Stellingen behorende bij het proefschrift

FUNDAMENTAL ASPECTS AND DESIGN OF AN FM UPCONVERSION RECEIVER FRONT-END WITH ON-CHIP SAW FILTERS

van

P.T.M. van Zeijl

1. In tegenstelling tot de gebruikelijke opvatting, wordt de optimale signaaloverdracht in SAW vertragingslijnen en SAW transversale filters verkregen door de ingangstransducent met een spanning te sturen en de kortsluitstroom van de uitgangstransducent te meten.
   Dit proefschrift, hoofdstuk 2.

2. Met behulp van ZnO technologie kunnen filters met uitstekende eigenschappen op één chip met elektronische circuits worden gerealiseerd.
   Dit proefschrift, hoofdstuk 2.

3. Voor de keuze van piezoëlektrische materialen voor het maken van SAW-devices is in verband met het maximaliseren van het dynamische bereik, niet alleen de effectieve koppelingsfactor κ^2 van belang, maar moet ook de effectieve dielektrische permittiviteit ϵ in beschouwing worden genomen.
   Dit proefschrift, hoofdstuk 4.

4. Door het minimaliseren van de breedte van een SAW transversaal filter kan het dynamisch bereik van een systeem waarin deze filters worden toegepast, bij een gegeven IC proces en stroomverbruik, worden gemaximaliseerd.
   Dit proefschrift, hoofdstuk 4.

5. Versterkers met een nauwkeurige ingangsimpedantie en een hoge uitgangsimpedantie zijn door middel van het toepassen van gebalanceerde versterkers met impedantie-tegenkoppelnetwerken te realiseren, ook in monolithische techniek.
   Dit proefschrift, hoofdstuk 5.

6. Resultaten die betrekking hebben op een combinatie van verschillende disciplines, zoals de combinatie van SAW-devices met elektronische circuits, leveren zowel problemen op bij de keuze van vakbladen voor een publicatie als bij de acceptatie in de industrie daar deze vakgebieden (ten onrechte) als onafhankelijk worden beschouwd.

7. De door promovendi gegenereerde wetenschap is van minimale betekenis als deze alleen door middel van een proefschrift wordt verspreid.

8. De milieuproblematiek kan serieuzer worden aangepakt als de industrie en de consument ervan overtuigd raken dat apparaten na hun economische en technische levensduur nog geld kosten.

9. De gelijke behandeling van mannen en vrouwen kan sterk worden bevorderd door de militaire dienstplicht af te schaffen.

10. Door het goedkoper worden van elektrische machines voor de doe-hetzelver neemt de kwaliteit van het geleverde werk af.

11. Door het toepassen van technieken zoals die in audio- en TV-apparatuur gebruikelijk zijn, kan de geluidsproduktie in computers en elektronische meetapparatuur worden geminimaliseerd en daarmee de verkoopbaarheid worden vergroot.

FUNDAMENTAL ASPECTS AND DESIGN OF AN FM UPCONVERSION RECEIVER FRONT-END WITH ON-CHIP SAW FILTERS

Paul T.M. van Zeijl

FUNDAMENTAL ASPECTS AND DESIGN OF AN FM UPCONVERSION RECEIVER FRONT-END WITH ON-CHIP SAW FILTERS

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus,
Prof. drs. P.A. Schenck,
in het openbaar te verdedigen
ten overstaan van een commissie
aangewezen door het College van Dekanen
op dinsdag 27 februari 1990 te 14.00 uur

door

Paulus Thomas Maria van Zeijl,

geboren te Monster,
elektrotechnisch ingenieur

Dit proefschift is goedgekeurd door de promotor
Prof.dr.ir J. Davidse.

Dr.ir. E.H. Nordholt heeft als toegevoegd promotor
in hoge mate bijgedragen aan het
totstandkomen van dit proefschrift.

It all looks fine to the naked eye
but it don't really happen that way at all

Naked Eye (Pete Townshend)

Contents

1. Introduction to FM receiver architectures

1.1 Introduction

The last few decades, the architecture of AM and FM consumer receivers has not changed. These downconversion receivers consist of a tuned filter which tracks the desired channel, followed by a frequency conversion to an intermediate frequency (455 kHz for AM and 10.7 MHz for FM). The receiver quality is impaired when varactor tuning is used in the tuned filter. As the received spectrum becomes increasingly occupied, the spurious responses may disturb the reception.

In general it can be said that there is an increasing demand for smaller receivers, preferably realised in a single chip. With the advent of new technologies, such as high-frequency processes and surface acoustic wave (SAW) filters, and better insight in the design of electronic circuitry, a revision of AM and FM receiver architectures is appropriate. In reference [1] such a revision has taken place for the long-wave band (150–365 kHz) and the medium-wave band (512–1625 kHz). In reference [2] the extension to the short-wave band (1.8–30 MHz) has been investigated.

In this thesis the results of the analysis and design of an FM receiver will be presented. Further, the realization of on-chip surface acoustic wave (SAW) filters has been investigated with respect to IC process compatibility and optimum signal transfer. In reference [3] the realization of an on-chip SAW delay line with bipolar electronic circuitry ($f_T = 400$ MHz) has been described. The combination of delay line and electronic circuitry forms a monolithic SAW physical-electronic system for sensors.

A surface acoustic wave delay line can easily be modified to become a transversal filter or resonator filter by means of applying an appropriate aluminum pattern. The high-frequency BIFET process ($f_T = 3$ GHz) that was developed at the Delft University of Technology [4] has been modified and used for the fabrication of the on-chip SAW filters. The combination of on-chip SAW filters and electronic circuitry results in an FM receiver with an on-chip IF filter.

Discrete SAW filters on piezoelectric lithium niobate or quartz are well-known and widely used in consumer TV receivers. With the new technique, on-chip SAW filters are also feasible for this application. Moreover, markets such as high-definition television (HDTV), satellite TV broadcasting receivers, digital radio receivers and measurement equipment, can benefit from the use

of these on-chip SAW filters. On-chip SAW resonators can be used in low phase-noise oscillators in several communication applications.

This research was financially supported by the Netherlands Technology Foundation (STW, Stichting voor de Technische Wetenschappen) and carried out by J.H. Visser and the author. J.H. Visser's research comprises the design of the on-chip SAW filters and the technological aspect [5] while this thesis reports the system design, electronic circuitry design and the optimum signal transfer modeling of the surface acoustic wave devices.

In the following section the function of the FM receiver and the transmission standards for transmitters in the FM broadcast band will be described. In Section 1.3 it will be shown that a direct-detection receiver is not a viable solution. The single-conversion receiver, as treated in Section 1.4, is more promising in this respect. Section 1.5 concludes this chapter with a general discussion of these aspects.

1.2 The characteristics of FM receivers

In this section, the function of the FM receiver and the transmission standards for transmitters in the FM broadcast band will be described. This leads to several receiver properties which will be discussed here in brief. A more complete description with some specifications will be given in Chapter 4.

The function of the receiver is to reproduce information that was originally modulated on a carrier in a specific transmitter. The signal-to-noise ratio or the signal-to-interference ratio of the reproduced information should be below certain specified levels. The transmitted carrier is present in an intensively occupied spectrum. The undesired signals in the received spectrum can be much larger than the desired signal.

In the present investigation, the receiver is assumed to be intended for the public FM broadcast band (87.5–108 MHz). The information present on the carrier contains (stereo) music or speech. Figure 1.1 shows the spectrum of the complete baseband signal [6]. The mono signal is represented by the Left-plus-Right (L+R) signal from 20 Hz to 15 kHz. The difference signal Left-minus-Right (L−R) is amplitude modulated on a 38 kHz carrier, which itself is suppressed. To make the demodulation of this difference signal possible, a 19 kHz pilot signal is added. A Radio Data System signal like RDS or ARI (Autofahrer Rundfunk Information (Broadcast information for motorists)) in Germany [7] or an improved stereo reception signal like in FMX in the USA

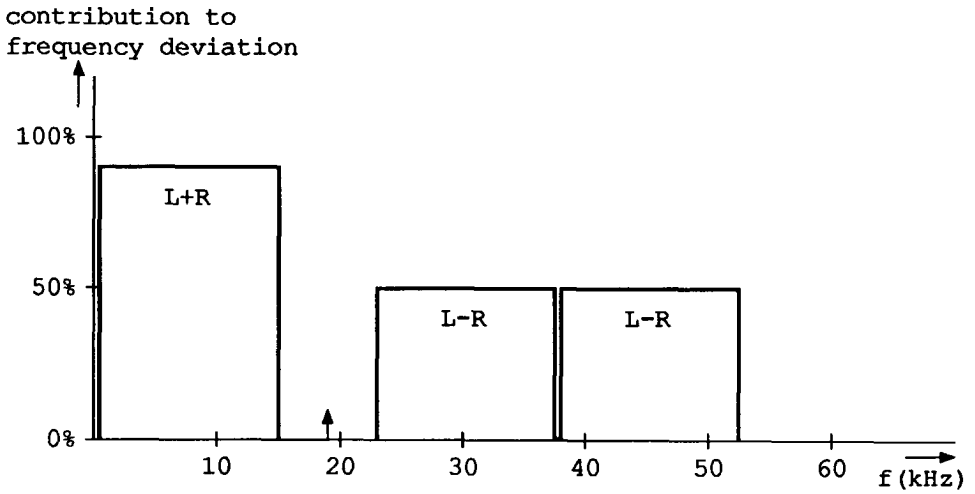

Fig. 1.1 The baseband signal.

[8] may be present, but will not be discussed here.

The total baseband spectrum as shown in Fig. 1.1 is frequency modulated on a carrier in the 87.5–108 MHz broadcast band with a 50 μs pre-emphasis (West-European standard) [6]. The electromagnetic field generated by the transmitting antenna is usually horizontally polarized, i.e. with the electrical vector in the horizontal plane [9].

The frequency deviation for the frequency modulation is maximally ± 75 kHz [6]. The bandwidth of the FM signal occupied in the broadcast band can be estimated by using Carsons rule and amounts to 180 kHz for mono or 256 kHz for stereo [10]. Measurements performed with program material in both mono and stereo show a somewhat smaller bandwidth of about 94 kHz for mono and 170 kHz for stereo (when colored noise instead of program material is used: 147 kHz for mono and 211 kHz for stereo) [11].

The frequency difference of the carrier frequencies of the FM transmitters is an integral multiple of 100 kHz. Transmitting stations, which are located within each others receiving ranges, have their transmitter frequencies at least 300 kHz apart. Frequencies of stations which are more remotely located may well be separated by a 100 kHz frequency difference. Therefore, a bandwidth of 180 kHz will be used throughout this thesis.

It must be kept in mind that the transmitted signal is determined by the specifications of the transmission standard and hence cannot be altered. The

receiver architecture, however, may be chosen to accommodate an available technology, which depends on the desired performance.

If only one carrier is present in the RF spectrum, the amplitude of this carrier signal is unknown or variable, which depends on the received station and/or the transmitter-receiver distance. Galactic noise or man-made noise determines the sensitivity or lower limit for the detection of the signal for very long transmitter-receiver distances. The receiver noise figure determines the sensitivity if these galactic and man-made noise sources are sufficiently small. When the receiver is close to the transmitter, the amplitude of the carrier signal is maximal and care has to be taken to handle the carrier signal without any resulting distortion in the demodulated signal. The ratio of maximum signal and noise level is called the dynamic range and may well exceed 100 dB for signals in the FM band.

When the RF spectrum becomes more fully occupied, the unwanted signals may be larger or smaller than the desired signal. These unwanted signals may interfere with the desired signal. Due to the mixing process in a single conversion receiver (see Section 1.4), there is an image channel at which the sensitivity of the receiver equals that of the desired channel. Mixers are preferably driven by square waves to improve the noise and intermodulation behavior [1]. The harmonics, however, give rise to spurious responses. The signals causing these image channel responses and these spurious responses should be attenuated as much as possible for a sufficiently high signal-to-interference ratio.

Selectivity is determined by the amount of attenuation applied to adjacent channels, which otherwise disturb the detected signal. A poor carrier-to-noise ratio in the local oscillator also limits the selectivity due to reciprocal mixing.

If the RF path or the IF path of the receiver shows any non-linear transfer, the intermodulation components of large undesired signals may dominate over a small desired signal. The receiver must therefore have a sufficiently high intermodulation-free dynamic range (IMFDR).

In the following section the above-mentioned general considerations concerning the FM receiver will be used to discuss the feasibility of the simplest receiver architecture. The improved receiver principle is treated in Section 1.4.

1.3 The tuned radio frequency receiver

In this section the simplest type of receiver will be discussed: the direct-detection receiver in the case of a selective detector and the tuned radio frequency receiver in the case of a non-selective detector. Figure 1.2 gives a block

diagram of such a tuned radio frequency receiver. The tuned radio frequency receiver consists of a channel-selective band-pass filter followed by an FM detector circuit. As different stations have to be received, both the band-pass filter and the FM detector have to be tuned to the desired station frequency.

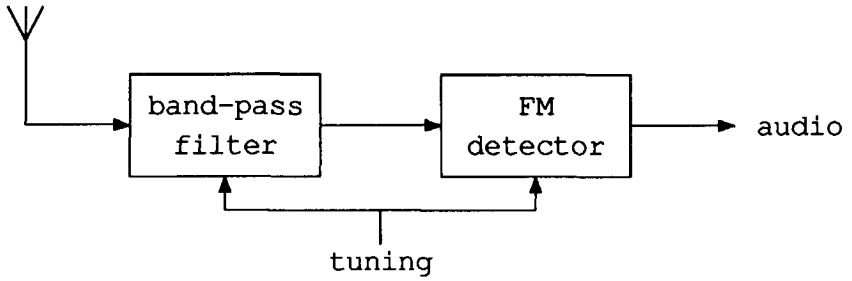

Fig. 1.2 The tuned radio frequency receiver.

The input filter must provide the required selectivity and the bandwidth of this filter must not increase during tuning to maintain this selectivity. The center frequency of the filter must track accurately with the center frequency of the FM detector to maintain the desired sensitivity. The only tunable filters applicable at these frequencies are LC filters. Filters constructed with mechanically tuned inductors and capacitors, when properly terminated, have a very high intermodulation-free dynamic range. However, a constant-bandwidth higher-order band-pass filter which tracks with the desired station frequency leads to unwieldy electro-mechanical constructions. These LC filters cannot be integrated. Tunable capacitors like varicaps are suitable for integration. However, they are non-linear devices and degrade the intermodulation-free dynamic range.

The question whether FM detectors can be tuned or if the properties of such detectors are constant over a certain range of input frequencies or input amplitudes, is more or less irrelevant because of the difficulty in realizing the proper channel-selective tunable input filters.

A phase-lock loop (PLL), however, is a selective FM detector, which may be coupled directly to the antenna without a band-pass filter. When a phase-lock loop is locked to the carrier, the disturbance caused by adjacent channels is lower than in the case of a non-selective FM detector. In Fig. 1.3 the basic block diagram of a phase-lock loop is shown. The phase-lock loop consists of a phase-detector or mixer, a low-pass filter and a controlled oscillator (VCO).

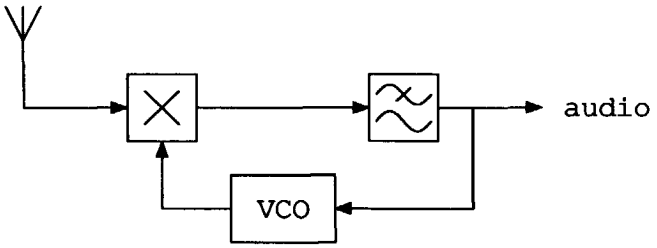

Fig. 1.3 The phase-lock loop.

Generally, the PLL locks at the largest signal available in the supplied spectrum. If the amplitude of the desired carrier signal is smaller than the amplitude of an adjacent channel signal, special measures, like the use of multiloop synthesizers instead of a single VCO, must be taken to ensure that the PLL locks at the desired carrier signal.

For a correct reproduction of the information, the bandwidth of the PLL must be higher than 53 kHz. A second-order PLL with a 53 kHz bandwidth attenuates the adjacent channel at 300 kHz about 15 dB. This adjacent channel attenuation may be increased by using a modified and complicated PLL [12].

As the loop gain of the PLL depends on the amplitude of the input signal, some kind of automatic gain control circuitry is needed to keep the behavior of the PLL constant over a large range of input signals [13], [14]. The offset, present in practical circuitry, then still limits the dynamic range of the PLL to about 40 dB. Techniques such as chopping [12] or auto-zeroing [13], [14] are required to increase this dynamic range. When a switching mixer is used or square wave signals are applied to the phase-detector, the PLL becomes sensitive at the harmonics of the in-lock VCO frequency. Signals present at the input at these harmonics generate DC-voltages at the phase-detector output which cannot be eliminated by auto-zeroing or chopping.

To obtain dynamic ranges in the order of 100 dB with a (modified) PLL, therefore, either becomes impractical or implies the use of very complex circuitry. In the following section the single-conversion receiver will be discussed, for which the desired dynamic range is feasible.

1.4 The single-conversion receiver

Problems like tracking channel-selective filters in tuned radio frequency receivers or offset and signal dependent behavior in phase-lock loops can be circumvented by converting the radio frequency (RF) signal at the input into an intermediate frequency (IF) signal by applying the heterodyne principle. In Fig. 1.4 the basic principle of this single-conversion receiver is sketched. The single-conversion receiver consists of a local oscillator (LO) with mixer

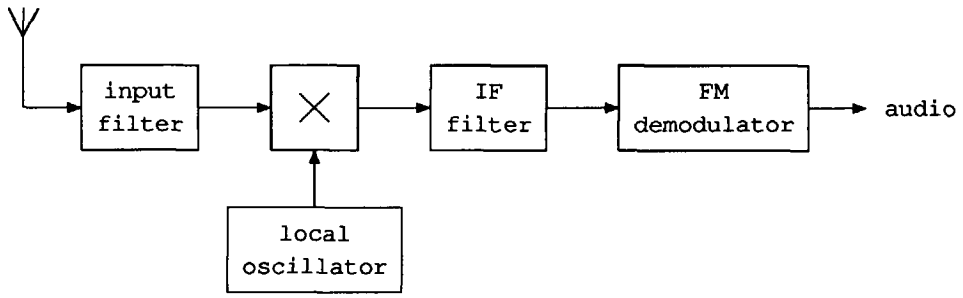

Fig. 1.4 The single-conversion receiver.

which converts the RF input signals into a fixed intermediate frequency. At the intermediate frequency a channel-selective fixed-frequency filter selects the desired channel after which demodulation can take place.

An ideal analog multiplying mixer delivers two signals, with frequencies $f_{RF} = |f_{LO} \pm f_{IF}|$, to the IF filter. One of these signals represents the desired signal while the other is the image response. Theoretically, the image response can be eliminated by using two mixers driven in quadrature. Practically, a maximum image suppression of only 40 to 60 dB at low frequencies seems feasible without adjustments [15].

The image rejection filter may be omitted if there are no signals to be expected at the image channel. This idea has been used to develop a single-chip FM receiver [16]. With a channel spacing of 300 kHz, there should be no signal halfway between two channels, i.e. at 150 kHz from the desired channel. Therefore, an intermediate frequency of about 75 kHz has been chosen. To prevent aliasing at 0 Hz, a frequency deviation reduction is mandatory and results in a complex receiver architecture. In practice, signals may be present at 150 kHz from the carrier and, if this is the case, they disturb the reception. When there are no signals present at the image channel and the input filter is

omitted, the noise of the image channel increases the receiver noise figure by 3 dB.

The frequency difference between desired response and image response is twice the intermediate frequency. In practical situations this means that, if the width of the RF band to be received is larger than twice the chosen intermediate frequency, the input filter must be tuned to the desired channel to obtain a sufficiently high image rejection.

The input filter can also be used to attenuate the LO harmonics-related spurious responses of switching mixers by the proper choice of the intermediate frequency, as now will be elucidated. Switching mixers perform much better with respect to noise and distortion than analog multiplying mixers [1]. In switching mixers, the RF input signals are not multiplied with a sinusoidal signal, but with a square-wave signal. The even-order (in the case of unbalance) and odd-order harmonics result in spurious RF frequencies at which the receiver is also sensitive. Although the conversion gain in the mixer is lower for these LO harmonics-related spurious responses, these responses cannot be neglected. There are six different situations, which depend on the choice of the local oscillator frequency f_{LO} and the intermediate frequency f_{IF}:

1.	$f_{IF} < f_{LO} < f_{RF}$	$f_{LO} = f_{RF} - f_{IF}$
2.	$f_{IF} < f_{RF} < f_{LO}$	$f_{LO} = f_{RF} + f_{IF}$
3.	$f_{LO} < f_{IF} < f_{RF}$	$f_{LO} = f_{RF} - f_{IF}$
4.	$f_{LO} < f_{RF} < f_{IF}$	$f_{LO} = f_{IF} - f_{RF}$
5.	$f_{RF} < f_{LO} < f_{IF}$	$f_{LO} = f_{IF} - f_{RF}$
6.	$f_{RF} < f_{IF} < f_{LO}$	$f_{LO} = f_{IF} + f_{RF}$

The first three situations are denoted as downconversion receivers because the intermediate frequency is below the lowest desired radio frequency f_{RF}. The last three situations are denoted as upconversion receivers.

Situation 2 represents the conventional downconversion receiver. The intermediate frequency is 10.7 MHz. The RF band ranges from 87.5 to 108 MHz and the local oscillator is tunable from 98.2 to 118.7 MHz. As the image channels range from 108.9 to 129.4 MHz, tracking filters are inevitable for suppressing the image and LO harmonics-related spurious channels.

It can be shown that the frequencies of all third- and higher-order LO harmonics-related spurious responses in situation 6 are higher than the image frequency [1]. Only the frequency of one of the spurious channels due to the second LO harmonic is lower than the frequency of the image channel. The mixer conversion gain associated with this spurious response is usually low. As the image frequency is far away from the RF band frequencies, a non-tuned low-pass filter of a sufficiently high order will attenuate the signals

at the image channel and the spurious channels below a specified level. To also attenuate all signals below 87.5 MHz a non-tuned band-pass filter may be applied.

The use of the upconversion principle results in an intermediate frequency higher than 108 MHz. To avoid problems such as pin-to-pin feedthrough at these high frequencies, all high frequency circuits and filters should be monolithically integrated on the same chip. Only the input signal, some bias voltages, the tuning voltage and the output baseband signal should occur at the pins of the IC. By using a silicon-integrated SAW filter as the IF filter, a monolithically integrated FM upconversion receiver is feasible.

A disadvantage of the upconversion receiver with a fixed input filter which passes the whole RF band, compared to the downconversion receiver with its tuned band-pass filter, is the required handling of the thus passed RF spectrum. Besides, as the intermediate frequency is higher than the frequency of the RF signal, more stringent demands are imposed on the electronic circuitry with respect to the high-frequency behavior. Properly designed amplifiers using negative feedback provide a solution to this problem [17].

1.5 Discussion

In the first section of this chapter it was proposed that a revision of FM receiver architectures is appropriate due to the advent of new technologies. On-chip surface acoustic wave devices is one of these new technologies and might be used not only in various applications such as AM, FM, TV and communication receivers, but in measurement equipment in both the consumer and professional field as well.

With the use of the receiver characteristics as given in Section 1.2, the tuned radio frequency receiver as discussed in Section 1.3 does not seem to be a viable solution for a single-chip FM receiver. Complex circuitry must be added to a phase-lock loop to extend its dynamic range, but even then it is unlikely that the desired dynamic range can be realized.

Most of the problems of the tuned radio frequency receiver and the phase-lock loop can be overcome by using the single-conversion receiver. By choosing the intermediate frequency and the local oscillator frequency higher than the frequency of the desired RF station frequency, all spurious responses of the receiver can be attenuated by a simple non-tuned low-pass or band-pass filter of a sufficiently high order. That such an upconversion receiver generally requires a larger effort in the design of the electronic circuitry at higher frequencies, and that more stringent demands are posed on the feedthrough in the IF filter, are disadvantages of the upconversion principle as compared to that of downcon-

version. By using properly designed amplifiers employing negative feedback and applying on-chip surface acoustic wave filters, these disadvantages can be surmounted.

In the present thesis only the front-end up to and including the IF filter of the upconversion FM receiver will be discussed. The models for the SAW transversal filter and the SAW resonator filter, which can both be applied as IF filters, are indispensable for the design of the FM upconversion receiver front-end. They will, therefore, be presented in Chapter 2.

As the frequencies of the image channels in the FM upconversion receiver are higher than 330 MHz ($f_{LO} > f_{IF} > f_{RF}$), the antenna behavior at these frequencies must be known in order to estimate the effect on the image rejection. The results of a literature study in combination with some measurements performed on a portable receiver will be presented in Chapter 3.

Based on the knowledge acquired in Chapters 2 and 3, the design of the FM upconversion receiver front-end for a portable receiver will be undertaken in Chapter 4. In Chapter 5 a new class of amplifiers with two feedback loops will be discussed. One of these amplifiers is realized in a 3 GHz BIFET process and can be applied as input amplifier in the FM front-end. In Chapter 6 the results obtained in this research will be summarized and some recommendations for further research will be given.

2. SAW filter modeling

2.1 Introduction

In Chapter 1 it was indicated that silicon-integrated surface acoustic wave (SAW) filters are essential in realizing an integrated upconversion receiver for the FM broadcast band. Hence, the modeling of SAW delay lines, SAW transversal filters and the behavior of SAW resonators and SAW resonator filters will be discussed in this chapter. The modeling will be done in such a way as to show the system and circuit designer the degrees of freedom with respect to e.g. dimensions and the transfer function. Moreover, the feasibility of a monolithic integration of SAW filters with high-frequency electronic circuitry is demonstrated. For a survey of the potentials of SAW devices in general, see reference [18].

In Fig. 2.1 the cross-section of an on-chip SAW device and bipolar transistor is shown. On a standard Si-substrate with epi-layer, as used in high-frequency

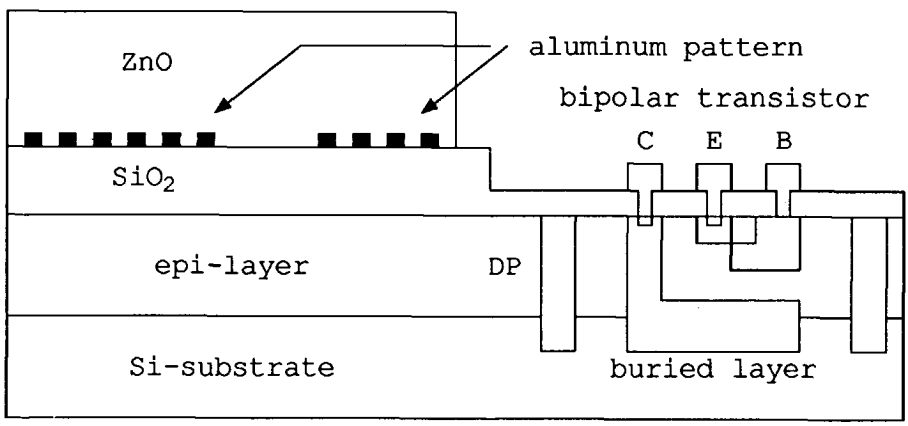

Fig. 2.1 Cross-section of an IC with a SAW device and a bipolar transistor.

bipolar processes, a silicon-dioxide layer is thermally grown. An aluminum layer, evaporated on the SiO_2 layer, is etched to form the metalization pattern for a delay line, a transversal filter or a resonator. On top of the aluminum pattern a (piezoelectric) ZnO layer is sputtered, which is necessary to obtain the conversion from electrical power to acoustic power and vice versa.

These on-chip SAW devices can be made in the frequency range from practically 10 MHz to about 3 GHz (limited by lithography and etching resolution). In theory, a temperature-independent filter center frequency can be obtained by choosing the proper layer thicknesses [19]. The feasibility of ZnO thin-film SAW devices on glass substrates which can be used as IF filters in TV receivers has already been demonstrated in reference [20]. The reader is referred to the Ph.D. Thesis of J.H. Visser [5] for a rigorous treatment of the technology, the optimum thicknesses of the SiO_2 and ZnO layers with respect to the transfer function, the alignment of the aluminum pattern with respect to the crystal orientation and the influence of the extra layers on the process.

In Fig. 2.2 the aluminum metalization pattern in the case of a SAW delay line is shown. The delay line consists of two uniform interdigital transducers

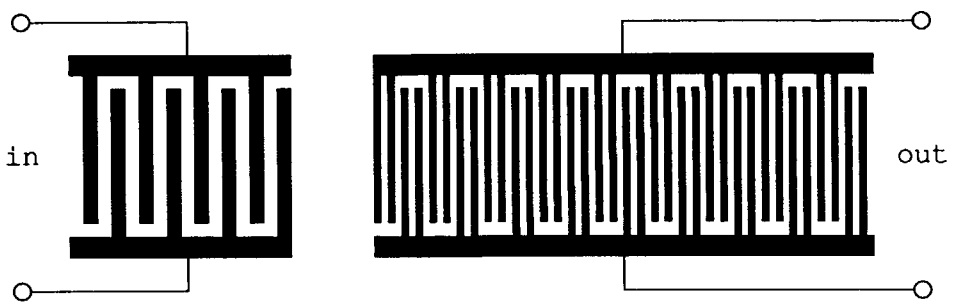

Fig. 2.2 A metalization pattern for a SAW delay line.

(IDTs), one converts the electrical power into acoustic power while the reverse process takes place in the other IDT. The modeling of the SAW delay line will be discussed in Section 2.2. It will be shown that the transfer function of the delay line has the property of a band-pass filter.

Each aluminum strip placed in the ZnO-SiO_2 interface forms a discontinuity for the passing acoustic wave and reflects therefore a small part of such a wave (usually $< 1\%$). For single electrodes, as used in the IDT on the left-hand side in Fig. 2.2, these reflected waves add in phase at the center frequency of the band-pass filter. The desired filter transfer function is seriously distorted when IDTs with a large number of these single electrodes are used. By using double or split electrodes, as used in the IDT on the right-hand side in Fig. 2.2, the reflected waves add in phase at twice the center frequency of the band-pass filter. The effect of the reflected waves can then be neglected for frequencies in the vicinity of the filter center frequency.

By modifying the overlap of the electrodes in one (originally uniform) IDT (apodization), according to the desired impulse response, the band-pass filter function of the delay line can be improved. In Fig. 2.3 the aluminum metalization pattern of such a transversal filter is shown. The overlap of the electrodes

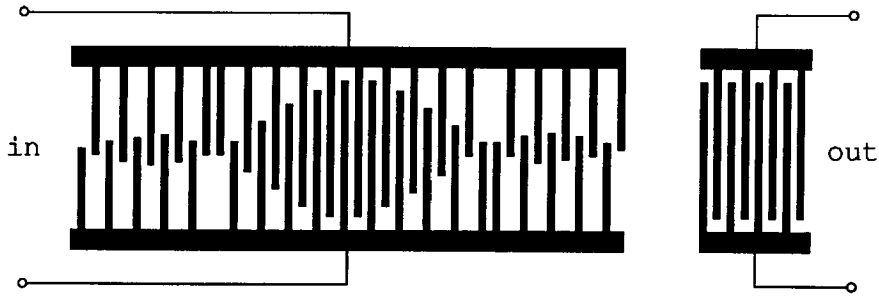

Fig. 2.3 A metalization pattern for a SAW transversal filter.

shows a $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$ function which has been truncated at the second null of the sinc function. The modeling of the SAW transversal filter will be discussed in Section 2.3.

The small reflection of an aluminum strip on the $ZnO\text{-}SiO_2$ interface has a detrimental influence on the band-pass filter function of the SAW delay line and the SAW transversal filter. This effect, however, can, when a large number of electrode pairs are employed, also beneficially be employed to construct resonators. In Fig. 2.4 the aluminum metalization pattern for a two-port SAW resonator is shown. The large number of strips on the left-hand and right-hand

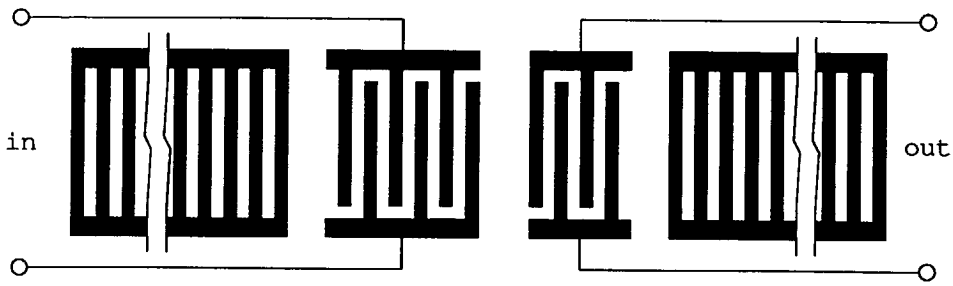

Fig. 2.4 A metalization pattern for a two-port SAW resonator.

sides of the two IDTs form two reflectors. Two IDTs are placed in the cavity which is formed by the reflectors. These IDTs couple electrical power into and out of this cavity. The reflectors in Fig. 2.4 can be seen as a large number of electrodes in an IDT. In Section 2.2 it will be shown that an IDT which is not short-circuited, regenerates acoustic waves. This regeneration of surface acoustic waves by the reflector array or reflector IDT might cause an undesired disturbance in the standing wave pattern in the cavity. By short-circuiting the reflector IDTs the disturbance of the standing wave pattern in the cavity is minimized. The quality factor Q of the resonator can be defined accurately by choosing the proper parameters for the SAW resonator, thus creating resonator filters. The behavior of the SAW resonator and the SAW resonator filter will be discussed in Section 2.4.

There is an essential difference in the properties of SAW delay lines and SAW transversal filters on the one hand and the SAW resonator filters on the other hand. Resonators are minimum phase devices, which means that the real and imaginary parts of the transfer function are coupled to each other via the Hilbert transform. The amplitude transfer function and the phase transfer functions can, therefore, not be designed independently. SAW delay lines and transversal filters are non-minimum phase devices. For these devices, the amplitude transfer function and the phase transfer function can be designed independently. How this property effects the performance of the FM receiver will be discussed in Chapter 4.

The feasibility of a monolithic integration of SAW filters with high-frequency electronic circuitry will be demonstrated in Section 2.5. This demonstration of the feasibility will be supported by several measurement results from a chip on which a number of SAW delay lines, in combination with a transimpedance amplifier and a bipolar double-balanced mixer, have been integrated. Finally, Section 2.6 concludes this chapter with a discussion. This chapter contains material previously published in references [21] and [22].

2.2 SAW delay line modeling

2.2.1 Introduction

In most articles on SAW devices, there is a misconception concerning the driving and loading of SAW devices, namely that a power transfer is indispensable for a signal transfer. To illustrate the results that arise from a maximal power transfer, the impulse response of a SAW transversal filter on a lithium niobate substrate has been measured. Figure 2.5 shows the measured impulse response for a maximal power transfer, i.e. the static IDT capacitances have

been tuned out by inductances and the filter is terminated at both the input and the output with an impedance equal to the input and output impedance of the SAW device, respectively. The three most important responses in this fig-

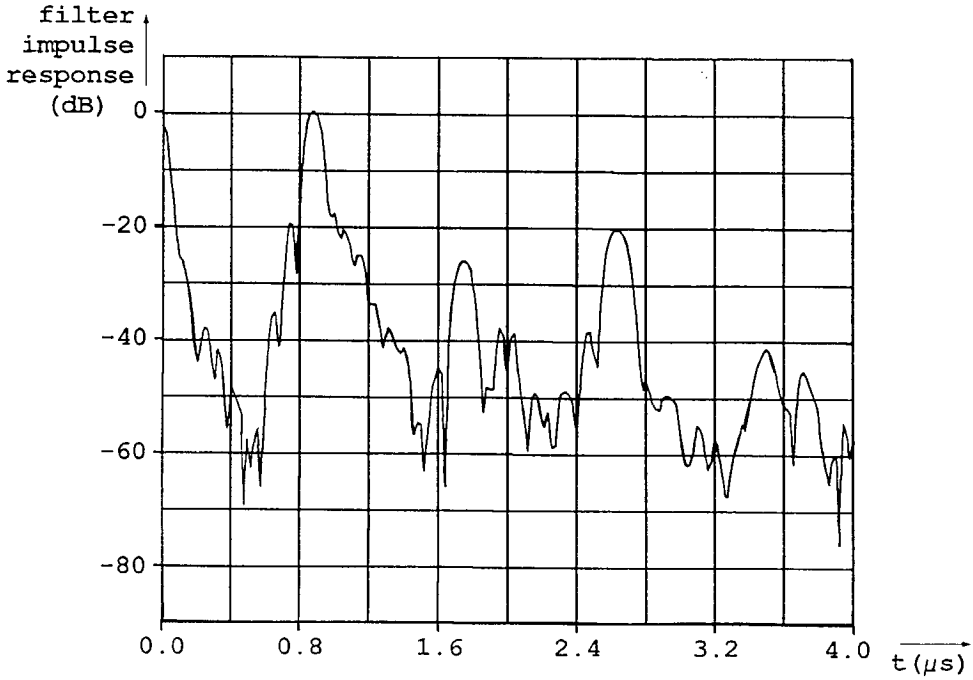

Fig. 2.5 The normalized impulse response for a SAW transversal filter on lithium niobate, in the case of maximum power transfer.

ure are the electromagnetic feedthrough (without significant delay), the main (desired) impulse response after 0.9 μs (corresponding to the center-to-center distance between the two IDTs) and the triple transit echo after 2.7 μs. The impulse response has been normalized, i.e. 0 dB for the desired response. It will be shown in Section 2.5 that although the electromagnetic feedthrough is just as high as the desired impulse response in this design, the electromagnetic feedthrough can be adequately suppressed, even at high frequencies.

The triple transit echo, however, is a direct consequence of the desire to maximize the power transfer [23]. A triple transit echo of -20 dB compared to the desired impulse response, gives rise to a ripple in the amplitude transfer of 3.5 dB peak-to-peak and a ripple of 360 ns peak-to-peak in the group delay for this wide-band filter ($B = 5$ MHz). Such a filter transfer function seriously

distorts the information that is FM modulated on a carrier [24]. It is obvious that such a filter transfer function cannot be tolerated in systems, like the present FM receiver, where a high-performance information transfer is desired.

To avoid the problems resulting from a triple transit echo, some authors have deliberately chosen for a power mismatch at both the input and the output of their SAW filters [25], [26], thus minimizing the power transfer. The optimum signal transfer, as will be shown in Section 2.2.2, however, is realized by driving the SAW device by a voltage and sensing the short-circuit current at the output of the SAW device. It should be noted that, although a certain power is required for driving the SAW devices, no power is extracted for the retrieval of the (filtered) information. Figure 2.6 shows the measured (normalized) impulse response for the voltage-driven and current-sensed transversal filter. As

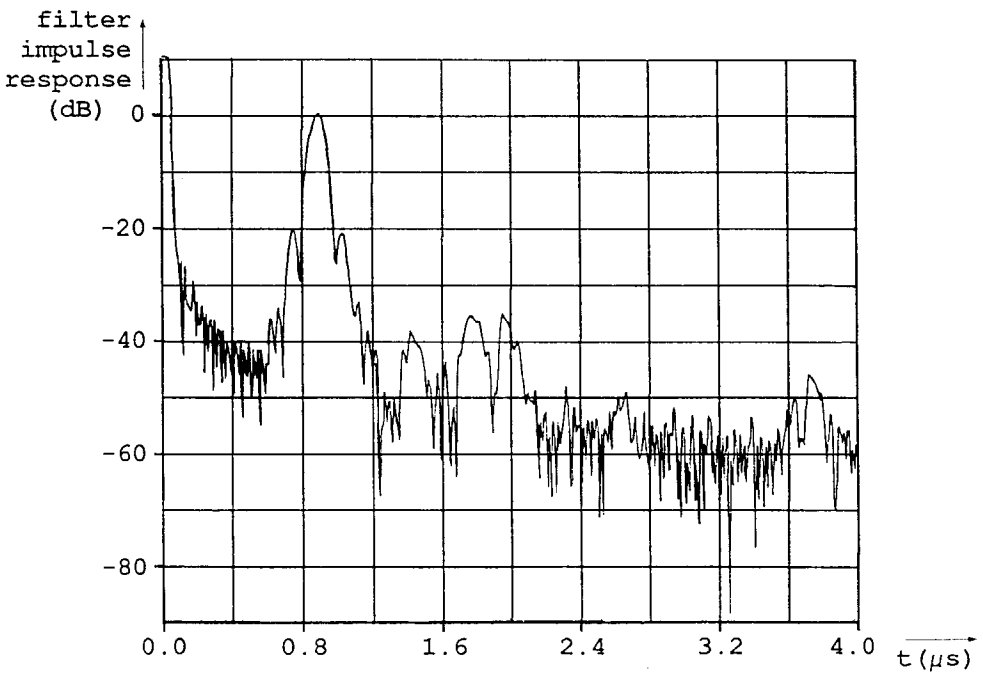

Fig. 2.6 The normalized impulse response for a SAW transversal filter on lithium niobate, in the case of optimum signal transfer.

can be seen from this figure, the triple transit echo is at least 50 dB attenuated compared to the desired response, resulting in negligible ripple in the amplitude transfer ($\approx$ 0.11 dB) and negligible ripple in the group delay ($\approx$ 11

ns) within the filter pass-band. Such a filter transfer function is excellently suited in systems for high-performance information transfer.

For a classification and a design procedure concerning amplifiers that can drive a load with a voltage or that can sense a current, while maintaining the desired quality of information transfer, the reader is referred to reference [17]. The measurements above were performed on a SAW transversal filter on a lithium niobate substrate. The conclusion concerning the driving and loading of the SAW devices is also valid for SAW devices in the ZnO-SiO_2-Si, three-layered structure, as will be shown in the next section.

