Doctoral Thesis

Low power RF-receiver front-end ICs for mobile communications

Author(s):
Piazza, Francesco

Publication Date:
2000

Permanent Link:
https://doi.org/10.3929/ethz-a-004064121

Rights / License:
In Copyright - Non-Commercial Use Permitted
Low Power RF-Receiver Front-End ICs for Mobile Communications

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Technical Sciences

presented by
FRANCESCO PIAZZA
Dipl. El.-Ing. ETH
born 05. January 1962
citizen of Olivone, Ticino

accepted on the recommendation of
Prof. Dr. Q. Huang, examiner
Prof. Dr. W. Bächtold, co-examiner

2000
Contents

Abstract ix

Riassunto xi

1 Introduction 1
   1.1 Structure of the Thesis ........................................... 2
   1.2 History and Evolution ............................................. 3
   1.3 Motivation ......................................................... 8

2 Superheterodyne and Direct Conversion 11
   2.1 Superheterodyne ................................................... 12
   2.2 Direct Conversion .................................................. 22

3 ERMES Pager Receiver Chip 33
   3.1 Introduction ....................................................... 33
   3.2 ERMES System Overview ......................................... 36
       3.2.1 Receiver implications ...................................... 39
   3.3 Antenna Configurations ......................................... 42
3.4 System Planning ........................................ 48
  3.4.1 The Superhet Receiver ............................ 49
  3.4.2 The Direct Conversion Receiver .................. 53
3.5 The ERMES Chip ....................................... 59
3.6 Low Current Operation ................................ 62
  3.6.1 Noise Performance .................................. 62
  3.6.2 Signal Handling Capability ......................... 63
3.7 The Low Noise Amplifier .............................. 64
  3.7.1 The Input Stage .................................... 65
  3.7.2 The Output Stage .................................. 67
  3.7.3 Interstage Filter .................................. 68
  3.7.4 LNA Biasing ........................................ 69
3.8 The Double Balanced Mixer ............................ 72
  3.8.1 Conversion Transconductance ....................... 73
  3.8.2 Mixer Noise ........................................ 74
  3.8.3 Mixer Design ....................................... 80
  3.8.4 Interface to the External Filters ................. 83
  3.8.5 Mixer Biasing ...................................... 84
3.9 The Local Oscillator ................................. 85
  3.9.1 Oscillation Amplitude ............................. 87
  3.9.2 Design of the VCO ................................. 89
  3.9.3 Amplitude Regulation .............................. 92
3.10 The IF Amplifier ..................................... 93
  3.10.1 Interface to the IF Filter ......................... 95
## Contents

3.10.2 IF Amplifier Biasing .......................... 96
3.11 The Quadrature Demodulator ....................... 97
3.12 The $\Delta V_{BE}$ Current Source and Bandgap Reference ... 99
3.13 The ESD Protection .............................. 102
3.14 Layout Considerations ........................... 104
3.15 Chip Mounting and Testing ......................... 106
3.16 Measurements .................................. 109
   3.16.1 Voltage Reference and Current Source ........ 109
   3.16.2 LNA Gain and AGC Range .................... 110
   3.16.3 LNA Input Impedance ....................... 111
   3.16.4 Mixer Conversion $g_m$ and Compression Point . 112
   3.16.5 RF Front-End Noise Figure .................. 113
   3.16.6 VCO Phase Noise ............................. 114
   3.16.7 IF Amplifier Gain and AGC Range .............. 115
   3.16.8 Demodulator Gain and Linearity ............... 116
   3.16.9 Output Filter Frequency Response ............. 117

4 64/65 ESCL Dual Modulus Prescaler 119
   4.1 The Frequency Synthesizer ....................... 121
   4.2 Prescaler Architecture .......................... 122
      4.2.1 The ESCL Technique ....................... 126
   4.3 Circuit Description ............................ 130
      4.3.1 Prescaler ................................. 130
      4.3.2 ESCL logic ............................... 131
4.3.3 ESCL to CMOS level converter ............................................. 133
4.3.4 CMOS logic ........................................................................ 134
4.3.5 $\Delta V_{BE}$ current source .................................................. 134
4.4 Simulations and Measurements .............................................. 136
  4.4.1 Bias Current Source ............................................................ 139
  4.4.2 Maximum Operating Frequency ......................................... 139
  4.4.3 Power Consumption and Minimum Operating Supply Voltage .................................................. 140
  4.4.4 Power Supply Noise ............................................................ 140
  4.4.5 Prescaler Performance at Different Bias Currents ............... 141

5 GPS Receiver Chipset .................................................................... 143
  5.1 Introduction ........................................................................... 143
  5.2 GPS System Overview ........................................................... 147
    5.2.1 Receiver implications ....................................................... 150
  5.3 The Active Patch Antenna ...................................................... 154
  5.4 System Architecture ............................................................. 156
    5.4.1 Block Diagram ............................................................... 157
    5.4.2 System Planning ............................................................ 159
  5.5 The Low Noise Amplifier ....................................................... 163
    5.5.1 The Input Stage ............................................................ 164
    5.5.2 The Output Stage .......................................................... 165
    5.5.3 LNA Biasing ................................................................. 167
  5.6 The RF Single Balanced Mixer .............................................. 168
<table>
<thead>
<tr>
<th>5.6.1 Interface to the IF Filter</th>
<th>170</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.7 The Local Oscillator</td>
<td>170</td>
</tr>
<tr>
<td>5.7.1 Design of the LC resonator</td>
<td>172</td>
</tr>
<tr>
<td>5.8 The 1st IF Amplifier and Mixer</td>
<td>174</td>
</tr>
<tr>
<td>5.9 The 2nd IF Strip</td>
<td>175</td>
</tr>
<tr>
<td>5.9.1 The 2nd IF Amplifier and Filter</td>
<td>176</td>
</tr>
<tr>
<td>5.9.2 The Differential Amplifier Stages</td>
<td>177</td>
</tr>
<tr>
<td>5.9.3 The Fully Differential OpAmp</td>
<td>179</td>
</tr>
<tr>
<td>5.9.4 The Latched comparator</td>
<td>180</td>
</tr>
<tr>
<td>5.10 The Frequency Synthesizer</td>
<td>182</td>
</tr>
<tr>
<td>5.10.1 The Divide-by-8 ECL Prescaler</td>
<td>183</td>
</tr>
<tr>
<td>5.10.2 The ESCL and CMOS dividers</td>
<td>184</td>
</tr>
<tr>
<td>5.10.3 The Phase Comparator</td>
<td>185</td>
</tr>
<tr>
<td>5.11 The Reference Crystal Oscillator</td>
<td>186</td>
</tr>
<tr>
<td>5.12 The $\Delta V_{BE}$ Current Source and Bandgap Reference</td>
<td>190</td>
</tr>
<tr>
<td>5.13 The ESD Protection</td>
<td>191</td>
</tr>
<tr>
<td>5.14 Layout Considerations</td>
<td>193</td>
</tr>
<tr>
<td>5.15 Chip Mounting and Testing</td>
<td>195</td>
</tr>
<tr>
<td>5.16 Measurements</td>
<td>196</td>
</tr>
<tr>
<td>5.16.1 $\Delta V_{BE}$ Current Source and Bandgap Voltage Reference</td>
<td>197</td>
</tr>
<tr>
<td>5.16.2 Front-End Gain and Linearity</td>
<td>197</td>
</tr>
<tr>
<td>5.16.3 Front-End (LNA) input Impedance</td>
<td>199</td>
</tr>
<tr>
<td>5.16.4 RF Front-End Noise Figure</td>
<td>199</td>
</tr>
</tbody>
</table>
Abstract

In this thesis, the design and implementation of two integrated RF receivers for wrist-watch applications is presented. Because of the applications inherent characteristics, achieving low power consumption, small size and low cost was obviously of major importance, as was ensuring that receiver performance was comparable to common hand-held receivers.

Such goals can only be achieved through careful design starting at system level. A "standard cell" approach cannot be applied in this case and thus, a conventional top-down design style has been preferred. The architectures and circuits were selected in response to the characteristics they needed to fulfill and not because of their novelty value. Particular value has been given to receiver performance which is predictable and can be reproduced, as well as its independence from temperature and process tolerances. Only widely available, relatively low cost CMOS and BiCMOS technologies have been used in this work.

The first chip described contains the circuits required to build an RF front-end for ERMES watch pagers and is implemented in 1.2\,\mu m BiCMOS. Its purpose was to demonstrate the feasibility and performance of such a pager. For this receiver, both the single superheterodyne and direct conversion architectures have been considered, while all critical receiver blocks, i.e., LNA, mixer, local oscillator, IF amplifier and demodulator, have been included. A dual modulus prescaler, implemented on a separated 1.2\,\mu m CMOS chip, complements this front-end.

The second circuit described is a chipset for a GPS receiver. Its immediate application is to provide GPS time reference to a wrist-watch. This chipset is a prototype of a complete triple conversion superhet receiver for the GPS L1 band. The superhet architecture was chosen mainly to achieve low power consumption, while the 2nd IF filter has been integrated to reduce complexity.
Seite Leer /
Blank leaf
Riassunto

In questa tesi sono presentati il progetto e l' implementazione di 2 ricevitori radio integrati per orologi da polso. A causa della particolare applicazione, l' ottenimento di un basso consumo, piccole dimensioni e basso costo era un requisito essenziale, oltre alla necessità di garantire delle prestazioni simili a quelle di comuni ricevitori portatili.

Dei simili obiettivi possono solo essere ottenuti con un progetto accurato a partire dalla pianificazione del sistema. Un approccio "standard cell" non può' essere usato in questo caso, quindi un convenzionale stile top-down è stato preferito. La scelta di architetture e circuiti è stata dettata solo dalle caratteristiche necessarie, e non e.g. novità o simili. Un valore particolare è stato dato alla prevedibilità e alla riproducibilità delle prestazioni, come pure alla loro indipendenza dalla temperatura e dalle tolleranze del processo. Per questo lavoro sono state utilizzate solo tecnologie CMOS o BiCMOS disponibili a un costo relativamente basso.

Il primo chip descritto contiene una collezione di circuiti necessari per la costruzione di un front-end RF per pagers ERMES per orologi ed è implementato in BiCMOS 1.2μm. Il suo scopo era dimostrare la fattibilità e le prestazioni ottenibili da un tale pager. Per questo ricevitore sono state considerate le architetture single superheterodyne e direct conversion, mentre tutti i blocchi critici, ossia LNA, mixer, oscillatore locale, amplificatore IF e demodulatore, sono stati inclusi. Un prescaler dual modulus, implementato su un chip CMOS 1.2μm separato completa questo front-end.

Il secondo circuito descritto é un chipset per un ricevitore GPS, la cui applicazione immediata é come riferimento di tempo GPS per orologi da polso. Questo chipset é il prototipo di un ricevitore superhet a tripla conversione completo, per la banda GPS L1. L' architettura superhet é stata scelta principalmente per ottenere un basso consumo di potenza, mentre il filtro per la seconda IF é stato integrato per ridurre la complessità del ricevitore.
Chapter 1

Introduction

The recent innovations in wireless communication technologies are totally revolutionising our telecoms industry. Indeed, huge leaps have been made since radio receivers were big wooden boxes filled with vacuum tubes and telephones were made of black bakelite and only available from Bell (USA) or domestic telephone companies which calmly profited from their monopolies (Europe).

Wireless communications have captured the public's imagination. They promise easy access to all manner of information, 24 hours per day, seven days a week anywhere in the world. One's location is no longer of any significance [1, 2]. Today's wireless communications support many different services. Apart from the more obvious examples such as broadcast radio or walkie-talkies, the most common are e.g. cellular and cordless phones, pagers and GPS navigation, while integration with computers or PDA's has made e.g. wireless fax and Internet access possible.

The concept of wireless communication is not new. Commercial radio services, as well as radio amateurs (hams) have been active since early 1900 [3, 4]. However, the high cost of the equipment and the difficulty of operating it - most people are not used to Morse code or the "over - over and out" protocol - restricted its use to only a few, specialized applications. Today, thanks to the evolution in microelectronics and digital computing, system performance, range of services,
user-friendliness and in particular the price of such equipment have improved substantially. As people became more aware of these benefits, demands for these services increased dramatically leading inevitably to a large and highly competitive market. Wireless phones and pagers are now commonplace, while the demand for newer services or smaller and more flexible terminals is growing steadily. In most Western countries the popularity of mobile communications is such that networks have almost reached saturation point. As a result, new frequencies are being allocated and a great deal of research is being performed on efficient multiple access schemes [1].

The necessity to introduce competitive products on the market has resulted in new opportunities but also in high pressure on radio (especially IC) designers, who are asked to satisfy the requirements of low cost and high performance. It is indeed possible that different techniques from the ones commonly used today may be required in the near future [2]. This does not mean that all old techniques are to be dismissed. Some (such as digital modulation, the direct conversion receiver or the balanced active mixer) that were impractical in discrete implementations, have recently been revived for IC design because of their high performance, simplicity or the high level of integration achievable by their use.

1.1 Structure of the Thesis

This thesis describes the design and implementation of two integrated communications receivers for consumer applications, namely the ER-MES pager and GPS, aimed at wrist-watch applications. Our mission was to achieve lower power, lower cost, smaller size and a smaller number of external components than current solutions, without compromising receiver performance.

In Chapter 1 a historical overview of the development of the radio and the motivation for this work is provided. This section is intended to place our work in its context, and lays no claim to be a complete or detailed treatise on the history of the radio.

The architectures considered for the receivers described in this thesis, i.e. Super-heterodyne and direct conversion, are described from a
1.2. History and Evolution

The invention and evolution of radio is a long story, spanning more than a century [3, 4]. It all began with the theoretical studies of Faraday, Ampere, Henry and especially J. C. Maxwell, who first predicted the existence of electromagnetic waves and stated the laws governing electromagnetic radiation (1864). It took 23 years before H. Hertz actually generated them in an experiment (1887), and another 8 years before G. Marconi put them into practical use by sending the first telegraphic signals over more than one mile (1895). In the following years, thanks to improvements in particular to antennas and detectors, the achievable range of wireless communications increased steadily. During these years, E. Branly also proposed the use of the word "Radiotelegraphy" instead of "Wireless telegraphy" (1897). In 1901, messages were transmitted from England to the USA making radiotelegraphy ready for commercial use.

One of the first, and perhaps most obvious commercial applications...
of radiotelegraphy was to enable communication with ships. Until that
time, ships had possessed no effective means to communicate with each
other or with land if they were out of sight. Their only option was to
follow their route and hope that nothing untoward would happen. In
the early days, radiotelegraphy was also seen by some radio operators
(especially G. Marconi) as a competitor to wired telegraphy. He believed
it would completely replace the latter within a few years. This belief
shows how at that time, operators were simply unaware about the huge
demand that would follow, and that - with bandwidth being a very
limited resource (only long waves were used at that time) - they would
never be able to fulfill that demand. Only years later were the two
services to complement each other, as we know today.

In the early days of radiotelegraphy, the only propagation mecha¬
nism known was the ground wave, meaning that generally only long
waves below 100kHz were used. This was not due to limitations of the
equipment available at that time - the first experiments by Hertz and
Marconi were carried out using UHF - but engineers simply believed
that attenuation would increase steadily with frequency, making higher
frequencies useless for anything other than laboratory experiments. The
role of ionospheric reflection as a propagation mechanism was not yet
understood, but nevertheless used unwittingly in intercontinental com¬
munications [5].

In the absence of any electronic device, the typical transmitter was
a quenched spark, a (Poulsen) arc transmitter or a high frequency al¬
ternator directly supplying the antenna. The first had the disadvantage
that it generated bursts of damped RF oscillations thus occupying a
bandwidth much larger than necessary, while the other two transmitters
generated continuous waves (CW). The receiver, too, was very simple.
No amplification was possible, and the typical receiver consisted of a
channel selecting tuned circuit and a detector. The most commonly used
detectors were the coherer (E. Branly, 1890) and semiconductor crystals
(Pickard, 1906) such as silicon, silicon carbide (carborundum) or lead
sulphide (galena). The heterodyning principle was already known, but
was not used in receivers for frequency conversion. One of its uses was
in alternator-based transmitters as a frequency multiplier (the mixer
was a saturated core transformer), while R. Fessenden proposed it in
1902 as a means to make the CW Morse signals audible as a beating
signal with a local oscillator. Shades of direct conversion!
1.2. History and Evolution

Typical modulation schemes, or keying (from the Morse code key, the term modulation was not yet in use) were necessarily digital. Amplitude shift keying (ASK) was the most common, and was obtained by breaking the DC supply of the spark. Frequency shift keying (FSK) was introduced later, partly because the high power Poulsen arc transmitters were difficult to key on and off, and partly as a means against intruders or pirate listeners. FSK was generated by shorting out a few windings of the antenna tank with the Morse key [5].

With the invention of the vacuum tube - the diode in 1904 by A. Fleming, the triode (formerly Audion) in 1907 by Lee DeForest - and the improvements to it by H. Arnold, who in 1912 recognized the need for a high vacuum for satisfactory operation, RF activities started to grow exponentially. An efficient detector was finally there, while amplification was eventually possible. This allowed sensitive receivers to be constructed [6] which, together with the invention of the electronic oscillator by E. Armstrong and L. DeForest, made the high powered but cumbersome and unstable arc transmitters obsolete. It was quite clear that tubes were here to stay. Even today (1990's) many types of tubes are still produced and used in some specialized applications such as high power and microwave transmitters [7], displays (CRTs, VFDs, FEDs [8]), microwave ovens and guitar amplifiers [9].

Analog modulation types, and thus radiotelephony, also became viable thanks to the vacuum tubes. Some radiotelephony experiments were already being made using arc transmitters, but with unsatisfactory results. AM was the most common modulation type. FM (developed by E. Armstrong) came only in the year 1933. The first commercial FM receiver was a GE set from 1938. A new variety of services (such as radio-teletype (RTTY), tele-printer, facsimile, etc.) was to be deployed, the most important of which being perhaps broadcast radio. The first radio-navigation experiments by R. Fessenden were also conducted during this period (1912-15).

Most (if not all) known receiver architectures and techniques we know today were invented before 1935. The oldest receiver architecture is the Tuned RF (TRF) receiver, which dates back to the origins of the radio and consists simply of narrow-band filters or amplifiers, tuned to the frequency to be received. It is still used today in some special, mainly long wave, applications such as radio controlled clocks and frequency standards. An improved TRF receiver came in the form of the
Neutrodyne (L. Hazeltine, 1918), essentially a TRF using neutralization to avoid spurious oscillations. The (Super)Regenerative receiver first comes to light in 1912 as a means of increasing sensitivity and selectivity of a TRF receiver. It is still used today in toys and simple remote controls because of its simplicity. The ubiquitous Superheterodyne had its origins in the early 1910s, but was fully developed (and patented) in 1920 by E. Armstrong. The first commercial superhet receivers came to the market in 1924. The Low-IF and Direct Conversion receivers also had their origins in the early 1910s. The first description of a modern direct conversion receiver, as we know it today, is from 1924. A patented commercial application followed in 1934 [10]. Many techniques commonly used today in receivers were also developed during these years. The automatic gain control (AGC), formerly automatic volume control, was invented in 1926 (H. Wheeler), Diversity reception in 1928, and the Phase Locked Loop (PLL) in 1932 [11], to name but a few.

Around 1935, the majority of RF activity seems to be concluded, and no substantial innovation on classical radio research can be seen during the following years. With the approach of World War II, and during the war itself, researchers and engineers were more concerned with the development of new weapons, or military applications of radio waves. Also companies that were little involved with the military followed that way, as can be seen e.g. in [12]. For the communications industry, cryptography was perhaps the main focus, while microwaves and microwave tubes (klystron, magnetron) became available to it as a by-product of radar development.

For the average person, perhaps the most visible evidence of these innovative years was television, which became widely used during the fifties. The single, biggest event of these years, which resulted in another even bigger revolution in electronics in general and communications in particular, was the invention of the transistor (W. Shockley, J. Bardeen, W. Brattain, 1948) [13]. Although the performance of early transistors was quite poor, small, low-power portable electronics eventually became a reality.

During the fifties and early sixties the first modern wireless activity, i.e. pagers and wireless phones can be seen. The first pager system (covering only a single building) was installed by Multitone Electric in 1956 in a hospital in London, while wide-range systems followed a few years later. Two such examples were the Swiss Autoruf CH (1958) or
1.2. History and Evolution

the AT&T Bellboy (1961). "Portable" telephones (they were the size of a large suitcase) came next. These cumbersome devices heralded the start of the "wireless revolution", as we know today.

Although the seventies and eighties saw the wireless communications market grow steadily, it was the nineties that brought wireless communication into the mainstream. As recently as 1988, engineers were still talking about the failure of cellular communications. Advances in IC technology, driven by the microprocessor industry, made possible the development of the smaller, lower power, lower cost solutions that are required to implement competitive products in a demanding market. Wireless telephones have been scaled down from large suitcases to sub-cigarette packet size, while pagers can easily fit into a watch.

Integrated communication circuits started to appear during the seventies. First it was simply a few integrated IF-strips, then complete AM/FM radios, and now it is the turn of single-chip GSM, DECT or GPS receivers.

The old term 'wireless' has also been revived, and replaces 'radio' in all modern communications applications. With the availability of powerful and cheap processors, radio is becoming digital - or should we perhaps say "Digital wireless is back"? Spark transmitters were indeed digital and Morse code (1835) is a digital code that uses entropy coding to achieve some data compression. These innovations precede modern digital signal processing by more than 100 years [14]. The word 'keying' is also back, to describe our modern digital modulation schemes.

Despite all this apparent innovation, the techniques currently used in communications receivers emerged pre-1935. Of course, the industry did not stand still, but innovation focussed specifically on offering higher quality, lower cost receivers or the introduction of new services, a situation that still persists today. During the sixties and seventies the decline of radio research was significant; few universities offered radio engineering courses, few books on RF were written, and - were it not for portable communications - RF could have become a lost art [15]. Even today, many universities are still not teaching RF courses while only a relatively small number of companies have the capability to do solid RF work.
1.3 Motivation

For most communications terminals, the quest for miniaturization, performance, features or low cost concerned mainly the digital or base-band parts. The RF part on the other hand, shows little evidence of change over the years, even considering what was said in section 1.2. Open a recent (mid nineties) pager or cellular phone and you will find at most a few VLSI digital chips controlling everything from signal processing to the user interface, and one base band processor. On the other side, the RF part is still made up of a selection of separate transistors and passive components, some of them requiring adjustments [2]. A closer inspection of the circuits, reveals that they have remained relatively unchanged during the last 40 years or so. Typically, modern components are used in rather conventional circuits (e.g. high Q tuned amplifiers), that have a good performance and are well understood but rely heavily on precision and tunability of separate components. Attempts at integrating anything more than is strictly necessary are rare. The integration of RF components is in most cases limited to IF chips, the PLL and pre-scalers (actually digital functions) and, in some cases, to some SSI RF functions (e.g. [33]). The most probable reasons for this are:

- The performance of the most widely used receiver architectures is still very good indeed and well-proven.

- The operating frequency of most modern communications equipment was too high for the low cost technologies needed for consumer applications.

- Downscaling passive components, e.g. SMD, has often solved the size problem, preventing new structures with low component counts to take over.

The requirements of most modern receivers, brought about by wireless telephony and data communications, are very different from those of standard RF applications such as AM/FM broadcast radios or TV tuners, where there is plenty of space and the cost of the tuner is only a small portion of the total. Conventional structures widely used today may no longer suffice for the new applications which require more compact and cheaper transceivers. Miniaturization required for example
by pocket cellular telephones and wrist-watch pagers exceeds the limits achievable by simply reducing the size of conventional structures, while the high pressure on the price of such devices calls for highly integrated RF circuits that are easy to adjust. On the other hand, to maintain user acceptance, the new devices must offer at least the same high performance as the conventional ones. A rethinking of the whole receiver architecture, not just its implementation may therefore be required. The operating frequency of most RF consumer applications is now within the reach of modern silicon technologies while digital transmission schemes allow for compensation of some receiver weaknesses using appropriate algorithms. To summarize, the most important problems to be solved during the design of RF front-ends for newer applications concern low power, small size and low cost.

The purpose of this thesis is the study and development of RF front-end circuits for two applications, namely ERMES pager and GPS. These front-ends shall be used in wrist-watch style receivers, thus they shall consume less power and use less external components than current implementations. Because of user acceptance and type approval requirements, these goals shall be achieved without compromising receiver performance. To avoid major failures, an evolutionary approach will be taken, that is by first studying and achieving a good understanding of system requirements and existing techniques and then to find ways for possible improvements. In general, circuits more suitable for integration than the ones commonly used in discrete designs will be used in conventional, but well performing pre-1935 receiver architectures. In other words, innovation will not be forced. The various circuits will be developed to achieve a given performance in the given system and not for the sake of having something new and untried. In recent publications, many radically new RF circuits have been described. In almost all cases, these circuits have very low performance, unsuitable for the applications envisaged in this thesis. This is sometimes due to intrinsic problems or errors (the sub-sampling mixer is a good example), and they will be discarded from the very beginning. No attempt at improving what are usually hopeless circuits will be made in this thesis.

For cost and availability reasons, no overly-advanced or developmental technologies will be used in this work. Only standard, easily available processes will be considered, provided their performance is adequate for our applications. No special technological characteristics or
non-standard processing steps will be exploited to improve the circuits - this is to avoid excessive costs, yield or reliability problems. Portability will not be hindered nor shall related difficulties be introduced by packaging or mounting. The principles shown here will still be valid also if porting to a better, more modern technology is desired. Such a technology may surely improve the performance of many of the circuits presented here, which can be easily adapted or redesigned to different specifications.
Chapter 2

Superheterodyne and Direct Conversion

The purpose of this chapter is to discuss the two receiver architectures that were considered for our designs - i.e. superheterodyne and direct conversion - from a technical point of view, and their implications if an IC implementation is envisaged.

The choice of an architecture for a given application depends on many factors, such as for example:

- required performance,
- cost,
- size and weight,
- complexity,
- power consumption.

Unfortunately no architecture exists that optimizes all the desired parameters simultaneously, and the best trade-off for the given application must be found. Among all possible architectures, the superheterodyne was chosen for the GPS receiver, while for the ERMES pager direct conversion has also been considered.
2.1 Superheterodyne

A superheterodyne receiver makes use of the heterodyne principle of mixing the incoming signal with a local oscillator, to obtain a third signal at a different frequency. The heterodyne principle was already known as a means to generate high RF frequencies (still is today) and to make CW signals audible (beat frequency oscillator, BFO), but it was Edwin Armstrong who fully developed the superheterodyne receiver in 1920. The idea behind superhet is to separate tuning from channel selection, and obtain selectivity and most of the amplification at a relatively low, fixed frequency.

The block diagram of a superhet receiver is shown in Fig. 2.1. The incoming RF signal, after being filtered to remove the image (see below) and amplified by the low noise amplifier (LNA), is converted to a fixed, usually lower intermediate frequency (IF), by mixing it with a sinewave $f_0$ generated by a local oscillator. At IF, channel filtering, as well as most amplification occur. The amplified signal is finally detected, to obtain the demodulated base-band signal.

In a mixer, the two frequencies $f_L = f_0 - f_{IF}$, and $f_U = f_0 + f_{IF}$ are both converted to the same IF (Fig. 2.2). The preselection filter in front of LNA and mixer is therefore required to select the desired frequency and suppress the undesired one, called image frequency. The same filter will also filter off strong out-of-band interferers, that could otherwise saturate the front-end of the receiver. Since the image lies at $f_{RF} \pm 2f_{IF}$, the preselection filter can be much broader than one channel. Relatively simple crystal, SAW or fixed LC filters covering the whole bandwidth of the receiver are sufficient in most cases. In wideband
receivers or where large signal power differences are likely, such as in broadcast AM, FM or TV receivers, variable LC filters ganged with the LO resonator are used. Since also in this case the filters do not need to be particularly selective, alignment of such ganged filters is easy and a simple two-points alignment will ensure sufficient tracking in most applications.

Tuning is accomplished by simply varying the local oscillator frequency. This is easily obtained with a variable capacitor or a programmable PLL. To reduce the tuning range, the LO frequency is usually chosen to be above the desired RF band, although reasons may exist to choose otherwise. An asymmetrical modulation spectrum that cannot be reversed (e.g. TV) or the known presence of strong interferers at the upper image frequency are two examples. Since any noise on the LO signal will be converted to IF with the modulated RF signal, and made indistinguishable from it, a clean and stable signal is required from the LO. A crystal, or a high Q LC oscillator are in general suitable, while ring oscillators, 1st order RC oscillators or other low Q oscillators are only suitable in wideband systems, or if very high power consumption can be allowed.

Channel selection, as well as most of the amplification, are provided by the IF strip. This is a decisive advantage over other architectures, especially if the receiver must cover a wide frequency range. Obtaining a well defined bandwidth in a fixed filter is in fact much easier than in a tunable filter, while because of the lower IF frequency, channel selection is achieved without the need of filters with very narrow relative
bandwidth, i.e. large Q. A variable bandwidth, e.g. to accommodate different data rates or modulation methods, can be easily accommodated by changing the IF filter. Depending on required performance, many different filter types can be used as the channel filter, the most common being ceramic or LC.

Obtaining a high and stable gain at a low fixed frequency is in general easier and more economical - especially in terms of power consumption - than at a high, variable RF. At low frequencies the IF amplifier can provide most of the gain required by a receiver while consuming only a small amount of power. AGC is also implemented by the IF amplifier, if required. Usually 60-80dB of AGC range is provided, while at RF the gain is only reduced in the presence of very strong signals, to prevent mixer overloading (delayed AGC). Alternatively, for constant envelope modulations and especially if a simple discriminator with poor AM rejection is used as the demodulator, amplitude limiters can be provided. These maintain a constant signal amplitude (thus a constant demodulator sensitivity) and remove most additive (AM) noise.

Several considerations must be made for the choice of a proper IF frequency for a given application. The proper choice must balance between image rejection and IF strip requirements. A low IF frequency, while simplifying channel selection and the design of a stable high gain IF strip, may complicate the design of the image filter because the image will lie very close to the desired RF channel, while the opposite is true if a high IF frequency is chosen. Choosing an IF frequency approximately between 1/5 and 1/20 of the desired RF frequency will result in a well balanced design in most cases. If a standard IF frequency such as 455kHz, 10.7MHz or 21.4MHz can be chosen, interferences at IF are avoided in a less than well shielded receiver. In fact these are protected frequencies, and no transmitter will send anything on them.

Many demodulator types can be used with a superhet receiver. In analog broadcast receivers, the most common types are an envelope AM detector or a quadrature FM discriminator, while quadrature demodulation is rarely used. The latter is the predominant demodulator type for all digital communications receivers. The AGC signal may be derived from the demodulator directly, from a separate detector with a long time constant, or generated by the base-band signal processing.

Although the superhet receiver introduces a spurious image response
not present in other receiver types, its performance is excellent. It has therefore superseeded other architectures in most applications. Nearly no technical problems exist in a well designed superhet receiver; the system has consistent performance, it has very good sensitivity and selectivity, its gain is partitioned over different frequencies thus a large overall gain can be obtained without incurring stability problems, while LO reradiation - a potential nuisance - lies well out of band and is easily rejected by the image filters and the LNA ($S_{12}$). If ceramic filters can be used throughout, alignment of the receiver is very simple. Laser trimming can be used in high volume productions to reduce alignment cost. Because of the relatively low Q of the image filters, alignment is not critical even if tunable filters are required. As said before a two-point alignment will yield sufficient tracking even in wideband systems despite the difference in resonant frequencies of image filters and local oscillator. Long-term stability is also quite easy to guarantee for the same reasons.

Power consumption in a receiver is always dominated by the parts operating at the highest frequency, especially the LNA and the mixer. A relatively large bias current is in fact necessary for these parts, to fulfill either bandwidth, noise or linearity requirements. By minimizing their number (one LNA, one mixer, a local oscillator and, in digitally controlled receivers, a prescaler), power consumption is also minimized when compared to the other architectures\(^1\). This may be a fundamental advantage in battery-powered portable equipment, since the largest, heaviest and most expensive item of such a receiver is very often the battery itself [23].

Excellent flexibility characterizes the superhet receiver, that finds applications in just almost every thinkable system, from near DC to GHz and more, from pocket radios to TV, HiFi, wireless communications, analog, digital, military, satellite, or deep space applications. Considerable experience in designing superhet receivers exists in most RF companies, and the architecture is well understood and well proven in billions of sold parts.

The biggest drawback of superhet lies in its complexity. Many filters and resonators are required, that will contribute substantially to the size and the total cost of the receiver. The size of ceramic and SAW

\(^{1}\)With the possible exception of the superregenerative receiver, whose performance is however low and inadequate for most applications.
filters cannot be reduced easily, while their price is already near to a minimum, with no real perspective for a further, substantial reduction. Since high quality filters and resonators are difficult to integrate in an IC, a low integration level has to be expected. In general, a single-chip superheterodyne receiver contains all the active parts, with the filters and resonators left off-chip. A large package is usually necessary to accommodate the many pins required to connect filters and other external parts to the integrated portion of the receiver, which may result in coupling between traces or difficulties in PCB placement. The benefits of a single-chip implementation of the receiver on miniaturization may therefore be less than expected, while a significant portion of PCB real estate may be wasted by the required traces. In some cases a different partitioning of the receiver (e.g. a separate LNA) may be a better choice, and may reduce such problems.

Many variants of the superheterodyne receiver exist. Some of these are:

**Multiple conversions:** In many receiver types such as wideband or VHF/UHF receivers, especially if densely packed narrow-band channels are used, a single conversion design may prove impractical, in terms of filter requirements, because of the high selectivities that would be required. In such cases, two or more conversions can be used, in order to convert progressively to the last, lower IF where channel filtering and large amplification are performed. Although multiple conversions will add new images to receiver response, these are well spaced from the desired RF channel and as such are easily removed. At the higher IFs, the filters need only remove the images generated by the subsequent conversions, thus low Q filters can be used. A well balanced design is obtained, also in this case, if the frequency is converted by a ratio approximately between 1/5 to 1/20 of the previous frequency. Some examples may be a dual conversion ERMES pager with 1st IF of 10.7MHz and 2nd IF of 455kHz or a DCS-1800 cellular phone with 1st IF of 246MHz and 2nd IF of 21.4MHz, as well as the GPS receiver described in this thesis, that uses triple conversion. A coherent demodulator is not considered a conversion in this sense, although it consists of two mixers that convert to base-band.

**Upconversion:** A wideband receiver requires both a local oscillator and image filters tunable over a large frequency range. Practical problems will very likely arise, and switchable filters and oscillator resonators may become mandatory. An elegant solution is the use of upconversion.
In such a receiver, the incoming RF signal is converted to an IF that is higher than the input RF signal. As the result, LO tuning range is greatly reduced, while the images will lie so far away from the desired RF, that the image filter is - at least formally - a simple, fixed, low-pass filter. In an actual design, tunable filters may still be required to remove strong out-of-band interferers, but their requirements will be quite easy to fulfill. Downconversion and detection as usual complete the receiver. Some applications of upconversion are TV tuners that cover from VHF band-I to UHF band-V, including CATV channels, in one step, or communications/ham receivers that have a continuous coverage from near DC to e.g. 30MHz (10m ham band). Superhet spectrum analyzers will also use upconversion to cover the lower frequency range (and frequency converters for the rest). The main disadvantage of upconversion is the increased power consumption caused by the high 1st IF.

Low IF, Offset Receiver: It is basically a superheterodyne receiver where a very low intermediate frequency just slightly bigger than the (maybe compressed) signal bandwidth is chosen. Channel filtering can therefore be made with a simple integratable lowpass filter, while AC coupling can be used throughout to obtain a bandpass characteristic and remove DC offsets. Because of the low IF frequency, the image lies in-band very near to the desired channel, thus image rejection can no longer be performed by a bandpass filter at RF as usual. An image reject mixer can be used, or if some performance degradation is acceptable, the IF can be chosen in such a way that the image lies exactly between two channels where ideally only noise exists. At best a 3dB noise figure increase will result, much more in a practical case because of the sidebands of the two channels adjacent to the image. Apart from noise figure worsening, not performing image rejection results in double reception of the same channel, that appears at two different LO frequencies (i.e. dial readings), making this principle only useful for very low performance systems (in principle only toys), or one-channel systems with moderate specs (e.g. POCSAG pagers). If image rejection is performed a low IF receiver is suitable also for higher performances communications applications, but its complexity and power consumption then becomes comparable to direct conversion.

Image reject mixer: Image rejection can be achieved, without using external filters, by the image reject mixer shown in Fig. 2.3. It consists of two mixers that receive quadrature local oscillator signals. If
\( \omega_0 \) is the local oscillator frequency, \( \omega_L \) is the desired RF signal and \( \omega_U \) is the image, after mixing (neglecting the sum terms) the output of the mixers is:

\[
I = k \cdot (\cos((\omega_0 - \omega_L)t) - \cos((\omega_U - \omega_0)t))
\]

for the I channel, and:

\[
Q = k \cdot (\sin((\omega_0 - \omega_L)t) + \sin((\omega_U - \omega_0)t))
\]

for the Q channel respectively. \( k \) is a factor that includes input signal amplitude and conversion gain of the mixers. After 90° phase shifting, the I and Q signals are summed together to form the IF signal:

\[
IF = k \cdot (\cos((\omega_0 - \omega_L)t) - \cos((\omega_U - \omega_0)t) + \\
\cos((\omega_0 - \omega_L)t) + \cos((\omega_U - \omega_0)t))
= 2k \cdot \cos((\omega_0 - \omega_L)t)
\]

Although in this ideal example the image is completely gone, normal mismatches between the I and Q channels and phase errors in the phase shifters limit image rejection to about 25-35dB. If the targeted application demands more, some additional filtering must be provided (usually before the LNA, which also protects it from strong out-of-band interferers), nevertheless the interstage filter is not needed anymore. A drawback of this architecture, apart from the higher complexity, is the larger power consumption resulting from the many parts working at the highest frequency. Since the image rejection principle is also used by the direct conversion receiver, the same design considerations described in section 2.2 apply also here.

**Frequency converter:** Strictly speaking every mixer stage is a frequency converter, but conventionally what is meant under this name is a superhet front-end that is used to convert a whole frequency range - not just single channels - to a wideband IF. In contrast to a usual superhet, the local oscillator is operated at a fixed frequency, or switched between a few discrete frequencies, while the IF is variable. Channel selection is obtained by choosing the IF where the desired channel appears. Frequency converters were once (\( \approx \) 30-60s) widely used as a means to extend the frequency range of existing receivers to cover new frequency bands as they were released for broadcast or ham usage. The frequency range of the existing receiver was therefore used as the variable IF. Additional benefits realized by the use of frequency converters
2.1. Superheterodyne

Figure 2.3: *Block diagram of an image reject mixer.*

were that - thanks to the extra gain added by the converter itself - they improved the performance of early TRF and superregenerative receivers considerably, and especially that they eliminated the highly annoying reradiation of the latter. Reradiation is in fact far off the received frequency, and as such is easily eliminated by the image filters and the reverse isolation of the converter. Today frequency converters are still widely used e.g. in spectrum analyzers and synthesized RF generators to cover the higher frequency bands, as well as in many microwave communications and broadcast applications, while in typical communications transceivers (in the low GHz bands) it is quite rare to see one in operation. Recently it has reappeared under the name *wideband IF*, or *variable IF* receiver, just a new name for an old device [16].

**IF AD conversion, digital radio**: Digital signal processing has always been an attractive method to perform operations usually done by analog circuits. Many operations used in radio work, such as mixing and detection can be made in the digital domain as well, while several others (e.g. correlation, selective linear-phase filters, noise removal or Fourier transform) are performed much easier as digital rather than analog operations. In a typical communications receiver, the use of DSP may allow the implementation of ideal demodulation methods and algorithms, thus achieving better performance than their analog counterparts, but will require AD conversion of the IF signal. Practical reasons thus dictate the use of a relatively low IF, or excessive power consumption will result. The receiver will likely be a double superhet, with 2nd IF from a few MHz to a few ten MHz, feeding a 8-12 bit AD...
converter through a sharp channel filter just in front of it. An AGC will maintain AD converter’s input near to full scale to make best use of the available dynamic range. Since the signal is bandpass, it can be subsampled without creating aliasing (see below). A possible sampling frequency can be $f_s = \frac{4f_{IF}}{3}$, the output will then appear at $f_s/4$. Because of the complexity involved, today IF AD conversion is used mainly in fixed receivers such as cellular base stations, while it is quite rare in portable applications. An exception may be GPS receivers (including the one presented here), that due to the large amount of signal processing involved, will use IF AD conversion.

**Subsampling mixer:** From a time-domain point of view, sampling an analog continuous-time signal corresponds ideally to multiplying it with a periodic sequence of Dirac impulses. As such, a sampler is therefore a mixer, and can be used for frequency conversions. A subsampling mixer is a sampler that is operated at a frequency below twice the Nyquist frequency. This process is shown in Fig. 2.4, in both time and frequency domains. Signal aliasing is prevented as usual by filtering the input signal, which in this case must have a bandpass characteristic. The most common application of subsampling is in sampling oscilloscopes, while it is found more and more in communications digital radios. As an example, the 3rd conversion of many GPS receivers (see [21]), as well as the one presented here, is obtained by subsampling. Subsampling implies that the bandwidth of the signal path is (much) larger than the Nyquist frequency, and cannot even be reduced with filters, therefore a very obvious problem arises: the broadband noise is aliased, thus degrading SNR. If this may be no problem for an oscilloscope, which generally requires only 50-60dB of dynamic range, it may become a serious challenge in a communications receiver. If subsampling is performed at IF, after channel filtering and large amplification (thus in-channel narrow-band noise becomes dominant), and a moderate subsampling factor is used, SNR degradation is minimized. On the other hand, using a large subsampling factor - especially if direct sampling at RF is tried before any large narrow-band amplification has occurred - and considering also signal attenuation due to $\sin(x)/x$ filtering into the non ideal sampler, will result in such poor performance to be unacceptable even for the least critical application.

As a conclusion, the main characteristics of the superheterodyne architecture can be summarized as follows.
2.1. Superheterodyne

Superheterodyne receiver, plus:

- Excellent performance and flexibility, adequate for most applications.
- Well understood, well proven architecture. Considerable design experience available in most RF companies.
- Low power consumption since few parts operate at highest frequency - easily outweighs most disadvantages in portable equipment.

Superheterodyne receiver, minus:

- Spurious image response: an image filter is needed to suppress it.
- Complex system, many filters and resonators tuned at different frequencies are required.
- Relatively expensive, requires good quality filters and alignment.
- Low integration level, high quality filters and resonators are difficult to integrate.
2.2 Direct Conversion

Direct conversion, although it may seem like derived from superhet (by reducing the IF to zero) and actually share the same principle, predates superhet somewhat - at least in its simplest form - and has in reality originated from the beat frequency oscillator (BFO). The original purpose of the BFO was to make CW signals of Poulsen arc transmitters audible. It is an RF oscillator (formerly: a mechanical switch (ticker), a tone wheel, or another, very small arc transmitter), whose frequency is adjusted a few hundreds Hz apart from the received frequency such that after detection (actually mixing), the difference frequency - or beat frequency - appears at the output of the receiver as audio signal. Since the detector was now used as a mixer, with the large BFO signal applied to it, there was no longer need for the received signal to be large enough to turn on the detector of early crystal receivers, therefore sensitivity was much improved, and the somewhat sensitive DC biasing of the detector removed.

If the BFO is adjusted to the same frequency as the RF input, the difference frequency is zero, and if the RF input is AM modulated, only the modulation remains, and can be recovered with a simple low pass filter. The direct conversion receiver uses exactly this principle. Today direct conversion in this simple form is used in detectors for suppressed carrier DSB or SSB AM modulations - where an envelope detector fails to detect the signal properly - while if angle modulations have to be detected, this principle must be expanded to the modern direct conversion receiver, with in-phase and quadrature channels.

