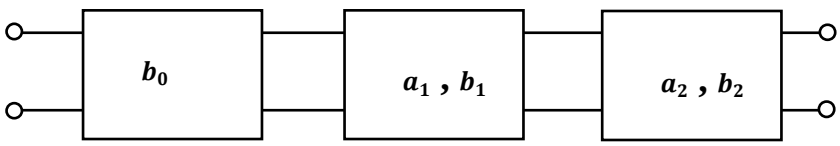


Fig. 7.6 An example of cascade filter realization – $n = 5$ (5th order filter)

Three commonly used filter types (approximation) are:

- Butterworth
- Chebyshev
- Bessel

7.2 Butterworth LP

Butterworth LP – (maximally flat magnitude); $b_k = 1$

$$a_k = 2 \cdot \sin\frac{(2k-1)\pi}{2n}; \quad Q_k = \frac{1}{a_k} \quad - \textit{ a quality factor}$$

This filter has the flattest possible pass-band magnitude response. Attenuation is – 3 dB at the design cutoff frequency, always. Attenuation above the cutoff frequency is moderately steep 20-dB per decade per pole (per "every one order") – fig.7.7. The pulse response of the Butterworth filter has moderate overshoot and ringing.

$$|P_{LP}| = \frac{1}{\sqrt{1+\left(\frac{\omega}{\omega_0}\right)^{2n}}}; \; n - \text{Given number of poles (filter order)}$$

$$|P_{LP}(\omega_0)| = \frac{1}{\sqrt{1+(1)^{2n}}} = \frac{1}{\sqrt{2}} \;\Rightarrow\; |P_{LP}(\omega_0)|_{dB} = 20\log(2)^{-0,5} = -3 \text{ dB} - \textbf{\textit{always}}$$

$$|P_{LP}(\omega > \omega_0)| = \frac{1}{\sqrt{\left(\frac{\omega}{\omega_0}\right)^{2n}}} = \left(\frac{\omega}{\omega_0}\right)^{-n} \;\Rightarrow\; |P_{LP}(\omega > \omega_0)|_{dB} = -n \cdot 20\log\frac{\omega}{\omega_0} \;\Rightarrow$$

slope is $-n \cdot 20$ dB/dec

Approximation LP function – for cascade realization

Even n (order $n = 2, 4, 6, \cdots$) – cascade realization by means $n/2$ second order filters.

$$P_{LP}(p) = P_{LP}\left(s = \frac{p}{\omega_0}\right) = \prod_{k=1}^{n/2} \frac{A_k}{\left(\frac{p}{\omega_0}\right)^2 + a_k\frac{p}{\omega_0} + 1}$$

$$P_{LP}(p) = \prod_{k=1}^{n/2} \frac{A_k\omega_0^2}{p^2 + a_k\omega_0 p + \omega_0^2} = \prod_{k=1}^{n/2} \frac{A_k\omega_0^2}{p^2 + p\frac{\omega_0}{Q_k} + \omega_0^2}$$

Odd n (order $n = 1, 3, 5, \cdots$) – cascade realization by means one first order filter and $(n-1)/2$ second order filters – analogously we get.

$$P_{LP}(p) = \frac{A_0\omega_0}{p + \omega_0} \cdot \prod_{k=1}^{(n-1)/2} \frac{A_k\omega_0^2}{p^2 + a_k\omega_0 p + \omega_0^2}$$

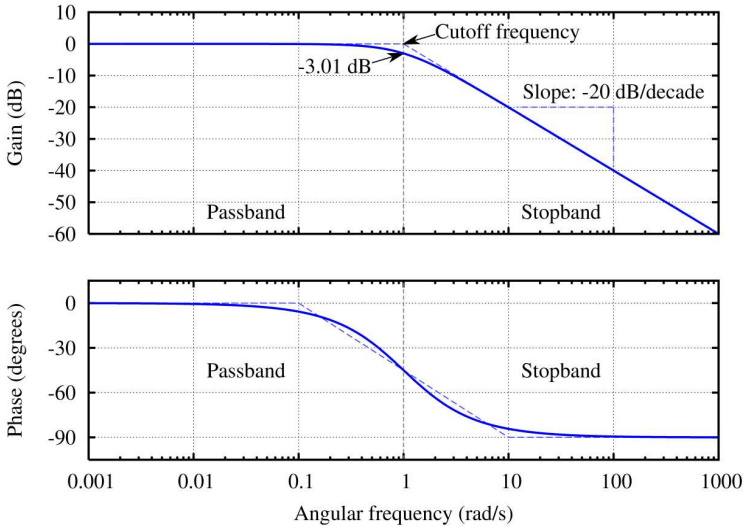


Fig. 7.7 The 1st order Butterworth lowpass

Example 7.1

Derive and if you need Butterworth LP, the 4th order.

☑ Solution:

It is even , thus to to , thus

:

:

Tab.1: Butterworth filter

2	-	1,414 214	1, 000 000	-	-
3	1, 000 000	1, 000 000	1, 000 000	-	-
4	-	0,765 367	1, 000 000	1,847 759	1, 000 000
5	1, 000000	0,618 034	1, 000 000	1,618 034	1, 000 000

7.2.1 Determining needed for the Butterworth lowpass

To determine , we start with equation - see Fig. 7.5. We require if . Attenuation on the frequency is (in dB). Further it means that . We substitute value of the and we get

⇒

$$10^{\alpha/10} = 1 + (\omega_1/\omega_0)^{2n}$$

$$10^{\alpha/10} - 1 = (\omega_1/\omega_0)^{2n}$$

$$\log\left(10^{\alpha/10} - 1\right) = 2n \cdot \log(\omega_1/\omega_0) \quad \Rightarrow$$

$$\boxed{n = \frac{\log\left(10^{\alpha/10} - 1\right)}{2 \cdot \log(\omega_1/\omega_0)}}$$

This is the needed Butterworth filter order (n) if we require the attenuation value α on the frequency ω_1 – Fig. 7.5.

7.3 Butterworth HP

Butterworth HP – (maximally flat magnitude) – Fig. 7.8

If we need Butterworth HP models we just substitute $s \to \omega_0/p$. We get

$$|P_{HP}| = \frac{1}{\sqrt{1+\left(\frac{\omega_0}{\omega}\right)^{2n}}}\ ;\ n \text{ – Given number of poles (filter order)}$$

$$|P_{HP}(\omega_0)| = \frac{1}{\sqrt{1+(1)^{2n}}} = \frac{1}{\sqrt{2}} \quad \Rightarrow \quad |P_{HP}(\omega_0)|_{dB} = 20\log(2)^{-0{,}5} = -3\ \text{dB} - \textbf{\textit{always}}$$

$$|P_{HP}(\omega_0 > \omega)| = \frac{1}{\sqrt{\left(\frac{\omega_0}{\omega}\right)^{2n}}} = \left(\frac{\omega}{\omega_0}\right)^{n} \quad \Rightarrow \quad |P_{HP}(\omega_0 > \omega)|_{dB} = n \cdot 20\log\frac{\omega}{\omega_0} \quad \Rightarrow$$

slope is $+n \cdot 20\ \text{dB/dec}$

$$P_{HP}(p) = P_{LP}\left(s = \frac{p}{\omega_0}\right) = \prod_{k=1}^{n/2} \frac{A_k}{\left(\frac{\omega_0}{p}\right)^2 + a_k\frac{\omega_0}{p} + 1} = \prod_{k=1}^{n/2} \frac{A_k p^2}{p^2 + a_k\omega_0 p + \omega_0^2}$$

and accordingly

$$P_{HP}(p) = \frac{A_0 p}{p+1} \cdot \prod_{k=1}^{n/2} \frac{A_k p^2}{p^2 + a_k\omega_0 p + \omega_0^2}$$

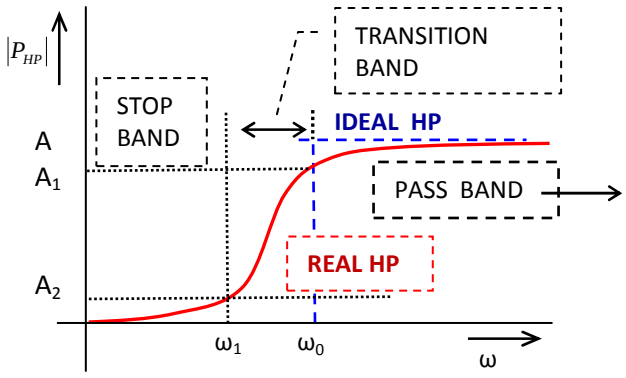


Fig. 7.8 Magnitude response $|P_{HP}|$ of a real highpass

7.4 Chebyshev LP

Chebyshev (equal ripple magnitude), Tab 2a, b, c, d

Chebyshev cutoff frequency is defined as the frequency at which the response falls below the ripple band. For a given number of poles (filter order), a steeper cutoff can be achieved by allowing more bandpass ripple – Fig. 7.9. The Chebyshev has even more ringing in its pulse response than the Butterworth.

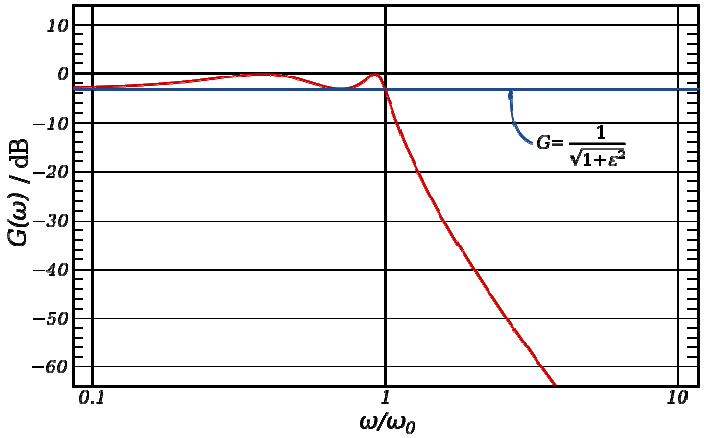


Fig. 7.9 An example of the Chebyshev magnitude response (

Tab.2a: Chebyshev filter – bandpass ripple

2	-	1,425 625	1,516 203	-	-
3	0,626 456	0,626 456	1,142 448	-	-
4	-	0,350 706	1,063 519	0,846 680	0,356 412
5	0,362 320	0,223 926	1,035 784	0,586 245	0,476 676

Tab.2b: Chebyshev filter – bandpass ripple

2	-	1,097 734	1,102 510	-	-
3	0,494 171	0,494 171	0,994 205	-	-
4	-	0,279 072	0,986 505	0,673 739	0,279 398
5	0,289 493	0,178 917	0,988 315	0,468 410	0,429 298

Tab.2c: Chebyshev filter – bandpass ripple

2	-	0,803 816	0,823 060	-	-
3	0,368 911	0,368 911	0,886 095	-	-
4	-	0,209 775	0,928 675	0,506 440	0,221 568
5	0,218 308	0,134 922	0,952 167	0,353 230	0,393 150

Example 7.2

Derive denormalized *Chebyshev* HP transfer functions, the 3th order, bandpass ripple 2 dB

☑ **Solution:**

We use Tab. 2c and substitution

Thus we must realize the cascade connection of the 1st order HP filter with a partial characteristic frequency and of the 2nd order HP filter with a partial characteristic frequency because the general model of the second order denominator is always - its quality factor we get from equation thus

It is valid that 3-dB frequency of LP Chebyshev filters is

where we know that bandpass ripple (in dB) is .

For HP Chebyshev is valid reciprocal formula .

Example 7.3

Determine the 3-dB frequency of the:

a) 3rd order LP Chebyshev filter with bandpass ripple dB.

b) 5th order LP Chebyshev filter with bandpass ripple dB.

c) 3rd order HP Chebyshev filter with bandpass ripple dB.

☑ **Solution:**

Ad a)

$$G = 10\log(1+\varepsilon^2) \;\Rightarrow\; \varepsilon = \sqrt{10^{G/10}-1} \;\Rightarrow\; \varepsilon = \sqrt{10^{0,1}-1} = 0{,}5088$$

$$\frac{\omega_3}{\omega_0} = \cosh\left(\frac{\operatorname{arg\,cosh}\left(\frac{1}{0{,}5088}\right)}{3}\right) = \cosh\left(\frac{\operatorname{arg\,cosh}(1{,}96541)}{3}\right) = \cosh\left(\frac{\ln 3{,}6574}{3}\right) = 1{,}095$$

Ad b)

$$G = 10\log(1+\varepsilon^2) \;\Rightarrow\; \varepsilon = \sqrt{10^{G/10}-1} \;\Rightarrow\; \varepsilon = \sqrt{10^{0,05}-1} = 0{,}3493$$

$$\frac{\omega_3}{\omega_0} = \cosh\left(\frac{\operatorname{arg\,cosh}\left(\frac{1}{0{,}3493}\right)}{3}\right) = \cosh\left(\frac{\operatorname{arg\,cosh}(2{,}8628)}{3}\right) = \cosh\left(\frac{\ln 5{,}5452}{3}\right) = 1{,}059$$

Ad c)

We do reciprocal calculus for the HP – we use result of a) solution:

$$\frac{\omega_3}{\omega_0} = 1{,}095 \;\Rightarrow\; \omega_3 = \frac{\omega_0}{1{,}095}$$

See Tab. 2d, too.

Tab. 2d 3-dB frequencies of some Chebyshev LP filters (HP is reciprocal)

Chebyshev LP filter order n	bandpass ripple G		
	0,5 dB	1 dB	2 dB
2	1,390	1,218	1,074
3	1,168	1,095	1,033
4	1,093	1,053	1,018
5	1,059	1,034	1,012
	ω_3/ω_0		

7.5 Bessel

Bessel (maximally flat time delay; also called Thomson), tab 3a, b

Due to its linear phase response, this filter has excellent pulse response (minimal overshoot and ringing) – its group delay $\tau = -d\varphi/d\omega$ is constant, ideally $\tau = 1/\omega_0$, actually see Fig. 7.10 – $\tau(\omega=\omega_0) = 0{,}9231/\omega_0$ for $n=2$; $\tau(\omega=\omega_0) = 0{,}9964/\omega_0$ for $n=3$; and $\tau(\omega=\omega_0) = 0{,}9999/\omega_0$ for $n=4$. For a given number of poles (filter order), its magnitude response is neither flat, nor its attenuation beyond the -3 – dB cutoff frequency as steep is as the Butterworth.

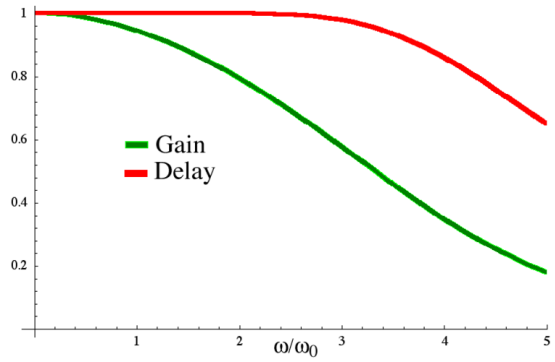


Fig. 7.10 An example of the Bessel magnitude response and group delay

Tab.3a Bessel filter

2	-	3,000 000	3,000 000	-	-
3	2,322 185	3,677 815	6,459 433	-	-
4	-	5,792 421	9,140 131	4,207 579	11,487 800
5	3,646 739	6,703 913	14,272 481	4,649 349	18,156 315

Denormalized Bessel transfer functions we get analogously as above.

Example 7.4

Derive denormalized *Bessel* LP transfer functions, the 4th order.

☑ **Solution:**

We use Tab. 3 and substitution

Thus we must realize the cascade connection of two 2nd orders LP filters. A generally model of the second order denominator is always thus:

The first second order filter:

;

The second second order filter:

$$\omega_{p_2} = \omega_0 \cdot \sqrt{b_2} = \omega_0 \cdot 3{,}3894; \quad \omega_0 a_2 = \frac{\omega_{p21}}{Q_{p_2}} \quad \Rightarrow \quad Q_{p_2} = \frac{\sqrt{b_2}}{a_2} = 0{,}8055$$

3-dB frequencies of LP Bessel filters are in Tab. 3b. For the HP Bessel filter we use reciprocal data.

Tab.3b 3-dB frequencies of Bessel filters

filter order n	2	3	4	5
ω_3/ω_0	1,36	1,75	2,11	2,42

7.6 Summary

Butterworth response

Advantages: It provides maximally flat magnitude response in the pass-band. It has good all-around performance. Its pulse response is better than Chebyshev. Its rate of attenuation is better than of Bessel.

Disadvantages: Some overshoot and ringing is exhibited in step response.

Chebyshev response

Advantages: It provides better attenuation beyond the pass-band than Butterworth.

Disadvantages: Ripple in pass-band may be objectionable. There is considerable ringing in step response.

Bessel response

Advantages: It provides best step response: very little overshoot or ringing.

Disadvantages: It exhibits slower rate of attenuation beyond the pass-band than Butterworth.

See Fig. 7.11, too.

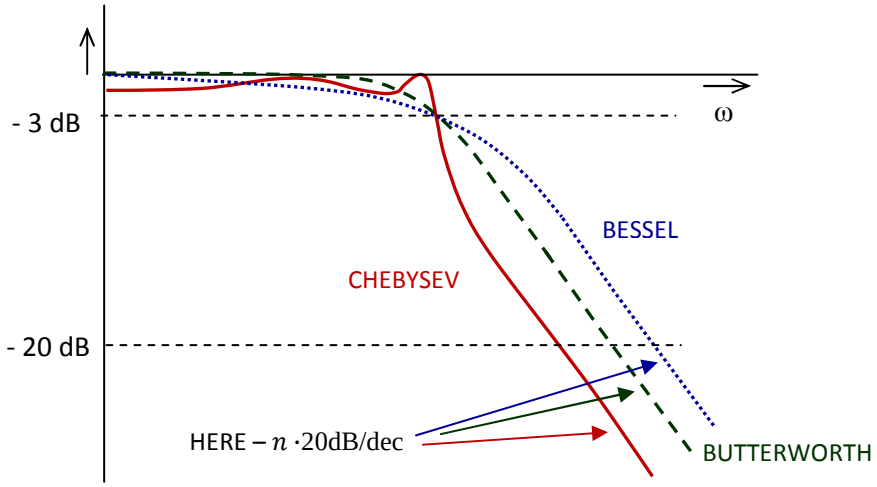


Fig. 7.11 Qualitative comparison of filter magnitudes for the same 3 dB frequency and the same filter order n

7.7 Network functions

Understanding of basic network functions (transfer) is very usefull.

7.7.1 Basic low-pass network functions

Single pole (1st order LP) – the single-pole low-pass transfer function in the complex frequency variables is

$$\hat{P}_{LP_1}(p = j\omega) = \frac{A_0\omega_p}{p+\omega_p}$$

ω_p – *is just the characteristic frequency of this 1st pole LP (not of the "all filter" if we use cascade realization).*

The **magnitude of the transfer** function for the response to sinusoidal steady-state excitation is

$$P_{LP_1}(\omega) = \sqrt{\frac{A_0^2\omega_p^2}{\omega^2+\omega_p^2}}$$

The **phase is**

$$\varphi(\omega) = \arg\frac{A_0\omega_p}{j\omega+\omega_p} = -\text{arctg}\left(\frac{\omega}{\omega_p}\right)$$

The **group delay** is generally (always) $\tau = -d\varphi(\omega)/d\omega$.

Complex Conjugate Pole Pair (2nd order LP) – the complex-conjugate-pole-pair LP transfer function and the sinusoidal steady-state magnitude function are

$$\hat{P}_{LP_2}(p = j\omega) = A_0\frac{\omega_p^2}{p^2+a\omega_p p+\omega_p^2} = A_0\frac{\omega_p^2}{p^2+p\frac{\omega_p}{Q}+\omega_p^2}$$

A_0; $a = \frac{1}{Q}$; ω_p – *just the properties of this 2nd pole LP (not of the "all filter" if we use cascade realization).*

$$P_{LP_2}(\omega) = \sqrt{\frac{A_0^2\omega_p^4}{\left(\omega_p^2-\omega^2\right)^2+\left(a\omega_p\omega\right)^2}} = \left|x = \frac{\omega}{\omega_p}\right| = \frac{A_0}{(1-x^2)^2+(ax)^2}$$

Differentiation with respect to *x* gives (it is enough to use denominator)

$$\frac{d}{dx}[(1-x^2)^2+(ax)^2] = 2(1-x^2)\cdot(-2x)+2a^2x = 2x\cdot(2x^2-2+a^2)$$

Now we can determine that magnitude maximum is if (function extreme)

$$2x_m^2-2+a^2 = 0$$

So we get

$$x_m^2 = 1-\frac{a^2}{2} \;\Rightarrow\; \left(\frac{\omega_m}{\omega_p}\right)^2 = 1-\frac{a^2}{2} \;\Rightarrow\; \omega_m = \omega_p\cdot\sqrt{1-\frac{a^2}{2}} = \omega_p\cdot\sqrt{1-\frac{1}{2Q^2}}$$

The ***magnitude has a peak*** at $\omega_m = \omega_p \cdot \sqrt{1 - a^2/2} = \omega_p \cdot \sqrt{1 - 1/2Q^2}$ for $\leq \sqrt{2}$, ***value of this peak*** is (Fig. 7.12)

$$P_{LP_2}(\omega_m) = P_{LP_{2max}}(x = x_m) = \frac{A_0}{\sqrt{(1-x_m^2)^2 + (ax_m)^2}} = \frac{A_0}{\sqrt{\left(1-\left(1-\frac{a^2}{2}\right)\right)^2 + (a)^2 \cdot \left(1-\frac{a^2}{2}\right)}} = \cdots$$

$$\cdots = \frac{A_0}{a \cdot \sqrt{1-\frac{a^2}{4}}} = \frac{A_0 \cdot Q}{\sqrt{1-\frac{1}{4Q^2}}}$$

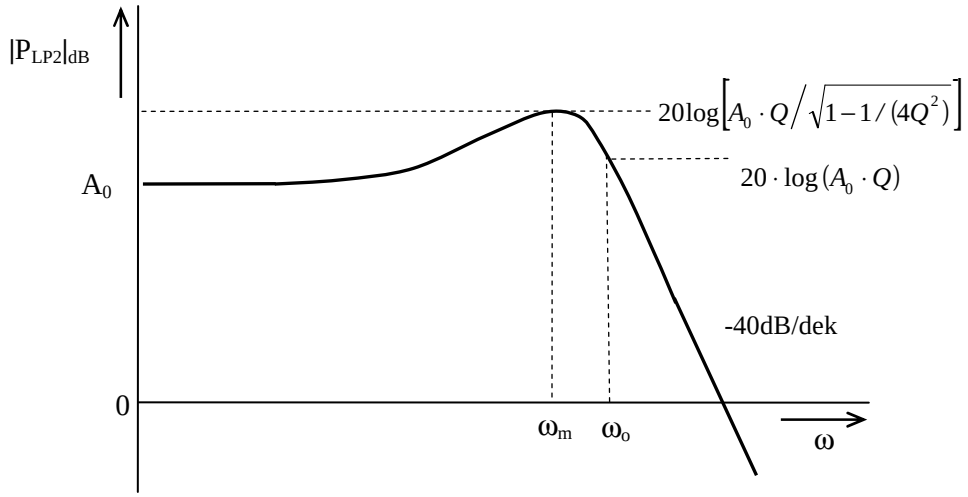


Fig. 7.12 Magnitude of the LP2 in dB: $P_{LP_2}\big|_{dB} = 20 \log P_{LP_2}$

7.7.2 Basic high-pass network functions

Single pole (1st order HP) – the single-pole high-pass transfer function in the complex frequency variables is

$$\hat{P}_{HP_1}(p = j\omega) = \frac{A_0 p}{p + \omega_p}$$

ω_p – *is just the characteristic frequency of this 1st pole HP (not of the "all filter" if we use cascade realization).*

The **magnitude of the transfer** function for the response to sinusoidal steady-state excitation is

$$P_{HP_1}(\omega) = \sqrt{\frac{A_0^2 \omega^2}{\omega^2 + \omega_p^2}}$$

The **phase is**

$$\varphi(\omega) = \arg \frac{A_0 j\omega}{j\omega + \omega_p} = \frac{\pi}{2} - \operatorname{arctg}\left(\frac{\omega}{\omega_p}\right)$$

The **group delay** is generally (always) $\tau = -d\varphi(\omega)/d\omega$.

Complex Conjugate Pole Pair (2nd order HP) – the complex-conjugate-pole-pair HP transfer function and the sinusoidal steady-state magnitude are

$$\hat{P}_{HP_2}(p=j\omega)=A_0\frac{p^2}{p^2+a\omega_p p+\omega_p^2}=A_0\frac{p^2}{p^2+p\frac{\omega_p}{Q}+\omega_p^2}$$

A_0; $a=\frac{1}{Q}$; ω_p – *just the properties of this 2nd pole HP (not of the "all filter" if we use cascade realization).*

$$P_{HP_2}(\omega)=\sqrt{\frac{A_0^2\omega^4}{\left(\omega_p^2-\omega^2\right)^2+\left(a\omega_p\omega\right)^2}}=\left|x=\frac{\omega_p}{\omega}\right|=\frac{A_0}{(1-x^2)^2+(ax)^2}$$

Now we easily get again that

$$x_m^2=1-\frac{a^2}{2}\ \Rightarrow\ \left(\frac{\omega_p}{\omega_m}\right)^2=1-\frac{a^2}{2}\ \Rightarrow\ \omega_m=\omega_p\cdot\sqrt{1-\frac{a^2}{2}}=\omega_p\cdot\sqrt{1-\frac{1}{2Q^2}}$$

The ***magnitude has a peak*** at $\omega_m=\omega_p\cdot\sqrt{1-\ a^2/2}=\omega_p\cdot\sqrt{1-1/2Q^2}$ for $a\le\sqrt{2}$, ***value of this peak*** is (Fig. 7.13)

$$P_{HP_2}(\omega_m)=P_{LP_{2max}}(x=x_m)=\frac{A_0}{\sqrt{\left(1-x_m^2\right)^2+(ax_m)^2}}=\frac{A_0}{\sqrt{\left(1-\left(1-\frac{a^2}{2}\right)\right)^2+(a)^2\cdot\left(1-\frac{a^2}{2}\right)}}=\cdots$$

$$\cdots=\frac{A_0}{a\cdot\sqrt{1-\frac{a^2}{4}}}=\frac{A_0\cdot Q}{\sqrt{1-\frac{1}{4Q^2}}}$$

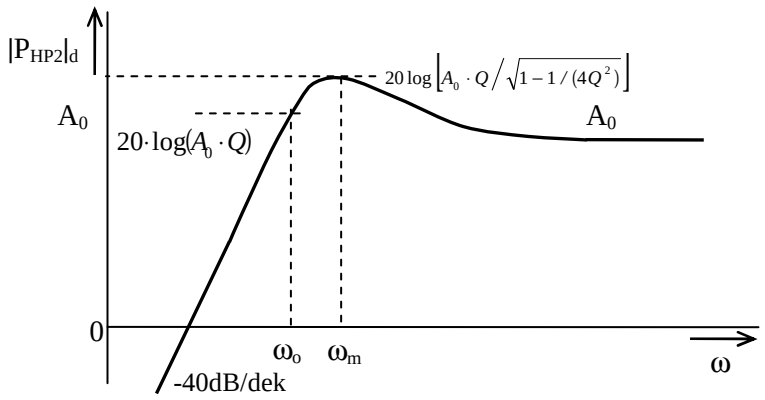


Fig. 7.13 Magnitude of the HP2 in dB: $P_{HP_2}\big|_{dB}=20\log P_{HP_2}$

Properties described above we use for tuning partial active filter stages to get needed properties of the "all cascade filter connection".

7.7.3 Band-pass network functions

Band-pass network function (2nd order), Fig. 7.14

The complex-conjugate-pole-pair BP transfer function and the sinusoidal steady-state magnitude are $p^2 + p\,\frac{\omega_p}{Q} + \omega_p^2$

$$\hat{P}_{BP}(p = j\omega) = \frac{A_0 \cdot p \cdot \frac{\omega_p}{Q}}{p^2 + p\,\frac{\omega_p}{Q} + \omega_p^2}$$

$$|P_{BP}| = \frac{\frac{A_0 \omega \omega_p}{Q}}{\sqrt{\left(\omega_p^2 - \omega^2\right)^2 + \left(\frac{\omega\omega_p}{Q}\right)^2}} = \frac{A_0}{\sqrt{\frac{\left(\omega_p^2 - \omega^2\right)^2 + \left(\frac{\omega\omega_p}{Q}\right)^2}{\left(\frac{\omega\omega_p}{Q}\right)^2}}} = \frac{A_0}{\sqrt{\left(\frac{\omega_p^2 - \omega^2}{\omega\omega_p}\right)^2 \cdot Q^2 + 1}}$$

$$= \left| x = \frac{\omega}{\omega_p} \right| = \frac{A_0}{\sqrt{\left(\frac{1}{x} - x\right)^2 Q^2 + 1}}$$

$$|P_{BP}(x = 1)| = \frac{A_0}{\sqrt{\left(\frac{1}{1} - 1\right)^2 Q^2 + 1}} = A_0$$

$$|P_{BP}(x \le 1)| = \frac{A_0}{\sqrt{\left(\frac{1}{x^2}\right)^2 Q^2}} = \frac{A_0 \cdot x}{Q} \quad \Rightarrow \quad \text{the first asymptote is + 20 dB/dec}$$

$$|P_{BP}(x \ge 1)| = \frac{A_0}{\sqrt{(-x^2)^2 Q^2}} = \frac{A_0}{x \cdot Q} \quad \Rightarrow \quad \text{the second asymptote is - 20 dB/dec}$$

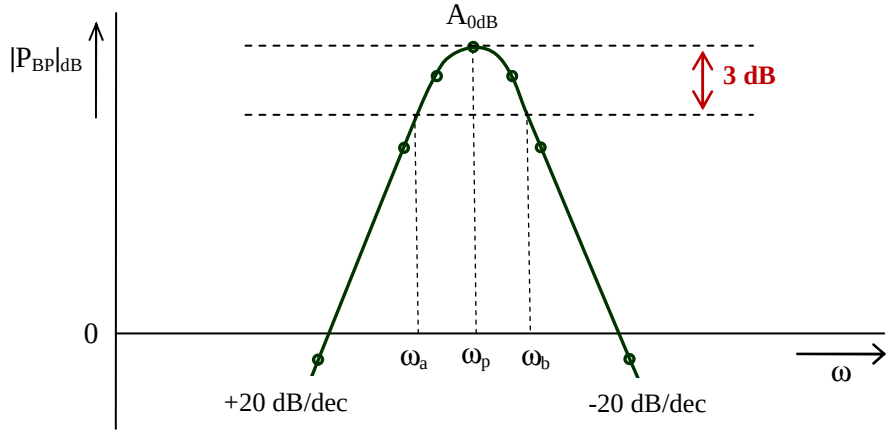


Fig. 7.14 Qualitative depiction of the BP magnitude

We next compute the 3-dB relative frequencies x_3 for the bandpass response. It means that it must be

$$|P_{BP}|_{3dB} = \frac{A_0}{\sqrt{\left(\frac{1}{x_3}-x_3\right)^2 Q^2+1}} = \frac{A_0}{\sqrt{2}} \;\Rightarrow\; \left(\frac{1}{x_3}-x_3\right)^2 Q^2 + 1 = 2$$

$$\left(\frac{1}{x_3}-x_3\right)^2 = \frac{1}{Q^2} \;\Rightarrow\; \frac{1}{x_3}-x_3 = \pm\frac{1}{Q} \;\Rightarrow\; x_3^2 \pm \frac{x_3}{Q} - 1 = 0 \;\Rightarrow$$

We get only two physically (positive) right roots (we neglect two negative roots):

$$x_{3_a} = \sqrt{1+\left(\frac{1}{2Q}\right)^2} - \frac{1}{2Q} \quad \text{and} \quad x_{3_b} = \sqrt{1+\left(\frac{1}{2Q}\right)^2} + \frac{1}{2Q}$$

Thus is valid:

$$\frac{\omega_a}{\omega_p} = \sqrt{1+\left(\frac{1}{2Q}\right)^2} - \frac{1}{2Q} \;\Rightarrow\; \omega_a = \omega_p\left(\sqrt{1+\left(\frac{1}{2Q}\right)^2} - \frac{1}{2Q}\right)$$

$$\frac{\omega_b}{\omega_p} = \sqrt{1+\left(\frac{1}{2Q}\right)^2} + \frac{1}{2Q} \;\Rightarrow\; \omega_b = \omega_p\left(\sqrt{1+\left(\frac{1}{2Q}\right)^2} + \frac{1}{2Q}\right)$$

These frequencies are identified in Fig. 7.14. We easily determine that

$$\omega_a \cdot \omega_b = \omega_p^2$$

and frequency difference of these frequencies defines the bandwidth (BW)

$$BW = \omega_b - \omega_a = \frac{\omega_p}{Q} \quad \Rightarrow \quad Q = \frac{\omega_p}{BW} = \frac{\omega_p}{\omega_b - \omega_a}$$

7.7.4 Band-stop network functions

Band-stop network function (2nd order), Fig. 7.15

The complex-conjugate-pole-pair BS transfer function and the sinusoidal steady-state magnitude are

$$\hat{P}_{BS}(p = j\omega) = \frac{p + \omega_o^2}{p^2 + \frac{p\omega_o}{Q} + \omega_o^2}$$

$$|P_{BS}| = \frac{A_0 \cdot |-\omega^2 + \omega_p^2|}{\sqrt{\left(\omega_p^2 - \omega^2\right)^2 + \left(\frac{\omega\omega_p}{Q}\right)^2}}$$

$$\left|P_{BS}(\omega \le \omega_p)\right| = A_0 \frac{-\omega^2 + \omega_p^2}{\sqrt{\left(\omega_p^2 - \omega^2\right)^2 + \left(\frac{\omega\omega_p}{Q}\right)^2}} = \left|x = \frac{\omega}{\omega_p}\right| = \frac{A_0 \cdot (1-x^2)}{\sqrt{(1-x^2)^2 + \left(\frac{x}{Q}\right)^2}}$$

$$|P_{BS}(x=1)| = A_0 \frac{1^2-1^2}{\sqrt{(1^2-1^2)^2+\left(\frac{1}{Q}\right)^2}} = 0$$

$$|P_{BS}(x=0)| = A_0 \frac{1^2-0^2}{\sqrt{(1^2-0^2)^2+\left(\frac{0}{Q}\right)^2}} = A_0$$

$$\left|P_{BS}(\omega \geq \omega_p)\right| = A_0 \frac{\omega^2-\omega_p^2}{\sqrt{\left(\omega_p^2-\omega^2\right)^2+\left(\frac{\omega\omega_p}{Q}\right)^2}} = \left|x=\frac{\omega}{\omega_p}\right| = \frac{A_0\cdot(x^2-1)}{\sqrt{(1-x^2)^2+\left(\frac{x}{Q}\right)^2}}$$

$$|P_{BS}(x\to\infty)| = A_0 \cdot \lim_{x\to\infty} \frac{(x^2-1)}{\sqrt{(1-x^2)^2+\left(\frac{x}{Q}\right)^2}} = A_0$$

We next compute the 3-dB relative frequencies x_3 for the bandstop response. It means that it must be

$$|P_{BS}|_{3dB} = \frac{A_0\cdot(1-x_3^2)}{\sqrt{\left(1-x_3^2\right)^2+\left(\frac{x_3}{Q}\right)^2}} = \frac{A_0}{\sqrt{2}} \Rightarrow \frac{1-x_3^2}{\sqrt{\left(1-x_3^2\right)^2+\left(\frac{x_3}{Q}\right)^2}} = \frac{1}{\sqrt{2}} \Rightarrow$$

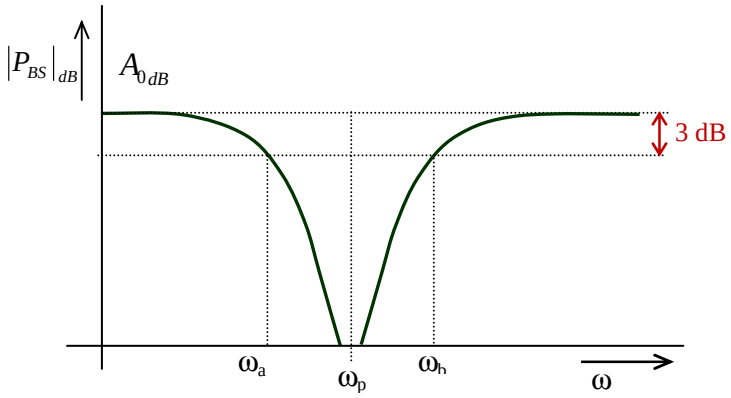


Fig. 7.15 Qualitative depiction of the BS magnitude

$$\frac{1}{\sqrt{\frac{\left(1-x_3^2\right)^2}{\left(1-x_3^2\right)^2}+\frac{(x_3/Q)^2}{\left(1-x_3^2\right)^2}}} = \frac{1}{\sqrt{2}} \Rightarrow \frac{(x_3/Q)^2}{\left(1-x_3^2\right)^2} = 1 \Rightarrow \frac{x_3/Q}{1-x_3^2} = \pm 1 \Rightarrow \pm x_3^2 + \frac{x_3}{Q} \pm 1 = 0$$

This equation gives the same physically right roots as it was for BP, so

$$\frac{\omega_a}{\omega_p} = \sqrt{1+\left(\frac{1}{2Q}\right)^2} - \frac{1}{2Q} \Rightarrow \omega_a = \omega_p\left(\sqrt{1+\left(\frac{1}{2Q}\right)^2} - \frac{1}{2Q}\right)$$

$$\frac{\omega_b}{\omega_p} = \sqrt{1+\left(\frac{1}{2Q}\right)^2} + \frac{1}{2Q} \Rightarrow \omega_b = \omega_p\left(\sqrt{1+\left(\frac{1}{2Q}\right)^2} + \frac{1}{2Q}\right)$$

These frequencies are identified in Fig. 7.15. We easy determine that

$$\omega_a \cdot \omega_b = \omega_p^2$$

and frequency difference of these frequencies defines the bandwidth (BW) of the bandstop filter now:

$$BW = \omega_b - \omega_a = \frac{\omega_p}{Q} \qquad \Rightarrow \quad Q = \frac{\omega_p}{BW} = \frac{\omega_p}{\omega_b - \omega_a}$$

7.8 Examples of second order filter realization

You can see a useful structure in Fig. 7.16.

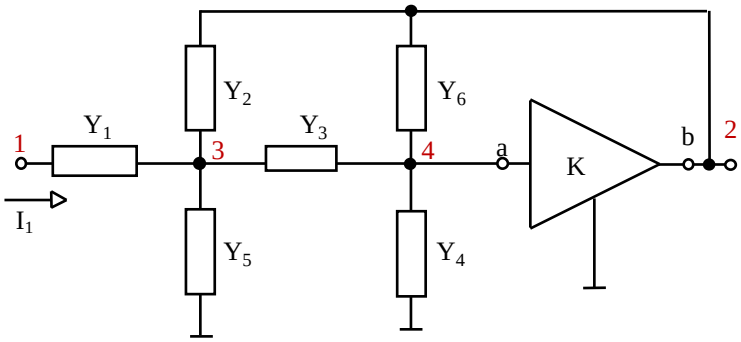


Fig. 7.16 The Bridgman-Brennar multiple-loop feedback biquad

We know the admittance model of the amplifier K – see Chapter 3:

	a	b				
a	0	0	U_a		I_a
b	$-KG_o$	G_o	U_a	=	I_b

So we can determine an admittance model of the Bridgman-Brennar multiple-loop feedback biquad:

	1	2(b)	3	4(a)				
1	Y_1	0	$-Y_1$	0		U_1		I_1
2(b)	0	$Y_2 + Y_6 + (\boldsymbol{G_o})$	$-Y_2$	$-Y_6 + (-\boldsymbol{KG_o})$		U_2		0
3	$-Y_1$	$-Y_2$	$Y_1 + Y_2 + Y_3 + Y_5$	$-Y_3$	$\cdot$	U_3	$=$	0
4(a)	0	$-Y_6$	$-Y_3$	$Y_3 + Y_4 + Y_6$		U_4		0

From this system of equations we derive, after some algebraic simplification, the ***basic transfer function*** (we suppose zero output resistance, thus $G_o \to \infty$):

$$\left.\frac{U_2}{U_1}\right|_{G_o \to \infty} = \frac{KY_1Y_3}{Y_4(Y_1 + Y_2 + Y_3 + Y_5) + Y_3(Y_1 + Y_5) + (1 - K) \cdot [Y_2Y_3 + Y_6(Y_1 + Y_2 + Y_3 + Y_5)]}$$

If we **choose** $K \to -\infty$ (an ideal inverting operational amplifier) we get

$$\left.\frac{U_2}{U_1}\right|_{\substack{G_o\to\infty \\ K\to-\infty}} = \frac{-Y_1Y_3}{[Y_2Y_3 + Y_6(Y_1+Y_2+Y_3+Y_5)]}$$

In this case the admittance Y_4 has no effect on the transfer function (zero voltage across it).

If we **choose** $K > 0$, we get filters with finite gain – ***Sallen and Key circuit***. If $Y_6 = Y_5 = 0$, we get

$$\left.\frac{U_2}{U_1}\right|_{G_o\to\infty} = \frac{KY_1Y_3}{Y_1Y_3 + Y_4(Y_1+Y_2+Y_3) + Y_2Y_3(1-K)}$$

If $Y_6 = 0$, we get only

$$\left.\frac{U_2}{U_1}\right|_{G_o\to\infty} = \frac{KY_1Y_3}{Y_4(Y_1+Y_2+Y_3+Y_5) + Y_3(Y_1+Y_5) + Y_2Y_3(1-K)}$$

An appropriate choice of admittances gives us desired transfer functions.

7.8.1 Sallen-Key LP (2nd order) filter

You can see the basic circuit in Fig. 7.17. It is evident that $Y_6 = Y_5 = 0$ and

$$Y_1 = G_1, Y_2 = pC_A, Y_3 = G_3, Y_4 = pC_B$$

An operational amplifier acts as a noninverting amplifier with a finite gain K.

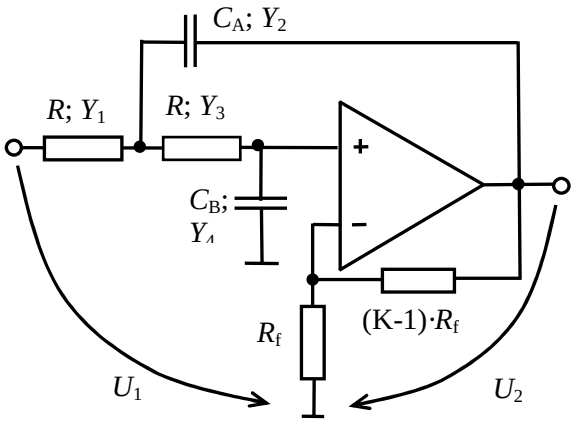


Fig. 7.17 One possible Sallen- Key LP configuration

Thus we can get (after some algebraic simplification)

$$\hat{P}_{LP_2}(p = j\omega) = A_0 \frac{\omega_p^2}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2}$$

where

$$\omega_p^2 = \frac{1}{R_1R_3C_AC_B}$$

$$2\xi = \frac{1}{Q} = \sqrt{\frac{R_3 C_B}{R_1 C_A}} + \sqrt{\frac{R_1 C_B}{R_3 C_A}} + \sqrt{\frac{R_1 C_A}{R_3 C_B}} \cdot (1 - K)$$

$$K = A_0$$

The first usual choice is: $R_1 = R_3 = R$ and $C_A = C_B = C$**.** In this case we get

$$\omega_p^2 = \frac{1}{R^2 C^2}$$

$$2\xi = \frac{1}{Q} = 3 - K$$

This filter is stable only if $K < 3$.

We can not define beforehand the filter gain, now. From these simple equations we can determine formulas suitable for this type ***filter design*** (we want ω_p and Q):

I. We chose the value of the C.
II. We determine $R = 1/(\omega_p C)$
III. We determine needed $K = 3 - 1/Q = 3 - 2\xi$. We chose suitable value of the R_f and determine corresponding value $(K - 1) \cdot R_f$

The second usual choice is: $R_1 = R_3 = R$ and $K = 1$; $(K - 1) \cdot R_f$ – short circuit; R_f – open.

In this case we get

$$\omega_p^2 = \frac{1}{R^2 C_A C_B}$$

$$2\xi = \frac{1}{Q} = 2 \cdot \sqrt{\frac{C_B}{C_A}}$$

If we chose *R*, we can determine that

$$C_A = \frac{2Q}{\omega_p R}; \qquad C_B = \frac{1}{2Q\omega_p R}$$

This filter is stable always – theoretically.

BOX

$$\omega_p^2 = \frac{1}{R^2 C_A C_B} \Rightarrow C_A = \frac{1}{R^2 \omega_p^2 C_B};$$

$$\frac{1}{Q} = 2 \cdot \sqrt{\frac{C_B}{C_A}} = 2 \cdot \sqrt{\frac{C_B}{\frac{1}{R^2 \omega_p^2 C_B}}} = 2 \cdot \sqrt{R^2 \omega_p^2 C_B^2} \Rightarrow C_B = \frac{1}{2QR\omega_p}$$

$$C_A = \frac{1}{R^2 \omega_p^2 C_B} = \frac{1}{R^2 \omega_p^2 \frac{1}{2QR\omega_p}} = \frac{2Q}{R\omega_p}$$

The *design procedure* is very simple

I. We chose the value of the R.
II. We determine $C_A = \frac{2Q}{R\omega_p}$ and $C_B = \frac{1}{2QR\omega_p}$

7.8.2 Iverting LP (2nd order) filter

If we choose $K \to -\infty$ and $Y_1 = G_1, Y_2 = G_2, Y_3 = G_3, Y_5 = pC_5, Y_6 = pC_6$ (Fig. 7.18) we get

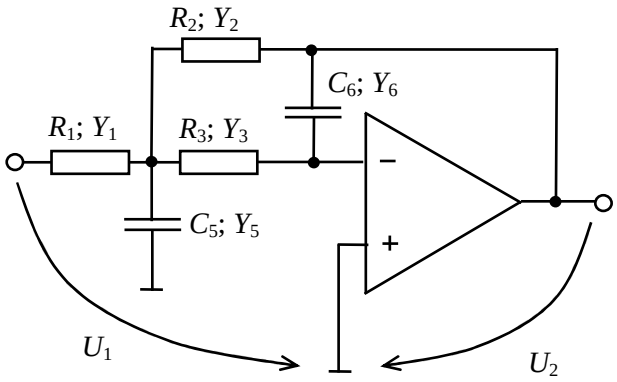


Fig. 7.18 One possible inverting LP configuration

$$P_{LP_2} = \frac{U_2}{U_1} = \frac{\frac{1}{R_1 R_3 C_5 C_6}}{p^2 + \frac{p}{C_5}\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right) + \frac{1}{R_2 R_3 C_5 C_6}} = \left|_{\text{rearranging}}\right| = \cdots$$

$$\cdots = -\frac{R_2}{R_1} \cdot \frac{\frac{1}{R_2 R_3 C_5 C_6}}{p^2 + \frac{p}{C_5}\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right) + \frac{1}{R_2 R_3 C_5 C_6}}$$

Now it is evident that – compare relation

$$\hat{P}_{LP_2}(p = j\omega) = A_0 \frac{\omega_p^2}{p^2 + p\,\frac{\omega_p}{Q} + \omega_p^2}$$

$$\omega_p^2 = \frac{1}{R_2 R_3 C_5 C_6}$$

$$2\xi\omega_p = \frac{\omega_p}{Q} = \frac{1}{C_5}\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right)$$

$$A_0 = -\frac{R_2}{R_1}$$

After rearranging we get

$$2\xi = \frac{1}{Q} = \sqrt{\frac{C_6}{C_5}} \cdot \left(\sqrt{\frac{R_3 R_2}{R_1^2}} + \sqrt{\frac{R_3}{R_2}} + \sqrt{\frac{R_2}{R_3}}\right)$$

From these equations we can determine formulas suitable for this type of ***filter design*** (we want A_o, ω_p and Q). We have just three equations. But we must determine five circuit elements. Thus we must choose two elements. Usually we choose $C_5 = C$ and $C_6 = mC$. Now we can determine that

$$\omega_p^2 = \frac{1}{R_2 R_3 m C^2}; \qquad \frac{1}{Q} = \sqrt{m}\cdot\left(\sqrt{\frac{R_3 R_2}{R_1^2}} + \sqrt{\frac{R_3}{R_2}} + \sqrt{\frac{R_2}{R_3}}\right); \qquad A_0 = -\frac{R_2}{R_1}$$

and from these equations we derive

$$R_1 = \frac{1}{2\cdot|A_o|\cdot QmC\omega_p}\cdot\left[1 \pm \sqrt{1 - 4m(|A_o|+1)Q^2}\right]$$

$$R_2 = |A_o|\cdot R_1$$

$$R_3 = \frac{1}{\omega_p^2 R_2 m C^2}$$

BOX

$$A_0 = -\frac{R_2}{R_1} \;\Rightarrow\; R_2 = |A_o|\cdot R_1; \qquad \omega_p^2 = \frac{1}{R_2 R_3 m C^2} \;\Rightarrow\; R_3 = \frac{1}{\omega_p^2 R_2 m C^2};$$

$$\frac{1}{Q} = \sqrt{m}\cdot\left(\sqrt{\frac{R_3 R_2}{R_1^2}} + \sqrt{\frac{R_3}{R_2}} + \sqrt{\frac{R_2}{R_3}}\right) = \sqrt{m}\cdot\left(\sqrt{\frac{\frac{1}{\omega_p^2 m C^2}}{R_1^2}} + \sqrt{\frac{1}{R_2\omega_p^2 m C^2}} + \sqrt{R_2\omega_p^2 m C^2}\right)$$

$$\frac{1}{Q} = \sqrt{m}\cdot\left(\sqrt{\frac{1}{\omega_p^2 m C^2 R_1^2}} + \sqrt{\frac{1}{|A_o|^2 R_1^2\omega_p^2 m C^2}} + \sqrt{|A_o|^2 R_1^2\omega_p^2 m C^2}\right) = \left|\text{ rearranging gives}\right|$$

$$R_1^2 - \frac{R_1}{|A_o| m\omega_p C Q} + \frac{|A_o|+1}{|A_o|^2\omega_p m C^2} = 0 \qquad \text{Now we easily determine that}$$

$$R_{1_{a,b}} = \frac{\frac{1}{|A_o| m\omega_p C Q} \pm \sqrt{\left(\frac{1}{|A_o| m\omega_p C Q}\right)^2 - \frac{4\cdot(|A_o|+1)}{|A_o|^2 m\omega_p C Q}}}{2} \;\Rightarrow\; R_{1_{a,b}} = \frac{1}{2|A_o| m\omega_p C Q}\left[1 \pm \sqrt{1 - 4m(|A_o|+1)Q^2}\right]$$

Physically right solution is $R_1 > 0$ (positive value of the R_1) – so it must be

$$1 - 4m(|A_o|+1)Q^2 > 0$$

We often choose just extreme case $1 - 4m(|A_o|+1)Q^2 = 0$, thus

$$m = \frac{1}{4\cdot(|A_o|+1)Q^2}$$

The *design procedure* is now:

I. We chose an appropriate value of the $C_6 = mC = mC_5$

II. We determine $C = C_5 = C_6/m = C_6[4(|A_o|+1)Q^2]$

III. We determine

$$R_1 = \frac{1}{2\cdot|A_o|\cdot QmC\omega_p} = \frac{1}{2\cdot|A_o|\cdot Q\omega_p C_6}$$

$$R_2 = |A_o|\cdot R_1 = \frac{1}{2\cdot Q\omega_p C_6}$$

$$R_3 = \frac{1}{\omega_p^2 R_2 m C^2} = \frac{1}{\omega_p^2 (R_2 m C) C} = \left|{}_{R_2 mC = R_2 C_6 = \frac{1}{2Q\omega_p}}\right| = \frac{1}{2Q(|A_o|+1)\omega_p C_6}$$

7.8.3 Iverting BP (2nd order) filter

If we choose $K \to -\infty$ and $Y_1 = G_1, Y_2 = Y_3 = pC, , Y_5 = 0, Y_6 = G_6$ (Fig. 7.19) we get

$$\hat{P}_{BP_2}(p = j\omega) = \frac{A_0 \cdot p\frac{\omega_p}{Q}}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2}$$

where

$$\omega_p^2 = \frac{1}{R_1 R_6 C^2}; \quad A_0 = -\frac{R_6}{2R_1}; \quad 2\xi = \frac{1}{Q} = 2 \cdot \sqrt{\frac{R_1}{R_6}}$$

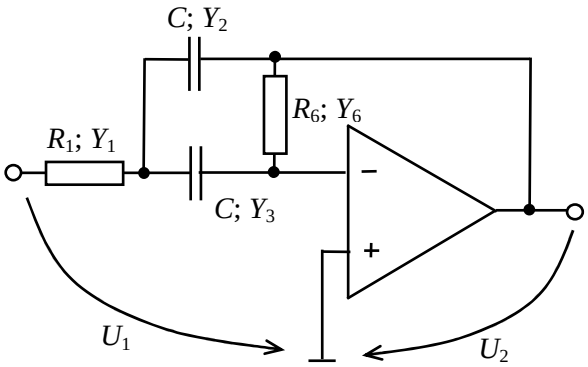


Fig. 7.19 One possible inverting BP configuration

The *design procedure* is now:

I. We chose an appropriate value of the R_1

II. We determine $R_6 = 4Q^2 R_1$

III. We determine $C = \frac{1}{2QR_1\omega_p}$;

we can not change the filter $A_o = -2Q^2$ in this simple circuit (Fig. 7.19)

BOX

$$\left.\frac{U_2}{U_1}\right|_{\substack{G_o \to \infty \\ K \to -\infty}} = \frac{-Y_1 Y_3}{Y_2 Y_3 + Y_6(Y_1 + Y_2 + Y_3)} = \frac{-G_1 pC}{p^2 C^2 + G_6(G_1 + pC + pC)} = \frac{-G_1 pC}{p^2 C^2 + 2pCG_6 + G_6 G_1} = -\frac{G_1 C}{C^2} \cdot \frac{p}{p^2 + p\frac{2G_6}{C} + \frac{G_6 G_1}{C^2}} = \cdots$$

$$\cdots = -\frac{G_1 C}{C^2} \cdot \frac{p\frac{2G_6}{C}\frac{C}{2G_6}}{p^2 + p\frac{2G_6}{C} + \frac{G_6 G_1}{C^2}} = -\frac{G_1 C}{C^2} \cdot \frac{C}{2G_6} \cdot \frac{p\frac{2G_6}{C}}{p^2 + p\frac{2G_6}{C} + \frac{G_6 G_1}{C^2}} = -\frac{R_6}{2R_1} \cdot \frac{p\frac{2}{CR_6}}{p^2 + p\frac{2}{CR_6} + \frac{1}{R_1 R_6 C^2}} \quad \Rightarrow$$

$$A_0 = -\frac{R_6}{2R_1}; \quad \omega_p^2 = \frac{1}{R_1 R_6 C^2}$$

$$\frac{\omega_p}{Q} = \frac{2}{CR_6} \quad \Rightarrow \quad Q = \frac{CR_6\omega_p}{2} = \frac{1}{2} \cdot \sqrt{\frac{R_6}{R_1}} \quad \Rightarrow \quad R_6 = 4Q^2 R_1 \quad \Rightarrow \quad A_o = -2Q^2 \quad \Rightarrow$$

we can not choose this parameter in this simple circuit

$$\omega_p^2 = \frac{1}{R_1 R_6 C^2} \quad \Rightarrow \quad C^2 = \frac{1}{R_1 R_6 \omega_p^2} \Rightarrow \quad C = \frac{1}{R_1 R_6 \omega_p}$$

If we need to define the A_o too, we must choose more complex circuit in Fig. 7.20. The other resistor R_5 allows us to do it.

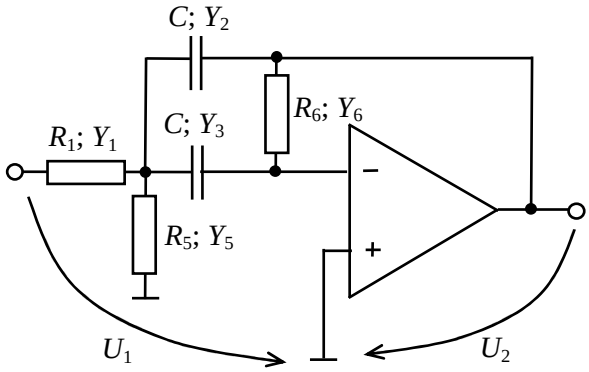


Fig. 7.20 More complex inverting BP configuration

In this case we get ($Y_5 = G_5$):

$$\omega_p^2 = \frac{1}{(R_1 \| R_5) R_6 C^2};$$

$$A_0 = -\frac{R_6}{2R_1};$$

$$2\xi = \frac{1}{Q} = 2 \cdot \sqrt{\frac{R_1 \| R_5}{R_6}}$$

This set of equations allows us determine formulas needed to *design procedure*, if we choose an appropriate value of C:

The *design procedure* is now:

I. We chose an appropriate value of the C (Fig. 7.20)

II. We determine

$$R_1 = \frac{Q}{\omega_p C \cdot |A_0|}; \quad R_5 = \frac{Q}{(2Q^2 - |A_0|)\omega_p C}; \quad R_6 = \frac{2Q}{\omega_p C}$$

It is evident that physically right solution is $R_5 > 0$ so that we may choose $2Q^2 > |A_o|$ only if we have $C_3 = C_2 = C$.

BOX

$$\frac{1}{Q} = 2 \cdot \sqrt{\frac{R_1 \| R_5}{R_6}} \Rightarrow R_1 \| R_5 = \frac{R_1 \cdot R_5}{R_1 + R_5} = \frac{R_6}{4Q^2} \Rightarrow$$

$$\omega_p^2 = \frac{1}{(R_1 \| R_5) R_6 C^2} = \frac{1}{\left(\frac{R_6}{4Q^2}\right) R_6 C^2} \Rightarrow R_6 = \frac{2Q}{\omega_p C}$$

$$|A_0| = \frac{R_6}{2R_1} \Rightarrow R_1 = \frac{R_6}{2 \cdot |A_0|} \Rightarrow R_1 = \frac{Q}{\omega_p C \cdot |A_0|}$$

$$R_1 \| R_5 = \frac{R_1 \cdot R_5}{R_1 + R_5} = \frac{R_6}{4Q^2} \Rightarrow 4Q^2 R_1 \cdot R_5 = R_6(R_1 + R_5) \Rightarrow 4Q^2 R_1 \cdot R_5 - R_6 R_5 = R_6 R_1 \Rightarrow$$

$$R_5 = \frac{R_6 R_1}{4Q^2 R_1 - R_6} = \frac{R_6}{4Q^2 - \frac{R_6}{R_1}} = \frac{\frac{2Q}{\omega_p C}}{4Q^2 - 2 \cdot |A_0|} \Rightarrow R_5 = \frac{Q}{(2Q^2 - |A_0|)\omega_p C}$$

7.8.4 Iverting HP (2nd order) filter

If we choose $K \to -\infty$ and $Y_1 = pC_1$, $Y_2 = pC_2, Y_3 = pC_3$, , $Y_5 = G_5$, $Y_6 = G_6$ (Fig. 7.21) we get

$$\hat{P}_{HP_2}(p = j\omega) = \frac{A_0 \cdot p^2}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2}$$

where

$$\omega_p^2 = \frac{1}{R_5 R_6 C_2 C_3}$$

$$\frac{\omega_p}{Q} = \frac{\left(\frac{C_1}{C_2 C_3} + \frac{1}{C_2} + \frac{1}{C_3}\right)}{R_6}$$

$$A_0 = -\frac{C_1}{C_2}$$

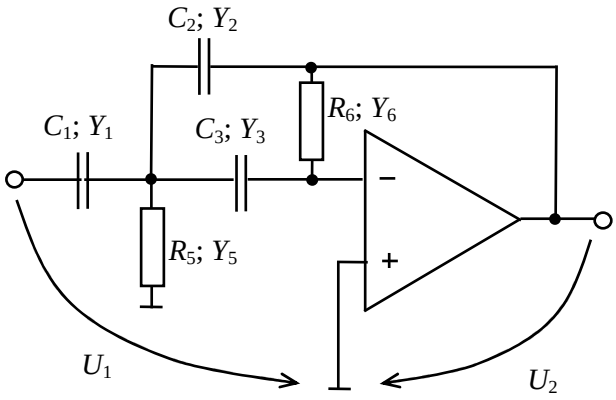


Fig. 21 Inverting HP circuit

Usually we choose $C_3 = C_1 = C$ and we can get from above equations next useful "design formulas":

$$C_2 = \frac{C}{|A_0|}$$

$$R_5 = \frac{|A_0|}{(2|A_0|+1)QC\omega_p}$$

$$R_6 = \frac{(2|A_0|+1)Q}{C\omega_p}$$

7.8.5 Sallen-Key HP (2nd order) filter

You can see the basic circuit in Fig. 7.22. It is evident that $Y_6 = Y_5 = 0$ and $Y_1 = pC_1$, $Y_2 = G_2, Y_3 = pC_3, Y_4 = G_4$.

An operational amplifier acts as noninverting amplifier with a finite gain K.

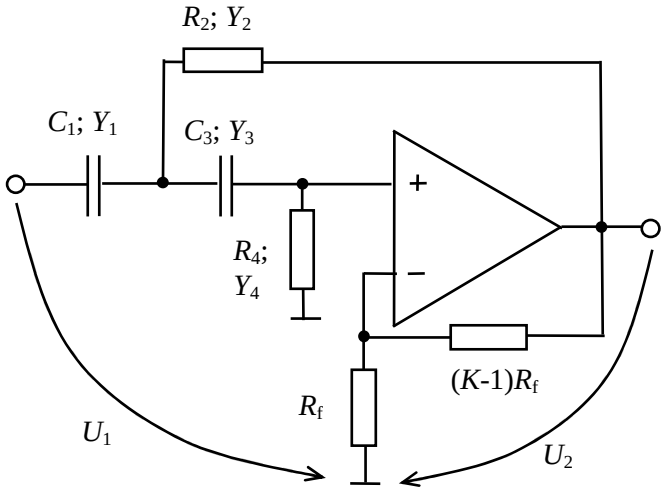


Fig. 7.22 One possible Sallen- Key HP

If $C_1 = C_3 = C$, $R_2 = R_4 = R$ than

$$\hat{P}_{HP_2}(p = j\omega) = \frac{A_0 \cdot p^2}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2}$$

where

$$\omega_p^2 = \frac{1}{(RC)^2}; \quad A_0 = K; \quad 2\xi = \frac{1}{Q} = 3 - K$$

We can not define beforehand the filter gain, now. From these simple equations we can easy determine formulas suitable for this type *filter design* (we want ω_p and Q):

I. We chose the value of the C.
II. We determine $R = 1/(\omega_p C)$
III. We determine needed $K = 3 - 1/Q = 3 - 2\xi$. We chose suitable value of the R_f and determine corresponding value $(K - 1) \cdot R_f$

This filter is stable only if $K < 3$.

The second usual choice is: $C_1 = C_3 = C$, and $K = 1$; $(K - 1) \cdot R_f$ – short circuit; R_f – open.

Then is valid

$$\omega_p^2 = \frac{1}{R_2 R_4 C^2}; \quad A_0 = 1; \quad 2\xi = \frac{1}{Q} = 2 \cdot \sqrt{\frac{R_2}{R_4}}$$

From these simple equations we can determine formulas suitable for this typeof *filter design* (we want ω_p and Q):

I. We chose the value of the C.
II. We determine $R_2 = \frac{1}{2Q\omega_p C}$; $R_4 = \frac{2Q}{\omega_p C}$

7.8.6 Notch filter (2nd order; bandstop)

The very simple realization of the notch filter is in Fig. 7.23. The upper operational amplifier creates LP (2nd order) filter. The bottom operational creates HP (2nd order) filter. These filters must have the same properties – it means ω_p, Q and K (thus R_d). The third operational amplifier creates the inverting adder amplifier.