2.2.2 The development of the SAW delay line model

In Fig. 2.7 the SAW delay line is shown again. We will assume that the variation in the surface wave velocity due to the metalization is small, that the Si-substrate is conducting and that plane surface waves are generated by the IDTs. Further, the dispersion is assumed to be negligible, which can be obtained by the proper choice of the SiO_2 and ZnO layer thicknesses [5]. The

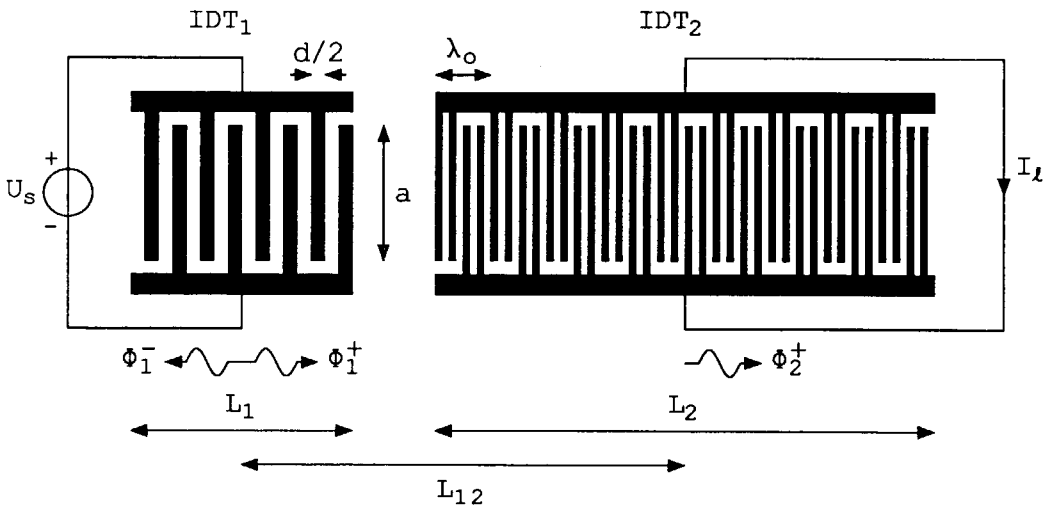

Fig. 2.7 The SAW delay line with source and load.

interdigital transducer on the left side of Fig. 2.7 (IDT_1, N_1 electrode pairs) is connected to the voltage source U_s. This voltage gives rise to two acoustic waves: ϕ_1^- traveling to the left and ϕ_1^+ traveling to the right. The wave ϕ_1^+ is

delayed by the time L_{12}/v_o (v_o denotes the velocity of the surface wave) and attenuated before arrival at the center of IDT_2. This corresponds to a change in phase ($-2\pi f L_{12}/v_o = -\omega L_{12}/v_o$) and a change in amplitude ($e^{-\alpha L_{12}}$, α is in the order of 0.5 per cm and depending on the layer thicknesses in the ZnO-SiO_2-Si structure and the frequency). The wave at the center of IDT_2 is then given by $\phi_2^+ = \phi_1^+ e^{-\alpha L_{12}} e^{-j\omega L_{12}/v_o}$ and induces a current I_l in IDT_2. The conversion from source voltage U_s into an acoustic wave is denoted by $\mu_1(f)$ and the conversion of acoustic wave to load current I_l is denoted by $g_2(f)$, respectively, as given by [27]:

$$\phi_1^+ = \mu_1(f) \cdot U_s \,, \tag{2.1a}$$

$$I_l = g_2(f) \cdot \phi_2^+ \,. \tag{2.1b}$$

By using these two equations the conclusion, that a voltage-driven and current-sensed SAW delay line gives the optimum signal transfer, will now be elucidated. Suppose that we have a delay line in which a surface acoustic wave travels from IDT_1 to IDT_2. This wave generates a current in IDT_2. If this IDT is not short-circuited, a voltage is generated over it. According to Equation 2.1a, this voltage generates an extra surface acoustic wave of which a part travels to IDT_1. If the same phenomenon takes place at IDT_1, the twice regenerated wave arrives at IDT_2. It has traveled the acoustic path three times and is therefore called the triple transit echo. By short-circuiting one of the two IDTs, or both IDTs, the regeneration of these triple transit echos can be minimized. Therefore, the input IDT must be driven by a voltage and at the output IDT the short-circuit current must be sensed (see also references [28] and [29]).

The transconductance from IDT_1 to IDT_2 can easily be calculated and is given by:

$$G_{12}(f) = \frac{I_l}{U_s} = \mu_1(f)\, g_2(f)\, e^{-\alpha L_{12}} e^{-j\omega L_{12}/v_o} \,. \tag{2.2a}$$

The transconductance from IDT_2 to IDT_1 can be calculated in the same manner and is given by:

$$G_{21}(f) = \mu_2(f)\, g_1(f)\, e^{-\alpha L_{12}}\, e^{-j\omega L_{12}/v_o} \,. \tag{2.2b}$$

For small signals and thus linear behavior of the device, the reciprocity principle can be applied: $G_{12}(f) = G_{21}(f)$ and so the ratio $\mu(f)$ over $g(f)$ is constant, independent of the IDT geometry. It can be shown that this ratio equals half the characteristic impedance Z_o of the acoustic transmission line. For substrates like quartz and lithium niobate [27] and for the present

ZnO-SiO_2-Si structure [30], [31], this ratio is given by:

$$\frac{\mu_1(f)}{g_1(f)} = \frac{\mu_2(f)}{g_2(f)} = \frac{Z_o}{2} = \frac{\kappa^2}{2\omega_o \epsilon a} \,. \tag{2.3}$$

The filter center frequency $f_o = \omega_o/(2\pi)$ is given by $f_o = v_o/\lambda_o$, where λ_o is defined in Fig. 2.7. The parameter κ^2 denotes the piezoelectric coupling coefficient of the ZnO-SiO_2-Si structure, ϵ denotes the effective dielectric permittivity of the three-layered ZnO-SiO_2-Si structure (in F/m) and a is the aperture (width) of the IDTs (in m). In the literature, the parameter ϵ is often denoted as C_t representing the capacitance of one electrode pair. It should be noted that κ^2 and ϵ depend on the thicknesses of the SiO_2 and ZnO layers for the present ZnO-SiO_2-Si structure [5]. The transconductance $G_{12}(f)$ of the delay line can now be written as:

$$G_{12}(f) = G_{21}(f) = \mu_1(f)\, \mu_2(f)\, \frac{2}{Z_o}\, \mathrm{e}^{-\alpha L_{12}}\, \mathrm{e}^{-j\omega L_{12}/v_o} \,. \tag{2.4}$$

So far we have determined the transfer function of the SAW delay line as a function of $\mu_1(f)$ and $\mu_2(f)$. For the design of the amplifiers driving and loading the SAW delay line, knowledge about the input and output impedances of the delay line is essential. These impedances will now be calculated for a voltage-driven and current-sensed SAW delay line. The impedance consists of two parts, namely the capacitance due to the electrodes of the IDT and the impedance due to the acoustic transmission line. The transducer capacitance can be considered as a static capacitance. For periodically spaced single electrodes with metalization ratio $d/\lambda_o = 0.5$ (see Fig. 2.7) the IDT capacitance C_1 can be approximated by [27]:

$$C_1 = \epsilon a N_1 \,, \tag{2.5a}$$

and the capacitance C_2 for double or split electrodes can be approximated by [27]:

$$C_2 = \sqrt{2}\, \epsilon a N_2 \,, \tag{2.5b}$$

where N_1 and N_2 denote the number of electrode pairs.

The acoustic transmission line gives rise to the admittance $Y_1(f) = G_1(f) + jB_1(f)$, where $G_1(f)$ is the radiation conductance and $B_1(f)$ is the radiation susceptance. The radiation susceptance $B_1(f)$ can be calculated from $G_1(f)$ by applying the Hilbert transform to $G_1(f)$. See, for instance, reference [32] for a definition of the Hilbert transformation and reference [33] for a tabulation of the most used functions with their Hilbert transformed functions. We will

first concentrate on the calculation of $G_1(f)$. When the electrical and acoustic losses are negligible and the output IDT is short-circuited, the power delivered to the IDT and the power of the acoustic wave are equal. The electrical power applied to the IDT is $|U_s|^2 \cdot G_1(f)$. The power of the surface acoustic wave is equal to $|\mu_1(f)|^2 \cdot |U_s|^2 \cdot (2/Z_o)$. The radiation conductance $G_1(f)$ is then given by [27]:

$$G_1(f) = |\mu_1(f)|^2 \frac{2}{Z_o} . \tag{2.6}$$

The result for $G_2(f)$ is similar. The radiation susceptance $B_1(f)$ and $B_2(f)$ can be calculated by applying the Hilbert transform to $G_1(f)$ and $G_2(f)$.

Thus far, the transconductance $G_{12}(f)$, the input admittance $Y_1(f)$ and the output admittance $Y_2(f)$ have been calculated as a function of $\mu_1(f)$ and $\mu_2(f)$, respectively. The functions $\mu_1(f)$ and $\mu_2(f)$ are determined by the charge distribution on the IDTs. The charge distribution on the electrodes at the end of an IDT is different from that on the electrodes in the middle of an IDT. For accurate analysis, computer simulations are necessary to obtain the charge distributions. For obtaining insight, we will neglect these end-effects. The delta model [34] will be used for the calculation of $\mu(f)$. The impulse response $\mu_1(t)$ (the inverse Fourier transformed function of $\mu_1(f)$) consists of an infinite row of delta pulses $\mu_{s1}(t)$ (a sampling function with fundamental frequency f_o) multiplied with an impulse response $h_1(t)$, describing the overlap of the electrodes. The transfer function $\mu_1(f)$ can then be written as:

$$\mu_1(f) = \mu_{s1}(f) * H_1(f) , \tag{2.7}$$

where $*$ denotes convolution, $H_1(f)$ is the Fourier transformed function of $h_1(t)$ and $\mu_{s1}(f)$ is the Fourier transformed function of $\mu_{s1}(t)$. The sampling function would result in equal-height filter pass-bands at multiples of f_o. To correct for this restriction in the delta model, a correction constant c_i will be introduced in $\mu_{s1}(f)$:

$$\mu_{s1}(f) = f_o \sum_{i=-\infty}^{\infty} c_i \, \delta(f - if_o) , \tag{2.8}$$

where c_i represents the frequency dependent response of the electrode pairs at frequency if_o. For metalization ratio $d/\lambda_o = 0.5$, the correction constant at $f = f_o$ is given by $c_1 = j\kappa^2\sqrt{2/\pi}$ [27]. The value of c_i is given in reference [35] for single electrodes and in reference [36] for split electrodes as a function of the metalization ratio d/λ_o.

The impulse response $h_1(t)$, describing the overlap of the electrodes of a non-apodized (uniform) IDT, when the end-effects are neglected, equals a

gate function as shown in Fig. 2.8a, its Fourier transform $H_1(f)$ is shown in Fig. 2.8b. These functions are described by the relations:

$$h_1(t) = \text{rect}\left(\frac{t}{L_1/v_o}\right) , \tag{2.9a}$$

$$H_1(f) = \frac{L_1}{v_o} \cdot \frac{\sin(\pi f L_1/v_o)}{\pi f L_1/v_o} = \frac{N_1}{f_o} \cdot \text{sinc}\left(\frac{f}{v_o/L_1}\right) , \tag{2.9b}$$

where $L_1 = N_1 \cdot \lambda_o$ and $v_o = f_o \cdot \lambda_o$ have been used. The gate function rect(x/y) is defined as 1 for $|x| < y/2$ and 0 elsewhere (also see Fig. 2.8a).

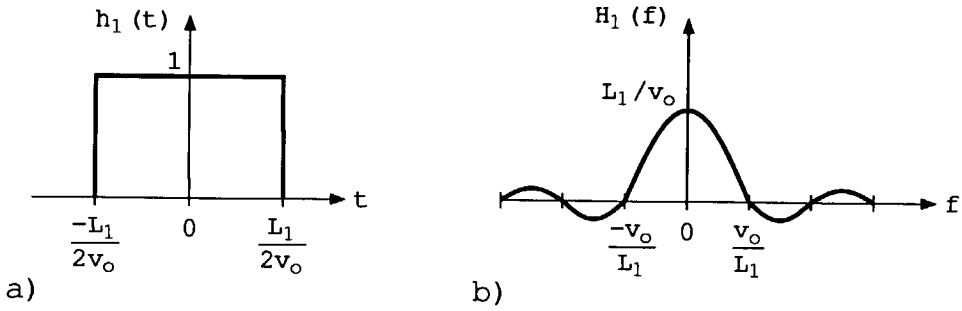

Fig. 2.8 The IDT envelope impulse response and frequency response.

From Fig. 2.8 and Equations 2.9a and 2.9b a very important conclusion can be drawn, namely the -3 dB bandwidth of the sinc response in the frequency domain is directly related to the total length of the transducer via the surface acoustic wave velocity v_o. The -3 dB bandwidth $B_{-3\,\text{dB}}$ is given by:

$$B_{-3\,\text{dB}} \approx 0.9\,\frac{v_o}{L_1} , \tag{2.10a}$$

or:

$$N_1 \approx 0.9\,\frac{f_o}{B_{-3\,\text{dB}}} . \tag{2.10b}$$

The -3 dB bandwidth is inversely proportional to the transducer length. Thus, short transducers give wide-band responses, while long transducers give narrow-band responses. The number of electrode pairs is a measure for the relative bandwidth while the length of the IDT is a measure for the absolute bandwidth.

The function $\mu_1(f)$ can be calculated by taking the convolution of the Fourier transformed sampling function $\mu_{s1}(f)$ and the frequency response $H_1(f)$. The

function $\mu_2(f)$ can be calculated in a similar way. With the two functions $\mu_1(f)$ and $\mu_2(f)$, all parameters of the SAW delay line can be calculated. They are given in the next two sections.

2.2.3 The model for the SAW delay line

We will now summarize the results obtained in the previous section and present the model with its parameters of interest. Figure 2.9 shows the model of the SAW delay line (also see reference [37]). The parameters for the model in the vicinity of the frequency f_o for the delay line with single electrodes in IDT_1 and split electrodes in IDT_2 are given in Equations 2.11a through 2.11e. The factors $\mu_1(f)$ and $\mu_2(f)$ are the transformer ratios in Fig. 2.9.

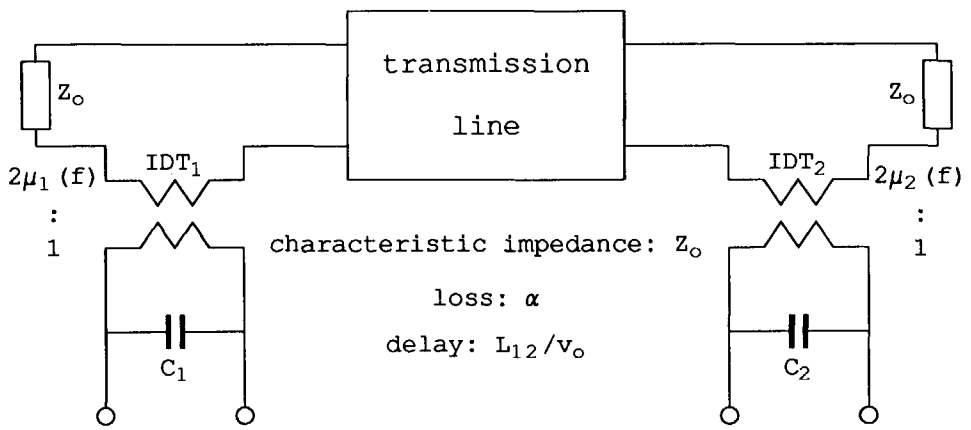

Fig. 2.9 The SAW delay line model.

$$Z_o = \frac{\kappa^2}{\omega_o \epsilon a}\,, \tag{2.11a}$$

$$\mu_1(f) = j\sqrt{\frac{2}{\pi}}\,\kappa^2 N_1 \,\mathrm{sinc}\left[\frac{f-f_o}{v_o/L_1}\right]\,, \tag{2.11b}$$

$$C_1 = \epsilon a N_1\,, \tag{2.11c}$$

$$\mu_2(f) = j\sqrt{\frac{2}{\pi}}\,\kappa^2 N_2 \,\mathrm{sinc}\left[\frac{f-f_o}{v_o/L_2}\right]\,, \tag{2.11d}$$

$$C_2 = \sqrt{2}\,\epsilon a N_2\,. \tag{2.11e}$$

From Fig. 2.9 it is clear that, in order to prevent the generation of echos, the surface acoustic wave traveling to the left of IDT_1 and the wave traveling to the right of IDT_2 should not reflect. This can be achieved by using an infinitely long crystal or applying absorbing pads.

In connection with the model in Fig. 2.9 some conclusions concerning the noise behavior of the SAW delay line can be drawn. A lossless transmission line ($\alpha = 0$) does not generate any noise. As some loss occurs in the transmission line, also some noise will be generated in the transmission line. The absorbing pads at the end of the crystal behave, in practice, as terminating resistances in order to minimize reflections in the acoustic line. Such a resistance generates a noise voltage, in series with the impedance Z_o, of which the power-density spectrum is given by:

$$S(u_z) = 4kT \cdot \frac{\kappa^2}{\omega_o \epsilon a} . \tag{2.12}$$

2.2.4 The unilateral model for the SAW delay line

According to the theory in Section 2.2.2 and supported by the measurements in Section 2.2.1, the optimum signal transfer is obtained by driving the SAW delay line by a voltage and sensing the output current of the SAW delay line. The delay line then essentially works as a frequency dependent transconductance. To simplify the noise and dynamic range calculations, as will be done in Chapter 4, a unilateral model for the SAW delay line as shown in Fig. 2.9 has been developed. This unilateral model is depicted in Fig. 2.10.

This model is only valid if the source impedance, including the effect of C_1, is much smaller than $|1/Y_1(f)|$ and if the load impedance, including the effect of C_2, is much smaller than $|1/Y_2(f)|$, which can be assured by employing negative-feedback amplifiers. If this is not the case, an extra source, due to the regeneration of surface waves, must be included in Fig. 2.10 and the input and output admittances have to be modified or the model in Fig. 2.9 has to be used.

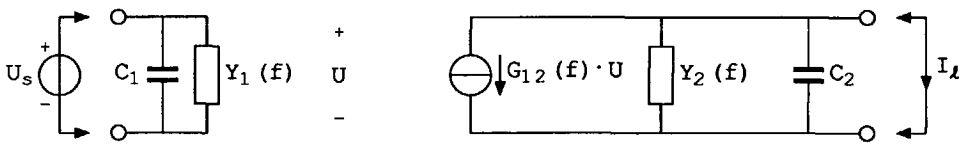

Fig. 2.10 The unilateral model for the SAW delay line.

The transconductance $G_{12}(f)$ for the response in the vicinity of f_o is given by:

$$G_{12}(f) = G_{21}(f) = G(f_o)\, e^{-\alpha L_{12}}\, e^{-j\omega L_{12}/v_o} \cdot N_1 \operatorname{sinc}\left[\frac{f-f_o}{v_o/L_1}\right] N_2 \operatorname{sinc}\left[\frac{f-f_o}{v_o/L_2}\right], \quad (2.13a)$$

$$G(f_o) = 8\kappa^2 f_o \epsilon a . \quad (2.13b)$$

The input admittance consists of a static capacitance C_1 (single electrodes) and an equivalent admittance $Y_1(f) = G_1(f) + jB_1(f)$ as a consequence of the acoustic transmission line:

$$G_1(f) = G(f_o)\, N_1^2 \operatorname{sinc}^2\left[\frac{f-f_o}{v_o/L_1}\right], \quad (2.14a)$$

$$B_1(f) = G(f_o)\, \frac{N_1^2}{\pi} \left\{\operatorname{sinc}\left[\frac{2(f-f_o)}{v_o/L_1}\right] - 1\right\} \frac{v_o/L_1}{f-f_o}, \quad (2.14b)$$

$$C_1 = \epsilon a N_1 , \quad (2.14c)$$

where $B_1(f)$ has been calculated by applying the Hilbert transform to $G_1(f)$. The output admittance consists of a static capacitance C_2 (split electrodes) and an effective admittance $Y_2(f) = G_2(f) + jB_2(f)$ as a consequence of the acoustic transmission line:

$$G_2(f) = G(f_o)\, N_2^2 \operatorname{sinc}^2\left[\frac{f-f_o}{v_o/L_2}\right], \quad (2.15a)$$

$$B_2(f) = G(f_o)\, \frac{N_2^2}{\pi} \left\{\operatorname{sinc}\left[\frac{2(f-f_o)}{v_o/L_2}\right] - 1\right\} \frac{v_o/L_2}{f-f_o}, \quad (2.15b)$$

$$C_2 = \sqrt{2}\, \epsilon a N_2 , \quad (2.15c)$$

where $B_2(f)$ has been obtained similarly to $B_1(f)$.

The noise due to both terminating impedances Z_o can either be represented by a voltage source in series with the input or a current source in parallel with the output. Their power-density spectra at the filter center frequency are given by:

$$S(u_z) = \frac{4kT}{G(f_o)N_1^2}, \quad (2.16a)$$

$$S(i_z) = 4kT\, G(f_o) N_2^2 . \quad (2.16b)$$

The functions $G_1(f)$ and $B_1(f)$ are drawn in Fig. 2.11. Notice that $B_1(f)$ behaves like a frequency dependent capacitance for $f < f_o$ and behaves like a frequency dependent inductance for $f > f_o$. It should also be noted that if we still would like to obtain a maximal power transfer, that such a maximal power transfer is very difficult to achieve over the whole filter pass-band due to the frequency dependent behavior of $G_1(f)$ and $B_1(f)$.

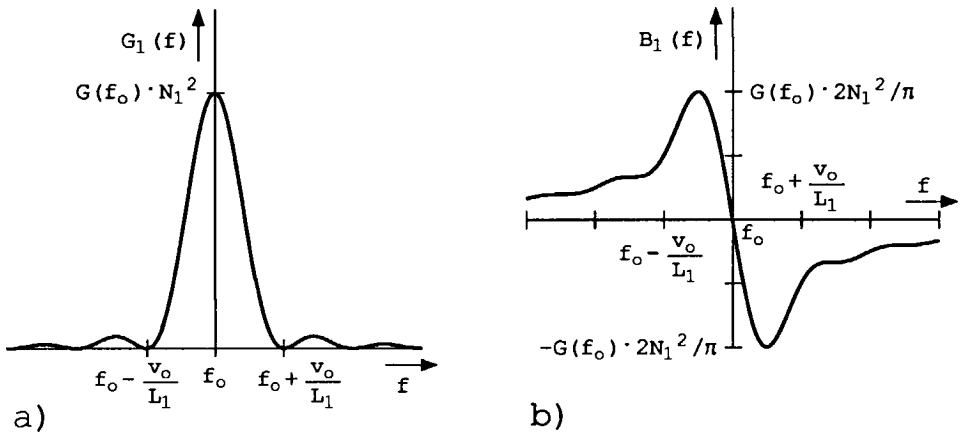

Fig. 2.11 The radiation conductance and the radiation susceptance for an unapodized IDT.

By comparing the radiation conductance, the radiation susceptance and the admittance of the static capacitance for IDTs with single electrodes within the filter pass-band ($f_o - v_o/L_1 < f < f_o + v_o/L_1$) it can easily be concluded that for

$$N_1 \ll \frac{\pi}{4\,\kappa^2}\,, \tag{2.17}$$

the admittance formed by the capacitance C_1 is larger than the radiation conductance $G_1(f)$ and the radiation susceptance $B_1(f)$. If this is the case, the admittance $Y_1(f)$ may be omitted in the calculations. For example, with $\kappa^2 = 0.0089$ (see Section 2.5) this number of electrodes is given by $N_1 \ll 88$. For the split electrodes as used in IDT_2 a similar relation results: $N_2 \ll 125$.

From Equation 2.13a, the phase shift ϕ occurring in the transfer $G_{12}(f)$ can easily be derived to be:

$$\phi = -\frac{\omega L_{12}}{v_o} + \pi \sum_i (-1)^i \, u(f - f_i) \,, \tag{2.18}$$

where u stands for the step function ($u(x) = 0$ for $x < 0$ and $u(x) = 1$ for $x > 0$) and f_i represent the frequencies at which the sinc functions have nulls. The first term in Equation 2.18 follows from the delay in the transfer function. The second term is a consequence of the sinc functions in Equation 2.13a, because the sinc functions change the phase over π rad at the frequencies f_i of the nulls.

In case the surface acoustic wave velocity v_o is independent of the frequency, which can be approached as close as possible by choosing the proper layer thicknesses, the group delay τ_g is given by:

$$\tau_g = -\frac{d\phi}{d\omega} = \frac{L_{12}}{v_o} - \pi \sum_i (-1)^i \, \delta(f - f_i) \,. \tag{2.19}$$

The group delay exhibits spikes at the frequencies f_i of the nulls of the sinc functions but is otherwise constant.

2.3 SAW transversal filter modeling

2.3.1 The model for the SAW transversal filter

The model for the SAW delay line from the previous sections will now be expanded to include the SAW transversal filter. A metalization pattern has already been shown in Fig. 2.3. The ideal low-pass filter is taken as a starting point. Figure 2.12a shows its impulse response $h_1(t)$ while Fig. 2.12b shows its frequency response $H_1(f)$, which are related to each other by means of the Fourier transform. The filter transfer function $H_1(f)$ is 1 within the filter pass-band B ($-v_o/L_1 < f < v_o/L_1$) and 0 elsewhere.

The functions $h_1(t)$ and $H_1(f)$ are given by:

$$h_1(t) = \operatorname{sinc}\left(\frac{2v_o t}{L_1}\right) \,, \tag{2.20a}$$

$$H_1(f) = \frac{L_1}{2v_o} \operatorname{rect}\left(\frac{f}{2v_o/L_1}\right) = \frac{N_1}{2f_o} \operatorname{rect}\left(\frac{f}{2v_o/L_1}\right) \,. \tag{2.20b}$$

It is clear from these equations that the bandwidth B of the filter is directly related to the length of the apodized transducer:

$$B = 2\,\frac{v_o}{L_1} \,, \tag{2.21a}$$

where L_1 represents the distance between the first nulls of the sinc function. This equation can also be written as:

$$N_1 = 2\,\frac{f_o}{B}\,. \tag{2.21b}$$

The smaller the bandwidth, the longer the impulse response and therefore the longer the length of the IDT. This phenomenon is called reciprocal spreading [10]. The number of electrode pairs N_1 in the main lobe of the sinc function in the apodized IDT is a measure for the relative bandwidth, while the length of the IDT is a measure for the absolute bandwidth.

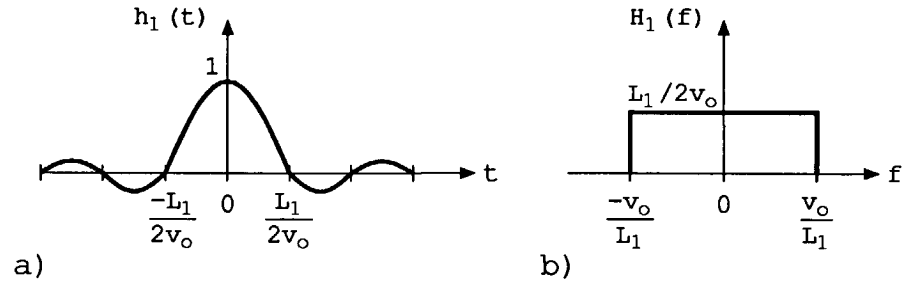

Fig. 2.12 The impulse response and frequency response of an ideal low-pass filter.

The impulse response of the ideal low-pass filter is non-causal and can therefore not be realized. By truncating the impulse response with a rectangular function rect(t/γ) with $\gamma > L_1/v_o$ and delaying the impulse response with a time L_{12}/v_o ($L_{12}/v_o > \gamma/2$), the impulse response becomes realizable. The frequency response is then given by:

$$H_1(f) = \frac{N_1}{2f_o}\,\mathrm{rect}\left(\frac{f}{2v_o/L_1}\right) * [\gamma \mathrm{sinc}(f\gamma)]\,, \tag{2.22}$$

where $*$ denotes convolution. The delay $e^{-j\omega L_{12}/v_o}$ corresponds with the center-to-center distance of the two IDTs. The function $H_1(f)$ is the Fourier transformed function of $h_1(t)$, the function that described the overlap of the electrodes. The low-pass filter is transformed into a band-pass filter by the Fourier transformed sampling function $\mu_{s1}(f)$ as given in Equation 2.8.

Instead of a rectangular window function, many other window functions are used to obtain a non-abrupt ending of the impulse response. In general

these window functions (see reference [38] for a survey of window functions) give rise to amplitude ripple in the pass-band, non-zero transfer in the out-of-band region and the roll-off slopes of the filter will become less steep. Applying computer programs like linear programming or the Remez Exchange Algorithm to find the best apodization of the IDT for the desired filter application is unavoidable.

As $H_1(f)$ has changed, the transformer ratio $\mu_1(f)$ has changed according to:

$$\mu_1(f) = j\sqrt{\frac{2}{\pi}}\,\kappa^2\,\frac{N_1}{2}\,\text{rect}\left(\frac{f-f_o}{2v_o/L_1}\right) * W(f-f_o)\,, \qquad (2.23a)$$

where $W(f)$ denotes the window function.

The capacitance of the apodized IDT can be approximated by introducing the effective electrode length a_n. When electrode n has a length a_n ($0 \leq a_n \leq a$, a denotes the maximum overlap), the capacitance of IDT_1 (single electrodes) can be approximated by:

$$C_1 = \epsilon \sum a_n\,, \qquad (2.23b)$$

where the summation has to be taken over all electrode pairs in the apodized IDT. The capacitance is $\sqrt{2}$ larger when split electrodes are used. The Equations 2.23a and 2.23b replace the Equations 2.11b and 2.11c in the model of Fig. 2.9.

2.3.2 The unilateral model for the SAW transversal filter

To simplify the noise and dynamic range calculations, as will be performed in Chapter 4, the unilateral model for the SAW transversal filter will now be presented. The validity boundaries for this model have been given in Section 2.2.4. The figure representing this model has already been depicted in Fig. 2.10, only the admittance equations for the apodized IDT and the transconductance have to be modified.

The transconductance $G_{12}(f)$ of the transversal filter is given by:

$$G_{12}(f) = G_{21}(f) = G(f_o)\,\mathrm{e}^{-\alpha L_{12}}\,\mathrm{e}^{-j\omega L_{12}/v_o}\,N_2\,\text{sinc}\left[\frac{f-f_o}{v_o/L_2}\right] \cdot \frac{N_1}{2}\,\text{rect}\left(\frac{f-f_o}{2v_o/L_1}\right) * W(f-f_o)\,. \qquad (2.24)$$

For very long IDTs, the acoustic loss influences the filter transfer function. The apodization should be modified accordingly to get the desired transfer

function. Equation 2.24 replaces Equation 2.13a in the unilateral SAW delay line model.

When the effect of the window function is small ($w(t) = 1$ or $W(f) = \delta(f)$), the admittance $Y_1(f) = G_1(f) + jB_1(f)$ and the capacitance C_1 of IDT_1 are given by:

$$G_1(f) = G(f_o)\,\frac{N_1^2}{4}\,\mathrm{rect}\left(\frac{f-f_o}{B}\right), \tag{2.25a}$$

$$B_1(f) = G(f_o)\,\frac{N_1^2}{4\pi}\,\ln\left|\frac{f-(f_o+B/2)}{f-(f_o-B/2)}\right|, \tag{2.25b}$$

$$C_1 = 0.6\,\epsilon a N_1\,, \tag{2.25c}$$

where $B_1(f)$ has been calculated by applying the Hilbert transform to $G_1(f)$ and B is defined in Equation 2.21a. The constant 0.6 used in the expression for C_1 has been estimated for an apodized IDT when only the main lobe of the sinc function is present in the apodization. Because the influence of the window function was assumed to be negligible, the radiation susceptance goes to infinity for $f = f_o \pm B/2$, the edges of the filter pass-band. The integrals from $f_o - B/2$ to $f_o + B/2$ of $B_1(f)$ or $B_1^2(f)$, however, are finite and amount to 0 and $3.3\,G^2(f_o)N_1^4B/(16\pi^2)$, respectively. Therefore, these functions can be used for noise calculations, where the integral of these functions and not the finiteness of these functions is important.

A radiation susceptance going to infinity would mean that the input impedance of the apodized IDT of the SAW transversal filter would become zero at these frequencies. In practice, due to end-effects, this behavior does not occur. Therefore, a modified radiation conductance $G_1'(f)$ and a correspondingly modified radiation susceptance $B_1'(f)$ will be introduced, which are closer to the practical situation. The essential modification is that the roll-off slope of $G_1(f)$ is not infinite but finite, see Fig. 2.13a. The corresponding $B_1'(f)$ can be calculated by applying the Hilbert transform to $G_1'(f)$. The input admittance of the SAW transversal filter is given in Equations 2.26a–2.26c. The functions $G_1(f)$, $G_1'(f)$, $B_1(f)$ and $B_1'(f)$ are drawn in Fig. 2.13. For $B_1'(f)$ the values $b_1 = 200$ kHz, $b_2 = 700$ kHz and $b_1 = 200$ kHz, $b_2 = 400$ kHz have been used.

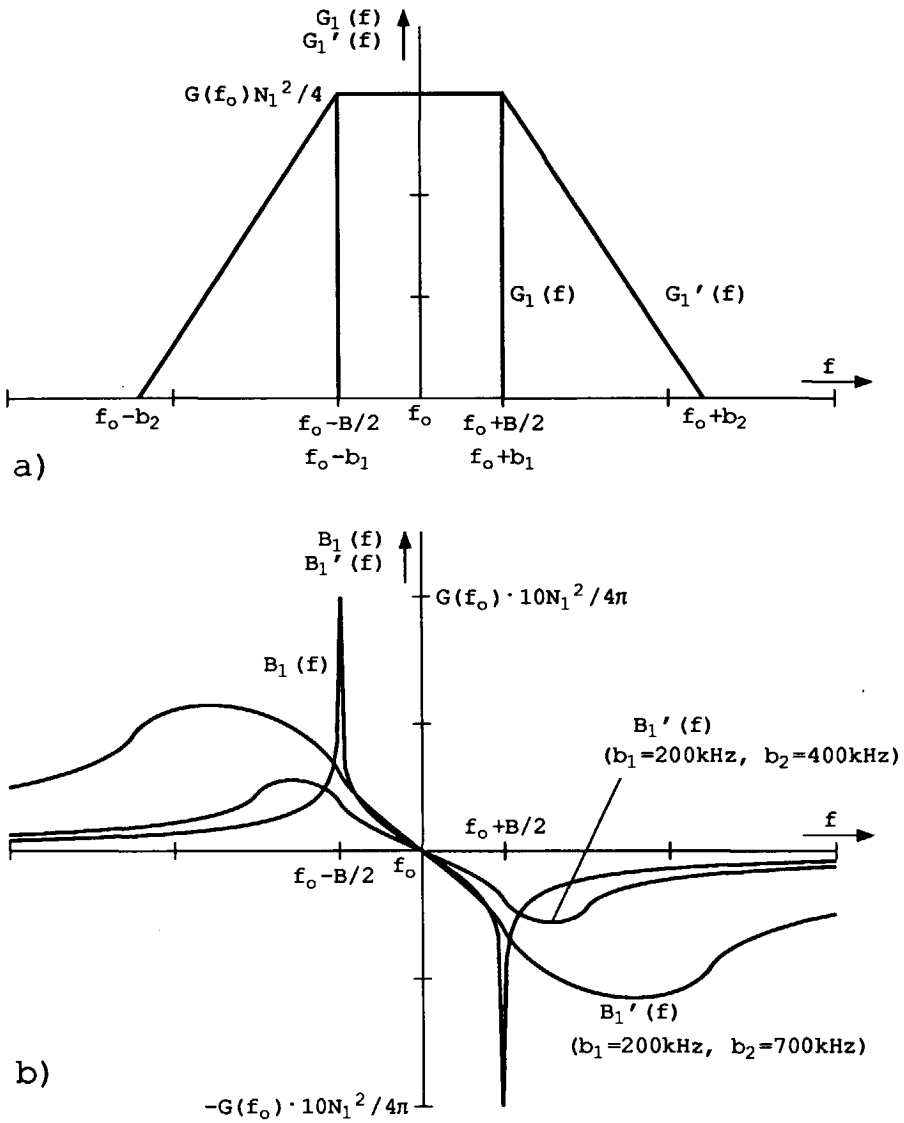

Fig. 2.13 The modified radiation conductance and radiation susceptance compared to their unmodified counterparts.

$$G_1'(f) = \begin{cases} 0 & \text{for } f \le f_o - b_2 \\ (f - (f_o - b_2))/\Delta b & \text{for } f_o - b_2 \le f \le f_o - b_1 \\ 1 & \text{for } f_o - b_1 \le f \le f_o + b_1 \\ ((f_o + b_2) - f)/\Delta b & \text{for } f_o + b_1 \le f \le f_o + b_2 \\ 0 & \text{for } f_o + b_2 \le f \end{cases} , \quad (2.26a)$$

$$\begin{aligned} B_1'(f) = G(f_o)\frac{N_1^2}{4\pi} \{ & [f - (f_o - b_1)] \ln |f - (f_o - b_1)| \\ & + [f - (f_o + b_1)] \ln |f - (f_o + b_1)| \\ & - [f - (f_o - b_2)] \ln |f - (f_o - b_2)| \\ & - [f - (f_o + b_2)] \ln |f - (f_o + b_2)| \} , \end{aligned} \quad (2.26b)$$

$$C_1 = 0.6\, \epsilon a N_1 , \quad (2.26c)$$

where $\Delta b = b_2 - b_1$. Depending on the application, Equation 2.25a, Equation 2.25b and Equation 2.25c (in the case of noise calculations) or Equations 2.26a, 2.26b and 2.26c (in the case of impedance calculations) replace Equations 2.14a, 2.14b and 2.14c in the model of Fig. 2.10.

For a window function $W(f)$ which is either completely real or completely imaginary, the phase shift ϕ occurring in the transconductance $G_{12}(f)$ within the pass-band B is given by:

$$\phi = -\frac{\omega L_{12}}{v_o} . \quad (2.27)$$

If v_o is frequency independent, which can be approached as close as possible by choosing the proper layer thicknesses, the group delay τ_g of the SAW transversal filter is given by:

$$\tau_g = -\frac{d\phi}{d\omega} = \frac{L_{12}}{v_o} . \quad (2.28)$$

The group delay is independent of the frequency. By introducing a window function that has both a real and an imaginary component, such as $e^{-bt}u(t)$ (causal exponential) or $e^{-b|t|}$ (symmetric exponential), the group delay can be designed according to a desired frequency dependency. Note that in this manner the amplitude and phase of the transversal filter can be designed independently. Even the group delay of a transversal filter with dispersion in the transmission line can thus be flattened.

2.3.3 The design procedure for SAW transversal filters

The design procedure for SAW transversal filters can be summarized as follows. From the desired bandwidth, out-of-band rejection and roll-off slopes of the filter, the distances between the nulls of the sinc impulse response in the apodized IDT can be estimated. The impulse response is directly related to the length of the IDT via the surface acoustic wave velocity. The length of the uniform IDT is chosen so short as not to influence the band-pass characteristic of the filter.

The number of electrode pairs in the main lobe of the apodized IDT can be calculated from the center frequency of the filter. The number of electrode pairs in the uniform IDT can be calculated in a similar way. The filter center frequency also determines the obtainable piezoelectric coupling constant κ^2 and the effective dielectric permittivity ϵ (see reference [5]).

The transconductance increases linearly with the aperture a. However, for a large aperture a, the input impedance and the output impedance of the transversal filter decrease, while also the occupied chip surface increases. Depending on the application a compromise must be sought. A further discussion on this aspect is given in Section 4.7 for the application of a SAW transversal filter as an IF filter in the FM upconversion receiver.

2.4 The behavior of SAW resonators

2.4.1 The model for the SAW resonator

In this section a model for the SAW resonator will be discussed. A metalization pattern for a two-port resonator has been shown in Fig. 2.4. In the case of a one-port resonator, there is only one IDT in the resonator cavity. The static capacitance of this IDT then limits the out-of-band rejection of the resonator filter. Therefore, only the two-port resonator will be discussed here.

An excellent electrical model for two-port resonators from the system and circuit designer's point of view has been derived in reference [2]. This model will be used for our application. The transfer $G_{12}(f)$ of the delay line, from one IDT to the second IDT, in absence of the reflection gratings, is given by Equation 2.13a.

The reflection grating is a distributed reflector. For frequencies close to the frequency of maximum reflection from the reflection grating, this distributed reflector can be modeled as a solid reflector with an effective reflection factor close to unity. The effective reflection factor Γ_o of the reflection grating is a function of the reflectivity r of a single strip and the total number of strips M

[2]:

$$\Gamma_o = \tanh(M \cdot |r|) \,. \tag{2.29}$$

The effective reflection factor Γ_o is usually in the order of 0.9–0.99. The solid reflector is at a distance L_p (the effective penetration depth) from the leading edge of the actual reflector [2]:

$$L_p = \frac{\lambda_o}{4 \cdot |r|} \,. \tag{2.30}$$

The acoustic waves are reflected by the reflection gratings resulting in a standing wave pattern in the cavity between the reflection gratings. The delay line transconductance $G_{12}(f)$ is enlarged by a factor comprising the effective reflection factor ($|\Gamma_o| \neq 1$) [2]:

$$G_s(f_o) = G_{12}(f_o) \, \frac{1 + 2\Gamma_o \cos(4\pi\Delta/\lambda_o) + \Gamma_o^2}{1 - \Gamma_o^2} \,, \tag{2.31}$$

where Δ represents the distance between the centers of the IDTs and the maxima of the standing waves within the cavity. The IDTs should be placed an integer number of half-wavelengths apart.