In a direct conversion receiver, as can be seen in Fig. 2.5, the RF signal is directly converted to base band where it is low-pass filtered for channel selection, amplified and detected (by a DSP in this example). By doing so, no image frequency is created. Since both sidebands of the modulated RF input are converted to the same base-band, a means must be provided to distinguish between them. Quadrature local oscillator signals with the same frequency as the RF signal, as well as the in-phase (I) and quadrature (Q) channels, are therefore required to recover the phase of the input signal, or distinguish between the two sidebands.

Since this receiver has no image response, no input filter is required, at least ideally, to guarantee proper operation. Nevertheless some kind
2.2. Direct Conversion

of RF filtering is still required, and should be retained to prevent strong out-of-band signals (e.g. FM broadcast or TV signals) from overloading the input of the receiver. Depending on the power of possible interferers and the desired receiver performance, a 2nd order, low Q LC RF filter may already be sufficient and give satisfactory results.

Given that downconverted signals now lie at base-band, channel selection is performed with lowpass filters. For a given selectivity, the requirements for these lowpass filters are relatively moderate, if compared to the channel filter of TRF or superhet receivers. Usually these filters are RC or $g_m$C active filters that can be easily integrated in an IC, thus saving the bulky and expensive external channel filter. Most amplification also occur at base-band, therefore it may be convenient to combine amplification and filtering in the same stages.

Depending on modulation type, several different detector types can be used. For simple binary FSK modulation an RS flip-flop will do. Since limiter-discriminator detection is not practical at base-band, for analog FM or multilevel digital modulations a Barber receiver structure can be used (see later). Alternatively the base-band signals can be digitized and detected with a DSP, which may implement an optimum demodulator for the given modulation. An AGC will be required to maintain the downconverted signals close to ADC full scale while avoiding overdrive, thus making the best use of available dynamic range.

Tuning is performed, like in superhet receivers, by changing the fre-
Figure 2.6: Image rejection vs. amplitude and phase mismatch.

quency of the local oscillator. The requirements for the local oscillator, regarding spectral purity and tuning range, are much the same as for superhet, thus the same comments apply. A major difference is that here quadrature LO signals are needed, whose precision depends on the required receiver performance, as explained later.

The performance of a well designed direct conversion receiver is in general good and adequate to most communications applications. It is basically comparable to superheterodyne, but thanks to the integrated channel filters and absence of image, direct conversion will enable a higher level of integration. The main parts which are likely left off-chip are the low Q RF filter and the local oscillator tank. Few adjustments, possibly none, may be required, thus production costs may be potentially low.

Similarly to superhet, direct conversion can be used in all common applications, while being more interesting where simplicity, high integration level or low receiver cost are required. Because of its simplicity, one of the first application has been in pagers (e.g. [24]), but receivers for more sophisticated standards such as GSM [25] or analog and digital TV [26] have also been developed. Many drawbacks, however, plague direct conversion, which can make its performance less than optimum, or render its design difficult.
2.2. Direct Conversion

One of the big disadvantages of direct conversion is its sensitivity to amplitude and phase mismatch in the quadrature channels. Very close matching is required in both the LO and the base-band I, Q channels to avoid performance degradation. To quantify these effects, the image rejection ratio (IRR) can be calculated with the formula:

\[
IRR = \sqrt{\sin(\frac{\Delta \phi}{2})^2 + \Delta A^2 \cos(\frac{\Delta \phi}{2})^2}
\]

Here the image is the leakage of one sideband into the other, \(\Delta A\) is the amplitude mismatch and \(\Delta \phi\) is the phase mismatch. Fig. 2.6 shows the result in graphical form. Assuming e.g. that for good performance an IRR of 30dB or better (the shaded area in Fig. 2.6) is required, this calculation shows that \(\Delta A < 0.3\text{dB}\) and \(\Delta \phi < 3^\circ\) are required for the combined errors of the LO and the quadrature base-band channels.

Taking advantage of the good matching characteristics of integrated components, well matched I, Q channels can be realized, provided that only few gain stages and low Q filter stages are used, and a careful symmetrical layout is drawn. In this case, the lowpass filters - whose phase is particularly sensitive to mismatch, not only can be integrated, but must be integrated for best performance. A discrete design would mandate a careful and expensive alignment.

LO phase shifters may also take advantage of IC implementation for the same reason. To obtain quadrature LO signals, various phase shifter types can be used, but unfortunately none of them is ideal, or equally suitable for IC implementation. Some examples, shown in Fig. 2.7, are:

![Figure 2.7: Some common phase shifters.](image-url)
• λ/4 line: This phase shifter consists of two transmission lines, one of which is λ/4 longer than the other. Phase shift thus results from the time differences the LO signal requires to reach the end of the transmission lines. To avoid reflections, the lines must be terminated, at least at the output, by their characteristic impedance. Provided that essential substrate characteristics such as thickness and εr of the dielectric are well controlled and known precisely, the delay depends basically on mechanical dimensions, thus this phase shifter can be quite accurate. Because of its principle of operation on the other hand, this phase shifter is necessarily narrow band and large, at least at low GHz frequencies, making it difficult to be economically integrated. Since this phase shifter is lossless, power consumption can be relatively low. It is essentially the power required to drive the terminations (Fig. 2.7a).

• Digital divider: A digital divider-by-2 can be used as phase shifter, since the output of the two latches in a MS flip-flop, or of two independent flip-flops clocked on the opposite edges of the clock signal (Fig. 2.7b), is in quadrature. This digital phase shifter offers the highest precision and widest bandwidth - ideally down to DC - and is very compact and best suited for an IC implementation. Its main disadvantage is the relatively high power consumption, not only within the phase shifter itself, but especially because a LO with twice the RF frequency is required. Nevertheless, if one considers the losses in the buffers and amplifiers that other phase shifters will require, the power consumption disadvantage is much less severe than expected, and this solution is surely competitive. The main source of error in this phase shifter is 2nd harmonic distortion of the LO signal, while jitter or delay time mismatch are in most cases negligible. By the use of differential structures throughout (e.g. ECL/ESCL dividers or emitter/source coupled LO), a symmetrical clock with low 2nd harmonic is guaranteed, while achieving other advantages such as reduced power consumption, substrate noise rejection, or making possible direct coupling to the LO or to the mixer. If a symmetrical clock signal cannot be provided, a further divider-by-2 could fix the problem, but at the expense of much higher power consumption, since a clock with 4 times the RF frequency would be needed.

• LC-filters: One of the most obvious phase shifter structures that may come to mind is surely an LC filter. One possible imple-
2.2. Direct Conversion

mentation is the lowpass filter shown in Fig. 2.7c. In this filter, a 90° phase shift occurs at resonance. Since sensitivity to phase is maximum at resonance, also this phase shifter is necessarily narrow band. A low Q filter is therefore best suited, to increase bandwidth somewhat and reduce sensitivity to component tolerances. Alternatively more complicated structures may be used if space and cost are of no objection. Since high quality, precision inductors and capacitors cannot be easily integrated, this phase shifter has usually to be implemented externally, and will possibly require alignment. Similarly to the λ/4 line, this phase shifter is lossless, and has the potential for low power consumption (termination resistor).

- **RC-filters**: Much better integratability is achieved if RC filters are used instead of LC. Fig. 2.7d shows a possible implementation of a RC phase shifter. Assuming well matched components, this phase shifter has a constant 90° phase shift at every frequency, that does not vary even when loaded, provided that the loads are well matched, too. It is therefore accurate and best suited for an IC implementation. Usable bandwidth is only limited by the amplitude dependence on frequency. By leveling the outputs with limiting amplifiers, an octave is easily reached, while no alignment is usually needed. On the negative side, the resistors necessarily introduce losses, thus high power consumption has to be expected. If more complex circuits are used e.g. to improve bandwidth, things are even worse. In some cases it may be interesting to use the input resistance of the leveling amplifiers as part of the filters. In such cases RC-RL filters will be used (L and C in series with the amplifier inputs) and no extra loss will be added. Lower power is therefore obtained, at the expense of some external components and the possible need for alignment.

Since a mixer is a commutative operator (it is just a multiplication), the RF signal can be shifted instead of the LO signal, if desired\(^2\). In such cases, to prevent noise figure degradation, only lossless phase shifters can be used, i.e. transmission lines, LC or RC-RL filters. Although the latter are the least suitable for integration, this option may actually simplify the LO path, thus reducing power consumption, while complexity can be maintained within reasonable limits if the phase shifter

\(^2\)See [24] for an example, which uses RC-RL filters in the RF path.
is combined with the filters and matching networks.

Apart from sensitivity to mismatch in the I and Q channels, other unique problems remain to be solved when designing a direct conversion receiver, the most important of which are its sensitivity to flicker noise, local oscillator reradiation, DC offset and AM detection. These are briefly discussed below.

- **Flicker noise:** In a direct conversion receiver most of the gain occurs at base band, while the gain of the RF front-end may be about 20dB or so. This means that at the input of the baseband strip the signal voltage may be only a few microvolts which is easily masked by noise. At base band the dominant noise is flicker noise, which can easily be orders of magnitude larger than thermal noise, especially for CMOS processes. Because of the small signals involved, this problem may be difficult to solve with the usual circuit tricks, and a technology with low flicker noise may be needed.

- **Local oscillator reradiation:** This has two serious effects on the receiver. The first is the DC offset generated by mixing of the LO signal with itself, and the second and more severe effect is that the reradiated signal becomes an in-band interferer that may disturb nearby receivers tuned on the same frequency. For this reason, in most communications standards, the maximum allowable reradiation is specified by type approval requirements. For
example, ERMES type approval [17] imposes a maximum limit of 2pW of radiated power within the ERMES channels, which is one of the hardest values known. Good circuit techniques such as differential structures, good PCB placement to prevent coupling, as well as good receiver shielding will certainly help solve reradiation problems. Measurements on an actual circuit [27] demonstrated that even the 2pW limit of ERMES can be achieved, and does not prevent the design of direct conversion ERMES pagers.

• **DC offset:** Ideally, the bandwidth of the base-band amplifier must extend all the way to DC, if signal losses around DC must be limited. Since the gain of the base-band strip can be large, even a small DC offset can be sufficient to saturate it. DC offset originates from 3 sources: transistor mismatch in the input stage, LO reradiation, and interference from a powerful transmitter (see Fig. 2.8). Offset from transistor mismatch is usually of the order of 1-5mV, which is already a thousand times the input signal, but it is usually constant (or varies very slowly) and can be removed with a calibration loop. Apart from applications using modulations with low DC contents (e.g. pagers), AC coupling is usually impractical because of the large capacitor values required. A more serious problem is the offset generated by self-mixing of the local oscillator. In fact the reradiated local oscillator signal is reflected back by the antenna and mixed down to DC, where it appears as a large offset. If the reflection coefficient of the antenna changes suddenly, for example when the user touches the antenna, this offset may fluctuate sufficiently to overload the base-band strip, thus reducing the sensitivity of the receiver or blocking it completely despite offset calibration. A large interferer will generate offset if its signal is directly rectified (see below) or if it is allowed to leak to the LO port, where it self mixes to DC. Its effect on the base-band strip is similar to that of LO reradiation.

• **AM detection:** Many transmitters, e.g. TV transmitters, amateur radio, etc. are amplitude modulated. Some of these interfering signals may pass through the RF filter and reach the input of the mixer. Here there are enough junctions and other (2nd order) nonlinearities to demodulate these signals and make them indiscernible from the desired signal, eventually destroying the latter completely. A mixer (and also an LNA if the latter is DC coupled to the mixer) with low 2nd order distortion, together with suffi-
cient RF filtering, are therefore needed to prevent both DC offset generation and AM detection.

If direct conversion is a simple architecture from a system designer point of view, it is actually a quite complex receiver, requiring many more parts than superhet. This will lead us to another important drawback of this architecture: large power consumption. As we have seen, an extra mixer and an accurate 90° phase shifter for the I and Q local oscillator signals, both operating at the highest RF frequency, are required. Although the lack of lossy filters in the RF path means that lower gains and higher noise figures can be tolerated for a given performance - thus reducing current consumption somewhat - current consumption in these parts, being usually dictated more by linearity or operating frequency requirements rather than gain or noise, is still substantial. The baseband strip is also partly responsible for increased power consumption. If a superhet requires a relatively simple single channel IF amplifier (limiting or with AGC, depending on modulation), direct conversion necessitates the two I and Q channels comprising linear non limiting AGC amplifiers and active filters, which may be quite power hungry, especially considering the close gain and phase matching required.

Experience has shown that, for a given technology and design style, a power consumption of roughly twice than superhet can be expected. A practical example could be the pager front-end in [24], which consumes 5.5mW, while readily available double superhet discrete front-ends [29] hardly consume 2-3mW despite the older technology. Since consumers see short battery life as one of the biggest nuisances of all nomadic electronics, the larger power consumption means that a larger battery - already one of the most expensive items [23] - must be used. The cost advantage of direct conversion may therefore not be reflected to the complete system, especially considering that progress in battery technology is much slower than in electronics, and major improvements are not expected [28].

Although there is a strong drive in the RF community to overcome the difficulties associated with direct conversion receivers, the trade-offs involved still make it difficult to a priori discard either superheterodyne or direct conversion architectures. One must consider the implications for the complete system - not just the cost of the receiver part - to make the proper choice, especially in systems where the power consumption
of the receiver is substantial compared to the total or if available space (for a larger battery) is limited.

A few common variants of the direct conversion receiver are:

**Barber receiver:** In many cases, especially when analog modulation types are involved, the use of DSP detection is impractical or not justified. In such cases one may want to use simple analog detectors such as the FM discriminator. Since at base-band an FM discriminator cannot be readily used, a Barber receiver will be implemented. In a Barber receiver, as shown in Fig. 2.9, the I and Q base-band signals are filtered and amplified in the usual way, before being remixed to a scalar IF, where amplitude limiting and detection by a conventional FM discriminator, a PLL or any other suitable detector is performed. Although this may first look like superheterodyne in disguise, one must realize that most advantages of direct conversion (with DSP detector) are retained, especially the easy channel filter requirements and the high integration level achievable. Some applications of the Barber receiver are in 1st generation direct conversion pagers or in TV tuners [26], or if coherent detection is desired (e.g. GPS).

**Homodyne:** A homodyne receiver is basically a direct conversion receiver, in which the local oscillator is precisely synchronized and in-phase with the incoming RF carrier. The input signal is therefore converted exactly to base-band, and no difference frequency (beat tone) that may corrupt the received signal is generated. In order to allow this
precise synchronization of the local oscillator, some means must be pro-
vided in the transmitted signal, such as a partially suppressed carrier,
a pilot tone or some other kind of synchronization signal. The homodyne receiver can be considered the precursor of the modern coherent
demodulator, commonly used for PSK, QAM or similar modulations in
digital radio, TV, modem, as well as in GPS receiver applications.

To conclude this section, we can summarize the main characteristics of the direct conversion architecture.

**Direct conversion receiver, plus:**

- High performance, comparable to superheterodyne (or slightly less). Suitable for communications applications.
- Simple channel filter requirements: low order, low Q, integratable lowpass filter.
- No image response, no image filters required. 2nd order, low Q, LC RF filter is in most cases sufficient.
- High integration level, small size, relatively simple PCB design.
- Low receiver cost: no ceramic filters, few adjustments (possibly none).

**Direct conversion receiver, minus:**

- In-channel local oscillator reradiation, may disturb other receivers.
- Close amplitude and phase matching required for the I and Q channels.
- Many problems (1/f noise, DC offset, AM detection, ...) remain to be solved.
- High power consumption due to complex RF part. Receiver size and cost advantage may be nulled by the requirement for a larger battery.
3.1 Introduction

A pager is a device meant to call persons on the move, or to send short messages to them. It is a receive-only system, and in contrast to wireless telephones an immediate answer is not expected nor possible. Typical of most modern pagers are the small, lightweight form factor and the long battery life, reaching a few thousand hours with a single LR6 alkaline cell. Subscription costs are also usually (much) lower than those for cellular telephones.

Early analog pagers had sizes spanning from that of a cigarette packet to that of a car radio, and the call signal was a simple beep or a blinking light whose meaning had to be defined beforehand by the caller. Alternatively the called person had to seek information from the caller by using a regular telephone. Since then, pagers have evolved to small and more powerful devices. A modern pager is usually smaller than a cigarette packet, and the smallest ones have the form of a key-holder or are integrated in a wrist-watch. By using digital modulation methods, higher performance and robustness is achieved, while much more data can be transmitted efficiently. This means that, apart from
the traditional tone-only call signal, a small message can be sent to the called person, which usually does no longer need to seek further information. The latter is not only useful as the caller can send small messages, but also to send information that needs to be updated frequently, such as weather reports, trade/financial or sports information, and other Teletext-like services.

The market for pagers has been traditionally dominated by business users, but consumer use may take over also this market, as it happened years ago with cellular telephones. Since a pager can hardly compete with the latter for what belongs to performance, other ways must be found to gain user acceptance, such as low purchase and operating costs or better convenience. An attractive way to obtain better convenience would be to integrate the pager in a wrist-watch. Provided the watch does not have a disproportionate size or weight, people will simply carry it at their wrist as they always used to. The pager will then become available at any time, people will have a box less to carry, while a pager worn at the wrist is much more difficult to forget or to lose.

Nevertheless, the integration of a pager in a wrist-watch is not a straightforward task and will pose many technical challenges. First of all, to keep a low price and to fit into the small package of a watch, a highly integrated solution with low external part count is highly desirable. Being a new system at the time of this work (1994), integrated receivers such as the POCSAG pager receiver in [24] are not available for the ERMES standard. Although application specific CMOS chips can be found for the digital and base-band analog parts, the front-end of a typical ERMES pager still consists of discrete transistors or a few chips and many passive components. Some of these parts are relatively expensive and require adjustments [18].

Another particularly important challenge for a pager to be integrated in a wrist-watch, where battery size must also be minimized, is achieving a very low power consumption. Although ERMES realizes battery economy with ON/OFF operation as explained later, this is not sufficient to ensure correct operation. A typical battery for this application will be a button type LiMnO₂ cell because of its high energy density. Nevertheless because of space constraints only a very small battery can be used in a watch, whose capacity is necessarily limited and even decreases at high drains. Such batteries also have usually a quite high internal resistance, and are only rated for a peak current not higher than a few mA.
3.1. Introduction

<table>
<thead>
<tr>
<th>Frequency band</th>
<th>169.425 to 169.8MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel spacing</td>
<td>25kHz</td>
</tr>
<tr>
<td>Number of channels</td>
<td>16</td>
</tr>
<tr>
<td>Symbol rate</td>
<td>6.25kbits/s</td>
</tr>
<tr>
<td>Modulation format</td>
<td>4-PAM/FM with $\Delta f=3.125$kHz</td>
</tr>
<tr>
<td>Modulation index</td>
<td>1.0</td>
</tr>
<tr>
<td>Data shaping filter</td>
<td>10th order Bessel, BW=3.9kHz</td>
</tr>
</tbody>
</table>

Table 3.1: ERMES basic air-interface parameters

If current consumption is not sufficiently low, battery overloading may occur which may result in an unacceptable power supply drop during the ON time. This means that to achieve both sufficient life and proper operation, power consumption must be reduced by about one order of magnitude compared to present solutions.

In a pager, the performance of the antenna is one of the most important factors in defining sensitivity. The small size means that an electrically small antenna with a relatively poor efficiency has to be used. The radiation resistance of such a small antenna is necessarily very low, of the order of 0.1–1Ω for a typical pager. Even lower values must be expected for an antenna that fits into a wrist-watch. To achieve acceptable performance, a front-end with higher gain and lower noise than the one required by larger systems may be necessary, which conflicts with the circuit design requirements for low power consumption.

The purpose of this work was to demonstrate the feasibility of an ERMES pager with sufficient performance to meet type approval requirements, that consumes little power and is sufficiently small to fit into a wrist-watch. Some aspects regarding the wrist-watch implementation are the following:

- **Power supply**: It is delivered by a 3V LiMnO$_2$ cell. The voltage range is 2–3.5V, while the total current consumption should be below 1mA.

- **Watch**: In order to be practical and aesthetically acceptable, the watch should have normal dimensions, i.e. approximately 25–30mm diameter and 6–10mm thickness.
This chip consists of a collection of the most critical circuits regarding both power consumption and performance, i.e. the RF front-end, the local oscillator and the prescaler. An IF amplifier and mixer (half a demodulator) complete the chip. For this work a complete, operating, pager receiver was not expected.

### 3.2 ERMES System Overview

The European Radio Messaging System (ERMES) is a new digital pager standard that was defined in 1988 by the ETSI. ERMES is a new generation of full featured pagers for Europe. It defines different pager types for European wide access from tone only to numeric and alphanumerics.

One of the first digital pager standards has been defined in 1984 in England by the Post Office Code Standardization Advisory Group, and is referred to as POCSAG. It is a one channel system which uses 2-PAM/FM modulation with data rates of 512, 1200 or 2400 bits/s. Its capacity, although adequate at the time of its development, is therefore rather limited for today’s needs. Different frequency bands are allocated in different countries, usually around 150-170, 290 or 470MHz, meaning that different receivers must be constructed to suit the local needs. ERMES solves this problems by specifying a single european-wide frequency band from 169.425MHz to 169.8MHz. Sixteen channels of 25kHz bandwidth each are allocated to ERMES, which uses both frequency and time division multiplexing for better efficiency.

Different pagers are addressed using 35 bits long radio identity codes; 13 bits are used for operator identity and 22 as local addresses, such that a single operator can address up to $2^{22}$ different pagers. Unlike older standards such as POCSAG, ERMES is synchronized to the UTC (Universal Time Coordinated), meaning that transmissions for a particular pager always start at fixed times. This time coordination allows e.g. pagers to switch easily between different networks, as well as to achieve a substantial battery saving by not turning on the receiver when not necessary. The On/Off ratio ranges from 1:14 (worst case) or 1:100 (typical) to 1:72000 (best case). For comparison, the On/Off ratio of POCSAG is only 1:8 typical.

The modulation is 4-PAM/FM, which is an extension of FSK, at
Frequency shift keying is a widely used modulation format, especially in low bit-rate applications where simplicity and low cost are more important than spectral efficiency [14]. It is obtained, in its most obvious form, as shown in Fig. 3.1: the data is represented by selecting one out of $N$ frequencies per time interval. The signalling frequencies $f_x$ are usually equidistant, i.e. $f_x = f_c \pm (0.5 + i)hf_s$, where $f_c$ is the carrier frequency, $h$ is the modulation index, $f_s$ is the symbol rate, and $i = \frac{0..(2^n-1)-1}{n}$ with $n$ the number of bits per symbol. Binary FSK, i.e. only 1 bit per symbol or 2 signalling frequencies, is used in most cases, while quaternary (or larger) FSK is less common. ERMES is one such example.

Obtaining FSK as in Fig. 3.1, i.e. using many independent oscillators with no phase constraints, will result in discontinuities at the interval boundaries. This is called incoherent FSK. Because of these discontinuities, the spectrum occupied is larger than necessary, which is undesirable since this limits the number of channels that can be transmitted in a given bandwidth. By constraining the phase to be continuous at any time, much better spectrum usage may be achieved. This is called continuous phase FSK, CPFSK. the CPFSK signal is continuous everywhere, including the symbol boundaries, and is usually obtained with an analog FM modulator, driven by more or less smoothed pulses.

The modulation index $h$ determines how much frequency change occurs at the appearance of a new symbol, much like the modulation depth.
in analog FM. A large $h$ implies a large frequency shift, that is a large occupied spectrum. For binary FSK, the minimum $h$ that guarantees orthogonality, i.e. the crosscorrelation coefficient $\rho_r = 0$, is $h = \frac{1}{2}$. This is called minimum shift keying, MSK, and is the most efficient binary FSK regarding spectrum usage. Despite the worse spectrum usage, a larger $h$ is not necessarily bad, since it permits reliable operation of simple limiter-discriminator-comparator detectors, thus allowing very cheap receivers to be constructed (early ERMES pagers were indeed constructed that way [18]). ERMES specifies $h=1$, while even larger $h$ are specified for POCSAG (depending on symbol rate).

In the case of rectangular CPFSK (immediate change between frequencies, no smoothing filter) the spectrum is peaked at the signalling frequencies, and for the special case $h = 1$ the spectrum contains only the signalling frequencies. Both are undesirable properties. To prevent this, a smoothing filter is almost always used. The purpose of the smoothing filter is to obtain a smooth change between the signalling frequencies, thus reducing or eliminating peaking. Such a modulation is called pulse amplitude modulated FM, PAM/FM, and is the modulation scheme used in both POCSAG and ERMES pagers. The structure of the modulator is shown in Fig. 3.2. POCSAG specifies 2-PAM/FM with $\Delta f = \pm 4.5$kHz, and ERMES specifies 4-PAM/FM with $\Delta f = \pm 1562.5$ and $\pm 4687.5$kHz. The smoothing filter is a 10th order Bessel low-pass

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Offset</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>$-4687.5$Hz</td>
</tr>
<tr>
<td>01</td>
<td>$-1562.5$Hz</td>
</tr>
<tr>
<td>11</td>
<td>$+1562.5$Hz</td>
</tr>
<tr>
<td>10</td>
<td>$+4687.5$Hz</td>
</tr>
</tbody>
</table>

Table 3.2: ERMES symbol mapping
filter, with a $-3\text{dB}$ frequency of 3.9kHz (ERMES). The resulting spectrum is no longer peaked, and must be contained within the masks shown in Fig. 3.3 for a base station to pass type approval. To minimize bit errors, ERMES uses Gray code mapping between frequencies and symbols. Table 3.2 shows this mapping.

To guarantee a minimum system performance, as required to get user acceptance, ERMES receivers - just like most communication equipment - must be submitted to conformance testing and achieve a minimum performance as specified by type approval requirements [17]. Type approval is very comprehensive and covers all technical aspects of the pager receiver, from the essential characteristics required from any pager (e.g. sensitivity, type of response to a call, or battery low indicator) to many optional features required only from some pager types (e.g. maximum message length or lost message indicator). In our case, the most relevant part of type approval are the RF requirements. A summary of the most important ones is shown in table 3.3.

### 3.2.1 Receiver implications

Since PAM/FM is a constant amplitude modulation, very simple, low power receivers with a FM discriminator as detector can be constructed,
Table 3.3: Minimum RF characteristics for type approval (summary).

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Limit value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average sensitivity</td>
<td>25dBμV/m</td>
</tr>
<tr>
<td>Co-channel rejection</td>
<td>-10dB</td>
</tr>
<tr>
<td>Adjacent channel selectivity</td>
<td>60dB</td>
</tr>
<tr>
<td>Spurious response immunity</td>
<td>76dBμV/m</td>
</tr>
<tr>
<td>Intermodulation immunity</td>
<td>70dBμV/m</td>
</tr>
<tr>
<td>Blocking or desensitization</td>
<td>84dBμV/m</td>
</tr>
<tr>
<td>Spurious emissions:</td>
<td></td>
</tr>
<tr>
<td>30MHz–1GHz</td>
<td>2nW</td>
</tr>
<tr>
<td>1GHz–4GHz</td>
<td>20nW</td>
</tr>
<tr>
<td>ERMES channels</td>
<td>2pW</td>
</tr>
</tbody>
</table>

provided a sub-optimum detector is acceptable. Such a receiver is likely a superheterodyne receiver very similar to an usual FM broadcast receiver. A 3-level comparator can be used to extract the digital data from the detected signal. Because of normal tolerances of the low-cost crystal used in the frequency synthesizer, some frequency offset must be anticipated. This results in a slowly varying DC offset at the output of the demodulator, which can be suppressed simply with AC coupling. A comparator with a higher number of levels (a 3–4bit AD converter) - together with AFC - is another possible solution which may lead to improved performance under such conditions.

The superheterodyne receiver will be one of the lowest power solutions possible, and, were it not for the drawbacks explained later, it would be very attractive for small form-factor pagers such as wrist-watch style, where power is at a premium. In fact the number of circuits working at the highest frequency, which are also the most power hungry, is minimized. They consist basically of an LNA, a mixer, a local oscillator and a prescaler.

Since considerable experience in designing superhet receivers is usually available in most RF companies, this solution represent an easy, low risk approach. Many problems typical of other solutions, e.g. local oscillator reradiation, DC offset, etc. that plague direct conversion receivers, simply do not exist here, while most components required
are easily available in quantities at relatively low cost. This includes ceramic filters and resonators, which are manufactured by several companies (provided the usual standards such as the 21.4MHz or 455kHz IF frequencies are retained).

The requirement for these high quality ceramic filters, however, is one of the major drawbacks of superhet. Ceramic filters are bulky and almost impossible to integrate, while their cost is already at a minimum, with no real perspective for further reduction. The need for several adjustments is also a major drawback, since it increases manufacturing costs, although the use of laser trimmed components alleviates the problem slightly. All this is no good and too complex for wrist-watch pagers, and other solutions may be desirable even at a cost of a higher power consumption.

The PAM/FM modulation, having no information at DC, lends itself almost naturally to a direct conversion receiver architecture. Not requiring high Q external ceramic filters, direct conversion will enable the highest level of integration, while many DC related problems, e.g. offsets of the I/Q channels or caused by self mixing may be simply removed by AC coupling.

Unfortunately, given the same technology, a direct conversion receiver will consume much more power than its superhet counterpart. A 2nd mixer, a 90° phase shifter and two I/Q channels are needed, the latter being linear non-limiting amplifiers requiring AGC. The digital interface is also more complex, consisting of two 8 bits (or so) AD converters. A simple comparator is no longer sufficient. The phase shifter, i.e. one of the most power hungry circuits in a direct conversion receiver, can be implemented externally using LC filters (as in [24]), but this partially neutralizes the simplicity advantage of direct conversion. Since type approval requires a maximum spurious emission of only 2pW within the ERMES channels, local oscillator reradiation is critical and must be addressed early in the design of the receiver.

A possible trade-off may be represented by a low-IF, or offset, receiver. Although considered obsolete by many, it is a viable "transitional" architecture, which realizes an integration level similar to direct

\[1\] A few low-IF receivers have been developed recently by the industry, and all but some very low performance FM broadcast receivers (e.g. Philips TDA7088T) were obsoleted after just a few years of production.
conversion with less risk, but also with less performance. The 2pW emission limit will still be a problem with this architecture. Some POCSAG pager receivers [31] (today obsoleted), as well as the old (1961) AT&T Bellboy used a low IF architecture.

3.3 Antenna Configurations

As one of the main factors that define the sensitivity of a pager, the intended antenna must be known before any receiver design is attempted. Because of its necessarily small electrical size, assuming that the antenna is a well matched 50Ω source would be wrong and almost certainly lead to an unusable receiver. In the absence of any better specifications, the first thing to do is to find out which kind of signal can be made available to the receiver chip.

Two main requirements for an antenna for pagers are small size and good performance when placed near to the user's body. Over the frequency range of interest the body is conductive, and acts predominantly as a reflector. The low impedance of the body causes a dip of the electric field and an increase of the magnetic field (up to 6dB at 170MHz) in the close vicinity of the body, with the field lines tangential to it [35]. It is therefore evident that a magnetic antenna should be used for good performance close to the body. In fact it is the most commonly used antenna for pagers, and consists of an electrically small loop. The radiation pattern of such an antenna is ideally figure-8 in the free space, but when body mounted the current distribution around the loop becomes sufficiently nonuniform to fill in the two zeros in the radiation pattern somewhat. Although it does not really become omnidirectional, a much better radiation pattern is realized. For best performance the antenna should be oriented with its magnetic axis horizontal and tangential to the body when the pager is in its normal operating position, to intercept the largest amount of field lines.

In order to properly design the input stage of the LNA and the antenna matching network we need to know impedance and signal level at the antenna. The equivalent schematic diagram of a loop antenna, with its matching network, is shown in Fig. 3.4. $V_s$ is the voltage developed by the loop antenna, $R_r$ is the radiation resistance, $L_{ant}$ is the loop antenna inductance and $R_L$ is the loss resistance associated to it.
Assuming that $E$ is the incident electric field, the voltage developed by the antenna can be calculated as [24]:

$$V_s = h_{eff} \cdot E, \quad h_{eff} = 2\pi \frac{AN}{\lambda}$$

with $h_{eff}$ the effective height, $A$ the area, $N$ the number of turns of the antenna and $\lambda$ the wavelength. The radiation resistance is:

$$R_r = 320\pi^2 \frac{A^2 N^2}{\lambda^4}$$

and is usually below 1Ω for typical antenna sizes. Both $V_s$ and $R_r$ can effectively be increased by increasing the loop area or the number of turns. Unfortunately also $L_{ant}$ increases (with the square of $N$) thus it is seldomly possible to use more than a couple of turns, or an impractical matching network between antenna and LNA may result.

$R_L$ is the loss resistance of the antenna and its effect is both a reduction of sensitivity and an increase of the noise figure already at the antenna. It should therefore be minimized. This can be done by using a thick or (better) a flat, smooth wire to reduce skin effect, as it is done in most pagers.

When it comes to the watch implementation, several other problems must be solved. First of all the small available space means that an antenna with a smaller than usual loop area must be used. This means that $V_s$ and $R_r$ are also smaller than usual, which make transformation to usable values more complicated. Placement of the antenna is also more complicated. Part of the problem originates because a watch,
different from a pager attached to a belt or placed into the shirt pocket, has no "normal" position and is continuously moved around during user’s activities. Therefore there is no way to guarantee that the antenna stays in the vertical position most of the time, as it would be necessary for optimum performance.

An even tougher placement problem is how to install the antenna in the round case of a watch, where one entire side is occupied by the dial. The first ideas that come to mind are shown in Fig. 3.5. In Fig. 3.5a, the antenna loop is placed parallel to the watch case. Here the maximum possible loop area is realized, while the dial remains free. Unfortunately this places the antenna with its magnetic axis perpendicular to the body, while the field lines are tangential to it. Nevertheless because of its small size, the field lines at the wrist will also not be exactly in the tangential direction. Tilting the antenna slightly on the side as shown in the figure (the gray loop) may also be of some help. The pager circuit will also be contained in the middle of the antenna, which will therefore become particularly efficient at picking up electrical noise generated within the pager circuit.

A more typical solution would place the antenna in the perpendicular position, as shown in Fig. 3.5b, to make it intercept most field lines. To maximize its area the whole thickness of the watch would be used. Unfortunately in this case the loop is very small and will pass over the middle of the dial. Although the antenna may become an aesthetical feature of the watch, it is still a disturbing element, the pager circuit would be still placed in the middle of it, and watch construction will be
### 3.3. Antenna Configurations

<table>
<thead>
<tr>
<th>Antenna</th>
<th>Round, $\phi$25mm</th>
<th>Rect. 25 x 5mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_{\text{eff}}$</td>
<td>1.7mm</td>
<td>0.44mm</td>
</tr>
<tr>
<td>$V_s$</td>
<td>29nV</td>
<td>7.5nV</td>
</tr>
<tr>
<td>$R_r$</td>
<td>75m$\Omega$</td>
<td>27m$\Omega$</td>
</tr>
<tr>
<td>$L_{\text{ant}}$</td>
<td>60nH</td>
<td>27nH</td>
</tr>
<tr>
<td>$R_L$</td>
<td>85m$\Omega$</td>
<td>65m$\Omega$</td>
</tr>
<tr>
<td>$F_{\text{ant}}$</td>
<td>3.3dB</td>
<td>5.3dB</td>
</tr>
</tbody>
</table>

**Table 3.4:** *Electrical parameters of the small loop antennas of Fig. 3.5 (one turn).*

complicated somewhat. To solve these problems, the antenna can be bent to the side of the case, as shown in the same figure, although this will increase losses ($R_L$). The radiation pattern, however, may benefit somewhat.

To increase its effective area the antenna can be loaded with a ferrite core. A sensitive antenna with tangential magnetic axis that does not interfere with the aspect of the watch could therefore be realized. Such a solution, however, can be discarded quite safely at the beginning, not only because of higher losses, higher inductance or similar technical problems a ferrite antenna may have, but especially because of the large space such an antenna would occupy within the watch. The antennas of Fig. 3.5, on the other hand, would typically be molded into the plastic case, and as such will occupy virtually no space.

A fourth solution could be to integrate the antenna into the watch strap. A bigger loop could therefore be realized while no space will be wasted within the watch. Unfortunately the loop will still be in the wrong position, while newer problems will be added because the strap, and thus the antenna, is deformable. The parameters of a deformable antenna will in fact be quite unstable and somewhat unpredictable, leading to large sensitivity variations. Alignment may also become difficult, while the Q of the antenna may need to be limited excessively thus possibly losing the advantage of a big loop on sensitivity. Such an antenna will also deteriorate with time, while strap replacement can no longer be made by a watchmaker as usual, because of the need for a new alignment of the antenna circuit that in most cases the watchmaker cannot do. This solution has therefore also been discarded.
Because of the aforementioned problems, the definitive solution will quite surely be similar to one of the two antennas shown in Fig. 3.5, possibly with 2 turns or so. It is clear that some antenna research is still needed, but the basic problem will certainly be the small space available, which will limit the available choices. Using a higher frequency pager standard is no real solution either. While simplifying the design of a small antenna, power consumption would be increased, leading to the need for a larger battery. This would therefore only shift the size problem. Such an option was anyhow not available in our case, since the only frequency band allocated to ERMES is at 170MHz.

The electrical parameters of these two antennas (approximated) are shown in table 3.4. The first is a round antenna of the type shown in Fig. 3.5a with a diameter of 25mm, and the second is a rectangular antenna as shown in Fig. 3.5b with dimensions 25mm x 5mm, both constructed with a φ1mm copper wire. The source voltage $V_s$ is calculated for an electric field of 25dB/μV/m, i.e. the minimum sensitivity required for type approval, impinging on the antenna in the direction of maximum sensitivity in the free space, while $F_{ant}$ is the noise figure of the antenna due to $R_L$. These sizes were chosen to fit in the smallest package envisaged. For comparison the antenna of a typical (small) pager - rectangular, 50mm x 10mm - will have essentially the same electrical characteristics as the φ25mm round antenna above, but with the advantage of being perpendicular to the body, thus achieving better efficiency. A watch able to contain it will nevertheless look huge and disproportionate.

As expected from such small antennas, $V_s$ and $R_r$ are both very small, especially in the 2nd example. The latter will definitely benefit from a second turn which will raise $V_s$ to 15nV and $R_r$ to 108mΩ respectively while improving $F_{ant}$ to 3.4dB.

It is quite clear that trying to impedance match the input of the front-end to these antennas is nearly impossible, and will also not lead to optimum performance. Since in this case there is no need for impedance matching, for maximum receiver sensitivity the antenna can be noise matched to the receiver. Even more important here is to boost the very small $V_s$ as much as it is practical. For this purpose the matching network shown in Fig. 3.4 can be used. It is an usual LC low pass matching network, where the inductance is provided by the antenna itself ($L_{ant}$). $C_s$ will be used to cancel out part of $L_{ant}$ if it is excessive.
3.3. Antenna Configurations

At this point the antenna-LNA interface is completely defined, and can be summarized as follows:

- A matching network between antenna and LNA is needed to boost \( V_s \) to a value that can be processed easily.

- The low \( V_s \) mandates a high Q matching network. Some variance must therefore be expected because of normal manufacturing tolerances, thus antenna alignment (the antenna trimmer, \( C_{tune} \)) is required.

- For easy alignment and long term stability, the Q must be limited to about 50 or so. This means that \( R_r \) is transformed to 100–300\( \Omega \) and \( V_s \) to 0.7–1.5\( \mu \)V, both good values that lead to a manufacturable receiver (although noise will be a problem).

- To get maximum voltage gain under these conditions, voltage matching can be used. This prevents the signal voltage at the antenna output from being attenuated right at the LNA input. The input impedance of the LNA must therefore be high. A couple k\( \Omega \) gives the most practical \( L_{ant} \) and \( C_{tune} \) (40–50nH, 15–20pF).

- Because of the high Q, the matching network already provides sufficient RF filtering to protect the LNA from strong interferers. No extra RF filter is therefore required.

Achieving a good noise figure in a system with such a small antenna is more difficult and requires some attention. In fact:

- The antenna already has a relatively high noise figure. The only way to reduce it is to increase \( R_r \), which requires a big loop or many turns, or reduce \( R_L \), which requires a thicker wire. Neither of them is particularly practical in a watch.

- Losses in the matching network must be minimized. Since inductors are lossy, no other inductor than \( L_{ant} \) must be used. The matching network shown in Fig. 3.4 is already optimum in this respect, since it uses only two (good) capacitors.

- A front-end with low noise figure is important. At best noise matching should occur for the transformed \( R_r + R_L \) (300–800\( \Omega \)).
As can be seen from the above considerations, the small antenna size is the major obstacle to good performance for this type of pagers. The noise generated by the antenna, referred to the input of the front-end, is in fact already 2-3nV/$\sqrt{Hz}$, which gives a C/N ratio of only 11dB for the case of the antenna in Fig. 3.5a (the one with the highest gain). This does not leave much margin for noise contributions from the front-end.

Defining how big this contribution can be is not easy for ERMES, because of the way sensitivity is specified. No precise numbers such as C/N ratio, BER or a similar quantity are in fact specified. Instead sensitivity is defined as the minimum signal level that guarantees proper pager operation. More precisely, the average usable sensitivity, as postulated by ERMES type approval, is the average of eight field strength measurements, where the pager is rotated in 45° increments after each measurement, that produces a message acceptance ratio of 80%. For this test a 55 character message is used, the full error correction capability of the code can be exploited, while a message is considered successful if no more than four contiguous characters are in error.

In the lack of any knowledge of parameters such as the performance of the antenna when body mounted (actually when mounted on the "simulated man" test fixture), BER performance of 4-PAM/FM, or error correction performance, no precise maximum noise figure of the front-end can be specified. In this work we have therefore tried to achieve the lowest noise figure which is compatible with low power consumption.

### 3.4 System Planning

Planning a receiver consists first of finding a suitable architecture that fulfills our requirements, then assigning the wanted performance figures, such as gain, operating frequency, linearity or noise specifications to each block. Apart from achieving sufficient performance to fulfill type approval specifications, many other requirements must be taken into consideration. Dictated by the wrist-watch implementation, the most stringent of them are certainly lowest power consumption and smallest size as said before.

Among all usual architectures, single conversion superhet and direct
conversion were considered for this design because of their good performances. Double superhet - as used by many pagers today - will have even better performance, at the expense of an extra IF filter. This 2nd IF filter, usually a 455kHz ceramic filter, very bulky (roughly a cube, with 5–6mm per side) and certainly difficult to place into a watch, was considered unacceptable for this design.

A first conflict of requirements already appears here: if power consumption has to be minimized the superhet receiver will be the best choice, while the direct conversion receiver will require the lowest number of external components. In order to get an idea of the trade-offs involved, the two architectures need to be studied somewhat more in detail.

### 3.4.1 The Superhet Receiver

In Fig. 3.6 the plan for a single superhet receiver is shown. This plan also shows the most important external components, while the parts that will be integrated on chip are shaded. The table below the block diagram contains the main specifications for every block, as well as an estimate for their current consumption. The gain shown for the IF amp actually represents the combined gain of IF amp and demodulator, while on this chip no prescaler has been implemented. Their values are

<table>
<thead>
<tr>
<th>Gain</th>
<th>25-30dB</th>
<th>20dB</th>
<th>-4dB</th>
<th>10dB</th>
<th>-1.5dB</th>
<th>0dB</th>
<th>50μA</th>
<th>6dB</th>
<th>15dB</th>
<th>6dB</th>
<th>105dBc</th>
<th>50μA</th>
<th>(150μA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise, F</td>
<td>3-5dB</td>
<td>3dB</td>
<td>4dB</td>
<td>15dB</td>
<td>1.5dB</td>
<td>6dB</td>
<td>50μA</td>
<td>15dB</td>
<td>10dB</td>
<td>50μA</td>
<td>50μA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idq</td>
<td>100μA</td>
<td>50μA</td>
<td>50μA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
therefore shown between brackets. An ESCL prescaler that fulfills the requirements of this system, although at higher power consumption, is described in section 4.