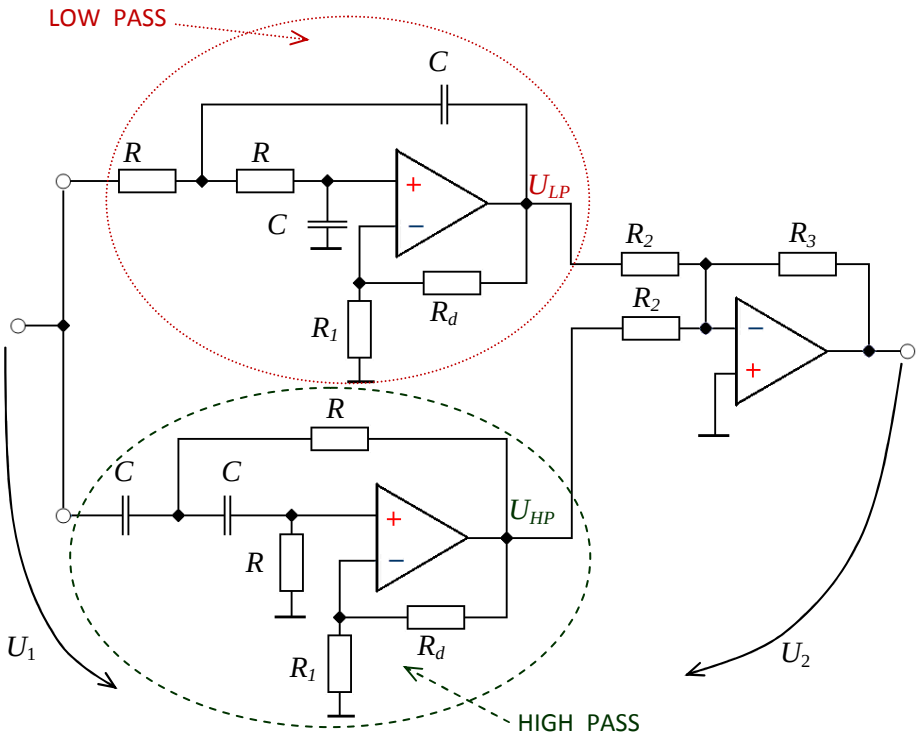


Fig. 7.23 One possible notch filter configuration – by means of Sallen-Key LP and HP filters – no cascade connection

It is evident that

$$U_{LP} = U_1 \cdot P_{LP_2} = U_1 \cdot A_0 \frac{\omega_p^2}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2}$$

$$U_{HP} = U_1 \cdot P_{HP_2} = U_1 \cdot A_0 \frac{p^2}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2}$$

where

$$\omega_p^2 = \frac{1}{(RC)^2}; \quad A_0 = K = 1 + \frac{R_d}{R_1}; \quad 2\xi = \frac{1}{Q} = 3 - K$$

$$U_2 = -\frac{R_3}{R_2}(U_{LP} + U_{HP}) = -\frac{R_3}{R_2} \cdot U_1 \cdot K \cdot \left(\frac{\omega_p^2}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2} + \frac{p^2}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2} \right) \Rightarrow$$

$$\frac{U_2}{U_1} = -\frac{R_3}{R_2} \cdot K \cdot \frac{p^2 + \omega_p^2}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2}$$

This is the notch filter – really.

7.9 Examples of first order filter realization

You can see basic first order circuits in Fig.7.24, 25, 26 and 27.

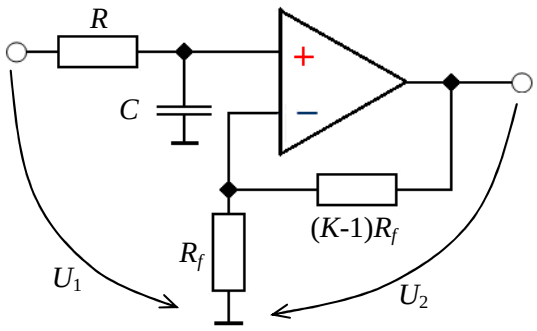


Fig. 7.24 Noninverting 1st order LP; an operational amplifier creates noninverting structure with gain K; a characteristic frequency is $\omega_p = 1/(RC)$; $A_o = K$

The transfer function of the circuit in Fig.7.24 is

$$\frac{U_2}{U_1} = K \cdot \frac{\frac{1}{pC}}{R + \frac{1}{pC}} = K \cdot \frac{\frac{1}{RC}}{p + \frac{1}{RC}} = K \cdot \frac{\omega_p}{p + \omega_p}$$

$$\omega_p = \frac{1}{RC}$$

The transfer function of the circuit in Fig.7.25 is

$$\frac{U_2}{U_1} = -\frac{\frac{R_2 \cdot \frac{1}{pC}}{R_2 + \frac{1}{pC}}}{R_1} = -\frac{R_2}{R_1} \cdot \frac{\frac{1}{R_2 C}}{p + \frac{1}{R_2 C}} = K \cdot \frac{\omega_p}{p + \omega_p}$$

$$\omega_p = \frac{1}{R_2 C};\ K = -\frac{R_2}{R_1};$$

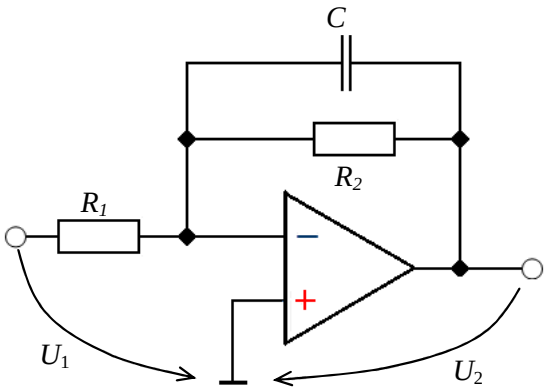


Fig. 7.25 Inverting 1st order LP; an operational amplifier creates inverting structure with DC gain $= -R_2/R_1$; a characteristic frequency is $\omega_p = 1/(RC)$; $A_o = K$

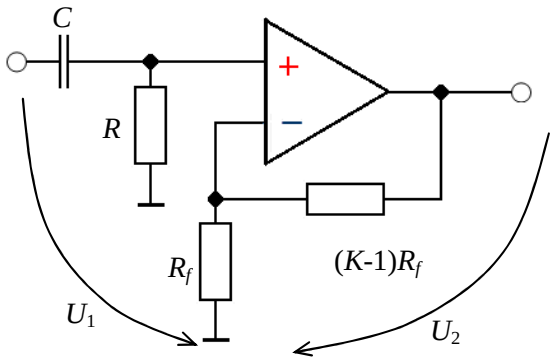


Fig. 7.26 Noninverting 1st order HP; an operational amplifier creates noninverting structure with gain K; a characteristic frequency is $\omega_p = 1/(RC)$; $A_o = K$

The transfer function of the circuit in Fig.7.26 is

$$\frac{U_2}{U_1} = K \cdot \frac{R}{R + \frac{1}{pC}} = K \cdot \frac{p}{p + \frac{1}{RC}} = K \cdot \frac{p}{p + \omega_p}$$

$$\omega_p = \frac{1}{RC}$$

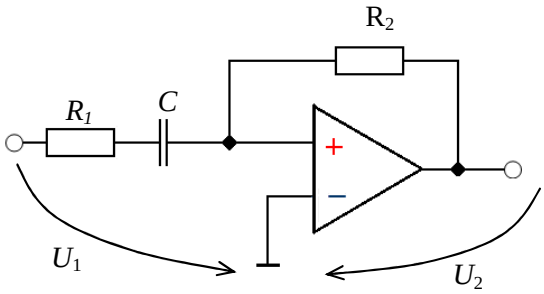


Fig. 7.27 Inverting 1st order HP; an operational amplifier creates inverting structure with AC gain $= -R_2/R_1$; a characteristic frequency is $\omega_p = 1/(RC)$; $A_o = K$

The transfer function of the circuit in Fig.7.27 is

$$\frac{U_2}{U_1} = -\frac{R_2}{R_1 + \frac{1}{pC}} = -\frac{R_2}{R_1} \cdot \frac{\frac{1}{R_1 C}}{p + \frac{1}{R_1 C}} = K \cdot \frac{\omega_p}{p + \omega_p}$$

$$\omega_p = \frac{1}{R_1 C}$$

$$K = -\frac{R_2}{R_1}$$

7.10 Basic cascade – connection

Cascade connections give us opportunity to create other transfer functions.

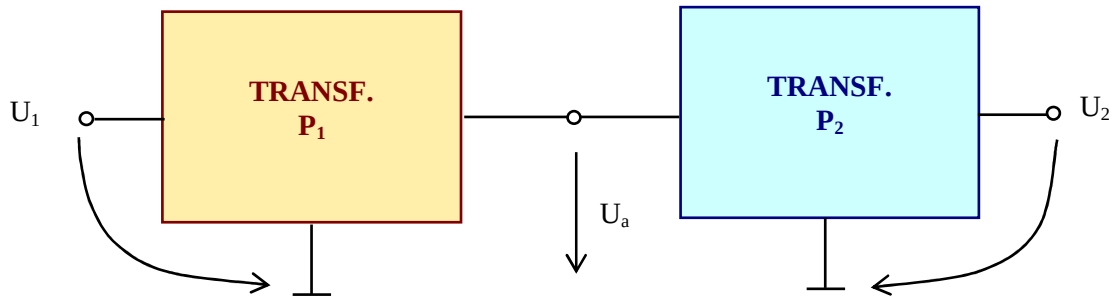


Fig. 7.28 Cascade connection of two filters with transfer functions
$P_1 = U_a/U_1$ **and** $P_2 = U_2/U_a$

It is evident that

$$P = \frac{U_2}{U_1} = \frac{U_a}{U_1} \cdot \frac{U_2}{U_a} = P_1 \cdot P_2$$

Further

$$|P| = |P_1 \cdot P_2| = |P_1| \cdot |P_2|$$

Thus in dB is valid

$$|P|_{dB} = 20\log(|P_1| \cdot |P_2|) = 20\log|P_1| + 20\log|P_2| = |P_1|_{dB} + |P_2|_{dB}$$

The last formula

$$\boxed{|P|_{dB} = |P_1|_{dB} + |P_2|_{dB}}$$

is very important. From this formula we can deduce that:

If we know the "decibel description" of the transfer function modulus $P_1(dB)$ and $P_2(dB)$, the resulting modulus of cascade connection we get as a sum $|P_1|_{dB} + |P_2|_{dB}$.

7.10.1 Two LP2 cascade

Two LP2s cascade is shown in Fig. 7.29

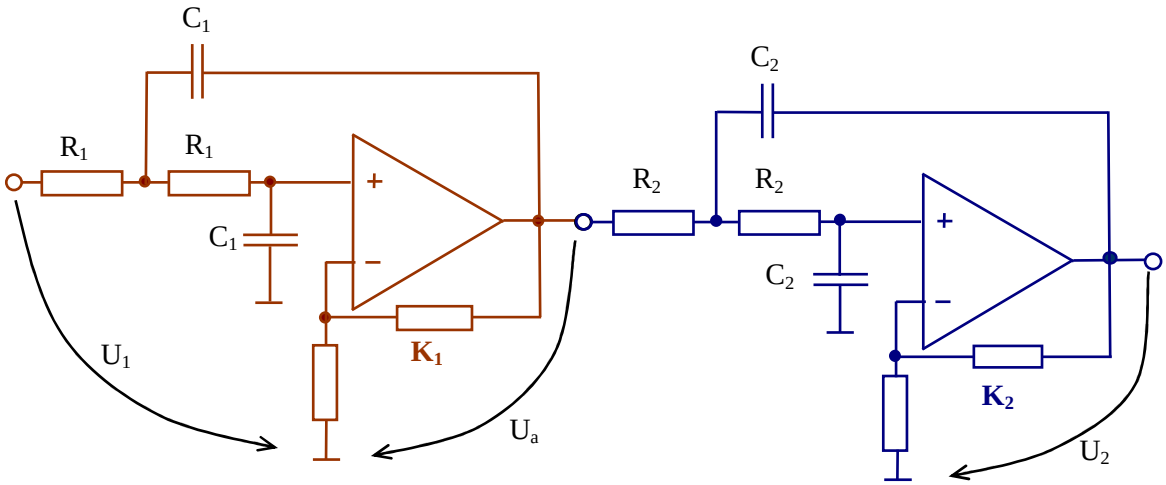


Fig. 7.29 Cascade connection of two LP2; LP2$_1$ ($\omega_{p_1} = 1/(R_1C_1)$; $Q_1 = 1/(3 - K_1)$) and LP2$_2$ ($\omega_{p_2} = 1/(R_2C_2)$; $Q_2 = 1/(3 - K_2)$)

If we choose suitable

$\omega_{p_1} = 1/(R_1C_1)$; $Q_1 = 1/(3 - K_1)$ and $\omega_{p_2} = 1/(R_2C_2)$; $Q_2 = 1/(3 - K_2)$

we can get different frequency responses. You can see a qualitative description in Fig. 7.30; for $\omega_{p_2} > \omega_{p_1}$ and $Q_2 > Q_1$.

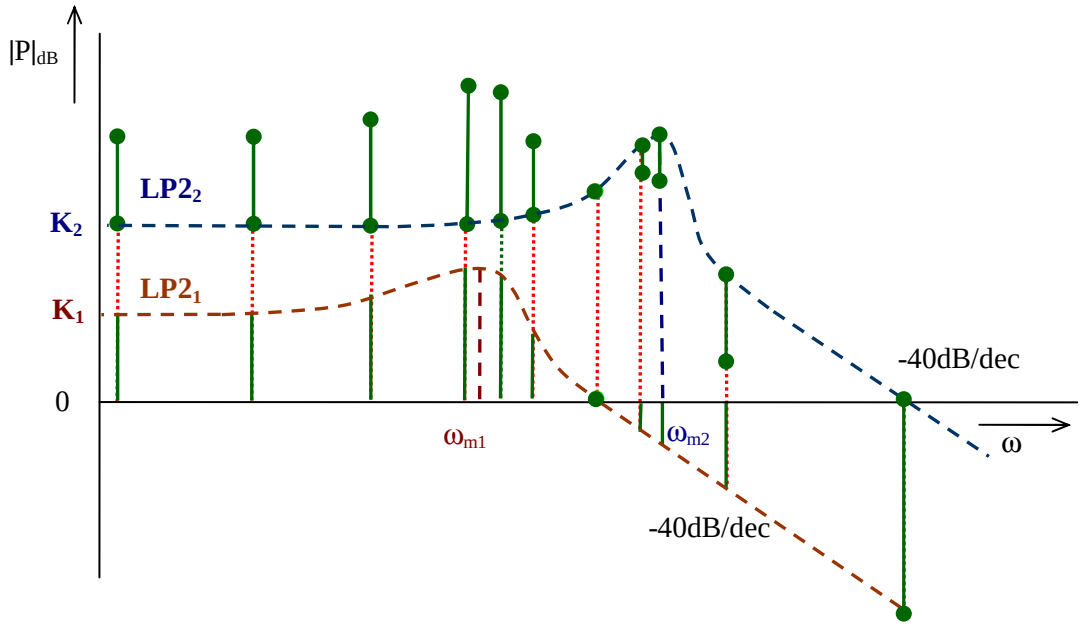


Fig. 7.30 A sum of modulus in dB

We can see the final modulus in fig. 31.

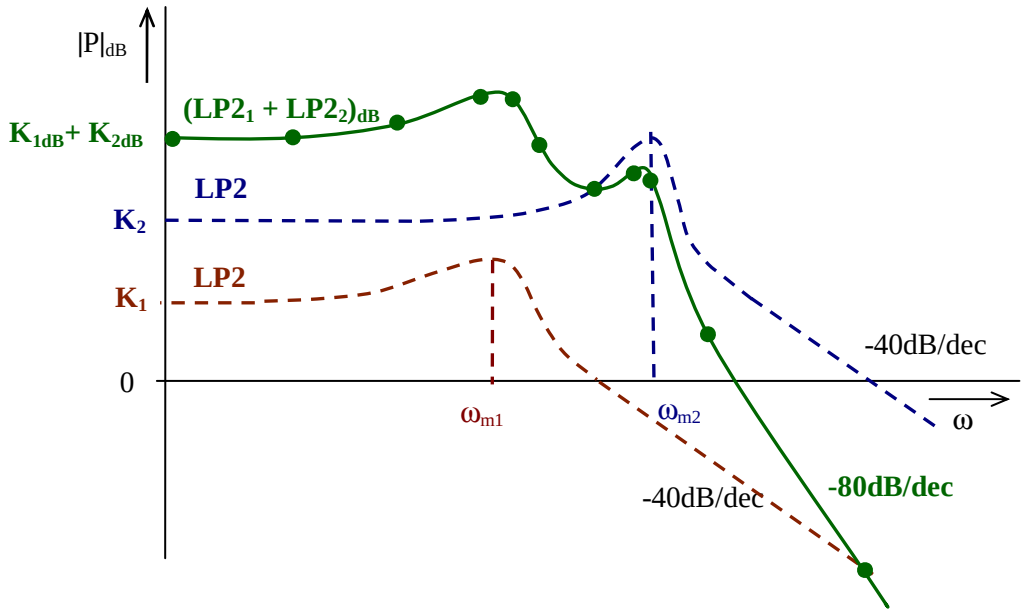


Fig. 7.31 A final sum of modulus in dB $|P|_{dB} = |P_{LP2_1}|_{dB} + |P_{LP2_2}|_{dB}$

If we increase ω_{o_1} and decrease quality factor Q_1 (K_1), we can get more friendly response – gray line depicts previous situation – Fig. 7.32.

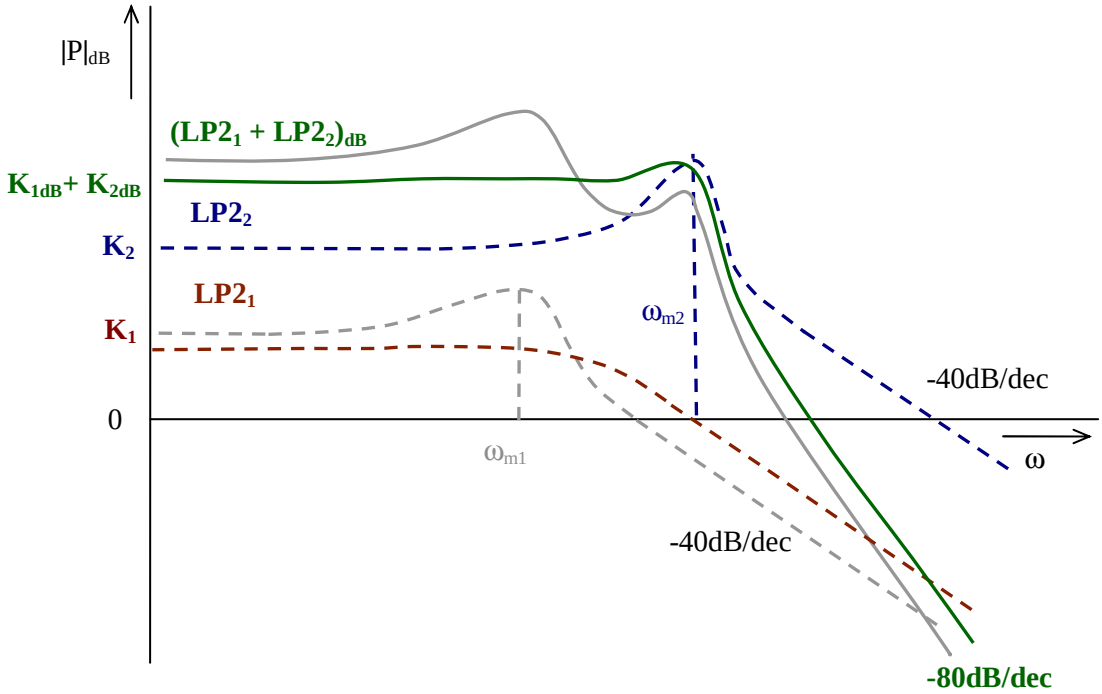


Fig. 7.32 A sum of modulus after increasing ω_{o1} and decreasing Q_1 (K_1),

If we decrease quality factor Q_2 (K_2), now, we can make another improvement of the modulus – gray line depicts previous situation – Fig. 7.33.

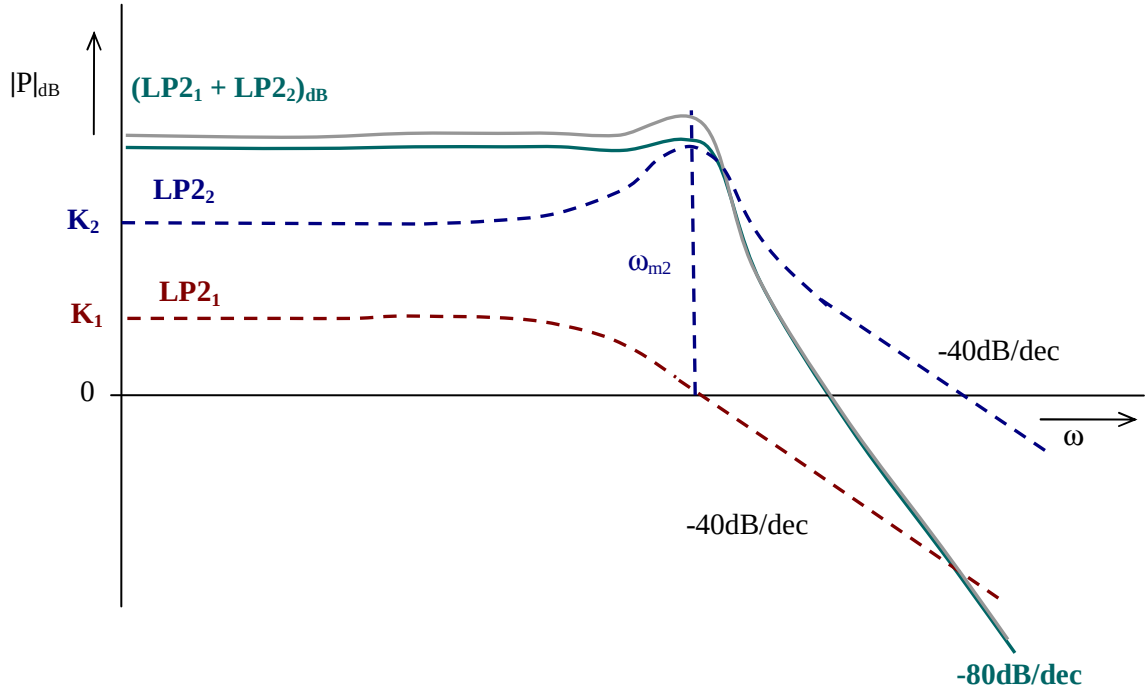


Fig. 7.33 A sum of modulus after decreasing Q_2 (K_2)

All the history of the filter adjustment you can see in Fig. 7.34.

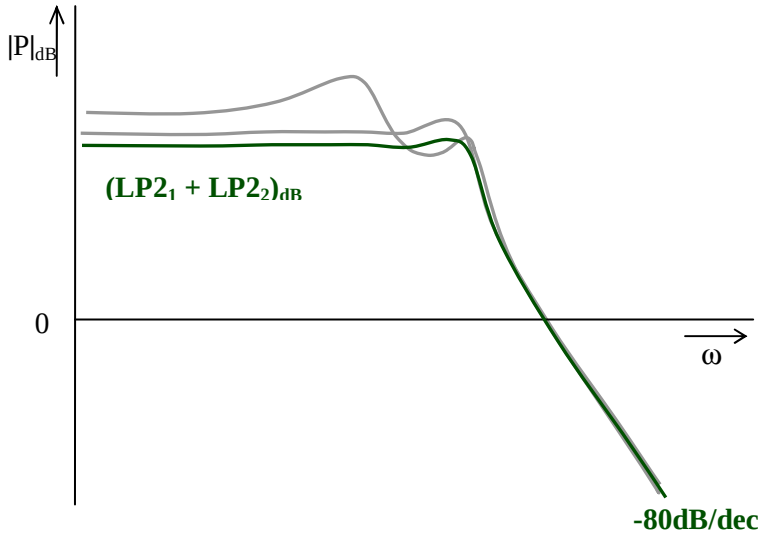


Fig. 7.33 A qualitative depiction of the filter adjustment history:
$$|\boldsymbol{P}|_{\boldsymbol{dB}} = \left|\boldsymbol{P}_{\boldsymbol{LP2_1}}\right|_{\boldsymbol{dB}} + \left|\boldsymbol{P}_{\boldsymbol{LP2_2}}\right|_{\boldsymbol{dB}}$$

We can see that a final filter slope is – 80 dB/dec – this means we have LP filter of 4th order. It was created intuitively. We can get needed $\omega_{p_1}, \omega_{p_2}$ and Q_1, Q_2 from the graphs in figures.

If we take these parameters from known approximation functions, we get filter properties in accordance with used approximation.

7.10.2 Two HP2 cascade

Two HPs cascade is depicted in Fig. 7.34.

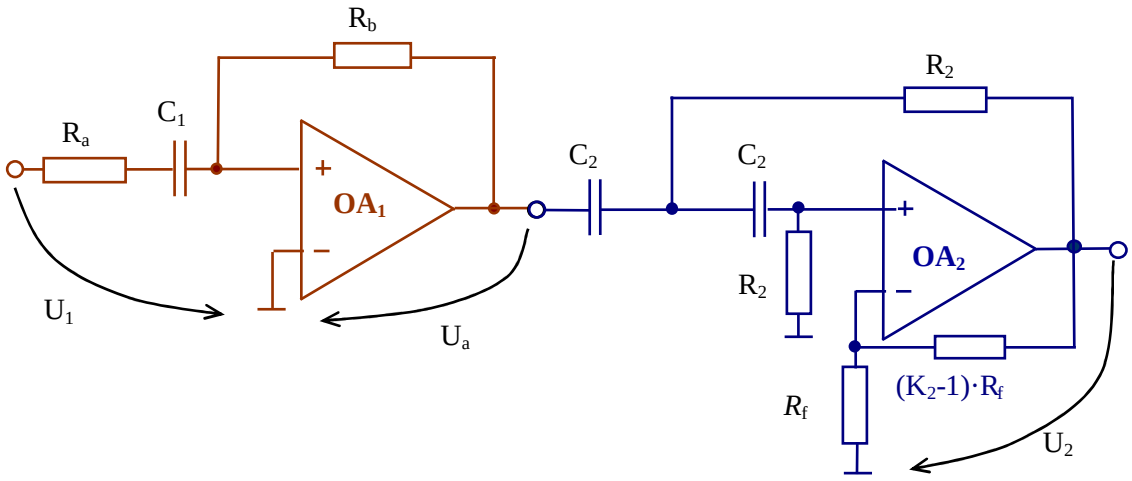


Fig. 7.34 Cascade connection of two HPs; HP1 (OA_1) – 1st order inverting high pass; HP2 (OA_2) – 2nd order noninverting highpass

It is evident that

$$P_1 = P_{HP_1} = \frac{U_a}{U_1} = -\frac{R_b}{R_a + \frac{1}{pC_1}} = -\frac{R_b}{R_a} \cdot \frac{p}{p + \frac{1}{R_a C_1}}$$

$\Rightarrow$

$$K_1 = -\frac{R_b}{R_a}\ ; \quad \omega_{p_1} = \frac{1}{R_a C_1}$$

$$P_2 = P_{HP_2};\ K_2;\ \omega_{p_1} = \frac{1}{R_a C_1}$$

A qualitative depiction of modulus summing is in Fig. 7.35. If we decrease quality factor Q_2 (it means K_2), we improve the final frequency response – Fig. 7.36. The filter slope is +60 dB/dec for low frequencies – it is 3rd order HP.

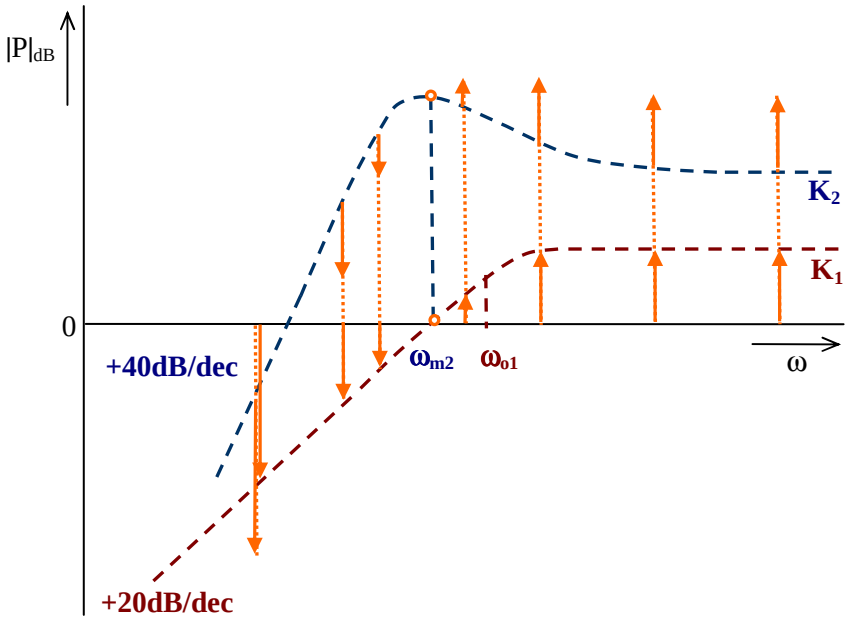


Fig. 7.35 A qualitative depiction of HP modulus summing

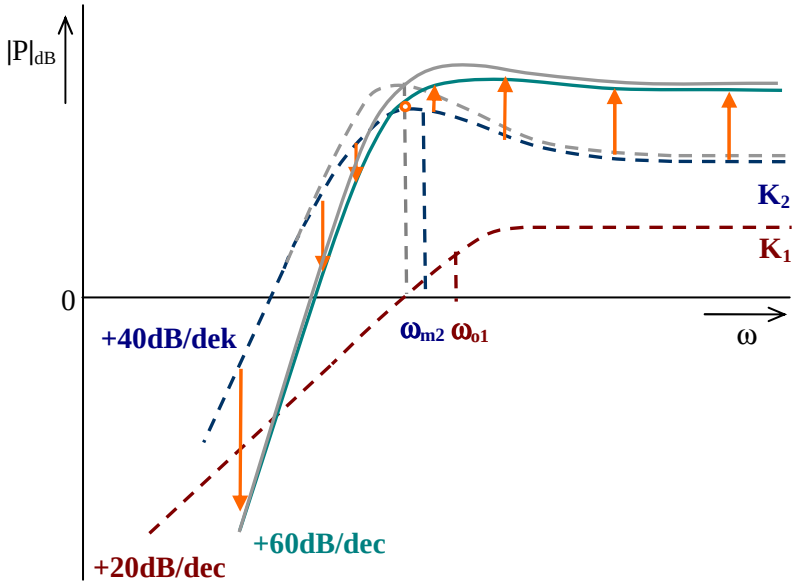


Fig. 7.36 A qualitative depiction of HP modulus summing $|P|_{dB} = |P_{HP2_1}|_{dB} + |P_{HP2_2}|_{dB}$ - after adjustment of HP2

7.10.3 Cascade connection of LP2 and HP2

Cascade connection of LP2 and HP2 is shown in Fig. 7.34

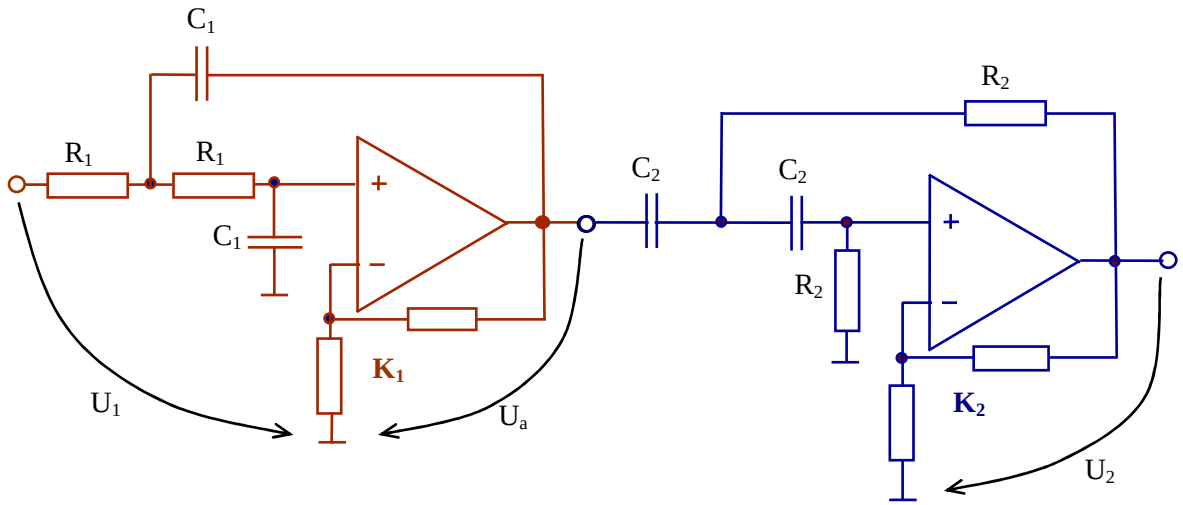


Fig. 7.37 Cascade connection of LP2 ($\omega_{p_1} = 1/(R_1C_1)$; $Q_1 = 1/(3 - K_1)$) and HP2 ($\omega_{p_2} = 1/(R_2C_2)$; $Q_2 = 1/(3 - K_2)$)

Let us choose $K_1 = K_2 = K$ (thus $Q_1 = Q_2 = Q$). A situation for $\omega_{p_1} = \omega_{p_2} = \omega_p$ is depicted in Fig. 7.38.

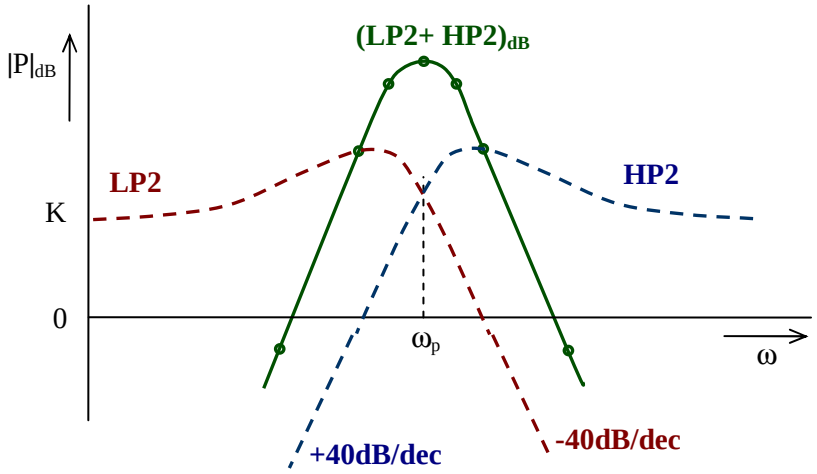


Fig. 7.38 A qualitative depiction of modulus summing: summing $|P|_{dB} = |P_{LP2}|_{dB} + |P_{HP2}|_{dB}$ if $\omega_{p_1} = \omega_{p_2} = \omega_p$; $K_1 = K_2 = K$ ($Q_1 = Q_2 = Q$)

This frequency response describes 4th order bandpass:

$$P = P_1 \cdot P_2 = P_{LP_2} \cdot P_{HP_2} = \frac{K\omega_p^2}{p^2 + p\,\frac{\omega_p}{Q} + \omega_p^2} \cdot \frac{Kp^2}{p^2 + p\,\frac{\omega_p}{Q} + \omega_p^2}$$

$\Rightarrow$

$$P = K^2Q^2 \cdot \left(\frac{p\frac{\omega_p}{Q}}{p^2 + p\frac{\omega_p}{Q} + \omega_p^2} \right)^2$$

$$P(\omega = \omega_p) = K^2Q^2 \cdot \left(\frac{j\omega\frac{\omega_p}{Q}}{-\omega_p^2 + j\omega_p\frac{\omega_p}{Q} + \omega_p^2} \right)^2 = K^2Q^2$$

$$P(\omega \ll \omega_p) \to K^2Q^2 \cdot \left(\frac{j\omega\frac{\omega_p}{Q}}{\omega_p^2} \right)^2 = -K^2\left(\frac{\omega}{\omega_p}\right)^2 \quad - \textit{ asymptote with slope +40 dB/dec}$$

$$P(\omega \gg \omega_p) \to K^2Q^2 \cdot \left(\frac{j\omega\frac{\omega_p}{Q}}{(j\omega)^2} \right)^2 = -K^2\left(\frac{\omega_p}{\omega}\right)^2 \quad - \textit{ asymptote with slope -40 dB/dec}$$

Let us choose $K_1 = K_2 = K$ (thus $Q_1 = Q_2 = Q$) and $\omega_{p_1} > \omega_{p_2}$ - we get wideband bandpass – Fig. 7.39.

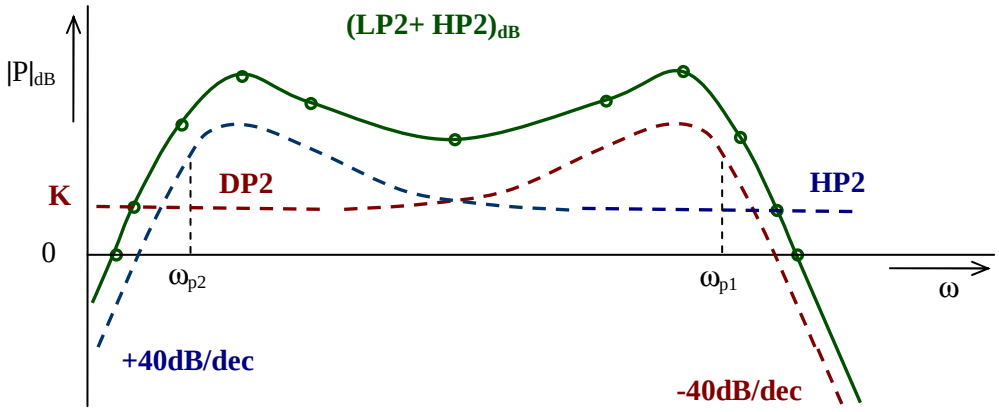


Fig. 7.39 A qualitative depiction of modulus summing: if $|P|_{dB} = |P_{LP2}|_{dB} + |P_{HP2}|_{dB}$ $\omega_{p_1} > \omega_{po_2}$ and $K_1 = K_2 = K$ ($Q_1 = Q_2 = Q$)

If we adjust quality factors we change frequency response, again.

7.10.4 Cascade connection of LP2 and BS2

Cascade connection of LP2 and BS2 (notch) is shown in Fig. 7.40

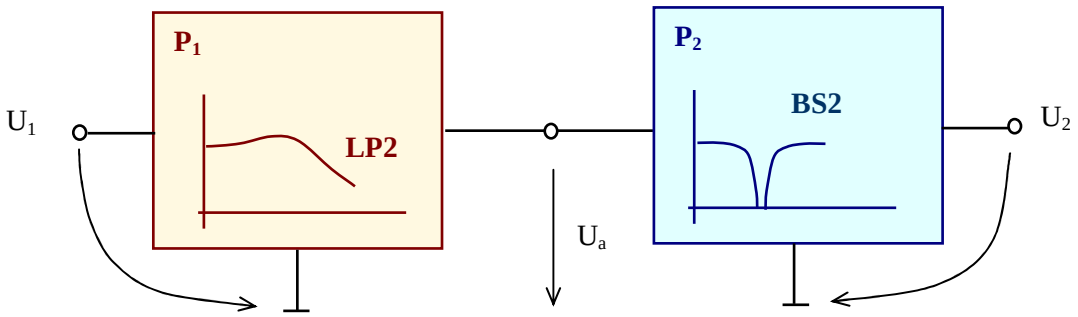


Fig. 7.40 Cascade connection of LP2 and bandstop filter

That way we get a **low-pass filter with „zero frequency”** - Fig. 7.41.

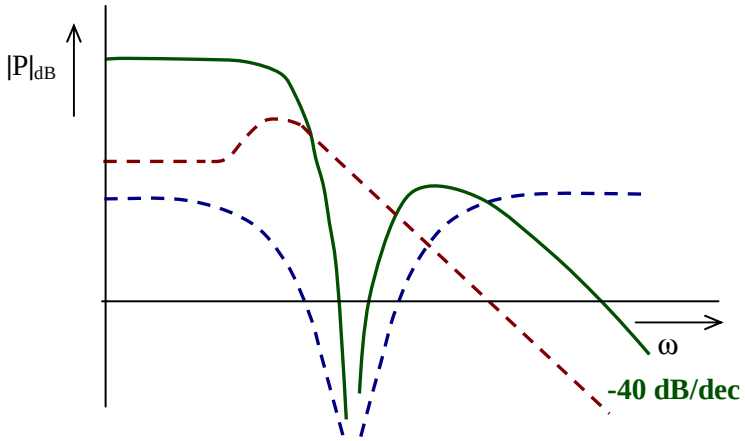


Fig. 41 A qualitative depiction of modulus of cascade connection LP2 and BS2

7.10.5 Cascade connection of HP2 and BS2

Cascade connection of HP2 and BS2 (notch) is shown in Fig. 7.42

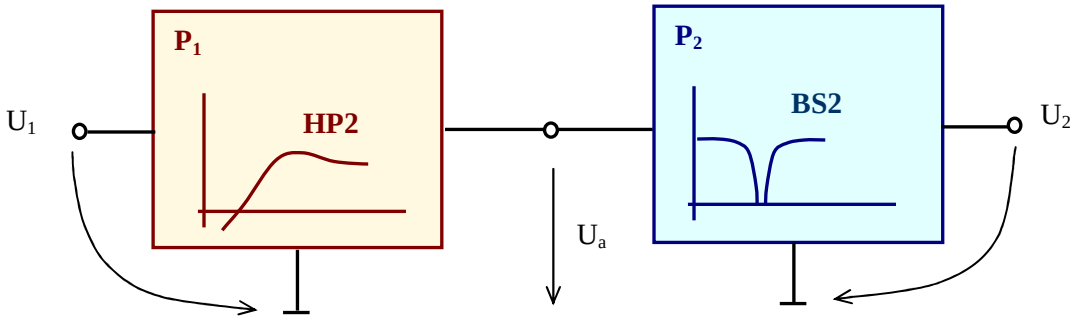


Fig. 7.42 Cascade connection of HP2 and bandstop filter

That way we get a **high-pass filter with „zero frequency”** - Fig. 7.43.

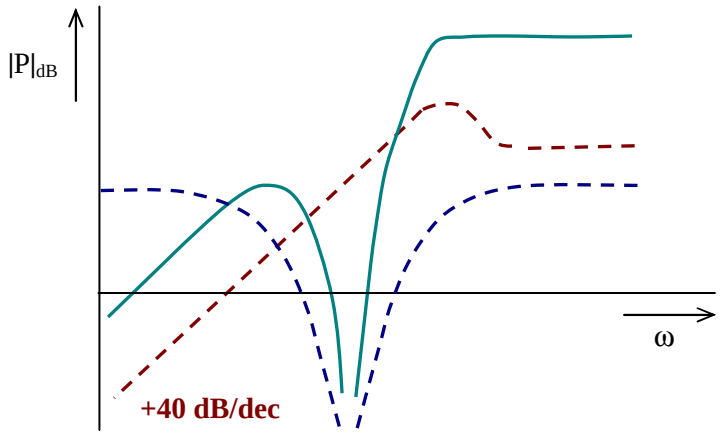


Fig. 41 A qualitative depiction of modulus of cascade connection HP2 and BS2

7.11 Non ideal operational amplifiers

An operational amplifier is not an ideal circuit component. The most degradation is a change of its gain as frequency increases – Fig. 7.42; a real operational amplifier modulus of gain A.

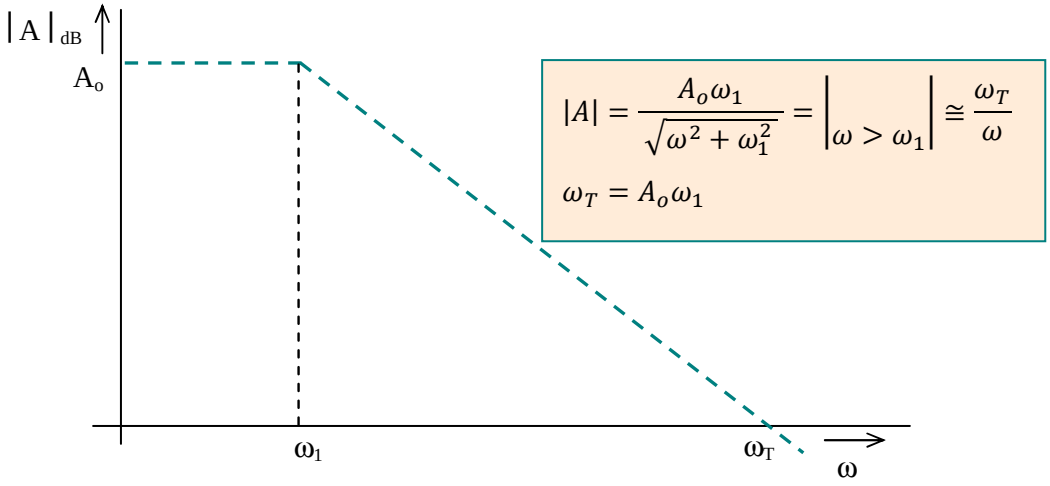


Fig. 7.42 A real operational amplifier modulus of gain A

All the above done outcomes are valid if needed modulus of transfer function is much less than the operational amplifier modulus $|A|$ - preferably ***for 40 dB (and more).***

In Fig.7.43 we can see solid lines – these meet the demands above. Dashed lines define transfers which do not meet the demands above ("graphical criterion"). It is evident that in real circuitry we are not able to construct an ideal HP or BS structures – high frequencies are problematic in any case (with any real operational amplifier).

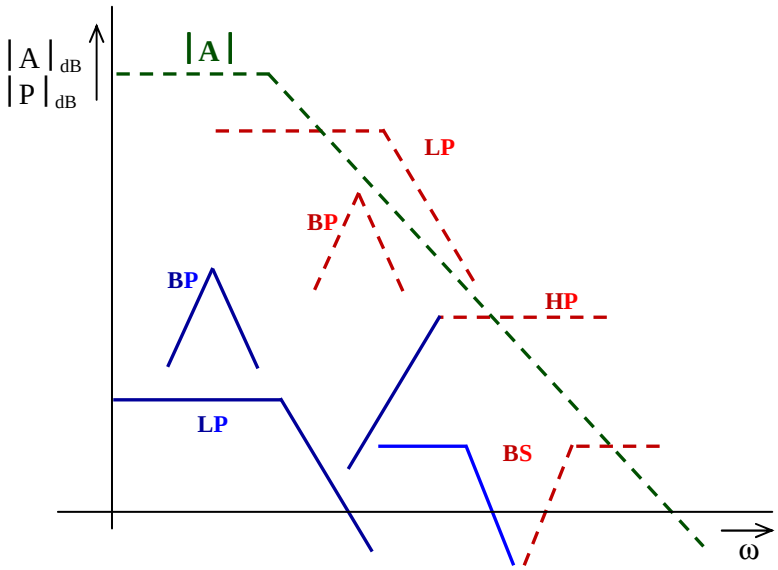


Fig. 43 A „Graphical criterion of operational amplifier usability": modulus of needed transfer function is much less than modulus of OPA gain – solid lines – satisfactory functions; dashed lines - not satisfactory functions

Summary

1) Filter is an electric circuit designed to implement a specific frequency response. Given a filter, obtaining the frequency response is just a matter of applying circuit theory.
2) A negative gain in decibels is reffered to as attenuation. For example, – 10 dB gain is the same as 10 dB attenuation.
3) For the ideal filter, the voltage remains fixed in amplitude until a critical frequency is reached (cutoff frequency). At that frequency, and for all higher frequencies, the output is zero. The ideal filter characteristic is the "brick wall" filter – but we cannot achieve the ideal. In practice we want a magnitude response to be constatnt as possible in the passband. In the stoppband we require n – polle rolloff – an approximation to the brick wall.
4) We can connect modules in cascade such that overall frequency response is of the needed form.
5) Butterworth filter
    a. The flattest possible pass – band magnitude response
    b. The pulse response has moderate overshoot and ringing
    c. Attenuation is – 3 dB at the design cutoff frequency
6) Chebyshev filter
    a. The equal ripple magnitude response

b. The cutoff frequency is defined as the frequency at which the response falls below the ripple band

c. Chebyshev filters provide better attenuation beyond the pass – band than Butterworth filters

d. The Chebyshev has even more ringing in its pulse response than the Butterworth

7) Bessel (Thomson) filter

a. The linear phase response (= maximally flat delay) → it provides best step response (very little overhoot or ringing)

b. It exhibits slower rate of attenuation beyond passband than Butterworth

8) It is useful to understand the netvork functions (1^{st} order, 2^{nd} order)

9) The Bridgman – Brennar structure is the starting point for the implementation of many different types of filter. Relations obtained by analysis are modified for the design of filter

10) All in the text done outcomes are valid if needed modulus of a transfer function is much less than the OPA modulus $|A|$ - preferably for 40 dB (and more)

11) If OPA is not ideal, it´s parameters can have a drastic effect on filter response – see [4, 5] for example

Questions 7

You can find the answers in this text.

1. Define mathematical models of 2^{nd} order filters (LP, HP, BP, BS – notch).
2. How many elements does an ideal LP filter need (brick-wall frequency response)?
3. What approximation has maximally flat magnitude?
4. How can we gat a BS (notch) filter from LP and HP filters (the same characteristic frequency and Q)?
5. What is a “peak” of magnitude (LP2, HP2) - how is it connected with quality factor Q?

Problems 7

Example 7.1

It is required that we design the second order low pass filter. However, only one op – amp is available and the stockroom has only 0,1 μF capacitors. It does have a good stock of resistors. You are to design a filter to meet the following specifications:

a) The response is to be Butterworth
b) The characteristic frequency f_0 is 159,15 Hz.
c) You can choose any low frequency gain

Example 7.2

It is required that we design the third order low pass filter. You are to design a filter to meet the following specifications:

a) The response is to be Butterworth
b) The characteristic frequency f_0 is 159,15 Hz.
c) We need low frequency gain just 1

Example 7.3

It is required that we design the third order low pass filter. You are to design a filter to meet the following specifications:

a) The response is to be Butterworth
b) The characteristic frequency f_0 is 1591,5 Hz.
c) We need low frequency gain just 1

Example 7.4

It is required that we design the third order low pass filter. You are to design a filter to meet the following specifications:

a) The response is to be Butterworth
b) The characteristic frequency f_0 is 15,915 Hz.
c) We need low frequency gain just 1

PROBLEMS KEY 7

Ad example 7.1)

See Tab.1: Butterworth filter; $a_1 = 1{,}414\,214$ thus $Q = 1/a_1 = 0{,}7071$; $b_1 = 1$ so $\omega_p = b_1\omega_o = 2\pi f_o = 1000$ rad·s^{-1}. We can choose the circuit in Fig. 7.17 for example – Fig. 7.44.

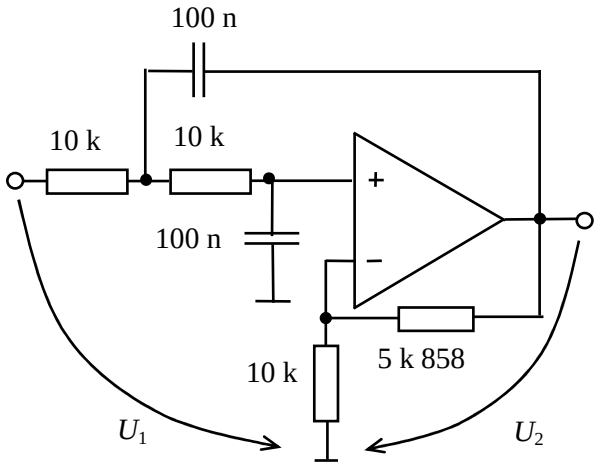


Fig. 7.44 One possible solution of the problem 1

Then

I. We chose the value of the $C = 100$ nF.

II. We determine $R = 1/(\omega_p C) = 1/(10^3 \cdot 10^{-7}) = 10$ kΩ.

III. We determine needed $K = 3 - 1/Q = 3 - 1/0{,}7071 = 1{,}5858$.
We chose suitable value of the $R_f = 10$ kΩ and determine corresponding value $(K-1)\cdot R_f = 0{,}5858 \cdot R_f = 5{,}858$ kΩ.

Ad example 7.2)

See Tab.1: Butterworth filter;

n	b_0	a_1	b_1
3	1, 000000	1, 000000	1, 000000

Thus we must use a cascade connection of one the second order LP filter: $a_1 = 1{,}000000$; $b_1 = 1{,}000000 \rightarrow Q_1 = 1/a_1 = 1$ and $\omega_p = b_1\omega_o = 2\pi f_o = 1000$ rad·s^{-1} and $Q = 0{,}7071$; and one the first order LP with $\omega_p = b_1\omega_o = 2\pi f_o = 1000$ rad·s^{-1}.

For the second order filter part we can choose the circuit in Fig. 7.17 for example – with $K = 1$ – see Fig. 7.45 – and a usual choice $R_1 = R_3 = R$.

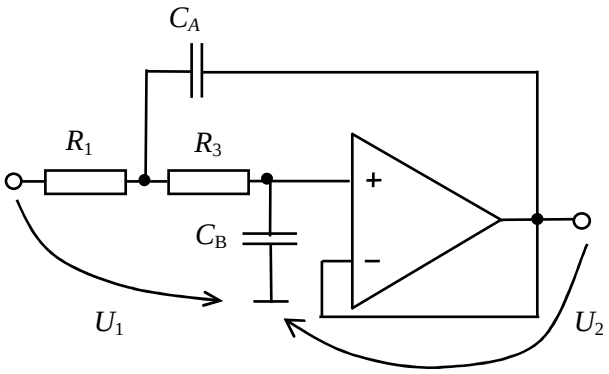


Fig. 7.45 The second order part of the all filter with low pass gain 1

Then

I. We chose the value of $R = 10$ kΩ.

II. We determine $C_A = \frac{2Q_1}{R\omega_P} = \frac{2}{R\omega_P} = \frac{2Q_1}{10^4 \cdot 10^3} = 2 \cdot 10^{-7} F$ (200) nF

$C_B = \frac{1}{2Q_1 R\omega_P} = \frac{1}{2R\omega_P} = \frac{C_A}{4} = 0{,}5 \cdot 10^{-7}\ F$ (50) nF

For the first order filter part we can choose the circuit in Fig. 7.24 for example – with $K = 1$ – see Fig. 7.46.

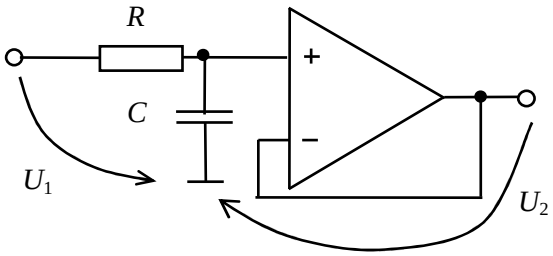


Fig. 7.46 The first order part of the all filter with low pass gain 1

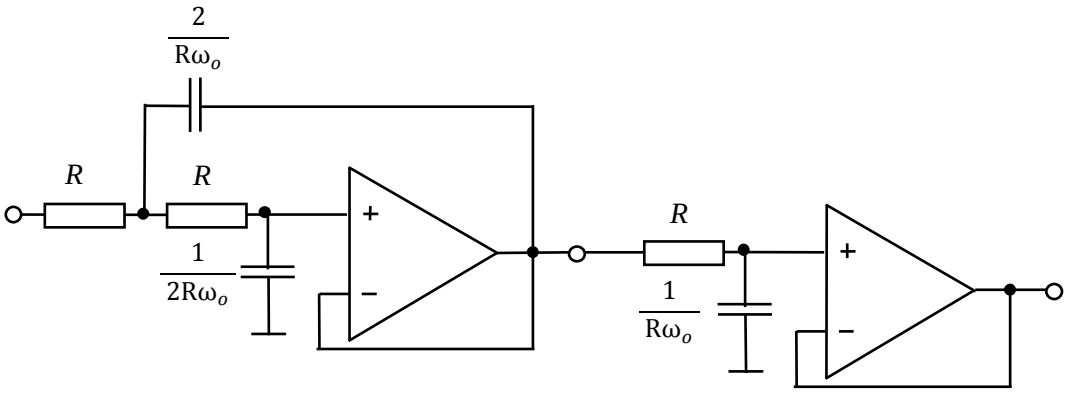


Fig. 7.47 The whole third order Butterworth LP with low pass gain 1; if $\omega_o = 10^3$ and $R = 10^4$ then $1/(R\omega_o) = 10^{-7}$

$$\frac{U_2}{U_1} = \frac{\frac{1}{pC}}{R + \frac{1}{pC}} = \frac{\frac{1}{RC}}{p + \frac{1}{RC}} = \frac{\omega_p}{p + \omega_p}$$

$$\omega_p = \frac{1}{RC}$$

I. We chose the value of $R = 10$ kΩ.

II. We determine $C = \frac{1}{R\omega_P} = 10^{-7}\ F$ (100) nF

Then we can see the "whole" 3rd order Butterworth LP in Fig. 7.47.

Ad example 7.3)

We can use the solution of the problem 7.2. It is evident that needed new characteristic frequency is ten times higher than 159,15 Hz. Thus it is enough to divide all capacitances by 10 (or it is enough to divide all resistances by 10) to get needed characteristic frequency. Now is valid $1/(R\omega_o) = 10^{-8}$. Or we can divide all capacitances by 2 and all resistances by 5, for example.

Ad example 7.4)

We can use the solution of the problem 7.2. It is evident that needed new characteristic frequency is ten times lawer than 159,15 Hz. Thus it is enough to multiple all capacitances by 10 (or it is enough to multiple all resistances by 10) to get needed characteristic frequency. Now is valid $1/(R\omega_o) = 10^{-6}$. Or we can to multiple all capacitances by 2 and all resistances by 5, for example.

Basic texts

[1] Punčochář, J.: Operační zesilovače v elektronice. BEN – Praha, 2002, ISBN 80-7300-059-8

[2] Punčochář, J.: Lineární obvody s elektronickými prvky. Skriptum, VŠB-TU Ostrava 2002, ISBN 80-248-0040-3

[3] Mohylová, J.: Lineární obvody s elektronickými prvky-Sbírka příkladů, VŠB-TU Ostrava 2002, ISBN 80-248-0098-5

[4] Punčochář, J.: Analysis of Sallen and Key Low – Pass filters with real operational Amplifiers. www.elektrorevue.cz

[5] Kolář, J., Punčochář, J.: Band stop filters with real operational amplifiers. Transactions of VŠB – TU Ostrava, VI, 1, pp. 59 – 63

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

[4] Wai-Kai Chen: Electronics Fundamentals of Circuits and Filters – third edition. Taylor & Francis Group, USA, 2009, ISBN-13:978-1-4200-5887-1

8. Rectifiers, current and voltage sources, logarithmic amplifier, analog multiplier

Time of study: 6 hours

Goals: the student should be able to

- define basic principles of rectifiers (active)
- define basic principles of electronic voltage sources
- define basic principles of electronic current sources
- define basic principles of a logarithmic amplifier
- define basic principles of an analog multiplier and divider

EXPLANATION

8.1 Rectifier

Rectification of signals smaller than a diode drop cannot be done with a simple diode-resistor combination. We can use op-amps with a diode in the feedback loop – Fig. 8.1.

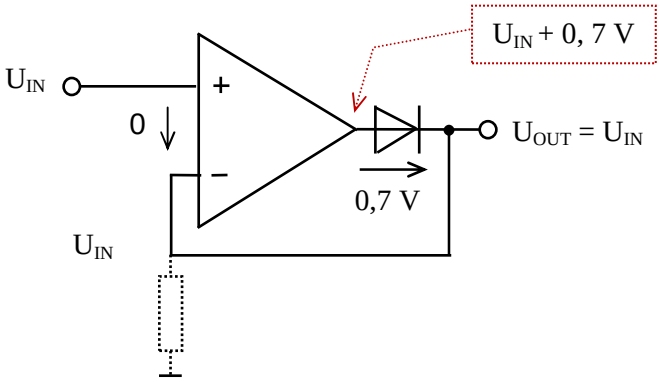


Fig. 8.1 Basic half-wave rectifier – ideal operational amplifier; $U_{IN} > 0$; Si – diode with voltage drop about 0,7 V

For $U_{IN} > 0$ the diode provides negative feedback (diode is open) – the output ideally follows the input (infinity op amp gain). If op amp gain A is finite – see fig. 8.2:

$$U_d = \frac{U_{IN}+U_d+0{,}7}{A} \quad \Rightarrow \quad U_d(1+A) = U_{IN} + 0{,}7$$

$$U_d = \frac{U_{IN}+0{,}7}{A+1}$$

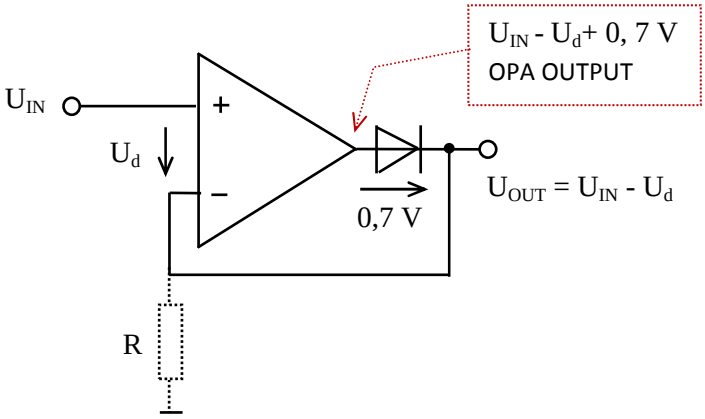


Fig. 8.2 Basic rectifier – real op amp; $U_{IN} > 0$; Si – diode with voltage drop about 0, 7 V

$$U_{OUT} = U_{IN} - U_d = U_{IN} - \frac{U_{IN}+0{,}7}{A+1} \quad \Rightarrow \quad U_{OUT} = \frac{A \cdot U_{IN}}{A+1} - \frac{0{,}7}{A+1}$$

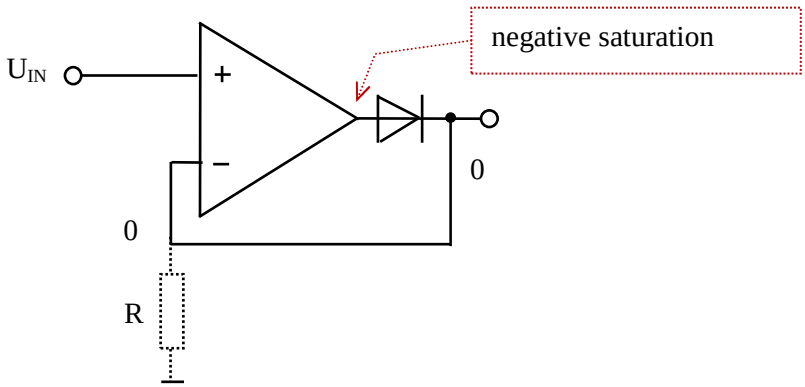


Fig.8.3 Basic rectifier; $U_{IN} < 0$

For $U_{IN} < 0$ the op amp goes to ***negative saturation*** and U_{OUT} is at ground (diode is off). R could be chosen smaller for lower output impedance (we must respect maximum op amp current) – Fig. 8.3. A better solution is to use a follower with another op amp – Fig. 8.4.