The distance L_c between the two distributed reflectors has to be chosen according to $L_c = \frac{1}{2}(n-1)\lambda_o$, where $n = 1, 2, 3 \ldots$. Usually an n of about 100 is chosen to avoid too much influence of acoustic wave losses and transverse modes in the cavity [2].

The quality Q of the resonator is given by [2]:

$$Q = \frac{2\pi}{\lambda_o} \, (L_c + 2L_p) \, \frac{\Gamma_o^2}{1 - \Gamma_o^2} \,, \tag{2.32}$$

where it has been assumed that the resonant frequency of the cavity is the same as the frequency of maximum reflection for the reflection grating.

Within the filter pass-band, the SAW resonator filter can now be described as an RLC network [39]. Figure 2.14 shows this model. It has implicitly been assumed that the resonator has to be driven by a voltage and that the short-circuit current should be sensed. The capacitances C_1 and C_2 and the admittances $Y_1(f)$ and $Y_2(f)$ are the delay line capacitances and admittances given in Section 2.2. The model in Fig. 2.14 is valid as long as Γ_o remains close to unity. Dependent on the number of reflectors M in the reflection grating, the reflection factor Γ_o remains close to unity for $0.995 < f/f_o < 1.005$ (M=250 [27]). For much larger deviations from f_o the transfer will approach the transfer of the delay line. The model in Fig. 2.14 is sufficiently accurate to obtain the transfer to about 10–20 dB down from the filter center frequency.

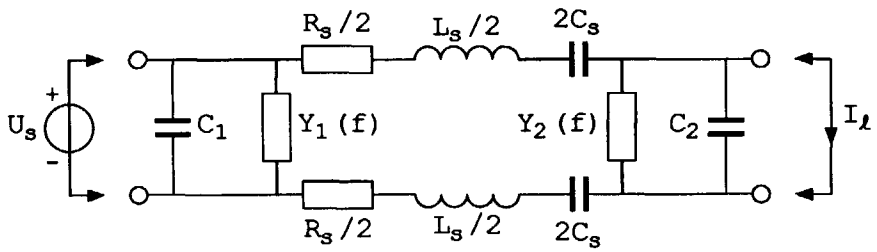

Fig. 2.14 The two-port resonator model.

The resonator series inductance L_s, series capacitance C_s and series resistance R_s are given by:

$$L_s = \frac{L_c + 2L_p}{G_{12}(f_o)\, v_o} \cdot \frac{\Gamma_o^2}{1 + 2\Gamma_o \cos(4\pi\Delta/\lambda_o) + \Gamma_o^2}\,, \tag{2.33a}$$

$$C_s = \frac{1}{\omega_o^2 L_s}\,, \tag{2.33b}$$

$$R_s = \frac{1}{G_s(f_o)}\,. \tag{2.33c}$$

The series resistance of the resonator produces noise with a power-density spectrum given by:

$$S(u_s) = 4kTR_s\,. \tag{2.34}$$

2.4.2 The group delay variation for a single resonator

The resonator should be driven by a voltage and the short-circuit current should be sensed to get the desired filter function. The transconductance $G(f)$ from input voltage U_s to output current I_l for frequencies near the filter center frequency, can be written as:

$$G(f) = \frac{I_l}{U_s} = \frac{1}{R_s} \cdot \frac{1}{1 + \mathrm{j}\nu Q}\,, \tag{2.35}$$

where the following abbreviations have been used:

$$\nu = \frac{f}{f_o} - \frac{f_o}{f}\,, \tag{2.36a}$$

$$f_o = \frac{1}{2\pi\sqrt{L_s C_s}}\,, \tag{2.36b}$$

$$Q = \frac{f_o}{B} = \frac{\omega_o L_s}{R_s} = \frac{1}{\omega_o R_s C_s}\,. \tag{2.36c}$$

The parameter ν denotes the detuning, f_o denotes the filter center frequency and B denotes the -3 dB bandwidth of the resonator. The amplitude transfer $|G(f)|$ shows a 6 dB/oct slope as a function of the detuning ν for frequencies near the filter center frequency.

The phase shift ϕ occurring in the resonator transfer $G(f)$ is given by:

$$\phi = -\arctan(\nu Q) \tag{2.37}$$

For frequencies near the resonance frequency f_o ($\nu \ll 1$) the group delay τ_g is given by:

$$\tau_g = \frac{1}{1+\nu^2 Q^2} \cdot \frac{2}{2\pi B} . \tag{2.38}$$

The peak-to-peak variation in group delay within the -3 dB bandwidth B is given by:

$$\Delta\tau_g = \frac{1}{2\pi B} \tag{2.39}$$

For a -3 dB bandwidth of 180 kHz, the variation in group delay is approximately 880 ns peak-to-peak. The implication of this result with respect to the FM receiver design will be dealt with in Chapter 4.

2.4.3 The electrically coupled resonator

Two SAW resonators can be coupled to each other by either an acoustic or an electric coupling. Only the, simpler, electrically coupled resonator will be discussed. The coupling of the two resonators, both with the same Q, is performed by a capacitor C_k, see Fig. 2.15. Because the number of electrode pairs

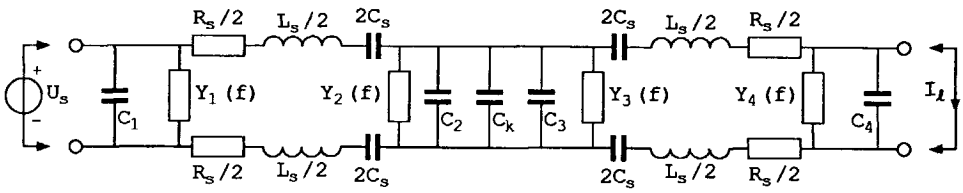

Fig. 2.15 The electrically coupled resonator.

in the cavity is limited, the output admittance $Y_2(f)$ of the first resonator and the input admittance $Y_3(f)$ of the second resonator are small compared with the admittance formed by the capacitances C_2, C_3 and C_k (see Equation 2.17).

The transconductance $G(f)$ from source voltage U_s to load current I_l can be written as:

$$G(f) = \frac{I_l}{U_s} = \frac{f_o}{\mathrm{j}fR_s} \cdot \frac{kQ}{(1+\mathrm{j}\nu Q)^2 + (kQ)^2(f/f_o)^2} , \tag{2.40}$$

where the following abbreviations have been used:

$$k = \frac{C_s}{C_s + C_2 + C_3 + C_k}, \tag{2.41a}$$

$$kQ = \frac{1}{\omega_o R_s (C_2 + C_3 + C_k)}, \tag{2.41b}$$

$$f_o = \frac{1}{2\pi} \cdot \sqrt{\frac{1}{L_s} \cdot \frac{C_s + C_2 + C_3 + C_k}{C_s(C_2 + C_3 + C_k)}}. \tag{2.41c}$$

The parameter k denotes the resonator coupling. The amplitude transfer $|G(f)|$ shows a 12 dB/oct slope as a function of the detuning ν for frequencies near the filter center frequency.

The -3 dB bandwidth B_k of the coupled resonator as a function of the coupling factor kQ ($kQ \leq 1$) is given by:

$$B_k = B_e \cdot \sqrt{k^2Q^2 - 1 + \sqrt{2(1 + k^4Q^4)}}, \tag{2.42}$$

where B_e denotes the bandwidth of a single resonator and $\nu \ll 1$ has been used.

The optimum coupling factor kQ for a minimal variation in the group delay within the -3 dB bandwidth of the coupled resonator will now be determined. For frequencies $f = f_o + \Delta f$ ($\Delta f \ll f_o$) near the resonance frequency f_o we may use $\nu = 2\Delta f / f_o$ as an approximation for the detuning. The phase shift ϕ occurring in the transfer function $G(f)$ can then be written as:

$$\phi = -\arctan\left[\frac{B_e^2(1 + k^2Q^2) - 4\Delta f^2}{4\Delta f B_e}\right]. \tag{2.43}$$

The group delay τ_g is given by:

$$\tau_g = \frac{4}{2\pi B_e} \cdot \frac{1 + k^2Q^2 + (2\Delta f/B_e)^2}{(1 + k^2Q^2)^2 + 2(1 - k^2Q^2)(2\Delta f/B_e)^2 + (2\Delta f/B_e)^4}. \tag{2.44}$$

In Fig. 2.16 the normalized group delay is shown as a function of the normalized frequency deviation from the filter center frequency Δf with kQ as parameter. The dots on the traces denote the -3 dB frequency of the coupled resonator. The group delay is maximally flat if all derivatives of τ_g with respect to Δf are zero at $f = f_o$. The first derivative is inherently zero. The second derivative is zero at $f = f_o$ if

$$kQ = \frac{1}{3}\sqrt{3}. \tag{2.45}$$

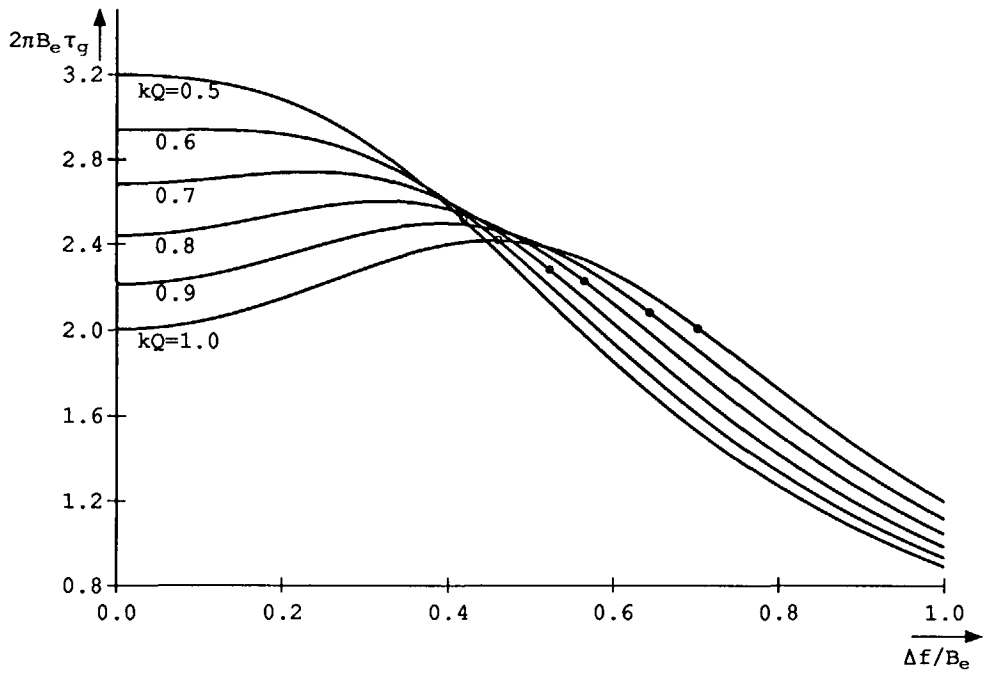

Fig. 2.16 The normalized group delay versus normalized frequency.

With this optimum value of kQ the variation in the group delay within the -3 dB bandwidth B_k ($B_k/B_e \approx 0.908$) amounts to:

$$\Delta\tau_g \approx \frac{0.520}{2\pi B_k} . \tag{2.46}$$

For a -3 dB bandwidth of 180 kHz, the variation in group delay is approximately 460 ns. The peak-to-peak variation in group delay for a coupled resonator is twice as small as that of a single resonator. From Fig. 2.16 it can be concluded that the peak-to-peak ripple in group delay can even be made smaller than for the case of the maximally flat group delay. If $kQ = 0.75$ is chosen, the peak-to-peak group delay is about 350 ns peak-to-peak for a coupled resonator with a 180 kHz bandwidth.

In the coupled resonator in Fig. 2.15, a capacitance C_k was added to the resonator to obtain the desired coupling factor kQ. To avoid the fabrication of this extra capacitance, the number of electrode pairs in the IDTs in the SAW resonator can be chosen in such a way that the desired coupling factor

kQ is achieved by the capacitances C_2 and C_3 only. Assuming the IDTs to be placed at a maximum of the standing waves in the cavity ($\Delta = 0$), using $N_1 = N_4$ and $N_2 = N_3$, Equation 2.41b can be rewritten to the number of electrode pairs in IDT_1 and IDT_4:

$$N_1 = N_4 = \frac{1-\Gamma_o}{1+\Gamma_o} \cdot \frac{\pi}{2} \cdot \frac{kQ}{\kappa^2} , \tag{2.47}$$

where single electrodes have been used for all IDTs. In case split electrodes are used in the second and third IDT, the number of electrode pairs in the first and fourth IDT should be $\sqrt{2}$ larger. Notice that the number of electrode pairs in IDT_1 and IDT_4 are independent of the aperture, the effective dielectric permittivity and the number of electrode pairs in the second and third IDT. The number of electrode pairs in the second and third IDT should be chosen as high as possible to minimize R_s or to maximize the transconductance at resonance, i.e. the cavity should be filled as much as possible.

2.5 Measurements on on-chip SAW delay lines in combination with electronic circuitry

2.5.1 Introduction

In this section the measurements performed on a test chip with on-chip SAW delay lines will be presented. A photograph of the chip (ELIS552), where several SAW delay lines, a bipolar double-balanced mixer and a transimpedance amplifier have been combined, is shown in Fig. 2.17. The chip has been mounted on a 68 pin SLAM. Absorbing material has been applied to the sides of the chip to prevent reflection of the surface waves at these sides. A number of delay lines is present, consisting of four small IDTs ($L = 0.42$ mm, $N = 15$, $\lambda_o = 28$ μm) and two large IDTs ($L = 2.8$ mm, $N = 100$, $\lambda_o = 28$ μm) as can be seen in the upper and lower part of the chip. Further, some aluminum plates have been applied as electrostatic shield for test purposes. The aperture of the electrodes is 692 μm. Single electrodes were used for all IDTs to avoid technological problems due to a limited etch resolution at the time of fabrication. The bipolar switching mixer and transimpedance amplifier can be seen in the acoustic path of the delay lines at the left-hand side of the chip and in the middle of the chip. The remaining part of the chip is used for several test devices and the process control modules. In the next sections, the measurement results of this chip will be presented.

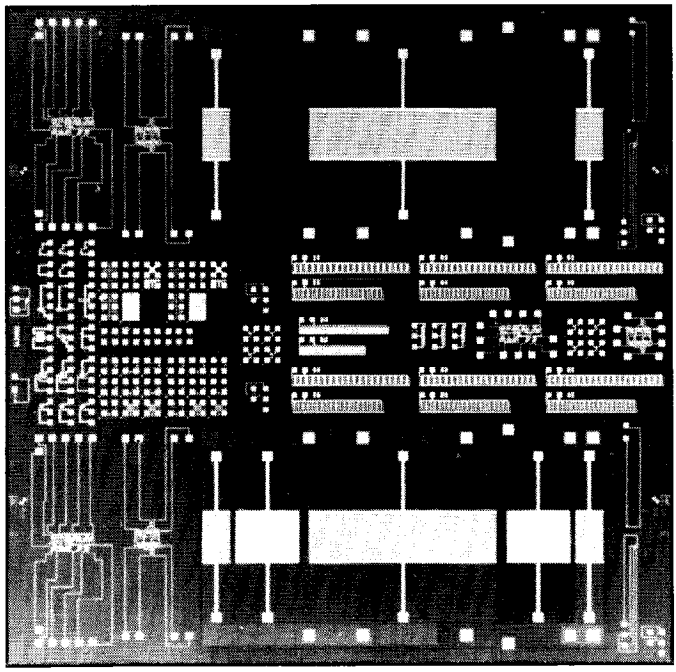

Fig. 2.17 The ELIS552 chip with on-chip SAW delay lines and electronic circuitry.

2.5.2 Capacitance and transfer measurements

Measurements on the IDT capacitance on 16 IDTs on different chips yields an effective dielectric permittivity of $\epsilon = 135 \pm 40$ pF/m. In a production-line process, the variation in ϵ should be much smaller.

In Fig. 2.18 the transconductance from a short IDT to a long IDT is shown as a function of frequency. As theory predicted (Equation 2.13a) the transconductance is the product of two sinc functions. The nulls of the sinc functions should be 2.02 MHz and 13.46 MHz apart. Due to dispersion, these nulls are somewhat closer together (15%). It can be concluded that the ZnO and SiO_2 layer thicknesses are not yet optimal.

The transfer from about 100 to 102 MHz is severely distorted by reflections of the surface waves against the electrodes. Split electrodes would have been used for the long IDT if technology had permitted this. Results on the improvement of this technology aspect are given in reference [40]. From the filter

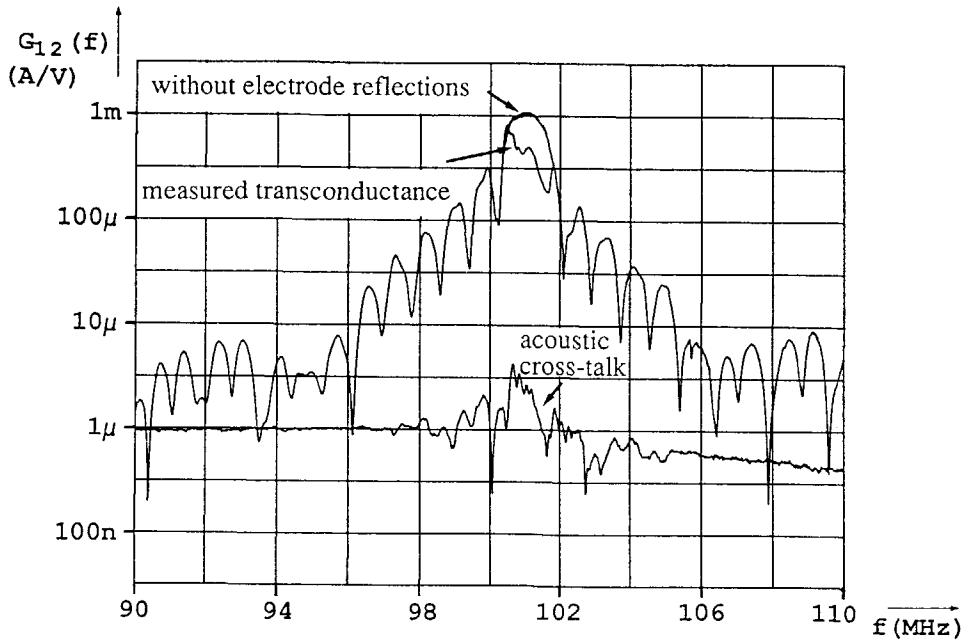

Fig. 2.18 The transconductance as a function of the frequency.

center frequency of 100.95 MHz, which has been calculated by averaging the frequencies of the nulls of the sinc function, the velocity of the surface acoustic waves can be calculated to be $v_o = 2827$ m/s. The piezoelectric coupling factor can be calculated from the transconductance of about 1 mA/V at the filter center frequency: $\kappa^2 = 0.0089$ ($\alpha = 0$).

Although there is an optimum direction in which the surface waves travel [5], these waves show diffraction. In Fig. 2.18 the transconductance from the short IDT in the lower part of the chip to the long IDT in the upper part of the chip is shown. In the middle of the chip some absorbing material has been applied. The resulting acoustic cross-talk in the vicinity of f_o is only 1000 times smaller than the transconductance from a long IDT to a short IDT which are placed in the surface acoustic wave propagation paths of each other. When no absorbing material is applied, the acoustic cross-talk is even 100 times larger. In case two SAW devices on one chip have the same center frequency and are used in different applications, for instance one as an IF filter and the second as a SAW delay line in an oscillator, some problems will arise due to this acoustic cross-talk. A large signal in the IF filter may pull the oscillator to another

frequency or the oscillator signal may be injected in the IF filter and so trouble the detection of the desired signal. This aspect complicates system design. In Chapter 4 this problem will be a point of discussion.

In Fig. 2.19 the transconductance from a short IDT to a long IDT is shown from DC to 330 MHz, measured in two different situations. The upper trace is the transconductance when a single-sided source and load are used. The lower trace represents the transconductance when 1:1 transformers are used to obtain a balanced source and load. The resulting feedthrough in terms of

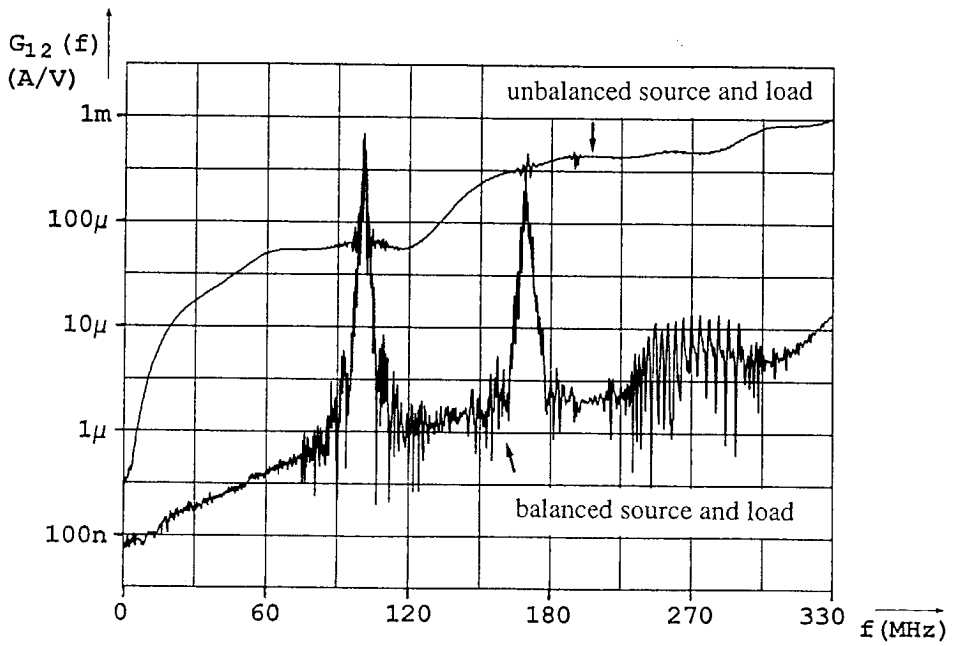

Fig. 2.19 The transconductance in the case of unbalanced and balanced source and load.

transconductance behaves more or less like it was caused by a capacitance, thus increasing proportionally with the frequency. The equivalent capacitance resulting from a feedthrough of 2 μA/V at 100 MHz is about 3 fF. It goes without saying that these measurement results can only be reached if the proper housing and bonding of the chip are chosen. Moreover, it can be concluded that a balanced source and load are essential in obtaining a low electromagnetic feedthrough [28]. The electromagnetic feedthrough can even be decreased by increasing the separation of the IDTs, however, at the cost of

a larger occupied chip area [29].

In the vicinity of 270 MHz an increase in output signal can be seen. This is caused by bulk waves i.e. waves that are generated at the sending IDT and which travel to the bottom of the substrate and reflect there. They are picked up by the receiving IDT of the delay line. Most of these responses can be eliminated by roughening the bottom of the substrate.

The second type of spurious response is a consequence of the three-layered ZnO-SiO_2-Si structure. This effect is not present in quartz or lithium-niobate SAW devices. In the ZnO-SiO_2-Si structure, an IDT generating a surface wave with velocity v_o also generates surface waves with higher velocities [41]. All waves have the same wavelength. For the used SiO_2 and ZnO layer thicknesses, this second-order Rayleigh-wave mode, due to a velocity of $1.7v_o$, causes the extra response at 170 MHz. These modes are inherent in the considered choice of materials. The piezoelectric coupling factor for these higher-order modes can hardly be diminished [5, Fig. 2.5]. In Chapter 4, criteria for the choice of the intermediate frequency of the FM upconversion receiver will be given; the influence of the response due to the mode with a velocity $1.7v_o$ is then minimal.

2.5.3 Distortion measurements

A SAW device is a passive device. The surface acoustic wave is a linear elastic change of the material at the surface only if the elasticity constants are linear. For large amplitudes of the surface waves, distortion occurs. Measurements on the delay line have been performed to obtain the distortion characteristics. In Fig. 2.20 the measurement result is shown. The signals applied to the large IDT of the SAW delay line have an amplitude of 1.0 V_{eff} each. The frequencies of the applied signals are 100.9 MHz and 101.1 MHz. At 100.7 MHz and 101.3 MHz, two third-order intermodulation components can be seen about 60 dB lower, as caused by the SAW delay line. Second-order intermodulation distortion components cannot be measured because the SAW device has a band-pass frequency transfer with a relatively narrow bandwidth.

The 1.0 V_{eff} applied to the large IDT gives rise to a calculated (Eq. 2.11b) surface wave corresponding to 0.71 V_{eff} where $\kappa^2 = 0.0089$ has been used. For the prediction of the distortion in future designs at 100 MHz, this amplitude of 0.71 V_{eff} must be used because the distortion is generated in the acoustic path and not in the conversion from voltage to surface wave. Different IDTs give therefore different levels of distortion when the same voltages are applied to the IDTs. This can be seen in the lower trace, which uses the same measurements

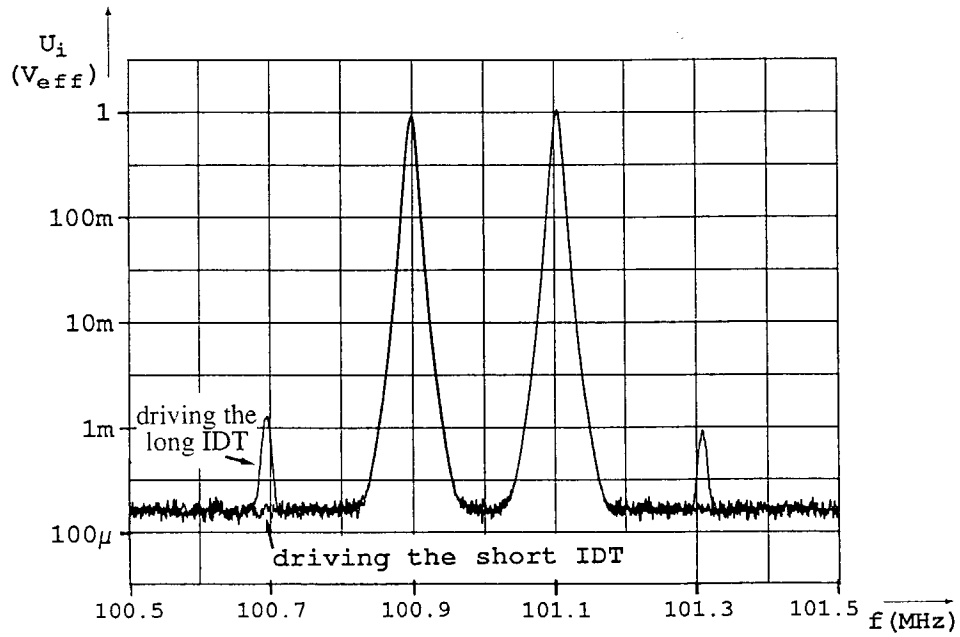

Fig. 2.20 The distortion in the SAW delay line.

setup, only the delay line has been reversed, the two signals are applied to the small IDT. No visible distortion is present in this situation because the acoustic power is much smaller.

The SAW device was capable of enduring 5 V_{eff} (one signal) for a short period of time without destruction of the device. The intermodulation distortion limits the use of SAW devices for large signals. The noise generated by the SAW device limits the use for small signals. The noise behavior of the SAW delay line has been discussed in Section 2.2. Using the power-density spectra given there leads to an effective noise voltage at the input of the delay line of 0.67 μV_{eff} in a 180 kHz bandwidth, in the case the long IDT is voltage-driven and the short IDT is current-sensed.

From the noise voltage, the maximum signal level and the intermodulation measurements, the dynamic range and the intermodulation-free dynamic range can be estimated for this SAW device. The dynamic range is at least 135 dB where the noise is measured in a 180 kHz bandwidth. The third-order intermodulation-free dynamic range amounts to at least 102 dB (B = 180 kHz).

2.5.4 Temperature behavior and matching

Because we have two large IDTs (one in the upper part and one in the lower part of the chip), which determine the frequency response, the matching of devices and the temperature dependence of the devices can be measured.

The difference in filter center frequency of the delay line in the upper part and in the lower part of the chip is in the range of -130 to -170 kHz for different chips. This difference is probably due to a difference in thickness of the SiO_2 or ZnO layers. With an industrial process, a better matching should be obtainable.

The temperature dependence of the filter center frequency can be obtained accurately by measuring the temperature dependence of the frequencies of the nulls in the frequency response. The frequency shift of the nulls depends linearly on the temperature, ranging from -15 to -18 ppm/K for different devices in the -40 °C to $+80$ °C temperature range. The difference in tracking of the center frequency of the two devices on one chip is smaller than 20 kHz.

2.5.5 Influence on the electronic circuitry

As three transimpedance amplifiers have been integrated on the test chip ELIS552 (see Fig. 2.17), two covered by a 10 μm thick zinc oxide layer and one not covered by zinc oxide, the influence of the zinc oxide layer on the amplifier performance can be measured.

Figure 2.21 shows the circuit diagram of the balanced transimpedance amplifier. The two feedback resistors R_{t1} and R_{t2} determine the current-to-voltage transfer of the transimpedance amplifier, which amounts to 0.94 V/mA. The active part of the amplifier consists of a differential pair (Q_1 and Q_2) followed by two emitter followers (Q_3 and Q_4). The remaining transistors and resistors determine the biasing of the amplifier. In the application in the front-end of the upconversion receiver, this transimpedance amplifier is directly connected to the output of the double-balanced bipolar switching mixer and hence the resistors R_{m1} and R_{m2} are then omitted. Note that in this way, the DC output current of the mixer is used as the bias current for the output stages of the transimpedance amplifier, thereby minimizing the power consumption in the receiver front-end. A 10 pF capacitor is used as a load for this amplifier. Two capacitors in parallel with R_{t1} and R_{t2} determine the amplifier stability.

Measurements performed on the low frequency accuracy in the amplifier transfer, the noise behavior of the amplifier and the amplifier bandwidth show differences smaller than 5% for the amplifiers covered by zinc oxide compared to those not covered by zinc oxide. The amplifiers covered by zinc oxide, however, show a nominally higher inaccuracy while also the noise produced

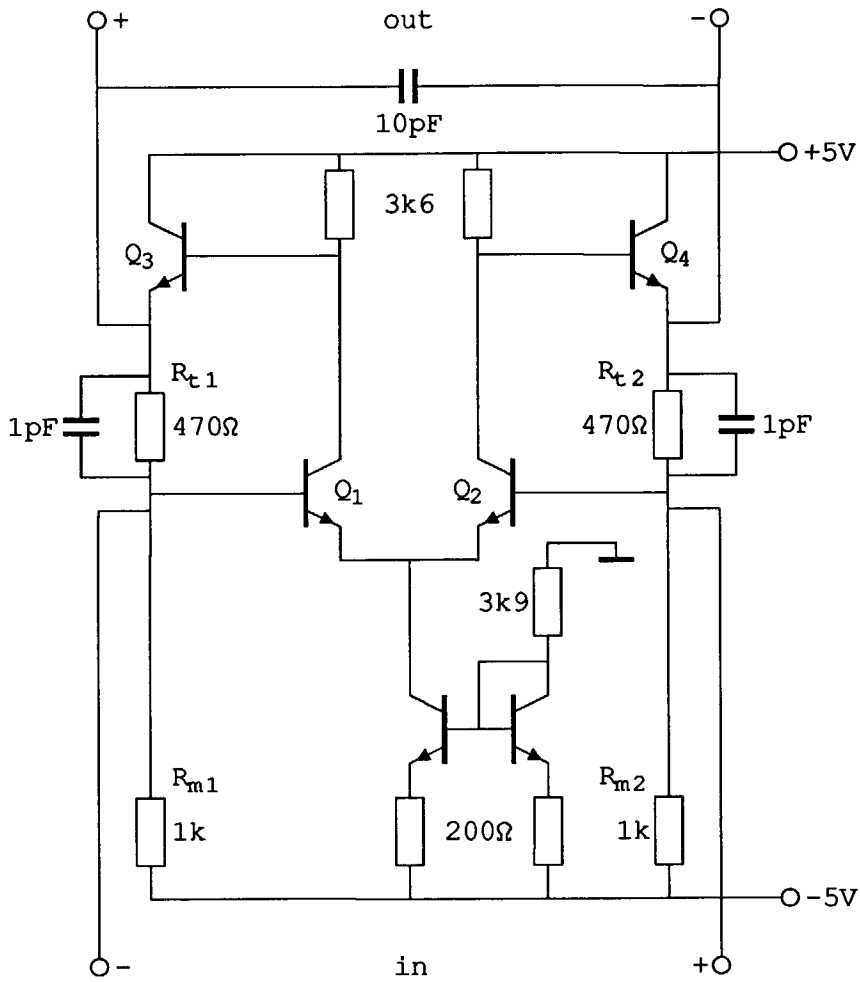

Fig. 2.21 The transimpedance amplifier.

by the amplifier is nominally somewhat larger. Further investigations are necessary to determine whether this difference is caused by the sputtering of the zinc oxide or if it is caused by process tolerances.

A problem that might arise when the zinc oxide is sputtered on electronic circuits is the higher capacitance from one part of the circuit to another part of the circuit compared to the situation where no zinc oxide is present. Usually,

the feedthrough of signals in electronic circuits is low as a consequence of the planar structure. However, a zinc oxide layer, which is several micrometers thick, gives rise to extra stray capacitances as the relative dielectric permittivity of zinc oxide is approximately 8. Especially in the case of circuits with a large amount of gain, such as limiters in FM detectors, these stray capacitances can lead to high-frequency instabilities. Therefore, it is recommended that the piezoelectric zinc oxide is only sputtered on those parts of the chip where the piezoelectricity is essential, i.e. on the SAW device itself.

2.6 Discussion

In this chapter we dealt with SAW devices. These devices can be made on the same chip as the (high-frequency) electronic circuitry by processing three extra layers; a thick SiO_2 layer (2 μm), a thin aluminum layer (0.3 μm) and a piezoelectric ZnO layer. By means of applying the appropriate aluminum pattern, SAW delay lines, SAW transversal filters, SAW resonators and SAW resonator filters can be realized.

The optimum signal transfer in SAW delay lines and SAW transversal filters is obtained by driving the delay line by a voltage and sensing the short-circuit current, resulting in a minimal triple transit echo. This conclusion is supported by measurements performed on a SAW transversal filter on a lithium niobate substrate.

The transconductance, the voltage-to-current transfer of a SAW delay line, has the property of a band-pass filter. The center frequency is defined by the electrode-to-electrode distance of the IDTs while the (absolute) bandwidth is determined by the length of the IDT or the number of electrode pairs determines the relative bandwidth.

The modeling of the SAW delay line and the SAW transversal filters in the ZnO-SiO_2-Si, three layered structure has been done in such a way as to show the system and circuit designer the available degrees of freedom in the design of these filters. The SAW delay line model can be modified into the model for a SAW transversal filter by apodization of the electrode lengths in one of the IDTs. The frequency response can, in this way, be changed to acquire the desired impulse response and thus the desired filter transfer function.

Two models for a SAW delay line and a SAW transversal filter have been presented. The most extensive model consists of a transmission line with frequency dependent transformers, the IDTs. In the case where the delay line or transversal filter is driven by a voltage and the short-circuit current is sensed (the optimum signal transfer), a simpler unilateral model may be used. Noise and dynamic range calculations with the use of this model will be performed

in Chapter 4.

A design procedure for transversal filters has been presented. For a certain specified frequency response, only the aperture of the IDTs can be enlarged to increase the transconductance at the filter center frequency. By enlarging the aperture, however, the input and output impedance of the SAW delay line or SAW transversal filter decrease thereby increasing the power consumption for driving the SAW device. A compromise for the aperture must therefore be sought.

When the proper apodization is applied in the transversal filter, a nearly perfect filter transfer function can be obtained: an amplitude transfer function of 1 within the filter pass-band and 0 elsewhere. The group delay is constant within the filter pass-band and even constant beyond the −3 dB frequencies. A resonator filter, however, exhibits a very different transfer function. The roll-off slopes for a single resonator is 6 dB/oct and 12 dB/oct for a coupled resonator. The variation in group delay in the transfer function of a single resonator is solely determined by the bandwidth. In a coupled resonator the resonator coupling factor can be chosen in such a way that a minimum variation in group delay is obtained. By choosing the proper number of electrode pairs in the IDTs at the input and at the output of an electrically coupled resonator, the desired resonator coupling factor can be obtained without any extra capacitances.

The feasibility of a monolithic integration of SAW filters with high-frequency electronic circuitry has been demonstrated by the test chip ELIS552. The results of the measurements support this demonstration of the feasibility. The electromagnetic feedthrough between input and output IDT can be made sufficiently low by increasing the distance between the two IDTs, however, at the cost of increased occupied chip area.

In the used ZnO-SiO_2-Si structure, higher-order Rayleigh-wave modes are present, resulting in spurious pass-bands. Criteria for the choice of the intermediate frequency of the FM upconversion receiver in order to minimize the disturbance of the desired signal will be given in Chapter 4. The intermodulation-free dynamic range of the on-chip SAW devices is so high that they hardly are a limiting factor in the performance of any system employing these devices. Further, it is recommended that the zinc-oxide layer is only sputtered onto those places on the chip where the piezoelectricity is essential, thus minimizing stray capacitances.

3. The dipole and monopole antenna

3.1 Introduction

The image channels for a conventional downconversion receiver, with 10.7 MHz as the intermediate frequency, range from 108.9 to 129.4 MHz ($f_{LO} > f_{RF} > f_{IF}$). The impedance and the transfer of the antenna at these image channels has not changed much compared with the broadcast channels (87.5–108 MHz). However, the image channels in an upconversion receiver are higher than 330 MHz ($f_{LO} > f_{IF} > f_{RF}$). The antenna impedance and the antenna transfer at these image frequencies must, therefore, be determined in order to calculate the image suppression. Should it be the case that the antenna transfer increases significantly at these image channels, then it may prove advantageous to choose a not too high intermediate frequency.

In this chapter the (75 cm long) whip antenna, as frequently used in portable radios and car-radios, and the closely related dipole antenna will be discussed. Due to a variable antenna environment in portable receivers, a rough estimation of the antenna impedance and transfer at the broadcast channels and at the image channels will be sufficient.

In Section 3.2 the antenna impedance will be treated. In Section 3.3 the transfer of the antenna will be discussed, both as a function of the frequency and as a function of the spatial coordinates (the radiation pattern). Section 3.4 concludes this chapter with a discussion.

3.2 The antenna impedance

In this section the impedance of the dipole antenna (see Fig. 3.1a) and the monopole or whip antenna (see Fig. 3.1b) will be discussed. The impedance of the dipole antenna in free space will be discussed first. The impedance of the whip antenna will be derived from the impedance of the dipole antenna. Due to a ground plane of limited dimensions, the impedance of the practical whip antenna is different from the impedance that is derived from the dipole antenna impedance. Measurements concerning the impedance of a whip antenna with a ground plane of limited size will be discussed.

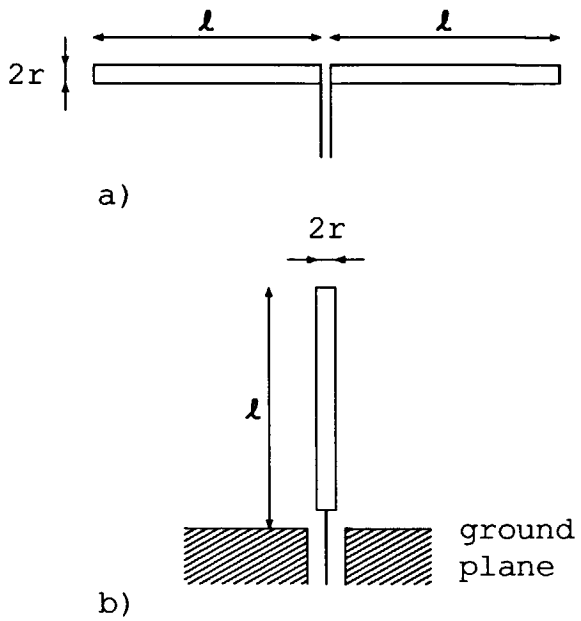

Fig. 3.1 The dipole and monopole antenna.

First the impedance of a dipole antenna in free space is given. The antenna impedance can be calculated by analyzing the current distribution along the antenna (see, for instance, reference [42]). The antenna impedance $Z_a = R+\mathrm{j}X$ of the dipole antenna is given in Fig. 3.2 [43]. The real part R is along the horizontal axis while the imaginary part X is along the vertical axis. Two spirals are shown. The inner spiral is valid for an antenna length-radius ratio of $l/r = 100$, while the outer spiral is valid for an antenna length-radius ratio of $l/r = 500$. The frequency, at which the antenna impedance is presented, is given along the spirals in multiples of $0.1 \cdot c/l$ (o for $l/r = 100$ and □ for $l/r = 500$), where c represents the velocity of light ($c = 3 \cdot 10^8$ m/s). This figure shows several frequencies at which the antenna impedance behaves like a resonant circuit; i.e. $X = 0$. It is easily seen that the antenna is in series resonance at $f \approx 0.25 \cdot c/l$, $f \approx 0.7 \cdot c/l$, etc, virtually independent of the value of l/r. The antenna is in parallel resonance at $f \approx 0.45 \cdot c/l$, $f \approx 0.95 \cdot c/l$, etc, virtually independent of the value of l/r. For low frequencies ($f \ll 0.25 \cdot c/l$) the antenna impedance behaves like a capacitance.