In this architecture the antenna is connected directly to the LNA through its high Q matching network, as defined before, for best noise figure. Because of its high Q (around 50) the matching network will also provide sufficient filtering to protect the LNA from strong out-of-band interferers, but will only provide a limited amount of image rejection (25–30dB). An image reject filter is therefore required before the mixer. This filter - a SAW device as [36] or similar - provides 60dB of image rejection, has an insertion loss of 4dB and requires 610Ω terminations. Packaged in a 9x7x2mm SMD case it is already one of the biggest parts of the receiver, and because of the relatively low operating frequency it is unlikely that much smaller parts will become available in the future.

After mixing to IF channel filtering occurs. The channel filter is a crystal filter as [36], with center frequency of 21.4 or 21.7MHz, insertion loss of 1.5dB and requiring terminating impedances of 1.5kΩ in parallel with 1pF. This filter, luckily, is somewhat smaller than the previous one, being packaged in a 7.5x5.1x1.3mm case. The IF strip must provide most of the gain and will typically feed a coherent demodulator directly. Roughly 80dB of gain must be provided here, possibly split between IF amp and demodulator to relax the requirements of the former. A very large gain at IF could in fact lead to instability. In this design the IF amp provides a gain of 26dB. A simple FM discriminator is better avoided in this receiver, partly because of its lower performance, and especially to avoid adding another bulky crystal resonator to the system. The better performance of a coherent detector is really needed here, to compensate for the somewhat weaker performance of the receiver dictated by the small antenna and low power consumption. AGC will be provided to the IF and the base-band amplifiers, while delayed AGC will be applied to the LNA for large input signals.

The front-end is completed with a VCO and a dual modulus prescaler. To obtain an adjacent channel rejection of 60dB, a phase noise figure of −105dBc/Hz at 25kHz offset will be required [18]. This specification, despite being quite moderate, will virtually exclude any fully integrated oscillator and dictates the use of an external high Q resonator. A usual Colpitts oscillator has therefore been implemented. To synthesize a local oscillator signal at 190MHz and 25kHz channel spacing, a dual
3.4. System Planning

modulus prescaler with 64/65 division ratio is required for the frequency synthesizer. Typically a bipolar ECL prescaler would be used. Since it has not been implemented in this chip, the prescaler is not discussed further here.

In order to define the gains of the various front-end blocks, the main parameters that have to be taken into consideration are noise figure, blocking and power consumption. Setting high gain to the LNA will help reduce the overall noise figure by minimizing noise contribution of mixer and interstage filter, but at the expense of higher power consumption of the LNA and risk of early mixer overloading. On the other hand, a lower LNA gain may improve front-end linearity and reduce LNA power consumption, but a very low noise mixer would be required to maintain an acceptable noise figure, which will very likely consume much power.

To get an estimate of the overall noise figure and the required gains, we need to first find the noise figures for all blocks. The low power consumption allowed for this design means that the noise figure cannot be very low. To summarize:

- The contribution of the LNA can be estimated from the shot noise of its input stage, provided this is the dominant noise source. Assuming a bipolar transistor biased at 50μA (half the budget for the LNA), an input referred noise voltage density of $2nV/\sqrt{Hz}$ results. The noise of the antenna corresponds to the noise of its (transformed) resistance $R_r + R_L$ of 300–800Ω, and is therefore $2-3.5nV/\sqrt{Hz}$. The noise figure referred to that resistance will therefore be typically around 1.3–3dB. The latter has been chosen as our objective specification shown in Fig. 3.6.

- Estimating the noise figure of the mixer is not straightforward, since many different noise sources contribute to it. Nevertheless experience has shown that the noise figure of a typical active mixer is around 12–15dB. The latter value has been assumed in our specification.

- For the IF amplifier, the noise figure can be estimated in a similar way as the LNA. Assuming 5–10μA of bias current, an $c_n = 4-6nV/\sqrt{Hz}$ will result, i.e. $F = 3-5$dB on the 1.5kΩ source impedance. With some margin a value of 6dB has been assumed.
The noise figure of a filter can be easily estimated by knowing that insertion loss in the passband is essentially resistive, thus $F$ corresponds to it. The interstage filter, with its insertion loss of 4dB, will therefore have $F = 4$dB, while for the IF filter, $F$ will be 1.5dB (both worst case).

When the noise figures are already quite high, it is good practice to make the LNA dominate the overall noise figure. This requires that its gain must be larger than the noise figure of the stages following it. A gain of 20dB will guarantee this and has therefore been chosen for our specification. The same reasoning can be made also for the mixer, whose gain has been set to 10dB. The overall noise figure for this receiver plan will therefore be 4.5dB for worst case filters, and around 4dB for typical ones (gives a SNR of $> 6$dB for the round antenna).

Having a high front-end gain may raise questions about linearity and blocking performance. Blocking and desensitization are specified in the type approval (table 3.3), which requires a minimum blocking immunity of 84dB$\mu$V/m. Using the data for antenna in Fig. 3.5a, $V_s = 27\mu$V results, which is transformed to 1.3mV, or $-44.5$dB$\mu$V\(^2\), after the matching network. Even considering a few dB increase in sensitivity when the antenna is body mounted, this value can still be assumed as a moderate $-40$dB$\mu$ for the LNA and $-24$dB$\mu$ for the mixer, both unproblematic values that can be realized also at low current.

Current consumption for this receiver plan is a very low 400$\mu$A, which is well within our objective of $< 1$mA. It is not difficult to imagine that because of the low operating frequencies, both the demodulator and the frequency synthesizer will consume much less than the remaining 600$\mu$A, and even the complete pager, with digital controller and display, will hardly consume more than 1.2–1.5mA, i.e. well within our expectations for a watch application.

The main objection to this plan is the large number of external components required and their size. Apart from the already mentioned RF and IF filters, this receiver will require the VCO tank, a matching network between mixer and IF filter (typically an RLC network, not shown in the plan), possibly the matching networks for the RF filter to cope with the low termination impedance of the latter, the loop

\[^2\text{0dBu} = 223.6\text{mV (Same voltage as 1mW on 50Ω).}\]
filter and the reference crystal oscillator for the PLL, some decoupling capacitors and the antenna matching network. A densely populated and somewhat expensive PC board must therefore be anticipated, while the antenna trimmer $C_{\text{tune}}$ must be mounted in a way to be accessible from outside without opening the watch to allow alignment, but normally covered to avoid water and dust entering the watch and to prevent tampering. On the lower side of the PC board, near to the battery case could be a good placement, such that the battery cover may cover also the trimmer.

3.4.2 The Direct Conversion Receiver

The same kind of study can be done also for the direct conversion receiver. In Fig. 3.7 the plan for such a receiver is shown. Again the main external components - reduced to antenna with its matching network and local oscillator tank in this case - are shown on white background while the integrated parts are shaded.

To make a plan for this receiver we must first choose a 90° phase shifter. Among the phase shifters shown in Fig. 2.7, only the analog RC-CR (Fig. 2.7d) and the digital divide-by-two (Fig. 2.7b) phase shifters are suitable for this design, transmission lines would simply be too long, while integrated coils would be way too bad at 170MHz. Plan 1 and plan 2 refer to the RC-CR and the digital phase shifters respectively.

For both plans, the front-end will be essentially similar to the superhet case - with the no longer needed filters removed and the mixers duplicated - thus most numbers have been retained. As previously discussed (section 2.2), many reasons, especially 2nd order linearity and local oscillator reradiation, dictate a fully differential signal path (all blocks in Fig. 3.7 are shown single-ended for simplicity). A fully differential LNA is therefore required in this design, which will consume about twice that of its single-ended counterpart especially if the 3dB noise figure has to be maintained. A value of 200$\mu$A has therefore been assumed. To guarantee that the noise figure will be dominated by the LNA, a gain of 16dB would be sufficient. Since a reduction of only 4dB of the gain will result in a negligible saving of current consumption, the gain of 20dB has been maintained. The overall noise figure will therefore benefit somewhat, and becomes 3.7dB typical for this plan.
Figure 3.7: Direct conversion receiver block diagram and planning. Plan 1 applies for an analog RC-CR phase shifter, while plan 2 applies for a digital divide-by-two phase shifter.

The mixers also need to be fully differential, not only for the same reasons as the LNA, but also to achieve a good local oscillator rejection at the IF port. Differently from the superhet receiver which uses an external crystal channel filter, now channel filtering must be performed by an integrated active filter, that could be easily saturated by a large local oscillator feedthrough. A double balanced Gilbert cell mixer identical to the one planned for the superhet receiver works best here, thus its parameters are the same.

The channel filters must perform various tasks in this design. Apart from the obvious function of separating the wanted channel from the
rest, the filters must provide a proper low impedance load to the mixers and provide antialiasing to the AD converters following the I,Q amplifiers. These filters will very likely be of active RC or $g_mC$ design, while SC filters can be ruled out because of aliasing. Although in the plan a gain of 0dB is shown, some voltage gain can be provided here to achieve a better noise figure and to relax the gain and F requirements of the I,Q amplifiers somewhat. The estimated current consumption of 150µA consists of 50µA for a transimpedance active load for the mixer and 100µA for a couple of filter stages.

The base-band I,Q amplifiers complete the signal path. Here most of the gain must be provided, while gain accuracy is necessary to achieve a good image suppression. To keep gain variability under control, it is better not to allocate a very large gain to the I,Q amplifiers, thus a total gain of 60dB has been assumed. This implies that the AD converters must have a high sensitivity (a few 10mV full scale) - which can be provided quite easily - to cope with the small signal that can be made available. A $g_m/g_0$ or a negative feedback amplifier can be used here, while gain control can be provided by switching in and out $g_m$ stages or feedback resistors. Precision, as well as current consumption, will be lower in the first case. 50µA has been assumed in this plan.

Where the two plans differ substantially is in the way the quadrature local oscillator signals are provided to the mixers. Plan 1 uses RC-CR filters and an oscillator running at 170MHz. The phase shifter consists of passive RC-CR filters driving two saturating amplifiers which restore equal amplitude to the two channels and provide a signal of sufficient amplitude (0dBu or so) to drive the mixers well in saturation. Their current consumption has been estimated at 50µA each. The filters will typically consist of resistors around 1kΩ and capacitors around 1pF. The latter value must be chosen to be much larger than the parasitic capacitances, which are not necessarily well matched, as required to obtain accurate I and Q signals with a phase error below 2° or so. Apart from being differential, the local oscillator and the prescaler will be essentially the same as in the superhet case, thus the same current consumption applies.

The part consuming the largest current in this plan is the local oscillator buffer required to drive the RC-CR filter with a clean sinewave at approximately 0dBu. To obtain the best efficiency as well as low distortion, an LC loaded class A amplifier can be used, which has a
theoretical efficiency of 50% (much less in a practical buffer) but requires two external inductors. The LC load, apart from filtering the buffered sinewave, will resonate away the capacitive part of the filter input impedance, leaving only 1kΩ to drive to 0dBu. Even so this load will still require 220μA RMS, meaning that a buffer able to drive it with sufficient linearity will easily consume 0.7–1mA, especially if worst case resistors with −25% tolerance or so are considered.

With the addition of two capacitors in series with the RC-CR filter, a matching network can be formed to increase the resistance that the buffer has to drive. Nevertheless, the low operating voltage - 2V worst case - does not leave much headroom for this. Assuming a 4:1 impedance transformation, the RMS current through the load is reduced to 110μA, but the buffer must already generate a 1.3Vpp differential voltage swing. Such a voltage swing is still unproblematic, but may be considered a kind of maximum practical limit. The current consumption for the buffer can therefore be estimated as 350–500μA. A value of 450μA, which gives a total of 500μA for the oscillator-buffer combination, has been assumed in our plan.

From the point of view of the external components required, this architecture is superior to the superhet one. Although roughly the same number of external parts are required, the bulky and expensive filters are eliminated. The most important parts required are in fact two capacitors and a trimmer for the antenna matching network, the LC tank for the local oscillator, the LC loads for the local oscillator buffer, and the loop filter and reference oscillator for the PLL.

On the other hand, the total current consumption for this direct conversion receiver plan is 1.45mA, which, while not disproportionate, is much higher than in the superhet case and is above the 1mA specification by a substantial amount. If the complete system is considered, this disadvantage is reduced somewhat, and a total current consumption for the whole pager of 2.3–2.6mA can be estimated. Although this is about twice that of the superhet pager, it is still a manageable value.

Somewhat better performance can be achieved by the 2nd receiver plan. In this plan the phase shifter consists of ECL digital dividers-by-two driven by a local oscillator of twice the required frequency, i.e. 340MHz. In this phase shifter, the local oscillator is required to drive the relatively light load represented by the divider inputs, while the
phase noise specification of $-105 \text{dBc/Hz}$ now applies for 50kHz offset. A relatively low current consumption can therefore be expected despite the higher frequency, and an estimate of $100 \mu A$ has been taken.

The actual phase shifter consists of two ECL D-latches, connected as a divider-by-two, and driving a mixer each. For best phase precision, highest mixer conversion gain and lowest noise figure, the output signals of these latches must be square waves with sufficiently large amplitude and fast slopes, again to drive the mixers well in saturation. A relatively high current consumption of $100 \mu A$ each has therefore been estimated. The prescaler is basically the same as for the other plans and thus the estimate for the current consumption has been maintained at $150 \mu A$. Nevertheless, connecting it to one of the 170MHz phase shifter outputs would cause a capacitive load imbalance, which in turn may lead to phase inaccuracy. To avoid this the prescaler may be preceded by a simple divider-by-two with no special accuracy requirements, and connected directly to the VCO. The current consumption of this divider has been estimated at $50 \mu A$, giving a total of $200 \mu A$ for the prescaler.

The total current consumption of the receiver plan 2 is a very reasonable 1.2mA, compatible with the intended application, but still much higher than the superhet receiver and already 20% higher than the specification despite the base-band circuits and the PLL are not yet included in this plan. An estimate for the complete pager is around 2-2.3mA, which is not that much larger than for the superhet pager. The number of external components is also minimized for this plan, the most important ones being just the antenna with its matching network, the local oscillator tank, the loop filter and reference oscillator for the PLL. A much lighter and cheaper PC board, small enough to fit in the smallest watch case envisaged, will be sufficient to accommodate all required parts quite easily.

Among the two direct conversion plans, plan 2 is certainly the most promising one. The slightly lower current consumption and especially the reduced number of external components make it preferable to plan 1. The local oscillator running at twice the required frequency is also an important advantage of this plan. Since the LO signal no longer lies within the received channel, reradiation from the oscillator - which occurs mostly from the LC tank - is much less critical than for plan 1, as up to 2nW of reradiated power is acceptable to get type approval. Provided that usual common sense is followed, placement of the LC
<table>
<thead>
<tr>
<th></th>
<th>Superhet</th>
<th>Direct-#1</th>
<th>Direct-#2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idd front-end</td>
<td>400μA</td>
<td>1.45mA</td>
<td>1.2mA</td>
</tr>
<tr>
<td>Idd Pager</td>
<td>1.2–1.5mA</td>
<td>2.3–2.6mA</td>
<td>2–2.3mA</td>
</tr>
<tr>
<td>Crystal filters</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td># of ext. RLC(^1)</td>
<td>9–14</td>
<td>9–12</td>
<td>5–6</td>
</tr>
<tr>
<td>F typical</td>
<td>4dB</td>
<td>3.7dB</td>
<td>3.7dB</td>
</tr>
</tbody>
</table>

**Table 3.5:** Characteristics of the 3 receiver plans. *(1-Only in the RF part, does not include the loop filter and decoupling capacitors.)*

Tank on the PC board is uncritical, and does not influence reradiation at 170MHz. In the case of plan 1, on the other hand, reradiation at 170MHz depends heavily from the placement of the LC tank and the LC load of the local oscillator buffer, i.e. parameters which are outside the reach of the IC designer which therefore has no means to guarantee the 2pW reradiation limit.

Differently from plan 1, the divided 170MHz local oscillator signal of plan 2 stays on-chip, where it is less susceptible to reradiation because of the small size of the chip, and where it is also under the control of the IC designer who can reduce reradiation with adequate layout techniques. The major source of reradiation at 170MHz will be in fact capacitive or substrate coupling from the phase shifter output to the LNA input, and thus to the antenna.

A comparison of the three plans studied here is shown in table 3.5. Basically all three plans lead to a practical and manufacturable receiver, although direct conversion will not respect the maximum current consumption specification of 1mA. Which plan to choose is now the responsibility of the customer, who must know if for him a long operating time - i.e. lowest power consumption at the price of a larger, more expensive receiver - is more valuable than a low cost receiver with short operating time or requiring a large battery. Since at this point it was not possible to favour either superhet or direct conversion, for this work we tried to accommodate both architectures at the same time, although with some preference for the superhet receiver because of its simpler structure and much lower power consumption.
3.5 The ERMES Chip

In order to prove which kind of performance can be achieved by the low power circuits required for this project, a chip containing all critical ones has been designed. The block diagram is shown in Fig. 3.8. It contains a 170MHz RF front-end and an IF-strip. The front-end comprises a low noise amplifier with gain control, a double balanced mixer and an LC local oscillator. The IF strip consists of an IF amplifier and a mixer with active output filter. This chip has been implemented in the AMS BAE 1.2μm BiCMOS process, with analog extensions such as poly resistors and poly-poly capacitors. The most important characteristics of this process appear in table 3.6. The design is largely bipolar, with the MOS transistors used only for biasing purposes.

The purpose of this chip was to demonstrate the feasibility and the performance of the receiver blocks required for both superhet and direct conversion receivers. Therefore our choice about what to integrate and how was directed more by practical considerations such as making characterization of the various parts easy rather than receiver completeness. Most important for us was that:

- All critical receiver blocks required to implement both superhet and direct conversion architectures must be integrated on this chip, even if they will typically not coexist on the same chip in a complete system.
### Table 3.6: Some important characteristics of the AMS BAE 1.2μm BiCMOS process.

<table>
<thead>
<tr>
<th>Process</th>
<th>Core Process</th>
<th>p-Epi, twin well</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design rule</td>
<td>1.2μm</td>
<td></td>
</tr>
</tbody>
</table>

| NPN bipolar   | Type         | Vertical, poly emitter |
|               | Peak $f_t$   | 8GHz                |
|               | Beta         | 100                 |
|               | IKF          | 1.7mA/μm emitter width |

| PNP bipolar   | Type         | Lateral parasitic  |
|               | Peak $f_t$   | 500MHz              |
|               | Beta         | 82                  |
|               | IKF          | 1.1μA/μm emitter width |

| n-MOS         | $V_t$        | 0.72V               |
|               | $\mu_0$      | 510cm²/Vs           |
|               | $K'_n$       | 80μA/V²             |

| p-MOS         | $V_t$        | −0.75V              |
|               | $\mu_0$      | 155cm²/Vs           |
|               | $K'_p$       | 24μA/V²             |

| Resistors     | Poly-1       | 24Ω/□               |
|               | Poly-2       | 85Ω/□               |
|               | N-well       | 5.8kΩ/□ (7kΩ/□)     |

| Capacitors    | Gate         | 1.57fF/μm²          |
|               | Poly–Poly    | 0.69fF/μm²          |
|               | Metal–Metal  | 0.035fF/μm²         |

- All parts should be measurable and all important nodes should be observable, at least indirectly.

The top part of the block diagram in Fig. 3.8 shows the LNA-mixer combination and the 21.4MHz IF strip. To avoid excessive capacitive loading, especially in the case an RF interstage filter is not used, the output node of both the RF and IF amplifiers has been connected internally to the mixers and therefore is not externally accessible for measurements. Thus, the circuit has been duplicated omitting the amplifiers to allow the latter to be characterized indirectly. This is shown in the bottom part of the same figure.

The IF strip includes an IF amplifier driving a single mixer, i.e. only
3.5. The ERMES Chip

half a demodulator. Although a complete receiver will necessarily include both mixers, no second mixer or phase shifters for the Q channel have been implemented on this chip, partly because they will not add significantly to the purpose of the chip, and partly to save pads and design time. At these low frequencies the phase shifter is quite uncritical, and will certainly consist of an ECL divider delivering very accurate quadrature signals.

An important parameter to which attention has been paid is independence of circuit characteristics from temperature, process and supply voltage variations. This is actually an obvious and very basic requirement for any circuit - including prototypes - especially at RF where typically open loop structures are implemented and overall negative feedback cannot be used to obtain stable and predictable gains. Supply voltage independent biasing which consists of a $\Delta V_{BE}$ current source and a 1.2V bandgap voltage reference has therefore been added. While providing all necessary bias currents internally, the on-chip $\Delta V_{BE}$ current source is also used to provide temperature and process independent gains, defined by resistor ratios only, as explained later.

ESD protection diodes have been provided on all pads. Although this may seem a bit excessive at first on a prototype chip like this, one must not forget that at RF ESD diodes will have a strong influence on circuit performance. Leaving the ESD diodes away may certainly lead to improved performance for most circuits, but such circuits will also be absolutely unusable in any industrial environment. In other words, not considering ESD protection - just like not considering process or temperature independence - does not demonstrate anything thus voiding the purpose of the chip, and showing results obtained that way can be considered a lie at best.

Considering the effects of ESD protection on the finished circuit only will also not necessarily lead to good results. In this circuit, where possible, we have tried to put at least the capacitive part of the ESD diodes impedance to good use, e.g. as part of the capacitor in LC resonators.

To reduce spurious coupling between blocks through the supplies, on-chip decoupling capacitors have been used on all supplies, except the voltage reference, in addition to the ordinary off-chip decoupling capacitors. These are poly-poly capacitors, for a total of 20pF.
Chapter 3. ERMES Pager Receiver Chip

3.6 Low Current Operation

In general if low power operation of an RF circuit is required a decrease in performance has to be conceded. When a bipolar transistor is operated at very low current levels the following problems occur:

- $f_t$ is decreased;
- The equivalent input noise voltage is increased;
- The signal handling capability is reduced.

3.6.1 Noise Performance

The noise voltage of a bipolar transistor consists mainly of thermal noise in the base and emitter resistances, shot noise due to DC current flow and 1/f flicker noise. At RF, flicker noise is negligible, although it may have to be considered for direct conversion receivers. At low current, transistor noise voltage is increased and shot noise current ($v_b^2 = 2qI_b$) is decreased, thus noise matching occurs for large source resistances $R_{opt}$, that in many cases cannot be obtained with practical matching networks. If the impedance of the signal source is well below $R_{opt}$, as it is in our case, the noise current becomes negligible and the overall noise will be dominated by noise voltage. The noise voltage of a transistor, referred to the base, is then expressed as:

$$e_n^2 = 4kT(R_B + R_E + (2g_m)^{-1})$$

where the first two terms represent thermal noise, and the last one represents shot noise. Since the $g_m$ of a bipolar transistor is a function of its collector current only, in a low current design shot noise becomes dominant. This shot noise is inherent to DC current flowing through a junction and cannot be decreased by careful design or the use of an advanced IC process. The best trade-off between noise and power consumption has to be found. It is not possible to reduce both simultaneously.

Thermal noise originates mainly from the base resistance and can be reduced if the size of the transistor is increased. Unfortunately, this
3.6. Low Current Operation

decreases the current density in the transistor and therefore $f_t$, while the junction capacitances are increased. The effect of junction capacitances on frequency response can be reduced by proper design of the amplifier as can be seen in the next section. However if the transistor is made too large, an unacceptable $f_t$ reduction may result. Thus, the transistor has to be sized just large enough to make thermal noise low compared to shot noise. Any further increase in size results only in decreased high frequency performance with almost no noise improvement.

3.6.2 Signal Handling Capability

Low current operation may cause an excessive increase in the input impedance of the amplifier as well as a reduction of its dynamic range, although indirectly. The small signal input impedance ($r_\pi$ and $C_b$) of a transistor is expressed as

$$r_\pi = \frac{\beta}{g_m}; \quad C_b = \tau F g_m$$

and is therefore increased by reducing the collector current, eventually becoming larger than the impedance of the parasitic elements. Matching a high impedance to a usual (low impedance) source is critical, the bandwidth of the matching network may become narrow and some form of alignment may become necessary. This difficulty may be potentially overcome if matching is done in two or more steps, but this increases losses, complexity and cost. Due to the large impedance ratio, a substantial voltage gain has to be expected in the input matching network. For example, for the one used in this work a voltage gain ($Q$) of 40–50 is required, while even the one used during the measurements to match the 50Ω instruments to the receiver (see section 3.16) has a voltage gain of 17dB.

Since the relationship between collector current and base-emitter voltage is exponential, the intercept point (IP3) is independent of the collector current, and is located at $A=87.5\text{mV}=-8.2\text{dBu}$. However this value is decreased by the voltage gain of the matching network, resulting in a decrease of the overall signal handling capability. In our case the IP3, as seen from the antenna, before the matching network (not the LNA input), can be no better than $-42\text{dBu}$ and will likely be around
Chapter 3. ERMES Pager Receiver Chip

---

Figure 3.9: Schematic diagram of the LNA.

---

-55—60dBu because of the LNA gain. Luckily such a low value is acceptable in our case because of the very low antenna efficiency.

### 3.7 The Low Noise Amplifier

Since for this design some preference has been given to the superhet architecture, a single-ended LNA as planned in Fig. 3.6 has been preferred. The target requirements are therefore the ones shown in the table below the same figure. In addition to that, 20dB of AGC range has been allocated to the LNA, which will be activated for strong signals only (delayed AGC) to prevent saturation. Since the output node of the LNA stays on chip and is connected directly to the mixer, no special requirement about its impedance level exists for this design, and the only need to be fulfilled is adequate driving capability to the (high impedance) mixer input. A complete superhet receiver will of course require the interstage filter. How to fulfill also this requirement will be discussed later.

In order to get 20dB of well controlled gain at low power, the two stage amplifier shown in Fig. 3.9 has been chosen. This design consists of a transconductance stage driving a transimpedance amplifier. Although the peak \( f_t \) of the available bipolar transistor is more than adequate for this design, some problems may be expected because of its high base resistance, which may limit both the frequency response and the noise performance.
3.7. The Low Noise Amplifier

3.7.1 The Input Stage

The most important consideration for the design of the input stage was to achieve a sufficiently low noise with low power consumption. To obtain a high input impedance, as it is required to voltage match the antenna, a common emitter bipolar transistor amplifier has been used. MOS transistors were simply discarded because of the much lower transconductance and higher gate and parasitic capacitances that will necessarily lead to poor RF performance. Transistors $Q_1$, $M_1$ and resistor $R_1$ form the transconductance amplifier. Roughly half of the current budget has been assigned to the input stage, thus a nominal collector current of only 48$\mu$A has been chosen for $Q_1$, which is just enough to prevent an unacceptable $f_t$ reduction.

Due to the low bias current of this transistor, a high input impedance
\[ r_i = \beta/g_m = 54k\Omega \] is realized. Although this represents a nearly ideal voltage matching to the antenna, it will lead to impractical matching network components and will very likely be dominated by rather unpredictable (capacitive) parasitics. $R_1$, apart from biasing the base of $Q_1$, will reduce the amplifier input impedance to a practical and sufficiently well controlled value by applying some negative feedback around $Q_1$.

Once the collector current is set, the minimum noise performance of $Q_1$ is also determined. At the chosen operating current, the shot noise of $Q_1$, i.e. the major noise contributor of the amplifier, becomes about $2.2nV/\sqrt{Hz}$. To get the best noise performance, noise matching to the antenna could be used, that, neglecting $r_B$, requires a source impedance of:

\[ R_{opt} = \frac{e_n}{i_n} = \frac{\sqrt{\beta}}{g_m} = 5.4k\Omega \]

for $\beta=100$. Since the antenna resistance $R_r + R_L$ is about 100–160m$\Omega$ (table 3.4), a matching network with a (loaded) Q of 180–230 would be required to noise match the antenna. Such a high Q matching network is certainly impractical for a low cost application and excessively sensitive to component tolerances, while alignment becomes critical. Once the noise of $r_B$ and the other noise sources within the LNA are considered, the required Q will become even higher. Noise matching therefore cannot be used, noise current becomes indeed negligible and the total noise of the amplifier will be dominated by noise voltage sources.
To prevent $r_B$ from contributing too much noise, $Q_1$ must be sized accordingly. As a trade-off between stray capacitances, effective transit frequency and base resistance, a transistor with a relatively large emitter size of $20\mu\text{m} \times 1.2\mu\text{m}$ has been used. The transistor has been laid out with two $10\mu\text{m}$ long emitter stripes and one base contact between them. Such layout has an almost square form factor, which yield the smallest area and therefore the smallest parasitic capacitances - roughly $260\text{fF}$ collector-substrate and $100\text{fF}$ collector-base. The transistor has an $r_B = 105\Omega$, which adds another $1.32\text{nV}/\sqrt{\text{Hz}}$, thus the total input referred noise voltage due to $Q_1$ becomes $2.56\text{nV}/\sqrt{\text{Hz}}$. Any further improvement, while being of nearly no benefit, would require a much bigger transistor with excessively poor RF performance at the chosen collector current.

Biasing for $Q_1$ is provided by transistor $M_1$ and resistor $R_1$. $M_1$ is a pMOS transistor with $W/L = 48/3$ ($\mu\text{m}$), whose gate is connected to another transistor, with $W/L = 12/3$ to form a 4:1 current mirror (not shown). To prevent noise from the current mirror and the associated circuitry (essentially the DA converter used to set the gain) from degrading the noise performance of the LNA, $M_1$'s gate is decoupled to VDD with an external capacitor. Due to its low $g_m$ and the large $g_m$ ratio between $Q_1$ and $M_1$, noise contribution from the latter, being only $770\text{pV}/\sqrt{\text{Hz}}$ referred to the LNA input, is not dominant. The size of this transistor was chosen as a compromise between a sufficiently low gate overdrive as required to guarantee performance also at the low voltage/worst technology/high temperature corner, and capacitive loading of $Q_1$'s collector.

$R_1$ is a $10k\Omega$ poly-2 resistor, whose purpose is to bias the base of $Q_1$ and to set the input impedance of the amplifier. Since the gain of $Q_1$ when loaded by the 2nd stage is $2.5\text{dB}$, an input impedance of $4k\Omega$ is realized. Because of the low (transformed) antenna impedance of a few $100\Omega$ and the fact that $R_1$'s noise voltage is not amplified by $Q_1$, the input referred noise due to $R_1$ is a low $700\text{pV}/\sqrt{\text{Hz}}$ (for zero $\Omega$ source impedance).

Gain control is achieved by varying the bias current of $Q_1$. Since the $g_m$ of a bipolar transistor is a linear function of its collector current, AGC action is linear. A dB scale can be easily realized in the AGC DA converter, e.g. using four binary weighted current sources the LNA gain can be varied from 2 to $20\text{dB}$ in $6\text{dB}$ steps. Since the input impedance
of the LNA will vary as a result of AGC action, some deviation from the ideal gain steps must be expected because of Q variations in the antenna matching network. Nevertheless $R_1$ will limit the maximum input impedance to below 10kΩ even if $Q_1$ is switched off completely. This, together with the losses in the antenna ($R_L$), will prevent excessive Q variations due to input impedance changes. Thus gain variations of a few dB at most can be expected.

### 3.7.2 The Output Stage

The transimpedance output stage consists of $Q_2$ and the feedback resistor $R_2$. Since the maximum $g_m$ of the input stage is 1.85mS (at 48μA), this stage must provide a transimpedance of 5.4kΩ together with a voltage gain of 17.5dB to obtain a total gain of 20dB. Since $Q_2$ hardly contributes to noise, it has minimum emitter area ($3\mu \text{m} \times 1.2\mu \text{m}$) to keep all parasitic capacitances low and is biased at 48μA by $M_2$, which is a pMOS transistor with $W/L = 48/3$. The required transimpedance is then obtained with a feedback resistor $R_2 = 10kΩ$ (poly-2). Just like $M_1$, $M_2$'s gate is also connected to a 12/3 transistor to form a 4:1 current mirror and decoupled to VDD to prevent noise contributions from the bias circuitry. Because of the high operating frequency AC coupling to the input stage is feasible with a small 2pF capacitor ($C_1$).

The stray capacitances connected to the base of $Q_2$ - mainly the collector-substrate capacitance of $Q_1$ and the bottom plate capacitance of $C_1$ - add a pole in the feedback around $Q_2$, which may cause peaking in the frequency response. The large size of $Q_1$ results in a large $C_{JS}$ of about 260fF, while the bottom plate of $C_1$ adds another 180fF. This peaking can be avoided with a small compensation capacitor ($C_2$) of 35fF in parallel to $R_2$. With such compensation the LNA is well behaved, showing a flat frequency response and no peaking at all.

The major noise contributors of the 2nd stage are $Q_2$ and the feedback resistor $R_2$. The noise contribution from $Q_2$, which consists of both $r_P$ noise and shot noise at the given bias current, appear as a total of 1.6nV/$\sqrt{Hz}$ referred to the LNA input. Feedback resistors always increase the noise of an amplifier, and $R_2$ is no exception. Its 12.6nV/$\sqrt{Hz}$ thermal noise appears unmodified at the LNA output, thus an additional equivalent input noise of 1.26nV/$\sqrt{Hz}$ is generated. Noise from $M_2$ is
in this case really negligible, being below 400pV/√Hz.

At this point all major noise contributors of the LNA are known and a total noise voltage density of 3.46nV/√Hz can be calculated, while a Spice simulation of the LNA gives 3.54nV/√Hz. A noise figure around 3–5dB can therefore be expected referred to the given antenna resistance (300–800Ω), which is somewhat above, or just fulfills our 3dB objective. Since in this LNA the dominant noise source is shot noise, improving this number really requires an increase of bias current above our 100μA objective. Some minor improvement can be achieved by using larger transistors to decrease their $r_B$ further, but if a significant reduction of noise figure is required, it can only be achieved at higher current.

Despite the low operating current, linearity is not an issue for this LNA since distortion of the mixer will clearly dominate. Knowing that the IP3 of a bipolar transistor is −8.2dBu and that the voltage gain of the input stage of the LNA is 2.5dB, a value of −10.7dBu can be calculated, which is well above our requirement of −30dBu (10dB higher than 1dB compression point).

### 3.7.3 Interstage Filter

Although the output of the LNA is connected internally to the mixer as is the case for a direct conversion receiver, the superhet receiver requires an interstage filter to achieve sufficient image rejection. This implies that the output of the LNA must go off-chip and provide the proper termination impedance of 610Ω to the filter.

Luckily the output impedance of the transimpedance amplifier is sufficiently well controlled and can be calculated, neglecting the output conductance of the active devices, as:

$$R_o = \frac{R'_1 + R_2}{1 + g_{m(Q_2)} R'_1}$$

where $R'_1$ is the transformed $R_1$, as it appears at the collector of $Q_1$ (8–9kΩ, depending on antenna impedance). Because of the relatively large resistors, the effect of stray capacitances on output impedance cannot be neglected. In our case the dominant stray capacitance is the 440fF ($C_{JS}$ of $Q_1$ and $C_1$ bottom plate) connected at the base of $Q_2$,
while the small compensation capacitance $C_2$ also has a slight effect on output capacitance. By properly choosing resistor values and the $g_m$ of $Q_2$ (bias current) any reasonable output impedance can be obtained. The LNA as designed happens to have an output impedance of about 2.5kΩ slightly inductive ($2700+j400$Ω simulated), although this was no design parameter.

An immediate way of achieving the 610Ω output impedance would be to recalculate the output stage, although this implies a somewhat higher collector current for $Q_2$. In this case some voltage gain reduction (smaller $R_2$) is perfectly acceptable to limit the increase of $Q_2$ collector current, as voltage gain will be restored in the matching network between the filter and the mixer. A feasible LNA design, with low power consumption and broadband output matching, can be obtained this way.

If a couple of external components are acceptable, the LNA can be left unmodified and its output can be matched to the filter with an uncritical low Q (1.8) external LC matching network. The low current consumption of 96μA will therefore be retained while the parasitic capacitance at the output of the LNA, mainly due to the output pad and its ESD protection, will be compensated by the matching network itself. Also in this case the voltage attenuation of the matching network is acceptable, as it is regained at the other side of the filter. Alternatively a single stage LNA, using the whole 100μA budget for better gain and noise figure could be used, although a somewhat higher Q (2.8) will be required in the matching network to maintain the 20dB voltage gain.

### 3.7.4 LNA Biasing

One of the main characteristics to be realized by the biasing circuitry, apart from the obvious function of establishing the desired bias voltages and currents for the amplifier transistors, is to achieve a sufficient independence of LNA parameters from normal VDD, temperature and process variations. In general such independence must be designed in from the beginning, since in the absence of any global negative feedback or similar means, the LNA circuit itself can hardly provide it without some help from outside (the bias). Simply making simulations at the corners hoping for good results, or trying to find solutions at a later
date once the problem is confirmed, is in most cases useless and will very likely result in failures and unusable circuits.

The solution to this problem lies in the almost magic property of bipolar transistors, whose $g_m$ is directly proportional to the collector current and inversely proportional to temperature, i.e. $g_m = I_c/V_t$, with $V_t = kT/q$, for whatever technology or even semiconductor material. Although for this LNA no actual biasing circuit has been implemented on chip, independent biasing has been at least studied and planned. A proposed circuit is shown in Fig. 3.10, whose operation is the following:

- **Temperature independence:** The gain of a voltage amplifier consisting of a bipolar transistor ($g_m$) driving a constant load $R_l$ (the transimpedance of the 2nd stage), as is essentially our LNA, is $G = g_m \cdot R_l$. Temperature independence is therefore achieved if $g_m$ is held constant by some means. This is best achieved by biasing the transistors with a current source proportional to the absolute temperature (PTAT), a simplified example of which is shown in Fig. 3.10. The PTAT current generated by the $\Delta V_{BE}$ cell is used directly to bias $Q_2$, while $Q_1$ is biased through a current mode DA converter with logarithmic characteristic (I-DAC) which sets the overall gain of the LNA. An I-DAC consisting of four current sources of $1.5\mu$A, $1.5\mu$A, $3\mu$A and $6\mu$A derived from the same PTAT current source allows the gain to be set precisely from 2 to 20dB in 6dB steps.
3.7. The Low Noise Amplifier

- **VDD independence:** Since the LNA is biased with current sources, VDD independence is intrinsic. Nevertheless some attention must still be paid when dimensioning the current sources ($M_1-M_4$), to guarantee that they will stay well in saturation also at the low voltage/high temperature/worst process corner, as said before.

- **Process independence:** First of all, to obtain process independence, the gain of the LNA must be determined by well controlled elements such as resistors and transistor transconductances, and not parasitic elements. This is obtained e.g. by designing an LNA with sufficient bandwidth (1.5–2 times the nominal frequency is a good trade-off), while the gain should not depend on the $Q$ of a resonator, the output conductance of a transistor or similar ill controlled parameters. Assuming that the gain is set mainly by resistors and $g_m$’s as is the case in this LNA, the problem to be solved is the high resistor tolerance of typical IC processes. The $g_m$ of a transistor is in fact uniquely defined by its collector current (for low $r_E$, which can be assumed for any reasonable layout). In this process the tolerance of poly-2 resistors was specified at ±18%. The gain of our amplifier is still $G = g_m \cdot R_i = R_i \cdot I_c/V_t$, while $I_c$ is defined by the $\Delta V_{BE}$ current source as $I_c = M \cdot \Delta V_{BE}/R_3$, with $\Delta V_{BE} = V_t \cdot \ln(A)$, $A$ the area ratio of $Q_3$ and $Q_4$, and $M$ the current mirror ratio. Collecting all terms, the gain becomes $G = M \ln(A) \cdot R_i/R_3$. Provided that the same layout and material is used for all resistors for good tracking, the LNA gain will therefore be determined by geometrical dimensions and a resistor ratio only, independently from process, temperature and VDD variations.

Since this LNA is the only circuit that uses poly-2 resistors in this chip - the rest uses N-well resistors - a separate bias source using a poly-2 resistor as reference, independent of the rest of the chip would be required. In order to save design time, and since such bias source would have added very little to the value of this test chip, it has not been implemented and the bias points (drains of $M_3$ and $M_4$) have been brought out on two separate pads. This will also allow testing of the LNA at different bias points than calculated, if desired.

The use of different resistor types was dictated by the available resistivities (85Ω/□ for poly-2 and 7kΩ/□ for N-well) and the values required
72 Chapter 3. ERMES Pager Receiver Chip

Figure 3.11: Gilbert mixer principle.

by the various circuits (from 10kΩ to 160kΩ). A complete receiver chip would typically use high resistive poly for all resistors - which is available in many processes including newer versions of AMS BAE and have typically a value around 2–3kΩ/□ - and a single biasing circuit.

For all other circuits, which use N-well resistors, an internal $\Delta V_{BE}$ current source is provided. Independent biasing, using the principles explained above can therefore be obtained and demonstrated.

3.8 The Double Balanced Mixer

The main requirement for the mixer of a communication receiver like this, is that it provides sufficient gain and noise figure, 10dB and 15dB respectively according to the plans for both superhet and direct conversion architectures, at low current consumption and with a relatively low local oscillator amplitude. In addition, the mixer for a superhet receiver must provide the proper terminations to image and channel filters, while the one for a direct conversion receiver must have low even order distortion and good LO-IF rejection.

Although many mixer types exist, very few guarantee all the requirements mentioned above at the same time. If an IC implementation is envisaged, the above requirements are best satisfied by single or double balanced mixers based on the Gilbert cell, which are virtually left as the only choice. A Gilbert mixer is essentially a transconductance amplifier followed by current steering switches that implement biphase multiplication, as shown in Fig. 3.11. A single balanced mixer is implemented by the parts drawn in solid black, while the shaded parts are
3.8. The Double Balanced Mixer

required for the double balanced mixer. The actual implementation of these mixers is shown in Fig. 3.12.

3.8.1 Conversion Transconductance

Assuming ideal switches, the conversion transconductance can be calculated by noting that the transconductor output current is multiplied by the switching function $S(t)$, which ideally alternates at the LO frequency $\omega_{LO}$ between 0 and 1 for a single-ended output - or $-0.5$ and $+0.5$ if the output is taken differentially - for the single balanced mixer, and between $-1$ and $+1$ for the double balanced mixer (independently on how the output is taken). The switching function $S(t)$ can be expanded in Fourier series, resulting in:

$$S_{SBM}(t) = \frac{1}{2} + \frac{2}{\pi} \left( \cos(\omega_{LO}t) - \frac{1}{3} \cos(3\omega_{LO}t) + \frac{1}{5} \cos(5\omega_{LO}t) + ... \right)$$

(3.1)

for the single balanced mixer, and in:

$$S_{DBM}(t) = \frac{4}{\pi} \left( \cos(\omega_{LO}t) - \frac{1}{3} \cos(3\omega_{LO}t) + \frac{1}{5} \cos(5\omega_{LO}t) + ... \right)$$

(3.2)

for the double balanced one. If the output is taken differentially, the DC term of (3.1) disappears, while (3.2) remains unchanged.
The output current at the IF frequency can then be found by multiplying the output current of the transconductor \((g_{m0}V_{RF}(t))\) with \(S(t)\), which yields

\[
I_{IF(SBM)}(t) = g_{m0}\hat{V}_{RF} \frac{1}{\pi} \cos((\omega_{LO} - \omega_{RF})t)
\]

for the single balanced mixer and

\[
I_{IF(DBM)}(t) = g_{m0}\hat{V}_{RF} \frac{2}{\pi} \cos((\omega_{LO} - \omega_{RF})t)
\]

for the double balanced one, with \(\hat{V}_{RF}\) the peak voltage of the RF signal. The conversion \(g_m\)'s are therefore

\[
g_{mc(SBM)} = \frac{1}{\pi} \cdot g_{m0} \tag{3.3}
\]

for the single balanced mixer and

\[
g_{mc(DBM)} = \frac{2}{\pi} \cdot g_{m0} \tag{3.4}
\]

for the double balanced mixer respectively. Since actual switching by real transistors - especially when driven by a small sinusoidal LO signal - is less than perfect, the expected \(g_{mc}\) is usually somewhat lower than the one predicted by (3.3) and (3.4). Values around 1/2 and 1/4 \(g_{m0}\) respectively, are quite common.