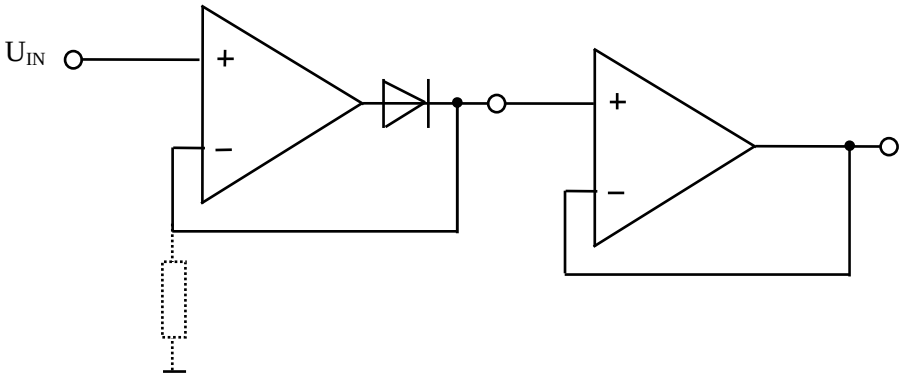


Fig.8.4 Basic rectifier with very small output resistance

There is a problem with this simple circuit. An op-amp cannot swing its output infinitely fast and the recovery from negative saturation takes some time, during which the output is incorrect – Fig. 8.5.

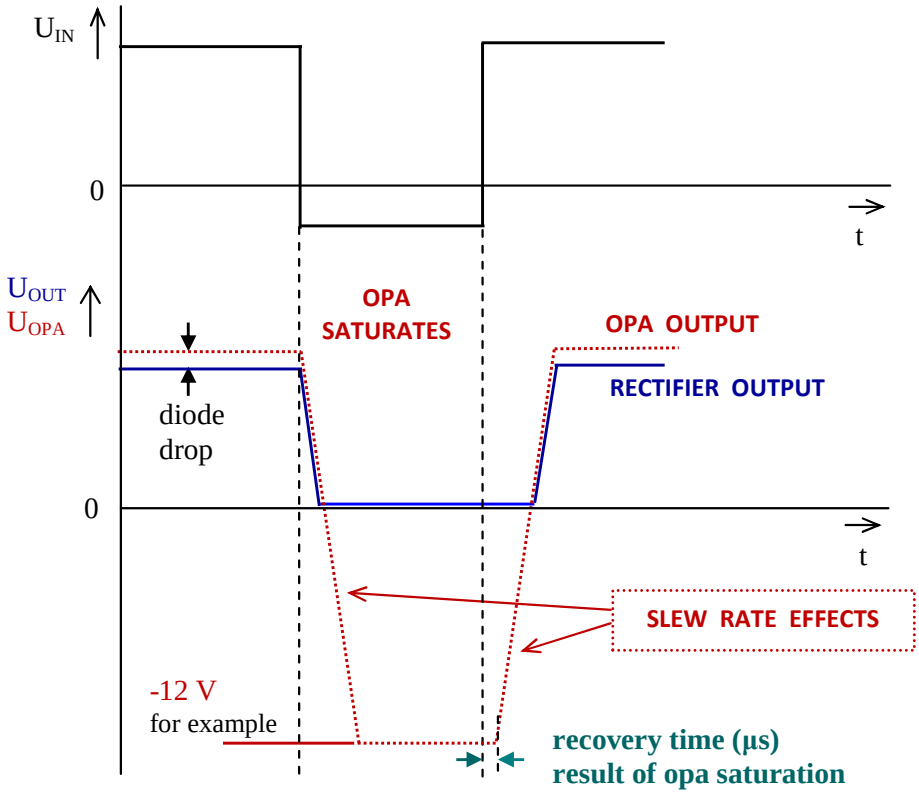


Fig. 8.5 Basic rectifier in dynamic mode

A circuit modification in Fig. 8.6 improves the situation considerably. D_1 makes the circuit a unity – gain invertor for $\boldsymbol{U_{IN} < 0}$. D_2 clamps the op-amp's output at one diode drop below ground for $\boldsymbol{U_{IN} > 0}$ - thus D_1 is back-biased and U_{OUT} sits at ground.

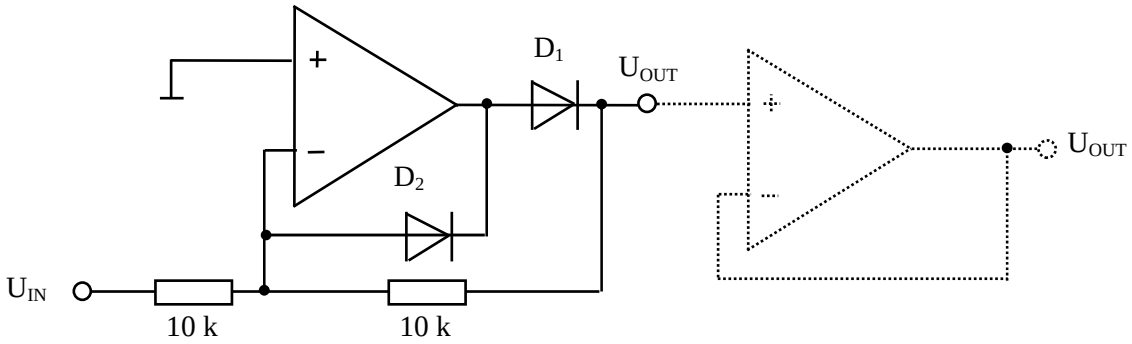


Fig.8.6 Improved half-way rectifier (inverting; with very small output resistance if the buffer is used)

The improvement comes because:

a) the op-amp's output swings only two diode drops as the input signal passes through zero (about 1,2 V) instead of – 12 V (see Fig. 8.5)

b) for $\boldsymbol{U_{IN} < 0}$ there exists feedback via D_2 – thus op amp does not saturate – thus recovery time is insignificant – Fig. 8.7

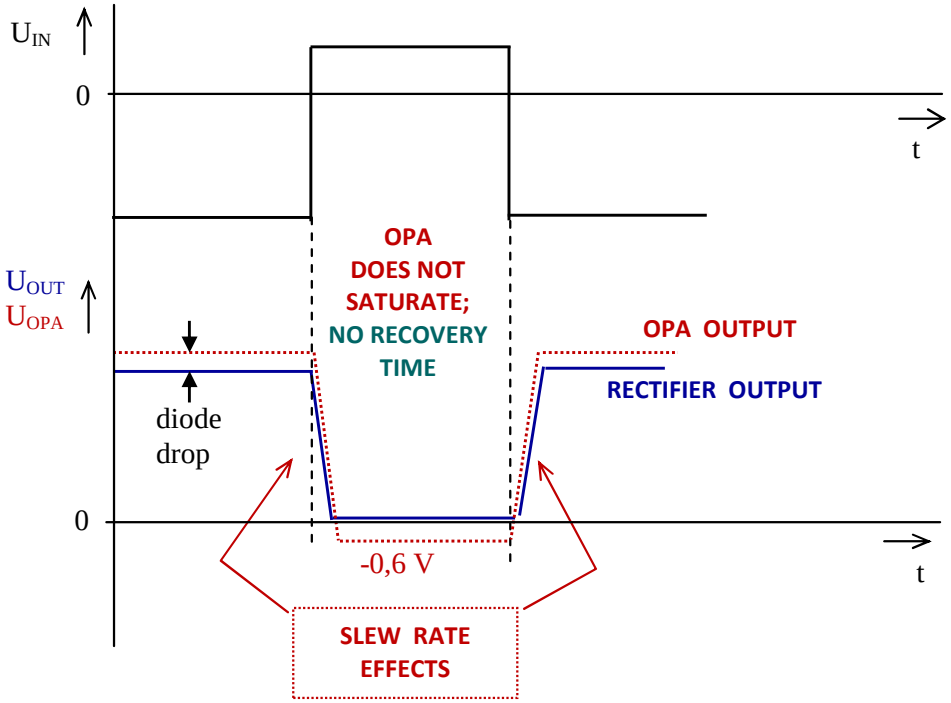


Fig. 8.7 Improved rectifier in dynamic mode

The performance of these circuits is improved if we choose an op-amp with a high slew rate and small recovery time.

The circuit shown in Fig. 8.8 gives a positive output equal to the magnitude of the input signal; it is a full-wave rectifier.

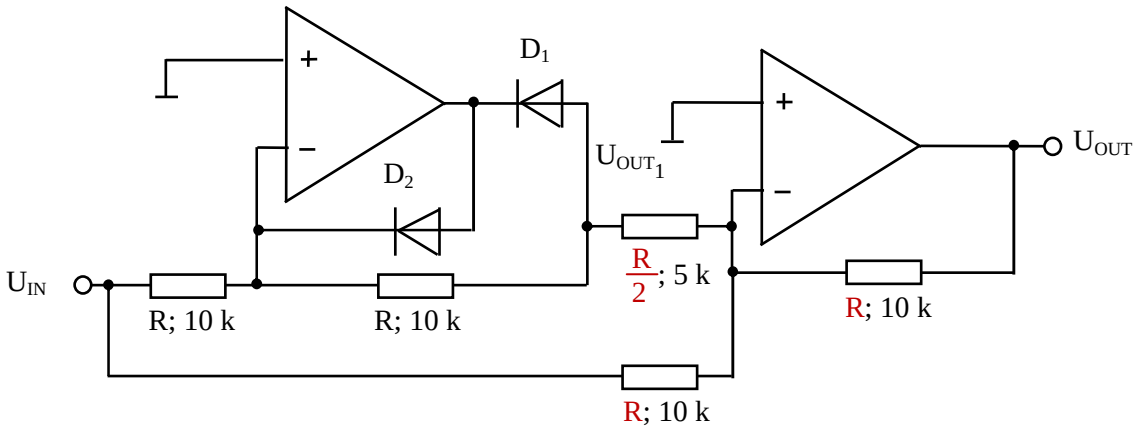


Fig.8.8 Improved full-way rectifier (noninverting; with very small output resistance)

The first opa creates improved half-way rectifier (*complementary to the circuit in Fig. 8.6*). But the second opa creates an inverting summing amplifier, there is always valid:

$$U_{OUT} = -2 \cdot U_{OUT_1} - U_{IN}$$

For $\boldsymbol{U_{IN} > 0}$ is valid $U_{OUT_1} = -U_{IN}$, thus

$$U_{OUT} = -2 \cdot U_{OUT_1} - U_{IN} = -2 \cdot (-U_{IN}) - U_{IN} = |U_{IN}|$$

For $\boldsymbol{U_{IN} < 0}$ is valid $U_{OUT_1} = 0$, thus

$$U_{OUT} = -2 \cdot 0 - U_{IN} = -(-U_{IN}) = |U_{IN}|$$

We can simply write that $U_{OUT} = |U_{IN}|$

8.2 Voltage sources

Simple DC voltage sources are in fact amplifiers (power) of reference voltage. Thus we need not to solve the „all“ the problems – amplifier problems (properties) we know, surely. We must choose just suitable voltage reference and power transistor – see Fig. 8.9, 8.10, 8.11, ... – to construct a voltage source with needed power (output voltages and currents). The feedback is negative, voltage (always) – thus output resistance is small (zero – ideally).

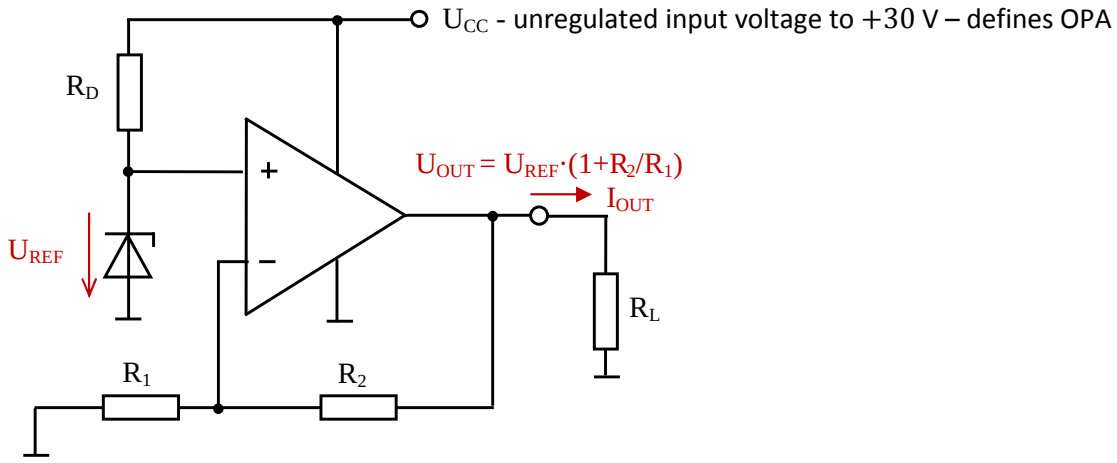


Fig. 8.9 Simple DC regulated voltage source with an OPA only

$I_{OUT_{MAX}}$ in Fig. 8.9 is defined by maximum output current of the used OPA – usually to 10 mA.

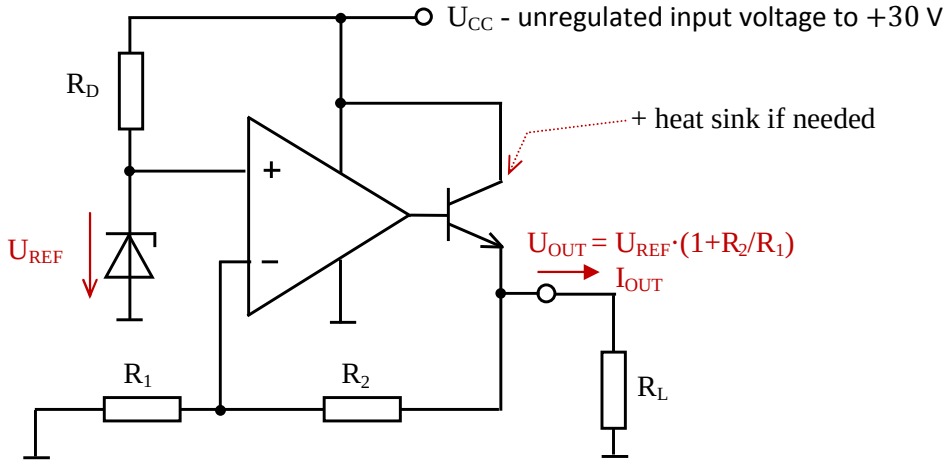


Fig. 8.10 Simple DC regulated voltage source with an OPA and transistor

$I_{OUT_{MAX}}$ in Fig. 8.10 is defined by maximum current of the used transistor – usually to 100 mA.

$I_{OUT_{MAX}}$ in Fig. 8.11 is defined by maximum current of the used transistor – usually to 1 A.

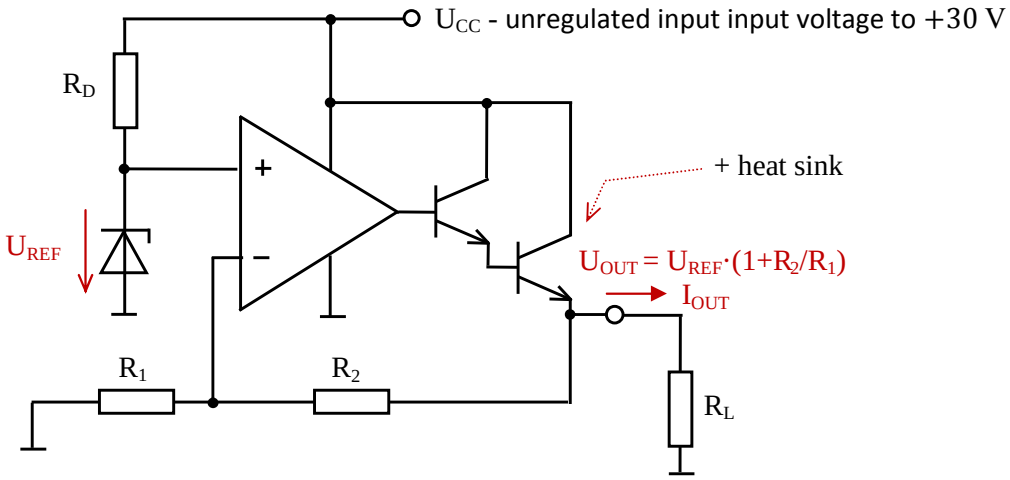


Fig. 8.11 Simple DC regulated voltage source **with an OPA and Darlington transistor configuration**

Transistors (Fig. 8.10 and 8.11) simply act as „current" buffer (common emitter connection) – their voltage gain is around +1 – they do not affect frequency stability (practically) of circuits. But maximum output voltage is (Fig. 8.11)

$$U_{OUT_{MAX}} = U_{CC} - (\text{drop in OPA}) - 2 \cdot U_{BE} \approx U_{CC} - 3{,}5\ \text{V}$$

If ***low drop structure*** is needed, we can use a circuit in Fig. 8.12. Resistors R_{BE} and R_B enable close (fully) PNP transistor (for low values I_{OUT}). Maximum output voltage is now

$$U_{OUT_{MAX}} = U_{CC} - (\text{saturation voltage of PNP})$$

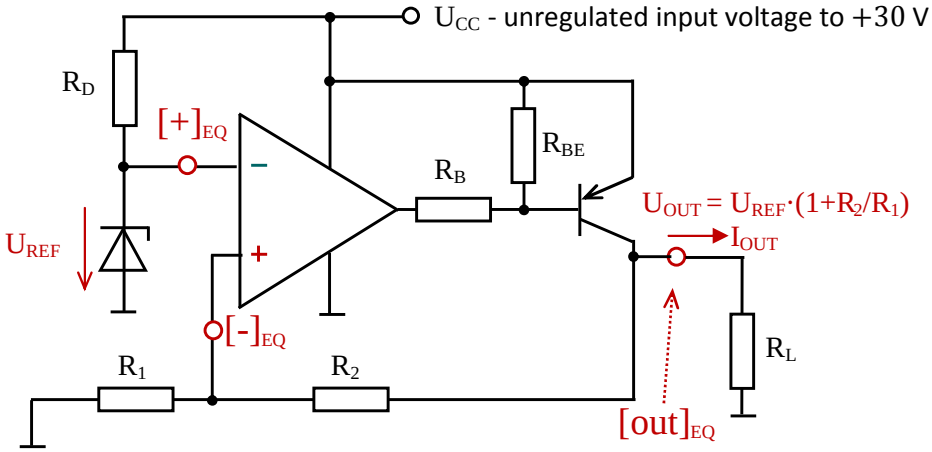


Fig. 8.12 Low drop DC regulated voltage source **with an OPA and PNP transistor**

We must realize, that the PNP transistor inverts signals (common emitter connection) – thus inverting and noninverting inputs of the OPA interchange – see Fig. 8.12; depiction of an equivalent OPA nodes is marked: equivalent noninverting input **[+]EQ**; equivalent inverting input **[-]EQ**; equivalent output **[out]EQ**. It is still negative voltage feedback. We must realize

that the PNP transistor amplifies signals, too. The equivalent OPA has a greater gain than OPA – it means that frequency stability can be worse.

If a *high voltage structure* is needed, we can use circuit in Fig. 8.13, for example.

Transistors amplify and inverse signals (T_1 - common base and T_2 - common emitter; T_3 and T_4 – complementary circuit) – thus inverting and noninverting inputs of the OPA interchange – see Fig. 8.13. OPA and transistors create in fact equivalent high voltage operational amplifier; depiction of an equivalent OPA nodes is marked: equivalent noninverting input **[+]EQ**; equivalent inverting input **[-]EQ**; equivalent output **[out]EQ**. It is still negative voltage feedback. Capacitors (all 100 pF typically) ensure frequency stability of the circuit.

Transistors separate voltages "15 V" and "100 V", too. The feedback circuit must ensure maximum voltage about 10 V (on the OPA input). Diodes (Si - silicon) protect emitter – base junctions of T_1 and T_2 - in reverse mode.

The circuit in Fig. 8.13 makes both polarities of U_{OUT} and I_{OUT} possible (***four quadrant source***).

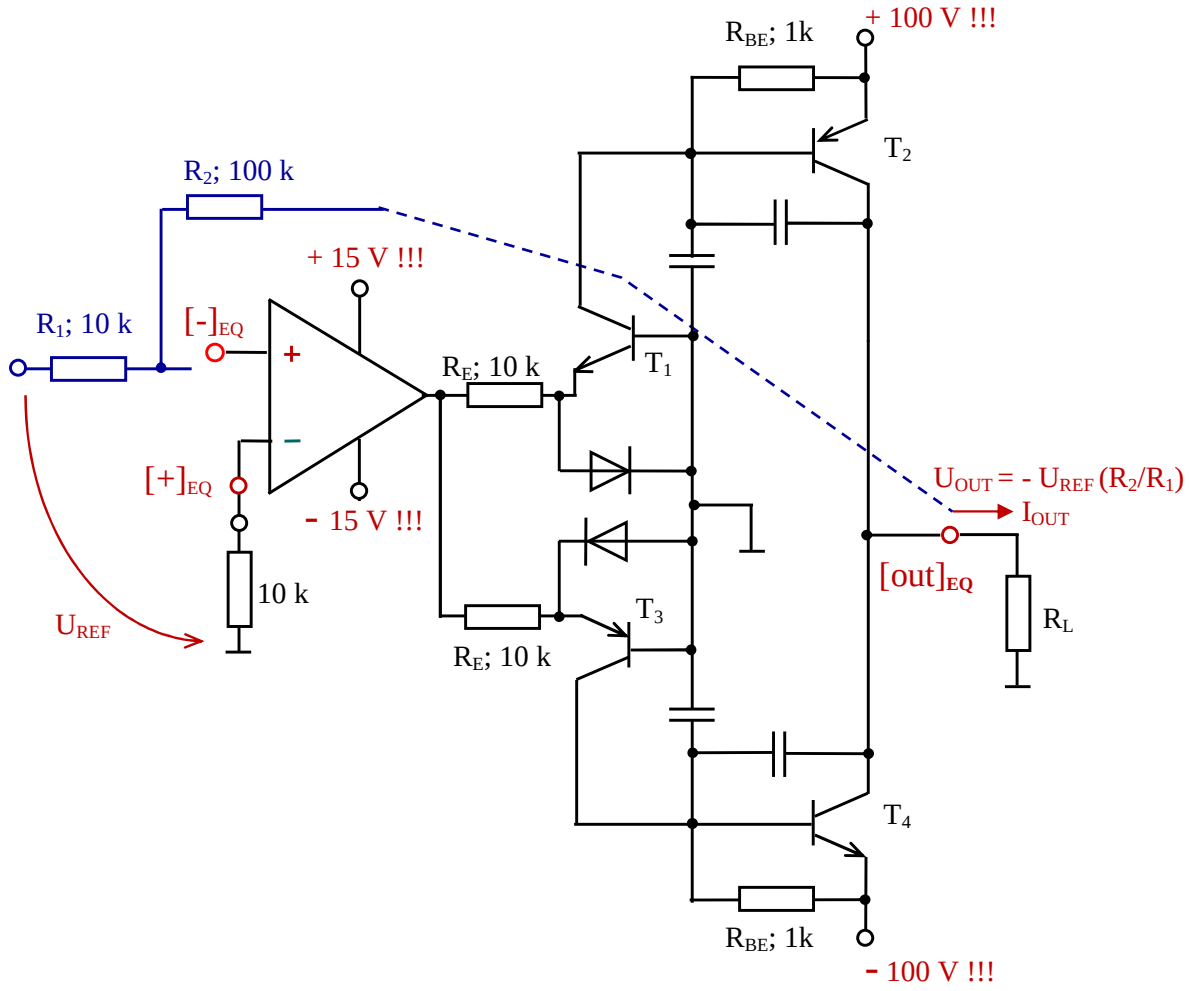


Fig. 8.13 High voltage DC regulated source (inverts U_{REF} , now) – four quadrant mode is possible

If we need only positive polarity of voltage, for example, we can modify the circuit in Fig. 8.13 – see Fig. 8.14; U_{REF} must be negative, of course.

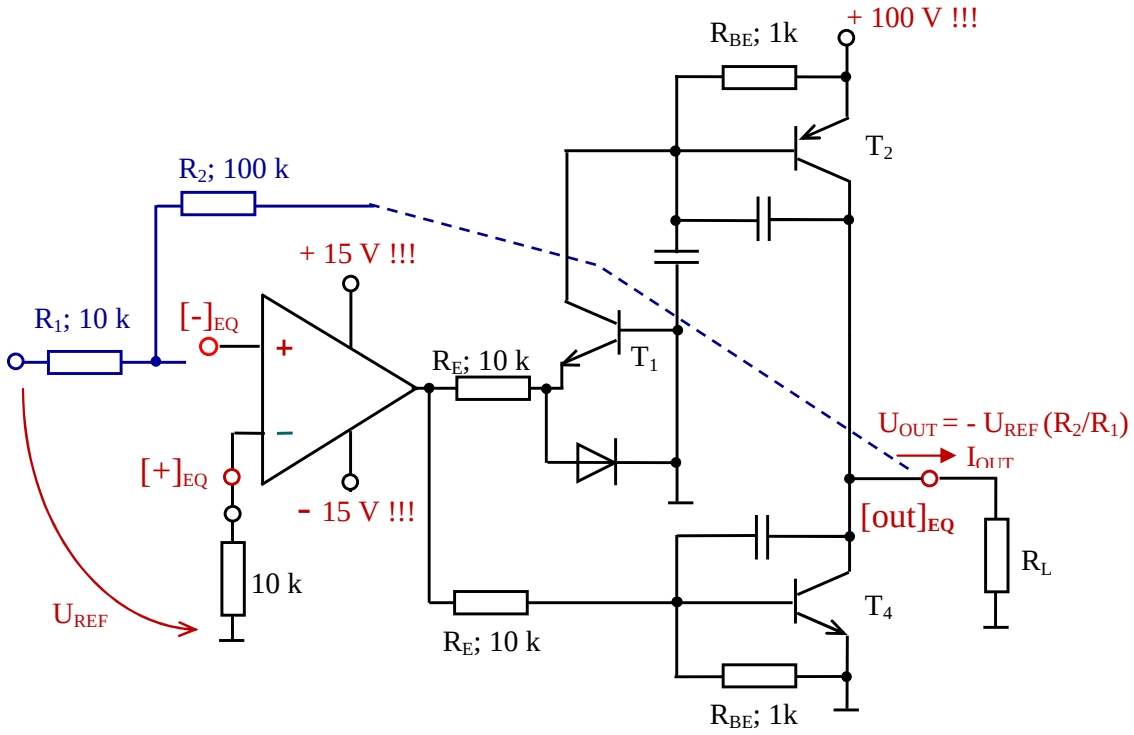


Fig. 8.14 High voltage DC regulated source (inverts U_{REF}); U_{OUT} – positive only; I_{OUT} – positive (source) or negative (sink) – two - quadrant mode is possible

The circuit in Fig. 8.14 makes both polarities of I_{OUT} possible, but only one polarity (here positive) of output voltage is possible (***two quadrant source***).

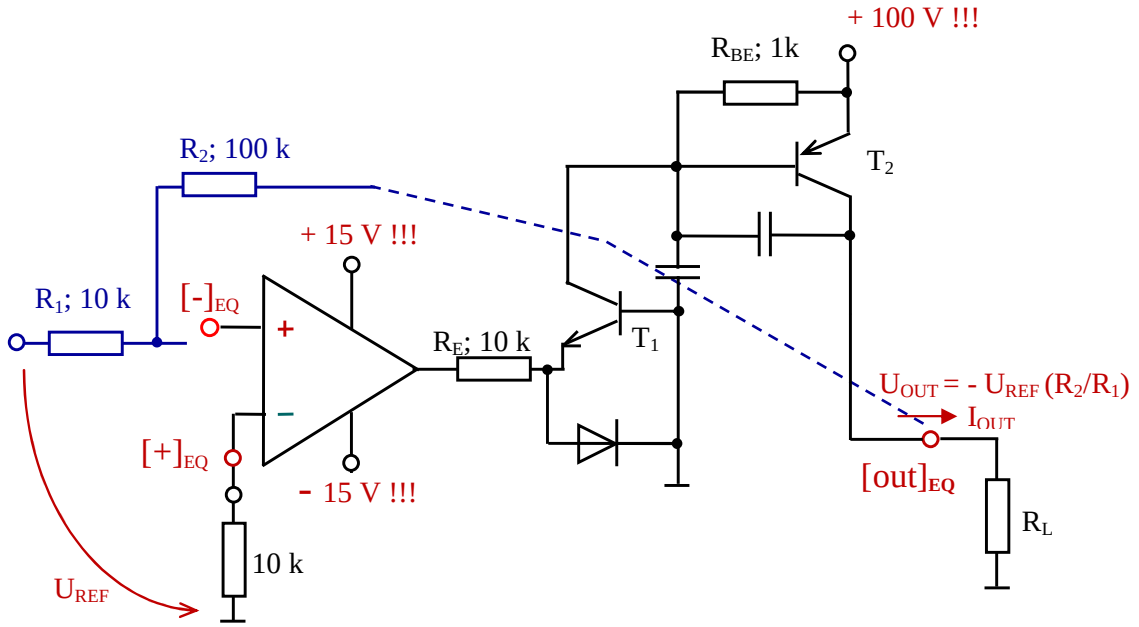


Fig. 8.15 High voltage DC regulated source (inverts U_{REF}); U_{OUT} – positive only; I_{OUT} – positive (source) only – one - quadrant mode is possible

If we need only positive polarity of voltage and current, we can modify the circuit in Fig. 8.14 too – see Fig. 8.15. This circuit is not able "sink" current – it is only "source" of current – it is ***one quadrant source***.

If we need ***digitally controlled voltage source*** we can use any of the above depicted circuits and use suitable DAC. For example we can modify the circuit in Fig. 8.10 – see Fig. 8.16.

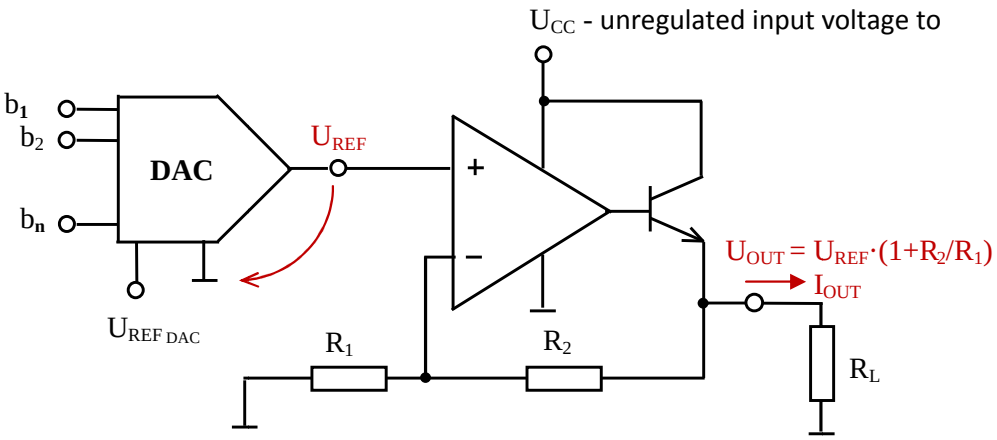


Fig. 8.16 Principle of digitally controlled DC voltage source

We can write

—

where

- a reference voltage of the DAC

– most significant bit (MSB)

– last significant bit (LSB)

Thus we can determine that the source output voltage is

— — —

Example 8.1

Let us have DAC: ; .

Determine an output voltage if and

☑ **Solution:**

— — — — — — — — —

— —

or

—

It is evident that the smallest step is now $\Delta U_{OUT} = \left(1 + \frac{R_2}{R_1}\right) \cdot 39{,}06\ \text{mV}$.

8.3 Current sources

DC current sources are in fact amplifiers (power) with negative current feedback. They are usually controlled by voltage (but not always). Ideally we need infinite output resistance. You can see a simple example in Fig. 8.17.

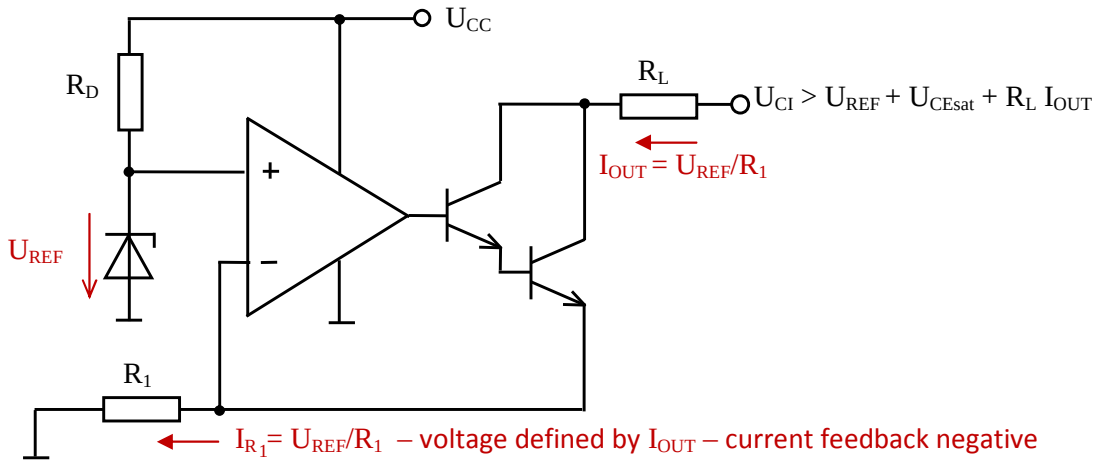


Fig. 8.17 DC regulated current source with an OPA and Darlington transistor configuration

Let us suppose an ideal OPA, then voltage on R_1 is just U_{REF} (zero differential input voltage of OPA, zero input currents) and

$$I_{R_1} = \frac{U_{REF}}{R_1} = I_{OUT} + \frac{I_{OUT}}{\beta_{DAR}}$$

We know that current gain β_{DAR} of Darlington configuration is very large, so we usually suppose that

$$\frac{U_{REF}}{R_1} = I_{OUT}$$

If we use one transistor only, an error is a bit greater – but in practice usually acceptable.

To ensure right quiescent transistor point we can use another voltage source

$$U_{C_I} > U_{REF} + U_{CE_{sat}} + R_L \cdot I_{OUT}$$

or it must be

$$R_L \cdot I_{OUT} < U_{C_I} - \left(U_{REF} + U_{CE_{sat}}\right)$$

We cannot connect a load resistance to ground terminal – it is disadvantage of this circuit (neither side grounded). This problem is solved in Fig. 8.18.

The first OPA_1 creates current source (see Fig. 8.17) – it is used to convert an input voltage U_{REF} (referenced to ground) to a U_{CC} – referenced input to the final current source (with OPA_2 – complementary circuitry to the circuit in Fig. 8.17 – no Darlington transistor configuration is used). The OPA_2 must be able to operate with its inputs and output near or at the positive supply voltage U_{CC} (RRIO - OPA_2) to be able to regulate low values of the output current (low values of U_{REF}), too. Maximum output voltage is limited now to a value:

$$U_{R_{L_{MAX}}} \cong U_{CC} - U_{REF} \cdot \frac{R_2}{R_1} - U_{CE_{sat}}$$

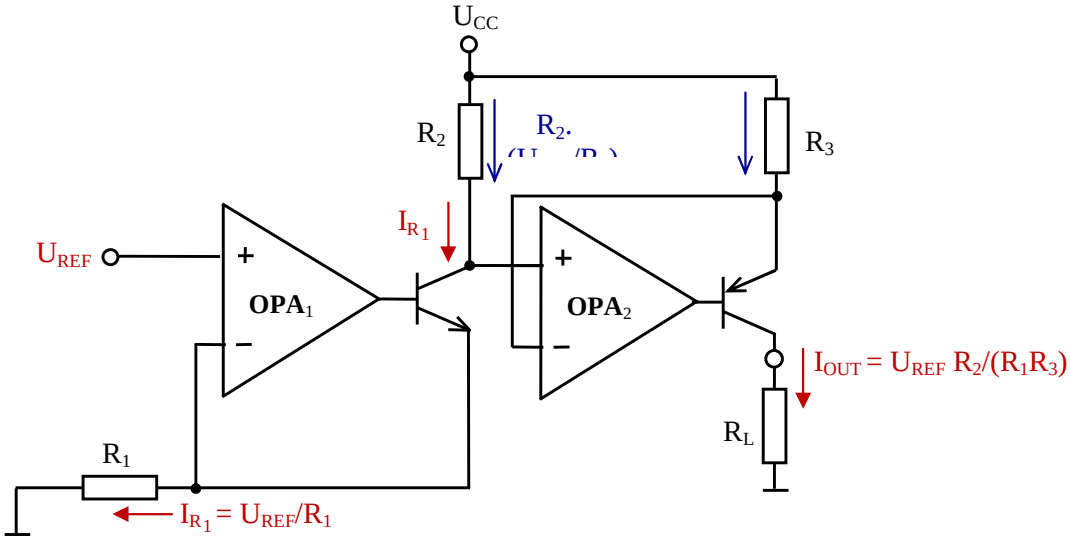


Fig. 8.18 DC regulated current source for grounded load; OPA_2 – rail to rail input/output - RRIO

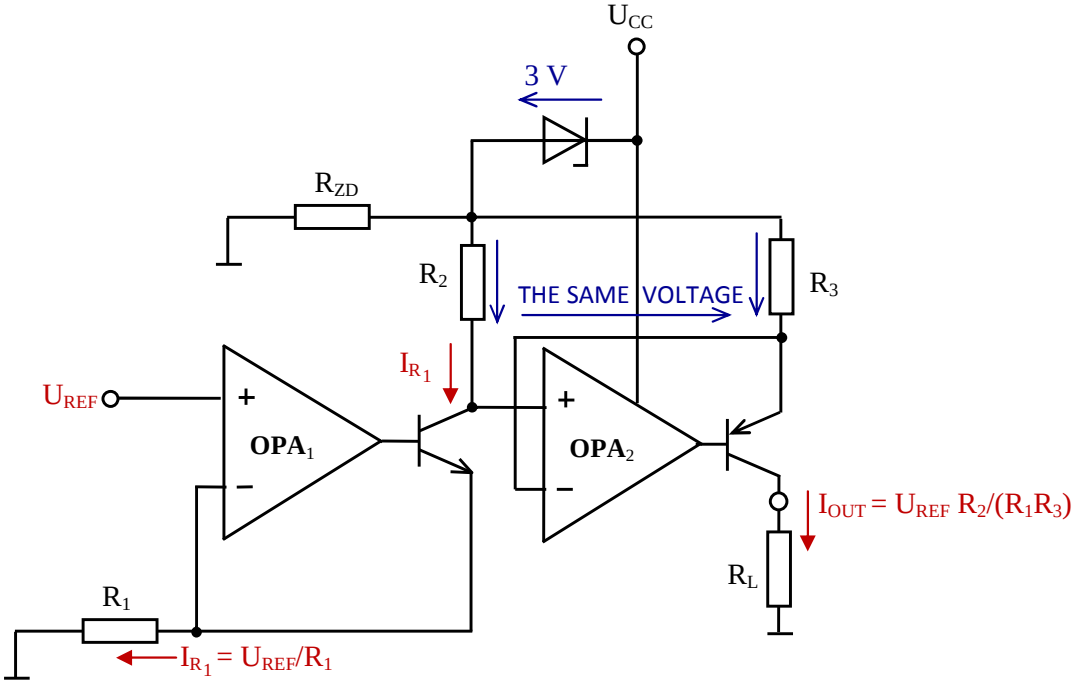


Fig. 8.19 DC regulated current source for grounded load; OPA_2 – NO rail to rail input/output

Alternatively, the OPA_2 could be powered from a higher value voltage than PNP transistor – see Fig. 8.19 for example. Voltage drop 3 V enables to use any "ordinary" OPA_2. But maximum output voltage is limited now to a value:

$$U_{R_{I_{MAX}}} \cong U_{CC} - 3\ \mathrm{V} - U_{REF} \cdot \frac{R_2}{R_1} - U_{CE_{sat}}$$

If OPA_2 is RRI (rail to rail input) only, we must use another solution – Fig. 8.20. The Zener diode ensures needed voltage drop for the OPA_2 output – it does not saturate now, it is able to fully shut PNP transistor if it is needed. Maximum output voltage is limited now to a value:

$$U_{R_{L_{MAX}}} \cong U_{CC} - U_{REF} \cdot \frac{R_2}{R_1} - U_{CE_{sat}}$$

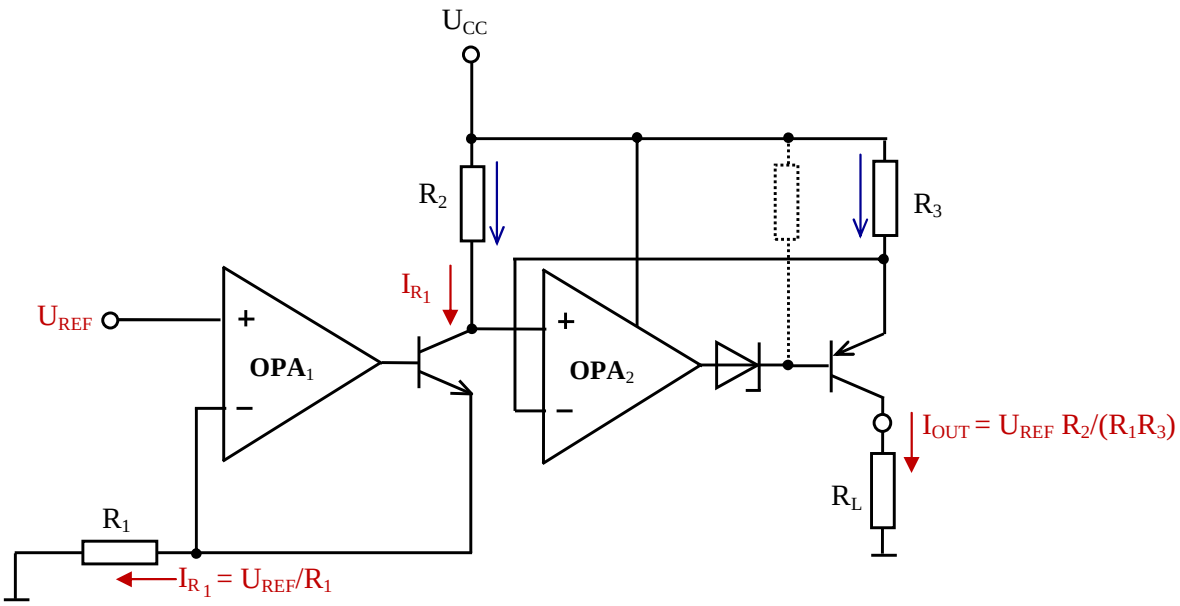


Fig. 8.20 DC regulated current source for grounded load OPA_2 – rail to rail input - RRI

8.4 Logarithmic amplifier

Logarithmic amplifiers (Fig. 8.21) use the nonlinear volt-ampere relationship of the PN junction itself. This relationship is given by

$$I_D \cong I_o \cdot \left(e^{U_D/U_T} - 1\right)$$

where

I_o is reverse saturation current

$$U_T = \frac{kT}{q} = \frac{T}{11000}\ \mathrm{V}\ (T \text{ in } ^\circ\mathrm{K})$$

if we have 25 °C then $T = 25 + 273 = 298\ ^\circ\mathrm{K}$; $U_T = 298/11000\ \mathrm{V} \approx 27\ \mathrm{mV}$.

More exactly the relationship is

$$I_D \cong I_o \cdot \left(e^{U_D/\eta U_T} - 1\right)$$

where $\eta = 2$ for small currents in silicon devices.

If we restrict the operation region of U_D so that

$$e^{U_D/U_T} \gg 1$$

the logarithmic relationship may be expressed as

$$\ln I_D = \ln I_o + \frac{U_D}{U_T} \quad \text{or} \quad U_D = U_T(\ln I_D - \ln I_o) = U_T \ln\left(\frac{I_D}{I_o}\right)$$

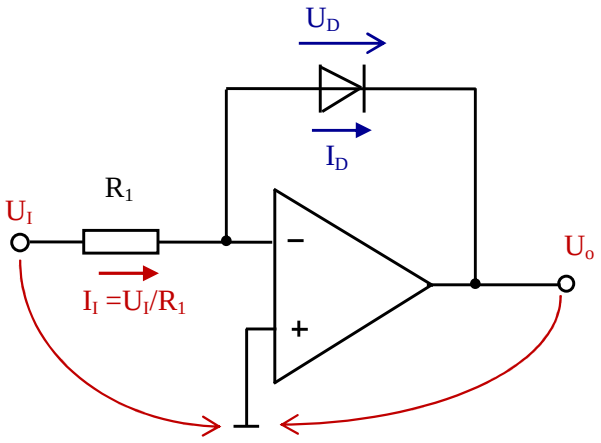


Fig. 8.21 Simple logarithmic diode amplifier

The derivation of the logarithmic function relation proceeds as follows:

$$I_1 = \frac{U_1}{R_1}$$

$$I_D = I_1$$

$$U_D = U_T \ln\left(\frac{I_D}{I_o}\right); \quad U_o = -U_D \quad \Rightarrow$$

$$U_o = -U_T \ln\left(\frac{U_I}{R_1 I_o}\right) = -U_T \ln\left(\frac{U_I}{R_1} - \ln I_o\right)$$

There are two temperature effects:

- a temperature – sensitive scale factor U_T
- a temperature – sensitive offset term $U_T \ln I_o$

The saturation current term can be reduced by the use of a current source I_R and a second (matched; the same I_o) diode D_2 – Fig. 8.22.

Thus we may write:

$$U_3 = U_2 + U_{D_2} = -U_T \ln\left(\frac{U_I}{R_1} - \ln I_o\right) + U_T(\ln I_R - \ln I_o) = -U_T \ln\left(\frac{U_I}{R_1 I_R}\right)$$

The only remaining temperature sensitivity in U_3 is that of the scale factor term (U_T). This can be compensated in the output amplifier (noninverting here) by making its gain temperature – sensitive and compensating for the U_T factor. This is most easily done by using a temperature – sensitive resistor R_T (positive temperature coefficient; temperature increases, U_T increases, output amplifier gain decreases – thus compensates effect of

temperature changes) in the feedback network as shown in Fig. 8.22. The output voltage is then given by:

$$U_o = U_3\left(1+\frac{R_F}{R_1+R_T}\right) = -\left(1+\frac{R_F}{R_1+R_T}\right)\cdot U_T \ln\left(\frac{U_I}{R_1 I_R}\right)$$

The output amplifier has usually a voltage gain of about 16, in order to give an output voltage of – 1 V per decade of input current.

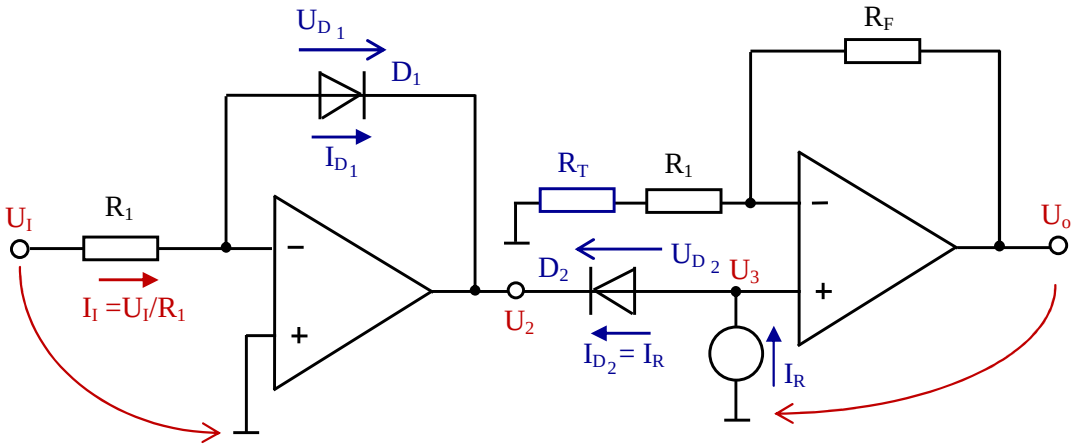


Fig. 8.22 Temperature compensated diode logarithmic amplifier

From the point of small signal properties we can represent D_1 as a signal resistor

$$r_{D_1} \cong U_T/I_1$$

and feedback transfer via D_1 to the noninverting input (Fig. 8.21) is then

$$\beta \cong R_1/(R_1 + r_{D_1}) \approx 1.$$

It means that circuit can be stable (see Nyquist criterion; Chapter 4) – see Fig. 8.23.

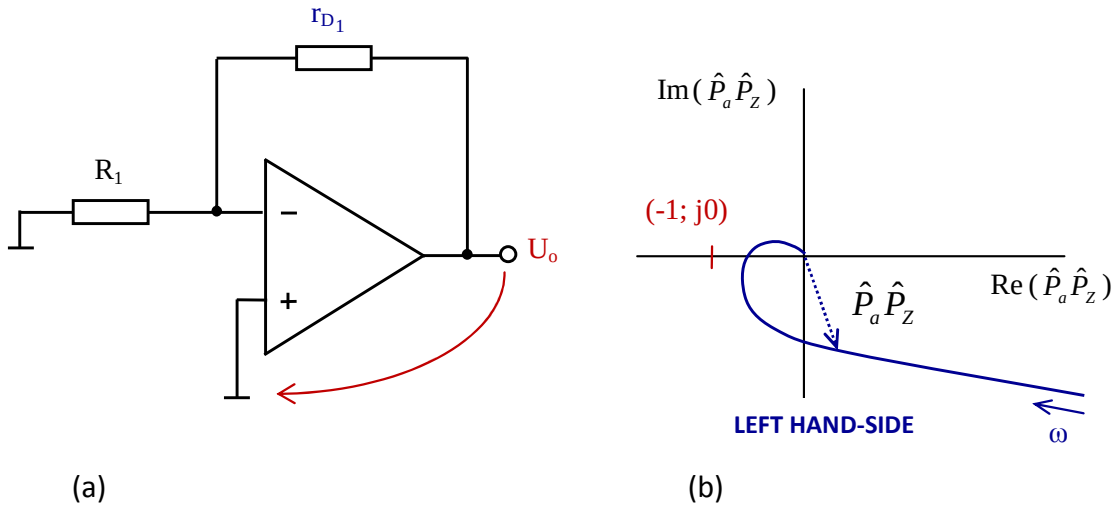


Fig. 8.23 a) A small signal model of the diode log. amp; b) Nyquist criterion: $\hat{P}_a$ - a real frequency stable OPA gain, $\hat{P}_Z = \beta \cong 1$ feedback via D_1 – stable system now.

The circuit in Fig. 8.24 ***exploits transistor***. We know that

$$I_C \cong I_o \cdot e^{U_{BE}/U_T};$$

$$U_{BE} = U_T(\ln I_C - \ln I_o)$$

Analogously to the circuit in Fig. 8.22:

$$U_3 = U_2 + U_{T_2} = -U_T \ln\left(\frac{U_I}{R_1} - \ln I_o\right) + U_T(\ln I_R - \ln I_o) = -U_T \ln\left(\frac{U_I}{R_1 I_R}\right)$$

$$U_o = U_3\left(1 + \frac{R_F}{R_1 + R_T}\right) = -\left(1 + \frac{R_F}{R_1 + R_T}\right) \cdot U_T \ln\left(\frac{U_I}{R_1 I_R}\right)$$

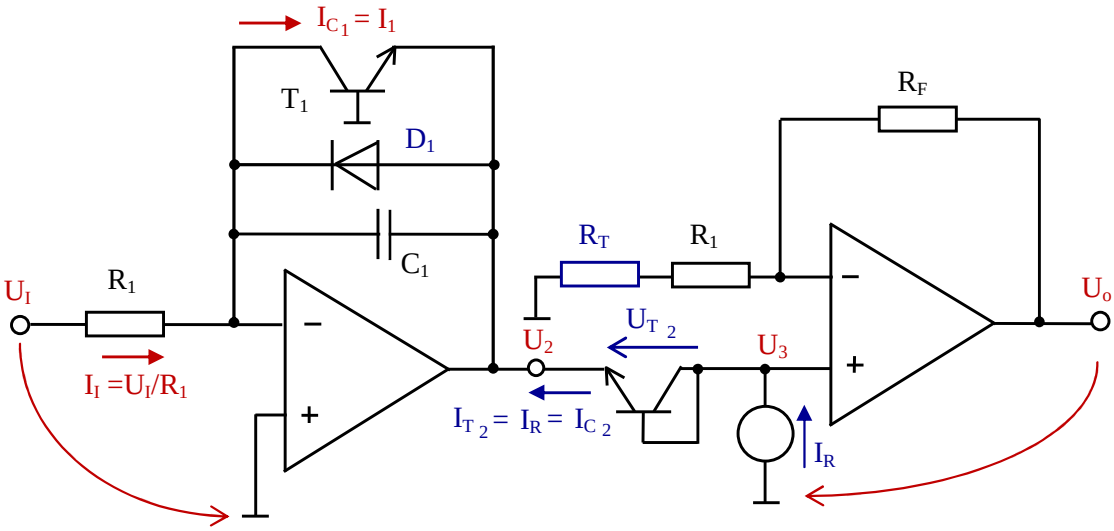


Fig. 8.24 [Temperature compensated transistor logarithmic amplifier](); $U_2 = -U_{BE_1}$; $U_{T_2} = U_{BE_2}$; T_2 – temperature compensation only – as D_2 in Fig. 8.22; D_1 – is necessary to prevent breakdown of T_1 (base-emitter); C_1 – stabilizes the circuit

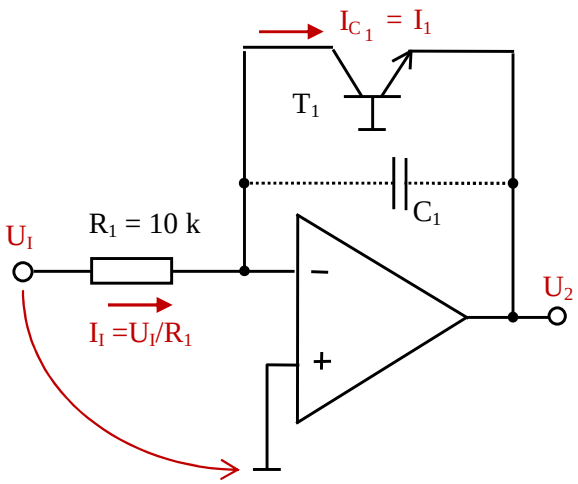


Fig. 8.25 A principle of transistor logarithmic amplifier;

In this circuit the base of T_1 is at the same voltage as the collector because of the virtual ground. T_1 and T_2 should be a matched pair – thermally coupled (ideally a monolithic pair). This circuit will give accurate logarithmic output over seven decades of current (approximately from $1\ \mathrm{nA}$ to $10\ \mathrm{mA}$). It is evident that we must use FET-input OPA which bias current is less than $1\ \mathrm{nA}$ (insignificant). The first OPA offset voltage must be trimmed for zero – U_I may be very small at the lower limit of current.

From the point of small signal properties T_1 represents a common base amplifier. T_1 contributes ***voltage gain inside the feedback loop*** – see Fig. 8.25 and 8.26 **!!!!!! Frequency stability of the system is problematic and dependent on the input voltage U_I value!!!! The capacitor is necessary to stabilize the circuit.**

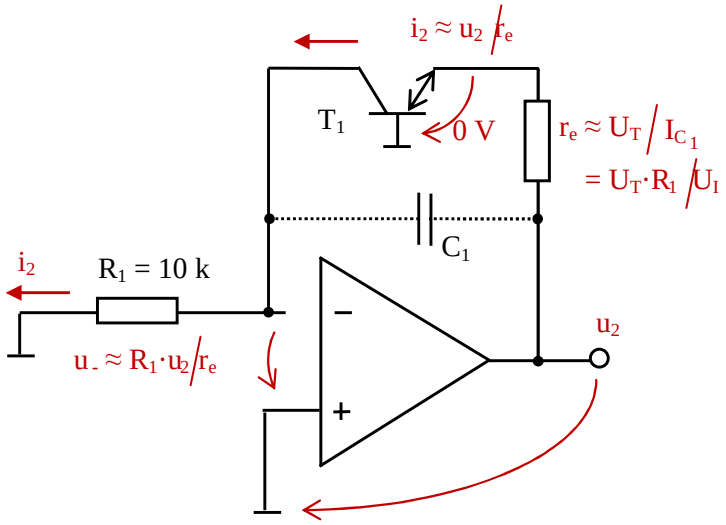


Fig. 8.26 A small signal equivalent circuit – the feedback is <u>disconnected</u> – see the inverting OPA input

Investigate now just a principle situation - the first OPA and the T_1 – Fig. 8.25 and 8.26. We use a T_1 model in accordance with Fig. 2.3 in Chapter 2.

We can easily determine that feedback from the output via T_1 is now (common base connection; we do not consider C_1):

$$\hat{P}_Z = \frac{u_-}{u_2} = \beta \cong \frac{R_1}{r_e} = \frac{R_1}{\frac{R_1 U_T}{U_I}} = \frac{U_I}{U_T}$$

It is a very interesting result. It is evident that for $U_I > U_T \approx 27\ \mathrm{mV}$ is $\beta > 1$ – results see in Fig. 8.27.

The base of T_1 could have been connected to its collector – then we do not need D_1 and C_1. But the base current would then have caused an error (only I_C is an accurate exponential function of U_{BE}) – this way we get nearly the circuit from Fig. 8.22, again. It is more stable – but it gives logarithmic output over "fewer" decades of current.

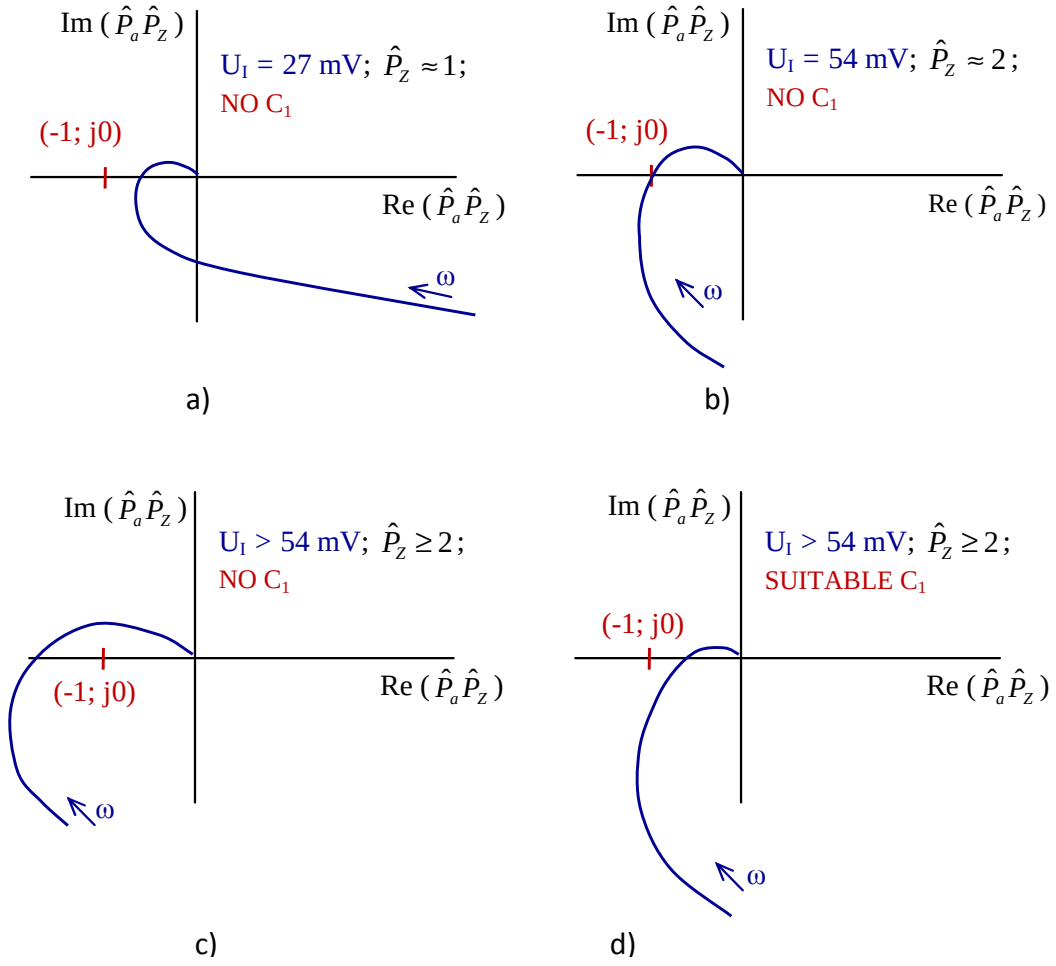


Fig. 8.27 Nyquist criterion: $\hat{P}_a$ - an example of a real frequency stable OPA gain;

a) $U_I = 27$ mV – "small" feedback, NO C_1;

b) $U_I = 54$ mV – critical feedback; NO C_1;

c) $U_I > 54$ mV – overcritical feedback; NO C_1 – the amplifier is unstable;

d) $U_I > 54$ mV – SUITABLE C_1 – the amplifier is stable, now

8.5 Antilogarithmic amplifier

By interchanging the position of the input and feedback elements of the principle logarithmic circuit of Fig. 8.25, we have an antilogarithmic amplifier (inverse-log amplifier), as shown in Fig. 8.28.

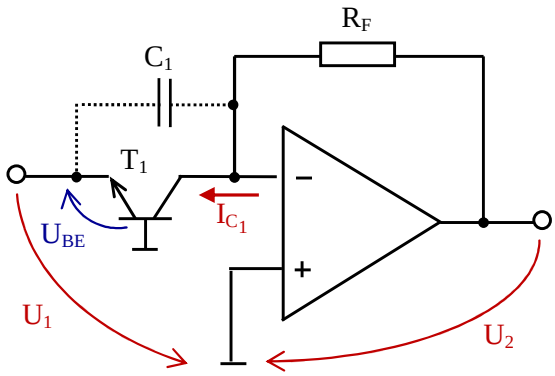


Fig. 8.28 A principle of antilogarithmic amplifier

We know that

$$I_C \cong I_o \cdot e^{U_{BE}/U_T};$$

We must ensure $U_I \leq 0$ in this case: $U_{BE} = -U_I \geq 0$ – thus the transistor quiescent point is right. Now it is evident that

$$U_2 = R_f \cdot I_{C_1} = R_f \cdot I_o \cdot e^{U_{BE}/U_T} = \Big|_{10^{1/2,3} = e} = R_f \cdot I_o \cdot 10^{U_{BE}/(2,3 \cdot U_T)}$$

Using both the logarithmic and antilogarithmic amplifier circuits, we can either multiply or divide input voltages.

8.6 Multiplier – divider

The basic multiplier relationship is

$$\log A + \log B = \log(A \cdot B)$$

We can see a principle structure in Fig. 8.29.

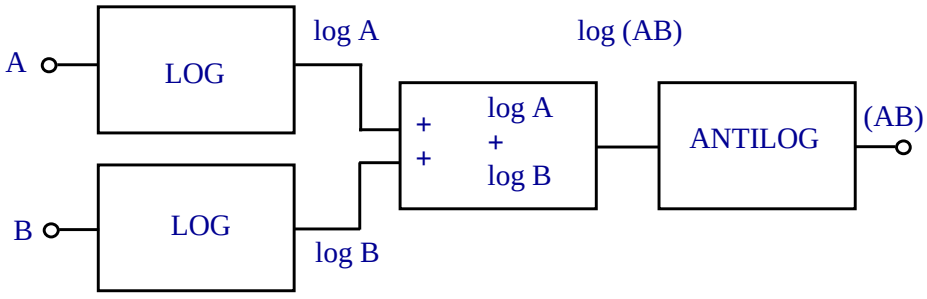


Fig. 8.29 A principle of multipliers

The basic divider relationship is

$$\log A - \log B = \log\left(\frac{A}{B}\right)$$

We can see a principle structure in Fig. 8.30.