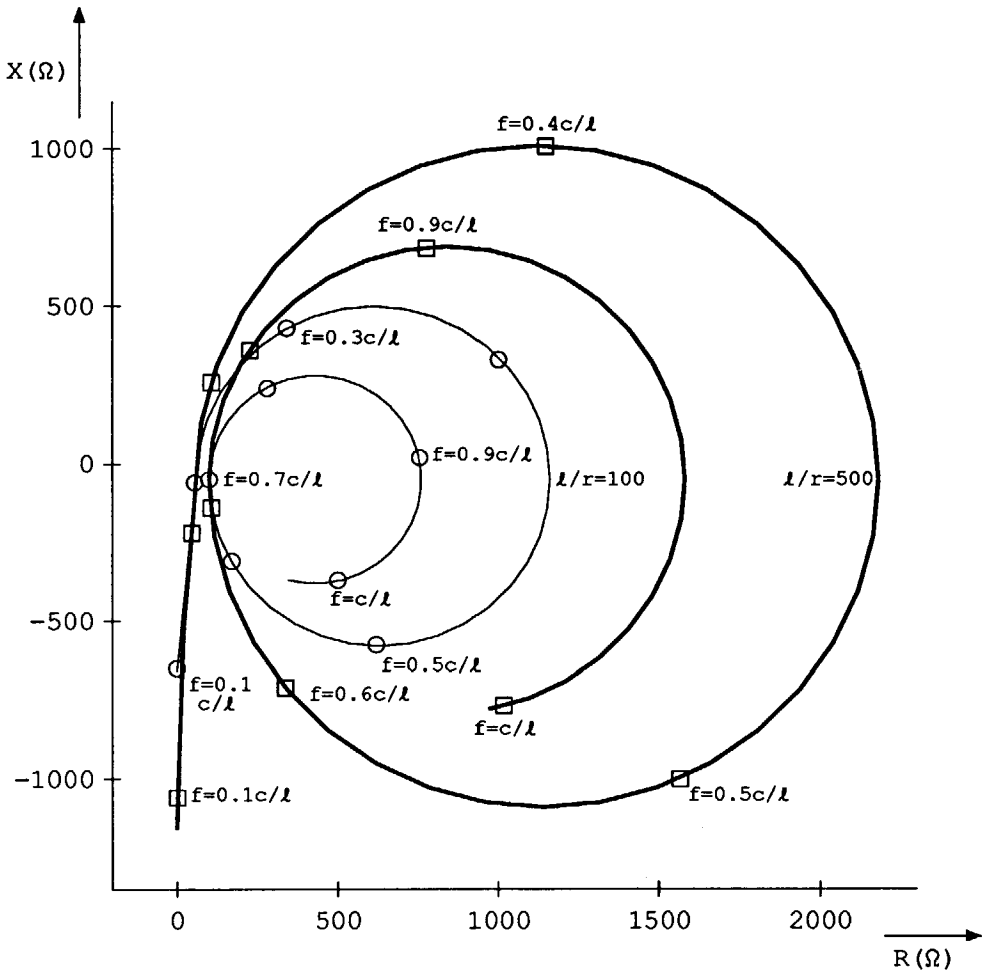

Fig. 3.2 The real and imaginary part of Z_a for a dipole antenna.

In Fig. 3.3 the resistance R_1 at the first series resonance ($f \approx 0.25 \cdot c/l$), the resistance R_2 at the first parallel resonance ($f \approx 0.45 \cdot c/l$), the resistance R_3 at the second series resonance ($f \approx 0.7 \cdot c/l$) and the resistance R_4 at the second parallel resonance ($f \approx 0.95 \cdot c/l$) are shown as a function of the length-radius ratio l/r for a dipole antenna [43]. The resonance resistance at the first two series resonances are nearly independent of the length-radius ratio l/r. The higher the length-radius ratio l/r gets, the higher the resistance at parallel

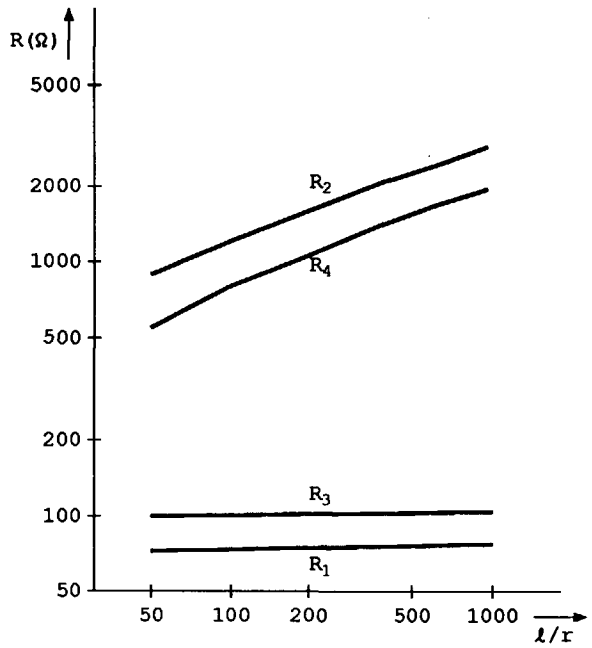

Fig. 3.3 The resistances at the series-resonance and parallel-resonance frequencies for a dipole antenna.

resonance gets.

A dipole antenna of length $2l = 150$ cm has its first series-resonance frequency at 100 MHz. In the vicinity of this first series-resonance frequency we can model the antenna impedance by a series-resonant network. This series-resonant network consists of a series resistance (approximately 75 Ω), a series capacitance (approximately 7.5 pF) and a series inductance (approximately 340 nH). This approximation of the antenna impedance is sufficiently accurate within the broadcast band (87.5–108 MHz) for our purposes. In the vicinity of the other resonant frequencies, the equivalent networks can also be obtained from Fig. 3.2 and Fig. 3.3.

The impedance of a whip antenna can easily be derived from the impedance of a dipole antenna by using the image theory [44]. This theory states that the impedance of a whip antenna of length l and radius r, operating against an infinitely-large and perfectly-conducting ground plane, is half the impedance

of the dipole antenna of total length $2l$ and radius r, operating in free space. Thus, the values given in Fig. 3.2 and Fig. 3.3 have to be halved to obtain the impedance of the whip antenna. Near the first series resonance ($f = 100$ MHz for $l = 75$ cm and $r = 3$ mm) we can approximate the impedance of the whip antenna by a series-resonant circuit. The series resistance is approximately 35 Ω, the series capacitance is approximately 15 pF and the series inductance is approximately 170 nH.

The given values for the equivalent series-resonant circuit of the whip antenna are only valid if the ground plane of the antenna is of infinite size and perfectly conducting. In reference [45] a comparison of calculations and measurements for a whip antenna mounted on a ground plane with limited dimensions is given. The data are limited to ground planes with dimensions in the order of one wavelength and larger. The variations in the impedance of the whip antenna are in the order of several ohms for a circular ground plane with a diameter of $0.25 \cdot \lambda$ (equivalent to 75 cm for a 75 cm long whip antenna at 100 MHz). A variation of a few ohms can usually be neglected. The variations in antenna impedance diminish for larger ground planes.

As the electromagnetic field is horizontally polarized, the antenna should be placed horizontal for a maximum signal reception (see Section 3.3) and hence the question arises as to whether this situation can be modeled as an antenna with a large ground plane. Besides, in practice, the ground plane of whip antennas mounted on portable radios is much smaller than 75 cm. To investigate the behavior of the antenna impedance in such a practical situation, some measurements on a whip antenna mounted on a portable receiver (dimensions: $80 \times 200 \times 320$ mm) have been carried out with the use of a network analyzer. The position of the whip antenna, relative to the portable radio and the measurement setup, was changed to obtain the variation in the antenna impedance.

The measured real part of the antenna impedance, at the first series resonance, ranged from 30 to 90 ohm. The first resonance frequency ranged from 100 to 110 MHz. The shift in resonance frequency indicates a decrease in the series inductance and capacitance of about 10% each. Further observations showed that the antenna impedance drawn in Fig. 3.2 does not match with the measurements for high frequencies ($f > 0.75 \cdot c/l$). Firstly, the resonance frequencies can differ about 30% from the values given in Fig. 3.2, and secondly, the impedance at the parallel-resonance frequencies can differ about 40% from the values given in Fig. 3.3. Also, small extra loops may be present within the large loops as given in Fig. 3.2. The consequences of this behavior and the influence on the design of the image rejection filter and the resulting

image rejection will be presented in the Chapter 4.

3.3 The antenna transfer

In this section the transfer from the electromagnetic field to the output voltage for a dipole antenna at its first series-resonance frequency will be discussed. The transfer for the whip antenna will be derived from the transfer for the dipole antenna. Next, the directional behavior of the dipole antenna transfer will be treated by means of the (idealized) radiation pattern for different operating frequencies. The transfer as a function of the frequency of the dipole and monopole antenna will be given by means of the radiation pattern. The radiation pattern and the antenna transfer for a whip antenna will be derived from that of the dipole antenna.

The open-terminal voltage V_a of a receiving dipole antenna can be expressed as [46]:

$$V_a = E_i \cdot h_{\text{eff}} \,, \tag{3.1}$$

where E_i denotes the incident electric field in the direction in which the antenna has the highest transfer and h_{eff} is the effective antenna height. The incident electric field is a function of the direction of observation. This aspect will be discussed later in this section with the use of the radiation pattern. The effective antenna height for a dipole antenna of total length $2l$ at its first series-resonance frequency is given by [46]:

$$h_{\text{eff}} = \frac{4l}{\pi} \,. \tag{3.2}$$

Thus, for a 150 cm long dipole antenna, the effective antenna height is $h_{\text{eff}} = 0.95$ m at 100 MHz.

The theory of images [44] states that the effective antenna height for a whip antenna of total length l on a perfectly-conducting ground plane of infinite size, is half the effective antenna height of a dipole antenna of total length $2l$. This means that the value given in Equation 3.2 must be halved to obtain the effective antenna height for a whip antenna. A 75 cm long whip antenna thus has an effective antenna height of 0.48 m at 100 MHz.

The transfer of the antenna depends on the orientation of the antenna with regard to the polarization of the electromagnetic field. This dependence can be represented by the radiation pattern of the antenna. In Fig. 3.4a the (idealized) sinusoidal current distribution along the dipole antenna for the first series-resonance frequency is shown with the resulting radiation pattern [47]. The

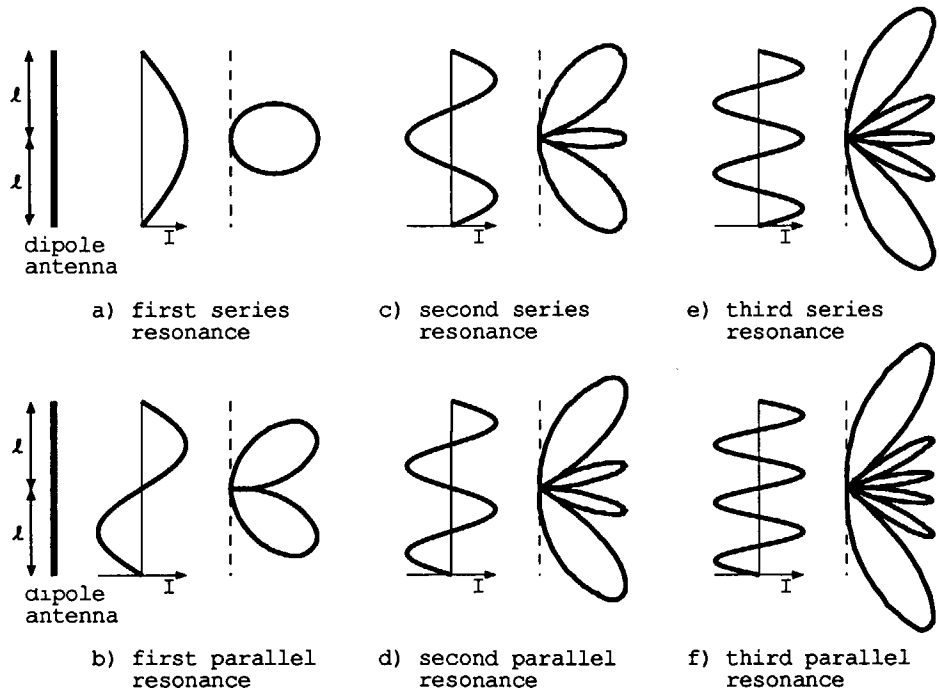

Fig. 3.4 The dipole radiation pattern for several resonance frequencies.

radiation pattern is normalized, i.e. the maximum value, perpendicular to the antenna, is 1. The radiation pattern is rotationally symmetrical around the antenna axis.

For higher resonance frequencies, each half wavelength of wire adds a lobe to the radiation pattern, see Fig. 3.4b to Fig. 3.4f. From Fig. 3.4 it is clear that the transfer perpendicular to the antenna axis is zero at the parallel-resonance frequencies. The antenna transfer, perpendicular to the antenna axis, is the same for all series-resonance frequencies. The antenna transfer perpendicular to the antenna axis at non-resonance frequencies is lower than the gain at the first series-resonance frequency [47].

In the direction nearly parallel to the antenna axis, the transfer of the antenna increases with increasing frequency. In Fig. 3.5 the angle of the main lobe with respect to the antenna axis and the gain of this main lobe are given as a function of the operating frequency [48]. The gain mentioned in Fig. 3.5 is normalized to the transfer of a dipole antenna at its first series-resonance

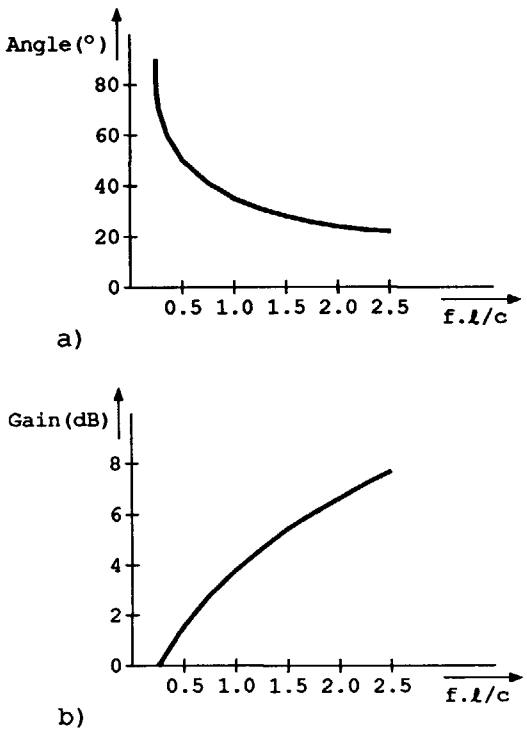

Fig. 3.5 The angle of the main lobe and its gain.

frequency. The increasing gain for higher frequencies can only be reached if the antenna wire remains a straight wire at these frequencies. The antenna radius should also be small when compared with the wavelength at these higher frequencies [49]. It can be concluded that the antenna gain increases slightly for increasing frequencies while the direction in which this maximum gain is reached gets closer to the antenna axis.

A whip antenna on an infinitely sized and perfectly conducting ground plane has only the upper half of the radiation pattern shown in Fig. 3.4. As the recommended polarization of the electromagnetic field is horizontal, i.e. the electrical vector is in the horizontal plane [9], the whip antenna should be placed horizontal to obtain a maximum received signal level. Taking the normalization in Fig. 3.5 into account, the gain at the image channels is the same

as with the dipole antenna. The radiation pattern changes when the ground plane is not perfectly conducting or when the ground plane has finite dimensions. For instance, the nulls in the radiation pattern are filled [48]. Only in special situations where waves, reflected against objects in the environment of the antenna, add in phase does the gain of the antenna increase. Generally this is not the case. The consequences of the above mentioned antenna gain on the design of the front-end of the FM receiver will be taken into account in Chapter 4.

3.4 Discussion

As the frequency difference of the image channels and the broadcast channels in an upconversion receiver is much larger than in a downconversion receiver, the impedance and the transfer of the antenna may change considerably. Due to a variable antenna environment in, for instance, portable receivers, only a rough estimation of the antenna behavior at the broadcast channels and at the image channels has been given.

The impedance of a whip antenna can be modeled by a series-resonant circuit ($R_a = 35\ \Omega$, $C_a = 15$ pF and $L_a = 170$ nH) for frequencies in the vicinity of the first series-resonance frequency (100 MHz for a 75 cm long whip antenna). Measurements show that, due to a ground plane of limited size in a portable receiver, the series resistance increases (to about 90 Ω) and the center frequency of the series-resonant network varies about 10%.

For a field strength of 1 V/m, the open-terminal voltage is about 0.48 V at 100 MHz for a 75 cm long whip antenna. For higher frequencies, the gain of the antenna increases slightly while the direction in which this gain is increased approaches the antenna axis. Therefore it may be concluded that the image suppression of the upconversion receiver is not seriously endangered. The information presented in this chapter will be used in Chapter 4 for the design of the front-end for the FM upconversion receiver.

4. Design of the FM upconversion receiver front-end

4.1 Introduction

In this chapter the design of the front-end for an upconversion receiver to be applied in a portable radio will be undertaken. Models for the SAW transversal filter, the SAW resonator filter and the whip antenna are indispensable for the design of this front-end. They have been given in Chapter 2 and Chapter 3 and will be summarized in this chapter where necessary. Parts of this chapter have previously been published in references [21] and [22].

A block diagram of the FM upconversion receiver is shown in Fig. 4.1. The signal from the whip antenna passes through the RF filter to the RF amplifier (amp.1). The RF filter is an LC or SAW filter for the suppression of signals at the image and spurious channels. The RF amplifier converts the output signal of the RF filter into the desired input signal for the mixer. By applying a local oscillator (LO) signal and the mixer, the radio frequent (RF) signal from the output of amplifier 1 is converted into a signal at the intermediate frequency. The second amplifier (amp.2) drives the SAW intermediate frequency (IF) filter with the amplified mixer output signal. The IF filter selects the desired channel. The SAW IF filter is loaded with a third amplifier (amp.3), which drives succeeding stages for the detection of the FM signal. The FM detector and the local oscillator (synthesizer) circuitry will not be considered in this thesis.

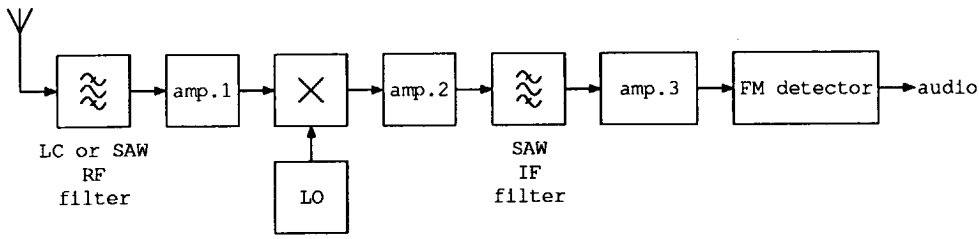

Fig. 4.1 A block diagram of the FM upconversion receiver.

In Section 4.2 the target specifications for the receiver front-end will be discussed. The choice of the intermediate frequency will be discussed in Section 4.3. In Section 4.4 the antenna-RF filter-amplifier configuration at the receiver input will be discussed with regard to the sensitivity of the receiver. The design of the RF filter and the influence of the antenna behavior on the image and spurious suppression will be discussed in Section 4.5. In Section 4.6 the choice for the IF filter, a SAW transversal filter or a SAW resonator filter, will be discussed. Section 4.7 deals with the noise and dynamic range optimization for an on-chip SAW transversal IF filter and an on-chip SAW resonator IF filter. In Section 4.8 some miscellaneous front-end considerations, such as the type of amplifiers to be used in the front-end and the gain distribution in the front-end, will be handled. Section 4.9 concludes this chapter with a discussion.

4.2 Target specifications

First, the sensitivity will be discussed. The quieting sensitivity of a receiver is the (FM modulated) carrier voltage applied to the receiver input for a 26 dB signal-to-noise ratio (measured unweighted) in a mono baseband audio signal at the output. As in the present work only the receiver front-end will be discussed, the noise floor at the receiver input is a more suitable specification for the sensitivity. A quieting sensitivity of 2 μV (mono, $B = 20$ kHz, $\Delta f = \pm 22.5$ kHz) as, for instance, given in reference [50] can be converted into a noise floor at the receiver input by using the FM signal-to-noise ratio improvement factor (as treated in, for instance, references [24] or [51]). This yields a receiver noise floor of approximately 1 μV, where the noise has been measured in a 180 kHz bandwidth.

The galactic noise can be represented by an antenna noise voltage $u_a = \sqrt{4kTF_aR_aB}$, where R_a is the antenna resistance and F_a is the antenna noise factor, which is about 10 at 100 MHz [52]. The antenna series resistance is 35 ohm at 100 MHz for a 75 cm long whip antenna. The open-terminal noise voltage at the output of the antenna then amounts to 1 μV in a 180 kHz bandwidth. This means that a receiver noise floor much lower than 1 μV does not lead to the reception of more stations or a better reception of the stations, certainly not when the median man-made noise in a business area is ample 10 dB higher than the galactic noise [52].

The maximum RF signal handling capability of a receiver can be defined as the maximum RF signal amplitude that can linearly be handled by the receiver, i.e. a compression below 1 dB. For signals with larger amplitudes than this

maximum signal handling amplitude, the front-end circuitry will show strong non-linearities. A maximum linear RF signal handling capability of about 200 mV_{eff} is a typical value for a conventional integrated downconversion receiver front-end. The dynamic range of the receiver is then in the order of 106 dB, where a 180 kHz noise bandwidth has been used.

The interference caused by the image channel depends on the image suppression of the receiver and the amplitude of the signal at the image frequency. The image suppression of a conventional portable downconversion receiver is usually in the order of 25–60 dB. The suppression of the spurious responses due to harmonics of the local-oscillator and the suppression of RF and LO signals in the IF part of the receiver due to feedthrough in the mixer should be in the same order as the image suppression.

The selectivity of the IF filter at the adjacent channel, in combination with the amplitude of the adjacent channel signal, determines the interference on the desired signal. The IF filter applied in the upconversion front-end should have a filter characteristic that is comparable with that of a ceramic filter as applied in a conventional portable downconversion receiver. Figure 4.2 shows the selectivity curve for a typical ceramic filter. The attenuation in the stopband is usually 40 dB. The selectivity curve for a coupled resonator is shown as a comparison, where a resonator coupling factor of $kQ = 1$ has been used.

Because the information is enclosed in the phase of the carrier signal, a phase transfer function which exhibits a non-linear function of the frequency results in non-linear distortion in the audio signal. In other words, the variation in group delay within the FM signal bandwidth should be as small as possible [24]. A variation of several hundred nanoseconds, as present in ceramic filters, gives a distortion in the audio signal which is usually acceptable ($< 1\%$).

A local oscillator signal with a poor carrier-to-noise ratio causes two effects: Firstly, the noise of the local oscillator is added to the desired signal by means of feedthrough or imperfect balancing, thus affecting the signal-to-noise ratio of the audio signal. Secondly, the out-of-band noise of the local oscillator signal may be converted into the frequency of the desired channel by an undesired signal with a large amplitude at an adjacent channel. This reciprocal mixing limits the receiver selectivity. The interference into the desired signal should be compared with the interference due to an IF filter with a limited suppression of the adjacent channels. Investigations on the local oscillator are beyond the scope of the present work. That they form no insurmountable problem may

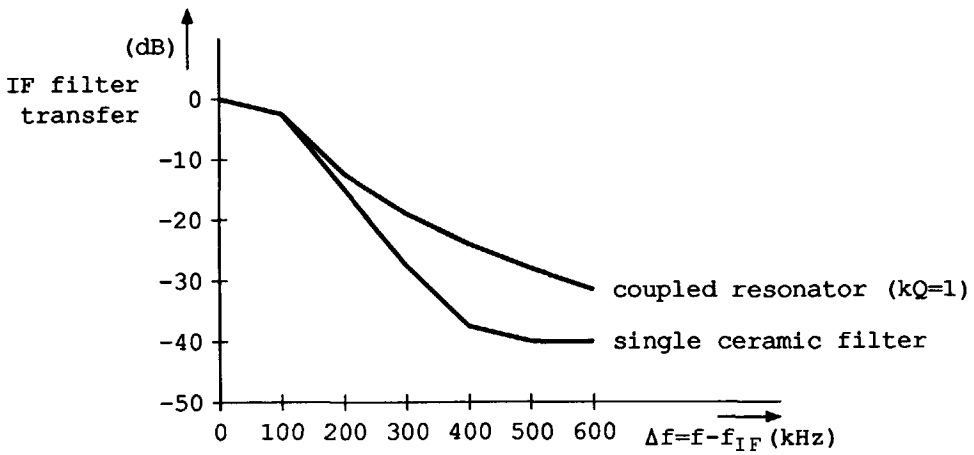

Fig. 4.2 The selectivity curves for a ceramic filter and a coupled resonator.

be illustrated by reference [53] where a fully integrated regenerative oscillator suitable for our application has been given.

When we receive two (or more) undesired signals with a large amplitude and a desired signal with a much smaller amplitude, the detection of the information can be disturbed due to non-linearities in the receiver. These non-linearities give rise to the generation of intermodulation distortion components at several frequencies, such as second-order intermodulation distortion components at $|\omega_1 \pm \omega_2|$ and third-order intermodulation distortion components at $|2\omega_1 \pm \omega_2|$ and $|\omega_1 \pm 2\omega_2|$, where ω_1 and ω_2 are the frequencies of the two undesired signals. When a band-pass filter is used as a preselection filter, only the third-order intermodulation components at $|2\omega_1 - \omega_2|$ and $|\omega_1 - 2\omega_2|$ fall within our receiving range. In Fig. 4.3 the frequency spacing of the desired signal and the undesired signals is sketched for a disturbed reception of the desired signal (see also reference [54]). The third-order intermodulation can be measured by applying two signals of equal amplitude but of different frequencies to the front-end. The intermodulation-free dynamic range (IMFDR) is defined as the ratio of the power of each of the two signals with equal amplitude but different frequencies, and the noise power in a given bandwidth, provided that the intermodulation products of the two signals have the same power as the noise in this given bandwidth. The intermodulation-free dynamic range should be about 80 dB in case the situation as shown in Fig. 4.3 oc-

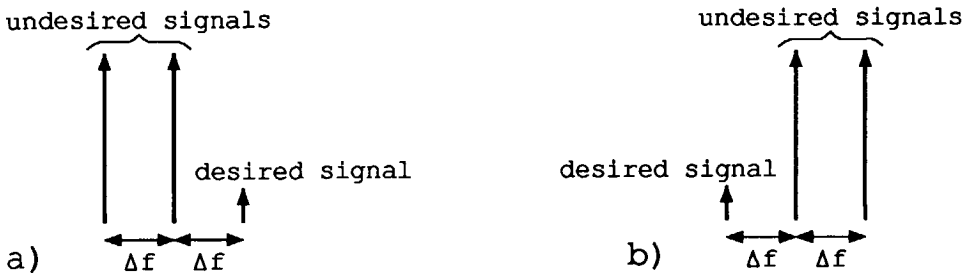

Fig. 4.3 The frequency spacing for a disturbed reception due to third-order intermodulation.

curs and the stopband rejection of the IF filter amounts to 40 dB. In practice, this situation does not occur frequently and a smaller intermodulation-free dynamic range is acceptable.

Aspects such as temperature-dependent behavior, reproducibility, power supply voltage and current consumption should also be taken into account in the design, depending on the specified application area of the receiver.

The target specifications for a receiver front-end in a portable radio can then be summarized as follows:

input frequency range	87.5-108 MHz
input noise floor (B = 180 kHz)	1 μV
maximum linear input level	200 mV
DR (B = 180 kHz)	106 dB
$IMFDR_3$ (B = 180 kHz)	80 dB
IF filter	comparable performance as a ceramic filter or a coupled resonator (see Fig. 4.2)
IF filter stopband rejection	40 dB
suppression of spurious responses	50 dB
temperature range	−30 °C to +70 °C

4.3 The choice of the intermediate frequency

The IF filter is an on-chip SAW filter, either a transversal filter or a resonator filter. Because the SAW filter is fabricated in the ZnO-SiO_2-Si, three-layered

structure, higher-order Rayleigh-wave modes are present. These modes result in extra pass-bands, as illustrated in the measurements section in Chapter 2.

The most dominant spurious pass-band in the filter transfer function is caused by the second-order Rayleigh-wave mode at approximately 1.7 times the fundamental pass-band frequency. The signals passing through this spurious pass-band may trouble the detection or interfere with the desired signal. For instance, FM detectors using phase-lock loops may be pulled out of lock by these extra signals. A low-pass filter of sufficiently high order between the IF SAW filter and the FM detector attenuates these signals. However, such a filter is incompatible with the desire of an integrated receiver.

The frequencies of the (spurious) signals at the antenna, which cause these extra signals at the output of the IF filter, can be placed far away from the desired FM broadcast band frequencies by choosing the proper intermediate frequency. In this way, these spurious signals are maximally attenuated by the RF filter at the input of the receiver. This filter therefore has three functions: Suppression of the image channels, suppression of the LO harmonics-related spurious channels and the suppression of the higher-order Rayleigh-wave mode related spurious channels.

Figure 4.4 shows the RF spectrum, the LO tuning range and the IF filter pass-bands. The RF band that we desire to receive ranges from 87.5 to 108 MHz. The IF filter has an extra pass-band at $1.7f_{IF}$ due to the second-order Rayleigh-wave mode. The LO frequency f_{LO} is given by: $f_{LO} = f_{IF} + f_{RF}$, where f_{RF} represents the frequency of the desired RF signal. Signals with frequencies $|1.7f_{IF} - f_{LO}|$ or $1.7f_{IF} + f_{LO}$ at the input of the receiver pass through the $1.7f_{IF}$ pass-band in the IF filter. Signals at the frequencies $1.7f_{IF} + f_{LO}$ are higher than the frequency of the image channel ($f_{image} = f_{IF} + f_{LO}$) and are attenuated at least as much as the signals at the image channels when a low-pass RF filter is used.

Signals at the frequency $|1.7f_{IF} - f_{LO}| = |0.7f_{IF} - f_{RF}|$ can be attenuated if $|0.7f_{IF} - f_{RF}| \gg f_{RF}$ or $f_{IF} \gg 2.9f_{RF}$ in case of a low-pass RF filter. For the present application this would result in an intermediate frequency higher than 310 MHz which would lead to impractical demands on the electronic circuitry when only a IC process with an f_T of 3 GHz is available. Besides, the stability and fabrication accuracy of the IF filter become very stringent.

In the case of a band-pass RF filter, a maximum attenuation occurs if $0.7f_{IF} = f_{RF}$ or $f_{RF} = 140$ MHz for the present application. Moreover, a band-pass filter at the input of the receiver has the advantage that all signals outside the FM broadcast band are attenuated. An intermediate frequency of

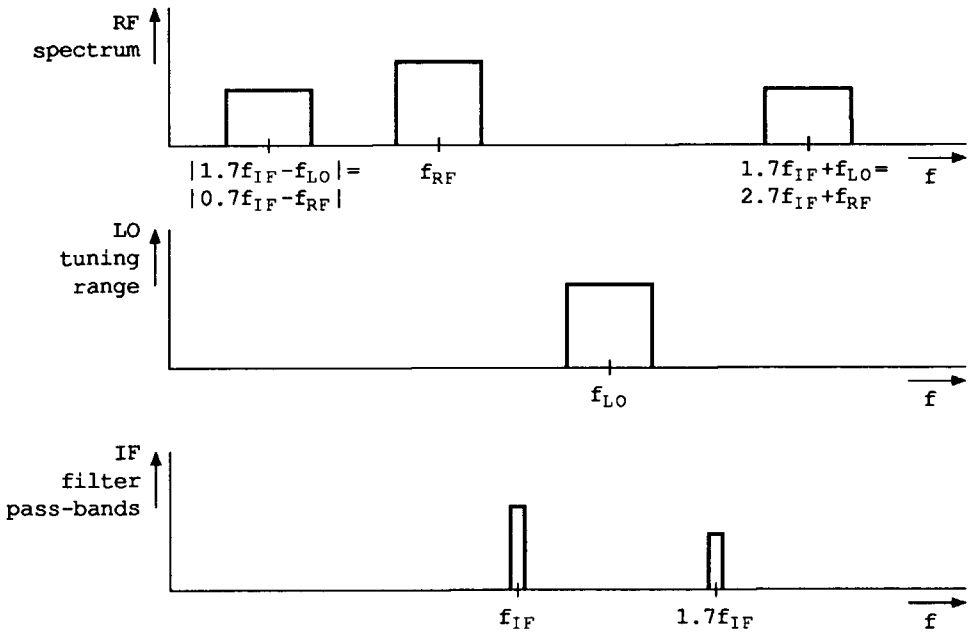

Fig. 4.4 The RF spectrum, the LO tuning range and the IF filter pass-bands.

140 MHz is not optimal with respect to the image suppression and the suppression of RF signals which pass through the front-end (feedthrough in the input filter, the input amplifier and the mixer) to the intermediate frequency part of the receiver. Depending on the order of the filter at the input, a somewhat higher intermediate frequency is preferable.

Here, we have chosen for 160 MHz as the intermediate frequency. The spurious pass-band in the IF filter then has a center frequency of 272 MHz. The local oscillator should be tunable from 247.5–268 MHz. The image channels range from 407.5 to 428 MHz, while the extra pass-band related spurious signals range from 4 to 24.5 MHz and 519.5 to 540 MHz.

The frequencies of the 519.5–540 MHz channels are higher than the frequencies of the image channels and are therefore at least as much attenuated as the signals at the image frequency. The input filter must be a band-pass filter for attenuating all signals at the image channels and at the Rayleigh-wave mode related spurious channels ranging from 4 to 24.5 MHz.

4.4 Antenna-RF filter-amplifier noise optimization

4.4.1 Introduction

At the input of the receiver we need a filter to attenuate the signals at the image channel, the LO harmonics-related spurious channels and second-order Rayleigh-wave mode related spurious channels. Four different filter configurations are possible.

There are two configurations using on-chip SAW filters. First of all, a SAW transversal filter may be connected between antenna and RF amplifier. The second configuration is an extension of the first configuration and uses a preamplifier between antenna and SAW filter. There are two configurations using LC filters, external to the chip. An LC filter may be connected between antenna and RF amplifier. The other configuration uses a preamplifier between antenna and LC filter.

In the following sections, the obtainable sensitivity of the receiver will be calculated for all four different configurations. Only amplifiers using bipolar transistors will be assumed to be usable at these frequencies.

4.4.2 Antenna-SAW filter-amplifier noise optimization

We will start with the noise optimization for a SAW transversal filter between antenna and RF amplifier. For the time being, a voltage source U_a is used as a representation for the antenna and we will omit the galactic noise picked up by the antenna. In Chapter 2 the model for the SAW transversal filter has been presented. Such a filter should be loaded with an amplifier with a low input impedance for an optimum signal transfer. Figure 4.5 shows the situation of interest for the noise optimization.

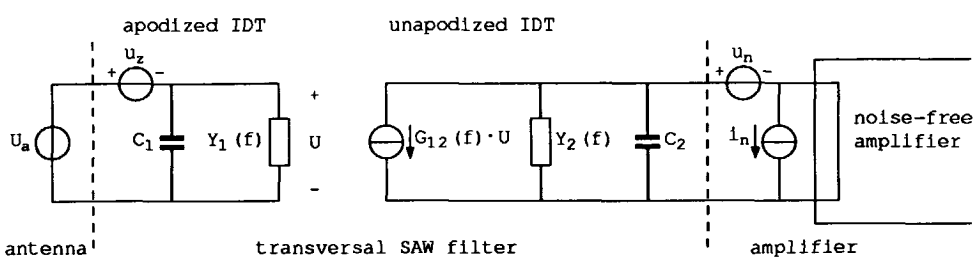

Fig. 4.5 The antenna-transversal SAW filter-amplifier configuration.

The number of electrode pairs in the main lobe of the sinc apodized IDT can be calculated from the desired bandwidth of 20.5 MHz at the center frequency of $f_o = 98$ MHz and amounts to $N_1 = 10$. If some influence of the unapodized IDT on the pass-band response is acceptable, $N_2 = 5$ may be chosen. Single electrodes are used for both IDTs. Because the number of electrode pairs is small, the admittances $Y_1(f)$ and $Y_2(f)$ may be neglected compared to the admittances formed by the static IDT capacitances C_1 and C_2, as has been shown in Chapter 2. The capacitances and the transconductance $G_{12}(f_o)$ within the pass-band of the transversal filter can be approximated by:

$$C_1 = 0.8\,\epsilon a N_1\,, \tag{4.1a}$$

$$C_2 = \epsilon a N_2\,, \tag{4.1b}$$

$$G_{12}(f) = G_{12}(f_o) = G(f_o)\,\frac{N_1}{2}\,N_2\,\mathrm{e}^{-j\omega_o L_{12}/v_o}\,, \tag{4.1c}$$

$$G(f_o) = 8\kappa^2 f_o \epsilon a\,, \tag{4.1d}$$

where the loss α in the acoustic path has been neglected. The constant 0.8 in Equation 4.1a is given for a sinc apodized transducer with four side lobes. The parameter a denotes the aperture (width) of the SAW device, ϵ denotes the effective dielectric permittivity of the three-layered ZnO-SiO_2-Si structure, κ^2 denotes the piezoelectric coupling constant of the ZnO-SiO_2-Si structure, v_o denotes the velocity of the surface acoustic wave and L_{12} denotes the distance between the centers of the two IDTs (see Chapter 2).

Notice that the unapodized IDT in Fig. 4.5 is used as the output IDT and that the apodized IDT in Fig. 4.5 is used as the input IDT. The capacitance of the unapodized IDT is smaller than the capacitance of the apodized IDT, which results in a lower effective noise floor at the input of the receiver.

The noise sources of the amplifier are represented by the voltage source u_n and the current source i_n. Only the noise of the base resistance r_b ($S(u_b) = 4kTr_b$), the base current shot noise ($S(i_b) = 2qI_B$) and the collector current shot noise ($S(i_c) = 2qI_C$) of a transistor at low frequencies ($f < f_T/\sqrt{h_{FE}}$) will be taken into account. By using $r_e = kT/(qI_C)$ and $I_B = I_C/h_{FE}$ (h_{FE} is the transistor DC current gain), the resulting power-density spectra of the noise sources u_n and i_n, for an amplifier with a single transistor input stage, can be approximated by:

$$S(u_n) = 4kT\,(r_b + \frac{r_e}{2})\,, \tag{4.2a}$$

$$S(i_n) = 4kT\,\frac{1}{2h_{FE}r_e}\,. \tag{4.2b}$$

The amplifier feedback network is assumed to be non-energic, thus contributing no noise. The noise caused by the SAW transmission line terminations is

represented by a voltage source u_z in series with the antenna voltage U_a. Its power-density spectrum is given by:

$$S(u_z) = 4kT \frac{1}{G(f_o)N_1^2} \cdot \tag{4.2c}$$

Both noise voltage u_n and noise current i_n can be transferred to a noise voltage in series with the antenna voltage. The power-density spectrum of the noise voltage $u_{n,\mathrm{tot}}$, in series with the antenna voltage U_a, is given by:

$$S(u_{n,\mathrm{tot}}) = \frac{4kT}{|G_{12}(f)|^2} \cdot \left[(r_b + \frac{r_e}{2}) \cdot |j\omega C_2|^2 + \frac{1}{2h_{\mathrm{FE}} r_e}\right] + \frac{4kT}{G(f_o)N_1^2} \cdot \tag{4.3}$$

Integrating this power-density spectrum from $f_o - B/2$ to $f_o + B/2$, where f_o represents the carrier frequency of the signal and B represents the signal bandwidth, and using $f_o \gg B$, yields:

$$\begin{aligned} u_{n,\mathrm{tot}}^2 &= \int_{f_o - B/2}^{f_o + B/2} S(u_{n,\mathrm{tot}})\, \mathrm{d}f \\ &= \frac{4kTB}{|G_{12}(f_o)|^2} \cdot \left[(r_b + \frac{r_e}{2}) \cdot \omega_o^2 C_2^2 + \frac{1}{2h_{\mathrm{FE}} r_e}\right] + \frac{4kTB}{G(f_o)N_1^2} \cdot \end{aligned} \tag{4.4}$$

A minimum value of $u_{n,\mathrm{tot}}^2$ can be obtained by differentiating $u_{n,\mathrm{tot}}^2$ with respect to r_e, hence an optimum value for r_e ($r_{e,\mathrm{opt}}$) is found:

$$r_{e,\mathrm{opt}} = \frac{1}{\omega_o C_2 \sqrt{h_{\mathrm{FE}}}} \cdot \tag{4.5}$$

The total noise voltage for this optimum value is then given by:

$$u_{n,\mathrm{tot}} = \sqrt{4kTB} \cdot \frac{1}{\kappa^2 N_1} \cdot \sqrt{\frac{\pi^2 r_b}{4} + \frac{\pi}{8 f_o \epsilon a N_2 \sqrt{h_{\mathrm{FE}}}} + \frac{\kappa^2}{8 f_o \epsilon a}} \cdot \tag{4.6}$$

The total noise voltage $u_{n,\mathrm{tot}}$ is shown in Fig. 4.6 as a function of the aperture a. It can be concluded that the aperture should be as large as possible for a minimization of the noise due to the base current shot noise, the collector current shot noise and the noise of the SAW transmission line terminations. The aperture has no influence on the noise contribution due to the transistor base resistance, because both the capacitance C_2 and the transconductance $G_{12}(f_o)$ are proportional to the aperture.