### 3.8.2 Mixer Noise

As we have seen from the gain calculations, mixing is an intrinsically inefficient process, which causes signal attenuation. As a result also noise performance cannot be very good. The main noise contributors of a Gilbert mixer are:

- Noise from the input transconductor.
- Noise folding due to mixing process.
- Noise from the commutating switches.
To get an estimate of total mixer noise, the above noise sources have to be analyzed.

The input transconductor will typically generate current noise at its output, which reaches the switches directly along with voltage noise (e.g. $r_B$ noise), and gets amplified and converted to current by the transconductor itself. The transconductor consists of transistors, whose noise contribution can be calculated in the usual way. This noise is usually wideband, since no filtering exists at the output of the transconductor (apart from 1st order roll-off at 2–3 times the RF frequency caused by parasitic capacitance at the transconductor output node). Its whole contribution must therefore be considered (noise folding).

To calculate the effect of noise folding, we must first consider that useful signal and noise are exactly the same thing, and they behave exactly the same way as they pass through the mixer. Any noise generated in front, or by the transconductor itself, is therefore downconverted by the mixer just like any other signal to the same IF and multiplied by the same gain factor ($1/\pi$ or $2/\pi$).

Since the noise generated after the image filter is wideband, all its contributions also at image and harmonics of the LO signal must be considered. This is best understood by looking at Fig. 3.13, which shows the effect of noise folding. Assuming white noise, the total noise power converted to IF can be calculated as the infinite sum of all contributions, which results in $(\pi/2)^2$. The noise power is therefore $(\pi/2)^2$ higher than that of the desired (narrow band) signal, which gives a noise figure of...
3.9dB. This is a lower bound for $F$, and an actual mixer will likely be higher than that.

In the case of a single balanced mixer, if the output is taken single-ended, the DC term in eq. (3.1) must be considered too. This actually results in a doubling of noise power, which becomes $2(\pi/2)^2$ yielding a minimum noise figure of 6.9dB.

At this point the noise figure due to these two contributions can be calculated. A typical (good) transconductor (one transistor or a differential pair) has a noise figure around 2–3dB. Adding this to the previously calculated $F$ yields 6–7dB, which becomes 9–10dB for a single balanced mixer whose output is taken single-ended.

The third noise contribution of the mixer is more difficult to characterize, and is the noise injected by the commutating switches directly into the IF port during the time when each pair of switches remains in the on-state simultaneously.

As shown in Fig. 3.14, the local oscillator signal is typically a sine-wave, as generated by an LC oscillator which is required to achieve sufficient spectral purity, whose amplitude is necessarily limited to about 0dBu because of power consumption considerations. When the LO volt-
age is above a given threshold \( V_{th} \) (e.g. a couple of \( V_T \), which gives already a 7:1 collector current ratio), the switching transistors \( Q_2, Q_3 \) and \( Q_5, Q_6 \) can be considered completely switched on or off, thus they will either be shut down or act as cascode transistors, contributing very little noise to the output.

However there are two time slots \( \tau \) per period \( T \), where all switching transistors are on at the same time. During that time they will form a differential amplifier with large \( g_m \), which will contribute its own noise - either shot noise or amplified \( r_B \) noise - to the output. The \( g_m \) of the switching differential pairs can in fact be as high as the \( g_m \) of the input transconductor for the double balanced mixer or half of that for the single balanced mixer, and even worse if emitter degeneration is used to linearize the input transconductor.

To calculate the noise figure of the mixer precisely, we should consider the exact \( g_m \) variation during the time \( \tau \). To simplify things somewhat, we can consider that during \( \tau \) the \( g_m \) of the switching pair is constant and equal to the maximum \( g_m \), which occurs at the zero crossing of the local oscillator, although this will overestimate noise somewhat. For typical LO amplitudes (\( \approx 0 \, \text{dBu} \)), \( \tau \approx V_{th}/\pi \sqrt{2}V_{LO} \) is about 5% of the cycle.

At this point, the total noise current at the output of the mixer can be calculated by summing all contributions, i.e. that of the input transconductor, the switches and the load resistors \( R_L \), usually required to obtain impedance matching to an external filter, or to convert the output current to a voltage. Referring to Fig.3.12, this output noise current is:

\[
\begin{align*}
\dot{i}_{n(SBM)}^2 & = 4kT \cdot \left( \frac{1}{4} (0.5 + r_B g_m) g_m + \frac{1}{2} (0.5 G_m^2) + \frac{1}{4} (2r_B^2) G_m^2 + \frac{1}{R_L} \right) \\
\dot{i}_{n(DBM)}^2 & = 4kT \cdot \left( \frac{1}{2} (0.5 + r_B g_m) g_m + (0.5 G_m^2) + (2r_B^2) G_m^2 + \frac{1}{R_L} \right)
\end{align*}
\]

for the single balanced mixer, and:

\[
\begin{align*}
\dot{i}_{n(DBM)}^2 & = 4kT \cdot \left( \frac{1}{2} (0.5 + r_B g_m) g_m + (0.5 G_m^2) + \frac{1}{R_L} \right)
\end{align*}
\]
for the double balanced mixer, with $G_{m2}^2 = g_{m2}^2 \frac{2\tau}{T}$ the average $g_m$ square of one switching transistor. The meaning of all terms in eq. (3.5) and (3.6) is the following:

- The 1st term is the contribution of the input transconductance, which includes both shot noise ($0.5g_{m1}$), and base resistance noise ($r_{B1}g_{m1}^2$).
- The 2nd term ($0.5G_{m2}$) is the shot noise contribution of the switching transistors.
- The 3rd term ($2r_{B2}G_{m2}^2$) is the thermal noise of the base resistance of the switching transistors, amplified and converted to current by the $g_m$ of the same transistors.
- The last term is simply the noise current of the load resistors.

The input referred noise voltage of the mixers is then obtained by dividing the output current noise by the appropriate conversion $g_m$.

Having identified the major noise sources, it is now easier to find tricks and methods to reduce their contributions to total mixer noise.

An immediate way of reducing switch noise contributions is to reduce $\tau$ by using as large as possible a local oscillator amplitude. This will improve switch performance, thus the best conversion $g_m$ and noise figure will be realized. Allowing poor switching because of an insufficient LO amplitude will always result in very poor mixer performance. Unfortunately there exist also reasons to limit the LO amplitude to relatively low values. First of all a large LO amplitude will in general impose a large power consumption to the local oscillator, or the LO buffer if required. Since the switch transistors should never be allowed to saturate, or the IF signal would be lost via the LO port because of the CB junction becoming forward biased, a further limit exists, especially in low voltage applications. Mixer debiasing, caused by the (rectified) LO signal at the emitters of the switches, is another reason to limit the LO amplitude. Switch saturation and debiasing are the most common reasons for mixer conversion $g_m$ and $F$ to collapse at high LO amplitudes. A practical LO amplitude is $\approx$ 0dBu or so for most applications, as previously stated, nevertheless if more than this can be allowed without any particular drawbacks, there is no reason not to use a larger LO amplitude to improve mixer performance.
3.8. The Double Balanced Mixer

A second possibility to reduce $\tau$ would be to use a square LO signal. In most cases, however, such a signal is not available, and generating one with sufficiently steep slopes from the LO sinewave will require a quite large amount of power. In some cases a square LO signal is available e.g. as the output of a digital phase shifter. Although the slopes of such a signal are rarely made steep in order to conserve power, some noise advantage may still be realized.

Apart from reducing $\tau$, the noise figure of the mixer can be improved by reducing the contribution of the switches. The $g_m$ of the switch transistors depends only on their collector current which is usually imposed by mixer requirements, thus their shot noise cannot be easily reduced. On the other hand $r_B$ noise can be made small by using transistors with large geometry and good base contacts. Similarly to the LNA, the size of these transistors should be chosen just big enough to make $r_B$ noise small compared to shot noise and no more, since an excessive size would result in nearly no noise improvement, while stray capacitances may become excessive. In higher frequency designs, the stray capacitances may impose a maximum size for the switch transistors, thus $r_B$ noise may become dominant.

Noise of the input transconductor can be reduced by making its $g_m$ as large as possible for the given collector current, which is usually a parameter imposed by mixer requirements. In fact, although its output noise current will increase with the square root of $g_m$, the input referred noise voltage will decrease by the same factor. This means that, unless linearity or similar reasons dictate otherwise, it is better not to use any emitter degeneration. As for the switches, also here the $r_B$ needs to be reduced by using sufficiently big transistors, while implementing the transconductor as a simple transistor or differential pair is still the best choice from the point of view of noise. Most variants such as the Micromixer [32] or the cross coupled pair will in fact incur in a noise penalty of $2-3$dB.

As a concluding remark, we notice that in these noise calculations no distinction has been made between SSB or DSB measurements. The reason for that is that the distinction is only helpful during the actual noise figure measurement of a mixer, but becomes irrelevant once the measurement is done. To measure the noise figure, a wideband noise source (e.g. the Hewlett Packard 346 series) is used as signal generator. When a mixer is connected directly to such wideband noise source, use-
ful signal (the wideband noise) is present also at the image frequency and at all harmonics of the local oscillator along with mixer noise, while in a receiver useful signal is only present in one sideband. Noise folding therefore occurs and an excessively good $F$ is measured. Such a measurement, which is called a $DSB$ measurement, is quite meaningless since it cannot be used straightaway to predict the final noise figure in a receiver application.

The good measurement is the $SSB$ measurement, and is obtained by connecting the noise source to the mixer through an image filter. Useful noise source signal will therefore be present only in the proper sideband, while mixer noise will obviously be still the same. The SSB measurement will then give us the correct noise figure of the mixer, i.e. the one calculated with the equations (3.5) and (3.6), and the one that allow us to calculate $F$ for a complete receiver.

In many cases it is impractical to use an image filter in front of the mixer, leaving a $DSB$ measurement as the only alternative. Luckily, assuming that the response of the mixer is the same for both sidebands and that the mixer is sufficiently bandlimited at its input to filter off the noise source contributions at the harmonics of the LO, the noise figure is simply found by adding 3dB to the $DSB$ measurement, to account for the 3 dB larger signal due to the image response. This is very often the case in downconversion mixers because of the relatively small offset $(2f_{IF})$ between useful signal and the image, and especially if broadband input matching is used. In upconversion mixers or if narrowband matching must be used, the response at the two sidebands may be very different, while even the noise source spectrum may no longer be flat. In such cases the "$+3$dB rule" does not work anymore and the $DSB$ measurement is of little practical use (except as a guess, with a couple dBs tolerance).

3.8.3 Mixer Design

As shown in the plans (Figs. 3.6 and 3.7), our mixer should realize $10$dB of conversion gain and $15$dB of noise figure at $50\mu$A bias current. Proper terminations to the external filters (superhet) or good local oscillator rejection (direct conversion) must also be provided where required, while linearity requirements are quite moderate and do not represent any problem.
3.8. The Double Balanced Mixer

<table>
<thead>
<tr>
<th></th>
<th>SBM</th>
<th>DBM</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversion $g_m$</td>
<td>588</td>
<td>294</td>
<td>$\mu S$</td>
</tr>
<tr>
<td>Input IP3</td>
<td>-8.2</td>
<td>-2.2</td>
<td>dBu</td>
</tr>
<tr>
<td>LO rejection</td>
<td>no</td>
<td>&gt; 20</td>
<td>dB</td>
</tr>
<tr>
<td>Load for 10dB gain</td>
<td>5.4</td>
<td>10.8</td>
<td>k$\Omega$</td>
</tr>
<tr>
<td>Output $i_n$</td>
<td>4.31</td>
<td>3.47</td>
<td>$pA/\sqrt{Hz}$</td>
</tr>
<tr>
<td>Input referred $e_n$</td>
<td>7.33</td>
<td>11.8</td>
<td>$nV/\sqrt{Hz}$</td>
</tr>
</tbody>
</table>

Table 3.7: Ideal characteristics of single (SBM) and double (DBM) balanced mixers at 48$\mu$A total bias current.

Having said that virtually only single and double balanced Gilbert mixers fulfill our requirements, the first thing to do is to choose one among the two. In Table 3.7 the ideal characteristics of the two mixers, when biased with 48$\mu$A, are shown. The noise values include the noise generated by the $r_B$ of minimum size transistors (700$\Omega$) and the load resistors given in the table. Since this is a bipolar design - MOS transistors were not considered even for the switches because of their high capacitances - most of these characteristics are uniquely defined by the bias current.

As can be seen from the table, the single balanced mixer has twice the conversion $g_m$ - thus the proper voltage gain when driving the 1.5k$\Omega$ channel filter can be achieved with a lower Q matching network - and about 3dB lower noise than the double balanced mixer, making it certainly the best choice for a superheterodyne receiver. On the other hand the large local oscillator feedthrough would make it useless for a direct conversion receiver, where the active channel filters can be easily saturated by the local oscillator signal, leaving the double balanced mixer as the only alternative. As a small bonus, the output of a double balanced mixer can be taken single-ended without any noise penalty.

Because the requirements of the two receiver types are in conflict, making a proper choice really requires the receiver architecture to be defined. Since the double balanced mixer will accommodate both - although with some difficulties - it has been chosen for this work.

The schematic diagram of the mixer and VCO combination, which includes the external parts for the VCO and the load for the mixer,
Figure 3.15: Schematic diagram of the mixer-VCO combination.

is shown in Fig. 3.15. Transistors $Q_1$ to $Q_6$ form the double balanced mixer core. The RF signal from the LNA is AC coupled to the base of $Q_1$ to obtain high input impedance, with a small 2pF poly-poly capacitor. The input impedance of the mixer is therefore sufficiently high not to load the LNA output significantly. The base of $Q_4$ is decoupled to ground with another 2pF capacitor.

In order to save power, the local oscillator signal is applied to the base of $Q_2$ and $Q_6$ directly, without any extra buffering. Transistors $Q_2$ and $Q_6$ will nevertheless also serve as buffers for the LO signal. The local oscillator, which comprises $Q_7$ and $M_1$ is discussed in the next section. DC biasing to VDD for the switching transistors is provided by the LO inductor ($L_1$) for one side, while the other side is directly connected to VDD. The switching transistors have been biased to VDD since this allows the lowest power supply voltage for the mixer, which starts operating correctly already at about 1.5V.

The transistors have been dimensioned taking into account noise and collector stray capacitance. The noise from the output of the LNA is already 35nV/$\sqrt{Hz}$, which is large if compared to the noise of the input differential pair of the mixer for whatever transistor size is chosen. Minimum transistors with a $3\mu m \times 1.2\mu m$ emitter, whose $r_B$ is 700\Omega, have therefore been chosen to minimize their collector capacitance. The resulting noise voltage of the differential pair $Q_1$ and $Q_4$, referred to the input of the mixer is 6.4nV/$\sqrt{Hz}$, which is indeed negligible when compared to the noise produced by the LNA. The noise from the switching
transistors is dominated by shot noise, which is about 5 times larger than the noise of the 700Ω base resistance of a minimum transistor. Minimum transistors have therefore been chosen also for the switches, to minimize their collector capacitances.

The linearity requirements, as planned in Figs. 3.6 and 3.7 which show 1dB compression points of -24dBu and -20dBu respectively, are easily fulfilled by this mixer. The input referred intercept point is defined by the input differential pair and is 6dB higher than the iIP3 of a single transistor. Its value is thus -2.2dBu, as shown in table 3.7. The iCP of a differential pair is 12.6dB below iIP3 (see appendix A), i.e. -14.8dBu, which fulfills the specification easily, while leaving a comfortable margin in case some non IM3 dependent compression occurs at the mixer output. No linearization is therefore required. Even the single balanced mixer, with its iIP3 of -8.2dBu and iCP of -17.8dBu will fulfill the specification without problems.

3.8.4 Interface to the External Filters

In a superhet receiver, the mixer must interface with two filters: the interstage and the IF channel filter. Although in this design the mixer input is connected on-chip to the LNA output, in a complete superhet receiver the mixer input should be connected to the interstage filter, which has a single-ended output requiring a 610Ω termination. This is best obtained if a single balanced mixer with a common base input stage, biased at 43μA for $1/g_m = 610Ω$, is used. The correct input impedance would then be realized with no external components.

Although a double balanced mixer is a bit impractical in this case, this mixer as designed would have a balanced input with 2.2kΩ impedance at the emitters of $Q_1$ and $Q_4$. A low Q matching network combined with a balun (a total of 2 capacitors and 2 inductors) will allow proper matching to the filter while requiring no alignment.

The IF output is open collector and can be connected to an external ceramic filter if a superheterodyne receiver is desired or to a transimpedance active load for direct conversion. In the first case a matching network, $L_2-C_5$ as shown in Fig. 3.15, is required to convert the 1.5kΩ input impedance of the IF filter to a higher impedance, as required to achieve the desired 10dB gain. $R_3$ will provide the neces-
necessary well controlled resistive part to properly set the gain and terminate the IF filter. In order to simplify the connection to the unbalanced IF filter, the output has been taken single-ended. This saves a balun and an output pad. Since this mixer is double balanced, noise performance is not affected by the single-ended connection, but dynamic range may be reduced somewhat because of clipping at mixer output. In the case of a direct conversion receiver, the mixer must be terminated to a low impedance load for best dynamic range and noise performance. The proper way of doing this is shown in section 3.11 (the demodulator).

### 3.8.5 Mixer Biasing

In order to stabilize conversion gain and input impedance over temperature, the mixer must be biased with a PTAT current. A $\Delta V_{BE}$ current source is therefore provided on-chip, with one of its outputs ($I_b$) mirrored to the mixer via the 4:1 current mirror $M_2$ and $M_3$. Although the bias current is connected internally to the mixer, the Bias pin has been brought off-chip to allow for measurements or adjustments. A bypass capacitor must also be connected externally to the Bias pin to prevent noise from the $\Delta V_{BE}$ current source from reaching the mixer and especially the local oscillator.

Since the reference resistor for the $\Delta V_{BE}$ current source is a N-well resistor, process independent gain would be obtained when the matching resistor $R_3$ is integrated on-chip as a N-well resistor. In this case we have preferred to leave this resistor off-chip to have the maximum freedom when configuring the output circuit.
3.9 The Local Oscillator

For this application, a local oscillator frequency between 190.825MHz and 191.2MHz has to be synthesized if the superhet architecture with 21.4MHz IF is chosen, while the range for the direct conversion receiver coincides with the channel frequencies, i.e. 169.425MHz to 169.8MHz. Independently from the architecture chosen, the local oscillator required in this application is therefore a VCO with a moderate tuning range. Since for the direct conversion receiver a digital phase shifter gives the best performance and is certainly preferable, the actual frequencies generated by the VCO are twice the channel frequencies, while even harmonic distortion must be kept low.

Apart from generating the required frequencies, the main characteristics to be realized by this local oscillator are:

- Deliver sufficient signal amplitude to drive the mixer well in saturation.
- Provide a sufficiently clean signal, i.e. low phase noise.

Since ERMES is a narrow band system, phase noise of the VCO is quite critical. The effect of phase noise - both in-band and out-of-band - is shown in Figs. 3.16 and 3.17 respectively. In the first case, as shown in Fig. 3.16, in-band noise of the oscillator is converted to IF exactly as the desired signal is, and as such it is made indistinguishable from the
desired signal itself. Because of this noise, the sensitivity of the receiver is degraded. In our case, since noise in the front-end is necessarily somewhat higher than usual because of the low power requirement, it is important not to degrade it further with a poor local oscillator signal.

The effect of out-of-band phase noise is shown in Fig. 3.17. In this case the phase noise of the VCO will convert the neighbouring channels to IF, where they appear as noise. Although phase noise within the neighbouring channels is necessarily lower than in-band phase noise, its effect can still be quite important since many channels (at least the two adjacent channels) are converted to IF at once and their respective noise contributions are cumulated. Depending on how the base stations are constructed, some difference in transmitted power for the various channels may also exist, as shown (exaggerated) in Fig. 3.17.

To achieve an adjacent channel selectivity of 60dB, as suggested by [18], a phase noise figure of $-105\text{dBc/Hz}$ at 25kHz offset from the carrier is required. Such phase noise specification - as already said in section 3.4 - is actually quite moderate and can readily be realized by many oscillator topologies at low power. It however virtually excludes any fully integrated oscillator (either 1st or 2nd order) and dictates the use of an external high Q resonator.

The PLL is also of little help at reducing phase noise. Since stability requires a loop bandwidth which is much less than the reference frequency (one channel, 25kHz), phase noise is only reduced for small offsets from the carrier. This is shown in Fig. 3.18, where the typical noise spectrum

![Figure 3.18: Typical phase noise spectrum of a PLL.](image-url)
of a PLL synthesized local oscillator is sketched. Near to the carrier, where loop gain is high, phase noise will be essentially dominated by the reference crystal oscillator and is thus excellent. On the other hand, for larger offsets where the loop gain is very low, the spectrum will simply coincide with the one of the free running oscillator, thus no benefit is achieved. Some spurious signals, in most cases falling in the middle of the channels or at the boundaries between channels (as shown), and originating usually from leakages through the loop filter, may appear. These spurious signals will also degrade sensitivity by converting down neighbouring channels, and must therefore be minimized.

Having said that the local oscillator will be a 2nd order oscillator with external tank, we must now find an adequate oscillator topology that fulfills our requirements. As it was the case for the mixer, also here the requirements are different for the two architectures chosen, thus two different oscillators are required. In fact:

- A superhet implementation requires a 190MHz oscillator, balanced or unbalanced as desired, possibly able to drive the mixer directly to save power.

- A direct conversion receiver requires a 340MHz oscillator, balanced for low even harmonic distortion, that interfaces to the digital phase shifter.

Since for this design we have not implemented a phase shifter, only the superhet receiver can be implemented. Therefore an unbalanced 190MHz Colpitts oscillator as shown in Fig. 3.15 was selected.

3.9.1 Oscillation Amplitude

For all practical amplitudes and at frequencies sufficiently below the cut-off of the active component as is our case, the oscillation amplitude of a Colpitts oscillator is uniquely determined by the losses of the (external) tank and the bias current of the active component, whatever it is.

In our case, the oscillation amplitude can be determined from the transfer function of a bipolar transistor [37]. Expressing the base-
emitter voltage as:

\[ V_{BE} = V_b + V_p \cos(\omega_0 t) \]

with \( V_b \) the (DC) bias voltage and \( V_p \) the peak oscillation voltage, the collector current of \( Q_7 \) becomes:

\[ I_C = i_s \cdot e^{\frac{V_{BE}}{V_T}} = i_s \cdot e^{\frac{V_b}{V_T}} \cdot e^{\frac{V_p}{V_T} \cos(\omega_0 t)} \]

Expanding the collector current into a Fourier series, \( I_C \) can be rewritten as:

\[ I_C = i_s \cdot e^{\frac{V_b}{V_T}} \cdot \left( I_0(\frac{V_p}{V_T}) + 2I_1(\frac{V_p}{V_T}) \cos(\omega_0 t) + \ldots \right) \]

where \( I_n \) are \( n \)-th order modified Bessel functions. In the expression above, the 1st term represents the DC bias current \( I_{DC} \) supplied from the bias circuit to the transistor, while the 2nd term is the current at the fundamental frequency \( I_{\omega_0} \).

From the collector current, the oscillation amplitude \( V_p \) can now be determined from the characteristics of the tank. Let \( g_{mc} = 1/R_0 \) be the critical \( g_m \) required to just start oscillation, with \( R_0 \) the transresistance of the tank, as seen from the LO-c node to the \( C_1 \) voltage (Fig. 3.19).

Knowing that \( I_C \approx I_E \), the oscillation amplitude results as:

\[ V_p = I_{\omega_0} \cdot R_0 = \frac{I_{\omega_0}}{g_{mc}} = \frac{2I_{DC}}{g_{mc}} \cdot \frac{I_1(\frac{V_p}{V_T})}{I_0(\frac{V_p}{V_T})} \]

A solution for this transcendental equation can be found by noting that for practical oscillation amplitudes (around 0dBu or larger), the ratio

\[ \frac{V_p}{V_T} \approx \frac{I_1(0)}{I_0(0)} \]
$I_1/I_0$ is around 0.9–0.95, which tends to 1 for very large oscillation amplitudes. With little error a value of 0.95 can be assumed for every practical purpose, thus the oscillation amplitude becomes:

$$V_p \approx \frac{1.9I_{DC}}{g_{mc}}$$

independently from the precise parameters of the transistor.

The reason for this can be easily understood by considering that in a Colpitts oscillator the transistor operates in class-C, i.e. essentially as a switch. Since the bias current of the oscillator is well defined and is not allowed to grow, when oscillation builds up, the DC bias point of the transistor (originally class-A) will rapidly shift to allow class-C operation. In steady state, the transistor will conduct current for only a small portion of the cycle, and the ratio $I_1/I_0$, which can assume values between 0 and 1 only, is related to the conduction time, which can go from 100% for $I_1/I_0 = 0$ (i.e. class-A, no oscillation) to 0% for $I_1/I_0 = 1$ (i.e. an ideal switch).

The fact that the transistor acts essentially as a switch, not only means that the oscillator is insensitive to transistor parameters, but also implies that any active component (JFET, MOS, tube, ...) can be used instead without significantly changing oscillator performance, provided it has proper performance at the oscillation frequency. The choice of the active component must therefore be based on considerations different from oscillator operation.

Although the RF performance of the available MOS transistors is sufficiently good for this oscillator, the latter were discarded essentially because of the large and quite lossy gate capacitance that would have loaded the tank excessively. Nevertheless there exist cases where an MOS transistor would have been the best choice. A micropower 32768Hz crystal oscillator is one such example.

### 3.9.2 Design of the VCO

The design of an oscillator for a communication receiver like this is dictated more by the interface to the rest of the receiver, rather than parameters such as power consumption or phase noise. Once the proper
topology is selected, the latter can in fact be considered almost as minor
details, since the current consumed by the oscillator is in general small
compared with the rest of the receiver, while phase noise is guaranteed
by the high Q of the external tank. To design this oscillator we must
therefore first define how to couple it with the remaining circuitry, i.e.
the mixer, the prescaler and the PLL (loop filter), and which oscillation
amplitude we need. The problems to be solved here are the possible
interferences from mixer and prescaler, and the low available voltage
(< 2V) for oscillator tuning.

To reduce power consumption, the oscillator has been coupled di-
rectly to the mixer without buffering. This means that the oscillator
must deliver sufficient signal to it, while some attention is required, as
oscillator pull-in or instability may occur, as explained later.

In order to be driven well into saturation while avoiding forward
biasing the CE junction of the switches, the mixer needs a local oscillator
amplitude around 300–600mVPP. The oscillator itself also has a limit
for the minimum practical amplitude. In order to ensure a safe start-up
at all conditions, the \( g_{m0} \) of the transistor must be a few times (2–
3 times or so) larger than \( g_{mc} \). A minimum oscillation amplitude of
200–300mVPP at the BE junction of the transistor therefore results. To
fulfill both requirements, we have set the typical amplitude to about
500mVPP. Although the exact amplitude is not very important, some
kind of stability and independence from the usual parameters is still
desirable, especially considering that the tank is external and that the
IC designer does not have any real control over it. The integrated
portion of this local oscillator is in fact limited to the transistor and its
biasing current source.

The schematic diagram of the local oscillator is shown in Fig. 3.15. It
comprises transistor \( Q_7 \) (on-chip), biased by \( M_1 \), and the external par-
allel resonator \( L_1, C_0, C_1, C_2 \) and the varactor \( D_1 \) for frequency tuning.
The mixer is directly connected to the base of transistor \( Q_7 \), where the
signal has the largest amplitude, and is biased to \( VDD \) through the
inductor \( L_1 \). Although here a straight Colpitts oscillator has been im-
plemented, the external tank can be easily reconfigured to form other
oscillator types (e.g. Clapp) if desired. In the latter case, it is important
to provide a low DC resistance connection between the pins \( LO-b \) and
\( VDD \) - ideally a DC short i.e. an inductor or a choke - to ensure proper
DC biasing of both the VCO and the mixer, as well as low flicker noise.
3.9. The Local Oscillator

Assuming that $C_0$ is a large coupling capacitor, the $g_{mc}$ of the tank as shown in Fig. 3.15 can be calculated as:

\[
g_{mc} = R_L \omega_0^2 \frac{(C_1C_2 + C_1C_{D1} + C_2C_{D1})^2}{C_1C_2}
\]

with $R_L$ a resistor in series with the inductor (Fig. 3.19) where all the losses of the tank (certainly dominated by the inductor losses at these frequencies) are lumped. To determine the values of the tank components we must now consider tuning range and oscillator pull-in.

Because of the direct connection to the mixer, the oscillator could become unstable or pulled away from its nominal frequency by the RF signal, should this become sufficiently large. Although $Q_2$ and $Q_7$ of the mixer will actually buffer the local oscillator signal providing some isolation from the RF port, to avoid oscillator pull-in the impedance of the tank components $L_1$ and $C_{eq}$ must be low at the resonance frequency. This will increase $g_{mc}$ and thus power consumption. Nevertheless by not using a buffer a power consumption benefit is still achieved. Since the use of LC circuits to tune out the parasitic capacitances is not practical for the buffer (too many external components), the bandwidth of the latter must be achieved with a large $g_m$, resulting in a substantially increased current consumption.

In the absence of any specific measurement a complete characterization of oscillator pull-in is very complicated, and some guesswork is therefore needed. After considering impedance and signal levels at the mixer LO input ($LO-b$), an impedance of around 100Ω was chosen for the tank inductor, i.e. 100nH after rounding to the nearest E6 value. Measurements on the finished circuit eventually demonstrated that this was overkill, as neither oscillator pull-in or instability has been observed during the measurements even for unrealistic large input signals.

The next problem to be considered is the small available tuning voltage, which is limited to below 2V because of the minimum supply voltage requirement. To tune the oscillator to the required frequencies for nominal component values, the total capacitance must be variable between 6.93pF and 6.96pF, which at first sight can be accommodated by any low capacitance varactor diode. Nevertheless, if alignment has to be avoided, the varactor must also accommodate all tolerances of the external components - easily 10% - including the varactor itself. The real capacitance range therefore goes from 6.3pF to 7.8pF for 10% inductor
tolerance, meaning that most of the required capacitance should be allocated to $D_1$. This introduces two problems: a $g_{mc}$ increase because the varactor shunts the tank and a high sensitivity of the control voltage input.

The tank can now be completely specified. $C_1$ and $C_2$ were set both at 3.3pF, which gives the required capacitance together with a 5pF low voltage hyperabrupt varactor (e.g. Siemens BBY53). To conclude the design, only the bias current of the oscillator remains to be defined. Assuming a Q around 40–50 for the tank, which is readily obtained with the lowest cost components, a bias current between $54\mu A$ ($Q=40$) and $44\mu A$ ($Q=50$) gives the desired 500mV$_{pp}$ amplitude. The oscillator has therefore been biased to $48\mu A$ by $M_1$, which forms a 4:1 current mirror with $M_3$. The resulting amplitudes are $440$mV$_{pp}$ ($Q=40$) and $550$mV$_{pp}$ ($Q=50$) respectively.

Although this oscillator is not really the design with the lowest possible power consumption, its current consumption of $48\mu A$ is small when compared with the rest and fulfills our specifications. Since the oscillator is very stable and no pull-in was noticed during the measurements, there is still some margin for improvements by using a different tank.

### 3.9.3 Amplitude Regulation

In order to guarantee the proper amplitude and maintain it constant, the oscillator needs to be biased with the current required by the particular tank, and with a temperature coefficient that matches the one of the tank (somewhat positive).

This means that the tank would need to be specified precisely, including its Q, which is obviously not possible. Inductors and capacitors with the required precision might be difficult to find, expensive, not exist at all or become obsolete during receiver production, while the manufacturer of these parts may decide to *improve* the quality of them (especially the Q of the inductor) without notice. Such a situation will be highly inconvenient especially if it happens in the middle of a running production, while if no substitute can be found the chip (thus the whole pager design) would become unusable. This means that we cannot rely upon the tank characteristics to keep the amplitude constant and that some other means has to be found.
3.10. The IF Amplifier

The best and most usual way to solve this problem is to use an (integrated) amplitude regulator that keeps the amplitude near to the optimum value independently from the actual tank performance. Parts from many different vendors can therefore be used without worries, provided that minimum characteristics are fulfilled. A small power consumption improvement may even result, should the tank be improved at a later stage.

Being a relatively uncritical circuit, the amplitude regulator was not implemented in our case and the amplitude was set to the proper value by adjusting $C_1$ and $C_2$ accordingly. A possible amplitude regulator can be based on [38] or [39], which has also been used for the reference crystal oscillator of the GPS receiver described in this thesis (chap. 5). Since the oscillator is biased with a PTAT current source, overcompensation of the tank losses will very likely occur, thus the amplitude will increase somewhat with temperature.

3.10 The IF Amplifier

The IF amplifier is the first part of the IF strip, which is uniquely required for the superhet receiver. It has therefore been integrated as a separate block and apart from the supplies it does not share any signal or bias line with the RF front-end. Its main requirements are the following:

- Operating frequency: 21.4MHz (or 21.7MHz).
- Sufficient gain to boost the signal to an easily processable amplitude and mask demodulator noise.
- Single-ended input, that connects to the (single-ended) IF filter with as few external components as possible.
- IF filter termination on the required impedance: 1.5kΩ and 1pF in parallel [36].

To maintain the signal for the demodulator to an optimum level and to avoid saturating it, an AGC input must be provided. The noise performance of the IF amp is not critical due to the high gain of the preceding
stages. In our plan (Fig. 3.6) we have specified a typical noise figure of 6dB at 1.5kΩ source impedance, although somewhat higher values do not degrade receiver performance substantially. In this design, the external IF filter provides channel filtering and sufficient suppression of out-of-band interferers. Linearity and intermodulation are therefore not an issue for the IF amp, as a sufficient value can be easily guaranteed.

The schematic diagram of the IF amplifier is shown in Fig. 3.20. It is a scaled down version of the LNA, thus the same design considerations apply. Just like the LNA, it consists of a transconductance stage to which gain control is applied, driving a transimpedance amplifier. The transconductance stage is realized by $Q_1$ and the transimpedance amplifier by $Q_2, Q_3$ and the feedback resistor $R_2$.

In order to reduce power consumption, an unusually low gain of 26dB has been chosen for the IF amp, which is nevertheless sufficient to mask demodulator noise adequately. The rest of the gain (40–60dB) can then be obtained more efficiently at base-band by the I/Q channels.

Since noise is quite uncritical here, all transistors, including the input stage, have minimum size to minimize $C_{js}$, while a resistive termination for the IF filter can be used. Although the noise figure will be increased by 3dB, a resistive termination has the advantage of being precise and broadband thus guaranteeing optimum IF filter performance. The termination impedance will also become independent from the actual input impedance of the IF amplifier (which varies with AGC).
The bias current for $Q_1$ is a trade-off between current consumption and $Q_1$'s total noise contribution. If resistive termination can be accepted, it makes no sense to reduce noise contribution of $Q_1$ to very low levels at the expense of high current consumption, thus $Q_1$ shot noise has been set to be roughly the same as the noise of the 1.5kΩ termination resistor ($5nV/\sqrt{Hz}$). A collector current of 8.7μA results, which has been set to 12μA since this current is directly available from the internal bias current source. At this current, the input referred noise voltage of $Q_1$, including the contribution of its 700Ω base resistor, is 5.4nV/√Hz. The noise figure due to $Q_1$ and the termination resistor then becomes 5dB, and is hardly increased by the contributions of the rest of the IF amp (mainly the $1.5nV/\sqrt{Hz}$ of $R_2$).

Due to the low frequency and especially the sensitivity of the transimpedance amplifier to stray capacitances at its input, AC interstage coupling is not practical here, thus the two stages are DC coupled. The voltage at the collector of $Q_1$ is therefore defined by the transimpedance stage, is equal to $V_{BE}$ of $Q_2$, and can no longer be used to bias $Q_1$. Biasing for the first stage has then been obtained with a replica circuit consisting of $Q_{1a}$ and resistors $R_1$ and $R_{1a}$ (50kΩ each). The collector current is the same for both $Q_1$ and $Q_{1a}$. Its nominal value is the previously calculated 12μA and can be varied for gain control, which, similarly to the LNA, is also linear.

Transistors $Q_2$ and $Q_3$ form the transimpedance amplifier. Buffer $Q_3$ prevents the collector of $Q_2$ from being loaded by $R_2$ and the demodulator, allowing more gain and better precision to be obtained at a given collector current. Thanks to the buffer stage, the transimpedance is almost uniquely defined by $R_2$, which has to be set to 50kΩ (N-well) to obtain the desired gain of 26dB. The high value of $R_2$, together with $C_{js}$ of $Q_1$ (40fF), forms a pole at 80MHz, that while not causing instability, will lead to a peaked frequency response. Capacitor $C_2$ (a small 35fF poly-poly capacitor) compensates the amplifier thus preventing peaking, and reduces its bandwidth to about 50MHz.

### 3.10.1 Interface to the IF Filter

As previously said, the IF filter requires a termination impedance of 1.5kΩ and 1.0pF in parallel on both sides. In this design the termination
for the IF filter has not been integrated, and must be provided with an external 1.5kΩ resistor. Because of its small value, the capacitive part will be provided partly by the input capacitance of the IF amplifier (mainly the capacitance of the input pad and its protection diodes), and partly by the stray capacitance of the PCB, thus some experimentation will be needed to set the capacitance to the required value.

Being typically above 40kΩ and relatively constant, the input impedance of the IF amplifier will hardly affect matching. In a final version of this chip, the required resistor can be easily integrated if desired, e.g. by making $R_1$ and $R_{1b}$ equal to 1.5kΩ (low resistance poly for best precision). Alternatively a common base input stage will achieve the same result, although at a slightly higher current ($17\mu A$) than the present solution and while complicating gain control that can no longer be achieved by varying the collector current of the input transistor.

### 3.10.2 IF Amplifier Biasing

Each transistor of the IF amp is biased to $12\mu A$ with the current sources $M_1$ to $M_3$, which are connected as 1:1 current mirrors (not shown). A total current consumption of $48\mu A$ results. Similar to the LNA, independent biasing can be obtained by the same biasing and gain control circuit as shown in Fig. 3.10. A finer gain step (3dB or so) may be advantageous.
3.11. The Quadrature Demodulator

Since here all gain defining components are integrated and the gain is essentially $g_m(Q_1) \cdot R_2$, independent biasing is achieved by simply using the internal PTAT current source. In our design the DA converter was not implemented, and only $Q_2$ and $Q_3$ receive the internal bias directly, although the Bias-P pin is brought on a pad to allow for adjustments. $Q_1$ can be biased with external resistors to allow for gain setting, or using the uncommitted $12\mu$A output provided by the biasing current source for gain independence.

3.11 The Quadrature Demodulator

The purpose of the quadrature demodulator is to convert the IF input into I and Q base-band signals, suitable for AD conversion. Sufficient amplitude and phase accuracy to achieve a good image rejection (chap. 2.2) is required, as well as some gain, to prevent the relatively noisy base-band channels from degrading overall noise performance. Low pass filtering must also be provided to protect the AD converters from aliasing, although this is best provided by the base-band amplifiers. In this demodulator only a coarse prefiltering is provided, mainly to suppress the IF and the local oscillator signals from the output.

Amplitude accuracy in such demodulator depends essentially on gain mismatch in the down-conversion mixers and the base-band channels, while phase accuracy, on the other hand, is mainly determined by the precision of the $90^\circ$ phase shifter, although mismatch in the base-band filters may affect phase somewhat. As previously said in this design only one demodulator channel has been implemented, and no phase shifter has been designed.

The schematic diagram of the designed demodulator channel, comprising a down-conversion mixer and an output filter, is shown in Fig. 3.21. It consists of a double balanced mixer driving a transimpedance active load, where low-pass filtering is implemented. Despite its lower gain, the double balanced mixer is required to achieve a good local oscillator rejection. This is necessary since the sensitive transimpedance load would be easily saturated if a large signal, such as the local oscillator leakage of a single balanced mixer, is admitted to it.

Given that the IF mixer and its load are fully integrated, there is
no requirement regarding input and output impedances or conversion $g_m$, provided the overall gain and sufficient linearity can be obtained. The mixer has been biased with $24\mu A$, which is more than sufficient to obtain the desired gain and bandwidth. Its conversion $g_m$ is therefore $147\mu S$ (ideally).

The input signal is AC coupled to the mixer. Unlike the IF amplifier, AC coupling here is feasible with a small $2pF$ capacitor. The input impedance of the mixer is in fact a sufficiently high $45k\Omega$, mainly determined by the $50k\Omega$ bias resistors.

Since no IF local oscillator or phase shifter has been integrated on this chip, the local oscillator signals have to be supplied externally. The mixer accepts signals of about $300-600mV_p$, referred to $VDD$, as typically generated by a digital phase shifter (ECL divider). With the proper choice of the IF local oscillator frequency, this mixer could also be used for conversion to a 2nd IF, possibly fully integrated, provided its frequency is within the bandwidth of the output filter.

In order to obtain a good linearity, the mixer should not be allowed to clip at its output. This means that the impedance of its load must be low, especially at the local oscillator frequency (and its harmonics), where because of imperfect LO rejection of the mixer a relatively large signal can be expected. At base-band a transimpedance active load is best suited for this purpose and has been used here, while at higher frequencies the low impedance is provided by the two $5pF$ capacitors between mixer output and $VDD$ (one is marked $C_1$ in Fig. 3.21).

Similarly to LNA and IF amp, the gain of the demodulator is defined by the (conversion) $g_m$ of the mixer and the transimpedance of the output stage. Its nominal value is $18dB$, which yields a gain of $44dB$ for the complete IF strip. To obtain the required overall gain, only another $20-30dB$ must be provided by the base-band I/Q channels. Such gain can be easily implemented at low power and with excellent linearity and precision using closed-loop opamps - which will also realize the rest of the base-band filtering - preceded by a (linear) two-quadrant Gilbert multiplier for AGC.

The output filter is a 2nd-order multiple-feedback low pass filter, and is part of the transimpedance active load. As gain stage a very simple structure consisting of the differential pair $Q_1$ and $Q_2$ has been
chosen, whose collectors directly drive the output pads and the feedback network $R_1$, $R_2$, $R_3$, $C_1$ and $C_2$. This means that the open loop gain of this stage is low and sensitive to output load, thus only very low Q filters can be built with this simple structure, while the output load must be kept sufficiently large to prevent a reduction of transimpedance and a change in filter characteristic.

On the other hand the simple structure is fast (relative to the bias current) thus providing proper filtering also at higher frequencies, and does not require any additional common mode feedback. Provided that the sum of the tail and the collector currents is (nearly) zero, the common mode voltage of the transimpedance stage is in fact simply defined by the voltage drop on $R_1$ (300mV), which is in turn defined by the bias current of the mixer. The filter has two real poles at 250kHz and 1MHz. The two $C_1$ capacitors also provide first order filtering for common mode signals.

Biasing for both the mixer and the filter is provided by MOS current sources derived from the internal $\Delta V_{BE}$ bias source. The gain is therefore temperature compensated and process independent. The bias current for the mixer is the aforementioned $24\mu A$. Since in the worst case the transimpedance load must be able to source at least the bias current of the mixer, it has been biased to $24\mu A$, too, which gives a total current consumption of $48\mu A$.