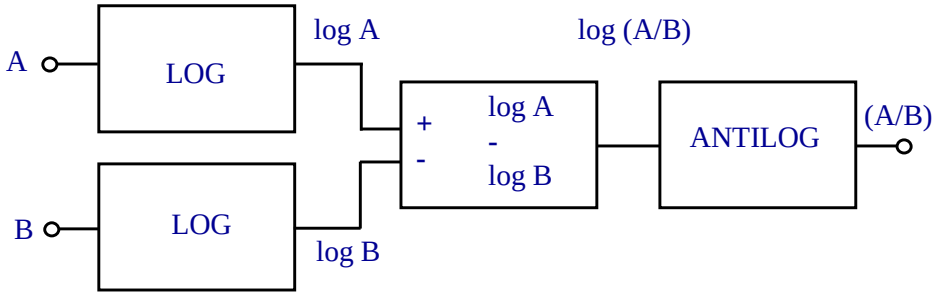


Fig. 8.30 A principle of dividers

A more complex multiplier/divider principle structure we can see in Fig. 31. T_1 to T_4 should be matched – thermally coupled (ideally monolithic structure).

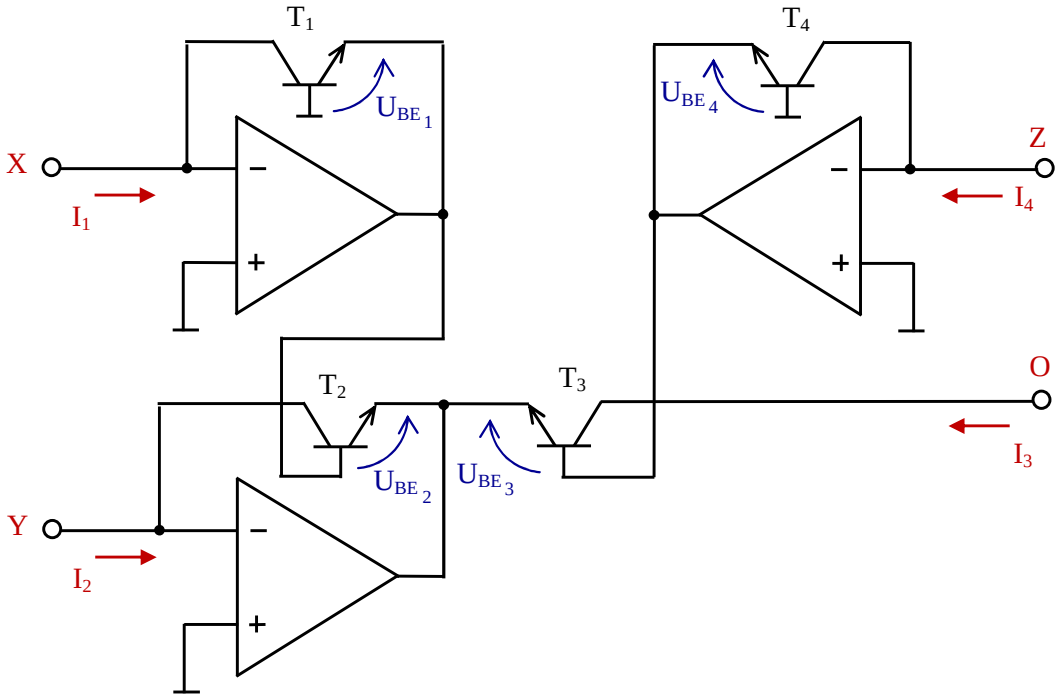


Fig. 8.31 A principle of multiplier/divider

The basic "structure relationship" is now:

$$U_{BE_1} + U_{BE_2} = U_{BE_3} + U_{BE_4}$$

Thus (see relations mentioned above):

$$U_T \cdot \ln\left(\frac{I_1}{I_o}\right) + U_T \cdot \ln\left(\frac{I_2}{I_o}\right) = U_T \cdot \ln\left(\frac{I_3}{I_o}\right) + U_T \cdot \ln\left(\frac{I_4}{I_o}\right);$$

$$U_T \cdot \ln\left(\frac{I_1 I_2}{I_o^2}\right) = U_T \cdot \ln\left(\frac{I_3 I_4}{I_o^2}\right);$$

$$I_1 I_2 = I_3 I_4 \quad \rightarrow \quad I_3 = \frac{I_1 I_2}{I_4}$$

This equation is „temperature independent“! ***We must realize that all input currents must be positive with respect to used arrows*** (always) – in an opposite case the structure will not function.

We can see a one-quadrant application in Fig. 8.32.

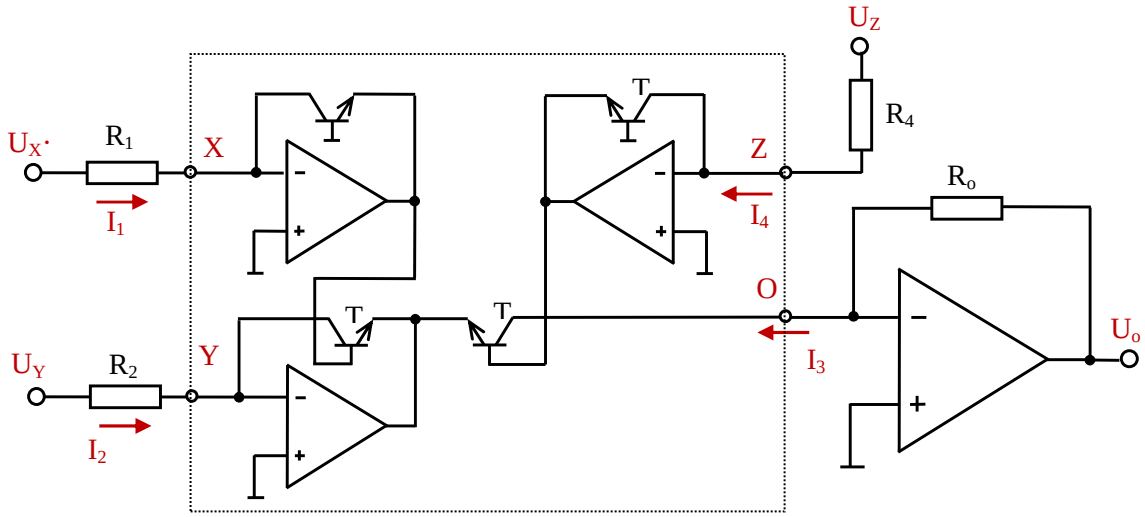


Fig. 8.32 A one – quadrant multiplier/divider

We may use positive input voltages only (one quadrant only). Then it is valid

$$I_1 = \frac{U_X}{R_1}$$

$$I_2 = \frac{U_Y}{R_2}$$

$$I_2 = \frac{U_Z}{R_4}$$

Hence

$$I_3 = \frac{I_1 I_2}{I_4} = \frac{U_X \cdot U_Y}{U_Z} \cdot \frac{R_4}{R_1 R_2}$$

and thus (the external OPA acts as current to voltage converter) the output voltage is

$$U_o = R_o I_3 = \frac{U_X \cdot U_Y}{U_Z} \cdot \frac{R_4 R_o}{R_1 R_2}$$

If both polarity inputs are needed, there imust be used more complex circuitry – see RC4200 data sheet for example.

8.7 Square root circuit; square circuit

A *square root circuit* we get quickly and easily from the circuit in Fig. 8.32 (one quadrant only) – see Fig. 8.33. We just interconnect U_Z and U_o.

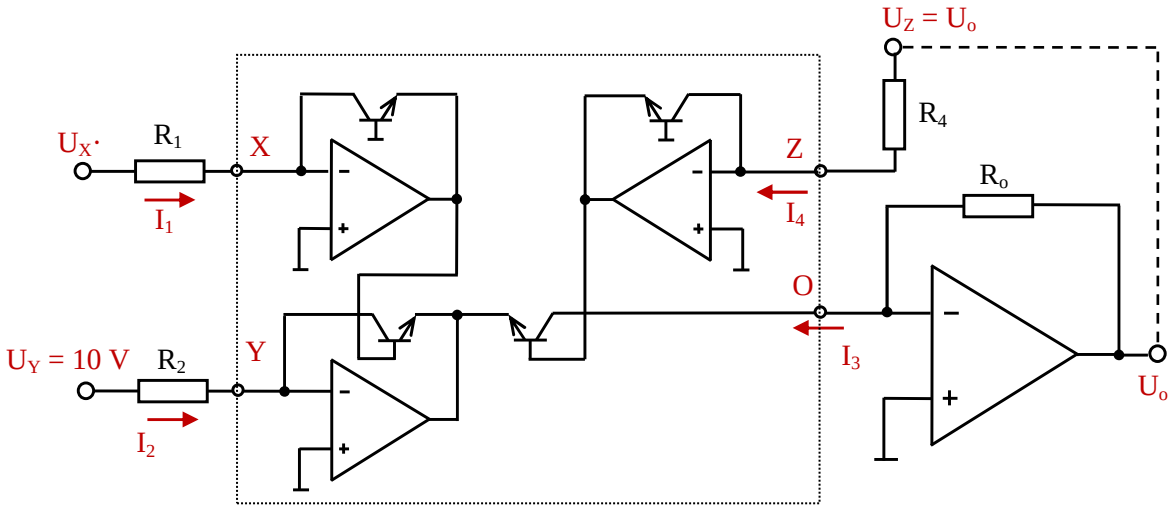


Fig. 8.33 A one – quadrant square root circuit

It is evident now that

$$U_o = \frac{U_X \cdot U_Y}{U_o} \cdot \frac{R_4 R_o}{R_1 R_2} = \Big|_{\text{all resistors are } R;\ U_Y = 10\ \text{V}} = \frac{U_X \cdot 10}{U_o}$$

Hence

$$U_o^2 = \Big|_{\text{all resistors are } R;\ U_Y = 10\ \text{V}} = U_X \cdot 10$$

Thus

$$U_o = \sqrt{10 \cdot U_X}$$

A ***square circuit*** we get quickly and easily from the circuit in Fig. 3.32 (one quadrant) – we simply interconnect inputs U_X and U_Y so that

$$U_o = \frac{U_X^2}{U_Z} \cdot \frac{R_4 R_o}{R_1 R_2} = \Big|_{\text{all resistors are } R;\ U_Z = 10\ \text{V}} = \frac{U_X^2}{10}$$

Summary

1) If an op amp does not saturate, recovery time is insignificant (an op amp´s recovery time measures how quickly it returns to a linear mode from a state of saturation).
2) Another source of error is a slew rate.
3) Simple DC voltage sources are in fact amplifiers (power) of a reference voltage.

4) If we need digitally controlled voltage source we can use suitable DAC.
5) DC current sources are in fact amplifiers (power) with negative current feedback.
6) Frequency stability of a logarithmic amplifier with transistor is problematic. The feedback (small signal) via the transistor (common base connection) is now considerably larger than one and dependent on the input voltage.
7) The basic analog multiplier relationship is $logA + \log B = \log(AB)$.
8) The basic analog divider relationship is $logA - \log B = \log(A/B)$.

Questions 8

To make sure that you have comprehended information in this chapter well in dept, we have a few theoretical questions for you here.

1. Describe a basic half-wave rectifier with an OPA.
2. Is it possible to think of a voltage source as an amplifier of reference voltage?
3. What type of feedback is used in current sources?
4. Describe small – signal conditions of a transistor logarithmic amplifier.
5. Describe basic principle of an analog multiplier.

Problems 8

Example 8.1

Analyze voltages of the circuit in Fig. 8.1 with an ideal OPA if : $U_{IN} = 10$ mV; 100 mV; 1 V; 5 V; suppose that diode drop is about 0, 65 V.

Example 8.2

Analyze voltages of the circuit in Fig. 8.1 with a no ideal OPA (OPA gain $\boldsymbol{A} = 20000$) if $\boldsymbol{U_{IN}} = \mathbf{10}$ mV; 100 mV; 1 V; 5 V; suppose that diode drop is about 0, 65 V.

Example 8.3

Determine output currents of the circuit in Fig. 8.18 with ideal OPAs and transistors (infinite current gain) if $U_{REF} = 100$ mV; 1 V; 5 V; suppose the same value of resistors R_1, R_2, R_3 and R_L – 1 kΩ for example.

Example 8.4

Determine output current and load voltages of the circuit in Fig. 8.18 with ideal OPAs and transistors (infinite current gain) if $U_{REF} = 1$ V; suppose the same value of resistors R_1, R_2, R_3 – 1 kΩ for example; and $R_L = 0$; 100; 1000 and 10 000 Ω.

Example 8.5

Determine (approximately) feedback from the output via T_1 in Fig. 8.25 if: $U_I = 10$ mV; 100 mV; 500 mV; 1 V; 5 V and 10 V.

Example 8.6

Determine output voltage of the circuit in Fig. 8.32 if $U_Y = 1$ V; 2 V and 5 V. We have $U_X = 2 + 1 \sin \omega t$; $U_Z = 10$ V; all resistors are equal – R.

Example 8.7

Determine output voltage (U_o) and allowable range of input voltages (U_X; U_Y) of the four - quadrant multiplier in Fig. 8.34 (remember – all input currents – I_1 to I_4 – of the multiplier/divider circuit must be positive).

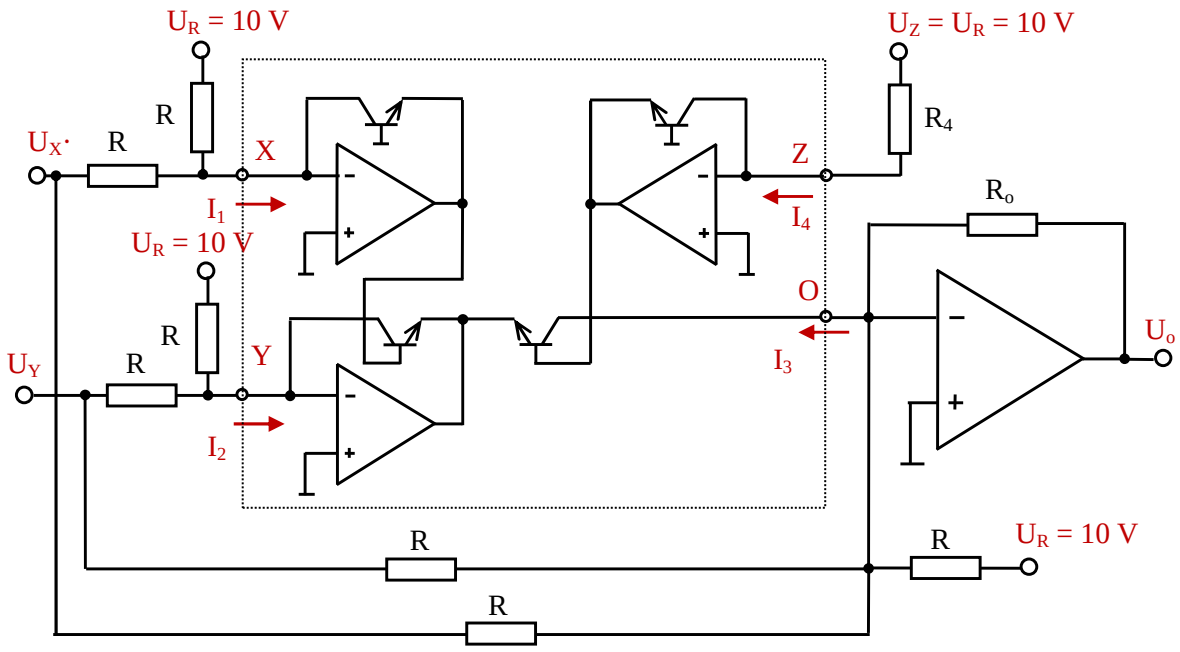


Fig. 8.34 A four – quadrant multiplier circuit

PROBLEMS KEY 8

Ad example 8.1)
Use findings in the Fig. 8.1

Ad example 8.2)
Use findings in the Fig. 8.2.

Ad example 8.3)
Use findings in the Fig. 8.18.

Ad example 8.4)
Use findings in the Fig. 8.18 and Ohm's law.

Ad example 8.5)
Use findings above Fig. 8.27.

Ad example 8.6)
We use findings:

$$U_o = R_o I_3 = \frac{U_X \cdot U_Y}{U_Z} \cdot \frac{R_4 R_o}{R_1 R_2} = \left| \text{all resistors are } R;\ U_Z = 10\ \text{V} \right| = \frac{U_X \cdot U_Y}{10} = \cdots$$

$$\cdots = \frac{(2 + \sin \omega t) \cdot U_Y}{10} = 0{,}2 \cdot U_Y + 0{,}1 \cdot U_Y \sin \omega t$$

We can say we control ac signal by voltage U_Y.

Ad example 8.7)
Remember – all inverting OPA inputs have zero voltages in our circuit – thus we can easy determine all currents – Fig. 8.35 (virtual ground).

It is evident that this easy situation gives:

$$I_1 = \frac{U_R}{R} + \frac{U_X}{R};$$

$$I_2 = \frac{U_R}{R} + \frac{U_Y}{R};$$

$$I_4 = \frac{U_R}{R};$$

$$I_3 = \frac{U_o}{R} + \frac{U_X}{R} + \frac{U_R}{R} + \frac{U_Y}{R}$$

Simultaneously is valid (if all resistors are equal)

$$I_3 = \frac{I_1 I_2}{I_4}$$

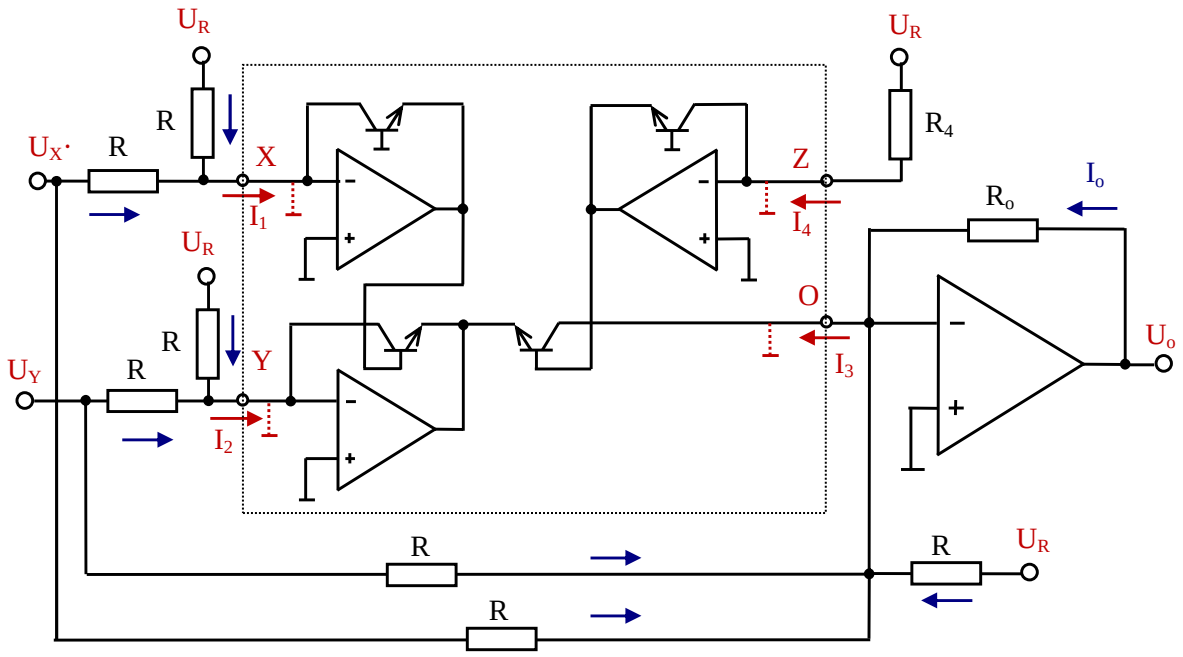


Fig. 8.35 A four – quadrant multiplier circuit – virtual zeroes and complete current depiction

Now we can determine

$$\frac{U_o}{R} = I_3 - \frac{U_X}{R} - \frac{U_R}{R} - \frac{U_Y}{R} = \frac{I_1 I_2}{I_4} - \frac{U_X}{R} - \frac{U_R}{R} - \frac{U_Y}{R} = \cdots$$

$$\cdots = \frac{\left(\frac{U_X}{R}+\frac{U_R}{R}\right)\cdot\left(\frac{U_R}{R}+\frac{U_Y}{R}\right)}{\frac{U_R}{R}} - \frac{U_X}{R} - \frac{U_R}{R} - \frac{U_Y}{R} = \frac{U_X}{R} + \frac{U_R}{R} + \frac{U_Y}{R} + \frac{U_X U_Y}{R U_R} - \frac{U_X}{R} - \frac{U_R}{R} - \frac{U_Y}{R}$$

Thus is valid

$$U_o = \frac{U_X U_Y}{U_R}$$

It is still the multiplier.

Always must be in this circuit:

$$I_1 = \frac{U_R}{R} + \frac{U_X}{R} > 0 \quad \text{thus} \quad \boxed{U_X > -U_R}$$

$$I_2 = \frac{U_R}{R} + \frac{U_Y}{R} > 0 \quad \text{thus} \quad \boxed{U_Y > -U_R}$$

We may use negative input voltages, too.

If we know that $I_{1,2_{max}} = 1\,\text{mA}$ and $U_{X,Y_{max}} = 10\,\text{V}$ for example, we determine needed resistor value:

$$I_{1,2_{max}} = \frac{U_R}{R} + \frac{U_{X,Y_{max}}}{R} = \frac{20}{R} \leq 10^{-3}$$

Thus we need

$\boxed{R > 20\ \mathrm{k\Omega}}$

We got this way ***four – quadrant multiplier***.

Basic texts

[1] Punčochář, J.: Operační zesilovače v elektronice. BEN – Praha, 2002, ISBN 80-7300-059-8

[2] Theory and Applications of Logarithmic Amplifiers. National Semiconductor, Application Note 311, April 1991

[3] Barrie, G.: Considering Multipliers (Part 1). Analog Dialogue 42 – 11, November 2008

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

9. Modulation, demodulation, signal sampling

Time of study: 6 hours

Goals: the student should be able to

- describe basic modulator principles
- describe basic demodulator principles
- describe basic signal sampling principles

EXPLANATION

9.1 Amplitude modulation circuits (AM), demodulation

Amplitude modulation of a sine or cosine carrier results in a variation of the carrier amplitude that is proportional to the amplitude of the modulating signal. A *modulating signal* should produce an AM wave of the form

$$S(t) = A_o[1 + m\{t\}] \cdot \cos 2\pi f_c t$$

where

$m\{t\}$ - modulating signal

$A\cos(2\pi f_c t)$ - carrier signal (A – amplitude; f_c – frequency; ϕ – phase)

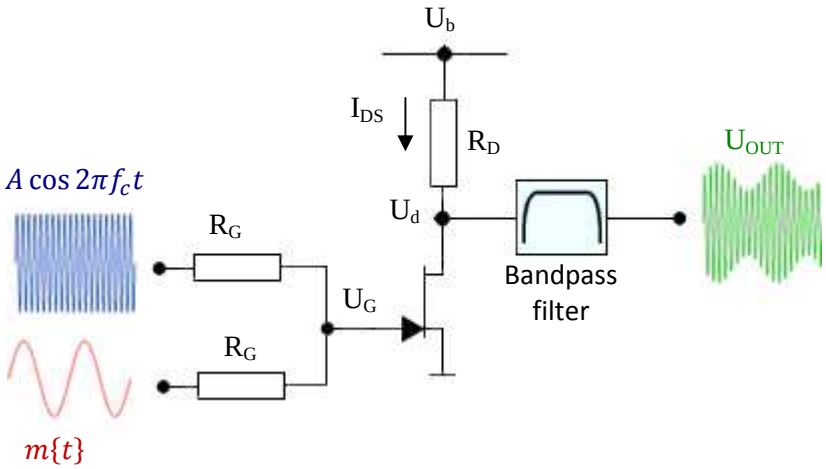


Fig. 9.1 Square law FET amplitude modulator – FET is nonlinear element

As an example let us describe *FET amplitude modulator* **– Fig. 9.1.**

A square law n-channel FET (**F**ield **E**ffect **T**ransistor) will pass a drain-source current

$$I_{DS} = K(U_G + U_P)^2$$

where

U_P is the FET's *pinch-off voltage* and U_G is the *gate voltage.* $U_P$is already absolute (positive) value here – we know that for N channel FETs is in the basic relationship negative sign and $U_P$is also negative; minus x minus gives plus sign.

It is evident that (superposition theorem)

$$U_P = \frac{m\{t\}+A\cos(2\pi f_c t)}{2}$$

thus

$$I_{DS} = \frac{K}{4}(m\{t\} + A\cos(2\pi f_c t) + U_P)^2$$

$$I_{DS} = \frac{K}{4}[m^2\{t\} + U_P^2 + 2U_P m\{t\}] + [m\{t\}A\cos(2\pi f_c t) + 2U_P \cdot A\cos(2\pi f_c t)] + \cdots$$

$$\cdots + \frac{K}{4}[A\cos(2\pi f_c t)]^2$$

This produces a drain voltage (output) of

$$U_D = U_b - R_D I_{DS}$$

We know that

$$\cos^2\alpha = \frac{1+\cos 2\alpha}{2}.$$

Hence the last term in expression for I_{DS} is a combination of a steady current and a fluctuation at the frequency $2f_c$. For simplicity we can arrange that the frequencies with which $m(t)$ fluctuate are all $\ll f_c$. This means that the first part of the expression consists of a steady current plus some fluctuations at frequencies well below f_c. We can now use a bandpass filter, designed to only pass frequencies f_c to strip away low and high frequencies and obtain an output

$$U_{OUT} = \frac{-R_D K}{4}[2m\{t\}A\cos(2\pi f_c t) + 2U_P \cdot A\cos(2\pi f_c t)] = \frac{-R_D K}{4}\left[\frac{m\{t\}}{U_P} + 1\right]\cos(2\pi f_c t)$$

which we can re-write in the form

$$U_{OUT} = A_0'[m'\{t\} + 1]\cos(2\pi f_c t)$$

where

$$m'\{t\} = \frac{m\{t\}}{U_P}; \quad A_0' = \frac{-R_D K U_P A}{2}$$

i.e. the output is a wave which unmodulated amplitude is A_0' and is amplitude modulated by an amount, $m'\{t\}$, proportional to the input modulating signal, $m(t)$. The circuit therefore behaves as an amplitude modulator.

If we have now:

$$m\{t\} = \sum_1^N a_m \cos(2\pi f_m t + \varphi_m)$$

then

$$m'\{t\} = \sum_{1}^{N} a'_m \cos(2\pi f_m t + \varphi_m)$$

$$a'_m = \frac{a_m}{U_p}$$

We get $\left(\cos\omega_1 t \cdot \cos\omega_2 t = \frac{\cos(\omega_1-\omega_2)+\cos(\omega_1+\omega_2)}{2}\right)$

CARRIER FR.

$$U_{out} = A'_o\left[1 + \sum_{1}^{N} a'_m \cos(2\pi f_m t + \varphi_m)\right] \cdot \cos(2\pi f_c t) = A'_o \cos(2\pi f_c t) + \cdots$$

$$\cdots + \frac{A'_o}{2}\sum_{1}^{N} a'_m \cos[2\pi(f_c + f_m)t + \varphi_m] + \frac{A'_o}{2}\sum_{1}^{N} a'_m \cos[2\pi(f_c - f_m)t - \varphi_m]$$

Upper Sideband – USB

Lower Sideband – LSB

....

You can see frequency spectrum of AM modulation in Fig. 9.2.

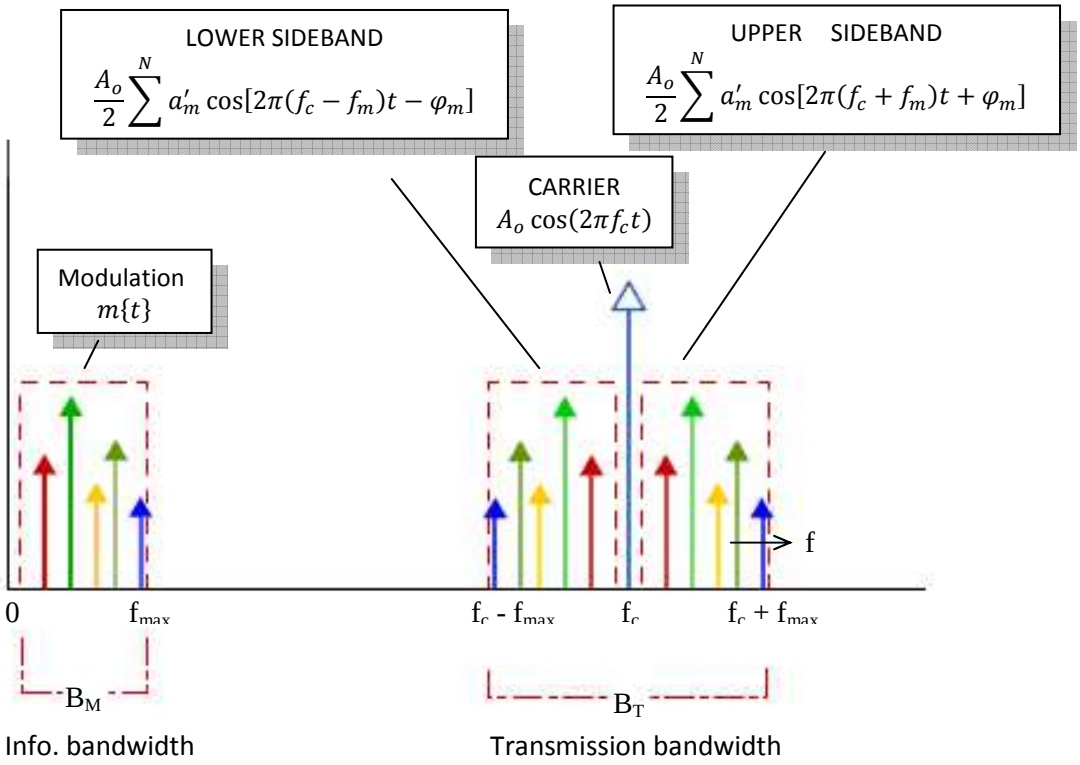


Fig. 9.2 Frequency spectrum of AM modulation

The carrier amplitude A must be large enough so that

$$A + m\{t\} \geq 0$$

at all times, or

$$A \geq |m\{t\}|_{MAX}$$

The extent of amplitude variation in AM about unmodulated carrier amplitude is measured in terms of factor called modulation index m (depth of modulation, degree of modulation, modulation factor). The baseband signal is preserved in the envelope of the AM signal if

$$A \geq |m\{t\}|_{MAX}$$

i.e. modulation factor

$$m = \frac{|m\{t\}|_{MAX}}{A} \leq 1$$

If $m > 1$, the signal is overmodulated (carrier phase inversion). This signal cannot be recovered well in most detection systems – Fig. 9.3.

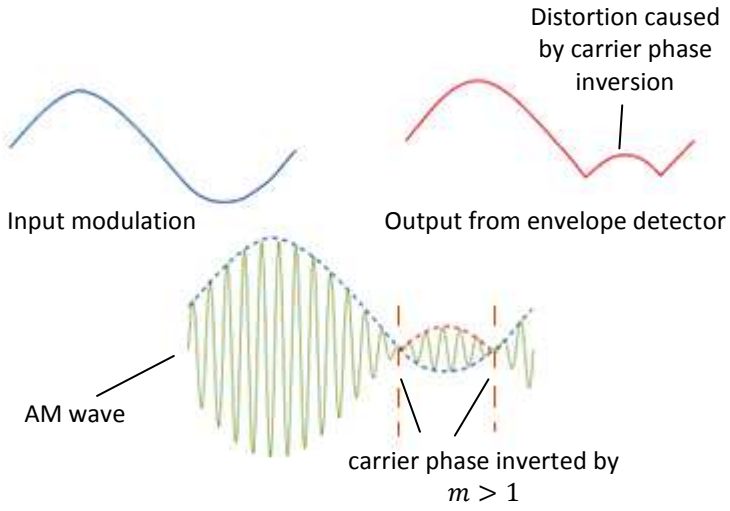


Fig. 9.3 Overmodulated AM signal

There are various ways to measure or *detect* the amplitude. We will consider one of the simplest, used by most portable radios, the ***Envelope Detector*** – Fig. 9.4.

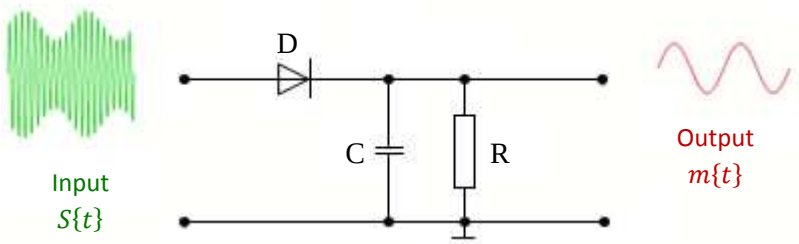


Fig. 9.4 Envelope detector

This is just a halfwave rectifier which charges a capacitor to a voltage ≈ to the peak voltage of the incoming AM waveform. When the input wave amplitude increases, the capacitor voltage is increased via the rectifying diode. When the input amplitude falls, the capacitor

voltage is reduced by being discharged by a 'bleed' resistor, R. The main advantage of this form of AM *Demodulator* is that it is very simple and cheap! It contains just one diode, and one capacitor, and one resistor. That is why it is used so often. However, it does suffer from some practical problems.

All real diodes are non-linear – the current they pass varies with the applied voltage – as a result, the demodulated signal is slightly distorted. This simple type of AM demodulator is not any good if we want the recovered waveform to be an accurate representation of the original modulating waveform (it is not Hi-Fi!!).

This circuit charges well the capacitor if the input voltage is greater than the capacitor voltage – it is the behavior of the diode. But this circuit blocks any current when the input voltage is below the capacitor voltage. The capacitor is discharged only via the resistor R - *Ripple* and *Negative Peak Clipping.* The ripple effect happens because the capacitor will be discharging a small amount in between successive peaks of the input AM wave – Fig. 9.5.

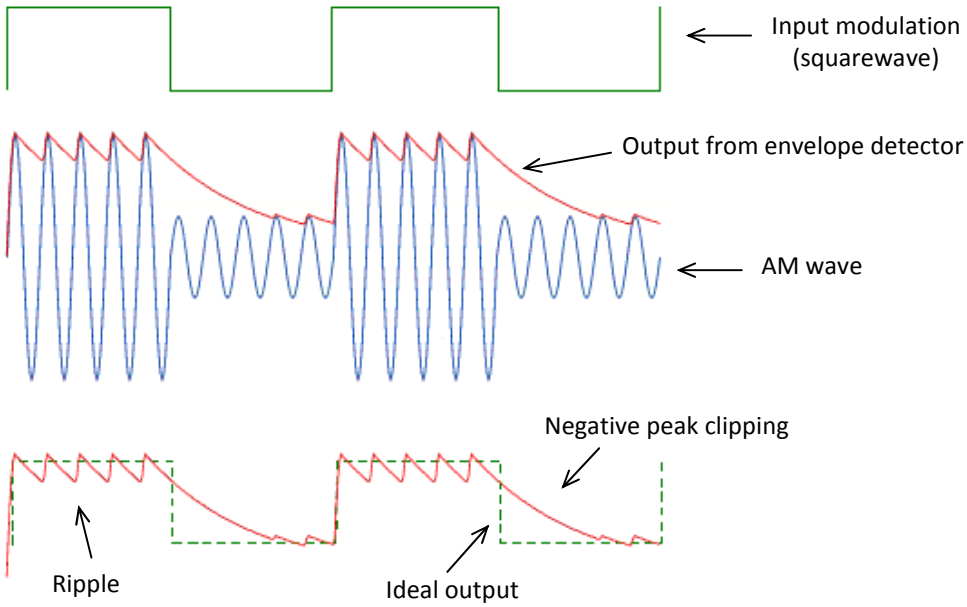


Fig. 9.5 Ripple and negative peak clipping effects

The illustration shows what happens in the worst possible situation where the modulating signal is a squarewave which frequency is not much lower than the carrier frequency.

The detector time constant (for discharging of capacitor) is $\tau = RC$. The time between successive peaks of the carrier will be $T = 1/f_c$. If we have $\tau \gg T$, we have discharging current between each peak and the next almost constant $\approx U_{max}/R$. Now we can determine the change of charge: $\Delta Q \approx T \cdot U_{max}/R$. Then the change of voltage between successive peaks (Ripple) is

$$\Delta U \approx \frac{\Delta Q}{C} \approx \frac{T \cdot U_{max}}{RC} = \frac{T \cdot U_{max}}{\tau} = \frac{U_{max}}{\tau \cdot f_c}$$

A sudden large reduction in the amplitude of the input AM wave means that capacitor charge isn't being 'topped up' by each cycle peak. The capacitor voltage therefore falls

exponentially until it reaches the new, smaller peak value. To assess this effect, consider what happens when the AM wave amplitude suddenly reduces from U_{max} to a much smaller value. The capacitor voltage then declines according to

$$U(t) = U_{max} \cdot exp^{-t/\tau}$$

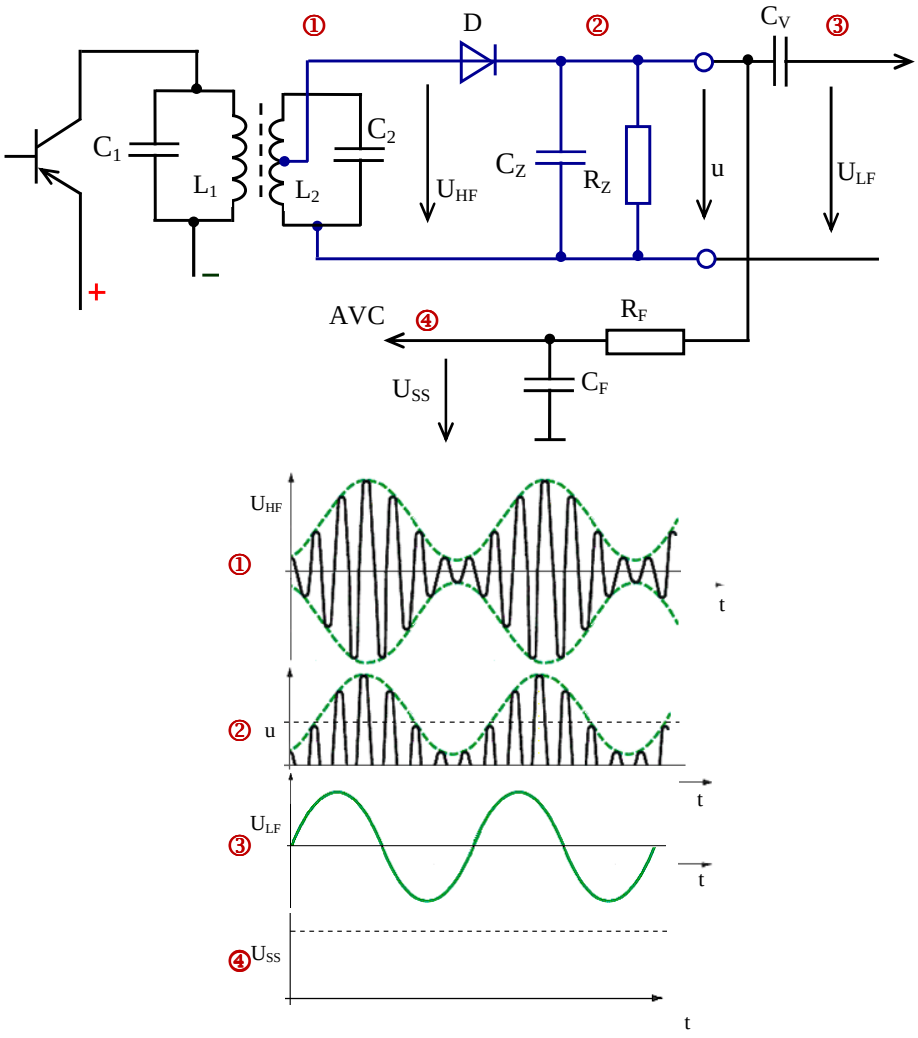


Fig. 9.6 Simple envelope detector with transistor amplifier – principle

This produces the negative peak clipping effect where any swift reductions in the AM wave amplitude are 'rounded off' and the output is distorted. Here we have chosen the worst possible case of squarewave modulation. In practice the modulating signal is normally restricted to a specific frequency range. This limits the maximum rate of falling of the AM wave amplitude. We can therefore hope to avoid negative peak clipping by arranging that the detector's time constant $\tau \ll t_m$ where

$$t_m = 1/f_m$$

and f_m is the highest modulation frequency used in a given situation.

The above implies that we can avoid negative peak clipping by choosing a small value of τ. However, to minimize the ripple we want to make τ as large as possible. In practice we should therefore choose a value

$$1/f_m \gg \tau \gg 1/f_c$$

to minimize the signal distortions caused by these effects. This is clearly only possible if the modulation frequency $f_m \ll f_c$. Envelope detectors only work satisfactorily when we ensure this inequality is true.

An example of connection with transistor amplifier is in Fig. 9.6.

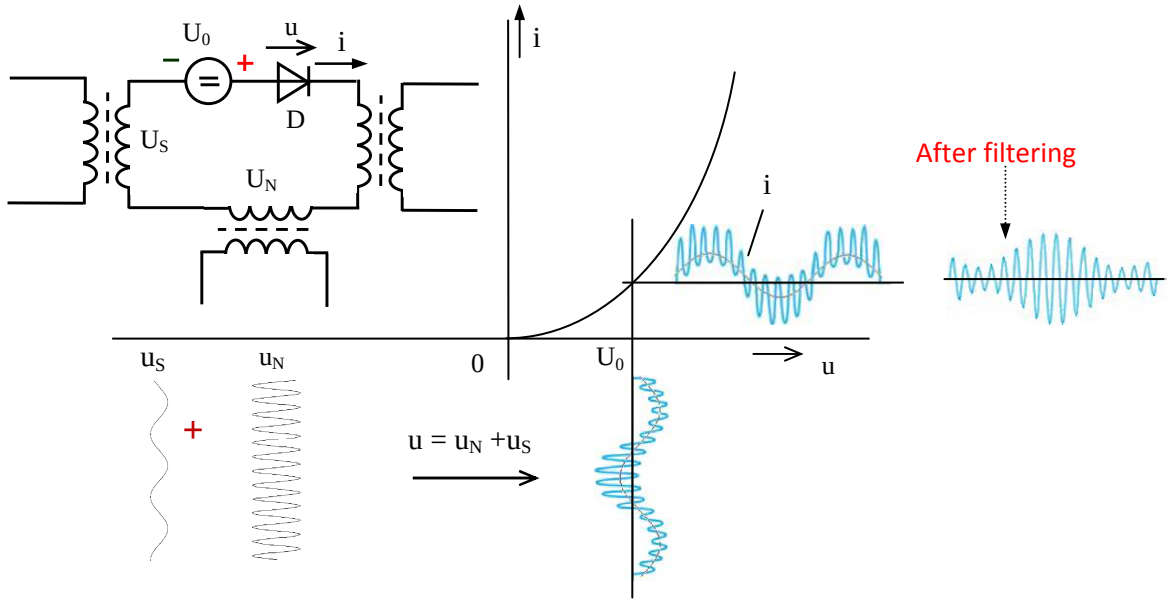


Fig. 9.7 Simple diode modulator – principle; u_N – carrier, u_s – modulating voltage, U_0 – sets quiescent point

Diode modulation consists of a mixing network, a diode rectifier, and an LC tuned circuit, often - Fig. 9.7. One diode is used as a nonlinear element – it "creates" needed frequency spectrum.

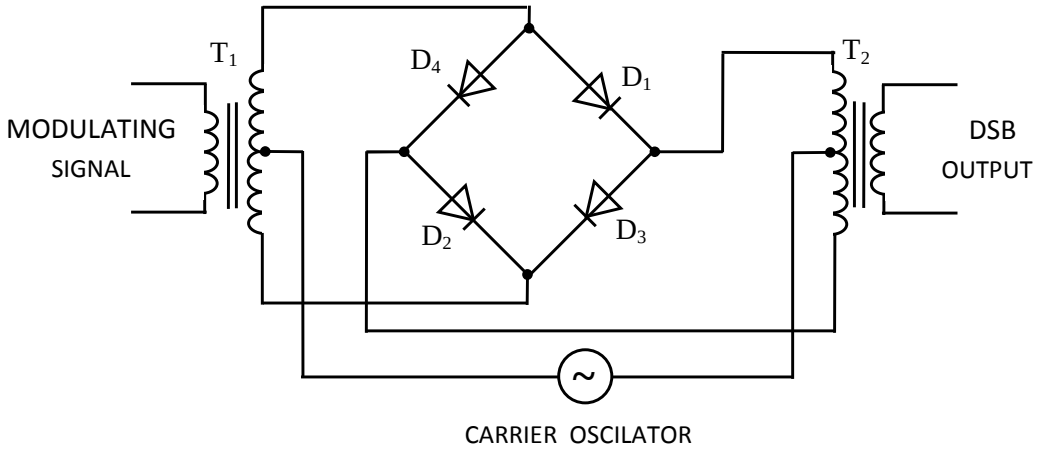


Fig. 9.8 Balanced (lattice) modulator - principle

A balanced modulator is a circuit that generates a DSB (double sideband) signal, suppressing the carrier and leaving only the sum and difference frequencies at the output - Fig. 9.8. The output of a balanced modulator can be further processed by filters or phase-shifting circuitry to eliminate one of the sidebands, resulting in a SSB (single sideband) signal.

The carrier sine wave is used as a source of forward and reverse bias for the diodes.

- The carrier turns the diodes off and on at a high speed rate.
- The diodes act like switches that connect the modulating signal at the secondary of T_1 to the primary of T_2.
- The carrier sine wave is considerably higher in frequency and amplitude than the modulating signal.

The AD633 can be used as a linear amplitude modulator with no external components. Figure 9.9 shows the circuit – principle. The carrier and modulation inputs to the AD633 are *multiplied to produce a double-sideband signal.* The carrier signal is fed forward to the AD633's Z input where it is summed with the double-sideband signal to produce a double-sideband with carrier output – Fig. 9.10.

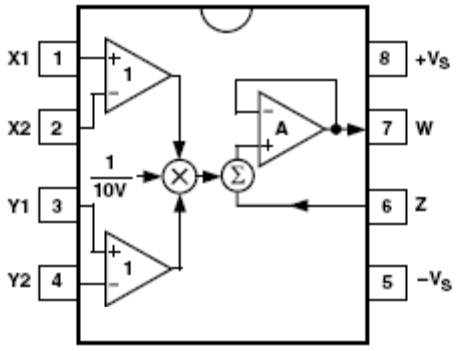


Fig. 9.9 AD633JN/AD633AN - a linear amplitude modulator with no external components

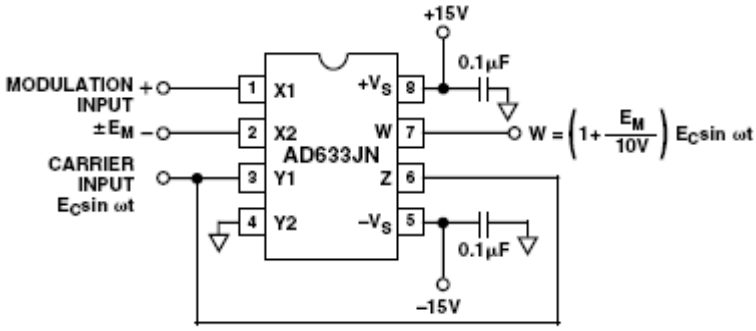


Fig. 9.10 Amplitude modulator with analog multiplier

It is evident that

$$\sin \omega t \cdot \sin \omega_M t = \frac{[\cos(\omega - \omega_M)t - \cos(\omega + \omega_M)t]}{2}$$

Thus all is clear, now.

9.2 Angular modulation – frequency (FM), and phase (PM)

The usual expression for a harmonic signal of angular frequency ω_C is

$$S(t) = \cos\big(\emptyset(t)\big) = \cos(\omega_C t + \varphi_o)$$

We define the instantaneous radian frequency ω_i to be the derivative of the angle $\emptyset(t)$ as a function of time:

$$\omega_i = \frac{d\emptyset}{dt} = \frac{d}{dt}(\omega_C t + \varphi_o)$$

If $\emptyset(t)$ varies in some manner with the modulating signal $f(t)$, the result is ***angular modulation.***

Phase modulation – we vary the phase of the carrier linearly with the modulation signal $m(t)$; k_1 is a constant of the system:

$$\emptyset(t) = \omega_C t + \varphi_o + k_1 \cdot m(t)$$

The phase of the carrier (PM) varies with the modulating signal (directly).

Frequency modulation – we wary the frequency of the carrier linearly with the modulation signal; k_2 is a constant of the system:

$$\omega(t) = \omega_C + k_2 \cdot m(t)$$

Thus

$$d\emptyset(t) = \omega_i \cdot dt \quad \Rightarrow$$

$$\emptyset(t) = \int_0^t [\omega_C + k_2 \cdot m(t)]dt = \omega_C t + k_2 \cdot \int_0^t m(t)dt + \varphi_o$$

The phase of the carrier (FM) varies with the integral of the modulating signal.

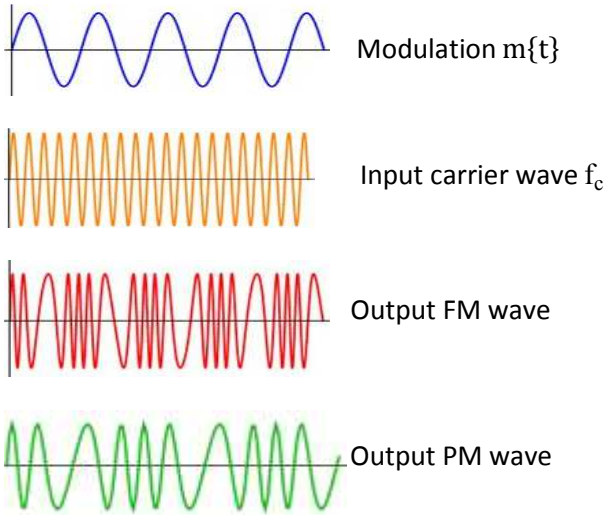


Fig. 9.11 FM and PM modulation

9.2.1 Frequency modulation

Let us analyze FM. We assume a harmonic modulation signal at the frequency f_m:

$$m(t) = a \cdot \cos(\omega_m t)$$

Let us have the instantaneous radian frequency ω_i

$$\omega_i = \omega_C + \Delta\omega_{PEAK} \cdot \cos(\omega_m t); \quad \Delta\omega_{PEAK} \ll \omega_C$$

$\Delta\omega_{PEAK}$ is a constant depending on the amplitude, a, of the modulating signal and on the properties of the modulating system. Now we easily get a relation

$$\emptyset(t) = \int \omega_i \, dt = \omega_C t + \frac{\Delta\omega_{PEAK}}{\omega_m} \cdot \sin(\omega_m t) + \varphi_o$$

We can take φ_o as zero. The frequency modulated carrier is then expressed by formula

$$S(t) = A \cdot \cos\left(\omega_C t + \frac{\Delta\omega_{PEAK}}{\omega_m} \cdot \sin(\omega_m t)\right) = A \cdot \cos(\omega_C t + m \cdot \sin(\omega_m t))$$

where

$m = \frac{\Delta\omega_{PEAK}}{\omega_m} = \frac{\Delta f_{PEAK}}{f_m}$ is the ***modulation index*** (representing the maximum phase shift of the carrier);

Δf_{PEAK} is the ***maximum frequency deviation*** of the carrier

For simplicity we assume that $m \ll \pi/2$ (usually $m < 0, 2$) – ***narrow band*** FM. We get

$$S(t) = A \cdot \cos(\omega_C t + m \cdot \sin(\omega_m t)) = \cdots$$

$$\cdots = A[\cos(\omega_C t)\cos(m \sin(\omega_m t)) - \sin(\omega_C t)\sin(m \sin(\omega_m t))]$$

For $m \to 0$ is true that $m \sin(\omega_m t) \to 0$. Thus we can use approximate formulas:

$$\cos(m \sin(\omega_m t)) \approx 1; \quad \sin(m \sin(\omega_m t)) \approx m \cdot \sin(\omega_m t)$$

Thus we get

$$S(t) = A \cdot [\cos(\omega_C t) - m \cdot \sin(\omega_c t) \cdot \sin(\omega_m t)]$$

Now we can get sideband form:

$$S(t) = A\cos(\omega_C t) + \frac{Am}{2}\cos(\omega_C + \omega_m)t - \frac{Am}{2}\cos(\omega_C - \omega_m)t$$

You can see spectra of AM and narrowband FM below.

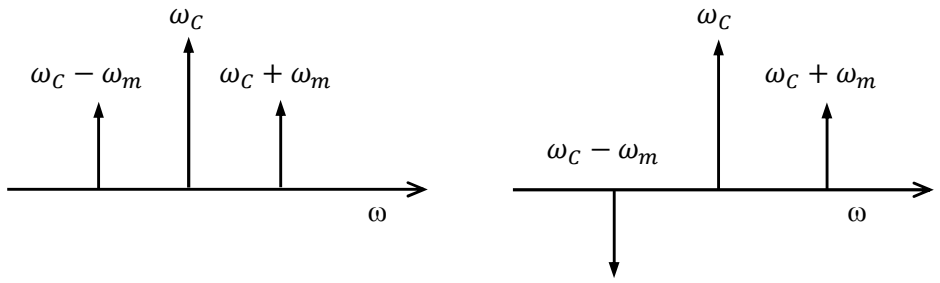


Fig. 9.12 Spectra of AM and narrowband FM

If m is not small – ***wideband FM*** – mathematic will be more complicated. We can find out that:

$$\cos(m\sin(\omega_m t)) = J_0(m) + 2\cdot J_2(m)\cos(2\omega_m t) + 2\cdot J_4(m)\cos(4\omega_m t) + \cdots$$

$$\sin(m\sin(\omega_m t)) = 2\cdot J_1(m)\sin(\omega_m t) + 2\cdot J_3(m)\sin(3t) + \cdots$$

where

$J_n(m)$ is the n[th] – order Bessel function of the first kind.

Then we get

$$S(t) = A\cos(\omega_C t + m\sin(\omega_m t)) =$$

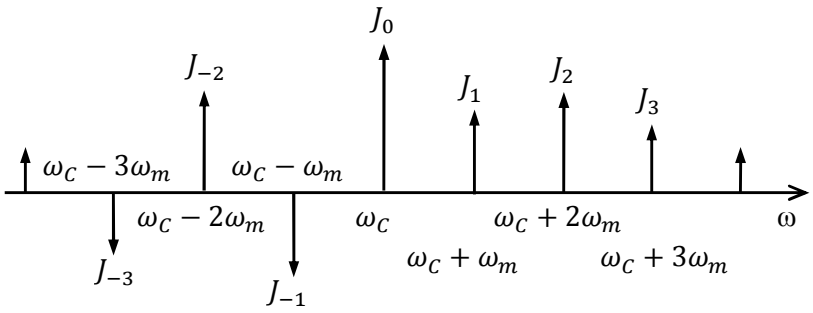
$$= A[\cos(\omega_C t)\cos(m\sin(\omega_m t)) - \sin(\omega_C t)\sin(m\sin(\omega_m t))] =$$

$$= J_0(m)\cos(\omega_c t) - J_1(m)[\cos(\omega_C - \omega_m)t - \cos(\omega_C + \omega_m)t] + \cdots$$

$$\cdots + J_2(m)[\cos(\omega_C - 2\omega_m)t - \cos(\omega_C + 2\omega_m)t] - \cdots$$

$$\cdots - J_3(m)[\cos(\omega_C - 3\omega_m)t - \cos(\omega_C + 3\omega_m)t] + \cdots$$

We thus have a time function consisting of a carrier and an infinite number of sidebands – amplitudes are proportional to $J_n(m)$. You can see the spectrum of the wideband FM in Fig. 9.13.



9.13 Spectrum of the wideband FM

The wideband FM wave has spectral components at all the frequencies, $\omega_C \pm n\omega_m$, where n can be any integer from $(-\infty)$ to $(+\infty)$. This result is rather startling since it means that, strictly speaking, a system has to provide an infinite bandwidth to carry an accurate FM (or PM) signal! Fortunately, there is a general tendency for $|J_n(m)| \to 0$ as $n \to \infty$. Moreover the higher order Bessel function values fall quickly with n when the modulation index m is small. In many practical situations we can arrange that and when this is true we find that

$$J_0(m) = 1; \quad J_1(m) \approx \frac{m}{2}; \quad \text{and} \quad J_n \approx 0 \ \text{ for } \ n \geq 1$$

We get narrowband FM again.

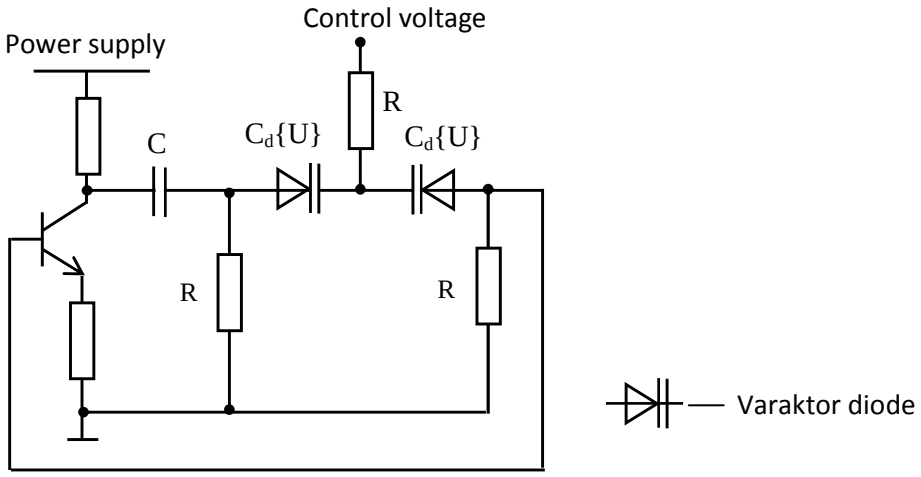
Narrowband FM only requires a bandwidth of $2f_m$. A "high-m; wideband" FM can be thought as a carrier whose frequency is varied over a relative wide range. It will therefore require a bandwidth of at least $2\cdot\Delta f_{PEAK}$

Combining these results with leads to *Carson's Rule*, that minimum practical bandwidth required to transmit an FM/PM signal will be

$$B = 2 \cdot (f_m + \Delta f_{PEAK})$$

This rule is a useful guide when we have to choose a system to carry an FM signal. It should be remembered, however, that in theory FM signals require an infinite bandwidth if we want to avoid any signal distortion during transmission.

As an example of FM modulator let us describe *Voltage Controlled Oscillator* (VCO) based upon the phase shift oscillator – Fig. 9.14.



9.14 Voltage Controlled Oscillator (VCO) based upon the phase shift oscillator.

In the VCO two of the original oscillator's capacitors have been replaced with *Varactor diodes*. When we apply a *Forward bias voltage* to a diode it will conduct. When we apply a *Reverse bias voltage* a *Depletion Zone* forms a junction at the diode. Charge carriers can't cross this zone, so the diode will not conduct in the 'reverse' direction. The reverse biased diode therefore has a capacitance of $C_d = \varepsilon A/d$, where d is the width of the depletion zone and A is the area of the each of diode junctions. Increasing the applied reverse bias pulls the carriers in the two halves apart, widening the depletion zone - the capacitance goes down. As a result the reverse biased diode has a capacitance which depends upon the applied voltage.

9.2.2 Basic demodulation FM principle

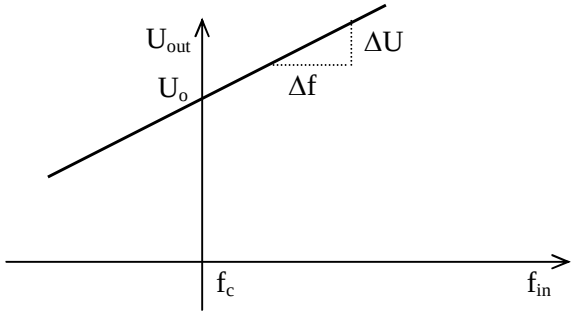
An FM demodulator or frequency discriminator is ***essentially a frequency-to-voltage converter (F/U)*** **– Fig. 9.15.** An F/U converter may be realised in several ways, including for example, tuned circuits and envelope detectors, phase locked loops *etc.* Demodulators are also called FM discriminators.

We define U_o as the output when $f_{in} = f_c$, the nominal input frequency. The gradient $K = \Delta U/\Delta f$ is called the voltage conversion factor. Then

$$U_{out} = U_o + K \cdot f_{in}$$

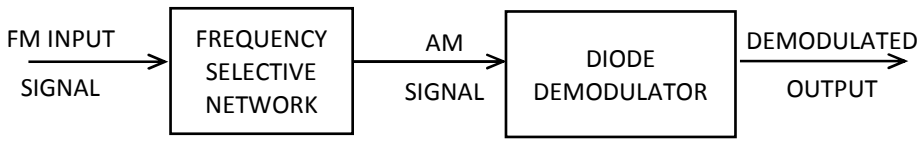
where

U_o represents a DC offset in U_{out}.

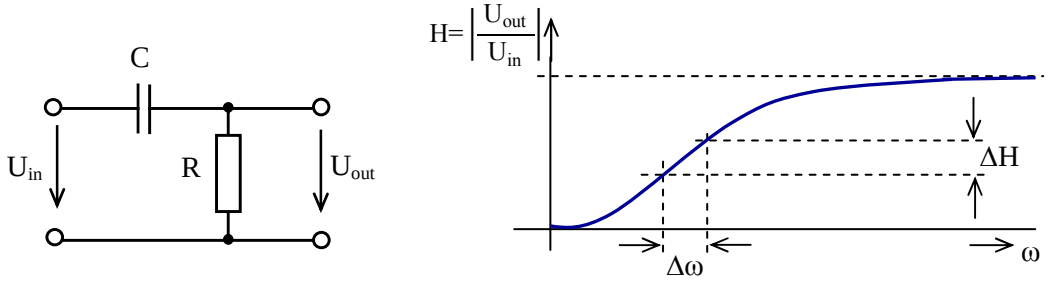


9.15 Basic characteristics of a F/U converter

The basic concept of FM demodulation is shown below – Fig. 9.16, Fig. 9.17. The FM input signal is input to a frequency selective network – it is any electronic circuit or element that gives a response (output) that varies with frequency. This circuit converts the frequency modulation into an amplitude modulation – it is then demodulated using any standard AM demodulator. There must be always ensured a constant amplitude of the FM input signal – input limiter is used. It eliminates the effects of any parasitic amplitude modulation of the FM signal.

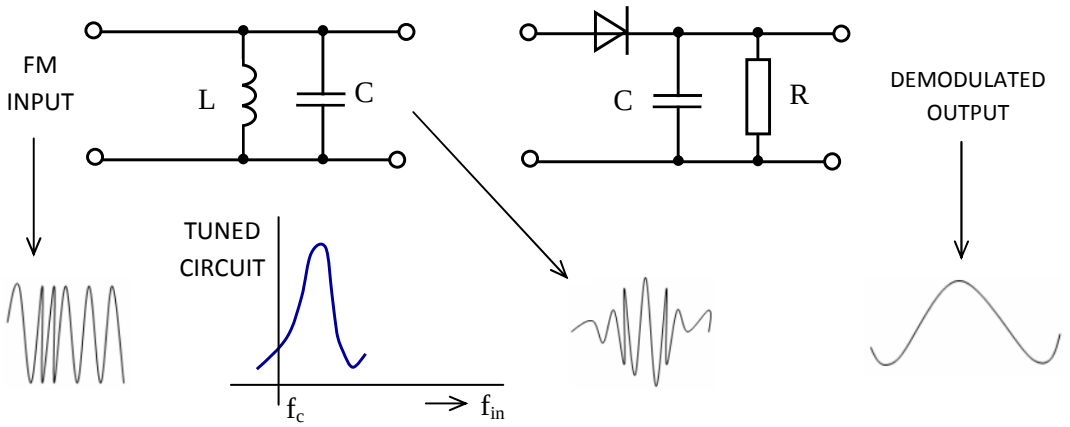


9.16 Basic principle of FM demodulation



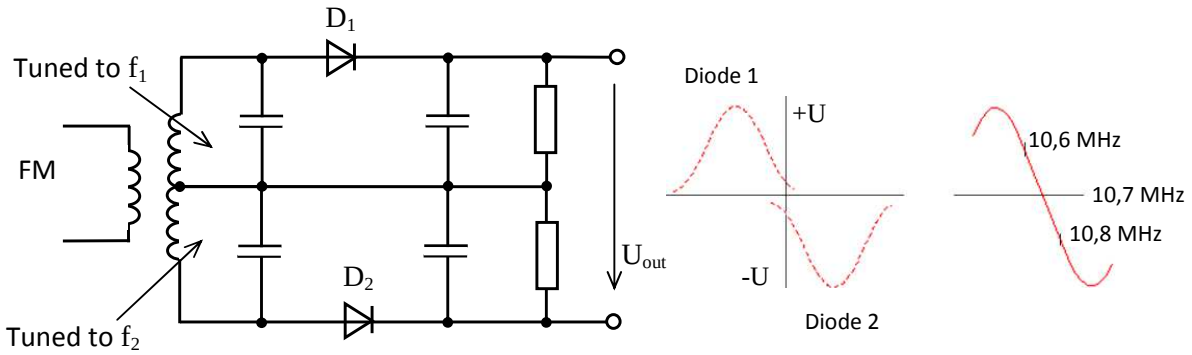
9.17 Simple RC high–pass filter

Tuned Circuit – One method (used in the early days of FM) is to use the slope of a tuned circuit in conjunction with an envelope detector.



9.18 Simple tuned circuit

The tuned circuit is tuned so that the f_c, the nominal input frequency, is on the slope, not at the centre of the tuned circuits. As the FM signal deviates about f_c on the tuned circuit slope, the amplitude of the output varies in proportion to the deviation from f_c. Thus the FM signal is effectively converted to AM. This is then envelope detected by the diode *etc* to recover the message signal. A better method is to use 2 similar circuits, this is known as a **Balanced Discriminator** – Fig. 9.19.



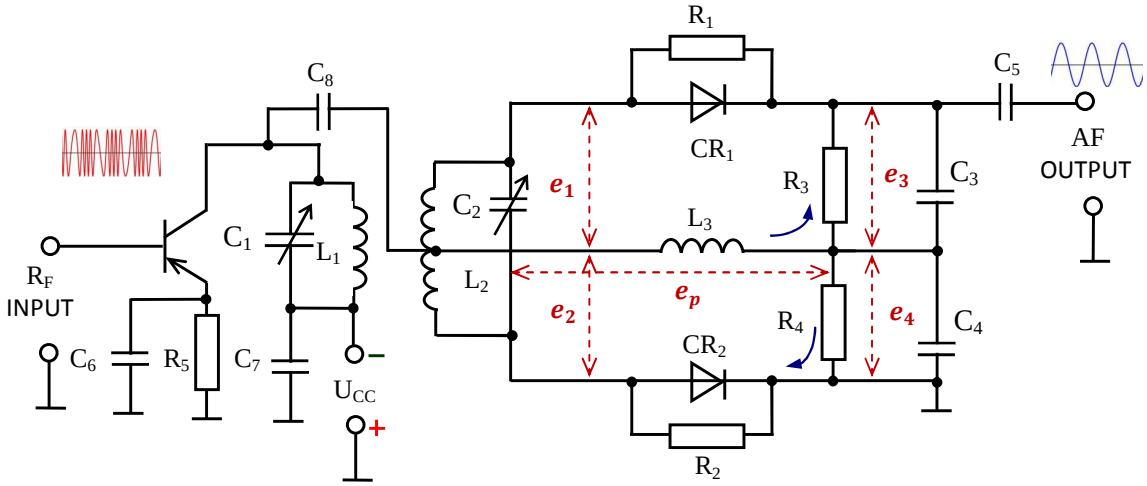
9.19 Balanced discriminator

9.2.3 Circuit Operation of a Foster-Seeley Discriminator

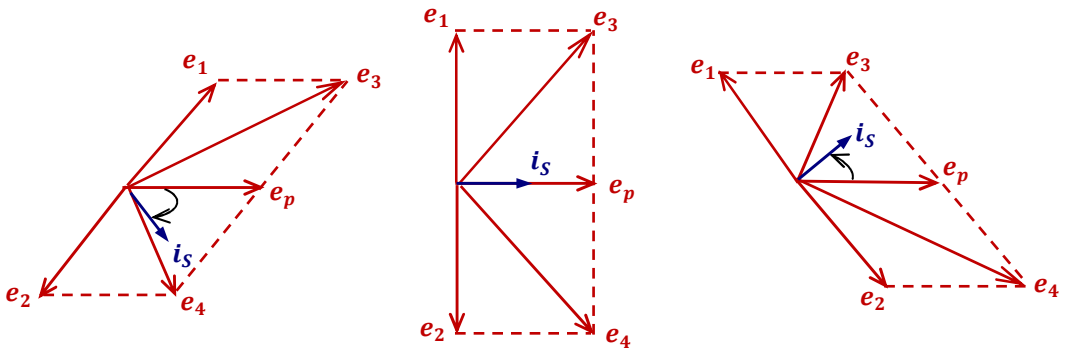
Fig. 9.20 shows a typical Foster-Seeley discriminator. The collector circuit of the preceding limiter/amplifier circuit (transistor) is shown. The limiter/amplifier circuit is a special amplifier circuit which limits the amplitude of the signal. This limiting keeps interfering noise low by removing excessive amplitude variations from signals. The collector circuit tank consists of C_1 and L_1. C_2 and L_2 form the secondary tank circuit (a tuned transformer circuit). Both tank circuits are tuned to the center frequency of the incoming FM signal. Choke L_3 is the DC return path for diode rectifiers CR_1 and CR_2. R_1 and R_2 are not always necessary but are usually used when the back (reverse bias) resistance of the two diodes is

different. Resistors R_3 and R_4 are load resistors and are bypassed by C_3 and C_4 to remove – RF (radio frequency). C_5 is the output coupling capacitor.

The operation of the Foster-Seeley discriminator can be best explained using vector diagrams – Fig. 9.21 – that show phase relations between the voltages and currents in the circuit. Let's look at the phase relations when the input frequency is equal to the center frequency of the resonant tank circuit. The input signal applied to the primary tank circuit is shown as vector e_p. Since coupling capacitor C_8 has negligible reactance at the input frequency (RF), choke L_3 is effectively in parallel with the primary tank circuit. Also, because L_3 is effectively in parallel with the primary tank circuit, input voltage e_p also appears across L_3. With voltage e_p applied to the primary of the tank circuit (L_1), a voltage is induced in the secondary one which causes current to flow in the secondary tank circuit. When the input frequency is equal to the center frequency, the tank is at resonance and acts resistive. Current and voltage are in phase in a resistance circuit, as shown by i_s and e_p. The current flowing in the tank causes voltage drops across each half of the balanced secondary winding of transformer. These voltage drops are of equal amplitude and opposite polarity with respect to the center tap of the winding. Because the winding is inductive, the voltage across it is 90 degrees out of phase with the current through it. Because of the center-tap arrangement, the voltages at each end of the secondary winding of the transformer are 180 degrees out of phase and are shown as e_1 and e_2 on the vector diagram. The voltage applied to the anode of CR_1 is the vector sum of voltages e_p and e_1, shown as e_3 on the diagram. Likewise, the voltage applied to the anode of CR_2 is the vector sum of voltages e_p and e_2 , shown as e_4 on the diagram. At resonance e_3 and e_4 are equal, as shown by vectors of the same length. Equal anode voltages on diodes CR_1 and CR_2 produce equal currents and, with equal load resistors, equal and opposite voltages will be developed across R_3 and R_4. The output is taken across R_3 and R_4 and will be 0 at resonance since these voltages are equal and of appositive polarity. The diodes conduct on opposite half cycles of the input waveform and produce a series of dc pulses at the rf rate. This rf ripple is filtered out by capacitors C_3 and C_4.



9.20 Forest - Seeley discriminator



9.21 Vectors of Forest - Seeley discriminator – a) operation above resonance, b) operation at resonance, c) operation below resonance

A phase shift occurs when an ***input frequency higher than the center frequency*** is applied to the discriminator circuit and the current and voltage phase relationships change. When a series-tuned circuit operates at a frequency above resonance, the inductive reactance of the coil increases and the capacitive reactance of the capacitor decreases. Above resonance the tank circuit acts like an inductor. Secondary current lags the primary tank voltage, e_p. Notice that secondary voltages e_1 and e_2 are still 180 degrees out of phase with the current (i_S) that produces them. The change to a lagging secondary current rotates the vectors in a clockwise direction. This causes e_1 to become more in phase with e_p while e_2 is shifted further out of phase with e_p. The vector sum of e_p and e_2 is less than that of e_p and e_1. Above the center frequency, diode CR_1 conducts more than diode CR_2. Because of this heavier conduction, the voltage developed across R_3 is higher than the voltage developed across R_4; the output voltage is positive.

When the ***input frequency is lower than the center frequency***, the current and voltage phase relationships change. When the tuned circuit is operated at a frequency lower than resonance, the capacitive reactance increases and the inductive reactance decreases. Below resonance the tank acts like a capacitor and the secondary current leads primary tank voltage e_p. This change to a leading secondary current rotates the vectors in a *counterclockwise* direction. From the vector diagram you should see that e_2 is brought nearer in phase with e_p, while e_1 is shifted further out of phase with e_p. The vector sum of e_p and e_2 is larger than that of e_p and e_1. Diode CR_2 conducts more than diode CR_1 below the center frequency. The voltage drop across R_4 is larger than that across R_3 and the output across both is negative.

Disadvantages **- t**hese voltage outputs can be plotted to show the response curve of the discriminator discussed earlier. When weak AM signals (too small in amplitude to reach the circuit limiting level) pass through the limiter stages, they can appear in the output. These unwanted amplitude variations will cause primary voltage e_p to fluctuate with the modulation and to induce a similar voltage in the secondary of the transfomer. Since the diodes are connected as half-wave rectifiers, these small AM signals will be detected as they would be in a diode detector and will appear in the output. This unwanted AM interference

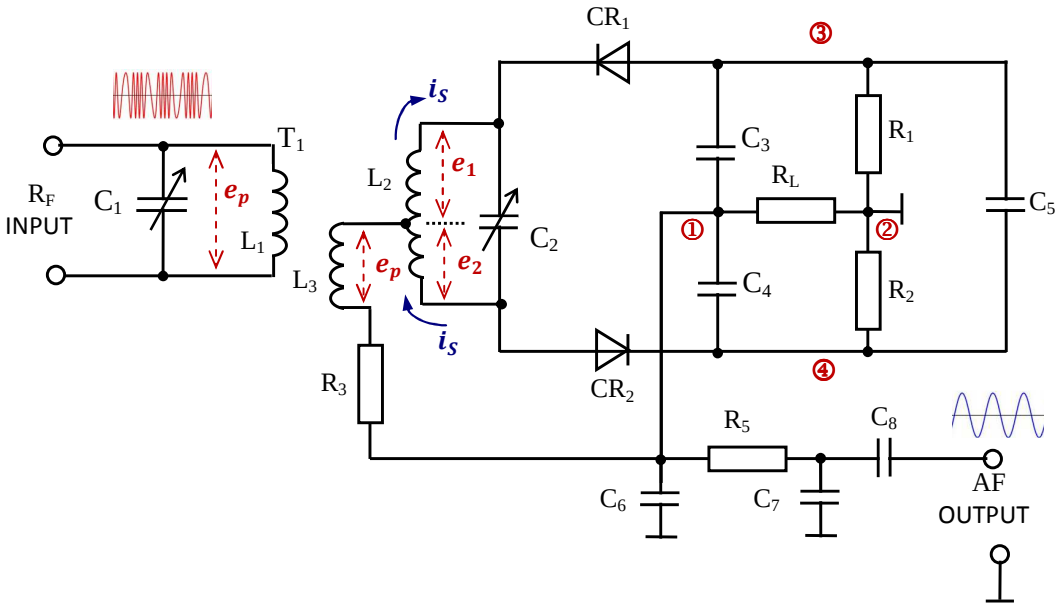
is cancelled out in the ratio detector (to be studied next in this chapter) and is the main disadvantage of the Foster-Seeley circuit.

9.2.4 Ratio detector

It uses a double-tuned transformer to convert instantaneous frequency variations of the fm input signal to instantaneous amplitude variations – Fig. 9.22. These amplitude variations are then rectified to provide a DC output voltage which varies in amplitude and polarity with the input signal frequency. This detector demodulates fm signals and suppresses amplitude noise without the need of limiter stages.

Circuit Operation

The input tank capacitor (C_1) and the primary of transformer T_1 (L_1) are tuned to the center frequency of the fm signal to be demodulated. The secondary winding of T_1 (L_2) and capacitor C_2 also form a tank circuit tuned to the center frequency. Tertiary (third) winding L_3 provides additional inductive coupling which reduces the loading effect of the secondary on the primary circuit. Diodes CR_1 and CR_2 rectify the signal from the secondary tank. Capacitor C_5 and resistors R_1 and R_2 set the operating level of the detector. Capacitors C_3 and C_4 determine the amplitude and polarity of the output. Resistor R_3 limits the peak diode current and furnishes a DC return path for the rectified signal. The output of the detector is taken from the common connection between C_3 and C_4. Resistor R_L is the load resistor. R_5, C_6, and C_7 form a low-pass filter to the output.



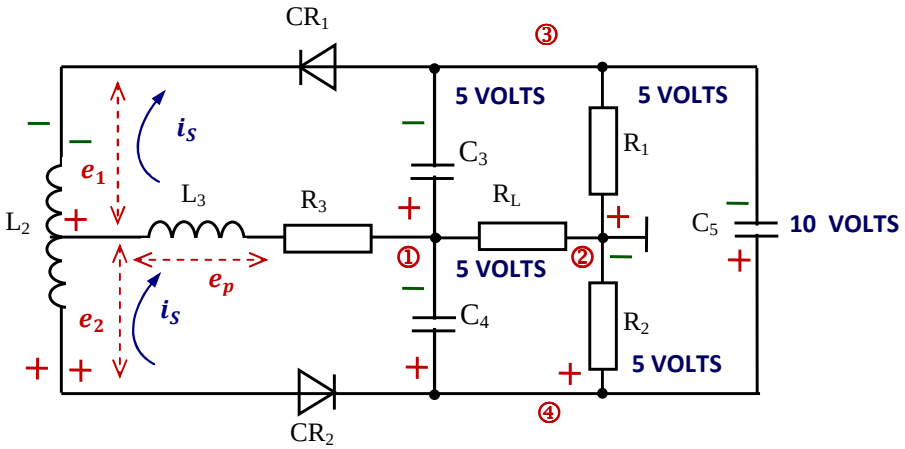
9.22 Ratio detector

This circuit operates on the same principles of phase shifting as did the Foster-Seeley discriminator. In that discussion, vector diagrams were used to illustrate the voltage amplitudes and polarities for conditions at resonance, above resonance, and below resonance. The same vector diagrams apply to the ratio detector but will not be discussed

here. Instead, you will study the resulting current flows and polarities on simplified schematic diagrams of the detector circuit.

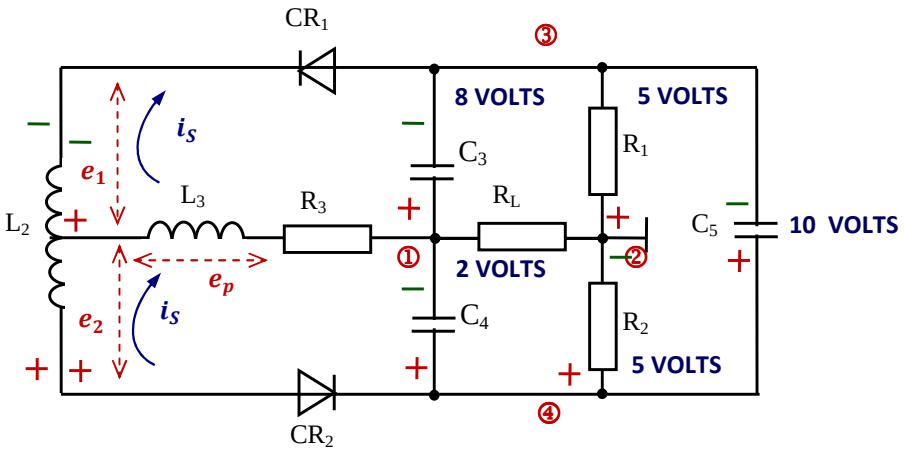
When the input voltage e_p is applied to the primary in Fig. 9.22 it also appears across L_3 because, by inductive coupling, it is effectively connected in parallel with the primary tank circuit. At the same time, a voltage is induced in the secondary winding and causes current to flow around the secondary tank circuit. At resonance the tank acts like a resistive circuit; that is, the tank current is in phase with the primary voltage e_p. The current flowing in the tank circuit causes voltages e_1 and e_2 to be developed in the secondary winding of T_1. These voltages are of equal magnitude and of opposite polarity with respect to the center tap of the winding. Since the winding is inductive, the voltage drop across it is 90 degrees out of phase with the current through it.

Figure 9.23 is a simplified schematic diagram of a ratio detector ***at resonance***. The voltage applied to the cathode of CR_1 is the vector sum of e_1 and e_p. Likewise, the voltage applied to the anode of CR_2 is the vector sum of e_2 and e_p. No phase shift occurs at resonance and both voltages are equal. Both diodes conduct equally. This equal current flow causes the same voltage drop across both R_1 and R_2. C_3 and C_4 will charge to equal voltages with opposite polarities. Let's assume that the voltages across C_3 and C_4 are equal in amplitude (5 volts) and of opposite polarity and the total charge across C_5 is 10 volts. R_1 and R_2 will each have 5 volts dropped across them because they are of equal values. The output is taken between points ① and ② - ground. To find the output voltage, you algebraically add the voltages between points ① and ② (loop ①③② or ①④②). Point ① to point ④ is -5 volts. Point ④ to point ② is $+5$ volts. Their algebraic sum is 0 volts and the output voltage is 0 at resonance. If the voltages on branch ①③② were figured, the same output would be found because the circuit branches are in parallel.



9.23 Current flow and polarities at resonance

When the input signal reverses polarity, the secondary voltage across L_2 also reverses. The diodes will be reverse biased and no current will flow. Meanwhile, C_5 retains most of its charge because of the long time constant offered in combination with R_1 and R_2. This slow discharge helps to maintain the output.



9.24 Current flow and polarities above resonance

When a tuned circuit operates at a ***frequency higher than resonance***, the tank is inductive – see fig. 9.24. The secondary current i_s lags the primary voltage e_p. Secondary voltage e_1 is nearer in phase with primary voltage e_p, while e_2 is shifted further out of phase with e_p. The vector sum of e_1 and e_p is larger than that of e_2 and e_p. Therefore, the voltage applied to the cathode of CR_1 is higher than the voltage applied to the anode of CR_2 above resonance.

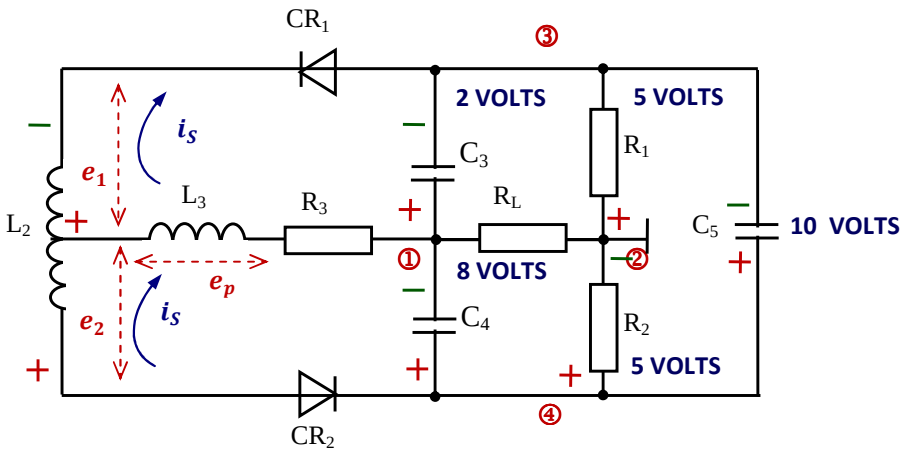
Assume that the voltages developed above resonance are such that the higher voltage on the cathode of CR_1 causes C_3 to charge to 8 volts. The lower voltage on the anode of CR_2 causes C_4 to charge to 2 volts. Capacitor C_5 remains charged to the sum of these two voltages, 10 volts. Again, by adding the voltages in loop ①③② or ①④② between points ① and ②, you can find the output voltage. Point ① to point ④ equals -2 volts. Point ④ to point ② equals $+5$ volts. Their algebraic sum, and the output, equals $+3$ volts when tuned above resonance. During the negative half cycle of the input signal, the diodes are reverse biased and C_5 helps maintain a constant output.

When a tuned circuit operates ***below resonance*** (Fig.9.25), it is capacitive. Secondary current i_s leading the primary voltage e_p and the secondary voltage e_2 is nearer in phase with primary voltage e_p. The vector sum of e_2 and e_p is larger than the sum of e_1 and e_p. The voltage applied to the anode of CR_2 becomes greater than the voltage applied to the cathode of CR_1 below resonance.

Assume that the voltages developed below resonance are such that the higher voltage on the anode of CR_2 causes C_4 to charge to 8 volts. The lower voltage on the cathode of CR_1 causes C_3 to charge to 2 volts. Capacitor C_5 remains charged to the sum of these two voltages, 10 volts. The output voltage equals -8 volts plus $+5$ volts, or -3 volts, when tuned below resonance. During the negative half cycle of the input signal, the diodes are reverse biased and C_5 helps to maintain a constant output.

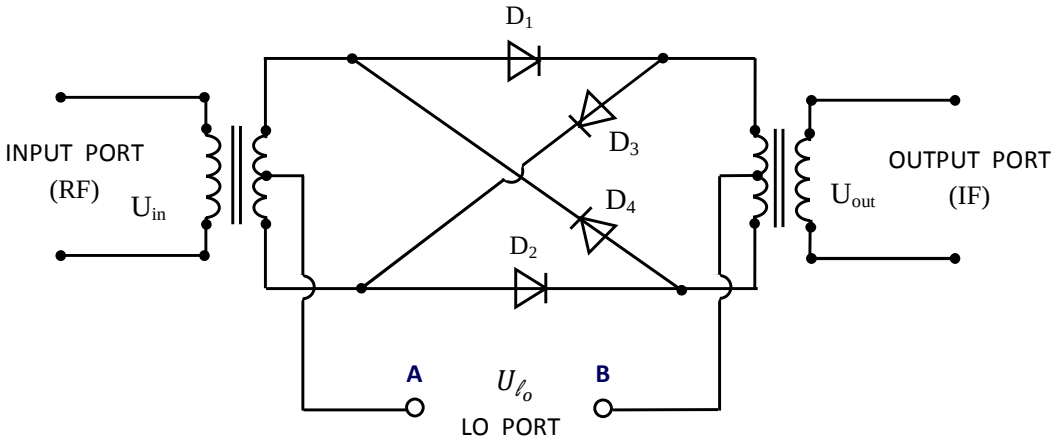
Advantage of a Ratio Detector - the ratio detector is not affected by amplitude variations on the FM wave. The output of the detector adjusts itself automatically to the average amplitude of the input signal. C_5 charges to the sum of the voltages across R_1 and R_2 and

because of its time constant, it tends to filter out any noise impulses. Before C_5 can charge or discharge to the higher or lower potential, the noise disappears. The difference in charge across C_5 is so slight that it is not discernible in the output. Ratio detectors can operate with as little as 100 millivolts of input. This is much lower than that required for limiter saturation and less gain is required from preceding stages.



9.25 Current flow and polarities below resonance

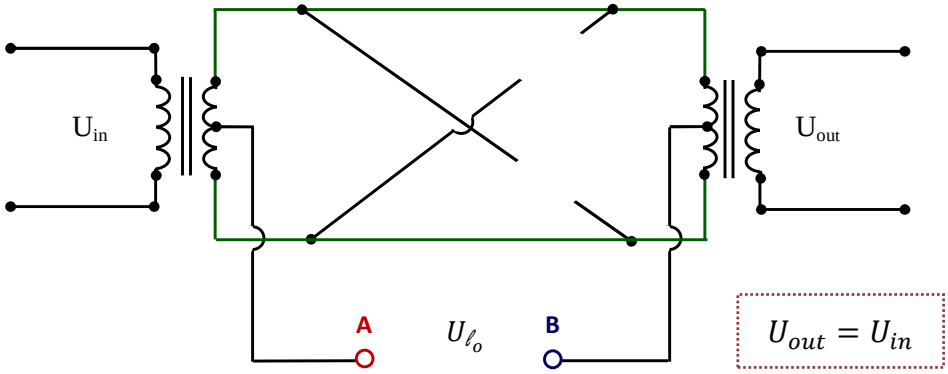
Double Balanced Mixer (DBM) as a ***phase detector*** is often used for FM/PM demodulation. Let us describe it once again. It consists of four diodes linking two transformers. There are three ways or *ports* by which signals can get in/out of the DBM. It is conventional to call these the RF, LO – *local oscillator*, and *IF* ports since DBM's are often used as mixers in heterodyne systems. To see how the system works we can start by considering the LO port. This lets us apply a voltage between the points A and B in the circuit – Fig. 9.26.



9.26 Double balanced mixer

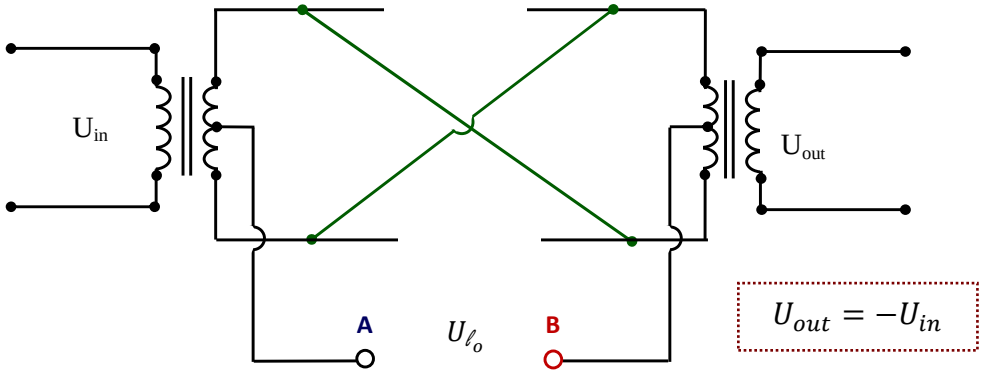
When we apply a voltage which makes A positive with respect to B (which we will assume occurs when $U_{\ell_o} \geq 0$) the diodes, D_1, D_2 will conduct and D_3, D_4 will not. This means that D_1 and D_2 will present a very low resistance to any signals and D_3, D_4 will present a very high resistance. As a result, the circuit will behave like below – Fig. 9.27. The two transformers

will be 'directly' connected together via the conducting diodes. Using 1:1 transformers and low-loss diodes we therefore find that, when $U_{\ell_o} \geq 0$ - is positive, $U_{out} = U_{in}$ (for a.c. signals, of course!).



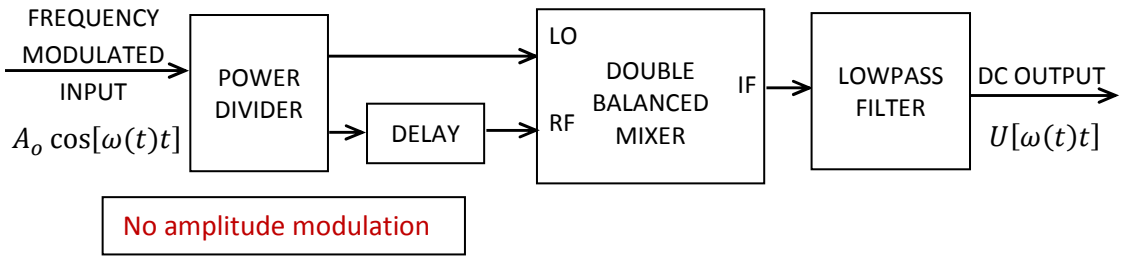
9.27 A positive referred to B

When we apply an LO voltage such that A is negative compared to B (Fig. 9.28), we can use a similar argument to show that, in this case, the circuit behaves like below. Hence when $U_{\ell_o} \leq 0$ - is negative, $U_{out} = -U_{in}$. We can therefore use the DBM as a sort of 'switch' to control whether we pass on the signal unaffected or invert it by applying the required polarity of $U_{\ell o}$.

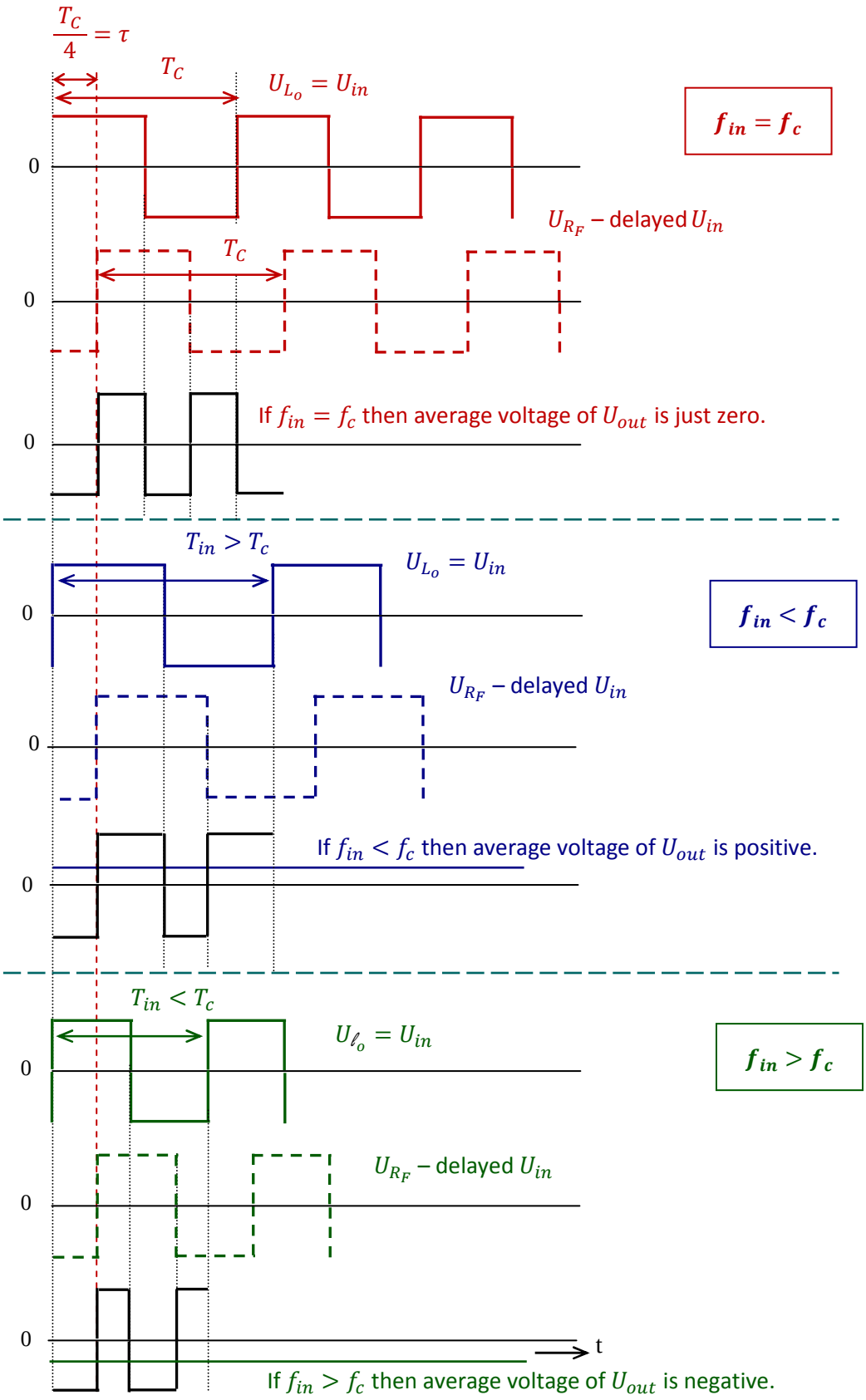


9.28 A negative referred to B

There is an example below – Fig. 9.29.



9. 29 Frequency discriminator based om a double-balanced mixer



9.30 Principle of function of DBM in Fig. 9.29

Box "delay" defines group delay (time delay) ***on carrier frequency*** ω_c: it must be just

$$\tau \cdot \omega_c = \frac{\pi}{2} \quad \Rightarrow \quad \tau = \frac{\pi}{2 \cdot \omega_c} = \frac{\pi}{2 \cdot 2\pi f_c} = \frac{T_c}{4}$$

In the ideal case we suppose ideally limited signals at inputs LO and R_F. Then we can easily derive Fig. 9.30.

Let us suppose that U_{out} takes values ± 1, only. It is evident it is enough to investigate just one half of the period.

$$U_{out} = -1 \quad \text{in the time interval} \quad \tau = \frac{\pi}{2 \cdot \omega_c}$$

$$U_{out} = +1 \quad \text{in the time interval} \quad \frac{T_{in}}{2} - \tau = \frac{T_{in}}{2} - \frac{\pi}{2 \cdot \omega_c}$$

Now we can easy determine the average value in the output:

$$U_{out_{AV}} = \frac{(-1) \cdot \frac{\pi}{2 \cdot \omega_c} + \left(\frac{T_{in}}{2} - \frac{\pi}{2 \cdot \omega_c}\right)}{\frac{T_{in}}{2}} = \left| \omega_{in} = 2\pi f_{in} = \frac{2\pi}{T_{in}} \Rightarrow T_{in} = \frac{2\pi}{\omega_{in}} \right|$$

$$U_{out_{AV}} = \frac{\frac{-\pi}{2 \cdot \omega_c} + \left(\frac{\pi}{\omega_{in}} - \frac{\pi}{2 \cdot \omega_c}\right)}{\frac{\pi}{\omega_{in}}} = 1 - \frac{\omega_{in}}{\omega_c}$$

If there is valid $\omega_{in} = \omega_c \pm \Delta\omega$ we get formula

$$U_{out_{AV}} = 1 - \frac{\omega_{in}}{\omega_c} = 1 - \frac{\omega_c \pm \Delta\omega}{\omega_c} = \pm \Delta\omega$$

This average voltage is separated by means of the low pass filter.

Delay circuit converts FM modulation into PM but preserves FM.

If we can use some multiplier, we can construct circuit as in Fig. 9.31 – **quadrature demodulator:**

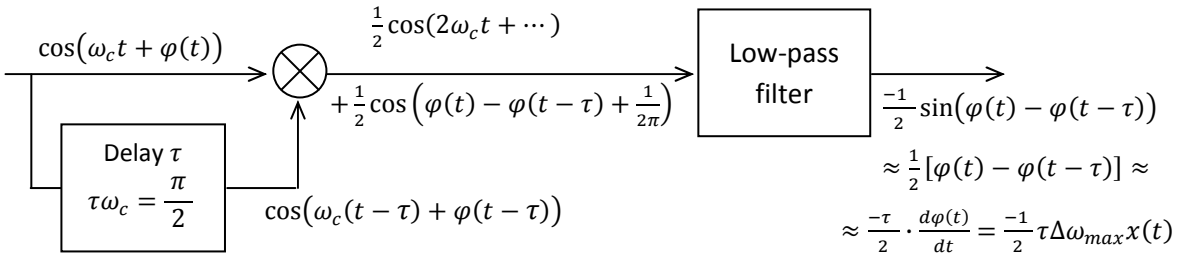


Fig 9. 31 Quadrature demodulator

The constant delay gives a phase shift of π/2 (quadrature) at the carrier frequency. The difference term from the multiplication

$$\cos(\omega_c t + \varphi(t)) \cos[\omega_c(t-\tau) + \varphi(t-\tau)] = \frac{\cos\{[\omega_c t + \varphi(t)] + [\omega_c(t-\tau) + \varphi(t-\tau)]\}}{2} + \cdots$$

$$\cdots + \frac{\cos\{[\omega_c t + \varphi(t)] - [\omega_c(t-\tau) + \varphi(t-\tau)]\}}{2}$$

is

$$\frac{\cos\{[\omega_c t+\varphi(t)]-[\omega_c(t-\tau)+\varphi(t-\tau)]\}}{2}=\frac{\cos\{\omega_c t+\varphi(t)-\omega_c t+\omega_c\tau-\varphi(t-\tau)\}}{2}=\left|\,_{\omega_c\tau=\frac{\pi}{2}}\right|=\cdots$$

$$\cdots=\frac{\cos\left\{\frac{\pi}{2}+\varphi(t)-\varphi(t-\tau)\right\}}{2}=\left|\,_{\cos\left(\frac{\pi}{2}+\alpha\right)=-\sin\alpha}\right|=-\frac{\sin[\varphi(t)-\varphi(t-\tau)]}{2}$$

This component passes through the low pass filter. If we suppose very small phase difference over the period τ, then we can write

$$-\frac{\sin[\varphi(t)-\varphi(t-\tau)]}{2}\approx-[\varphi(t)-\varphi(t-\tau)]$$

and it is valid

$$\frac{\varphi(t)-\varphi(t-\tau)}{\tau}\approx\frac{d\varphi(t)}{dt}\quad\Rightarrow\quad-[\varphi(t)-\varphi(t-\tau)]\approx-\tau\cdot\frac{d\varphi(t)}{dt}$$

The sum term from the multiplication gets twice the carrier frequency and it is removed by the low pass filtering. Then the output signal is

$$-\tau\cdot\frac{d\varphi(t)}{dt}=\left|\,_{FM:\ \varphi(t)=k_2\int_0^t m(t)dt+\varphi_o}\right|=\frac{-\tau}{2}\cdot k_2\cdot m(t)$$

- it approximates the original message.

We can see the constant delay circuit τ in Fig. 9.32, its small signal model:

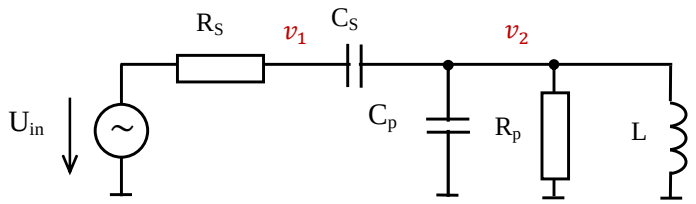


Fig 9. 32 Small signal model of the „delay τ" block

The ratio of U_2 over U_1 is the ratio of impedances Z_p (it is parallel combination of L, R_p and C_p) over $(Z_p+1/(pC_s))$, $p=j\omega$ – harmonic steady state – it is just an impedance divider. By simplifying this ratio we get

$$\frac{U_2}{U_1}=\frac{C_s}{C_s+C_p}\cdot\frac{p^2}{p^2+p\cdot\frac{1}{R_p(C_s+C_p)}+\frac{1}{L(C_s+C_p)}}=\frac{C_s}{C_s+C_p}\cdot\frac{p^2}{p^2+p\cdot\frac{\omega_o}{Q}+\omega_o^2}$$

From this formula it is evident that

$$\omega_o=\frac{1}{\sqrt{L(C_s+C_p)}}\quad\text{and}\quad\frac{\omega_o}{Q}=\frac{1}{R_p(C_s+C_p)}\quad\Rightarrow\quad Q=\frac{R_p}{\omega_o L}$$

If we substitute $p=j\omega$, we get

$$\frac{U_2}{U_1}=\frac{C_s}{C_s+C_p}\cdot\frac{-\omega^2}{\omega_o^2-\omega^2+j\omega\cdot\frac{\omega_o}{Q}}$$

Now we can derive phase responses as

$$\Theta(\omega) = -\pi - \operatorname{arctg}\frac{\omega\omega_o}{Q(\omega_o^2-\omega^2)}$$

Next we can get time delay of the circuit:

$$\tau = -\frac{d\Theta}{d\omega} = \frac{d}{d\omega}\left(\operatorname{arctg}\frac{\omega\omega_o}{Q(\omega_o^2-\omega^2)}\right) = \left|_{\frac{d}{dx}(\operatorname{arctg})=\frac{1}{1+x^2}}\right| = \cdots$$

$$\cdots = \frac{1}{1+\left[\frac{\omega\omega_o}{Q(\omega_o^2-\omega^2)}\right]^2}\cdot\frac{d}{d\omega}\left[\frac{\omega\omega_o}{Q(\omega_o^2-\omega^2)}\right] = \cdots$$

$$\cdots = \frac{\omega_o^3+\omega^2\omega_o}{Q(\omega_o^2-\omega^2)^2+\frac{\omega_o^2\omega^2}{Q}}$$

If we suppose

$$\omega = \omega_o + \Delta\omega = \omega_o\left(1+\frac{\Delta\omega}{\omega_o}\right) = \omega_o(1+x)$$

we can get

$$\tau = \frac{\omega_o^3+\omega^2\omega_o}{Q(\omega_o^2-\omega^2)^2+\frac{\omega_o^2\omega^2}{Q}} = \frac{\omega_o^3[1+(1+x)^2]}{\omega_o^4\{Q[1-(1+x)^2]\}^2+\frac{\omega_o^4(1+x)^2}{Q}} = \left|_{x\to 0}\right| \cong \cdots$$

$$\cdots \cong \frac{2Q}{\omega_o}\cdot\frac{1+x}{1+2x}$$

Similarly we can determine the modulus

$$\left|\frac{U_2}{U_1}\right| = \frac{C_s}{C_s+C_p}\cdot\frac{\omega^2}{(\omega_o^2-\omega^2)^2+\left(\omega\cdot\frac{\omega_o}{Q}\right)^2} = \left|_{x\to 0}\right| \cong \frac{C_s}{C_s+C_p}\cdot\frac{Q(1+2x)}{1+x}$$

So if we have $\Delta\omega/\omega_o \ll 1$, the delay and modulus are constant nearly:

$$\tau = \frac{2Q}{\omega_o} \qquad \left|\frac{U_2}{U_1}\right| = \frac{QC_s}{C_s+C_p}$$

9.2.5 Phase – locked loop (PLL)

- Commonly used for carrier synchronization and indirect FM demodulation
- Consists of (Fig. 9.33)
  - VCO – performs frequency modulation
  - Multiplier
  - Loop Filter – low-pass kind, removes the high-frequency components contained in the multiplier output.

To explain the operation, let assume that the VCO frequency is initially set to the carrier frequency of the incoming FM signal and the VCO's output has a 90- degrees phase-shift.

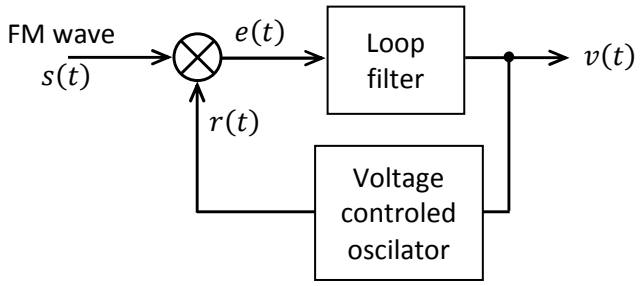


Fig 9. 33 Block diagram of the phase-locked look

We have FM signal

$s(t) = A_c \sin[\omega_c t + \varphi_1(t)]$; $\varphi_1(t) = 2\pi k_1 \int m(t)dt$; $m(t)$ – modulation signal.

$r(t) = A_v \cos[\omega_c t + \varphi_2(t)]$ – it describes the 90 –degree phase shift and the carrier frequency too.

But VCO is a frequency modulator, too – so it must be $\varphi_2(t) = 2\pi k_2 \int v(t)dt$.

$$e(t) = k_m A_c \sin[\omega_c t + \varphi_1(t)] \cdot A_v \cos[\omega_c t + \varphi_2(t)] =$$

$$= k_m A_c A_v \frac{\sin[2\omega_c t + \varphi_1(t) + \varphi_2(t)] + \sin[\varphi_1(t) - \varphi_2(t)]}{2}$$

k_m defines a multiplier constant.

Now we can determine a phase difference, low-pass loop filter removes the high-frequency components contained in the multiplier output:

$$\Delta\Theta = \varphi_1(t) - \varphi_2(t) = 2\pi k_1 \int m(t)dt - 2\pi k_2 \int v(t)dt$$

If we can ensure $\Delta\Theta \to 0$, we get

$$\Delta\Theta = \varphi_1(t) - \varphi_2(t) = 2\pi k_1 \int m(t)dt - 2\pi k_2 \int v(t)dt = 0$$

thus

$$2\pi k_1 \int m(t)dt - 2\pi k_2 \int v(t)dt = 0$$

thus the output voltage

$$v(t) = \frac{k_1}{k_2} m(t)$$

the output voltage approximates the original message $m(t)$.

Now we differentiate the phase difference

$$\frac{d}{dt}\Theta = 2\pi k_1 m(t) - 2\pi k_2 v(t) \qquad (*)$$

and will investigate the problem in more detail.

Let we have

$$v(t) = k_{LP}e(t) = \Big|_{\text{low frequency only}}\Big| = k_{LP}k_m A_c A_v \frac{\{\sin[2\omega_c t + \varphi_1(t) + \varphi_2(t)] + \sin[\varphi_1(t) - \varphi_2(t)]\}}{2}$$

$$v(t) = \frac{k_{LP}k_m A_c A_v}{2} \cdot \sin[\varphi_1(t) - \varphi_2(t)] = \Big|_{\Delta\Theta \to 0}\Big| = \frac{k_{LP}k_m A_c A_v}{2} \cdot \Delta\Theta$$

k_{LP} - defines a transfer constant of the loop (lowpass) filter. Hence for the small phase difference

$$\frac{d}{dt}\Theta = 2\pi k_1 m(t) - 2\pi k_2 v(t) = 2\pi k_1 m(t) - 2\pi k_2 \frac{k_{LP}k_m A_c A_v}{2} \cdot \Delta\Theta$$

We can rewrite this formula as below $\frac{d}{dt}\Theta = 2\pi k_1 m(t) - K \cdot \Delta\Theta$

where $K = 2\pi k_2 \frac{k_{LP}k_m A_c A_v}{2}$ is **phase lock loop gain**.

Now is valid too

$$v(t) \cong \frac{k_{LP}k_m A_c A_v}{2} \cdot \Delta\Theta = \frac{2\pi k_2}{2\pi k_2} \cdot \frac{k_{LP}k_m A_c A_v}{2} \cdot \Delta\Theta = \frac{K \cdot \Delta\Theta}{2\pi k_2}$$

It is supposed the "signal" output voltage of $v(t) = U_o \sin(\omega_m t)$, now. We easily determine corresponding (needed) phase difference

$$v(t) = U_o \sin(\omega_m t) = \frac{K \cdot \Delta\Theta}{2\pi k_2} \;\Rightarrow\; \Delta\Theta = \frac{2\pi k_2}{K} U_o \sin(\omega_m t)$$

Now from eq. (*) we determine

$$2\pi k_1 m(t) = \frac{d}{dt}\Delta\Theta + 2\pi k_2 v(t) = \frac{2\pi k_2}{K} U_o \omega_m \cos(\omega_m t) + 2\pi k_2 U_o \sin(\omega_m t)$$

Thus

$$m(t) = \frac{k_2 \cdot \omega_m}{k_1 \cdot K} \cdot U_o \cos(\omega_m t) + \frac{k_2}{k_1} \cdot U_o \sin(\omega_m t)$$

If the K is infinite (large enough), we get again

$$m(t) = \frac{k_2}{k_1} \cdot U_o \sin(\omega_m t) = \frac{k_2}{k_1} \cdot v(t)$$

For the K large enough, very small phase difference can generate appropriate output signal $v(t)$ – this approximates original message.

The PLL has some interesting features. The most useful of these is that it largely ignores fluctuations in the amplitude of the input FM wave. This means that the PLL FM demodulator tends to ignore any unwanted interference which appears alongside the input signal. The loop is said to have the property of *AM Rejection*.

Of course, the system can't work perfectly for any input signal size, no matter how small. The loop gain does depend upon A_c, so if the FM wave's amplitude is too small we can't assume that the gain is very large. For this reason we usually try to ensure that the input wave's amplitude is 'big enough' to avoid these problems.

The analysis above assumed that the VCO could always adjust its output to 'track' any changes in the FM wave's frequency. This is the same as assuming that the low pass filter isn't affecting the DBM's output. In reality, a low pass filter will tend to attenuate (and phase shift) any swiftly changing signals. As a result, the above arguments only apply strictly for a system which is *locked* (the signal and VCO maintain a steady phase relationship) when the modulation frequency is 'low' (i.e. passes through the filter without being affected).
Consider for a moment what happens when we start with the PLL not locked to an input signal. The input (f_s) and VCO (f_o) frequencies will differ by some amount $\Delta f = f_s - f_o$ so the output voltage (averaged over a few cycles of carrier) will be

$$v(t) \approx \frac{k_2}{k_1} \cdot \sin(2\pi\Delta f)$$

i.e. the output voltage tends to oscillate at the *difference frequency*, Provided that Δf is low enough this variation can pass through the low pass filter the loop can use this to help it *lock onto* the input wave. This brings the two frequencies together, and removes this 'beating' effect. However, if the difference frequency is too high to get through the low pass filter the multiplier's output will not be communicated to the VCO. The PLL then cannot lock onto the input FM wave and essentially ignores it! As a result, the system can only lock onto an input and demodulate it if the signal frequency is such that $|\Delta f| \leq B$ where B is the bandwidth of the filter. The system is said to have a *lock in range* of $\pm B$ set by the choice of the *loop filter*. A loop '*free running*' at a frequency, f_o will therefore ignore signals outside the range $f_o \pm B$.

9.2.6 Phase modulation

The usual expression for a harmonic signal is

$$S(t) = \cos[\emptyset(t)] = \cos(\omega_c t + \varphi_o)$$

The instantaneous radian frequency ω_i is

$$\omega_i = \frac{d\emptyset}{dt} = \frac{d}{dt}(\omega_c t + \varphi_o)$$

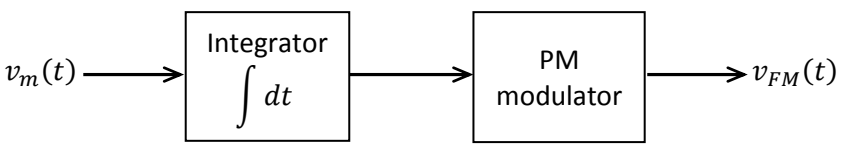
Frequency modulation: $\omega(t) = \omega_c + k_2 \cdot m(t)$;

$$\emptyset(t) = \omega_c t + k_2 \int_0^t m(t)dt + \varphi_o$$

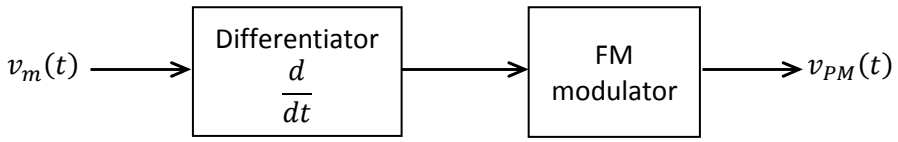
Phase modulation: $\emptyset(t) = \omega_c t + \varphi_o + k_1 \cdot m(t)$;

$$\omega(t) = \omega_c + k_1 \cdot \frac{d}{dt} m(t)$$

The phase of the carrier (PM) varies with the modulating signal (directly). The phase of the carrier (FM) varies with the integral of the modulating signal.

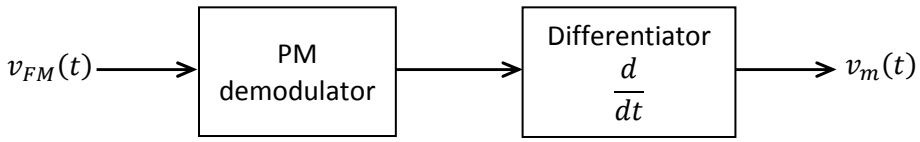


9.34 Generation of FM

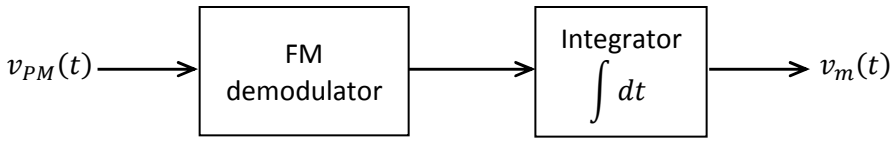


9.35 Generation of PM

Analogously for demodulation we can get



9.36 FM demodulator

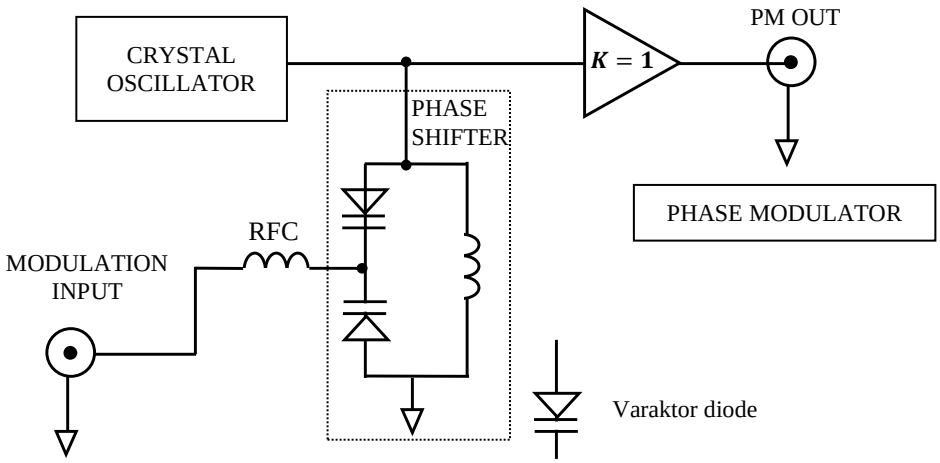


9.37 PM demodulator

Comparing formulas above we find that — unless we know something about the modulation in advance — it may not be obvious whether the signal is FM or PM modulated. In both cases the wave's frequency and phase vary from moment to moment. Mathematically speaking, FM & PM are almost identical twins. The only difference is that one corresponds to a modulation pattern which is the differential of that produced by the other. The ‘good news’ is, that this means we can mix FM & PM arguments and most of our conclusions about one apply to the other. The ‘bad news’ is that, in practice we often have to know in advance which type of modulation is being used if we want to recover the modulated information correctly. In general, however, both FM & PM waves are obviously different to an AM wave.

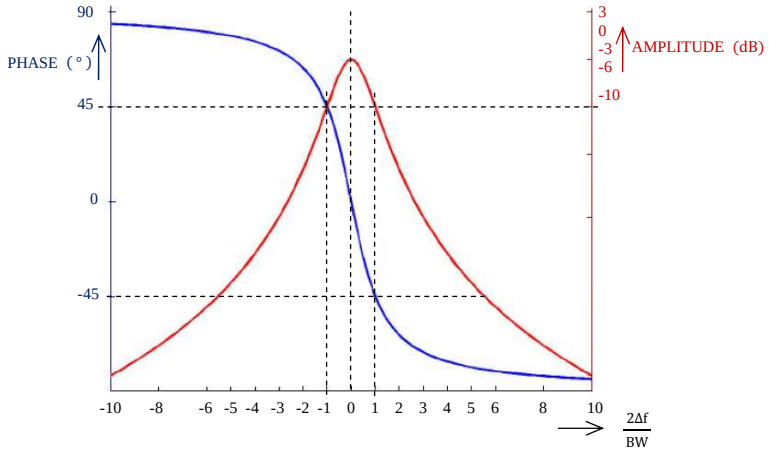
I tis generally valid for modulation (Fig. 9.34 to Fig. 9.37):

The earliest analog phase modulators are of the type in Fig. 9.38. A crystal oscillator (XO) generates the desired frequency and voltage controlled phase shifter does the modulating. Typically, the phase shifter is a tuned circuit with voltage controlled capacitors as a tuning element.



9.38 Phase modulator

Figure 9.39 shows the amplitude and phase of a single tuned resonant circuit. The plot is normalized so units are independent of Q. The plot shows the phase shift of 45 degrees (0.785 radians) at the –3 dB points on the amplitude curve. The curve for phase is a tangent curve and goes from +90 to –90 degrees with zero degrees at resonance. The modulation of the voltage to the varicap shifts the resonant frequency. It results in phase shift of the output – in accordance with the phase response curve.



9. 39 Response of a LC circuit

9.2.7 PM demodulator

The basic concept upon which phase detection rests is that the application of two identical frequencies, constant amplitude signals to a mixer results in a dc output which is proportional to the phase difference between the two signals. While it is true that even a single diode can be used as a mixer, most phase detectors involve the use of double balanced mixers – see description of the DBM above.

Very often the ***Gilbert-cell*** (see Fig. 9.40) is used as a multiplier circuit (modulators, demodulators, mixers). It includes a tail current source (Q_7), a differential transconductance stage (Q_1, Q_2), and a switching quad ($Q_3 - Q_6$). The output can be driven to a resistive or reactive tuned load.

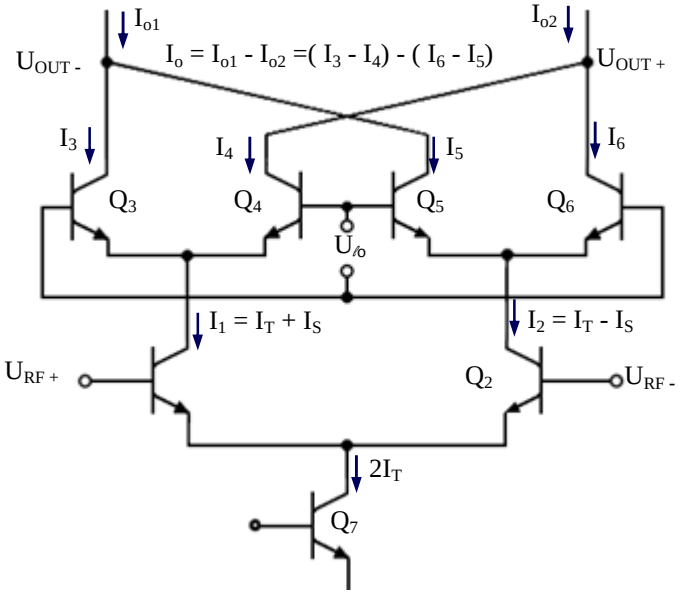
If we suppose small signals only, we can write (see small signal BJT model, Chapter 2):

$$r_{e_3} \cong r_{e_4} \cong \frac{U_{th}}{I_T+I_S}; \qquad r_{e_5} \cong r_{e_6} \cong \frac{U_{th}}{I_T-I_S}$$

$$I_3 = \frac{I_T+I_S}{2} - \frac{\frac{U_{\ell o}}{2}}{r_{e_3}}; \qquad I_4 = \frac{I_T+I_S}{2} + \frac{\frac{U_{\ell o}}{2}}{r_{e_4}}; \qquad I_5 = \frac{I_T-I_S}{2} - \frac{\frac{U_{\ell o}}{2}}{r_{e_5}}; \quad I_6 = \frac{I_T-I_S}{2} - \frac{\frac{U_{\ell o}}{2}}{r_{e_6}}$$

$$I_{o_1} = I_3 + I_5 = \frac{I_T+I_S}{2} - \frac{U_{\ell o}}{2r_{e_3}} + \frac{I_T-I_S}{2} + \frac{U_{\ell o}}{2r_{e_5}} = I_T + \frac{U_{\ell o}}{2} \cdot \left(\frac{-1}{r_{e_3}} + \frac{1}{r_{e_5}}\right)$$

$$I_{o_2} = I_4 + I_6 = \frac{I_T+I_S}{2} - \frac{U_{\ell o}}{2r_{e_4}} + \frac{I_T-I_S}{2} + \frac{U_{\ell o}}{2r_{e_6}} = I_T + \frac{U_{\ell o}}{2} \cdot \left(\frac{1}{r_{e_4}} - \frac{1}{r_{e_6}}\right)$$



9. 40 BJT Gilbert cell

An output current we define as a difference

$$I_o = I_{o_1} - I_{o_2} = I_T + \frac{U_{\ell o}}{2} \cdot \left(\frac{-1}{r_{e_3}} + \frac{1}{r_{e_5}}\right) - \left[I_T + \frac{U_{\ell o}}{2}\right] \cdot \left(\frac{-1}{r_{e_4}} + \frac{1}{r_{e_6}}\right) \cong U_{\ell_o} \cdot \left(\frac{1}{r_{e_{5,6}}} - \frac{1}{r_{e_{3,4}}}\right)$$

It is evident that for small signals is

$$I_S \cong \frac{U_{RF}}{2r_e} \quad \text{where} \quad r_e = \frac{U_{th}}{I_T}.$$

Now it is valid

$$I_o \cong U_{\ell_o} \cdot \left(\frac{1}{r_{e_{5,6}}} - \frac{1}{r_{e_{3,4}}}\right) = U_{\ell_o} \left(\frac{1}{\frac{U_{th}}{I_T+I_S}} - \frac{1}{\frac{U_{th}}{I_T-I_S}}\right) = U_{\ell_o} \left(\frac{I_T+I_S}{U_{th}} - \frac{I_T-I_S}{U_{th}}\right) = U_{\ell_o} \cdot \frac{2I_S}{U_{th}}$$

$$I_o \cong U_{\ell_o} \cdot \frac{2I_S}{U_{th}} = U_{\ell_o} \cdot \frac{\frac{2U_{RF}}{2r_e}}{U_{th}} = \frac{U_{\ell_o} \cdot U_{RF}}{r_e\, U_{th}}$$

We can investigate ***large signal properties***, too. The mixing principle of a Gilbert cell is based on the so called controlled transconductance mixer; see simplified (basic) circuit in Fig. 9.41. For a down-conversion application, the information carrying high frequency (RF) signal is applied as a voltage that modulates the current source with quiescent current $2I_o = I_1 + I_2$. First, assuming no RF signal, the current is divided up in the currents I_1 and I_2 depending on the applied voltage U_{ℓ_o} through (see the BOX 1 below)

$$I_1 = \frac{2I_o}{1 + e^{-U_{\ell_o}/U_{th}}} \qquad I_2 = \frac{2I_o}{1 + e^{U_{\ell_o}/U_{th}}}$$

from the exponential voltage to current relation of the bipolar transistor. Here, $U_{th} = kT/q$ denotes the thermal voltage ($U_{th} = 26$mV at room temperature).

BOX 1

$I_E \cong I_{E_0} \cdot e^{U_{BE}/U_{th}} \longrightarrow I_{E_1} \cong I_{E_0} \cdot e^{U_{BE_1}/U_{th}}; \quad I_{E_2} \cong I_{E_0} \cdot e^{U_{BE_2}/U_{th}} \quad \Rightarrow$

$U_{BE_1} = U_{th} \cdot \ln \frac{I_1}{I_{E_0}}; \quad U_{BE_2} = U_{th} \cdot \ln \frac{I_2}{I_{E_0}}$

It is evident that: $U_{\ell_o} = U_{BE_1} - U_{BE_2} \quad \Rightarrow$

$$U_{BE_1} - U_{BE_2} = U_{th} \cdot \ln \frac{I_1}{I_{E_0}} - U_{th} \cdot \ln \frac{I_2}{I_{E_0}} = U_{th} \cdot \ln \frac{I_1}{I_2}$$

$$\frac{I_1}{I_2} = e^{U_{\ell_o}/U_{th}} \quad \Rightarrow$$

$$I_1 = I_2 \cdot e^{U_{\ell_o}/U_{th}}; \qquad I_2 = I_1 \cdot e^{-U_{\ell_o}/U_{th}}$$

It is evident that:

$$I_1 + I_2 = 2I_o \quad \Rightarrow$$

$$I_1 + I_2 = I_1 + I_1 \cdot e^{\frac{-U_{\ell_o}}{U_{th}}} = 2I_o \quad \Rightarrow \quad I_1 = \frac{2I_o}{1 + e^{\frac{-U_{\ell_o}}{U_{th}}}}$$

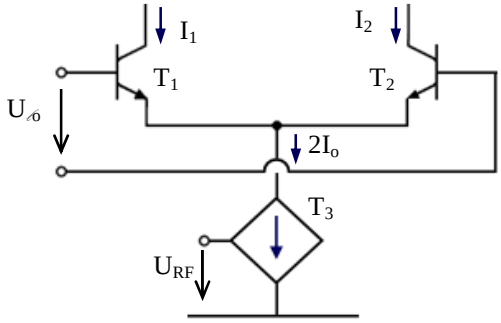
$$I_1 + I_2 = I_2 + I_2 \cdot e^{\frac{U_{\ell_o}}{U_{th}}} = 2I_o \quad \Rightarrow \quad I_2 = \frac{2I_o}{1 + e^{\frac{U_{\ell_o}}{U_{th}}}}$$

The differential output current i_{out} is then given by (see the BOX 2), Fig. 9.42:

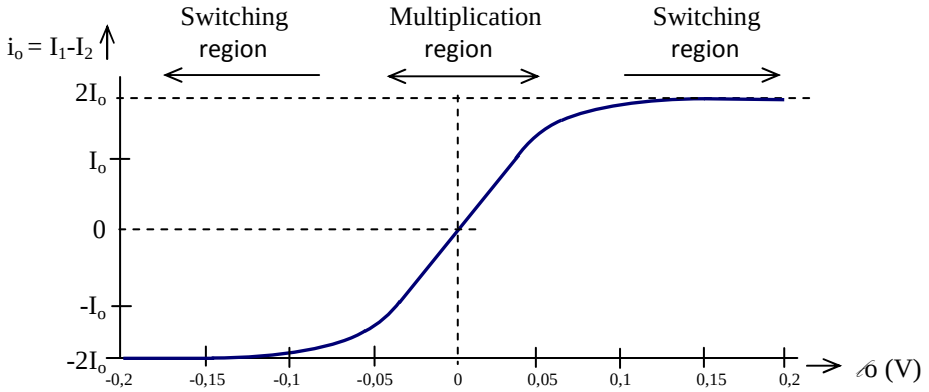
$$i_{out} = I_1 - I_2 = 2I_o \tanh\left(\frac{U_{\ell_o}}{2U_{th}}\right)$$

If U_{ℓ_o} is small ($\ll U_{th}$), $\tanh\left(U_{\ell_o}/\,2U_{th}\right) \cong U_{\ell_o}/\,2U_{th}$ and so the output current i_0 is approximately linearly proportional to U_{ℓ_o}(the mixer is said to be working in the"***multiplication region***"; in the opposite case $\tanh \rightarrow \pm 1$ – it means ***switching region***):

$$i_{out} = I_o \frac{U_{\ell_o}}{2U_{th}}$$



9.41 Basic Gilbert cell



9.42 Output current of the basic Gilbert cell (diference)

BOX 2

$$i_{out} = I_1 - I_2 = \frac{2I_0}{1+e^{-\alpha}} - \frac{2I_0}{1+e^{\alpha}} = \left| \alpha = \frac{U_{\ell o}}{U_{th}} \right| = 2I_0 \cdot \frac{1+e^{\alpha}-(1+e^{-\alpha})}{(1+e^{\alpha})\cdot(1+e^{-\alpha})} = 2I_0 \cdot \frac{e^{\alpha}-e^{-\alpha}}{e^{\alpha}+e^{-\alpha}+2} = \cdots$$

$$\cdots = 2I_0 \cdot \frac{(e^{\alpha/2})^2-(e^{-\alpha/2})^2}{(e^{\alpha/2})^2+(e^{-\alpha/2})^2+2} = \left| e^{\alpha/2}\cdot e^{-\alpha/2}=1 \right| = 2I_0 \cdot \frac{(e^{\alpha/2})^2-(e^{-\alpha/2})^2}{(e^{\alpha/2})^2+(e^{-\alpha/2})^2+2(e^{\alpha/2}\cdot e^{-\alpha/2})} = \cdots$$

$$\cdots = 2I_0 \cdot \frac{(e^{\alpha/2}+e^{-\alpha/2})\cdot(e^{\alpha/2}-e^{-\alpha/2})}{(e^{\alpha/2}+e^{-\alpha/2})^2} = 2I_0 \cdot \frac{e^{\alpha/2}-e^{-\alpha/2}}{e^{\alpha/2}+e^{-\alpha/2}} = 2I_0 \cdot \tanh\frac{U_{\ell o}}{2U_{th}}$$

Now, if U_{RF} is a small signal voltage and the transconductance of the current source (T_3) is g_m, $2I_0$ is replaced by $2I_0 + g_m U_{RF}$ and

$$i_{out} \cong (2I_0 + g_m U_{RF}) \cdot \frac{U_{\ell o}}{2U_{th}} = I_0 \cdot \frac{U_{\ell o}}{U_{th}} + \frac{g_m U_{RF} U_{\ell o}}{2U_{th}}$$

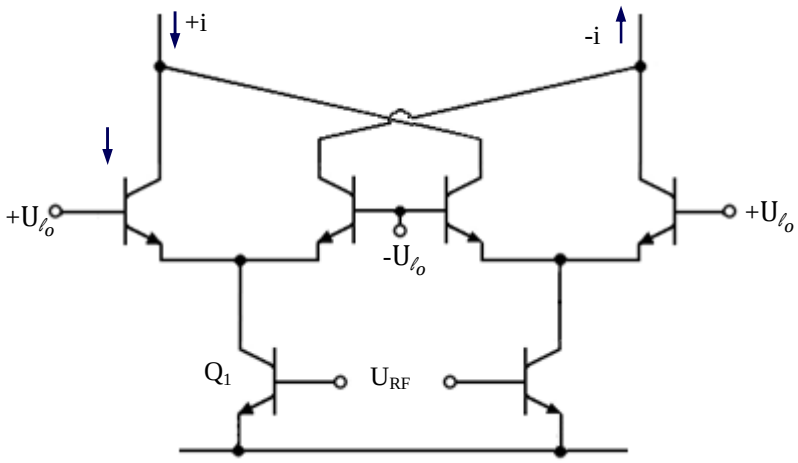
The first term is called LO leakage or feed-through and the second term is the wanted one. The first term can be cancelled by employing a second, identical, circuit, driven by $-U_{\ell o}$, which output current is

$$i_{out} \cong I_0 \cdot \frac{U_{\ell o}}{U_{th}} - \frac{g_m U_{RF} U_{\ell o}}{2U_{th}}$$

and the output current is taken as the difference between i_{out} and i_{out_2}:

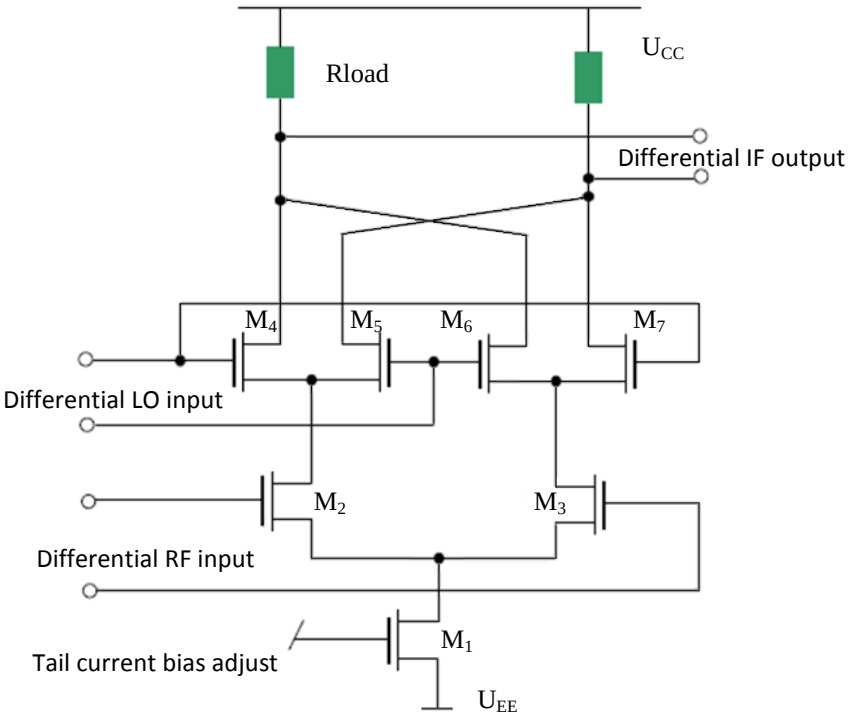
$$i_{IF} = i_{out} - i_{out_2} = \frac{g_m U_{RF} U_{\ell o}}{2U_{th}}$$

See figure 9.43 – and compare with Gilbert cell described above.



9.43 Employing a sekond circuit

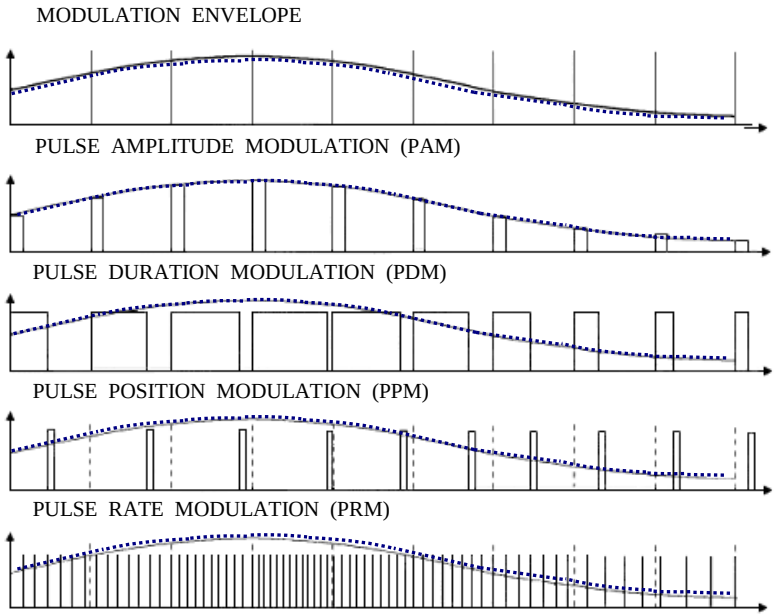
Although the Gilbert-cell was initially designed with bipolar transistors, its operation principle is similar using CMOS technology – see figure 9.44 below.



9.44 MOSTEF Gilbert-cell

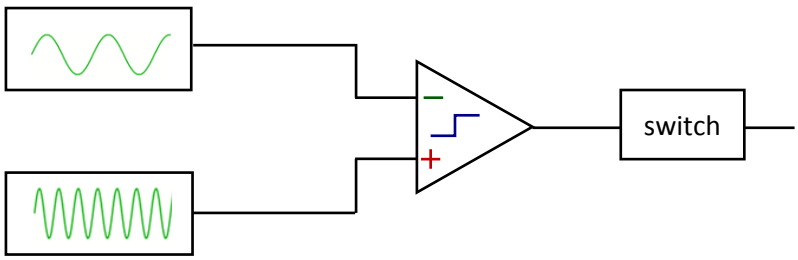
9.3 Pulse modulations

Another class of modulation, called pulse modulation, is applicable in sampled waveform scenarios. The information in a sampled signal can be represented (or encoded) by varying the amplitude, duration (or width), position (relative to a time reference), or repetition rate of the transmitted pulses. The way in which information is encoded by several types of pulse modulation is illustrated in Fig. 9.45. The most common forms of pulse modulation are essentially forms of amplitude modulation. However, frequency or polarization modulation could also be used to encode the signal amplitude modulation.



9.45 Types of pulse modulation

As an example let us see a ***PDM (PWM) modulator principle***. In the analogue domain a PWM signal can be generated by comparing the signal to a triangle or sawtooth waveform. This technique, called natural sampling, is the basis of almost all analogue modulators. See Fig. 9.46 below.



9.46 PDM principle

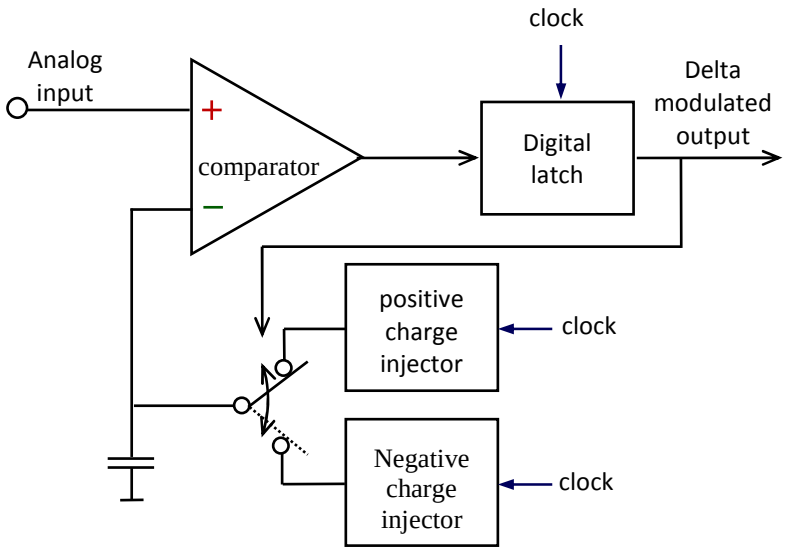
When the momentary value of the input signal is larger than the triangle, the output of the switch is high. It is easy to see that in this way the pulse width at the output is proportional to the input voltage.

Related to but not a true form of pulse modulation is **pulse code modulation (PCM).** PCM is a truly digital modulation technique. In PCM the signal is sampled, the sampled signal is quantized into two more discrete levels, and the quantized level is transmitted as a unique string of bits, the bits having been encoded into pulses. Pulse amplitude, position, duration, pulse rate, frequency, polarization, phase, and polarity among others are capable of being used to encode the bits.

Analog-over-digital methods:

- Pulse-code modulation (PCM)
- Differential PCM (DPCM)
- Adaptive DPCM (ADPCM)
- Delta modulation (DM)
- Sigma-delta modulation
- Continuously variable slope delta modulation (CVSDM), also called Adaptive-delta modulation (ADM)
- Pulse-density modulation (PDM)

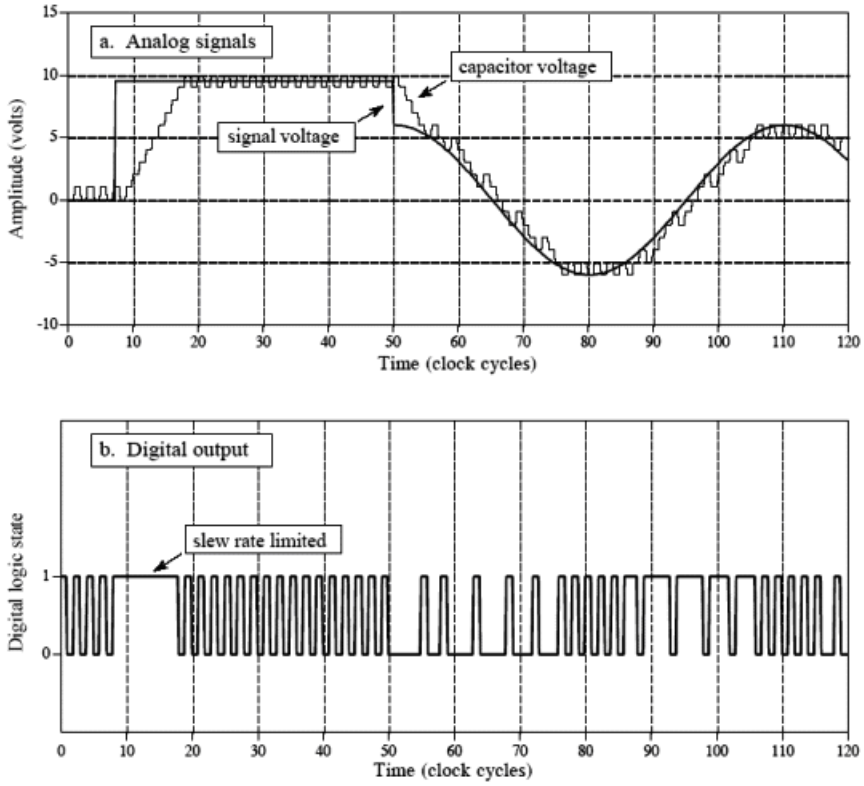
As an example let us see a ***delta modulator*** – see figure 9.47. The analog input is a voice signal with amplitude of a few volts, while the output signal is a stream of digital ones and zeros. A comparator decides which has the greater voltage, the incoming analog signal, or the voltage stored on the capacitor. This decision, in the form of a digital one or zero, is applied to the input of the latch. At each clock pulse, typically at a few hundred kilohertz, the latch transfers whatever digital state appears on its input, to its output. This latch insures that the output is synchronized with the clock, thereby defining the sampling rate, i.e., the rate at which the 1 bit output can update itself.



9.47 Delta modulator

Figure 9.48 illustrates the signals produced by this circuit. At time equal to zero, the analog input and the voltage on the capacitor both start with a voltage of zero. As shown in (a), the

input signal suddenly increases to 9,5 V on the eighth clock cycle. Since the input signal is now more positive than the voltage on the capacitor, the digital output changes to a *one*, as shown in (b). This results in the switch being connected to the positive charge injector, and the voltage on the capacitor increasing by a small amount on each clock cycle. Although an increment of 1 V per clock cycle is shown in (a), this is only for illustration, and a value of 1 mV is more typical. This staircase increase in the capacitor voltage continues until it exceeds the voltage of the input signal. Here the system reached ***equilibrium*** with the ***output oscillating between a digital one and zero***, causing the voltage on the capacitor to oscillate between 9 V and 10 V.



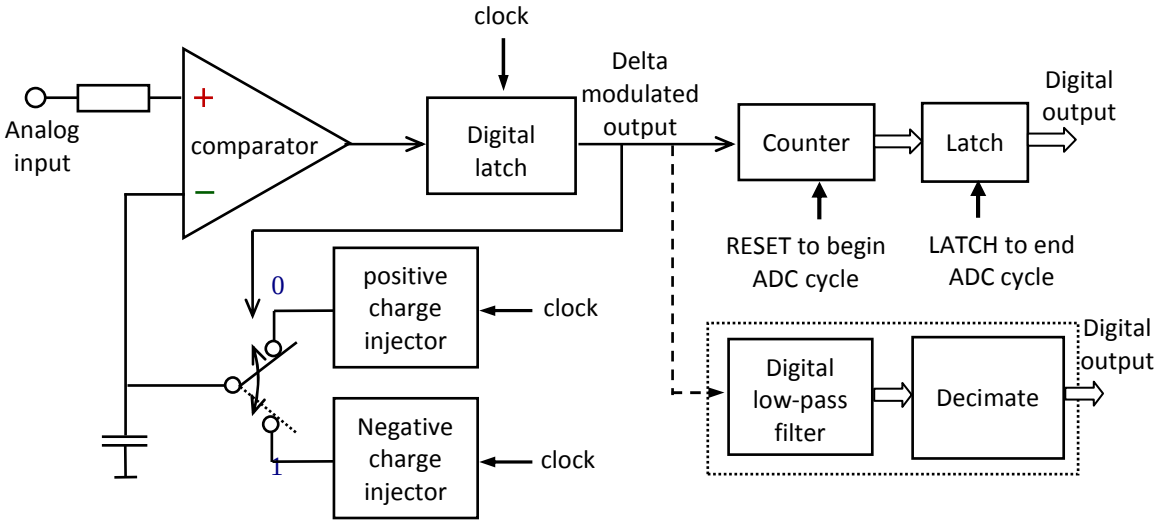
9.48 Example of signals produced by the delta modulator

In this manner, the feedback of the circuit forces the capacitor voltage to track the voltage of the input signal. If the input signal changes very rapidly, the voltage on the capacitor changes at a constant rate until a match is obtained. This constant rate of change is called the **slew rate**, just as in other electronic devices such as op amps.

Now, consider the characteristics of the delta modulated output signal. If the analog input is increasing in value, the output signal will consist of more ones than zeros. Likewise, if the analog input is decreasing in value, the output will consist of more zeros than ones. If the analog input is constant, the digital output will alternate between zero and one with an equal number of each. Put in more general terms, the relative number of ones versus zeros is directly proportional to the slope (derivative) of the analog input.

Output information of the delta modulator is related to the *derivative* of the input signal. It does not contain “DC information". The ***delta-sigma converter (modulator)***, shown in Fig.

9.49 – delta modulated output, eliminates these problems by cleverly combining analog electronics with DSP algorithms. Notice that the voltage on the capacitor is now being compared with ground potential. The feedback loop has also been modified so that the voltage on the capacitor is *decreased* when the circuit's output is a digital *one*, and *increased* when it is a digital *zero*. As the input signal increases and decreases in voltage, it tries to raise and lower the voltage on the capacitor. This change in voltage is detected by the comparator, resulting in the charge injectors producing a *counteracting* charge to keep the capacitor at zero volts.



9.49 Block diagram of a delta-sigma analog-to-digital converter

In the Fig. 9.49 is block diagram of a delta-sigma analog-to-digital converter, too. In the simplest case, the pulses from a delta modulator are counted for a predetermined number of clock cycles. The output of the counter is then latched to complete the conversion. In a more sophisticated circuit, the pulses are passed through a digital low-pass filter and then resampled (decimated) to a lower sampling rate.

If the input voltage is positive, the digital output will be composed of more ones than zeros. The excess number of ones is needed to generate the *negative* charge that cancels with the *positive* input signal. Likewise, if the input voltage is negative, the digital output will be composed of more zeros than ones, providing a net positive charge injection. If the input signal is equal to zero volts, an equal number of ones and zeros will be generated in the output, providing an overall charge injection of zero.

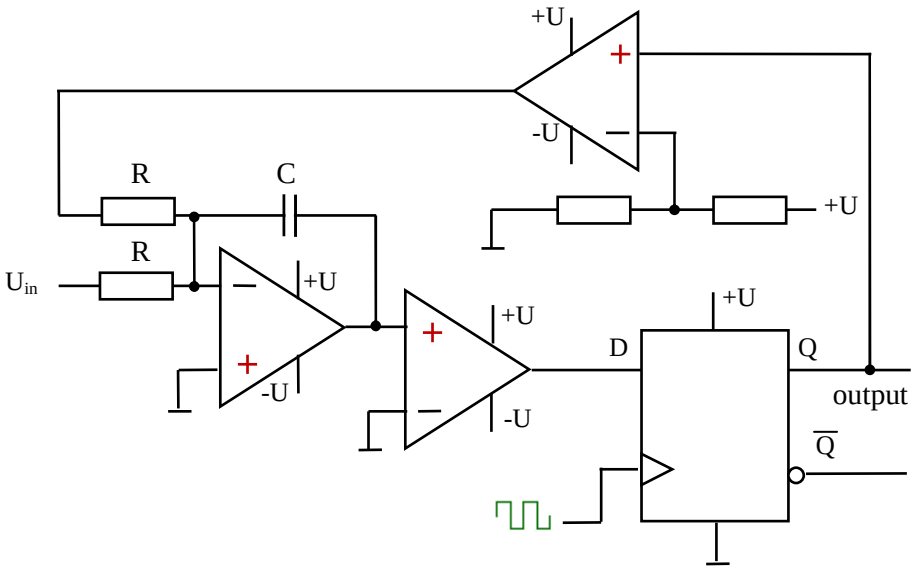
The relative number of ones and zeros in the output is now related to the *level* of the input voltage, not the *slope* as in the previous circuit. This is much simpler. For instance, you could form a 12 bit ADC by feeding the digital output into a counter, and counting the number of *ones* over 4096 clock cycles. A digital number of 4095 would correspond to the maximum positive input voltage. Likewise, digital number 0 would correspond to the maximum negative input voltage, and 2048 would correspond to an input voltage of zero. This also shows the origin of the name, *delta-sigma*: delta modulation followed by summation (sigma).

The ones and zeros produced by this type of delta modulator are very easy to transform back into an analog signal. All that is required is an analog lowpass filter, which might be as simple as a single RC network. The high and low voltages corresponding to the digital ones and zeros average out to form the correct analog voltage.

This method of transforming the single bit data stream back into the original waveform is important for several reasons. First, it describes a slick way to replace the counter in the delta-sigma ADC circuit. Instead of simply counting the pulses from the delta modulator, the binary signal is passed through a *digital* low-pass filter, and then *decimated* to reduce the sampling rate. For example, this procedure might start by changing each of the ones and zeros in the digital stream into a 12 bit sample; ones become a value of 4095, while zeros become a value of 0. Using a digital low-pass filter on this signal produces a digitized version of the original waveform, just as an analog lowpass filter would form an analog recreation. Decimation then reduces the sampling rate by discarding most of the samples. This results in a digital signal that is equivalent to direct sampling of the original waveform.

This approach is used in many commercial ADC's for digitizing voice and other audio signals. An example is the National Semiconductor ADC16071, which provides 16 bit analog-to-digital conversion at sampling rates up to 192 kHz. At a sampling rate of 100 kHz, the delta modulator operates with a clock frequency of 6,4 MHz. The low-pass digital filter is a 246 point FIR. This removes all frequencies in the digital data above 50 kHz, ½ of the eventual sampling rate. Conceptually, this can be viewed as forming a digital signal at 6,4 MHz, with each sample represented by 16 bits. The signal is then decimated from 6,4 MHz to 100 kHz, accomplished by deleting every 63 out of 64 samples. In actual operation, much more goes on inside of this device than described by this simple discussion.

Another very simple example of Sigma Delta Circuit you can see in Fig. 9.50:



9.50 Simple Sigma Delta Circuit

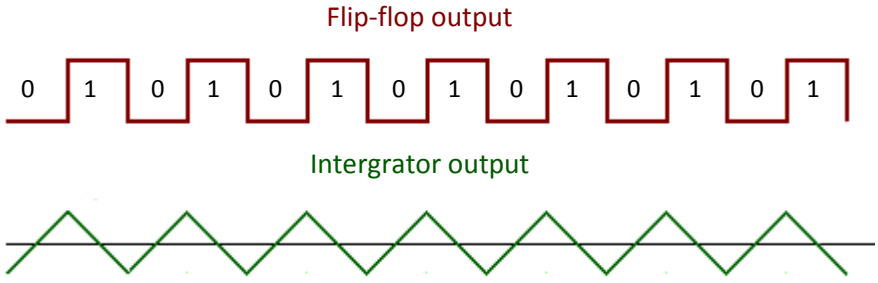
In a ΔΣ converter, the analog input voltage signal is connected to the input of an integrator, producing a voltage rate-of-change, or slope, at the output corresponding to input

magnitude. This ramping voltage is then compared against ground potential (0 volts) by a comparator. The comparator acts as a sort of 1-bit ADC, producing 1 bit of output ("high" or "low") depending on whether the integrator output is positive or negative. The comparator's output is then latched through a D-type flip-flop clocked at a high frequency, and fed back to another input channel on the integrator, to drive the integrator in the direction of a 0 volt output.

The leftmost op-amp is the (summing) integrator. The next op-amp the integrator feeds into is the comparator, or 1-bit ADC. Next comes the D-type flip-flop, which latches the comparator's output at every clock pulse, sending either a "high" or "low" signal to the next comparator at the top of the circuit. This final comparator is necessary to convert the single-polarity 0V/5V logic level output voltage of the flip-flop into a +U/-U voltage signal to be fed back to the integrator.

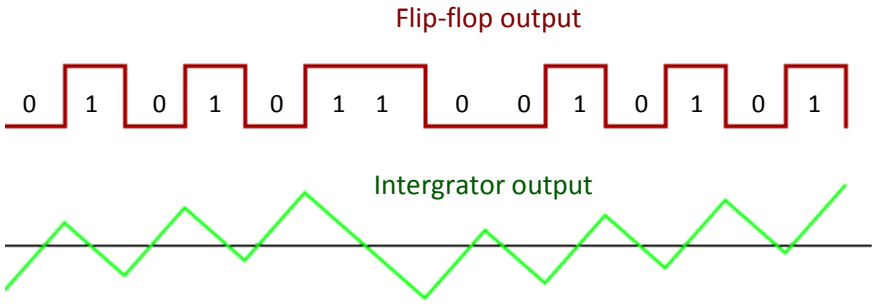
If the integrator output is positive, the first comparator will output a "high" signal to the D input of the flip-flop. At the next clock pulse, this "high" signal will be output from the Q line into the noninverting input of the last comparator. This last comparator, seeing an input voltage greater than the threshold voltage of (1/2)(+U), saturates in a positive direction, sending a full +U signal to the other input of the integrator. This +U feedback signal tends to drive the integrator output in a negative direction. If that output voltage ever becomes negative, the feedback loop will send a corrective signal (-U) back around to the top input of the integrator to drive it in a positive direction. This is the delta-sigma concept in action: the first comparator senses a difference (Δ) between the integrator output and zero volts. The integrator sums (Σ) the comparator's output with the analog input signal.

Functionally, this results in a serial stream of bits output by the flip-flop. If the analog input is zero volts, the integrator will have no tendency to ramp either positive or negative, except in response to the feedback voltage. In this scenario, the flip-flop output will continually oscillate between "high" and "low," as the feedback system "hunts" back and forth, trying to maintain the integrator output at zero volts – see Fig. 9.51.



9.51 ΔΣ converter operation with 0 volt analog input

If, however, we apply a ***negative analog input*** voltage, the integrator will have a tendency to ramp its output in a positive direction. Feedback can only add to the integrator's ramping by a fixed voltage over a fixed time, and so the bit stream output by the flip-flop will not be quite the same– see Fig. 9.52.

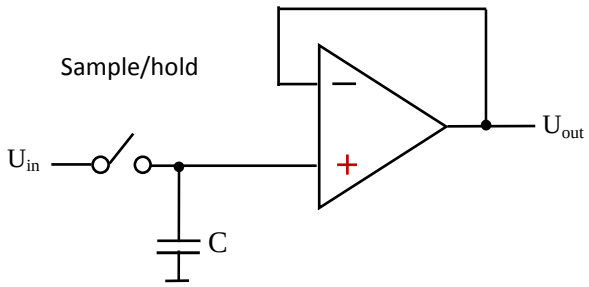


9.52 ΔΣ converter operation with small negative analog input

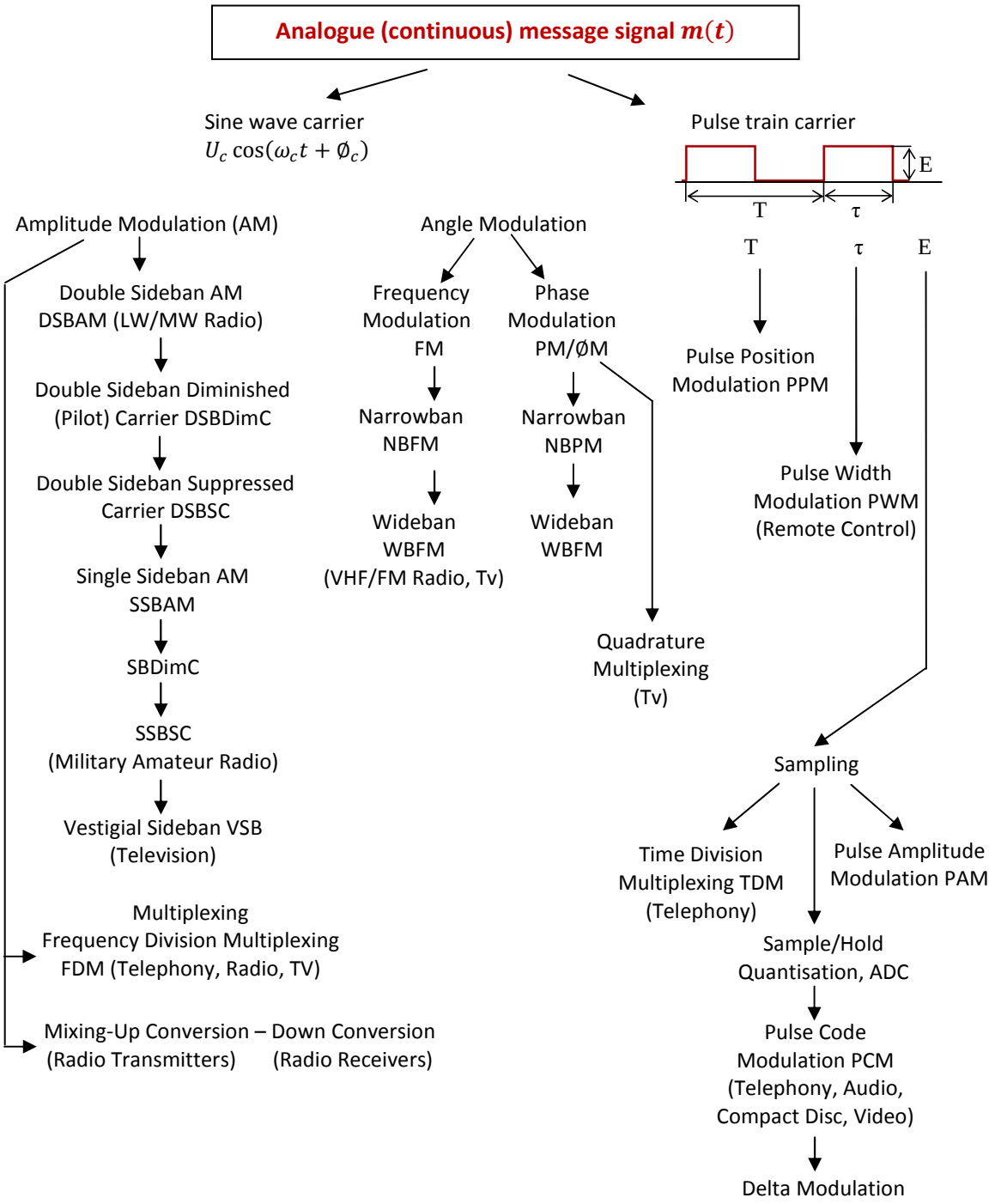
9.4 Sampling theorem

Frequently this is called the *Shannon* sampling theorem, or the *Nyquist* sampling theorem, after the authors of 1940s papers on the topic. The sampling theorem indicates that a continuous signal can be properly sampled, ***only if it does not contain frequency components above one-half of the sampling rate***. For instance, a sampling rate of 2,000 samples/second requires the analog signal to be composed of frequencies below 1000 cycles/second. If frequencies above this limit *are* present in the signal, they will be aliased to frequencies between 0 and 1000 cycles/second, combining with whatever information that was legitimately there.

Sample-and-hold circuits are used to sample an analog signal and hold it so that it can be analyzed or converted into, say, a digital signal at one's leisure. In fig. 9.53 is shown example sample/hold circuit – a switch acts as a sample/hold control. Sampling begins when switch is closed and ends when the switch is opened. When the switch is opened, the input voltage present at that exact moment will be stored in C. The op amp acts as unity-gain amplifier (buffer), relaying the capacitor's voltage to the output but preventing the capacitor from discharging (recall that ideally, no current entrs the inputs of an op amp.) The lenght time a sample voltage can be held varies depending on how much current leaks out of the capacitor. To minimize leakage currents, use op amps with low input bias-current (e.g. FET op amp).



9.53 Sample-and-hold circuit



Summary

1) The sine (cosine) wave is the basic for all complex waveforms. Without modulation the carrier (sine wave) is constant. Amplitude modulation is the process of mixing two different frequencies across a nonlinear resistance (a nonlinear functional block).
2) The frequency spectrum of a modulated wave can be conveniently illustrated in graph form as frequency versus amplitude (phase).

3) In frequency modulation (FM, fm) the instantaneous frequency is varied in accordance with the modulating signal; the amplitude of the “radio – frequency” wave is kept constant.
4) In phase modulation the carrier´s phase is caused to shift at the rate of the modulating signal.
5) Demodulation (detection) is the process of re – creating original modulating frequencies.
6) The slope detector is the simplest form of frequency detector. It is essentially a tank circuit tuned slightly away from the desired fm carrier.
7) The Foster – Seeley discriminator uses a double tuned rf transformator to convert frequency changes of the received fm signal into amplitude variation of the rf wave.
8) The ratio detector uses a double – tuned transformer connected so that the instantaneous frequency variations of the fm input signal are converted into instantaneous amplitude variations.
9) PLL is basically a closed loop frequency control system, which functioning is based on the phase sensitive detection of phase difference between the input and output signals of the voltage controlled oscillator (VCO).
10) A delta modulator (convertor) – the relative number of ones versus zeros is directly proportional to the slope (derivative) of the analog input. It does not contain “DC information”.
11) A sigma delta converter – the relative number of ones and zeros is now related to the level of the input voltage (not the slope).

Questions 9

You can find the answers in this text.

1. Define amplitude and frequency modulations - basic properties.
2. Write the general mathematical expression for modulation.
3. Draw the frequency spectrum of AM.
4. How do you get FM using PM system?
5. Describe basic demodulation principles (AM, FM, PM) – circuitry.
6. Define delta and sigma delta modulations – basic properties.
7. What is aliasing – sampling theorem?
8. What is PCM?

Problems 9

Example 9.1

If a 100 kHz signal and a 30 kHz signal are applied to a balance modulator, what frequencies will appear on the output?

Example 9.2

If a 100 kHz signal and 30 kHz signal are applied to a standard modulator, what frequencies will appear on the output?

Example 9.3

A voice signal ranging from 300 Hz to 3 kHz amplitude modulates a 600 kHz carrier. Develop the frequency spectrum.

Example 9.4

A PLL has free running frequency $\boldsymbol{f_c = 500}$ kHz. Bandwidth of low pass filter $\boldsymbol{B = 10}$kHz. Suppose an input signal of frequency $\boldsymbol{f_c = 600}$ kHz is applied. Will the loop acquire lock? What is VCO output frequency?

Example 9.5

What sampling frequency would we use to sample the following signal:

$$x(t) = 3 \cdot sin(2\pi \cdot 90 \cdot t) - 6 \cdot cos(2\pi \cdot 900 \cdot t)$$

PROBLEMS KEY 9

Ad example 9.1)

Balanced modulation produces no carrier frequency on the output, whereas standard AM does.

Ad example 9.2)

Balanced modulation produces no carrier frequency on the output, whereas standard AM does.

Ad example 9.3)

[(600-3)kHz to [(600-0,3)kHz; 600 kHz; [(600+0,3)kHz to [(600+3)kHz]

Ad example 9.4)

$f_s - f_C = 100$ kHz; A difference frequency comparator is outside the passband of lowpass filter. Hence loop will not acquire lock. VCO frequency will be free running frequency.

Ad example 9.5)

$2\pi \cdot 90 = 2\pi \cdot f_1$; $2\pi \cdot 900 = 2\pi \cdot f_2$; $f_1 = 90$ Hz, $f_2 = 900$ Hz

When we sample at the Nyquist, we need to sample at twice the highest frequency component of the signal → the sampling frequency $f_s \geq 2 \cdot f_2 = 1800$ Hz.

Basic texts

[1] Widkjaer, J.: Transmission, and demodulation. Class Notes, 31415 RF – Communications Circuits, Autumn 2005

[2] Application Note 150 – 1. Spectrum analysis Amplitude and Frequency Modulation, Hewlet Packard, November 1981

[3] Jarmen, D.: A Brief Introduction to Sigma Delta Conversion. Intersil, AN9504, May 1995

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[2] Stuerly, K., R.: Frequency modulation. 2nd Edition, Hulton Press, LTD, London 1946

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

10. A/D and D/A converter principles

Time of study: 6 hours

Goals:

- define basic principles of D/A converters
- define basic principles of A/D converters
- describe aliasing
- use multiply D/A converter for a control of circuit properties

EXPLANATION

10.1 Basic considerations – ADC

The devices which perform the interfacing function between analog and digital worlds are analog-to-digital (A/D or ADC) and digital-to-analog (D/A or DAC) converters, which together are known as data converters.

The input to the system is a physical parameter such as temperature, pressure, flow, acceleration and position, which are analog quantities. The parameter is first converted into an electrical signal by means of a transducer, once in electrical form, all further processing is done by electronic circuits – $x_{IN}(t)$ – Fig. 10.1.

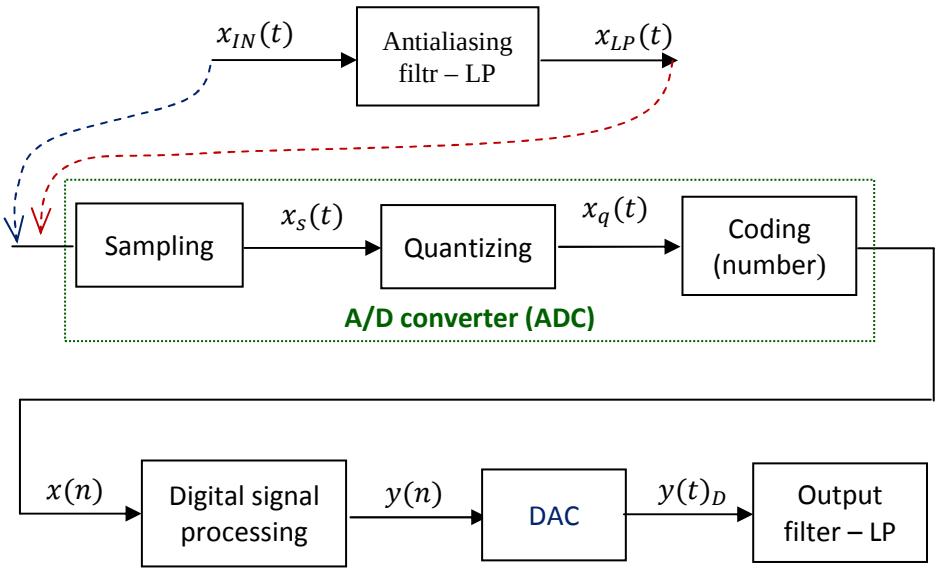
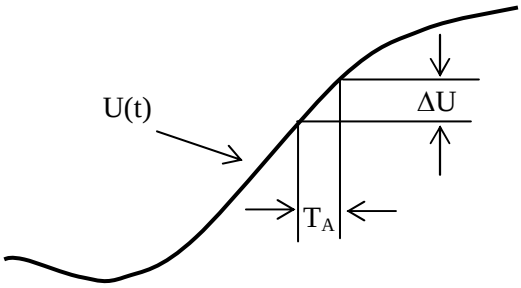


Fig. 10.1 Possible block diagram of digital signal processing

An Analog to Digital converter is an electronic circuit which accepts an analog input signal (usually a voltage $U(t)$) and produces a corresponding digital number at the output – see Fig. 10.1. The resultant digital word goes to a computer data bus or to the input of a digital circuit.

The analog-to-digital converter requires a small amount of time to perform the quantizing and coding operations. The time required to make the conversion depends on: the converter resolution, the conversion technique, and the speed of the components employed in the converter. The conversion speed required for a particular application depends on the time variation of the signal to be converted and on the accuracy desired.



$$\Delta U = \frac{dU(t)}{dt} \cdot T_A$$

T_A – aperture time

ΔU – amplitude uncertainty

Fig. 10.2 Aperture time T_A and amplitude uncertainty

Aperture time **-** refers to the time uncertainty (or time window) in making a measurement and results in an amplitude uncertainty (or error) in the measurement if the signal is changing during this time – Fig. 10.2.

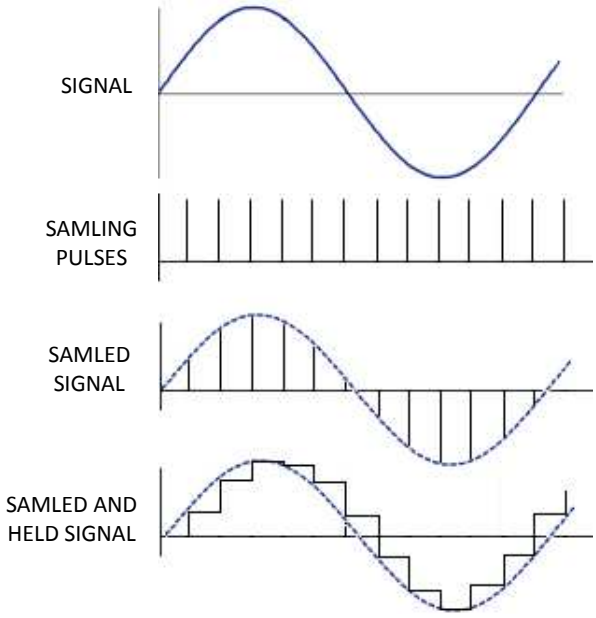


Fig. 10.3 Signal sampling

For the specific case of a sinusoidal input signal, the *maximum rate of change* occurs at the zero crossing of the waveform, and the amplitude error is

$$\Delta U = T_A \cdot \frac{d}{dt}(U_m \sin \omega t) = T_A \cdot (U_m \omega \cos \omega t) = \Big|_{\omega=0}\Big| = T_A \cdot U_m \cdot \omega$$

The resultant error as a fraction of the peak to peak full scale value is

$$\varepsilon = \frac{\Delta U}{2U_m} = \frac{T_A U_m \omega}{2U_m} = \frac{T_A U_m 2\pi f}{2U_m} = \pi f T_A$$

From this result the aperture time required to digitize a 1 kHz signal to 10 bits resolution can be found. The resolution required is one part in 2^{10} or approximately 0.001, thus

$$T_A = \frac{\varepsilon}{\pi f} = \frac{0{,}001}{3{,}14 \cdot 10^3} \cong 318 \cdot 10^{-9}$$

The result is a required aperture time of just 318 ns! It is evident that it is hard to find a 10-bit A/D converter to perform this conversion at any price! Fortunately, there is a simple and inexpensive way around this dilemma by using a sample-hold circuit – Fig. 10.4. The aperture time of the A/D converter is therefore greatly reduced by the much shorter aperture time of the sample-hold circuit. In turn, the aperture time of the sample-hold is a function of its bandwidth and switching time. We can see all "signal sampling" in Fig. 10.3.

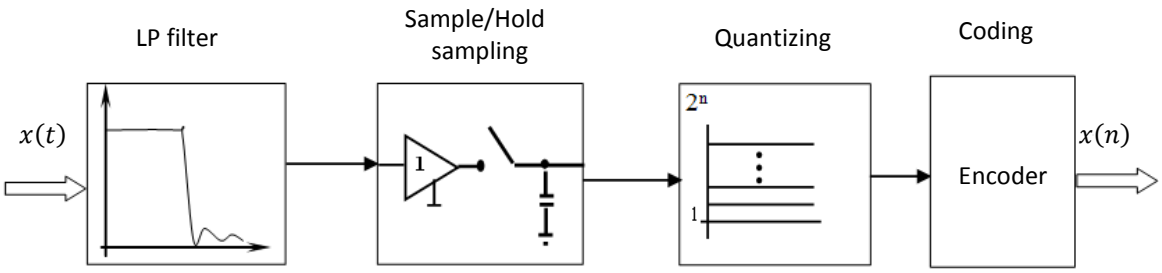


Fig. 10.4 Once more ADC

- **The electrical signal is sampled** (a **sample-hold circuit** acquires the signal voltage and then holds its value while an analog-to-digital converter converts the value into digital form)
- **Quantizing is the process** of transforming an analog signal into a set of discrete output states - there are $2^n - 1$ analog decision points (or threshold levels) in the transfer function (n – number of the quantizer bits)
- Each threshold level corresponds to a number (**code**).

LP filter **–** a low pass active filter which reduces high frequency signal components, unwanted electrical interference noise, or electronic noise from the signal; so-called antialiasing filter; its characteristics frequency must be ½ of sampling frequency f_s.

Sampling **– (sampling theorem, Nyquist theorem):**
If a continuous bandwidth-limited signal contains no frequency components higher than f_c, then the original signal can be recovered without distortion if it is sampled at a rate of at least $2f_c$ samples per second, thus

$$f_s \geq 2f_c$$

Aliasing in the time domain

There is an example of aliasing in Fig. 10.5.

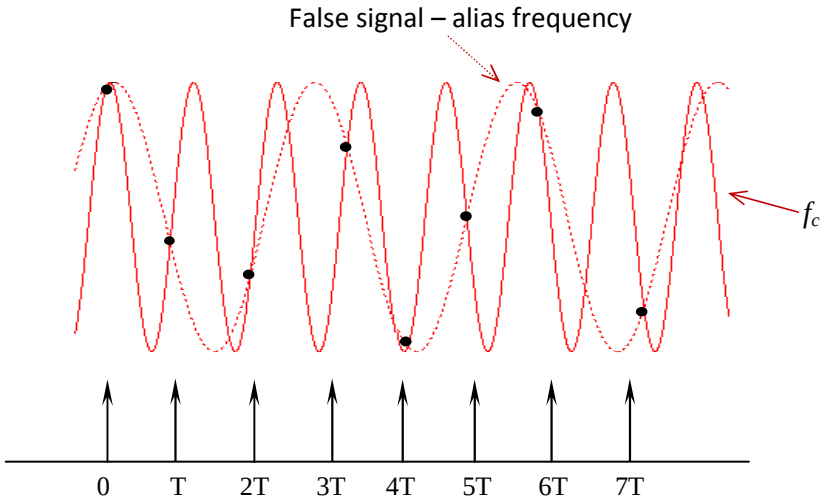


Fig. 10.5 Aliasing – in the time domain; $f_s < f_c$

Aliasing in the frequency domain – spectrum

We can see aliasing in the frequency domain in Fig. 10.6.

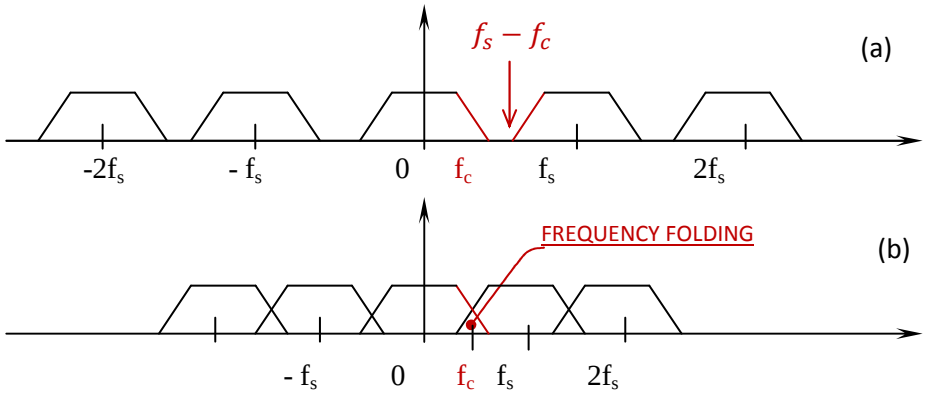


Fig. 10.6 a) right sampling – $f_s > 2f_c$; practically at least 5x higher
b) bad sampling – $f_s < 2f_c$

From the figure 10.6, if the sampling rate is increased such that $f_S - f_C > f_C$, then the two spectra are separated and the original signal can be recovered without distortion. This demonstrates the result of the ***Sampling Theorem*** that $f_S > 2f_C$. Frequency folding (aliasing) can be eliminated in two ways: first by using a high enough sampling rate, and second by filtering the signal before sampling to limit its bandwidth to $f_S/2$ – antialiasing filter.

Quantizer, coding

At any part of the input range of the quantizer, there is a small range of analog values within which the same output code word is produced. This small range is the voltage difference between any two adjacent decision points and is known as the analog quantization size, or quantum, q – it is found in general by dividing the full scale analog range (FSR) by the number of output states. FSR is defined by the applied reference voltage $U_{REF.}$ q is the smallest analog difference which can be resolved, or distinguished by the quantizer (the quantization step; quantum - analog).

$$q = \frac{U_{REF}}{2^n - 1} \approx \frac{U_{REF}}{2^n}$$

For a given analog input value to the quantizer, the output error will vary anywhere from 0 to $\pm q/2$; the error is zero only at analog values corresponding to the code center points. This error is also frequently called quantization uncertainty or quantization noise.

The quantizer output can be thought of as the analog input with quantization noise added to it. The noise has a peak-to-peak value of q, but as with other types of noise, the average value is zero. Its RMS value, however, is useful in analysis and can be computed from the triangular waveshape to be $q/(2 \cdot \sqrt{3})$.

The most popular **code** is natural binary, or straight binary, which is used in its fractional form to represent a number

$$N = a_1 \cdot 2^{-1} + a_2 \cdot 2^{-2} + \cdots + a_n \cdot 2^{-n}$$

where each coefficient "a_i" assumes a value of zero or one and N has a value between zero and one.

The **binary code word** 110101 therefore represents the decimal fraction

$$1 \cdot 0{,}5 + 1 \cdot 0{,}25 + 0 \cdot 0{,}125 + 1 \cdot 0{,}0625 + 0 \cdot 0{,}03125 + 1 \cdot 0{,}015625 = 0{,}828125$$

or $82{,}8125\%$ of full scale for the converter. If full scale (U_{REF}) is $+10$ V, then the code word represents $+8{,}28125$ V.

The leftmost bit has the most weight, $0{,}5$ of full scale, and is called the **most significant bit**, or **MSB**; the rightmost bit has the least weight, $2 - n$ of full scale, and is therefore called the **least significant bit**, or **LSB**. The bits in a code word are numbered from left to right from 1 to n.

The LSB has the same analog equivalent value as q discussed previously, namely

$$\textbf{LSB (analog value)} = q = \frac{U_{REF}}{2^n - 1} \approx \frac{U_{REF}}{2^n}$$

Table 1 and Table 2 are useful summaries of the resolution, number of states, and LSB weights.

Table 1:

U_{REF} (V)	n	$LSB(q)$
1.00	8	3.9062 mV
1.00	12	244.14 µV
2.00	8	7.8125 mV
2.00	10	1.9531 mV
2.00	12	488.28 µV
2.048	10	2.0000 mV
2.048	12	500.00 µV
4.00	8	15.625 mV
4.00	10	3.9062 mV
4.00	12	976.56 µV

Table 2:

$Resolution = 2^n - 1$	$[n = number\ of\ bites]$	$error\ (FSR)$
n	2^n	1 bit ppm $[1 \cdot 10^{-6}]$
8bits	256	3906
10bits	1 024	976
12bits	4 096	244
14bits	16 384	61
16bits	65 536	15
18bits	262 144	3.8
20bits	1 048 576	0.95
22bits	4 194 304	0.24
24bits	16 777 216	0.06

The dynamic range $DR(dB)$ of a data converter in dB is found as follows:

$$DR(dB) = 20 \log 2^n = 20n \log 2 = 20n(0{,}3010) = 6{,}2n$$

A 12-bit converter, for example, has a dynamic range of $72{,}2\ dB$.

10.2 Basic considerations - DAC

A Digital to Analog converter is an electronic circuit which accepts a digital number at its input and produces a corresponding analog signal (usually a voltage) at the output.

There are different DACs realizations.

10.2.1 Summation of binary weighted currents

Summation of binary weighted currents is shown in Fig. 10.7.

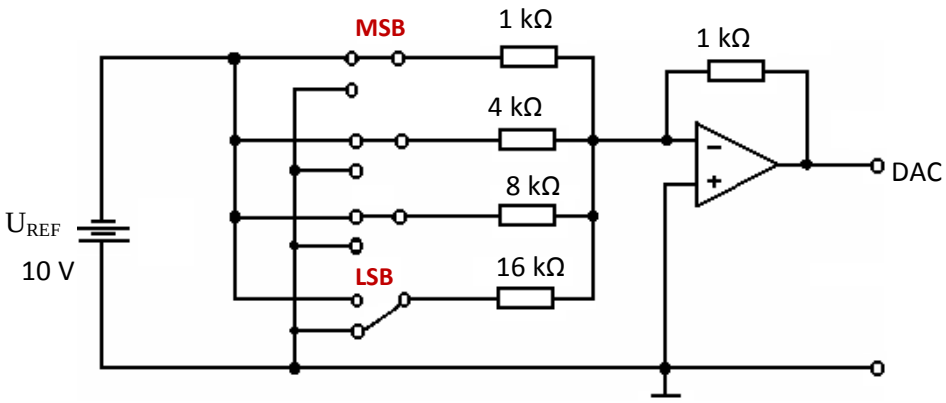


Fig. 10.7 DAC – summation of weighted currents

- It is a summing connection (inverting) of an operational amplifier. For given ratios the output voltage U_{DAC} is:

$$U_{DAC} = -10 \cdot \left(1 \cdot \frac{10^3}{2 \cdot 10^3} + 1 \cdot \frac{10^3}{4 \cdot 10^3} + 1 \cdot \frac{10^3}{8 \cdot 10^3} + 0 \cdot \frac{10^3}{16 \cdot 10^3}\right) = -8{,}75\ \text{V}$$

- For example, in a 12-bit converter you would need a range of resistor values of $2000:1$, with corresponding precision of the small resistor values – an elegant solution is $R-2R$ ladder.

- The switch resistance must be smaller than $1/2^n$ of the smallest resistor.

- This principle is used only in fast, low-precision DACs.

10.2.2 R–2R ladder

$R-2R$ ladder is shown in Fig. 10.8

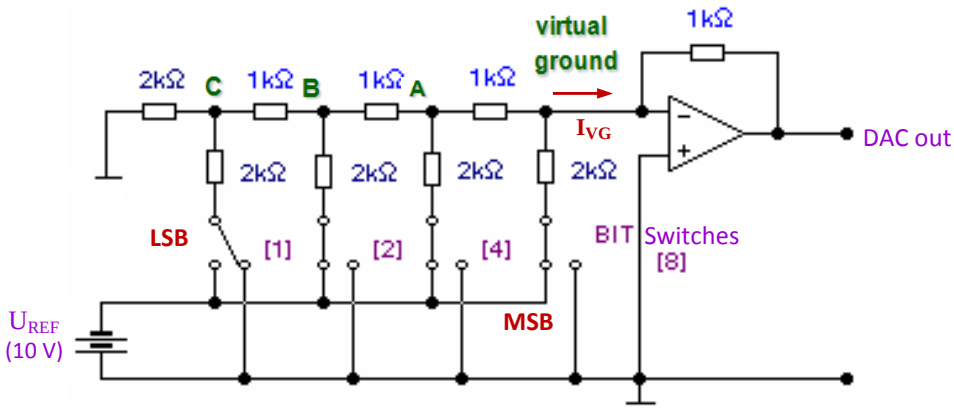


Fig. 10.8 $R-2R$ ladder

- Only two resistor values are needed (R and $2R$).
- The resistor must be precisely matched, though the actual resistor values are not critical.
- Electronic switches connect resistors to either ground or the U_{REF} line. The result is binary weighted current I_{VG} flowing down VIRTUAL GROUND (The Thevenin resistance of an $R/2R$ ladder is always R – regardless of the number of bits in the ladder. Since an $R/2R$ ladder is a linear circuit, we can apply the principle of superposition to calculate I_{VG}.).
- Operational amplifier – current – to – voltage converter (inverting).
- The circuit shown generates an output of zero to – 10 V (the maximum input count is 15, with output voltage -10x15/16).
- The operational amplifier tends to be the slowest part of the DAC – we can use a converter with current output.

10.2.3 DAC with current output

DAC with current output (current switched DAC) is in Fig. 10.9.

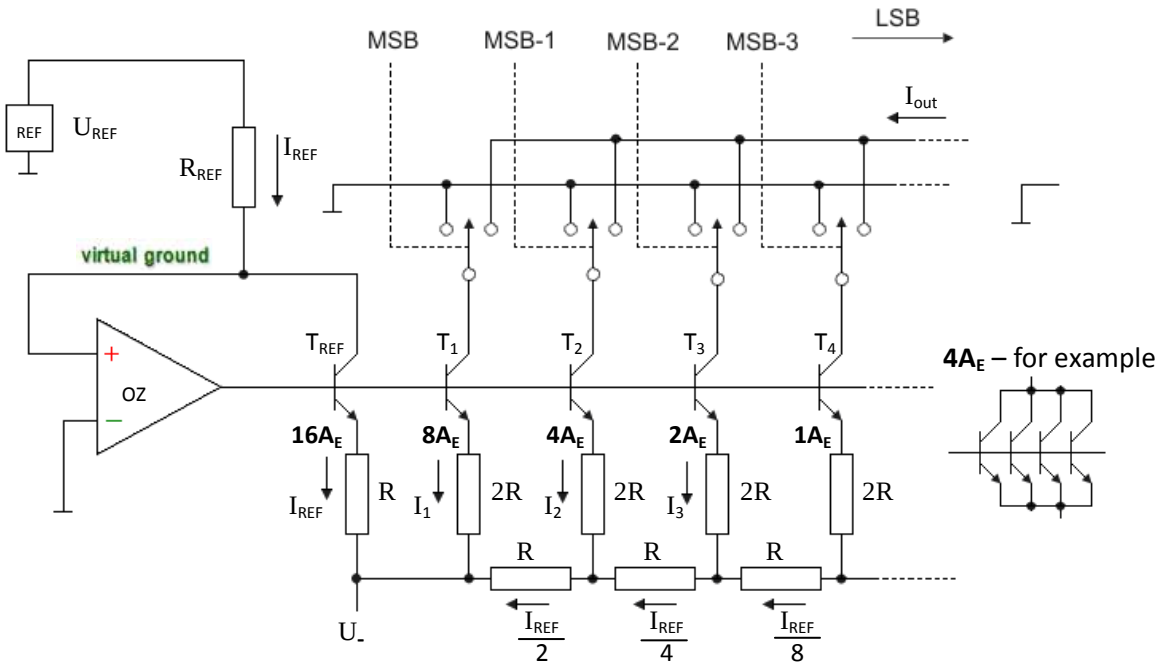


Fig. 10.9 $R-2R$ ladder; current output

- The current sources are ON all the time, and their output current is switched to the output terminal or to the ground (very fast and cheap).

- The transistor areas are scaled ($16A_E$: $8A_E$: $4A_E$: $2A_E$:$1A_E$), thereby ensuring equal current densities in all the transistors for optimum U_{BE} matching (the emitter areas of the BJT devices must be proportional to the emitter current).
- $OZ + T_{REF}$ – create current $I_{REF} = U_{REF}/R_{REF}$ (T_{REF} is an inverting structure, thus feedback is negative, virtual ground).
- $I_1 = \frac{(R \cdot I_{REF} + U_{BE}) - U_{BE}}{2R} = \frac{I_{REF}}{2}$
- $I_2 = \frac{(R \cdot I_{REF} + U_{BE}) - U_{BE} - \frac{R \cdot I_{REF}}{2}}{2R} = \frac{I_{REF}}{4}$
- $I_3 = \frac{(R \cdot I_{REF} + U_{BE}) - U_{BE} - \frac{R \cdot I_{REF}}{2} - \frac{R \cdot I_{REF}}{4}}{2R} = \frac{I_{REF}}{8}$
- etc.

There are a few ways to generate an output voltage from a current DAC. To generate large swings, or to buffer into small load resistances, a current to voltage amplifier (with an op amp) can be used.

10.2.4 Multiplying DACs (MDAC)

Multiplying DACs (MDAC) - can be made from DACs that have no internal reference by using the reference input for the analog input signal. A DAC with good multiplying properties (wide analog input range, high speed, etc.) will be called a “multiplying DAC”.

When used like this, MDAC behaves as a digitally controlled audio attenuator because the output U_o is a fraction of the voltage representing the input digital code and the attenuator setting can be controlled by digital logic. If followed by an op amp integrator, the MDAC provides digitally programmable integration which can be used in the design of digitally programmable oscillators, filters.

The output voltage generally is

$$U_{DAC} = \pm U_{REF} \cdot N = \pm U_{REF} \cdot (a_1 \cdot 2^{-1} + a_2 \cdot 2^{-2} + \cdots + a_n \cdot 2^{-n})$$

where

Sign $+$ means a noninverting DAC.
Sign $-$ means an inverting DAC.
a_1 – MSB; a_n – LSB.

10.3 Basic principles od ADCs

There are different ADCs realizations.

10.3.1 Flash conversion (parallel encoder)

Of all the conversion techniques, one of the fastest is direct conversion, better known as "flash" conversion. ADCs based on this architecture are extremely fast and perform their multibit conversion directly, but they require intensive analog design to manage the large number of comparators and reference voltages required. As shown in Fig. 10.10, a converter with N-bit resolution has $2^N - 1$ comparators connected in parallel, with reference voltages set by a resistor network and spaced $U_{REF}/2^N$ ($\sim$1 least significant bit, or LSB) apart. Input voltage offset of operational amplifiers must be less than the "$LSB/2$".

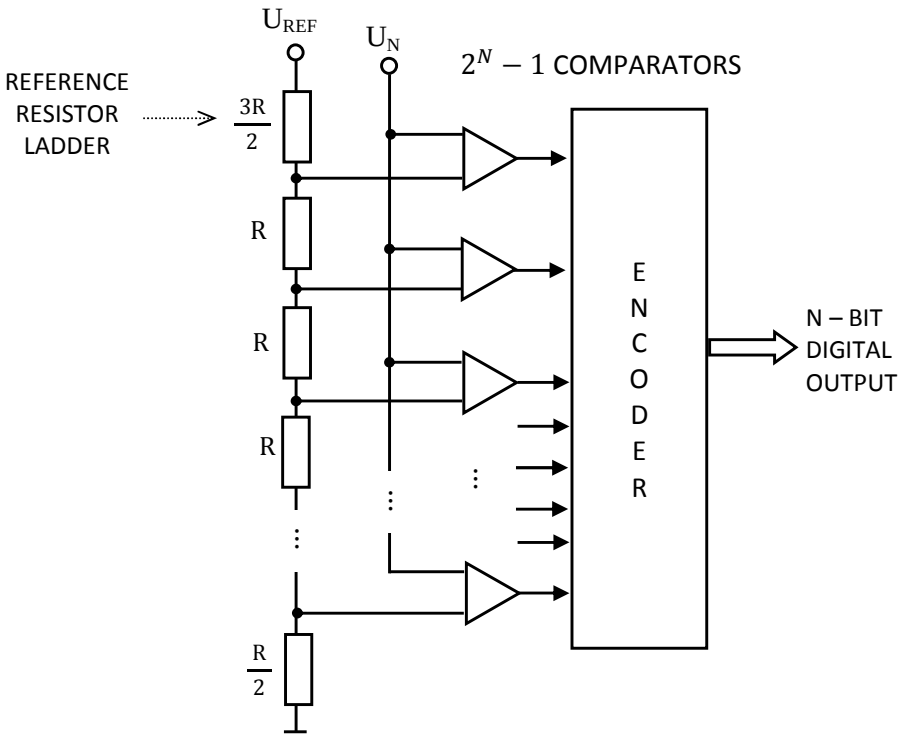


Fig. 10.10 ADCs based on the direct-conversion architecture (better known as flash converters) include $2^N - 1$ comparator banks and a reference resistor-divider network

A change of input voltage usually causes a change of state in more than one comparator output. These output changes are combined in a decoder-logic unit that produces a parallel N-bit output from the converter.

10.3.2 Dual slope ADCs

The dual-slope ADC architecture was truly a breakthrough in ADCs for high resolution applications such as digital voltmeters (DVMs), etc. A simplified diagram is shown in Fig. 10.11 and the integrator output waveforms are shown in Fig. 10.12.

The input signal is applied to an integrator; at the same time a counter is started, counting clock pulses. After predetermined amount of time (T), a reference voltage having opposite

polarity is applied to the integrator. At that instant, the accumulated charge on the integrating capacitor is proportional to the ***average value of the input over the interval T***. By choosing that time interval to be a multiple of the power-line period, the converter becomes insensitive to 50 Hz "hum" (and its harmonics) – Fig. 10.13. The integral of the reference is an opposite-going ramp having a slope of U_{REF}/RC. At the same time, the counter is again counting from zero. When the integrator output reaches zero, the count is stopped, and the analog circuitry is reset. Since the charge gained is proportional to $U_{IN} \cdot T$, and the equal amount of charge lost is proportional to $U_{REF} \cdot t_x$, then the number of counts relative to the full scale count is proportional to t_x/T, or U_{IN}/U_{REF}. If the output of the counter is a binary number, it will therefore be a binary representation of the input voltage.

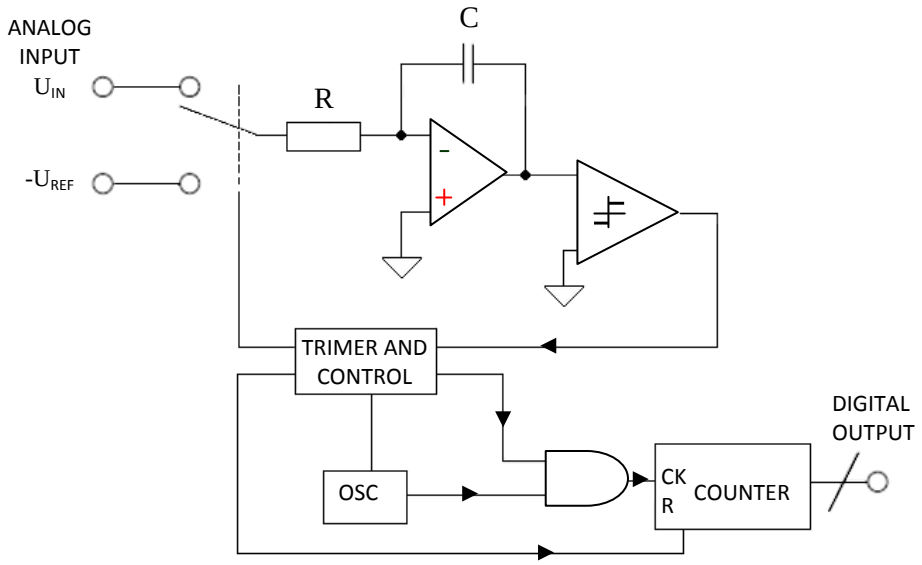
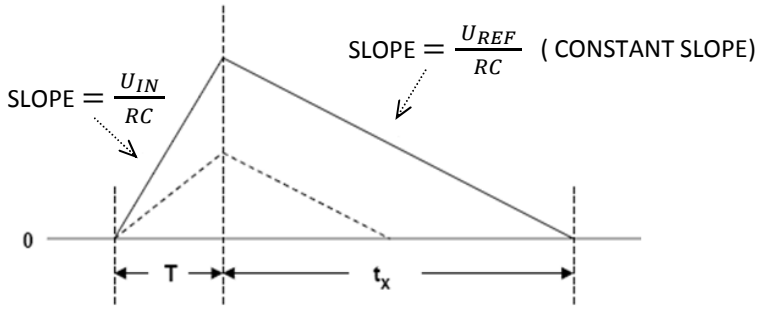


Fig. 10.11 Dual slope ADC



$$\frac{U_{IN}}{RC} \cdot T = \frac{U_{REF}}{RC} \cdot t_x$$

$$t_x = \frac{U_{IN}}{U_{REF}} \cdot T$$

Fig. 10.12 Dual slope ADC – output waveforms

Dual-slope integration has many advantages. Conversion accuracy is independent of both the capacitance and the clock frequency, because they affect both the up-slope and the down-slope by the same ratio.

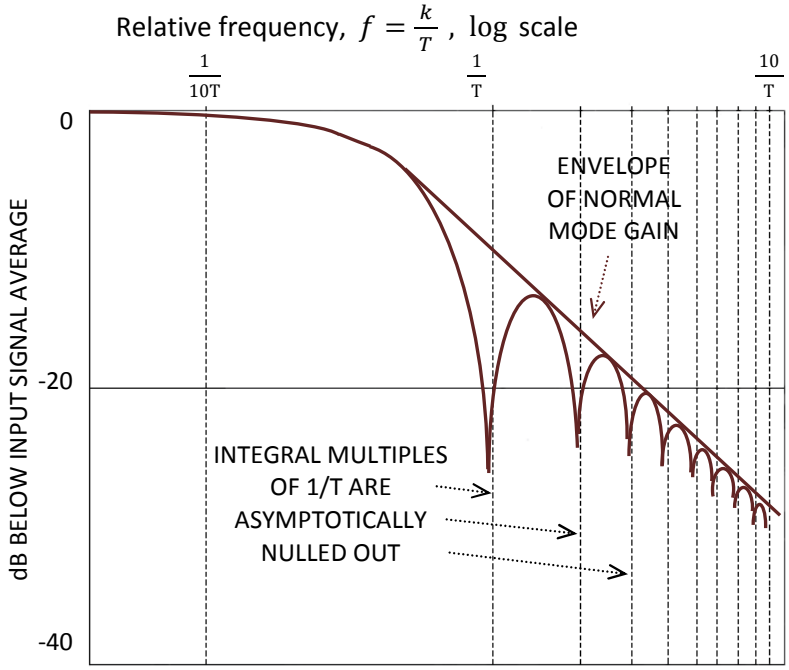


Fig. 10.13 Frequency response of integrating ADC

10.3.3 SAR ADCs

Although there are many variations in the implementation of a SAR (successive approximation register) ADC, the basic architecture is quite simple (see Fig. 10.14). The analog input voltage (U_{IN}) is held on a track/hold. To implement the binary search algorithm, the N-bit register is first set to midscale (that is, 100... .00, where the MSB is set to '1'). This forces the DAC output (U_{DAC}) to be $U_{REF}/2$, where U_{REF} is the reference voltage provided to the ADC. A comparison is then performed to determine if U_{IN} is less than or greater than U_{DAC}. If U_{IN} is greater than U_{DAC}, the comparator output is a logic high or '1' and the MSB of the N-bit register remains at '1'. Conversely, if U_{IN} is less than U_{DAC}, the comparator output is a logic low and the MSB of the register is cleared to logic '0'. The SAR control logic then moves to the next bit down, forces that bit high, and does another comparison. The sequence continues all the way down to the LSB. Once this is done, the conversion is complete, and the N-bit digital word is available in the register.

Fig. 10.15 shows an example of a 4-bit conversion. The y-axis (and the bold line in the figure) represents the DAC output voltage. In the example, the first comparison shows that $U_{IN} < U_{DAC}$. Thus, bit 3 is set to '0'. The DAC is then set to 0100_2 and the second comparison is performed. As $U_{IN} > U_{DAC}$, bit 2 remains at '1'. The DAC is then set to 0110_2, and the third comparison is performed. Bit 1 is set to '0', and the DAC is then set to 0101_2 for the final comparison. Finally, bit 0 remains at '1' because $U_{IN} > U_{DAC}$.

Notice that four comparison periods are required for a 4-bit ADC. Generally speaking, an N-bit SAR ADC will require N comparison periods and will not be ready for the next conversion until the current one is complete. This explains why these types of ADCs are power- and space-efficient, yet are rarely seen in speed-and resolution combinations beyond a few Msps at 14 to 16 bits. The two critical components are the comparator and the DAC.

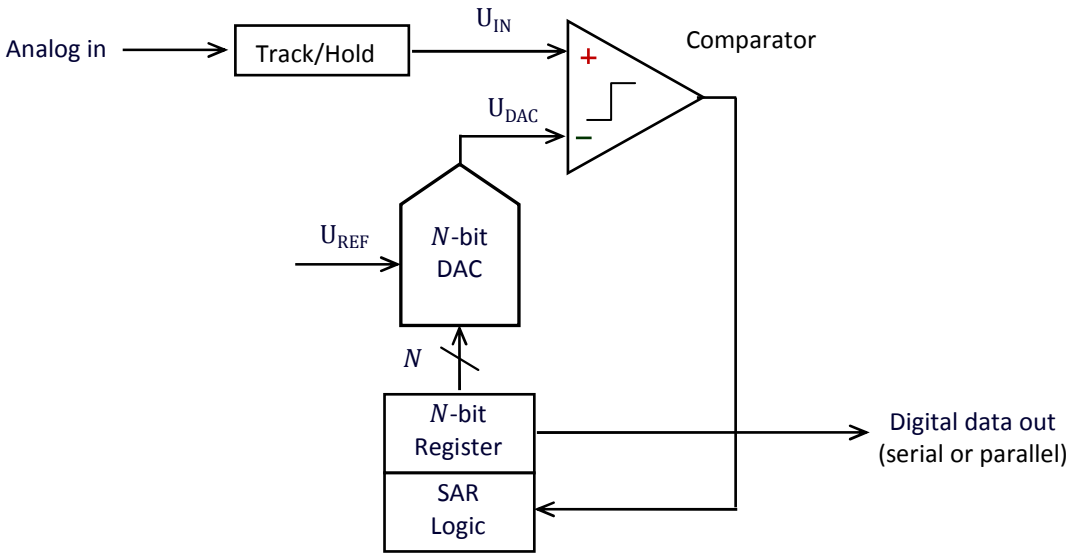


Fig. 10.14 Simplified N − bit SAR architecture

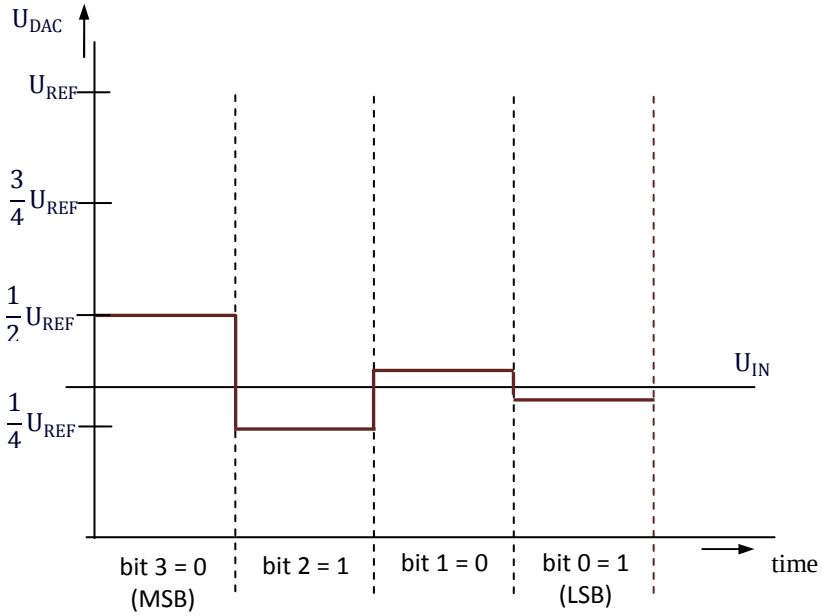


Fig. 10.15 SAR operation – 4 – bit example

10.3.4 Voltage – to – frequency ADC

Voltage-to-frequency ADCs convert the analog input voltage to a pulse train with the frequency proportional to the amplitude of the input (see Fig. 10.16). This can be done simply by charging a capacitor with a current proportional to the input level and discharging it when the ramp reaches a preset threshold. The pulses are counted over a fixed period to determine the frequency, and the pulse counter output, in turn, represents the digital voltage.

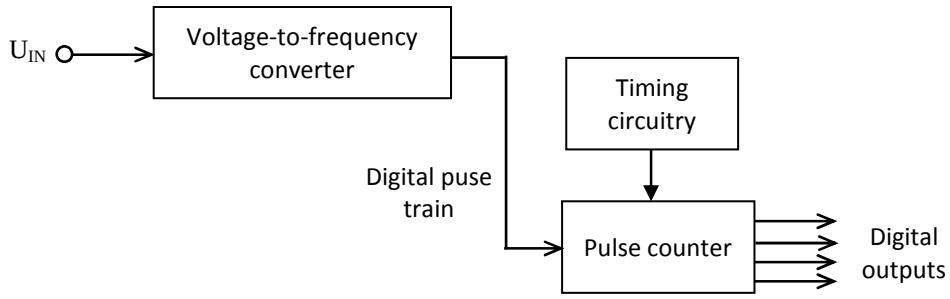


Fig. 16 Principle of voltage – to – frequency DACs

Voltage-to-frequency converters inherently have a high noise rejection characteristic, because the input signal is effectively integrated over the counting interval. Voltage-to-frequency conversion is commonly used to convert slow and noisy signals. Voltage-to-frequency ADCs are also widely used for remote sensing in noisy environments. The input voltage is converted to a frequency at the remote location and the digital pulse train is transmitted over a pair of wires to the counter. This eliminates noise that can be introduced in the transmission lines of an analog signal over a relatively long distance.

10.3.5 Delta sigma converter

Delta sigma converter – see the 9. Chapter

10.3.6 Pipelined analog-to-digital converters

The pipelined analog-to-digital converter (ADC) has become the most popular ADC architecture for sampling rates from a few megasamples per second (Msps) up to 100Msps. Resolutions range from eight bits at the faster sample rates up to 16 bits at the lower rates. These resolutions and sampling rates cover a wide range of applications, including CCD imaging, ultrasonic medical imaging, digital receivers, base stations, digital video, etc.

Fig. 10.17 shows a block diagram of a 12-bit pipelined ADC. In this schematic, the analog input, U_{IN}, is first sampled and held steady by a sample-and-hold (S&H), while the flash ADC in stage one quantizes it to three bits. The 3-bit output is then fed to a 3-bit DAC (accurate to about 12 bits), and the analog output is subtracted from the input. This "residue" is then gained up by a factor of four and fed to the next stage (Stage 2). This gained-up residue

continues through the pipeline, providing three bits per stage until it reaches the 4-bit flash ADC, which resolves the last 4LSB bits. Because the bits from each stage are determined at different points in time, all the bits corresponding to the same sample are time-aligned with shift registers before being fed to the digital-error-correction logic. Note when a stage finishes processing a sample, determining the bits, and passing the residue to the next stage, it can then start processing the next sample received from the sample-and-hold embedded within each stage. This pipelining action is the reason for the high throughput.

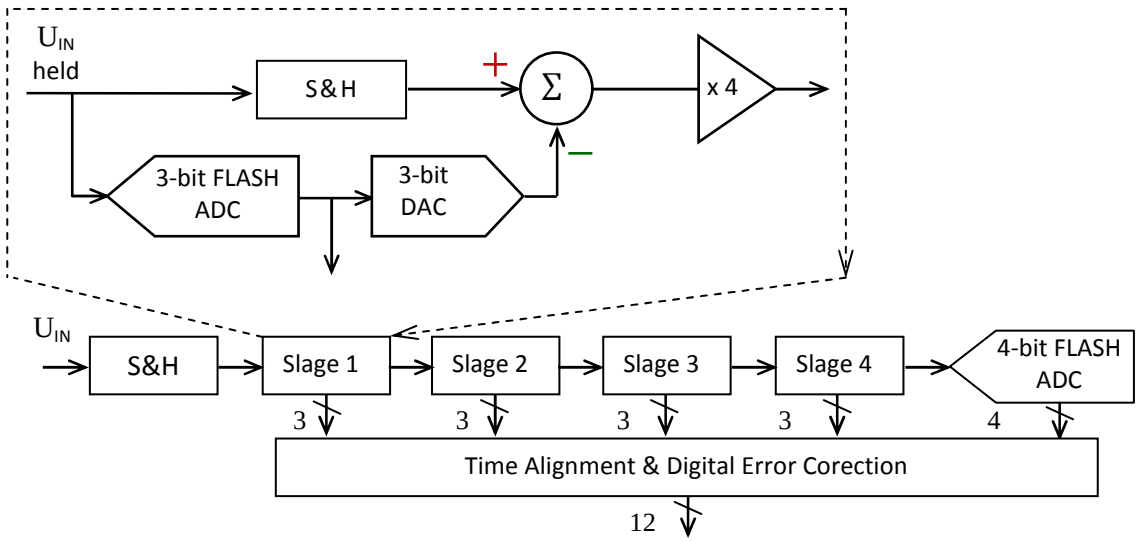


Fig. 10.17 Pipelined ADC with four 3-bit stages (each stage resolves two bits)

Although each stage generates three raw bits in the Figure 10.17 example, because the interstage gain is only 4, each stage (Stages 1 to 4) effectively resolves only two bits. The extra bit is simply to reduce the size of the residue by one half, allowing extra range in the next 3-bit ADC for digital error correction, as mentioned above. This process is called "1-bit overlap" between adjacent stages. The effective number of bits of the entire ADC is therefore $2 + 2 + 2 + 2 + 4 = 12$ bits.

10.3.7 Comparison of ADCs

As ADCs can consume a large percentage of power in a device, it is of vital interest to minimize ADC power consumption. We can see optimal power consumption for different sampling rates and resolutions in Fig. 10.18 (qualitative description only).

When selecting an ADC, some of the factors to consider are

- Precision
- Speed
- Accuracy (external trimming required? monotonicity?)
- Required supply voltages and power dissipation
- Reference (internal or external? if internal, is it accessible externally?)

- Input impedance and analog voltage range (unipolar, bipolar, or both?)
- etc.

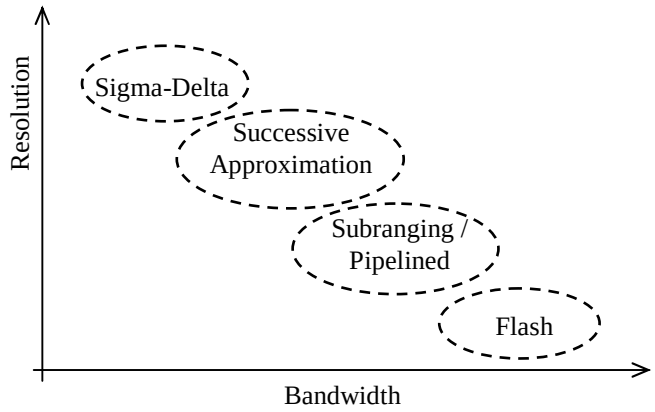


Fig. 10.18 A/D Converter technologies, resolution and bandwidth

Summary

1) An Analog to Digital converter is an electronic circuit which accepts an analog input signal and produces a corresponding digital number at the output.
2) A Digital to Analog converter is an electronic circuit which accepts a digital number on its input and produces a corresponding analog signal at the output.
3) Almost all converters need some “analog conditioning” which is application dependent.
4) A waveform is sampled at a constant rate T.
5) Sampling rate
    a. Nyquist rate = 2x highest frequency of interest
    b. Practically – always sample at least 5x
    c. Ensure ADCs have input filtering (antialiasing) where necessary (large hf signals)
    d. Filter DAC outputs to remove higher frequencies and switching “glitches”
6) An aperture time of ADCs is reduced by the much shorter aperture time of a sample – hold circuit.
7) LP filter reduces high frequency signal components.
8) Performance cannot be better than the REFERENCE.

Questions 10

You can find the answers in this text.

1. Explain basic operating principles of DACs.
2. Explain basic operating principles of ADCs.
3. Can we use DAC as a variable controlled attenuator?
4. Explain the operating principle of an ADC, usually referred as a SAR.
5. Explain what aliasing is, how it happens, and what may be done to prevent it from happening to an ADC circuit.

Problems 10

Example 10.1

Determine the required sampling frequency f_s if the maximum signal frequency is 16 kHz

Example 10.2

Suppose an analog-digital converter inputs a voltage ranging from 0 to 5 volts DC and converts the magnitude of that voltage into an 8-bit binary number. How many discrete "steps" are there in the output as the converter circuit resolves the input voltage from one end of its range (0 volts) to the other one (5 volts)? How much voltage does each of these steps represent?

Example 10.3

Determine the output voltage of a multiplying DAC (inverting), which $\boldsymbol{U_{REF} = 10}$ V, a binary word is 10001001

Example 10.4

Determine the maximum input voltage offset of comparators in parallel converter according to problem 3.

PROBLEMS KEY 10

Ad example 10.1)

Use the equation $f_s \geq 2f_c$

Ad example 10.2)

This ADC (Analog-to-Digital Converter) circuit has 256 steps in its output range, each step representing 19.61 mV

Ad example 10.3)

Use the equation

$$U_{DAC} = \pm U_{REF} \cdot N = \pm U_{REF} \cdot (a_1 \cdot 2^{-1} + a_2 \cdot 2^{-2} + \cdots + a_n \cdot 2^{-n});\ \ n = 8$$

Ad example 10.4)

We know $n = 8,\ \mathrm{LSB}(\text{analog value}) = q = \frac{U_{REF}}{2^n - 1} \approx \frac{U_{REF}}{2^n}$; thus comparators input voltage offset must be less than $q/2$.

Basic texts

[1] Gray, N.: ABCs of ADCs. National Semiconductor, November 24, 2003

[2] Koster, W.: Basic DAC Architectures II: Binary DACs. ANALOG DEVICES, MT – 015 TUTORIAL, 2009

[3] Principles of Data Acquisition and Conversion. Intersil, Application Note AN002, October 1986

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). pp. 612÷636 Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

11. Amplitude stabilization of the harmonic oscillator

Time of study: 3 hours

Goals: the student should be able to

- define basic principles of oscillator amplitude stabilization
- describe low distortion oscillator (spot sinus)

EXPLANATION

You can see two basic oscillator structures in the 6th chapter. A positive-feedback loop is formed by an amplifier A and a frequency-selective network β. The amplifier produces a 0° or 180° voltage phase shift, as does the feedback network. This results in a combined 0° voltage phase shift around the loop, which is the same thing as a 360° phase shift. In order to oscillate, the loop gain (return ration) $A\beta$ must be equal to unity.

- Because circuit components and transistors change characteristics (drift) with age, temperature, voltage, etc., the $A\beta = 1$ condition could not be permanently satisfied. The product will become either less or larger than unity. In the former case the oscillation simply stops, and in the latter case nonlinearity is required in order to limit the amplitude.

- An oscillator in which the loop gain is exactly unity is an abstraction completely unrealizable in practice.

- In every practical oscillator the loop gain is slightly larger than unity, and the amplitude of the oscillations is limited by the onset of nonlinearity. The distortion is low if the amplitude of oscillation remains within the linear region of the amplifier. It must not be allowed to go into a full-swing oscillation.

An amplitude-limiting mechanism is basically an automatic gain control (AGC) circuit that forces the amplifier gain to decrease when the amplitude of the oscillation increases.

11.1 Wien bridge oscillator

The positive feedback network ($RC + RC$) is frequency – selective, and at the most favored frequency it passes a maximum of 1/3 of the output swing back to the + input – Fig. 11.1. The negative feedback (560 Ω + lamp; R_{lamp} grows with current - positive temperature coefficient (PTC) thermistor) adjusts the gain (noninverting amplifier, $1 + 560/R_{lamp}$). The incandescent lamp is used as a variable-resistance element (with a long time-constant of response; the lamp is rated at 14 mA and 10V).

The initial gain (not yet oscillation, the lamp is cold, R_{lamp} is small) is greater than 3 – the oscillator begins to oscillate. As the output level rises, the lamp heats slightly, reducing the gain to 3.

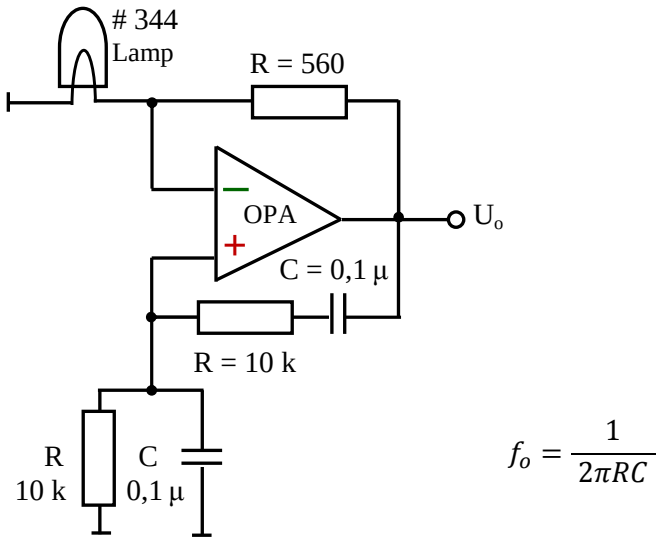


Fig.11.1 Wien bridge oscillator – amplitude stabilization (LAMP)

Fig. 11.2 shows Wien-bridge oscillators with diode amplitude limiting mechanisms; when the diodes are off, the gain is $1 + R_2 \| R_1$; and when the diode is on, the gain is reduced to $1 + (R_2 \| R_3)/R_1$. The start up condition requires a gain slightly greater than 3 or

$$\frac{R_2}{R_1} > 2$$

The inequality above can be satisfied by making it equal to a value between 2.1 to 2.2.

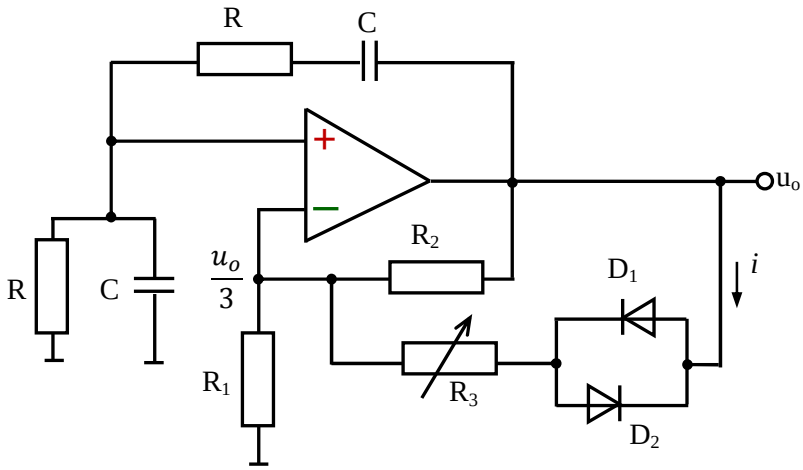


Fig. 11.2 Wien bridge oscillator – amplitude stabilization (diodes)

When a diode is on, the gain should be slightly less than 3, or

$$\frac{R_2 \| R_3}{R_1} < 2$$

The inequality above can be satisfied by making it equal to a value between 1.8 to 1.9.

When a diode is conducting, the amplitude of the output voltage is limited. Since $u_+ = u_- = u_o/3$, a nodal equation gives

$$\frac{u_o}{3R_1} = \frac{u_o - \frac{u_o}{3}}{R_2} + \frac{u_o - \frac{u_o}{3} - U_D}{R_3}$$

or

$$u_o = \frac{3U_D}{2\left(1+\frac{R_3}{R_2}\right) - \frac{R_3}{R_1}}$$

where $U_D > 0{,}5$ V for an actual diode.

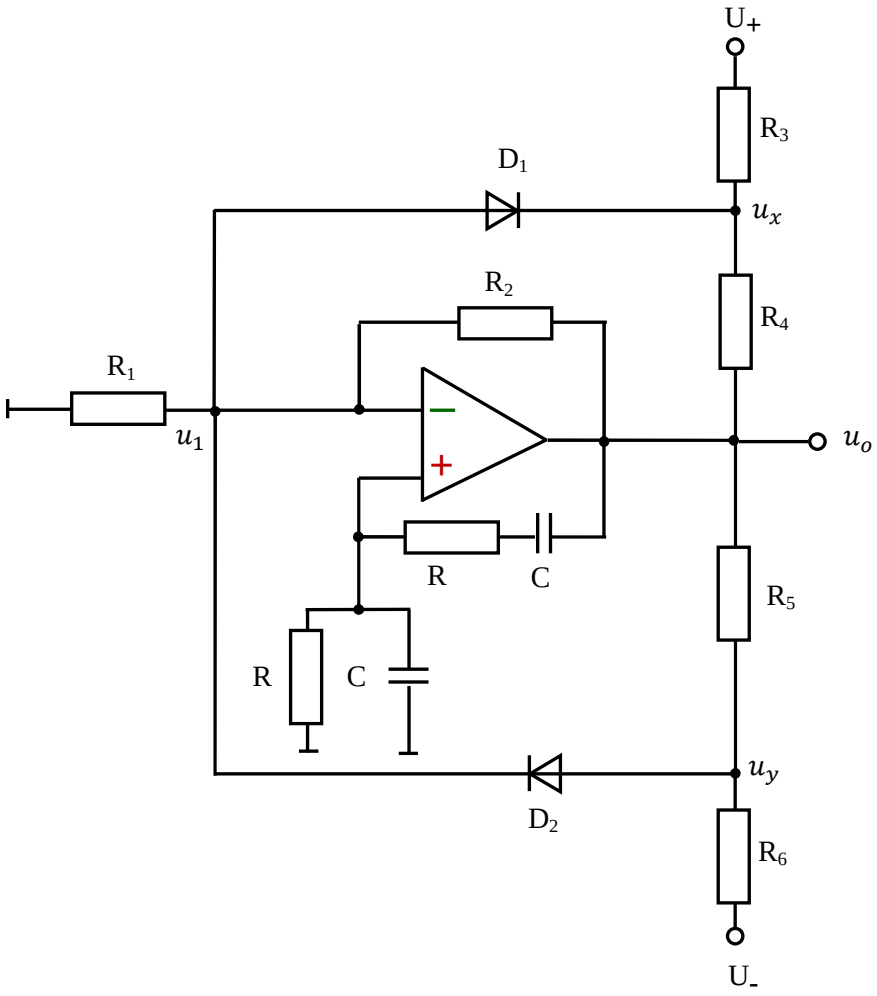


Fig. 11.3 Wien bridge oscillator – another amplitude stabilization (diodes)

Another Wien-bridge oscillator with a diode amplitude-limiting circuit is shown in Fig. 11.3. In this oscillator the amplitude-limiting circuit consists of the diodes D_1 and D_2, and the resistors R_3, R_4, R_5 and R_6. To understand the operation of the amplitude-limiting circuit, observe that as u_o increases, the voltage at node u_y will exceed the voltage u_1, forcing D_2 to conduct. When D_2 conducts, the value of u_y is $u_y = u_1 + 0{,}7$, and u_o is clamped at the value $u_{o(max)}$, given by (superposition theorem)

Since is approximately , it follows from equation above that

Similarly, as decreases, the voltage will drop below , forcing to conduct. When conducts, the voltage is and is clamped at the value , given by

The simultaneous solution of equations gives the value of the resistors that limit the output voltage to . In order to obtain a symmetrical sinusoidal voltage, the selection and is usually made.

Example 11.1

a) Design the Wien-bridge oscillator shown in Fig. 11.3 to oscillate at 5 kHz.

b) Design an amplitude-limiting circuit. The amplitude of the sinusoidal output voltage is to be limited to .

☑ Solution:

A practical value of 0,01 µF for the capacitors can be selected. Then, using

the value of is

kΩ

A practical value of 3 kΩ can be used in series with a trimming potentiometer to set the frequency of oscillation at 5 kHz.

To start the oscillation, a value of is used. From

a gain of 3,2 is obtained with kΩ and kΩ. The supply voltages of the op amp can be selected as V and V. The output voltage will reach saturation producing a clipping in the output waveform and, therefore, a significant amount of distortion. This occurs because the starting condition requires , and the gain of the amplifier changes when its output reaches saturation. Some sort of amplitude-limiting mechanism is needed to reduce the harmonic distortion.

With and , it follows from (;)

$$\frac{u_{o(max)}}{3} + 0{,}7 = \frac{u_{o(max)}R_6}{R_5+R_6} + \frac{(U_-)R_5}{R_5+R_6} \implies \frac{5}{3} + 0{,}7 = \frac{5R_3}{R_5+R_3} + \frac{-12R_5}{R_5+R_3}$$

$$\frac{u_{o(min)}}{3} - 0{,}7 = \frac{u_{o(min)}R_3}{R_3+R_4} + \frac{(U_+)R_4}{R_3+R_4} \implies -\frac{5}{3} - 0{,}7 = \frac{-5R_3}{R_5+R_3} + \frac{+12R_5}{R_5+R_3}$$

(the same equation as above)

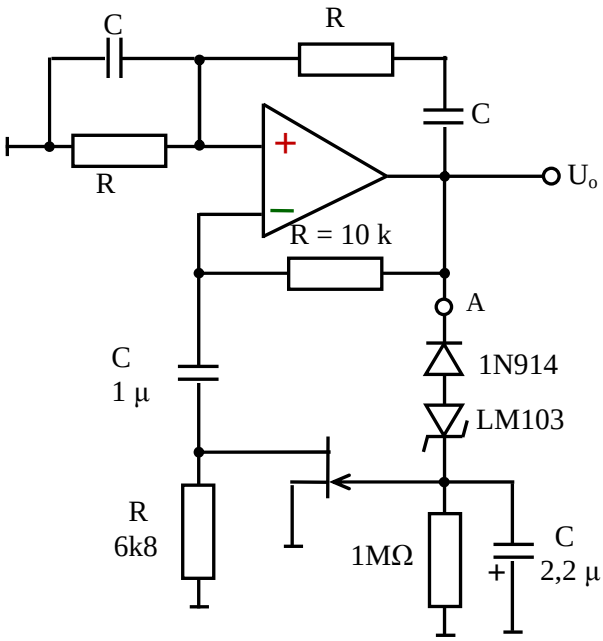
that

$$R_3 = \frac{431}{79} \cdot R_5$$

and if we choose $R_4 = R_5 = 2$ kΩ

we get $R_3 = 10{,}9$ kΩ.

In the circuit in Fig. 11.4, an amplitude discriminator consisting of the diodes and RC adjusts the AC gain by varying the resistance of the JFET, which behaves like a voltage-variable resistance for small voltages. The LM103 is a two-terminal monolithic reference diode electrically equivalent to a breakdown diode. The long time constant is used (2s) to avoid distortion, since fast feedback will distort the wave by attempting to control the amplitude within the time of one cycle.



$$f_o = \frac{1}{2\pi RC}$$

Fig. 11.4 Wien bridge oscillator – JFET amplitude stabilization

The small amplitude generates a small voltage (negative) – JFET is "open" (small resistor in parallel with 6,8 kΩ) – small JFET resistance determines the maximum gain of more than 3. As the amplitude grows, the voltage (negative, on RC) grows too and JFET closes – the minimum amplifier gain is about $1 + 10/6{,}8 = 2{,}47$.

We can add a buffer (voltage gain of 1) to isolate the non-linear effects of rectifier (diodes) from the oscillator output – Fig. 11.5 – and thus reduce the distortion of the oscillator.

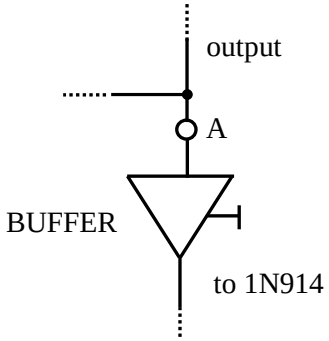


Fig. 11.5 Wien bridge oscillator – JFET amplitude stabilization; buffer reduces distortion of the oscillator

11.2 Phase shift oscillator

Phase-shift oscillators usually use RC networks in the feedback path. The op amp is used in an inverting configuration with a gain of $-R_2/R_1$. Thus, the signal experiences a phase shift of -180° through the amplifier, and the phase shift from each RC section is 60° at the frequency of oscillation, for a total phase shift in the feedback path of 180°.

In this oscillator the RC sections are connected without isolation and, therefore, there is loading. In the last stage the resistors R and R_1 appear in parallel. The loading of R_1 can be neglected if $R_1 \| R \gg R$, or in some cases by removing R in the third stage and letting $R_1 = R$.

To summarize, the phase-shift oscillator in Fig. 11.6 will oscillate at the frequency ω_o given by

$$\omega = \omega_o = \frac{1}{\sqrt{6}RC}$$

if the gain is

$$\left|A_{u_o}\right| = \frac{R_2}{R_1} > 29$$

The loading of the op amp is minimized by making $R_1 > 10R$.

The harmonic distortion can be significantly reduced with an amplitude-limiting circuit. The amplitude-limiting circuit is designed using (see Fig. 11.3)

$$u_y = u_1 + 0{,}7 = \frac{u_{o(max)} R_6}{R_5 + R_6} + \frac{(U_-) R_5}{R_5 + R_6}$$

and

$$u_x = \frac{u_{o(min)}}{3} - 0{,}7 = \frac{u_{o(min)} R_3}{R_3 + R_4} + \frac{(U_+) R_4}{R_3 + R_4}$$

with u_1 set equal to zero (i.e., $u_y = 0{,}7$ V and $u_x = -0{,}7$ V).

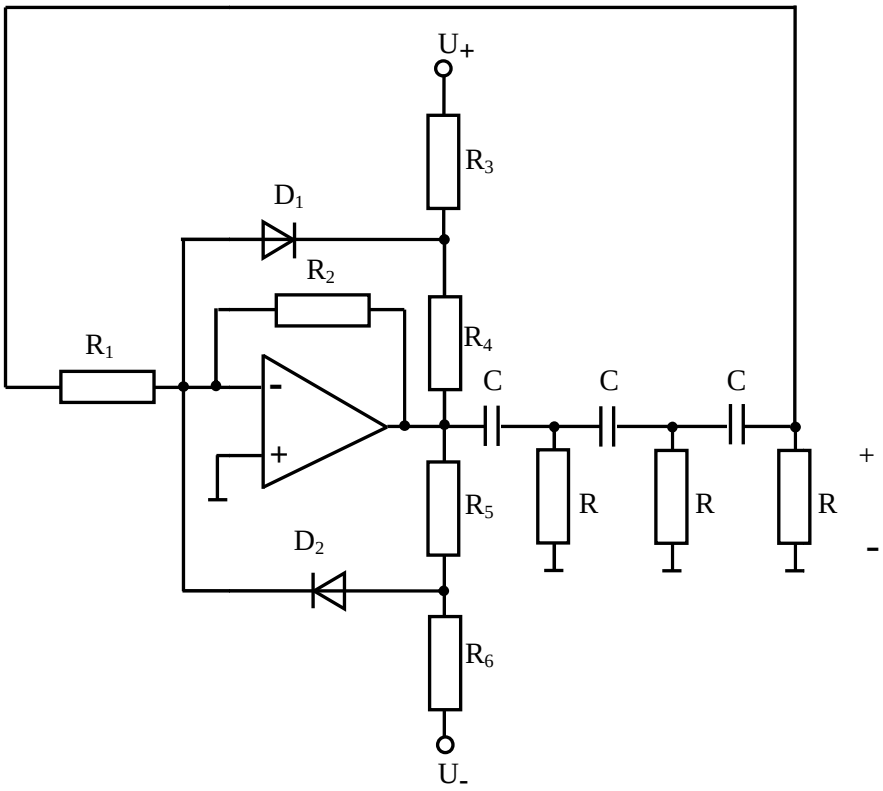


Fig. 11.6 Phase shift oscillator – amplitude stabilization (diodes)

Example 11.2

a) Design the phase shift oscillator shown in Fig. 11.3 to oscillate at 1 kHz.

b) Design an amplitude-limiting circuit. The amplitude of the sinusoidal output voltage is to be limited to .

☑ **Solution:**

A practical value of 0.01 μ F for the capacitors can be selected. Then, using

Ω

The resistors and must provide the gain in order to prevent loading . Letting 15 kΩ, then 29 435 kΩ. A 495 kΩ resistor was used to implement . This will allow for some extra gain to satisfy the start of oscillation condition (i.e.,).

From (supply voltage ± 12 V)

with and 5 V, we obtain

$$0{,}7 = \frac{5R_6}{R_5+R_6} - \frac{12R_5}{R_5+R_6} \Longrightarrow R_6 = \frac{127}{43} \cdot R_5$$

which can be satisfied with $R_5 = 2$ kΩ and $R_6 = 5{,}9$ kΩ.

From

$$-0{,}7 = \frac{u_{o(min)}R_3}{R_3+R_4} + \frac{(U_+)R_4}{R_3+R_4}$$

for symmetry, we obtain $R_3 = R_6 = 5{,}9$ kΩ and $R_4 = R_5 = 2$ kΩ.

11.3 Band pass filter – comparator oscillator (spot sinus)

The basic idea of low THD BPF-based oscillators is to incorporate a bandpass filter (BPF) along with a limiter and a comparator, in a positive feedback loop – Fig. 11.7. The oscillation frequency is set by the center frequency of the filter while the amplitude is set by the limiter (this filter has independent control of frequency, amplitude and distortion of the output).

Input of BPF is roughly a square wave. According to its Fourier series, a 50% duty-cycle square wave consists of odd order harmonic sine waves with the fundamental at the same frequency as the square wave.

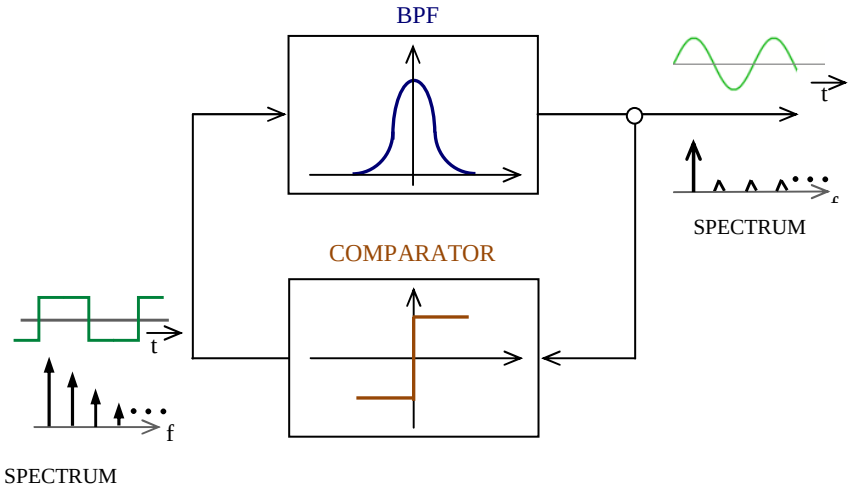


Fig. 11. 7 Bandpass filter – comparator oscillator

Fourier Series for a Square Wave

$$\frac{4k}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \cdots\right)$$

where k = peak amplitude of the square wave.

Thus ***THD is dominated by lower order harmonics***.

The ***THD is directly proportional to the quality factor of the loop filter*** – Fig.11.8. Fig. 11.9 shows the THD of the oscillator versus the quality factor (Q) of a second-order filter.

Achieving linearity better than 56 dB requires very high-Q filter ($Q > 70$). Implementing such high-Q filter requires large op amp gain–bandwidth product as well as a large spread of the capacitor values, and will end up with larger silicon area.

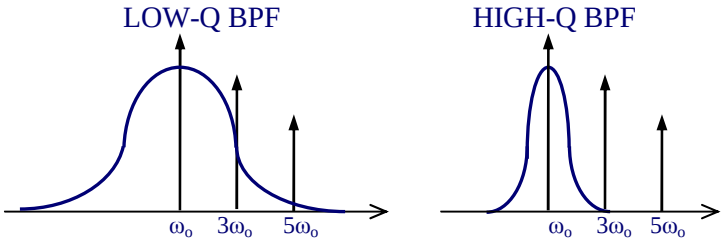


Fig. 11.8: The Fourier series versus Q of BPF

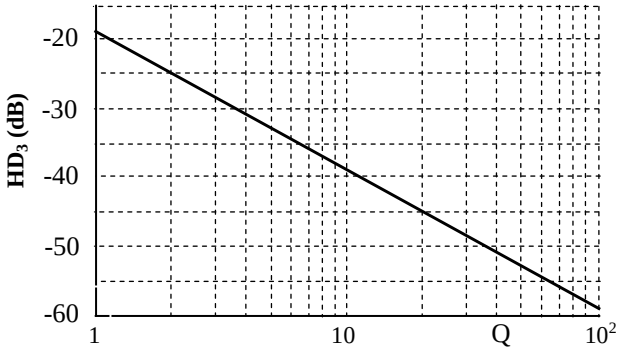


Fig. 11.9 The third harmonic versus Q

Notice that although the filter has unity gain, the amplitude of the sine wave output signal is greater than that of the square wave. This is because the fundamental has an amplitude of $4/\pi$ times that of the square wave as shown by the Fourier series. The bandpass filter will also filter out any DC component of the square wave input.

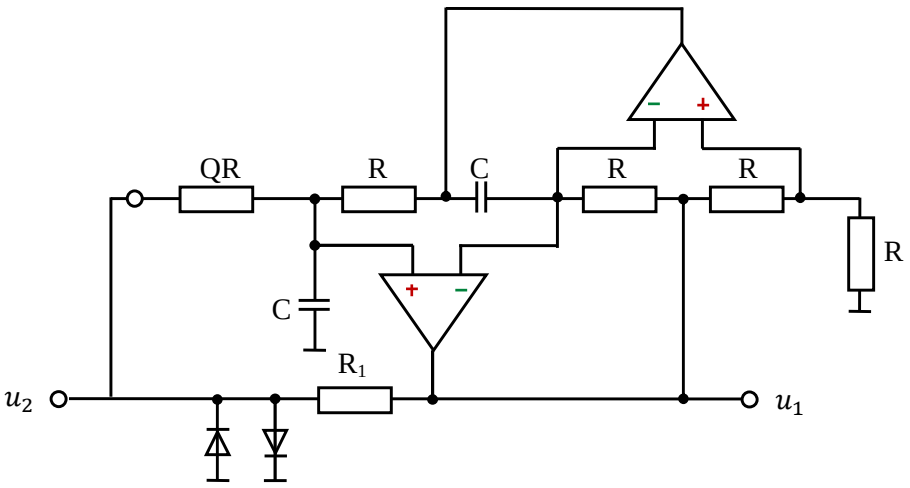


Fig. 11.10 Bandpass filter – the limiter used only

The very simple (only limiter is used) circuit is shown in Fig. 11.10, for example. The limiter is a pair of diodes (and R_1) to have a squarewave at u_2. The active filter (other circuit elements) selects the fundamental frequency and provides the sinus output at u_1 (any filter circuit with positive gain can be used to implement the bandpass filter).

Summary

1) An oscillator in which loop gain is exactly unity is an abstraction completely unrealizable in practice. In every practical oscillator the loop gain is larger than unity, and the amplitude of the oscillation is limited by means of automatic gain control circuit – AGC.
2) The result waveforms are never exactly sinusoidal – the closer the value βA is to "exactly 1", the more nearly sinusoidal is the waveform.
3) An oscillator is started by amplifying noise voltage, which is always present.
4) A lamp is a positive temperature coefficient resistor (PTC).
5) Thermistors (another simple AGC system) have the disadvantage of being sensitive to ambient temperature conditions.
6) JFETs work very well when the signal level controlled by the channel resistance is less than a few hundred milivolts.
7) BPF – based oscillator
    a. Oscillation frequency is set by BPF (band pass filter)
    b. Oscillation is guaranteed by high gain of comparator
    c. Linearity is heavily dependent on Q – factor of BPF (requires very high Q – factor)

Questions 11

You can find the answers in this text.

1. Why isn't an input signal (to the oscillator) needed to obtain an output voltage signal?
2. Compare the operation of the described oscillator circuits.
3. Why does a harmonic oscillator need an amplitude control circuit?
4. Explain the function of the buffer on Fig. 11.4
5. How can we get the sinusoidal voltage from the square wave voltage?
6. Why we need high-Q band pass filter on Fig. 11.7?

Problems 11

Example 11.1

Redesign the circuit of Fig. 11.3 for operation at 1 kHz. The amplitude of the sinusoidal output voltage is to be limited to $|u_o| = 2$ V.

Example 11.2

Redesign the circuit of Fig. 11.3 for operation at 5 kHz. The amplitude of the sinusoidal output voltage is to be limited to $|u_o| = 2$ V.

Example 11.3

Determine the needed $\boldsymbol{Q}$ of BPF (Fig. 11.7) if we need $\boldsymbol{HD_3 \approx -40}$ dB

PROBLEMS KEY 11

Ad example 11.1 ÷ 11.3

See example 11.1 and example 11.2 and Fig.11.9.

Basic texts

[1] Edgar Sánchez-Sinencio, Kilby, J.: High Linearity Oscillator Architectures Band – Pass Based. http://ansc.tamn.edu/journals.htm

[2] Hickman Ian: Analog Circuits Cookbook. Oxford 1999, ISBN 0 7506 4234 3

[3] Franco, S.: Design with operational amplifiers and analog integrated circuits. Mc Graw Hill, New York 2002, ISBN 0-07-232084-2

[4] Punčochář, J.: Operační zesilovače v elektronice. BEN – Praha, 2002, ISBN 80-7300-059-8

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). pp. 612÷636 Cambridge University Press, Cambridge 1982

[2] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

12. Frequency and time domain considerations

Time of study: 3 hours

Goals: the student should be able to

- define the basic behavior of amplifiers in the time domain
- define the basic behavior of the second order filters in the frequency and time domain
- assess the effect of slew rate on the behavior of the amplifier

EXPLANATION

One of the most important limitations of practical op amps is gain rolloff with frequency. This limitation affects both the frequency-domain and the time-domain behavior of circuits built around op amps. We have linear effects, such as finite small-signal bandwidth and nonzero rise time, and nonlinear effects, such as slew-rate limiting and finite full-power bandwidth. Additional effects are the settling time and intermodulation distortion.

The *open-loop voltage gain* of OPA (see 6.Chapter) is

$$A(p) = \frac{A_o\omega_1}{p+\omega_1} = \frac{\omega_T}{p+\omega_1} = \Big|_{\omega \gg \omega_1} \cong \frac{A_o\omega_1}{p} = \frac{\omega_T}{p} = \frac{-j\omega_T}{\omega}$$

where

$p = j\omega$ (harmonic steady state)

A_o is DC OPA gain

ω_1 is the first pole (frequency) of an OPA (dominant)

$\omega_T \cong A_o\omega_1$ is gain-bandwidth product of the OPA (GBW).

The ***closed-loop small-signal bandwidth*** is (inverting and noninverting amplifier; see 6.Chapter)

$$\omega_3 = \frac{\omega_T}{1 + \frac{R_2}{R_1}}$$

$$\Rightarrow \quad f_3 = \frac{f_T}{1 + \frac{R_2}{R_1}}$$

and the error function (inverting, noninverting, differential amplifier) is always

$$E = \frac{1}{1 + j\frac{\omega}{\omega_3}} \rightarrow \frac{\omega_3}{p + \omega_3}$$

It is also characteristic of a low pass filter and one can in fact model an amplifier as a low pass filter followed by a very wideband amplifier (ideal) with ideal gain G_{ID} (see Fig. 12.1):

$-\frac{R_2}{R_1}$ $\rightarrow$ inverting structure

$1 + \frac{R_2}{R_1}$ $\rightarrow$ noninverting structure

$K = \frac{R_2}{R_1}$ $\rightarrow$ differential structure

see 6.Chapter.

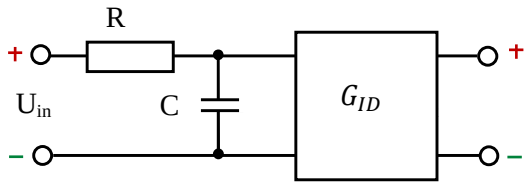


Fig. 12.1: RC low-pass model

In response to an applied voltage step, the output voltage can be calculated using equation

$$U_{out} = U_{in} \cdot G_{ID} \cdot (1 - e^{-\omega_3 t}) = U_{in} \cdot G_{ID} \cdot \left(1 - e^{-t/\tau}\right)$$

where

$$\tau = \frac{1}{\omega_3} = \frac{1}{RC}$$

The ***rise time*** t_r in response to a step is defined as the time it takes for the output to go from 10% (t_{10}) to 90% (t_{90}) of the step amplitude.

We can derive

$$U_{out} = U_{in} \cdot G_{ID} \cdot (1 - e^{-\omega_3 t}) \quad \Rightarrow \quad 0{,}1 = 1 - e^{-\omega_3 t_{10}} \quad \Rightarrow \quad t_{10} = \frac{-\ln 0{,}9}{\omega_3}$$

$$\Rightarrow \quad 0{,}9 = 1 - e^{-\omega_3 t_{90}} \quad \Rightarrow \quad t_{90} = \frac{-\ln 0{,}1}{\omega_3}$$

thus

$$t_r = t_{90} - t_{10} = \frac{-\ln 0{,}1}{\omega_3} - \left(\frac{-\ln 0{,}9}{\omega_3}\right) = \frac{\ln 0{,}9 - \ln 0{,}1}{\omega_3} = \frac{\ln 9}{\omega_3} = \frac{2{,}2}{\omega_3} = \frac{2{,}2}{2\pi f_3} = \frac{0{,}35}{f_3}$$

The rise time of the 1st order structure is related to the bandwidth (f_3 now) by the relation

$$t_r = \frac{0{,}35}{f_3}$$

For the „brick wall“ filter is

$$t_r = \frac{0{,}5}{f_0}$$

f_0 - cutoff frequency.

The rate at which U_{out} changes with time is the highest at the beginning of the exponential transition, when its value is (see the box below)

$$U_{in} \cdot G_{ID} \cdot \omega_3$$

BOX

$$\frac{dU_{out}}{dt} = U_{in} \cdot G_{ID} \cdot \omega_3 \cdot e^{-\omega_3 t} = \Big|_{t=0}\Big| = U_{in} \cdot G_{ID} \cdot \omega_3$$

Increasing the of step magnitude U_{in} increases this initial rate, until a point is reached beyond which the rate saturates at a constant value called the slew rate (SR). The transition is now a ramp, rather than an exponential.

The step magnitude corresponding to the onset of ***slew-rate limiting*** in such that

$$U_{in} \cdot G_{ID} \cdot \omega_3 = SR$$

or

$$U_{in_{max}} \cdot G_{ID} = \frac{SR}{\omega_3}$$

$$U_{in_{max}} = \frac{SR}{\omega_3 \cdot G_{ID}}$$

This means that as long as the input step is less than $U_{in_{max}}$, a follower ($G_{ID} = 1$) will respond with an exponential transition governed by $\tau = 1/\ \omega_3$. For a greater input step, however, the output will slew at a constant rate of SR, and it will do so until it comes within $U_{in_{max}}$ of the final value, after which it will complete the transition in exponential fashion.

In certain applications it is important to know the **settling time** t_s, defined as the time it takes for the output to settle within a specified band around its final value, usually for a full-scale output transition. It is apparent that slew-rate limiting plays an important role in the settling-time characteristic of a circuit.

Slew-rate limiting affects also *the* FPBW (full power bandwidth), defined as the maximum frequency at which the circuit still yields an undistorted full-power output. Letting

$$U_{out} = U_m \sin \omega t = U_m \sin 2\pi f t$$

we have

$$\frac{dU_{out}}{dt} = U_m\, 2\pi f \cos 2\pi f t \quad \Rightarrow \quad \frac{dU_{out}}{dt}\Big|_{max} = U_m 2\pi f$$

Equating it to the slew rate SR and solving for f, which value is the $FPBW$, we get

$$FPBW = \frac{SR}{2\pi U_m}$$

For instance, for $U_m = 10$ V and $SR = 0{,}5$ V/µs $= 0{,}5 \cdot 10^6$ V/s (µA741), the op amp has

$$FPBW = \frac{0{,}5 \cdot 10^6}{20\pi} = 7\,961 \text{ Hz}$$

Recovery time

When an amplifier is pushed into saturation, it takes time for it to recover and then slew to its new final output value - depending on the output structure and the compensation circuitry. Because of the time it takes to come out of saturation, an amplifier is slower when used as a comparator than when it is used under control in a closed-loop configuration. One can find saturation-recovery information in many amplifier data sheets. Figure 12.2 shows saturation recovery plots for the popular amplifier AD8061.

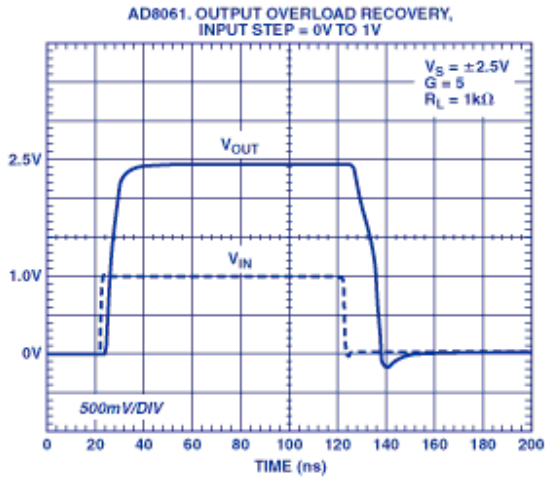


Fig. 12.2: Saturation recovery plots for the AD8061

12.1 Step Response of Second-order Systems and Damping Ratio

Transfer function for a second-order system (**Low Pass**) can be written as,

$$P(p) = \frac{\omega_o^2}{p^2 + p\frac{\omega_o}{Q} + \omega_o^2} = \frac{\omega_o^2}{p^2 + p\omega_o 2\xi + \omega_o^2}$$

$\xi = \frac{1}{2Q}$ is ***damping factor***.

The unit step response of this low-pass transfer function can be found by multiplying equation $1/p$ to get (Laplace transform)

$$U_{out}(p) = \frac{P(p)}{p} = \frac{\omega_o^2}{p(p^2 + p\omega_o 2\xi + \omega_o^2)} = \frac{1}{p} \cdot \frac{\omega_o^2}{(p - p_1)(p - p_2)} = \cdots$$

$$\cdots = \frac{1}{p} - \frac{p_2}{(p_2 - p_1)} \cdot \frac{1}{(p - p_1)} + \frac{p_1}{(p_2 - p_1)} \cdot \frac{1}{(p - p_2)}$$

where

$$p_1 = -\omega_o \cdot \left(\xi - \sqrt{\xi^2 - 1}\right)$$

$$p_2 = -\omega_o \cdot \left(\xi + \sqrt{\xi^2 - 1}\right)$$

This gives

$$u_{out}(t) = 1 - \frac{1}{2}\left(\frac{\xi}{\sqrt{\xi^2-1}} + 1\right) exp\left[-\omega_o \cdot \left(\xi - \sqrt{\xi^2-1}\right) \cdot t\right] + \cdots$$

$$\cdots \frac{1}{2}\left(\frac{\xi}{\sqrt{\xi^2-1}} - 1\right) exp\left[-\omega_o \cdot \left(\xi + \sqrt{\xi^2-1}\right) \cdot t\right]$$

$$u_{out}(t) = 1 - \tfrac{1}{2} exp\left[-\omega_o\left(\xi - \sqrt{\xi^2-1}\right)t\right] - \tfrac{1}{2} exp\left[-\omega_o \cdot \left(\xi + \sqrt{\xi^2-1}\right)t\right] - \cdots$$

$$-\tfrac{1}{2}\left(\tfrac{\xi}{\sqrt{\xi^2-1}}\right) exp\left[-\omega_o\left(\xi - \sqrt{\xi^2-1}\right)t\right] + \tfrac{1}{2}\left(\tfrac{\xi}{\sqrt{\xi^2-1}}\right) exp\left[-\omega_o\left(\xi + \sqrt{\xi^2-1}\right)t\right]$$

$$u_{out}(t) = 1 - exp(-\xi\omega_o t)\tfrac{1}{2}\left\{exp\left[\omega_o\left(\sqrt{\xi^2-1}\,\right)t\right] + exp\left[-\omega_o\left(\sqrt{\xi^2-1}\,\right)t\right]\right\} - \cdots$$

$$-\left(\tfrac{\xi}{\sqrt{\xi^2-1}}\right) exp[-\xi\omega_o t]\tfrac{1}{2}\left\{exp\left[\omega_o\left(\sqrt{\xi^2-1}\,\right)t\right] - exp\left[-\omega_o\left(\sqrt{\xi^2-1}\,\right)t\right]\right\}$$

We use

$$\cosh x = \tfrac{e^x + e^{-x}}{2}; \quad \sinh x = \tfrac{e^x - e^{-x}}{2}$$

$$u_{out}(t) = 1 - exp(-\xi\omega_o t)\left\{\cosh\left[\omega_o\left(\sqrt{\xi^2-1}\,\right)t\right] + \left(\tfrac{\xi}{\sqrt{\xi^2-1}}\right)\sinh\left[\omega_o\left(\sqrt{\xi^2-1}\,\right)t\right]\right\}$$

We choose a substitution

$$\left.\begin{array}{l} 1 = A \cdot \sinh\beta \\ \dfrac{\xi}{\sqrt{\xi^2-1}} = A \cdot \cosh\beta \end{array}\right\} \tfrac{A\cdot\sinh\beta}{A\cdot\cosh\beta} = \tanh\beta = \tfrac{\sqrt{\xi^2-1}}{\xi}$$

$$(A \cdot \cosh\beta)^2 - (A \cdot \sinh\beta)^2 = A^2[(\cosh\beta)^2 - (\sinh\beta)^2] = A^2 \cdot 1 = 1 - \tfrac{\xi^2}{\sqrt{\xi^2-1}} = \tfrac{1}{\xi^2-1}$$

$$\Rightarrow \quad A = \tfrac{1}{\sqrt{\xi^2-1}}; \quad \tanh\beta = \tfrac{\sqrt{\xi^2-1}}{\xi}; \quad \text{or} \quad \sinh\beta = \tfrac{1}{A} = \sqrt{\xi^2-1}$$

Now we can rewrite

$$u_{out}(t) = 1 - exp(-\xi\omega_o t)\left\{A\sinh\beta\cosh\left[\omega_o\left(\sqrt{\xi^2-1}\,\right)t\right] + \cdots\right.$$

$$\left.\cdots + A\cosh\beta\left(\tfrac{\xi}{\sqrt{\xi^2-1}}\right)\sinh\left[\omega_o\left(\sqrt{\xi^2-1}\,\right)t\right]\right\}$$

The output voltage is

$$u_{out}(t) = 1 - \frac{exp(-\xi\omega_o t)}{\sqrt{\xi^2-1}} \cdot \sinh\left[\omega_o\left(\sqrt{\xi^2-1}\right)t + \beta\right]$$

An overdamped system; $\xi > 1$ $(Q < 0{,}5)$

The step-response never goes above 1, hence the classification of overdamped; the poles p_1 and p_2 are real (see Fig. 12.3a).

Critically damped case; $\xi = 1\ (Q = 0{,}5)$

We can use equation (see above)

$$u_{out}(t) = 1 - exp(-\xi\omega_o t)\left\{\cosh\left[\omega_o\left(\sqrt{\xi^2 - 1}\,\right)t\right] + \left(\frac{\xi}{\sqrt{\xi^2-1}}\right)\sinh\left[\omega_o\left(\sqrt{\xi^2 - 1}\,\right)t\right]\right\}$$

If $\omega_o\left(\sqrt{\xi^2 - 1}\right)t \to 0$ then:

$$\cosh\left[\omega_o\left(\sqrt{\xi^2 - 1}\right)t\right] \approx 1$$

$$\sinh\left[\omega_o\left(\sqrt{\xi^2 - 1}\right)t\right] \approx \omega_o\left(\sqrt{\xi^2 - 1}\right)t$$

Thus we can easy derive that

$$u_{out}(t, \xi = 1) = 1 - exp(-\omega_o t) \cdot \{1 + \omega_o t\}$$

Underdamped system; $\xi < 1\ (Q > 0{,}5)$

The poles are complex. It is known that

$$\sinh jx = j \sin x$$

$$\sqrt{\xi^2 - 1} = j\left(\sqrt{1 - \xi^2}\right)$$

thus

$$u_{out}(t, \xi < 1) = 1 - \frac{exp(-\xi\omega_o t)}{\sqrt{1-\xi^2}} \cdot \sin\left[\omega_o\left(\sqrt{1 - \xi^2}\,\right)t + \beta\right];\quad \sin\beta = \sqrt{1 - \xi^2}$$

ω_o – ***natural*** frequency

$\omega_d = \omega_o\left(\sqrt{1 - \xi^2}\right)$ – ***damped natural frequency***

The same algorithm we can use for high pass (HP), band pass (BP) and band stop (BS; notch) filters – see Fig. 12.3:

HP2: $u_{outHP_2}(t) = \frac{exp(-\xi\omega_o t)}{\sqrt{\xi^2-1}} \cdot \sinh(\beta - \omega_v t)$

BP2: $u_{outBP_2}(t) = \frac{2\xi exp(-\xi\omega_o t)}{\sqrt{\xi^2-1}} \cdot \sinh(\omega_v t)$

BS2: $u_{outPZ_2}(t) = 1 - \frac{2\xi exp(-\xi\omega_o t)}{\sqrt{\xi^2-1}} \cdot \sinh(\omega_v t)$

$$\omega_v = \omega_o \cdot \sqrt{\xi^2 - 1};\quad \sinh\beta = \sqrt{\xi^2 - 1}$$

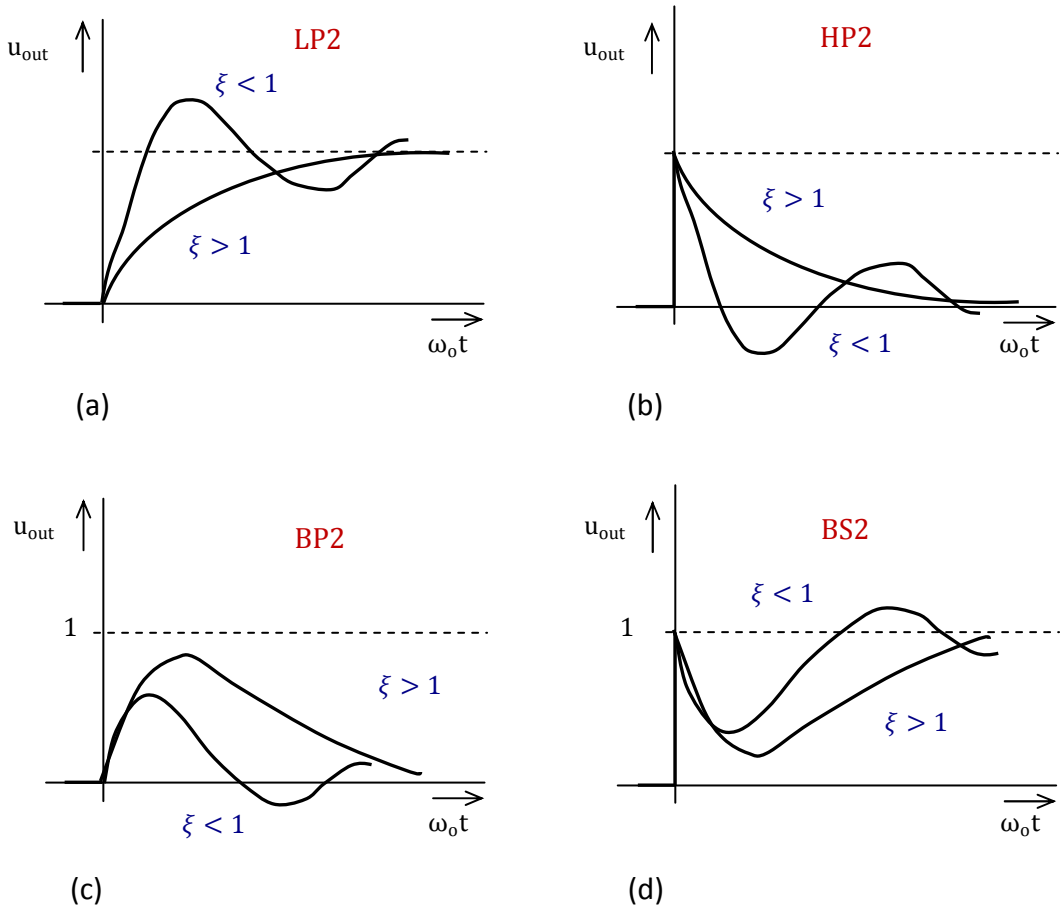


Fig. 12.3 Step response of second – order system; a) Low – pass; b) High – pass; c) Band – pass; d) Band – stop

Summary

1) The rise time of an electronic structure is related to the bandwidth f_0 by the relation

$$t_r = \frac{0{,}35 \div 0{,}5}{f_0}$$

2) The equation above is valid if (small signals)

$$U_{in} \leq U_{in_{max}} = \frac{SR}{\omega_3 \cdot G_{ID}}$$

3) The FPBW is related to the slew rate (SR) by the relation

$$FPBW = \frac{SR}{2\pi U_m}$$

4) An amplifier is slower when used as comparator (slew rate, saturation – recovery time) than when it is used in a closed – loop configuration (no saturation).
5) It is very useful to know step responses of second – order systems.

Questions 12

You can find the answers in this text.

1. What is the definition of a ***rise time***?
2. What is the definition of a ***full-power bandwidth***?
3. What is the ***recovery time***?
4. What is the ***settling time***?
5. Sketch and name the form of the response of the $2n$ order low pass filter (as a function of ξ).

Problems 12

Example 12.1

Calculate ξ if $Q = 0{,}5\ (1;\ 4;\ 10)$.

Example 12.2

Calculate t_r if $f_3 = 1\ (10;\ 40;\ 100;\ 1000)$ kHz

Example 12.3

Calculate FPBW if

a) $U_m = 10$ V and $SR = 0{,}5\ (1;\ 5;\ 10)$ V/µs
b) $U_m = 1$ V and $SR = 0{,}5\ (1;\ 5;\ 10)$ V/µs

Example 12.4

Calculate $U_{in_{max}}$ (slew - rate limiting, step) if

a) $G_{ID} = 1;\ \omega_3 = 2\pi \cdot 10^3;\ SR = 0{,}5\ (1;\ 5;\ 10)$ V/µs
b) $G_{ID} = 1;\ \omega_3 = 2\pi \cdot 10^5;\ SR = 0{,}5\ (1;\ 5;\ 10)$ V/µs

PROBLEMS KEY 12

Ad example 12.1 ÷ 12.4

You can find the answers in the text above.

Basic texts

[1] Punčochář, J.: Lineární obvody s elektronickými prvky. Skriptum, VŠB-TU Ostrava 2002, ISBN 80-248-0040-3

[2] Punčochář, J.: *Operační zesilovače v elektronice*. BEN, Praha 2002 (5. vydání), kap. IV, čl.40.3

[3] Punčochář, J.: Step response of an operational amplifier http://www.elektrorevue.cz

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). pp. 612÷636 Cambridge University Press, Cambridge 1982

[2] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

13. Relaxation oscillators

Time of study: 4 hours

Goals: the student should be able to

- describe basic relaxation structures
- change time parameters of relaxation structures
- assess the effect of slew rate and final recovery time
- provide a stable oscillation amplitude

EXPLANATION

Nearly every electronic instrument needs an oscillator or waveform generator of some sort. A very simple kind of oscillator (astable multivibrator, relaxation oscillator) can be made by charging and discharging a capacitor through a resistor – Fig. 13.1. The operation is simple.

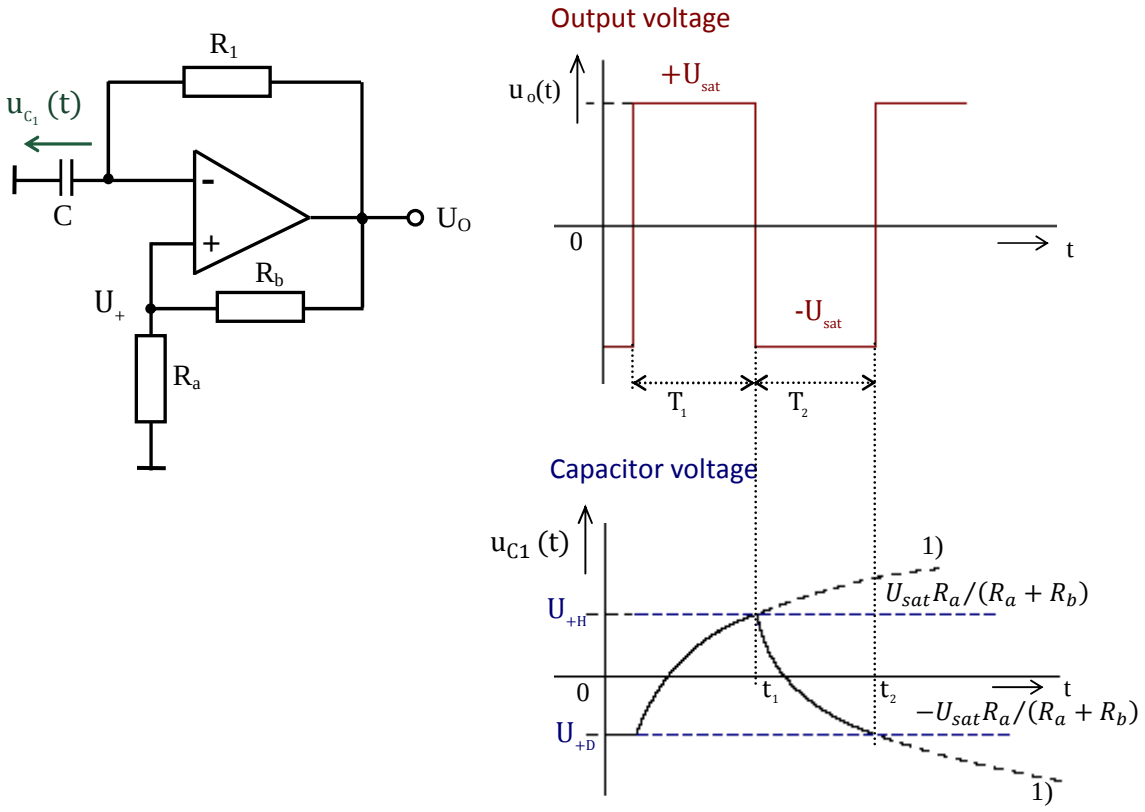


Fig. 13.1 Op amp relaxation oscillator; output and capacitor voltages

Note 1) Final value of the capacitor voltage would be theoretically equal to $\pm U_{sat}$, but the comparator ganges its state, thus charging (discharging) of the capacitor.

Assume that when power is first applied, the op-amp output goes to positive saturation $(+U_{sat})$. The capacitor C_1 begins charging up toward $+U_{sat}$, with time constant R_1C_1. When it reaches value $U_{+H} = U_{sat}R_a/(R_a + R_b)$, the op-amp switches into negative saturation $(-U_{sat})$ – it is a Schmitt trigger – and the capacitor begins discharging toward $(-U_{sat})$ with the same time constant. When it reaches value $U_{+D} = -U_{sat}R_a/(R_a + R_b)$, the op amp switches into positive saturation $(+U_{sat})$, the capacitor C_1 begins charging up toward $(+U_{sat})$. The cycle repeats indefinitely.

And thus charging the capacitor in the interval T_1 can be described by the equation

$$u_{C_1}(t) = \left[u_{C_1}(t=0) - u_{C_1}(t \to \infty)\right] \cdot exp(-t/\tau_1) + u_{C_1}(t \to \infty)$$

where (see Fig. 13.1c)

$$\tau_1 = R_1C_1$$

$$u_{C_1}(t=0) = U_{+D} = -U_{sat}R_a/(R_a + R_b)$$

$$u_{C_1}(t \to \infty) = +U_{sat}$$

When $u_{C_1}(t)$ reaches value $U_{+H} = U_{sat}R_a/(R_a + R_b)$, the op amp switches into negative saturation and $t = T_1$. From this we can determine

$$U_{+H} = [U_{+D} - U_{sat}] \cdot exp(T_1/\tau_1) + U_{sat} \quad \Rightarrow$$

$$U_{+H} - U_{sat} = [U_{+D} - U_{sat}] \cdot exp(T_1/\tau_1) \quad \Rightarrow$$

$$exp(T_1/\tau_1) = \frac{U_{+D} - U_{sat}}{U_{+H} - U_{sat}} \cdot \quad \Rightarrow$$

$$T_1 = R_1C_1 \cdot \ln\left(1 + \frac{2R_a}{R_b}\right) = \Big|_{R_a = R_b} = 1{,}0986 \cdot R_1C_1$$

Similarly, we determine that

$$T_2 = R_1C_1 \cdot \ln\left(1 + \frac{2R_a}{R_b}\right) = \Big|_{R_a = R_b} = 1{,}0986 \cdot R_1C_1$$

and so the period is

$$T = T_1 + T_2 = 2 \cdot R_1C_1 \cdot \ln\left(1 + \frac{2R_a}{R_b}\right) = \Big|_{R_a = R_b} \approx 2{,}2 \cdot R_1C_1$$

If saturation levels are unequal, we must (can) stabilize the output voltage amplitude – Fig. 13.2.

Still there is valid

$$T = T_1 + T_2 = 2 \cdot R_1C_1 \cdot \ln\left(1 + \frac{2R_a}{R_b}\right) = \Big|_{R_a = R_b} \approx 2{,}2 \cdot R_1C_1$$

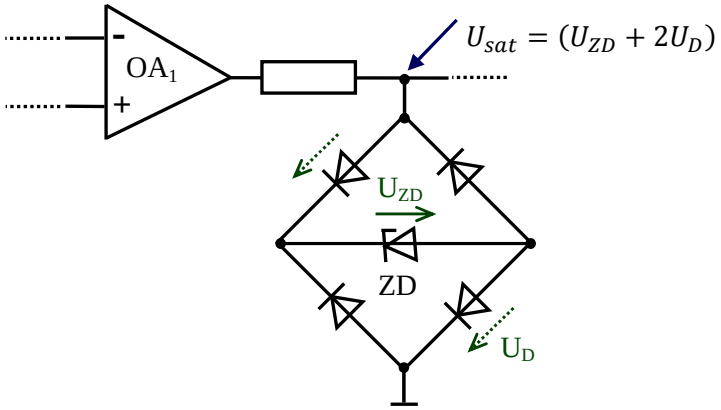


Fig. 13.2 Stabilization of the output voltage amplitude

If we need $T_1 \neq T_2$, we can modify the circuit in Fig. 13.1 – see Fig. 13.3.

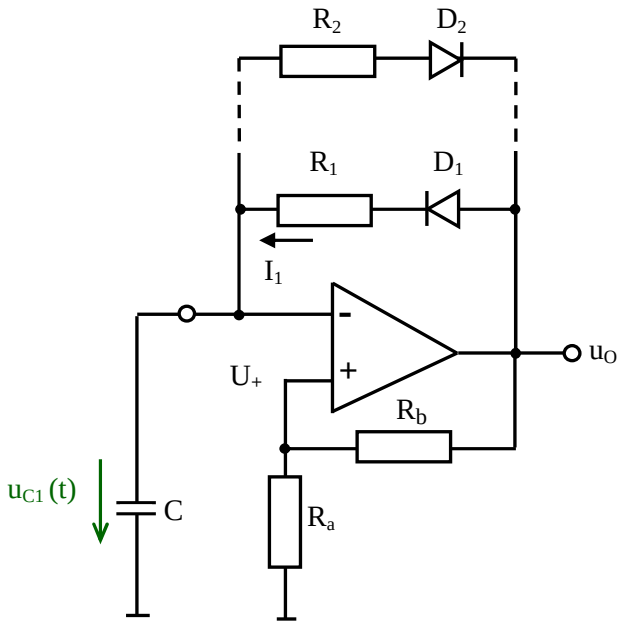


Fig. 13.3 A modification of the circuit on Fig. 13.1 by means of diodes $D_{1,2}$

The diode D_1 is switched for positive output voltage $(u_o = +U_{sat})$ and we can determine that

$$T_1 \cong R_1 C_1 \cdot \ln\left(1 + \frac{2R_a}{R_b}\right) = \Big|_{R_a = R_b}\Big| = 1{,}0986 \cdot R_1 C_1$$

The diode D_2 is switched for negative output voltage $(u_o = -U_{sat})$ and we can easy determine that

$$T_2 = R_2 C_1 \cdot \ln\left(1 + \frac{2R_a}{R_b}\right) = \Big|_{R_a = R_b}\Big| = 1{,}0986 \cdot R_2 C_1$$

Thus

$$T = T_1 + T_2 = (R_1 C_1 + R_2 C_1) \cdot \ln\left(1 + \frac{2R_a}{R_b}\right) = \Big|_{R_a = R_b}\Big| \approx 1{,}1 \cdot (R_1 C_1 + R_2 C_1)$$

A better oscillator can be made by charging a capacitor through a current source – see Fig. 13.4. The multivibrator consists of an integrator (OA_1) and a comparator (OA_2). The comparator compares the voltage U_d continuously with the inverting input that is at 0 V. When U_d goes above or below 0 V, the output of OA_2 is at the positive or negative saturation ($\pm U_{sat}$), respectively. At this time $U_d \approx 0$ V and $u_p(t) = U_{PK}$ and

$$I_1 = \frac{U_{PK}}{R_1}$$

$$-I_2 = \frac{-u_{o\,(t)}}{R_2}$$

where

$u_p(t)$ – is the triangular wave

$u_o(t)$ – is the square wave (see Fig. 13.4 or 13.5).

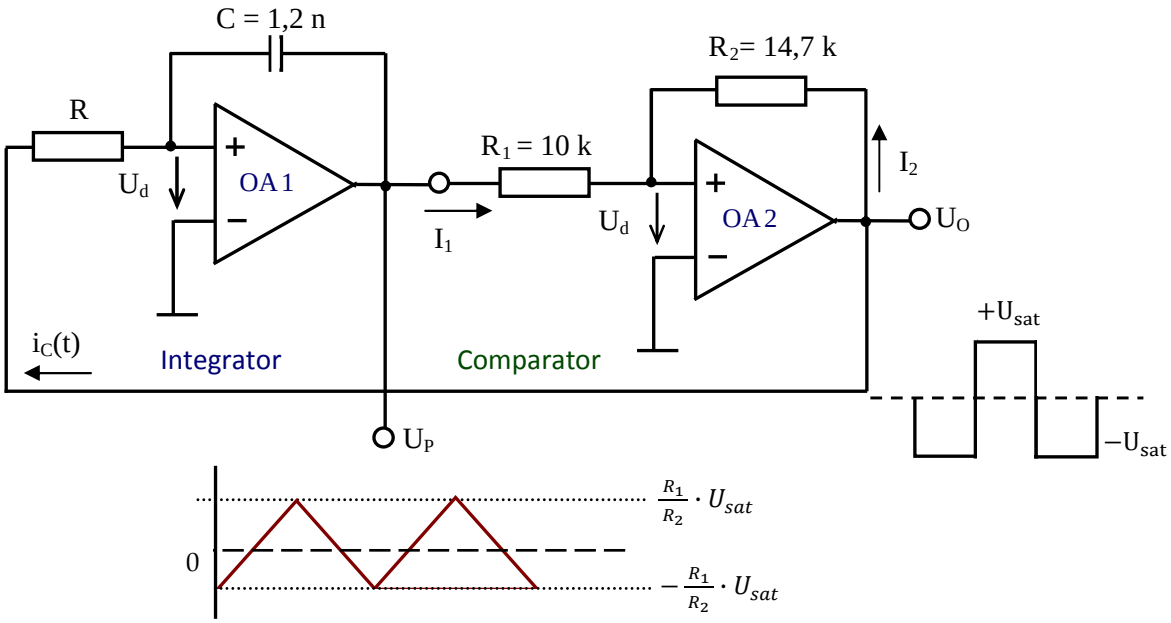


Fig. 13.4 Square and triangular wave generator (multivibrator, supply voltage U_{CC} is ±15 V)

It is evident that for the ideal OA_2 (zero input currents)

$$\frac{U_{PK}}{R_1} = \frac{-u_o(t)}{R_2}$$

thus

$$U_{PK} = -\frac{R_1}{R_2} \cdot u_o(t) = \pm \frac{R_1}{R_2} \cdot U_{sat}$$

To work in the oscillation, the condition $R_1/R_2 < 1$ is necessary. On the contrary case the needed magnitude $|U_{PK}|$ will be greater than U_{sat}, the circuit will not oscillate (if we suppose the same U_{sat} of the OA_1 and the OA_2).

Let us suppose that $u_o(t) = +U_{sat}$. This $+U_{sat}$ is an input of the integrator (inverting) OA_1 and a capacitor C current (for the ideal OA_1) is

$$i_C(t) = \frac{U_{sat}}{R} = I_C$$

The output of the OA_1, therefore, will be a negative-going ramp.

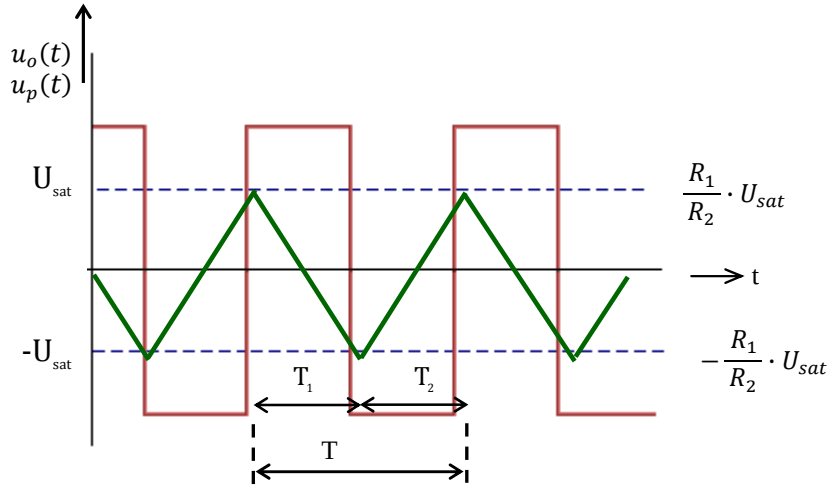


Fig. 13.5 Square and triangular wave of the circuit in Fig. 13.4

If $u_o(t) = -U_{sat}$ the capacitor current

$$i_C(t) = -\frac{U_{sat}}{R} = -I_C$$

the output of the OA_1, therefore, will be a positive-going ramp.

The time it takes for the output waveform $u_p(t)$ to swing from $-U_{sat}R_1/R_2$ to $+U_{sat}R_1/R_2$ is equal to half the time period $T/2$. Peak-to-peak amplitude of the triangular wave is

$$\Delta U_{PK} = \frac{R_1}{R_2}U_{sat} - \left(-\frac{R_1}{R_2}U_{sat}\right) = 2 \cdot \frac{R_1}{R_2}U_{sat}$$

A change of a capacitor charge ΔQ_C (during $T/2$) is

$$\Delta Q_C = C \cdot \Delta U_{PK}$$

and

$$\Delta Q_C = I_C \cdot \frac{T}{2}$$

Hence

$$I_C \cdot \frac{T}{2} = C \cdot \Delta U_{PK}$$

thus

$$\frac{U_{sat}}{R} \cdot \frac{T}{2} = C \cdot 2 \cdot \frac{R_1}{R_2}U_{sat}$$

or

$$T = 4RC \cdot \frac{R_1}{R_2}$$

The T is the time period of oscillations (ideal op amps). We can easily determine the frequency, too:

$$f = \frac{1}{T} = \frac{R_2}{4RC \cdot R_1}$$

The integrator (OA_1) is "nearly ideal" if its slew rate SR_I is greater than the maximum slope of the voltage $u_p(t)$ - Fig. 13.4:

$$\frac{2 \cdot \frac{R_1}{R_2} U_{sat}}{\frac{T_{min}}{2}} \leq SR_I$$

where T_{min} is the needed minimum time period.

Hence the needed SR_I is given by ($1/T_{min} = f_{max}$ - the maximum needed frequency)

$$\boxed{SR_I \geq 4 \cdot \frac{R_1}{R_2} U_{sat} f_{max}}$$

And on the contrary for the given SR_I the "maximum ideal integrator frequency" is

$$f_{max} \leq \frac{SR_I}{4 \cdot \frac{R_1}{R_2} U_{sat}}$$

The comparator (OA_2) is ideal if its slew rate SR_C is infinite and op amp recovery time t_r is zero. In practice we must choose:

$$\boxed{SR_C \geq 10 \cdot SR_I}$$

See [2], too.

Theoretical frequencies are in the Tab. 1. Some practical measurements are in the Tab. 2.

Poor dynamic properties of the comparator lead to overcharge of the capacity C (generate voltage ΔU above U_{PK}) if R decreases, now

$$\pm U_{PK} = \frac{R_1}{R_2} U_{sat} + \Delta U$$

This phenomenon leads to a decrease of frequency f with respect to its ideal value – Tab.1. To overcome (partly) the effect of the op amp OA_2 slew rate (SR_C) and overload recovery time (t_r) we can modify the multivibrator - see a compensating resistor R_K and a compensating capacitor C_K in parallel to R_1 -Fig. 13.6.

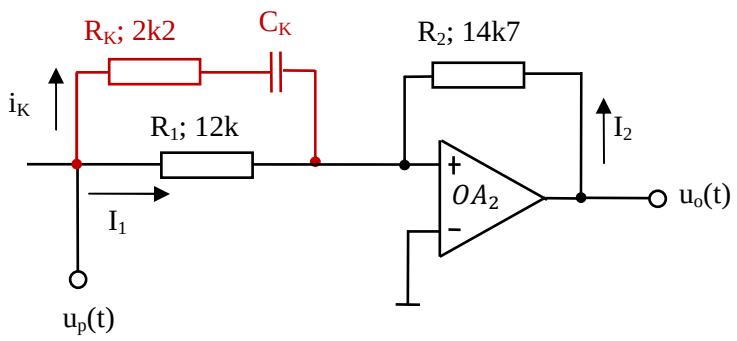


Fig. 13.6 A compensating circuit C_K, R_K

Tab. 1: Results from eq. $f = \frac{1}{T} = \frac{R_2}{4RC \cdot R_1}$ and given conditions

R [kΩ]	f [kHz]
2,2	139,18
10	30,620
100	3,062
1000	0,306

Tab. 2: Measured properties of the circuit in Fig. 13.4 ($U_{sat} \approx 13$ V; $U_{CC} = \pm 15$ V; t_n - rise time; t_S - decay time of OA_2; slew rate $SR_C \approx 26/t_n$ or $26/t_S$); + - optimum; ! - no good solution

OA_1	MAC156+	MAC156	MAA741!	MAA741!	MAA741!
OA_2	MAC157+	MAC156	MAC157+	MAC156	MAA741!
t_n [µs]	0,4	2	0,4	2	68
t_S [µs]	0,4	0,6	0,4	0,6	80
2,2	106,62	80,63	15,30	14,69	6,027
10	28,648	23,95	13,58	13,07	5,230
100	3,061	2,942	2,940	2,931	1,845
1000	0,301	0,301	0,301	0,301	0,300
R [kΩ]	f [kHz]				

Tab. 3. Measured real peak-to-peak capacitor amplitudes (U_{pp}) and frequencies of oscillation (f) as a function of R and C_K - Fig. 13.6. [OA_1 - MAC156, OA_2 - MAC157; the MAC157 as an integrator is frequency unstable; $*$ - the op amps are hot due to a frequency dependent dissipation - switching losses - any touch of a finger can show up this "burning" problem].

R (kΩ)	f (kHz)	U_{pp} (V)	f (kHz)	U_{pp} (V)	f (kHz)	U_{pp} (V)
2,2*	106,6	23,5	131,6	18,5	257,0	10
10	28,65	19	29,98	18,5	33,81	16
100	3,061	18,5	3,073	18,2	3,108	17
1000	0,301	18	0,301	18	0,301	18
C_K	not		100 pF		330pF	

If we chose the right value of the C_K the peak-to-peak amplitude of the triangular wave will be constant (with frequency; reduces effective value of R_1 – it compensates ΔU). Tab.3

summarises the measured properties of the real multivibrator for $C_K = 0$ (≡ no capacitor C_K), $C_K = 100$ pF (nearly optimum) and $C_K = 330$ pF ("overcompensation").

If $f >$ 200 kHz the MAC157 can temperature is above 90°C (8 pin metal can, TO-99). A junction-to-case thermal resistance of the TO-99 is 45°C/W (150 °C/W junction-to-ambient thermal resistance). It is evident that frequency 257 kHz will push the operating junction temperature high enough to concern. The instantaneous power dissipation during transition (in an op amp) may be quite large because voltage and current have high values simultaneously. With high repetition rate of f, the transistors (in the op amp) undergoes $2f$ transitions per second and so they dissipate much power, so some care should be made to ascertain if any power is frequency dependent.

The lifetime of all semiconductors is inversely related to their operating junction temperature. The cooler semiconductors can be kept during operation, the more closely they will approach to maximum useful live. To maintain low junction temperature, either thermal resistance (°C/W) or the power dissipated (or both) must be kept low. The useful output power (and the maximum useful frequency) is always limited by the cooling efficiency of op amp case.

Even simple circuits with op amps can be expensive, if they are intended for processing fast signals specified in the frequency domain. Therefore we must decide which solution is the best one for our purposes (heat sink, forced air movement, very high-speed op amps). And what about fast signals? It must be decided in agreement with "data sheets" for any chosen op amp.

Summary

1) A simple relaxation oscillator can be made by charging and discharging a capacitor through a resistor. We get square wave only.
2) If saturation levels are unequal, we must stabilize the output voltage amplitude.
3) A better relaxation oscillator can be made by charging and discharging a capacitor through a current source. We get square and triangular outputs.
4) The integrator is “nearly ideal” if its slew rate is larger than the maximum slope of the capacitor voltage.
5) The comparator is ideal if its slew rate is infinite and op amp recovery time is zero.
6) The instantaneous power dissipation during transitions may be quite large because voltage and current have high vakues simultaneously. With high repetition rate f amplifiers dissipate a much power. To maintain low junction temperature either thermal resistance or the power dissipated must be kept low.

Questions 13

You can find the answers in this text.

1. How it is possible to suppress the effect of different op amp saturation voltages on function of relaxation circuit?
2. What is the amplitude of the integrator voltage depending on the finite slew rate and recovery time?
3. How can we correct the effect of finite slew rate (of comparator)?
4. Is a voltage to frequency converter a relaxation circuit?

Problems 13

Example 13.1

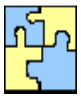
Design the circuit of Fig. 13.1 for operation at 1 kHz.

Example 13.2

Design the circuit of Fig. 13.1 for operation at $\boldsymbol{T_1 = 0,5}$ ms and $\boldsymbol{T_2 = 4,5}$ ms.

Example 13.3

Redesign the circuit of Fig. 13.4 for operation at 5 kHz. Select appropriate operational amplifiers (needed slew rates), suppose supply voltages ± 15 V and U_{sat} = ± 13 V.



PROBLEMS KEY 13

Ad example 13.1

Select for example $R_1 = R_a = R_b = 10$ kΩ, then

$$f = \frac{1}{T} = \frac{1}{2{,}2 \cdot C_1 \cdot R_1} \Rightarrow C_1 = \frac{1}{T} = \frac{1}{2{,}2 \cdot f \cdot R_1} = 45{,}5 \text{ nF}$$

Ad example 13.2

Select for example $R_1 = R_a = R_b = 10$ kΩ, then

$$T_1 = 1{,}1 \cdot C_1 \cdot R_1 \Rightarrow C_1 = \frac{T_1}{1{,}1 \cdot R_1} = 45{,}5 \text{ nF}$$

and

$$T_2 = 1{,}1 \cdot C_1 \cdot R_2 \Rightarrow R_2 = \frac{T_2}{1{,}1 \cdot C_1} = \frac{T_2 \cdot R_1}{1{,}1 \cdot C_1 \cdot R_1} = \frac{T_2}{T_1} \cdot R_1 = 9 \cdot R_1 = 90\ \text{k}\Omega$$

Ad example 13.3

Use eqs.:

$$f = \frac{1}{T} = \frac{R_2}{4RC \cdot R_1} \rightarrow \text{needed } R = \frac{R_2}{4fC \cdot R_1} = 61\,250\ \Omega$$

$$SR_I \geq 4 \cdot \frac{R_1}{R_2} U_{sat} f_{max} \rightarrow \text{needed } SR_I > 4 \cdot \frac{10}{14{,}7} 13 \cdot 5 \cdot 10^3 = 0{,}177\ \text{V/µs}$$

$$SR_C = 10 \cdot SR_I = 1{,}77\ \text{V/µs}$$

(See tables 2 and 3, too).

Basic texts

[1] Punčochář, J.: Operační zesilovače v elektronice. BEN, Praha 2002 (5. vydání), čl. 76, čl. 79

[2] Punčochář, J.: Astable multivibrator with real op amps. http://www.elektrorevue.cz

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). pp. 612÷636 Cambridge University Press, Cambridge 1982

[2] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

14. Thermal consideration, thermal resistance

Time of study: 4 hours

Goals: the student should be able to

- describe the effect of temperature on reliability
- work with the concepts of power dissipation, thermal resistance – static, dynamic
- specify the requirements for heatsink (cooling)

EXPLANATION

The reliability of semiconductor devices is represented by the failure rate curve (called the "bathtub curve"), which is illustrated in Fig. 14.1. The curve can be divided into the three following regions: (1) initial failures, which occur within a relatively short time after a device starts to be used, (2) random failures, which occur over a long period of time, and (3) wear-out failures, which increase as the device nears the end of its life.

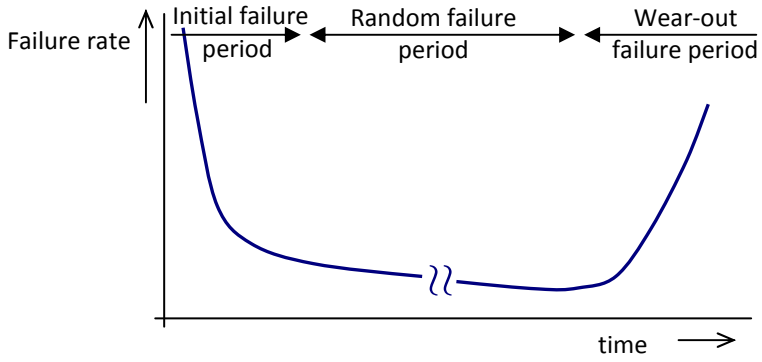


Fig. 14.1: Failure Rate Curve (Bathtub Curve)

"Initial failures" are considered to occur when a latent defect is formed, for example, during the device production process and then becomes manifest under the stress of the operation.

"Random failures" occur once devices having latent defects have already failed and been removed. In this period, the remaining high-quality devices operate stably. The failures that occur during this period can usually be attributed to randomly occurring excessive stress.

"Wear-out failures" occur due to aging of devices from wear and fatigue. The failure rate tends to increase rapidly in this period. Semiconductor devices are therefore designed so that wear-out failures will not occur during their guaranteed lifetime.

Semiconductor devices stop working for a variety of reasons, but most failure mechanisms do seem to have one factor in common - the hotter the device runs, the sooner it will fail. The high-temperature failure rate can then be used to assess the likelihood of failure at a more normal temperature from the Arrhenius equation:

$$AF_t = exp\left[\frac{E_A}{k} \cdot \left(\frac{1}{T_u} - \frac{1}{T_a}\right)\right]$$

where (if you care)

AF_t is the acceleration factor due to temperature,

E_A is the activation energy (eV),

k is Boltzmann's constant,

T_u and T_a (^{o}K) are the temperatures in normal use and in the accelerated test, respectively.

So running a semiconductor 10^{o}C cooler could double its life. The point is: **cooler devices survive longer**.

But all electronic components dissipate heat. Therefore, long life and reliable performance of a component may be achieved by effectively controlling the device operating temperature within the limits set by the device design engineers.

Almost all the power is dissipated at the base-collector junction, deep inside the transistor. The cutaway drawing of a TO-220 power transistor on the Fig. 14. 2 shows a tiny chip stuck onto the block of metal that forms the main body of the device. The metal tab acts as a built-in heatsink

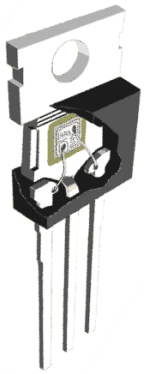


Fig. 14.2: Simplified structure of TO-220 power transistor

Heat sinks (external) are devices that enhance heat dissipation from a hot surface, usually the case of a heat generating component, to a cooler ambient, usually air. The primary purpose of a heat sink is to maintain the device temperature below the maximum allowable temperature specified by the device manufacturer.

14.1 Thermal Circuit

Before discussing the heat sink selection process, it is necessary to define common terms and establish the concept of a thermal circuit. The objective is to provide basic fundamentals of heat transfer for those readers who are not familiar with the subject. Notations and definitions of the terms are as follows:

P: total power or rate of heat dissipation in W, it represents the rate of heat dissipated by the electronic component during the operation. For the purpose of selecting a heat sink, the maximum operating power dissipation is used.

T_j**:** maximum junction temperature of the device in ^{o}C. Allowable T_j values range from 115°C in typical microelectronic applications to as high as 180° C for some electronic control devices, in special and military applications, 65°C to 80°C are not uncommon.

T_c**:** case temperature of the device in ^{o}C. Since the case temperature of a device depends on the location of measurement, it usually represents the maximum local temperature of the case.

T_s**:** sink temperature in ^{o}C. Again, this represents the maximum temperature of a heat sink at the location closest to the device.

Using temperatures and the rate of heat dissipation, a quantitative measure location of a thermal of thermal resistance of heat transfer efficiency across two components can be expressed in terms ***R,*** defined as

$$R = \frac{\Delta T}{P}$$

where ΔT is the temperature difference between the two locations. The unit of thermal resistance is in ^{o}C/W (= K/W), indicating the temperature rise per unit rate of heat dissipation. This thermal resistance is analogous to the electrical resistance R_e**,** given by Ohm's law:

$\Delta U \rightarrow \Delta T$; $I \rightarrow P$; voltage difference; current

Consider a simple case where a heat sink is mounted on a device package as shown in Fig. 14. 3. Using the concept of thermal resistance, a simplified thermal circuit of this system can be drawn, as also shown in the figure. In this simplified model, heat flows serially from the junction to the case then across the interface into the heat sink, and finally dissipated from the heat sink to the air stream.

The thermal resistance between the junction and the case of the device is defined as

$$R_{jc} = \frac{\Delta T_{jc}}{P} = \frac{T_j - T_c}{P}$$

This resistance is specified by the electronic device manufacturer – it is beyond the user's ability to alter or control.

Similarly, case-to-sink and sink-to-ambient resistances are defined as

$$R_{cs} = \frac{\Delta T_{cs}}{P} = \frac{T_c - T_s}{P}$$

$$R_{sa} = \frac{\Delta T_{sa}}{P} = \frac{T_s - T_a}{P}$$

Here, R_{cs} represents the thermal resistance across the interface between the case and the heat sink and is often called the interface resistance. This value can be improved substantially depending on the quality of mating surface finish and/or the choice of interface material. R_{sa} is the heat-sink thermal resistance.

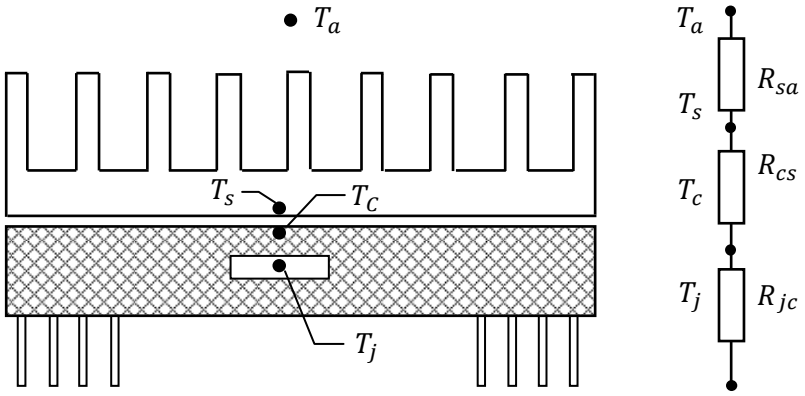


Fig. 14.3 Thermal Resistance Circuit

Semiconductor devices are nearly always electrically insulated from the heatsink. This means that some material must be used between the case of the device and the heatsink surface. This will invariably increase the thermal resistance - there is no known material which is a perfect electrical insulator and thermal conductor. Heatsink compound ("grease") *must* be used on both sides of mica or Kapton to ensure minimum thermal resistance. There are actually a number of alternatives for electrically insulating the transistor from the heatsink while still allowing heat transfer. The basic list – Table 1:

Table 1 - Thermal Resistance of Various Mounting Methods (TO-220 Case)

Material	Thermal	Electrical	Thermal Resistance	Other Properties
mica	Good	Excellent	~ 0,75 – 1,0	Fragile
Kapton	Good	Excellent	~ 0,9 – 1,5	Very robust
aluminium oxide	Excellent	Very Good	~ 0,4	Fragile - easily damaged
beryllia (beryllium oxide)	Excellent	Excellent	~ 0,25	Toxic
Sil-Pads	Fair +	Excellent	~ 1,0 – 1,5	Convenient

The figures shown in Table 2 for the TO-3 and TOP-3 case styles are highly optimistic, and are unlikely to be achieved in practice. In general silicone rubber pads are used only where devices operate at low power, regardless of manufacturer and claims.

Table 2 - Comparison of Case Types – thermal resistances

Package	Direct (Dry)	Direct+ Grease	Mica+ Grease	Sil-Pad™
TO-3	0.5 - 0.7	0.3 - 0.5	0.4 - 0.6	0.4
TOP-3	0.8 - 1.1	0.5 - 0.7	0.6 - 0.9	0.65
TO-126	1.5 - 2.0	0.9 - 1.2	1.2 - 1.7	NA
TO-220	1.1 - 1.3	0.9 - 1.1	1.0 - 1.6	1.5

The importance of a thermal compound cannot be overstated. The drawing – Fig. 14.4 - shows the surface of the transistor and heatsink at a microscopic level. The surfaces only touch in a few places. This is why the thermal resistance is so much higher without the use of any thermal compound.

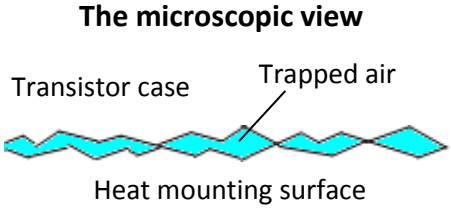


Fig. 14.4 Surface of the transistor and heatsink at a microscopic level

Obviously, the total junction-to-ambient resistance is the sum of all three resistances:

$$R_{ja} = R_{jc} + R_{cs} + R_{sa} = \frac{T_j - T_a}{P}$$

14.2 Required Heat-Sink Thermal Resistance

To begin the heat sink selection, the first step is to determine the heat-sink thermal resistance required to satisfy the thermal criteria of the component. By rearranging the previous equation, the heat-sink resistance can be easily obtained as

$$R_{sa} = \frac{T_j - T_a}{P} - R_{jc} - R_{cs}$$

In this expression, T_j and R_{jc} are provided by the device manufacturer, and T_a, P and R_{cs} are the user defined parameters. The ambient air temperature for cooling electronic equipment depends on the operating environment in which the component is expected to be used. Typically, it ranges from 35 to 45°C, if the external air is used and from 50 to 60°C, if the component is enclosed or is placed in a wake of another heat generating equipment.

The interface resistance R_{cs} depends on the surface finish, flatness, applied mounting pressure, contact area, and, of course, the type of interface material and its thickness.

With all the parameters on the right side of the above expression identified, R_{sa} becomes the required maximum thermal resistance of a heat sink for the application. In other words,

the thermal resistance value of a chosen heat sink for the application has to be equal to or less than the above R_{sa} value for the junction temperature to be maintained at or below the specified T_j.

When selecting a heat sink, it is necessary to classify the air flow as natural, low flow mixed, or high flow forced convection - the next step is to determine the required type and volume of a heat sink.

14.3 Heat-Sink Types

The thermal resistance for a flat square plate heat sink may be approximated by

$$R_{sa} = \frac{3{,}3}{\sqrt{\lambda b}} \cdot C_f^{0{,}25} + \frac{650}{A} \cdot C_f \qquad [\mathrm{K/W, W/K{\cdot}cm, mm, cm^2}]$$

Typical values of heatsink thermal conductance λ in W/K cm at 350 K, are shown in Table 3 and

b is the thickness of the heat sink, mm

A is the area of the heat sink, cm^2

C_f is a correction factor for the position and surface emissivity of the heat-sink orientation according to Table 4.

Table 3 – Thermal conductance λ

material	λ [W/K cm]
copper - Cu	3,8
aluminium - Al	2,1
brass - CuZn	1,1
steel (low carbon)	0,46

Table 4 – Heatsink correction factor

Surface position C_f	shiny	blackened
vertical	0,85	0,43
horizontal	1,0	0,5

The correction factorC_f illustrates the fact that black surfaces are better heat radiators and that warm air rises, creating a 'chimney' effect. Equation is valid for one power-dissipating device, in the centre of the sink, at a static ambient temperature up to about 45°C, without other radiators in the near vicinity. In order to decrease thermal resistance, inferred by the equation, finned-type heat sinks are employed which increase sink surface area – Fig. 14.5.

The effective sink thermal resistance can be significantly reduced by *forced air cooling*, as indicated in Fig. 14.6.

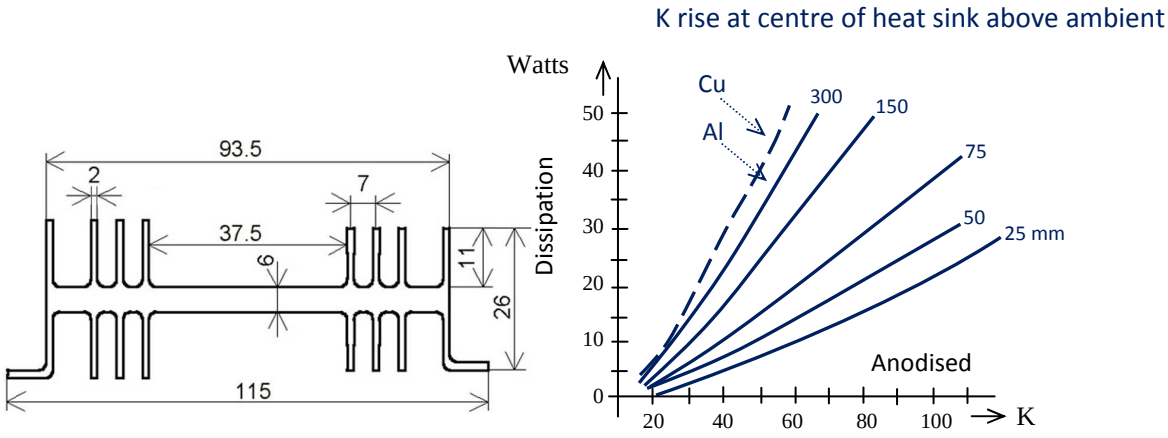


Fig. 14.5 Heat-sink typical data (for aluminium and copper) - cross-section view; temperature rise versus dissipation for an anodised finish and different lengths (qualitative description only)

The effective sink thermal resistance can be significantly reduced by *forced air cooling*, as indicated in Fig. 14.6.

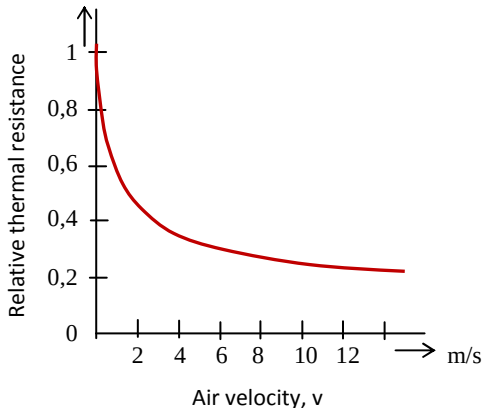


Fig.14.6 Improved cooling with - forced air cooled heat-sink - relative thermal resistance improvement with surface airflow

14.4 Junction temperature under pulsed conditions

Under surge conditions the junction temperature rises exponentially, according to its thermal inertia – Fig. 14.7. Rather than using the thermal resistance, that is appropriate for steady state operation, we use the Transient Thermal Impedance (or, more correctly, Thermal Response Curve), as the one shown in Fig. 14.8. For a surge of given duration (x axis), this curve gives a thermal response factor (y axis). The peak junction temperature due to the surge condition can be calculated as indicated in the figure itself. The power dissipation P_{DM} is normally calculated from the voltage and current across the device during the surge.

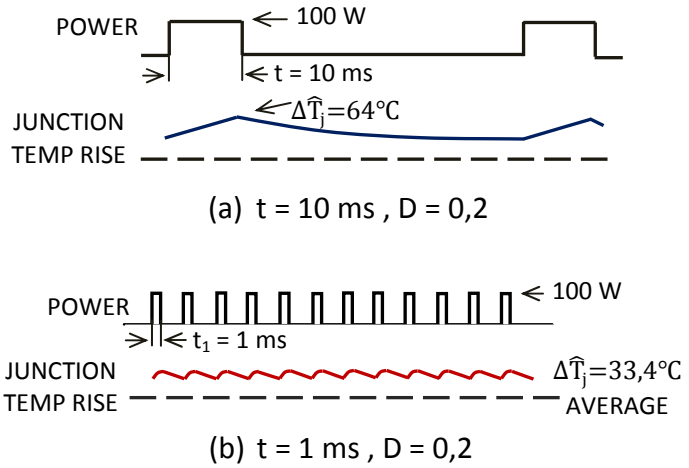


Fig. 14.7 Waveforms illustrating that peak junction temperature is a function of switching frequency: (a) lower switching frequency with 10 ms pulse and a 20 per cent duty cycle and (b) high frequency and 1 ms pulse with a duty cycle the same as in (a).

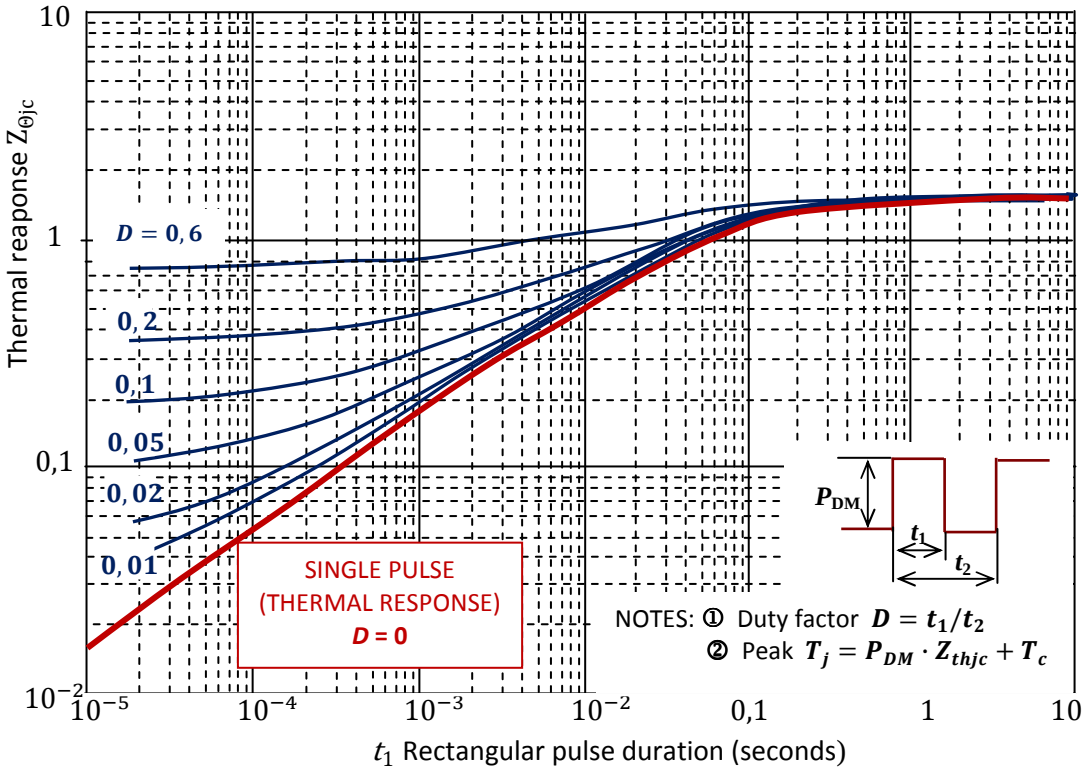


Fig. 14.8 Transient Thermal Impedance Curves (IRF530 HEXFET)

To determine the absolute value of the peak junction temperature, it is, of course, necessary to know the case temperature T_c under steady-state operating conditions. Because of thermal inertia, the heatsink responds only to average power dissipation (except at extremely low frequencies which generally will not be of practical interest). T_c is therefore given by:

where:

ambient temperature

case-to-sink thermal resistance

sink-to-ambient thermal resistance

average power dissipation also,

peak power x duty cycle for rectangular pulses of power

The transient thermal response curve ***assumes constant case temperature***. This is generally valid for pulses shorter than 10 ms. For longer surges the case temperature starts to rise and the results are of questionable accuracy. For operation in free air, case temperature starts to rise within few milliseconds and this curve does not provide any useful information. More sophisticated analytical methods that take the entire thermal system into account are normally used to calculate temperature rise under these conditions.

Example 14.1

Suppose the transistor ( ) must dissipate 10 Watts into still air at 60°C.
How big a heatsink would be needed to keep its junction temperature below 120°C, if ?

☑ Solution:

The temperature difference between the hot junction and the transistor's case will be

.

The temperature difference between the hot case and the heatsink will be

.

So if the junction is at 120°C, the case will be at 105°C and the heatsink at 97°C. So the heatsink must move 10 W across a temperature difference of

.

The heatsink's thermal resistance must then be no more than

.

Or simply:

The designer simply looks through the catalogues for a suitable heatsink of convenient size and shape. If it turns out there isn't one - if the only suitable heatsink is specified at (say) 4°C/watt - then the junction would have to run slightly hotter.

The case temperature would be

,

so the junction would run at , but that should make no practical difference.

Example 14.2

Design flat square plate heat sink with .

☑ Solution:

We choose black Al, vertical, thus and ; mm.

⇒ ⇒

If we use profile in Fig. 14.5, we must realize (see example 14.1) at 10 W; thus needed length is about 75 mm.

Example 14.3

A semiconductor with a junction to case thermal resistance of 1,5 absorbs a single 100 W power pulse for 20 µs. Based on the thermal impedance characteristics in Fig. 14.8, what is the expected junction temperature rise, assuming the case-mount temperature does not respond to this short pulse?

☑ Solution:

The period for a single power pulse is infinite t_2, therefore the duty cycle

.

From Fig. 14.8, for a single 20 µs s pulse . The peak junction temperature is therefore

PEAK

.

The peak junction temperature will rise to 2,2 K above the case mount temperature at the end of the 100 W rectangular power pulse.

Summary

1) Running a semiconductor 10°C cooler could double its life – cooler devices survive longer.

2) The interface resistance depends on the surface finishing, flatness, mounting pressure and the type of interface material.
3) The heat – sink resistance depends on the thickness, area, position (orientation), surface emissivity (shiny/black) and material (Cu, Al, …).
4) The effective heat – sink resistance can be significantly reduced by forced air cooling.
5) Under pulse conditions the junction temperature rises exponentially, according ti its thermal inertia. Because of thermal inertia, the heatsink responds only the awerage power dissipation.
6) The transient thermal response curve assumes constant case temperature. This is generally valid for pulses shorter than 10 ms.

Questions 14

You can find the answers in this text.

1. How does temperature affect the reliability of the elements?
2. What is thermal resistance?
3. What is “Transient Thermal Impedance”?
4. How does airflow affect thermal resistance?

Problems 14

Example 14.1

Suppose the transistor ($R_{jc} = 1,8$ °C/W) must dissipate 8 Watts into still air at 40 °C. How big a heatsink would be needed to keep its junction temperature below 100°C, if $R_{cs} = 0{,}8$ °C/W?.

Example 14.2

Design flat square plate heat sink with $R_{sa} = 4{,}9$ K/W. Choose black Al, vertical, $b = 2$ mm.

Example 14.3

A semiconductor with a junction to case thermal resistance of 1,5 K/W absorbs a single 100W power pulse for $t_1 = 200$ µs. Based on the thermal impedance characteristics in Fig. 14.8, what is the expected junction temperature rise, assuming the case-mount temperature does not respond to this short pulse?.

PROBLEMS KEY 14

Ad example 14.1

4,9 °C/W

Ad example 14.2

$A = \frac{279,5}{4,9-1,304} = 77,7$ cm^2 (it is 8,8x8,8 cm^2)

Ad example 14.3

For a single 200 µs = 2·10^{-4} s pulse $Z_{thjc} \approx 0,075$;

thus PEAK

$$T_j = Z_{thjc} \cdot 100 + T_c = 0,075 \cdot 100 + T_c = 7,5 + T_c$$

Basic texts

[1] Philips Semiconductors, Discrete Semiconductor Packages, Thermal Considerations, Chapter 5, May 1999

[2] International Rectifier: Application Note AN – 949, Current Rating of Power Semiconductor

[3] Seri Lee: How to select a heat sink.

Other texts

[1] Horowitz, P.- Hill, W.: The art of electronics (second edition). pp. 612÷636 Cambridge University Press, Cambridge 1982

[2] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005, ISBN 80-730-161-6

[3] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

References

[1] Mohylová, J.: Lineární obvody s elektronickými prvky -Sbírka příkladů, VŠB-TU Ostrava 2002

[2] Punčochář, J.: Lineární obvody s elektronickými prvky. Skriptum, VŠB-TU Ostrava 2002

[3] Punčochář, J.: Astabilní obvod s reálnými operačními zesilovači. www.elektrorevue.cz

[4] Punčochář, J.: Dolní propusti Sallen - Key s reálnými operačními zesilovači www.elektrorevue.cz

[5] Mohylová, J.: Vliv vektorové chyby invertoru na přenos souhlasné složky signálu diferenčního zesilovače. www.elektrorevue,cz

[6] Mohylová, J.: Sylaby Teorie obvodů I, II a III. Katedra teoretické elektrotechniky FEI, VŠB –TU Ostrava, 1997 – 2001

[7] Punčochář,J.: Operační zesilovače v elektronice. BEN, Praha 2002 (5. vydání)

[8] Čermák, J.: Kurz polovodičové techniky, SNTL, Praha 1976

[9] Huelsman, P. L.: Basic Circuit Theory (3 rd edition). Prentice - Hall, Inc., 1991

[10] Lurje, O. B.: Integralnyje mikroschemy v usilitelnych ustrojstvach. Radio i svjaz, Moskva, 1988

[11] Mikulec, M., – Havlíček, V.: Basic circuit theory II. Vydavatelství ČVUT, Praha, 1996

[12] Čajka, J. - Kvasil, J.: Teorie lineárních obvodů. SNTL, Praha, 1979

[13] Dostál, J.: Operační zesilovače. BEN, Praha, 2005

[14] Angot, A.: Užitá matematika pro elektrotechnické inženýry. SNTL, Praha, 1971

[15] Žalud, V.: Moderní radioelektronika. BEN, Praha, 2000

[16] Vobecký, J. - Záhlava, V.: Elektronika (součástky a obvody, principy a příklady), Grada, Praha 2001

[17] Belza, J.: Operační zesilovače pro obyčejné smrtelníky. BEN, Praha 2004

[18] Horowitz, P.- Hill, W.: The art of electronics (second edition). Cambridge University Press, Cambridge 1982

[19] Frohn, M. – Siedler, H.-J. – Wiemer, M. – Zastrow, P.: Elektronika, polovodičové součástky a základní zapojení. Ben, Praha 2006

[20] Beneš, O. – Černý, A. – Žalud, V.: Tranzistory řízené elektrickým polem, SNTL, Praha 1972

[21] Neumann, P. – Uhlíř, J.: Elektronické obvody a funkční bloky, ČVUT, Praha 1999

[22] Foit, J. – Hudec: Součástky moderní elektroniky, ČVUT, Praha 1996

[23] Lawless, B.: Fundamentals Analogy Electronics, Prentice Hall 1996

[24] Schubert, T. – Kim, E.: Active and non-linear electronics, John Wiley & Sons, Inc. 1996

[25] AN 211A: Field effect transistors in theory and practice, Motorola Semiconductor Applications Note, Motorola, Inc. 1993

[26] Kuphaldt, Tony R.: Lessons In Electric Circuits, www.ibiblio.org/kuphaldt/

[27] Doleček, J.: Moderní učebnice elektroniky 2. díl, BEN, Praha, 2005

[28] Doleček, J.: Moderní učebnice elektroniky 4. díl, BEN, Praha, 2006

[29] Boylestad, R., Nashelsky L.: Electronics Devices and Circuit Theory – seventh edition. Prentice Hall, Ohio, 1998, ISBN-13:978-0137692828

[30] Wai-Kai Chen: Electronics Fundamentals of Circuits and Filters – third edition. Taylor & Francis Group, USA, 2009, ISBN-13:978-1-4200-5887-1

Index