By using the values given in Chapter 2 ($\kappa^2 = 0.0089$, $f_o = 98$ MHz, $\epsilon = 135$ pF/m) and assuming some transistor parameters ($r_b = 10\ \Omega$ and $h_{\mathrm{FE}} = 100$),

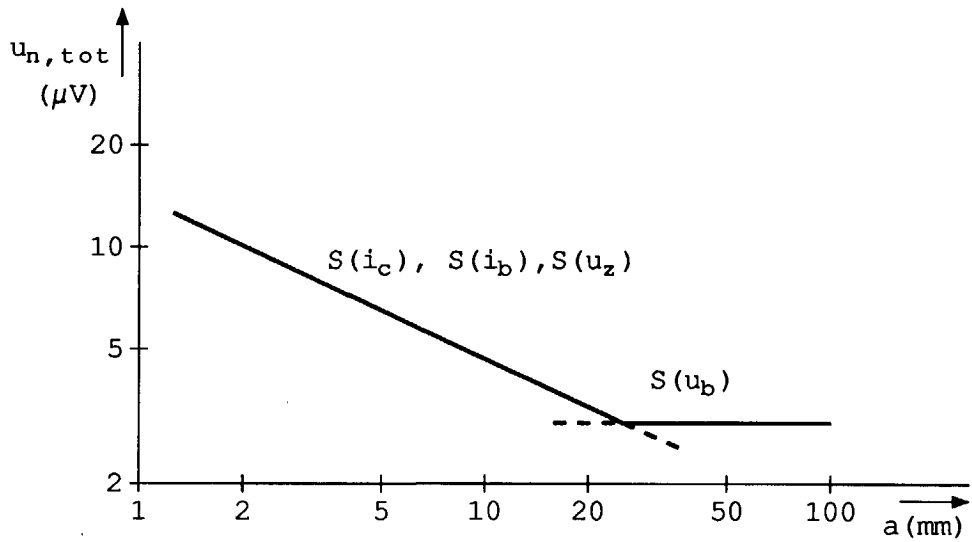

Fig. 4.6 The total noise voltage as a function of the aperture.

it follows that the intersect point of both curves in Fig. 4.6 is given by $u_{n,tot} = 18\sqrt{4kTr_bB} = 3.0$ μV and $a = 27$ mm. Although the wide-band SAW filter is short (about 1 mm), it should be very wide (27 mm) to obtain a low noise floor at the receiver front-end, thus occupying a large chip area. A minimum sensitivity of 3 μV even increases when the galactic noise is also taken into account. This figure implies that a sensitivity of 1 μV can hardly be obtained.

The use of inductances in parallel with C_2 will yield a much lower noise voltage. Such an inductance must then be placed off-chip because of its size ($\approx$ 150 nH), thereby increasing the number of pins of the IC. Also the feedthrough from pin-to-pin will then limit the ultimate rejection of the filter, leading to a poor image and spurious suppression. Therefore, we will refrain from inductances for this application.

4.4.3 Antenna-preamplifier-transversal SAW filter

In the previous section it was shown that the desired noise floor cannot be reached with a SAW transversal filter connected between antenna and RF amplifier. A receiver noise floor of 1 μV can be obtained by inserting a preamplifier between the antenna and the SAW transversal filter. This amplifier must then not only handle all signals within the FM broadcast band, but also

the signals outside this band. To obtain a small occupied chip area, the voltage amplification of the amplifier must be large, which might become incompatible with a desired supply voltage and the given maximum voltage delivered by the antenna.

4.4.4 Antenna-preamplifier-LC filter noise optimization

In this configuration, the LC image and spurious rejection filter follows a preamplifier, which is directly connected to the antenna. The model for the 75 cm long whip antenna at 100 MHz consists of a voltage source U_a and a series-resonant network (see Chapter 3). The series resistance R_a is nominally 35 ohm. The series-resonant network is in resonance at 100 MHz.

Figure 4.7 shows the situation of interest for the noise optimization of the preamplifier. The influence of the impedance formed by the inductance L_a

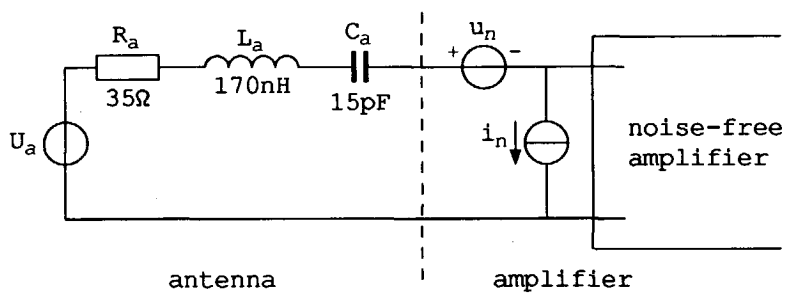

Fig. 4.7 The antenna-preamplifier configuration.

and the capacitance C_a will be neglected in the following noise calculations, because it marginally changes the result from the noise optimization. The spectral densities of the noise voltage source u_n and the current source i_n of the preamplifier have already been given in Equation 4.2a and Equation 4.2b. The total noise can be summed in one source in series with the signal source U_a. By following the same procedure as that for the SAW device, an optimum value for r_e ($r_{e,\mathrm{opt}} = R_a/\sqrt{h_{\mathrm{FE}}}$) is found, leading to a minimum noise voltage of:

$$u_{n,\mathrm{tot}} = \sqrt{4kTB \cdot \left[F_a R_a + r_b + \frac{R_a}{\sqrt{h_{\mathrm{FE}}}}\right]}, \tag{4.7}$$

where the factor $F_a R_a$ represents the galactic noise [52]. For $h_{\mathrm{FE}} = 100$, it can be concluded that the contribution of the base current and collector current

shot noise may be neglected. The noise due to the base resistance can usually be made sufficiently small. The total noise voltage then amounts to 1 μV, where $r_b = 10\ \Omega$, $F_a = 10$, $R_a = 35\ \Omega$ and $B = 180$ kHz have been used.

4.4.5 Antenna-LC filter-amplifier noise optimization

In Chapter 3 it has been shown that the antenna series resistance varies from 30 to 90 ohm, depending on the antenna environment. To minimize the variation in the filter transfer function, the filter must be terminated with an accurately-known resistance R_l. The filter will be designed in Section 4.5.

An amplifier employing non-energic negative feedback will be assumed here to realize the accurately-known impedance R_l [17]. Figure 4.8 shows the situation of interest. The expressions for the noise voltage u_n and the noise current

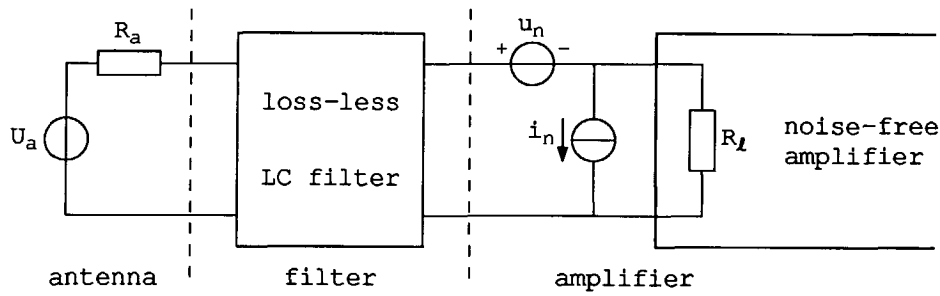

Fig. 4.8 The antenna-LC filter-amplifier configuration.

i_n of the amplifier have already been given in Equation 4.2a and Equation 4.2b.

As a consequence of the varying output impedance of the filter, a special noise optimization procedure must be used in order to keep the power-density spectrum of the effective input noise voltage flat over the filter pass-band. Assuming $r_b \ll r_e$ and $R_l \gg R_a$, the optimum value for r_e ($r_{e,opt}$) is given by $r_{e,opt} = R_l/\sqrt{h_{FE}}$ [1]. The total noise voltage in series with the antenna voltage can then be written as [1]:

$$u_{n,tot} = \sqrt{4kTB \cdot \left[F_a R_a + r_b + \frac{R_l}{2\sqrt{h_{FE}}}\right]}\ . \tag{4.8}$$

For $R_l > 2R_a$, the noise floor at the input of the receiver is somewhat higher than in the case of a preamplifier directly coupled to the antenna. However,

for $R_l < 2F_aR_a\sqrt{h_{FE}}$, the effective input noise voltage has not increased much compared to the antenna-preamplifier configuration. For instance, for $R_l = 500\ \Omega$, $r_b = 10\ \Omega$, $h_{FE} = 100$, $R_a = 35\ \Omega$ and $F_a = 10$ the noise voltage amounts to $u_{n,tot} = 1.1\ \mu V$.

4.4.6 Comparison and conclusions

An on-chip transversal SAW filter, between antenna and amplifier, leads, when no inductances are allowed, to a relatively high equivalent noise floor at the input of the receiver and occupies a considerable chip area. A preamplifier may lower this equivalent noise floor and lessen the occupied chip area at the cost of extra power consumption.

A problem associated with the on-chip SAW filters in the ZnO-SiO_2-Si, three-layered structure, is the presence of higher-order Rayleigh-wave modes (see Chapter 2). These higher-order Rayleigh-wave modes result in extra pass-bands in the frequency response. The pass-band at 1.7 times the frequency of the fundamental pass-band frequency, due to the second mode, is dominant. The signals passing through this extra band have also to be handled by the front-end, thereby raising the demands posed on the front-end.

There are two ways of configuring an LC filter: directly coupled to the antenna or coupled to the antenna via a preamplifier. In the first case, a slight, usually negligible, increase in the effective noise floor is observed compared to the second case. Further, an amplifier with an accurately-known input impedance is then needed to terminate the filter correctly. However, this configuration has the advantage that all signals outside the FM broadcast band are attenuated before any amplification takes place.

An LC filter cannot be integrated. However, as the receiver IC has to have a pin for the RF input signal, it does not make much difference whether the antenna is connected directly to the input of the receiver or via an LC band-pass filter. From the above discussion, it can be concluded that the configuration with an off-chip LC filter between antenna and amplifier yields the best solution in the present application.

4.5 Design of the image and spurious rejection filter

4.5.1 The filter behavior at the broadcast channels

Before we discuss the effects of the antenna impedance on the transfer function of the image and spurious rejection filter, the model for the antenna, as developed in Chapter 3, will be summarized. The 75 cm long whip antenna at

its first series resonance can be represented by a voltage source U_a and a series-resonant network with a resonance frequency of 100 MHz. Due to variations in the environment of the antenna, the real part of the antenna impedance ranges from 30 to 90 ohm. The resonance frequency of the series network may shift about 10% around a nominal value of 100 MHz. Figure 4.9 sketches this model.

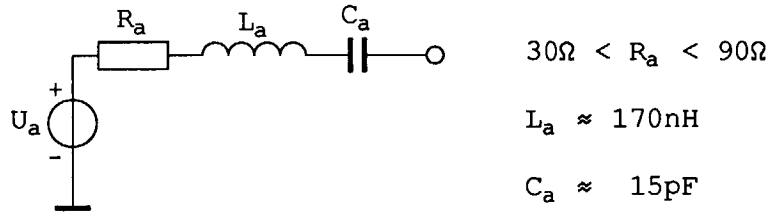

Fig. 4.9 The model for a 75 cm long whip antenna in the vicinity of 100 MHz.

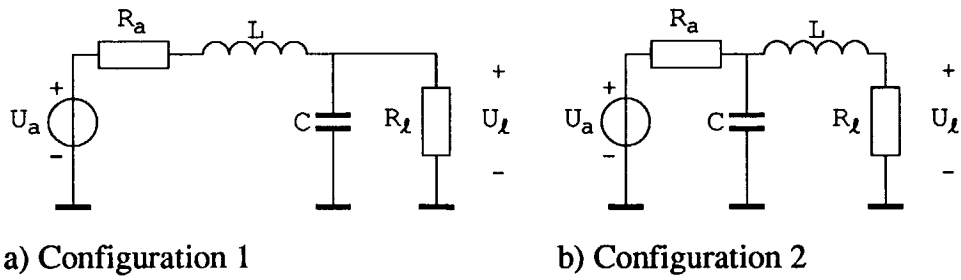

Fig. 4.10 The second-order low-pass filters with their sources and loads.

First the influence of the varying real part of the antenna impedance will be elucidated with the use of a second-order low-pass filter, hence the series inductance L_a and the series capacitance C_a will be omitted. Figure 4.10 shows the two possible second-order low-pass filters with their sources and loads. The transfer function $H(s)$ from antenna voltage U_a to the voltage U_l over the load resistance R_l is given by:

$$H(s) = \frac{U_l}{U_a} = \frac{1}{s^2 + 2\xi\omega_n s + \omega_n^2} \cdot \frac{1}{LC}, \tag{4.9a}$$

$$H(s) = \frac{U_l}{U_a} = \frac{1}{s^2 + 2\xi\omega_n s + \omega_n^2} \cdot \frac{1}{LC} \cdot \frac{R_l}{R_a}, \tag{4.9b}$$

for configuration 1 and 2, respectively. The parameter s denotes the Laplace variable. The damping ξ and the -3 dB bandwidth ω_n for configuration 1 are given, respectively, by:

$$\xi = \frac{1}{2\omega_n}\left[\frac{1}{R_l C} + \frac{R_a}{L}\right], \tag{4.10a}$$

$$\omega_n = \sqrt{\frac{R_a + R_l}{R_l LC}}. \tag{4.10b}$$

For configuration 2 they are given by:

$$\xi = \frac{1}{2\omega_n}\left[\frac{1}{R_a C} + \frac{R_l}{L}\right], \tag{4.11a}$$

$$\omega_n = \sqrt{\frac{R_a + R_l}{R_a LC}}. \tag{4.11b}$$

At low frequencies ($s \rightarrow 0$) the transfer $H(s)$ approaches 1 for $R_l \gg R_a$ and is then nearly independent of R_a for both configurations. The -3 dB bandwidth ω_n approaches $1/LC$ for $R_l \gg R_a$ and is then also nearly independent of R_a in configuration 1 while $R_l \ll R_a$ should be chosen in configuration 2 to obtain the same result. It may be clear that in configuration 2 no suitable compromise for R_l can be found. Thus only configuration 1, with $R_l \gg R_a$ is applicable in our situation.

The damping ξ in configuration 1 is determined by two terms: $1/(R_l C)$ and R_a/L. The Equations 4.10a and 4.10b can be solved for a desired ω_n, ξ, R_a and R_l. Two solutions result for L and C each: In one case the damping is mainly determined by the factor $1/(R_l C)$ and in the other case the damping is mainly determined by the factor R_a/L. For a damping independent of R_a, the term $1/(R_l C)$ should be much larger than the term R_a/L or $R_a R_l \ll L/C$.

Using the above knowledge and consulting a filter handbook, for instance reference [55], we can design the filter transfer function. Using $\omega_n/(2\pi) = 30$ MHz, $R_l = 350\ \Omega$ and calculating the band-pass filter component values ($f_o = 98$ MHz) leads to the filter shown in Fig. 4.11. The transfer function U_l/U_a for three different values of R_a (30 ohm, 50 ohm and 90 ohm) is given in Fig. 4.12. As can be concluded from Fig. 4.12, the variation in the transfer function U_l/U_a, for a variation in R_a from 30 to 90 ohm, is limited to approximately 15%. This variation is sufficiently small for our application.

Notice that a value of 350 ohm for R_l is sufficiently high to not increase the receiver input noise floor, as calculated in the previous section, significantly. Because the damping of the filter is effectuated at the output of the filter,

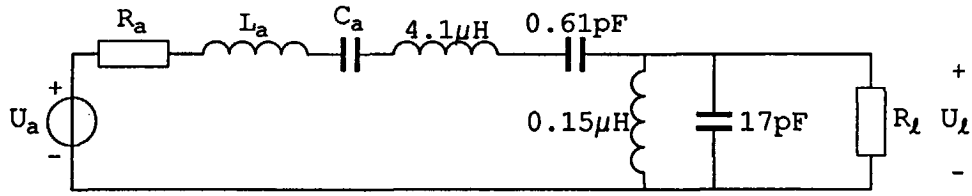

Fig. 4.11 The *LC* Butterworth band-pass filter realization.

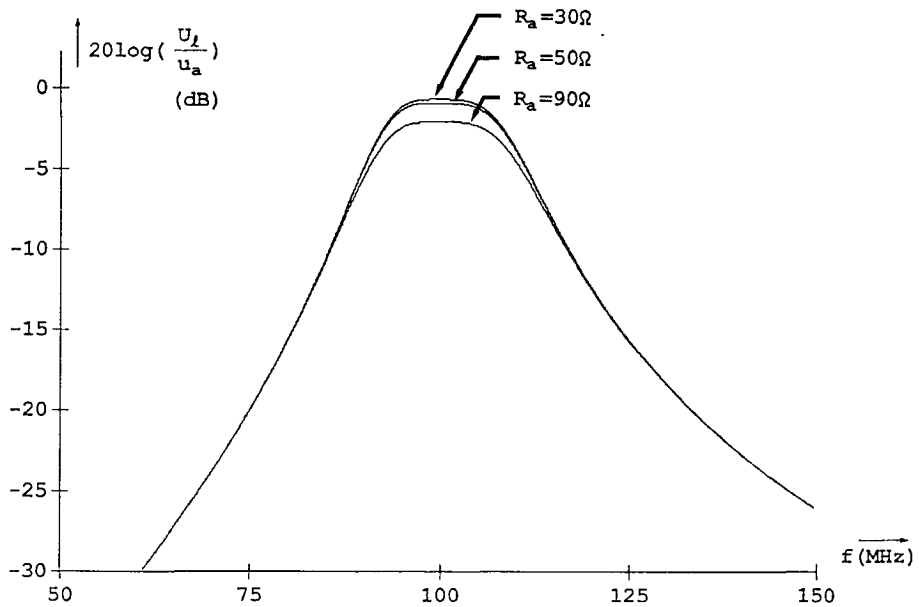

Fig. 4.12 The *LC* filter transfer function U_l/U_a for three different values of R_a.

the output impedance of the filter shows a very undamped behavior (an undamping of $\xi = R_a/(2\omega_n L)$). This has its influence on the noise optimization procedure for the antenna-LC filter-amplifier configuration and has already been treated in the previous section.

4.5.2 The filter behavior at the image channels

Next, the influence of the antenna impedance for frequencies higher than the FM broadcast band frequencies on the filter transfer function will be treated.

In Chapter 3 it has been shown that the antenna impedance exhibits several series-resonance frequencies (100 MHz, 300 MHz, etc.) and several parallel-resonance frequencies (200 MHz, 400 MHz, etc.), alternating with one another. The impedance at the series-resonance frequencies is approximately 35 ohm. The impedance at the parallel-resonance frequencies is in the order of 500 ohm to 3000 ohm and decreasing for higher frequencies. The exact frequencies of the parallel and series resonances are unpredictable due to variations in the antenna environment.

The lossless RF band-pass LC filter consists only of inductances and capacitances. The first element of the low-pass equivalent of this band-pass filter is either a series inductance or a parallel capacitance. In combination with a high antenna impedance at one of the parallel-resonance frequencies, the filter transfer function shows small dips in the frequency response. The depth of these dips decreases with the frequency. The frequencies at which these dips occur are the frequencies of the parallel resonances and are therefore unpredictable due to variations in the antenna environment. Anyhow, there is no increase in the transfer function from the antenna voltage source to the voltage over the load resistance due to the parallel-resonance effects in the antenna impedance.

To attenuate the spurious responses, which are a consequence of the second-order Rayleigh-wave mode in the on-chip SAW IF filter, a band-pass filter has to be used instead of a low-pass filter. For frequencies above the center frequency of the band-pass filter, the band-pass filter transfer function approaches the low-pass filter transfer function. Therefore, the above conclusion, concerning the possible deterioration of the filter transfer function due to parallel-resonance effects in the antenna impedance, is also applicable to band-pass filters.

As has been explained in Chapter 3, the antenna transfer increases slightly for increasing frequencies. For instance, the antenna gain at 400 MHz (the frequencies of the image channels for an intermediate frequency of 160 MHz and $f_{LO} > f_{IF}$) may be about 4 dB higher than the gain at 100 MHz. The above designed second-order band-pass filter (see Fig. 4.11) attenuates the signals at 400 MHz about 50 dB. An antenna gain of 4 dB, whether it actually occurs or not, can therefore be neglected.

4.6 Considerations regarding the IF filter

In Section 4.2 the selectivity curves for a ceramic filter and a coupled resonator filter, both with a -3 dB bandwidth of 180 kHz, have been shown (Fig. 4.2).

The adjacent channel suppression (Δf = 300 kHz) amounts to 28 dB and 18 dB, respectively. The variation in group delay of the ceramic filter transfer function is in the order of several hundred nanoseconds.

Either a SAW transversal filter or a SAW resonator filter may be applied as an IF filter. There is an essential difference in the properties of transversal filters and resonator filters. The phase-frequency transfer function (and therefore also the group delay) and the amplitude-frequency transfer function in a SAW transversal filter can be designed independently. This leads to a minimal distortion in the audio signal in combination with a maximum rejection of adjacent channels. Resonator filters are minimum phase devices and show a variation in group delay which is inversely proportional to the filter bandwidth. Thus in the design of a resonator filter, a compromise between allowable distortion and desired filtering must be sought, while in a transversal filter no such compromise exists.

A single resonator with a −3 dB bandwidth of 180 kHz has an adjacent channel suppression (Δf = 300 kHz) of about 11 dB, which is not very much. The filter slope is 6 dB/oct in the detuning. The variation in group delay for a 180 kHz bandwidth is 880 ns peak-to-peak. A coupled resonator filter shows a more than doubled adjacent channel suppression (18 dB @ 300 kHz). The filter slope is 12 dB/oct in the detuning. The variation in group delay is minimal for a resonator coupling factor $kQ = 0.75$ and amounts to 350 ns for a −3 dB bandwidth of 180 kHz.

Preliminary calculations show that the length of an on-chip SAW resonator is about 6 mm, while the width has to be chosen according to noise and/or dynamic range specifications, as will be shown in Section 4.7. If the selectivity of a coupled resonator is insufficient, either a selective FM detector or an extra coupled resonator should be applied.

The bandwidth B of a SAW transversal filter is directly related to the distance L between the nulls of the sinc function in the apodized IDT through the surface acoustic wave velocity v_o: $B = 2v_o/L$. In the case of an ideal, infinitely long, sinc apodized IDT, the frequency response shows a rectangular pass-band: 1 for $f_o - B/2 < f < f_o + B/2$ and 0 elsewhere. In the case of a finite IDT length, the −3 dB bandwidth will be somewhat smaller, the out-of band rejection lower and the ripple in the pass-band larger. For instance, by using a rectangular apodization in the IDT (also called unapodized or uniform IDT), the −3 dB bandwidth $B_{-3\,\mathrm{dB}}$ is given by $B_{-3\,\mathrm{dB}} = 0.9\ v_o/L$, where L denotes the IDT length. The amplitude-frequency response then shows a sinc

response, where the first side lobe levels are only -13.5 dB down at $2.2B_{-3\,\mathrm{dB}}$ Hz distance from the center frequency of the filter.

For a -3 bandwidth of 180 kHz and using $v_o = 2827$ m/s, the minimum length of the IDT is 14 mm while a larger IDT is preferable for a higher out-of-band rejection and steeper filter slopes. The chip then becomes impractically long. For a maximum length of the chip of 10 mm, and so about 7 mm for the longest IDT, preliminary calculations, using an FFT program, show that a -3 dB bandwidth of 600 kHz can be obtained. The selectivity curve is shown in Fig. 4.13. The width of the device has to be chosen according to noise and

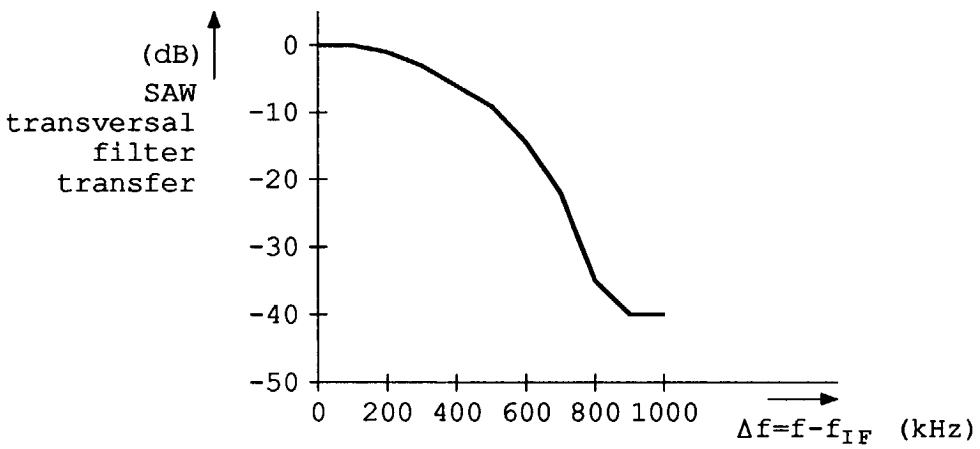

Fig. 4.13 The selectivity curve for 10 mm long transversal SAW filter.

dynamic range calculations, as will be done in Section 4.7. The final choice whether a SAW resonator filter or a SAW transversal filter will be applied, is left to the manufacturer.

A further important aspect is the reproducibility in fabrication and the temperature stability of the center frequency of the on-chip SAW filter. The definite answer concerning the manufacturing reproducibility of the filter center frequency can only be determined in a production line fabrication of these SAW devices.

For a maximum deviation of the filter center frequency from the intermediate frequency of 50 kHz, the drift in the center frequency should be smaller than 3 ppm/K over a 100 K temperature range. In Chapter 2 a value of -15 ppm/K to -18 ppm/K has been measured for the drift in the center frequency of a

100 MHz delay line. For a resonator filter with a 180 kHz bandwidth, this drift is too large. For a transversal filter with a bandwidth of 600 kHz (to comply with the 10 mm chip length), however, this drift is sufficiently low.

In general, the temperature drift is not sufficiently low for narrow-band filters, such as our application of a coupled resonator filter in the FM upconversion receiver. Counter measures are necessary to relieve the temperature stability demands and probably also the fabrication reproducibility demands. A short inventory of such measures will now be presented.

A possible solution consists of an on-line measurement of the center frequency of the filter. In reference [56] such a method has been presented, which is based on the minimization of the distortion that a pilot signal ($f = 90$ kHz), injected on the local oscillator, experiences in passing through the IF filter. This technique can only be applied for resonator type filters, because it is based on a symmetrical group delay as a function of the frequency, with a minimum of distortion at the filter center frequency. This generally is not the case in transversal filters.

An off-line measuring of the filter center frequency can be performed by temporarily inserting the IF filter in a measuring system. Disadvantages of this method are the use of several switches and the temporarily lack of signal at the audio output.

Compensation requires an extra SAW device on the same chip. The center frequencies of both devices should be accurately matched and track over the whole temperature range. In Chapter 2 it was already found that the difference in tracking of two SAW devices on the same chip is smaller than 20 kHz. The difference in the center frequencies of two devices, placed 6 mm apart, amounts to 170 kHz at 100 MHz, which should be much smaller for adjacent devices. A problem with this method is the acoustic crosstalk between two different SAW devices on the same chip. This may result in injection locking in the oscillator using the reference SAW device or oscillator signals injected in the IF part of the receiver. To minimize this acoustic crosstalk, the pass-bands of both devices must be chosen in such a way that they do not overlap one another. Such a difference in center frequency is only accurate if the dispersion in the ZnO-SiO_2-Si structure is sufficiently low.

Further research will have to show which of the above-mentioned counter measures to the problem associated with fabrication reproducibility and temperature stability leads to the simplest and best solution.

4.7 Noise and dynamic range calculations for SAW IF filters

4.7.1 Introduction

Both SAW transversal filters and SAW resonator filters can be used as IF filters in the FM upconversion receiver. For the SAW transversal filter, two configurations can be distinguished: First, the apodized IDT is used as the input IDT and the unapodized IDT is used as the output IDT. In the second situation, the transversal filter is used reversely, the unapodized IDT is used as the input IDT and the apodized IDT is used as the output IDT. As a coupled resonator filter is symmetrical, only one resonator filter configuration can be distinguished.

As has been shown in Chapter 2, the optimum signal transfer in a SAW transversal filter is obtained when the SAW transversal filter is driven by a voltage and the short-circuit current is sensed. The desired filter transfer function in a SAW resonator can only be obtained in the voltage-driven and current-sensed transfer, the transconductance. The noise and dynamic range calculations for the three configurations, two SAW transversal filter configurations and one SAW resonator filter configuration, will therefore only be performed for a voltage-driven and current-sensed SAW filter.

In all three configurations, the noise due to the amplifier that loads the SAW device and the noise generated by the SAW device will be transferred to an effective noise voltage in series with the voltage source that drives the SAW device. Only amplifiers with bipolar transistors will be assumed to be applicable at these frequencies. By choosing the proper bias current for the bipolar transistor in the input stage of the amplifier, the effective noise voltage is minimized. In addition, the dependency of the noise voltage on the aperture (width) of the SAW device, will be presented.

The driving current in the amplifier that drives the SAW device is limited. Therefore, the voltage at the input IDT of the SAW device is limited. The dynamic range, the ratio of the maximum voltage associated with a certain bias current and the minimum effective noise voltage at the input of the SAW device will be calculated for all three configurations. The dependency of the dynamic range on the aperture will be presented.

As our goal is the realization of the IF filter and the driving and loading amplifiers including their interconnection on the same chip, no inductors are allowed for tuning out the static IDT capacitances. At the end of this section a comparison with respect to noise and dynamic range for the three configurations will be presented.

4.7.2 Configuration 1

The calculation of the noise and the dynamic range for the SAW transversal filter configurations will be presented first. We will assume an acceptable chip length of 10 mm. The apodized IDT determining the bandwidth has been taken 7 mm long while the unapodized IDT is about 1 mm long. Therefore, the number of electrode pairs can be calculated as $N_1 = 400$ for the apodized IDT and $N_2 = 60$ for the unapodized IDT, where a filter center frequency of 160 MHz has been used. Split electrodes will be used for both IDTs to minimize the effect of acoustic wave reflections on the electrodes at the filter center frequency.

Because the number of electrode pairs in the uniform IDT is small, only the static capacitance of this IDT has to be taken into account. The number of electrode pairs in the apodized IDT is so large that the radiation conductance and the radiation susceptance are in the same order of magnitude as the admittance of the static capacitance. Thus, for the apodized IDT, both the static capacitance, the radiation conductance and the radiation susceptance have to be taken into account.

First we will discuss the configuration in which the apodized IDT (IDT_1) of the SAW transversal filter is driven by a voltage source U_s and the unapodized IDT (IDT_2) is loaded by an amplifier with a low input impedance (configuration 1). Figure 4.14 shows this configuration. The parameters of the SAW transversal filter within the filter pass-band can be approximated by:

$$Y_1(f) = G_1(f) + jB_1(f)\,, \tag{4.12a}$$

$$G_1(f) = G(f_o) \cdot \frac{N_1^2}{4}\,, \tag{4.12b}$$

$$B_1(f) = G(f_o) \cdot \frac{N_1^2}{4\pi} \cdot \ln \left| \frac{f - (f_0 + B/2)}{f - (f_0 - B/2)} \right|\,, \tag{4.12c}$$

$$C_1 = 0.6\,\sqrt{2}\,\epsilon a N_1\,, \tag{4.12d}$$

$$C_2 = \sqrt{2}\,\epsilon a N_2\,, \tag{4.12e}$$

$$G_{12}(f) = G_{21}(f) = G(f_o) \cdot \frac{N_1}{2} \cdot N_2\, \mathrm{e}^{-j\omega_o L_{12}/v_o}\,, \tag{4.12f}$$

$$G(f_o) = 8\kappa^2 f_o \epsilon a\,, \tag{4.12g}$$

where the loss α in the acoustic path has been neglected. The constant 0.6 in Equation 4.12d has been estimated for a sinc apodized IDT where only the main lobe is present. The parameter a denotes the aperture (width) of the SAW device, ϵ denotes the effective dielectric permittivity of the three-layered

ZnO-SiO_2-Si structure, κ^2 denotes the piezoelectric coupling constant of the three-layered ZnO-SiO_2-Si structure, v_o denotes the velocity of the surface acoustic wave, N_1 denotes the number of electrode pairs in the main lobe of the the sinc apodized IDT, N_2 denotes the number of electrode pairs in the unapodized IDT and L_{12} denotes the distance between the centers of the two IDTs.

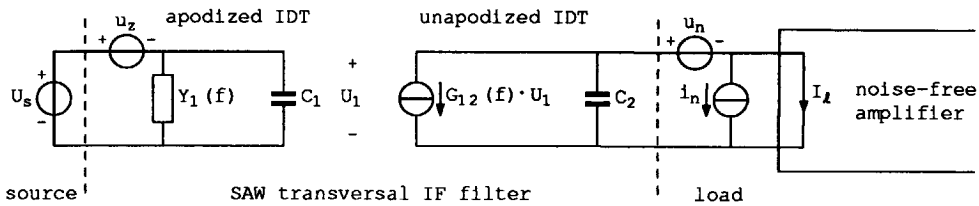

Fig. 4.14 Configuration 1, driving the apodized IDT and loading the unapodized IDT of a SAW transversal IF filter.

The noise sources of the amplifier are represented by the voltage source u_n and the current source i_n. The noise caused by the SAW transmission line terminations is represented by the voltage source u_z. The power-density spectra of these sources have been given in Section 4.4 (Equation 4.2a, Equation 4.2b and Equation 4.2c). The amplifier feedback network is assumed to be non-energic, thus contributing no noise. The loss in the transmission line is neglected.

Both noise voltage u_n and noise current i_n can be transferred to a noise voltage in series with the source voltage U_s. By following the same procedure as in Section 4.4, this total noise can be optimized with respect to r_e. The optimum r_e ($r_{e,\mathrm{opt1}}$) is given by:

$$r_{e,\mathrm{opt1}} = \frac{1}{\omega_o C_2 \sqrt{h_{\mathrm{FE}}}} \,. \tag{4.13}$$

The total noise voltage $u_{n,\mathrm{tot1}}$ in series with the source voltage U_s for this optimum value is given by:

$$u_{n,\mathrm{tot1}} = \frac{\sqrt{4kTB}}{\kappa^2 N_1} \cdot \sqrt{\frac{\pi^2 r_b}{2} + \frac{\pi\sqrt{2}}{8 f_o \epsilon a N_2 \sqrt{h_{\mathrm{FE}}}} + \frac{\kappa^2}{8 f_o \epsilon a}} \,. \tag{4.14}$$

The maximum voltage U_{max1} associated with a bias current I_{bias} in the output stage of the filter driving amplifier, for signals with frequencies within the filter pass-band, is given by:

$$U_{\text{max1}} = \frac{I_{\text{bias}}}{\sqrt{2} \cdot |Y_1(f) + j\omega_o C_1|_{\text{max}}} , \tag{4.15}$$

where the index max indicates the maximum value of the admittance $|Y_1(f) + j\omega_o C_1|$.

The radiation susceptance $B_1(f)$ of the admittance $Y_1(f)$ goes to infinity at the edges of the filter pass-band for a (theoretically) ideal filter. In practice, due to end effects, this ideal behavior cannot be realized. Besides, the apodized IDT is too short for a good approximation of this ideal behavior. Therefore, we will use the modified radiation susceptance $B_1'(f)$ as defined in Chapter 2, Section 2.3. Within the filter pass-band, the radiation susceptance can then be estimated by:

$$|B_1'(f)| \leq G(f_o) \cdot \frac{N_1^2}{4\pi} \cdot 6 , \tag{4.16}$$

where $b_1 = 200$ kHz, $b_2 = 700$ kHz, $\Delta b = 500$ kHz (see Section 4.6 and Chapter 2) have been used.

The maximum voltage U_{max1} can then be written as:

$$U_{\text{max1}} = \frac{I_{\text{bias}}}{\sqrt{2} \cdot \omega_o C_1 m_1} , \tag{4.17a}$$

$$m_1 = \sqrt{1 + 1.4 \cdot \kappa^2 N_1 + 0.65 \cdot [\kappa^2 N_1]^2} . \tag{4.17b}$$

The correction factor m_1 represents the influence of $Y_1(f)$ on the input admittance of the SAW device compared to $\omega_o C_1$. The ratio of the maximum voltage U_{max1} and the total noise voltage $u_{n,\text{tot1}}$ is then given by:

$$DR_1 = \frac{U_{\text{max1}}}{u_{n,\text{tot1}}} = \frac{\kappa^2 I_{\text{bias}}}{\sqrt{4kTB}} \cdot \left[280\, r_b f_o^2 \epsilon^2 a^2 m_1^2 + \frac{32\, f_o \epsilon a m_1^2}{N_2 \sqrt{h_{\text{FE}}}} + 7.1\, \kappa^2 f_o \epsilon a m_1^2 \right]^{-1/2} . \tag{4.18}$$

4.7.3 Configuration 2

The second configuration also concerns the SAW transversal filter. The unapodized IDT is now driven by a voltage source U_s and the apodized IDT is now loaded by an amplifier with a low input impedance (configuration 2). Figure 4.15 shows this configuration. The parameters of the SAW transversal filter within the filter pass-band have been given in Equation 4.12a–4.12g. The

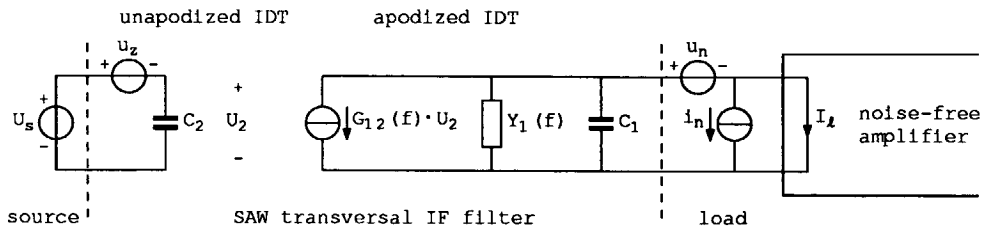

Fig. 4.15 Configuration 2, driving the unapodized IDT and loading the apodized IDT of a SAW transversal IF filter.

power-density spectra of the noise voltages u_n and u_z and the noise current i_n have been given in Equation 4.2a, Equation 4.2c and Equation 4.2b in Section 4.4. The amplifier feedback network is assumed to be non-energic, thus contributing no noise. The loss in the transmission line is neglected.

Both noise voltage u_n and noise current i_n can be transferred to a noise voltage in series with the source voltage U_s. The power-density spectrum of the total noise voltage $u_{n,\mathrm{tot2}}$ is then given by:

$$S(u_{n,\mathrm{tot2}}) = \frac{4kT}{|G_{21}(f)|^2} \cdot \left[(r_b + \frac{r_e}{2}) \cdot |Y_1(f) + j\omega C_1|^2 + \frac{1}{2h_{\mathrm{FE}} r_e} \right] + \frac{4kT}{G(f_o) N_2^2} \,. \tag{4.19}$$

Integrating this power-density spectrum from $f_o - B/2$ to $f_o + B/2$, where f_o represents the carrier frequency of the signal and B represents the signal bandwidth, using $f_o \gg B$ and numerically evaluating the integrals involving $B_1(f)$, yields:

$$\begin{aligned} u_{n,\mathrm{tot2}}^2 &= \frac{4kTB}{|G_{21}(f_o)|^2} \cdot \left[(r_b + \frac{r_e}{2}) \cdot \omega_o^2 C_1^2 \cdot m_2 + \frac{1}{2h_{\mathrm{FE}} r_e} \right] \\ &\quad + \frac{4kTB}{G(f_o) N_2^2} \,, \end{aligned} \tag{4.20a}$$

$$m_2 = 1 + 0.19 \left[\kappa^2 N_1 \right]^2 \,. \tag{4.20b}$$

The correction factor m_2 represents the influence of $Y_1(f)$ on the total noise voltage compared to $\omega_o C_1$.

A minimum value of $u_{n,\mathrm{tot2}}^2$ can be obtained by differentiating $u_{n,\mathrm{tot2}}^2$ with respect to r_e, hence an optimum value for r_e ($r_{e,\mathrm{opt2}}$) is found:

$$r_{e,\mathrm{opt2}} = \frac{1}{\omega_o C_1 \sqrt{m_2} \sqrt{h_{\mathrm{FE}}}} \,. \tag{4.21}$$

The noise voltage in series with the source voltage U_s for this optimum value is given by:

$$u_{n,\mathrm{tot2}} = \frac{\sqrt{4kTB}}{\kappa^2 N_2} \cdot \sqrt{1.8\, r_b m_2 + \frac{0.33\sqrt{m_2}}{f_o \epsilon a N_1 \sqrt{h_{\mathrm{FE}}}} + \frac{\kappa^2}{8 f_o \epsilon a}} \,. \tag{4.22}$$

In the same way as in configuration 1, the current delivered by the amplifier, represented by U_s, has a maximum value. The maximum voltage U_{max2} associated with a bias current I_{bias} in the output stage of the filter driving amplifier, for signals with frequencies within the filter pass-band, is given by:

$$U_{\mathrm{max2}} = \frac{I_{\mathrm{bias}}}{\sqrt{2} \cdot |j\omega_o C_2|} \,. \tag{4.23}$$

The ratio of the maximum voltage U_{max2} and the total noise voltage $u_{n,\mathrm{tot2}}$ is then given by:

$$DR_2 = \frac{U_{\mathrm{max2}}}{u_{n,\mathrm{tot2}}} = \frac{\kappa^2 I_{\mathrm{bias}}}{\sqrt{4kTB}} \cdot \left[280\, r_b f_o^2 \epsilon^2 a^2 m_2 + \frac{52\, f_o \epsilon a \sqrt{m_2}}{N_1 \sqrt{h_{\mathrm{FE}}}} + 2\pi^2 \kappa^2 f_o \epsilon a \right]^{-1/2} . \tag{4.24}$$

4.7.4 Configuration 3

The third configuration consists of two SAW resonator filters which are coupled to each other. The noise and dynamic range will only be calculated for the situation in which no external capacitors are used.