### 3.12 The $\Delta V_{BE}$ Current Source and Bandgap Reference

In order to maintain the various gains and transconductances constant over temperature, process and supply voltage changes, an independent current source with PTAT characteristic is needed. For this purpose a $\Delta V_{BE}$ circuit has been used.

The schematic diagram of the biasing circuit, which comprises a $\Delta V_{BE}$ current source and a band-gap voltage reference, is shown in Fig. 3.22.

The PTAT current source consists of $Q_1$, $Q_2$, $M_6$, $M_7$ and the N-well
resistor $R_1$ of 29kΩ (should have been 36kΩ). Since an emitter area ratio $A = 4:1$ has been chosen, the $\Delta V_{BE}$ becomes:

$$\Delta V_{BE} = V_t \ln A = 36mV$$

which, together with $R_1$ of 29kΩ, defines a nominal current of 1.25μA (should have been 1μA). This current is then mirrored out by P-channel transistors (not shown) to the various receiver blocks. One of them, $M_5$, whose drain is directly connected to a pad, forms an uncommitted 1.25μA current source that can be used for characterizing the $\Delta V_{BE}$ cell or to obtain independent biasing of the 1st IF amp stage.

Start-up is guaranteed by leakage currents and no explicit start-up circuit was needed to ensure correct operation. Nevertheless in the final implementation a start-up circuit would be in order, to guarantee a quick start also at low temperatures. The one used for the voltage reference is suitable also here.

Both the RF mixer and the demodulator also require a voltage to bias their input transistors. This voltage is uncritical and its value should be about 1.2V, and has thus been realized as a band-gap voltage reference.

---

3The nominal value of the N-well sheet resistance was lowered just before the delivery of the layout. Because of time constraints the biasing circuit was left unchanged.
3.12. The $\Delta V_{BE}$ Current Source and Bandgap Reference

The band-gap reference uses a Brokaw cell, which consists of $Q_5$ and $Q_6$ whose emitter area ratio is 4:1. The value of $R_2$, where the $\Delta V_{BE}$ voltage of 36mV appears, is 21kΩ, which gives a bias current for $Q_5$ and $Q_6$ of 1.7μA. $R_3$, where the amplified $\Delta V_{BE}$ appears, is 159kΩ, which gives:

$$k\Delta V_{BE} = 2 \frac{R_1}{R_2} \Delta V_{BE} = 545mV$$

This, together with the $V_{BE}$ of $Q_6$ (615mV) yields an output voltage of 1.16V. Although it seems a bit low, in the absence of any better data this voltage was made equal to the nominal voltage of the standard cell bandgap voltage reference, as specified by the foundry. The latter could not be used for this project since it was not delivered to us. A simulation seems to prove this voltage, but a measurement over temperature of the finished circuit has not been made, thus the actual temperature coefficient is unconfirmed. The collectors of $Q_5$ and $Q_6$ are loaded by $R_4$ and $R_5$ of 134kΩ and connected to the error amplifier.

The error amplifier comprises $Q_3$ and $Q_4$ and the output stage $M_{11}$. The latter is P-channel for low drop-out operation, thus the whole circuit starts operating at a supply voltage of about 1.4V. When a P-channel output stage is used, Miller compensation is no longer practical as this would naturally lead to a rather poor supply rejection, especially at higher frequencies, as well as sensitivity to load capacitance. Since the latter is a decoupling capacitor, it should possibly have no restriction on its value range. Frequency compensation is therefore provided by the lead capacitor $C_1$ and the lag capacitor $C_2$, both 1.5pF. With this compensation scheme no external capacitor is needed for stability, but a capacitor of at least 10nF is certainly recommended for low noise and good high frequency line/load rejection.

This voltage reference can source up to 20μA or sink up to 5μA, which is adequate as the load (the mixers) sinks almost no current. To allow it to be switched off completely (e.g. for measurement purposes), its output is brought to the mixers through two external pads, while a separate $VDD$ has been used. Start-up is ensured by $M_3$ and $M_4$.

To obtain sufficient precision, matching between components is important, thus many good layout techniques have been used, as shown in Fig. 3.23. All bipolar transistors have an octagonal emitter shape with an area of 15μm² in place of the more usual 1.2μm wide stripe,
and are arranged in a common centroid fashion. For the N-well resistors nonminimum design rules have been used. $R_2$ is a single stripe, with no bends or the like and a width of 10$\mu$m, and is surrounded by $R_3$, which consists of seven $R_2$'s in series plus a 12k$\Omega$ chunk with the same layout. $R_4$ and $R_5$ consist of two 67.2k$\Omega$ 10$\mu$m wide resistors in series each, with common centroid for best matching. All MOS transistors also have rather large dimensions, $W \times L = 20 \times 5\mu m^2$, and are interleaved where required. And to conclude, the boundary of all components requiring good matching, especially the resistors (whose edges are unprotected), has been made sufficiently similar although dummy components were not used for this purpose.

### 3.13 The ESD Protection

Providing ESD protection to a chip is in general no trivial task. The layout of the protection structures must be carefully studied and verified, and it is better not to fiddle too much with it unless precise know-how is available to avoid ending up with ineffective protection. In most cases a recommended layout is delivered by the foundry, although in many cases only for digital I/O pads. Apart from providing protection it is important that the ESD structures do not add excessive capacitance to the protected pad, especially when RF pads are concerned. recommend. from the foundry.
For our technology, two protection structures were recommended. The first structure consists of large diodes connected between the pads and the supplies, as shown in Fig. 3.24. The $N$ diode is a 102$\mu$m x 6$\mu$m N-diffusion in the P substrate, while the $P$ diode consists of a P-diffusion with the same size in an N-well with N+ buried layer, connected to $VDD$. This structure should provide an ESD protection of 2kV (HBM) and has the lowest stray capacitance of the two (1pF).

The second structure, intended especially for digital I/O, consists of two large MOS transistors whose gates are connected alternatively to the supplies (for an input) or to a suitable driver (for an output). This structure has a larger stray capacitance than the previous one (about 4–5pF), while the transistors were strangely fingered, raising some doubt about their effectiveness. In an ESD event it is in fact very likely that only one finger will have to carry the whole ESD current. Nevertheless this structure should also provide 2kV (HBM) ESD protection.

In our chip, all signal pads and the power supply for the voltage reference are ESD protected by the first structure (Fig. 3.24). The capacitance of a complete protected pad, metal-1 + metal-2, square with 110$\mu$m per side, is 1pF consisting of about 450fF for the pad metal, 260fF for the $N$ diode and 290fF for the $P$ diode. The second structure has not been used at least on the basis of the higher stray capacitance.

The power supply pads also need to be protected, not only against ESD discharges on the pad itself, but also because the power supply must absorb the positive ESD discharges on all other pads. Typical protection schemes range from nothing at all or simple clamps (consisting of a bipolar NPN or NMOS transistor connected across the supplies and with the base (gate) grounded), to active clamps (usually based on
a bandgap voltage reference) or Zener diodes (very robust, but requiring extra masks).

In our case no protection structure was specified or recommended, and clamping was left to the core itself. This seems dubious at least, considering the number of small junctions typically connected to \textit{VDD} which can be easily damaged by an ESD event. Newer versions of this technology will in fact use a power supply clamp based on lateral parasitic NPN transistors. Since the development of proper ESD protection structures for \textit{VDD} was not the purpose of this work, no internal protection has been used and the \textit{VDD} pin was left unprotected. It is quite clear that this would be strictly unacceptable in a final design. Once the chip is mounted on its test PCB, thanks to the large external decoupling capacitors, the \textit{VDD} pin is sufficiently well protected and ESD damage is quite unlikely to occur.

### 3.14 Layout Considerations

Although operating frequencies and gains are not particularly high in this chip, placement and routing of the various parts must still be done with some care, as overall performance may be affected negatively. A photomicrograph of the chip is shown in Fig. 3.25.

The chip is clearly divided into 5 separated blocks, whose placement corresponds roughly to the one of the block diagram of Fig. 3.8. The left side contains the RF parts, the LNA-mixer combination on top and the RF mixer alone just below it. On the right side the IF and baseband parts are placed, the IF amp and demodulator combination above and the demodulator alone below. The biasing circuit has been placed in the middle. Apart the supplies and some biasing, all the parts are completely independent.

To simplify the layout most of the circuits have the same 80\textmu m height and connect together by abutment. Many signal connections are also made by abutment, to keep sensitive lines (e.g. the output of the LNA) as short as possible. Every row is terminated by a few 2pF decoupling capacitor blocks evenly distributed between all circuits, which also connect to the supplies by abutment, for a total of 20pF.
The LNA and the mixer-VCO combination are 202\(\mu m\) and 200\(\mu m\) long, while the length of the IF amp is 192\(\mu m\). The layout of the demodulator consists of two blocks, one containing the mixer and the active components of the filter whose length is 273\(\mu m\), and the other containing the filter RC network whose area is 346\(\times\)146\(\mu m^2\) (the two big blocks side-by-side on the right side of the chip). The size of the biasing circuit is 243\(\times\)137\(\mu m^2\), while the whole chip, including the pads, measures 2240\(\times\)980\(\mu m^2\).

In order to maintain a certain degree of precision and matching between components, nonminimum design rules were used where appropriate. All N-well resistors are 10\(\mu m\) wide and use the same layout as in the bias current source. The LNA uses 4\(\mu m\) wide poly-2 resistors instead of N-well resistors like the rest of the circuit. This was partly due to the relatively low value of the resistors and partly because of the much lower stray capacitance of poly-2 when compared to N-well.

Given that RF signals are frequency converted, avoiding spurious coupling between inputs and outputs that could cause instabilities or oscillation is quite easy. The input of LNA and IF amp has been kept as far as practical from the local oscillator inputs to avoid signal pick-up, while the respective output connections have simply been kept as short as possible and well separated from the inputs.
Since the substrate is relatively high ohmic, an abundance of substrate contacts and guard rings has been placed to protect all circuits from spurious coupling and achieve as low a substrate contact resistance as possible.

The placement of the pads has been made not only considering parasitic coupling between pads or bonding wires, but ease of placement of the external components on a single-sided board and coupling between tracks were considered as well. The sensitive inputs have been placed near to at least one pad at AC ground, while the supply pads are central and adjacent to simplify decoupling. The large number of pads was required to give access to most bias points, including the ones that receive bias internally. The pads consist of standard metal-1 metal-2 pads of □110μm size, while no special trick (such as shielding) has been used to reduce their RF losses. The pin-out is shown in Fig. 3.25.

3.15 Chip Mounting and Testing

Given the application of this chip in a watch, it is quite clear that the only practical choice for mounting is chip-on-board (COB). Until chip-
scale packages will become available, all common packages will in fact very likely be too big and quite impractical in this application, even if QFP or TSSOP are considered. For our measurements the same mounting technique has therefore been used.

Small PCBs containing all the needed external components have been constructed for this purpose. For practical reasons (mainly external component placement and number of available connectors) two different PCBs have been constructed, one for testing of the RF parts, and the other for the IF and base-band parts. A drawing of these PCBs, including component placement, is shown in Fig. 3.26 and 3.27 respectively. These are 2 in x 1 in PCBs, with one signal plane and one ground plane on opposite sides, while the base material is 0.8mm thick standard FR4. The bare die has been directly bonded to it, and the assembled PCBs have been mounted in an RF test fixture with (maximum) six 50Ω SMA connectors on the four sides. To keep parasitic inductances and capacitances low, SMD passive components have been used, mainly of size 1206 (resistors) and 0805 (ceramic capacitors).

Unless otherwise noted, the chip has been tested under typical conditions. The supply voltage has been set to its nominal value of 3V and the various bias currents have been set to their nominal 12µA using ex-
Since the chip has been tested in a 50Ω system, matching to 50Ω was required. For best noise performance, the input of the LNA has been matched using an high pass matching network consisting of a 2.5pF capacitor and a 175nH hand wound inductor, which has been carefully adjusted for precise matching at 170MHz. A schematic diagram of this matching network and the resulting input impedance are both shown in Fig. 3.28. Its voltage gain is 17dB, thus overall gain and compression point will be scaled by that value. In our measurement data this has been already compensated. All other inputs have been resistively matched for simplicity. Resistive matching always increases noise, but this was unimportant as no noise measurements other than the one on the complete front-end have been made.

The local oscillator signal for the RF front-end was generated by the internal local oscillator. The external components used are $L_1=100\text{nH}$ and $C_2=2.2\text{pF}$ - which gives about the desired 3.3pF together with the

![Matching Network:](image)

**Figure 3.28:** Matching network and matched input impedance (for measurements only).
3.16. Measurements

1pF pad capacitance - while $C_1$ has been increased to 4.7pF to obtain the desired oscillation amplitude of about 500mV_{PP}. $C_0$ and $D_1$ have been substituted by a 3–10pF variable capacitor for tuning. The IF strip receives a 500mV_{PP} differential square wave local oscillator signal, similar to the one generated by an ECL divider, from an external generator. Its frequency was 21.4MHz for gain and linearity measurements and 10.7 to 12.7MHz for the measurement of the frequency response of the output filter.

Since at the time of this work the desired 21.4MHz IF filter was only deliverable in OEM quantities and no alternative part could be found, an already available 10.7MHz ceramic filter for FM radio applications was used instead. The open collector output of the RF mixer has therefore been loaded with a parallel LRC tank consisting of $L_2=6.8\mu H$, $R_3=5.6k\Omega$ and a 7–50pF trimmer adjusted for a center frequency of 10.7MHz. This load has a Q of about 12 (10 when loaded by the output buffer) for some prefiltering. A single stage buffer amplifier has then been used to match the 330Ω IF filter. The output of the IF filter - or the buffer when the filter was not used - has been matched to the measuring instruments with a wideband transformer. The local oscillator has also been adjusted to 180.7MHz to yield the proper IF.

3.16 Measurements

3.16.1 Voltage Reference and Current Source

Since many parameters of the receiver depend on the actual bias current, the bias circuit has been characterized first. The measured data for a temperature of approximately 25°C appears below.
The $\Delta V_{BE}$ current source is fully operational already at about 1V, and its operating current - 1.540\(\mu\)A at 3V supply voltage - is 28% higher than designed. This means that the N-well resistivity is about 4.5k\(\Omega/\square\), or 22% lower than the (already lowered) nominal 5.8k\(\Omega/\square\) (was 7.0k\(\Omega/\square\)).

The band-gap reference is in regulation from 1.4V at full load, and thanks to the closed loop design its output voltage is precisely 1.205V for VDD varying between 1.4V and 4.0V. Since the equipment for temperature testing was not available, no temperature coefficient has been measured. Loop stability has been verified for various capacitive loads (from open to several \(\mu F\)) and no instability or oscillation was noticed.

### 3.16.2 LNA Gain and AGC Range

Since the LNA uses resistors of different type than the on-chip bias source, it makes no sense to use the internal biasing. Bias for the LNA must therefore be provided externally. In the following table the gain versus the bias current of the first stage is shown. The current shown is the one supplied to the AGC pad. The current in the 1st stage is 4 times higher.
3.16. Measurements

<table>
<thead>
<tr>
<th>Current</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0μA</td>
<td>12.3dB</td>
</tr>
<tr>
<td>6.0μA</td>
<td>16.0dB</td>
</tr>
<tr>
<td>8.0μA</td>
<td>18.5dB</td>
</tr>
<tr>
<td>10.0μA</td>
<td>20.2dB</td>
</tr>
<tr>
<td>12.5μA</td>
<td>22.3dB</td>
</tr>
</tbody>
</table>

Useful AGC range (approx.)  : 20dB
1dB compression point       : −39dBu at Iagc=12.5μA
                          : −30dBu at Iagc=4.0μA

The gain shown is the voltage gain of the LNA referred to its input, not the matching network’s input. Its maximum value is 22.3dB, meaning that the poly-2 resistors are higher than nominal (but within tolerance). Scaling the bias current (to 10μA) to fit poly-2 resistor tolerance yields a gain of 20.2dB which is very close to the nominal value. Gain control is linear by principle, with a useful range of about 20dB. If the gain is reduced much below this value, linearity is first lost, then AGC inversion occurs because of the signal path over $R_1$. The input referred −1dB compression point is dominated by the mixer and is located at −39dBu at maximum gain, and improves to −30dBu if the LNA gain is reduced by 10dB to 12.3dB. Although these values are somewhat lower than expected (−37.1dBu and −27.1dBu respectively), our specification of −40dBu is already fulfilled. In a typical application the maximum LNA gain will be 20.2dB and the image filter will be used. The −1dB compression point of the complete receiver will therefore be about −33dBu, which is well above our spec.

3.16.3 LNA Input Impedance

The input impedance of the LNA, with the input stage biased to 48μA, is shown in the Smith chart below. Because of the high input impedance of the LNA, compared to the 50Ω of the network analyzer, this measurement is rather delicate and prone to error (magnitude). It has therefore
been repeated many times using two different instruments, and the curve shown is the one that was closest to the average of all measurements. This measurement was also used to determine the type and component values of the impedance matching network used for the other measurements. In this figure the values of $C_s$ and $L_p$, as calculated by the network analyzer, are displayed.

The equivalent circuit of the input of the LNA at 170MHz is 2.6kΩ and 2.5pF in parallel. This is lower than the calculated 4kΩ, possibly because of losses in the input pad and its ESD protection, but is fully adequate for this application and results in a practical antenna matching network.

3.16.4 Mixer Conversion $g_m$ and Compression Point

This measurement has been taken on the separated RF mixer directly at its output, with no filter and buffer, for better precision. The bias current was set externally to 12µA. The results were then used to determine the parameters of the LNA indirectly.

\[
\begin{align*}
\text{Conversion } g_m & : 260\mu\text{S} \\
1\text{dB compression point} & : -17.5\text{dBu}
\end{align*}
\]

The conversion $g_m$ is very close to the ideal value of 294µS. In case of a superhet receiver, because of the matching resistor for the IF filter, only half of that will be available. The larger $g_m$ of the single
balanced mixer will really help in that case, and is therefore recommended. The compression point, although somewhat lower than the expected $-15\text{dBu}$, is coherent with the one measured for the complete front-end (LNA) and sufficiently well within specifications.

### 3.16.5 RF Front-End Noise Figure

The noise figure has been measured on the whole front-end only. The 170MHz input is matched to 50$\Omega$, while only the parallel LCR prefilter ($Q=10$) has been used on the IF output, to avoid limiting signal bandwidth excessively (the noise figure meter has a bandwidth of 4MHz). The active output buffer has been used. Because of the way this measurement was done, it is referred to 50$\Omega$ and not to the antenna impedance, thus the final value must be calculated from this measurement. A noise voltage density measurement has been made to double check the noise figure measurement, and no significant difference has been noticed.

![170 MHz RF Front End](image)

<table>
<thead>
<tr>
<th>Freq.</th>
<th>SSB F</th>
</tr>
</thead>
<tbody>
<tr>
<td>10MHz</td>
<td>6.1dB</td>
</tr>
<tr>
<td>11MHz</td>
<td>6.2dB</td>
</tr>
<tr>
<td>12MHz</td>
<td>6.2dB</td>
</tr>
<tr>
<td>13MHz</td>
<td>6.25dB</td>
</tr>
<tr>
<td>14MHz</td>
<td>6.5dB</td>
</tr>
<tr>
<td>15MHz</td>
<td>6.8dB</td>
</tr>
</tbody>
</table>

Because of the large LNA gain, the input referred noise voltage of the front-end is dominated by the $3.5\text{nV}/\sqrt{Hz}$ of the LNA, which is increased to only $3.7\text{nV}/\sqrt{Hz}$ by the contribution of the mixer (if noise folding is not considered). The matching network transforms the 50$\Omega$ source impedance to 2.6k$\Omega$, whose voltage noise is $6.6\text{nV}/\sqrt{Hz}$. The noise figure is therefore 1.2dB, which becomes 5.2dB because of noise folding. The measurement is 6.2dB, which is very close to the expected value. This small deviation is absolutely normal, the most likely reasons for it being losses in the matching network (0.5dB or so), higher
than expected noise at the image because of mismatching, somewhat underestimated mixer noise and normal measurement error.

### 3.16.6 VCO Phase Noise

Since the VCO has no buffered output, its spectrum has been measured indirectly. The signal of the free running VCO, tuned at 180.7MHz has been mixed down to 10.7MHz using the internal mixer, whose input was connected to a 170MHz oscillator. Here a very clean, very stable PLL synthesized signal generator has been used, to avoid degrading the measurement. The following measurements show the phase noise of the local oscillator for various frequency offsets from the carrier.

![VCO Phase Noise vs. Freq. Offset](image.png)

<table>
<thead>
<tr>
<th>Offset (kHz)</th>
<th>Noise (dBc/Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5kHz</td>
<td>-51.1dBc/Hz</td>
</tr>
<tr>
<td>1.0kHz</td>
<td>-63.2dBc/Hz</td>
</tr>
<tr>
<td>2.0kHz</td>
<td>-74.9dBc/Hz</td>
</tr>
<tr>
<td>5.0kHz</td>
<td>-86.3dBc/Hz</td>
</tr>
<tr>
<td>25.0kHz</td>
<td>-97.9dBc/Hz</td>
</tr>
</tbody>
</table>

This curve shows two measurement anomalies. For low offsets the slope is about 12dB/octave in place of the expected 6 or 9dB/octave (the latter if flicker noise dominates). This may be due principally to normal frequency instability of the VCO (which is not regulated by a PLL) which tends to increase the measured values near to the carrier. Some phase noise of the local oscillator of the spectrum analyzer or leakage through its IF filter (whose bandwidth cannot be reduced below about 100Hz for the same stability problem) may also have contributed somewhat. The early flattening of the curve is due to white noise coming from the test setup. This limits the resolution of the measurement to -98dBc/Hz, that becomes apparent for frequency offsets larger than 5kHz, which is not sufficient to check the required limit of -105dBc/Hz at 25kHz offset. Nevertheless sufficient mid-band data is available to attempt a meaningful characterization of the oscillator. Extrapolating these measurements to 25kHz, using the 2kHz value and assuming that
3.16. Measurements

Flicker noise dominates, an approximate phase noise of $-108\text{dBc/Hz}$ is calculated. Despite this number is rather pessimistic, the ERMES requirement is already fulfilled. If extrapolation from the first 3 measured points is made, a value of $-118\text{dBc/Hz}$ is obtained. The true value will very likely lie between these two limits. The power spectrum of the VCO is shown below.

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{vco_spectrum.png}
\caption{Power spectrum of the VCO.}
\end{figure}

3.16.7 IF Amplifier Gain and AGC Range

The measurement shown below is the voltage gain of the IF amplifier versus the bias current of the input stage. For this measurement the output stage of the IF amplifier is biased with the internal current source, while the input stage is externally biased with a variable resistor to allow for gain variation. Since the IF amp uses resistors of the same N-well material as the current source, gain independence can be achieved by biasing also the input stage with the internal source, using the output provided for that purpose. To avoid making modifications to the circuit, however, that output has not been used directly and the external bias current has been simply set by hand to make it track precisely the value of the latter (10 times higher). The input has been matched to 50$\Omega$ with an external resistor.
For nominal bias current (12\(\mu\)A) the gain is lower than nominal because of the lower value of the N-well resistors. Nevertheless when the IF amp is biased with the same value as the internal current source would do (15.4\(\mu\)A) the nominal 26dB gain is precisely achieved, demonstrating that with proper design high gain precision independent from technology parameters can be achieved also by an amplifier without global negative feedback. Again, gain control is linear by principle and its useful range is 20–25dB if good linearity is desired. Since the only direct path to the 2nd stage is the (small) Miller capacitance of \(Q_1\), the gain can be reduced much more without any AGC inversion, but linearity is worsened. As it was the case for the LNA, also here linearity is dominated by the mixer (demodulator), and has therefore not been measured.

### 3.16.8 Demodulator Gain and Linearity

The measurement shown below has been performed on the isolated demodulator mixer, and has been used to determine the performance of the IF amp indirectly. Similarly to the IF amp, the demodulator has been measured at nominal current (actually 12.5\(\mu\)A in place of 12\(\mu\)A) and at 10 times the current of the bias current source.

<table>
<thead>
<tr>
<th>Current</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0(\mu)A</td>
<td>14.4dB</td>
</tr>
<tr>
<td>6.0(\mu)A</td>
<td>17.9dB</td>
</tr>
<tr>
<td>8.0(\mu)A</td>
<td>20.5dB</td>
</tr>
<tr>
<td>10.0(\mu)A</td>
<td>22.3dB</td>
</tr>
<tr>
<td>12.5(\mu)A</td>
<td>24.3dB</td>
</tr>
<tr>
<td>15.4(\mu)A</td>
<td>26.1dB</td>
</tr>
</tbody>
</table>

As it was the case for the IF amp, also here good gain accuracy is achieved when the demodulator is biased at 10 times the current...
generated by the internal current source. Again, the lower value of the N-well resistors reduces this gain to 16.7dB at nominal bias current. The compression point is as expected and is determined by the voltage swing limit at the output of the transimpedance active load.

3.16.9 Output Filter Frequency Response

For this measurement a 10.7MHz signal was supplied to the resistively matched IF input of the demodulator, while the local oscillator signal was swept from 10.7MHz to 12.7MHz. The differential output of the filter was recorded and scaled relatively to the 1kHz value. This measurement and a plot of the frequency response are shown below.

<table>
<thead>
<tr>
<th>f (Hz)</th>
<th>G</th>
<th>f (Hz)</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0k</td>
<td>0.0dB</td>
<td>500k</td>
<td>−7.43dB</td>
</tr>
<tr>
<td>10k</td>
<td>0.0dB</td>
<td>600k</td>
<td>−9.12dB</td>
</tr>
<tr>
<td>100k</td>
<td>−0.44dB</td>
<td>700k</td>
<td>−10.5dB</td>
</tr>
<tr>
<td>150k</td>
<td>−1.16dB</td>
<td>800k</td>
<td>−12.0dB</td>
</tr>
<tr>
<td>200k</td>
<td>−1.94dB</td>
<td>900k</td>
<td>−13.0dB</td>
</tr>
<tr>
<td>250k</td>
<td>−2.79dB</td>
<td>1.0M</td>
<td>−14.0dB</td>
</tr>
<tr>
<td>300k</td>
<td>−3.74dB</td>
<td>1.2M</td>
<td>−16.5dB</td>
</tr>
<tr>
<td>350k</td>
<td>−4.81dB</td>
<td>1.4M</td>
<td>−18.1dB</td>
</tr>
<tr>
<td>400k</td>
<td>−5.60dB</td>
<td>1.6M</td>
<td>−20.0dB</td>
</tr>
<tr>
<td>450k</td>
<td>−6.47dB</td>
<td>2.0M</td>
<td>−22.5dB</td>
</tr>
</tbody>
</table>
Although the frequency response of the output filter depends on the absolute value of RC components, little deviation is revealed by the measurements. The attenuation at 250kHz is 2.8dB and the 1st pole is slightly higher than designed because of the lower than nominal N-well resistors, possibly combined with higher poly1-poly2 capacitors (which reduces the error). In production much bigger deviations must be expected, since the fortunate combination of low resistors with high capacitors is certainly not reproducible. Nevertheless this deviation can be allowed (e.g. if the base-band AD converters have a sufficiently high sampling frequency to avoid aliasing), while the Q of the filter will still be well controlled and process independent, being it only determined by component matching.

A summary of the main characteristics of this chip is shown in table 3.8.

<table>
<thead>
<tr>
<th>Process</th>
<th>1.2µm BiCMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chip size</td>
<td>2240 x 980µm²</td>
</tr>
<tr>
<td>Power supply voltage</td>
<td>3V (typical)</td>
</tr>
<tr>
<td>Front-end supply current</td>
<td>230µA</td>
</tr>
<tr>
<td>LNA gain (12.5µA)</td>
<td>22.3dB</td>
</tr>
<tr>
<td>RF mixer conversion $g_m$</td>
<td>260µS</td>
</tr>
<tr>
<td>Front-end noise figure (SSB)</td>
<td>6.2dB</td>
</tr>
<tr>
<td>VCO phase noise (2kHz offset)</td>
<td>$-74.9$dBc/Hz</td>
</tr>
<tr>
<td>IF amplifier gain</td>
<td>26.1dB</td>
</tr>
<tr>
<td>IF mixer conversion gain</td>
<td>18.5dB</td>
</tr>
<tr>
<td>Output filter bandwidth</td>
<td>250kHz</td>
</tr>
</tbody>
</table>

Table 3.8: ERMES chip characteristics.
Chapter 4

64/65 ESCL Dual Modulus Prescaler

Among the circuits presented here, the frequency synthesizer and especially the prescaler are the last to be critical to the performance of a low power portable receiver like ours. Operating at a frequency orders of magnitude higher than the rest of the digital part, it not only has to be efficient in terms of current consumption, but is also required to operate reliably with the limited voltage supply of a "low on power" battery.

The prescaler - which is a simple frequency divider - is a kind of bridge between the RF analogue chip which contains the actual receiver, and the digital chip which together with signal processing and control functions will contain the main part of the PLL frequency synthesizer. The connection between the local oscillator and the prescaler input is somewhat delicate for reasons that will be explained later, while output and modulus switching are simple low frequency digital connections that require no particular precautions.

As we have previously seen, to achieve a good performance especially in terms of power consumption, the higher transconductance and lower stray capacitances of bipolar transistors were required. Therefore, the RF part is designed in 1.2μm BiCMOS. CMOS technologies with suffi-
ciently short gate length to be competitive with the chosen technology were simply not available at the time of this work. The digital part on the other hand will be a rather large chip working at a relatively low frequency. Therefore, a low cost CMOS technology is the most appropriate solution. Since the requirements for the RF and digital parts of the pager receiver are very different, it is likely that a two chip solution is the most successful in terms of performance, power consumption and cost, at least when using the technologies that were available to us.

Regarding the PLL, there are basically two ways to partition it between the two chips. While the VCO will certainly reside on the RF chip and the programmable divider on the digital chip for obvious reasons, the prescaler can be placed on either of the two chips. Each solution results in differing performances and poses particular problems.

One possibility is for the prescaler to be placed on the RF chip. In this case, it will enjoy the high speed offered by the bipolar transistors and the critical connection to the VCO will stay on-chip. While the interface to the digital chip will consist of the low frequency connections as said previously, the result may be interference between it and the sensitive RF receiver. In the second case the prescaler can reside on the digital chip. Interference with the RF receiver will be avoided in this case, but a rather high power consumption must be anticipated, while the VCO signal must be buffered and routed off-chip, resulting in a further increase of power consumption and likelihood of reradiation. Placing the prescaler on the RF chip is therefore certainly preferable and will result in the best performance. However, this option was not available to us at the time of this work, so a digital CMOS compatible prescaler was designed. Because of the chosen technology, a rather high power consumption, compared to the rest of the receiver, has to be anticipated.

On the other hand, for several years the industry has tried to adopt CMOS for the RF analogue part of communications applications, mainly for cost reasons. Provided good performance can be achieved, a full CMOS prescaler would be an interesting future-oriented building block. The same CMOS technology could then be used for both the RF and the digital parts, potentially leading to a reliable, very low cost single-chip solution.

This design is from early work dating back to around December 1992.
4.1. The Frequency Synthesizer

In this application a local oscillator frequency from 190.825MHz to 191.200MHz in 25kHz steps (i.e. the channel spacing) will be generated for the superhet receiver. In communications applications, the most common frequency synthesizer architecture uses the pulse swallowing technique, as shown in the block diagram of Fig. 4.1. Pulse swallowing combines a fast but simple frequency divider (the dual modulus prescaler) with two fully programmable but slow counters (the swallowing and the reference counters) to obtain any integer division ratio larger than or equal to \( P(P - 1) \), where \( P \) is the division ratio of the prescaler.

In the first phase the VCO signal is divided by \( P + 1 \). This division is carried out \( N_S \) times, i.e. until the swallowing counter reaches 0. Afterward the swallowing counter is inhibited and the prescaler is

![Block diagram of a pulse swallowing frequency synthesizer](image-url)
switched to divide by \( P \), until the remaining counts reach \( N_R \) of the reference counter. At the end of the cycle, both counters are reset to their respective \( N_S \) and \( N_R \) and a new cycle is started. The output of the divider is this reset signal, which is sent to the frequency and phase comparator. The output frequency is therefore

\[
f_{out} = \frac{f_{VCO}}{N_S(P + 1) + (N_R - N_S)P} = \frac{f_{VCO}}{N_S + N_R P}
\]

In order for the system to work, the division ratio of the reference counter must be larger or equal to that of the swallowing counter. For simplicity the latter is usually chosen to be between 1 and \( P \) (programmed value = 0 to \( P - 1 \)). Provided that \( P \) is a power of two (it is in most cases), this can be concatenated with \( N_R \) (as the LSB's) to form a single binary number. In this way, programming the frequency synthesizer is transparent and does not require any knowledge about how the divider is implemented.

At this point, the frequency synthesizer is completely defined. To generate the VCO frequencies required for the superhet receiver, a division ratio programmable between 7633 and 7648 is required. This means that \( P = 64 : 65 \) - which allows division ratios above 4032 - and \( N_R = 119 \), while \( N_S \) should be programmable between 17 and 32. In our case therefore, things are particularly simple, as a fixed divider is required for the reference counter while channel selection is achieved by writing a fixed 1 for the MSB followed by the 4 bit hex numbers 0x0 to 0xF for the rest in the 5 bit swallowing counter.

A similar calculation for the direct conversion receiver also yields \( P = 64 : 65 \), with \( N_R = 105 \) and \( N_S = 57 \) to 72, or \( N_R = 106 \) and \( N_S = 1 \) to 8 for the eight higher channels if desired.

## 4.2 Prescaler Architecture

The main considerations that drive the design of the prescaler for a communications application such as ours are low power consumption, operation on a wide range of supply voltages (battery) and particularly the interface with the local oscillator (taking into account the possibility of interference). Therefore, despite being just a simple digital divider,
4.2. Prescaler Architecture

the design of a prescaler is no straightforward task. In this work we have tried to design a prescaler that will cover a much larger frequency range than that required just by ERMES - thereby enabling it to suit many other applications.

As we have seen in the previous section, the prescaler receives the signal from the local oscillator as its input. As in virtually every communications application, this is a varicap-tuned LC oscillator. This is dictated by phase noise requirements, which preclude full swing first order VCOs (as e.g. a ring oscillator) for the local oscillator. Unlike typical digital circuits which receive rail-to-rail digital inputs, the input to the prescaler is therefore a low amplitude (200–500mV_{pp}) sinusoidal signal. For our receiver the signal at the pin LO-e, i.e. the emitter of transistor Q_{7} in Fig. 3.15, would be fed to the prescaler, possibly via a unity gain buffer. Here the nominal amplitude is 250mV_{pp}.

The requirements for our prescaler can therefore be summarized as follows:

- Dual modulus prescaler, division ratio 64/65.
- High input sensitivity, 200–500mV_{pp} from an LC local oscillator.
- Low power, independent from VDD variations (3V battery supply).
- Low noise, sensitive local oscillator (and RF receiver) connected to it.
- Low cost technology, same as RF receiver or digital decoder chip.

The most established prescaler technique that fulfills these requirements is bipolar ECL. ECL circuits already accept low amplitude logic signals, consume low power, are independent from the power supply and generate very low noise if properly designed. Such a prescaler could be integrated with the RF receiver chip in the same 1.2μm BiCMOS technology, would in all probability fulfill the 150μA estimate of Fig. 3.6 and 3.7 quite easily and would, therefore, be the way to proceed in this case. As previously said, this option was not available and an alternative had to be found.
The straightforward method of designing a CMOS prescaler with the required characteristics is to boost the clock signal to rail-to-rail levels in order to drive standard static or dynamic CMOS flip-flops, as shown in Fig. 4.2. Using the technique described in [40] a prescaler with good performance and very low power consumption can be designed. However, several drawbacks strongly discourage such a solution.

First of all, the clock amplifier, requiring at least about 20dB of gain, will tend to consume too much power even at a couple of hundred megahertz if designed in the chosen 1.2µm CMOS technology. Since the speed of dynamic logic is highly dependent on the supply voltage, this technique may not be adequate for battery operation where a large supply voltage variation has to be expected. The transistors must be sized for the minimum supply voltage, and since the power consumption of a CMOS circuit is proportional to \( V_{DD}^2 \), excessive power consumption may result at high \( V_{DD} \). A further problem arises since the unbalanced rail-to-rail logic signals may generate too much noise, perhaps to the point of preventing the integration of such a prescaler on the same chip with sensitive analogue circuitry.

An alternative approach is to use flip-flops that run directly with the low amplitude VCO signal. Several well-known current mode logic families described in recent literature, such as Folded Source Coupled Logic (FSCL [41]) and Enhancement Source Coupled Logic (ESCL [42]) have this characteristic. A basic schematic diagram of both FSCL and ESCL appears in Fig. 4.3. Source coupled logic is not a new concept, having already been used in the early seventies e.g. as a read amplifier in RAM memory ICs.

Based on differential stages biased with constant current, source coupled logic uses low amplitude signals (which lead to low power consumption) and is insensitive to power supply variations. Thanks to the fully differential structure, source coupled logic generates very lit-
4.2. Prescaler Architecture

Figure 4.3: Basic schematic diagram of (a) FSCL and (b) ESCL logic families.

ttle noise and has a high noise immunity, making it more suitable than
standard CMOS logic for low power battery operated radio applica-
tions. More complex logic functions than the inverter/buffer shown in
Fig. 4.3 can be constructed by implementing the logic function in the
input stage (shaded in Fig. 4.3) in the same way as complex ECL gates,
sometimes resulting in fewer transistors than its standard CMOS coun-
terpart. ESCL has been preferred over FSCL in this work, since its
power consumption is lower (roughly halved) for the same speed and
minimum supply voltage.

The most obvious drawback of these logic families is static power
consumption. Static power (intrinsic to any current mode logic) may
discourage its use in a typical logic circuit but is of no consequence
here, since the flip-flops always toggle at full speed. Other less obvious
problems concern the size factor (much larger if compared to standard
CMOS but irrelevant in our case, since total size is clearly dominated by
the rest of the digital circuit) and the need for a biasing current source.
In most digital CMOS technologies - and ours was no exception - the
analogue components are either specified only as a maximum value (e.g.
resistors and capacitors), ill specified (MOS transistors are specified at
high gate overdrive and have no weak inversion) or are not specified at
all (parasitic PNP bipolar transistors). This may imply some risk while
some guesswork is certainly needed.
4.2.1 The ESCL Technique

Enhancement source coupled logic was initially proposed [42] for the design of complex mixed analogue-digital circuits due to its good noise performance. It is similar to bipolar ECL but uses enhancement NMOS transistors. The principle of the ESCL technique is illustrated in Fig. 4.4a. It is a differential stage biased with constant current and loaded with 2 resistors. The operation of this gate is based on steering the constant current $I_0$ through one of the two load resistors by applying a differential voltage at the input.

The choice of the amplitude of the logic signal is a balance between speed and power consumption (which improve at low logic swings) and the need to guarantee complete switching of the input stage, which requires a logic swing of at least the $V_{gs} - V_t$ of the input transistors.

Speed is mainly determined by the RC time constant formed by the load resistors and the capacitance of the input transistors (both gate and source/drain). To optimize speed, one must consider that for a given gate length and bias current:

- Transistor capacitance is directly proportional to gate width, unless the side-wall capacitance dominates (when the transistor is small) and capacitance becomes about constant.

- The overdrive voltage $V_{gs} - V_t$ is inversely proportional to the square root of gate width, unless narrow channel effects take over and the dependence of the overdrive on gate width becomes almost linear.
4.2. Prescaler Architecture

It is quite clear that although small transistors yield higher speed, an optimum size exists and must be found by considering the aforementioned problems. At lower frequencies, the minimum design rules will very likely prevent the optimum from being reached. In that case, a sub-optimum performance must be accepted.

Although the highest speed is achieved when using the minimum logic swing which guarantees complete switching of the input pair, in general a significantly larger swing is required, partly to guarantee a fast switching and partly to compensate the swing reduction caused by feed-through via the stray capacitances of the transistors (mainly Miller capacitance).

The effect of signal feed-through can be seen very well in traces $V(3)$ to $V(5)$ of the prescaler simulation in Fig. 4.12. Usually a logic swing of the order of $0.5 - 1V_{pp}$ is chosen, which is obtained by sizing the load resistors and the current source accordingly.

Further, an important requirement of ESCL is the minimum gain of the ESCL stage. Because of noise margin considerations, the gain of each stage should be maintained $> 1$ (typically about 1.5–2) even when transistors are connected in series (e.g. AND/OR function). In extreme cases, the logic signal would be lost completely, for example if many stages which do not respect the gain condition are cascaded.

The load resistors may pose some design problems. Indeed, as has been previously stated, in most digital technologies, resistors are specified as maximum values, if specified at all. The materials typically available for resistors are the poly, the diffusions (around 30–100$\Omega/\square$) or the N-well (usually unspecified, typical values around 1–3k$\Omega/\square$). In all cases, large size can pose a further problem as it can be associated with large stray capacitance. To save area, it is therefore customary to replace the load resistors with N-channel transistors. In old three supply-IC’s, these were the normal enhancement loads, with the drain connected to the 5V and the gate to the 12V supply, while in newer technologies, diode-connected N-MOS transistors as $M_3$ and $M_4$ of Fig. 4.3b were used. In order to fulfill the gain requirement, the channel widths of the latter should be narrower than that of the input transistors.

This approach is not applicable for low voltage applications such as this work in which the minimum supply voltage is 2V. Since the
threshold voltages of the differential pair and the load are added, the minimum operating voltage would be in the range 3–3.5V (worst case), which is definitely outside the desired value. To circumvent this problem, "real" resistors should be used as load, to reduce the voltage drop from \(\approx 1.5V\) to just 0.2–0.5V. Since real resistors are too large, PMOS transistors operating in the triode region are used instead [43], as shown in Fig. 4.4b. Apart from the size or voltage drop advantage, the loads are now voltage controllable, a definite advantage which is required in our case since we have decided to target a wide range of speeds and bias currents.

In order to fit a wide range of bias currents, the logic swing must be adapted to the actual operating conditions and should by no means be maintained at a constant. Since the \(V_{gs}-V_t\) of a transistor is proportional to the square root of its drain current, the logic swing must be increased by the same factor to guarantee that complete switching occurs also at higher currents. In the ESCL gate shown in Fig. 4.3 this occurs automatically.

An immediate solution could be to use a servo amplifier as e.g. in [43], driven by a square root generator, as shown in Fig. 4.5. In this circuit, transistor \(M_6\) is identical to \(M_1\) and is biased with the same current. Therefore, the \(V_{gs}-V_t\) will be the same. Being a large transistor, \(M_7\) will be biased at about \(V_t\), so that the square root term \(V_{gs}-V_t\) will appear on \(R_1\), or - amplified - on \(R_2\) where it represents the reference for the servo loop. The logic swing is therefore equal to \(R_2\)
4.2. Prescaler Architecture

voltage, or $R_2(V_{gs} - V_t)/R_1$ as desired. This solution, although guaranteeing some matching between the reference transistor ($M_6$) and the input stage ($M_1$ and $M_2$), will very likely occupy as much area as the prescaler itself, requires a compensation capacitor ($C_C$, unspecified in most digital technologies) and offers no particular advantages. It has therefore been discarded.

A much better solution from the point of view of size and complexity is shown in Fig. 4.4b, where the sizes of all transistors are also given. The resistance of the load transistors $M_3$ and $M_4$ in the triode region is

$$R_l = \frac{1}{K'P \frac{W}{L}(V_{gs} - V_t)}$$

i.e. inversely proportional to the overdrive voltage. The gate bias voltage of these transistors is derived from a current source which in turn is mirrored from the bias current of the differential input transistors. The overdrive voltage of $M_5$,

$$V_{gs} - V_t = \sqrt{\frac{2}{k'P} \cdot \frac{L}{W} I_d}$$

is proportional to the square root of the bias current. Therefore, the resistance is inversely proportional to that resulting in a logic swing which is directly proportional to the square root of bias current, as required. This also enables the gain of the ESCL structure to be quite insensitive to power supply and bias current variations (which can be either intentional (to increase the speed) or due to process or temperature change). Proper sizing of the load bias transistor ($M_8$) with respect to the gate bias transistor ($M_6$) also ensures that the load always stays in the triode region.

A possible problem arising from this is that the logic swing is defined by P-channel transistors while the required swing is imposed by N-channel transistors which obviously do not match. In all truth, this is a minor problem, partly because sensitivity to parameter variations is low ($\sqrt{K'P}$), partly because limits such as slow-N/fast-P do not exist in practice since the gate oxide is the same for both devices and partly because the process tolerance is swamped by the need to design a swing which is greater than the minimum. Nevertheless, the prescaler has been designed by taking these tolerances into account, while to be
safe, its performance at the process, voltage and temperature corners has been verified extensively (with Spice simulations) before submitting the design for production. Measurements made on the finished prescaler have proven this circuit's proper functioning and appropriate performance. In fact, the prescaler works correctly for bias currents in a range of more than 10/1, without any external intervention to modify the logic swing (it would nevertheless be impossible to modify it, since all nodes of the biasing circuit for the load transistors are internal and as a result, unaccessible).

4.3 Circuit Description

4.3.1 Prescaler

The architecture chosen for this design is the classical dual modulus prescaler, a logic diagram of which is shown in Fig. 4.6. It is a dual modulus prescaler dividing by 64/65 as specified before, and consists of a 4/5 synchronous divider (Johnson counter) driving an asynchronous chain of 4 toggle (divide-by-two) flip-flops. The high frequency stages have been implemented with ESCL logic, while standard CMOS logic is used at low frequency mainly to save power. Although in Fig. 4.6 unbalanced lines are shown everywhere, all ESCL signals are obviously balanced.

Figure 4.6: Logic diagram of the ESCL prescaler.
4.3. Circuit Description

Two ESCL flip-flop types are used in the first divider stage. The flip-flops (1) and (3) in Fig. 4.6 include a built-in NAND and OR gate respectively, while flip-flop (2) is a simple D-flip-flop. These flip-flops are driven directly by the VCO signal, without any clock buffer (\(ck\) input).

The toggle flip-flop (4) is also implemented with ESCL logic despite running at only 1/4 of the input frequency. This is done for three reasons: the first is to maintain low switching noise, the 2nd is to ensure that the delay introduced by an ESCL-to-CMOS level converter in front of this flip-flop does not limit the maximum speed of the prescaler and the 3rd is to postpone the level converter in order to alleviate its gain-bandwidth requirement and power consumption. Conversion from ESCL to CMOS levels is performed after this flip-flop with a differential amplifier. At this point, the frequency is about 24MHz. Therefore, an amplifier with a gain bandwidth product of only \(\approx 240\text{MHz}\) is required. Such a moderate requirement may be obtained from almost any amplifier topology at low power consumption.

The following 3 stages are implemented with standard CMOS static flip-flops, connected in toggle configuration. Since the frequency here is low, small transistors with minimum size or just above the design rule can be used to reduce power and switching noise without compromising the maximum toggle frequency of the prescaler. The output of the prescaler is then buffered with a pad driver which comprises two inverters and ESD protection (not shown in the schematic diagram). The NAND gate for modulus switching is also implemented with standard CMOS logic.

4.3.2 ESCL logic

Firstly, the design of the ESCL part of the prescaler requires that we define logic swing, bias current and transistor dimensions. The requirements here are to guarantee a sufficient sensitivity (about 250mV\(_{pp}\)) and proper operation at least at 190MHz, for all worst case conditions combined.

The calculation of the required parameters is an iterative process. Since speed - thus the RC time constant - is known, we must first calculate a tail current, size the transistors for the desired overdrive
and verify that the speed requirement is fulfilled. If necessary, the tail current must be modified and the calculation repeated. To achieve a good compromise between our various requirements, a differential logic swing of 500mV$_{PP}$ and a bias current of 30µA per differential pair was found. The transistors were then sized for an overdrive voltage of about 250mV, required to achieve a good sensitivity with sufficient noise margin, and resulting in W=9.6µm. Under these conditions, a worst case speed of 200MHz is achieved, while under typical conditions, it is about 300MHz. These values are somewhat less than the optimum because of the sensitivity specification.

The parts designed using ESCL are marked in Fig. 4.6 and consist of the fast flip-flops and logic gates.

A schematic diagram of the ESCL flip-flop with an input gate (1 and 3) is shown in Fig. 4.7. It consists of two cascaded level-sensitive static latches, with the input gate (C=A∩B) embedded in the first latch. Indeed, the motivation for building logic gates into the flip-flops is to save power. The nominal current consumption is in fact just 60µA for both flip-flop types. Here, only one flip-flop type is required to implement both 1 and 3, since in swapping the two outputs, the NAND function may be obtained, while the (N)OR function is achieved if true
and complement are swapped at the inputs.

Due to its static design, no minimum operating frequency exists, but a minimum rise and fall time limit still remains. If a sinusoidal input is used (as expected), this limit corresponds to a minimum frequency requirement, dependent on the amplitude of the input signal.

The ESCL circuits are completed by an OR/NOR gate. This gate is identical to the input stage of the flip-flop of Fig. 4.7. A separate gate was needed here, since implementing it into flip-flop 3 would have loaded its master latch with excessive stray capacitance and reduced the gain below unity. One of its inputs is driven directly by the modulus switching NAND, i.e. with a CMOS signal. This is possible since ESCL accepts CMOS levels without the need for level conversion.

### 4.3.3 ESCL to CMOS level converter

ESCL to CMOS level conversion is performed with a differential amplifier, the schematic diagram of which appears in Fig. 4.8. A folded cascode transconductance amplifier has been used for this function to achieve wide bandwidth, near rail-to-rail output swing and low voltage operation. The amplifier has enough gain at 24MHz to drive its output into clipping (see simulation, trace V(6)). As in any transconductance amplifier, capacitive loading reduces amplifier bandwidth. Therefore
the output is buffered with a small CMOS inverter.

4.3.4 CMOS logic

These are standard static CMOS functions. The D flip-flop is the common implementation with inverters and transmission gates drawn in Fig. 4.9, while the NAND gate driving the ESCL input is followed by an inverter to also generate the necessary complement. The N-channel transistors have a width of 2.4\(\mu\)m, which is just above the minimum design rule, since this was the layout with the minimum drain-substrate capacitance. The P-channel transistors have twice this width to obtain symmetrical waveforms.

4.3.5 \(\Delta V_{BE}\) current source

The bias current source is an important part of an ESCL circuit since speed and power consumption of the complete circuit are affected by it. Sufficient precision is required as is supply-independence as the battery supply is expected to diminish substantially during its lifetime. In our case, the nominal value is 30\(\mu\)A. This is implemented with a \(\Delta V_{BE}\) current source as shown in Fig. 4.10, making good use of the lateral parasitic bipolar transistors, a characteristic of every CMOS technology [44]. A schematic layout of one such transistor is shown in Fig. 4.11.

The \(\Delta V_{BE}\) cell is a modified Brokaw cell that does not use the emitter current as a reference. The latter is in fact ill controlled due to the action of the vertical PNP associated to the lateral one. The emitter
4.3. Circuit Description

Figure 4.10: Bipolar current source with lateral parasitic PNPs.

area ratio has been set to 8:1, while resistor $U_8$ establishes a current of $6\mu A$ per arm. The latter is then increased to the desired $30\mu A$ by the 5:1 current mirror $U_6$ and $U_7$. The resulting bias current has a PTAT characteristic, which helps to compensate for the $g_m$ reduction of the MOS transistors with temperature increases, thus maintaining stage gain and speed. Start-up of this current source is guaranteed by leakage currents, while its output is connected to a pad to allow for measurements, adjustments or to disable the prescaler when in stand-by mode.

A common problem when designing analogue circuits like this current source in a digital CMOS technology is that both the resistors and the bipolar transistors are not characterized. The technology we used does not represent an exception to the rule. In our case, only a maximum resistance value for gate poly and S/D diffusions was known, while for the BJTs no data existed.

Since the resistance per square was highest for P-diffusion resistors, the latter was chosen for this design. Their typical value had to be estimated from the given maximum value, and 25% less, or $95\Omega/\square$ was assumed. The biasing circuit which was designed for $30\mu A$ nominally, generates $40\mu A$, indicating that the guessed value for the P-diffusion resistor was somewhat excessive.

For lateral bipolar transistors, things are slightly more complicated since no data or design rule existed. The parameters to be defined are
emitter size and especially base (poly) width. This choice is somewhat critical: if the base width is too small, the output conductance of the transistor will increase, thereby compromising precision and line regulation of the current source. If an excessively large width is chosen, the $\beta$ is reduced and the parasitic current through the vertical PNP will dominate. Based on the experience of others, an octagonal emitter of 3.6\(\mu\)m per side and a base width of 2.4\(\mu\)m was chosen. Measurements later demonstrated that this was the correct choice, as the $\beta$ is 40 for both the lateral and vertical transistor, while line regulation is good, as the measurements have shown.

4.4 Simulations and Measurements

The simulation of the prescaler appears in Fig. 4.12. This simulation refers to the prescaler core only. The output pad and its driver have not been included. Indeed, in a typical application, where the prescaler is embedded in a complete frequency synthesis chip, the output would be connected internally to the programmable counters and an output pad would not be required. The conditions are the following:
4.4. Simulations and Measurements

Division ratio : 65
Input signal : 250MHz, 400mV<sub>pp</sub>
Power supply : 3V, common mode voltage: 2.8V
Bias current : 30μA
Model : Typical

The signals appearing in the simulation are:

V(2,1) : input signal.
V(3) : Q output of the 3<sup>rd</sup> ESCL flip-flop.
V(4) : The output signal of the 4/5 divider.
V(5) : Input to the level shifter.
V(6) : Level shifter output (drain of U9).
V(7) : Qcmos output of the level shifter.
V(8) : Output of the 1<sup>st</sup> CMOS divider by two.
V(9) : Prescaler output.
V(10) : Output of the ESCL NOR gate.
V(11) : Output of the CMOS NAND gate.
i(vdd) : Power supply current.

Figure 4.12: Spice simulation of the prescaler.
The nodes where these signals appear are also shown in Fig. 4.6, with the exception of $V(6)$ (which is an internal signal). The input signal is differential, node 2 being the true (non inverting) input. All other signals have been taken single ended, referenced to ground.

A photomicrograph of this chip is shown in Fig. 4.13. The chip has been completely laid out by hand and no standard cells have been used. It has 8 pads and its dimensions are 677µm x 630µm (including the pads). The ESCL prescaler is the block at the centre, and its size is 377µm x 272µm. The bias current source is the block on the right of the chip. Its area, 196µm x 265µm, is quite substantial equalling almost half that of the prescaler.

The measurements have been made on an unpackaged chip bonded to a small printed circuit board containing all the required external components. The input is resistively terminated to 50Ω and receives a single-ended signal with an amplitude of 0dBu. The common mode voltage has been set to 2.5V (unless otherwise noted). The power supply is the nominal 3V, unless otherwise noted, and is decoupled to ground with a 100nF ceramic capacitor in parallel with a 4.7µF tantalum capacitor. A small resistor has been introduced to allow for power supply noise measurements. The Bias pin has been decoupled to ground with 10nF, while external resistors have been used to inject or extract some
current to change the bias point when required.

4.4.1 Bias Current Source

The first part of the measurements has been made with the prescaler biased with its internal current source. The latter has therefore been characterized first. In the following table the output current for a temperature of about 25°C is shown.

<table>
<thead>
<tr>
<th>VDD (V)</th>
<th>IDD (µA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5V</td>
<td>35.7µA</td>
</tr>
<tr>
<td>2.0V</td>
<td>38.0µA</td>
</tr>
<tr>
<td>2.5V</td>
<td>39.1µA</td>
</tr>
<tr>
<td>3.0V</td>
<td>40.0µA</td>
</tr>
<tr>
<td>3.5V</td>
<td>40.7µA</td>
</tr>
<tr>
<td>4.0V</td>
<td>41.3µA</td>
</tr>
<tr>
<td>4.5V</td>
<td>41.9µA</td>
</tr>
<tr>
<td>5.0V</td>
<td>42.3µA</td>
</tr>
<tr>
<td>5.5V</td>
<td>42.8µA</td>
</tr>
<tr>
<td>6.0V</td>
<td>43.2µA</td>
</tr>
</tbody>
</table>

At the nominal supply voltage of 3V the current generated is 40µA, or about 30% higher than nominal. As stated previously, this means that the estimated value for the P-diffusion resistivity was still too high. A better than nominal maximum operating frequency can therefore be expected. Line regulation is quite good for such a simple design.

4.4.2 Maximum Operating Frequency

The following table shows the maximum operating frequency versus division ratio of the prescaler when biased with its internal current source.

<table>
<thead>
<tr>
<th>Division Ratio</th>
<th>Max. Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>393MHz</td>
</tr>
<tr>
<td>65</td>
<td>392MHz</td>
</tr>
</tbody>
</table>
Thanks to the higher than nominal bias current, the operating frequency is also higher than nominal by a factor of about $4/3$. Nevertheless the linear relationship between speed and bias current cannot be expected at speeds that differ greatly from the nominal.

### 4.4.3 Power Consumption and Minimum Operating Supply Voltage

The following table shows the current consumption at VDD=3V and the minimum operating voltage vs. frequency with division ratio=65. The current consumption measurement includes the current drawn by the output pad driver, driving about 1pF of external load.

<table>
<thead>
<tr>
<th>Freq.</th>
<th>IDD</th>
<th>Vmin</th>
</tr>
</thead>
<tbody>
<tr>
<td>65MHz</td>
<td>565µA</td>
<td>1.35V</td>
</tr>
<tr>
<td>130MHz</td>
<td>588µA</td>
<td>1.45V</td>
</tr>
<tr>
<td>195MHz</td>
<td>611µA</td>
<td>1.60V</td>
</tr>
<tr>
<td>260MHz</td>
<td>634µA</td>
<td>1.85V</td>
</tr>
<tr>
<td>325MHz</td>
<td>657µA</td>
<td>2.10V</td>
</tr>
<tr>
<td>390MHz</td>
<td>681µA</td>
<td>2.95V</td>
</tr>
</tbody>
</table>

The maximum operating frequency is 392MHz, and the current consumption is 681µA, or 1.75µA/MHz, a good value for a 1.2µm CMOS technology. At 390MHz, the CMOS part of the prescaler, including the pad driver, consumes about 1/3 of the total, i.e. 211µA. This is also shown in the next measurement.

### 4.4.4 Power Supply Noise

The following figure shows the supply noise generated by the chip. The power supply is 3.0V, the input frequency 150MHz and the division ratio has been set to 64. This measurement has been done taking the voltage across a 47Ω (46.6Ω measured) resistor connected in series with the power supply, resulting in about 430µA per division.
The ESCL part of the prescaler draws about 470\mu A or 4/3 of the nominal value, as expected from the previous measurements. The CMOS part of the prescaler shows peaks of about 850\mu A while the large 2.1mA and 2.5mA current peaks are generated by the output pad driver. Apart from these large current peaks - the simulation does not include the pad driver - this measurement corresponds very closely to the simulation.

### 4.4.5 Prescaler Performance at Different Bias Currents

Since access to the main bias point is given (pin Bias), the bias current can be varied by sourcing or sinking some current from it. The next table shows the best measured performance achieved by one prescaler chip chosen at random under different bias current and supply voltage conditions. Being its best performance, it obviously cannot be guaranteed in production. The input signal is 0dBu, or +4dBu (1).

<table>
<thead>
<tr>
<th>VDD/VB</th>
<th>IB</th>
<th>f. max</th>
<th>IDD</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/2.5V</td>
<td>14.5\mu A</td>
<td>195MHz</td>
<td>0.289mA</td>
<td>1.48\mu A/MHz</td>
</tr>
<tr>
<td>3/2.5V</td>
<td>127\mu A</td>
<td>650MHz</td>
<td>1.835mA</td>
<td>2.82\mu A/MHz</td>
</tr>
<tr>
<td>5/3.0V</td>
<td>188\mu A</td>
<td>910MHz</td>
<td>3.07mA</td>
<td>3.37\mu A/MHz</td>
</tr>
</tbody>
</table>

The next two figures show the input and output waveforms of the prescaler at 195MHz and 910MHz. The conditions for these measurements are given in the previous table. The ringing in the output waveform in the 910MHz measurement was generated by some inductance in
the ground connection of the test fixture, while the input signal appears too small due to the 400MHz bandwidth limit of the oscilloscope. Its amplitude was $+4\text{dBu}$, as stated in the table.

![Image](image.png)

Although the current consumption at 195MHz is still about twice the projected value, both current consumption and efficiency can be considered very good for this technology. To show the significance of these figures, a comparison between this prescaler and other existing prescalers with similar functions and performance can be made. The prescaler [45] guarantees a minimum toggle frequency of 1.1GHz, typically consumes 4.5mA at 3V power supply and is implemented in ECL using QUBiC, an advanced 1.0μm BiCMOS process. To make a fair comparison, we need to know the maximum operating frequency which is not specified in the data sheet. Assuming 1.4GHz, a current consumption of 3.2μA/MHz will result, which is comparable with our measurements although our prescaler is implemented in a cheaper technology. Superior numbers are shown by the prescaler [46]. With a typical 1.4GHz toggle frequency (specified, guaranteed minimum is 1.1GHz) at 1.0mA, the current consumption is 0.71μA/MHz, but this excellent result is achieved using MOSAIC V, an advanced submicron bipolar process which is far more expensive than the CMOS process we used.

Thanks to the voltage controlled loads, a further advantage of ESCL over ECL is its ability to operate over a wide range of speeds and bias currents, provided the biasing for the load transistors is properly designed. As these measurements have clearly demonstrated, the bias current can be varied over more than a factor 10, and therefore, the same design can be tailored to different applications over a wide frequency range by simply adapting the bias current to fit the given application.
Chapter 5

GPS Receiver Chipset

5.1 Introduction

Despite its military origins, Global Positioning System (GPS) civilian use is growing steadily. Common civilian applications range from ship and aircraft navigation to precise surveying and geological studies, as well as a means to spread a precise world-wide coordinated universal time (UTC) that can be used in place of cesium or rubidium clocks, as reference for measuring instruments or as an aid in synchronizing digital communications. On the consumer side GPS is also proliferating. Many different hand-held or car mounted GPS navigation systems are available today to the consumer market. Some of them are quite simple and give the coordinates only, while some other give their position on maps drawn on a graphics display and usually read from a CD-ROM or a similar means. Location and tracking of mobile phones is another soon to become widespread consumer application of GPS.

The consumer market is notoriously very price sensitive, thus in order to make such consumer applications possible, cheap silicon must be available to enable the construction of low cost miniaturized receivers requiring little or no alignment. Highly integrated receiver chips with low external part count are therefore desirable, and today complete single-chip GPS receivers have become commonplace [21, 22]. These
single-chip receivers however, are usually targeted towards relatively big hand-held, in-vehicle or fixed applications where the power consumption or the number and size of external components is less a problem. Although performing well in their intended applications, these chips are therefore not satisfactory for the consumer market.

If GPS receivers can be made much less power hungry and much smaller than those on the market today, however, many applications outside navigation and surveying can be envisaged. The latter include receivers that are small enough to be carried in the pocket or worn on the wrist for better convenience. The chipset reported in this section of the thesis is targeted towards such applications, and forms a low power RF front-end for a GPS receiver for the 1575.42MHz civilian L1 band. The immediate application of such an integrated receiver is to provide GPS time reference to a wrist-watch, in a way similar to a DCF-77 watch, but operating world-wide. GPS positioning will then be used to set the correct time zone. To complete the watch, a small LC display may also be added at very little cost, to show the coordinates of the actual position of the watch. A possible aspect of the GPS watch is shown in Fig. 5.1.

Similarly to the ERMES watch pager presented before, also this application only makes sense if some conditions can be fulfilled, such as:

- Sufficiently good and reliable performance is realized.
5.1. Introduction

- The watch has no disproportionate size or weight.
- A sufficiently long battery life can be achieved.

Since these requirements are obviously in conflict with each other, the design of a GPS receiver fulfilling all three can be quite challenging.

Differently from most communications receivers, consumer GPS receivers for uncritical applications like this one are not requested to fulfill any type approval requirements, specific standards or anything similar (apart from general requirements such as not to cause interference to GPS itself or other services, which are usually covered by EMC regulations). Some trade-off in performance is therefore possible, provided a sufficiently reliable performance can still be achieved. Obviously a user expects to get the desired data from his GPS receiver when asked for, no matter if performance is guaranteed by type approval or not.

Size, weight and power consumption are certainly among the most important requirements for this application. The number and size of required ICs and external components must be minimized not only due to cost reasons, but mainly because of the small space available in a watch. Low power consumption is also a primary requirement. Although the GPS function will typically be used quite sparsely (say once a day or on request, especially if only time is desired), a large power consumption may easily result in power supply sag and battery overloading, leading to shortened battery life. The problem here is even more acute than for the ERMES watch pager, since the high operating frequency implies high power consumption. The specified power source is a small lithium battery (or 2-3 NiMH cells), with an operating voltage range of 2.4V to 3.6V.

For this project the objectives to be reached were all well defined, which meant that:

- A complete receiver plan, whose performance was validated through both simulations and experiments, was delivered to us.
- Sufficiently accurate documentation was made available.
- The design of antenna and digital correlator chip were scheduled in parallel (at other institutes), while the finished antenna - mounted into a Swatch case - was characterized and shown to us.
The purpose of this project was to demonstrate feasibility and performance of a low cost integrated GPS receiver that fits into a watch of regular size. Although a chip ready for mass production was not expected, a receiver chip(set), complete with all necessary functions, was desirable.

For this design no particular technology was proposed or recommended, and the choice of a suitable technology was left to us. Despite being somewhat expensive, Philips QUBiC 1µm BiCMOS technology was chosen, especially for its good high frequency performance. The NPN bipolar transistors have an $f_t=13$GHz, with low parasitic capacitances thanks to recessed oxide isolation. The most important characteristics of this process are shown in table 5.1.

<table>
<thead>
<tr>
<th>Process</th>
<th>Core Process Design rule</th>
<th>n-Epi, recessed oxide</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPN bipolar (non walled)</td>
<td>Type</td>
<td>Vertical, poly emitter</td>
</tr>
<tr>
<td></td>
<td>Peak $f_t$</td>
<td>13GHz</td>
</tr>
<tr>
<td></td>
<td>Beta</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>IKF</td>
<td>16mA/Emitter</td>
</tr>
<tr>
<td>PNP bipolar</td>
<td>Type</td>
<td>Lateral parasitic</td>
</tr>
<tr>
<td></td>
<td>Peak $f_t$</td>
<td>250MHz</td>
</tr>
<tr>
<td></td>
<td>Beta</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>IKF</td>
<td>40µA/Emitter</td>
</tr>
<tr>
<td>n-MOS</td>
<td>$V_t$</td>
<td>0.85V</td>
</tr>
<tr>
<td></td>
<td>$\mu_0$</td>
<td>450cm$^2$/Vs</td>
</tr>
<tr>
<td></td>
<td>$K'_n$</td>
<td>78µA/V$^2$</td>
</tr>
<tr>
<td>p-MOS</td>
<td>$V_t$</td>
<td>-0.82V</td>
</tr>
<tr>
<td></td>
<td>$\mu_0$</td>
<td>160cm$^2$/Vs</td>
</tr>
<tr>
<td></td>
<td>$K'_p$</td>
<td>28µA/V$^2$</td>
</tr>
<tr>
<td>Resistors</td>
<td>N-buried layer</td>
<td>33Ω/□</td>
</tr>
<tr>
<td></td>
<td>low R P-poly</td>
<td>250Ω/□</td>
</tr>
<tr>
<td></td>
<td>high R P-poly</td>
<td>1kΩ/□</td>
</tr>
<tr>
<td></td>
<td>salicide N-poly</td>
<td>8Ω/□</td>
</tr>
<tr>
<td>Capacitors</td>
<td>Poly-N-plug</td>
<td>1.5fF/µm$^2$</td>
</tr>
</tbody>
</table>

Table 5.1: Some important characteristics of the QUBiC 1µm BiCMOS process.
5.2 GPS System Overview

Positioning and navigation rely upon our knowledge of the precise position of some reference points, that may be located either in the sky, sea or ground. In the past centuries, stars and planets have been widely used as landmarks to locate objects (ships, observatories) accurately or as time references. Stars are in fact so far away that they appear fixed to an observer on earth, while the orbits of some planets may be used as chronometers. Navigation using stars works quite well, but cannot guarantee 24-hour availability, since the stars are not visible during the day and when the sky is cloudy.

Ground-based radionavigation systems, such as e.g. the Loran C, have been introduced to solve the availability problems caused by poor visibility. While guaranteeing 24-hour all-weather availability, the coverage of ground-based radionavigation is usually quite limited (shores or ports with heavy traffic).

In order to overcome this problem, the global positioning system (GPS) was introduced. Based on satellites, the GPS allows a continuous world-wide coverage, providing accurate 3-dimensional co-ordinates and velocity data as well as transmitting a coordinated universal time code (UTC). It is similar to navigation with stars, except that the stars are replaced by satellites. Although originated from the military, the GPS has evolved into a dual-use system, and is available to everyone at no charge. Typical applications may range from ship/spacecraft/aircraft/vehicle positioning, missile guidance to surveying and leisure hiking.

The GPS is based on 24 satellites (space vehicles, SV) that are located in 6 orbital planes at height of 20200km and circle the Earth every 12 hours. Each of these planes is inclined at 55 degrees to the earth's equator and contains 4 satellites. The GPS uses one way time-of-arrival ranging. Each satellite sends the UTC and navigation data using a direct-sequence spread-spectrum CDMA technique. A receiver can calculate its own position and speed using the data coming from any 4 satellites and the delays required in the correlators to correlate the signals of these satellites, as will be explained later. As a receive-only system, GPS may serve an unlimited number of users.

Two services are provided by GPS: a precise positioning (P-code) and a standard positioning service (coarse acquisition, C/A-code). The first
is primarily intended for military applications and provides a positioning accuracy better than 22m and a time accuracy of 200ns. Non military use of PPS is permitted, but requires approval by DoD. The standard positioning service is available to everyone, but its precision is somewhat reduced by dithering the satellite's clocks with a pseudorandom sequence, as well as the addition of offset errors to the transmitted navigation data (selective availability). The positioning accuracy is 100m while the clock accuracy is 340ns. Authorized PPS users can realize full system accuracy by removing these (encrypted) pseudorandom errors. Access to the P-code can be denied completely to standard positioning users, by activating an antispoofing (AS) mode in the satellites. AS has been provided to defeat deception jamming. In the latter case the P-code is encrypted, and forms the Y-code. Because of this characteristic, the precise positioning code is also known as the P(Y)-code [19, 20].

All 24 satellites send on the same two frequencies, L1 and L2, derived from a 10.23MHz atomic frequency standard. L1 is the primary frequency, 1575.42MHz (154 times the atomic clock) and L2 is the secondary frequency, 1227.6MHz (120 times the atomic clock). Interference between signals of different satellites is avoided using direct sequence pseudorandom noise (PRN) codes with low cross-correlation for
the CDMA modulation. The modulation used is binary phase shift keying. L1 carries both the C/A-code and the P(Y)-code, in quadrature and attenuated 3dB, as well as a 50bps navigation message, L2 carries usually the P(Y)-code and the 50bps navigation message. Alternatively the navigation message can be suppressed from L2 to improve performance against jamming, or the C/A-code can be transmitted instead. A block diagram of the signal structure of GPS is shown in Fig. 5.2.

The direct sequence PRN code is generated using linear feedback shift registers. The C/A-code uses Gold codes with a length of 1023 bits, obtained by XORing the output of two 10 bits shift registers that generate maximum length pseudo noise (Fig. 5.3). Since the chip rate of C/A-code is 1.023MHz, the repetition rate of the PRN sequence is 1ms. The short repetition rate results in less than optimum power spectrum which has 1kHz spaced lines, that deviates, sometimes substantially, from the ideal sin(x)/x envelope.

For the generation of the P(Y)-code, four 12 bit shift registers are used. Two of them are XORRed together to form X1, while the other two form X2 in the same way. After delaying and XORing, as shown in Fig. 5.3, the P(Y)-code is formed. With this architecture, a very long code is generated, whose length would be more than 38 weeks, if it were not truncated after 7 days and repeated. 37 unique sequences,
6.1871 \cdot 10^{12} \text{ chips long (7 days) are thus formed, and 24 of them are assigned to each of the 24 GPS satellites. Even if the P(Y)-code is not a maximum length code, the very long period and the fast chipping rate result in essentially optimal spectrum characteristics.}

In order to determine its position, a GPS receiver must find the value of 4 unknown variables: the X, Y, Z coordinates, and time, the latter is required because the receiver’s clock is not synchronized with satellite’s clock. The receiver, therefore, requires the data from at least 4 satellites, with which it can set up a system of 4 independent equations, whose solutions are the desired coordinates and time. Basically, to acquire the data of a satellite, the receiver uses cross correlation, that is, it must replicate the satellite’s pseudorandom code and shift it until maximum correlation with the received signal is found. The delay required corresponds to a time-of-arrival measurement, as measured by the receiver’s clock.

Since the P(Y)-code is very long, its acquisition may be very slow, unless the receiver is accurately initialized in position or time, to reduce the number of delays to search. With the 7-day code, that consists of more than $6 \cdot 10^{12}$ bits, the receiver must know with some precision where to start trying to correlate its replicated code with the received signal. To assist P(Y)-code acquisition, the C/A-code is available to the receiver. Being the C/A-code only 1ms long, its acquisition is very fast. C/A-code data can be used to initialize the receiver for P(Y)-code acquisition, or can be used alone for standard positioning users that do not need or are not allowed to access PPS.

5.2.1 Receiver implications

As previously said, a GPS receiver needs the signals from 4 satellites to have enough data to calculate its position and time. This means that 4 tracking channels are required. To guarantee uninterrupted service, at least one more tracking channel should be added to the receiver, which will track one more satellite to be used as replacement in case the signal of a satellite gets corrupted or lost. More channels are usually made available to provide all-in-view navigation solutions and speed up search. Typical GPS receivers (e.g. [22]) use 6-12 correlator channels.

In order to acquire and track a satellite, the receiver must not only
replicate its PRN sequence, but must replicate its carrier exactly to bring the satellite's signal to base-band, while maintaining the phase of the signal. A high precision, high resolution (mHz) frequency synthesizer is therefore required for every tracking channel, which is adjusted away from nominal to compensate for Doppler errors (different for every SV, and caused by SV and receiver movement) and reference TCXO error. Because of the 50bps data, that may occasionally reverse the phase of the carrier every 20ms, a usual PLL cannot be used for the frequency synthesizers. Typically only Costas loops are used in GPS receivers, because they can be made insensitive to 180° phase reversals. Because of these requirements, an analog implementation of the tracking channels is rather impractical (it is made e.g. to achieve a high resilience to interference), and a digital receiver architecture with sampling at IF will be the most natural choice. A block diagram of such a receiver is shown in Fig. 5.4.

A typical GPS RF front-end [21] consists of a double or triple conversion superhet, which is required to convert the incoming signals to a relatively low frequency IF (a few MHz) for efficient sampling and AD conversion, without requiring critical image and channel filters. Simple LC filters are in most cases sufficient (and used in some of the front-end ICs in [21]), although SAW filters may be used to provide somewhat better performance, or smaller dimensions (for filters above \( \approx 200\text{MHz} \)). At the last IF, because of the relatively low performance required, the filter may even be fully integrated, to reduce complexity and size of the receiver.

The filtered signal will then be sampled - sometimes subsampled, thus generating a new, digital IF - and processed by the digital correlator chip. In case of subsampling, only very small factors can be used, to
prevent excessive noise degradation due to the noise at the image frequencies. The large gain in front of the sampler, as well as the IF filters, will also alleviate the noise problem. In some cases, the IF frequency can be chosen to be a small integer multiple of $f_0$ ($f_0 = 1.023\text{MHz}$), although this is not a requirement, and any other frequency between $\approx 1-10\text{MHz}$ will do.

The digital portion of the receiver will consist of 6-12 tracking modules, each containing an I/Q demodulator and a correlator. A block diagram of a typical tracking module for the C/A-code [21] is shown in Fig. 5.5. The demodulator contains two mixers (multipliers) driven by a DCO, through $\sin$ and $\cos$ maps (ROM). The DCO is simply an accumulator, which is loaded with a given value and complete a carrier cycle when it overflows. The length of the accumulator is set to have a very fine resolution, below 100mHz or so, which is required to keep the DCO in phase with the SV signal. The base-band signals are then passed through the correlators, consisting of multipliers driven by the C/A-code replica and integrate and dump blocks (accumulators). The latter integrate the mixer's output over a complete code period, nominally 1ms, before dumping the result and start again. Before being dumped, the result is read and evaluated by the receiver processor, which searches the maximum value - positive or negative depending on transmitted bit - which indicates highest correlation, by adjusting the DCOs.
5.2. GPS System Overview

A straight implementation of a direct conversion architecture will be very impractical for a GPS receiver. Because of the need of replicating carrier frequency and phase precisely, which are different for every SV, a complete front-end, with its high resolution VCO and frequency synthesizer, would be needed for every tracking module. A possible solution here, if direct conversion is really desired, may be to use a Barber receiver architecture. In such architecture, the RF signal is first converted to base-band and low-pass filtered - like in any conventional direct conversion receiver - using I/Q channels, then it is upconverted again to a convenient IF, with two mixers and a quadrature oscillator. The output of the two mixers, after being summed together, can then be fed to the tracking modules, as in the superhet case. See also sec. 2.2.

From the point of view of the analog part of the receiver, the GPS signal consists entirely of noise. This is not only due to the PRN signals, emitted by all SV on the same frequency. Considering that the minimum power emitted by a SV on the L1 frequency is only 32.6W, spread over an entire hemisphere, there is not much power left to be collected by a receiver's antenna. Unless a big antenna is used, the GPS signals are in fact so small to be buried into the thermal (cosmic) noise. Because of this characteristic, the resolution of the AD converter can be very low. Two (sign/magnitude) to 3 bits are commonly used, while in many cases where best performance is not needed, 1 bit is sufficient. The large thermal noise associated with the GPS signal effectively dithers the AD converter, that will be well-behaved despite its low resolution. Using more than 3 bits has usually no performance advantage, while the width of the accumulators of the correlators may grow unnecessarily. If a 1 bit AD converter - i.e. a comparator - can be used, no AGC is required in the receiver. A simpler, lower power receiver can thus be designed.

Non linearity and blocking performance of the RF front-end also benefits from the very low signal amplitude. For the GPS signal itself, which is buried into noise, front-end nonlinearity is not much a problem. The received signal level is in fact very low and relatively constant, so that the linearity specification is dictated by the required system performance in the presence of external interfering signals rather than GPS signal distortions. For non critical applications, where relatively low blocking performance can be accepted, there is potential for power consumption saving in the front-end.

Since GPS is a spread spectrum system, phase noise of the local
oscillator is in general not a great concern regarding detectability of the GPS signal itself. Within the (wide) band considered, even an oscillator with moderate specifications will in fact contribute a negligible amount of noise when compared to the rest. On the other hand, phase noise of the local oscillator is a basic limitation on narrowing the loop bandwidth of the carrier-tracking Costas loop, and therefore on the achievable C/N ratio for a GPS receiver tracking a single satellite. When tracking multiple satellites, however, the phase noise is common to all tracking loops and can be removed by a common-mode rejection scheme if desired. No particularly good local oscillator is therefore necessary to achieve a sufficient receiver performance, and even values as poor as $-90$–$95$ dBc/Hz at $100$ kHz offset are still acceptable.

A GPS receiver needs only one local oscillator frequency. Fixed dividers with simple division ratios can therefore be used. The lowest power consumption can be achieved if a chain of divide-by-2 toggle flip-flops is used. Differential current mode logic such as ECL or ESCL can be used here, while dynamic CMOS logic should be avoided. The large, unbalanced rail-to-rail CMOS logic levels, and the consequent need for an amplifier to boost the small local oscillator signal to valid CMOS levels may result in excessive power consumption, as well as excessive substrate contamination, perhaps up to the point to prevent the integration of the prescaler on the same chip as the remainder of the receiver.

5.3 The Active Patch Antenna

The antenna plays an important role in defining the performance of a system like this one. The sensitivity of the receiver depends in fact to a great extent on antenna gain and directivity pattern. Thanks to the high operating frequency, an efficient antenna with small dimensions can be readily designed, but many constraints peculiar to this application must still be fulfilled. These are:

- The antenna must have a right-hand circular polarization.
- The axis of maximum sensitivity must be perpendicular to the dial. Looking at the watch will then orient the antenna towards the satellites.
5.3. The Active Patch Antenna

- Small size, low profile antenna. Must fit into the watch case without occupying too much internal space and without protruding outside it.

- The antenna should not interfere with the features of the watch (e.g. hands and LC display) excessively.

The design of the watch should of course also not hinder the performance of the antenna. If features (such as the dial or the hands) are required to lie in front of it, a suitable design and choice of materials for the latter is necessary. Since the design of the antenna was done at a different location and was not under our responsibility, only a brief description is given here.

Microstrip antennas are one of the most common low profile antennas. They are derived from microstrip resonators, and use the radiation loss of the resonator in a positive manner. They are constructed by printing conductors of the desired shape on a dielectric substrate. The radiator may take the form of a line (blade antenna) or of a typically rectangular, circular or annular patch (patch antenna) [35]. Being based on a resonator, a patch antenna is necessarily narrow band, a definite advantage in our case since the necessary prefiltering is already provided by the antenna itself.

For our project, an active patch antenna with a circular radiator has been chosen since it best fits the watch application. The antenna will be placed just behind the dial, or replace it completely thus becoming an aesthetical feature of the watch. If the display (Fig. 5.1) is not required the only hindrances will be the hands and the top glass. An appropriate choice of the materials for the latter will minimize absorption and thus sensitivity loss. In case the display is desired, the available area for the patch is reduced and its form must be modified to fit the remaining area. Placing only a small display on the lower part of the dial, as shown in Fig. 5.1, will help maintaining a relatively large area for the antenna.

The basic structure of the patch antenna and how it is mounted into the watch case is shown in Fig. 5.6a and 5.6b respectively. It consist of a round radiator printed on a substrate with a ground plane under it. The center of the round radiator is grounded. A tubular connection will be used, such that the hub of the hands can pass through it. The feed point of the antenna is offset from the center. Its position determines
antenna impedance and is chosen to (noise) match the input of the active antenna preamp.

Right-hand circular polarization can be obtained with one of the several usual means, such as e.g. two feed points at $90^\circ$ fed in quadrature, an elliptical patch or a tab oriented at $45^\circ$ from the feed point. The radiation pattern of this antenna has its maximum on the axis perpendicular to the radiator as required. In that direction a gain of 0dBi was given for the unmounted antenna, which was reduced by a couple dB when mounted in a regular Swatch plastic case - not the best material regarding RF absorbtion.

The antenna preamplifier is constructed with a discrete bipolar transistor and a few passive components. It has a gain of 10dB and a noise figure of 1.5dB. Thanks to the selectivity of the antenna no particularly good filter is necessary to achieve an acceptable rejection of out-of-band interferers, and a simple low Q LC filter (or no filter at all) is sufficient. The output of the preamplifier is connected to an image reject SAW filter. It has therefore been impedance matched to $50\Omega$ as required by the latter. In the next section the complete receiver plan, including the active antenna, is given and discussed.

### 5.4 System Architecture

As it was already the case for the ERMES watch pager, when planning a receiver like this several parameters somewhat different than usual
must be considered. Because of the rather unconventional application, small size and low power consumption tend to be more important than overall performance and even price. It is not much worth to have a good receiver if it and especially the battery do not fit into a watch case of usual dimensions.

### 5.4.1 Block Diagram

A block diagram of this receiver, including the most important external components is shown in Fig. 5.7. Although a single-chip implementation is envisaged, due to time constraints the receiver has been split into 2 blocks. The 1st is the 1.57GHz RF front-end to the left of the dashed line in Fig. 5.7. It consists of an LNA, a single balanced mixer, a Colpitts LC oscillator and an ECL divide-by-8 prescaler (shown in dark gray). The IF and base-band components (light gray) form the 2nd chip, which has been designed in the same BiCMOS technology as the front-end, such that both chips can be easily merged into a single chip solution at a later stage.

To reduce power consumption, the most obvious thing to do is to reduce the number of components operating at the highest frequency and providing them with a fast recovery power-on capability for ON/OFF operation. A superheterodyne receiver is therefore the most natural choice. High gain IF amplifiers must also work at the lowest practical frequency. On the other hand, filter size and selectivity requirements
prevents the intermediate frequency from being too low. SAW filters for frequencies lower than roughly 100–150MHz are typically too big and incompatible with our application. As a trade-off between these requirements, a triple conversion architecture has therefore been chosen.

The 1.57GHz GPS L1 signal is received by the active patch antenna, amplified by 10dB and filtered with a SAW filter to remove signals at the image frequency and other strong out-of-band signals which may overload the front-end. This filter, a SAW device [47], provides a typical out-of-band attenuation of 40–45dB (minimum 35dB), has an insertion loss of 2.8dB and requires 50Ω terminations. Its SMD package, with a size of only 3x3x1.1mm³, occupies little space within the watch and is fully compatible with the intended application.

After prefiltering, the GPS signal enters the receiver chip, where it is amplified by about 15dB and mixed down to 1st IF with a single balanced mixer. The 1st IF signal is then routed to an off-chip 179MHz SAW filter where channel filtering takes place. By limiting the bandwidth of the receiver to the minimum necessary, the use of a SAW filter gives a significant improvement in correlated SN ratio compared to a simple LC filter. The performance of this filter is somewhat critical to system performance, due to the rather low 2nd IF chosen. Since a filter with the desired frequency and passband characteristics could not be found as standard component, a custom design was considered. The filter has single-ended input and output, is specified for a center frequency of 179.2MHz and a bandwidth of 1.9MHz, and requires a termination impedance of 500Ω on both sides. Although a relatively high IF frequency has been chosen to minimize its size, the dimensions of this SAW filter will already be around 9x7x2mm³. It will therefore be one of the biggest parts of the receiver, and little hope exists that its size can be reduced substantially in the future.

The filtered signal is then amplified by 12dB and down-converted to the 2nd IF of 4.7MHz with a double balanced mixer. Since channel filtering occurs at 1st IF, the 2nd IF filter need only remove the higher frequency mixing products and the local oscillator feedthrough of the 1st mixer, which are both well above 4.7MHz. The required performance is therefore quite moderate, allowing the filter to be fully integrated on chip. The main benefit of integrating the 2nd IF filter is the elimination of an external, possibly bulky filter and its associated passive components, thus reducing receiver complexity to a level comparable to that
of a single superheterodyne receiver. A 5th-order Butterworth active RC filter has been chosen. After 68dB of gain, the 2nd IF signal is amplitude limited and converted to digital with a 1bit AD converter, i.e. a comparator. Using a 1bit converter results in slightly degraded performance compared to that of a multi-bit converter, but it allows the design of a simpler, lower power 2nd IF strip without automatic gain control. Sampling at 3.6MHz, the AD converter also down-converts the 2nd IF signal to the 3rd IF of 1.05MHz.