As a consequence of the short IDTs within the resonator cavity we may neglect the radiation admittances compared to the admittances of the static capacitance of the IDTs. Split electrodes will be used to minimize the effect of acoustic wave reflections at the filter center frequency.

The configuration is shown in Fig. 4.16. The relevant parameters of the SAW resonator filter, in the case that the IDTs are placed on the maxima of the standing wave pattern in the resonator cavity, are given by:

$$C_1 = \sqrt{2}\, \epsilon a N_1 \,, \tag{4.25a}$$

$$C_2 = \sqrt{2}\, \epsilon a N_2 \,, \tag{4.25b}$$

$$R_s = \frac{1}{G(f_o) N_1 N_2} \cdot \frac{1 - \Gamma_o}{1 + \Gamma_o} \,, \tag{4.25c}$$

$$G(f_o) = 8\kappa^2 f_o \epsilon a \,, \tag{4.25d}$$

$$G_s(f) = \frac{I_l}{U_s} = \frac{f_o}{jfR_s} \cdot \frac{kQ}{(1+j\nu Q)^2 + (kQ)^2(f/f_o)^2}, \tag{4.25e}$$

$$kQ = \frac{1}{\omega_o R_s 2C_2}, \tag{4.25f}$$

$$\nu = \frac{f}{f_o} - \frac{f_o}{f}, \tag{4.25g}$$

$$Q = \frac{f_o}{B_e}. \tag{4.25h}$$

The parameter Γ_o denotes the effective reflection factor of the reflection grating, kQ denotes the coupling factor of the two resonators, ν denotes the detuning, Q denotes the quality factor of a single resonator and B_e denotes the -3 dB bandwidth of a single resonator (see Chapter 2).

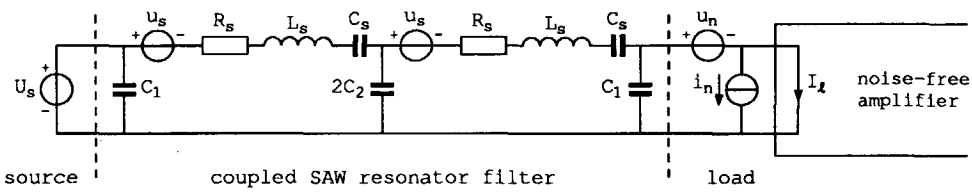

Fig. 4.16 Configuration 3, a coupled SAW resonator filter as IF filter.

The noise sources of the amplifier are represented by the voltage source u_n and the current source i_n. The power-density spectra of these sources have been given in Section 4.4 (Equation 4.2a and Equation 4.2b). The amplifier feedback network is assumed to be non-energic, thus contributing no noise. The resonator noise is represented by the voltage sources u_s. The power-density spectra of these voltage sources are given by:

$$S(u_s) = 4kTR_s \tag{4.26}$$

All noise sources can be transferred to a noise voltage in series with the source voltage U_s. The power-density spectrum of this noise voltage $u_{n,\mathrm{tot3}}$ is given by:

$$\begin{aligned} S(u_{n,\mathrm{tot3}}) = 4kT \{ & R_s \left[1 + |2j\omega R_s C_2(1+j\nu Q)|^2\right] \\ + & (r_b + \frac{r_e}{2}) \frac{f^4 C_1^2}{4f_o^4 C_2^2 k^4 Q^4} \left|(1+j\nu Q)^2 - \frac{2jkQf_oC_2}{fC_1}(1+j\nu Q) + \frac{k^2Q^2f_o^2}{f^2}\right|^2 \\ + & \frac{1}{2h_{\mathrm{FE}}r_e} \cdot \frac{1}{|G_s(f)|^2} \} . \end{aligned} \tag{4.27}$$

Integrating this power-density spectrum from $f_o - B/2$ to $f_o + B/2$, where f_o represents the carrier frequency of the signal and B represents the signal bandwidth and using $f_o \gg B$ leads to:

$$u_{n,\text{tot3}}^2 = 4kTB \cdot \left\{ R_s \left[1 + \frac{1}{k^2Q^2}(1 + \frac{B^2}{3B_e^2}) \right] + \left(r_b + \frac{r_e}{2} \right) m_3 + \frac{R_s^2}{2h_{\text{FE}} r_e} m_4 \right\} , \quad (4.28a)$$

$$m_3 = \left(\frac{N_1}{2N_2} \right)^2 \cdot \frac{1}{k^4Q^4} \cdot \left\{ \left(1 + k^2Q^2 \right)^2 + \frac{B^4}{5B_e^4} + \frac{2B^2}{3B_e^2} \left[1 - k^2Q^2 + \frac{k^2Q^2N_2^2}{2N_1^2} \right] + k^2Q^2 \left(\frac{2N_2}{N_1} \right)^2 \right\} , \quad (4.28b)$$

$$m_4 = \frac{1}{k^2Q^2} \left\{ \left(1 + k^2Q^2 \right)^2 + \frac{2B^2}{3B_e^2}(1 - k^2Q^2) + \frac{B^4}{5B_e^4} \right\} , \quad (4.28c)$$

A minimum value of $u_{\text{n,tot3}}^2$ can be obtained by differentiating $u_{\text{n,tot3}}^2$ with respect to r_e, hence an optimum value for r_e ($r_{e,\text{opt3}}$) is found:

$$r_{e,\text{opt3}} = \frac{R_s}{\sqrt{h_{\text{FE}}}} \cdot \sqrt{\frac{m_4}{m_3}} . \quad (4.29)$$

The minimum noise voltage $u_{n,\text{tot3}}$ in series with the source voltage U_s, for this optimum value is then given by:

$$u_{n,\text{tot3}} = \sqrt{4kTB} \cdot \sqrt{R_s \left[1 + \frac{1}{k^2Q^2}(1 + \frac{B^2}{3B_e^2}) + \sqrt{\frac{m_3 m_4}{h_{\text{FE}}}} \right] + r_b m_3} . \quad (4.30)$$

The dependency of $u_{n,\text{tot3}}$ on the aperture can be found by substituting Equation 4.25c in the above equation.

As in the previous configurations, the current delivered by the amplifier, represented by U_s, has a maximum value. The maximum voltage U_{max3} associated with a bias current I_{bias} in the output stage of the filter driving amplifier, for signals with frequencies within the filter pass-band, is given by:

$$U_{\text{max3}} = \frac{I_{\text{bias}}}{\sqrt{2}} \cdot R_s \cdot m_5 , \quad (4.31a)$$

$$m_5 = kQ \cdot \frac{2N_2}{N_1} \cdot \sqrt{\frac{4k^2Q^2 + 1}{(k^4Q^4 + k^2Q^2)(2N_2/N_1)^2 + 2k^2Q^2(2N_2/N_1 + 2) + 1}} , \quad (4.31b)$$

where $R_s m_5$ is an approximation of the (minimum) input impedance of the coupled resonator filter at the frequency $f = 1/(2\pi\sqrt{L_s C_s})$. The ratio of the maximum voltage $U_{\max 3}$ and the total noise voltage $u_{n,\mathrm{tot}3}$ is then given by:

$$DR_3 = \frac{U_{\max 3}}{u_{n,\mathrm{tot}3}} = \frac{m_5 I_{\mathrm{bias}}}{\sqrt{2}\sqrt{4kTB}} \cdot \left[\frac{1}{R_s}\left\{1 + \frac{1}{k^2Q^2}(1 + \frac{B^2}{3B_e^2}) + \sqrt{\frac{m_3 m_4}{h_{\mathrm{FE}}}}\right\} + \frac{r_b m_3}{R_s^2}\right]^{-1/2} . \quad (4.32)$$

4.7.5 Comparison and conclusions

We will now compare the noise voltages $u_{n,\mathrm{tot}1}$, $u_{n,\mathrm{tot}2}$ and $u_{n,\mathrm{tot}3}$ for all three configurations. The following values for the parameters are used: $\kappa^2 = 0.0089$, $v_o = 2827$ m/s, $\epsilon = 135$ pF/m, $r_b = 10\ \Omega$, $h_{\mathrm{FE}} = 100$, $B = 180$ kHz, $f_o = 160$ MHz, $kQ = 1/\sqrt{3}$ and $B/B_e = 0.908$. With $\Gamma_o = 0.9$ it follows that $N_1 = 8$ and $N_2 = 50$ for the number of electrode pairs in the coupled resonator for a nearly filled cavity (only the static IDT capacitances are used as coupling capacitances). The correction factors for the noise voltages are then: $m_2 = 3.4$ and $m_3 = 3.3$, $m_4 = 6.8$.

The noise voltages are drawn in Fig. 4.17 as a function of the aperture a. From Fig. 4.17 it can be concluded that the aperture has no influence on the noise due to the transistor base resistance. This is caused by the fact that both transconductance and output admittance of the filters are proportional to the aperture.

The aperture should be chosen large to minimize the noise contribution of the base current shot noise, the collector current shot noise and the noise of the SAW transmission line terminations (in the case of the SAW transversal filter) or the resonator noise (in the case of the SAW resonator filter). The noise voltage is inversely proportional to the square root of the aperture for these noise sources. In comparison to the situation with the RF filter, as mentioned in Section 4.4, the IF transversal filter in configuration 1 does not have to be very wide to obtain a low equivalent noise voltage at the input of the SAW device because the SAW device is very long. The noise voltage in configuration 1 is lower than that in configuration 2 because the output impedance of the SAW device in configuration 1 is much higher.

With the use of Fig. 4.17, the necessary gain of the stages preceding the IF filter can be determined for each configuration. Not only the noise is then of interest but also the dynamic range of the configuration, which will be compared next for the three configurations.

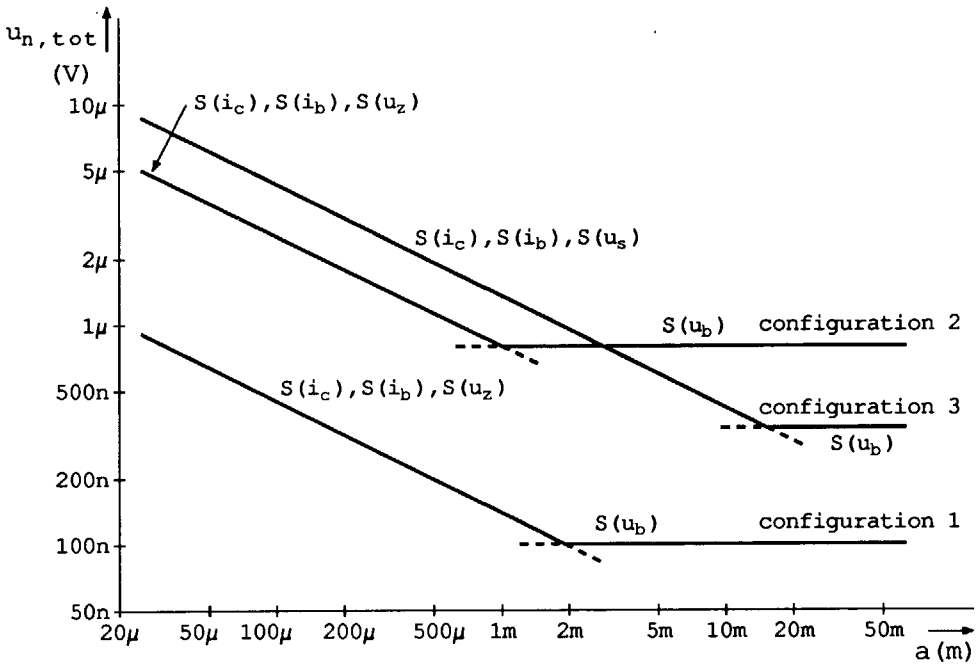

Fig. 4.17 The noise voltages as a function of the aperture for the three configurations.

For the calculation of the dynamic ranges DR_1, DR_2 and DR_3, a bias current of $I_{\text{bias}} = 5$ mA has been used. The correction constants are then $m_1 = 3.8$ and $m_5 = 1.2$. The dynamic ranges are drawn in Fig. 4.18 as a function of the aperture a. The aperture should be as small as possible for a maximal dynamic range. This can be understood by the fact that both input admittance, output admittance and transconductance of the SAW devices are proportional to the aperture.

The dynamic range is inversely proportional to the aperture when the noise of the transistor base resistance dominates. The dynamic range is inversely proportional to the square root of the aperture when the base current shot noise, collector current shot noise and the noise generated by the SAW transmission line terminations (in the case of the SAW transversal filter) or the resonator noise (in the case of the SAW resonator filter) dominate. For small apertures, the dynamic range of configuration 2 is about 10 dB larger com-

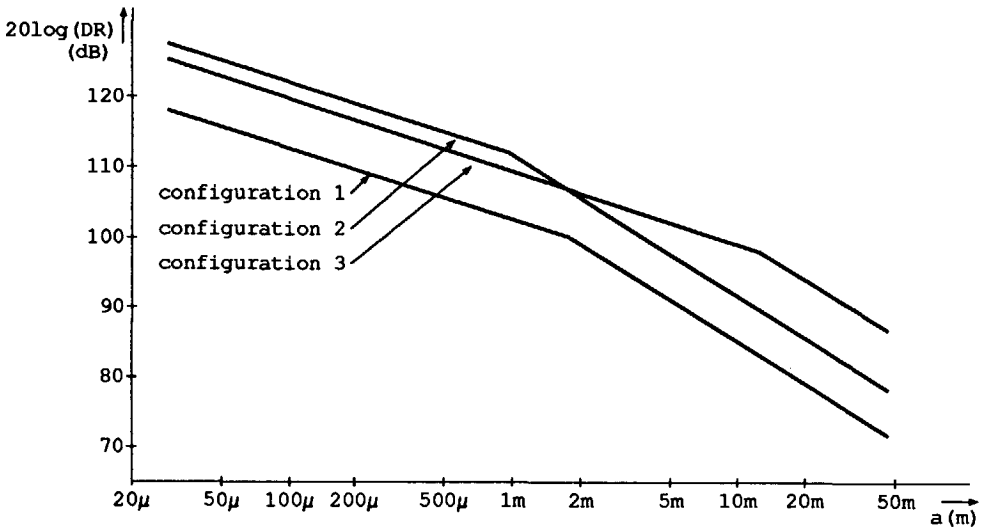

Fig. 4.18 The dynamic ranges as a function of the aperture for the three configurations.

pared to the dynamic range of configuration 1. Note that in the resonator configuration (configuration 3), the aperture for a single resonator is used. A coupled resonator will be twice as wide.

The noise voltages and the dynamic ranges as shown in Fig. 4.17 and Figure 4.18, respectively, are essential for the determination of the gain distribution in the receiver front-end. An example of this gain distribution will be worked out in the following section.

Some comments have to be given on the dynamic range and noise calculations. First of all, the noise optimization is only performed for noise within the filter pass-band. For a selective FM detector, such as a phase-lock loop, this optimization is sufficient. When a wide-band FM detector is used, the out-of-band noise should also be considered.

A similar remark must be made for the consideration regarding the calculations for the maximum signal level. For out-of-band signals with large amplitudes, the equations should be revised. It should also be kept in mind that for a very large voltage $U_{\max}$ the power supply voltage is the restricting

factor and not the bias current.

In the calculation of the maximum voltage, the current through the feedback network of the amplifier that drives the SAW device, has been neglected. In practice, this current may also be a substantial part of the total current, especially when the input impedance of the SAW device becomes very high.

The aperture of the SAW filter cannot be made extremely large or extremely small. The lower limit is determined by the fact that for a SAW device with a small aperture, the IDTs behave as a point source. To avoid this effect, usually a minimum aperture of $40\lambda_o$ is taken. Further research to determine whether any corrections in the apodized IDT can counteract this effect, are necessary. The upper limit for the aperture is determined by the maximum allowable chip area or chip housing.

The noise sources were only considered for low frequencies. For higher frequencies ($f > f_T/\sqrt{h_{FE}}$), the collector current shot noise adds substantially to the power-density spectrum of the amplifier equivalent input noise current source i_n. The noise and dynamic range equations should be modified when the low frequency assumption is no longer valid.

4.8 Miscellaneous front-end considerations and gain distribution

First some general front-end considerations will be treated briefly. All circuits between the antenna and the IF filter must have a dynamic range of at least 106 dB and an intermodulation-free dynamic range of 80 dB, because a dynamic range reduction can only take place in the circuits following the IF filter.

A mixer is used to perform the frequency conversion of signals from the FM broadcast band to signals at the intermediate frequency. A bipolar current switching mixer is the best circuit as far as noise and distortion are concerned [1]. The noise and distortion analysis have been presented in references [1] and [2]. A double-balanced mixer configuration is preferable to minimize the feedthrough of LO signals and RF signals. Both input signals and output signals of the mixer are (balanced) currents.

The input amplifier should have an equivalent input noise level of 1 μV (B = 180 kHz). A maximum input level of at least 200 mV_{eff} for linear operation is desirable. Further, this amplifier has to terminate the LC RF filter at the input of the receiver. The output signal of the amplifier must be a (balanced) current for driving the double-balanced bipolar switching mixer. Amplifiers employing two feedback loop are to be preferred for a low-noise, linear termination of the filter and driving the bipolar switching mixer with a current. Such an amplifier will be presented in Chapter 5.

The IF filter is either a SAW transversal filter or a SAW resonator filter.

For a maximum suppression of electromagnetic feedthrough in these filters, and thus a maximum out-of-band rejection, these filters should be driven by a balanced amplifier with a low output impedance and loaded by a balanced amplifier with a low input impedance. The amplifier, which drives the SAW IF filter, must therefore be a balanced transimpedance amplifier as the input signal of this amplifier is the (balanced) output current of the bipolar switching mixer.

Amplifier 3 in Fig. 4.1 loads the SAW IF filter and drives succeeding stages for the detection of the FM signal. Dependent on the type of load, this amplifier must deliver either a voltage or a current to the load. A balanced amplifier is preferable for a minimization of the electromagnetic feedthrough in the SAW device. An amplifier with a low (differential-mode) input impedance is desirable for an optimum signal transfer in the SAW device. The noise optimization criteria for this amplifier have been discussed in Section 4.7. When a channel-selective filter is used as the IF filter, even a (symmetric) limiter can be used instead of amplifier 3, provided that the noise optimization is not degraded.

By using the noise and dynamic range equations and figures in Section 4.7, the gain distribution over the front-end can be calculated. For an extensive discussion of the contribution of the mixer to the noise, the reader is referred to reference [1]. Only the situation with the coupled SAW resonator filter used as an IF filter will be discussed here. The configurations with the SAW transversal filters can be treated in a similar way.

From Fig. 4.18 it is clear that the aperture (the width of the SAW device) should be chosen as small as possible for a maximum dynamic range. For a filter center frequency of 160 MHz, it follows that $\lambda_o = 17.7$ μm. To prevent the IDTs behaving as point sources, a minimum aperture of $40\lambda_o = 0.70$ mm is recommended. The dynamic range is then DR = 110 dB and the equivalent noise voltage at the input of the IF filter amounts to $u_n = 1.4$ μV. A maximum utilization of the receiver power consumption is obtained if this noise voltage amounts to a noise voltage of about 1 μV at the input of the receiver. Therefore, the voltage gain from input of the receiver to filter input voltage should be 1.4. The maximum filter voltage is then 450 mV_{eff}.

4.9 Discussion

In this chapter the architecture of the FM upconversion receiver front-end has been discussed. The choice for the intermediate frequency has been made on four different grounds. For a maximum image and LO harmonics-related spurious suppression, the intermediate frequency should be chosen as high as

possible. The gain of the antenna increases slightly with increasing frequency. When non-selective FM detectors are used, the second pass-band in the IDT filter, due to a second-order Rayleigh-wave mode, may pass signals that trouble the detection of the desired signal. The amplitude of the signals that pass through the $1.7f_{IF}$ pass-band can be minimized by either choosing an intermediate frequency higher than 330 MHz or in the vicinity of 140 MHz. The first choice ($f_{IF} > 330$ MHz) puts impractical demands on the electronic circuitry, the temperature stability of the IF filter and on the filter manufacture reproducibility. The second choice for the intermediate frequency ($f_{IF} = 140$ MHz) is not yet optimal with respect to the feedthrough of RF signals in the input filter, RF amplifier and mixer. Hence, an intermediate frequency of 160 MHz has been chosen. To attenuate the signals which might pass through the extra IF filter pass-band, a band-pass filter should be used at the input of the receiver.

Both SAW transversal filters and simple LC filters have been investigated with respect to applicability as RF filter at the input of the receiver. Noise calculations show that a sensitivity of 1 μV can, in practice, hardly be reached with a SAW transversal filter. Besides, a SAW filter with a bandwidth of 20 MHz would have to become very wide if a sensitivity in the order of several microvolts is to be reached. A simple LC filter, external to the chip without any increase in IC pins, provides a suitable alternative. The desired sensitivity can be obtained, and simultaneously the filter transfer function can be made independent of the varying antenna impedance by the appropriate termination of the filter, which must be designed accordingly.

Both on-chip SAW transversal filters and on-chip SAW resonator filters can be applied as IF filter. A SAW transversal filter of sufficient length (longer than 14 mm) gives the best selectivity and the lowest (ideally no) distortion of the desired information. A coupled SAW resonator filter exhibits less selectivity and distorts the desired information, but is shorter than a SAW transversal filter. The ultimate choice has to be based on a trade-off between production costs and the desired specifications.

The reproducibility in fabrication and the temperature stability of the center frequency of the on-chip SAW filters are, at the moment, not good enough for the application as IF filter in the FM upconversion receiver. Further research on these aspects and/or techniques such as compensation, are essential for a proper functioning of the FM upconversion receiver.

By using models for the SAW devices, as presented in Chapter 2, the effective noise voltage at the input of the SAW device, due to noise of the amplifier that loads this device and the noise generated by the SAW device, has been calculated. Criteria for the optimization of the bias current of the bipolar

transistor at the input stage of the amplifier that loads the SAW device, have been given. The two possible ways of configuring the SAW transversal filter, driving the unapodized IDT and loading the apodized IDT and vice versa, give essentially different effective noise voltages at the input of the SAW device, due to different output impedances of both configurations. The contributions of the base current shot noise, the collector current shot noise and the noise added by the SAW device can be minimized by maximizing the width of the SAW device.

Driving the SAW device requires a certain bias current in the output stage of the amplifier used for this purpose. The dynamic range, the ratio of the maximum voltage associated with this bias current and the effective noise voltage at the input of the SAW device, has been calculated for all three configurations. By minimizing the aperture of the SAW devices, the dynamic range can be maximized or in other words, the demands posed on the electronic circuitry can be relieved within the restrictions of the available IC process and supply current. The gain of the stages preceding the SAW filter can then be chosen in such a way that the desired dynamic range of the whole front-end is reached.

5. A new class of balanced dual-loop amplifiers

5.1 Introduction

The front-end of the FM upconversion receiver consists of a RF filter followed by an amplifier, of which the output signal is delivered to the mixer (see Chapter 4). It was concluded that a band-pass filter consisting of inductances and capacitances, external to the chip, gives the best performance with respect to the sensitivity of the receiver while simultaneously all undesired responses are suppressed. The filter should then be terminated correctly, without introducing extra noise or distortion. Further, an amplifier is required to transfer signals (voltage, current or a combination of both) from the output of the filter to a current for the bipolar switching mixer.

A very simple solution to terminating the filter is the "brute-force" method: the filter is terminated with a resistor and the voltage across the resistor or the current through the resistor is amplified. As the output signal of the amplifier must be a current for the mixer, a transadmittance or current amplifier, respectively, have to be employed. However, this brute-force method degrades noise performance: the noise figure of the total filter-amplifier configuration is at least 3 dB [17], [57].

A more sophisticated method is the use of multi-loop amplifiers [17], [57] [58]. By employing multiple feedback networks, the input impedance of the amplifier can be made linear and accurately known, while the noise added by the feedback loops can be minimized. The resulting noise figure may be well below 3 dB. Some specific examples of these dual-loop amplifiers will be treated in Section 5.2. The practical realizations of these dual-loop amplifiers reveal several disadvantages.

In Section 5.3 two new amplifier configurations will be presented. Most of the disadvantages have been surmounted. The key to this new technique is the use of two balanced amplifier circuits with two overall feedback loops. Similar techniques can be used to realize amplifiers with linear, accurately-known input impedances and low output impedances. In addition, amplifiers with either high or low input impedances and linear, accurately-known output impedances can also be realized.

In Section 5.4 the common-mode behavior of the new, balanced dual-loop amplifier configurations will be discussed. In Section 5.5 a practical realization of one of these new amplifier configurations will be presented, which will be directly applicable as the input amplifier in the front-end of FM receivers.

Finally, Section 5.6 concludes this chapter with a discussion. This chapter contains material previously published given in references [57], [59], [60] and [61].

5.2 Definitions and standard configurations

In this section we will discuss the standard dual-loop amplifier configurations that have a linear, accurately-known input impedance and simultaneously have a high output impedance. The amplifier configurations will be described on the basis of classical two-port theory in terms of transmission parameters. These parameters, therefore, are defined first. In Fig. 5.1 a two-port is shown. The

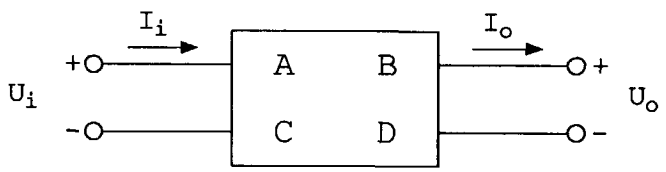

Fig. 5.1 A two-port.

general form of the transmission matrix of the two-port is given by:

$$\begin{bmatrix} U_i \\ I_i \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} U_o \\ I_o \end{bmatrix}, \tag{5.1}$$

with A, B, C and D as transmission parameters. The transfer parameters μ, γ, ζ and α of the two-port are the reciprocal values of the transmission parameters and can be calculated as follows:

$$\text{voltage-gain factor} \qquad \mu = \frac{1}{A} = \left[\frac{U_o}{U_i}\right]_{I_o=0}, \tag{5.2a}$$

$$\text{transadmittance} \qquad \gamma = \frac{1}{B} = \left[\frac{I_o}{U_i}\right]_{U_o=0}, \tag{5.2b}$$

$$\text{transimpedance} \qquad \zeta = \frac{1}{C} = \left[\frac{U_o}{I_i}\right]_{I_o=0}, \tag{5.2c}$$

$$\text{current-gain factor} \qquad \alpha = \frac{1}{D} = \left[\frac{I_o}{I_i}\right]_{U_o=0}. \tag{5.2d}$$

When the source impedance Z_s and the load impedance Z_l are known, then the input impedance Z_i and the output impedance Z_o of the two-port are given,

respectively, by:

$$Z_i = \frac{AZ_l + B}{CZ_l + D}, \tag{5.3a}$$

$$Z_o = \frac{B + DZ_s}{A + CZ_s}. \tag{5.3b}$$

The active parts of the amplifiers to be mentioned will be modeled as nullors. All transmission parameters of a nullor are zero; all transfer parameters are infinite.

The best theoretical solution to the problem of terminating a filter with a low-noise, linear, accurately-known impedance and driving the load with a current, independent of the load impedance, is by fixing two of the four transfer parameters of a nullor by applying two feedback loops. For this application the transmission parameters B and D should be fixed. This means that the output current must be converted into a voltage, and this voltage is compared with the input voltage, thus fixing B. Also the output current must be compared with the input current, thus fixing D. Figure 5.2 shows such an amplifier, the transformer fixes parameter D and the gyrator fixes parameter B [17]. The parameters are given in Equations 5.4a–5.4f.

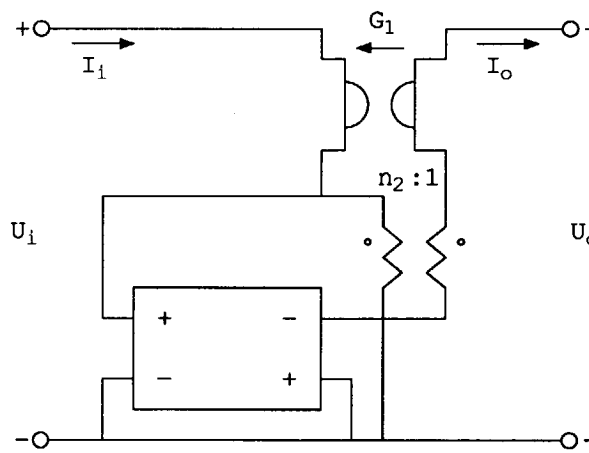

Fig. 5.2 Theoretical power-to-current amplifier.

$$A = 0. \tag{5.4a}$$

$$B = \frac{1}{G_1} . \tag{5.4b}$$

$$C = 0 . \tag{5.4c}$$

$$D = \frac{1}{n_2} . \tag{5.4d}$$

$$Z_i = \frac{n_2}{G_1} . \tag{5.4e}$$

$$Z_o = \infty . \tag{5.4f}$$

Because both the input voltage and the input current are converted into an output current, these configurations are called power-to-current amplifiers. Although gyrators can be approximated by a certain combination of active devices and resistances to fix the gyrator constants, the realization is very often impractical at high frequencies. A practical approximation of the gyrator is a combination of a transformer and a resistor [17]. Transformers are only practical devices in limited bandwidth situations and are not suitable for monolithic integration. In spite of their impracticality for integration, the amplifiers constructed with ideal transformers and ideal gyrators theoretically give the best result with respect to noise and distortion behavior.

In Fig. 5.3 a power-to-current amplifier with indirect feedback is sketched [17]. The active part AP is a nullor. All transfer parameters of this active part are stabilized by the two feedback loops. The transmission parameters A and C from the amplifier parts I and II are unimportant since the output voltage of amplifier parts I and II is zero. Parameter D should be equal to zero while parameter B of amplifier part I and II should be matched and have ratio $1 : 1/n$. The parameters of the amplifier are given in Equations 5.5a–5.5f.

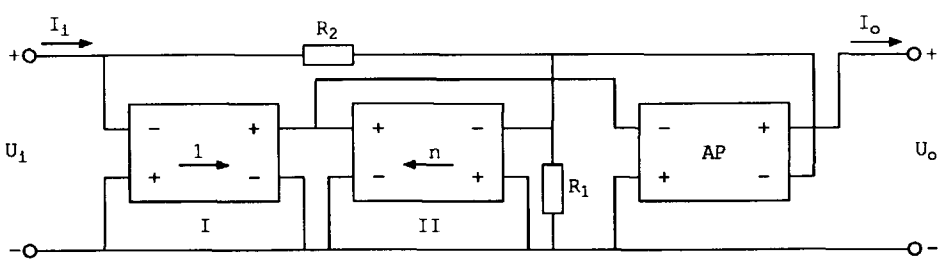

Fig. 5.3 A power-to-current amplifier with indirect feedback.

$$A = 0 . \tag{5.5a}$$

$$B = \frac{R_1 R_2}{(n+1)R_1 + R_2} . \tag{5.5b}$$

$$C = 0 . \tag{5.5c}$$

$$D = \frac{(n+1)R_1}{(n+1)R_1 + nR_2} . \tag{5.5d}$$

$$Z_i = \frac{R_2}{n+1} . \tag{5.5e}$$

$$Z_o = \infty . \tag{5.5f}$$

Parameters B and D are fixed by R_1, R_2 and the device ratio n. Any mismatch in parts I and II represents itself as a non-ideal transfer in the feedback loop and therefore in the amplifier. The errors caused by the mismatch cannot be decreased by increasing the loop gain of the amplifier since the mismatch is part of the feedback network.

A power-to-current amplifier with an active feedback network is shown in Fig. 5.4 [17]. The arrows in the nullors indicate the direction of signal transfer in the nullors. In this configuration, amplifier part I is stabilized by local feedback (R_3). The transimpedance so constructed is used as part of the feedback network in the amplifier. Together with the remaining resistances R_1 and R_2, the parameters B and D are fixed as given in Equations 5.6a–5.6f.

$$A = 0 . \tag{5.6a}$$

$$B = \frac{R_1 R_3}{R_1 + R_2 + R_3} . \tag{5.6b}$$

$$C = 0 . \tag{5.6c}$$

$$D = \frac{R_2 + R_3}{R_1 + R_2 + R_3} . \tag{5.6d}$$

$$Z_i = \frac{R_1 R_3}{R_2 + R_3} . \tag{5.6e}$$

$$Z_o = \infty . \tag{5.6f}$$

Thus if the active part AP is a nullor, the transfer is determined by the active feedback network. In that case distortion, due to the active devices in this feedback network, determines the distortion behavior of the amplifier. Noise contributions from the active feedback network degrade the dynamic range even further.

So far only solutions have been shown where the power-to-current transfer is realized in a single amplifier. It is also possible to use a cascade of amplifiers. The first amplifier can then be optimized to terminate the source. A dual-loop,

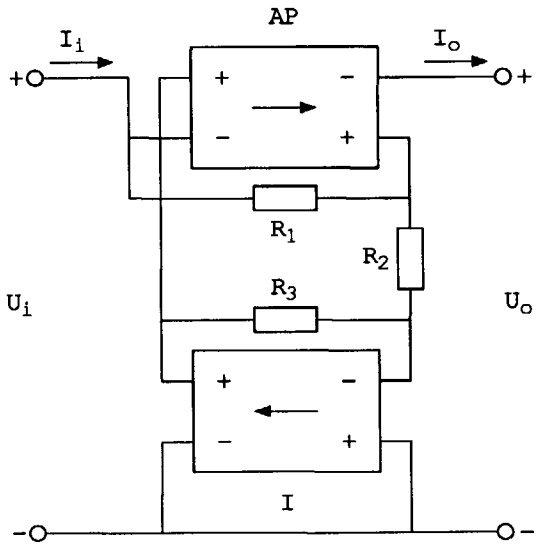

Fig. 5.4 A power-to-current amplifier with active feedback.

a three-loop or a four-loop amplifier should be used to in order to avoid a brute-force termination of the source. The last amplifier can be optimized to deliver the desired current to the load. In general, this cascade of amplifiers consumes more power to obtain an intermodulation-free dynamic range comparable with that of a a single amplifier. A cascade of amplifiers is then only preferable if a very high gain and a very large bandwidth have to be achieved.

As has been shown in this section, there are four possibilities for realizing a power-to-current transfer. All these amplifier configurations have several disadvantages. In the next section two new power-to-current amplifier configurations with impedance feedback will be presented; the mentioned problems of the standard configurations have been surmounted. In addition, the new technique used for these two new configurations can also be used to construct two power-to-voltage, two current-to-power and two voltage-to-power amplifiers as well.

5.3 New amplifier configurations

5.3.1 New dual-loop amplifier configurations

In the previous section, several known solutions to the problem of terminating a source and driving the load with a current were given. All of the configurations described are unbalanced amplifiers. By using balanced amplifiers, not only an "in-phase" signal but also an "anti-phase" (inverse) signal is present. This extra signal gives the possibility of realizing a dual-loop amplifier with impedance feedback. Figure 5.5 depicts a basic configuration of such an amplifier. The voltage U_i is amplified to $[1+(R_{3a}+R_{3b})/R_4]\cdot U_i$, in the same way as it is done in a balanced voltage amplifier. The input current I_i is sensed by resistors R_{2a} and R_{2b} and the voltage drop over R_{2a} and R_{2b} is compared with the input voltage U_i and the amplified input voltage $[1+(R_{3a}+R_{3b})/R_4]\cdot U_i$. The input voltage is sensed by R_4, R_{3a} and R_{3b}. The output current I_o is the sum of current through resistors R_{2a} and R_{3b} or R_{2b} and R_{3a}.

The cross-coupling of resistors R_{2a} and R_{2b} in Fig. 5.5a and the cross-coupling of R_{3a} and R_{3b} in Fig. 5.5b is necessary to maintain negative feedback for differential-mode signals. The transfer parameters can be calculated by using the definitions given in Section 5.2. The parameters in Equations 5.7a–5.7f are given as differential mode parameters (index d). Which of the two amplifiers drawn in Fig. 5.5 yields the best results with respect to noise, biasing and/or high-frequency behavior with practical devices instead of nullors depends on the application. In Section 5.4 a difference with respect to biasing will be indicated. In Section 5.5 the noise behavior will be treated.

Although in Fig. 5.5 only resistors are used in the feedback network, inductors and capacitors can also be employed. For example, in a limited bandwidth, an inductive input impedance is achieved by connecting a capacitor in series or in parallel with R_{3a} and R_{3b}.

In the light of the classification mentioned in references [17] and [58], it must also be possible with a similarly balanced structure to realize an amplifier with a linear, accurately-known input impedance and a low output impedance. As both the input voltage and the input current are converted into an output voltage, these configurations are called power-to-voltage amplifiers. Fig. 5.6 shows the two versions of the amplifier. As can be seen by comparing Fig. 5.5 and Fig. 5.6, the basic structure of the scheme is the same. The only difference is the way the load is connected. The parameters are given in Equations 5.8a–5.8f.

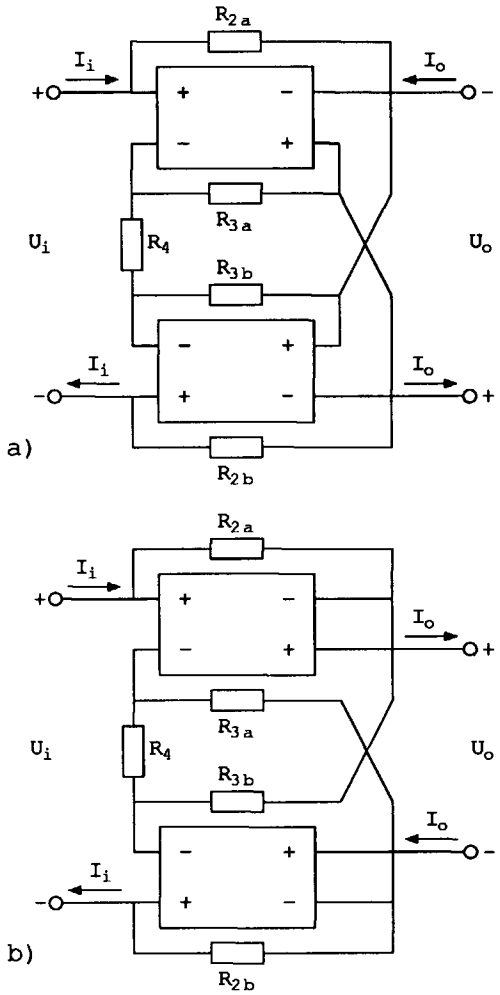

Fig. 5.5 The new power-to-current amplifiers.

$$\mu_{\mathrm{d}} = \infty . \tag{5.7a}$$

$$\gamma_{\mathrm{d}} = \frac{1}{R_4}\left[1 + \frac{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4}{R_{2\mathrm{a}} + R_{2\mathrm{b}}}\right] . \tag{5.7b}$$

$$\zeta_{\mathrm{d}} = \infty . \tag{5.7c}$$

$$\alpha_{\mathrm{d}} = 1 + \frac{R_{2\mathrm{a}} + R_{2\mathrm{b}}}{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4} . \tag{5.7d}$$

$$Z_{\mathrm{id}} = \frac{(R_{2\mathrm{a}} + R_{2\mathrm{b}})R_4}{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4} . \tag{5.7e}$$

$$Z_{\mathrm{od}} = \infty . \tag{5.7f}$$

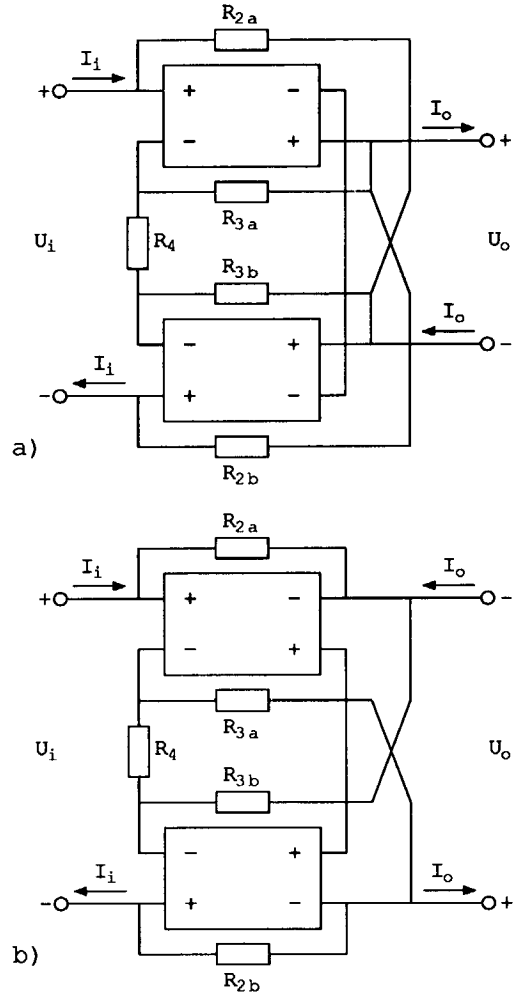

Fig. 5.6 The new power-to-voltage amplifiers.

$$\mu_{\mathrm{d}} = 1 + \frac{R_{3\mathrm{a}} + R_{3\mathrm{b}}}{R_4} . \tag{5.8a}$$

$$\gamma_{\mathrm{d}} = \infty . \tag{5.8b}$$

$$\zeta_{\mathrm{d}} = \frac{(R_{2\mathrm{a}} + R_{2\mathrm{b}})(R_{3\mathrm{a}} + R_{3\mathrm{b}} + R_4)}{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4} . \tag{5.8c}$$

$$\alpha_{\mathrm{d}} = \infty . \tag{5.8d}$$

$$Z_{\mathrm{id}} = \frac{(R_{2\mathrm{a}} + R_{2\mathrm{b}})R_4}{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4} . \tag{5.8e}$$

$$Z_{\mathrm{od}} = 0 . \tag{5.8f}$$

Thus far amplifiers with a linear, accurately-known input impedance and a high or low output impedance (i.e. a current or voltage output) have been shown. These amplifiers are especially useful because of their good noise behavior and their suitability for monolithic integration. On the other hand, an amplifier with a high or low input impedance and a linear, accurately-known output impedance can be very useful, particularly when driving loads with a characteristic impedance. For example, the brute-force method leads, in the case of characteristic impedance matching, to a degradation in power efficiency: a 3 dB power loss in the series or parallel connected resistance.

Balanced amplifiers can also be used for a minimization of this power loss. In Fig. 5.7 an example of an amplifier with a low-input impedance and a linear, accurately-known output impedance is given. As both the output current and the output voltage are directly related to the input current, these configurations are called current-to-power amplifiers. Equations 5.9a–5.9f describe the parameters. This amplifier can very easily be derived from the one in Fig. 5.6: interchange the input and output of the amplifier and the nullors in the amplifier and Fig. 5.7 results.

Figure 5.8 can be deduced in the same way from Fig. 5.5 and constitutes an amplifier with a high input impedance and a linear, accurately-known output impedance. As both output current and output voltage are directly related to the input voltage, these configurations are called voltage-to-power amplifiers. Equations 5.10a–5.10f describe the parameters.

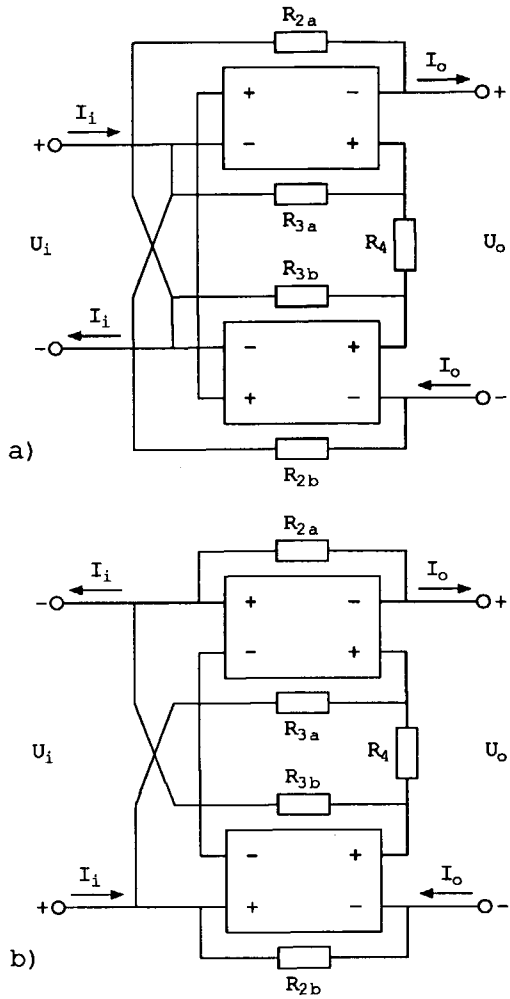

Fig. 5.7 The new current-to-power amplifiers.

$$\mu_{\mathrm{d}} = \infty \,. \tag{5.9a}$$

$$\gamma_{\mathrm{d}} = \infty \,. \tag{5.9b}$$

$$\zeta_{\mathrm{d}} = \frac{(R_{2\mathrm{a}} + R_{2\mathrm{b}})(R_{3\mathrm{a}} + R_{3\mathrm{b}} + R_4)}{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4} \,. \tag{5.9c}$$

$$\alpha_{\mathrm{d}} = 1 + \frac{R_{3\mathrm{a}} + R_{3\mathrm{b}}}{R_4} \,. \tag{5.9d}$$

$$Z_{\mathrm{id}} = 0 \,. \tag{5.9e}$$

$$Z_{\mathrm{od}} = \frac{(R_{2\mathrm{a}} + R_{2\mathrm{b}})R_4}{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4} \,. \tag{5.9f}$$

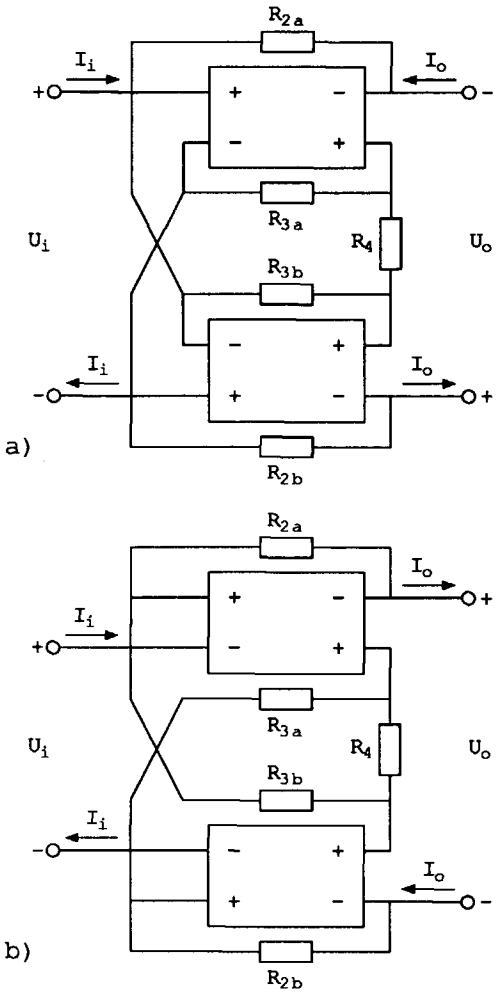

Fig. 5.8 The new voltage-to-power amplifiers.

$$\mu_{\mathrm{d}} = 1 + \frac{R_{2\mathrm{a}} + R_{2\mathrm{b}}}{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4} . \tag{5.10a}$$

$$\gamma_{\mathrm{d}} = \frac{1}{R_4}\left[1 + \frac{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4}{R_{2\mathrm{a}} + R_{2\mathrm{b}}}\right] . \tag{5.10b}$$

$$\zeta_{\mathrm{d}} = \infty . \tag{5.10c}$$

$$\alpha_{\mathrm{d}} = \infty . \tag{5.10d}$$

$$Z_{\mathrm{id}} = \infty . \tag{5.10e}$$

$$Z_{\mathrm{od}} = \frac{(R_{2\mathrm{a}} + R_{2\mathrm{b}})R_4}{R_{3\mathrm{a}} + R_{3\mathrm{b}} + 2R_4} . \tag{5.10f}$$

5.3.2 Balanced single-loop amplifier configurations

To complete the extensive classification of balanced amplifiers, the dual-loop amplifiers sketched in Fig. 5.5, 5.6, 5.7 and 5.8 can be degenerated to single-loop amplifiers. This can be achieved by making certain impedance(s) in the feedback network either zero or infinite. This results in balanced voltage amplifiers, balanced current amplifiers, balanced transadmittance amplifiers and balanced transimpedance amplifiers, all in two versions. These configurations are sketched in Fig. 5.9, 5.10, 5.11 and 5.12, respectively. In Equation 5.11, 5.12, 5.13 and 5.14 the relevant parameters are given for all balanced single-loop amplifiers.

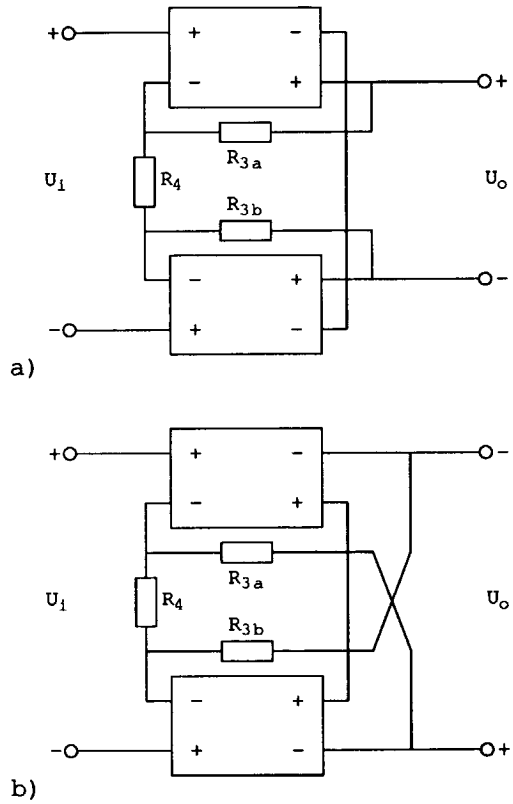

Fig. 5.9 The balanced voltage amplifiers.

$$\mu_d = 1 + \frac{R_{3a} + R_{3b}}{R_4} \,. \tag{5.11}$$

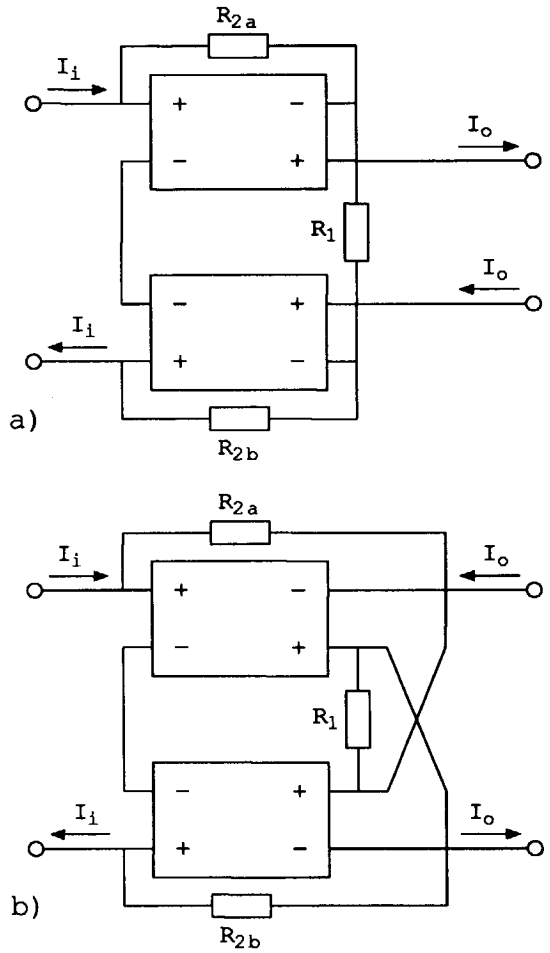

Fig. 5.10 The balanced current amplifiers.

$$\alpha_{\mathrm{d}} = 1 + \frac{R_{2\mathrm{a}} + R_{2\mathrm{b}}}{R_1} \,. \tag{5.12}$$

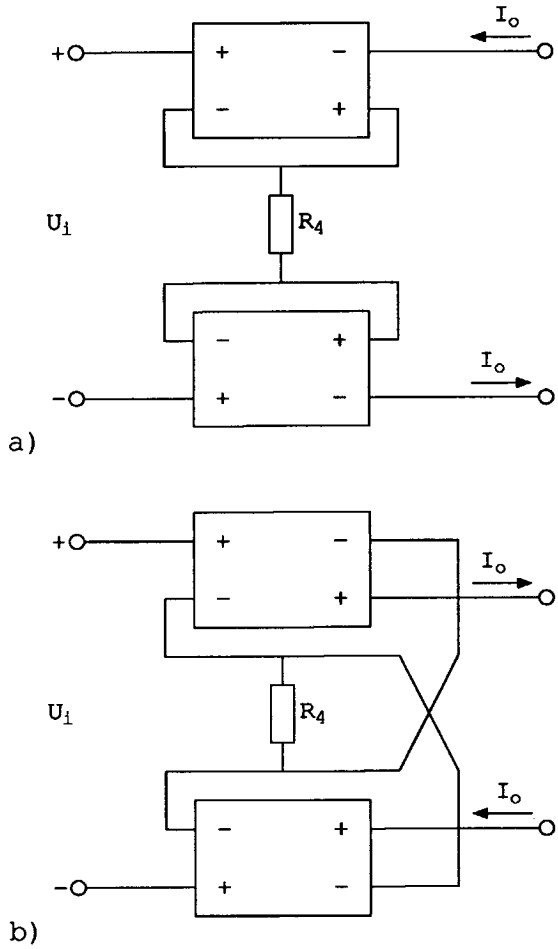

Fig. 5.11 The balanced transadmittance amplifiers.

$$\gamma_d = \frac{1}{R_4} \,. \tag{5.13}$$

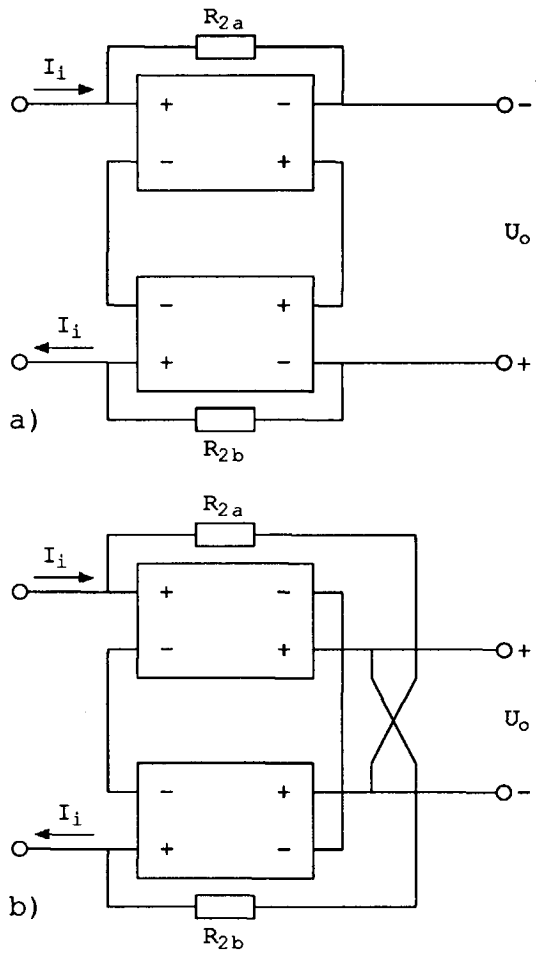

Fig. 5.12 The balanced transimpedance amplifiers.

$$\zeta_d = R_{2a} + R_{2b} \,. \tag{5.14}$$

One of the two versions in Fig. 5.9, 5.10 and 5.11 is new while the other amplifier is the commonly used amplifier configuration. For the sake of completeness it should be mentioned that by using balanced amplifiers, and having both signal and inverted signal available, the class of multi-loop amplifiers can be extended. By using the new technique shown in this section, three-loop and four-loop amplifiers with impedance feedback are feasible. In reference [17] the practical value of these circuits has been discussed.

In this section the differential-mode behavior of several new balanced amplifiers was discussed. It was shown that each balanced amplifier can be realized in two different versions which have the same differential-mode transfer parameters. By eliminating certain impedances, balanced single-loop amplifiers are obtained, leading to three new balanced single-loop amplifiers. Common-mode analysis can reveal differences between the two balanced dual-loop amplifier configurations. This common-mode behavior will therefore be the point of discussion in the next section.

5.4 Common-mode behavior

In this section the common-mode behavior of the new balanced amplifiers will be discussed. The common-mode transfer parameters are necessary for calculating the output signals (or unbalance in output signals) in case of unbalanced sources. Further, a striking difference between the two versions of each of the new balanced dual-loop amplifiers will be revealed.

First it must be made clear why the common-mode transfer parameters must be used in case an unbalanced source is connected to a balanced amplifier. This can be clarified by Fig. 5.13a. In this figure a balanced amplifier with voltage inputs is assumed for simplicity. All other kinds of input signals (current or a combination of voltage and current) can be treated similarly. The amplifier is connected to an unbalanced source, so one of the inputs is grounded. The differential-mode and common-mode voltages are defined by:

$$U_d = U_1 - U_2 , \tag{5.15a}$$

$$U_c = \frac{U_1 + U_2}{2} , \tag{5.15b}$$

where U_1 and U_2 are the voltages of node 1 and 2 to ground, respectively. Thus the differential-mode voltage is equal to the signal voltage: $U_d = U_s$. The common-mode voltage is equal to half the source voltage: $U_c = U_s/2$. In Fig. 5.13b the same circuit is again presented with the differential-mode

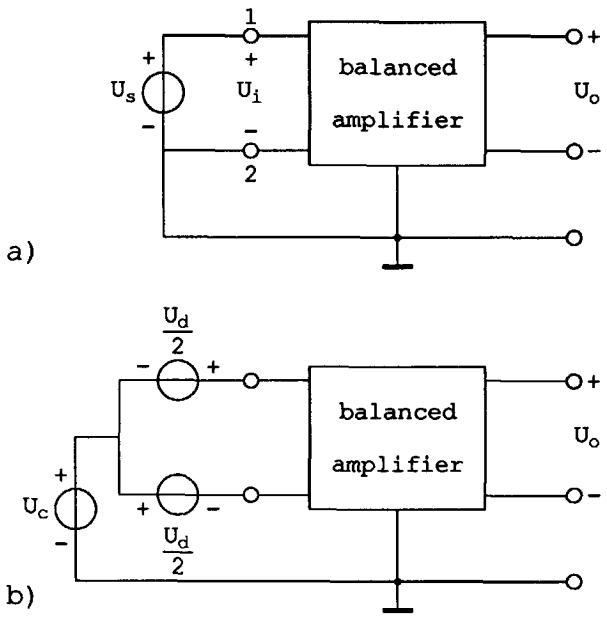

Fig. 5.13 A balanced amplifier connected to an unbalanced source.

and common-mode sources drawn explicitly. In order to calculate the output signals of a balanced amplifier when an unbalanced source is used and consequently a common-mode signal is applied to the input, both differential-mode and common-mode parameters must be taken into account.

The common-mode transfer parameters of balanced amplifiers can be calculated much more easily if Bartlett's bisection theorem (see reference [62] or reference [63]) can be applied (see reference [64] for the application). According to this theorem only one-half of a balanced circuit has to be reckoned with if either differential-mode or common-mode behavior is of interest. For this purpose a line of symmetry has to be established in the balanced circuit. As an example the power-to-current dual-loop amplifier of Fig. 5.5a is used, see Fig. 5.14.

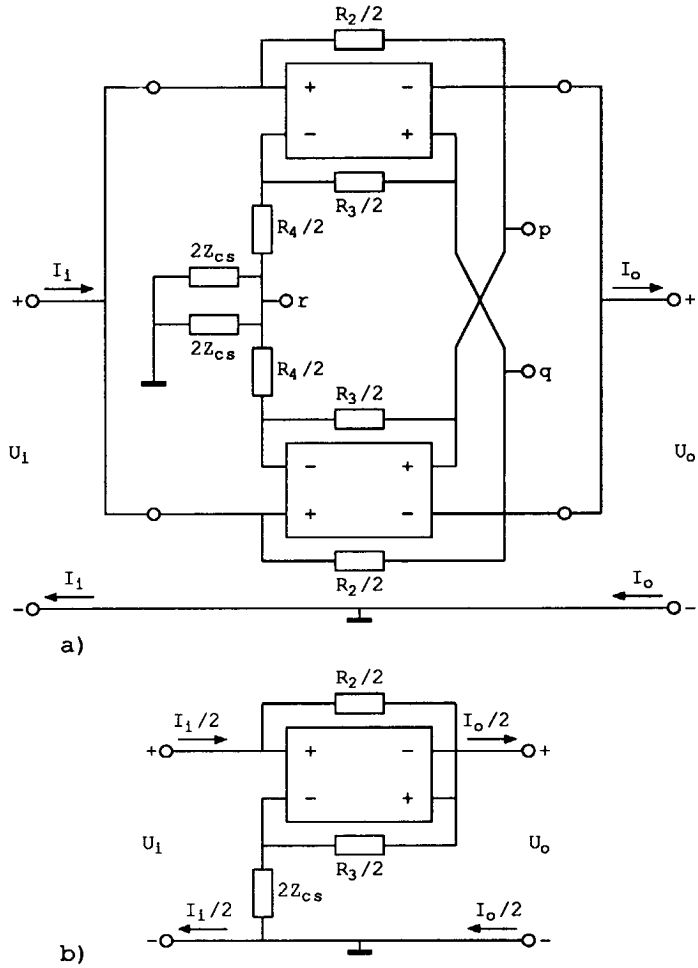

Fig. 5.14 The bisection theorem applied to a balanced dual-loop amplifier.

$$\mu_c = \infty . \tag{5.16a}$$

$$\gamma_c = -\frac{R_2 + R_3}{R_2 Z_{cs}} . \tag{5.16b}$$

$$\zeta_c = \infty . \tag{5.16c}$$

$$\alpha_c = 1 + \frac{R_2}{R_3} . \tag{5.16d}$$

$$Z_{ic} = -\frac{R_2 Z_{cs}}{R_3} . \tag{5.16e}$$

$$Z_{oc} = \infty . \tag{5.16f}$$

Since we are only looking at common-mode behavior, the resistances R_{2a} and R_{2b} have been replaced by $R_2/2$. The resistances R_{3a} and R_{3b} have been replaced by $R_3/2$. The cross-coupling of R_2 can be eliminated. As can easily be seen by considering only common-mode voltages, there is no voltage between nodes p and q, so we may connect them. In Fig. 5.14b the resulting common-mode circuit is drawn. To account for leakage paths to ground, a finite impedance Z_{cs} ($Z_{cs} \gg R_4$) from node r to ground is assumed. Equations 5.16a–5.16f give the common-mode parameters (index c) for both versions of the power-to-current amplifiers. The resulting input impedance Z_{ic} is negative, which can lead to common-mode instability. We will look into this phenomenon at the end of this section.

Only the knowledge that the transmission parameters of a nullor are zero and that the output voltage and current of a nullor are limited, are necessary for the calculation of the transfer parameters. Therefore we may omit the + and − signs within the nullors to include both versions of the amplifier. In Fig. 5.15 the common-mode circuit for both versions of the power-to-voltage amplifier, as drawn in Fig. 5.6, is sketched. Equations 5.17a–5.17f describe the common-mode parameters.

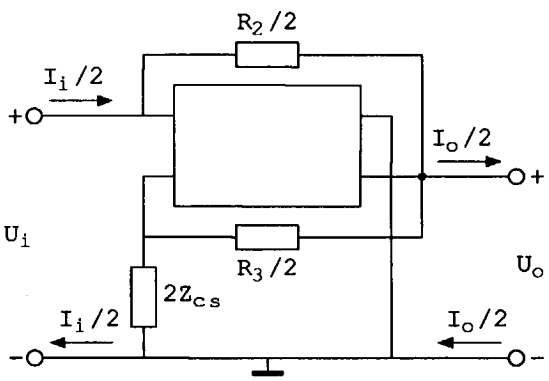

Fig. 5.15 The power-to-voltage common-mode circuit.

$$\mu_c = 1 + \frac{R_3}{4Z_{cs}} . \tag{5.17a}$$

$$\gamma_c = \infty . \tag{5.17b}$$

$$\zeta_c = -R_2 \left[\frac{1}{4} + \frac{Z_{cs}}{R_3}\right] . \tag{5.17c}$$

$$\alpha_c = \infty . \tag{5.17d}$$

$$Z_{ic} = -\frac{R_2 Z_{cs}}{R_3} . \tag{5.17e}$$

$$Z_{oc} = 0 . \tag{5.17f}$$

Figure 5.16 sketches the common-mode circuit for both versions of the current-to-power amplifier as drawn in Fig. 5.7. Equations 5.18a–5.18f describe the common-mode parameters.

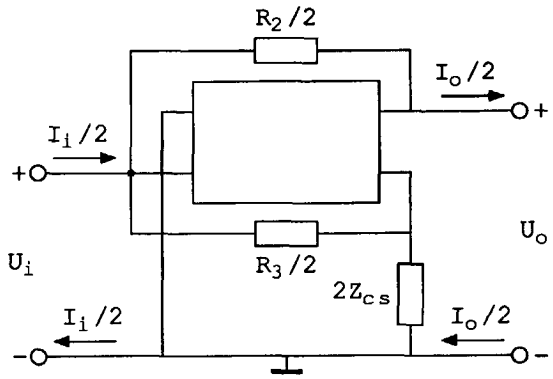

Fig. 5.16 The current-to-power common-mode circuit.

$$\mu_c = \infty . \tag{5.18a}$$

$$\gamma_c = \infty . \tag{5.18b}$$

$$\zeta_c = -R_2 \left[\frac{1}{4} + \frac{Z_{cs}}{R_3} \right] . \tag{5.18c}$$

$$\alpha_c = 1 + \frac{R_3}{4Z_{cs}} . \tag{5.18d}$$

$$Z_{ic} = 0 . \tag{5.18e}$$

$$Z_{oc} = -\frac{R_2 Z_{cs}}{R_3} . \tag{5.18f}$$

Here the output impedance is negative, which can lead to common-mode instability. This phenomenon will be treated at the end of this section.

Figure 5.17 sketches the common-mode circuit for both versions of the voltage-to-power amplifier as drawn in Fig. 5.8. Equations 5.19a–5.19f describe the common-mode parameters.

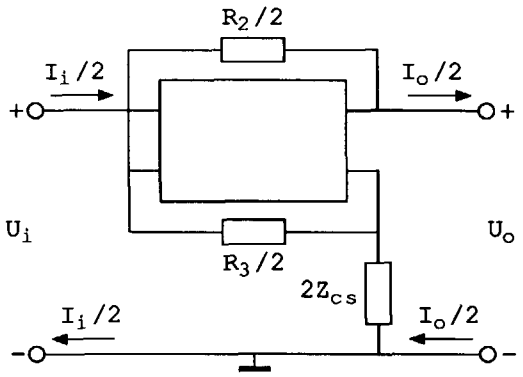

Fig. 5.17 The voltage-to-power common-mode circuit.

$$\mu_c = 1 + \frac{R_2}{R_3} . \tag{5.19a}$$

$$\gamma_c = -\frac{1}{Z_{cs}} \left[1 + \frac{R_3}{R_2} \right] . \tag{5.19b}$$

$$\zeta_c = \infty . \tag{5.19c}$$

$$\alpha_c = \infty . \tag{5.19d}$$

$$Z_{ic} = \infty . \tag{5.19e}$$

$$Z_{oc} = -\frac{R_2 Z_{cs}}{R_3} . \tag{5.19f}$$

So far the common-mode parameters of the dual-loop amplifiers have been given. It was shown that the common-mode input impedance or the common-mode output impedance is negative. The problem caused by this negative impedance will now be discussed. In Fig. 5.18 the common-mode circuits (including signs) for the two power-to-current versions are drawn separately. From Fig. 5.18a it can be deduced that the feedback caused by R_2 is positive while the feedback caused by R_3 is negative. In Fig. 5.18b the situation is the reverse; R_2 is part of the negative feedback loop while R_3 is part of the positive feedback loop. Depending on the value of the source impedance either the negative or positive feedback loop dominates. The amplifier of Fig. 5.18a is stable for low common-mode source impedances and unstable for high common-mode source impedances (also called "short-circuited stable"). The amplifier of Fig. 5.18b is unstable for low common-mode source impedances while it is stable for high common-mode source impedances (also called "open stable"). It must be emphasized that only the common-mode source impedance determines stability, the differential-mode source impedance has no influence on

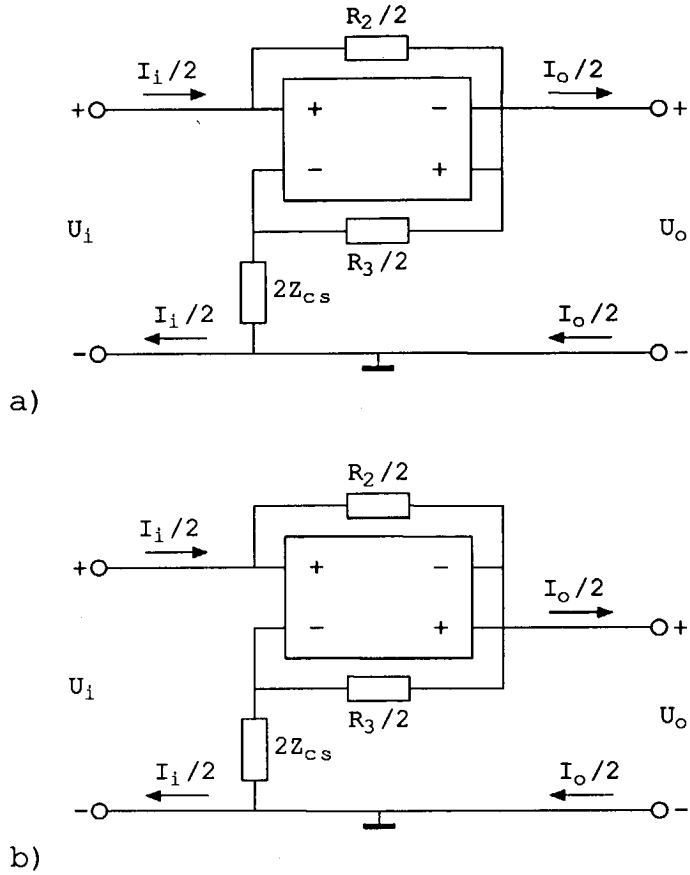

Fig. 5.18 The common-mode circuit for both power-to-current amplifiers.

this stability. For the power-to-voltage, the current-to-power and voltage-to-power amplifiers comparable deductions can be made. For the latter two, the load impedance instead of the source impedance determines stability.

Balanced single-loop amplifiers also can be hindered by this phenomenon. For instance, the balanced transimpedance amplifier in reference [65] and as shown in Fig. 5.12b here, has a common-mode positive feedback loop in the amplifier. Apparently this poses no problem in this case due to a low common-mode loop gain.

In this section the common-mode transfer parameters of the new balanced dual-loop amplifiers have been calculated. The common-mode transfer parameters of the single-loop balanced amplifiers are easily derived by eliminating certain impedances. It was shown that one of the transfer parameters was negative, so common-mode instabilities may arise and thereby prevent the circuit from being properly biased, because biasing entails common-mode currents and voltages. By calculating the common-mode loop gain and lowering this loop gain below unity and/or choosing the proper configuration in connection with the available source, stability is ensured. The stability problem is essentially caused by the cross-coupling of impedances from one side of the circuit to the side of the circuit where the differential-mode signal is inverted.

5.5 A high-dynamic range power-to-current amplifier

5.5.1 Introduction

In this section the design of a practical example of a dual-loop power-to-current amplifier will be presented. The realized amplifier should be directly applicable as the input amplifier for the FM front-end.

When a whip antenna is used, we have an unbalanced source with a low common-mode source impedance. The desired differential-mode input impedance of the amplifier is 330 ohm. The noise figure of the amplifier must be lower than 3 dB. The dynamic range should be at least 106 dB. The intermodulation-free dynamic range should be about 80 dB. The voltage-to-current transfer should be about 20 mA/V.

We will start by choosing an amplifier configuration. The resistances in the feedback network will be chosen according to noise considerations and transfer requirements. Next a physical implementation for the nullors will be given. With this implementation the noise of the active devices can be calculated and minimized. The bias circuitry is designed in such a way that no extra noise is introduced. The amplifier is integrated on a 3 GHz semi-custom chip in the Delft high-frequency BIFET process [4]. This section will be concluded with the measurement results of the amplifier.

5.5.2 Design of the amplifier

The choice of amplifier configuration (Fig. 5.5a or 5.5b) depends solely on the common-mode source impedance. As this impedance is low the configuration of Fig. 5.5a is chosen.

Now the noise contributions of the feedback network and the source resistance will be discussed. By using the Blakesley transformation for noise voltage sources and an equivalent transformation for current sources (see reference [17] for instance), the total noise due to the source and feedback resistances can be summed in one source in series with the signal source. Under the assumption that true differential source and load are used, the power-density spectrum of the input noise voltage is given by:

$$\begin{aligned} S(u) &= 4kT\left\{R_g + R_4\left[\frac{R_2+R_3+2R_g}{R_2+R_3+2R_4}\right]^2 + \right. \\ &\quad \left(\frac{1}{R_{3a}}+\frac{1}{R_{3b}}\right)\left[\frac{R_3(R_4-R_g)}{2(R_2+R_3+2R_4)}\right]^2 + \\ &\quad \left.\left(\frac{1}{R_{2a}}+\frac{1}{R_{2b}}\right)\left[\frac{R_2(R_4-R_g)}{2(R_2+R_3+2R_4)}\right]^2\right\}, \end{aligned} \tag{5.20}$$

where $R_2 = R_{2a} + R_{2b}$ and $R_3 = R_{3a} + R_{3b}$ have been used. The successive terms represent the contributions of R_g (the source resistance), R_4, R_{3a}, R_{3b}, R_{2a} and R_{2b}.

The contributions of R_2 and R_3 can be made zero by choosing $R_4 = R_g$. This leads to $S(u) = 4kT(R_g + R_4)$. The minimum noise figure is $10\log[(R_g + R_4)/R_g] = 3$ dB, where only the noise of the feedback network is taken into account. This choice is optimal with respect to the noise contributions of R_2 and R_3 but not optimal with respect to the noise contribution of R_4.

A much better noise figure can be obtained by also optimizing for the noise due to R_4. Using the usual single-loop amplifier choices for feedback resistors ($R_2 \gg R_g \gg R_4$) leads to a power-density spectrum of the input noise voltage of $S(u) = 4kT(R_g)$ or a minimum noise figure of 0 dB. In practice, this is not completely feasible because R_2, R_3 and R_4 must be chosen in such a way as to obtain the desired transfer. Moreover, the voltage amplification factor $1 + R_3/R_4$ in the amplifier cannot be made too large because of a limited supply voltage. Loop gain and high-frequency behavior further compromises the choice of the resistors in the feedback network.

To accommodate the transfer parameter requirements $R_2 = 1320\ \Omega$, $R_3 = 110\ \Omega$ and $R_4 = 55\ \Omega$ have been chosen. This leads to $Z_i = 330\ \Omega$, $\gamma = 21.2$ mA/V and $\alpha = 7$. The noise contribution of the feedback network amounts to $S(u) = 4kT(R_{eq})$ with $R_{eq} = 147\ \Omega$. The minimum noise figure then amounts to 1.6 dB, where $R_g = 330\ \Omega$ has been used.

In Fig. 5.5 the amplifier configuration was shown with nullors. A physical implementation with practical devices must be found for the nullor. For mono-

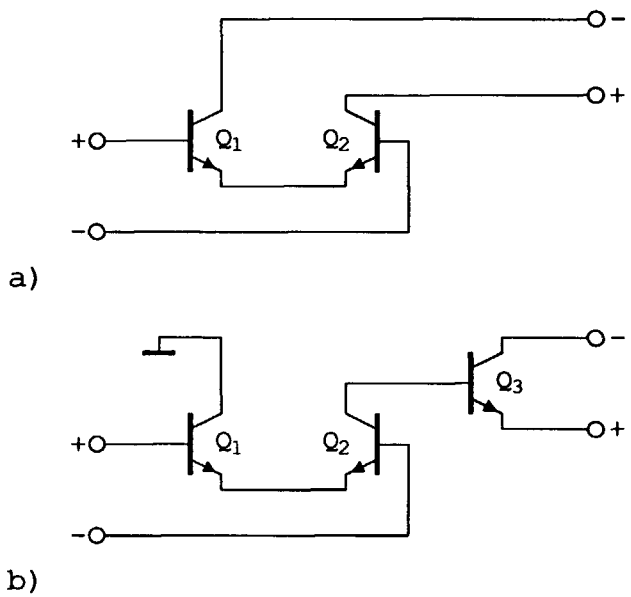

Fig. 5.19 The nullor implementation.

lithic integration an all-NPN bipolar transistor design is desirable because of the high-frequency behavior. Figure 5.19a shows a simple approximation of a nullor, a differential pair consisting of transistor Q_1 and Q_2. The loop gain of the resulting amplifier is so low that accuracy is endangered. The loop gain can be enlarged by adding an extra transistor Q_3 at the output; see Fig. 5.19b. To prevent current mismatch in the output currents of the nullor, the collector of Q_1 should preferably be connected to the collector or emitter of Q_3. However, connection to the collector of Q_3 increases the output capacitance while connection to the emitter of Q_3 conflicts with the high-frequency stability in this design. Therefore the collector of Q_1 is connected to signal ground.

Now that the nullor has been designed, the extra noise contributed by the active devices can be calculated. Only the noise of the base resistance r_b ($S(u_b) = 4kTr_b$), base current shot noise ($S(i_b) = 2qI_B$) and collector current shot noise ($S(i_c) = 2qI_C$) at low frequencies is taken into account. Again, the noise sources are transferred to one source in series with the signal source. By using $r_e = kT/(qI_C)$ and $I_B = I_C/h_{\mathrm{FE}}$ (h_{FE} is the transistor DC current gain)

the resulting power-density spectrum of the input noise voltage is given by:

$$S(u) = 4kT\left\{\left[4r_b + 2r_e + \frac{4r_b^2}{2h_{\mathrm{FE}}r_e}\right]\left[\frac{R_2+R_3+R_4+R_g}{R_2+R_3+2R_4}\right]^2 + \frac{R_g^2}{4h_{\mathrm{FE}}r_e} + \frac{R_4^2}{4h_{\mathrm{FE}}r_e}\left[\frac{R_2+R_3+2R_g}{R_2+R_3+2R_4}\right]^2\right\}. \quad (5.21)$$

For a low noise contribution, the base resistance r_b should be as low as possible, so large transistors must be used. By using the approximation $R_2 \gg R_g \gg R_4$ a minimum value of $S(u)$ can be obtained by differentiating $S(u)$ with respect to r_e and hence an optimum value for r_e ($r_{e,\mathrm{opt}}$) and thus an optimum bias current $I_{e,\mathrm{opt}}$ can be found:

$$r_{e,\mathrm{opt}} = \sqrt{\frac{1}{4h_{\mathrm{FE}}}\left[4r_b^2 + \frac{R_g^2}{2} + \frac{R_4^2}{2}\right]} = \frac{kT}{qI_{e,\mathrm{opt}}}. \quad (5.22)$$

For $r_b = 10\ \Omega$ and a transistor current gain of $h_{\mathrm{FE}} = 100$, the optimum bias current is 2.5 mA. The total noise caused by the active devices then amounts to $S(u) = 4kT(R_{\mathrm{eq}})$ with $R_{\mathrm{eq}} = 105\ \Omega$. The noise due to the feedback network and the active devices together amounts to $S(u) = 4kT(R_{\mathrm{eq}})$ with $R_{\mathrm{eq}} = 237\ \Omega$ or a noise figure of 2.5 dB.

So far noise calculations of the feedback network and the active devices have been made. An optimum bias current for a minimum noise contribution of the input stage was found. For a dynamic range in the order of 106 dB the output stage must be biased at about 5 mA.

Figure 5.20 gives a possible biasing scheme for the total amplifier. Q_1–Q_4 are large transistors (485 μm × 2 μm emitter in finger structure) for their low base resistance. The transistors Q_5 and Q_6 are sufficiently large (76 μm × 2 μm emitter) to have their optimum performance at 5 mA. Transistors Q_7–Q_{10} act as current sources with large emitter series resistances to keep their noise contributions low. The two 1.8 kΩ resistors determine the base voltage of Q_5 and Q_6. SPICE simulations show a slight peaking in the transfer parameter α which can be counteracted by a small capacitance (0.7 pF) over R_{2a} and R_{2b} (phantom zero).

To decrease the influence of low-frequency changes in the positive power supply voltage, p-channel JFETs can be used instead of the 1.8 kΩ resistors for biasing the transistors Q_5 and Q_6. Figure 5.21 gives the FET replacement for the 1.8 kΩ resistors. Current sources with PNP transistors seriously degrade the noise figure. As the f_T of the PNPs is low (about 20 MHz) the collector current shot noise significantly increases the noise figure at 100 MHz.

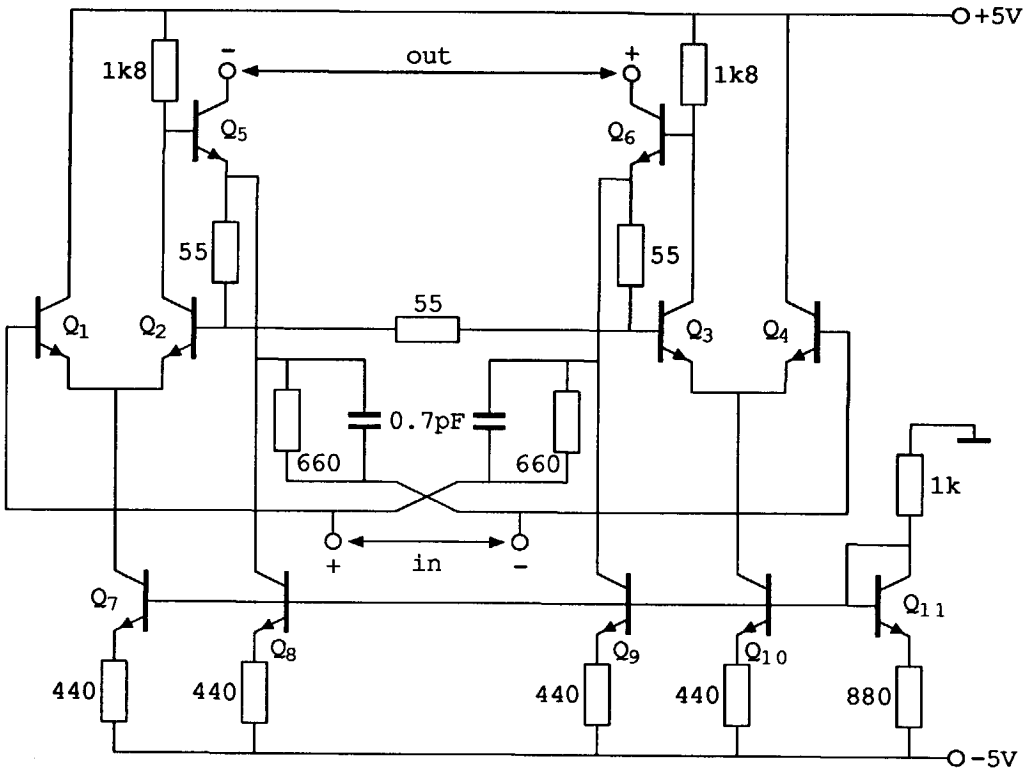

Fig. 5.20 The biased circuit.

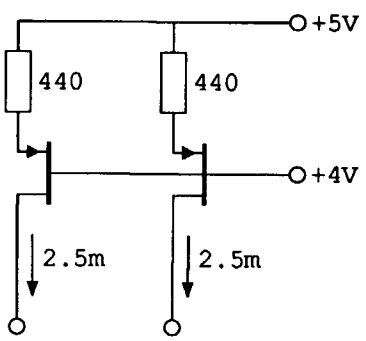

Fig. 5.21 The FET replacement.

The circuit shown in Fig. 5.20 has been integrated on a semi-custom chip in the high-frequency BIFET process that has been developed at the Delft University [4], [66]. Figure 5.22 shows a photograph of the chip. The four large transistors for the input stage can be found in the middle of the chip at the top and at the bottom.

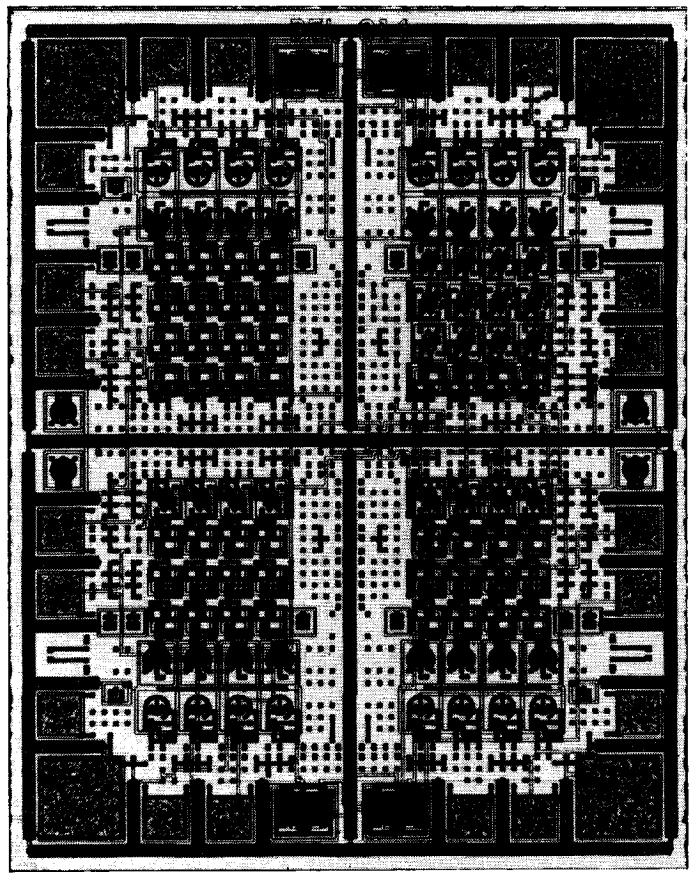

Fig. 5.22 A photograph of the integrated amplifier.

5.5.3 Amplifier measurements

Some measurements were performed on the IC with the use of transformers to obtain a balanced source and load. At 10 MHz the transadmittance γ is 19 mA/V while the current-gain factor α is 6.7. The input impedance is 350 Ω. The noise figure at 10 MHz is 2.6 dB which is equivalent to a noise source at the input of $S(u) = 4kT(R_{eq})$ with $R_{eq} = 250\ \Omega$. At 100 MHz the noise figure is 3 dB for the versions with the resistor and FET current sources. For the version with the PNP current sources, the noise figure is 4 dB at 100 MHz.

The transadmittance -3 dB bandwidth is 400 MHz. The current gain bandwidth is 125 MHz. The input impedance has a -3 dB bandwidth of 125 MHz, which is obviously determined by the limited bandwidth of the current gain factor. All bandwidth measurements were performed with a single-sided source and load.

In Fig. 5.23 the output I_o current is plotted as a function of the input voltage U_i. The noise is measured in a 180 kHz bandwidth. For a 100 MHz input signal the 1 dB compression level is 99 dB higher, i.e. the dynamic range is 99 dB. In case two signals with equal amplitude but of different frequencies (90 and 100 MHz) are applied to the amplifier, the second-order and third-order intermodulation components can be measured. For the frequency of the second-order intermodulation component 10 MHz was taken while 80 MHz was taken for the frequency of the third-order intermodulation component. The $IMFDR_2$ is 76 dB when a balanced source and load are used, while 70 dB is measured for an unbalanced source. The $IMFDR_3$ in both cases amounts to 74 dB.

The specifications of the amplifier can be summarized as follows:

transadmittance:	$\gamma = 19$ mA/V	$f_{-3\,dB} = 400$ MHz
current gain:	$\alpha = 6.7$	$f_{-3\,dB} = 125$ MHz
input impedance:	$Z_i = 350\ \Omega$	$f_{-3\,dB} = 125$ MHz
noise figure:	NF = 2.5 dB @ $f = 10$ MHz	
	NF = 3.0 dB @ $f = 100$ MHz	
dynamic range:	DR = 99 dB @ $f = 100$ MHz ($B = 180$ kHz)	
intermodulation-free dynamic range:		$IMFDR_2 = 76$ dB
		$IMFDR_3 = 74$ dB

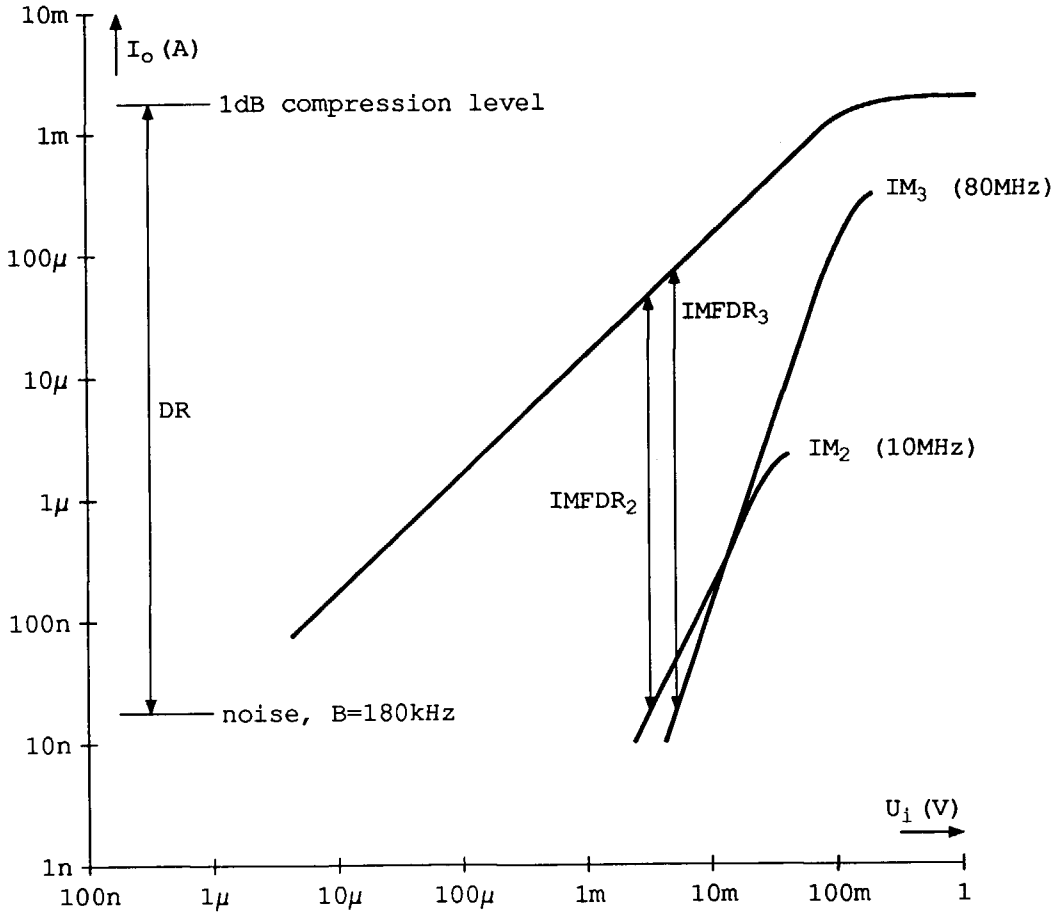

Fig. 5.23 Distortion characteristics of the power-to-current amplifier.

5.6 Discussion

In the first section of this chapter the question was raised: How should a filter be terminated without introducing excessive noise. The brute-force approach is to be abandoned because of its high noise figure. With the transmission matrix as the point of reference, the standard dual-loop amplifier configurations were discussed. They appear to be impractical for integrated circuit realizations.

Section 5.3 presented a new class of high-performance dual-loop negative-feedback amplifiers which use impedance feedback networks to accomplish a

linear and accurately-known input impedance combined with a high or low output impedance. A noise figure smaller than 3 dB is feasible. The key to this new technique is the use of two balanced amplifier circuits with two overall feedback loops.

Similar techniques can be used to realize amplifiers with an accurately-known output impedance without degrading the power efficiency. With the use of these techniques, the input impedances are either high or low.

In Section 5.4 the common-mode transfer parameters of the new class of amplifiers have been given. It has been shown that, although the differential-mode feedback is negative feedback, the common-mode feedback may be positive, depending on the source or the load impedance. Stability is ensured by lowering the common-mode loop gain below unity and/or choosing the proper configuration in connection with the available common-mode source impedance. This common-mode positive feedback is not solely a problem of dual-loop amplifiers but can also be found in balanced single-loop amplifiers.

In Section 5.5 the design and performance of one particular amplifier, belonging to the new class of amplifiers, have been presented. This amplifier has a linear, accurately-known input impedance and a high output impedance. The noise behavior of the feedback network and the active devices is discussed extensively: a noise figure lower than 3 dB is feasible. An all-NPN amplifier design is presented. Biasing is implemented without deteriorating the amplifier noise figure. The amplifier has been realized in a 3 GHz f_T BIFET process on a semi-custom chip. A bandwidth well above 100 MHz, low distortion and a noise figure of 2.5 dB at 10 MHz have been achieved.

6. Conclusions and recommendations

6.1 Results obtained in this research

The first result to be discussed is the demonstration of the feasibility of an integrated FM upconversion receiver front-end with on-chip SAW filters. The conventional FM consumer receivers are basically single-(down)conversion receivers. The radio frequency signals are, with the use of a mixer and a local oscillator, downconverted into an intermediate frequency of 10.7 MHz. After the IF filtering, and thus attenuation of the adjacent channels, the desired signals are demodulated by standard FM detectors.

Due to the mixing process in this single-conversion receiver, the receiver is also sensitive at a second channel, the image channel. The mixer is preferably driven by a square-wave signal for increasing the intermodulation-free dynamic range of the mixer, resulting in the LO harmonics-related spurious channels at which the receiver is also sensitive. To attenuate the signals that may be present at all these spurious channels, a tuned filter which tracks with the desired channel at the input of the receiver is indispensable.

The signals at the spurious channels can more easily be attenuated while simultaneously the problems with the tuned and tracking filter can be circumvented by choosing an intermediate frequency that is higher than the highest frequency of the RF signal. In this case, a simple band-pass filter of sufficiently high order with a bandwidth equal to that of the FM broadcast band attenuates the signals at all spurious channels.

Such an upconversion receiver is only successful when a number of conditions are met. For instance, the transfer of the antenna from electromagnetic field strength to antenna voltage should not significantly increase at the high frequencies of the image channels. The intermediate frequency filters must have a sufficiently high intermodulation-free dynamic range at the high intermediate frequency and they must, preferably, be made on-chip to avoid problems arising from pin-to-pin feedthrough. Also, the electronic circuitry should exhibit the desired properties at these high intermediate frequencies.

In Chapter 3 it was shown that the antenna transfer does not significantly increase for the image channels. Also the behavior of the antenna impedance does not unfavorably influence the attenuation of the signals at the image channels. The on-chip SAW filters do have a more than sufficient intermodulation-free dynamic range while the electromagnetic feedthrough can, even at high frequencies, be adequately suppressed. In addition, in Chapter 5 a new type

of amplifier has been presented which is excellently suited as input amplifier in an FM receiver front-end.

The dimensions of on-chip SAW resonator filters are somewhat smaller than those of ceramic filters. This means, however, that these resonator filters occupy a chip area which is, when compared to that of today's consumer radio ICs, quite large. An on-chip SAW transversal filter in a ZnO-SiO_2-Si structure, has a length that is approximately equal to a comparable filter realized on the piezoelectric materials quartz or lithium niobate. For a channel-selective IF filter in an FM receiver, such a SAW transversal filter is longer than 10 mm, which exceeds the present standards in IC manufacturing. However, a transversal filter may have a filter shape that is much better than that of a coupled resonator filter. The final choice of the IF filter, either a SAW transversal filter with minimal audio distortion at the cost of making a long chip for a sufficient selectivity or a SAW resonator filter with higher audio distortion but smaller chip area, is left to the manufacturer.

Some problems that need further investigation are the temperature stability and the reproducibility in fabrication of the filter center frequency of both the SAW resonator filter and the SAW transversal filter. Nevertheless, it can be concluded that it is feasible to realize an FM upconversion receiver front-end, where the IF selectivity is obtained by either a SAW transversal filter or a SAW resonator filter.

The second result obtained in the present research is the development of a simple, yet sufficiently accurate model, for the on-chip SAW delay lines and SAW transversal filters from the system and circuit designer's point of view. The modeling has been performed in such a way as to give the insight necessary for understanding the impact of several significant parameters on the filter design. In practice, computer programs are indispensable for an adequate calculation and design of the exact band-pass characteristics of the filter.

It has been concluded from the theory and illustrated by measurements that the signal transfer most suited for systems that require a high-performance information transfer, is the voltage-driven and current-sensed SAW delay line and SAW transversal filter. The usefulness of the developed models has been demonstrated in Section 4.7 where the noise and dynamic range of the amplifier-IF filter-amplifier configurations in the FM upconversion receiver have been calculated.

The third result obtained in the present research is the demonstration of the feasibility of silicon-integrated SAW delay lines, SAW transversal filters

and SAW resonators in combination with high-frequency electronic circuitry. The electromagnetic feedthrough of the silicon-integrated SAW devices can be adequately suppressed for most applications, even at high frequencies. The dynamic range and intermodulation-free dynamic range of the on-chip SAW devices is so high that they hardly are a limiting factor in the performance of any system employing these devices.

For further information of the results of SAW filters obtained in the ZnO-SiO_2-Si structure, the Ph.D. Thesis of J.H. Visser should be consulted [5]. These on-chip SAW devices cannot only be used in the FM upconversion receiver front-end but also in various other applications such as intermediate frequency filters in television receivers, broadcast satellite receivers, high-definition television receivers and various other communication systems in both the consumer and the professional field.

The fourth result obtained in the present research concerns the new class of (balanced) dual-loop amplifiers. Several years ago, an extensive classification of multi-loop amplifiers in combination with a structural design procedure for these amplifiers has been made by E.H. Nordholt [17]. The classification lacks high-performance dual-loop amplifiers that are suitable for monolithic integration. In Chapter 5, balanced amplifier structures were used to obtain dual-loop amplifiers that are suitable for monolithic integration. This idea leads to four new types of dual-loop amplifiers of which each version can be realized in two different ways.

By eliminating certain impedances in the feedback network of these amplifiers, these dual-loop amplifiers can be degenerated into single-loop amplifiers, of which some have been well-known while others are new. The use of the balanced amplifier structures can also be used to obtain amplifiers with three or four feedback loops with impedance feedback networks. The realization of one of the new dual-loop amplifiers in a 3 GHz process demonstrates the feasibility of the new dual-loop amplifier configurations.

6.2 Recommendations for further research

An improvement in processing technology is essential for commercializing the on-chip SAW filters. The ZnO-SiO_2-Si technology for allowing the realizing of the on-chip SAW filters combined with electronic circuitry has only been in use for a couple of years. More process technology research will undoubtably lead to a significant improvement in reproducibility and matching of the resulting filters. If executed in cooperation with an IC factory, such research will no doubt result in the production of integrated circuits and systems with on-chip

SAW filters.

To relieve the demands posed on the reproducibility, temperature stability and matching in the first products, it is recommended that a wide-band filter at not too low or too high frequencies in combination with filter driving and loading amplifiers, mixers and/or detectors, is used as a starting point for the fabrication of these ICs. A wide-band filter also has the advantage that the length of the chip is more acceptable to IC manufacturers. For instance, a transversal filter with a 5 MHz bandwidth is approximately 4 mm long.

It must, however, be kept in mind that a reproducibility and stability of the filter center frequency as on quartz is very hard to obtain with the present ZnO-SiO_2-Si structure. Further research should, therefore, also be concentrated on the determination of the long term-stability of the on-chip SAW filters.

For the more distant future, it is recommended that other piezoelectric materials be investigated with respect to compatibility with silicon integration. Materials like Tl_3VS_4 and Tl_3TaSe_4 have a surface acoustic wave velocity in the order of 1000 m/s [67]. A transversal filter realized with one of these materials has a length that is three times less than a transversal filter realized on quartz, lithium niobate or ZnO-SiO_2-Si, as this velocity is the only factor that determines the length of a SAW transversal filter for a given (absolute) bandwidth. These materials are, therefore, especially interesting for the realization of narrow-band filters.

Research should be undertaken to find piezoelectric materials that are compatible with silicon technology and in which higher-order Rayleigh-wave modes are not present. Filters without extra pass-bands give the system designer more freedom in the choice of the intermediate frequency.

It is also advantageous to find piezoelectric materials of which the material constants are such that the effective noise and the dynamic range, as calculated in Section 4.7, are respectively, lower and higher. This relieves the demands posed on the electronic circuitry that drives and loads the SAW device.

The last recommendation concerns the new class of balanced dual-loop amplifiers. As these amplifiers are both balanced and have two feedback loops, their design is quite complicated. The common-mode behavior of these new amplifiers also needs extra attention. It is recommended that research be directed to the problem of calculating the low-frequency accuracy and the high-frequency behavior in a simple, yet sufficiently accurate, way.

References

[1] H.C. Nauta, "Fundamental aspects and design of monolithically integrated AM radio receivers", Ph.D. Thesis, Delft University of Technology, Delft, the Netherlands, Feb. 1986.

[2] J.W.Th. Eikenbroek, "Development of an integrated AM short-wave upconversion receiver front-end", Ph.D. Thesis, Delft University of Technology, Delft, the Netherlands, March 1989.

[3] M.J. Vellekoop, E. Nieuwkoop, J.C. Haartsen and A. Venema, "A monolithic SAW physical-electronic system for sensors", IEEE Ultrasonics Symposium Proceedings, Chicago, 1987, pp. 641–644.

[4] L.K. Nanver, "High-performance BIFET process for analog integrated circuits", Ph.D. Thesis, Delft University of Technology, Delft, the Netherlands, Oct. 1987.

[5] J.H. Visser, "Surface acoustic wave filters in ZnO-SiO_2-Si layered structures: design, technology, and monolithic integration with electronic circuitry", Ph.D. Thesis, Delft University of Technology, Delft, the Netherlands, Dec. 1989.

[6] "Transmission standards for FM sound broadcasting at VHF", Recommendations and Reports of the CCIR, Vol. X, Part 1, Recommendation 450–1, Geneva, 1986, pp. 237–239.

[7] "System for automatic tuning and other applications in FM radio receivers for use with the pilot-tone system", Recommendations and Reports of the CCIR, Vol. X, Part 1, Recommendation 643, Geneva, 1986, pp. 249–255.

[8] "Stereophonic or multi-dimensional sound in frequency-modulation sound broadcasting", Recommendations and Reports of the CCIR, Vol. X, Part 1, Report 300–6, Geneva, 1986, pp. 239–242.

[9] "Polarization of emissions in frequency-modulation broadcasting in band 8 (VHF)", Recommendations and Reports of the CCIR, Vol. X, Part 1, Report 464–4, Geneva, 1968, pp. 227–231.

[10] A.B. Carlson, "Communication systems, an introduction to signals and noise in communication", McGraw-Hill Book Company, New York, third edition, 1986.

[11] "The RF spectrum of frequency-modulation sound-broadcasting transmitters", Recommendations and Reports of the CCIR, Vol. X, Part 1, Report 1065, Geneva, 1986, pp. 218–223.

[12] M. Kreutzer, "Ein integrierbares Empfängerkonzept für FM realisiert durch einen synchronen Homodynempfänger", Ph.D. Thesis, Die Universität des Saarlandes, Saarbrücken, Germany, 1987.

[13] J. van der Plas, "Fasevergrendelde lus met lusversterking- en nulpuntscompensatie (Phase-lock loop with loop gain compensation and zero compensation)", patent pending, the Netherlands, Dec. 1988.

[14] J. van der Plas and E.H. Nordholt, "A novel extended dynamic range synchronous detector for AM shortwave upconversion receivers", IEEE Transactions on Consumer electronics, Vol. CE-35, no. 3, Aug. 1989, pp. 390–396.

[15] E. Langer, "A new high-frequency receiver front-end with inherent image frequency suppression", Frequenz, Vol. 33, July-Aug. 1979, pp. 236–239.

[16] W.G. Kasperkovitz, "FM receivers for mono and stereo on a single chip", Philips Technical Review, Vol. 41, no. 6, 1983–1984, pp. 169–182. Also: W.G. Kasperkovitz, "FM ontvangers voor mono en stereo op één 'chip'", Philips Technisch Tijdschrift, jaargang 41, no. 6, 1983, pp. 177–191.

[17] E.H. Nordholt, "Design of high-performance negative-feedback amplifiers", Elsevier, Amsterdam, 1983.

[18] R.J. Murray and P.D. White, "Surface acoustic wave devices", Wireless World, March 1981, pp. 3–6, April 1981, pp. 7–10.

[19] S. Ono, K. Wasa and S. Hayakawa, "Surface acoustic wave properties in ZnO-SiO_2-Si layered structures", Wave Electronics, no. 3, 1977, pp. 35-49.

[20] O. Yamazaki, T. Mitsuyu and K. Wasa, "ZnO thin-film SAW devices", IEEE Transactions on Sonics and Ultrasonics, Vol. SU-27, no. 6, Nov. 1980, pp. 369–379.

[21] J.H. Visser, P.T.M. van Zeijl and L.K. Nanver, "On-chip SAW filters in integrated radio systems", IEEE International Conference on Consumer Electronics Proceedings, Chicago, 1989, pp. 90–91.

[22] P.T.M. van Zeijl, J.H. Visser and L.K. Nanver, "FM radio receiver front-end circuitry with on-chip SAW filters", IEEE Transactions on Consumer Electronics, Vol. CE-35, no. 3, Aug. 1989, pp. 512–519.

[23] W.S. Ishak, H.E. Karrer and W.R. Shreve, "Surface acoustic wave delay lines and transversal filters", Hewlett-Packard Journal, Dec.1981.

[24] P.F. Panter, "Modulation, noise and spectral analysis", McGraw-Hill Book Company, New York, 1965.

[25] J. Yamada, A. Yuhara, T. Shiba and T. Toyana, "High-frequency SAW filters stretched over non-overlapped electrodes for satellite broadcast receivers", IEEE Transactions on Consumer Electronics, Vol. CE-31, no. 4, 1985, pp. 673–679.

[26] R. LaRosa, "Optimum coupling circuits for SAW transversal filters", IEEE Ultrasonics Symposium Proceedings, Denver, 1987, pp. 73–78.

[27] S. Datta, "Surface acoustic wave devices", Prentice-Hall, Englewood Cliffs, New Jersey, 1986.

[28] J.C. Haartsen, "Analysis and design of surface acoustic wave delay lines applied in oscillators", Masters thesis, Delft University of Technology, Delft, the Netherlands, 1986.

[29] H.A. Mol, "Phasevectorsum-oscillator with SAW delay lines", Masters thesis, Delft University of Technology, Delft, the Netherlands, 1987.

[30] A. Venema, "Transduction and propagation of surface acoustic waves in three-layered media with an electrically conductive substrate", Ph.D.Thesis, Delft University of Technology, Delft, the Netherlands, 1980.

[31] A. Venema, Private Communication, Dec.1989.

[32] D.F. Tuttle, Jr., "Network synthesis, Vol.1", John Wiley and Sons, New York, 1958.

[33] M.M. Alavi-Sereshki and J.C. Prabhakar, "A tabulation of Hilbert transforms for electrical engineers", IEEE Transactions on Communications, Vol. COM-20, no. 6, Dec. 1972, pp. 1194–1198.

[34] T.L. Szabo, K.R. Laker and E. Cohen, "Interdigital transducer models: their impact on filter synthesis", IEEE Transactions on Sonics and Ultrasonics, Vol. SU-26, no. 5, Sept. 1979, pp. 321–333.

[35] H. Engan, "Excitation of elastic surface waves by spatial harmonics of interdigital transducers", IEEE Transactions on Electron Devices, Vol. ED-16, no. 12, Dec. 1969, pp. 1014–1017.

[36] L.S. Hartmann and B.G. Secrest, " End effects in interdigital surface wave transducers", IEEE Ultrasonics Symposium Proceedings, Chicago, 1972, pp. 413–416.

[37] D.A. Leedom, R. Krimholtz and G.L. Mathaei, "Equivalent circuits for transducers having arbitrary even- or odd-symmetry piezoelectric excitation", IEEE Transactions on Sonics and Ultrasonics, Vol. SU-18, no. 3, July 1971, pp. 128–141.

[38] A.R. Reddy, "Design of SAW band-pass filters using new window functions", IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol. UFFC-35, no. 1, Jan. 1988, pp. 50–56.

[39] W.R. Shreve, "Surface wave twoport resonator equivalent circuit", IEEE Ultrasonics Symposium Proceedings, 1975, pp. 295–298.

[40] B.M. Blok, J.H. Visser, E. van der Drift, and J. Romijn, "High resolution technology for silicon-integrated surface acoustic wave devices", Journal of Vacuum Science and Technology, Dec. 1989.

[41] W.J. Ghijsen, "The acousto-electric field analysis of multilayered surface acoustic wave devices", Ph.D. Thesis, Delft University of Technology, Delft, the Netherlands, Sept. 1987.

[42] R. King and C.W. Harrison Jr., "The distribution of current along a symmetrical center-driven antenna", Proceedings of the I.R.E., October 1943, pp. 548–567, pp. 697.

[43] J.D. Kraus, "Antennas", Mc. Graw-Hill Book Company, New York, 1950.

[44] S.A. Schelkunoff and H.T. Friis, "Antennas, theory and practice", John Wiley and Sons, New York, 1952.

[45] G.A. Thiele and T.H. Newhouse, "A hybrid technique for combining moment methods with the geometrical theory of diffraction", IEEE Transactions on Antennas and Propagation, Vol. AP-23, no. 1, Jan. 1975, pp. 62–69.

[46] R.C. Johnson and H. Jasik, editors, "Antenna engineering handbook", Mc Graw-Hill Book Company, New York, second edition, 1984, Chapter 4.

[47] E.A. Laport, "Radio antenna engineering", Mc Graw-Hill Book Company, New York, 1952.

[48] R.C. Johnson and H. Jasik, editors, "Antenna engineering handbook", Mc Graw-Hill Book Company, New York, second edition, 1984, Chapter 11.

[49] W. Smith, "Antenna manual", Editors and Engineers limited, Santa Barbara, California, 1948.

[50] "Radio, audio and associated systems. Bipolar, MOS", Philips data handbook, book IC01, 1988.

[51] P.Z. Peebles Jr., "Communication system principles", Addison-Wesley Publishing Company, London, 1976.

[52] "Worldwide minimum external noise levels, 0.1 Hz to 100 GHz", Recommendations and Reports of the CCIR, Vol. I, Report 670, Geneva, 1982, pp. 224–229.

[53] J.G. Sneep and C.J.M. Verhoeven, "Design of a low-noise 100 MHz balanced Schmitt-trigger oscillator", Proceedings of the 15th European Solid-State Circuits Conference, Vienna, Sept. 1989, pp. 284–287.

[54] "A procedure for modeling receiver intermodulation characteristics", Recommendations and Reports of the CCIR, Vol. IV, Part. 1, Report 552–1, Geneva, 1986, pp. 217–224.

[55] A.I. Zverev, "Handbook of filter synthesis", John Wiley and Sons, New York, 1967.

[56] H. Akiya, S. Matsuoka and M. Naito, "Using distortion products to effect automatic frequency control in FM tuners", Journal of the Audio Engineering Society, Vol. 28, June 1980, no. 6, pp. 422–428.

[57] P.T.M. van Zeijl, "Dual-loop negative-feedback amplifiers with impedance feedback suitable for monolithic integration", Proceedings of the 30th Midwest Symposium on Circuits and Systems, Syracuse, New York, Aug. 1987, pp. 925–928.

[58] E.H. Nordholt, "Classes and properties of multi-loop negative feedback amplifiers", IEEE Transactions on Circuits and Systems, Vol. CAS-28, no. 3, March 1981, pp. 203–211.

[59] P.T.M. van Zeijl, "Tegenkoppelversterker met nauwkeurig bepaalde ingangs- resp. uitgangsimpedantie in combinatie met hoge of lage uitgangs- resp. ingangsimpedantie" (Negative-feedback amplifiers with accurate input- or output-impedance in combination with high- or low-output- or input-impedance), patent pending in Europe, the USA and Japan.

[60] P.T.M. van Zeijl, "A new high-performance dual-loop amplifier", Proceedings of the Fourteenth European Solid-State Circuits Conference, UMIST, Manchester, UK, Sept. 1988, pp. 150–153.

[61] P.T.M. van Zeijl, "A new high-dynamic range dual-loop power-to-current amplifier", IEEE Journal of Solid-State Circuits, Vol. 24, no. 3, June 1989, pp. 646–650.

[62] A.C. Bartlett, "The theory of electrical artificial lines and filters", Chapman & Hall's, London, 1930.

[63] M.E. van Valkenburg, "Introduction to modern network synthesis", John Wiley and Sons, New York, 1960.

[64] R.D. Middlebrook, "Differential amplifiers, their analysis and their applications in transistor d-c amplifiers", John Wiley and Sons, New York, 1962.

[65] F.W. Taen, "Het ontwerp van een geïntegreerde voorversterker voor glasvezelcommunicatie in het Thomson proces", internal report, Laboratory of Electronics, Delft University of Technology, Delft, the Netherlands, Dec. 1986.

[66] J.W.Th. Eikenbroek, L.K. Nanver, C.J.M. Verhoeven and P.T.M. van Zeijl, "BEL semi-custom chip design manual", internal report, Laboratory of Electronics, Delft University of Technology, Delft, the Netherlands, 1986.

[67] A. Jhunjhunwala, J.F. Vetelino and J.C. Field, "Temperature compensated cuts with zero power in Tl_3VS_4 and Tl_3TaSe_4", Electronic Letters, no. 25, Vol. 12, Dec. 1976, pp. 683–684.

Summary

In this thesis, the results of a study concerning the fundamental aspects and the design of an FM upconversion receiver front-end with on-chip surface acoustic wave (SAW) filters are presented.

In Chapter 1 several architectures for FM receivers are treated. It is concluded that an upconversion receiver, where the intermediate frequency and the local oscillator frequency are higher than the highest frequency of the received signal, is most suited for the realization of an integrated receiver front-end. Moreover, the realization of silicon-integrated SAW filters is essential to prevent problems associated with pin-to-pin feedthrough.

In Chapter 2 the modeling of the on-chip SAW delay lines and transversal filters is performed from the system and circuits designer's point of view. A voltage-driven and current-sensed SAW device exhibits the optimum signal transfer. In the case of a SAW transversal filter, noise is generated as a consequence of the (acoustic) terminations of the transmission line. In case of a SAW resonator, the series resistance generates noise.

The amplitude and phase transfer functions of SAW transversal filters can be designed independently of one another. In resonators, however, the phase shift occurring in the filter transfer function is directly coupled to the amplitude transfer function. The length of a SAW transversal filter is inversely proportional to the (absolute) bandwidth. A test chip with a number of SAW delay lines, a transimpedance amplifier and a bipolar mixer demonstrates the feasibility of the silicon-integrated SAW devices.

In Chapter 3 the behavior of the dipole and monopole antenna is discussed. It is shown that the antenna impedance exhibits, alternatingly, series and parallel resonances. The antenna impedance can be modeled as a series-resonant network at the first series-resonance frequency. Measurements show that the antenna impedance varies due to changes in the antenna environment. The transfer function from electromagnetic field strength to open-terminal antenna voltage becomes only slightly more efficient for increasing frequency. Therefore, it is concluded that the image rejection in the upconversion receiver is not significantly endangered.

In Chapter 4 several aspects of the FM front-end are discussed. The prob-

lems associated with the second-order Rayleigh-wave mode, as present in the SAW devices in a ZnO-SiO_2-Si structure, can be minimized by the proper choice of the intermediate frequency. A simple LC filter, external to the chip, gives the best suppression of all undesired responses while still a sensitivity of 1 μV can be obtained. In this situation, an amplifier with two feedback loops is essential in realizing the proper filter transfer function without adding significant noise, independent of variations in the antenna impedance.

A SAW transversal filter, used as an IF filter, yields the best receiver specifications. However, this solution requires a very long chip. The models developed for the SAW devices are used to obtain and compare the effective noise voltage and dynamic range of the various IF filter configurations. For a desired minimum effective noise voltage or maximal dynamic range at a given supply current, the width of the SAW devices in a given IC process can be determined.

In Chapter 5 a new class of dual-loop amplifiers is described. In these new amplifiers the disadvantages such as a high noise contribution, high distortion or unsuitability for monolithic integration, usually encountered in the already known dual-loop amplifiers, have been circumvented. Balanced amplifier structures with impedance feedback networks are used to construct these new dual-loop amplifiers. Common-mode stability is ensured by choosing the proper configuration with respect to the available common-mode source or load impedance. The design of one of these new amplifiers, usable as input amplifier in the FM front-end, and the measurements performed on the integrated version, demonstrate the realizable specifications for such an amplifier.

Finally, the results obtained in the present research are summarized in Chapter 6. Besides, some recommendations for further research are given.

Samenvatting

In dit proefschrift worden de resultaten van een studie betreffende de fundamentele aspecten en het ontwerp van een FM upconversie ontvanger front-end met geïntegreerde akoestische oppervlakte-golf filters (SAW filters) besproken.

In hoofdstuk 1 zijn verscheidene architecturen voor FM ontvangers besproken. De conclusie is dat een upconversie-ontvanger, met de middenfrequentie en de lokale oscillator frequentie hoger dan de hoogste frequentie van het ontvangen signaal, het meest geschikt is voor de realisatie van een geïntegreerd ontvanger front-end. Bovendien is de realisatie van geïntegreerde SAW-filters essentieel om problemen ten gevolge van overspraak te vermijden.

In hoofdstuk 2 is de modelering van de geïntegreerde SAW vertragingslijnen en transversale filters uitgevoerd vanuit het standpunt van de systeem en circuit ontwerper. Een spannings gestuurd en kortgesloten SAW device waarbij de kortsluitstroom als uitgangssignaal gebruikt wordt, vertoont de optimale signaal overdracht. In het geval van een transversaal SAW filter wordt de ruis gegenereerd door de (akoestische) afsluitingen van de transmissielijn. In het geval van de SAW resonator genereert de serieweerstand ruis.

De amplitude- en fase-karakteristiek van transversale SAW filters kunnen onafhankelijk van elkaar worden ontworpen. Dit in tegenstelling tot resonatoren waarbij de fase-karakteristiek van het filter direct gekoppeld is aan de amplitude-karakteristiek van het filter. De lengte van een transversaal SAW filter is omgekeerd evenredig met de (absolute) bandbreedte. Een gerealiseerde testchip met een aantal SAW vertragingslijnen, een transimpedantie versterker en een bipolaire mengschakeling toont de haalbaarheid van de in silicium geïntegreerde SAW devices aan.

In hoofdstuk 3 is het gedrag van de dipool en de monopool antenne besproken. De antenne-impedantie vertoont serie- en parallelresonanties die elkaar afwisselen. De antenne-impedantie kan bij de eerste serieresonantie-frequentie worden gemodelleerd door een serieresonant circuit. Metingen tonen aan dat de antenne-impedantie varieert ten gevolge van veranderingen in de antenne-omgeving. De overdracht van de elektromagnetische veldsterkte naar de open klemspanning van de antenne neemt slechts weinig toe bij toenemende frequentie. Daarom mag worden geconcludeerd dat de spiegelonderdrukking van

de upconversie-ontvanger niet significant in gevaar wordt gebracht.

In hoofdstuk 4 worden verschillende aspecten van het FM front-end besproken. De problemen die worden veroorzaakt door de tweede-orde Rayleigh-golf mode, zoals die aanwezig is in een SAW device in een ZnO-SiO_2-Si structuur, kunnen worden geminimaliseerd door een juiste keuze van de middenfrequentie. Een simpel LC-filter dat extern op de chip wordt aangesloten geeft de beste onderdrukking van alle ongewenste responsies terwijl nog steeds een gevoeligheid van 1 μV kan worden bereikt. In deze situatie is een versterker met twee terugkoppellussen noodzakelijk om, zonder extra ruis toe te voegen, de juiste filter-overdrachtsfunctie te realiseren, onafhankelijk van de antenne impedantie.
Een transversaal SAW-filter, dat als middenfrequent filter wordt gebruikt, geeft de beste ontvangerspecificaties. Dit vereist echter wel een lange chip. De modellen die voor de SAW devices zijn ontwikkeld, worden gebruikt om van de diverse middenfrequent filters de effectieve ruisspanning en het dynamische bereik te berekenen en te vergelijken. Voor een gewenste minimale effectieve ruisspanning of maximaal dynamisch bereik kan, bij een gegeven voedingsstroom en IC proces, de breedte van het SAW device worden bepaald.

In hoofdstuk 5 is een nieuwe klasse van versterkers met twee tegenkoppellussen beschreven. In deze versterkers zijn de gewoonlijk in de bekende twee-lussen versterkers optredende nadelen, zoals een hoge ruisbijdrage, hoge vervorming of ongeschiktheid voor integratie, omzeild. Om deze nieuwe twee-lussen versterkers te realiseren worden gebalanceerde versterkerstructuren met impedantie-terugkoppelnetwerken gebruikt. De common-mode stabiliteit wordt verzekerd door de keuze van de juiste versterkerconfiguratie bij een gegeven common-mode bron- of belastingsimpedantie. Het ontwerp van één van deze nieuwe versterkers tezamen met de metingen die aan een geïntegreerde versie zijn verricht, laten de specificaties zien die met een dergelijke versterker haalbaar zijn. Deze versterker is bruikbaar als ingangsversterker in het FM front-end.

Tenslotte worden in hoofdstuk 6 de resultaten van het verrichte onderzoek samengevat. Tevens zijn enige aanbevelingen voor verder onderzoek gedaan.

Acknowledgements

This research was a combined project of the Electronics Research Laboratory and the Electronic Instrumentation Laboratory and was supervised by Prof.Dr.ir J. Davidse, Dr.ir. E.H. Nordholt and Dr.ir. A. Venema. I am much indebted to all of them for their support, which included many fruitful discussions and critical evaluations of this text. My colleague J.H. Visser designed the SAW filters and took care of the technological aspects as described in his Ph.D. Thesis "Surface acoustic wave filters in ZnO-SiO_2-Si layered structures: design, technology, and monolithic integration with electronic circuitry". I am much indebted to him for many discussions and a cooperation which made it possible to complete this project successfully.

I wish to thank all the members of the Electronics Research Laboratory, my fellow Ph.D. students, the graduate students and the members of the Radio Group. Without the support and this stimulating environment, the completion of this thesis would not have been possible.

Further, I would like to thank C.J.M Verhoeven, J.W.Th. Eikenbroek and L.K. Nanver with whom I had the pleasure of developing the semi-custom chip which subsequently was used to demonstrate the feasibility of the new balanced dual-loop amplifiers. J.W.Th. Eikenbroek is thanked for several discussions on radio receivers in general and SAW devices in particular. L.K. Nanver is thanked for her cooperation in the realization of the chip on which the SAW delay lines were combined with the electronic circuitry (ELIS552) and for her final review of the text. I am grateful to L.P. de Jong for his critical evaluations of several versions of this text.

Further acknowledgement goes to W.G.M.M. Straver for making the photographs, the Delft Institute for Microelectronics and Submicron Technology (DIMES) for processing several devices, especially H.W. van Zeijl for processing the ELIS552. A.F.P van Schie expertly bonded these ICs. The 68 pin SLAM houses were exceptionally fast delivered by G.J. Schouten from Semi Dice International. All the figures in this thesis were realized by R.J.H. Janse. My English was corrected by J.B. Zaat-Jones.

This research was financially supported by the Netherlands Technology Foundation (Stichting voor de Technische Wetenschappen - STW) under contract DEL 44.0592.

About the author

Paul T.M. van Zeijl was born in Monster on the 10th of November 1961. He started his studies in 1980 at the Delft University of Technology, Delft, the Netherlands, where he received his Master's degree from the Department of Electrical Engineering on November 21th, 1985. Subsequently he joined the Electronics Research Laboratory at this University and was engaged in a Ph.D. research on the subject of integrated radio receivers with on-chip SAW filters. Several publications, a patent and this Ph.D. Thesis are the result of this Ph.D. research.