Signal detection is performed digitally by the correlator chip that contains all the digital processing and controlling parts of the receiver. The 1.05MHz signal, sampled at 3.6MHz, is converted to base-band by a quadrature demodulator and fed to the tracking modules (see Fig. 5.5). Thanks to the 1bit design, the mixers consist of simple XOR arrays, which helps reducing chip size and power consumption somewhat. A more important place where substantial power can be saved and chip area reduced is the tracking module array. If best performance is required, using 16 tracking modules will allow all-in-view measurements, but at the cost of chip area and power consumption. In our case a lower number of tracking modules (8–12) was considered - with perhaps only 5 or 6 active after at least 4 satellites have been acquired - to save power.

Since GPS uses CDMA the receiver needs only to receive one channel, which means that only a single frequency has to be generated by the local oscillator. The frequency synthesizer is therefore a very simple design using fixed dividers with simple division ratios, which results in a further reduction of power consumption. Most important here is the use of an asynchronous divider by 8 - i.e. 3 cascaded toggle flip-flop - at the highest frequency. As reference, a precision crystal oscillator is used, whose nominal frequency is 10.907875MHz.

### 5.4.2 System Planning

The plan for this receiver is shown in Fig. 5.8. In the block diagram, the parts integrated in the developed chipset are shown in gray. Below it a table with voltage gains, noise figures (F), compression points (1dB CP) and current consumption (Idd) is shown, together with the expected signal amplitudes for the GPS signal, thermal noise and interference given in dBm or dBu (50Ω) as appropriate.
Setting the gains for the various blocks of a receiver is always a trade-off, especially at the front-end. The main parameters that have to be taken into consideration are noise figure, intermodulation (IP3) and power consumption. The trade-off is essentially the following: choosing high gain for the active antenna and the LNA will help reduce the noise figure by minimizing noise contribution of the mixer, but at the expense of higher power consumption in these blocks and risk of early mixer overloading. Lower LNA gain may improve receiver linearity and LNA power consumption, but a low noise mixer that will very likely consume much power would be required to maintain an acceptable noise figure, possibly resulting in no power consumption advantage.

To determine gain, noise figure and linearity requirements of the various receiver blocks, one has to first find which kind of signal will be available from the antenna. Although the signal at the antenna output will consist mainly of noise - given that the GPS signal is much weaker and well buried into it - a good noise figure is still needed. The processing gain is in fact much better used to extract the GPS signal (maybe weakened or disturbed) from cosmic noise or interference rather than from the noise of a bad receiver.
5.4. System Architecture

**Noise figure:** The very small patch antenna required by the watch application has necessarily low gain. In the plan delivered to us this was given as 0dBi, which is not a bad value for such a small antenna. This tends to relax the linearity requirement somewhat, but imposes a low noise figure for the receiver. A further - luckily wrong - assumption that played a role in the planning of this receiver is that the noise figure of the on-chip LNA and mixer cannot be good because of (usually) quite high $r_B$ of integrated BJTs. A front-end with relatively high gain has therefore been chosen. With the gains and noise figures shown, the total SSB noise figure, referred to the input of the on-chip LNA is 7.5dB, while at the receiver input (antenna) it is 3.9dB, which is adequate considering the application of this receiver. Since noise contributions of on-chip LNA and image filter are quite substantial in this plan, a small gain increase (e.g. to 13dB) of the preamp could easily improve the overall noise figure (e.g. to 2.9dB).

**Linearity:** Having a large front-end gain may lead to linearity problems, thus linearity must be analyzed first to validate this plan. As previously stated, for the GPS signal itself front-end nonlinearity is not much a problem. The linearity specification of the receiver is therefore dictated by the required system performance in the presence of external interfering signals. The most troublesome interferers identified are powerful radars for aeronautical radiolocation (around 1.3–1.4GHz), which may appear at up to $-10$dBm or so at the antenna preamp input. A 1dB compression point of $-10$dBm has therefore been assumed. The RF interstage filter [47] has a somewhat peaked response, with two (relatively uncritical) $-15$dB peaks around 1.1GHz and 1.9GHz and an attenuation of about $35–40$dB in the range 1.3–1.4GHz. Assuming a $33$dB attenuation ($30$dB below in-band insertion loss), the radar signal will appear at $-43$dBm at the input of the chip. This is a very moderate value that can be easily achieved at low power despite the high total front-end gain, thus no particular problem is expected regarding linearity.

**Power consumption:** Power consumption for this plan is quite low considering the high operating frequency and technology used. The front-end - which will certainly dominate power consumption - consumes 13mA including the antenna preamp. Assuming another 2mA for the remainder of the chip, a total current of 15mA, or 45mW of power at 3V supply will be consumed by the RF analog part of the GPS receiver. A
correlator chip with 5 or 6 (active) tracking modules can be estimated to consume around 10mA and a microcontroller another 10–15mA. A total of 35–40mA can be expected, which is sufficiently low for our application and compares very well to GPS chipsets existing at the time of this work (1994). ON/OFF operation is an obvious means to further reduce power consumption. By chopping the power supply at intervals either below 1ms\(^1\) (at the cost of C/N ratio) or at intervals of several 100ms or even seconds (at the cost of GPS updates), the average current consumption can be reduced easily to around 1mA or below. A battery supporting high drain for short periods will nevertheless be still necessary.

A possible objection to this plan is the requirement for many external parts, some of them quite bulky. The most important are certainly the discrete antenna preamp and the two SAW filters, but the VCO tank and the unavoidable 10.9MHz quartz are not negligible as well. Nevertheless the situation is not that bad, and many solutions exist to this objection, especially if some performance can be traded-off in favor of a reduced size or part count.

At the time this design was planned (1994), low cost integrated LNAs with the required performance, size and operating voltage and current (3V, 3mA) were not yet available, thus a discrete design (GaAs or Si-Ge HBT) was the only possible choice. Today a device such as [48] could be used, that while having the required small size (6 leads SOT23), has 15.7dB of gain and 1.2dB of noise figure, thus leading to not only a reduction of size and required parts, but - apart from signal handling capability - also to improved performance. Alternatively an LNA with sufficiently low noise figure (about 2.5dB) could be integrated on-chip, although this will not necessarily result in a simpler system.

The RF SAW filter - which is small and relatively low cost - is better kept, since it will improve performance by limiting receiver bandwidth and guarantees a good attenuation of image and out of band interferers. On the other hand, if some performance degradation is acceptable, the bulky and expensive (custom design) 1st IF filter could be replaced by an LC bandpass filter. The chip, as designed, will accommodate it without modifications.

\(^1\)This requires that the integrated values of all tracking modules are saved during the OFF period. The carrier and code NCOs, as well as the C/A-code generator, must operate normally to avoid the loss of satellite tracking.
5.5. The Low Noise Amplifier

As it will be explained later, the local oscillator has been designed around an external tank partly to save power and partly to have a direct access to the mixer, and not because of phase noise considerations. In the final design the external tank could be suppressed at the expense of some increase of power consumption. It is up to the customer to decide whether he prefers lower power or lower part count.

Apart from the 1st IF SAW filter, the 10.9MHz quartz will be one of the biggest parts of the receiver. A quartz with the required precision (±5ppm or so) is usually packaged in a ceramic case with a size of 7x5x1.2mm³ or similar, which should not pose any particular placement problems even in a small watch case. If needed, much smaller units are becoming available today, with at least one² as small as 5 x 3.2 x 0.9mm³.

Differently from the ERMES watch pager, this receiver does not require trimming. Bulky trimmers, as well as expensive alignment operations are therefore avoided, while PCB placement is simplified.

5.5 The Low Noise Amplifier

The performance of the LNA required by this application, as shown in Fig. 5.8, is quite moderate, especially in terms of noise and signal handling capability. Again, parameter independence from temperature, process tolerances or VDD variation is important, as is a 50Ω input impedance to adequately terminate the RF SAW filter, possibly without the need for an external matching network.

To obtain a gain of 15dB at 1.57GHz a 2 stage amplifier is needed, the schematic diagram of which is shown in Fig. 5.9 together with the mixer. Similarly to the LNA for the ERMES receiver described earlier, a transconductance amplifier (transistor Q1) driving a transimpedance load (Q3 and related parts) has been used, thus the same design methodology applies. The main difference is the much higher current levels required by this design, dictated by the higher operating frequency. This requires a different transistor sizing strategy.

²Tokyo Denpa Co., TSS6 SMD quartz crystal units
5.5.1 The Input Stage

The main parameters to be considered while designing the input stage are bandwidth and noise performance. Linearity is not an issue here, since the output stage and especially the mixer will determine it.

Achieving high frequency operation at the lowest possible operating current would dictate a relatively small transistor biased near to peak $f_T$. At these conditions shot noise becomes small and noise is dominated by $r_B$, which can only be lowered by increasing transistor size. This is in conflict with the previous requirement, thus a balance between the two must be found. Using $r_B$ data (400Ω per emitter) and a measured curve for $f_T$ vs. $I_C$ (delivered with the process documentation), a transistor with 24 emitters in parallel biased at 3mA was chosen. Under these conditions, $f_T$ is about 8GHz while $r_B$ is 17Ω, which contributes a noise voltage density of $526\mu V/\sqrt{Hz}$. The contribution of $Q_1$’s $I_C$ shot noise is as expected a much smaller $280\mu V/\sqrt{Hz}$.

Since the source impedance is not negligible, $I_B$ shot noise is not negligible either. Assuming an $f_T$ of 8GHz, the noise current density at the base of $Q_1$ is $14pA/\sqrt{Hz}$ at 1.57GHz, or $700pV/\sqrt{Hz}$ assuming a 50Ω source resistance.

The small-signal input impedance of a bipolar transistor becomes small at high frequency because of the shunting effect of the input capacitance $C_e$ and the Miller capacitance $C_m$. The ultimate value of the input impedance will approach $r_B$, i.e. 17Ω, which is much lower than...
the desired 50Ω. A common technique to obtain impedance matching without using an external matching network is to add an inductor in series to the emitter. The input impedance therefore becomes

\[ Z_{\text{in}} \approx \beta(j\omega) \cdot Z_E = \frac{\omega_T}{j\omega} \cdot j\omega L_E = \omega_T L_E \]

which is (ideally) purely resistive. By substituting \( Z_m = 50\Omega \) and \( f_T = 8\text{GHz}, \) \( L_E = 1\text{nH} \) is obtained, which can readily be constructed, without using external components, with a short bonding wire. Since this value is quite small, the emitter connection of \( Q_{1a} \) has been brought out on 2 separate pads, such that the emitter inductor may be realized as 2 bonding wires in parallel, effectively halving the total inductance. In our test circuit, however, the best results were obtained with a single, short bonding wire.

Since the N-plug–Poly capacitors have a large bottom plate capacitance, capacitive coupling to the 2nd stage has not been implemented to avoid excessive capacitive loading to the collector of \( Q_{1a} \). The two stages have therefore been DC coupled. The collector voltage of \( Q_{1a} \) is therefore not available for biasing, thus \( Q_{1a} \) has been biased to 3mA with the replica circuit \( Q_{1b}, R_{1b} \) and \( R_{2b} \). For the same reason no current source has been used as load, and a low capacitance poly resistor has been used instead. The 1st LNA stage, with its bonding wire matching network, has a voltage gain of 12dB.

### 5.5.2 The Output Stage

The design of the output stage of this LNA is similar to the one used in the ERMES pager receiver, thus it is described only briefly here. The transimpedance stage is formed by \( Q_3 \) and \( R_3 \), and has been dimensioned for a gain of 4dB, which yields a gain of 16dB for the complete LNA. Since linearity is not critical here, a small bias current of 2mA can be used for this transistor. The gain of 4dB has been set by dimensioning \( R_3 \) accordingly (110Ω). A small 12Ω degeneration resistor has been inserted in \( Q_3 \) emitter. Although this improves linearity somewhat, the main purpose of it is to achieve a well controlled bias current in both \( Q_3 \) and the single balanced mixer DC coupled to it and biased by \( Q_3 \) itself. Resistive biasing has been chosen again to minimize stray capacitances. Thanks to the low stray capacitances, the LNA has a
well behaved frequency response, with no peaking and a well controlled 6dB/octave roll-off.

As it was already the case for the ERMES pager LNA, the main noise contributor of the 2nd stage is the feedback resistor $R_3$. Its $1.35\text{nV/}\sqrt{\text{Hz}}$ noise appears unmodified at the output of the LNA, thus contributing $210\text{pV/}\sqrt{\text{Hz}}$ to the LNA input. The size of $Q_3$ has been chosen to make its noise contribution lower than $R_3$'s. Using four emitters, the input referred $Q_3$ noise is only $170\text{pV/}\sqrt{\text{Hz}}$.

All main noise contributors are now known and LNA noise, thus noise figure, can be calculated. The first stage contributes $526\text{pV/}\sqrt{\text{Hz}}$ from $r_B$ and $280\text{pV/}\sqrt{\text{Hz}}$ shot noise, while the 2nd stage contributes another $210\text{pV/}\sqrt{\text{Hz}}$ ($R_3$) and $170\text{pV/}\sqrt{\text{Hz}}$ ($Q_3$). This gives a total input referred noise voltage density of $654\text{pV/}\sqrt{\text{Hz}}$. A Spice simulation, which calculates the contributions of all elements, gives $673\text{pV/}\sqrt{\text{Hz}}$.

Noise matching occurs for $R_n = e_n/i_n = 47\Omega$, which is close to the source impedance. Although having noise matching equal to the source impedance does not guarantee that this is the design with the lowest possible noise figure, near optimum noise performance as well as minimum noise sensitivity to input mismatch is obtained for this particular design. The noise figure can then be calculated as:

$$F = 10 \log \left( \frac{e_n^2(R_s) + e_n^2(LNA) + R_s^2 i_n^2(LNA)}{e_n^2(R_s)} \right)$$  \hspace{1cm} (5.1)$$

which - resulting in $3.2\text{dB}$ (Spice: $4.1\text{dB}$) - is well within our specification of $7\text{dB}$.

Since noise matching is at least in first approximation already realized and contributions from noise sources other than $Q_{1a}$ are small, improving the noise figure of this LNA is not easy. Reducing the $e_n$ produced by $r_B$ of $Q_{1a}$ - i.e. the largest noise voltage contributor - would require a larger transistor, which in turn results in increased $i_n$ due to $f_T$ reduction or increased base current if $I_C$ is increased to maintain the same $f_T$. Vice-versa an improvement of $i_n$ can only be achieved with a reduction of $Q_{1a}$ size and bias current, both leading to degraded noise voltage performance. A single stage LNA could be a way to achieve some $F$ improvement by removing the contributions of the 2nd stage if a lower gain can be accepted.

In the discussion above a purely resistive behavior has been assumed,
which - although reasonable in first approximation - is unlikely at GHz frequencies for the technology used in this work. Some noise figure improvement could be achieved by matching the LNA input to the optimum noise (complex) impedance, but since the RF interstage filter needs 50Ω terminations, such noise matching is not possible. In the process documentation there was also no sufficient data to find $\Gamma_{opt}$.

The input referred 1dB compression point of this LNA is dominated by the 2nd stage, and is calculated to be $-22.7$dBm which is much better than the specification. The overall front-end 1dB CP, however, will be dominated by the mixer. Measurements on the whole front-end, as shown later, confirm these estimates.

### 5.5.3 LNA Biasing

As in most circuits that do not use global feedback - and this LNA is no exception - the necessary independence from temperature, supply voltage and process is not intrinsic, and must therefore be designed-in from the very beginning. Apart from establishing the desired bias point for the LNA, it is the important task of the biasing circuit to achieve such independence. Although due to time constraints no independent bias has been implemented in this prototype, a suitable circuit has nevertheless been designed well in advance.

Given that the $g_m$ of a bipolar transistor is inversely proportional to the absolute temperature, temperature compensation of the LNA gain requires that its bias current have a PTAT characteristic. As said
in the previous section, to maintain a low capacitive loading, resistive biasing has been used in place of current sources within the LNA, thus the required PTAT bias current can only be obtained by regulating the supply voltage accordingly. This can be easily achieved with the circuit shown in Fig. 5.10.

The operation of this circuit is quite straightforward. A $\Delta V_{BE}$ current source supplies a PTAT current to the branch comprising $R_{tc}$ and $Q_{tc}$. The voltage across $R_{tc}$ is therefore also PTAT. The voltage on top of $R_{tc}$ is buffered, decoupled ($C_d$) and becomes the supply voltage of the LNA. Since the $V_{BE}$ of $Q_{tc}$ cancels out the $V_{BE}$ of $Q_{1a}$ and $Q_3$, the voltage across $R_{1a}$ and $R_4$ is identical to the one across $R_{tc}$ and has the desired PTAT characteristic, which imposes the required PTAT bias current to $Q_{1a}$ and $Q_3$.

Since the LNA supply voltage is regulated, independence from VDD is also realized, provided an LNA supply somewhat below the minimum VDD specified (2.4V) is chosen. A value around 1.8–2V will guarantee sufficient dropout for the buffer together with sufficient voltage drop on $R_{1a}$ and $R_4$, preventing the latter from becoming too low. Process independence is obtained by using resistors of the same type for the $\Delta V_{BE}$ current source, the bias voltage generator and the LNA.

5.6 The RF Single Balanced Mixer

According to the GPS receiver plan of Fig. 5.8, the RF mixer should have a gain of 10dB and a noise figure of 12dB, with a compression point of just $-28$dBu. A proper 500$\Omega$ termination to the IF filter is also required, while its input - as an internal node - must only have a sufficiently large impedance to prevent overloading the LNA.

In this case the architecture was clearly specified in advance as superhet with external 1st IF filter. In order to minimize power consumption a single balanced RF-mixer has therefore been chosen. The conversion $g_{m0}$ of a single balanced mixer is in fact roughly twice that of a double balanced mixer for the same bias current, at the expense of linearity. The large local oscillator feedthrough typical of single balanced mixers is of no consequence in this application, since the oscillator signal is removed by the external IF filter. This mixer, also shown in Fig. 5.9,
5.6. The RF Single Balanced Mixer

consists of the input transconductor $Q_4$ with its 12Ω emitter degeneration resistor and the chopping differential pair $Q_5$ and $Q_6$.

In order to obtain a gain of 10dB with a 250Ω load (i.e. the 500Ω 1st IF SAW filter in parallel with its 500Ω matching resistor), a conversion $g_m$ of 12.6mS is required. Taking into account the effect of the emitter degeneration resistor, a collector current of 2mA is required for $Q_4$.

Since the capacitors in this technology have a large bottom plate capacitance, DC coupling between LNA and mixer was used to prevent excessive capacitive loading of the LNA output. DC biasing at 2mA of the mixer is provided by the LNA 2nd stage, $Q_3$, with which $Q_4$ forms a 1:1 current mirror. Since the collector currents of $Q_3$ and $Q_4$ are equal and - provided the LNA is biased as shown in Fig. 5.10 - have PTAT characteristic, temperature compensation of the conversion $g_m$ of the mixer is also provided. Emitter degeneration (12Ω) was added principally to improve DC bias stability, but a linearity improvement with only a slight degradation of noise figure also results.

In order to save pins, both the local oscillator and the IF ports are unbalanced. The LO in pin is connected directly to the base of the local oscillator transistor. DC bias to this pin is provided by a 5kΩ resistor to VDD, which also biases the local oscillator.

An undesired effect of having an unbalanced IF port is a 3dB increase of the mixer noise figure. Assuming a switching time $\tau$ of 5% of the cycle (which is usual), the current noise density at the output of the mixer can be calculated with equation 3.5 for the single balanced mixer. Because of the unbalanced output, the noise power will be doubled, thus a noise current density of 44.6pA/$\sqrt{Hz}$ results, where 41pA/$\sqrt{Hz}$ are contributed by $Q_4$ and its $r_B = 100Ω$. The noise voltage density at the input of the mixer is therefore 3.53nV/$\sqrt{Hz}$, which corresponds to 561pV/$\sqrt{Hz}$ at the input of the LNA. This value is sufficiently low compared to the LNA contribution not to dominate the front-end noise figure, thus validating our design choices.

The noise figure of the complete front-end can now be calculated from the noise voltage densities, but since no filter exists between LNA and mixer, the LNA contribution at the image frequency must be considered. Using equation 5.1 and doubling the LNA contribution to consider the image, a noise figure of 6.7dB is calculated (7.6dB if the Spice
simulated LNA noise is used) which is still within the specifications.

5.6.1 Interface to the IF Filter

The IF output of the mixer is high impedance open collector, thus it does not interface directly to the IF filter. In order to properly bias this output, a DC short to VDD must be provided, while a 500Ω output impedance must be realized over a sufficiently large bandwidth around the IF frequency to terminate the IF filter on its characteristic impedance. At the end, a low impedance - ideally a short circuit - must be provided by the load at RF and especially at the frequency of the local oscillator to absorb the large feedthrough of the latter.

Since the mixer has sufficient $g_m$ to drive a 250Ω load while providing the required gain, no impedance transformation is needed and the simplest way to interface the IF filter to the mixer is via a parallel RCL network. The resistor will obviously have a value of 500Ω, while the LC will provide some IF prefiltering and guarantee low impedance at both DC and the local oscillator frequency. A Q around 2–3 can be set to achieve sufficient prefiltering, IF filter matching over a large bandwidth and to avoid the need for alignment. If desired, the IF SAW filter and its RCL termination can be replaced by a suitable LC filter, provided that the requirements on the mixer side are fulfilled.

5.7 The Local Oscillator

The local oscillator must provide a precise 1396.208MHz signal (128 times the reference crystal oscillator frequency) to the mixer, with an amplitude around −3dBu, or 450mV$_{pp}$, for best performance. Since only this frequency must be generated, a relatively small tuning range is required. The latter must nevertheless be sufficiently large to cover all possible tolerances due to normal deviations of component values as well as variations of supply voltage or temperature changes cumulated. Because of the application envisaged, also here power consumption is an obvious major concern.

To define a phase noise specification for this oscillator, a few con-
5.7. The Local Oscillator

Figure 5.11: *Schematic diagram of the local oscillator.*

Considerations must be made. Since, as previously said, phase noise is no critical parameter regarding detectability of the GPS signal itself, no particularly good local oscillator is necessary to achieve a sufficient receiver performance. Fully integrated oscillators such as ring, relaxation, or LC oscillators with integrated low-Q inductor could be used in this application without degrading its performance.

Integrated oscillators, however, tend to consume much more power than those based on an external high-Q resonator, while the local oscillator port of the mixer would become inaccessible from outside. If the latter is a definite advantage in a production part, in a prototype like this having an inaccessible local oscillator port will only complicate testing and debugging.

The local oscillator chosen for our application is therefore the varactor-tuned Clapp oscillator (a variant of a Colpitts oscillator) with external LC tank, whose schematic diagram appears in Fig. 5.11. Transistor \( Q_1 \) is the active part of the oscillator, while \( Q_2 \) to \( Q_4 \) form a 10:1 biasing current mirror. Capacitor \( C_d \) has the important function of decoupling the base of \( Q_2 \) to ground, thus filtering off high frequency noise from the bias circuit and preventing that \( Q_2 \) Miller capacitance forms a negative feedback loop around \( Q_2 \) that would result in decreased impedance at the collector, thus unnecessarily loading the tank.

As it was already the case for the ERMES receiver chip, also here the characteristics of the local oscillator are dictated more by the interfaces to mixer and prescaler rather than phase noise, power consumption or
similar considerations. To save power the oscillator has been connected directly to mixer and prescaler without any buffer, thus proper signal levels, DC bias and impedances must also be provided. The mixer LO port is connected to the base of $Q_1$, which has been biased to VDD as required by the mixer. The divide-by-8 prescaler is AC coupled to $Q_1$ emitter, where the impedance is lowest.

Similarly to the ERMES pager local oscillator, only bipolar transistors were used. MOS transistors were discarded on the basis of the large, lossy stray and gate capacitances. Since the available integrated capacitors were also not the best quality regarding losses and bottom plate capacitances, the complete resonator has also been kept off-chip, although stray inductances at base and emitter of $Q_1$ may become problematic. If better capacitors would have been available, the tank could have benefitted from the integration of at least $C_1$ and $C_2$, while bonding wire inductances would simply have constituted part of $L_1$.

5.7.1 Design of the LC resonator

Since nearly all characteristics of an LC oscillator are determined by the type and quality of the tank circuit, this part must be designed first. Apart from achieving the desired performance, the tank must obviously use only components with practical values and size. Given the good results achieved by the oscillator of the ERMES receiver described in this work regarding stability and frequency pulling despite the direct connection to the mixer, for this receiver a more aggressive design consuming relatively little power has been chosen.

As the most critical component of the tank, the inductor $L_1$ was determined first. Because of the high operating frequency, the use of a transmission line instead of an inductor for $L_1$ could be advantageous and help solve parasitic inductance problems. Because of size considerations the transmission line was rejected and a 10nH inductor was preferred. This value is sufficiently large to keep parasitics under control and to guarantee a self resonance frequency well above 1.4GHz.

The Q of the tank, which is dominated by the Q of $L_1$, determines many oscillator parameters, such as phase noise and especially power consumption, thus an estimate was needed. Datasheets values as well as measurements of a few commercial low cost and hand-wound 10nH
inductors gave a Q around 30–50 at 1.4GHz. The corresponding equivalent series resistance is therefore 1.8 to 2.9Ω.

Assuming that all tank losses are lumped to $L_1$ (which is reasonable), oscillator pulling, stability and power consumption are mainly determined by $C_1$ and $C_2$. By lowering the impedance at $Q_1$ base and emitter, i.e., using a large value for $C_1$ and $C_2$, a better oscillator stability can be achieved, but since oscillation amplitude is still $V_p \approx 1.9I_{DC}/g_{mc}$ with $g_{mc} = \omega^2C_1C_2R_L$, a larger power consumption results for a given signal amplitude. As a trade-off between the two requirements, $C_1 = C_2 = 6\mu$F has been chosen.

The tank circuit is now completely defined and only a suitable varactor diode must be found. With the chosen values a nominal capacitance of 1.3pF is required, which is realized by an hyperabrupt varactor\textsuperscript{3}. The bias current of the oscillator can now be determined. The loss resistance of $L_1$ is as previously said around 1.8 to 2.9Ω. To account for the losses in all capacitors, connections and bonding pads, a total series resistance of 4Ω has been assumed. Stray capacitances (around 1pF) appear at base and emitter of $Q_1$ and must be accounted for. Based on this estimate we can derive that a $Q_1$ bias current of approximately 1mA (900µA has been chosen for practical reasons) is required, which is less than 9% of the total current consumption of the front-end. The size of $Q_1$ (10 emitters) is determined more by ESD considerations rather than oscillator performance.

Again, to guarantee the proper amplitude we cannot rely on the losses of the tank alone, thus some kind of regulation will be required. Thanks to the availability of excellent Schottky diodes in this technology, using a rectifier to derive a regulation voltage will certainly be one of the simplest and most effective ways to obtain amplitude regulation. Alternatively, since phase noise is not too critical here, a simple limiter may do the job. In order to save design time no amplitude regulator has been implemented in this prototype, and the bias current for the oscillator is derived from the PTAT system bias. Overcompensation of temperature dependence of oscillation amplitude will therefore occur, while the oscillation amplitude - which was excessive - has been reduced to an appropriate value by increasing $C_2$.

\textsuperscript{3}Such as e.g. Siemens BBY52 or MA-COM MA4ST551
This is the last circuit which has an external analog connection whose characteristics must be adapted to the requirements of external components, in this case the 1st IF SAW filter. Apart from test points (that will disappear in a final integration), the GPS signal will in fact leave the chip as a digital 1 bit signal. The only real requirement for this circuit is therefore to provide a 500Ω, possibly broadband, input impedance. Although a quite uncritical parameter at this point, the circuit must have a sufficiently low noise figure (around 6–10 dB) to prevent the 1st IF amplifier/mixer from degrading overall noise figure.

The schematic diagram of the 1st IF mixer, preceded by a preamplifier, is shown in Fig. 5.12. To prevent the mixer from contributing excessive noise, a gain of 12 dB has been allocated to the preamp. Such gain can be easily obtained with a single transistor, thus a single-stage common base amplifier has been chosen. By biasing it at 80μA, an input impedance of 300Ω is realized, which is increased to the desired 500Ω by R₁. Apart from providing the proper input impedance, R₁ is required to protect Q₁ against ESD, together with the standard ESD protection of the pad. The small transistor used (2 emitters, as a balance between noise and stray capacitances) will in fact not withstand ESD sufficiently unless extra protection is added. The gain of 12 dB is obtained by setting the load resistor R₅ accordingly, while temperature independence is achieved with a PTAT bias current as usual.
The rest of the circuit is internal and no parameter is forced by external component requirements, thus the design is much simpler, more freedom exists in the choice e.g. of gains or bias currents and a more optimal design can be done. Since the 2nd IF filter is active and could be saturated by excessive local oscillator feedthrough, a double balanced Gilbert mixer is required to achieve a high local oscillator rejection. Mixer design has been already discussed before and is not repeated here. Since noise is no longer a problem here, small one emitter transistors have been used everywhere, while the mixer is biased at 40μA per branch which together with the 200Ω emitter degeneration resistor \( R_6 \) gives a conversion \( g_m \) of 400μS. DC bias of the input transconductor \( Q_8 \) and \( Q_9 \) is provided by the preamp transistor \( Q_1 \), while the local oscillator port receives DC bias and a nearly square signal from the ECL divider, to which the mixer interfaces directly. The output of this mixer is connected to a transimpedance active load which also implements the real pole of the 2nd IF filter. Temperature independence is achieved with a PTAT bias current as before, but since here all components are integrated and match well, a better accuracy is obtained.

## 5.9 The 2nd IF Strip

After conversion to the 2nd IF, the GPS signal must be further filtered and amplified before being converted to digital. The 2nd IF strip must therefore provide:

- Large gain, to guarantee an "usable" signal level at its output.
- Proper low impedance loading of the 1st IF mixer.
- Bandpass characteristic to improve channel filtering, remove local oscillator feedthrough and DC offsets, and as antialiasing for the AD converter.
- Amplitude limiting to a sufficiently well controlled value.

Thanks to the 1bit design no AGC is required, and simple limiting is sufficient. To avoid excessive AM to PM conversion in the limiting stages however, it must not be driven too hard. The gain must therefore be sufficient to guarantee limiting and nothing more.
5.9.1 The 2nd IF Amplifier and Filter

The schematic diagram of the 2nd IF strip is shown in Fig. 5.13. It consists of a transimpedance stage, two 2nd order filter stages and three differential amplifiers, two of them including an amplitude limiter.

Most of the gain of this receiver is realized as usual at the last IF. Since the signal level (actually noise) at the input of the 1st IF amplifier is $-82\text{dBu}$ (Fig. 5.8), a gain around 75–80dB is necessary to reach the limiting level ($\approx 800\text{mV}_{PP}$) of the last amplifier stage. To account for possible gain tolerances of the whole receiver chain, a gain of 86dB (85dB simulated) from the 1st IF input to the 2nd IF output has been allocated. The chosen distribution of the gains among the different stages is 12dB for the 1st IF amp (previous section), 8dB for the mixer and its transimpedance load and 66dB for the 2nd IF amplifier.

The 2nd IF strip begins with the transimpedance active load for the mixer. The structure of this circuit is the same as used in the demodulator for the ERMES pager (Fig. 3.21). Also here a 2nd order filter structure has been used, with 2 real poles (because of the low gain of the differential pair) where the one with the lowest frequency is part of the following 5th order filter. The two 1pF capacitors at the input also provide a passive low impedance path for the local oscillator signal leaking through the mixer. The transimpedance of this stage (6.3kΩ) is chosen to yield the desired gain of 8dB. The output of this stage is then buffered and fed to the first filter stage.

The 2nd IF filter is a 5th order low-pass Butterworth active RC fil-
Figure 5.14: Schematic diagrams of a) transimpedance and b) differential amplifiers (1st and 2nd stages).

The 2nd IF Strip

This consists of two biquad sections implementing the complex pole pairs plus the aforementioned real pole in the transimpedance stage. The cut-off frequency of 6.6MHz has been chosen to obtain sufficient filtering while allowing a 20% margin for component tolerances. AC interstage coupling (3 real poles) has been used to remove offset and low frequency signals, resulting in the filter's characteristic being bandpass. A fully differential structure has been used to minimize pick-up of substrate noise and the potential instability due to parasitic feedback. The amplified and limited output signal is then fed to a latched comparator (not shown in Fig. 5.13) for AD conversion.

In order to gain access to the internal analog IF signal (for test purposes), the output of the 2nd IF strip is buffered and brought to two pads. In the final integration these two test points are no longer necessary and will be deleted. Pad count and power consumption will therefore be reduced. Temperature and process independence are obtained with a PTAT bias current as usual. Thanks to the fully integrated structure and well matched components, good accuracy is achieved.

5.9.2 The Differential Amplifier Stages

The gain of the 2nd IF amplifier is obtained with the differential amplifiers shown in Fig. 5.14. The transimpedance stage is constructed
around the non degenerated differential amplifier of Fig. 5.14a. Since this amplifier must be able to source at least the same current delivered by the mixer if saturation has to be avoided, it has been biased to 80\mu A. The buffers \( Q_3 \) and \( Q_4 \) provide a low impedance (1300\Omega differential) output for the active filter. Unless otherwise noted, minimum transistors (one emitter) have been used everywhere to minimize stray capacitances.

The two amplifiers with a gain of 20dB are implemented with the degenerated differential amplifier shown in Fig. 5.14b. Biased at 40\mu A per branch, the 1st amplifier requires a degeneration resistor of 1.7k\Omega to obtain the desired gain of 20dB, while the 2nd one needs 200\Omega because of the lower load (15k\Omega in place of 30k\Omega). The 2nd amplifier includes also two limiting Schottky diodes. The diodes have been used instead of the natural clipping to obtain a lower limiting level (about 800mVpp) without the need to reduce \( R_5 \) and \( R_6 \) and thus gain.

The 3rd amplifier, whose schematic diagram appears in Fig. 5.15, is similar to the 2nd but has 26dB of gain thanks to the 30k\Omega load. Its output (\( AD_+ \), \( AD_- \)) is DC connected to the latched comparator. The buffers for the IF test points are simple emitter followers biased at 160\mu A each. Their output impedance (160\Omega each) results in negligible losses when these test points are loaded with a typical 1M\Omega, 25pF oscilloscope probe.
5.9.3 The Fully Differential OpAmp

The active 2nd IF filter requires fully differential opamps to implement the biquad sections. These sections have unity gain and relatively low Q, thus an opamp with relatively moderate characteristics is sufficient. Negligible deviations from the ideal filter frequency response are already obtained if the opamp has a gain bandwidth product of a couple 100MHz and a DC gain of 40dB or more. Since the load is resistive a buffered opamp is necessary. Because of the small signal amplitude no particularly critical specifications exist regarding output swing.

The opamp used is the single stage differential amplifier with emitter-follower output buffers shown in Fig. 5.16. This simple structure in fact fits best the requirements of wide bandwidth, low output impedance and low power consumption.

The input differential pair, $Q_1$ and $Q_2$, is biased at $40\mu A$, and thus a $g_m$ of $380\mu S$ is obtained. Transistors with two emitters are used to improve matching somewhat. $C_1$ and $C_2$ frequency compensate the opamp. With the values chosen (300fF plus 50fF stray capacitance) a gain bandwidth product of about 350MHz (290MHz simulated) is obtained. The simple structure with only two poles (compensation and output poles) also means that an excellent phase margin of $75^\circ$ can be achieved.
The output buffers $Q_4$ and $Q_5$ are biased at $40\mu$A, thus an output impedance of $1300\Omega$ (differential) is achieved. The value of the bias current has been chosen to obtain sufficient swing at the given load (15k$\Omega$ resistors) rather than bandwidth. With the small capacitive load that the opamp has to drive, the bandwidth is already a couple GHz.

Transistors $Q_9$, $Q_{10}$ and associated parts form the common mode feedback. The output common mode voltage is set to the voltage applied to the $vcm$ pin, which receives a nominal 1.2V from a bandgap voltage reference. Since the common mode path is much slower than the main opamp, $C_1$ and $C_2$ are not sufficient to frequency compensate also the common mode feedback loop, thus the compensation network $R_5, C_3$ has been added. The common mode amplifier is biased at $20\mu$A, thus a current consumption of $140\mu$A for the complete opamp results.

### 5.9.4 The Latched comparator

After filtering, amplification and amplitude limiting, the 2nd IF signal is converted to digital with a latched comparator (1bit AD converter). The sampling frequency is 3.6MHz, thus the comparator also converts the 2nd IF signal to the 3rd IF of 1.05MHz, which is fed to the digital correlator chip.

The latched comparator, shown in Fig. 5.17, is the conventional structure consisting of a differential amplifier ($M_1$, $M_2$) driving a regenerative latch ($M_{10}$, $M_{11}$ and $M_{12}$). Comparators of this type are widely used because of the excellent precision and sensitivity that can be achieved at low power consumption.

Since the power consumption of this comparator is insignificant compared to the rest of the chip, there was no need to minimize it. The input pair has been biased to $40\mu$A to make it fit the existing bias tree rather than to achieve speed or precision. Both are in fact guaranteed at the chosen bias current. Because of the large physical distance between the comparator and the 1st IF amplifier - that could have lead to significant errors because of insufficient matching - the biasing for the P-channel transistors generated within the 1st IF amplifier and used also for the transimpedance amplifier has not been used here. A bias generated locally has been used instead ($Q_2$, $M_9$ and associated parts).
A possible problem of this comparator type is the possible lack of sufficient voltage to drive the switch transistor ($M_{12}$) gate when the comparator is operating at low supply voltages. Since the minimum specified voltage (2.4V) is not sufficient to guarantee sufficient drive for $M_{12}$, a charge pump has been used to boost the drive signal of the latter above the supply.

The charge pump, consisting of $M_{14}$ (which resides in an isolated well together with $M_{13}$) and $C_1$, operates by charging $C_1$ during the latch phase ($cp1$ high) and supplying the stored voltage to the drive inverter $M_{13}$ and $M_{15}$ during the sample phase ($cp1$ low). The body diode of $M_{14}$ clamps the boosted voltage to one diode drop above VDD to avoid damages to $M_{13}$ and $M_{15}$ at high VDD (3.6V specified, but the chip operates up to 5.5V). For the same reason non minimum gate length has been used for these transistors. With this charge pump in place operation from below 2V to 5.5V is guaranteed even for worst-case process (high $V_{TO}$).

In order to obtain a stable and synchronized output signal for the correlator chip, the digital output of the comparator is buffered with a flip-flop before going to the output pad. The current consumption of the comparator is 90μA including the digital part operating at 3.6MHz but not the pad driver.
5.10 The Frequency Synthesizer

This receiver requires 3 different local oscillator frequencies derived from the same 10.907875MHz crystal oscillator. The 1st local oscillator frequency is 1397.1MHz (128 times the reference), the 2nd one is 174.5MHz (16 times the reference) and the sampling clock is 3.64MHz (1/3 of the reference). Apart from the frequency synthesizer, the reference frequency itself is used as the main clock source for the correlator chip.

The block diagram of the frequency synthesizer is shown in Fig. 5.7. It consists of the already discussed 1.4GHz oscillator which generates the 1st local oscillator frequency, a chain of dividers to derive the 2nd local oscillator and the PLL frequency, a phase and frequency comparator and an amplitude regulated crystal oscillator. The sampling clock is derived from the reference oscillator directly by dividing its frequency by three.

Since only one frequency has to be synthesized, fixed dividers with simple division ratios can be used. To minimize power consumption in the dividers, a receiver plan that uses a divider-by-8 at the highest frequency, i.e. an asynchronous chain of 3 D flip-flops, has been designed. This divider consists of three cascaded bipolar ECL flip-flops. The 2nd divider, also a divider-by-8, uses ESCL flip-flops while the rest of the frequency synthesizer uses standard CMOS logic.
5.10. The Frequency Synthesizer

5.10.1 The Divide-by-8 ECL Prescaler

This divider performs the first division by 8 to generate the 2nd local oscillator frequency. It is connected directly to the 1.4GHz local oscillator, therefore this is the fastest divider in the chain and the most critical one regarding power consumption and switching noise. In order to reduce both an ECL implementation was the obvious choice.

The schematic diagram of the 3 flip-flops used in this divider is shown in Fig. 5.18. Different from the analog part of the receiver walled transistors have been used, especially because of the very low stray capacitances that these transistors have. The smallest available walled transistors, with emitter size of 1x2μm², have been used for this design, and have a collector-substrate capacitance of only 6fF. The poor matching of such transistors is of no consequence in this circuit.

The structure is the classical ECL master-slave D flip-flop. The output swing has been set to 400mVpp as a trade-off between noise margins and power consumption, while the bias current has been scaled to fit the frequency at which every flip-flop operates. A bias current of 200μA, 100μA and 50μA per latch has been set for the 1st, 2nd and 3rd flip-flop respectively by setting the resistors to the appropriate value. The 2nd and 3rd flip-flops are preceded by buffers/level shifters biased at 50μA and 25μA respectively, while the 1st flip-flop has no buffer, is AC coupled to the emitter of the oscillator transistor and is biased at about two diode drops from ground. The schematic diagram of the bias circuit is shown in Fig. 5.19.

The output of this divider is brought off-chip for testing purposes.
with an ECL buffer consisting of a differential pair biased at 400μA and loaded with 500Ω resistors. In the final implementation this output will be connected on-chip to the 2nd mixer and the ESCL divider, thus the buffer will no longer be needed.

### 5.10.2 The ESCL and CMOS dividers

The 174MHz 2nd local oscillator signal as generated by the ECL divider must be further divided by 48 to obtain the PLL frequency. For this purpose a combination of an ESCL divider-by-8 and a CMOS divider-by-6 has been used.

This divider chain begins with an asynchronous ESCL divider consisting of three cascaded D flip-flops, again to save power. ESCL has been used here in place of ECL mainly to reduce area and complexity somewhat. A schematic diagram of these flip-flops is shown in Fig. 5.20. Apart from the use of bipolar current sources to fit the existing bias tree, these flip-flops are similar to the ones used in the ESCL dual modulus prescaler presented in chapter 4. The bias current has been scaled to fit the speed requirements of the first two flip-flops, while because of the negligible power saving possible the 3rd flip-flop has not been scaled and is identical to the 2nd one. Biasing of the P-channel transistors is obtained with a circuit similar to the one shown in Fig. 4.4 (M5 size:
5.10. The Frequency Synthesizer

Figure 5.21: Schematic diagram of the ESCL to CMOS level converter.

10/1.4), while conversion to CMOS levels is performed by the level converter shown in Fig. 5.21, that apart from the use of bipolar transistors is identical to the one shown in Fig. 4.8.

The following divider-by-6 consists of a toggle D flip-flop followed by a divider-by-3. For these functions normal CMOS parts from the standard cell library have been used.

5.10.3 The Phase Comparator

In this application a phase and frequency comparator is obviously necessary, therefore the circuit shown in Fig. 5.22 has been used [49]. It consists of a digital phase/frequency comparator driving a current source output stage. The operation of this circuit is quite self-explanatory, while the input signals \( F_1 \) and \( F_2 \) are the 3.6MHz reference frequency and the (divided) VCO frequency respectively.

The current source output is connected via the loop filter to the VCO’s varactor diode (see also Fig. 5.7). A rail-to-rail current source has been used to exploit the full VDD range for tuning. With the loop filter shown, a PI regulation loop is realized. To save current and reduce the size of the loop filter components, the output current has been set to 10\( \mu \)A. This current is derived from the system bias and has therefore about 20% tolerance and a PTAT characteristic. This must be accounted for when designing the loop filter, to guarantee that the
PLL will remain stable and well behaved over the whole process and temperature range.

Because of the low operating current, the current sources are much slower than the phase comparator and cannot be switched on at the same speed as the latter. This will result in a deadband around the mid point, similar to crossover distortion in class B amplifiers, and leading to poor PLL performance (mainly jitter). To counteract this problem, the phase comparator operating frequency has been reduced to 3.6MHz, i.e. 1/3 the reference frequency, and a 15ns delay line (gate $U_6$) has been added to the reset signal of the latter. A 15ns delay time has been chosen to guarantee that both the Up and Down current sources will always be switched on simultaneously around the mid point at all process and temperature corners, thus removing the deadband completely.

The main part of the phase comparator has been designed using standard cells from the library, while full custom layout has been used for the delay line and the current source.

### 5.11 The Reference Crystal Oscillator

This oscillator generates the 10.907875MHz reference frequency for the PLL and the clock for the digital correlator chip. The precision requirement of this oscillator depends mainly on the characteristics of the correlators, in particular on the algorithm used to search and acquire the satellites.
If a large tolerance is allowed for the crystal oscillator, the search field is widened, which will necessarily lead to long search times unless some countermeasure is taken. To obtain reasonable search times, a high precision TCXO (a couple ppm or so) over the temperature range will therefore be required in such case. A simple and obvious countermeasure would be to correlate over significantly less than 1023 chips during search. A correlation over 100–200 chips will be sufficient - at least for the satellites with the strongest signal - and will speed up search by about 5–10 times.

In our case no particular details were known about the implementation of the digital chip and no precision specification regarding the reference oscillator was given explicitly. Datasheets of a few possible crystals (TEW, NDK, Toyocom) were nevertheless found or delivered by our partner. These were all AT cut fundamental mode crystals for telecom applications, with 2 or 5ppm temperature characteristic (over the range 0–40°C and −10–60°C respectively), requiring 20pF of load capacitance and a drive level of 50 or 100μW.

The equivalent circuit diagram of a crystal, including the fundamental oscillation mode only, and the values specified for one of them (the others were similar anyway) is shown in Fig. 5.23. $L_1$, $C_1$ and $R_1$ are the motional inductance, the motional capacitance and the equivalent series resistance, are mainly determined by mechanical characteristics of the crystal such as cut type, cutting angle, electrode structure or crystal size, and determine uniquely the series resonance of the crystal. $C_0$ is the parallel capacitance of the crystal, and results from the sum of the capacitance of the crystal element and the package stray capacitance. $C_0$, together with the external load capacitance, contributes to the determination of the parallel resonant frequency. In parallel res-
Chapter 5. GPS Receiver Chipset

Figure 5.24: Schematic diagram of the amplitude regulated crystal oscillator.

Onant oscillators it is therefore fundamental to use the specified load capacitance, otherwise the nominal oscillation frequency of the crystal cannot be obtained. A fine adjustment of the latter is required in precision oscillators.

The drive level determines the amplitude of the vibration of the crystal, and corresponds to the active power dissipated by resistor $R_1$. It is chosen by the manufacturer as a trade-off between oscillator performance - which improves at high drive levels - and reliability (aging) - which worsens at high drive levels. The value is given for a typical (bogey) crystal, and is used to derive the required oscillation amplitude (in volts) that the oscillator must approximately generate with all crystals of a given type. Although a good crystal is quite tolerant in this respect, the drive level must be respected if best overall performance is desired.

The design of a crystal oscillator is no different from the design of any LC oscillator. The performance is almost uniquely determined by the quality of the resonator as usual, thus the task of the oscillator - apart from sustaining oscillation - is to provide the correct load capacitance and drive level to the crystal. The schematic diagram of the reference crystal oscillator is shown in Fig. 5.24. It consists of a Pierce oscillator, an amplitude regulator and an output buffer amplifier.
5.11. The Reference Crystal Oscillator

The Pierce oscillator is the most commonly used crystal oscillator structure. It is a parallel resonant 3-point oscillator, and is the variant of a Colpitts oscillator in which the emitter of the transistor and the mid point of the load capacitor are grounded. It has been used here partly because of the lower signal swing that suits well low voltage operation and partly because of the lower stray capacitances, which appear mainly in parallel to the load capacitors $C_{x1}$ and $C_{x2}$.

Since the Pierce oscillator is a parallel resonant oscillator, it is important that the load capacitance is well defined, has the correct value and is stable to obtain the necessary precision. Suitable load capacitors must therefore be chosen, while internal capacitances must be either made small or well defined. Junction and gate capacitances are notoriously voltage and temperature dependent, and must therefore be minimized. At this frequency both MOS and bipolar transistors would be suitable as active element. The latter has been preferred because of the much lower stray capacitances.

The Pierce oscillator is built around $Q_1$ and its biasing components $M_{12}$ and $R_2$. The tank is external and consists of the crystal and the load capacitors. Our crystal requires a load of 20pF, thus the value of $C_{x1}$ and $C_{x2}$ is 40pF. These are actually high stability 39pF NPO (COG) ceramic capacitors, in parallel with about 1pF of stray capacitance. A trimmer can be used to fine tune the load capacitance, thus the oscillator frequency, as shown in Fig. 5.24.

In order to be driven to 50μW, our crystal - whose parameters appear in Fig. 5.23 - requires an oscillation amplitude of about 2.2V_{pp}. Assuming $C_{x1} = C_{x2} = C_x = 40$ pF, the $g_{mc}$ can be calculated as:

$$g_{mc} = R_1\omega_0^2(C_x + 2C_0)^2 = 387\mu S$$

thus a typical bias current of 110μA is necessary. This current will be set to the appropriate value by the amplitude regulator.

The amplitude regulator used in this oscillator is based on [39], which is derived from the amplitude detector presented in [38]. When $M_8$ and $M_9$ are operated in weak inversion, the critical voltage $V_C$, i.e. the regulated oscillation AC peak voltage, is:

$$I_0 \left( \frac{V_C}{nV_T} \right) = A$$
where \( I_0 \) is the 0-order modified Bessel function and \( A \) the width ratio of \( M_8 \) and \( M_9 \). For our regulator \( A = 4 \) has been chosen, thus \( V_C = 109\text{mV} \). The capacitive divider \( C_4 = 0.5\text{pF} \) and \( C_5 = 2\text{pF} \) increases the latter to 550mV which corresponds to the desired drive level of 50\( \mu\text{W} \) (2.2Vpp). Other drive levels can be easily obtained by adjusting the capacitor ratio to a different value.

The start-up current must be limited and adjusted to an appropriate value well above the operating current of 110\( \mu\text{A} \). Resistor \( R_1 \) sets the start-up current of the regulator loop to 15\( \mu\text{A} \), which corresponds to about 500\( \mu\text{A} \) for the oscillator. A buffer amplifier, AC coupled to the oscillator and consisting of transistor \( M_{15} \) and a CMOS inverter, amplifies the sinusoidal oscillator signal to rail-to-rail CMOS levels. This CMOS signal is routed to a pad and is used by the receiver chip as reference for the PLL, as sampling clock, and as clock for the digital correlator chip.

### 5.12 The \( \Delta V_{BE} \) Current Source and Bandgap Reference

The GPS receiver requires various current sources with PTAT characteristic to bias the various blocks, as well as a 1.2V voltage source to set the common mode voltage of the IF strip. A \( \Delta V_{BE} \) current source is
therefore a natural choice. Its schematic diagram is shown in Fig. 5.25.
The $\Delta V_{BE}$ cell is formed by $Q_1$ (4x) and $Q_2$ (1x), and is cascaded by
$Q_4$ and $Q_5$ for better precision. $R_1$ sets the operating current to 10$\mu$A,
which is amplified and mirrored out by $M_4$. Several current sources
are required by the receiver: two with 10$\mu$A (PLL and buffer for the
crystal oscillator), three with 40$\mu$A (IF strip, ECL and ESCL dividers)
and one with 90$\mu$A (RF oscillator) output. These are implemented as
many $M_4$'s scaled as shown in the schematic diagram.

The 1.2V voltage is obtained from a bandgap voltage reference.
Since it connected to the base of transistors only, this source drives
virtually no load, while its absolute value is uncritical since it is only
used to set the common mode voltage of the IF strip. It has therefore
been implemented as an open loop unbuffered reference. In order to
make it independent from the $\Delta V_{BE}$ current source, a separate branch
consisting of $Q_3$ ($V_{BE}$) and $R_3$ ($k\Delta V_{BE}$) has been used. This output is
connected internally to the common mode input of the IF-strip and is
brought to a pad to allow for measurements and external decoupling.

Current sources like this can be designed such that start-up is guar¬
anteed by leakage currents. However, especially at low temperatures
start-up may be slow, while a negative transient on the power supply
may switch off the current source for substantial time (up to seconds).
A start-up circuit has therefore been added. This circuit is formed by
transistors $Q_6$, $Q_7$ and related parts, that inject some current into $Q_5$.
Once the output current approaches its nominal value, this circuit is
shut down by $M_6$.

5.13 The ESD Protection

This technology came with a library containing well designed and tested
ESD protection structures in layout and, where applicable, also in sche¬
matic form. If these structures are used, effective ESD protection is
therefore guaranteed. The library contained the following protection
devices:

- **Collector-substrate diodes**, 2$\mu$m length and 50$\mu$m or 100$\mu$m width.
These diodes are used to protect against negative ESD pulses up to
at least 1kV or 2kV (HBM), and have a stray capacitance of 160fF or 310fF respectively. They are also used to connect together split VSS buses.

- **Collector-base diodes.** These are vertical parasitic PNP transistors with $\beta = 1$ (because of the buried layer) and are available with a length of 2µm and a width of 50µm or 100µm. They are used to protect against positive ESD pulses up to at least 1kV or 2kV (HBM), and have stray capacitances of 220fF and 440fF respectively.

- **ESD clamp.** This is a large 120 x 10µm$^2$ NPN bipolar transistor with base and emitter shorted together. It is used to protect the VDD pads against both positive and negative ESD pulses up to at least 2kV and has a typical breakdown voltage of 19V.

- **Active clamp circuit,** consisting of a large NPN transistor driven by a bandgap voltage reference. Its purpose is the same as the ESD clamp but the breakdown voltage is lower and better controlled.

All signal pads of this GPS receiver are protected with collector-substrate and collector-base diodes. The LNA input and the local oscillator use the 50µm diodes while all other pads, where stray capacitance is less critical, use the 100µm ones.

Since the rugged base-emitter junction of the LNA’s input transistor should already be sufficiently protected against positive ESD pulses, only one diode has been used on this pad, thus reducing parasitic capacitance to the minimum. The emitter of this transistor, which is at ground potential, has its positive ESD protection diode returned to VSS in place of VDD. This is to guarantee that a positive ESD pulse between this pad and the RF input cannot harm the LNA input transistor by reverse biasing its base-emitter junction. The base of the crystal oscillator transistor is also protected by a single diode for the same reasons as the LNA. At the nominal oscillation amplitude the base swings about 200mV negative, thus forward biasing the ESD diode. Given the very small voltage, however, no significant current can flow through it, thus the oscillator is not disturbed by ESD protection.

The bipolar ESD clamp has been used on the VDD bus. Despite its better performance, the active clamp circuit has not been used because of its large size that would have occupied a substantial part of the core.
5.14 Layout Considerations

The photomicrographs of the two GPS chips are shown in Fig. 5.26. The size of these two chips is $1134 \times 982 \mu\text{m}^2$ and $1404 \times 1404 \mu\text{m}^2$ respectively.

In a chip operating at 1.57GHz or where high gain amplifiers must coexist with digital logic, many layout issues may affect the performance, as for example component and pad placement, coupling through the substrate and between parallel lines, etc.

During the placement of the various blocks, attention has been paid to parasitic coupling that could degrade performance or cause oscillation. The LNA and mixer are the block at the top left of the RF chip, the VCO is placed top right, while the pads of LNA and VCO are on opposite sides of the chip. The ECL prescaler has been placed bottom right, with the low speed stages and the output buffer at the bottom margin, as far as possible from LNA input and VCO.

The same kind of placement has been used for the IF chip. The IF-strip, in the middle of the chip, is oriented with its sensitive IF input near to the bias circuit and as far as possible from the digital logic block (top left). Its input pad, on the right edge, is at the opposite side from the digital pads. The crystal oscillator is oriented with the oscillator transistor at the bottom left corner, away from both the IF input and the digital CMOS logic. Because of the balanced low amplitude signals, the
ESCL logic is quite harmless even if it lies near to the crystal oscillator.

Particular attention is paid to layout symmetry of the IF-strip, the ECL and the ESCL dividers, while the lines carrying true and complement have been routed in parallel to minimize signal injection into the substrate and maximize common mode rejection. Large substrate contacts and guard rings separate all stages. Since the substrate is high ohmic, guard rings are an effective means to prevent parasitic coupling. More than 30pF of on-chip decoupling capacitors have been placed between the supply lines, to prevent coupling through them. Most bias lines, especially in the VCO and IF-strip, are also decoupled with on-chip capacitors. The bias circuit, located bottom left for the RF chip and bottom right for the IF chip, is identical for both chips and has its own supply decoupling capacitor. Its outputs are connected to pads and decoupled externally to ground with low ESL ceramic SMD capacitors.

Thick metal2 or metal1-metal2 lines have been used for the ground connections, while metal1 has been used for VDD. The RF connections are mainly metal2 and have been kept short wherever possible, routed away from other sensitive lines and partially shielded with metal1.

All analog signal pads on both chips are placed between ground pads or pads at AC ground (vdd, bias, etc.) to provide some shielding and prevent spurious signal coupling. Metal layers 1 and 2 have been used in most pads, while the pads carrying high frequency signals are metal2.
5.15 Chip Mounting and Testing

In this application, due to its stringent space constraints, chip-on-board is the only practical mounting technique for the same reasons as for the ERMS pager chip. For these measurements the same technique has therefore been used. The naked dies have been directly bonded to small PCBs which contain all the necessary external components. The schematic diagram of the test circuit for the RF front-end is shown in Fig. 5.27, while a drawing of these PCBs, with component placement, appears in fig. 5.28 and 5.29. The PCBs, measuring 1in x 1in, have one signal plane and one ground plane on opposite sides and are mounted on a test fixture that holds PCB and connectors in place. The base material is 0.8mm thick FR4.

Gain and input impedance of the LNA are both determined by the inductance of the bonding wire connecting the emitter of the input
transistor to ground, thus some experimentation was needed to find the proper way to make that connection. The best performance was achieved using a single bonding wire, slightly longer than 1mm, that despite the short length also gave reproducible results on all (working) chips measured. The inductor for the VCO has been constructed with a loop of copper wire.

The chips have been tested under typical conditions, and the supply voltage was set to its nominal value of 3V. The bias currents were generated by the internal current source. Although no temperature and supply voltage independent biasing has been implemented in this prototype for LNA and mixer, their performances were nevertheless verified over the full specified supply voltage range of 2.4V to 3.6V. For all measurements the local oscillator signal was generated by the internal VCO.

5.16 Measurements

Nine RF front-ends and two IF-strips were measured. The yield for the RF front-end was quite low, as one chip had very poor performance while three did not work at all. The non working chips consumed about 3mA less than normal, while for the one with poor performance the current consumption was somewhat erratic. The reason for this low yield was not searched. As can be seen in Fig. 5.26 the metal2 layer is rather rough and uneven, thus processing problems have been assumed as a possible cause. The five working chips had consistent performance to
5.16. Measurements

within a few percents, showing that the low yield is not caused by design problems. The IF-strips had smooth metallizations and no processing problems were found. The following measurements were made on a single chip per type chosen at random.

5.16.1 $\Delta V_{BE}$ Current Source and Bandgap Voltage Reference

When open loop circuits like in this receiver are used, it is the important task of the biasing circuit to define and maintain constant many receiver parameters such as gains and I/O impedances. The biasing circuit has therefore been characterized first. The measured data for one $40\mu A$, the $90\mu A$ and the voltage reference outputs at a temperature of approximately $20^\circ C$ appears below. In the graph the curves have been normalized to $40\mu A (\Diamond)$, $90\mu A (\square)$ and $1.2V (\bigtriangleup)$. No $10\mu A$ output is externally available for measurement.

<table>
<thead>
<tr>
<th>VDD</th>
<th>Ib1</th>
<th>Ib2</th>
<th>Vbg</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4V</td>
<td>42.2\mu A</td>
<td>94.5\mu A</td>
<td>1.1828V</td>
</tr>
<tr>
<td>2.6V</td>
<td>42.5\mu A</td>
<td>95.2\mu A</td>
<td>1.1870V</td>
</tr>
<tr>
<td>2.8V</td>
<td>42.8\mu A</td>
<td>95.8\mu A</td>
<td>1.1909V</td>
</tr>
<tr>
<td>3.0V</td>
<td>43.0\mu A</td>
<td>96.4\mu A</td>
<td>1.1945V</td>
</tr>
<tr>
<td>3.2V</td>
<td>43.3\mu A</td>
<td>97.0\mu A</td>
<td>1.1977V</td>
</tr>
<tr>
<td>3.4V</td>
<td>43.6\mu A</td>
<td>97.6\mu A</td>
<td>1.2008V</td>
</tr>
<tr>
<td>3.6V</td>
<td>43.8\mu A</td>
<td>98.2\mu A</td>
<td>1.2038V</td>
</tr>
</tbody>
</table>

The output current is about 7% higher than nominal, which is well within the process tolerances. Despite the very simple design the voltage reference is only about 0.5% apart from the nominal value. Due to the open loop design, the line regulation of the latter cannot be good. It is nevertheless well within the requirements of the receiver. Both the current source and the voltage reference start operating at about 1.9V.

5.16.2 Front-End Gain and Linearity

Due to the internal connection, no measurement can be performed on the LNA or the mixer alone. The entire front-end has therefore been
measured as one unit. For this measurements the IF output was connected to the 50Ω input of the measuring instrument via a 3:1 transformer. The load of the RF mixer, together with the 560Ω matching resistor mounted on the PCB, has therefore the nominal value of 250Ω. As local oscillator both the internal VCO and an external generator were used. The following measurement shows the gain versus input signal power, measured using the internal VCO. The numbers have been scaled to show the actual voltage gain of the front-end.

The nominal gains of the LNA and the mixer are 16.7dB and 10dB respectively, which give a voltage gain for the complete front-end of 26.7dB. The actual measurement is very close to the nominal value. The acceptable local oscillator range is −15dBm to +3dBm. At these two extremes the gain of the front-end decreases by 1dB. The input referred 1dB compression point is about −28dBm. Despite the lack of supply independent biasing for the front-end, the performance over the whole 2.4V to 3.6V voltage range has been measured. This measurement appears below.
5.16. Measurements

The gain variation of 7dB versus supply voltage is large, but in the absence of any independent biasing must obviously be expected. It is nevertheless not representative of a final, complete implementation. Thanks to the large gain at the 2nd IF, sufficient receiver performance can nevertheless be guaranteed, thus - although not optimal - the chip can be used as is at least for testing purposes.

5.16.3 Front-End (LNA) input Impedance

The reflection coefficient of the input of the front-end (LNA) has been measured over the band from 1GHz to 2GHz. The $S_{11}$ is $-10.6\text{dB}$ at 1575.42MHz and $-8.4\text{dB}$ at the image frequency of 1217MHz. Sufficiently good matching to the RF filter is therefore guaranteed over a large bandwidth without any external components. Since no image filter exists between LNA and mixer, good matching at the image frequency is required to guarantee proper RF filter performance and to maintain a sufficient noise figure. The $S_{11}$ (log magnitude) measurement appears in the following figure.

![Graph showing $S_{11}$ log magnitude measurement.](image)

5.16.4 RF Front-End Noise Figure

This measurement has been made over a bandwidth of 1575.42±10MHz, limited by the (low Q) LC IF filter. No image filter has been used at the input of the front-end, thus a DSB measurement has been done. The
latter has been scaled to obtain the SSB value, assuming equal front-end gains at the input and at the image frequency (+3dB).

SSB Noise figure: F=8.1dB

The measured value is close to the predicted 7.6dB (Spice simulation), and because of the wide bandwidth of the front-end and its (bonding wire) matching network does not vary significantly over the 20MHz bandwidth of the measurement. The slightly larger measured value may be due to losses on the PCB, the bonding wires (matching network) and normal measurement uncertainty.

5.16.5 VCO Phase Noise

Since the local oscillator has no buffered output, the downconverted noise spectrum of the free running oscillator has been measured at IF (150MHz for simplicity). A clean PLL synthesized RF generator has been used as input signal. Its output power has been adjusted to obtain a power at the IF output of −20dBm. The resolution of this measurement is therefore somewhat limited by the noise of LNA and mixer.

![Phase Noise Graph]

The phase noise at 100kHz offset is −95.1dBc/Hz, which fully satisfies our application. Spurious signals possibly originating from the prescaler were searched by varying the frequency of the RF generator accordingly. Despite the unbuffered connection to the prescaler however,
no such spurious signals have been noticed, while oscillator performance was not affected substantially by switching the prescaler on and off.

5.16.6 ECL Prescaler Maximum Frequency

Since the local oscillator was not fast enough to measure the maximum operating frequency of the prescaler, the self oscillating frequency with the oscillator stopped has been measured, in addition to the regular divide-by-8 operation at the local oscillator frequency. Self oscillation with no input signal is a normal behavior of any ECL divider. In fact with no input signal all transistors from $Q_1$ to $Q_8$ (see Fig. 5.18) are biased and form, together with the stray capacitances, a relaxation oscillator.

Self oscillating frequency: 1728MHz

The maximum operating frequency is expected to be about 10..20% higher.

5.16.7 Front-End Current Consumption

The current consumption has been measured on the 5 good front-end chips at the nominal supply voltage of 3V, while for one of them the measurement has been done over the 2.4V to 3.6V range. The measurements are shown in the following table.

<table>
<thead>
<tr>
<th>Chip No.</th>
<th>Supply current</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (2.4V)</td>
<td>8.1mA</td>
</tr>
<tr>
<td>1 (3.0V)</td>
<td>10.50mA</td>
</tr>
<tr>
<td>1 (3.6V)</td>
<td>12.8mA</td>
</tr>
<tr>
<td>2</td>
<td>10.53mA</td>
</tr>
<tr>
<td>3</td>
<td>10.41mA</td>
</tr>
<tr>
<td>4</td>
<td>10.21mA</td>
</tr>
<tr>
<td>5</td>
<td>10.11mA</td>
</tr>
</tbody>
</table>
5.16.8 IF-Strip Frequency Response and Noise

The IF-strip is fully integrated, thus the only externally accessible ports are the 1st IF input and the filtered, amplified and limited 2nd IF output. In this measurement the mid-band gain and the frequency response from the 500Ω 1st IF input to the differential 2nd IF output is given. The noise figure has been derived from noise voltage density measurements. In the following curve the measured frequency response is shown along with the simulated one.

This measurement is in good agreement with the simulation at frequencies below 10MHz. At higher frequency an excessively good filter attenuation is measured. This is due to a 7MHz parasitic pole at the output of the 2nd IF amplifier, caused by its output impedance (2 x 163Ω), the protection resistors (2 x 820Ω in series with the outputs) and the capacitances of the oscilloscope probes (25pF). Since the precise frequency of this pole cannot be measured in our test fixture no attempt has been made to compensate this error, and the raw measurement data is presented instead.

5.16.9 Crystal Oscillator Amplitude

The peak-to-peak oscillation amplitude of the reference crystal oscillator versus the supply voltage has been measured. Since the specified crystal (a custom part) was not yet available, a standard 12.288MHz crystal for digital clock applications has been used instead. Thanks to the amplitude regulation the oscillator accepts such crystal without problems. The measurement has been done with an oscilloscope, thus
the absolute accuracy is no better than 5% or so. The following table shows the peak-to-peak signal voltage between the collector of the oscillator transistor and ground. The oscillation amplitude across the crystal is about \textit{twice} the value shown in the table.

<table>
<thead>
<tr>
<th>Supply voltage</th>
<th>Osc. amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>2V</td>
<td>1.04V_{PP}</td>
</tr>
<tr>
<td>3V</td>
<td>1.10V_{PP}</td>
</tr>
<tr>
<td>5V</td>
<td>1.12V_{PP}</td>
</tr>
</tbody>
</table>

The oscillation amplitude has the correct value, and is regulated quite well despite the simple regulator structure. For this measurement the supply voltage had to be varied over a range larger than nominal. In the 2.4V to 3.6V range in fact, the variation of the oscillator amplitude was too small to be measured with sufficient accuracy.

\section*{5.16.10 IF-Strip Current Consumption}

The current consumption has been measured on two IF-strips. The value given here is the worst value between the two chips measured. For this measurement the IF input was terminated to a 50Ω resistor but received no signal, all outputs where floating while the local oscillator port received a 174.5MHz, 400mV_{PP} signal. The crystal oscillator was operating with the aforementioned 12.288MHz crystal.

Supply current: 1.81mA

A summary of the main characteristics of this GSM receiver chipset appears in table 5.2.
### RF front-end:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chip size</td>
<td>$1134 \times 982 \mu m^2$</td>
</tr>
<tr>
<td>Voltage gain</td>
<td>26.5 dB</td>
</tr>
<tr>
<td>Input impedance</td>
<td>500 $\Omega$</td>
</tr>
<tr>
<td>SSB noise figure</td>
<td>8.1 dB</td>
</tr>
<tr>
<td>1dB compression point</td>
<td>$-28 \text{dBm}$</td>
</tr>
<tr>
<td>VCO phase noise (100kHz offset)</td>
<td>$-95.1 \text{dBc/Hz}$</td>
</tr>
<tr>
<td>Supply current (3V)</td>
<td>10.5 mA</td>
</tr>
</tbody>
</table>

### IF-strip:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chip size</td>
<td>$1404 \times 1404 \mu m^2$</td>
</tr>
<tr>
<td>Overall voltage gain</td>
<td>84.7 dB</td>
</tr>
<tr>
<td>Input impedance</td>
<td>500 $\Omega$</td>
</tr>
<tr>
<td>SSB noise figure</td>
<td>8.8 dB</td>
</tr>
<tr>
<td>Limiting level</td>
<td>6dBu</td>
</tr>
<tr>
<td>2nd IF filter cut-off frequency</td>
<td>6.6 MHz</td>
</tr>
<tr>
<td>Crystal oscillator amplitude</td>
<td>$2.2 V_{pp}$ (50$\mu W$)</td>
</tr>
<tr>
<td>Supply current (3V)</td>
<td>1.81 mA</td>
</tr>
</tbody>
</table>

Table 5.2: Main characteristics of the GPS receiver chipset.
Chapter 6

Conclusion

This thesis discusses the design and implementation, as integrated circuits, of two RF receivers intended for wrist-watch applications. The first circuit described is a front-end for ERMES pagers. Since pagers can hardly compete with cellular phones in terms of performance, market penetration must be achieved by offering a cheaper or more convenient product. With convenience in mind, the wrist-watch was identified as an attractive base for the application, provided that the receiver performs reliably and the watch is of regular size. The second circuit described is a receiver for the GPS L1 band. In this case, the immediate application of this receiver is to provide GPS time reference to the watch which will automatically self-adjust to the correct date and time and match the appropriate time zone worldwide.

Although the RF performance required by these applications is not very high in absolute terms, achieving this kind of performance from a receiver small enough to fit in a watch whilst incorporating its antenna is quite challenging. A given minimum performance had to be guaranteed primarily to satisfy the user who obviously expects his receiver to perform reliably, but in the case of the ERMES pager, to also ensure that type approval requirements were satisfied. Maintaining sufficiently low power consumption in accordance with the capacity of a typical (small) button watch battery was a particularly demanding requirement from both system and circuit design points of view.
It is clear that to implement such a combination of conflicting characteristics, a "standard cell" design approach - i.e. an approach where the system and the various circuits are designed independently from each other - is guaranteed to fail. Therefore, a more conventional top-down design style has been used in this thesis. The RF receiver has always been seen as a complete system, spanning at least from the antenna to the base-band processing and never as separate circuits. Firstly, careful planning was required. Possible alternatives and optimizations at system level were studied. The various circuits were then designed as a single unit, with the focus more on their interfaces than on the circuits themselves. Where problems arose, modifications were made without hesitation, until a good receiver design was achieved.

In this thesis a top-down systematic design flow, together with the trade-offs made, is outlined. The main steps consisted of:

1. An understanding of the system and its requirements, so that clear specifications could be drafted.

2. The design of the system, as a function of the given application with its needs and constraints (and not vice-versa), such as:
   - Dimensions, packaging and chip mounting;
   - Power consumption, battery size;
   - External components, (in particular) filters;
   - Characteristics of the IC technology.

3. The design of the actual circuits only after full system specifications are stated.

Since circuit design offers the highest degree of freedom and adaptability with existing external parts, it has been considered the least of our priorities.

With operating frequencies falling well within the parameters of the chosen technologies, a somewhat different design style which is closer to base-band low frequency rather than to classical RF circuits could be adopted with the view of achieving better integration.

The circuits were designed to achieve the performance required by the given application. Their selection was not influenced by any desire
to be particularly innovative or the like. Predictable and repeatable performance, independence from temperature, supply voltage and process tolerances obviously represented one of our main priorities. Indeed, any circuits lacking the above characteristics are utterly unusable.

The receiver circuits, as designed, offered predictable performances and fulfilled all expectations by a sufficient margin. The RF front-end for the ERMES receiver has about 32dB of gain and a linear AGC. The SSB noise figure is 6.2dB (impedance matched) - which improves to about 3.7-4dB when matched to the loop antenna - for the direct conversion receiver or if the image filter between LNA and mixer is used. The current consumption is as little as 230μA from 3V supply.

A complete pager, consuming about 1.2–2.2mA (12–22μA at the average ON-OFF ratio of 100:1), could be built using these circuits. With the proper choice of external components and assembly technique, such a pager could fit with a small battery into a watch case of regular size and allow continuous operation for up to a couple of months, as shown in table 6.1.

The GPS receiver also performs well, requires few external components and is compatible with the intended watch application. However, because of the higher operating frequency, the GPS receiver chipset will obviously consume much more than the pager. With the previous estimate of 40mA for the complete GPS watch, the device can operate autonomously and continuously for a fairly respectable total of 4 hours from a small battery, as shown in table 6.1. ON/OFF operation could improve operating times further. A duty cycle of 10:1 or so can be readily implemented, while still providing one GPS update per second,

<table>
<thead>
<tr>
<th>Battery (IEC)</th>
<th>Life (approx.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERMES Pager (22μA)</td>
<td></td>
</tr>
<tr>
<td>CR1216, 25mAh</td>
<td>47 days</td>
</tr>
<tr>
<td>CR1620, 60mAh</td>
<td>110 days</td>
</tr>
<tr>
<td>GPS Receiver (40mA)</td>
<td></td>
</tr>
<tr>
<td>CR2025, 170mAh</td>
<td>4.2 hrs</td>
</tr>
<tr>
<td>CR2032, 230mAh</td>
<td>5.7 hrs</td>
</tr>
</tbody>
</table>

Table 6.1: Approximate battery life for various battery types.
which is more than sufficient for a watch application. A full 40 hours operation would be the result. If the GPS is used rarely or as a time reference only, i.e one minute per day or so, 250 days of operation can be achieved from the smaller of the proposed batteries, even without any ON/OFF operation.

The technologies used in this work were production 1–1.2μm technologies, available around 1994–1996 at low cost. This work has therefore demonstrated that wrist-watch radios with acceptable performance and sufficiently long battery life are not complex gadgets from the future but were already feasible in the early/mid 1990's. It is unfortunate that our partner had no interest in the ERMES watch/pager, which was therefore never implemented as a product. Now it is probably too late: today, pagers are virtually obsolete in most developed countries, having been supplanted by cellular telephones. This trend is sure to continue in the rest of the world. While a niche professional market will certainly persist for some time to come, such a market may be too small and specific to provide the impetus for new developments, with the risk that it will eventually fade away as we see today with the consumer market.

On the other hand, the GPS watch is a far more likely candidate for commercial introduction. According to recent market analyses [50] the GPS market is growing fast and steadily, and will soon be dominated by car navigation and consumer applications, each totalling about 25% of the GPS market by the early 2000s. A bright future can be forecast for applications such as GPS watches. A GPS watch has already appeared on the market [51], although - given its size and weight (65.5 x 66.6 x 29.6 mm$^3$, 138g) - may still be slightly inconvenient. Apart from GPS watches, an interesting and surely bigger market requiring small, low power GPS receivers will be for cellular phones, where GPS can be used for both navigation and pinpointing one's location.

In the years to come, many more small form factor wireless consumer applications will appear on the market. Be they for watches, smart cards or the like. All such applications will operate with low power, be extremely compact and cost very little. Although the RF front-end will only make up a small fraction of such digital wireless systems, it will always be critical to achieving the aforementioned goals and a great deal of know-how must be drawn upon during the development process. With this work, we hope that although there is still a long way to go, Dick Tracy's two-way wrist radio is now one step closer.
Appendix A

Interference Resilience

The interference resilience of a receiver is usually described as intermodulation, compression and blocking. In this appendix, these characteristics are briefly discussed and the mathematical calculation for the case of a bipolar transistor amplifier is stated. This can obviously be applied also to mixers, differential amplifiers and the like and can be easily adapted to any amplifying device. This information (once widely drawn upon) seems to appear quite rarely today in modern books and papers. It is given here for completeness.

A.1 Intermodulation

Intermodulation occurs when at least 2 signals at different frequencies are applied to a non linear network (the receiver), and is the production of sum and difference frequencies of various orders. The second and third orders are generally the strongest and are therefore the most likely to cause trouble. Although it exists at any signal level and is strongest with large signals, intermodulation describes low level non-linearities.

In radio receivers, the 3rd order or IM3 intermodulation is most likely to interfere with the received signal. This is best shown in fig. A.1, which depicts the location of the IM2 and IM3 products in the simple
Appendix A. Interference Resilience

Case of two interfering signals \( f_1 \) and \( f_2 \). These two signals, represented here as two unmodulated signals for simplicity, can be either out-of-band signals or two in-band channels serving other users. IM3 will produce two new signals of equal amplitude, lying at the frequencies \( 2f_1 - f_2 \) and \( 2f_2 - f_1 \), and carrying a mixture (the convolution) of the modulations spectra of the signals \( f_1 \) and \( f_2 \) - thus occupying a bandwidth equal to twice the spectrum of one interferer plus the spectrum of the other. The channels (within a given service) being usually equally spaced, one of the two IM3 products may fall exactly on the desired channel (represented here by a small modulation spectrum).

Second order intermodulation, IM2, will produce two new signals lying at \( f_1 + f_2 \) and \( f_1 - f_2 \) (fig. A.1). They too have equal amplitude and will carry a modulation spectrum mixture whose width is the sum of the two individual spectrum widths. IM2 is less troublesome than IM3, since for most receivers either at least one of the two interferers or IM2 itself will lie out of band and get filtered away by the RF filters in the front-end. One exception is the direct conversion receiver. As previously discussed, 2nd order non-linearities will cause DC offset and AM detection.

Intermodulation products of higher orders are much lower in amplitude and as such, they will not interfere with the received channel as much as IM2 and IM3. Even order intermodulation generates relatively harmless products away from the interfering signals, while odd order IM generates products near to the interferers that may fall on the desired channel.

By plotting the powers of the output signals and the IM products
A.1. Intermodulation

on a log-log chart, a graph similar to that in fig. A.2 is obtained. Since the amplitude of the \( n^{th} \) intermodulation product is proportional to the input amplitude raised to the \( n^{th} \) power \( P_{IMn} = c_n P_{in}^n \), with \( c_n \) a proportionality constant) the slope of the \( n^{th} \) IM curve is \( n \). When extended, they will therefore intercept the fundamental. These intercept points (IPn) are the most commonly given intermodulation figures, and usually refer to the input of the device. Since IPn expresses a low level non-linearity, a measurement must be done at a relatively low input level, while the slope must be forced to the proper value \( n \) and best fitted to the points as measured. In a well executed measurement, the measured points will show very little deviation from the ideal slope.

Intermodulation distortion originates in the active devices of a receiver because of their non-linearities. Assuming a non-degenerated common emitter bipolar amplifier stage (other configurations or devices can calculated in a similar way), the input referred IP3 (iIP3) can be calculated by expressing the I-V relationship for the collector current of the BJT:

\[
I_c = I_s \cdot \exp\left(\frac{V_{be}}{V_t}\right)
\]
and the transconductance:

\[ g_m = \frac{\partial I_c}{\partial V_{be}} = \frac{I_{c0}}{V_t} \]

with \( I_{c0} \) the collector DC bias current. Since we are considering low level non-linearities, no bias shift will occur in the amplifier, and \( I_{c0} \) can be considered constant and independent from the input signal. The total output current of the transistor (expanded in power series) becomes:

\[ I_c = I_{c0} + g_m \cdot V_{be} + \frac{1}{2!} \cdot \frac{\partial g_m}{\partial V_{be}} \cdot V_{be}^2 + \frac{1}{3!} \cdot \frac{\partial^2 g_m}{\partial V_{be}^2} \cdot V_{be}^3 + \ldots \]

When two signals of equal amplitude \( A/\sqrt{2} \) and frequencies \( \omega_1 \) and \( \omega_2 \) are applied to the transistor, i.e. \( V_{be} = A(\sin(\omega_1 t) + \sin(\omega_2 t))/\sqrt{2} \), the 3rd order intermodulation becomes:

\[ I_{c3} = \frac{3}{4} \cdot \frac{1}{6} \cdot \left( \frac{A}{\sqrt{2}} \right)^3 \cdot \frac{I_{c0}}{V_t^3} \cos((\omega_1 \pm 2\omega_2)t) \]

The response at the fundamental for a single signal having the same power as the two tone signal used to calculate IM - i.e. a signal with peak amplitude \( A \) - and frequency \( \omega_0 \) is:

\[ I_{c1} = g_m \cdot A \sin(\omega_0 t) = A \frac{I_{c0}}{V_t} \cdot \sin(\omega_0 t) \]

The IP3 occurs when the IM3 intercepts the fundamental, that is:

\[ I_{c3} = I_{c1} \]

\[ \frac{1}{8} \cdot \left( \frac{A}{\sqrt{2}} \right)^3 \cdot \frac{I_{c0}}{V_t^3} = A \frac{I_{c0}}{V_t} \]

\[ A^2 = 8 \cdot 2^{3/2} \cdot V_t^2 \]

Since at ambient temperature \( V_t \) is 26mV, the IP3 becomes:

\[ A = 87.5mV = -8.2dBu \]

independent from the collector current \( I_{c0} \).
A.2. Gain compression, blocking

Likewise, the IP2 can be calculated by taking the IM2 term and equating it with the response for the fundamental:

\[ I_{c2} = I_{c1} \]

\[ I_{c2} = \frac{1}{2} \cdot \left( \frac{A}{\sqrt{2}} \right)^2 \cdot \frac{I_{c0}}{V_t^2} \cdot \cos((\omega_1 \pm \omega_2)t) \]

\[ \frac{1}{2} \cdot \left( \frac{A}{\sqrt{2}} \right)^2 \cdot \frac{I_{c0}}{V_t^2} = A \cdot \frac{I_{c0}}{V_t} \]

\[ A = 4V_t = -9.7dBu \]

again independent from \( I_{c0} \).

A.2 Gain compression, blocking

Gain compression and blocking are related to high level non-linearities. The most common way to describe gain compression is to use the 1dB compression point, which corresponds to the signal power that causes a 1dB gain reduction in the DUT. It is usually referred to the input, as shown in fig. A.2.

Gain compression occurs because of overload effects in the amplifier stages. In BJT amplifiers, the most common compression mechanism is related to signal rectification due to device non-linearities (bias shift), but clipping at the output may be a likely cause in amplifiers using more linear devices (e.g. MOS transistors), negative feedback, or biased in class AB.

A direct consequence of gain compression is receiver desensitization or blocking. It describes the loss of sensitivity that occurs when a strong interferer is applied to the receiver along with the desired signal. The performance parameter considered is usually a 1 or 3dB attenuation of the desired signal, a 3dB decrease of SNR or (equally) the fulfillment of a specified SINAD or BER for a desired signal which is 3dB larger than the sensitivity of the receiver.
To get an estimate of gain compression and blocking, it is no longer sufficient to set up the equation for the collector or drain current of the active device, as this will not take into account bias shift. In many cases, particularly if the bias current is signal dependent or when dealing with multistage amplifiers whose stages compress more or less simultaneously, a direct calculation of compression and blocking may become unduly complicated. Provided that:

- The amplifier does compress,
- Compression is caused by the same mechanism as intermodulation, i.e. (3rd order) device non-linearities,

compression and blocking can be calculated by expanding the amplifier transfer function in power series [34], regardless of amplifier topology:

\[ V_{out} = a_1 \cdot V_{in} + a_2 \cdot V_{in}^2 + a_3 \cdot V_{in}^3 + ... \]

with the small-signal gain of the amplifier \( a_1 > 0 \) and \( a_3 < 0 \) to realize compression. Gain compression due to 3rd order non-linearities can then be calculated by setting \( V_{in} = A_1 \sin(\omega_1 t) + A_2 \sin(\omega_2 t) \). Simple algebra allow us to calculate the apparent gain of the amplifier, which becomes:

\[ a'_1 = a_1 \cdot (1 + \frac{3a_3}{2a_1} A_2^2) \]

The coefficients \( a_1 \) and \( a_3 \) can be derived from IP3:

\[ IP3 = 1.155 \sqrt{\frac{a_1}{|a_3|}} \]

which is set equal to the IP3 calculated using device equations. The 1dB compression point then becomes:

\[ CP = 0.383 \sqrt{\frac{a_1}{|a_3|}} \]

while the amplitude of the interferer which causes 1dB blocking becomes:

\[ A_2 = 0.269 \sqrt{\frac{a_1}{|a_3|}} \]
A.2. Gain compression, blocking

The 1dB compression point is therefore 9.6dB below the IP3.

If compression does not originate from the same mechanism as intermodulation, (e.g. compression due to clipping at the output), this calculation is no longer valid and the difference between CP and IP3 can easily be much more or less than 9.6dB. In a differential pair, signal rectification at the emitter nodes (which causes unequal signal distribution between the two transistors) reduces the compression point by about 3dB, that is 12.6dB below the IP3. Very linear amplifiers, particularly those using global negative feedback, have a IP3 that is much higher (many 10dBs) than the compression point. In these cases the reason for gain compression is almost exclusively output clipping. In class AB amplifiers, gain expansion rather than compression is observed. If sufficiently large signals are applied, however, compression will eventually appear, again because of some form of clipping. In these cases the compression point tends to be closer to, or even higher than the IP3. A sharply rising IM3 curve for large signals (slope $\gg 3$, see Fig. A.2) is a clear sign that compression and low level intermodulation are not generated by the same mechanism, and that the most probable cause of compression is clipping.
Bibliography


[21] GPS RF front-end chips, such as Motorola MRFIC1502, Mitel GP2010, GP2015, NEC UPB1004GS, or Sony CXA1951AQ.

[22] GPS digital correlators, such as Motorola MCS38140, Mitel GP1020, GP2021, Philips SC1575 or Sony CXD2930BR.
Bibliography


[29] *Information from SWISSPHONE Telecommunications, Switzerland (internal colloquium).*


[48] –, *MAX2640/MAX2641, 400MHz to 2500MHz SiGe Ultra-Low-Noise Amplifiers*, Maxim Integrated Products.


Curriculum Vitae

Francesco Piazza
Born 5th January, 1962 in Locarno
Citizen of Olivone, Canton Ticino


June 1987  Matura Type C

Oct. 1987 - May 1992  Student at the Electrical Engineering department of the Swiss Federal Institute of Technology (ETH) Zurich

May 1992  Degree in Electrical Engineering from ETH Zurich (dipl. El. Ing. ETH)

May 1992 - Dec. 1999  Ph.D. student and research assistant at the Analogue Design group of the Integrated Systems Laboratory, headed by Prof. Dr. Q. Huang. Worked on Ph.D. thesis and various RF and high speed communications ICs.
Publications

Semester/Diploma Thesis:


Conference papers:


Journal Papers:


