DISS. ETH NO. 15234

Frequency Synthesis for Wireless Transceivers

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Technical Sciences

presented by

DIRK PFAFF

Dipl. Ing. ETH
born 4. 4. 1970
citizen of Liestal, Baselland

accepted on the recommendation of

Prof. Dr. Qiuting Huang, examiner
Dr. Roland Best, co-examiner

2003

Content

Abstract

Frequency synthesizers are indispensable to any wireless transceiver to accommodate frequency translation from the radio frequency (RF) band to baseband and vice versa. This work describes analysis, integrated circuit implementation and experimental characterization of frequency synthesizers for GSM-type mobile stations. Both architectural and circuit level issues are considered. After a review of GSM/DCS/PCS frequency synthesizer specifications, which reveal the stiff requirements on accuracy and purity of the synthesized signal as well as on the dynamic behavior, the systematic design of phase locked loops is presented. Particular attention receives the optimum choice of loop filter singularities of a third order loop. Besides this, undesired effects of charge pump phase locked loops, such as the generation of spurious tones in the spectrum of the synthesized signal, are considered.

A large fraction of the presented work is dedicated to circuit design. Low power consumption, which is among the most important performance parameters of wireless terminals where the available power is limited, is targeted besides low cost. All proposed circuits are realized with cost effective standard CMOS technologies, instead of more suited, but more expensive bipolar technologies. Low power consumption is achieved by rigorous re-thinking of the building blocks which dominate the consumption. Application of discrete resonator devices, which are either surface mounted devices in the 1GHz range or printed circuit board based planar transmission lines in the 4GHz range, are found to be extremely useful to lower the consumption of RF oscillators. The design of a low power 1GHz voltage controlled oscillator using only 0.25mA is presented. A 3.6GHz oscillator providing quadrature outputs at half the frequency, accommodating receivers with low or zero intermediate frequency, is presented as well. A quadrature demodulator including highly linear downconversion mixers exhibits excellent 40dB of unwanted sideband rejection despite the low consumption of only 10mA.

Besides the oscillator, programmable dividers tend to consume considerable amount of power due to the high input frequency. While these blocks are often

realized with bipolar technologies to benefit from the large transconductance to current ratio, the design of CMOS programmable dividers is more demanding.

Nevertheless, competitive CMOS dual modulus prescalers are presented which even outperform bipolar solutions. While a $0.25\mu m$ 1GHz prescaler dissipates 0.9mA, a $0.18\mu m$ 4GHz prescaler consumes 2.5mA. Beyond experimental results, some low power design guidelines for current mode logic circuits are proposed.

Frequency synthesizer design involves circuits with operating frequencies ranging from several GHz to some hundred kHz. While the design of the high frequency part is in performance and consumption critical, the low frequency section is challenging as well. Reduction of spurious tones requires careful design of the charge pump. A novel charge pump topology is proposed to lower the spurious tones level of integer-N frequency synthesizers. The capability of the novel charge pump is verified by a 4GHz frequency synthesizer realized in $0.18\mu m$ CMOS. The third order loop with a loop bandwidth of 40kHz exhibits a reference frequency spurious tone, 400kHz away from the carrier, of only -68dBc.

Zusammenfassung

Frequenzgeneratoren sind unverzichtbare Komponenten eines jeden drahtlosen Sende- und Empfängersystems um die Uebersetzung vom Radio- zum Basisband vorzunehmen. Diese Arbeit behandelt die Analyse, den Entwurf und die experimentelle Charakterisierung von Integrierten Schaltungen für die Frequenzsynthese in batteriebetriebenen GSM Endgeräten. Sowohl Architektur- als auch Schaltungsbelange werden behandelt. Nach einer Uebersicht der durch den GSM Standard auferlegten harten Anforderungen an die Genauigkeit und spektrale Reinheit der erzeugten Frequenzen, sowie den Anforderungen an das Einschwingverhalten, wird der systematische Entwurf von Phasen-Regelkreisen behandelt. Grosse Aufmerksamkeit wird der optimalen Wahl der Filter-Singularitäten eines Regelkreises dritter Ordnung gewidmet. Des weiteren werden Effekte wie die unerwünschte Entstehung von diskreten Tönen im Spektrum der erzeugten Signale untersucht.

Ein grosser Teil der Arbeit befasst sich mit dem Entwurf der Integrierten Schaltungen. Minimale Leistungsaufnahme, eine absolute Notwendigkeit in batteriebetriebenen Geräten, wo die verfügbare Leistung begrenzt ist, sowie tiefe Herstellungskosten werden angestrebt. Alle vorgestellten Schaltungen sind daher in kostengünstiger CMOS Technik realisiert, obwohl bipolare Technologie geeigneter, aber auch teurer wäre. Geringe Leistungsaufnahme wird erzielt durch gründliches Ueberdenken sämtlicher Baublöcke, die wesentlich zur Leistungsaufnahme beitragen. Die Anwendung von diskreten Schwingkreisen mit entweder oberflächenbestückten Komponenten im 1GHz Bereich oder angepassten Leitungen im 4GHz Bereich, hat sich als sehr nützlich erwiesen, den Verbrauch von Hochfrequenz-Oszillatoren markant zu reduzieren. Ein 1GHz spannungsgesteuerter Oszillator mit einer Aufnahme von lediglich 0.25mA wird vorgestellt. Ein 3.6GHz Oszillator mit Quadratur-Ausgängen für Empfänger mit tiefer Zwischenfrequenz ist des weiteren beschrieben. Ein vollständiger Quadratur-Demodulator, der hochlineare Mischer enthält, zeigt eine ausgezeichnete Unterdrückung des unerwünschten Bandes von 40dB, obwohl lediglich 10mA verbraucht werden.

Neben den Oszillatoren nehmen die programmierbaren Teiler aufgrund der hohen Frequenz merklich Leistung auf. Daher werden diese Blöcke oft in bipolarer Technologie realisiert, welche sich durch ein vorteilhaftes Transkonduktanz zu Strom Verhältnis auszeichnet. Die Realisierung in CMOS Technik ist anspruchsvoller und tendiert mehr Leistung aufzunehmen.

Trotzdem kann diese Arbeit Vorteiler in CMOS mit attraktiver Leistungsaufnahme vorstellen, welche den Verbrauch von bipolaren Lösungen sogar unterbieten. Ein 0.25µm 1GHz Vorteiler benötigt lediglich 0.9mA, ein 0.18µm 4GHz Vorteiler 2.5mA. Neben experimentellen Resultaten wird eine Entwurfsprozedur für tiefe Leistungsaufnahme vorgeschlagen.

Frequenzgeneratoren beinhalten Schaltungen mit Betriebsfrequenzen von einigen GHz bis zu einigen hundert kHz. Der Entwurf der kHz Baublöcke ist auch entscheidend für die Qualität der erzeugten Signale. Eine Reduktion der unerwünschten diskreten Töne verlangt einen sehr sorgfältigen Entwurf der Ladungspumpe. Eine neuartige Schaltung wird vorgeschlagen, welche die diskreten Töne eines Integer-N Frequenzgenerators nachhaltig reduziert. Um das Konzept zu verifizieren, wurde ein 4GHz Frequenzgenerator in 0.18µm CMOS Technik realisiert. Der Phasen-Regelkreis dritter Ordnung mit einer Bandbreite von 40kHz zeigt einen Referenzfrequenz-Ton von lediglich -68dBc, welcher 400kHz vom Träger entfernt auftritt.

Chapter 1

Introduction

The tremendous growth of wireless systems all over the word, with hundreds of millions units sold per year, is driven on the one hand by the consumer demand for personal communication everywhere and at any time, and also by the overwhelming improvement of the underlying technology. Indeed, it is hard to imagine that near 100% penetration of the population with cellular phones, which is a reality today within a growing number of countries, would have been possible with the first working cellular phone presented by Motorola in 1973. Despicably called the "brick", the prototype cellular phone measured 25cm, weighed more than 1kg and provided a too short standby and talk time to be successful on the market. Fig. 1.1 shows this bulky device, which hardly resembles today's wireless gadgets. The first commercial handheld cellular phone, available 10 years later, still deserved the same blemishing nick-name. Until now ongoing process of product enhancement towards lower cost, weight, size and extended autonomy and functionality empowered the unprecedented success of personal wireless communications.

Enabling factor for these achievements is the progress of integrated circuit (IC) technology, providing more and faster transistors at lower price. While very large scale integration (VLSI) technology provided the required signal processing power to replace analog by more advantageous digital communication, the integration of the radio frequency (RF) part, responsible for signal conditioning, turned out to be more delicate than expected. The reason for this lies in the nature of RF circuits which rely widely on passive components which are incompatible with VLSI technology. Great effort was invested into monolithic solutions, however, the highest level of integration does not necessarily represent the optimum solution with respect to cost, size and weight. Integration of passive devices is often paid with a higher power consumption which must be compensated by a larger battery capacity to maintain autonomy. Since the battery contributes significantly to the overall cost, weight and size of the cellular handset [1], highly inte-
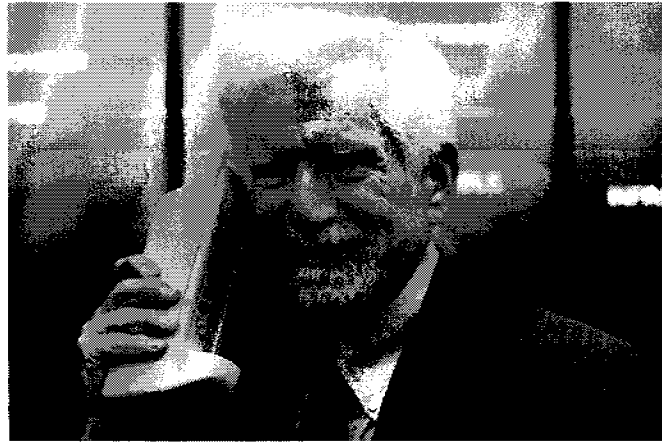
Figure 1.1: The "brick", the world's first handheld cellular phone presented by Motorola in 1973.

grated solutions which employ a limited number of passive components where performance is critical, may be more economical.

The frequency synthesizer is an indispensable and critical part of a cellular handset's RF sections. Stiff analog performance, high frequency and incompatibility with IC technology of certain building blocks, combined with the limited available power and cost efficiency are the most eminent challenges of any frequency synthesizer design. This thesis explores the various design aspects of frequency synthesizers targeted to GSM and WCDMA.

1.1. Direct and Indirect Frequency Synthesis

Division of the RF-band into channels of constant width commands the synthesis of a large number of equally spaced frequencies in order to enable up- and down-conversion from the baseband to the RF channel and vice versa respectively. Direct synthesis by means of oscillators at the channel frequency is not practical due to the enormous number of required oscillators. An oscillator can accommodate more than one single channel frequency by mixing with other oscillator frequencies, however, a still considerable number of oscillators is required to synthesize all channel frequencies of a cellular service. On top of this, costly filters are required to suppress unwanted mixing products.

Quartz crystals are mandatory to meet the stiff requirements on frequency purity and stability. They are likely to remain the predominant source of stable, low noise frequency source for many years to come. Direct frequency synthesis by

quart crystal banks combined with mixers, an approach of frequency synthesis used in early mobile radio systems [2], is both too bulky and expensive, since crystals cannot be integrated. Cost considerations require to keep the number of used crystal oscillators at an absolute minimum, preferably at a single unit. A virtually unlimited number of channel frequencies can be realized with a single reference crystal oscillator and a tunable RF oscillator embedded in a negative feedback control loop which forces the tunable RF oscillator to a digitally programmable multiple of the reference frequency. This indirect frequency synthesis technique is the underlying working principle of virtually any RF frequency synthesizer, independent on whether the synthesizer is used in cellular transceivers, TV receivers, wireless entry systems etc.

The negative feedback loop compares most often the phase of the tunable RF oscillator with the phase of the reference oscillator, rather than their frequencies. While one would expect a frequency control loop rather than a phase control loop, phase comparison is beneficial to the accuracy of the synthesized frequencies. A practical control loop provides a finite DC-gain, leading to a finite error, a finite steady-state error phase in case of phase comparison. Since the output frequency is the derivative of the output phase, the steady-state output frequency is flawless despite the finite gain of the control loop. A residual frequency error would result from a frequency control loop. Hence, a phase control loop, referred as phase locked loop (PLL), is mostly always favored.

Indirect frequency by means of a PLL represents the most powerful frequency synthesis concept, since it combines digital programmability with a minimum number of analog circuitry. A couple of different loop architectures have been established. The two most promising ones, the integer-N PLL and the Sigma-Delta fractional-N PLL, are reviewed in Section 1.2.

1.2. Indirect Synthesis by Phase Locking

Indirect frequency synthesis by phase locking can be accomplished by a system shown in Fig. 1.2. Besides the two oscillators, a phase detector, a frequency divider, both digital circuits, and an analog lowpass filter complete the frequency synthesizer. Phase locking commands the output frequency to the reference frequency multiplied by the divider's division ratio. Since the latter is limited to integer values, the spacing of the synthesized frequencies is identical to the reference frequency. Fine frequency resolution requires therefore a low reference fre-
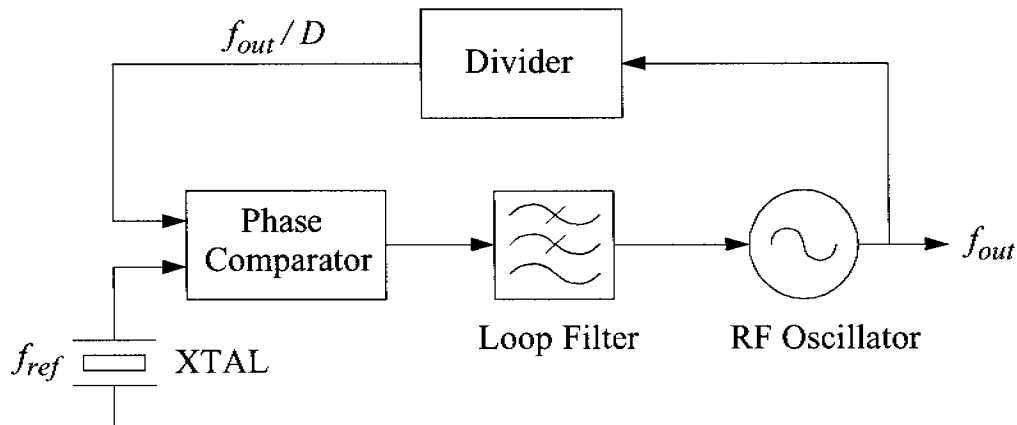
Figure 1.2: A principal schematic of a PLL frequency synthesizer.

quency. It is further necessary to limit the loop bandwidth, as a rule of thumb, to one tenth of the reference frequency, in order to keep the loop stable. Since the dynamic behavior of the synthesizer is bound by the loop bandwidth, the switching time, i.e. the time required to establish an alternate output frequency, is large to frequency synthesizers with a fine frequency resolution. This is the fundamental limitation of the system according to Fig. 1.2, which is referred as integer-N frequency synthesizer to express that the output is an integer multiple of the reference.

Numerous topologies have been proposed to overcome the bandwidth to resolution limitation. Some of them are based on a multi-loop approach [3]. While a wideband, fast switching loop provides fast settling and a coarse frequency resolution, a second loop refines the resolution. The main drawback of multi-loop topologies lies in the increased complexity, especially in the need for a second tunable oscillator. The most promising approach to break the resolution to bandwidth limitation of the integer-N synthesizer lies in fractional-N technique. Instead of adding a second loop, a single loop is employed with a modified frequency divider. The latter provides two division ratios D and $D + 1$. A digital control logic switches between the two ratios, so that an average division ratio $D + F$, with F being a fractional number between 0 and 1, is established. Fig. 1.3 shows the principal schematic of fractional-N frequency synthesizer. Since the frequency resolution is not limited by the reference frequency anymore, a large reference frequency combined with a wideband loop can be used. The faster switching time can be realized by a moderate circuit overhead, which is limited
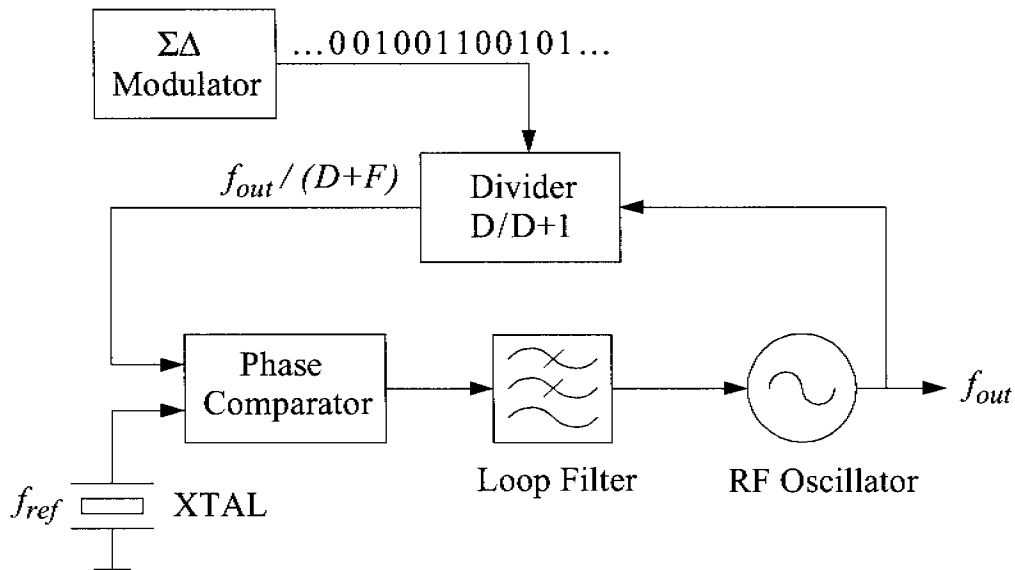
Figure 1.3: A fractional-N frequency synthesizer with Sigma-Delta modulator.

to the digital part of the synthesizer, causing thereby only moderate increase of cost. However, this is only half of the story. The instantaneous deviation of the division ratio from the average value causes imperfections of the spectral purity of the synthesized frequency. If these deviations are periodic, which would be the case if the divider is switched periodically between D and $D+1$, strong spurious tones are generated in the spectrum of the synthesized frequency. To remove these spurious tones, the divider control signal can be obtained from a Sigma-Delta modulator, which efficiently randomizes the division error and moves the error power to higher frequencies due to the noise shaping characteristic of the modulator [4]. While the spurious tones can be removed by the help of the Sigma-Delta modulator, the error power, although randomized, is still existent, causing an increase of synthesizer phase noise. To cope with this problem, the loop bandwidth must be lowered to suppress the additional noise. This, however, annihilates the inherent advantage of fractional-N. Besides the noise problematic, it was observed, that the Sigma-Delta modulator's randomizing ability fails at division ratios close to an integer division ratio, giving rise to spurious tones.

1.3. Trends and Achievements in Synthesizer Design

The last few years have seen an increasing activity in the field of RF frequency synthesizer design, expressed also by a respectable number of publications [5]-[10]. Tremendous effort has been spent to realize highly integrated synthesizers, leading to monolithic synthesizers, which are even combined with the cellular transceiver in single integrated circuit [10]. Single chip cellular solutions may be seen within the next years.

Modern monolithic RF frequency synthesizers use either integer-N, fractional-N or dual-loop architectures as can be seen from Table 1.1 which lists some selected designs. The last two or three years has seen a strong trend towards fractional-N with Sigma-Delta error randomizers. While it was speculated a few years ago that Sigma-Delta fractional-N synthesizers will provide wideband frequency synthesizers with frequency switching times of only 30-40µs, today's realized fractional-N synthesizers exhibit the same loop bandwidth as integer-N synthesizers and share hence the same switching time of about 300µs. Reason for the large reference frequency to loop bandwidth ratio of fractional-N, which is found around 200, lies in the additional noise caused by the Sigma-Delta modulator which deserves strong filtering to avoid an increase of synthesizer phase noise. Remaining advantages of fractional-N synthesis are the smaller division ratio, which is beneficial to inband phase noise performance, and the randomization of phase errors, which helps to reduce spurious tones. However, it is often overseen, that integer-N synthesizers represent inherently low spurs systems. Residual spurious tones, caused by artifacts of the employed building blocks, can be minimized by advanced circuit technique. The superiority of the fractional-N concept over integer-N, which is often tacitly presumed, is hence more than questionable.

Since the handset's standby time is determined by the receiver and the RF synthesizer, power consumption is doubtlessly a key figure of an RF frequency synthesizer. Synthesizers targeting GSM/DCS/PCS tend to consume around 100mW (compare with Table 1.1). Significant progress stays away despite more advanced IC technologies. The steadied power consumption level is mainly caused by the low quality of integrated passive devices, which stayed rather constant over the years. Reconsideration of the application of external passive devices may be necessary to reduce the power consumption noticeably. Reduction of frequency synthesizer power consumption must be one of the main objectives of future designs.

	Architecture:	Carrier:	Phase Noise at 600kHz offset:	Consumption:
Ali [5]	Integer-N	0.9GHz	-116dBc/Hz	50mW
Parker [6]	Integer-N	1.6GHz	-115dBc/Hz	90mW
Craninckx [7]	Fractional-N	1.8GHz	-123dBc/Hz	51mW
Khan [8]	Dual Loop	1.8GHz	-112dBc/Hz	95mW
De Muer [9]	Fractional-N	1.8GHz	-120dBc/Hz	70mW
Magoon [10]	Fractional-N	1.2GHz	-124dBc/Hz	~100mW

Table 1.1: Recently published RF frequency synthesizers.

A recently reported GSM transceiver consumes in receive mode only 20mA [11]. Future synthesizers must target the same level of consumption.

1.4. About this Thesis

This thesis explores the various design aspects of RF frequency synthesizers for cellular applications with focus on solutions with ultimate low power consumption, rather than on the highest integration level. Low cost is targeted as well, circuit implementations rely hence on cost effective standard CMOS technologies. Particular attention is paid to the power optimization of circuits operated at the highest frequency, which are the RF oscillator and the frequency divider. These blocks dominate due to the high frequency, which is several orders of magnitude larger than the frequency of the remaining synthesizer circuits, the power consumption and so deserve careful optimization. While frequency divider optimization is tackled by the introduction of a rigorous design approach, the oscillator needs a complete re-evaluation to reduce the consumption significantly. Printed circuit board microstrip lines are found to be extremely useful to overcome the limitations of integrated passive devices. Tuned microstrip lines are analyzed at length, revealing interesting properties which open the way to low power multi-GHz oscillators. Low cost is considered again by choice of low cost printed circuit board materials used in consumer electronics.

Besides contributions to low power, high frequency circuit design, the integer-N frequency architecture is reviewed. The question, whether the simple integer-N

architecture is qualified to meet the stringent requirements of GSM/DCS/PCS cellular standard is asked. It was found, that the level of spurious tones created by the integer-N topology is critical. The main mechanisms responsible for the introduction of spurious tones are identified and the analog performance of the involved blocks, required to meet the specifications, is quantified. It is shown, that the level of spurious tones reacts very sensitively to phase detector nonidealities. A novel charge pump topology is proposed to overcome some limitations of existing solutions.

Several integrated circuits have been realized to validate the proposed concepts. Among these are a very low power dual modulus prescaler which divides a 1GHz input by 64 and 65. Very low power consumption of only 0.9mA at 2.5V supply voltage was achieved. An ultra-low power 0.25mA, 900MHz voltage controlled oscillator has been presented which meets the stringent GSM phase noise requirements. The tremendous reduction of power consumption by roughly a factor 10 compared to fully integrated solutions, was achieved by careful design and the use of a few external devices.

Low IF and zero-IF receivers require separation of the input signal in an I- and Q-path at RF. An RF oscillator with quadrature phases becomes indispensable. Doubling the oscillator frequency and generation of quadrature phases by a digital divider was considered in the past inefficient in terms of power. This approach, however, has been reconsidered and lead to a 1.8GHz quadrature demodulator driven by a 3.6GHz voltage controlled oscillator. The consumption of the demodulator, which has been realized in 0.18μm CMOS, is only 10mA. The proposed downconversion mixers exhibit excellent linearity and decent noise performance. The measured accuracy of the quadrature phases is excellent despite the very low power consumption.

Finally, an integer-N frequency synthesizer has been implemented with the proposed novel charge pump topology. Despite the low reference frequency to loop bandwidth ratio of 10, the reference frequency spurious tone is measured as -68dBc, which is not only low enough the meet GSM/DCS/PCS specifications with a decent margin, but represents also a very low spurs level compared to other integer-N synthesizer designs. The proposed 4GHz synthesizer is accompanied by a low-power 64/65 dual modulus prescaler with a current consumption of only 2.5mA.

The thesis is organized in the following way. Chapter 2 derives frequency synthesizer requirements from GSM/DCS/PCS type approval tests. Upcoming third generation WCDMA standards, which are less demanding to frequency synthesizers, are considered as well. Chapter 3 analyzes the integer-N synthesizer architecture. Optimum choice of the loop filter, noise and phase detector sampling effects are treated. On top of this, the impact of PLL building block imperfections on the level of spurious tones, such as phase detector dead-zone and charge pump mismatch, is investigated. The following Chapters describe the circuit design of synthesizer blocks: Chapter 4 covers frequency dividers, Chapter 5 RF mixers, Chapter 6 RF oscillators. The complete frequency synthesizer is presented in Chapter 7 and Chapter 8 summarizes the achievements and draws some final conclusions.

References

[1] S. Rogerson, *"Where less is more"*, Mobile Europe, Vol. 7, pp. 27-30, Feb. 1997

[2] M. Underhill, *"The Frequency Control in Radio - From Quartz to Direct Digital Synthesis"*, Proc. IEE 100 Years of Radio, 1995, pp. 167-176

[3] C. Vaucher, *"Synthesizer Architectures"*, Proc. Advances in Analog Circuit Design, 1997

[4] T. Ridley, M. Copeland, T. Kwasniewski, *"Delta-Sigma Modulation in Fractional-N Frequency Synthesis"*, IEEE J. of Solid State Circuits", Vol. 28, May 1993, pp. 553-559

[5] A. Ali, J. Tham, *"A 900MHz Frequency Synthesizer with Integrated LC Voltage-Controlled Oscillator"*, Proc. IEEE International Solid State Circuit Circuit Conference, 1996, pp. 390-391

[6] J. Parker, D. Ray, *"A 1.6-GHz CMOS PLL with On-Chip Loop Filter"*, J. IEEE J. of Solid State Circuits, Vol. 33, No. 3, March 1998

[7] J. Craninckx, M. Steyaert, *"A Fully Integrated CMOS DCS-1800 Frequency Synthesizer"*, IEEE J. of Solid State Circuits, Vol. 33, No. 12, Dec. 1998

[8] T. Kan, G. Leung, H. Luong, *"A 2-V 1.8-GHz Fully Integrated CMOS Dual-Loop Frequency Synthesizer"*, IEEE J. of Solid State Circuits, Vol. 37, No. 8, Aug. 2002

[9]   B. De Muer, M. Steyaert, *"A CMOS Monolithic ΣΔ-Controlled Frac-tional-N Frequency Synthesizer for DCS-1800"*, IEEE J. of Solid State Cir-cuits, Vol. 37. No. 7, July 2002

[10] R. Magoon, A. Molnar, J. Zachan, G. Hatcher, W. Rhee, *"A Single-Chip Quad-Band (850/900/1800/1900MHz) Direct Conversion GSM/GPRS RF Transceiver with Integrated VCOs and Fractional-N Synthesizer"*, IEEE J. of Solid State Circuits, Vol. 37, No 12, Dec. 2002

[11] P. Orsatti, F. Piazza, Q. Huang, *"A 20-mA-receive, 55-mA transmit, sin-gle-chip GSM transceiver in 0.25μm CMOS"*, IEEE J. of Solid State Cir-cuits, Vol. 34, No. 12, Dec. 1999, pp. 1869-1880

Chapter 2

Synthesizer Requirements

A frequency synthesizer as seen from an architectural transceiver view is simply the building block providing a periodic local oscillator signal with programmable frequency, enabling up- and downconversion. However, all kinds of frequency synthesizers rely on analog circuits, which introduce due to their analog nature various kinds of nonidealities. A real synthesizer differs hence significantly from the simplified architectural view.

A first limitation lies in the programmability of the synthesizer. Not any, but only discrete, equally spaced frequencies can be synthesized. This limitation implies fortunately no further issues because modern wireless systems are channelized, i.e. the allocated band is divided into channels of identical bandwidth. More critical is the fact, that the absolute frequency of the synthesized steady state signal differs from the wanted frequency by a residual error which is dependant on component aging and temperature.

A frequency synthesizer will be in general reprogrammed to different frequencies during operation of the transceiver, e.g. when the systems hops to another channel or if the transceiver switches from receive to transmit mode or vice versa. Instead of reacting immediately to a frequency change command, all phase locked loop based synthesizers need some time until the new frequency is established at the output. Direct digital synthesizers (DDS) [1] do not exhibit a switching delay, but this class of synthesizers is limited to frequencies in the MHz range. Synthesizer switching time is a crucial performance parameter, because the transceiver is stopped from transmitting and receiving information during the switching transient.

As every analog circuit, a frequency synthesizer is affected by noise. Various noise sources influence the quality of the synthesized signal. Most critical impact of noise are fluctuations of the instantaneous frequency. The spectral purity of the synthesizer output may be further degraded by spurious tones around the carrier.

It is not surprising that the contamination of the synthesizer output spectrum influences the performance of the transceiver both in the transmit and the receive path. The impact of phase noise at the LO port of the transmitter is twofold. First, synthesizer phase noise is superposed to the upconverted phase modulated signal, degrading thereby the information to be transmitted. Second, a baseband signal is not only upconverted by the mixer, but its spectrum is spread by the noisy local oscillator signal. The up-converted signal may become so wide, that it leaks into the adjacent channels. Blocking of any communication in the adjacent channel may result. At the receiver side, a noisy LO determines the ability of the receiver to withstand blocking signals. Further, a reduction of the received signal quality is caused by superposed phase noise.

This brief overview on frequency synthesizer nonidealities demonstrates, that frequency synthesizer performance is critical to transceiver performance although the synthesizer is not directly part of the signal path. Synthesizer performance with the most prominent parameters being frequency accuracy, switching time, spectral purity and spurious tones, needs to be evaluated carefully. This Chapter identifies the requirements for wireless handset frequency synthesizers targeted to GSM and WCMDA standards. The investigations are focussed on GSM/PCS/DCS standards since those lead to the more stringent requirements. This is a more than prudent strategy since handsets dedicated to the upcoming third generation WCDMA standards will likely be multi-standard devices covering second and third generation standards.

GSM and WCDMA type approval test cases provide a mix of direct synthesizer specifications, e.g. spurious tones suppression specified by GSM standard, and test cases which formulate synthesizer specifications indirectly, e.g. synthesizer phase noise must be derived from blocking test cases. Other test cases involve synthesizer performance as well as nonidealities of transceiver blocks, e.g. residual phase error of the transmitted signal is constituted of synthesizer phase noise, baseband circuitry and RF front-end phase errors. The synthesizer requirements must hence be jointly formulated with transceiver requirements.

If not explicitly stated otherwise, all synthesizer requirements apply to the GSM standard defined by the *European Communication Standards Institute* (ETSI). The frequency synthesis relevant specifications are taken from standardization documents [2], [3], [4]. Frequency synthesizer requirements for third generation WCDMA wireless handsets are derived from the standard proposed by the *Third*

Generation Partnership Project known by the acronym 3GGP. The relevant specifications are taken from [5].

2.1. Frequency Resolution and Frequency Accuracy

Synchronization of the mobile station with the base transceiver station requires an accuracy of the mobile station clock timing and RF frequencies within 0.1 parts per million (ppm). Clock timing and RF frequencies must be derived from a single quartz crystal oscillator which has to provide the mentioned frequency accuracy. Although quartz crystals are known for their exceptional frequency stability, their resonance frequency varies slightly with temperature and aging [6]. Variations over the temperature range have been significantly reduced by the introduction of the AT cut quartz, resulting in a cubic frequency versus temperature function. Excellent 35ppm frequency variation can be guaranteed over an industrial temperature range with AT cut crystals, but the required extremely small tolerance of 0.1ppm is definitely beyond what can be expected from quartz manufactures. One approach to tackle this problem is to tune the reference crystal oscillator during synchronization of the mobile with the base station. A CMOS crystal oscillator tunable by capacitor banks is presented in [7]. It was demonstrated that a quartz crystal oscillator can be fine tuned to the required accuracy of 0.1ppm.

Besides the need for frequency accuracy, the synthesizer must provide a reasonable fine resolution of programmable output frequencies. The minimum resolution is bound by the channel spacing of 200kHz in case of GSM. Hence, the frequency resolution of a GSM frequency synthesizer must be of the same 200kHz. Channel spacing of WCDMA is much wider due to the code division multiple access scheme requiring spreading of the baseband signal and is specified as 5MHz. Despite the larger channel spacing, a frequency resolution of 200kHz is specified also for WCDMA systems to provide maximum flexibility in frequency allocation within the available RF band.

2.2. Switching Speed

GSM is a half-duplex communication system. This implies that the transceiver operates either in transmit or receive mode, but never in transmit and receive at the same time. Besides being a half-duplex system, GSM separates transmission

and reception not only in time but also in frequency. In order to relax the receiver's blocking requirements, the GSM frequency band is divided into a lower transmit band and a higher receive band with a gap in between of a couple of MHz. The half-duplex operation of the transceiver allows the use of a single frequency synthesizer for the generation of both the transmit and the receive LO signal. Most GSM handsets employ therefore a single frequency synthesizer, since besides being attractive in terms of cost and power consumption, mutual pulling of two independent oscillators can be avoided. However, the single synthesizer must switch periodically between transmit and receive frequencies. Required switching speed is determined by the applied time division multiple access (TDMA) and frequency division multiple access (FDMA) schemes. Switching speed requirements are more precisely presented in Section 2.2.1. Section 2.2.2 outlines the more demanding switching requirements for high-speed data services embedded in the GSM standard.

2.2.1. Switching Speed Requirement for Standard GSM

A physical GSM channel consists of a receive and a transmit frequency pair (called RF channel) whereby up to eight users share the same RF channel. This is made possible by splitting the time axis in frames of approximately 4.62ms length which themselves are split into eight time slots of approximately 577μs length. A mobile handset occupies during each time frame a single transmit and a single receive slot when connected to the base transceiver station. Fig. 2.1 outlines the described timing.

The same Figure shows that transmit and receive time slot are separated by two free time slots or time slots allocated by other users. Switching from receive to transmit frequency must be performed within these two slots. The frequency step is common to all RF channels and measures 45MHz for GSM and 90MHz for PCS. The frequency must have settled at the end of the available approximately 1.1mS within 90Hz in case of GSM and 180Hz in case of PCS systems.

Unfortunately, the transceiver must not only transmit during a single time slot and receive during another one, it must further monitor the signal strength of the neighboring base transceiver station of the cellular network to prepare the handover to other base stations. Again, Fig. 2.1 outlines the timing. The switching from transmit to monitor frequency is further aggravated by the fact that the frequency step may be larger than the one from receive to transmit. The worst case
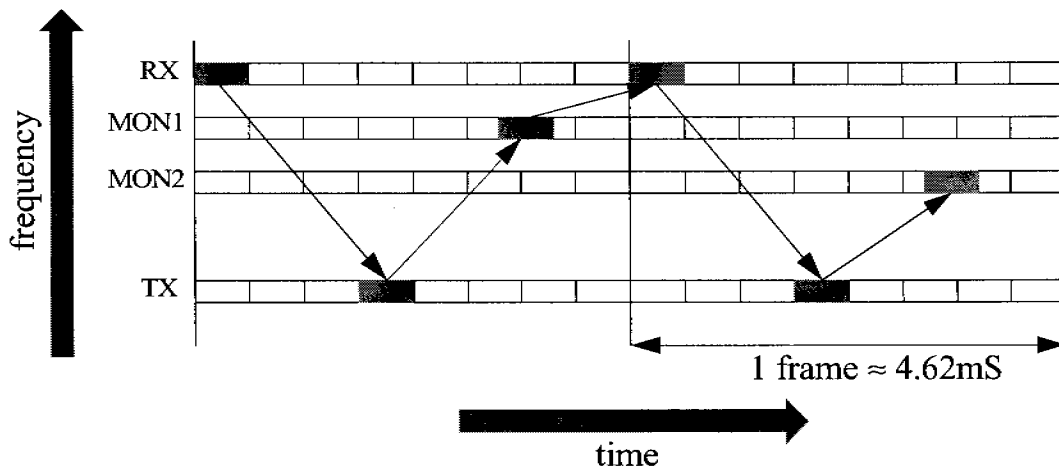
Figure 2.1: Receive, transmit and monitoring mode switching scheme of a standard GSM communication link.

scenario requires switching from a transmit frequency at the low transmit band to a monitor frequency at the upper edge of the receive band. Fortunately, the required frequency accuracy for monitoring is below the required accuracy for reception. Nevertheless, monitoring the neighboring cells causes a slightly more stringent switching behavior. It can be shown that the available switching time is reduced to 1.5 time slots or to 866μs with the other parameters unchanged [8].

2.2.2. Switching Speed Requirement for GSM High Speed Data Services

Standard GSM provides a maximum date rate of 14.4kb/s. While sufficient for the transmission of audio information with decent quality, an increasing market demand for high-speed data communications has been established over the last few years. Third generation WCDMA cellular will serve this purpose, but network operators are confronted by the immense cost to set up the infrastructure for the new service. Modifications of the predominant GSM system (and its higher frequency replicas PCS and DCS) providing higher data rate, but allowing reuse of existing infrastructure is highly desirable. The introduction of *High-Speed Circuit Switched Data* with acronym HSCSD and *General Packed Radio Service* with acronym GPRS must be seen in this context.

HSCSD and GPRS break with the convention of receiving and transmitting in only one time slot per frame. Instead, more than one time slot can be allocated to
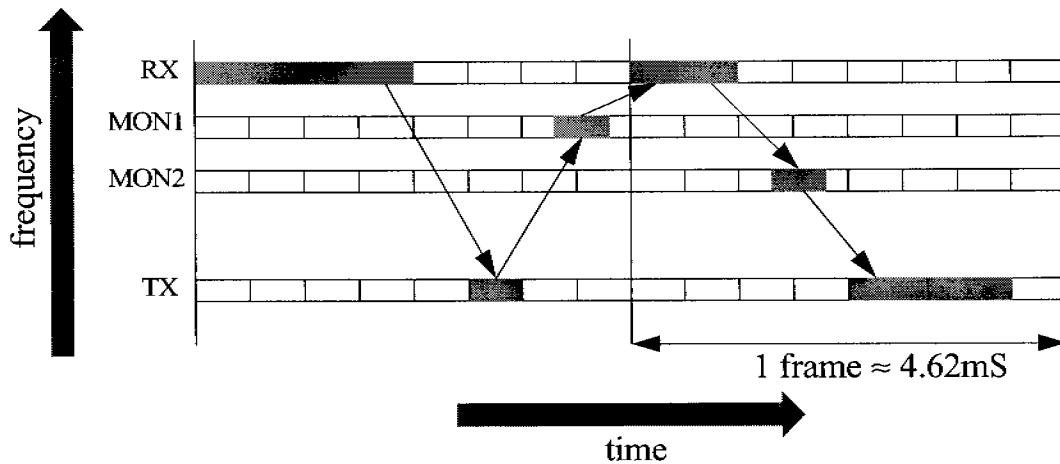
Figure 2.2: Frequency switching scheme of a GSM multi-slot communication link with a total of five allocated time slots per frame.

increase the data rate. A set of classes has been introduced with different numbers of time slots allocated in receive and transmit mode. The sum of receive and transmit time slots may even exceed eight, implying full-duplex operation. This has a major impact on the transceiver and is therefore no longer considered.

The multi-slot standards increase the maximum data rate to 57.6kb/s for HSCSD and to 171.2kb/s for GPRS. Besides reducing the number of users in a cell, the available switching time of the synthesizer is reduced. Fig. 2.2 outlines a timing diagram of two multi-slot time frames with five allocated time slots. It can be seen that only a single time slot is available for switching from transmit to receive. This implies a reduction of the switching time by a factor of at least two compared to standard GSM. A more detailed discussion of HSCSD/GPRS synthesizer switching time requirements can be found in [8]. Only final results are presented here. Available switching time is derived as 310μs and as 288μs for GPRS and HSCSD respectively, which is about three times faster than required for standard GSM.

2.3. Spectral Purity

Spectral purity of the synthesizer signal is crucial since it affects transceiver operation by various effects. Two type of spectrum imperfections can be distinguished. First, the synthesizer spectrum exhibits a carrier surrounded by a noise skirt, which decays with increasing offset from the carrier. Second, the spectrum

is often accompanied by discrete, unwanted tones, often called spurious tones. Section 2.3.1 describes the random modulation process which generates the noise skirt. Section 2.3.2 and Section 2.3.3 deal with the effects of the noise skirt within and outside the channel bandwidth respectively. Finally, Section 2.3.4 examines the impact of spurious tones.

2.3.1. Phase Noise Foundation

A rigorous description of phase noise spectra, voltage spectra and their relationship must be constituted before noise requirements can be derived from type approval test cases. The output voltage $V_{synth}(t)$ of a synthesizer can be formulated as

$$V_{synth}(t) = A(t) \cdot \cos\left(2\pi \cdot f_{LO} \cdot t + \varphi(t)\right)$$ (2.1)

with $A(t)$ and $\varphi(t)$ being random processes modulating amplitude and phase. Most practical synthesizers provide a virtually constant amplitude. Amplitude modulation can be therefore neglected in most spectral purity considerations, revealing phase noise as the dominant contributor of spectral imperfections. The random phase modulation creates sidelobes in the voltage spectrum of the synthesizer, which are symmetrically arranged around the carrier frequency. The sidelobes are most often characterized by the single sideband (SSB) power in a 1Hz bandwidth, normalized to the power of the carrier signal, measured at an offset frequency f_m from the carrier f_{LO}:

$$L(f_m) = 10 \cdot \log\left(\frac{\text{noise power in 1Hz bandwith at } f_{LO} \pm f_m}{\text{carrier power}}\right)$$ (2.2)

The unit of $L(f_m)$ is dBc/Hz, indicating that noise power is expressed relative to the carrier. The chosen variable f_m reminds that the sidelobes are caused by a "baseband" modulation process rather than a linear superposition of noise at the RF band. $L(f_m)$ can be displayed by a spectrum analyzer, but the relation to the random phase modulation process $\varphi(t)$, which cannot be directly measured, is not obvious. While some impacts of spectral imperfections are best described by sidelobe power expressed by $L(f_m)$, e.g reciprocal mixing, others require insight into the spectrum of the phase modulating signal. Fortunately, the power spectral

density $S_\phi (f_m)$ of the phase random process $\varphi(t)$ and the voltage power spectral density (PSD) $S_V (f)$ of the synthesizer output are directly linked by the identity

$$S_\phi(f_m) = 2 \cdot \frac{S_V(f_{LO} \pm f_m)}{P_{CR}} \qquad (2.3)$$

with P_{CR} the carrier power. A derivation of the latter identity is presented in Appendix 2.A. Eq. 2.3 can be rearranged by replacing the ratio of the right hand side and using the definition of $L(f_m)$ instead:

$$S_\phi(f_m) = 2 \cdot 10^{L(f_m)/10} \qquad (2.4)$$

The last two Equations demonstrate that the voltage spectrum seen in the side-lobes around the carrier is a direct measure of the phase noise spectrum. Fig. 2.3 visualizes the identity of spectral phase noise power and power in the sidelobes. The factor two in Eq. 2.3 can be explained by the fact that phase noise power at frequency f_m is divided into two spectral components located at $f_{LO} + f_m$ and $f_{LO} - f_m$. The relation between phase noise and voltage spectra enables later the computation of residual phase error.

Random phase modulation is not only observed in frequency, but also in time domain. Here, phase noise reveals deviations of the synthesizer output voltage zero-crossings from the average period. While this measure called period jitter is not very significant in the context of transceivers, it is a crucial performance parameter of PLL's aimed to the generation of clock signals and clock recovery. Period jitter is mentioned here for the sake of completeness. Appendix 2.B confirms that RMS period jitter Δt_{rms} is fully determined by the phase noise spectrum. The identity is derived as:

$$\Delta t_{trms} = \frac{T}{\pi} \cdot \sqrt{\int_0^\infty S_\phi(f) \cdot \sin^2(\pi \cdot T \cdot f) \, df} \qquad (2.5)$$

T stands for the average zero-crossing period which is equal to the inverse of the synthesizer carrier frequency f_{LO}.
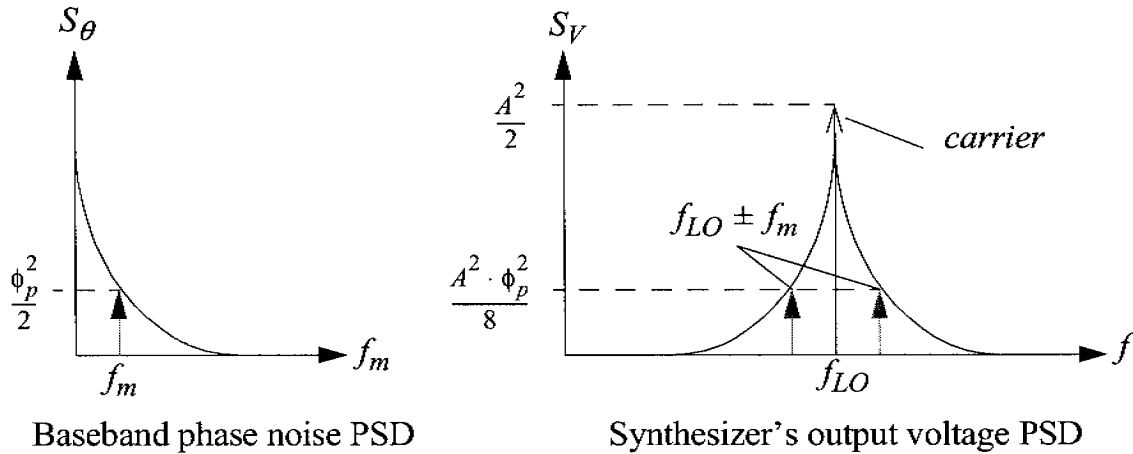
Figure 2.3: PSD of the baseband phase modulating process and the resulting voltage PSD surrounding the RF carrier signal.

2.3.2. Phase Trajectory Deviation and Residual Phase Error

Digital phase and frequency modulated signals carry the information in the phase. Any deviation from the intended, ideal phase trajectory of the transmitted signal increases the probability of erroneous symbol detection and increases accordingly the bit error rate (BER). In general, an increasing number of symbols in the constellation diagram reduces the maximum acceptable phase deviation. Tolerable phase deviation is besides the bit error rate strongly dependant on the used modulation scheme. The RMS deviation from the ideal phase trajectory is called residual phase error. Among other transceiver components, synthesizer phase noise contributes to overall residual phase error $\Phi_{e,rms}$. Residual phase error due to synthesizer imperfections can be given as

$$\Phi_{e,rms} = \sqrt{\langle \varphi^2(t) \rangle} \tag{2.6}$$

with $\varphi(t)$ the random phase noise introduced in Eq. 2.1 and $\langle . \rangle$ indicating the time average operator:

$$\langle [.] \rangle = \lim_{T \to \infty} \frac{1}{T} \cdot \int_{-T/2}^{T/2} [.] dt \tag{2.7}$$

Assuming wide-sense stationarity, phase noise in time domain $\varphi(t)$ and its power spectral density $S_\phi(f_m)$ are linked by the following formula [9]:

$$\langle \varphi^2(t) \rangle = \int S_\phi(f_m) df_m \qquad (2.8)$$

The lower integration limit f_1 is roughly equal to the bandwidth of the receiver's carrier recovery loop. The upper integration limit f_2 is determined by the bandwidth of the modulated baseband signal [10]. Combination of Eq. 2.4 and Eq. 2.8 reveals the final term for the RMS residual phase error $\Phi_{e,rms}$:

$$\Phi_{e,rms} = \sqrt{\int_{f_1}^{f_2} 2 \cdot 10^{L(f_m)/10} df_m} \qquad (2.9)$$

Phase deviation from the ideal phase trajectory is specified for GSM transmitters as 5° RMS and 20° peak deviation. The overall transmitter residual phase error is composed of phase errors contributed by the baseband modulator, the up conversion mixer, the power amplifier and finally by synthesizer phase noise. To leave some margin to the former building blocks, 2° RMS synthesizer residual phase error is often postulated. An analytical residual phase error estimation is aggravated by the lower integration limit of Eq. 2.9 which lies within the PLL loop bandwidth. Synthesizer phase noise is determined within the PLL loop bandwidth by various loop blocks rather than the free running oscillator. Fortunately, a PLL's phase noise spectrum can be approximated by a flat region within the loop bandwidth, and by a $1/f^2$ decaying region outside the loop bandwidth. This simple model is well qualified to represent phase noise power spectral density in a band relevant for residual phase error computation. Fig. 2.4 sketches a synthesizer phase noise spectrum and its above described approximation. Residual phase noise, integrated from DC to f_2, with the proposed phase noise approximation and a PLL loop bandwidth f_{LBW} smaller than f_2, can be found with a few manipulations as:

$$\Phi^2_{e,rms} = 2 \cdot 10^{L_0/10} \cdot f_{LBW} \cdot \left(2 - \frac{f_{LBW}}{f_2} \right) \qquad (2.10)$$

with L_0 being the constant phase noise level within the loop bandwidth. Inband phase noise level is found by solving Eq. 2.10 for L_0:
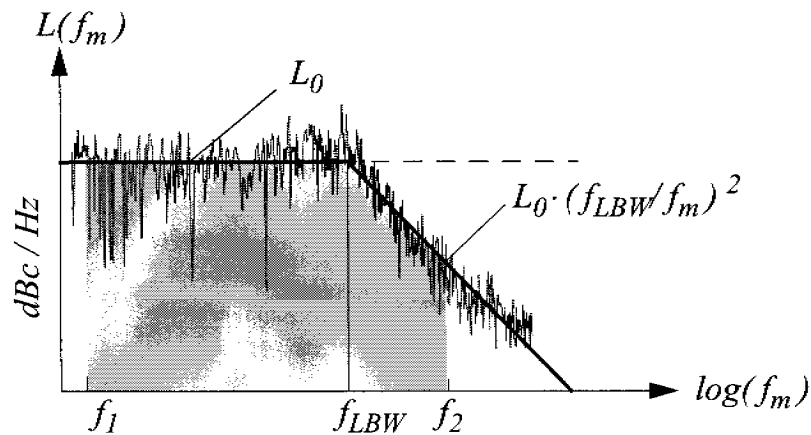
Figure 2.4: Approximation of the synthesizer phase noise spectrum by a region of constant phase noise and a region with $1/f^2$ decaying phase noise.

$$L_0 = 20 \cdot \log(\Phi_{e,\,rms}) - 10 \cdot \log(f_{LBW}) - 6dB - 10 \cdot \log\left(1 - \frac{f_{LBW}}{2f_2}\right) \quad (2.11)$$

Eq. 2.11 allows the calculation of the acceptable inband phase noise for a given residual phase error. Numerical values for GSM are 35mrad residual phase error (which corresponds to 2°) and an upper integration limit f_2 of 200kHz. The loop bandwidth of the PLL is assumed to be 10% of the channel spacing frequency, or 20kHz in absolute numbers. This value is determined by PLL stability considerations. Applying these numbers to Eq. 2.11 reveals inband phase noise of -78dBc/Hz.

2.3.3. Receive Signal Degradation by Reciprocal Mixing

The previous Section described how synthesizer phase noise close to the carrier deteriorates the bit error rate by addition of noise to the ideal phase trajectory. Besides this effect, phase noise at large offset frequencies from the carrier affects the quality of the received signal negatively. Assume a wanted channel which is downconverted to an intermediate frequency (IF) by a noisy local oscillator. An interfering signal located Δf Hertz above (below) the wanted signal will be downconverted to the same IF by mixing with phase noise present at the same Δf Hertz above (below) the LO carrier. Fig. 2.5 illustrates the effect, which is most often called reciprocal mixing. It is unnecessary to mention that the unwanted power
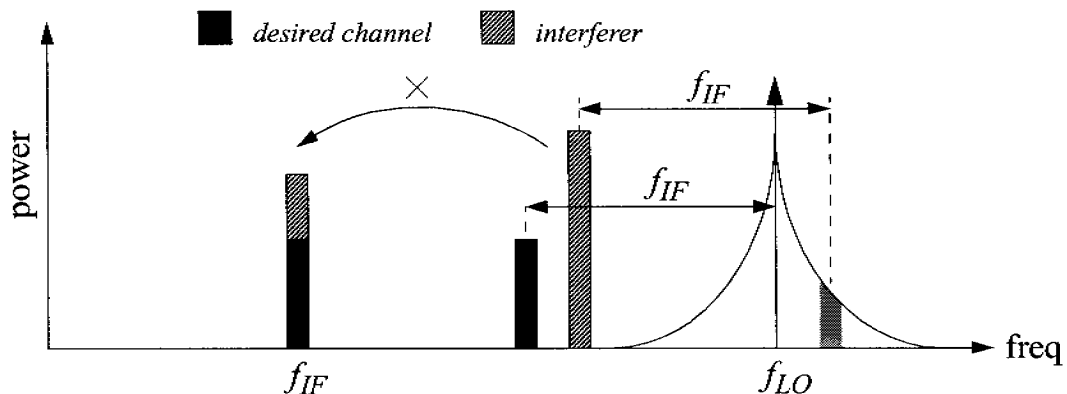
Figure 2.5: Signal degradation at IF by reciprocal mixing.

present at IF, which scales with interferer power at RF and phase noise power at the relevant frequency, will deteriorate the bit error rate of the receiver.

Reciprocal mixing may represent a serious obstacle to cellular services. The near-far problem reveals the worst case situation of a mobile station connected to a remote base station providing only a weak signal, while the mobile station can be blocked by a nearby strong transmitter in the same frequency band. GSM and WCDMA prescribe therefore stiff blocking cases with the blockers being up to 76dB and 70dB stronger than the wanted signal. Fig. 2.6 shows a blocking signal template for GSM and WCDMA. Notice that reciprocal mixing is less severe to communication systems with equidistant receiver to transmitter constellations, e.g. satellite communications.

Once the power level of the interferes is bound by the type approval tests, minimization of synthesizer phase noise remains the only option to keep the signal to interferer ratio moderate. This leads directly to free running oscillator phase noise requirements, because the relevant phase noise lies at frequencies beyond the loop bandwidth of the PLL. Since oscillator phase noise cannot be suppressed by the PLL at these frequencies, employing an oscillator with low free running phase noise is mandatory.

Computation of the acceptable free running oscillator phase noise with respect to reciprocal mixing requires a closer look at the test cases defined by the wireless standards. Blocking test cases in GSM require detection of a -99dBm wanted channel with a bit error rate of 2% in the presence of a single interferer. In order to demodulate the GMSK modulated GSM signal with the mentioned BER, the signal to noise and interferer ratio (SNIR) must measure 6.2dB [11]. The small
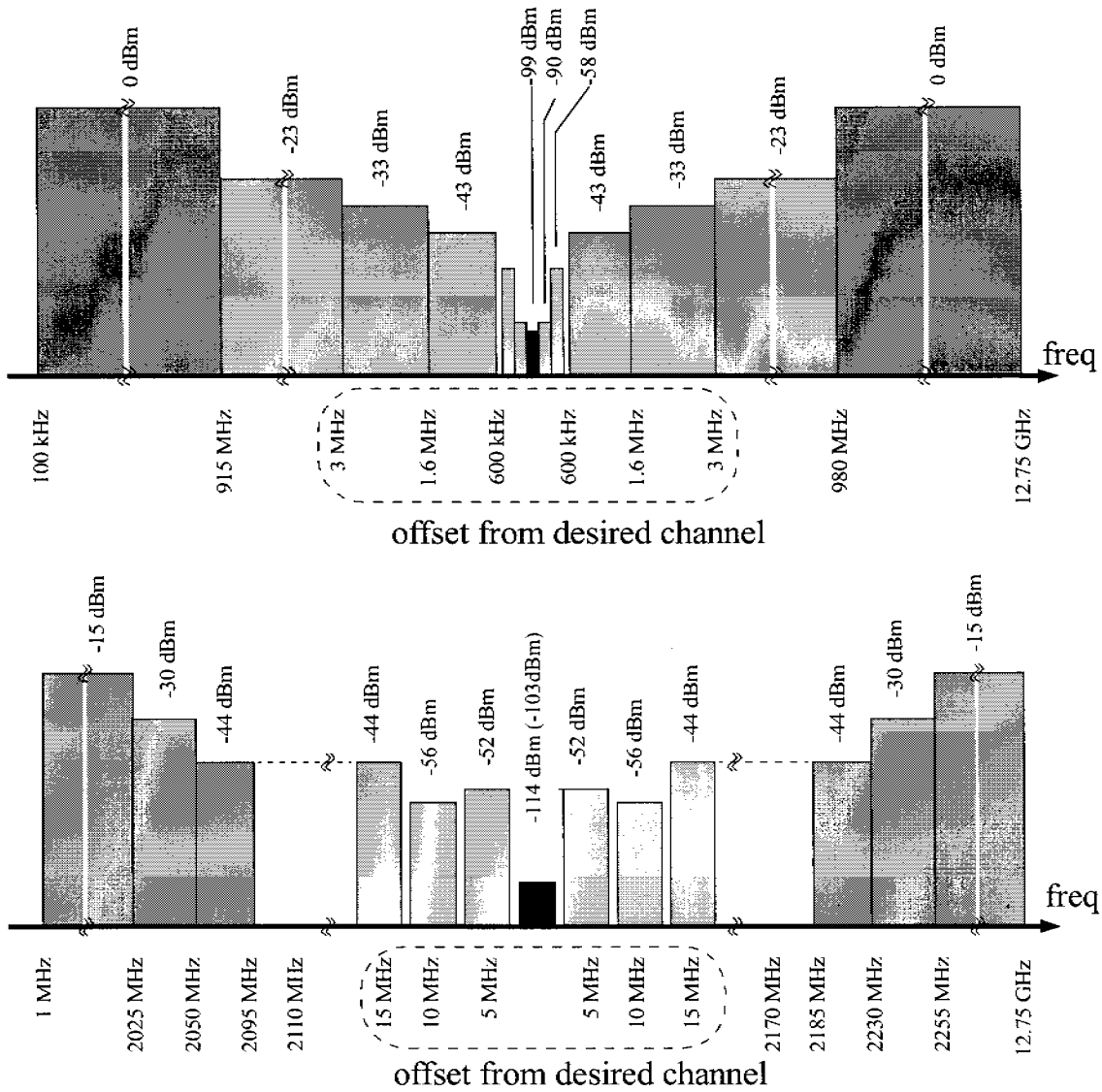
Figure 2.6: Blocking signal template for GSM (upper Figure) and WCDMA (lower Figure).

available signal power, which is only 3dB above the receiver's reference sensitivity, implies that half of the unwanted power at IF stems from thermal noise. Therefore, the signal to interferer ratio must measure 9.2dB. Acceptable phase noise with a blocker $\Delta(f)$ Hertz away from the wanted channel can be formulated as

$$L(\Delta f) = P_S - P_I - S/I - 10 \cdot \log(f_{CHBW}) \tag{2.12}$$

with P_S the power of the wanted channel, P_I the power of the interferer, S/I the signal to interferer ratio and f_{CHBW} the bandwidth of wanted channel. According to GSM specifications, the blocker revealing the most stringent phase noise requirement is located 3MHz away from the wanted channel and delivers -23dBm power. Eq. 2.12 evaluated with the numbers for this specific blocker signal leads to oscillator phase noise requirement at 3MHz offset as:

$$L(3MHz) = -99dBm - (-23dBm) - 9.2dB - 10 \cdot \log(200kHz)$$
$$= -138.2dB/Hz$$
(2.13)

Fig. 2.7 shows in stippled lines a phase noise template based on GSM blocking test cases. Assuming an $1/f^2$ oscillator phase noise dependency, free running oscillator phase noise at 100kHz offset from the carrier must not exceed -110dBc/Hz to pass type approval.

It is interesting to compare GSM phase noise requirements with WCDMA standards, which are believed to be less demanding. Like in GSM, most interferer test cases are defined with a signal 3dB above reference sensitivity. Demodulation must yield a BER of 1%, which translates into an SNIR of 7dB. Analogous to GSM, half of the unwanted power arises from thermal noise sources rather than from interferers, leading to a required signal to interferer ratio of 10dB. The interferer leading to the most stringent phase noise requirement is found as an unmodulated continuos wave (CW) tone present at 10MHz offset from the wanted channel providing -46dBm power. Since the interferer is defined in the intermodulation test case, the total interferer power at IF is composed of a part due to recipcroal mixing, and a second part due to third order intermodulation. Assuming that these two parts provide the same amount of unwanted power, a signal to interferer ratio of 13dB is required, if reciprocal mixing is considered as the only source of interferer power. A formula determining oscillator phase noise based on the intermodulation test can be found as

$$L(\Delta f) \approx G_P + P_S - P_I - S/I - 10 \cdot \log(f_{CHBW})$$
(2.14)

with G_P the SNIR enhancement factor resulting from despreading of the WCDMA signal. The other symbols are identical to those already introduced in the GSM phase noise calculation. Maximum oscillator phase noise can be evaluated by applying numbers for the CW interferer to the latter Equation:
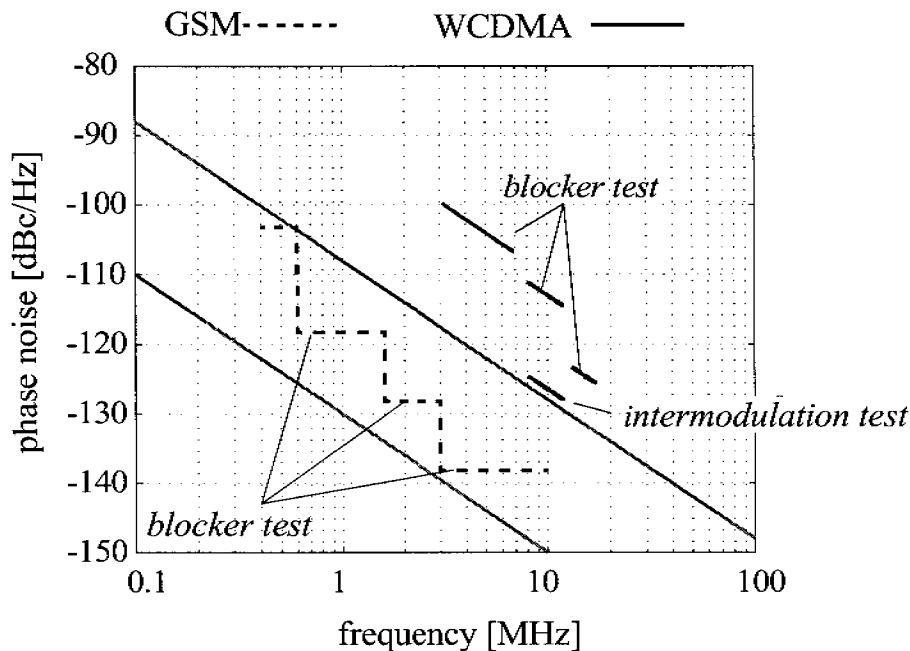
Figure 2.7: Synthesizer phase noise template for GSM and WCDMA as determined by reciprocal mixing of blocking signals.

$$L(10MHz) = 21dB - 114dBm + 46dBm - 13dB - 10 \cdot \log(3.84MHz)$$
$$= -126dBc/Hz$$

(2.15)

Notice the SNIR enhancement factor of 21dB and the CDMA channel bandwidth of 3.84MHz. Fig. 2.7 shows in bold lines a phase noise template for WCDMA. The graph shows clearly that the intermodulation test yields in the most demanding phase noise requirement.

Assuming again a $1/f^2$ phase noise dependency of the free running oscillator, phase noise at 100kHz offset must measure at least -87dBc/Hz to fulfill WCDMA type approval. Compared to the required -110dBc/Hz at the same offset frequency for GSM, free running oscillator phase noise requirements for WCDMA standards are much less stringent, as assumed earlier. This is basically a consequence of the much wider channels, moving the blockers away from the wanted channel.

2.3.4. Spurious Tones

Practical implementations of frequency synthesizers exhibit often spurious tones in their spectrum besides the previously discussed noise sidelobes. While all kind

of upconverted signals may cause spurious tones, e.g. picked up digital clock sig-nals etc., the strongest spurious tones are most often caused by the PLL's phase comparison. With the phase comparison frequency being identical to the channel spacing, a characteristic which applies to integer-N frequency synthesizers, spu-rious tones must be expected at offset frequencies identical to integer multiples of the channel spacing, i.e. $f_{LO} \pm n \cdot f_{CHSP}$. While the levels of spurious tones are hardly predictable, their impact on transmitter and receiver performance is appar-ent. Spurious tones in the transmitter LO upconvert the baseband signal to fre-quencies other than the wanted one. Unwanted power outside the allocated RF channel is generated, exceeding potentially the power level defined by the trans-mit power mask. On the receiver side, spurious tones present in the receive LO downconvert interferer to the same IF, revealing an unrectifiable signal to inter-ferer ratio. The effect resembles reciprocal mixing and requires therefore no fur-ther explanation.

Spurious tones specifications for the receive LO are determined by the power level of the wanted and interfering signal and the required signal to interferer ratio to achieve symbol detection with the previously specified BER. The spuri-ous tone level relative to the power of the synthesizer carrier P_{SP} is expressed as:

$$P_{SP} = P_S - P_I - S/I \qquad (2.16)$$

With interferer power of -90dBm, -58dBm and -43dBm at offset frequencies of 200kHz, 400kHz and 600kHz from the wanted channel, spurious tones are found for the receive LO as -18.2dBc, -50.2dBc and -65.2Bc. Spurious tones required to meet the specified GSM transmit power mask can be found as -30dBc, -60dBc and -66dBc for the same offset frequencies. Acceptable level of spurious tones is hence determined by the transmit LO, which are slightly more demanding than the requirements found for the receive LO.

2.4. Conclusions

Frequency synthesizers providing local oscillator signals enabling up- and down-conversion in cellular transceivers must satisfy extraordinary requirements. The most demanding ones are accuracy and spectral purity of the synthesized signal. While the frequency needs to be in the low GHz range, a deviation from the desired frequency of only some dozens Hz is acceptable. Needless to mention, that no kind of tunable oscillator in the GHz range can provide this accuracy.

Frequency band:	GSM	880-960 MHz
	DCS	1710 - 1880 MHz
	PCS	1850 - 1990 MHz
Frequency resolution:		200 kHz
Frequency accuracy:	GSM	90 Hz
	PCS	180 Hz
RX-TX frequency step:	GSM	45MHz
	PCS	90 MHz
Switching speed:	standard	866μS
	data service	~300μS
Free running oscillator phase noise:		-110dBc/Hz @100kHz
Oscillator tuning range:		~ 10%
Peak phase noise: (within PLL loop bandwidth)		-78dBc/Hz
Spurious tones:		-30dBc @200kHz
		-60dBc @400kHz
		-66dBc @600kHz

Table 2.1: Summary of GSM/DCS/PCS frequency synthesizer requirements.

Phase locking to a highly accurate, but low frequency reference oscillator is an efficient method to tackle the accuracy problem.

Besides frequency accuracy, spectral purity is one of the key parameters. Any kind of spectral imperfections of the local oscillator signal affects finally the bit error rate of the transceiver. Excellent spectral purity of the free running RF oscillator is mandatory, since the extraordinary clean reference oscillator determines the synthesizer output spectrum only within a small bandwidth. The most stringent spectral requirements are in general found in cellular standards with narrow bandwidth channels. Therefore, phase noise requirements for GSM are more demanding than requirements for WCDMA.

Besides the static parameters frequency accuracy and purity, the synthesizer must satisfy dynamic requirements. Switching between receive and transmit frequency must be performed within some 100µs. Recently introduced high speed data services tighten the dynamic requirement. Switching performance must be tackled with a reasonable fast phase locking control loop.

Table 2.1 summarizes the requirements imposed to a frequency synthesizer for a mobile transceiver station. However, frequency synthesizer design has matured over the years and the focus moved from satisfying only the constraints to the search for more economical solutions. Reduction of cost, weight, size and power consumption become more important design goals.

References

[1] J. Crawford, *"Frequency Synthesizer Design Handbook"*, Artech House, 1994, pp. 308-346

[2] ETSI EN 300 910 (GSM 05.05 version 8.5.1 Release 1999), *"Digital Cellular Communication System (Phase 2+); Radio Transmission and Reception"*, European Telecommunications Standards Institute, 2000

[3] ETSI EN 300 908 (GSM 05.02 version 8.5.1 Release 1999), *"Digital Cellular Communication System (Phase 2+); Multiplexing and Multiple Access on the Radio Path"*, European Telecommunications Standards Institute, 2000

[4] ETSI TS 100 912 (GSM 05.10 Version 8.4.0 Release 1999), *"Digital Cellular Communication System (Phase 2+); Radio Subsystem Synchronization"*, European Telecommunications Standards Institute, 2000

[5]   3G TS25.101 version 3.2.0 Release 1999, *"Technical Specification Group Radio Access Networks: UE Radio Transmission and Reception (FDD)"*, Third Generation Partnership Project 3GPP, 2000

[6] D. Salt, *"The Hy-Q Handbook of Quartz Crystal Devices"*, Chapman & Hall, Feb. 1987

[7] P. Basedau, Q. Huang, *"Design Considerations for High-Frequency Crystal Oscillators Digitally Trimmable to sub-ppm Accuracy"*, IEEE T. on Very Large Scale Integration (VLSI) Systems, Vol. 5, No. 4, Dec. 1997, pp. 408-416

[8] C. Wuensch, *"Fractional-N Synthesizer Meet HSCSD and GPRS Demands"*, Microwave & RF, Aug. 1999

[9] L. Couch II, *"Digital and Analog Communication Systems"*, 5th edition, Prentice-Hall, 1997, pp. 381-395

[10] C. Voucher, *"Architectures for RF Frequency Synthesizers"*, Ph.D. Thesis, Philips Electronics, 2001, pp. 91

[11] B. de Muer, *"Monolithic CMOS Fractional-N Frequency Synthesizers"*, Ph.D Thesis, Katholieke Universiteit Leuven, 2002, pp. 45

Appendix 2

2.A Relation between Sidelobe Power and Phase Noise Power Spectral Density

Phase noise is often characterized and measured by the voltage PSD of the noise skirt around the carrier. However, a random process which modulates the phase of the synthesizer output is responsible for the introduction of a symmetric voltage noise skirt. The relation between the PSD of the phase modulating random process and the voltage PSD of the synthesizer output is derived now. The presented calculation is based on a deterministic harmonic phase modulating signal in order to bypass stochastic mathematics. Since noise can be understood as a superposition of harmonic signals of random amplitude, the outcome of the harmonic analysis can directly be applied to noise.

The synthesizer output voltage of which the phase is modulated by a harmonic signal can be formulated as

$$V_{synth}(t) = A \cdot \cos\left[2\pi \cdot f_{LO} \cdot t + \phi_p \cdot \sin\left(2\pi \cdot f_m \cdot t \right) \right] \qquad (2.A.1)$$

with ϕ_p the amplitude of the phase modulating signal. The synthesizer output voltage can be rearranged by some trigonometric manipulations as:

$$V_{synth}(t) = A \cdot Re\left\{ e^{j2\pi \cdot f_{LO} \cdot t} \cdot e^{j\phi_p \sin 2\pi \cdot f_m \cdot t} \right\}$$

$$= A \cdot Re\left\{ e^{j2\pi \cdot f_{LO} \cdot t} + \frac{\phi_p}{2} e^{j2\pi(f_{L0}+f_m)t} - \frac{\phi_p}{2} e^{j2\pi(f_{L0}-f_m)t} \right\} \qquad (2.A.2)$$

$$= A\cos\left(2\pi \cdot f_{LO} \cdot t\right) + A\frac{\phi_p}{2}\cos\left(2\pi(f_{L0}+f_m)t\right) - A\frac{\phi_p}{2}\cos\left(2\pi(f_{L0}-f_m)t\right)$$

Notice that e^x is approximated by a Taylor series as $e^x \cong 1 + x$, which holds only for small angle modulation such as noise. The phase modulated synthesizer contains three harmonic tones with the voltage PSD:

$$S_V(f) = \qquad\qquad (2.A.3)$$

$$\frac{A^2}{2} \cdot \delta(f-f_{LO}) + \frac{A^2 \cdot \phi_p^2}{8} \cdot \delta(f-f_{LO}-f_m) + \frac{A^2 \cdot \phi_p^2}{8} \cdot \delta(f-f_{LO}+f_m)$$

Notice, that the PSD of the phase modulating signal is given by:

$$S_\phi(f) = \frac{\phi_p^2}{2} \cdot \delta(f-f_m) \qquad\qquad (2.A.4)$$

The combination of Eq. 2.A.3 and Eq. 2.A.4 reveals the relation between phase PSD and voltage PSD

$$S_V(f) = P_{CR} \cdot \delta(f-f_{LO}) + \frac{P_{CR} \cdot S_\phi(f-f_{LO})}{2} + \frac{P_{CR} \cdot S_\phi(f_{LO}-f)}{2} \qquad (2.A.5)$$

with P_{CR} the carrier power. Solving of Eq. 2.A.5 for the phase PSD leads finally to:

$$S_\phi(f) = 2 \cdot \frac{S_V(f_{LO} \pm f)}{P_{CR}} \qquad\qquad (2.A.6)$$

2.B Period Jitter versus Phase Noise Power Spectral Density

Period jitter is defined as the RMS deviation of the time difference of two consequent zero-crossings from the average period. Consider a periodic signal with random phase modulation:

$$V(t) = A \cdot \cos\left(2\pi \cdot \frac{t}{T} + \varphi(t)\right) \qquad (2.B.1)$$

T represents the average time difference of the zero crossings and $\varphi(t)$ the phase random process. Variables t_a and t_b are introduced to represent the occurrence of zero-crossings. The argument of the cosine function at t_a and t_b must be equal to 0 and 2π respectively. This leads to a pair of Equations:

$$2\pi \cdot \frac{t_a}{T} + \varphi(t_a) = 0$$

$$2\pi \cdot \frac{t_b}{T} + \varphi(t_b) = 2\pi \qquad (2.B.2)$$

The instantaneous deviation Δt from the average period is defined as:

$$\Delta t = (t_b - t_a) - T \qquad (2.B.3)$$

Combination of Eq. 2.B.2 and Eq. 2.B.3 and solving for Δt leads to:

$$\Delta t = \frac{T}{2\pi} \cdot \left(\varphi(t_a) - \varphi(t_b)\right) \qquad (2.B.4)$$

Assuming an ergodic random phase process $\varphi(t)$, the RMS period jitter is found by computation of the time averaged square of the instantaneous period deviation:

$$\Delta t_{rms}^2 = \langle \Delta t^2 \rangle = \left(\frac{T}{2\pi}\right)^2 \cdot \left\langle \left(\varphi(t_a) - \varphi(t_b)\right)^2 \right\rangle$$

$$= \left(\frac{T}{2\pi}\right)^2 \cdot \left(\langle \varphi^2(t_a) \rangle + \langle \varphi^2(t_b) \rangle - 2 \cdot \langle \varphi(t_a) \cdot \varphi(t_b) \rangle\right) \qquad (2.B.5)$$

Time average of a squared wide-sense stationary process is identical to the integral of its power spectral density [9]. Since $\varphi(t)$ is presumed to be wide-sense stationary, the following identity applies:

$$\langle \varphi^2(t_a) \rangle = \langle \varphi^2(t_b) \rangle = \int_0^\infty S_\phi(f) df \qquad (2.B.6)$$

Further, the relation between the autocorrelation function $R_\phi(\tau)$ and phase noise PSD is according to the Wiener-Khintchine Theorem [9]:

$$R_\phi(t_a, t_b) = R_\phi(T) = \langle \varphi(t_a) \cdot \varphi(t_b) \rangle = \int_0^\infty S_\phi(f) \cdot \cos(2\pi \cdot T \cdot f) df \qquad (2.B.7)$$

Combination of Eq. 2.B.5, Eq. 2.B.6 and Eq. 2.B.7 leads to a period jitter formula:

$$\begin{aligned} \Delta t_{trms} &= \frac{T}{\pi} \cdot \sqrt{\int_0^\infty S_\phi(f) \cdot \frac{1 - \cos(2\pi \cdot T \cdot f)}{2} df} \\ &= \frac{T}{\pi} \cdot \sqrt{\int_0^\infty S_\phi(f) \cdot \sin^2(\pi \cdot T \cdot f) df} \end{aligned} \qquad (2.B.8)$$

Period jitter is, according to Eq. 2.B.8, fully determined by oscillator phase noise. Notice that the latter Equation provides an indirect method to measure period jitter. Rather than measuring zero-crossings by the help of a sampling oscilloscope, jitter can be obtained alternatively from spectrum measurements. This may be an attractive alternative since time based jitter measurements are limited by the jitter accuracy of the oscilloscope's internal time base.

Chapter 3

Phase Locked Loop Fundamentals

Phase locked loops (PLLs) provide the frequency stability of a highly stable, crystal based reference oscillator to an RF oscillator, the frequency of which would drift away without the regulation by the PLL. Besides the desired frequency stabilization, a number of unwanted effects occur. Among these, the generation of spurious tones in the spectrum of the RF oscillator must be considered as the most serious one. The root of spurious tones is found in the operation of the phase locked loop's phase detector, which represents, besides the required oscillators, the heart of the PLL. The sole task of the phase detector is to translate the phase difference of two signals into a proportional voltage. However, the wanted voltage is mostly superposed by fundamental and harmonic tones of the input signals. Under locked condition, when the phase difference remains constant, these harmonics generate a periodic voltage ripple on the otherwise constant RF oscillator tuning voltage. This voltage ripple is translated into spurious tones by frequency modulation.

A filter is required in general to suppress the voltage ripple, not only in order to reduce the level of spurious tones, but also to keep the distortion of the RF oscillator tuning voltage moderate. Unwanted nonlinear effects could otherwise deteriorate the dynamic behavior of the phase locked loop. The election of a loop filter is driven by conflictive considerations. While a small filter bandwidth is advantageous to the suppression of spurious tones, fast PLL settling, i.e. fast settling of the output frequency, is accomplished by a large bandwidth.

The choice of the phase detector is of great significance besides the identification of an optimum loop filter. Different kinds of phase detectors exhibit different levels of harmonics. Phase locked loops employing a phase-frequency-detector (PFD) charge pump (CP) combination are well qualified for high performance

Phase-Frequency Detector (PFD)
Charge Pump (CP)
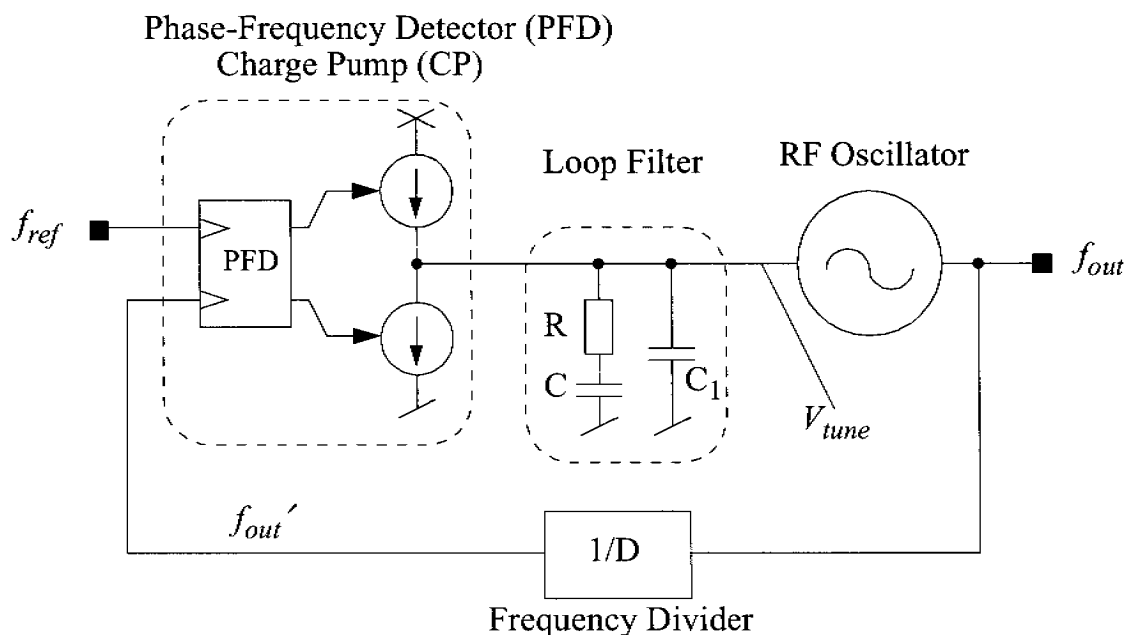
Figure 3.1: *A generic diagram of a charge pump phase locked loop.*

frequency synthesizers. Besides a wide locking range limited only by the RF oscillator and the available tuning voltage range, the PFD-CP PLL produces only weak harmonics. However, this holds only for type II loops.

This Chapter is aimed to a particular PLL species: The *type II, PFD-CP phase locked loop.* After highlighting the outstanding advantages of this particular topology, the underlying mathematics is developed. Many aspects such as transient behavior and noise can be characterized by a linearized model. On top of this, sampling effects are studied to consider the nature of the phased detector more accurately. PLL artifacts caused by charge pump nonidealities, responsible for the introduction of spurious tones in the synthesizer spectrum, are studied in detail and formulas are provided which allow the prediction of the level of spurious tones. The considerations and derived Equations at large provide the mathematical foundation for a design of a high performance, low power frequency synthesizer presented in Chapter 7.

3.1. Elementary Charge Pump PLL Operation

Fig. 3.1 shows a principal schematic of a PFD-CP phase locked loop, which will be called charge pump PLL from now on for convenience reasons. While most building blocks are self-explanatory, the PFD-CP combination requires a closer

look in order to understand PLL operation. The PFD represents an edge triggered device with three logical states. The state transition diagram with the three states named as OFF, UP and DOWN is shown in Fig. 3.2. Contrary to a synchronous state machine, the PFD is triggered by two signals f_{ref} and f_{out}'. Assuming that the triggering signals have the same frequency but f_{out}' is lagging f_{ref}, the PFD moves from the OFF state to the UP state with arising edge of f_{ref} and will fall back to the ZERO state with appearing edge at f_{out}', but the PFD will never enter the DOWN state. Periodic positive current pulses are generated at the output of the PFD-CP combination through wiring of the PFD with the charge pump in such a way that the upper current source is activated in the UP state. While the strength of the current pulses is determined by the current source, the length of the pulses is identical to the time lag of the f_{out}' to the f_{ref} signal. Analogous considerations with f_{out}' leading f_{ref} and the lower current source of the charge pump being activated in the DOWN state, lead to periodic negative current pulses with the length given by the time difference of the two triggering signal edges. Consequently, an average current I_{av} is established at the output of the PFD-CP combination which can be formulated as:

$$I_{av} = \frac{\varphi_e}{2\pi} \cdot I_{CP} \quad with \quad -2\pi \leq \varphi_e \leq 2\pi \qquad (3.1)$$

I_{CP} represents the current of the two current sources and φ_e the error phase which is constituted by the phase difference of the reference signal f_{ref} and the phase of the signal at the divider output f_{out}'. Eq. 3.1 demonstrates that the average current seen at the output of the PFD-CP combination is a perfect measure of the error phase as long as the latter remains within the 4π wide linear range.
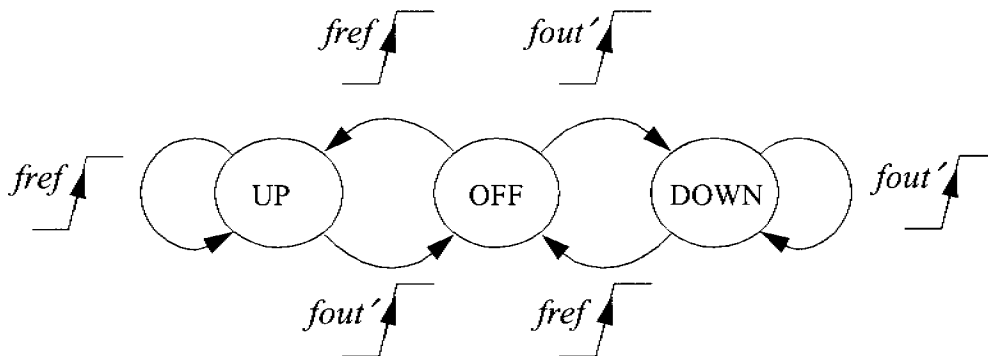

Figure 3.2: State transition diagram of the elementary phase-frequency detector (PFD).

The frequency stabilization capability of the PLL can be framed now. A slightly too small RF oscillator frequency causes accumulation of error phase which is translated by the charge pump into an increased average current. The increased current raises the RF oscillator tuning voltage resulting in a higher output frequency. This process continues until the desired output frequency is reached and the phase of the RF oscillator is locked to the phase of the reference oscillator.

Frequency stabilization can be achieved instead by PLLs employing other phase detector circuits. Charge pump type PLLs became the unrivaled workhorse in frequency synthesis and in other areas despite the availability of alternatives. The reason for the predominance of the PFD-CP combination lies in a couple of advantages which make the charge pump PLL superior to other solutions. One of the upper hands of the PFD lies in its built-in capability to operate not only as phase, but also as frequency detector. Detection of frequency differences rather than phase differences becomes crucial when the initial frequency error is so big, that the error phase accumulated in a single cycle exceeds the linear range of the phase detector. PLLs without frequency detection capabilities experience under such a condition a slow pull-in process [1] which requires a considerable amount of time until the frequencies are adjusted, time which is often not available. Even worse, it may happen that the PLL is not able to regain phase locking. Fortunately, none of this applies to the charge pump PLL thanks to the frequency detection capability of the PFD. The PFD-CP combination produces a strong average current which counteracts frequency differences. Fig. 3.3 shows the PFD-CP average current versus input frequency ratio. The plot shows that an average current identical to approximately one half of the current source strength is established even at virtually identical input frequencies. This average current, which converges quickly to the full current source strength, is responsible for a fast change of RF oscillator tuning voltage, shifting the latter's frequency efficiently towards the wanted value. This mechanism provides fast frequency acquisition, even if the initial RF oscillator frequency is far away from the desired frequency. Notice that the locking range is limited only by the RF oscillator tuning range, but not by the phase detector.

Unfortunately, the excellent frequency acquisition is accompanied by an undesired property. The pulse shaped nature of the charge pump current is responsible for strong reference frequency harmonics. Other phase detectors with weaker harmonics have been proposed, but they lack frequency discrimination. The fre-
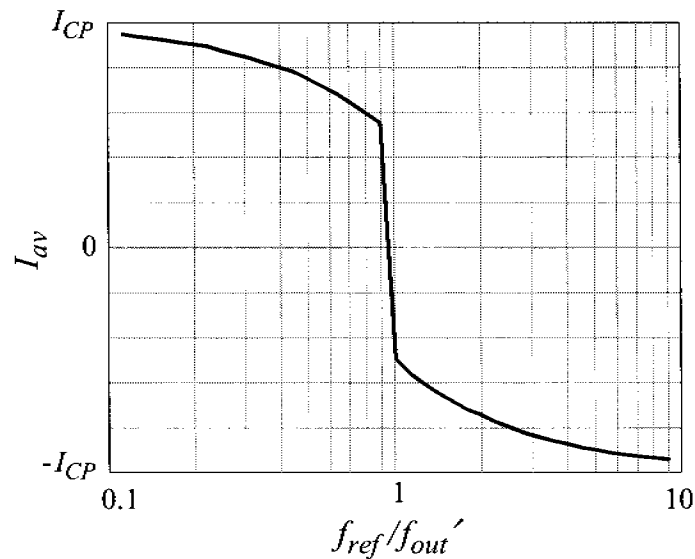
Figure 3.3: Average charge pump current of a PFD-CP combination operated with differing input frequencies.

quency synthesis most promising representative is the sample-and-hold phase detector, which generates no harmonics at all in locked condition [2]. Additional circuitry may be necessary to overcome this shortage.

Fortunately, the generation of harmonics in a charge pump PLL can be restricted to the locking transient, so that steady-state is completely free of harmonics. This is possible because the level of harmonics depends on the error phase. They vanish if the steady-state error phase becomes zero, i.e. if the PFD never leaves the ZERO state. A locked charge pump PLL can be forced to zero error phase if the charge pump is combined with a loop filter providing infinite DC impedance. Such a loop, which generates no spurious tones as the sample-and-hold phase detector PLL, must be mandatory of type II, as shown in Section 3.2. Nevertheless, even a type II charge pump PLL is not completely free of spurious tones. Residual spurious tones can be caused by PFD-CP artifacts. Section 3.6.3 discusses this problem in more detail.

Frequency detection capability and good spurious tones performance, if combined with an appropriate loop filter, reveals the charge pump PLL as the most capable phase locking structure of frequency synthesis.

3.2. Linearized Model and its Limitations

The nature of the charge pump PLL outlined in Section 3.1 is discrete in time and nonlinear. Nonlinearity and sampling, both caused by the PFD, burden the analysis to an extent which makes the introduction of a simplified model mandatory. Obviously, the neglect of nonlinearity and sampling affects the model accuracy. Nonlinearity behavior caused by the nonlinear phase detector gain can be simply avoided by restricting the error phase to the linear region of the phase detector. RF oscillators covering only a small band hardly exceed the linear region. Nonlinearity represents therefore a minor obstacle in the context of cellular applications which fulfill the small RF band requirement. Sampling of the error phase on the other side can impact loop stability and step response significantly. The in general negative repercussion of sampling can be lowered by choice of a reference frequency much larger than the bandwidth of the PLL. The charge pump current can then be approximated by its average current, revealing a time continuous system. Limiting further the application of the model to frequency synthesizers covering a small RF band, the charge pump PLL can be approximated by a linear, time invariant (LTI) system, allowing the application of the full wealth of analytical methods available to these systems.

Fig. 3.4 shows a block diagram of the time continuous, linearized model which will be used to analyze the charge pump during the entire course of this Chapter. Table 3.1 provides a listing of the employed blocks, the used symbols and the units of the involved measures. The model provides besides the reference frequency input f_{ref} a second input f_1 which will be used to model RF oscillator phase noise. The introduction of the PLL's open loop gain G_{OL} as

$$G_{OL}(s) = f_{out}'(s)/f_{ref}(s) \tag{3.2}$$

provides more insight into the loop behavior. Two transfer functions can be defined as

$$G_{CL}(s) = \frac{f_{out}'(s)}{f_{ref}(s)} = \frac{G_{OL}(s)}{1 + G_{OL}(s)} \tag{3.3}$$

$$G_1(s) = \frac{f_{out}(s)}{f_1(s)} = \frac{1}{1 + G_{OL}(s)} \tag{3.4}$$
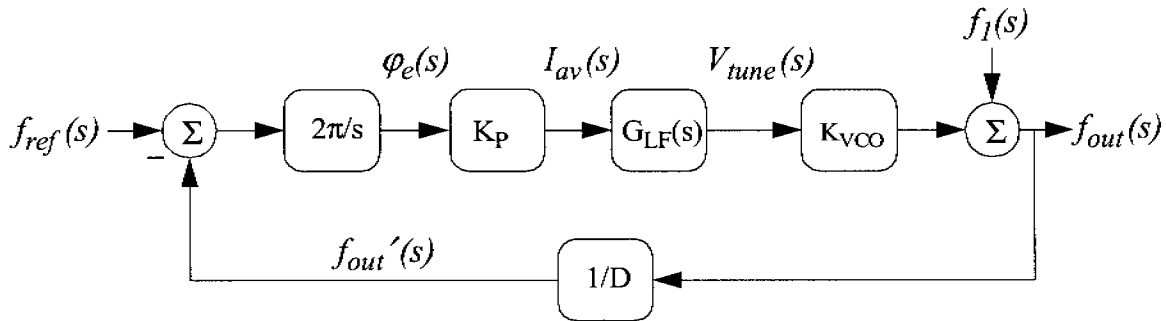
Figure 3.4: Linear PLL model.

with G_{CL} and G_I called closed loop and error transfer function respectively. The open loop gain can be expressed by the loop's linearized building block parameters. Inspection of the block diagram provided by Fig. 3.4 reveals the open loop gain as

$$G_{OL}(s) = \frac{I_{CP} \cdot K_{VCO}}{D} \cdot \frac{1}{s} \cdot G_{LF}(s) \tag{3.5}$$

with the symbols described in Table 3.1. Eq. 3.5 indicates that any PLL represents inherently a type I system due to the ideal integrator which stems from frequency to phase conversion. Consequently, the open loop gain has low pass characteristic and infinite DC gain. Keeping this in mind, fundamental PLL properties can be formulated which apply to any phase locked loop:

- *Frequency Tracking*:
  The output signal tracks the reference signal as long as the modulation of the reference is kept within the loop bandwidth (Eq. 3.3).

- *Reference Oscillator Noise Amplification*:
  Phase noise of the reference signal is replicated to the RF oscillator due to the frequency tracking property. Even worse, the replicated phase noise is multiplied by the frequency division ratio D (Eq. 3.3).

- *RF Oscillator Noise Suppression*:
  Phase noise of the RF oscillator is suppressed by the open loop gain (Eq. 3.4). Synthesizer phase noise of the PLL output outside the loop bandwidth is determined by free running RF oscillator phase noise (Eq. 3.3 and Eq. 3.4).

The frequency synthesizer's frequency stability is consequently determined by the stability of the used reference oscillator, an effect which originally motivated the introduction of a PLL. On top of this, the frequency purity is divided into two

Block:	Equation:	Unit:
	Frequency → Phase:	
$2\pi/s$	$\varphi_e(s) = \dfrac{2\pi}{s} \cdot (f_{ref} - f_{out}')$	[rad/Hz]
	Phase → Current:	
K_P	$I_{av}(s) = \dfrac{I_{CP}}{2\pi} \cdot \varphi_e$	[A/rad]
	Current → Voltage:	
$G_{LF}(s)$	$V_{tune}(s) = \begin{cases} R \cdot \dfrac{s + \omega_z}{s} \cdot I_{av}(s) \\[2ex] R \cdot \dfrac{s + \omega_z}{s} \cdot \dfrac{\omega_z \cdot (b - 1)}{s + \omega_z \cdot b} \cdot I_{av}(s) \end{cases}$ $\omega_z = 1/RC \qquad b = 1 + C/C_1$	[Ω]
	Voltage → Frequency:	
K_{VCO}	$f_{out} = f_{fr} + K_{VCO} \cdot V_{tune}$	[Hz/V]
	Frequency → Frequency:	
$1/D$	$f_{out}' = \dfrac{1}{D} \cdot f_{out}$	[.]

Table 3.1: The building blocks of the linear PLL model, Equations related to the building blocks and the involved units.

regions. It is determined close to the carrier by amplified reference oscillator phase noise and far away from the carrier, where the loop provides no gain any more, by free running RF oscillator phase noise.

Another issue of great importance is the time required to establish a new frequency at the synthesizer output. A change of output frequency is initiated in most cases by reprogramming the division ratio of the frequency divider. The transient response can be obtained from an equivalent step at the reference fre-

quency. However, this holds only for small frequency steps, i.e. if the change of division ratio does not alter the open loop gain significantly. Again, the small RF band of cellular applications assists the accuracy of the dynamic behavior as predicted from the linear model.

The next Sections analyze the performance of PLLs employing different loop filters. Section 3.3 starts with a simple second order type II loop and emphasizes its limitations which motivate the introduction of a third order loop. Section 3.4 proposes a particular third order type II loop of great significance.

3.3. Second Order Loop

The second order loop must be considered as the simplest implementation of a charge pump PLL. However, a couple of drawbacks limit the application to low performance systems. Nevertheless, the second order is often found in literature [3] due to its simplicity which eases the mathematical treatment. A short review is given here for the reason of completeness.

A second order loop must be mandatory of type II to avoid excessive spurious tones. The loop filter must provide hence an ideal integrator. Besides the integrator, the filter must provide a zero to stabilize the feedback loop. Such an arrangement is found in the series combination of a resistor and a capacitor, leading to the following loop filter impedance:

$$G_{LF} = R \cdot \frac{s + \omega_z}{s}$$

$$\text{with} \quad \omega_z = \frac{1}{RC}$$

(3.6)

The closed loop transfer function can be formulated in the standard notation of second order LTI systems:

$$G_{CL}(s) = \frac{\omega_n}{q_p} \cdot \frac{s + \omega_n \cdot q_p}{s^2 + s \cdot \frac{\omega_n}{q_p} + \omega_n^2}$$

(3.7)

with ω_n the natural frequency and q_p the pole quality factor. Both can be expressed by loop parameters as

$$\omega_n = \sqrt{\frac{I_{CP} \cdot K_{VCO}}{D \cdot C}}$$

$$q_p = \sqrt{\frac{D}{I_{CP} \cdot K_{VCO} \cdot C}} \cdot \frac{1}{R}$$

(3.8)

Assuming undercritical damping, the second order loop's step response can be found as:

$$f_{out}'(t) = 1 + \frac{e^{-\frac{\omega_n}{2q_p} \cdot t} \cdot \sin\left(\omega_n \cdot a \cdot t - \Theta\right)}{a}$$

(3.9)

$$\text{with} \quad a^2 = 1 - \left(\frac{1}{2q_p}\right)^2 \quad \text{and} \quad \tan(\Theta) = 4 \cdot q_p^2 - 1$$

Eq. 3.6 to Eq. 3.9 can be used as a foundation for a second order type II charge pump PLL design. However, as mentioned before, the application of the second order loop is restricted to low performance systems. The handicaps of the second order loop are mainly twofold. First, the tuning voltage of the RF oscillator suffers from strong voltage jumps each time one of the charge pump current sources is activated. The voltage jumps with strength $R \cdot I_{CP}$ may become so large that they get clipped by the supply lines. The second and more serious handicap lies in thermal noise of the loop filter resistor that arrives unfiltered at the tuning node of the RF oscillator and superposes to oscillator phase noise. In order to avoid a substantial degradation of the synthesizer spectrum, the loop filter resistor must be kept small. Unfortunately, the small resistor value must be compensated by a larger charge pump current I_{CP} in order to maintain the pole quality. Especially low phase noise synthesizers require a so large charge pump current that the latter may substantially increase the overall power consumption of the frequency synthesizer. An additional shunt capacitor C_1 as shown in the loop filter in Fig. 3.1 tackles both problems. Voltage jumps during charge pump switching are suppressed to a large extent by the second capacitor. Further, thermal noise of the loop filter resistor is filtered by a first order lowpass section. Consequently, a larger resistor and smaller charge pump current combination can be tolerated without affecting the synthesizer spectrum.

The modification of the loop filter turns the second order loop into a third order one. While rules-of-thumb such as keeping C_1 at one tenth of the value of the main capacitor are sometimes proposed, and optimum solution requires an in-depth study of the third order transfer functions.

3.4. The Third Order Loop

A third order type II loop is constituted by a loop filter shown in Fig. 3.1. Its impedance, which provides an additional pole, can be found as

$$G_{LF}(s) = R \cdot \frac{b-1}{b} \cdot \frac{s+\omega_z}{s} \cdot \frac{\omega_p}{s+\omega_p} \qquad (3.10)$$

with parameter b defined as:

$$b = 1 + \frac{C}{C_1} \qquad (3.11)$$

Pole and zero frequency can be expressed by the loop filter passive components:

$$\omega_z = \frac{1}{RC} \qquad (3.12)$$
$$\omega_p = b \cdot \omega_z$$

The open loop gain can be found by combination of Eq. 3.5 and Eq. 3.10, leading to the following identity:

$$G_{OL}(s) = \frac{I_{CP} \cdot K_{VCO} \cdot R}{D} \cdot \frac{b-1}{b} \cdot \frac{s+\omega_z}{s^2} \cdot \frac{\omega_p}{s+\omega_p} \qquad (3.13)$$

So far, the linear model of the third order type II loop is described by the help of Eq. 3.11 to Eq. 3.13. The remaining and more demanding task lies in the identification of a favorable loop by clever choice of the four free loop parameters, which are the values of the passive components R, C, C_1 and the charge pump current I_{CP}. Notice that the RF oscillator gain K_{VCO} is in general not under control of the PLL designer. Some publications [4] treat the charge pump current as a fixed parameter. This must be seen in the context of standard product IC application which most often do not provide programmable charge pumps. The lack of a fourth design parameter may lead to different conclusion than those presented later and may cause inferior synthesizer performance.

The identification of optimum loop parameters is a multi-step procedure and involves the following considerations:

- *Stability*
- *Step response*
- *Suppression of spurious tones*
- *Noise*

Section 3.4.1 proposes the choice of zero and pole frequencies optimizing both stability and spurious tones suppression. Consideration of noise and sampling effects described in Section 3.4.3 and Section 3.5 contribute another two Equations so that the four design parameters can be analytically determined without the need for any trial-and-error iterations.

3.4.1. Pole and Zero Placement

Open loop pole and zero must be evaluated carefully to accomplish decent stability margins. An awkward choice can cause excessive ringing or excessive over-damping leading both to longer than necessary settling time. Fig. 3.5 shows the Bode plot of the third order loop. It can be seen that the open loop phase reaches 180° for low but also for high modulating frequencies, indicating that excessive ringing must be expected if the loop bandwidth is not sized with care. Decent phase margin (PM) can be expected only if the loop bandwidth ω_{LBW} lies somewhere between the zero ω_z and the pole ω_p. This can be achieved by choice of an appropriate charge pump current, since open loop gain scales with charge pump current. Notice that a change of charge pump current does not affect the open loop phase. It can be shown by a couple of arithmetic manipulations that the maximum phase margin occurs if the loop bandwidth is placed at the geometric mean of the pole and zero frequency:

$$\omega_{LBW} = \sqrt{\omega_z \cdot \omega_p} \tag{3.14}$$

The loop bandwidth is then equally separated from the zero and the pole, if the frequency axis is plotted on a logarithmic scale. Again, Fig. 3.5 sketches this fact. The obtainable phase margin *PM* which can be achieved by such a pole-zero-bandwidth-constellation depends on the separation of the pole and the zero. A small phase margin is gained from an arrangement with the pole and zero being placed close together, since pole and zero tend to cancel each other if spaced closely. Some trigonometric manipulations demonstrate that the phase
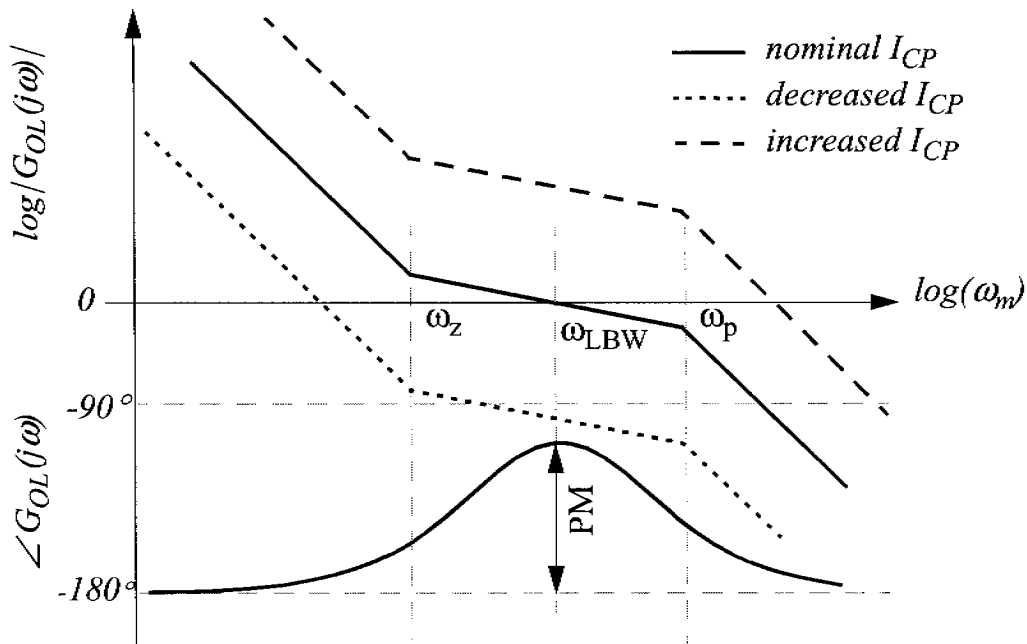
Figure 3.5: Bode plot of a third order type II PLL with the loop bandwidth chosen equidistant from zero and pole frequency, measured on a logarithmic plot. The stippled lines highlight the impact of a modified charge pump current.

margin is completely determined by the ratio of pole and zero frequencies. Hence, the previously introduced b parameter alone determines the phase margin as:

$$b = \left[\tan(PM) + \frac{1}{\cos(PM)} \right]^2 \qquad (3.15)$$

The arrangement of the loop bandwidth located equidistant from the pole and zero represents a special case of great significance since it represents an effective way to maximize the phase margin. Although this pole-zero-bandwidth-constellation represents common design technique [4], one might argue, that phase margin can be increased by moving the pole to a higher frequency or the zero to a lower frequency, revealing a different constellation with larger phase margin. There are, however, strong reasons against doing so. Moving the pole to higher frequencies deteriorates high frequency suppression and causes thereby stronger spurious tones. Moving the zero to a lower frequency may also impact the loop in a negative way because the low frequency gain is reduced thereby. Besides the need for a larger capacitance to realize the low frequency zero, the loop looses its

ability to suppress RF oscillator phase noise within the loop bandwidth because of the lack of a large open loop gain. One can conclude therefore that the proposed pole-zero-bandwidth-constellation combines effectively decent loop stability, spurious tones as well as RF oscillator phase noise suppression. The ongoing third order loop analysis is focussed therefore on this special case of a third order type II loop. Before advancing to the particular design aspects, the locations of pole and zero are reformulated as:

$$\omega_z = \frac{\omega_{LBW}}{\sqrt{b}}$$

$$\omega_p = \omega_{LBW} \cdot \sqrt{b}$$

(3.16)

As mentioned before, the loop bandwidth can be set by the charge pump current. The dependency on the latter current as well as on other loop parameters can be found as:

$$\omega_{LBW} = \frac{I_{CP} \cdot K_{VCO} \cdot R}{D} \cdot \frac{b-1}{b}$$

(3.17)

This identity can be gained from an inspection of the open loop gain according to Eq. 3.13 under consideration of the described relationship of pole, zero and loop bandwidth.

3.4.2. Step Response

The evaluation of the third order loop's step response with a pole-zero-bandwidth-constellation determined by Eq. 3.14 requires a closer look at the closed loop transfer function which can be established by the help of some arithmetic manipulations as:

$$G_{CL}(s) = \omega_{LBW}^2 \cdot \sqrt{b} \cdot \frac{s + \dfrac{\omega_{LBW}}{\sqrt{b}}}{(s + \omega_{LBW}) \cdot \left[s^2 + s \cdot \omega_{LBW} \cdot (\sqrt{b} - 1) + \omega_{LBW}^2 \right]}$$

(3.18)

The transient behavior is determined by the position of the closed loop singularities, which can be effortlessly identified by means of the latter Equation:

- A closed loop zero sits on the negative real axis at $\omega_{LBW} / \sqrt{b}$.
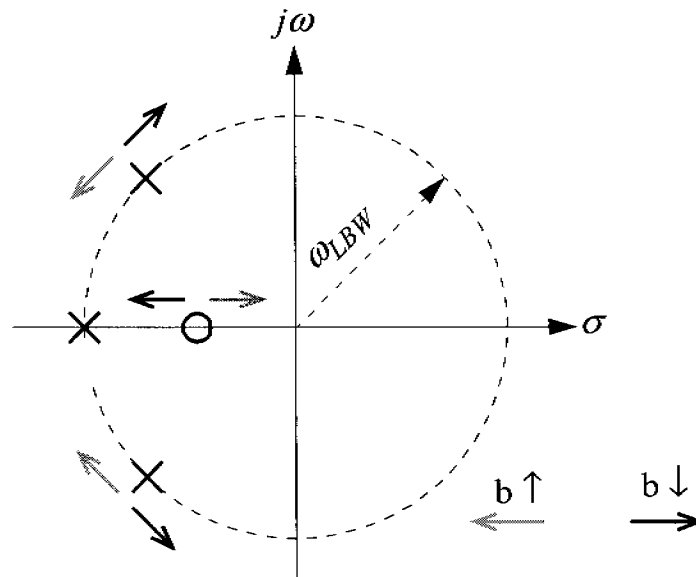
Figure 3.6: Closed loop pole and zero constellation of the particular third order type II PLL with the b factor being approximately 5.8. Arrows indicate the locus of the poles and zero with changing b factor.

- A closed loop pole is present on the negative real axis at ω_{LBW}.

- Two more closed loop poles exist with their location being dependant on the b parameter. They form a conjugate complex pole pair located on a negative half plane circle with radius identical to the loop bandwidth. Prerequisite for the existence of the conjugate pole pair is a b parameter smaller than 9. The pole quality factor pair is equal to the inverse of $\sqrt{b} - 1$.

Fig. 3.6 sketches the location of the closed loop zero and poles. Arrows indicate the locus of the singularities with changing b parameter. It is interesting to note that the location of the singularities is, besides scaling by the loop bandwidth, fully determined by the b parameter. Determination of an optimum b parameter is therefore the key to a fast step response. Before heading to a quantitative description of step response, some qualitative considerations may help to gain more insight into the optimum choice. Choosing the b parameter very large results in an overdamped system with a slow exponential settling. Besides slow settling, the system will also suffer from weak spurious tones suppression since the large b factor moves the open loop pole towards higher frequencies, resulting in poor high frequency filtering capability. On the other side, strong undercritical damping by choice of a very small b factor provides good filtering of the charge pump

current pulses, but causes strong ringing with an accordingly long settling time. Critical to moderate undercritical damping with *b* factors slightly smaller than 9 are therefore highly preferred.

The unity step response of an underdamped third order system can be derived by applying Laplace theorems as:

$$
f_{out}' = 1 - \frac{(\sqrt{b}-1)\cdot e^{-\omega_{LBW}\cdot t}}{\sqrt{b}-3}
$$

$$
- \frac{2\cdot e^{-\frac{\sqrt{b}-1}{2}\cdot \omega_{LBW}\cdot t}\cdot \cos\left(\frac{\sqrt{3-b+2\cdot\sqrt{b}}}{2}\cdot \omega_{LBW}\cdot t\right)}{\sqrt{b}-3} \tag{3.19}
$$

The frequency error after a step at the reference input is composed of an exponential part caused by the negative real pole and a damped sinusoidal component due to the conjugate complex pole pair, whereas the latter dominates the settling error. More important than the kind of frequency settling is the required time until the output reaches the wanted frequency within a specified tolerance. Based on the envelope of the frequency error, the required settling time τ_s can be computed as a function of the relative settling error. Fig. 3.7 shows three graphs with the *b* parameter ranging from 3.8 to 7.8. It can be seen that the required settling time for a particular settling error increases with decreasing *b* factor, which is a result of stronger ringing. A 45° phase margin is often considered as a balanced compromise between settling and spurious tones suppression. Fig. 3.7 provides therefore a curve with the *b* parameter set to *b=5.8*, the value which corresponds to 45° phase margin.

It is time to check whether the third order type II loop suffices the settling time requirements of GSM systems, as they were postulated in Chapter 2. Remember the required 90Hz settling accuracy after a 45MHz frequency step which must be fulfilled within 200µS if high speed data services are targeted. The curves provided in Fig. 3.7 can be used to determine the minimum loop bandwidth required to satisfy the latter specification. It can be seen that a loop bandwidth of 20kHz is wide enough to meet the specification, assuming a loop with a phase margin of 45°. However, nonlinear frequency acquisition effects, tolerances of the loop components, phase sampling effects as well as nonlinearity of the phase detector
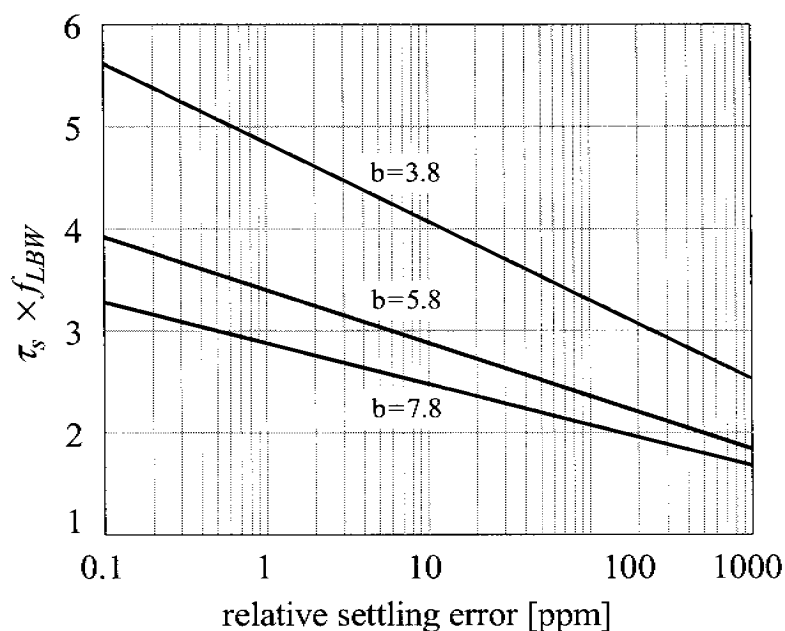
Figure 3.7: Normalized settling time versus relative settling error of the particular third order PLL.

gain may slow down the settling. An increase to 30kHz or 40kHz is appropriate to regard these effects.

3.4.3. Noise Considerations

The synthesizer phase noise spectrum is raised mainly by three noise sources:

- *RF oscillator noise*
- *Reference oscillator noise*
- *Loop filter noise*

While the impact of the former two noise sources is determined by the closed loop and error transfer function, the amount of phase noise caused by the third contributor can be controlled by the loop filter resistor. In order to reduce the latter's noise contribution, the resistor value can be lowered, but the charge pump current needs to be increased at the same time to maintain the loop bandwidth according to Eq. 3.17. Besides the goal to keep loop filter noise moderate, a no larger than necessary charge pump current is favored. A thorough understanding of how resistor noise impacts the synthesizer phase noise spectrum is required to find a balanced resistor-current combination.

In addition to the two oscillators, the phase locked loop contains a bunch of active devices required to setup the frequency divider and the charge pump. Although both building blocks could potentially degrade synthesizer noise performance, experiments show that frequency divider noise is significantly below the RF oscillator phase noise level. Charge pump noise may be more critical, but since the current sources are disconnected from the loop filter most of the time, the average noise injected into the loop filter is strongly reduced by a low duty cycle. Thus, synthesizer phase noise is in general neither affected by the divider nor the charge pump.

The mechanism which converts the resistor's voltage noise into phase noise requires a closer look. Resistor noise causes a contamination of the RF oscillator tuning voltage, which is transferred to the oscillator output by frequency modulation. Assuming a tuning voltage consisting of a DC part superposed by noise as

$$V_{tune} = V_{DC} + n(t), \tag{3.20}$$

the output of the RF oscillator driven by the latter voltage can be formulated as

$$V_{synth}(t) = A \cdot \cos\left[2\pi \cdot K_{VCO} \cdot \left(V_{DC} \cdot t + \int n(t)\, dt\right)\right] \tag{3.21}$$

with A being the RF oscillator's constant amplitude. Hence, the excess phase, i.e. the phase deviation from the linear phase trajectory, is proportional to the integrated voltage noise at the tuning node. RF oscillator phase noise power spectral density and the voltage noise power spectral density are thus related by the following identity

$$S_\phi = S_{Vtuning} \cdot \left|\frac{2\pi \cdot K_{VCO}}{s}\right|^2 = S_{Vtuning} \cdot \frac{K_{VCO}^2}{f_m^2} \tag{3.22}$$

with $S_{Vtuning}$ being the voltage PSD at the tuning node. The relationship between SSB phase noise $L(f_m)$ and phase noise PSD S_ϕ was derived in Chapter 2 (compare with Eq. 2.4). The identity is repeated here for the reader's convenience:

$$L(f_m) = 10 \cdot \log\left(\frac{S_\phi}{2}\right) = 10 \cdot \log\left(\frac{S_{Vtuning} \cdot K_{VCO}^2}{2 \cdot f_m^2}\right) \tag{3.23}$$
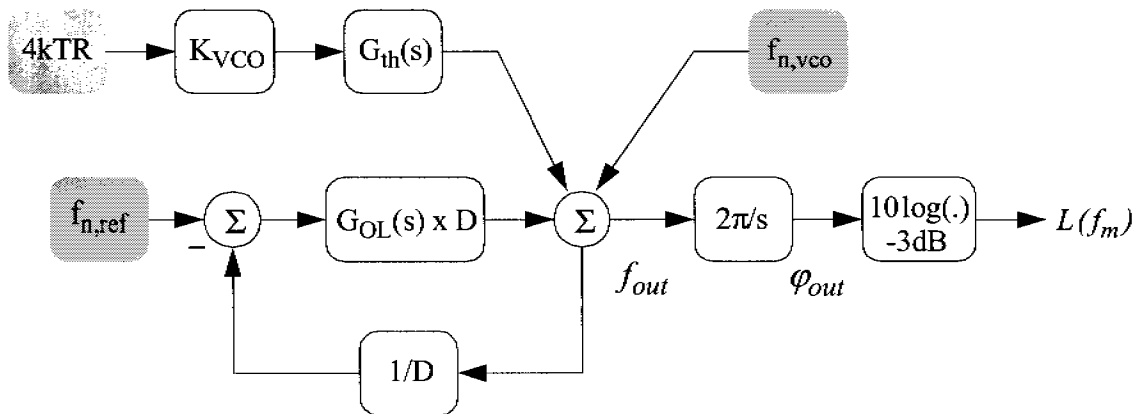
Figure 3.8: Linear PLL noise model with the three main contributor of phase noise.

Eq. 3.23 predicts the RF oscillator phase noise in the presence of noise at the tuning input. Of significance is the response to white noise, which appears as phase noise with a slope of -20dB/dec. Since most oscillator's phase noise exhibit a slope of -20dB/dec over a wide range, a noisy oscillator can be modelled by an ideal, noise free oscillator and a white noise source at the tuning input.

Considering these concepts, the linearized phase locked loop model of Fig. 3.4 can be extended to evaluate the frequency synthesizer's phase noise. Fig. 3.8 shows the proposed PLL noise model, which contains the three dominant noise sources and displays directly phase noise in dBc/Hz at the output. The model provides a powerful instrument to validate quickly the synthesizer phase noise spectrum and visualizes the contribution of the individual noise sources. Notice, that RF and reference oscillator phase noise is modelled by noise sources $f_{n,vco}$ and $f_{n,ref}$ with the somewhat bizarre unit *"Hertz squared per bandwidth"*. Conversion from the common dBc/Hz measure to the latter unit seems to be a unnecessary step, however, it allows the direct comparison of resistor caused noise with oscillator noise. Synthesizer phase noise of the third order type II phase locked loop is discussed in the remaining text in more detail by the help of the proposed noise model.

One of the reasons for the introduction of a second order loop filter is the superior noise performance. Thermal noise of the loop filter resistor is attenuated by a lowpass section, labeled in Fig. 3.8 as G_{th}. The latter can be found as:

$$G_{th}(s) = \frac{b-1}{b} \cdot \frac{\omega_p}{s + \omega_p} \qquad (3.24)$$

The suppression of loop filter noise at modulating frequencies above the pole frequency allows the use of a larger resistor compared to the second order loop. Notice that loop filter noise is also suppressed within the loop bandwidth by the loop gain. The remaining band from ω_{LBW} to ω_p lacks any suppression, phase noise is therefore dominated by loop filter noise in this region. This can be tolerated in most cases.

The discussion of noise will be closed by an analysis of the individual roll-off factors of the main noise contributors. It is assumed that free running oscillator phase noise shows a slope of -20dB/dec, an assumption which holds for crystal and LC-tank oscillators over a wide range. Phase noise contribution of the individual noise sources can be found by inspection of the proposed noise model as

$$L_{filt}(f_m) = 20 \cdot \log \left| \sqrt{4kTR} \cdot G_{th} \cdot G_1 \cdot K_{VCO} \cdot \frac{2\pi}{s} \right|_{s = j \cdot 2\pi \cdot f_m} - 3dB$$

$$L_{vco}(f_m) = 20 \cdot \log \left| \sqrt{f_{n,vco}} \cdot G_1 \cdot K_{VCO} \cdot \frac{2\pi}{s} \right|_{s = j \cdot 2\pi \cdot f_m} - 3dB \qquad (3.25)$$

$$L_{ref}(f_m) = 20 \cdot \log \left| \sqrt{f_{n,ref}} \cdot G_{CL} \cdot D \cdot K_{VCO} \cdot \frac{2\pi}{s} \right|_{s = j \cdot 2\pi \cdot f_m} - 3dB$$

with $L_{filt}(f_m)$, $L_{vco}(f_m)$ and $L_{ref}(f_m)$ representing phase noise caused by the loop filter, the RF oscillator and the reference oscillator respectively. In order to understand the individual roll-off factors, the error transfer function G_1 of the special case third order loop is needed, which can be found as:

$$G_1(s) = \frac{s^2 \cdot (s + \omega_{LBW} \cdot \sqrt{b})}{(s + \omega_{LBW}) \cdot \left[s^2 + s \cdot \omega_{LBW} \cdot (\sqrt{b} - 1) + \omega_{LBW}^2 \right]} \qquad (3.26)$$

Notice the two ideal zeros in the error transfer function. Keeping further the nature of the transfer functions G_{CL} and G_1 in mind, the qualitative growth and decline of the individual phase noise contributor can be stated:

- *Reference oscillator*:
  Reference oscillator phase noise is multiplied by D and attenuated by -40dB/dec outside the loop bandwidth (due to three poles and one zero
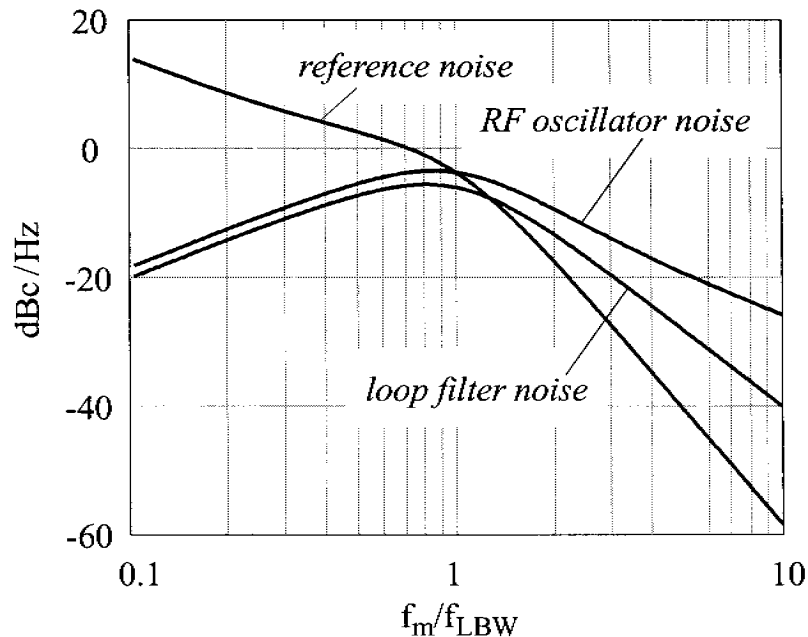
Figure 3.9: Phase noise at the PLL output produced by reference oscillator, RF oscillator and loop filter noise.

present in the closed loop transfer function). Hence, closed loop phase noise declines within the loop bandwidth with -20dB/dec and with -60dB/dec outside the loop bandwidth.

- *RF oscillator phase noise:*
  RF oscillator phase noise is suppressed within the loop bandwidth by the open gain, but is growing with 20dB/dec. This behavior results from a 40dB/dec increase of error transfer function (due to the two zeros ideal zeros). Roll-off outside the loop bandwidth is -20dB/dec as stated earlier.

- *Loop filter noise:*
  Behaves within the loop bandwidth like the RF oscillator, but falls at modulated frequencies larger than ω_p with -40dB/dec due to the additional lowpass section provided by the shunt capacitor.

Fig. 3.9 shows a qualitative plot of the individual phase noise contributors with the *b* parameter set to *b=5.8*.

3.5. Sampling Effects of the Charge Pump PLL

So far, the reference frequency harmonics of the charge pump current have been ignored. This simplification is valid as long as the low pass characteristic of the

loop strongly suppresses the high frequency components of the charge pump current. The linearized PLL model holds therefore only for loops with a small loop bandwidth to reference frequency ratio. As a rule of thumb, the loop bandwidth must be limited to one tenth of the reference frequency or lower to circumvent sampling effects [5]. A large loop bandwidth however is often required to meet the synthesizer settling time specification. A model which takes the switching nature of the charge pump into account is therefore required to characterize fast settling PLLs.

Sampling of the error phase causes in general a reduction of loop stability. The sampling process is accompanied by folding of high frequency parts of the open loop gain to lower frequencies, an unwanted effect known as aliasing. The additional low frequency gain is responsible for a reduction of the stability margins, which finally alter the step response of the loop. Stronger ringing with an accordingly increased settling time must be expected due to the reduced phase margin. The PLL settling time drops therefore over a large range of increasing loop bandwidth, reaches a minimum and starts to increase once the loop bandwidth is chosen too large. The increase of settling time takes place because the reduction by wider loop bandwidth is overwhelmed by the reduction of stability margins caused by aliasing. Optimum step response can be expected from a PLL with a loop bandwidth to reference frequency ratio of around one tenth. A synthesizer optimized for fast settling operates therefore in a region where sampling effects start to play a minor role, but where the loop can be still characterized by the linear model for the most part. The concepts derived from the linear model are therefore still relevant for fast switching synthesizers, but the expected moderate change of loop behavior should be qualified by the help of an appropriate model extension.

The characterization of sampling effects of the charge pump PLL is more demanding than e.g. sampling effects of the sample and hold phase detector PLL. The analysis is impeded by an error phase dependant sampling rate and error phase dependant duration of the current pulses. The error phase dependent sampling rate stems from the fact that the PFD can be triggered by both the reference oscillator and the frequency divider. A constant sampling rate determined by the period of reference frequency signal occurs only, if the charge pump is activated by the reference oscillator, i.e. if the error phase is positive. On the other side, non periodic sampling results from charge pump activation by the non periodic
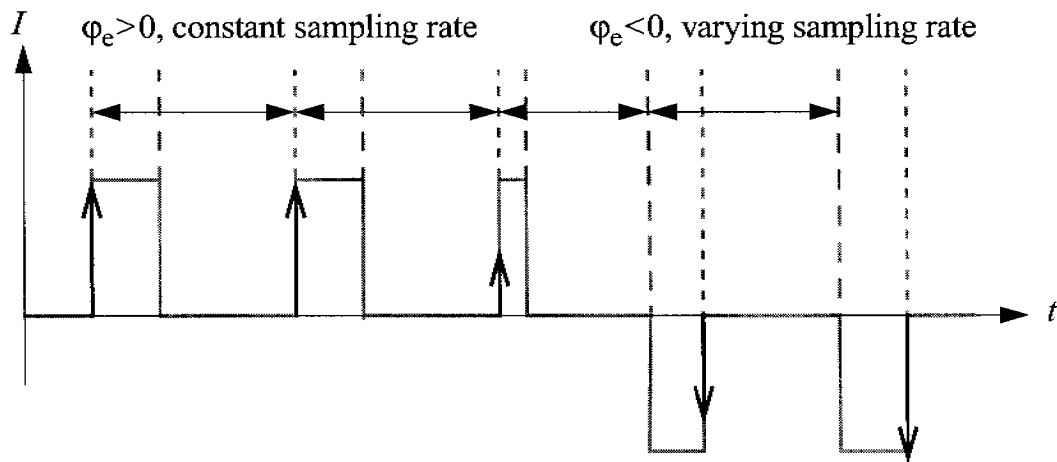
Figure 3.10: Rectangular shaped charge pump current (grey lines) and approximation by a train of weighted dirac impulses.

frequency divider output. This fact is illustrated in Fig. 3.10. Fortunately, the lack of constant sampling can be most often ignored without major loss of accuracy. A sampled model of a third order type II charge pump PLL was presented by Gardner [5]. While constant sampling is presumed, error phase dependant length of the current pulses is considered by computation of the exact occurrence of the PFD triggering events. Gardner derived difference equations of the output frequency and error phase at the sampling instants. Stability limits can be obtained from taking the z-transforms of the difference equations and applying standard methods of z-domain analysis. While this approach provides the ultimate sampling stability limit, calculation of the step response is more difficult. The difficulty stems from the fact that the z-transform is prone to numerical problems, a fact which holds especially for higher order systems [6].

Although the approach proposed by Gardner represents the most rigorous treatment of charge pump sampling effects, the derived Equations are not very meaningful to PLL designers because most PLLs are not operated close to the sampling stability limit. An extension of the PLL characterization by means of s-domain transfer functions is more beneficial in most cases. Such an approach is outlined in [6]. The analysis is based on a constant sampling rate. On top of this, the charge pump current is approximated by ideal current impulses rather than rectangular current pulses of finite length, as shown in Fig. 3.10. The sampled error phase can then be described by a train of weighted Dirac impulses as

$$[\varphi_e(t)]^* \;=\; \varphi_e(t) \cdot T_s \cdot \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \tag{3.27}$$

with $[.]^*$, $\delta(.)$ and T_s representing impulse sampling, dirac impulse and sampling period respectively. The spectrum of the sampled error phase can be derived as:

$$[\varphi_e(s)]^* \;=\; \sum_{k=-\infty}^{\infty} \varphi_e\!\left(s - jk \cdot \frac{2\pi}{T_s}\right) \tag{3.28}$$

The underlying mathematics leading to Eq. 3.28 are presented in Appendix 3.A The latter Equation says that the sampled error phase spectrum is identical to the spectrum of the time continuous error phase plus frequency shifted replicas. Consideration of this outcome as well as properties of sampling allow the calculation of the impulse sampled closed loop transfer function as

$$\frac{f_{out}{}'(s)}{[f_{ref}(s)]^*} \;\approx\; \frac{G_{OL}(s)}{1 + [G_{OL}(s)]^*} \tag{3.29}$$

with

$$[G_{OL}(s)]^* \;=\; \sum_{k=-\infty}^{\infty} G_{OL}\!\left(s - jk \cdot \frac{2\pi}{T_s}\right) \tag{3.30}$$

The interested reader finds a derivation of Eq. 3.29 and Eq. 3.30 in Appendix 3.B. Comparison of Eq. 3.29 with the time continuous counterpart Eq. 3.3 reveals that sampling primary affects the denominator of the transfer function. This fact unquestionably underlines the postulation that sampling affects loop stability and step response behavior.

Stability can be assessed by numerical computation of Eq. 3.30 using the standard Bode plot method without the need for computation of z-transforms. The infinite sum can be truncated without effectual loss of accuracy due to the inherent band limitation of the open loop gain. Although the proposed method does not directly provide information about the dynamic behavior of the PLL, transient response can be estimated from a a comparison of the continuous open loop gain with the sampled open loop gain described by Eq. 3.30. A substantial degradation of phase margin is an indication for a too large loop bandwidth. Minor

degradation of phase margin can be compensated by a slight readjustment of loop parameters.

3.6. Charge Pump Artifacts

Virtually every charge pump suffers from a couple of artifacts deteriorating synthesizer performance. Probably the most prominent artifact is the drop of loop gain. The drop is located at crossover of the error phase and affects hence type II loops. Although the loop gain can be linearized by appropriate circuits, the linearization is in general accompanied by an increase of spurious tones. A thorough understanding of charge pump artifacts is therefore coercive to PLL designers.

3.6.1. The Dead Zone

Once a type II charge pump PLL reaches steady state, the charge pump stops producing current pulses and a constant output frequency without any spurious tones is established at the synthesizer output. In reality, however, the PLL continuously tends leave locked condition due to disturbances caused by the various noise sources. In order to maintain phase locking, the loop has to generate correction current pulses to force the RF oscillator back to the desired frequency. The length of the required pulses lies in the sub-nanosecond range, a value which is due to circuit limitations beyond the capabilities of the charge pump. It can be observed that the charge pump injects too few charge into the loop filter to counteract the disturbances. Consequently, the PFD-CP gain K_P drops significantly below the desired value $I_{CP}/2\pi$ or even vanishes completely at crossover. The region of very low or zero phase detector gain is often called dead zone or phase detector crossover distortion [7].

The appearance of a phase detector dead zone harms the PLL operation. The loop looses all its open loop gain in steady state and accordingly its ability to control the RF oscillator. The output frequency can fluctuate randomly within a range determined by the width of the dead zone. Synthesizer phase noise is expected to increase to the level of the free running RF oscillator within this range [8].

The width of the dead zone, which is given by the shortest producible current pulse, is strongly related to the charge pump implementation and the used IC technology. Unfortunately, even charge pumps implemented in deep sub-micron

CMOS technologies enabling very fast circuits do not eliminate the dead zone completely. Modifications of the basic charge pump topology are necessary to tackle the problem.

One approach is to inject a constant current into the loop filter. This current pushes the steady-state error phase away from crossover, moving thereby the operating point into the linear region of the charge pump. While the injected current removes the dead zone if the injected current is chosen large enough, strong spurious tones are generated due to the resulting non-zero steady state error phase. The inherent low spurious tone characteristic of the type II loop is annihilated by this approach.

The dead zone problem can be hurdled without a strong deterioration of spurious tone performance. Virtually all implementations of charge pumps which claim dead zone free operation are based on the same underlying idea: A small average charge pump current can be produced not only by very short current pulses, but also by concurrent activation of both charge pump current sources for a much longer period and shutting down one of them slightly earlier. No net charge is injected into the loop filter during the simultaneous activation of the current sources, since the two current sources cancel each other. Charge transferred to the loop filter is determined by the time difference of the current source activation durations, a difference which can be made arbitrarily small without the need for fast charge pump circuits. Extremely short current pulses seen by the loop filter can be generated by this simple trick without the need for fast circuits.

In order to enable simultaneous activation of the charge pump current sources, the state transition diagram of the elementary PFD shown in Fig. 3.2 must be extended by a fourth state. Fig. 3.11 shows two possible arrangements, whereby the additional state is labeled as SHORT. While the SHORT state may follow the OFF state, an arrangement which was implemented by a PFD circuit proposed in [8], most implementations introduce the SHORT state prior to entering the OFF state. Neither of the two solutions are superior, but the latter arrangement is less hardware intensive and is therefore most often the preferred choice.

The steady-state spectrum of a type II charge pump PLL with the described modified PFD is inherently free of spurious tones. However, non perfect cancellation of the two charge pump current sources during simultaneous activation is responsible for the generation of residual spurious tones. Besides keeping the current sources mismatch small, a no longer than necessary simultaneous current source
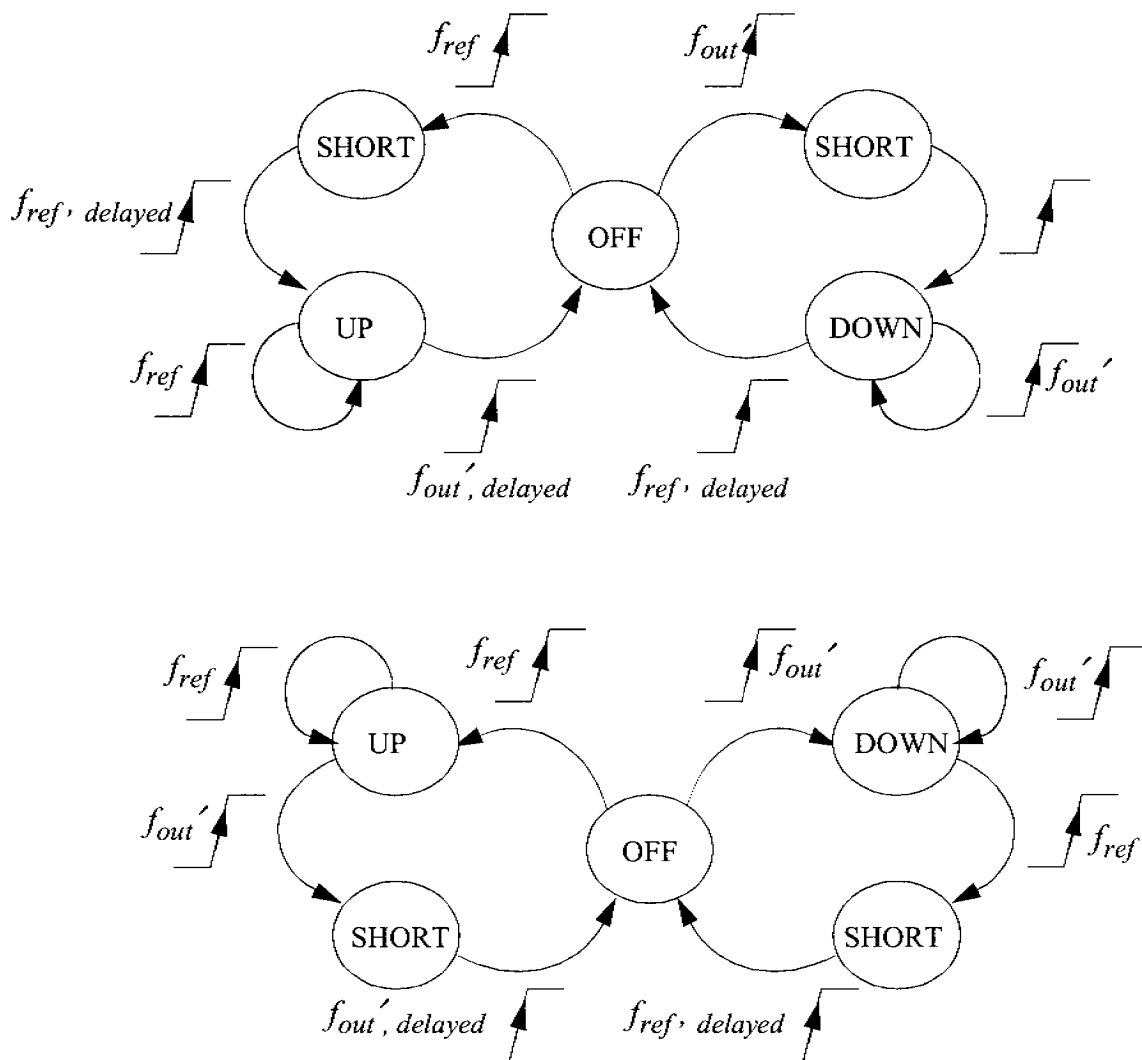
Figure 3.11: State transition diagram of phase frequency detectors with additional SHORT state.

activation helps to minimize spurious tones further. While Section 3.6.2 identifies the minimum duration of the SHORT state required to linearize the charge pump, Section 3.6.3 investigates the residual spurious tones to be expected from current source mismatch as well as from leakage.

3.6.2. Optimum Charge Pump Linearization

The occurrence of the dead zone is strongly related to the shape of the charge pump current. The charge pump current was modelled so far by rectangular shaped current pulses with the duration t_{pulse} determined by the PFD triggering

signals. These current pulses produce an average current which scales linearly with the length of the current pulse and consequently with the error phase. No gain nonlinearity must be expected.

Real charge pump circuits however are not able to provide rectangular shaped currents. Slewing of PFD signals responsible for on- and off-switching as well as other effects prevent an abrupt change of the charge pump current. A finite rise and fall time is required to change the charge pump current from zero current to the full current I_{CP} and vice versa. A more realistic model of a charge pump delivers therefore trapezoidal rather than rectangular shaped current pulses. The average current of such a charge pump scales linearly with the error phase as long as the occurrence of the PFD trigger events are separated by a duration longer than the rise time of the charge pump current. However, if the second triggering signal forces the PFD back to the OFF state before the current has reached the final value I_{CP}, the current shape changes from trapezoidal to triangular. The average current scales then quadratically with the error phase as:

$$I_{av} = f_{ref} \cdot I_{CP} \cdot \frac{t_{pulse}^2}{t_{r,f}} \qquad (3.31)$$

$t_{r,f}$ represents the rise and fall time which are assumed to be identical. The upper half of Fig. 3.12 illustrates the charge pump current if the time gap of the PFD triggering signal is smaller than the rise and fall time. Case I, which presumes a charge pump circuit without additional SHORT state, shows the triangular current responsible for a gain drop illustrated in the lower half of Fig. 3.12. Notice that the gain collapses completely at crossover. Case II and case III of Fig. 3.12 illustrate the current shapes and phase detector characteristics of a charge pump with a modified PFD. The SHORT state is chosen prior to the entrance of the OFF state. Notice the activation of the second current source which is shown in stippled lines. While case II remains in the SHORT state for a duration t_{act} which is smaller than the rise and fall time $t_{r,f}$, case III presents curves of an arrangement with t_{act} chosen larger than $t_{r,f}$. The former case extends the activation time of current sources, but the current shapes are still triangular. The region of vanishing phase detector gain can be avoided, but the phase detector characteristic is still strongly nonlinear. Contrary case III, which provides a linear phase detector characteristic. This is achieved by extension of the current pulses so that their shape remains trapezoidal despite the small error phase.
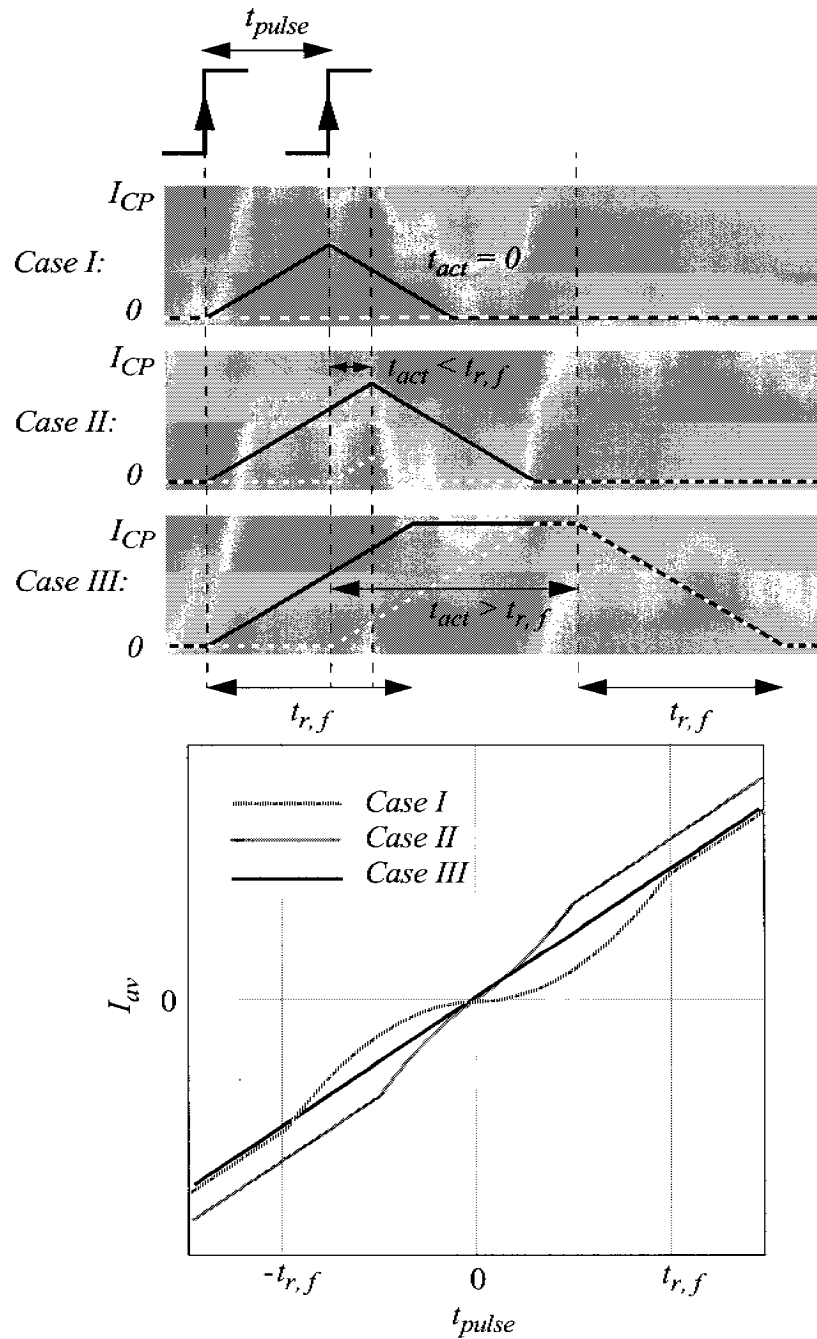
Figure 3.12: Finite raise/fall time charge pump currents at small error phase (upper Figure) and according phase detector characteristic (lower Figure).

Case I: Elementary PFD suffering from dead zone.
Case II: Modified PFD, but duration of the SHORT state is chosen too short.
Case III: Modified PFD providing linear phase detector characteristic.

The plots shown in Fig. 3.12 demonstrate that the dead zone can be removed by extension of the current pulses so that they become larger than the current sources rise time. The optimum duration of the SHORT state must be chosen identical to the rise time. A longer activation time does not further improve linearity.

3.6.3. Spurious Tones due to Mismatch and Leakage

Prediction of the level of spurious tones is probably the most challenging part of synthesizer analysis because these tones originate from a variety of circuit artifacts rather than from fundamental PLL operation. Despite the importance of the subject, spurious tones are often excluded from analytical considerations. This is not surprising, since any analysis is confronted with the parasitic nature of the processes responsible for the generation of spurs. Nevertheless, some fundamental rules on how charge pump nonidealities influence spurious tones can be found despite the difficulties.

While any PLL circuit may contribute to spurious tones by parasitic coupling, measurable spurs are most often exclusively caused by switching of the PFD/CP combination. Spurious tones are hence located at the reference frequency and reference frequency harmonics. The most dominant effects responsible for spurious tones generation in the context of charge pump switching can be identified as current source mismatch, component leakage, clock-feed-trough (CFT) and current sharing. While the last two effects must be tackled by appropriate circuit techniques, issues which will be discussed in Chapter 7, the impact of mismatch and leakage depends strongly on the duration of the simultaneous activation of the charge pump current sources t_{act}. This section is aimed to the identification of the upper limits of mismatch, activation time and leakage regarding to spurious tones specifications.

Any type of charge pump artifact generates spurious tones by the same underlying principle. The artifacts generate error currents which flows into the loop filter. The DC component of the error current shifts the tuning voltage away from steady-state. In order to maintain phase locking, the PLL counteracts the tuning voltage drift through activation of a charge pump, leading to a non zero steady state error phase. While the counteracting charge pump current removes the DC part of the error current and stabilizes thereby the tuning voltage, reference frequency and higher order components remain at the charge pump output and cause
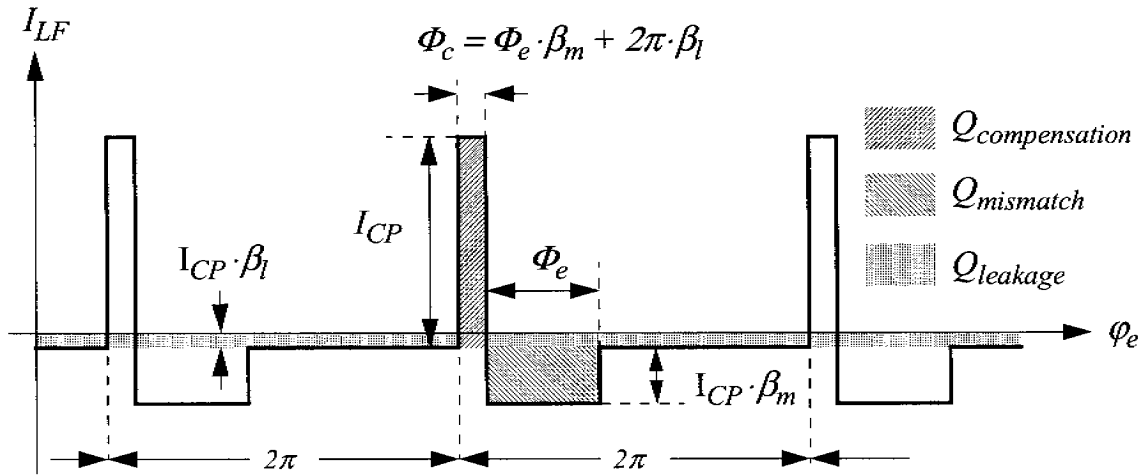
Figure 3.13: Steady state loop filter current driven by a charge pump suffering from current source mismatch and leakage.

a voltage ripple on the tuning voltage. The sideband to carrier ratio SB/C of the spurious tones depends on the tuning voltage ripple according to the following formula

$$\frac{SB_n}{C} = \left(\frac{V_{tuning,\, n \cdot f_{ref}} \cdot K_{VCO}}{2 \cdot n \cdot f_{ref}} \right)^2 \qquad (3.32)$$

with $V_{tuning,n \cdot f_{ref}}$ representing the amplitude of the (n-1)-th harmonic of the tuning voltage ripple.

Computation of spurious tones according to Eq. 3.32 requires the study of the charge pump current in the course of a cycle. Fig. 3.13 sketches the current flowing into the loop filter in the presence of charge pump mismatch and leakage. Three regions with different currents can be distinguished which corresponds to the PFD's OFF, SHORT and UP (DOWN) states. Current during the SHORT state is dominated by current source mismatch which produces in each cycle an error charge

$$Q_{mismatch} = \frac{\Phi_e}{2\pi} \cdot \frac{\beta_m \cdot I_{CP}}{f_{ref}} \qquad (3.33)$$

with β_m being the current source mismatch parameter and Φ_e the activation angle determined by t_{act}. Another part of the overall error charge is generated by a con-

stant leakage current which determines the current during OFF state. The error charge due to leakage can be found as

$$Q_{leakage} = \frac{\beta_l \cdot I_{CP}}{f_{ref}}$$ (3.34)

with β_l representing the ratio of leakage to charge pump current. The overall error charge is compensated by the charge pump current during the UP (DOWN) state. The UP respectively DOWN state activation angle can be found as:

$$\Phi_c = \Phi_e \cdot \beta_m + 2\pi \cdot \beta_l$$ (3.35)

Spurious tones caused by mismatch and leakage can be found numerically by the help of Eq. 3.32 and the tuning voltage spectrum, which can be obtained from the fast fourier transform (FFT) applied to the current according to Fig. 3.13. The levels of spurious tones for the special case third order type II PLL are shown in Fig. 3.14. It's not surprising that the strongest spurious tone appears at the reference frequency since higher order tones are more attenuated by the loop filter. The graph provided in Fig. 3.14 shows further, that a high performance charge pump is required if the strongest spurious tone shall be kept below -50dBc, i.e. the matching of the current source must be better than 5%, the simultaneous activation of both current sources must be limited to 5ns and the leakage current must be limited to 5nA. These numbers indicate that charge pump artifacts are very critical to the generation of spurious tones if the mentioned sideband to carrier ratio is targeted. Besides the required extraordinary high performance, the numerical computation exhibits further a strong sensitivity of the spurious tone level to β_m, β_l and Φ_e. An analytical rather than numerical computation of sideband to carrier ratio is highly desirable to gain more insight on how the latter parameters are linked to spurious tones.

The dominant spurious tone at the reference frequency can be found by extensive algebraic manipulations which are outlined in Appendix 3.C. The analysis, which considers the special case third order type II charge pump PLL in the presence of current source mismatch, yields in a compact formulation of the sideband to carrier ratio at the reference frequency. The result is repeated here:

$$\frac{SB_1}{C} = \frac{\Phi_e^4 \cdot \beta_m^2 \cdot D^2 \cdot b}{4} \cdot \left(\frac{f_{LBW}}{f_{REF}}\right)^4$$ (3.36)
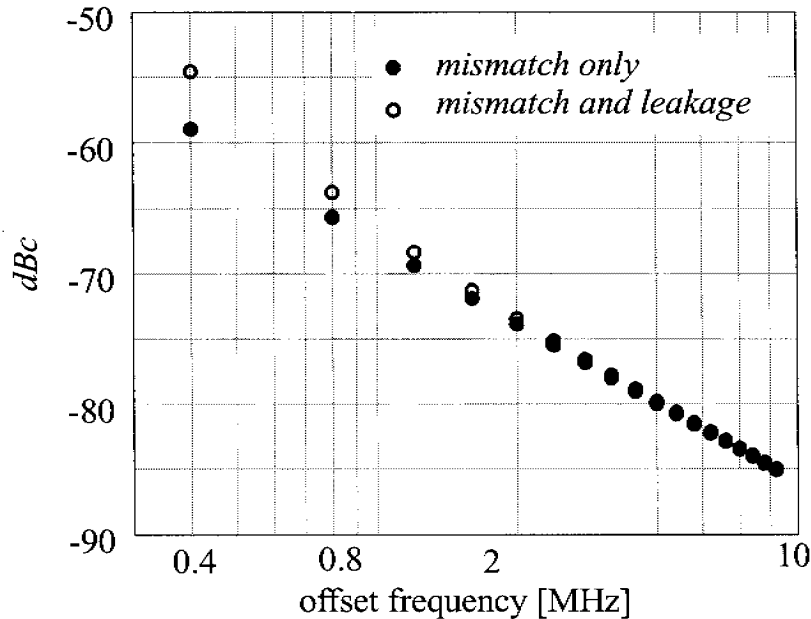
reference frequency	400kHz
loop bandwidth	40kHz
output frequency	4GHz
current source mismatch	5%
SHORT state activation time	5ns
leakage to charge pump current ratio	5nA / 1.15mA

Figure 3.14: Simulation of the spurious tones caused by charge pump current source mismatch and leakage. Parameters of the employed third order type II loop are listed in the Table.

The compactness of the above formula permits to draw a couple of fundamental conclusions:

- *Loop bandwidth dependency:*
  A reduction of loop bandwidth suppresses the reference frequency spurious tone. Although the loop filter provides only a single pole, a fact which would presume a 6dB per octave decline of spurious tones versus loop bandwidth, the reference frequency spurious tone versus loop bandwidth scales with -12dB/oct. Fast settling conflicts hence with spurious tone performance.

- *Activation time dependency*:
  Doubling the pulse length raises the reference spurious tone by 12dB. The strong sensitivity makes small dead zone charge pumps mandatory for low spurious tones PLLs.

- *Mismatch dependency:*
  The reference spurious tone improves by 6dB if the current source mismatch can be halved.

- *Division ratio dependency:*
  A large frequency division ratio D raises the level of reference frequency spurious tone. Doubling the division ratio increases the spur by 6dB. Even worse, if the doubling of division ratio is accompanied by halving the reference frequency in order to improve the frequency resolution of the synthesizer by a factor two, an overall increase of 18dB must be expected. Synthesizers frequency resolution conflicts therefore with spurious tone performance.

The presented investigations provide the information to specify the charge pump with regard to mismatch and leakage. As mentioned in the introduction of this Section, spurious tones depend on a variety of unwanted effects. Designing the charge pump with sufficient low mismatch, activation time etc. does not guarantee compliance of the specification since other charge pump artifacts may overwhelm the considered effects.

3.7. Summary

Phase locked loops differ in the employed phase detector and the order and type of the loop. The infinite locking range as well as good spurious tones performance if embedded in a type II loop qualifies the charge pump PLL to high performance synthesizers. Decent step response, noise and spurs performance can be expected from a third order loop with optimum placed singularities. The four free loop parameters can be derived from the settling time and noise specifications so that the design procedure does not require any trial-and-error iterations.

The circuit implementation is challenged by two key components. First, the RF oscillator must provide low free running phase noise since the PLL does not provide any noise suppression outside the loop bandwidth. The second component critical to performance is the charge pump. The strong sensitivity of spurious tones to the analog performance of this block requires a careful design. The evaluation of this particular block is endeavoured by many parasitic effects which are

hard to estimate. Spurious tones performance remains often the troublesome unknown of many frequency synthesizer design.

References

[1] R. Best, *"Phase-Locked Loops, Design, Simulation and Applications"*, McGraw-Hill, 1999, pp. 323-336

[2] T. Laopoulos, C. Karybakas, *"A phase locked motor speed control system with sample-and-hold phase detector"*, IEEE Trans. on Industrial Electronics, Vol. 35, May 1988, pp. 245-252

[3] B. Razavi, *"Design of Monolithic Phase-Locked Loops and Clock Recovery Circuits - A Tutorial"*, appeared in *"Monolithic Phase-Locked Loops and Clock Recovery Circuits"*, IEEE Press, 1996, pp. 29-30

[4] Application Note AN-1001, *"An Analysis and Performance Evaluation of a Passive Filter Design Technique for Charge Pump PLL's"*, National Semiconductor, July 2001

[5] F. Gardner, *"Charge-Pump Phase-Locked Loops"*, IEEE Trans. on Communications, Vol. 28, Nov. 1980, pp. 1849-1858

[6] J. Crawford, *"Frequency Synthesizer Design Handbook"*, Artech House, 1994, pp. 203-244

[7] W. Egan, *"Test Your Charge-Pump Phase Detectors"*, Electronic Design, June 1978, pp. 134-137

[8] R. Ahola, S. Lindfors, J. Routama, K. Halonen, *"A Novel Phase Detector with No Dead Zone and Charge-Pump with Very Wide Output Range"*, Proc. IEEE European Solid State Circuits Conference, 1998, pp. 352-355

Appendix 3

3.A Impact of Impulse Sampling on the Spectrum

A time continuous signal $f(t)$ which undergoes impulse sampling can be described by a the summation of weighted dirac impulses:

$$[f(t)]^* = f(t) \cdot T_s \cdot \sum_{n = -\infty}^{\infty} \delta(t - nT_s) \qquad (3.A.1)$$

The in T_s periodic sampling function can be expanded by a Fourier series as:

$$T_s \cdot \sum_{n = -\infty}^{\infty} \delta(t - nT) = 1 + 2 \cdot \sum_{k = 1}^{\infty} \cos\left(kt \cdot \frac{2\pi}{T_s}\right) \qquad (3.A.2)$$

Remember Euler's famous formula:

$$\cos(x) = \frac{1}{2} \cdot \left(e^{jx} + e^{-jx}\right) \qquad (3.A.3)$$

Usage of Euler's formula and a rearrangement of indexes yields in the following identity:

$$T_s \cdot \sum_{n = -\infty}^{\infty} \delta(t - nT) = \sum_{k = -\infty}^{\infty} e^{-kt \cdot \frac{2\pi}{T_s}} \qquad (3.A.4)$$

Combination of Eq. 3.A.1 and Eq. 3.A.4 leads to an alternative representation of the impulse sampled time domain signal $f(t)$:

$$[f(t)]^* = f(t) \cdot \sum_{k = -\infty}^{\infty} e^{-kt \cdot \frac{2\pi}{T_s}} \qquad (3.A.5)$$

The spectrum of the sampled signal can be obtained by computation of the Laplace transform of Eq. 3.A.5:

$$[f(s)]^* = \int_0^\infty [f(t)]^* \cdot e^{-st} dt$$

$$= \int_0^\infty \left(f(t) \cdot \sum_{k=-\infty}^{\infty} e^{kt \cdot \frac{2\pi}{T_s}} \cdot e^{-st} \right) dt$$

$$= \sum_{k=-\infty}^{\infty} \int_0^\infty f(t) \cdot e^{-st + kt \cdot \frac{2\pi}{T_s}} dt$$

$$= \sum_{k=-\infty}^{\infty} f\left(s - jk \cdot \frac{2\pi}{T_s}\right)$$

(3.A.6)

The final results underlines the fact that impulse sampling causes folding of high frequency parts of the spectrum to lower frequencies.

3.B Impulse Sampled Closed Loop Transfer Function

Inspection of the PLL diagrams of Fig. 3.1 and Fig. 3.4 reveal the following two Equations:

$$f_{out}'(s) \cdot \frac{2\pi}{s} = [\varphi_e(s)]^* \cdot G_{OL}(s) \qquad (3.B.1)$$

$$[\varphi_e(s)]^* = \left[\frac{2\pi}{s} \cdot \left(f_{ref}(s) - f_{out}'(s)\right)\right]^* \qquad (3.B.2)$$

Combination of the latter two Equations and application of fundamental properties of a sampled signals lead to the following identities:

$$[\varphi_e(s)]^* = \left[\frac{2\pi \cdot f_{ref}(s)}{s} - [\varphi_e(s)]^* \cdot G_{OL}(s)\right]^*$$

$$= \left[\frac{2\pi \cdot f_{ref}(s)}{s}\right]^* - \left[[\varphi_e(s)]^* \cdot G_{OL}(s)\right]^* \qquad (3.B.3)$$

$$= \left[\frac{2\pi \cdot f_{ref}(s)}{s}\right]^* \cdot \frac{1}{1 + [G_{OL}(s)]^*}$$

$$f_{out}'(s) \cdot \frac{2\pi}{s} = \left[\frac{2\pi \cdot f_{ref}(s)}{s}\right]^* \cdot \frac{G_{OL}(s)}{1 + [G_{OL}(s)]^*} \qquad (3.B.4)$$

Assuming that the reference frequency signal is strictly band limited, the following approximation opens the way to the characterization of the sampled closed loop transfer function:

$$\left[\frac{2\pi \cdot f_{ref}(s)}{s}\right]^* \approx [f_{ref}(s)]^* \cdot \frac{2\pi}{s} \qquad (3.B.5)$$

The closed loop transfer function can be finally found by combination of Eq. 3.B.4 and Eq. 3.B.5 as

$$\frac{f_{out}'(s)}{[f_{ref}(s)]^*} \approx \frac{G_{OL}(s)}{1 + [G_{OL}(s)]^*} \qquad (3.B.6)$$

or alternatively by applying the result found in Appendix 3.A as:

$$\frac{f_{out}'(s)}{[f_{ref}(s)]^*} \approx \frac{G_{OL}(s)}{1 + \sum_{k = -\infty}^{\infty} G_{OL}\left(s - jk \cdot \frac{2\pi}{T_s}\right)} \qquad (3.B.7)$$

Notice that the sampling process affects only the denominator of the closed loop transfer function.

3.C Spurious Tones caused by Current Source Mismatch

The computation of the reference frequency spurious tone of a third order type II loop involves several algebraic steps. Required prerequisite is the knowledge of the current flowing into the loop filter. The habit of the current was discussed in Section 3.6.3 and is sketched in Fig. 3.13. The current can be described by a fourier series expansion since the current is periodic in 2π:

$$I(\varphi_e) = \sum_{k=1}^{n} I_{Ik} \cdot \cos(k\varphi_e) + I_{Qk} \cdot \sin(k\varphi_e) \tag{3.C.1}$$

The fundamental coefficients I_{I1} and I_{Q1} can be found as

$$I_{I1} = \frac{1}{\pi} \int_{-\pi}^{\pi} I(t) \cdot \cos(\varphi) \, d\varphi_e = \frac{I_{CP}}{\pi} \cdot \left[\sin(\Phi_e \beta_m) - \beta \cdot \sin(\Phi_e) \right]$$

$$\tag{3.C.2}$$

$$I_{Q1} = \frac{1}{\pi} \int_{-\pi}^{\pi} I(t) \cdot \sin(\varphi) \, d\varphi_e = -\frac{I_{CP}}{\pi} \cdot \left[1 + \beta_m - \cos(\Phi_e \beta) - \beta_m \cdot \cos(\Phi_e) \right]$$

yielding in the amplitude of the fundamental current component as:

$$I_{fund} = \sqrt{I_{I1}^2 + I_{Q1}^2} = \frac{2 \cdot I_{CP}}{\pi} \cdot \left[(1 + \beta_m + \beta_m^2) + \beta_m \cdot \cos(\Phi_e + \Phi_e \beta) \right. \tag{3.C.3}$$

$$\left. - (1 + \beta_m) \cdot \langle \cos(\Phi_e \beta_m) + \beta_m \cdot \cos(\Phi_e) \rangle \right]$$

The fundamental component of the tuning voltage ripple can be found by multiplying the latter result with the loop filter impedance. Before doing so, it is necessary to expand the cosine function by a Taylor series in order to replace the transcendental Equation by an algebraic term. The simplest Taylor series expansion of the cosine function yielding in a non zero fundamental current is found as a fourth order series:

$$\cos(x) \approx 1 - \frac{x^2}{2} + \frac{x^4}{24} \tag{3.C.4}$$

Combination of Eq. 3.C.3 and Eq. 3.C.4 as well as the earlier derived loop bandwidth Equation (compare with Eq. 3.17) yields in a term for the fundamental current:

$$I_{fund} \approx \frac{I_{CP}}{2\pi} \cdot \Phi_e^2 \cdot \beta_m \cdot (1+\beta) \approx \frac{I_{CP}}{2\pi} \cdot \Phi_e^2 \cdot \beta_m$$

$$= \frac{D \cdot \omega_{LBW}}{R \cdot K_{VCO}} \cdot \frac{b}{b-1} \cdot \frac{\Phi_e^2 \cdot \beta_m}{2\pi} \qquad (3.C.5)$$

The loop filter impedance can be found by evaluation of Eq. 3.10 at the reference frequency:

$$\left| Z_{LPF}(j2\pi f_{REF}) \right| \approx R \cdot \frac{b-1}{b} \cdot \frac{\omega_{LBW}}{\omega_{REF}} \cdot \sqrt{b} \qquad (3.C.6)$$

The fundamental component of the tuning voltage can be formulated effortlessly by combination of Eq. 3.C.5 and Eq. 3.C.6, leading to the following identity:

$$V_{fund} = I_{fund} \cdot \left| G_{LF}(j\omega_{ref}) \right| = \frac{\Phi_e^2 \cdot \beta_m}{2\pi} \cdot \frac{\sqrt{b} \cdot \omega_{LBW}^2 \cdot D}{K_{VCO} \cdot \omega_{REF}} \qquad (3.C.7)$$

The latter result allows the computation of the sideband to carrier ratio as:

$$\frac{SB}{C} = \left(\frac{V_{fund} \cdot 2\pi \cdot K_{VCO}}{2 \cdot \omega_{REF}} \right)^2 = \frac{\Phi_e^4 \cdot \beta_m^2 \cdot D^2 \cdot b}{4} \cdot \left(\frac{\omega_{LBW}}{\omega_{REF}} \right)^4 \qquad (3.C.8)$$

Chapter 4

Frequency Divider

The frequency divider forms together with the voltage controlled oscillator (VCO) the radio frequency section of the synthesizer. Power consumption of the latter is dominated by these two blocks due to the high frequency, which may be several decades larger than those of the remaining parts. In contrast to the VCO, where power consumption is of good use for the carrier to noise ratio, the frequency divider represents a pure digital circuit without impact on synthesizer performance as long as frequency divider phase noise is negligible compared to oscillator phase noise. Experiments show that this hold true for virtually any practical frequency divider implementation. The main design focus lies hence on power consumption minimization besides keeping interference with other sensitive analog building blocks as small as possible.

One of the main reasons why the phase locked loop based frequency synthesizer became the unchallenged leader in frequency synthesis lies in its digital programming capability. A new frequency can be commanded by modifying a digital register without the need for any analog tuning. Obviously it is the frequency divider which must provide the programmability since it is the only digital synthesizer part. Unfortunately, the programmability increases the complexity of the divider significantly. It is the high frequency combined with the need for programmability, embedded with limited available power, that makes the divider design a challenging task.

The strategy to find competitive solutions is twofold. First, architectures which shift as much programmability as possible to lower frequencies, simplifying thereby the high frequency part of the divider, must be found. Second, the remaining high frequency part must be power optimized by the help of appropriate logic styles and optimum transistor scaling. This Chapter explores both design aspects. The architectural and circuit investigations are accompanied by two design examples.

4.1. Frequency Division by Digital Counters

Frequency dividers translate an input signal of frequency f_{in} into a signal of lower frequency f_{out}. Digital counters are mostly suited to perform the translation. The input signal of the divider, which is driven by the VCO output if embedded in a PLL, can be fed (after being converted into an appropriate logic signal) to a counter clock input, decrementing the counter's value each active clock edge, until the counter reaches zero. The consequent clock will cause a preset value D to be loaded. Using this technique, the frequency at the divider output is lowered by $D+1$. Programmability is achieved by replacing the fixed-wired preset value by a freely programmable digital register. The operation principle of this first frequency divider is visualized in Fig. 4.1.
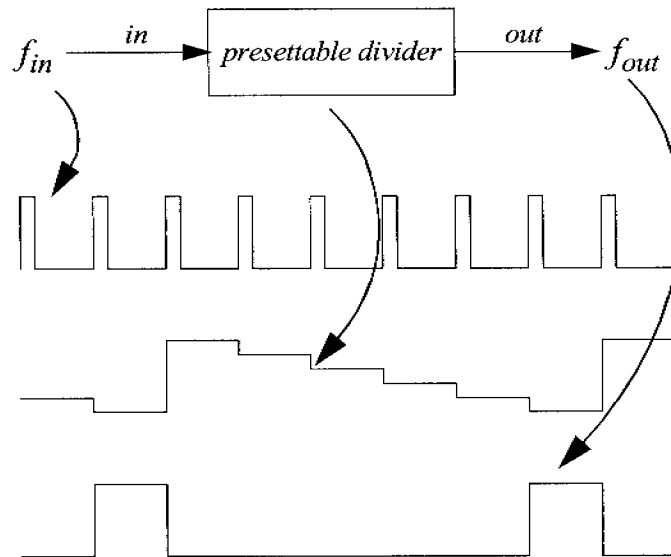
Figure 4.1: Principle of frequency division by a presettable counter. The staircase curve shows the divider's actual counter value.

The frequency divider of Fig. 4.1 is most often not suited for direct integration as a CMOS standard cell based design, especially if power consumption is critical. The high input frequency causes an excessive power consumption or even precludes the use of CMOS standard cells because of speed limitations. The frequency divider architecture shown in Fig. 4.1 must be replaced by a more advanced structure. The idea is to split the frequency divider into several circuits, mostly into two. A first circuit, the prescaler, is able to operate at the highest fre-

quency with moderate power consumption and lowers the speed requirements of the second circuit, allowing application of the full wealth of digital standard cells at reasonable power consumption. The question arises how to distribute the division ratios and the programmability among the two sub circuits. A large division ratio of the prescaler is highly desirable in order to lower the speed of the consequent (standard cell based) circuit as much as possible. More important, little or no programmability is desired in the prescaler, since all programmability increases complexity and power consumption. A minimum complexity prescaler provides no programmability at all (i.e. fixed division ratio), especially if its modulus is a binary power 2^N. A frequency divider built from a fixed modulus prescaler followed by a low frequency, programmable counter would be a very attractive solution in terms of consumption, but suffers from a dire drawback. While the first frequency divider can be programmed to any arbitrary division ratio, the programmability of the solution employing a fixed modulus prescaler is limited to division ratios with increments equal to the prescaler modulus. This, however, is not acceptable to most applications.

As conclusion, it is necessary to add some programmability to the prescaler. Section 4.2 demonstrates, that a 'seamless' (i.e. division ratios starting from a minimum value to a maximum values with arbitrary values in between) frequency divider can be set up with a slightly more complex prescaler providing two division ratios and a set of programmable counters.

4.2. The Pulse Swallower Divider

The pulse swallower principle, which has its roots in the beginning of the 1970ties [1], combines a prescaler and two programmable counters in a clever way, such that the maximum input frequency of the programmable counters is reduced by the prescaler modulus, but the fine resolution of the prescaler-free implementation is maintained. This is achieved by commanding the prescaler to an alternate, by one incremented modulus. Each time when the prescaler is operated in the alternate modulus, an input clock edge is said to be swallowed because the result is the same as if the prescaler would have operated with the normal modulus and one input clock edge was eliminated.

Fig. 4.2 shows a principal diagram of the pulse swallower divider. A dual modulus prescaler (DMP) divides the input signal by M or alternatively by $M+1$, depending on the logic level of the modulus control input MC. The by M lowered
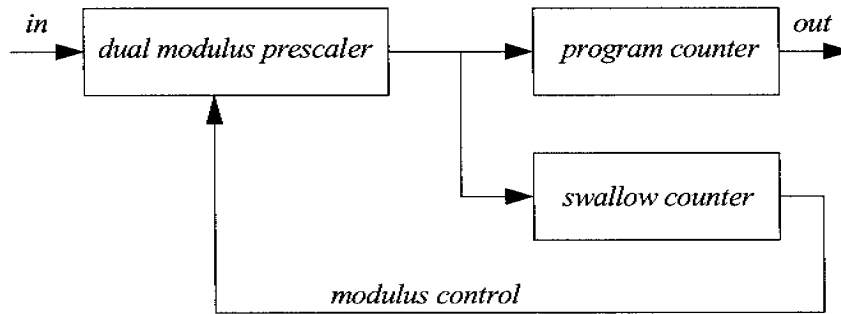
Figure 4.2: Principal diagram of a pulse swallower divider.

frequency is fed to two programmable counters which are named swallow and program counter with preset values S and P respectively. The swallow counter provides a feedback signal which controls the prescaler modulus. The operation principle can be summarized as:

- *The swallow and program counter are decremented by each clock at the DMP output.*

- *Once the swallow counter reaches zero, zero is hold during the consequent DMP clocks.*

- *When the program counter reaches zero as well, it will dispose the preset values to be loaded into both counters during the subsequent DMP clock.*

- *The DMP divides the input by M+1 as long as the swallow counter has not reached zero, otherwise by M.*

Based on these operation rules, the division ratio divides the input frequency to a lower frequency with the ratio given as:

$$\begin{aligned} D &= (M+1) \cdot S + M \cdot (P+1-S) \\ &= M \cdot (P+1) + S \end{aligned} \tag{4.1}$$

Eq. 4.1 indicates that the pulse swallower technique allows programming of division ratios in unity steps. However, a few conditions must be fulfilled to guarantee 'seamless' ratios starting from a minimum D_{min} to a maximum D_{max}. The required conditions can be are formulated as:

$$\begin{aligned} P &\geq S \\ S &= \{0, 1, \ldots, M-1\} \\ P &= \{M-1, M, \ldots, P_{max}\} \end{aligned} \tag{4.2}$$

While the first condition is necessary to allow prescaler operation in both moduli, the second condition defines the minimum required capacity of the swallow counter to fill all gaps if the program counter value is incremented or decremented. The last condition defining the boundaries of the program counter is a direct consequence of the first and second condition. Eq. 4.1 and Eq. 4.2 bound the minimum and maximum programmable division ratios:

$$D_{min} = M^2$$
$$D_{max} = M \cdot (P_{max} + 2) - 1$$

(4.3)

The last Equation states that the prescaler modulus is determined by the square root of the minimum division ratio. Table 4.1 lists pulse swallower parameters of frequency synthesizers for GSM/DCS/PCS direct conversion transceivers. It can be seen that input frequency of the swallow and program counter is lowered by the DMP to around 30MHz. This allows implementation by CMOS standard cells.

	EGSM	DCS	PCS
f_{min} (D_{min})	880 (4400)	1715 (8575)	1850 (9250)
f_{max} (D_{max})	960 (4800)	1880 (9400)	1990 (9950)
M/M+1	32/33	64/65	64/65
S	0 ... 31	0 ... 63	0 ... 63
P	31 ... 151	63 ... 149	63 ... 155

Table 4.1: Pulse swallower parameters for GSM/DCS/PCS.

Other divider architectures have been proposed to further reduce the frequency of the programmable dividers. Two series connected DMPs result in a lowered programmable counter input frequency [2]. Other authors propose to replace the dual modulus prescaler by a multi modulus prescaler. Again, the input frequency of the consequent programmable counters can be lowered. Finally, an approach which removes the need for programmable counters completely has been proposed in [4]. It was demonstrated that an arbitrarily programmable frequency divider can be implemented by a chain of series connected 2/3 DMPs. However,

a single DMP, combined with two programmable dividers, represents in most cases an adequate programmable divider structure.

4.3. Dual Modulus Prescalers on Gate Level

The DMP is the most critical part of the frequency divider and deserves great attention, both on gate level and on transistor level. Virtually all DMPs are based either on synchronous counters or, less often applied, on phase switching technique. Both topologies are studied here and speed requirements of the used logical devices are derived. On cost of power consumption, the phase switching technique permits fractional division ratios which allow a wider PLL loop bandwidth. Section 4.3.3 compares the different topologies and investigates wether fractional division ratios obtained from a phase switching prescaler justify the larger power consumption.

4.3.1. Synchronous Counter based Dual Modulus Prescaler

A principal schematic of a synchronous counter based 64/65 DMP is shown in Fig. 4.3. The prescaler splits up into a synchronous 4/5 divider (consisting of flip-flops FF1-FF3) and an asynchronous divider chain (consisting of toggle flip-flops FF4-FF7). The synchronous counter divides most of the time by four, but increases the division ratio to five if commanded from the asynchronous divider chain. This happens every 16th clock seen by the asynchronous divider chain, if the modulus control input is set to 65 division ratio. Consequently, a division ratio of $15 \times 4 + 1 \times 5 = 65$ is realized. The same division ratio could be achieved alternatively with a 2/3 synchronous divider triggering an asynchronous divider chain with one additional toggle flip-flop. While a full-speed flip-flop could be saved, more power is required in the asynchronous part due to the doubled clock frequency of the divider chain. The proposed circuit in Fig. 4.3 represents a well balanced compromise between asynchronous divider chain requirements and synchronous counter consumption.

The key to power consumption minimization of the prescaler lies in the identification of flip-flop speed requirements. This provides the basis for circuit optimization on transistor level. The synchronous nature of the 4/5 counter determines the requirements for FF1-FF3. The sum of the propagation delay of a flip-flop

synchronous 4/5 divider

asynchronous divide by 16
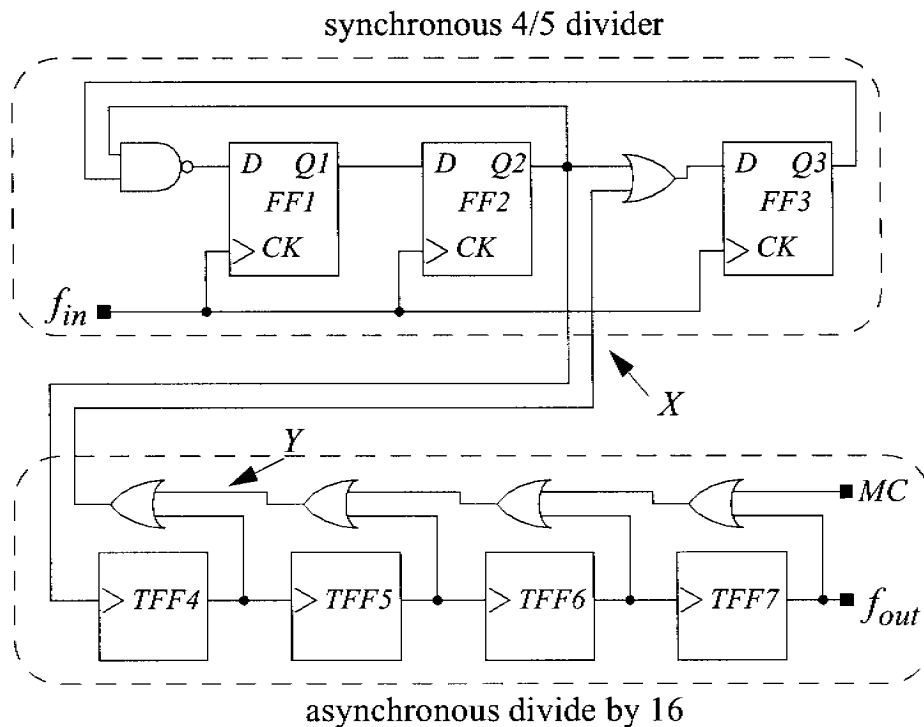
Figure 4.3: Principal schematic of a 64/65 DMP using a synchronous 4/5 divider.

and an AND or OR gate must be not longer than the input clock period $1/f_{in}$, or saying in other words, the flip-flops of the synchronous counter must be able to operate in toggle configuration at slightly more than full speed.

The speed requirement of the remaining toggle flip-flops and gates is less obvious. Critical to the correct operation of the prescaler is the timing of the feedback signal which determines the division ratio of the synchronous counter. The latter signal, which is marked in Fig. 4.3 as node X, is delayed by FF2, FF4 and two OR gates. Fortunately, node X must not settle within a single input clock period. This can be understood by looking at the 4/5 counter as a synchronous finite state machine (FSM) with a single control input X. Fig. 4.4 shows the FSM transition diagram. It can be seen that, as long as X is logical one, the FSM loops in states '111', '011', '001' and '101', with the three digits representing the Q outputs of FF1-FF3. This loop forms divide-by-four mode. Assuming that all flip-flops trigger on the positive clock edge, node X changes to zero during the transition from '101' to '111'. The next two following transitions are unconditional, i.e. the logical level at node X has no influence on the transitions. It is finally the transition

Figure 4.4: State transition diagram of a synchronous 4/5 divider

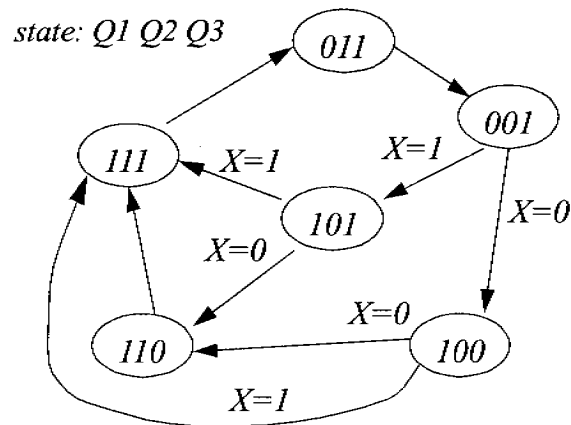leaving '001' which evaluates the feedback signal. Hence, the sum of the delays of FF2, FF4 and two OR gates must be smaller than three input clock periods. So, FF4 can be designed with halved speed compared to FF1, FF2 or FF3. The logical zero at node X leads to a '111', '011', '001', '100', '110' transition sequence which is the divide-by-five mode. However, an alternate sequence '111', '011', '001', '101', '110' can be identified by inspection of the state transition diagram shown in Fig. 4.4. This alternate sequence will further relax the settling requirement of node X from three input clock periods to four, reducing the speed requirement of FF4 to one third of a full speed flip-flop. The delay of FF4 and its OR gate determines which of the two divide-by-five sequences is used.

An important question, which has not been addressed yet, is: Is the Q output of FF2 the optimum signal to trigger the asynchronous divider chain? Indeed, the asynchronous divider chain could be triggered by any flip-flop output of the synchronous divider including also their inverted outputs $\overline{Q}$, except FF3 which never changes its logical value change during divide-by-four mode. However, the available time to settle the feedback signal is reduced by one input clock period if FF4 is triggered by $\overline{Q1}$ and by two clock periods if triggered by $\overline{Q2}$ respectively. The Q output of flip-flop FF2 represents therefore the optimum triggering signal for the asynchronous divider chain.

To complete the speed requirement analysis of the prescaler, toggle flip-flops FF5-FF8 must be considered as well. A logical zero must be generated at node Y if the asynchronous divider reaches zero state, i.e. if FF5 to FF7 provide logical zero at their outputs. Since the divider chain is reverse counting from '1111' to

'0000', the output of FF7 holds a zero 32 input clock cycles, FF6 16 input clock cycles and FF5 8 clock cycles before zero state is reached. Hence, FF7 demands the loosest speed requirement of one sixteenth of a full speed flip-flop. FF6 and FF5 demand doubled and quadrupled speed compared to FF7. Table 4.2 summarizes the speed requirements of the synchronous counter based, 64/65 dual modulus prescaler.

FF1-FF3	FF4	FF5	FF6	FF7
$1 \times f_{in}$	$1/3 \times f_{in}$	$1/4 \times f_{in}$	$1/8 \times f_{in}$	$1/16 \times f_{in}$

Table 4.2: Flip-flop speed requirements of the synchronous counter based DMP.

4.3.2. Fractional Dual Modulus Prescaler

Synchronous divider based prescalers represent the legitimate workhorse in prescaler design due to their robustness and proven topology. Nevertheless, alternate techniques should be reviewed from time to time.

Such an alternative is found in the phase switching technique. The idea is illustrated in Fig. 4.5. Signal A is initially wired to the output, while a by one quarter period delayed replica B, i.e. the quadrature signal of A, is waiting to be switched to the output. If the switching is performed at the right time, the output signal is delayed by one fourth of the input period T. Fig. 4.5 shows that the instantaneous period during phase switching is increased to $1.25 \times T$. This concept was proposed to be used in conjunction with two toggle flip-flops dividing the input clock by four prior to phase switching [5]. The second flip-flop serves not only as divide-by-two, but also as quadrature generator. The resulting instantaneous output period during phase switching is then raised to $1.25 \times 4/f_{in} = 5/f_{in}$, giving the configuration the capability to replace a synchronous 4/5 divider. It is claimed that less power is consumed compared to the synchronous counter solution, because only a single full speed flip-flop is required. However, besides the two flip-flops and the phase-switching network, additional logic is necessary to control the switching transient. This control logic may contribute significantly to the overall power consumption, increasing the power consumption to the level of the synchronous 4/5 divider solution.
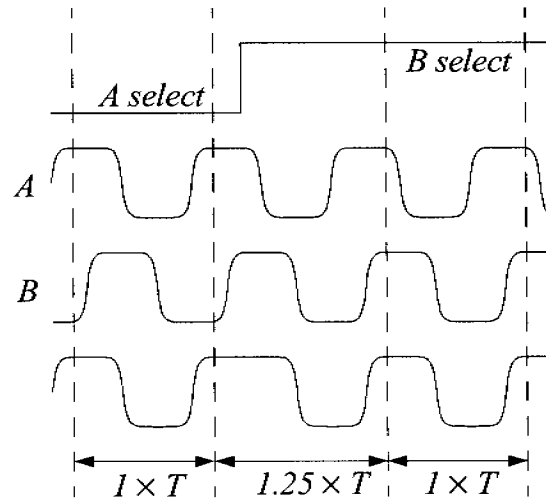
Figure 4.5: Principle of phase switching.

The timing of the phase select signal is very critical. A more than by one quarter period delayed phase select signal produces a glitch at the output, causing defective triggering of the following divider stages.

Phase switching technique is an interesting concept despite the fact, that no or only little power can be saved. A fractional $M/M+0.25$ prescaler can be implemented by skipping the proposed divider chain prior to the phase switching circuit and feeding the phase switching circuit directly with the quadrature LO signals. It can be shown that the fractional dual modulus prescaler combined with a program and a swallow counter can generate fractional division ratios D of the form

$$D = M \cdot (P + 1) + \frac{S}{4} \qquad (4.4)$$

D_{min} and D_{max} can be found as

$$D_{min} = 4 \cdot M^2$$
$$D_{max} = M \cdot (P_{max} + 2) - \frac{1}{4} \qquad (4.5)$$

and the counter capacities, required for 'seamless' programming to any arbitrary division ratio within the boundaries, are similarly found as:

Crossbar:
S=1: Aout=Ain, Bout=Bin
S=0: Aout=Bin, Bout=Ain

asynchronous divider

Buffer/Inverter:
S=1: O=I
S=0: O=!I

switching control network
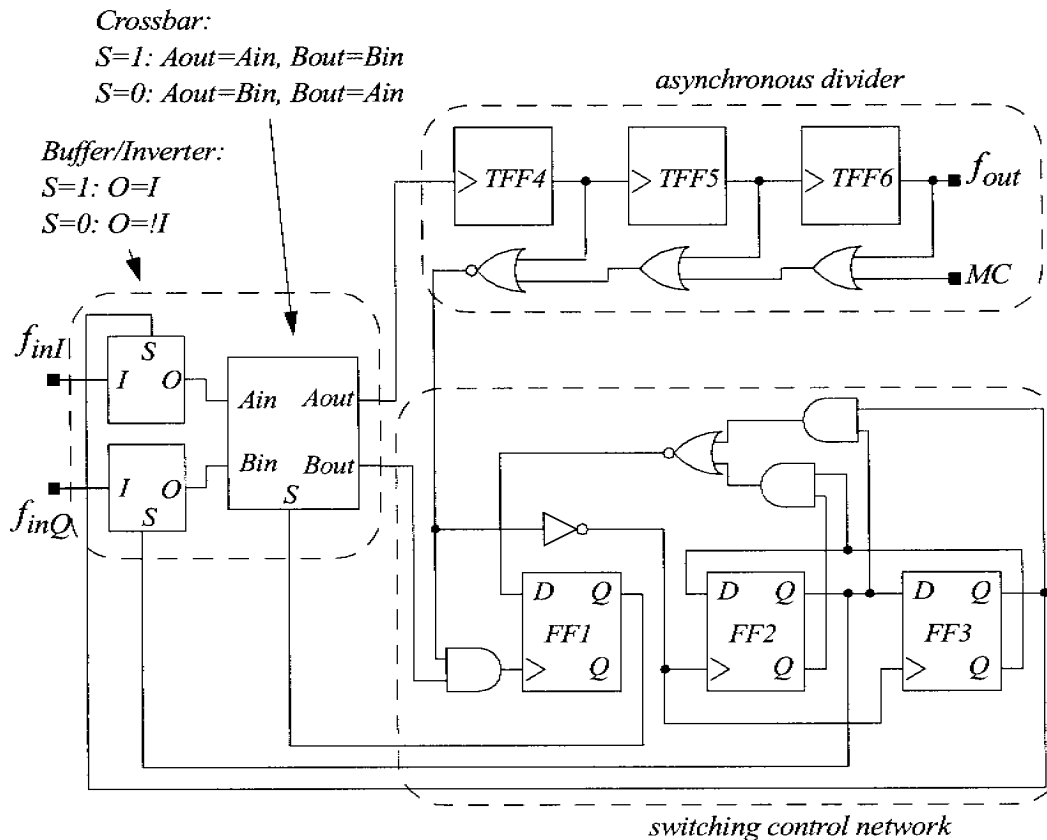
Figure 4.6: Architecture of a fractional, 8/8.25 DMP.

$$S = \{0, 1, ..., 4 \cdot M - 1\}$$
$$P = \{4 \cdot M - 1, 4 \cdot M, ..., P_{max}\}$$
(4.6)

The remaining part of this Section proposes a fractional DMP architecture and determines the flip-flop speed requirements. The prescaler splits up into three parts: A phase switching network, an asynchronous divider chain, and a switching control network. Fig. 4.6 shows a schematic of the proposed circuit. The phase switching network, which is a pure combinational logic, is able to provide any $0°$, $90°$, $180°$, $270°$ phase shifted version of the input signal f_{in} at its outputs A_{out} and B_{out}. To perform this task, it requires two input signals f_{inI} and f_{inQ} being in quadrature and the combinational building blocks named inverter/buffer and crossbar network. The former building block works as a buffer if the control input S is set to logical one and as inverter otherwise. The crossbar network wires inputs and outputs either directly or cross wisely, depending on its control input

S. Changing the logical state of the control input causes a 90° phase shift at the outputs of the phase switching network.

The operation of the 8/8.25 prescaler is basically a direct application of the phase switching principle outlined in Fig. 4.5. An asynchronous divider set up by FF4-FF7 receives its triggering input clock from A_{out}, which is wired through the switching network to one of the input signal phases. Consequently the input frequency is lowered by 8. The instantaneous frequency at A_{out} is reduced by one fourth if the asynchronous divider reaches zero state with activated modulus control input. The clock input of FF1 is then enabled by the help of an AND gate and the output of the latter flip-flop will change with the active clock edge at B_{out}, which is a 90° delayed replica of the signal at A_{out}. A_{out} is delayed now by one fourth input period because the output of FF1 is connected to the crossbar network. The division ratio is consequently $7 \times 1 + 1 \times 1.25 = 8.25$.

After delaying the clock by one quarter input period, the clock phase at B_{out} must be readjusted to be prepared for the next phase switching event. While B_{out} was lagging A_{out} by 90° before phase switching, the situation is vice versa after the phase switching event, since the crossbar has simply swapped A_{out} and B_{out} against each other. To regain the required phase relations, the phase at B_{out} must be switched by 180°. This occurs when the asynchronous divider chain leaves the zero state and FF2 and FF3 are triggered, modifying the control inputs of the two buffer/inverter blocks.

Since the operation of the 8/8.25 prescaler of Fig. 4.6 is clarified now, a closer look on the speed requirement is necessary. It was mentioned before that the delay seen at the crossbar network control input must be smaller than one quarter of an input clock period, revealing an FF1 propagation delay of less than one fourth of the input signal period. Hence, FF1 must be able to operate at quadrupled input frequency. The same speed is imposed to FF4 to ensure that the clock of FF1 is enabled in time. Fortunately, the tolerable delay to the control inputs of the buffers/inverters is much lower, relaxing the FF2/FF3 speed requirement significantly.

Recapitulating these considerations, the design of the 8/8.25 fractional DMP using the phase switching technique is endeavoured by the demand for two flip-flops which must be four times faster than the prescaler input frequency.

4.3.3. Prescaler Comparison

If a frequency synthesizer employs an RF oscillator of twice the LO frequency followed by a digital divide-by-two circuit to generate quadrature phases, a the synchronous counter based DMP may be driven either by one of the quadrature phases or by the RF oscillator. While the former approach consumes less current, the latter allows doubling of the PLL bandwidth. Consequently, the synthesizer switching speed is doubled. Besides the faster switching time, no unbalanced load is imposed to the quadrature phases which may affect the phase accuracy. An even faster synthesizer can be realized with the proposed fractional divider hooked to the quadrature phases. A quadrupled PLL bandwidth compared to the first approach can be realized. However, the consumption is expected high due to the required fast switching flip-flops

To gain more insight into the power consumption level, three prescalers have been roughly dimensioned and verified by the help of a circuit simulator. All prescalers accommodate a divider for DCS/PCS. Notice, that all solutions achieve the same frequency resolution of 200kHz. Transistor data are taken from a standard 0.18µm CMOS technology. Table 4.3 summarizes the divider parameters and lists the expected power consumption. As suspected, the fractional prescaler based divider exhibits the highest consumption. The frequency synthesizer consumption is dominated by the consumption of prescaler and RF oscillator. A solution for a low phase noise 3.6GHz oscillator with 3mA consumption is proposed in Chapter 6, which results in a competitive synthesizer consumption of 6mA if combined with the 3.6GHz divider. This combination represents an optimum solution to DCS/PCS, since synthesizers relying on the other two dividers either collide with the DCS/PCS switching time requirement or consume more power.

	Solution A	Solution B	Solution C
Prescaler Modulus	64/65	64/65	16/16.25
Divider Input Frequency	~ 1.8GHz	~ 3.6GHz	~ 1.8GHz
Divider Output Frequency	200kHz	400kHz	800kHz
Prescaler Consumption	< 1mA	~ 3mA	> 6mA

Table 4.3: Estimated 0.18µm CMOS DMP current consumption.

4.4. Transistor Level Implementation

Current mode logic (CML) represents the preferred CMOS logic style to implement high frequency prescalers due to various advantages compared to other logic styles. A strong argument for CML lies in its low logic swing which is only a fraction of the full supply voltage. The low swing allows the prescaler to be driven directly from the VCO without the need for clock amplification, which would cost a considerable amount of power. Another argument for CML lies in the constant current drawn from the power supply alleviating spikes on the supply lines which are often found in digital voltage mode circuits. The low swing combined with the fully differential operation reduces further noise injection into the IC substrate. Current mode logic eases therefore the integration of the prescaler with other sensitive analog circuits on the same die.

4.4.1. Current Mode Logic Flip-Flops

Fig. 4.7 shows the schematic of a CML flip-flop which is mainly an ECL-type flip-flop with the bipolar transistors replaced by NMOS transistors. The flip-flop is divided into a master and a slave latch, both consisting of a tail current source, a current switching network implemented by NMOS transistors and resistive loads. Depending on the logical level at the clock input, the latches are either in sample or hold mode. The signal at the data input D of the flip-flop is evaluated during sample mode and a voltage drop is established across one of the load resistors. The voltage, which is further amplified by a cross coupled transistor pair during hold mode, is passed to the input of the slave latch. With sample and hold mode of the slave latch interchanged compared to the master latch, the data input of the flip-flop is never transparent to the flip-flop output Q. Hence, the flip-flop of Fig. 4.7 acts as a pure edge triggered device, a mandatory prerequisite for application as divider flip-flop.

While the operation of CML flip-flops is unquestionably understood, designing a CML flip-flop for minimum power consumption is a delicate task. Due to the lack of a rigorous design strategy, designers rely on simulators and size the tail current source, the load resistors and switches on a trial and error basis. Although common design practice, this approach is not satisfactory at all, since neither minimum consumption is guaranteed, nor guidelines are provided to the novice designer, who cannot fall back on figures of previous designs as a starting point
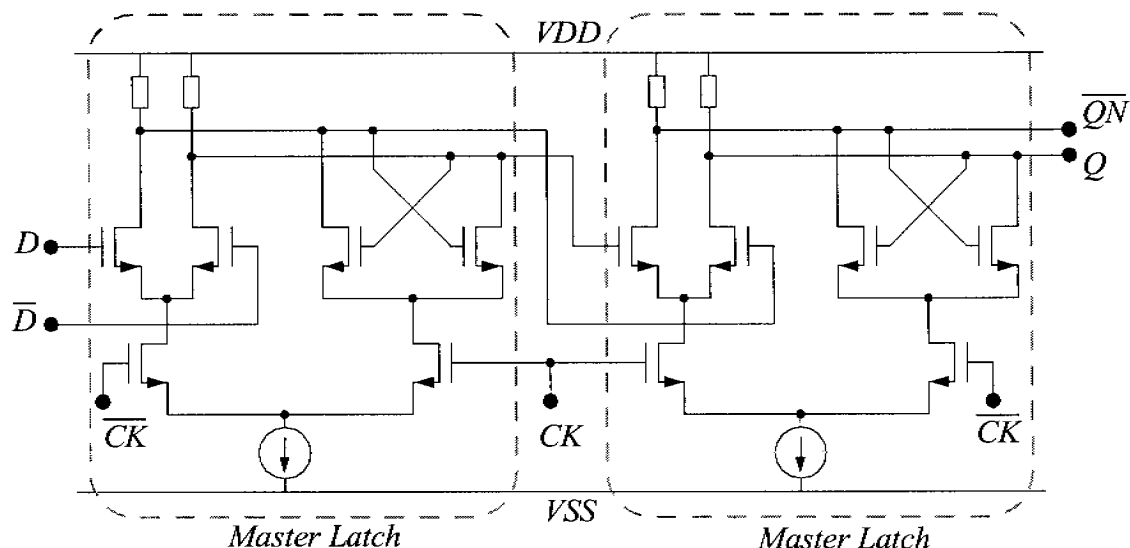
Figure 4.7: A CML master-slave flip-flop.

for simulations. Although a more analytic approach would be highly desirable, the reason for the lack of any analytic design methodology is widespread. First, the analysis of any digital circuits by analog techniques is complicated by the inherent nonlinearity of digital circuits which use the MOS transistors in any operation region between off state to strong inversion. Second, CML design is a more delicate task than the design of their bipolar counterpart due to the limited transconductance compared to bipolar. While ECL circuits mostly use minimum size transistors [6], the channel width of the MOS switches of CML circuits needs to be designed significantly above minimum width, revealing one more parameter to be optimized. Despite these difficulties, Section 4.4.2 provides design guidelines for CML circuits founded on a current consumption minimization analysis outlined in Appendix 4.A.

4.4.2. Power Optimization of CML Circuits

Optimization of current mode logic flip-flops represents basically the pursuit for an optimum combination of logic swing, load resistor, switch size and bias current, so that the differential outputs settle within a given period. Moreover, the signals involved in CML circuits must be self-sustaining, i.e. the CML circuits must provide a gain larger than unity. Consisting essentially of CMOS differential pairs with resistive loads, the maximum differential gain of a CML circuit is expressed as
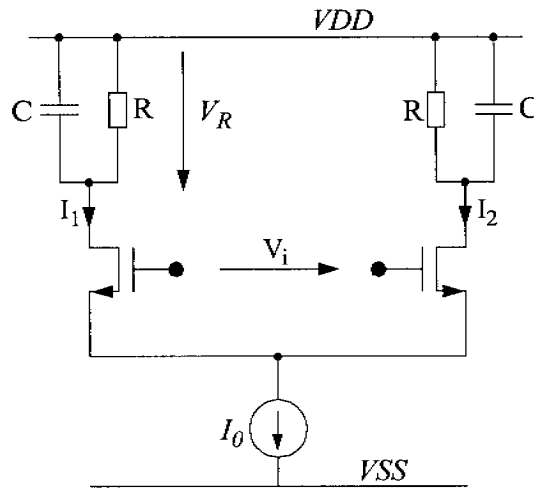
Figure 4.8: A CML inverter used as vehicle for the power optimization analysis.

$$A_V = g_m \cdot R \tag{4.7}$$

with g_m the switching transistor transconductance, while biased with one half of the tail current source, and R the load resistor.

The first parameter to be studied is the size of the switching transistors. It is crucial to power optimization that the switching transistors act as switches conducting most of the time the full bias current rather than operating as linear amplifier. To allow a single transistor of a differential pair to conduct all bias current, the transistors must be made large enough, so that the transistors overdrive voltage is smaller than the amplitude of the signal driving the differential pair input, i.e.

$$\frac{V_{ov}}{V_m} \leq 1 \tag{4.8}$$

with V_{ov} the switching transistor's overdrive voltage when conducting the full bias current and V_m the logic swing at the input. The ratio expressed in Eq. 4.8 is best suited to characterize the quality of the switches. A small V_{ov}/V_m ratio is desirable to speed up the output settling yielding in a larger output swing, but larger switches are required to keep the latter ratio small, increasing thereby the capacitive loading at the output. It can be shown, that the peak-to-peak output voltage is increased only by 15% with the V_{ov}/V_m ratio reduced from 1 to 1/3. The increase must be paid by nine times larger switching transistors, raising the time constant of the output nine times. Keeping the V_{ov}/V_m ratio very small by

using very big switches represents hence an unfavorable strategy. The V_{ov}/V_m ratio should be therefore designed around unity.

The second parameter to be studied is the voltage swing. One might expect that the smaller the voltage swing is, the less power is consumed, since the output swing scales with the product of bias current and load resistance. Indeed, this would be true if no next CML stage would have to be driven by the differential voltage established at the output nodes. An extremely small output swing would require to reduce the overdrive voltage of the succeeding input stage to keep its V_{ov}/V_m ratio equal to unity. Again, this can be done only by increasing the size of the switching transistors. The output providing only a very small swing would have to drive therefore a very large capacitance caused by the large transistors of the succeeding stage. Minimizing the logic swing is thus definitely not the way to go. On the other hand, a very large swing needs to be biased with more current, increasing thereby the power consumption. The existence of an optimum logic swing must be therefore presumed.

A CML inverter with the schematic shown in Fig. 4.8 has been analyzed to find the optimum logic swing. The inverter has been chosen to carry out the analysis due to its simplicity. A couple of assumption are further presumed to keep the mathematics compact. First, the MOS transistor drain current characteristic is modelled as

$$I_d = \begin{cases} 0 & V_{gs} < V_{th} \\ \dfrac{\beta}{2} \cdot V_{ov}^2 & V_{gs} \geq V_{th} \end{cases} \tag{4.9}$$

$$with \ \beta = \frac{W}{L} \cdot K' \quad and \quad V_{ov} = V_{gs} - V_{th}$$

with V_{gs}, V_{th}, V_{ov} and β the gate to source voltage, the threshold voltage, the strong inversion overdrive voltage and the transconductance parameter respectively. Second, all capacitance is assumed to be concentrated into two capacitors C in parallel to the load resistors R. These capacitors are composed of a fixed part (which represents interconnect and resistor capacitances) and a part which is proportional to the width of the switching transistors. The capacitor can be expressed therefore as

$$C = C_0 + k \cdot \beta \tag{4.10}$$

with C_0 being the fixed part and k a technology dependant constant. With these assumptions and some algebraic manipulations outlined in Appendix 4.A, the optimum logic swing can be found as

$$V_{0opt} \approx \frac{10.3}{T} \cdot \frac{1+\varepsilon}{1-\varepsilon} \cdot \ln\left(\frac{1}{\varepsilon}\right) \cdot k \qquad (4.11)$$

with T the period of the signal applied to the inverter input and ε a settling error of the exponential settling at the RC loaded output. It is interesting to note, that the optimum logic swing scales linearly with the operating frequency. The optimum switching transistor width is found as:

$$\beta_{opt} = \frac{C_0}{k} \qquad (4.12)$$

Switching transistor, load resistance and bias current can be properly sized by the help of the design guidelines and the Equations derived in Appendix 4.A. The obtained parameters need to be further refined by the help of SPICE based simulation.

4.5. A 1GHz 64/65 Prescaler in 0.25μm CMOS

A 64/65 prescaler for an input frequency of 1GHz has been designed and laid out in a standard, double metal, single polysilicon 0.25μm CMOS process. An arrangement similar to Fig. 4.3 has been chosen, but negative instead of positive edge triggered flip-flops are employed [7]. Accordingly, the OR gates are replaced by AND gates since the asynchronous divider is no longer reverse counting. The schematic of the flip-flops is identical to Fig. 4.7 with CK and $\overline{CK}$ interchanged.

Flip-flops of different switching speed are employed. Fast type flip-flops biased with 100μA current sources form the 4/5 divider. The asynchronous divider is mainly implemented by slow type flip-flops biased with 25μA current sources. The first flip-flop of the asynchronous divider chain is biased with 50 μA current sources. Table 4.4 summarizes the parameters of all flip-flop types. The gates of the synchronous 4/5 divider are built into FF1 and FF3 to re-use the bias current. While the bias current and the propagation delay of the logic gates can be saved, one more stacked transistor is required in FF1 and FF3 resulting in a reduction of the available per transistor voltage headroom. Nevertheless, the overall current

	fast	medium	slow
I_0	100μA	50μA	25μA
R	3kΩ	6kΩ	12kΩ
W/L [μm]	8/0.25	5/0.25	2.5/0.25

Table 4.4: Parameters of the flip-flop employed in the 1GHz DMP.
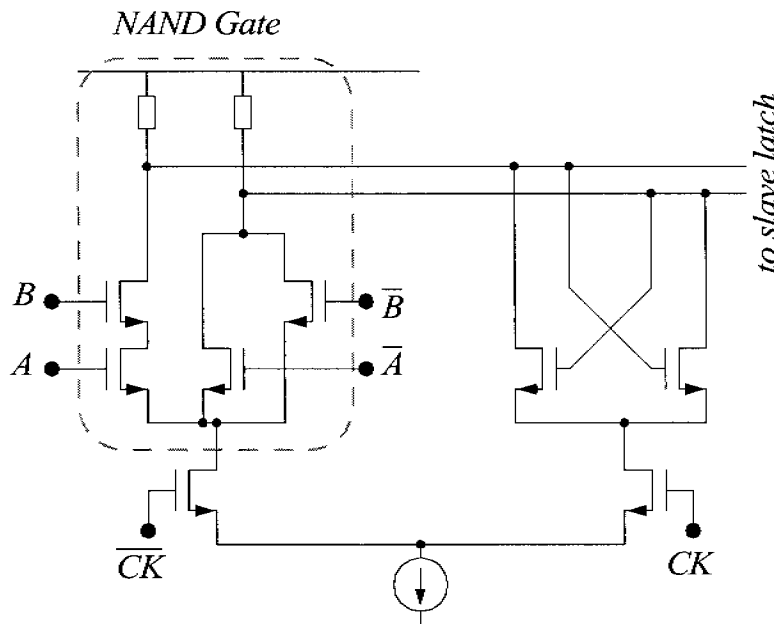
Figure 4.9: Master latch with built in NAND gate.

consumption can be lowered by the help of built-in gates despite the voltage headroom issue. Fig. 4.9 shows a schematic of the master latch with integrated NAND gate. The clock input of the toggle flip-flops is preceded by source follower stages to lower the common mode voltage at the flip-flop clock input. Otherwise, the switching transistors would by pushed into triode region, causing a reduction of switching speed.

Three different kinds of resistances are at disposal to implement the load resistors: Diode connected PMOS transistors, into triode region biased PMOS transistors and linear resistors. The first solution must be ruled out due to the loss of voltage headroom imposed by the transistor's threshold voltage. Triode resistors are favoured over linear resistors because of their compactness compared to

Figure 4.10: Bias circuit of the triode resistors.

polysilicon resistors and due to the lack of well controlled high resistivity poly-silicon resistors. Series and parallel combinations of identically sized transistors constitute the required 3kΩ, 6kΩ and 12kΩ resistors, as can be seen in Fig. 4.10. The operation region of the PMOS transistors must be strictly limited to the tri-ode region. Even a slight shift into the saturation r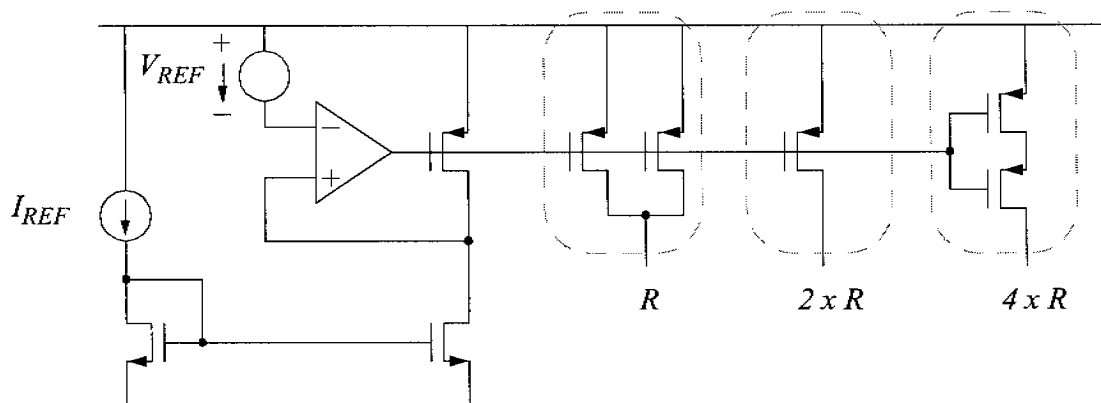egion would change the imped-ance dramatically with fatal impact on the dynamic behavior of the flip-flop. The PMOS transistor overdrive voltage is therefore designed significantly larger than the logic swing. Diode connected transistors in parallel to the triodes acting as voltage clamps are therefore no longer necessary and are skipped to remove the additional loading [8]. The gate voltage of the triode biased transistors is gener-ated by a replica circuit shown in Fig. 4.10. The triode resistance is determined by the reference voltage and current rather than the technology and temperature dependant transconductance parameter, which would determine the triode resis-tance if biased by a simple current mirror. A cute side effect of the proposed bias-ing circuit lies in the fact that the logic swing and the load resistance can be controlled independently. The regulating amplifier of the replica circuit is imple-mented as a single stage operational transconductance amplifier with a DC gain of 45dB. The circuit measures only 100μm×400μm due to the MOS transistor only design.

4.5.1. Experimental Results

The positive input of the prescaler is driven by a signal generator while a DC voltage is applied to the inverted prescaler input. Simulations have shown that the unbalanced input causes only a small degradation of the maximum input fre-

Input Frequency	1GHz	2GHz	3GHz
Division Ratio	65		
Supply Voltage	1.7V	2.5V	3.5V
Current Consumption	0.5mA	1.25mA	2.1mA
Logic Swing	230mV	300mV	340mV

Table 4.5: Best measured 0.25µm DMP performance with 1GHz, 2GHz and 3GHz input.

quency. Extensive measurements with different combinations of bias current and voltage swing settings, which are both adjustable by external sources, showed that the prescaler divides a 1GHz input by 65 while consuming only 500µA at 1.7V supply voltage, excluding the power required to drive the measurement instrument. Although designed for a 1GHz input, the prescaler divides a 2GHz input signal with the power supply increased to 2.5V at a current consumption of 1.25mA. Even a 3GHz input can be handled, however, the supply must be increased beyond save limits. Table 4.5 summarizes best measured prescaler performance for 1GHz, 2GHz and 3GHz input. As expected from the theoretical analysis, the optimum logical swing increases with higher frequencies. However, the optimum swing found in the experiments scales less than linearly, as it could be expected from Eq. 4.11. The reason lies in the supply voltage limitations, which force the swing below the optimum value.

The above numbers represent the absolute operating limits of the prescaler. Nominal operating condition for a 1GHz input signal is at 2.5V supply voltage, 900µA current consumption and 300mV logic swing. The prescaler proves to be very robust under this condition, e.g. input frequency ranges from 0.15GHz to 1.5GHz while the input power ranges from -13dBm to 12dBm. Reference current and voltage may vary by ±35% and ±20% respectively.

The input sensitivity shown in Fig. 4.11 was measured under nominal operating condition. It is measured as -13dBm in the 800MHz to 1400MHz band, which is low enough to be driven by a VCO. An interesting property of the prescaler is self-oscillation which is found at 1300MHz. The self-oscillation is caused by the pure combinational behavior of the flip-flops if no differential signal is applied to the clock inputs. The feedback of FF2 to the NAND gate of FF1 forms then a ring
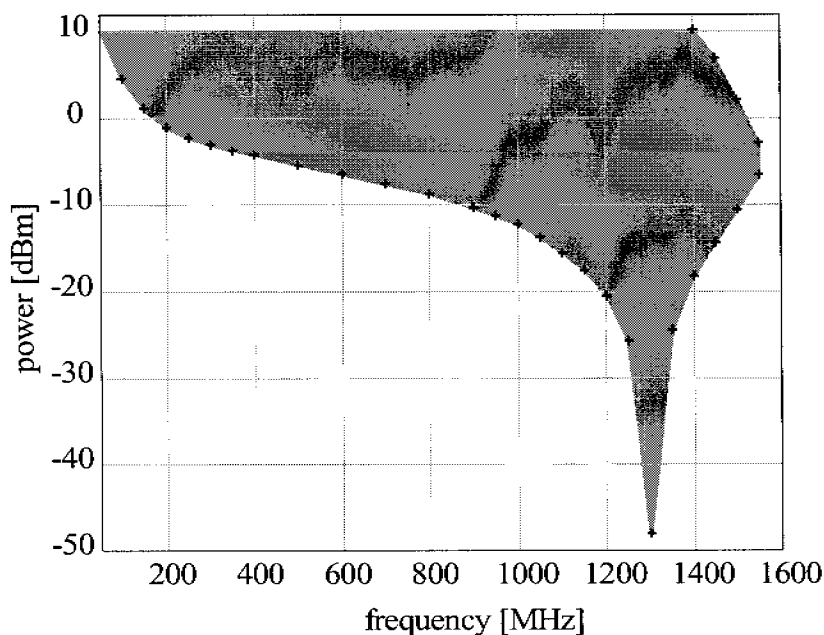
Figure 4.11: Prescaler input sensitivity and region of operation.

oscillator. The minimum input amplitude raises towards very low input frequencies to maintain a minimum input signal slope. At very high input frequencies, the input amplitude must be raised as well to speed up the switching transistors. However, a too large input amplitude stops the prescaler since the switching transistors are pushed into the triode region. This leads to a region of operation outlined by the grey area in Fig. 4.11.

It is time to compare the proposed solution with other recently published prescaler designs. It is remarkable that the 900μA consumption can compete with advanced bipolar prescalers which consume 1 to 2.5mA [9]. Fig. 4.12 shows a comparison with other CMOS prescaler designs. The proposed prescaler outperforms all other CMOS prescalers. This is the result of a careful design based on an analytical design methodology.

4.6. A 4GHz, 0.18μm CMOS Frequency Divider

A frequency divider for a DCS/PCS synthesizer with the divider connected to a VCO of doubled LO frequency has been implemented in a standard, 6 metal, single polysilicon CMOS technology. The divider will be used in the frequency synthesizer described later in Chapter 7. Since the 0.18μm 4GHz prescaler
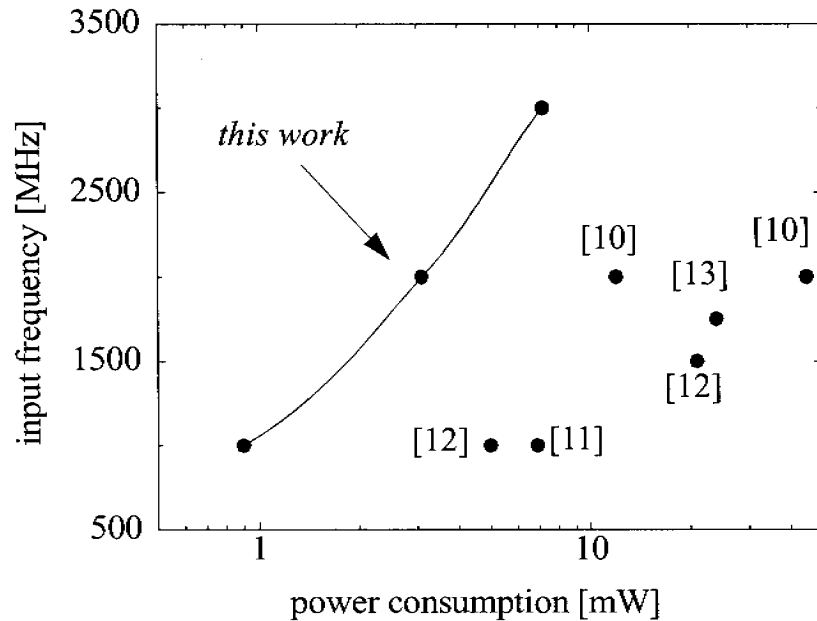
Figure 4.12: Power consumption comparison of CMOS prescalers.

resembles strongly the 0.25μm 1GHz prescaler, the description is kept less detailed. Section 4.6.1 confines the discussion of the prescaler on the most important issues. Section 4.6.2 outlines the design of the programmable counters which accompany the prescaler to form a programmable frequency divider.

4.6.1. A 4GHz, 2.5mA Dual Modulus Prescaler

The topology of the 64/65 prescaler and a schematic of the flip-flops are shown in Fig. 4.3 and Fig. 4.7 respectively. Gates are built into FF1 and FF3 and level shifters precede the toggle flip-flops for the same reason described in Section 4.5. Again, three differently sized flip-flops of fast, medium and slow speed are employed. Table 4.6 summarizes the design parameters of the three flip-flop types. The logic swing is the same 400mV to all flip-flops. The availability of well controlled, high resistivity resistors combined with the smaller load resistance value proffers the use of linear resistors instead of triodes. The load resistors are therefore realized by combinations of unity resistors of 0.9μm width and 12μm length. A maximum deviation of the resistor values is expected to be less than ±15% of the nominal value.

The prescaler will be driven by an off-chip VCO. Matching to 50Ω is crucial to establish a well controlled voltage swing at the prescaler clock input, as well as to

	fast	medium	slow
I_0	240μA	80μA	40μA
R	1.5kΩ	4.5kΩ	9kΩ
W/L [μm]	6/0.18	4/0.18	4/0.18

Table 4.6:   Typical bias current, load resistance and switch size of the three flip-flop types used in the 4GHz 64/65 prescaler.

50Ω transmission line   35fF pad capacitance   2pF MIM capacito
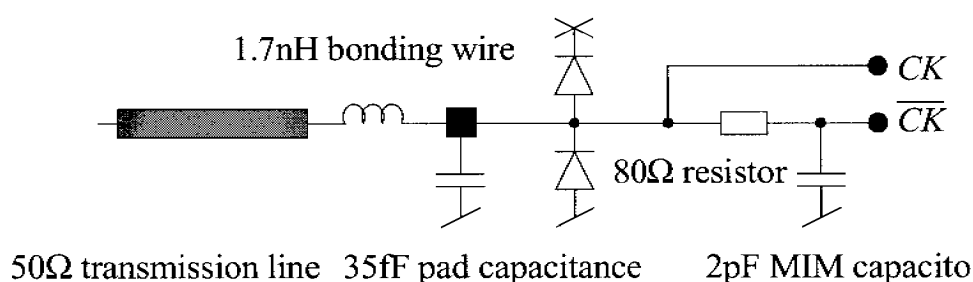
Figure 4.13:   Input matching network of the 4GHz 64/65 prescaler.

provide a well controlled load to the VCO. The input matching network can be seen in Fig. 4.13. The on-chip capacitances caused by the pad, the ESD clamps and the capacitance providing a DC-ground to the inverted clock input are cancelled by the bonding wire inductance. A 80Ω polysilicon resistor is transformed to 50Ω seen at the interface between the bonding wire and the PCB stripline, resulting in a wideband 50Ω matching. The chosen 2pF capacitor at the inverted clock input represents a compromise between remaining signal at the inverted clock input and bonding wire length. Conversion from the differential CML level to single ended rail-to-rail level is performed at the output of the last flip-flop of the asynchronous divider chain. A folded cascode amplifier has been chosen to accomplish the conversion because of the near rail-to-rail output swing of the OTA, which is required to drive CMOS logic [14]. The overall consumption of the prescaler is 2.5mA at a nominal supply of 1.8V. Experiments showed, that the prescaler operates with a 4GHz input signal.

4.6.2. Programmable Counters

The 4GHz divider is completed by an 8-bit program and a 6-bit swallow counter. These reverse counters are fully implemented by CMOS standard cells. Synchronous reverse counters can be set up by a chain of identical counter slices containing a single counter bit each. A counter slice must toggle its bit each active clock edge if all preceding bits are zero or must hold the bit value otherwise. A counter slice providing this function can be realized by a flip-flop, an OR gate and two multiplexers with two inputs, as shown in Fig. 4.14. The second multiplexer is required to load the preset value. Additional OR gates keep the swallow counter in the zero state until the program counter initiates loading of the preset value stored in two registers. Fig. 4.14 shows the schematic of the programmable counters. The registers containing the preset values can be programmed through a three wire, serial, synchronous interface. The serial interface is synchronized to the divider clock to avoid metastability in the counter flip-flops.

Interference of the digital divider with other synthesizer building block may degrade the synthesizer performance. The standard cells operate at a lowered supply voltage of 1.2V to reduce noise injected into the chip substrate and spikes on the supply voltage.
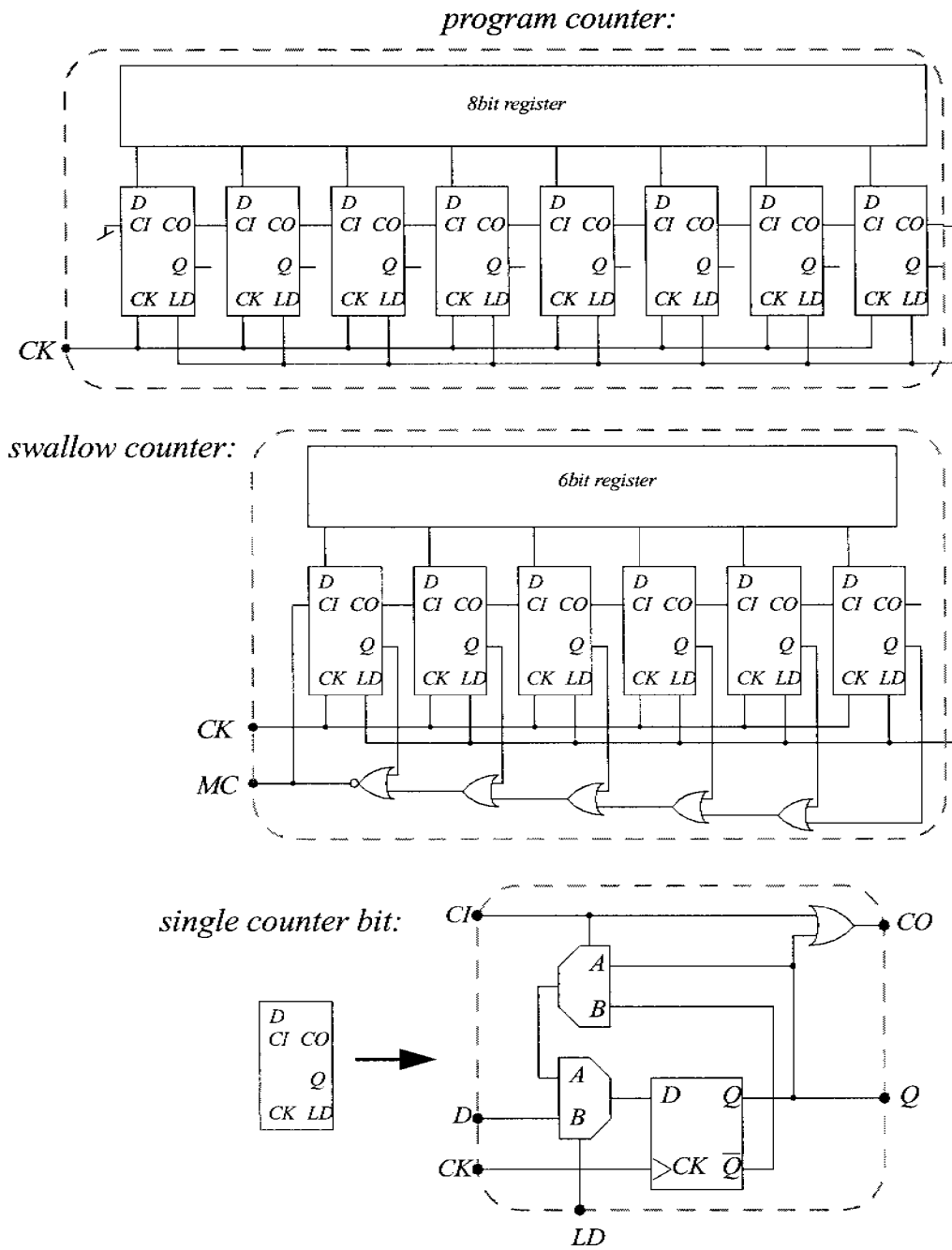
program counter:

swallow counter:

single counter bit:

Figure 4.14: Schematic of the program counter, swallow counter and a single counter bit.

References

[1] J. Nicholds, C. Shinn, *"Pulse Swallowing"*, EDN Magazine, pp. 39-42, October 1970

[2] Wiliam F. Egan, *"Frequency Synthesis by Phase Lock"*, 2000, John Wiley, pp. 157-160

[3] Roland Best, *"Phase-Locked Loops"*, 3rd edition, 1997, McGraw-Hill

[4] C. Voucher, D. Kasperkovitz, *"A Wide-Band Tuning System for Fully Integrated Satellite Receivers"*, IEEE J. of Solid-State Circuits, Vol. 26, no. 10, October 1991

[5] J. Craninckx, M. Steyaert, *"Wireless CMOS Frequency Synthesizer Design"*, Kluwer Academic Publishers, 1998, pp. 164-168

[6] A. Deiss, D. Pfaff, Q. Huang, *"A 200-MHz sub-mA RF front end for wireless hearing aid applications"*, IEEE. J of Solid-State Circuits, Vol. 35, no. 7, July 2000, pp. 977-986

[7] D. Pfaff, Q. Huang, *"A Quarter-Micron CMOS, 1GHz VCO/Prescaler-Set for Very Low Power Applications"*, Proc. IEEE Custom Integrated Circuits Conference, pp. 649-652, San Diego, CA, USA, May 16-19, 1999

[8] J. Parker, D. Ray, *"A Low-Noise 1.6GHz CMOS PLL with On-Chip Loop Filter"*, Proc. IEEE Custom Integrated Circuits Conference, pp. 407-410, 1997

[9] Motorola Semiconductor Data Sheet MC12053A, 1997

[10] J. Craninckx, M. Steyaert, H. Miyakawa, *"A Fully Integrated Spiral-LC CMOS VCO Set with Prescaler for GSM and DCS-1800 Systems"*, Proc. of the Custom Integrated Circuits Conference 1997, pp. 403-406

[11] B. Park, P. Allen, *"A 1GHz, Low -Phase-Noise CMOS Frequency Synthesizer with Integrated LC VCO For Wireless Communications"*, Proc. of the Custom Integrated Circuits Conference 1998, pp. 567-570

[12] H. Cong et. al., *"Multigigahertz CMOS Dual-Modulus Prescaler IC"*, IEEE J. of Solid State Circuits*", Vol. 23, 1988, Oct. 1988, pp. 1189-1193

[13] J. Craninckx, M. Steyaert, *"A 1.75-GHz/3V "Dual-Modulus Divide-by-128/129 Prescaler in 0.7μm CMOS"*, IEEE J. of Solid State Circuits, Vol. 31, July 1996, pp. 890-897

[14] F. Piazza, Q. Huang, *"A Low Power CMOS Dual Modulus Prescaler for Frequency Synthesizers"*, IEICE Trans. on Electronics, Vol. 80, No. 2, Feb. 1997, pp. 314-319

Appendix 4

4.A Power optimization of a CML inverter

Sizing of a CML inverter according to Fig. 4.8 for minimum power consumption is aggravated by the nonlinear current flowing into the RC-load. The nonlinear current I_1 flowing into the latter relates to the differential input voltage V_i as:

$$I_1 = \begin{cases} 0 & \text{if } V_i < -V_{ov} \\[2mm] \dfrac{I_0}{2} + V_i \cdot \dfrac{\beta}{2} \cdot \sqrt{\dfrac{1}{2} \cdot V_{ov}^2 - \left(\dfrac{V_i}{2}\right)^2} & \\[4mm] I_0 & \text{if } V_i > V_{ov} \end{cases} \qquad (4.A.1)$$

Assuming a sinusoidal input with period T and amplitude V_m, I_1 can be approximated in the interval $0...T$ by a piecewise linear function:

$$I_1 = \begin{cases} I_0 \cdot \dfrac{t}{T_R} & 0 < t < T_R \\[3mm] I_0 & T_R < t < T/2 \\[3mm] I_0 \cdot \left[1 - \dfrac{t - T/2}{T_R}\right] & T/2 < t < T/2 + T_R \\[3mm] 0 & T/2 + T_T < t < T \end{cases} \qquad (4.A.2)$$

The raise and fall time T_R depends on the switching transistor transconductance and the input amplitude. It can be found as:

$$T_R = \frac{V_{ov}}{V_m} \cdot \frac{1}{\sqrt{2} \cdot \pi} \cdot T \qquad (4.A.3)$$

The approximation of the RC-load current holds as long as the differential pair acts as current distributing network, i.e. if V_{ov}/V_m is kept below unity. A closer inspection reveals that the peak deviation of the piecewise linear current approximation from the nonlinear current is less than 10% if the V_{ov}/V_m ratio set to unity. Computation of the voltage drop across the RC-load excited by the piecewise linear current according to Eq. 4.A.2 is easily performed by application of Laplace theory. The voltage across the load V_R is found in time domain as

$$V_R = \begin{cases} I_0 \cdot R \cdot \left[\dfrac{RC}{T_R} \cdot e^{-\frac{t}{RC}} + \dfrac{t}{RC} - \dfrac{RC}{T_R} \right] & \text{if } 0 < t < T_R \\[3em] I_0 \cdot R \cdot \left[1 - e^{\frac{-(t-d)}{RC}} \right] & \text{if } T_R < t < T/2 \end{cases}$$

(4.A.4)

with

$$d = RC \cdot \ln \left[\frac{RC}{T_R} \cdot \left(e^{T_R/RC} - 1 \right) \right]$$

(4.A.5)

This result requires further interpretation. The Equation shows, that the voltage settling of the CML output driven by a ramped, piecewise linear current is identical to the settling of an RC-load driven by a current step, which is delayed by a duration d. Fortunately one can show, that under prerequisite of a small raise time compared to the RC-time constant, the delay d can be approximated by a much simpler expression:

$$d \approx \frac{T_R}{2}$$

(4.A.6)

With this result in mind, the identification of a parameter set yielding in minimum consumption is a straight forward task. The required time constant to settle the output voltage with a relative error ε at time $T/2$ is found as

$$RC = \frac{T - T_R}{2} \cdot \frac{1}{\ln(1/\varepsilon)}$$

(4.A.7)

and the single ended peak-to-peak voltage V_0 across the RC-load as:

$$V_0 = I_0 \cdot R \cdot \frac{1 - \varepsilon}{1 + \varepsilon} \qquad (4.A.8)$$

The capacitance loading the output consists of a fixed part and a part which scales linearly with the width of the switching transistors. Hence, the load capacitance can be modelled as

$$C = C_0 + k \cdot \beta \qquad (4.A.9)$$

with k being a parameter depending on the used technology and the number of input stages which need to be driven by the inverter. Finally, the overdrive voltage of the switching transistors when conducting the full tail current is related to the bias current as:

$$V_{ov} = \sqrt{\frac{I_0}{\beta/2}} \qquad (4.A.10)$$

Eq. 4.A.3 to Eq. 4.A.10 can be combined to express the bias current. It can be found as

$$I_0 = \frac{8}{T^2} \cdot \left(\frac{1 + \varepsilon}{1 - \varepsilon} \cdot \frac{\ln(1/\varepsilon)}{x \cdot (1 - 0.225 \cdot x)} \right)^2 \cdot \left(k^2 \cdot \beta + 2 \cdot C_0 + \frac{C_0^2}{\beta} \right) \qquad (4.A.11)$$

with $x = V_{ov}/V_m$. Single ended peak-to-peak output voltage V_0 and differential amplitude of the driving input signal V_m are assumed to be identical.

Eq. 4.A.11 shows that the tail current source, and accordingly the current consumption, depends on the design parameters x, ε, and β. Minimization of the current consumption is achieved by sizing the design parameters according to the following rules:

- *Size the switching transistors for optimum transconductance, defined by*

$$\beta_{opt} = \frac{C_0}{k} \qquad (4.A.12)$$

- *Keep the overdrive to driving amplitude ratio close to unity:*

$$V_{ov}/V_m = 1 \qquad (4.A.13)$$

- *Keeping the settling error large helps to save power. However, limited supply voltage restricts the settling error to small values, e.g. ε=0.1.*

Eq. 4.A.11 can be evaluated with the minimum power constraints. The resulting minimum current formula can be found as:

$$I_{0min} \approx 53.3 \cdot \frac{C_0 \cdot k}{T^2} \cdot \left[\frac{1 + \varepsilon}{1 - \varepsilon} \cdot \ln\left(\frac{1}{\varepsilon}\right) \right]^2 \qquad (4.A.14)$$

As postulated earlier, the minimum bias current is accompanied by a well defined optimum logic swing:

$$V_{0opt} \approx \frac{10.3}{T} \cdot \frac{1 + \varepsilon}{1 - \varepsilon} \cdot \ln\left(\frac{1}{\varepsilon}\right) \cdot k \qquad (4.A.15)$$

It is interesting to note that the optimum logic swing scales with operating frequency. However, if increased clock speed is not accompanied by a better technology, i.e. small k factor, the optimum swing may exceed the available voltage headroom. Excessive power consumption may result or even the preclusion of a CMOS technology for the intended speed.

Seite Leer /
Blank leaf

Chapter 5

RF Mixers

It seems at a first look that mixers are not relevant to frequency synthesis since they form part of the receiver or transmitter chain rather than the frequency synthesizer. Synthesizer designers are for this reason rarely involved in mixer design despite the fact that the synthesizer's RF oscillator must drive the mixer LO port which represents in most practical cases a considerable capacitive load. The conventional approach of strict separation of transceiver and synthesizer design yields often in an unnecessary high power consumption, caused often by the need for power wasting interstage buffers. An approach of joint mixer and RF oscillator design on the other side promises more power efficient solutions. It is therefore prudent to study the mixer's needs in a first step and optimize the RF oscillator accordingly to drive the mixer at its best. For these reasons, the discussion of RF oscillators is preceded by a minute study of downconversion mixer design.

To be useful to wireless transceivers, mixers must satisfy a couple of key requirements besides providing the primary frequency translation. The former are determined during transceiver planing which essentially tries to distribute gain, linearity and noise onto the transceiver blocks in such a way that all blocks can be realized with decent effort. While transceiver planing is definitely beyond the scope of synthesizer design and can be found elsewhere [1], typical downconversion mixer specification for cellular applications are summarized in Table 5.1. Eminent characteristic of mixers in general is the large noise figure (NF), a consequence of the ineludible noise mixing process which accompanies the wanted signal band translation. The large NF requires mostly signal pre-amplification prior to mixing by a low noise amplifier (LNA) in order to lower the mixer noise contribution to the receive chain.

This Chapter investigates common mixer topologies for downconversion and describes how transistor size and current are linked to mixer performance parameters such as linearity and noise. Beyond this, the RF oscillators relevant mea-

sures, LO port capacitance and required LO amplitude, are determined. The Chapter is closed by experimental results of a 2GHz, deep submicron CMOS downconversion mixer.

Gain	SSB-NF	CP	IP3	Input Imp.
10dB	15dB	-10dBm	-5dBm	50Ω

Table 5.1: Typical downconversion mixer specifications for cellular receivers.

5.1. Mixer Topologies in Comparison

Active RF mixers can be split basically into three parts. A *low noise input stage* converts the RF signal from the voltage into the current domain and establishes further a controlled input impedance, typically 50Ω, to match the mixer to a filter or to the LNA. The signal in the current domain undergoes thereafter frequency translation by the help of the second part, which is a *current switching network*. The desired frequency translation is accomplished by multiplication of the mixer current by ±1. The *output load* finally converts the current domain signal back into a voltage to allow processing by the subsequent baseband blocks. This third mixer part may also provide filtering of unwanted high frequency components.

Three topologies for downconversion are at disposal. The schematics of the considered mixers are shown in Fig. 5.1, Fig. 5.2 and Fig. 5.3. The topologies are referred to as

* *Single balanced mixer (SBM)*,
* *Double balanced mixer with common gate input stage (CG-DBM)*
* *Single ended input, double balanced mixer (SE-DBM)*

In order to allow a fair comparison, it is assumed that all mixers provide the same single ended input impedance. The differential input of the CG-DBM is therefore extended by an ideal transformer which splits the RF signal into two anti-phase signals of half the amplitude each. A short noise analysis shows that the input stages of the three mixers share the same noise figure. The latter can be found as

$$NF_{input} = 1 + \gamma \qquad (5.1)$$

with γ the channel noise factor. The three mixers are hence believed to exhibit the same noise figure. Mixer linearity and conversion gain, which is defined as the
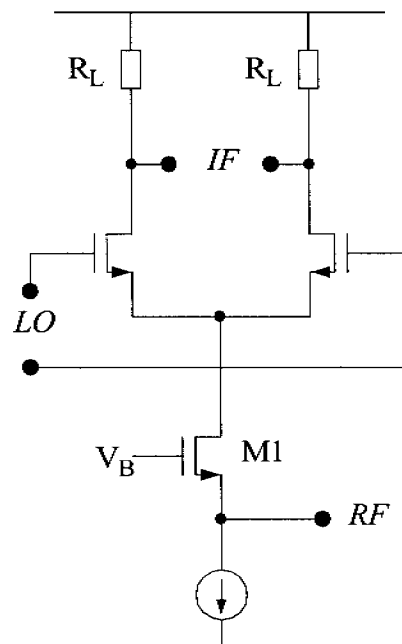
Figure 5.1: The single balanced mixer (SBM).

ratio of the signal amplitude at IF and signal amplitude at RF, are therefore the properties of interest for a comparison. Input referred third order intercept point (IP3) of CMOS mixers was found as [2]

$$IP3 = 10 \cdot [2.78 - \log(\theta) - \log(g_m/I_0)] \qquad (5.2)$$

with g_m and I_0 representing the transconductance of the input stage transistors and their bias current respectively. The parameter θ, or THETA in the various mutations of the SPICE circuit simulator, expresses mobility degradation due to the transverse electrical field in a MOS channel. Notice that the simple quadratic current characteristic of MOS transistors exhibits no third order distortion. Estimation of third order distortion requires therefore consideration of mobility degradation effects expressed by θ.

Although it is known that IP3 of mixers implemented in deep submicron CMOS departs strongly from the by Eq. 5.2 predicted value, an observation explained by transistor biasing close to the weak inversion region [2], the formula is nevertheless a valuable instrument for first order linearity comparison of the considered mixer architectures. The last term in Eq. 5.2 predicts that IP3 scales with the input transistor overdrive voltage. This outcome forms the basis for a power con-
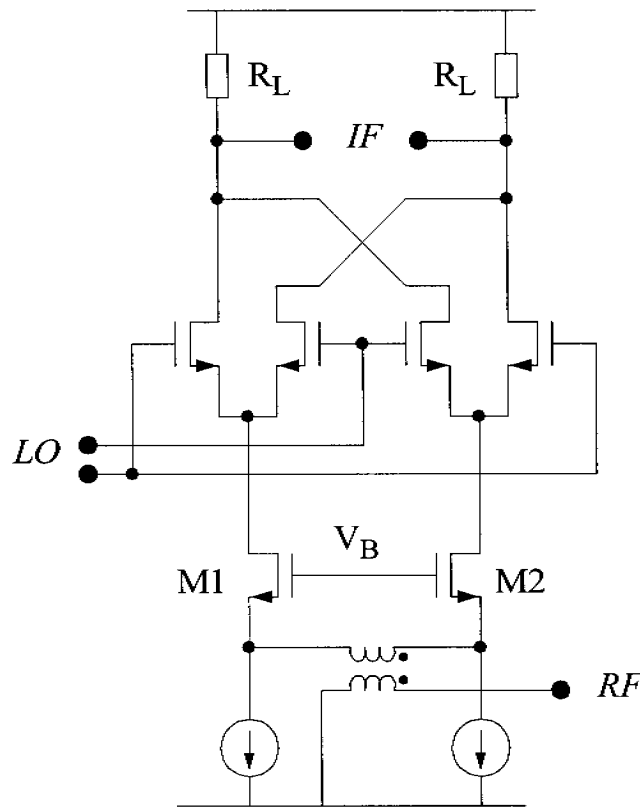
Figure 5.2: The common gate input, double balanced mixer (CG-DBM).

sumption comparison carried out now. Let us compare first the SBM with the CG-DBM topology. Both mixers share the same conversion gain formula

$$G_{SBM} = G_{CGDBM} = \frac{2}{\pi} \cdot g_m \cdot R_L \tag{5.3}$$

However, since the CG-DBM requires twice the g_m in the input stage transistors (i.e. 40mS for 50Ω matching), the CG-DBM provides twice the gain of the SBM. This holds true of course only for identical load resistors R_L. One could suspect that the CG-DBM consumes four times the current of the SBM since two 40mS transistors are needed compared to a single 20mS transistor. The CG-DBM consumption, however, can be lowered by taking nonlinearity into account too. Since the CG-DBM input transistors must handle only half of the signal swing each, they require only half the overdrive voltage for the same IP3. This allows quadrupling the transistor W/L ratio while maintaining the bias current of the SBM. The CG-DBM consumes therefore only twice the current of the SBM rather than the suspected quadrupled consumption. The larger consumption compared to the
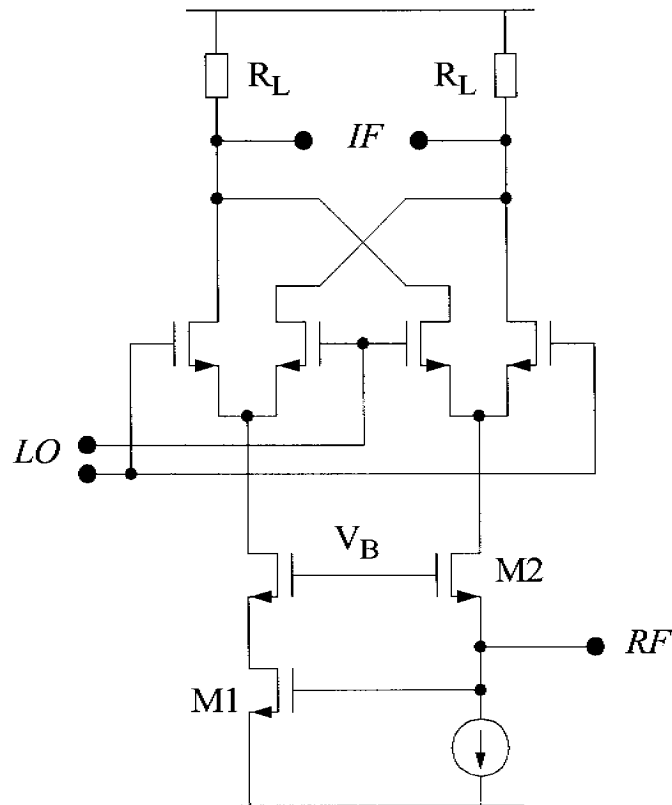
Figure 5.3: The single ended input, double balanced mixer (SE-DBM).

SBM is justified by a doubled conversion gain and, more important, by the double balanced mixing operation.

The SE-DBM, a structure inspired by the bipolar micromixer [3], combines a single ended input with double balanced mixing operation. The conversion gain of

$$G_{SEDBM} = \frac{4}{\pi} \cdot g_m \cdot R_L \qquad (5.4)$$

is identical to the CG-DBM since a single 20mS common gate connected transistor accommodates 50Ω matching. As the full swing applies to both the common source and the common gate transistor of the class-AB input stage, the transistors must be sized and biased as the input stage transistor of the SBM, if the mixer is to exhibit the same IP3. Power consumption is therefore twice the consumption of the SBM. Again, double balanced mixing is paid by twice the consumption. However, the advantage of the SE-DBM is the single ended input which removes the need for a balun. The SE-DBM eases the interface to single-ended LNAs or filters, but it does not reduce power consumption.

This comparison demonstrates that the mixer consumption can be lowered only by using a single balanced mixer or by compromising the mixer linearity. Single balanced mixing is in general only an option to super-heterodyne receivers which do not demand suppression of LO feedthrough.

5.2. Noise in Mixers

The SNR at the mixer IF port is rather low due to downconversion of uncorelated noise from various frequencies to the same IF, while signal power is converted to IF from a single frequency band only (or from two bands in case of double side-band mixing). Consequently, the single sideband noise figure (SSB-NF) of an ideal mixer, i.e. a mixer which is free of any internal noise sources, measures already 3dB since mixing of noise from the image band halves the SNR at IF. This holds only if the conversion gain from the image band to IF is identical to the conversion gain of the RF band to IF, an assumption which holds in most cases. This ideal 3dB SSB-NF is even pessimistic. It may be increased by mixing of noise above the RF band to IF. This can be understood by the fourier series expansion of the ideal ±1 switching function as:

$$V(t) = \frac{4}{\pi} \cdot \left[\cos(\omega_{LO} \cdot t) + \frac{\cos(3 \cdot \omega_{LO} \cdot t)}{3} + \frac{\cos(5 \cdot \omega_{LO} \cdot t)}{5} + \dots \right] \quad (5.5)$$

The odd order harmonics at $3\omega_{LO}$, $5\omega_{LO}$ etc. are responsible for downconversion of additional noise. Fig. 5.4 illustrates the noise mixing process. Fortunately, the amplitudes of the harmonics drop quickly, so that the SSB-NF increases only slightly to $\pi^2/4$ or 3.9dB.

The noise figure of a more realistic mixer employing one of the considered input stages is hence expressed as

$$NF_{SSB, mixer} = 10 \cdot \log(1 + \gamma) + 3.9dB \quad (5.6)$$

or 6.1dB assuming a channel noise factor of 2/3 which applies to large channel transistors. "Real world" mixers however exhibit SSB-NF in the region of 10...15dB. Explanation for the raised noise figure are on the one hand the increased channel noise factor observed in deep submicron CMOS transistors and, on the other hand, additional noise introduced by the switching transistors [4]. The NF at low IF frequencies may even exceed the mentioned range due to

upconverted flicker noise. While flicker noise of the input stage does not contribute to noise at IF, flicker noise of the switching transistors leaks to IF by random modulation of the switching function. It was shown that flicker noise at IF can be minimized by using large switching transistors driven by a large LO swing [4].
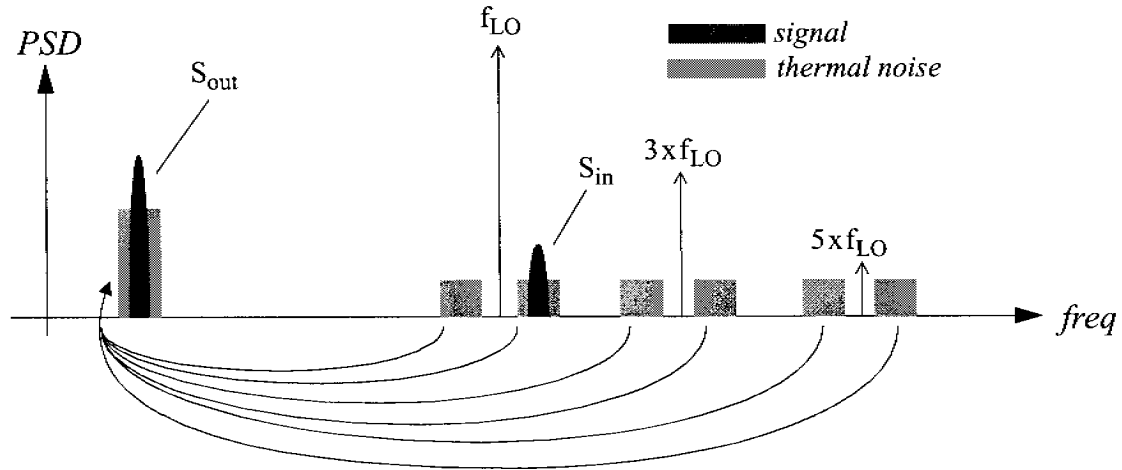
Figure 5.4: Illustration of the noise mixing process in downconversion mixers.

5.3. Switching Loss

The previous Sections have elaborated so far the dependency of input impedance, conversion gain, linearity and noise from the mixer input stage. Besides the input stage, the switching transistors impact conversion gain and noise as well. While too small sized switching transistors or a too small LO amplitude cause an undesired degradation of conversion gain and noise figure, too large sized transistors burden the RF oscillator unnecessarily by increased capacitive loading besides other negative effects. A thorough understanding of the current switching process by means of MOS differential pairs is hence crucial. The conversion gain loss due to imperfect switching is studied now for this reason.

Ideal switching of the signal current by ± 1 causes not only the desired frequency translation, but also a signal loss equal to $2/\pi$ or 3.9dB. The switching loss stems from the $4/\pi$ fundamental of a unity square wave and the trigonometric identity

$$
\begin{aligned}
A \cdot \cos(\omega_1 \cdot t) \cdot \cos(\omega_2 \cdot t) \\
= \frac{A}{2} \cdot \left[\cos(\omega_1 \cdot t - \omega_2 \cdot t) + \cos(\omega_1 \cdot t + \omega_2 \cdot t) \right]
\end{aligned}
\tag{5.7}
$$
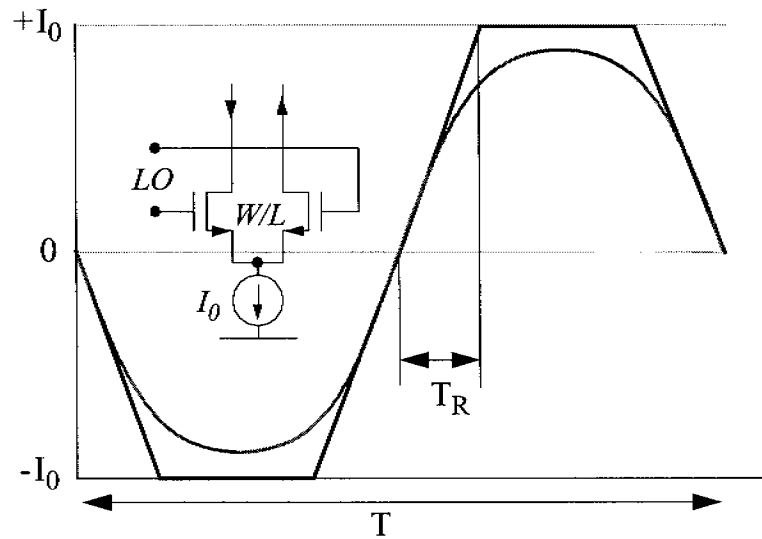
Figure 5.5: Differential, nonlinear switching current and its piecewise linear approximation.

underlying that the amplitude of the downconverted signal scales linearly with the input amplitude besides reduction by a factor two. Imperfect switching by MOS transistor differential pairs raises the switching loss above the ideal value of 3.9dB. The nonlinear switching function provided by a MOS differential pair driven by a sinusoidal LO is shown in Fig. 5.5. The LO amplitude and switching transistor transconductance dependant switching function can be accommodated by a piecewise linear approximation also shown in Fig. 5.5. Remember that the same approximation was used in the CML inverter analysis carried out in Chapter 4.

The switching loss can be refined by replacement of the rectangular switching function by the piecewise linear approximation of the nonlinear switching function, leading to a transconductance and LO amplitude dependant switching loss *SL* as

$$SL = \frac{2}{\pi} \cdot \frac{\sin(\alpha)}{\alpha}$$
$$with \quad \alpha = 2\pi \cdot \frac{T_R}{T} = \frac{\sqrt{2}}{V_{LO}} \cdot \frac{I_0}{g_{m,sw}} \tag{5.8}$$

with V_{LO} and $g_{m,sw}$ representing the LO amplitude and the peak transconductance of the switching transistor, i.e. while conducting the full bias current. Fig. 5.6
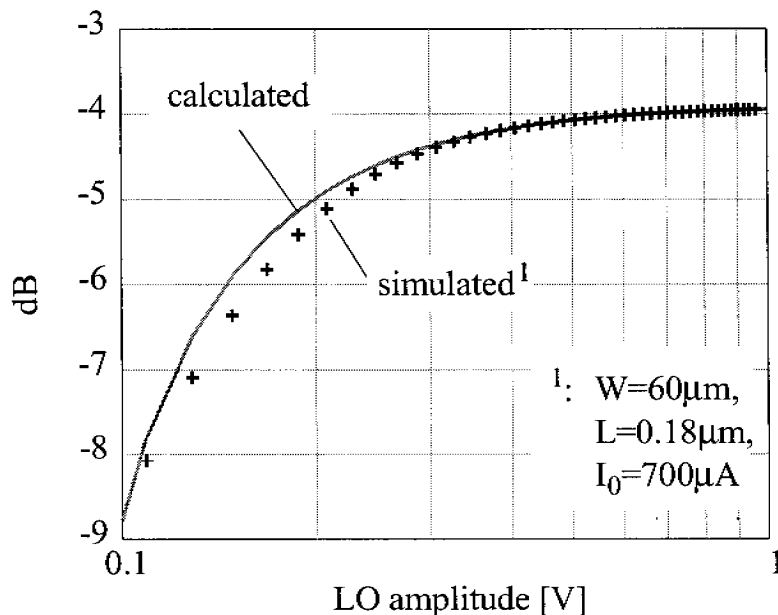
Figure 5.6: Calculated and simulated switching loss versus LO amplitude.

visualizes calculated switching loss as well as values obtained from a transient simulation. Good agreement is observed despite the simplicity of the approximation.

According to Eq. 5.8, the switching loss can be minimized by a large LO amplitude. However, the LO amplitude is limited by the available voltage headroom, which is already occupied to a large extent by the input stage. Remember that good linearity requires a large overdrive voltage of the input stage transistors. This reduces the headroom available to the switching transistors and limits thereby the LO amplitude. To keep the switching loss low, typically at 5dB or less, a small LO amplitude must be compensated by larger switching transistors. Eq. 5.8 evaluated at the largest possible LO amplitude which keeps the input stage still in saturation bounds a lower limit for the size of the switching transistors.

The quality of the switching function does not only affect switching loss, but also the level of flicker noise at IF, as mentioned before. Minimization of noise may require to increase the switching transistors size beyond a level dictated by switching loss minimization. The mixer implementation described in Section 5.4 underlines this fact.

5.4. Mixer Implementation

A single ended input, double balanced mixer according to Fig. 5.3 has been implemented in a standard 0.18μm CMOS technology. The mixer is designed for a nominal input frequency of 2GHz. The input stage, set up by a 60μm wide minimum length common source connected transistor M1 and a 40μm common gate connected transistor M2, both biased at 1.6mA, provides broadband matching to 50Ω and excellent linearity. Input referred IP3 is expected around 10dBm. The transistors of the input stage provide the same g_m of 20mS. This required slight shrinking of the common gate transistor which would provide otherwise too much effective transconductance due to the back gate effect. Switching could be performed by 120μm wide transistors driven by a differential 400mV LO. This would result in an acceptable switching loss of 4.7dB. Minimization of flicker noise at IF, however, commands doubling of the switching transistor width. The SSB-NF outside the flicker noise regime is expected at around 13dB.

The mixer is finally completed by 400Ω polysilicon load resistors and a biasing network responsible for the generation of two bias voltages required to define the operating point of the mixer input stage. Estimations of mixer noise and linearity are obtained from simulations carried out by the Spectre-RF circuit simulator [5].

5.4.1. Experimental Results

Input matching is verified by measurement of the S_{11} scattering parameter at the mixer RF port. Less than -12dB S_{11} is observed in the 1GHz to 3GHz range, attesting excellent broadband matching. The mixer's linearity is evaluated with a two tone test of which the result is shown in Fig. 5.7. The Figure shows the linear as well as second and third order intermediation tones as appearing at the IF port. The LO frequency and the frequency of the tones applied to the RF port are chosen as 2GHz, 2.004GHz and 2.005GHz respectively. The graph displays an input referred IP3 of +8.5dBm which is only slightly below the expected value. Further, input referred second order intercept point IP2 and the 1dB compression point CP are measured as +45dBm and -3dBm respectively.

The mixer SSB-NF is derived from measurements of the conversion gain and noise at the IF port. Noise figure numbers are computed from the following formula
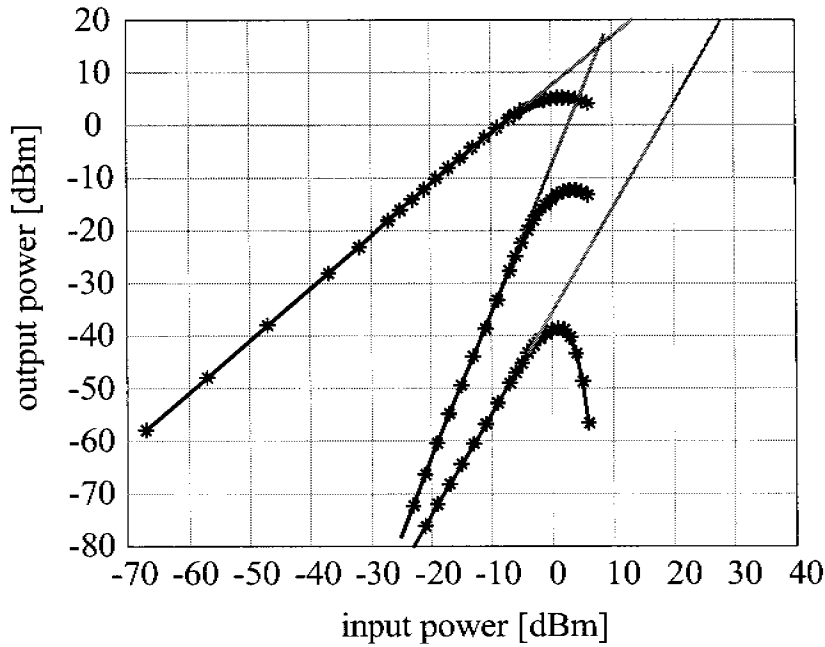
Figure 5.7: Mixer linearity evaluated by a two tone test. The plot shows the linear, second and third order intermodulation tones.

$$NF_{SSB,\,mixer} \;=\; \frac{SNR_{RF}}{SNR_{IF}} \;=\; \frac{S/\left(kTR_S \cdot f_{BW}\right)}{\left(S \cdot G^2\right)/N_{IF}} \;=\; \frac{N_{IF}/G^2}{kTR_S \cdot f_{BW}} \qquad (5.9)$$

with SNR_{RF} and SNR_{IF} the signal to noise ratio at the RF and the IF port respectively, R_S the mixer input impedance, S the signal power delivered to the mixer RF port, N_{IF} the noise power at IF, f_{BW} the bandwidth of interest and finally G the conversion gain. Input referred noise, i.e. noise PSD at IF normalized by the squared conversion gain, is measured outside the flicker noise regime as 2.1nV/√Hz. The noise spectrum at low IF frequencies up to a couple of 100kHz suffers from flicker noise as can be seen from the measured noise spectrum shown in Fig. 5.8. Again, the noise power spectral density is referred to the mixer input. Flicker noise however could be kept moderate by the help of the large switching transistors and the large LO amplitude. This explains why the averaged input referred noise in the band from 10kHz to 200kHz raises only slightly to 3.2nV/√Hz. These numbers correspond according to Eq. 5.9 to an SSB-NF of 13dB for the flicker noise free region and to 17dB for low IF in the mentioned frequency band. The two additional shadowed noise spectra of Fig. 5.8 illustrate
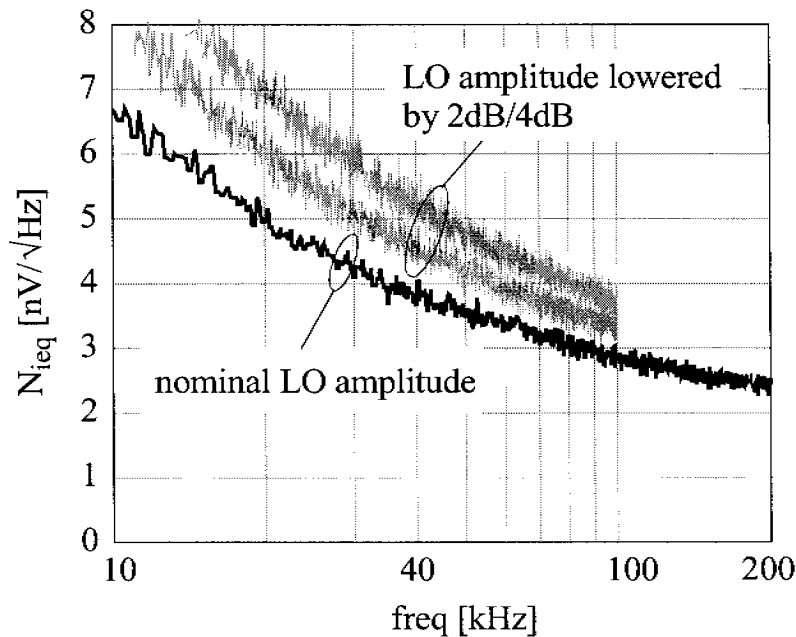
Figure 5.8: At the IF port measured noise spectrum, referred to the input. The black spectrum is taken from a measurement with nominal LO amplitude while the grey shadowed spectra display mixer noise at lowered LO amplitudes.

the dependency of flicker noise on LO amplitude. Not surprising at all, it can be seen that flicker noise raises as the LO amplitude is lowered. Table 5.2 finally summarizes the mixer performance and Fig. 5.9 presents a micrograph of the realized mixer chip.

5.5. Summary

Power is needed in CMOS downconversion mixers primarily for linearity. The required power consumption is accompanied by the need for large sized switching transistors to commute the mixer current. The size of these switches, and thereby indirectly the input capacitance of the mixer LO port, is determined by switching loss which must be kept small to maintain the mixer conversion gain. Even larger switches are in general required if the mixer must accommodate low IF.

One can conclude therefore that the demanding linearity and noise requirement of cellular applications impact not only the downconversion mixer consumption, but also the consumption of the RF oscillator due to enlarged capacitive loading

Input impedance:		50Ω
Matching:	S_{11}	< -12dB
Conversion gain:		9dB
Noise:	SSB-NF	13dB[1]/ 17dB[2]
Linearity:	IP3	+8.5dBm
	IP2	+45dBm
	CP	-3dBm
Consumption:	Supply	1.8V
	Current Drain	3.2mA
[1] *IF >1MHz* [2] *IF in 10kHz - 200kHz band*		

Table 5.2: Summary of the measured mixer performance.
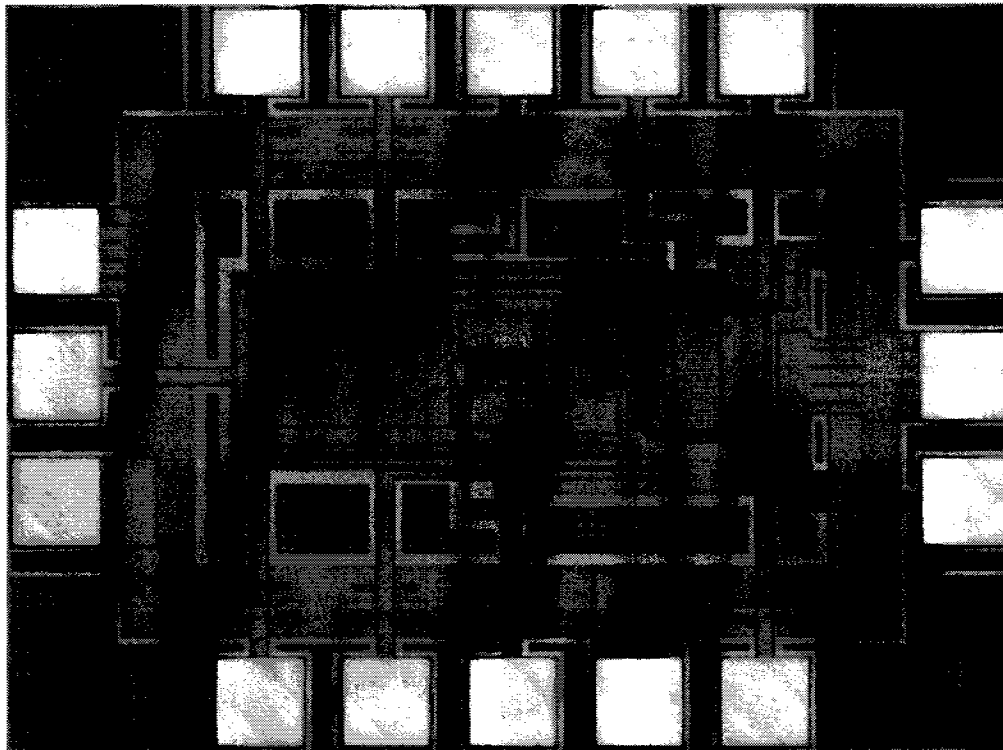
Figure 5.9: Die micrograph of the 0.18µm CMOS mixer. The integrated circuit measures 700µm x 900µm.

by the mixer LO port. This statement is underlined by the presented 2GHz CMOS downconversion mixer which exhibits excellent measured linearity and decent noise performance, but also a rather large LO port capacitance of 500fF. This definitely burdens the design of a low power RF oscillator.

References

[1] P. Orsatti, *"A Low Power CMOS GSM Transceiver for Small Mobile Stations"*, Hartung-Gorre, 2000, Series in Microelectronics, Vol. 92, pp. 34-41

[2] Q. Huang, F. Piazza, P. Orsatti, T. Ohguro, *"The Impact of Scaling Down to Deep Submicron on CMOS RF Circuits"*, IEEE J. of Solid State Circuits, Vol. 33, No. 7, July 1998, pp. 1023-1036

[3] B. Gilbert, *"The MICROMIXER: A Highly Linear Variant of the Gilbert Mixer Using a Bisymmetric Class-AB Input Stage"*, IEEE J. of Solid State Circuits, Vol. 32, No. 9, Sept. 1997, pp. 1412-1423

[4] H. Darabi, A. Abidi, *"Noise in RF-CMOS Mixers: A Simple Physical Model"*, IEEE J. of Solid State Circuits, Vol. 35, No. 1, Jan. 2000, pp. 15-25

[5] Cadence Design Systems, *"Datasheet Spectre-RF Simulation Option"*, 2003, see http://www.cadence.com/datasheets/spectrerf.html

Chapter 6

RF Oscillators

RF oscillators provide a periodic output signal with a tunable frequency. The latter is determined in general by a resonator (or tank), formed by an inductive device, mostly a lumped inductor, and a capacitive device. The resonant frequency can be tuned by alteration of the resonator capacitance. This is most often achieved by changing the reverse bias voltage of a varactor diode, revealing a voltage controlled oscillator (VCO). Oscillator performance is strongly related to the quality factor of the used resonator. The design of RF oscillators is therefore mainly driven by the search for better resonators. Besides careful choice of the resonator, the active devices of an oscillator must be sized deliberately to accommodate the resonator at best.

This Chapter investigates the low power design aspects of low phase noise RF oscillators. Solutions for oscillators in the 1 to 4GHz range are proposed and underlined with experimental results. Quadrature generation, i.e. the generation of 90° phase shifted oscillator output replicas, is considered as well.

6.1. LC-Tank Oscillator Fundamentals

Although different classes of oscillators exist, virtually only LC-tank oscillators and oscillators with other resonant devices are used for today's RF frequency synthesizers because of the tough phase noise requirements. Various topologies can be distinguished among the class of LC-tank oscillators. They are based on the same underlying principles despite their differing appearance. Oscillation is started by thermal disturbances of the circuit's DC-operation point, causing the oscillation to grow exponentially until steady state amplitude is reached. The frequency of the oscillation is determined by the tank's parallel resonance frequency. Oscillation growth is conceived by the active part of the oscillator of which a positive feedback composes an equivalent negative resistance. Evidently, the equivalent negative resistance, which scales with the transconductance of the

involved transistors, must overwhelm the tank loss to sustain the oscillation. Critical transconductance g_{mc}, defined as the transconductance which just compensates the tank loss, is an important oscillator measure. To ensure reliable oscillation, the transconductance of the transistors is chosen in general at least three times the critical transconductance.

The growth of amplitude is stopped by a counteracting effect, e.g. by clipping at the supply voltage rails. Clipping at the supply rails, however, is not power efficient and is further accompanied by an undesired increase of phase noise. Limiting the amplitude by a different mechanism is desirable. LC-tank oscillators provide inherently amplitude regulation, exempting the need for a dedicated amplitude regulation circuit. Once the amplitude is growing, the DC-operation of the active device shifts towards the transistor's off-state. The DC-operation point shift forces the amplitude growth to slow down until steady state amplitude is reached. Depending on the used topology, the active devices operates at steady state in class-B or class-C, i.e. the shifted DC-operation point allows the active device to conduct current during 50% or less of a cycle.

Various topologies are known to generate a negative resistance. They differ in the number of active devices and how often current is injected into the tank within a cycle. Tapping of the tank, either at the inductor or the capacitor, is often required to connect the transistors to the resonator. Tank tapping is mandatory if the active part is setup by a single transistor. Single transistor oscillators are the preferred choice of discrete oscillator modules where the number of devices must be minimized. Differential oscillators with two or four transistors are often used in integrated circuits since the differential output eases the connection to mixers and other differential devices.

Critical transconductance and steady state amplitude of single transistor LC-tank oscillators are elaborated at length in Section 6.1.1. The derived solutions can be effortlessly applied to differential oscillators which are covered in Section 6.1.2. Phase noise of LC-tank oscillators is shorty reviewed in Section 6.1.3. Besides this, low power design considerations are provided.

6.1.1. Single Transistor LC-Tank Oscillators

A negative resistance can be formed by a single transistor and a tapped capacitor. Fig. 6.1 shows the bipolar circuit as well as its small signal equivalent circuit. An inductor and a bias current source complete the oscillator. While the topology
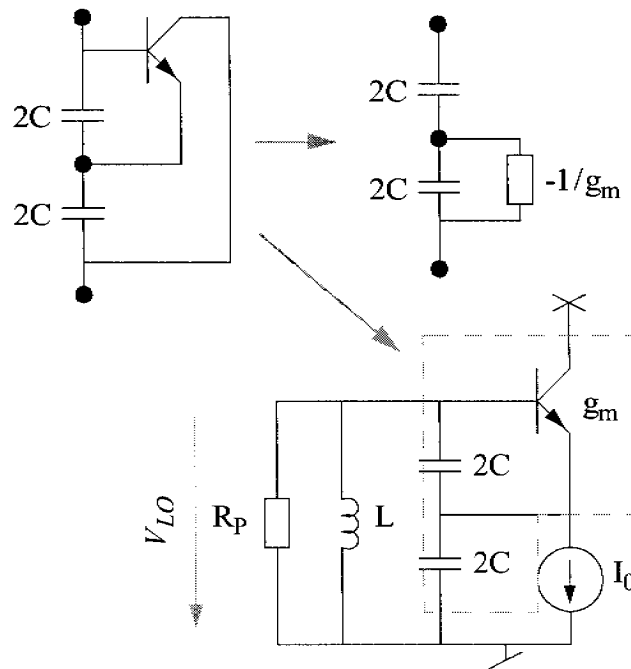
Figure 6.1: Negative resistance generation by a bipolar transistor and a tapped capacitor (upper left), small signal equivalent circuit (upper right) and schematic of the bipolar Colpitts oscillator (lower Figure).

with the collector node grounded is known as Colpitts oscillator, grounding of the emitter node leads to another popular oscillator referred to as Pierce oscillator. Fig. 6.1 shows the schematic of a bipolar Colpitts oscillator. Tank loss is represented by the shunt resistor R_P. Critical transconductance can be obtained by transformation of the small signal negative resistance $-1/g_m$ into a resistor parallel to the tank. The latter transformation is determined by the ratio of the two tank capacitors. To achieve best negative resistance, the tank capacitor must be tapped in the middle, i.e. the effective tank capacitance C must be split into two identical capacitors of twice the tank capacitance each. The critical transconductance of this optimum configuration can be found as:

$$g_{mc} = 4/R_p \qquad (6.1)$$

Steady state oscillator amplitude of the bipolar Colpitts is elaborated now. It is assumed that the bipolar transistor provides enough g_m to start the oscillation. The growing oscillation amplitude modulates the base-emitter voltage. Consequently, the emitter current is modulated as well. The growth of oscillation

amplitude is accompanied by a growth of the emitter DC-current. This phenomenon is caused by the transistor's nonlinear current characteristic, which is exponential in case of the bipolar transistors.

The excess DC-current flows into the tank capacitors. Thus, the DC-operation point of the oscillator is modified. A closer look shows that the average base-emitter voltage is lowered, counteracting thereby the excess current and the amplitude growth. In steady state, the oscillation amplitude and the DC-operating point are constant. The base-emitter voltage can then be formulated as

$$V_{BE}(t) = V_B + V_m \cdot \cos(\omega_0 t) \tag{6.2}$$

with V_B and V_m the constant steady state bias voltage and amplitude respectively. The emitter current delivered from the bipolar transistor can be found by Fourier series expansion as

$$I_E(t) = I_S \cdot e^{\frac{V_B}{U_t}} \cdot \left[J_0(V_m/U_t) + 2 \cdot J_1(V_m/U_t) \cdot \cos(\omega_0 t) + \ldots \right] \tag{6.3}$$

with U_t the thermal voltage and J_0 and J_1 modified Bessel functions of zero and first order respectively. By setting the DC part of Eq. 6.3 equal to the bias current I_0, the fundamental component of the emitter current can be derived as:

$$I_{fund} = 2 \cdot I_0 \cdot \frac{J_1(V_m/U_t)}{J_0(V_m/U_t)} \tag{6.4}$$

The ratio of the modified Bessel functions of Eq. 6.4 converges quickly towards unity for base-emitter modulating amplitudes larger than 100mV. Hence, the fundamental current is roughly twice the bias current. The AC part of the emitter current flows into the tank and forms with the (real) tank impedance the oscillator steady state amplitude. To compute the amplitude, the transimpedance from the emitter to the base node must be figured out. A short circuit analysis identifies the latter as $R_P/2$. The steady state oscillator amplitude is hence given by the simple identity:

$$V_{LO} \approx I_0 \cdot R_P \tag{6.5}$$

The latter Equation shows, that the steady state amplitude is primarily determined by the bias current and the tank parallel impedance.

A Colpitts oscillator can be constructed alternatively with a CMOS transistor. Computation of the CMOS Colpitts oscillator steady state is however more costly. The fundamental current component of the source current was derived as [1]

$$I_{fund} = I_0 \cdot \frac{5+x}{3} \tag{6.6}$$

with x a function of the g_m/g_{mc}-ratio. It was shown that x converges for large g_m/g_{mc} to unity and drops to 0.2 for a $g_m/g_{mc} = 3$. Since the g_m/g_{mc}-ratio must be chosen equal to 3 or larger to ensure reliable oscillation, the influence of the transistor g_m on the oscillator amplitude is less than 15%. Hence, the steady state amplitude of the CMOS colpitts oscillator is mainly determined by the product of bias current and the tank impedance, but is virtually independent of transistor g_m. It is worth to note, that the amplitude of the bipolar and the CMOS Colpitts oscillator are practically identical, if the same bias current and tank is applied. The choice of the semiconductor technology, CMOS or bipolar, is therefore not relevant to the performance of the oscillator. This rule holds for any LC-tank oscillator.

6.1.2. Differential LC-Tank Oscillators

Constitution of negative resistance by means of positive feedback is feasible with transistor only circuits. Diode connected transistors provide an equivalent small signal resistance $1/g_m$. If the gate could be connected to a node with the inverted drain voltage, a negative resistance $-1/g_m$ would be formed. This concept can be realized in differential circuits where inverted signals are available at the opposed half-circuit. A differential oscillator based on these considerations is shown in Fig. 6.2. Critical transconductance and steady state amplitude are now derived and compared with the Colpitts oscillator. The same bias current and total transconductance, composed of the transconductances of the cross coupled pair transistors, is presumed for a fair comparison. On top of this, except for inductor rather than capacitor tapping, the same tank with loss resistance R_P is presumed. The negative small signal resistance, realized by the cross coupled pair, is equal to $-4/g_m$. Since this resistance is parallel to R_P, the critical transconductance can be directly found as
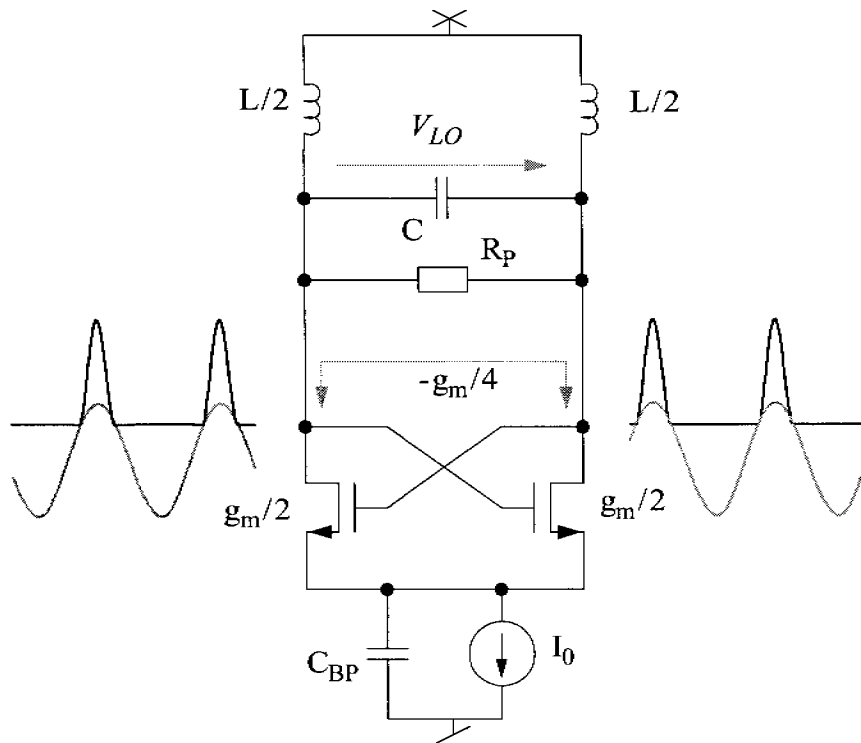
Figure 6.2: Differential oscillator topology.

$$g_{mc} = 4/R_p \qquad (6.7)$$

which is identical to the Colpitts g_{mc}. Computation of the steady state amplitude can be obtained from an analysis of the currents injected into the tank. The cross couple pair steers the bias current during one half-cycle to the left branch of the tank and to the right branch during the other half-cycle. Rectangular shaped drain currents with a fundamental component

$$I_{fund} = \frac{2}{\pi} \cdot I_0 \qquad (6.8)$$

determine together with the tank loss the steady state amplitude as:

$$V_{LO} = \frac{2}{\pi} \cdot I_0 \cdot R_P \qquad (6.9)$$

Different to the Colpitts oscillator, the transistors of the differential oscillator work in class-B due to the 50% duty cycle of the transistor current. This can be changed by introduction of a bypass capacitor C_{BP} at the common source node. As in the Colpitts oscillator, excess transistor DC-current charges this capacitor.
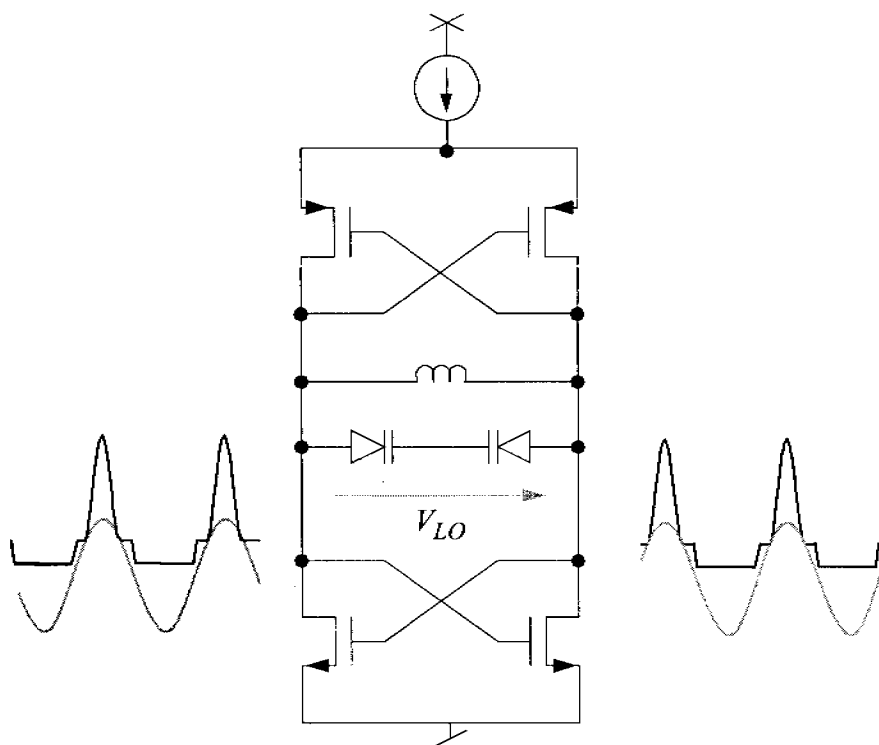
Figure 6.3: Double cross coupled pair oscillators

The raised voltage at the common source node pushes the transistor into class-C operation. Since the fundamental component of the class-C distorted current exceeds the fundamental component of the rectangular shaped current, the oscillation amplitude is moderately increased by the introduction of the bypass capacitor. With the fundamental current described by Eq. 6.6, the steady state amplitude of the CMOS differential oscillator is determined by the identity

$$V_{LO} = \frac{5+x}{6} \cdot I_0 \cdot R_P \qquad (6.10)$$

which is identical to the CMOS colpitts oscillator. Hence, the into class-C forced differential oscillator is in terms of critical transconductance and steady state amplitude equivalent to the Colpitts oscillator.

An advanced oscillator topology with power saving potential was proposed in [2]. The oscillator, shown in Fig. 6.3, introduces a second cross coupled pair which increases the fundamental current component. While the single cross coupled pair oscillator injects current into a tank branch only once each cycle, the double cross coupled pair oscillator provides current in both half-cycles. This is

illustrated in Fig. 6.2 and Fig. 6.3. Assuming class-B operation, the fundamental current component is exactly doubled, resulting in twice the steady state amplitude. The double cross coupled pair oscillator may be pushed into class-C as well with either both cross coupled pairs working in class-C or only one of them. The case with the PMOS pair being in class-B and the NMOS transistors in class-C is considered here. Class-C operation of the NMOS pair requires a lowered oscillator common mode voltage. This necessitates grounding of the tank capacitors, which are replaced in Fig. 6.3 by two common cathode varactor diodes to form a voltage controlled oscillator. The fundamental current is then given by:

$$I_{fund} = I_0 \cdot \left(\frac{2}{\pi} + \frac{5+x}{6} \right) \tag{6.11}$$

The larger fundamental current component permits lower bias current for the same oscillation amplitude. However, flaws of the double cross coupled pair oscillator may defeat its advantages, as will be shown later.

6.1.3. Phase Noise in LC-Tank Oscillators

Noise in oscillators is well known to modulate the oscillation frequency. Experiments exhibit phase noise power spectral inversely proportional to the squared offset from the carrier frequency, revealing an over several decades observable -20dB/dec slope of the oscillation noise spectrum. The process responsible for modulation of the oscillator phase was not fully understood for a long time. Designers relied in the past on Leeson's observation, which predicts a 6dB phase noise improvement by doubling the bias current or the resonator Q. The influence of the active devices on phase noise remained mysterious and had to be figured out experimentally.

Reported phase noise of oscillators with similar tank quality factors and biasing varies significantly, indicating that inadequately sized transistors strongly degrade phase noise. A recently published article [3] analyzes phase noise of the CMOS Colpitts oscillator rigorously. The relationship between thermal noise sources and phase noise was derived as

$$L(\Delta\omega) = \frac{kT}{4I_0^2 Q^2} \cdot \left(\frac{\omega_0}{\Delta\omega} \right)^2 \cdot \left[\frac{2}{3}g_{mB} + g_{mc} + \frac{2}{3}g_m \sqrt{\frac{acos(x)}{\pi}} \right] \tag{6.12}$$

with g_{mB} the transconductance of the bias current source, Q the tank quality, $\Delta\omega$ the offset from the carrier ω_0 and x the previously introduced function of the g_m/g_{mc}-ratio. A well designed oscillator distinguishes itself by moderately sized transistors. The active devices contribute then to approximately two thirds of the total transconductance expressed by the bracketed term in Eq. 6.12. The resulting phase noise, realized by choosing $g_m \approx 3g_{mc}$ and $g_{mB} \approx 0.2g_m$, can be rewritten as

$$L(\Delta\omega) = \frac{3kT}{R_P I_0^{\,2} Q^2} \cdot \left(\frac{\omega_0}{\Delta\omega}\right)^2 \tag{6.13}$$

whereby g_{mc} is replaced by R_P according to Eq. 6.1. Tank quality factor Q and the parallel loss R_P are interchangeable parameters. They are linked through the following identity:

$$R_P = \sqrt{L/C} \cdot Q \tag{6.14}$$

Combination of Eq. 6.14 and Eq. 6.15 leads to the enlightening identity:

$$L(\Delta\omega) \cong \frac{3kT}{I_0^{\,2} Q^3 \cdot \sqrt{L/C}} \cdot \left(\frac{\omega_0}{\Delta\omega}\right)^2 \tag{6.15}$$

The last Equation states that doubling the tank quality factor improves phase noise by 9dB rather than 6dB as postulated by Lesson's formula. However, the doubling of the quality factor must be accompanied by accordingly downsized transistor transconductances to achieve this improvement.

This review of phase noise, based on the analysis carried out in [3], emphasizes the importance of careful transistor sizing besides the overwhelming significance of high-Q oscillator tanks. Although the presented Equation is derived for the CMOS Colpitts oscillator, the important outcome, which is the proportionality of phase noise to the inverse of Q cubed, is valid for any LC-tank oscillator.

6.2. An Ultra-Low Power 1GHz CMOS Oscillator

Recently published fully integrated 1GHz CMOS oscillators with below -100dBc/Hz phase noise at 100kHz offset tend to consume 2 to 5mA [2] [4]. The main reason for the high power consumption, which may be even larger if GSM

phase noise requirements have to be fulfilled, is the low quality factor of the integrated LC-tank. Quality factor limiting device is in general the planar spiral inductor. While the design of a 10nH integrated inductor with a quality factor of 5 to 8 does not represent a great challenge anymore, significant quality factor improvement by more advanced spiral geometries stays away despite intensive research ongoing on this subject. Minor improvement was reported by the introduction of a patterned ground shield aimed to the minimization of substrate losses [5]. However, it seems that superior inductors can be expected only from modification of the IC technology, e.g. from addition of thicker top layer metallization and thicker oxide. These modifications raise the silicon die cost, annihilating thereby the inherent low-cost advantage of CMOS. The quality factor of integrated LC-tank resonators is further reduced by loss of the integrated varactor, which is implemented either as reverse biased pn-junction or as MOS gate capacitor. The overall resonator Q of a standard CMOS integrated LC-tank at 1GHz must be expected as low as 4 to 5.

Where power consumption is critical, the tank can be setup by a few discrete surface mounted devices (SMD) of much larger quality factor. Indeed, many commercial oscillators for portable wireless applications use external resonators [6]. Low cost, small size 10nH inductors with a quality factor of 80 are available, while hyperabrupt junction varactors enable a wide tuning range. Power optimization requires identification of the optimum inductor and capacitor combination for a given resonant frequency $\omega_0 = \sqrt{1/LC}$. It can be observed, that Q of discrete inductors is slightly increasing with the inductance value while at the same time Q of the junction varactors raises with decreasing capacitance. Maximum tank quality factor is hence achieved with a large inductance, small capacitance combination. The lower limit for the varactor capacitance is determined by parasitic capacitances, caused by package, ESD protection etc. Reliability and tuning range considerations command a varactor capacitance considerably larger than the parasitic capacitance.

A differential oscillator has been setup with a discrete tank resonator. A high-Q resonator is established by a 10nH SMD inductor and two common cathode connected varactors, housed in a single SMD package. The measured resonator quality factor is approximately 20. This is roughly 4 times the Q of an equivalent integrated CMOS tank. According to Eq. 6.15, the higher quality factor allows reduction of the bias current by one decade while phase noise remains

unchanged. Oscillators in the 1GHz range with below -100dBc/Hz phase noise at 100kHz offset, biased with only some 100uA should be therefore feasible.

6.2.1. Experimental Verification

As differential oscillator with a double cross coupled pair according to the schematic shown in Fig. 6.3 was chosen to demonstrate the low power, low phase noise capabilities of discrete tank oscillators. The double cross coupled pair topology and the discrete tank with a quality factor of 20 permit a a low bias current of only 250μA. The circuit, which is laid out in a standard 0.25μm CMOS technology, provides besides the oscillator a buffer for measurement purposes. The pads, including those to the external tank, are ESD protected. All measurements are carried out with the integrated circuit directly glued onto the PCB.

Fig. 6.4 shows the spectrum of the oscillator. Phase noise was measured as low as -112dBc/Hz at 100kHz offset from the 1GHz carrier despite the very low bias current. Measured phase noise of -118dBc/Hz at 200kHz offset indicates that no upconverted flicker noise degrades the performance of the oscillator. The oscillator can be tuned from 850MHz to 1020MHz with only 2V tuning voltage range. Phase noise stays below -110dBc/Hz over the full tuning range, but increases due to varactor forward biasing as the oscillator is tuned below 850MHz. Fig. 6.5 shows the oscillator tuning range. The same Figure provides measured phase noise at 100kHz offset from the carrier at different tuning voltages.

RL -10.7dBm   10dB/

CENTER 1.0218642GHz   SPAN500.0kHz
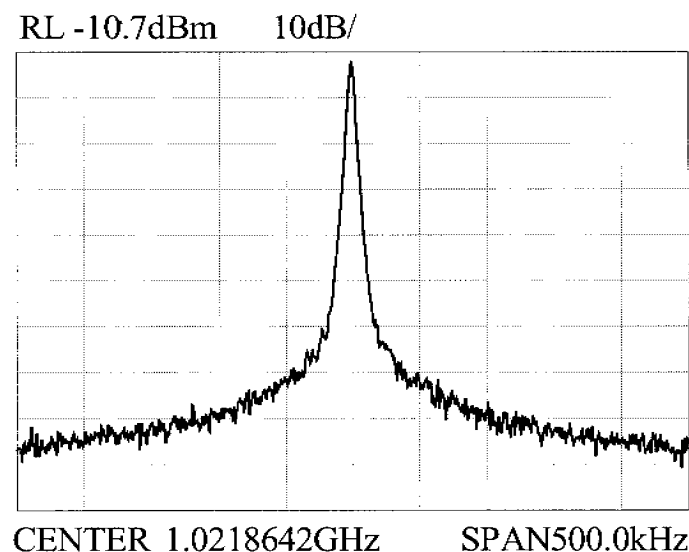
Figure 6.4: Spectrum of the 1GHz 250μA discrete tank oscillator.
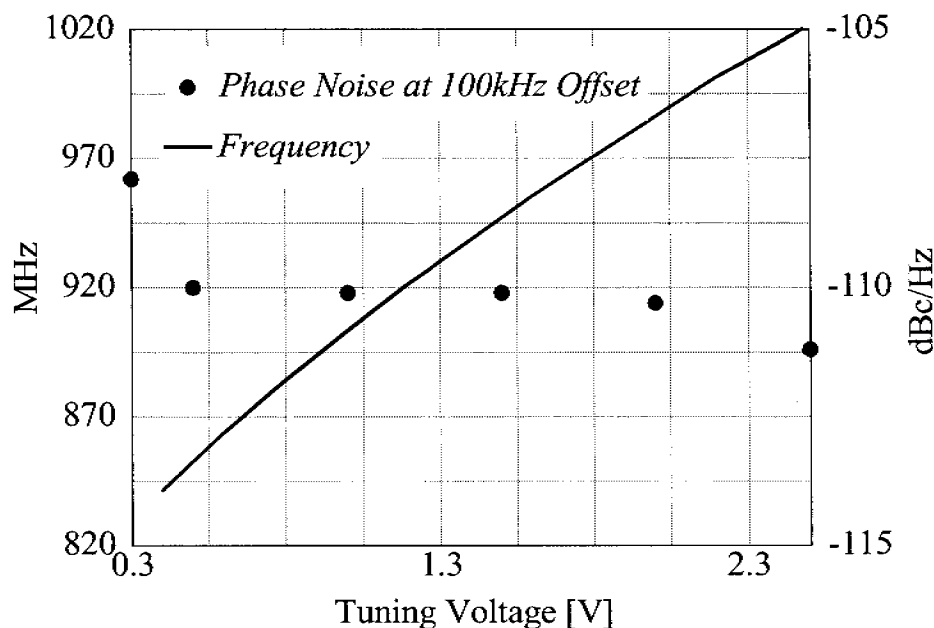
Figure 6.5: 1GHz oscillator tuning range and phase noise at 100kHz offset from the carrier at varying tuning voltages.

Due to the large oscillator tuning gain of 85MHz per Volt and the low phase noise level, the oscillator is susceptible to low frequency alterations in the DC common mode level, which generate additional FM noise around the carrier. To overcome this limitation, the anodes of the varactors must be decoupled from the oscillator. Two RF chokes provide a low ohmic path to ground. However, this increases the number of discrete components. The differential oscillator with a single cross coupled pair shown in Fig. 6.2 suffers less from FM noise since the varactors are wired to the positive supply voltage through the tank inductors. A similarly implemented oscillator with a single NMOS cross coupled pair consumes 60% more current, but does not require additional decoupling devices. Measured phase noise is close to the double cross coupled pair oscillator.

6.3. Quadrature Generation Concepts

Quadrature signals can be generated from a local oscillator in several different ways. The most common is the RC-CR network. The passive network, which represents a combination of a first order low-pass and a first order high-pass section, provides 90° phase difference between the I and Q outputs. The phase difference is insensitive to the absolute RC time constants, but requires good

matching of the R and C elements. Differences of the output amplitudes must be equalized by limiting amplifiers. At 900MHz 30-35dB rejection has been achieved by such an implementation in 0.25μm CMOS technology, although the 12mA current consumption is still too high for a receiver [7]. Poly-phase filters are another popular way to realize quadrature signals. Similar to the RC-CR network, matching of R and C elements is important for good image rejection, which dictates low impedance levels for the network. High power consumption is often reported for such filters due to required buffers. A pair of oscillators can be coupled together in such a way that they oscillate at exactly the same frequency, but maintain 90° phase difference between the two [8]. Performance of such quadrature oscillators has not been very consistent, however.

An approach often used for quadrature generation at low frequencies is to double the LO frequency and then divide its output with a digital divider, such as a master-slave flip-flop in toggle configuration. If the LO signal has 50% duty cycle, the phase difference between the master and slave output is exactly a quarter of the output period. Previously such an approach has been considered unsuitable for low power design at RF frequencies, because the consumption of the divider tended to be high. With scaling of CMOS technologies, however, it has become considerably easier to implement low power digital circuits at several GHz, therefore the digital approach to quadrature generation deserves re-evaluation. It is well known that the oscillator current consumption raises rapidly with the oscillation frequency. It is therefore important that both oscillator and the divider are designed for low power, so that the combined solution is competitive with other approaches to quadrature generation. This urges the need for high-Q resonators in the 2GHz range, respectively 4GHz range to accommodate DCS/UMTS services. The design of high-Q resonators in this frequency range is a challenging task.

6.4. Multi-GHz Planar Transmission Line Resonators

Low device self-resonances restrict the application of discrete LC-tank resonators to frequencies up to 2-3GHz. Printed circuit board based planar transmission lines are an attractive LC-tank replacement at higher frequencies. Their operation frequency is limited mainly by parasitics of the transmission line to chip interface. Standard wire bonding technique permits transmission line resonators in the low GHz range. Resonators up to 10GHz can be realized with flip-chip tech-

nique. Different to discrete LC-tank resonators, the application of transmission line resonators is limited at low frequencies as well. They are considered too bulky at frequencies where discrete LC-tanks are doing well. However, the physical dimensions of a transmission line resonator reaches the tolerable 1cm region at 4GHz. Planar transmission lines combine besides the high frequency operation a couple of interesting properties. An eminent advantage is the simple implementation by means of printed circuit board (PCB) copper tracks on a single or double ground plane, referred to as microstrip or striplines respectively. Further, transmission lines are known to provide very high quality factors. Less obvious properties of transmission lines tuned by varactors are derived in the course of this Section.

The investigations are focussed on microstrip resonators with open-circuit termination and varactor tuning at the input. While short-circuit termination could be used too, the realization of a PCB short-circuit at several GHz may cause serious problems. The impedance seen at the input of an open-ended transmission line is sketched in Fig. 6.6. Depending on the length of the transmission line, the impedance varies from capacitive to inductive and forms a parallel resonance at a physical length equal to half the wavelength. Evidently, the microstrip input impedance must behave inductively to resonate with the varactor, limiting the useful transmission line length from one fourth to one half of the wavelength λ_g. While a line length close to $\lambda_g/4$ commands an excessively large varactor, a line length chosen near $\lambda_g/2$ would result in a poor tuning range. Microstrip length at around $0.35 \cdot \lambda_g$ can be considered as a balanced compromise.

The microstrip resonator must be optimized for highest quality factor in order to minimize oscillator power consumption and phase noise. Evaluation of the transmission line resonator quality factor is however not a straightforward task. Q factors of microstrip lines are reported in the literature only for $\lambda_g/2$ resonators [9], but not for arbitrary wavelength. Further, the high-Q microstrip line is tuned by a varactor with an in general much lower quality factor. A rigorous quality factor analysis taking wavelength and varactor loading into account is presented in Section 6.4.1. On top of this, the optimum microstrip track width for highest quality factor is worked out in Section 6.4.2. Section 6.4.3 finally investigates more advanced transmission line structures.
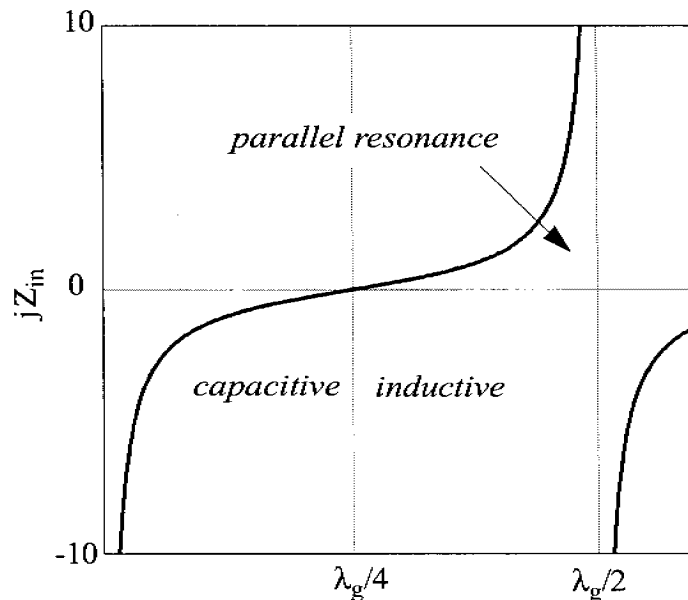
Figure 6.6: Input impedance of an open-circuit terminated transmission line versus physical length of the transmission line.

6.4.1. Tuned Transmission Line Resonator Quality Factor

Investigation of the quality factor of an open circuit terminated, varactor tuned transmission line resonator requires to touch base with the term quality factor or Q in general. Although quality factor is widely used in different fields, e.g. filtering, it is often a source of misinterpretation. Different definitions have established in the past, which unfortunately are not necessarily self-consistent. The most universal interpretation of quality is concerned with the efficiency of energy storage. Storage of energy in physical systems is always accompanied by an undesired dissipation of a fraction of the stored energy. A Q definition based on the ratio of stored and dissipated energy is intuitive and has further the advantage that it is not limited to a particular system or to systems of a particular order. A fundamental definition of Q based on energy ratios is given as [10]

$$Q = 2\pi \frac{E_{stored,pk}}{E_{loss}} \tag{6.16}$$

with $E_{stored,pk}$ the peak energy stored in the system and E_{loss} the during a single cycle dissipated energy. Notice that this definition coincides with Q definitions by means of bandwidth, if the considered system is of second order.

An analytical derivation of the quality factor according to Eq. 6.16 requires insight into the energies stored in a varactor tuned transmission line resonator. Since the energies scale quadratically with the applied voltage, the resonator voltage must be defined in a first step. Assuming that the resonator is embedded in an oscillator which stabilizes the amplitude, the voltage at the resonator input can be formulated by the Equation

$$V(t) = V_0 \cdot \cos(2\pi f_r t) \tag{6.17}$$

with V_0 the steady state amplitude and f_r the resonant frequency. The resonant frequency is determined by the varactor capacitance C_{var}, the transmission line electrical length θ and the transmission line's characteristic impedance Z_0. The latter measures are linked by the following transcendental Equation:

$$\tan\theta = -2\pi f_r \cdot C_{var} \cdot Z_0 \tag{6.18}$$

Computation of the time varying energy stored in the transmission line requires rather exhaustible mathematical manipulations which are presented in Appendix 6.A. The transmission line energy E_{ms} was found there as:

$$E_{ms}(t) = \frac{V_0^2}{8\pi \cdot f \cdot Z_0} \cdot \left[\frac{2\theta + \sin(2\theta) \cdot \cos(2\omega t)}{2\cos^2\theta}\right] \tag{6.19}$$

Computation of the varactor energy is obtained effortlessly due to the lumped nature of this resonator part. The varactor voltage defined by Eq. 6.17 and the resonance condition expressed by Eq. 6.18 lead directly to a formulation of the time varying varactor energy as:

$$E_{var} = \frac{V_0^2}{8\pi \cdot f \cdot Z_0} \cdot \left[-\tan(\theta) \cdot \left(1 + \cos(2\omega t)\right)\right] \tag{6.20}$$

To gain more insight into the stored energies, transmission line and varactor energies are visualized in Fig. 6.7. Based on Eq. 6.19 and Eq. 6.20, the Figure illustrates the energy of a $0.35 \cdot \lambda_g$ sized transmission line resonator at the resonant frequency. Unlike LC-tank resonators, the energy of the inductive resonator part, i.e. the transmission line, never vanishes, and more important, is roughly three times larger than the energy of an equivalent inductor of which the energy is also shown in Fig. 6.7 for comparison purposes. The three times larger peak resonator
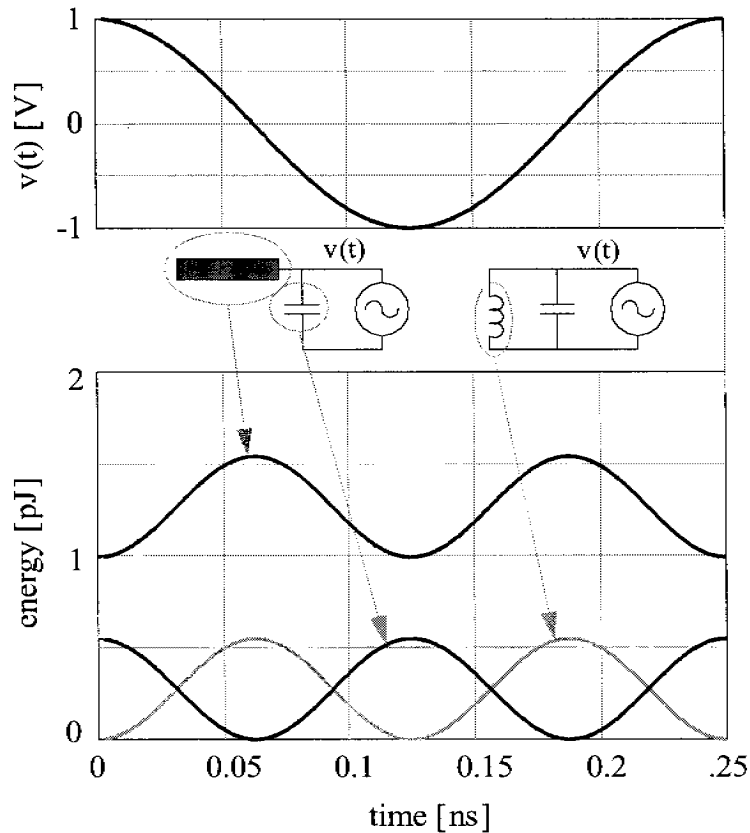
Figure 6.7: Energies of an input connected, varactor tuned open-circuit termi-nated transmission line resonator with 0.35·λ$_g$ length. The grey line illustrate the inductor energy of an equivalent LC-tank resonator.

energy level compared to LC-tank resonators has, not surprising at all, impact on the resonator quality factor, as can be seen in the following investigation.

According to the definition of Q applied to the varactor tuned resonator, the reso-nator Q can be found as

$$Q_{res} = 2\pi \cdot \frac{E_{ms,pk}}{E_{loss,ms} + E_{loss,var}} \qquad (6.21)$$

with $E_{loss,ms}$ and $E_{loss,var}$ being the microstrip and varactor losses per cycle. A rearrangement of Eq. 6.21, with the help of Eq. 6.19 and Eq. 6.20, leads to the more meaningful expression:

$$\frac{1}{Q_{res}} = \frac{1}{Q_{ms}} + \frac{1}{x \cdot Q_{var}} \qquad \text{with} \quad x = \frac{\sin(2\theta) - 2\theta}{2\sin(2\theta)} \qquad (6.22)$$

Q_{ms} and Q_{var} represent the quality factors of the unloaded transmission line and varactor respectively and x depicts the ratio of transmission line and varactor peak energy, a measure which is larger than unity, as can be seen in Fig. 6.7. The vantage of the transmission line resonator gets more evident by a comparison with the LC-tank resonator quality Q_{LC}. The latter is given by the following formula

$$\frac{1}{Q_{LC}} = \frac{1}{Q_{ind}} + \frac{1}{Q_{var}} \tag{6.23}$$

with Q_{ind} the unloaded inductor quality factor. While the Q of both structures depends on the unloaded quality factor of the individual resonator devices, the LC-tank Q is always smaller than the unloaded Q of the individual tank devices. The varactor tuned microstrip resonator Q, on the other hand, may exceed the unloaded Q of the varactor. It is not only the varactor Q boosting property of the transmission line resonator that makes this resonator type superior to the integrated LC-tank resonator, but also the exceptional large Q of the transmission line. The unloaded Q of a PCB based planar transmission line tends to be not only one to two orders of magnitude larger than the unloaded Q of an integrated spiral inductor, it is further considerably larger than the unloaded Q of varactors. Consequently, the transmission line Q is mainly determined by the boosted varactor Q:

$$Q_{res} \approx x \cdot Q_{var} \tag{6.24}$$

One can conclude therefore, that a tuned transmission line resonator is able to provide a quality factor considerably larger than that of an integrated LC-tank. The quality enhancement factor x depends on the electrical length of the transmission line. Fig. 6.8 illustrates the dependency. It measures approximately 3 for a transmission line length in the useful range.

6.4.2. Unloaded Q of Planar Transmission Lines

Microstrip and striplines represent the most suitable devices to constitute a varactor tuned transmission line resonator. Compatibility with printed circuit boards and a high quality factor, achievable even with low-cost laminated substrates, are their eminent advantages. Nevertheless, the transmission line must be optimized carefully since non-optimum transmission line dimensions may significantly
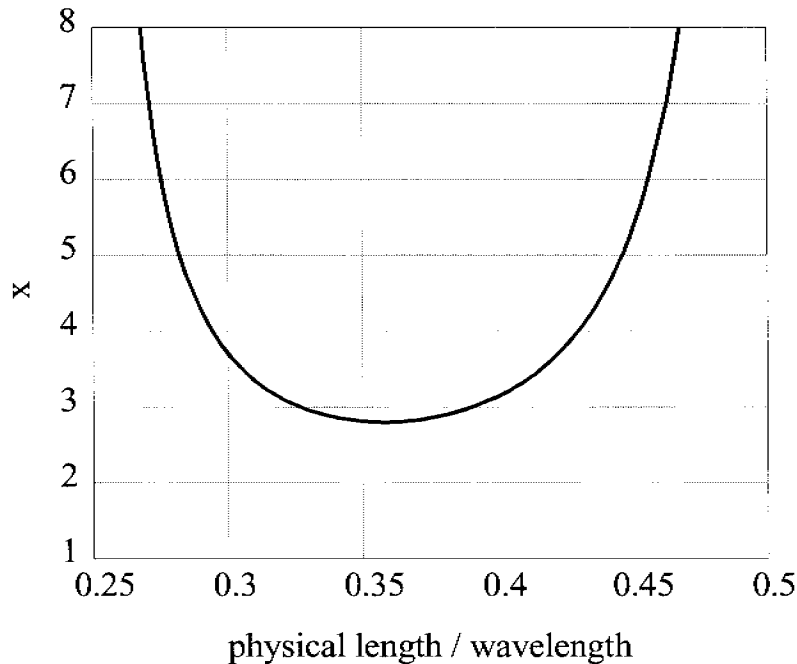
Figure 6.8: The factor x expresses the radio of the peak transmission line energy and the peak varactor energy.

degenerate the exceptional quality factor. The unloaded Q depends on a variety of parameters of which most are determined by other than quality factor considerations. Among these parameters are the kind of substrate material, which is restricted to FR4 or similar materials for cost reasons, the thickness of the substrate, the electrical length and the resonant frequency. Optimization of the planar transmission line is therefore limited to a single parameter, which is the conductor track width, or the characteristic impedance which directly relates to the track width. Since neither very wide nor very narrow tracks are advantageous, an optimum width must be presumed. A closer look at the processes responsible for the introduction of loss is crucial to understand this fact.

Planar transmission line loss is dominated in the low GHz range by three sources: Conductor, dielectric and radiation loss. Conductor loss is caused by the finite resistivity of the conductor tracks which are made of thin copper cladding. This loss increases due to the skin effect with the square root of the frequency. Conductor loss is minimized by wide tracks, i.e. by a low characteristic impedance. Dielectric loss is caused by the finite conductivity of the substrate material. It is mainly determined by material properties and to a much lesser extent by the

transmission line geometry. Narrow microstrip tracks slightly reduce the dielectric loss since a considerable fraction of the electrical field penetrates the air above the track. This approach lessens dielectric loss by a factor two at best. Dielectric loss is further frequency independent. The third source of loss is radiation. Both ends of the transmission line radiate energy which is lost if the transmission line is not perfectly shielded. The lack of shielding applies in general to microstrip lines. Narrow tracks combined with a thin substrate reduce radiation. Radiation loss raises quickly at higher frequencies since it scales with the square of the frequency.

The above quantitative formulations are underlined by a rigorous analysis. Energy dissipated in the conductor, the dielectric material and energy lost by radiation are formulated in Appendix 6.A. The provided results, combined with Eq. 6.16 and Eq. 6.19, submit the formulation of the conductor Q_{cond}, dielectric Q_{diel} and radiation Q_{rad} quality factor which belong each to one of the sources of loss. The three individual Q factors are found as:

$$Q_{cond} = \frac{\pi}{\alpha_c \lambda_g} \tag{6.25}$$

$$Q_{diel} = \frac{\pi}{\alpha_d \lambda_g} \cdot \left[\frac{2\theta - \sin 2\theta}{2\theta + \sin 2\theta} \right] \tag{6.26}$$

$$Q_{rad} = \frac{1}{4 Z_0 G_R} \cdot \left[\frac{2\theta - \sin 2\theta}{1 + \cos^2 \theta} \right] \tag{6.27}$$

The latter Equations are generic so far, i.e. they apply to any transmission line geometry. The geometry dependency of the transmission line is carried in the conductor loss constant α_c, the dielectric loss constant α_d and the equivalent radiation shunt conductance G_R. Notice that the unity of the loss constants is Neper per meter. Semi-empirical approximations for the loss parameters as well for the characteristic impedance are listed in Appendix 6.B. Notice that the approximations apply to microstrip lines, but not to striplines.

In order to underline the quantitative considerations stated earlier, the unloaded Q of a 4GHz microstrip line of $0.35 \cdot \lambda_g$ length is evaluated. The outcome is visualized in Fig. 6.9. Electrical length and substrate material correspond to the oscillator which will be discussed later in Section 6.5. It can be seen that the optimum
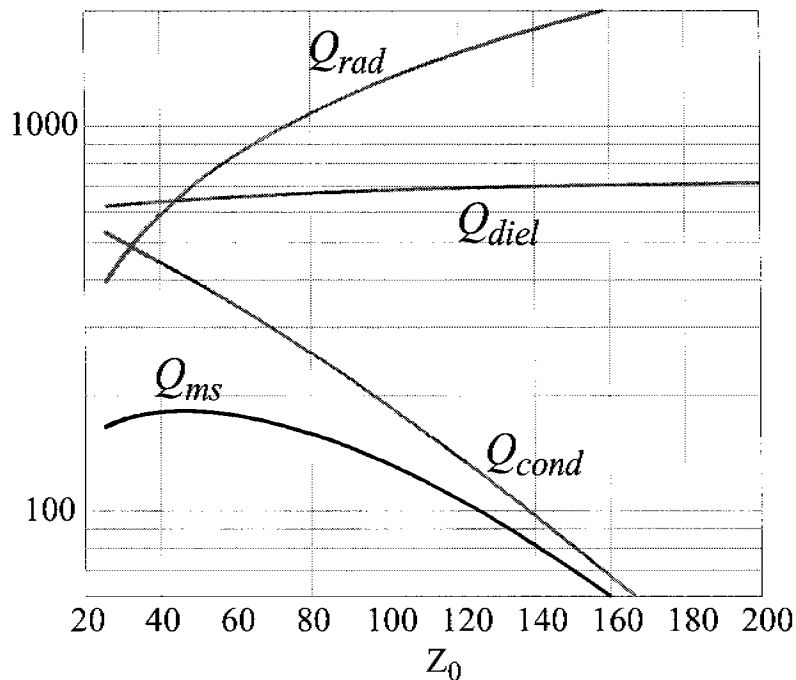
Figure 6.9: Unloaded quality factor a 0.35·λ_g length microstrip line at 4GHz versus characteristic impedance. The microstrip line is implemented using a 0.8mm thick RO4003 substrate material.

characteristic impedance represents a balance between a wide track width to improve the conductor quality Q_{cond} and a narrow track width to improve the radiation quality Q_{rad}. The optimum characteristic impedance of the considered configuration is found at around 50Ω. Notice that the optimum width lowers at higher frequencies.

The microstrip unloaded quality factor Q_{ms} is determined by the sum of all losses and is hence given as.

$$1/Q_{ms} = 1/Q_{cond} + 1/Q_{diel} + 1/Q_{rad} \qquad (6.28)$$

It reaches a maximum of slightly less than 200, as can be seen in Fig. 6.9 as well. The presented analysis demonstrates that a high quality factor can be realized with optimum dimensions and low-cost substrate materials. The achievable Q is definitely beyond any spiral inductor Q. On top of this, the microstrip unloaded Q is also much larger than the unloaded varactor Q. Hence, the Q of the varactor tuned resonator is roughly three times the unloaded varactor Q.

6.4.3. Radial Stub Lines

The rectangular shaped microstrip line discussed so far is well qualified to provide an inductive impedance at the input port. This structure bears, however, some disadvantages. First, the locus of the input is not well defined due the 1-2mm wide conductor track width. A more accurately localized input port is highly desirable. The second and eventually more critical drawback lies in the low tuning gain of the varactor tuned microstrip resonator, which is below the tuning gain of an equivalent LC-tank resonator. An increase of tuning gain commands either a varactor with a larger C_{max}/C_{min} ratio, or a flatter reactance versus frequency characteristic of the inductive resonator part. The reactance versus frequency curve of the rectangular line depicted in Fig. 6.6 exhibits a decreasing slope towards $\lambda_g/4$ length. A decent tuning range could be expected, but the large varactor capacitance required to resonate with a transmission line close to $\lambda_g/4$ length precludes this approach. Modification of the line shape is the only way out to raise the resonator tuning range if no improved varactor is available.

A radial shape shown in Fig. 6.10 overcomes both handicaps of the rectangular geometry. The input port can be as small as desired by keeping the inner radius R_1 of the device small. On top of this, the input reactance shows a flatter characteristic. This characteristic is used in microwave circuit design where the structure is used to provide a broadband point of low impedance [11]. Another interesting property in the context of tunable resonator design lies in the fact, that the reactance versus frequency slope can be controlled by the device's opening angle α. This is reflected by a formulation of the reactance at the inner radius proposed by Vinding [12]

$$X \propto \frac{h}{2\pi \cdot R_1} \cdot Z_0 \cdot \frac{360}{\alpha} \tag{6.29}$$

indicating that the resonant frequency tuning gain can be raised by a wide opening angle. Vinding's formula, however, was found to be inaccurate for the determination of the device radii. Atwater [11] proposed to model the device as a cascaded chain of rectangular shaped microstrip transmission line segments with varying width. It is claimed that the input impedance can be accurately predicted by division of the device into a moderate (20-30) number of transversal segments.
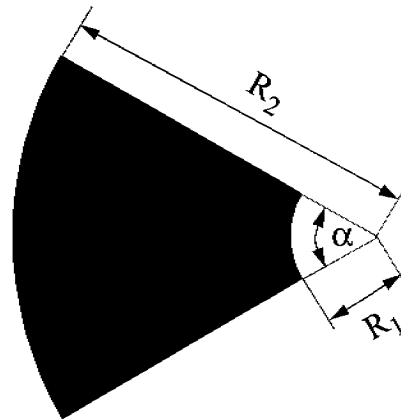
Figure 6.10: Geometry of the radial line characterized by the inner radius R_1, the outer radius R_2 and the opening angle α.

$R_1 = 0.2$mm

$R_2 = 15.4$mm

$\alpha = 60°$

$l = 16.4$mm

$w = 1.8$mm
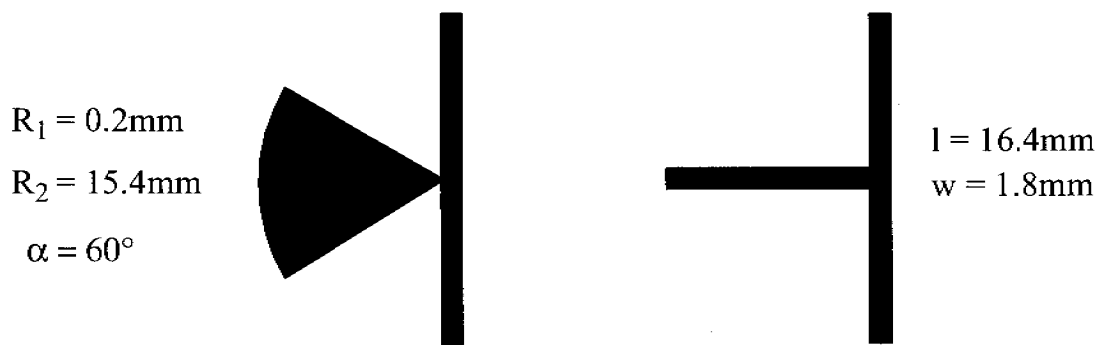
Figure 6.11: Layout of the microstrip experimental boards. The test structure on the left shows the radial line test setup, the structure on the right a test structure of an open-circuit terminated rectangular microstrip line.

The interesting properties of the radial line urge the replacement of the rectangular line. Determination of the radii is however not as straightforward as the determination of the rectangular line length. The lack of an accurate design formula requires the involvement of simulators. The outer radius R_2 is determined by the proposed model consisting of cascaded microstrip lines. The accuracy is further refined by a planar EM field simulation [13]. The simulations are finally validated experimentally. Two test PCBs containing a rectangular and a radial microstrip have been setup on a low cost RO4003 substrate material [14]. Fig. 6.11 shows the layout of the two microstrip devices and lists their dimensions. The devices are characterized by a two-port measurement carried out by a net-
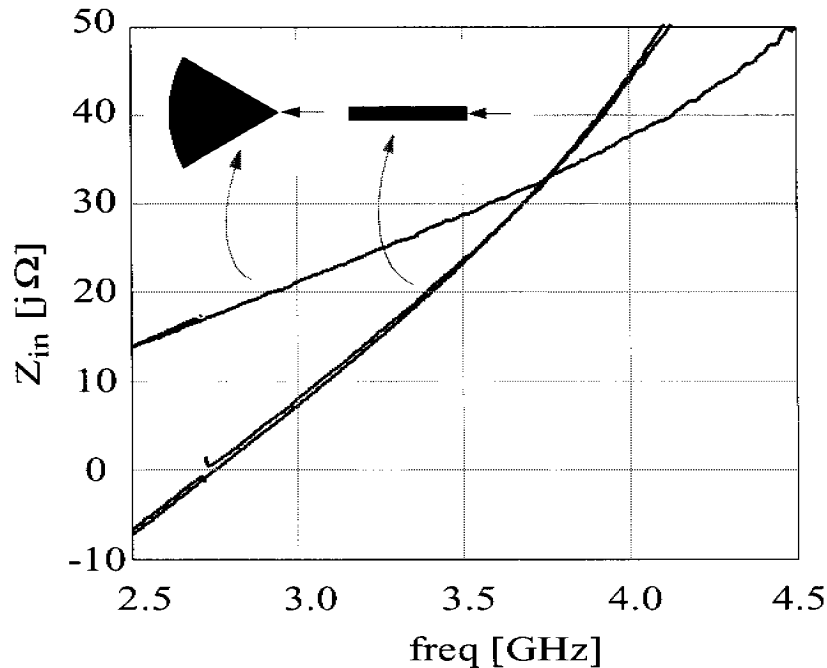
Figure 6.12: Measured input reactance of the radial line and the rectangular microstrip line.

work analyzer connected to the two feeder microstrip lines. The input reactance is obtained after deembedding of the feeder lines. This procedure leads to the input reactance shown in Fig. 6.12. The measured input reactances match well with the simulated values. As expected, the slope of the radial line is noticeably smaller than that of conventional microstrip line. The two microstrip devices form part of a set of 4GHz oscillators which are thoroughly discussed in Section 6.5.

6.5. A 4GHz Oscillator with 2GHz Quadrature Outputs

Depending on the shape, the striplines of a 4GHz oscillator require 2 to 6cm^2 PCB area, which corresponds to about 5 to 15% of the total PCB area of a modern, small-sized cellular handset [15]. Despite this large area, which is about one order of magnitude larger than the required size of the RF front-end integrated circuit, the transmission line resonator does not necessarily increase the handset's dimension. About half of the overall PCB area of such a handset is used for digital components with a high density of PCB interconnections, making a multilayer PCB indispensable [15]. The remaining half contains the radio section with bulky
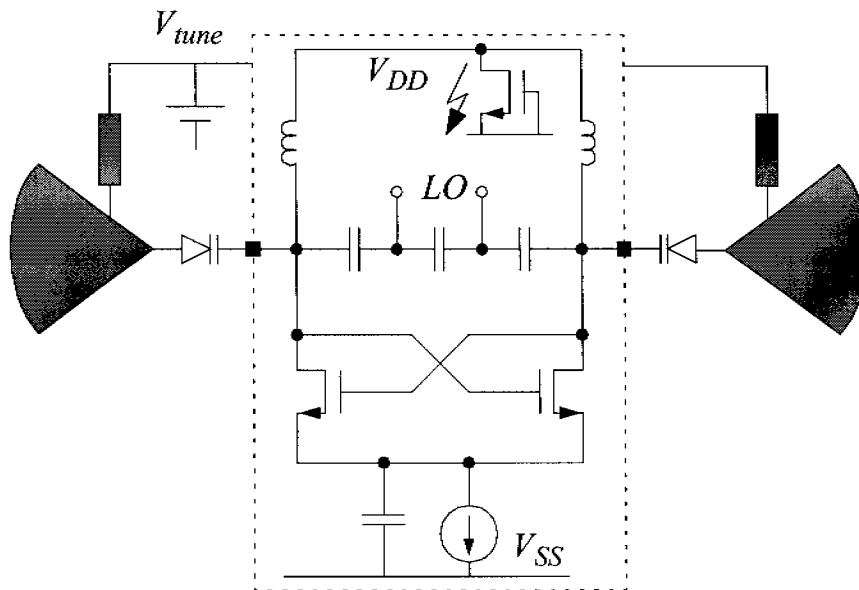
Figure 6.13: Transistor level schematic of the 3.6GHz stripline oscillator.

components and only a few interconnects. The idle metal layers of the radio section can be put to good use to realize transmission line circuits. Neither the PCB size nor its cost is so affected by the resonator. This requires however integration of the transmission lines into one of the inner layers of the PCB so that the surface is free for other RF front-end components.

An 1800MHz quadrature demodulator, consisting of a 3.6GHz stripline oscillator, a quadrature divider and two downconversion mixers, has been realized to demonstrate the aptness of stripline resonators. Low phase-noise of the oscillator and accuracy of the LO quadrature phases, required to support the rejection of unwanted sidebands, receive particular attention besides ultimate low power consumption. The discussion is mainly focussed on the oscillator and divider. The employed mixers have been presented in Chapter 5 and are therefore precluded from the discussion.

6.5.1. Voltage Controlled Oscillator

The 50% duty cycle requirement of the oscillator output commands a differential oscillator of which the transistor level schematic is shown in Fig. 6.13. The heart of the oscillator, the resonator, consists of radial stripline devices with input connected discrete varactors besides on-chip metal-insulator-metal capacitors.
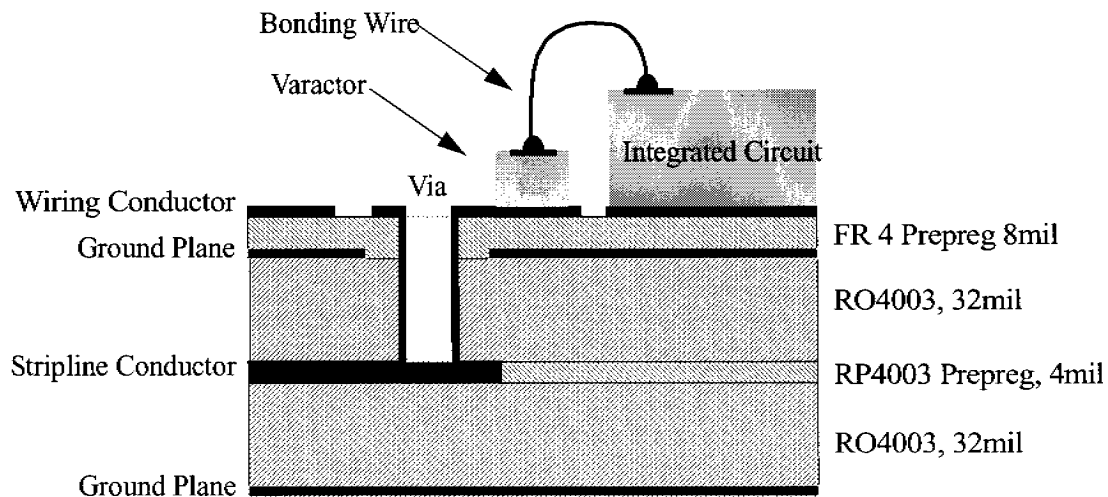
Figure 6.14: Cross-section of the four layer PCB.

Another oscillator employing rectangular microstrip lines has been realized as well [17]. This design is not further described here. The used varactor provides a capacitance ranging from 0.85pF to 1.5pF. They are accompanied by effective series resistances of 1.2Ω each [16]. Four 1.2pF on-chip capacitors close the resonator loop. The capacitors form a voltage divider which provides the appropriate amplitude to the quadrature divider. Two bonding wires connect the varactor cathodes with the internal resonator part. The radial transmission lines are based on the measured structure presented in Section 6.4.3. The geometry is slightly modified to maintain the input reactance in a stripline setup with two ground planes. Fig. 6.14 sketches the cross-section of the four layer PCB. It is composed of two compounded RO4003 cores, used to setup the striplines, and an additional layer of FR4 dielectric material to provide a fourth conductor layer, used for low frequency interconnects. A wide dead-end via provides the stripline input to the PCB surface. The bare die varactors are directly glued onto the PCB. A bonding wire, contributing estimated parasitic 0.7nH and 0.2Ω, connects the varactor cathode with the integrated circuit.

The varactors must be biased by a low-ohmic path to avoid any additional FM interference which may degrade the phase noise performance. Such a low-ohmic path is provided on the anode side by thin transmission lines which are attached to the radial stripline. Optimum locus of the attachment point was found by EM simulations. Low-ohmic bias on the cathode side is provided by two on-chip RF chokes, realized by 9nH planar spiral inductors and 50Ω series resistors. The lat-
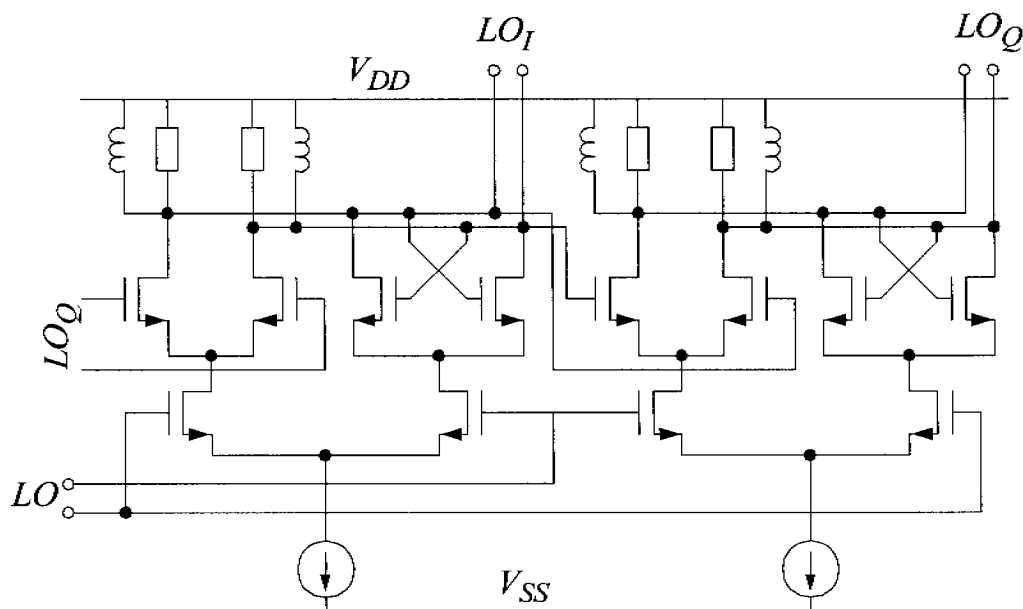
Figure 6.15: Transistor level schematic of the quadrature divider flip-flop with tuned output loads

ter are required to suppress the unwanted resonance formed by the RF chokes and the internal capacitors. On top of this, the RF chokes decouple the ESD clamps from the oscillator pads which would otherwise add too much capacitance to the resonator.

The total differential resonator loss, estimated as 2.5mS, is compensated by a cross coupled pair consisting of two 120μm/0.18μm sized NMOS transistors. Biased with 3mA, these transistors provide enough transconductance to ensure reliable start-up and a large amplitude.

6.5.2. Quadrature Divider

The purpose of the quadrature divider, implemented as a current mode logic flip-flop, is twofold. It generates primary quadrature signals, signals which are tapped at the flip-flop's master and slave latch output as shown in Fig. 6.15. On top of this, the divider isolates the large mixer LO port capacitance from the oscillator. While the divider must drive a total of 1pF, it stresses the oscillator by only one fifth compared to its own load. Besides the moderate load, the divider can be driven by a smaller amplitude compared to the mixer LO port amplitude.
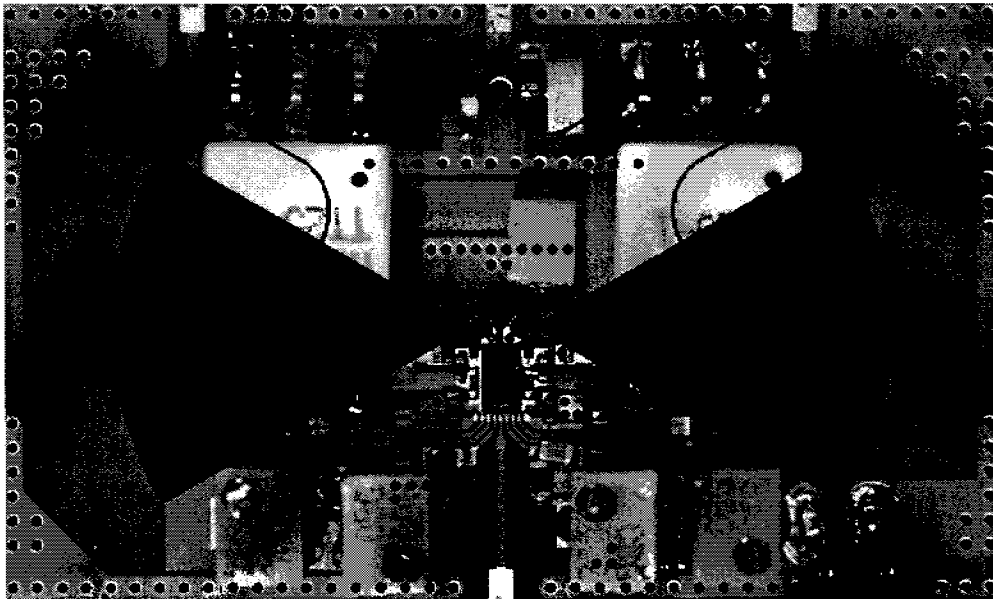
Figure 6.16: Quadrature demodulator multilayer test board with cross fading stripline devices.

It is hence fair to state that the quadrature divider lowers the oscillator load by one order of magnitude.

The load of 250fF at each divider output requires large sized switching transistors biased in the mA region. The consumption of the divider is lowered by the introduction of on-chip spiral inductors at the divider outputs to partially tune out the load capacitances. The quality factor of the tuned load is limited by 580Ω polysilicon resistors in parallel to the 17nH spiral inductors. The resulting Q of 2 guarantees broadband operation and robustness against process tolerances. The spiral inductors allow reduction of the total bias current to 3.8mA and moderate switching transistor size of 30μm/0.18μm. The realized output swing is approximately 400mV.

6.5.3. Experimental Results

The 1.8GHz quadrature demodulator with a 3.6GHz oscillator has been implemented in a 0.18μm CMOS technology. The striplines are embedded in a four layer PCB with the previously specified cross-section. Fig. 6.16 shows a photograph of the experimental board. The photograph displays besides the integrated circuit and some passive components a cross fade of the stripline layout which of

course is invisible in reality. As postulated before, the area above the stripline circuits can be reused for the placement of passive components.

Fig. 6.17 shows a block diagram of the implemented demodulator. No power wasting buffers are required between the building blocks. The downconversion mixers are identical to the circuit presented in Chapter 5 except down-scaling of all currents and transistor sizes by factor 2. Consequently, their input impedance doubles to 100Ω. The two mixer RF inputs are connected together to form a single 50Ω demodulator RF input. The layout of the mixers is symmetrically arranged to match the gain of the in- and quadrature path. Each mixer consumes 1.6mA.
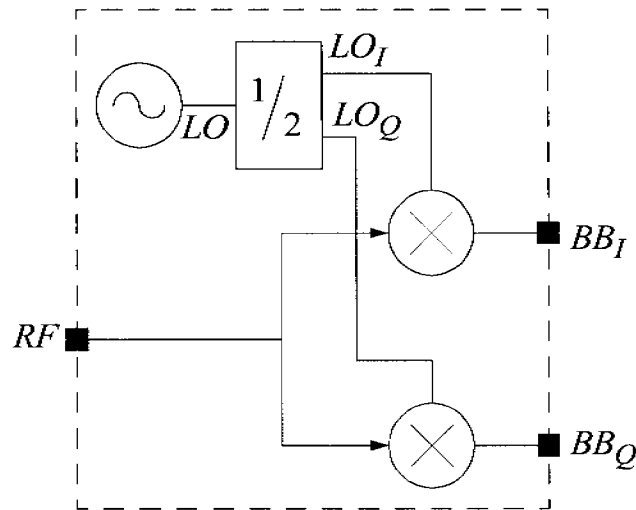
Figure 6.17: Block diagram of the quadrature demodulator.

The voltage controlled oscillator can be tuned from 3600MHz to 3950MHz. A plot of the tuning characteristic can be seen in Fig. 6.18. Phase noise performance is measured indirectly after downconversion to 20MHz using the downconversion mixers. Measurement plots of spectrum and phase noise, obtained from a spectrum analyzer and a HP E5500 phase noise measurement system running the delay line measurement method, are shown in Fig. 6.19 and Fig. 6.20 respectively. Phase noise at 100kHz offset from the carrier is measured as -110dBc/Hz. This number includes oscillator, divider and mixer noise. Phase noise of the involved RF signal source (Marconi 2042 in low noise mode) was removed numerically.
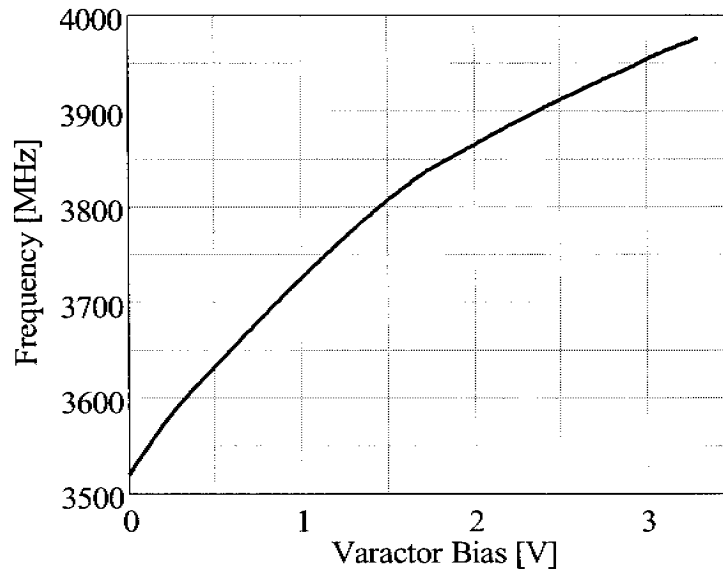
Figure 6.18: Frequency tuning characteristic of the voltage controlled stripline oscillator.
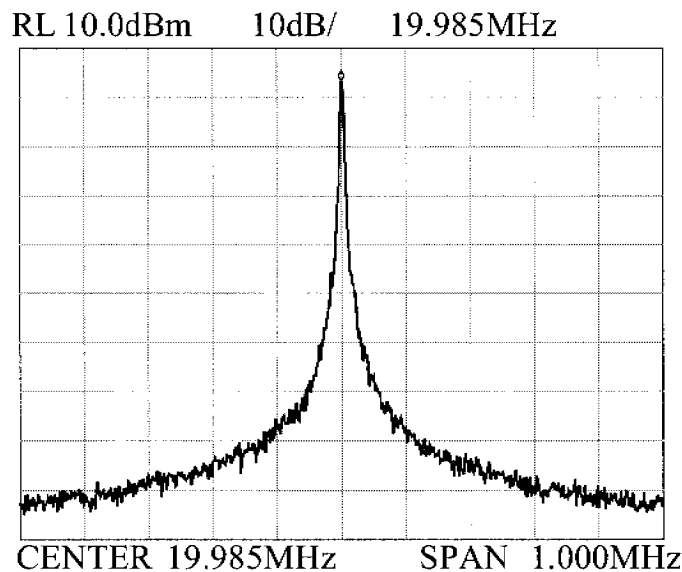
Figure 6.19: Quadrature oscillator spectrum measured after downconversion to 20MHz.

The ability of the demodulator to reject unwanted sidebands is mainly determined by the accuracy of the quadrature phases as generated by the divider and to a lesser extent by gain matching of the downconversion mixers. Suppression of the unwanted sideband, required to measure 30-35dB to preserve the phase information in cellular receivers [18], was evaluated by downconversion of a
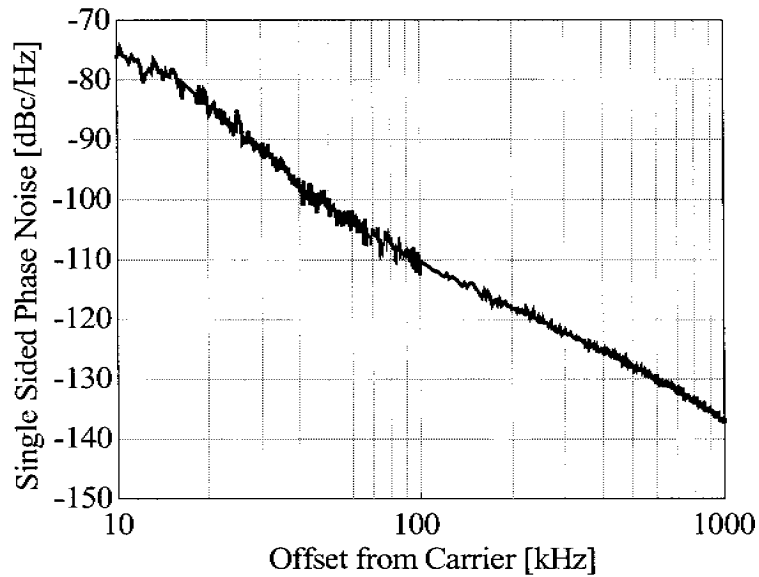
Figure 6.20: Single sided phase noise measurement taken at base band.
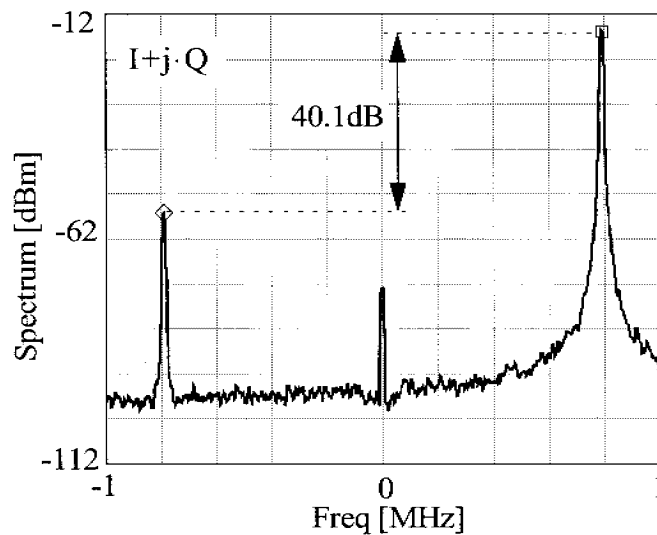
Figure 6.21: Complex spectrum measured at the mixer outputs exposes the unwanted sideband rejection.

2GHz sinusoidal to 1MHz. Fig. 6.21 shows the complex I+jQ spectrum as obtained from a vector signal analyzer with the input channels connected to the outputs of the two mixers. The demodulator achieves excellent 40dB of unwanted sideband rejection.

Measured phase noise and sideband rejection demonstrate the demodulator's suitability for 1800MHz cellular handsets. The good performance is accompa-

nied by an ultimate low power consumption which is a consequence of a rigorous low power design approach. The demodulator consumes a total of 10mA at 1.8V supply voltage. Fig. 6.22 shows a die micrograph of the 1.5mm×2.3mm measuring integrated circuit. Table 6.1 summarizes the demodulator performance.
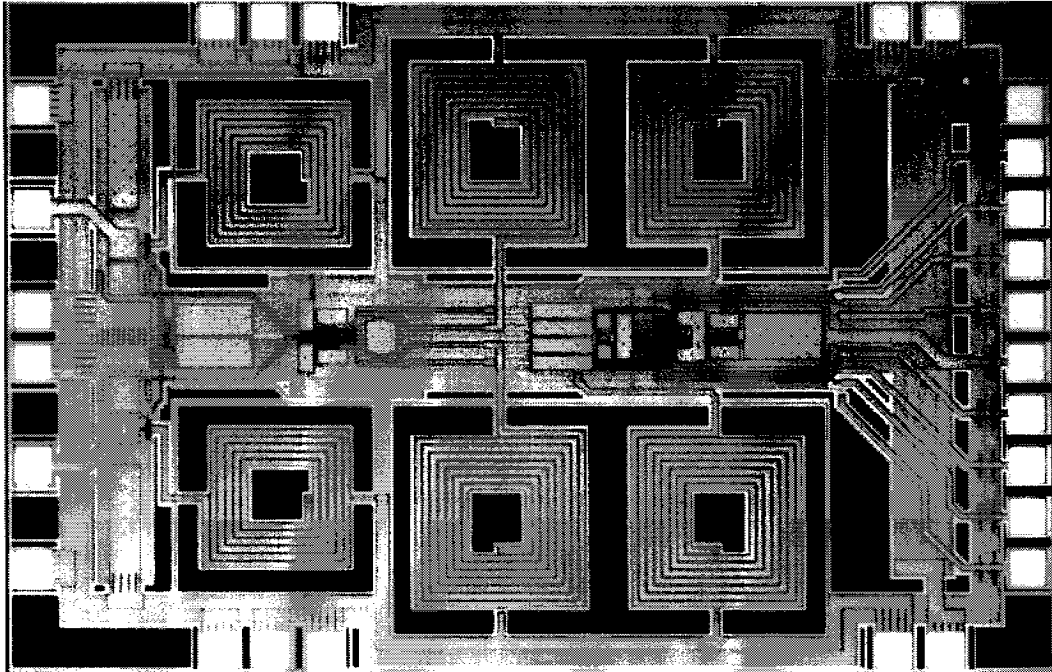
Figure 6.22: Chip micrograph of the 0.18μm CMOS quadrature demodulator.

Chip	Technology	0.18μm CMOS
	Supply voltage/current	1.8V/10mA
Striplines	Substrate	RO4003
	Substrate thickness	0.8mm
Oscillator	Tuning range	3600MHz-3950MHz
	Phase Noise	-110dBc/Hz@100kHz
	Consumption	3mA
Divider	Sideband rejection	40dB
	Consumption	3.8mA
Mixers	Consumption	3.2mA

Table 6.1: Performance summary of the 1800MHz quadrature demodulator.

References

[1] Q. Huang, *"Power Consumption vs. LO Amplitude for CMOS Colpitts Oscillators"*, Proc. IEEE Custom Integrated Circuit Conference, 1997, pp. 255-258

[2] J. Craninckx, M. Steyaert, H. Miyakawa, *"A Fully Integrated Spiral-LC CMOS VCO Set with Prescaler for GSM and DCS-1800 Systems"*, Proc. IEEE Custom Integrated Circuit Conference, 1997, pp. 403-406

[3] Q. Huang, *"Phase Noise to Carrier Ratio in LC Oscillators"*, Trans. on Circuits and Systems I, Vol. 47, No. 7, July 2000, pp. 965-980

[4] B. Park, P. Allen, *"A 1GHz, Low -Phase-Noise CMOS Frequency Synthesizer with Integrated LC VCO For Wireless Communications"*, Proc. IEEE Custom Integrated Circuits Conference 1998, pp. 567-570

[5] C. Patrick, S. Wong, *"On-Chip Spiral Inductors with Patterned Ground Shields for Si-Based RF IC's"*, IEEE J. Solid State Circuits, Vol. 33, No. 5, May 1998

[6] J. Tham, *"A 2.7V 900MHz/1.9GHz Dual-Band Transceiver IC for Digital Wireless Communication"*, Proc. IEEE Custom Integrated Circuits Conference 1998, pp. 559-562

[7] P. Orsatti, F. Piazza, Q. Huang, *"A 20mA-Receive, 55-mA-Transmit, Single-Chip GSM Transceiver in 0.25μm CMOS"*, IEEE J. Solid State Circuits, Vol. 34, December 1999, pp. 1869-1880

[8] A. Rofougaran, J. Rael, M. Rofougaran, A. Abidi, *"A 900MHz CMOS LC-Oscillator with Quadrature Outputs"*, Digest of Technical Papers ISSCC, 1996, pp. 392-393

[9] A. Gopinath, *"Maximum Q-Factor of Microstrip Resonators"*, IEEE Trans. Microwave Theory and Techniques, Vol. 29, No. 2, Feb. 1981, pp. 128-131

[10] L. Couch II, *"Digital and Analog Communication Systems"*, Prentice Hall, 1997, pp. 242

[11] H. Atwater, *"The Design of the Radial Line Stub: A Useful Microstrip Circuit Element"*, Microwave Journal, Dec. 1985, pp. 149-156

[12] H. Atwater, *"Microstrip Reactive Circuit Elements"*, IEEE Trans. Microwave Theory and Techniques, Vol. 31, No. 6, June 1983, pp. 488-491

[13] Agilent, *"Agilent EEsoft Momentum Product Overview"*, Agilent, Nov. 2000

[14] Rogers Corporation, *"Datasheet RO4000 Series High Frequency Circuits Materials"*, Rogers microwave products, 1998

[15] C. Lassen, M. Christensen, *"The Via Squeeze"*, IEEE Spectrum, October 1999, pp. 36-41

[16] Metelics, MSV34,067-C12 Abrupt Junction Varactor Diode

[17] D. Pfaff, J. Rogin, Q. Huang, *"A 14mA 0.25μm CMOS Quadrature Demodulator Including a Low Noise Local Oscillator"*, Proc. IEEE European Solid-State Circuits Conference, 2001, pp. 56-59

[18] P. Orsatti, *"A Low Power CMOS GSM Transceiver for Small Mobile Stations"*, Hartung-Gorre, 2000, Series in Microelectronics, Vol. 92, pp. 37

[19] G. Gonzalez, *"Microwave Transistor Amplifiers"*, Prentice-Hall, 1984, pp. 14

[20] T. Edwards, *"Foundations for Microstrip Circuit Design"*, John Wiley, 1992,

[21] L. Lewin, *"Radiation from discontinuities in striplines"*, Proc. IEE, Vol. 107, 1960

[22] J. James, A. Henderson, *"High-frequency behaviour in microstrip open-circuit termination"*, IEE Journal on Microwaves, Optics and Acoustics, Vol. 3, No. 5, 1979

[23] E. Hammerstad, O. Jensen, *"Equations for Microstrip Circuit Design"*, Proc. European Microwave Conference, 1975, pp. 268-272

Appendix 6

6.A Energy and Losses in Open Ended Transmission Lines

An electromagnetic field is formed between and around the conductors of a transmission line, once its input is connected to a source. Consequently, energy is stored in the transmission line. The electromagnetic field distribution depends on transmission line geometry as well as on the termination impedances. The longitudinal field components of many transmission line geometries are much smaller than the transversal components, leading to the concept of quasi TEM fields.

TEM fields of many technically relevant transmission lines have been worked out by solving Maxwell's Equations or by semi-analytical approaches. Considering these results and knowing further that any transmission line geometry can be modelled by an infinite number of lumped elements, energies and losses can be carried out by passive circuit rather than by full TEM wave analysis.

A lossless transmission line can be characterized by the series connection of an infinite number of identical line sections, each consisting of a two-port device [19]. Each two-port consists of two lumped components L' and C' called the primary line constants. Energy stored in the transmission line is obtained by summation of the energies stored in all two-port devices. The model can be enhanced to accommodate lossy transmission lines by adding loss to the reactive devices, leading to another two primary line constants R' and G'. Energy loss is obtained analogously by summation of the energy dissipated in all two-port devices. Fig. 6.23 shows the schematic of a transmission line section.
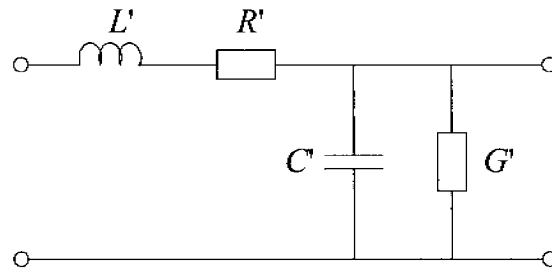
Figure 6.23: Lossy transmission line section

Transmission line energy and losses are derived now for open-circuit terminated transmission lines. The results are functions of the primary line constants or measures which are directly linked to the primary line constants.

Stored Energy:

It is assumed that the input is connected to a harmonic voltage source $V_0 \cdot \cos(\omega t)$. Voltage and current amplitudes vary along a the transmission line and can be found as [19]

$$v(x, t) = V_0 \cdot \frac{\cos(\beta \cdot x)}{\cos\theta} \cdot \cos\omega t \qquad (6.A.1)$$

and

$$i(x,t) = -\frac{V_0}{Z_0} \cdot \frac{\sin\beta x}{\cos\theta} \cdot \sin\omega t \qquad (6.A.2)$$

with $Z_0 = \sqrt{L'/C'}$ the characteristic impedance, β the phase propagation constant and x the physical distance measured from the open end. The electrical length θ of the transmission line corresponds to the phase shift a harmonic signal experiences if it travels from the source to the end of the transmission line. The latter relates to the physical length l and to the wavelength λ_g as:

$$\theta = \beta l = 2\pi \cdot \frac{l}{\lambda_g} \qquad (6.A.3)$$

The knowledge of the voltages and currents along the transmission line allows computation of the energy stored in a line section residing at position x. The energy in a section of infinitesimal length can be found as:

$$E_{ms}'(x,t) = \frac{V_0^2}{\cos^2(\theta)} \cdot \frac{C'}{4} \cdot \left[1 + \cos(2\omega t) \cdot (\cos^2\beta x - \sin^2\beta x) \right] \qquad (6.A.4)$$

Total microstrip energy is found by integration of the section energies. This leads to

$$E_{ms}(t) = \frac{V_0^2}{8\pi \cdot f \cdot Z_0} \cdot \left[\frac{2\theta + \sin(2\theta) \cdot \cos(2\omega t)}{2\cos^2(\theta)} \right] \qquad (6.A.5)$$

and to the peak energy stored in the transmission line:

$$E_{ms,pk} = \frac{V_0^2}{8\pi \cdot f \cdot Z_0} \cdot \left[\frac{2\theta - \sin(2\theta)}{2\cos^2(\theta)} \right] \qquad (6.A.6)$$

The last Equation shows that the maximum stored energy is as a strong function of the electrical length. Eq. 6.A.6 evaluated at $\theta = \pi$ leads to the well known peak energy of a $\lambda_g/2$-resonator [9]:

$$E_{ms,pk,\lambda/2} = \frac{V_0^2}{8 \cdot f \cdot Z_0} \qquad (6.A.7)$$

The remaining part of this Appendix considers energy losses. It is assumed that the energy loss is small compared to the stored energy. Consequently, voltages and currents along the transmission can be approximated by the voltages and currents of the lossless transmission line. This assumption simplifies the mathematics significantly without a relevant deterioration of accuracy.

Conductor and Dielectric Loss:

The average power dissipated in a line section resident at position x can be approximated as

$$P'_{cond} = \frac{V_0^2}{2Z_0^2} \cdot \left(\frac{\sin\beta x}{\cos\theta} \right)^2 \cdot R' \qquad (6.A.8)$$

and

$$P'_{diel} = \frac{V_0^2}{2} \cdot \left(\frac{\cos\beta x}{\cos\theta} \right)^2 \cdot G' \qquad (6.A.9)$$

The dissipated energy per cycle can be found by integration along the transmission line. This leads to conductor and dielectric energy loss as

$$E_{loss, cond} = \frac{V_0^2 \alpha_c}{2 Z_0 \beta f} \cdot \left(\frac{2\theta - \sin 2\theta}{2\cos^2\theta} \right) \qquad (6.A.10)$$

and

$$E_{loss, diel} = \frac{V_0^2 \alpha_d}{2 Z_0 \beta f} \cdot \left(\frac{\sin 2\theta + 2\theta}{2\cos^2\theta} \right) \qquad (6.A.11)$$

Notice that the primary line constants R' and G' are replaced in Eq. 6.A.10 and Eq. 6.A.11 by the more meaningful conductor loss constant α_c and dielectric loss constant a_d. The latter constants are linked by the following identities:

$$\alpha_c = \frac{R'}{2Z_0}$$

$$\alpha_d = \frac{G'Z_0}{2} \qquad (6.A.12)$$

Radiation Loss:

Radiation loss of planar transmission lines is modelled by conductors G_R terminating the input and output of the transmission line. As before, it is assumed that radiation loss is small compared to the stored energy. Then, energy lost by radiation can be formulated as:

$$E_{loss,rad} = \frac{V_0^2}{2f} \cdot \left(1 + \frac{1}{\cos^2(\theta)}\right) \cdot G_R \qquad (6.A.13)$$

6.B Microstrip Design Equations

The broad range of microstrip applications in microwave circuit design and digital interconnect urged the development of design supporting Equations in the past. Characteristic impedance, losses and line discontinuities, e.g. open ends, are well expressed by semi-analytical formulas. Those relevant to the presented microstrip resonator are listed now.

Character Impedance

A closed-form characteristic impedance formulation was presented by Hammerstad [23]. The characteristic impedance is reported as

$$Z_0 = \begin{cases} \dfrac{60}{\sqrt{\varepsilon_{eff}}} \ln\left(\dfrac{8h}{w} + \dfrac{w}{4h}\right) & \text{if } \dfrac{w}{h} \le 1 \\[4mm] \dfrac{120\pi}{\sqrt{\varepsilon_{eff}}} \left\{\dfrac{w}{h} + 1.393 + 0.667\ln\left(\dfrac{w}{h} + 1.444\right)\right\}^{-1} & \text{if } \dfrac{w}{h} > 1 \end{cases} \qquad (6.B.1)$$

with w the conductor width and h the thickness of the substrate material. The effective dielectric constant ε_{eff} used in Eq. 6.B.1 is different from the relative dielectric constant ε_r of the used dielectric material since the electric field of the microstrip line penetrates only partially the dielectric material while the remaining field resides in the neighboring air above the microstrip line. The portion of the field captured by the dielectric material is expressed by the so called filling factor q, a microstrip geometry dependant factor between 0.5 and 1. The filling factor is given as:

$$q = \begin{cases} \dfrac{1 + (1 + 12h/w)^{-1/2} + 0.04(1 - w/h)^2}{2} & \text{if } \dfrac{w}{h} \leq 1 \\[4mm] \dfrac{1 + (1 + 12h/w)^{-1/2}}{2} & \text{if } \dfrac{w}{h} > 1 \end{cases} \qquad (6.B.2)$$

The effective dielectric constant is determined by:

$$\varepsilon_{eff} = q\varepsilon_r + (1 - q) \qquad (6.B.3)$$

Microstrip loss:

Microstrip losses cause an exponential signal drop along an impedance matched transmission line. The total loss constant α, expressed in Neper per meter, can be separated into a conductor loss constant α_c and a dielectric loss constant α_d respectively. Approximation for the loss constants are reported as [20]:

$$\alpha_c = \frac{0.072}{8.686} \cdot \frac{\sqrt{F}}{w \cdot Z_0} \qquad (6.B.4)$$

$$\alpha_d = \frac{27.3}{8.686} \cdot \frac{\varepsilon_r}{\varepsilon_{eff}} \cdot q \cdot \frac{\tan\delta}{\lambda_g} \qquad (6.B.5)$$

F represents the frequency in GHz and $\tan\delta$ the loss tangent, a measure specified by the substrate manufacturer.

Open end radiation:

Radiation loss was investigated by Lewin [21] and James [22]. The radiation loss can be modelled by an equivalent radiation admittance G_R residing at the ends of a lossless transmission line. The radiation admittance derived by James is repeated here as:

$$G_R = 160 \cdot \left(\frac{\pi \cdot h}{Z_0 \cdot \lambda_g \cdot \sqrt{\varepsilon_{eff}}} \right)^2 \qquad (6.B.6)$$

Seite Leer /
Blank leaf

Chapter 7

Frequency Synthesizer Implementation

Integer-N frequency synthesizers are well qualified for GSM and WCDMA standards but, compared to their fractional-N counterparts, the loop bandwidth with respect to the reference frequency is much larger. Less filtering of charge pump current harmonics is obtained. In order to compete with fractional-N, integer-N synthesizers must compensate the lack of strong filtering by improved loop circuits which produce inherently less charge pump harmonics. This Chapter explores the low spurious tone circuit design aspects of integer-N synthesizers. The synthesizer building blocks responsible for the introduction of spurs are identified and requirements for these blocks are formulated. Different charge pump topologies are analyzed and their deficiencies named. A novel topology based on self-calibration technique is proposed to overcome the limitations of existing solutions.

Besides discussion of charge pump and PFD, a step-by-step PLL design procedure is presented. The procedure allows systematic determination of third order PLL design parameters. Nonlinear effects are taken into account by means of behavioral simulations.

Finally, an integer-N frequency synthesizer with 4GHz outputs targeted to GSM and WCDMA has been implemented using the proposed self-calibrating charge pump. The Chapter is finished by experimental results which validate the integrity of the proposed concepts and circuits.

7.1. Low Spurious Tone Circuit Design

Circuits directly wired to the RF oscillator tuning node are most critical to spurious tone performance. Eventualities causing spurs have been already named in Chapter 3 as charge pump current source mismatch, charge sharing,

clock-feed-though and leakage. While the last of these effects can be mainly eliminated by usage of high-Q loop filter capacitors, the remaining effects require application of sophisticated charge pump topologies.

The level of spurs does not depend on the charge pump performance alone. The PFD as well as the interface circuitry between PFD and charge pump, which in general is required for reasons given later, strongly affect the spurious tones as well. Remember that the spurs level reacts sensitively to the duration of the con-current activation of both charge pump current sources (SHORT state), a duration which is under control of the PFD. To keep the duration short without the intro-duction of phase detector crossover distortion, the interface circuitry driving the charge pump switches must to be designed strong to provide a large slew rate.

By recalling the outcome of the spurs analysis presented in Chapter 3, a switch-ing time (the time required to change the status of the charge pump switches from on to off and vice versa) needs to be as short as 1ns to allow reduction of the SHORT state duration to the order of a few nano-seconds. Besides strong switch drivers, the PFD must be able to provide fast transitions to keep the SHORT state duration in the required range.

Circuit design for low spurious tone level involves hence optimization of several synthesizer building blocks. Requirements of the involved blocks can be summa-rized as:

- *Charge Pump*:
  Current source mismatch, charge sharing and clock-feed-through must be tackled individually by appropriate circuit techniques. The charge pump rep-resents therefore the most challenging building block with regard to spurious tone performance.

- *Loop Filter*:
  The loop filter's passive nature reduces the criticalness of this building block. The sole nonideality of the filter affecting the spurious tone level is compo-nent leakage. Leakage can be lowered to an acceptable level by usage of high-Q capacitors. Thin film capacitors with insulation resistances of several thousands of MΩ reduce loop filter leakage to negligible values [1].

- *Phase-Frequency Detector*:
  This device must provide a fast transition from the SHORT to the OFF state. An appropriate architecture combined with careful transistor sizing must be

used. Proven and robust topologies have been proposed [2] [3] which satisfy this requirement.

- *Switch Driver*:

  The switch driver must be designed to change the state of the charge pump quickly. Notice that a larger switching time widens the dead zone and phase detector gain linearization requires an extended duration of the SHORT state. The latter extension aggravates charge pump current source matching. In order to avoid too demanding matching requirements, the switching time must be limited to 1ns or even below.

Different charge pump topologies have been proposed in the past [4]-[8], however, the achieved suppression of the reference frequency spurious tone in the region of -45dBc to -55dBc is often not satisfactory, especially if rather strong spurs occur despite a low loop bandwidth to reference frequency ratio. Recall that weak PFD/CP performance can be compensated by a reduction of loop bandwidth with the remedy of slowed dynamics.

Charge pump design becomes even more challenging if another design aspect is taken into account. The RF oscillator tuning range is limited by the available charge pump output voltage range, which is often scarified for better matching performance. The limited tuning gain of LC-tank oscillators combined with the low supply voltage of less than 2V if deep sub-micron CMOS technologies are used, commands rail-to-rail output charge pumps. A novel charge pump combining rail-to-rail operation with improved spurious tones performance is presented in the following Section.

7.1.1. A Self-Calibrating Rail-To-Rail Charge Pump

A circuit providing charge pump functionality can be set up by a current sink, a current source and two switches. The principal schematic of a basic charge pump is shown in Fig. 7.1. Positive and negative loop filter current injection is gained by selective activation of the two switches which are under control of the PFD. Although the circuit provides full functionality, its application is restricted to low performance synthesizers since strong spurious tones must be expected. Charge sharing and current source mismatch must be considered as the dominant sources of error current leaving incidentally the charge pump. The main handicap of the charge pump according to Fig. 7.1 lies in the fact that the drain node of the cur-
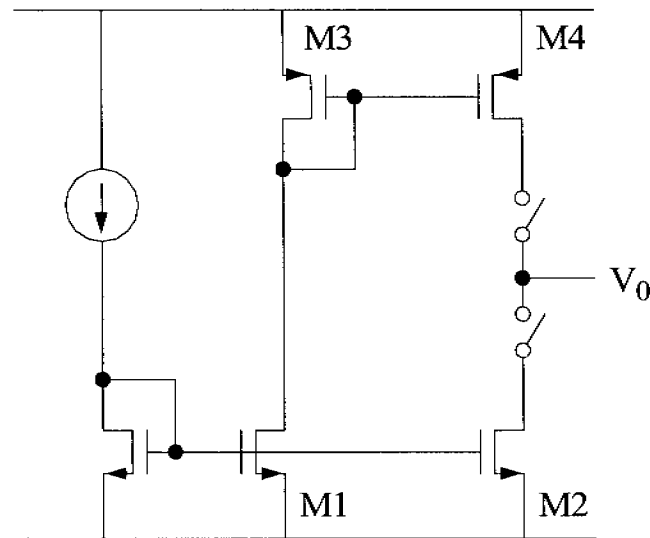
Figure 7.1: Elementary charge pump circuit.

rent sink M2 drops to the negative supply voltage and the drain node of the current source M4 raises to the positive supply voltage during OFF state. Capacitance present at these nodes is completely discharged at the instant of closing switches. Consequently, charge stored in the loop filter capacitors is transferred to the drain node capacitance once a switch is closed. The sharing of charge between the loop filter capacitors and the capacitance associated with the drain nodes continues until either the voltage drop across the switches is equalized or the drain to source voltage of M2 and M4 is raised above the saturation voltage so that further charging of the drain node capacitance is overtaken by the transistor drain current. Charging of the drain node capacitances by loop filter capacitors causes an undesired excursion of the tuning node with the strength of the excursion being dependant on the ratio of the drain node capacitance to the total loop filter capacitance and on the initial tuning voltage before switching has taken place. This effect is often called charge sharing in order to express that charge is shared among the loop filter capacitors and the drain node capacitances. The second main eventuality increasing the level of spurious tones, current source mismatch, is caused by copier errors of the current mirrors. Underlying effects responsible for the introduction of these errors are channel length modulation as well as mismatch of the current mirror transistors. Increase of the physical size of the current mirror transistors mitigates both effects. The use of transistors with channel length larger than the minimum length reduces the impact of chan-

nel length modulation as well as the impact of threshold voltage mismatch [9] since the increase of channel length is accompanied by a larger transistor overdrive voltage. An increase of transistor gate area minimizes further the impact of transconductance parameter mismatch [9]. Unfortunately, the larger transistors increase the capacitance at the drain nodes since this capacitance scales with transistor channel width. The reduction of spurious tone level gained by improved matching is so potentially neutralized by increased charge sharing. Besides this, the large drain node capacitance may alter the loop filter impedance and charge pump output range is reduced by the large transistor overdrive voltage.

The dilemma of the simple charge pump according to Fig. 7.1 - large transistors required to improve matching, but small transistors obligatory to reduce charge sharing - requires modification of the topology. It was proposed to swap the position of the transistors M2 and M4 with the position of the switches so that the switches are connected to the transistor source and the drains are directly wired to the loop filter [4] [5]. Of course, the exchange requires also a modification of the biasing circuitry. Less charge sharing can be expected since the critical nodes, which have moved from the transistor drain to the source, are not pushed to the supply rails any more when the current sink or source is disabled. However, the desired elimination of charge sharing is only partial. The source nodes still experience a voltage excursion which is identical to the overdrive voltage of M2 and M4, responsible for residual charge sharing. Besides this, matching of the charge pump sink current with the source current is more difficult due to the more complex biasing network. Reported spurious tones in the order of only -50dBc [4] must be considered as a result of the deficiencies related to this topology.

Charge sharing is prevented more efficiently by a current bypass path allowing the current to flow even when the current sink or source is switched off [6] [7]. Fig. 7.2 presents the improved charge pump solution. Two additional switches driven by complementary clock phases bypass the currents to a replica of the tuning node. A regulating amplifier makes sure that the voltage at the replica node is always identical to the voltage present at the loop filter. Hence, the capacitances at the drain nodes of M2 and M4 are precharged during the OFF state. Virtually no charge sharing occurs if one or both drain currents are steered to the loop filter since no voltage resides across the switches. Residual charge sharing, however, may happen due to the finite switching time required to commute the drain cur-
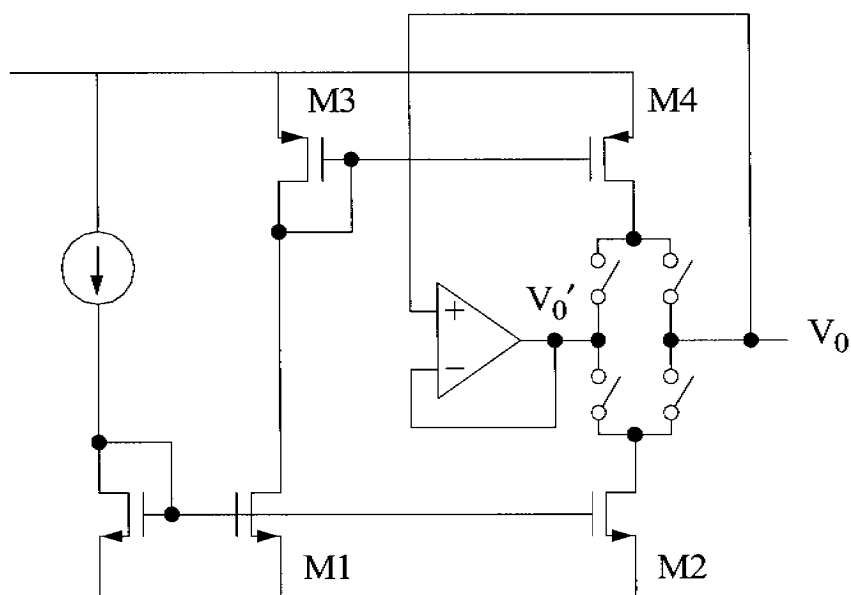
Figure 7.2: Extension of the elementary charge pump circuit by a current bypass path to avoid charge sharing.

rent from one switch to the adjacent one. Careful design of the clock phases helps to reduce the unwanted voltage excursion to an acceptable level.

The charge pump providing a regulated current bypass removes efficiently charge sharing, but suffers from mismatch to the same extend as the elementary charge pump. As discussed earlier, large sized current mirror transistors would improve matching, but reveal a reduced charge pump output voltage range as well as loop filter loading. Again, mismatch can be prevented by modification of the circuit topology. Self-calibration is a widely used technique to overcome circuit limitations caused by component mismatch. It is especially attractive to sampled data systems which often allow periodic calibration during normal operation of the circuit. A well known example is found in amplifier input offset cancellation as applied in switched capacitor circuits.

The concept of self-calibration can be adopted to a charge pump providing a current bypass. The current bypass can be of good use to perform self-calibration of the M4 drain current during the OFF state. Circuit prerequisite for self-calibration is a high impedance at the replica node to guarantee that the full current of the current sink transistor M2 is flowing into the current source transistor M4. Further, the gate voltage of M4 must be under control of a regulating loop rather than a static bias network. These considerations lead to a novel charge pump cir-
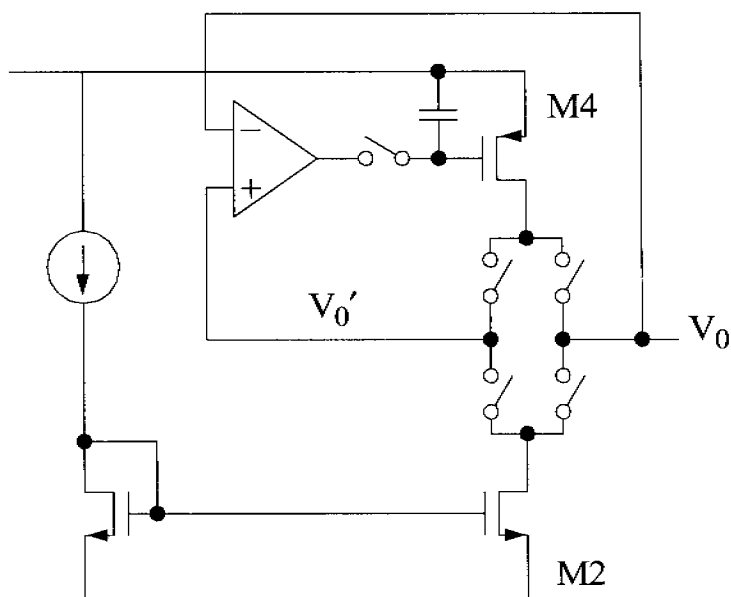
Figure 7.3: Self-calibrating charge pump circuit to avoid current source mismatch.

cuit shown in Fig. 7.3. Assuming for the moment that the additional switch at the output of the regulating amplifier is closed and limiting further the analysis of this new topology to the OFF state, i.e. the two switches at the left are closed, the drain current of M4 is a perfect replica of the M2 drain current. Besides this, the drain nodes remain precharged as in the topology of Fig. 7.2. It is quite amazing that a simple rearrangement of the topology according to Fig. 7.2 removes efficiently current source mismatch. Neglecting the additional switch, the self-calibrating charge pump requires even less devices compared to Fig. 7.2 since current mirror transistors M1 and M3 are not needed any more.

However, the control loop adjusting the drain current of M4 is opened if the charge pump leaves the OFF state. The additional switch at the regulating amplifier output combined with the capacitance at the gate of M4 'freezes' the gate voltage to keep the drain current constant. The additional circuitry may be discarded because the duration of the UP, DOWN and SHORT takes only a small fraction of the OFF state. Keeping the bandwidth of the control loop small helps to preserve the M4 drain current during a short breakage of the control loop even without additional switch. Nevertheless, the switch is drawn in Fig. 7.3 to reveal a principal schematic without any control loop bandwidth restrictions.

Current bypass technique combined with a self-calibrating current source represents a great step toward a low spurious tones charge pump. Extension to rail-to-rail output is the next logical step of the pursuit to a high performance charge pump topology. While a large loop filter voltage close to the positive rail does not affect the operation of the current source formed by M4 and the regulating amplifier, a low loop filter voltage close to the negative rail pushes M2 into the triode region causing a drop of charge pump current. The output voltage range is hence limited at the lower boundary while operation up to the positive rail is inherently provided.

The voltage range of the NMOS current mirror can be significantly extended by introduction of a second regulating amplifier. Fig. 7.4 shows the final schematic of the proposed high performance charge pump extended by another amplifier and switch. Assuming that the charge pump resides in the OFF state and assuming that the additional switch at the amplifier output is closed, the second control loop duplicates the drain voltage of M2 to the drain of the NMOS current mirror transistor. Identical gate to source voltage as well as identical drain to source voltage guarantees perfect replication of the reference current into M2 even with both transistors being forced to triode region. Charge pump current drop close to the negative supply is therefore avoided. But notice that the control loop raises the gate voltage of the NMOS transistors when pushed into triode region. The increased gate voltage affects the operation of the PMOS current mirror providing the reference current. Consequently, the ultimate operation limit is shifted from the NMOS to the PMOS current mirror. However, the overdrive voltage of the latter can be kept small without affecting the critical capacitance of the drain node of M2.

The additional switch at the output of the second amplifier is required to break the control loop during DOWN state. This is necessary to avoid oscillation of the control loop caused by the positive feedback formed by the connection of the M2 drain node with the positive input of the regulating amplifier. However, breakage of the loop can be neglected without any impact if the control loop bandwidth is kept small.

Remaining eventuality affecting spurious tone performance is clock-feed-through. Unwanted injection of charge stored in the MOS switch channels into the loop filter can be cancelled to some extent by introduction of dummy switches of half the size of the main switches [8].
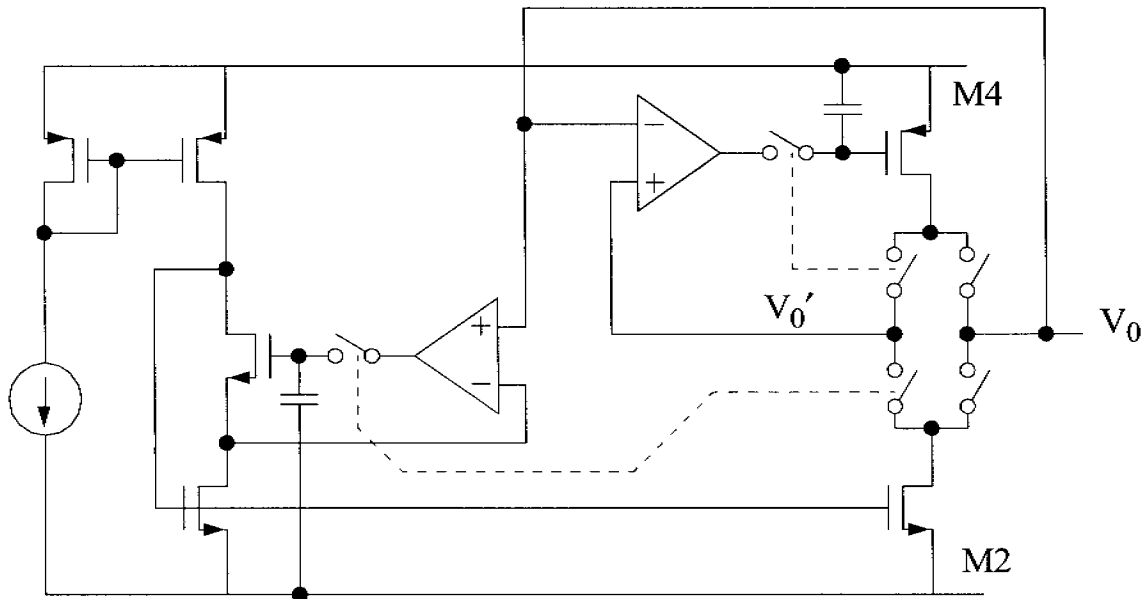
Figure 7.4: Improved self-calibrating charge pump circuit with rail-to-rail output range.

7.1.2. Phase Frequency Detector

A phase-frequency detector can be established by resettable flip-flops and additional combinational logic responsible for the generation of a reset pulse [10]. A conceptual schematic which serves as basis for a variety of PFD implementations is shown in Fig. 7.5. The operation of the circuit is quite apparent. Assuming that both flip-flops are initially reset, an active edge at one of the PFD clock inputs sets the corresponding flip-flop while the second flip-flop remains reset. Depending on which flip-flop was set, the PFD resides after clocking in the UP or DOWN state. While repetitive clock edges appearing at the same input do not alter the PFD, a clock edge at the adjacent input sets also the second flip-flop. The state with both flip-flops set, which corresponds to the SHORT state, is only momentary. After a short duration determined by the propagation delay of the AND gate, the SHORT state is decoded and a pulse appears at the reset input of both flip-flops. Consequently, the circuit recaptures the initial condition, which is the OFF state. The conceptual circuit enacts hence exactly the state transition diagram of the modified PFD introduced in Chapter 3 and shown in Fig. 3.11.

Although the conceptual schematic could be realized directly with standard cell flip-flops and gates, the PFD is in general optimized on gate level with the ulti-
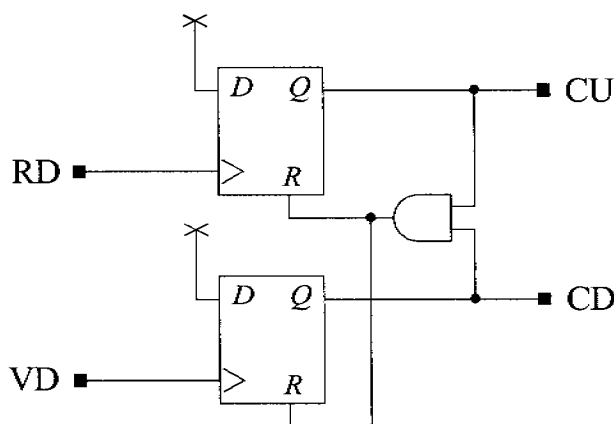
Figure 7.5: Conceptual phase-frequency detector schematic.

mate goal to identify solutions with reduced SHORT state duration. Room for optimization is given by the fact that the flip-flops must not provide the full functionality of standard flip-flops since their data inputs are fixed wired. A couple of different gate level PFD implementations have been proposed which are founded all on the conceptual PFD despite their differences. An early implementation dates 30 years back and was sold as a standard product IC [2]. The simplicity and proven robustness of the circuit guaranteed ongoing success, so that this PFD is still used three decades after its invention in modern VLSI frequency synthesizer ICs. Nevertheless, a weakness of this implementation lies in the required 4-input NAND gate responsible for decoding the reset condition. The gate level schematic of this particular implementation is shown in Fig. 7.6. Since multi-input gates are inherently slow due to the large number of stacked transistors, the duration of the SHORT state, which is determined by the propagation delay of the latter NAND gate, is quite long.

SHORT state duration as short as a few nano-seconds or less can be achieved more easily with an implementation employing a standard, two-input gate in the reset path. Such a circuit was proposed by Von Kaenel [3]. Fig. 7.7 shows the gate level schematic. A NOR gate based set-reset (RS) latch formed by gates G5 and G6 frames the heart of each flip-flop. The OFF state sees the RS-latch reset with the outputs CU and CD being low. Further, the output of gate G1, which will provide later the reset pulse, is low as well as the two PFD clock inputs. Notice that nodes A1 and A2 must be high under these conditions. Any change at the clock inputs can propagate therefore through the AND gate G4 and appears at the
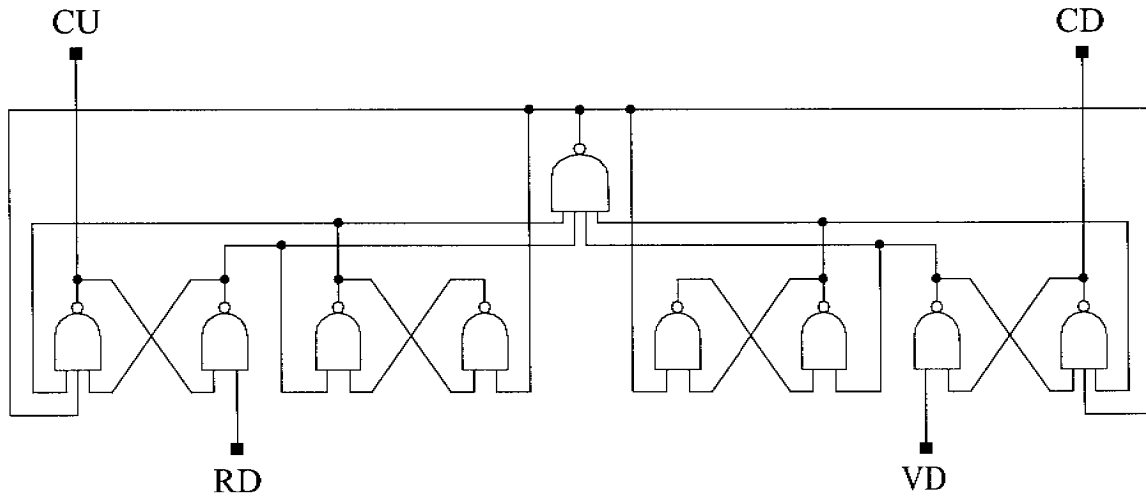
Figure 7.6: Gate level schematic of the PFD introduced in the MC4044 standard product IC.
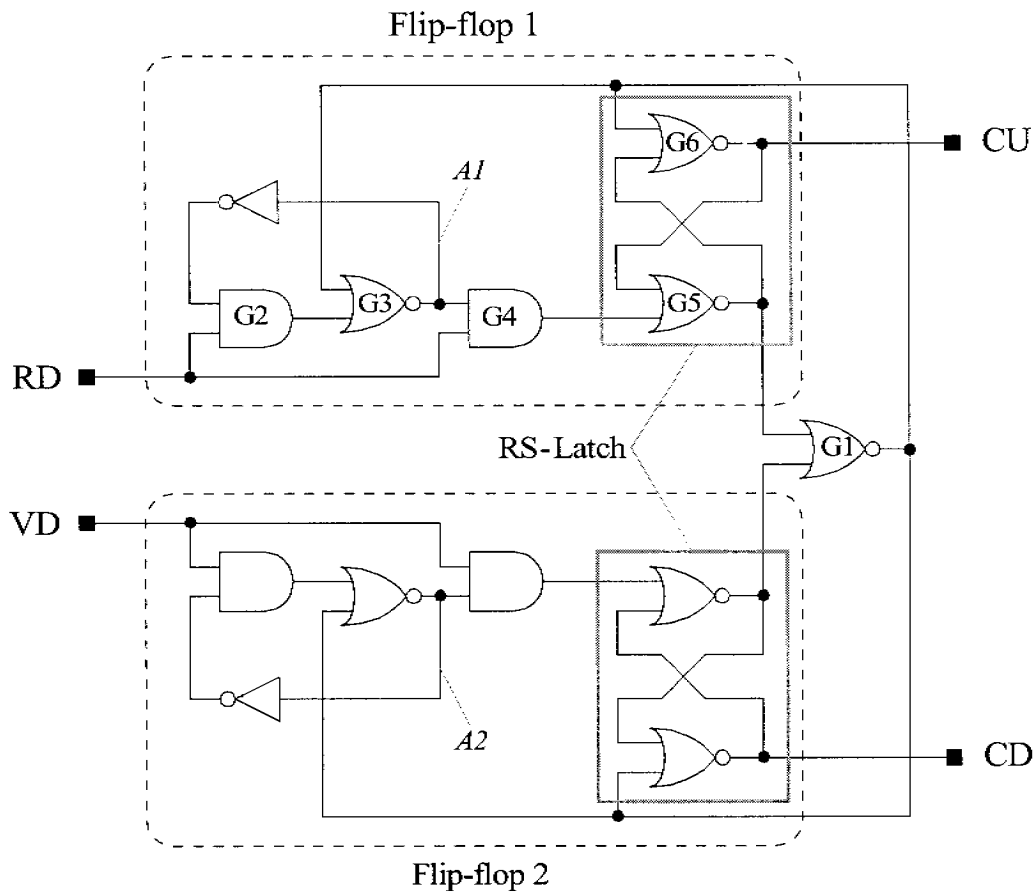
Figure 7.7: Gate level schematic of a PFD with speed improved reset logic.

set input of the RS-latch. Consequently, a positive edge at any clock input sets one of the RS-latches. Resetting of the latch is only possible by the appearance of a positive clock edge at the adjacent clock input, causing momentary setting of the second RS-latch until the reset pulse is generated by gate G1. The aim of the remaining gates is to prevent the clock input from setting the RS-latch during the reset procedure. This is achieved by disconnecting the clock from the RS-latch set input. Notice that the latches cannot be set if nodes A1 and A2 respectively are kept low. As long as the reset pulse from gate G1 is present, node A1 is kept low by G3. Later, when the reset pulse has disappeared, the clock input may be still high. To prevent the RS-latch from being set again, the A1 node is kept low by the help of G2 and the inverter. The gates accompanying the RS-latch guarantee thereby, that the PFD clock must perform a negative edge before the latches are ready to be set again.

The duration of the SHORT state is determined by the propagation delay of a single two-input gate, which is NOR gate G1. Further optimization of the PFD will be carried out later on transistor level.

7.2. PLL Design Procedure

RF band, frequency resolution and switching speed requirements postulated in Chapter 2 for GSM and WCDMA systems determine the PLL design parameters. Quadrature local oscillator signal generation by a digital divider flip-flop requires doubling the synthesizer output frequency to around 4GHz as well as doubling the reference frequency to 400kHz. Since frequency switching time is critical, the loop bandwidth must be chosen large. Keeping the latter at one tenth of the reference frequency is adequate to meet the switching time requirement for GSM and avoids undesired sampling effects. Further optimization requires to keep the PLL phase margin at around 45°.

The special case third order loop, identified in Chapter 3 as an optimum solution, is mainly determined by the above criteria. Determination of the loop components, however, requires further consideration of the synthesizer phase noise performance. Remember that PLLs with identical dynamic behavior and frequency resolution can be set up with different loop filter resistor and charge pump current combinations, but phase noise is deteriorated at low charge pump current due to excessive loop filter noise contribution. In order to quantify the additional phase noise caused by the loop filter, it is conducive to introduce the angular

intercept frequency ω_{ic} at which loop filter noise has dropped to a level equal to the intrinsic RF oscillator noise. Keeping the intercept frequency small mitigates loop filter caused phase noise but is penalized by a large charge pump current. A balanced compromise must be found depending on the synthesizer phase noise requirement.

Synthesis of the special case third order loop can be condensed into a seven step design procedure leading to the determination of all loop components. Parameters required to determine R, C, C_1 and I_{CP} values are the output frequency, the reference frequency, the loop bandwidth (which can be initially set to one tenth of the reference frequency), the loop filter noise intercept frequency, the RF oscillator gain and phase noise. Table 7.1 lists all steps with the according Equations. Notice that RF oscillator phase noise is specified by an equivalent noise source $f_{n,vco}$ present at the tuning node. Table 7.2 provides a numerical example which relates to the frequency synthesizer design example discussed in more detail in Section 7.3. Fig. 7.8 shows a plot of the expected synthesizer phase noise spectrum as well as the contribution of the individual noise sources.

Remember that all Equations underlying the PLL design procedure are founded on a linearized PLL model which was thoroughly analyzed in Chapter 3. Validation of the loop parameters with a more accurate model taking sampling and preferably also loop nonlinearities into account, is more than prudent, although the impact of these effects is expected to be low due to the moderate loop bandwidth to reference frequency ratio and the small RF band compared to the output frequency. A first validation is already carried out by design procedure step VII which proposes the computation of the impulse sampled open loop gain in order to estimate the loss of phase margin due to sampling of the error phase. Fig. 7.9 shows the Bode plot of the linearly modeled PLL as well as the Bode plot of the impulse sampled PLL model. As guessed before, the phase margin is only slightly reduced by aliasing. The computed loss of less than $4°$ is too small to have significant impact on the PLL's dynamic behavior. Readjustment of loop parameters is hence not required. The last step of the design procedure may become more important if the loop bandwidth is raised beyond one tenth of the reference frequency.

Validation of the PLL is completed if nonlinearities caused by non constant sampling and phase detector gain discontinuities are considered too. These effects are most easily treated by the help of transient simulations based on a behavioral

Ia.	*Loop bandwidth (settling time optimized)*	$\omega_{LBW} \approx \dfrac{1}{10} \cdot 2\pi \cdot f_{ref}$
Ib.	*Loop bandwidth (spurs optimized)*	ω_{LBW} *as small as possible*
II.	*Calculate b-factor determined by phase margin:*	$b = \left[\tan(PM) + \dfrac{1}{\cos(PM)} \right]^2$
III.	*Calculate pole and zero:*	$\omega_z = \dfrac{\omega_{LBW}}{\sqrt{b}} \qquad \omega_p = \omega_{LBW} \cdot \sqrt{b}$
IV.	*Calculate resistor*	$R = \dfrac{f_{n,\,vco}}{4kT} \cdot \left(\dfrac{b}{b-1}\right)^2 \cdot \left(1 + \dfrac{\omega_{ic}^2}{\omega_{LBW}^2 \cdot b} \right)$
V.	*Calculate capacitor*	$C = \dfrac{1}{\omega_z \cdot R} \qquad C_1 = \dfrac{C}{b-1}$
VI.	*Calculate charge pump current:*	$I_{CP} = \dfrac{D \cdot \omega_{LBW}}{R \cdot K_{KVCO}} \cdot \dfrac{b}{b-1}$
VII.	*Compute sampled gain:*	$[G_{OL}(s)]^* = \displaystyle\sum_{k=-\infty}^{\infty} G_{OL}\left(s - jk \cdot \dfrac{2\pi}{T_s}\right)$

Table 7.1: Seven step design procedure leading to the PLL design parameters R, C, C_1 and I_{CP}

PLL model. Fig. 7.10 shows the simulated frequency switching transient. Loop parameters correspond to those shown in Table 7.2. The PFD-CP combination as well as the RF oscillator are modeled using an analog description language [11]. The graphs show that the sampled error phase and the RF oscillator tuning voltage of the behavioral PLL follow closely the transients of the linear PLL model, which are shown in grey lines. This is not surprising since the small RF oscillator tuning range limits the error phase to the linear region of the phase detector even if the frequency step covers the full RF band. A strong deviation of the behavior-

Input Parameter		
VCO gain	K_{VCO}	70MHz/V
VCO phase noise	$L(f_m)$	-110dBc/Hz @ 100kHz
Reference phase noise	$L(f_m)$	-150dBc/Hz @ 100Hz
Output frequency	f_{out}	4GHz
Reference frequency	f_{ref}	400kHz
Loop bandwidth	f_{LBW}	40kHz
Noise intercept frequency	f_{ic}	200kHz
Computed component values		
	C	252pF
	C_1	52pF
	R	38kΩ
	I_{CP}	1.15mA

Table 7.2: Component values as obtained from the proposed design procedure.
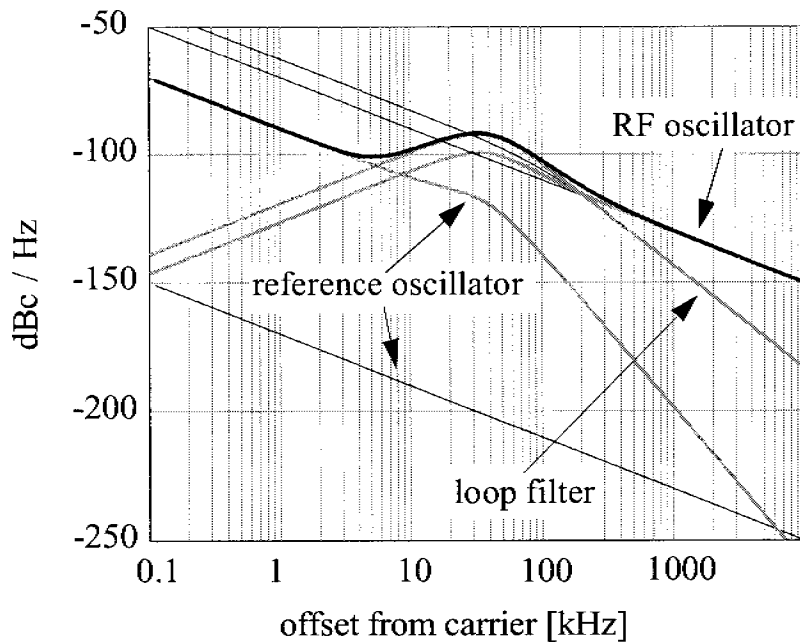
Figure 7.8: Total phase noise (bold), individual phase noise contributors (grey) and individual phase noise contribution with the loop opened (black).
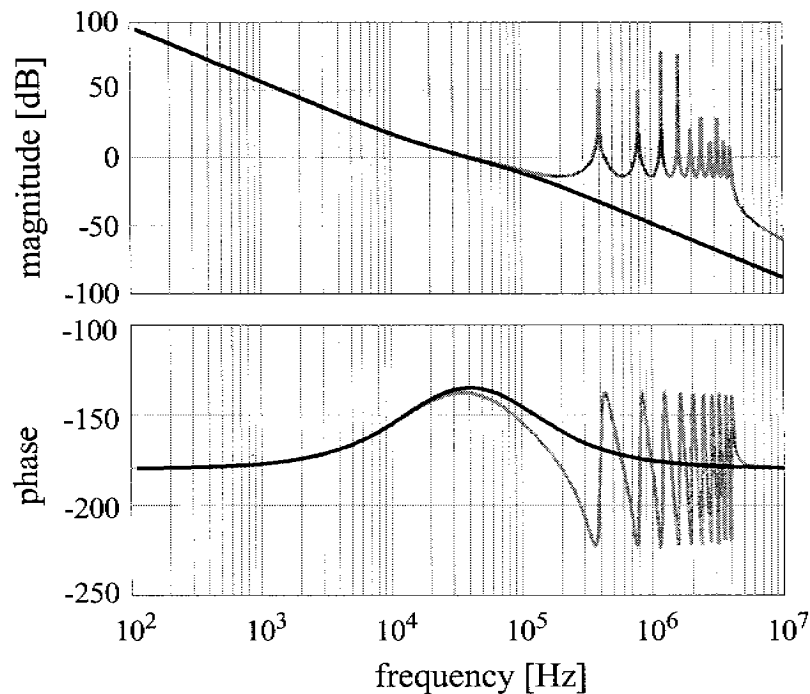
Figure 7.9: Open loop gain of the impulse sampled PLL (grey) and linear modelled PLL (black).

ally modeled PLL from the linear model may occur due to clipping of the tuning voltage at the supply rails. Clipping effects are not further considered.

7.3. Synthesizer Implementation and Experimental Verification

A 4GHz frequency synthesizer based on the proposed charge pump has been implemented in a 0.18μm CMOS technology to demonstrate the capabilities of the novel topology. The implemented circuits are more precisely described in the next Section followed by measurement results presented in Section 7.3.2.

7.3.1. A 4GHz, 0.18μm CMOS Synthesizer

The frequency synthesizer integrated circuit contains the frequency divider, PFD, charge pump as well as a three wire serial interface providing access to the internal registers. The RF and reference oscillator are off-chip devices as well as the loop filter [12] [13]. Fig. 7.11 shows a principal diagram of the frequency synthesizer. A resistive power splitter provides a 50Ω port for measurement purposes as

reference clock   reference clock
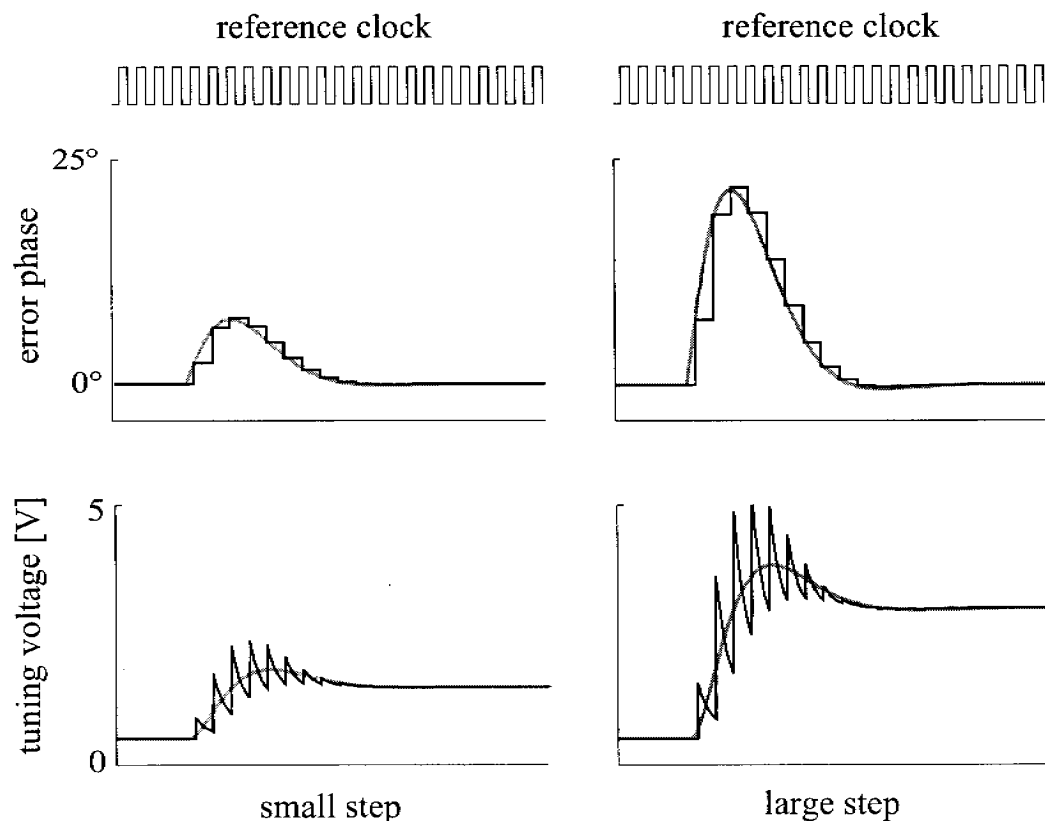
Figure 7.10: PLL switching transient as obtained from a behavioral simulation. The grey lines show the step response of the linear PLL model.

well as the RF signal to the divider input which is matched to 50Ω. Frequency division ratios ranging from 4096 to 16447 can be obtained by programming of a 14bit wide control word. Phase locking of the RF oscillator to a 400kHz reference signal, obtained from a 10MHz temperature compensated crystal oscillator (TXCO) module, accommodates theoretically the synthesis of frequencies ranging from 1.6GHz to 6.5GHz. Certainly, the output frequency is limited by the used RF oscillator rather than the frequency divider. The nominal supply voltage of 1.8V limits seriously the tuning range of the synthesizer. However, at least 3V tuning range is necessary to accommodate a band of 250MHz. Fortunately, the used technology provides transistors with a thicker gate oxide. They can be operated at a nominal supply voltage of 3.3V without risk of gate breakdown. The charge pump and all its auxiliary circuits are therefore implemented exclusively by thick gate oxide transistors. The remaining circuits, which are much more speed critical than the charge pump, use 1.8V transistors to benefit from the sig-

$$\circledcirc\ f_{out}$$

-6dB Power Splitter

VCO

4GHz

V_{tune}

$38k\Omega$

$252pF$

$52pF$

50Ω Strip Line

64/65 Prescaler

8b P-counter

6b S-counter

D Q

R

Delay

D Q

R

÷25 divider

3-Wire Interface

10MHz

TCXO

3bit programmable delay

level converter and switch driver
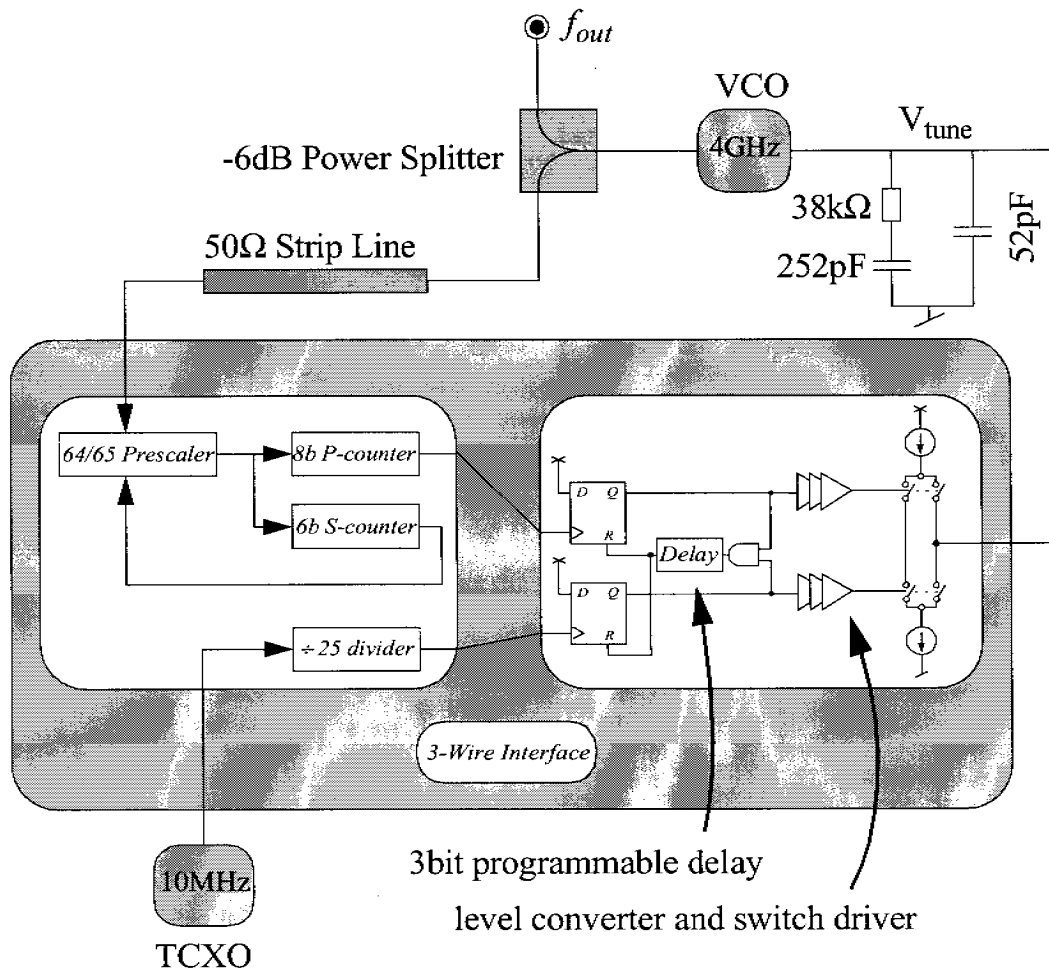
Figure 7.11: A simplified diagram of the implemented frequency synthesizer.

nificantly larger transit frequency. Interfacing of the two supply domains is managed by the switch drivers which are located between the PFD and the charge pump.

The remainder of this Section describes implementation specific issues which have been omitted in the principal circuit description carried out earlier. Divider related design issues are not considered since a detailed description of the low power 0.18μm CMOS frequency divider has been presented in Section 4.6.

PFD implementation:

The PFD is implemented on transistor level as proposed by Von Kaenel [3]. The flip-flop complexity is reduced by the help of 'complex' gates leading to a 19 transistor per flip-flop solution. The transistor size of selected gates is increased

to reduce their propagation delay. This is done for gates G6 and G1 whose size is doubled and quadrupled respectively with respect to the regularly sized gates (which is 3μm/0.18μm for NMOS and 4.5μm/0.18μm for PMOS). On top of this, the propagation delay of the flip-flop resetting gate G1 can be extended by selective addition of delay stages. The additional delay is under control of a register that allows to program the duration of the PFD SHORT state in steps of approximately 500ps starting from 1ns. Finally, the circuit is laid out symmetrically to achieve good matching of the two flip-flops.

Switch driver implementation:

Mainly three reasons command an interface circuit linking the PFD with the charge pump switches. The first one is isolation. While the transistors of the MOS switches are chosen large (30μm/0.34μm for NMOS, 90μm/0.34μm for PMOS) to minimize the voltage drop across the switches, the transistors used in the PFD are optimized for speed and are hence significantly smaller. An isolating buffer is required to drive the large input capacitance of the switches. The second reason commanding a switch driver lies in the necessity for slightly overlapping clock phases to keep residual charge sharing small. The last purpose of the switch driver is level conversion of the 1.8V signals delivered from the PFD to 3.3V signals required by the charge pump. Determined by the three purposes, the complete switch driver consists of three sections. The first section converts the signal level to 3.3V by the help of a circuit shown in Fig. 7.12. The next section shifts the clock edges to provide an overlap of around 100ps. The third and final section contains buffer stages which limit output slewing to approximately 300ps. The shortest activation of the charge pump current source or current sink amounts therefore to around 1ns. Besides clock phases for the main switches, the switch driver provides also clock phases driving the dummy switches introduced to reduce clock-feed-through.

Charge pump implementation:

In order to complete the description of the high performance charge pump circuit proposed in Section 7.1.1, the regulating amplifiers are presented in more detail now. The design is challenged by the fact that the voltage at the inputs of the amplifier responsible for calibration of the M4 drain current can vary from the negative to the positive supply voltage. The amplifier must hence provide a rail-to-rail input stage. A constant g_m, constant slew rate rail-to-rail operational

Figure 7.12: Digital level converter.

transconductance amplifier (OTA) was chosen with the schematic shown in Fig. 7.13 [14]. Depending on the input common mode level, two out of six differential pairs provide gain. The control loop consisting of the single-stage OTA and a second gain stage formed by the common source connected charge pump transistor M4, which is renamed in Fig. 7.13 as MCP, provides an extraordinary large gain. The gain of the first stage is enhanced by the large impedance determined by the two cascodes M14/M16 and M18/M20. A well characterized dominant pole is set up by the cascode impedance and the two compensation capacitors C_{CA} and C_{CB}, of which the effective capacitance is enlarged by the gain of the common source arrangement. The dominant pole at 10Hz, requiring a total of 28pF compensation capacitance, is chosen low enough so that an undesired high frequency pole pair as well as a right half plane zero cannot affect the stability of the control loop. Design Equations for the two-stage cascoded OTA are taken from [15]. Unity gain bandwidth of the control loop, determined only by the DC gain and the dominant pole, is at around 1MHz.

The implementation of the second amplifier, required to extend the charge pump output voltage range to the negative supply, can be kept more simple since this control loop must be active only at low charge pump voltages. A standard folded cascode OTA with a single PMOS input pair is employed. Since feedback is formed by a source follower rather than a common source gain stage, the gain of this control loop is significantly smaller. Stabilization is achieved therefore by a much smaller compensation capacitor of 6pF. Unity gain bandwidth is kept at the same 1MHz.

Figure 7.13: Rail-to-rail common mode input range OTA.

Noise of the charge pump potentially affects phase noise. However, the low duty cycle, determined by the ratio of the SHORT state duration and the reference frequency period, lowers the noise contribution of the charge pump significantly. Simulations have shown that the synthesizer peak phase noise is increased by roughly 3dB due to charge pump noise.

7.3.2. Experimental Results

The frequency synthesizer has been laid out in a standard 0.18μm CMOS technology. Fig. 7.14 shows a micrograph of chip which measures 1.5mm x 1.5mm. An evaluation board containing all required devices has been set up to characterize the performance of the frequency synthesizer. Fig. 7.15 shows a photograph of the board with the integrated circuit directly mounted onto the PCB substrate. The experimental setup is completed by a PC based graphical control software allowing convenient programming of the output frequency as well as of the PFD reset path delay. Loop filter parameters, charge pump current and parameters of the RF oscillator are listed in Table 7.2.

The synthesizer proved to be fully functional. Phase locking of the RF oscillator with the reference oscillator is obtained over the full tuning range. The frequency
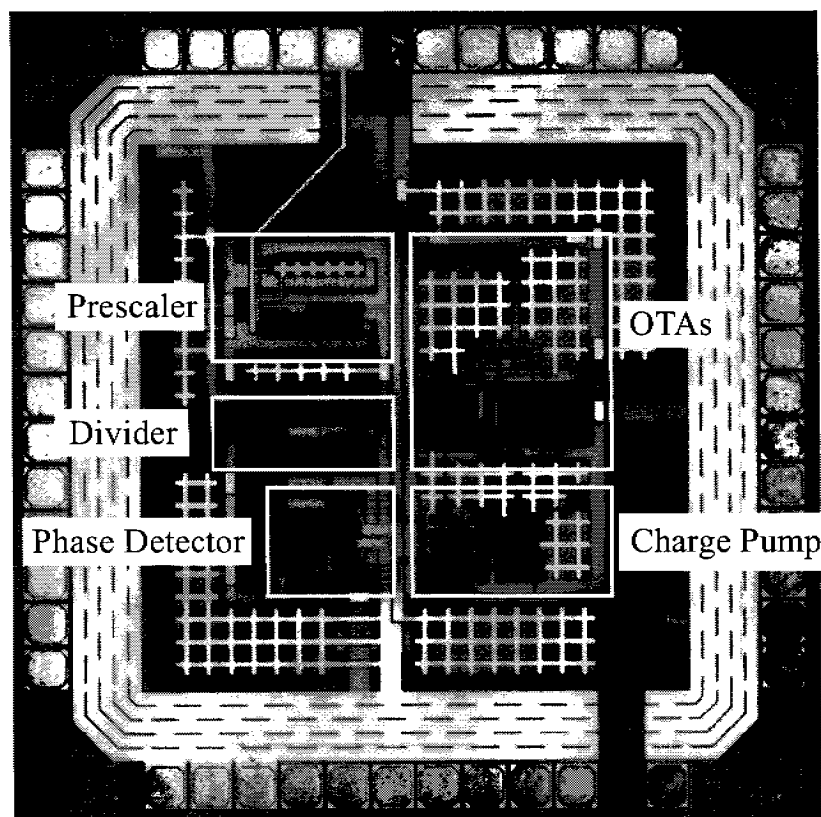
Figure 7.14: Chip micrograph of the 0.18µm CMOS frequency synthesizer.

switching transient can be observed most easily by measuring the RF oscillator tuning voltage. Fig. 7.16 shows the latter voltage, measured with a digital sampling oscilloscope, during switching of the output frequency from 3950MHz to 4050MHz.

The upper plot provides an oversight of the switching transient. The measured curve matches with that obtained from a simulation of the behavioral PLL model, indicating that both determination of the PLL parameters as well as implementation of the synthesizer circuits have been correctly carried out.

Validation of switching time compliance is a difficult measurement task and can be performed only indirectly. Required frequency stabilization after frequency switching in the order of a couple of dozens of Hertz would require measurement of the tuning voltage with an accuracy in the µV region. Only the beginning of the settling curve, which has the shape of a damped sine, can be observed due to the limited dynamic range of digital sampling oscilloscopes. To improve the measurement accuracy to some extent, the DC part of the tuning voltage was
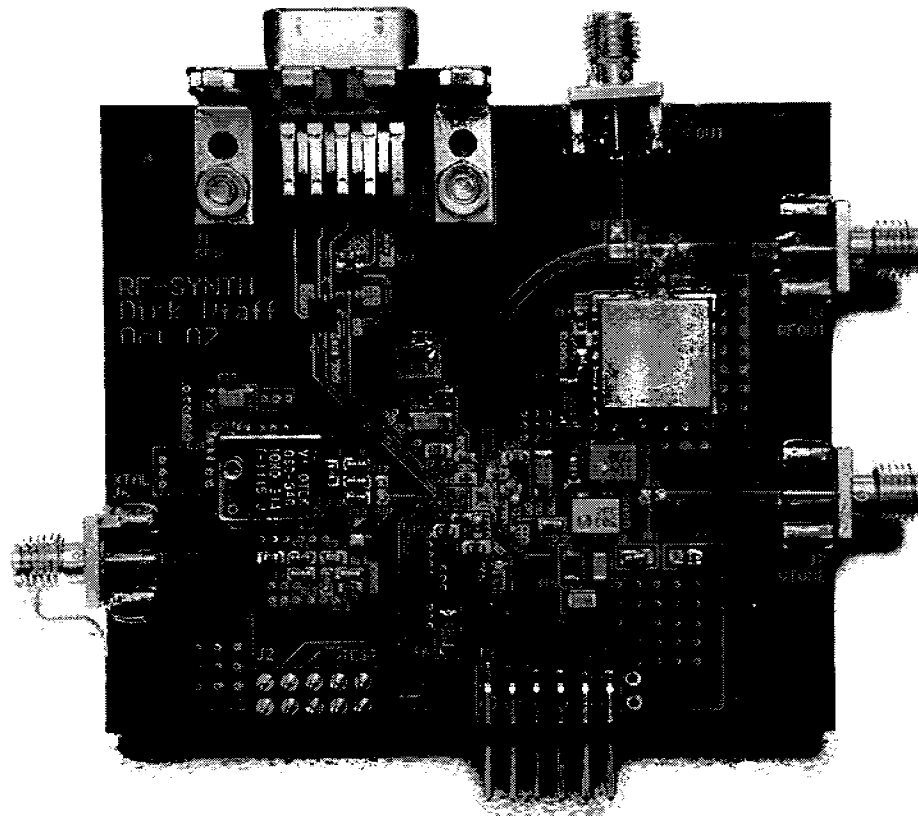
Figure 7.15: Frequency synthesizer evaluation PCB.

removed by an instrumentational amplifier. The lower plot of Fig. 7.16 shows a close-up view of the measured settling curve. Notice that the scaling differs from the upper plot. According to the lower plot, the period of the damped sine measures 30μs. This is 5μs above the expected value which can be obtained from Eq. 3.19. The discrepancy can be explained by additional loading through the instrumentational amplifier. Since the correct period of the sine is verified by measurements, one can postulate that the synthesizer meets the switching time requirement.

Eq. 7.17 shows the frequency synthesizer spectrum measured with an HP 8563E spectrum analyzer. The reset path delay of the PFD is set to minimum. The reference frequency spurious tone at 400kHz offset from a 4GHz carrier is measured as -68.5dBc/Hz. This number represents a significant improvement over previously reported synthesizers. The level of spurious tones versus PFD short state duration was predicted to scale with 12dB/oct. However, the measured increase is smaller, as can be seen in Fig. 7.18. Nevertheless, a strong increase of spurs is
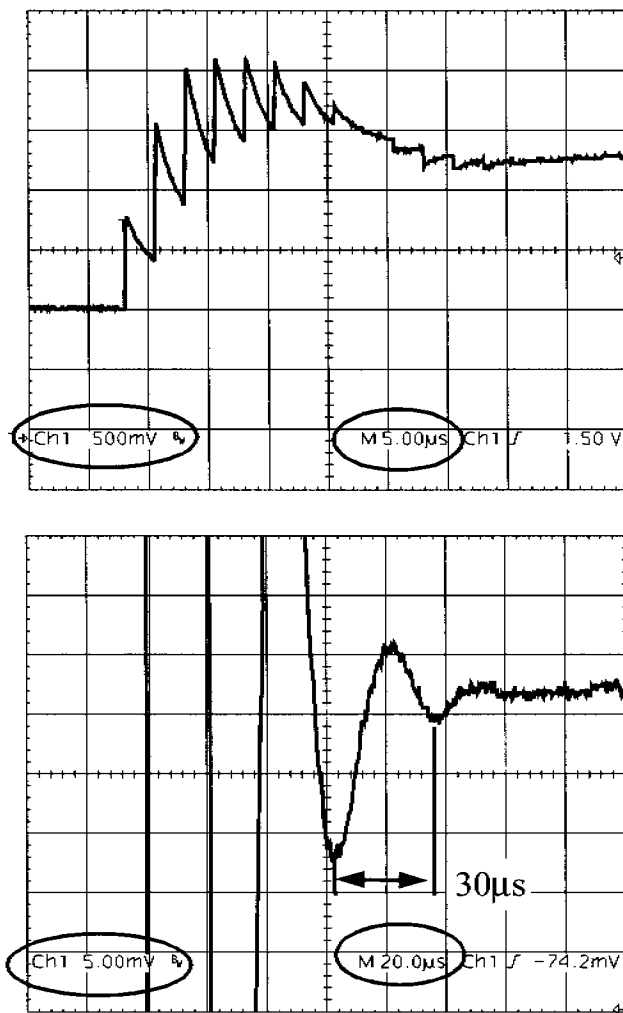
Figure 7.16: Measured RF oscillator tuning node voltage during an output frequency step from 3950MHz to 4005MHz.

found at extended PFD SHORT state durations. The smaller than 12dB/oct slope suggests that other effects, which are not dependent on the duration of the SHORT state, contribute significantly to the spurious tone power. This observation underlines the ability of the self-calibrating charge pump to remove efficiently current source mismatch.
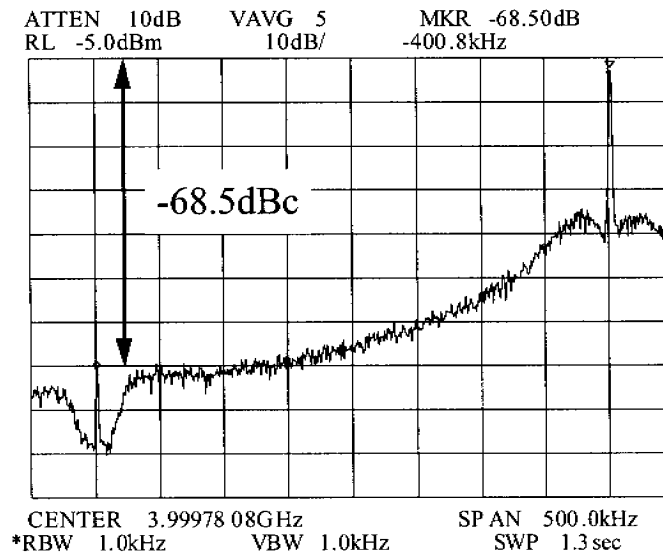
Figure 7.17: Frequency synthesizer spectrum measured with an HP 8563E spectrum analyzer.
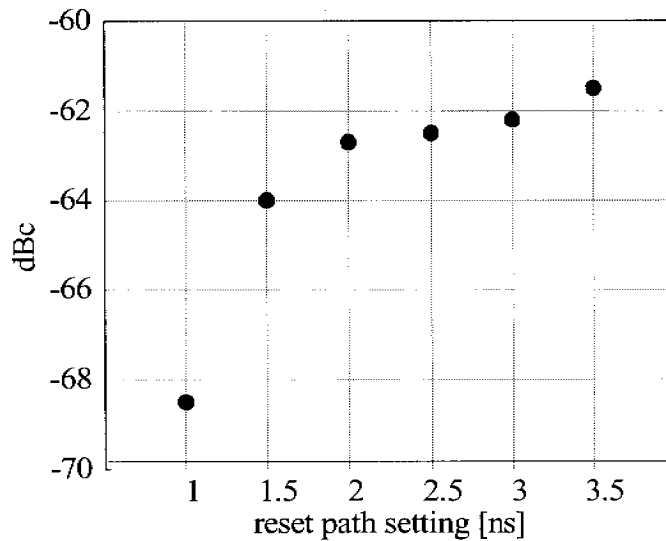

Figure 7.18: Reference frequency spurious tone at different reset path delay settings.

References

[1] Panasonic Components, *"Data Sheet PPS Film Chip Capacitor Type ECHU"*, Panasonic Components, 2002

[2] Motorola, *"Data Sheet MC4044 Phase-Frequency Detector"*, Motorola Semiconductor, 1972

[3] V. Kaenel, D. Aebischer, C. Piguet, E. Dijkstra, *"A 320MHz, 1.5mW at 1.35V CMOS PLL for Microprocessor Clock Generation"*, IEEE J. Solid State Circuits, Vol. 31, No. 11, Nov. 1996

[4] C. Lam, B. Razavi, *"A 2.6GHz/5.2GHz Frequency Synthesizer in 0.4μm CMOS Technology"*, IEEE J. Solid State Circuits, Vol. 35, No. 5, May 2000

[5] W. Rhee, B. Song, A. Ali, *"A 1.1GHz CMOS Fractional-N Frequency Synthesizer with a 3-b Third-Order DS Modulator"*, IEEE J. Solid State Circuits, Vol. 35, No. 5, Oct. 2000

[6] M. Johnson, E. Hudson, *"A Variable Delay Line PLL for CPU-Coprocessor Synchronization"*, IEEE J. of Solid State Circuits, Vol. 23, No. 5, Oct. 1988

[7] H. Rategh, H. Samavati, T. Lee, *"A CMOS Frequency Synthesizer with an Injection-Locked Frequency Divider for a 5-GHz Wireless LAN Receiver"*, IEEE J. of Solid State Circuits, Vol. 35, No. 5, May 2000

[8] T. Lin, W. Kaiser, *"A 900MHz CMOS Frequency Synthesizer with an Automatic SC Tuning Loop"*, Proc. CICC, 2000, pp. 375-378

[9] M. Pelgrom, A. Duinmaijer, A. Welbers, *"Matching Properties of MOS Transistors"*, IEEE J. Solid State Circuits, Vol. 24. No. 5, Oct. 1989

[10] C. Sharpe, *"A 3-state Phase Detector Can Improve Your Next PLL Design"*, Electronic Design News, Sept. 1976, pp. 55-59

[11] D. Fitzpatrick, I. Miller, *"Analog Behavioral Modeling With the Verilog-A Language"*, Kluwer Academic Publisher, 1997

[12] *"Product Data Sheet V844ME05 Voltage Controlled Oscillator"*, Z-Communications Inc., 2002

[13] *"Product Data Sheet OSC Series Ultra Miniature TCXOs"*, Vectron International, 1999

[14] W. Redman-White, *"A High Bandwidth Constant gm and Slew-Rate Rail-to-Rail Input Circuit and its Application to Analog Cells for Low Voltage VLSI Systems"*, IEEE J. Solid State Circuits, Vol. 32, No. 5, May 1997

[15] D. Ribner, M.Copeland, *"Design Techniques for Cascoded CMOS OpAmps with Improved PSRR and Common-Mode Input Range"*, IEEE J. Solid State Circuits, Vol. 19, No. 6, Dec. 1984

Chapter 8

Summary and Conclusions

This thesis started with two observations. First, the standby time of GSM/DCS/PCS cellular handsets is more and more limited by the power consumption of the frequency synthesizer which tends to consume typically around 100mW, while a GSM receiver with only 50mW consumption has been recently presented by Orsatti. Reduction of the synthesizer consumption was one of the main objectives of the thesis. Second, a strong trend towards fractional-N frequency synthesis has been observed, despite the larger complexity of the topology compared to the integer-N and despite the fact that faster frequency switching speed, which is an potential advantage of fractional-N synthesis, stays away. Reconsideration of the classical integer-N topology was the second main objective of the thesis.

Reduction of the frequency synthesizer power consumption requires optimization of the power-hungry prescaler and RF oscillator. While low power prescalers are traditionally realized with bipolar technologies, cost pressure demands for CMOS solutions. Although deep-submicron technologies ease the integration of power efficient CMOS prescalers, their design is more demanding than that of their bipolar counterparts due to the limited available transconductance of CMOS. Design guidelines for CMOS current mode logic have been presented, leading to low power prescaler solutions which even outperform bipolar solutions, at least if quarter micron CMOS or better are available. A 1GHz and a 4GHz dual modulus prescaler have been presented with aggressive consumption of only 0.9mA and 2.5mA respectively.

While prescalers profit directly from advanced CMOS due to their digital nature, the situation is completely different for RF oscillators. Indeed, over the years, constant power consumption level of frequency synthesizers is mainly caused by the lack of significant improvement of RF oscillators. Underlying reason for this lies in the fact, that the performance of RF oscillators is mainly determined by passive devices. Since the quality of integrated passive components is not

expected to improve with the down-scaling of CMOS, significant reduction of RF oscillator power consumption can be expected only from hybrid solutions.

Power in RF oscillator is primary used to meet the stringent phase noise requirement of GSM/DCS/PCS. A low power, low phase noise, 1GHz oscillator with a discrete resonator, meeting the GSM phase noise requirements, has been presented. The oscillator consumes only 0.25mA. This extremely low current level, which is roughly one decade below the consumption of oscillators with fully integrated resonators, is empowered by a roughly 4 times larger quality factor of the discrete resonator.

The accurate generation of quadrature oscillator signals, required in low-IF and zero-IF receivers, is another part that may raise the frequency synthesizer power consumption considerably. Generation of quadrature phases by the help of a digital flip-flop in toggle configuration profits also from down-scaling of CMOS. However, doubling of the RF oscillator frequency is required. Printed circuit board microstrip lines have been found extremely useful to implement low power, low phase noise multi-GHz oscillators. It has been demonstrated that a 3.6GHz oscillator meeting DCS/PCS phase noise requirements can be constructed with a microstrip resonator without affecting cost and size of the cellular handset. The stripline oscillator and quadrature divider flip-flop combination is able to drive highly linear, low noise downconversion mixers without any additional buffers. A very competitive consumption of 10mA including oscillator, quadrature divider and downconversion mixers, combined with excellent sideband rejection, has been reported.

The doubled RF oscillator frequency does not only permit power efficient generation of quadrature phases, it further allows to double the loop bandwidth of the phase locked loop. This enables frequency switching within 300µs with a final accuracy of 0.1ppm. Hence, no fractional-N topology is necessary to accommodate GSM high speed data services. However, the low reference frequency to loop bandwidth ratio requires careful design of the low frequency building blocks to avoid excessive generation of spurious tones in the spectrum of the frequency synthesizer. The building blocks which are critical to the generation of spurious tones are identified as phase-frequency detector and charge pump. Previously proposed charge pump topologies report weak spurious tones performance due to various artifacts. A novel circuit has been proposed to overcome the limitations of existing solutions. An integrated 4GHz frequency synthesizer

has been realized to demonstrate the capabilities of the new charge pump. The employed third order phase locked loop with an optimum constellation of the loop filter pole and zero frequency achieves a spurious tone suppression of -68dBc, despite the low reference frequency to loop bandwidth ratio of 10. One can conclude therefore, that the simple integer-N topology is capable to provide fast switching speed and decent spurious tone performance to meet GSM/DCS/PCS requirements.

This thesis demonstrates that the power consumption of GSM/DCS/PCS frequency synthesizers can be significantly reduced by a rigorous low power design approach. The power consumption of the critical building blocks has been successfully lowered to 3mA for a 3.6GHz RF oscillator, to 3.5mA for a quadrature generator and to 2.5mA for a 3.6GHz prescaler. This opens the way to frequency synthesizers with an overall consumption of 20mW while meeting GSM/DCS/PCS requirements. Reduction to this level, however, requires an external high-Q resonator to meet the stiff phase noise requirements at low biasing level. Upcoming third order WCDMA standards accept higher RF oscillator phase noise, enabling eventually fully integrated oscillators at a decent consumption. Finally, chip area, and consequently cost, can be lowered by relying on the simple integer-N topology. This, however, requires advanced phase locked loop building blocks to tackle the spurious tones problem.

Publications

[1] G. Brenna, D. Tschopp, D. Pfaff, Q. Huang, *"A 2GHz Direct-Conversion WCDMA Modulator in 0.25µm CMOS"*, Digest of Technical Papers of the IEEE International Solid-State Circuits Conference, pp 244-245, San Francisco, USA, February 3-7, 2002

[2] D. Pfaff, Q. Huang, *"An 18mW 1800MHz Quadrature-Demodulator in 0.18µm CMOS"*, Digest of Technical Papers of the IEEE International Solid-State Circuits Conference, pp. 242-243, San Francisco, USA, February 3-7, 2002

[3] D. Pfaff, J. Rogin, Q. Huang, *"A 14mA 2GHz 0.25µm CMOS Quadrature Demodulator Including a Low Phase Noise Local Oscillator"*, Proc. of IEEE European Solid State Circuits Conference, pp. 56-59, Villach, Austria, September 18-20, 2000

[4] A. Deiss, D. Pfaff, Q. Huang, *"A 200MHz Sub-mA RF Front-End for Wireless Hearing Aid Applications"*, IEEE Journal of Solid State Circuits, pp. 977-986, Vol. 35, July 2000

[5] D. Pfaff, Q. Huang, *'Low Power Circuits for RF-Frequency Synthesizers in the Low GHz Range"*, 9th Workshop on Advances in Analog Circuit Design, Tegernsee, Germany, April 11-13, 2000

[6] A. Deiss, D. Pfaff, Q. Huang, *"A 200MHz Sub-mA RF Front-End for Wireless Hearing Aid Applications"*, Proc. of IEEE European Solid State Circuits Conference, pp. 274-277, Duisburg, Germany, September 21-23, 1999

[7] D. Pfaff, Q. Huang, *"CMOS VCO/Prescaler Design for Very Low Power Applications"*, Proc. of Int. Workshop on Low Power RF Integrated Circuits, Lausanne, Switzerland, October 19-20, 1999

[8] D. Pfaff, Q. Huang, *"A Quarter-Micron CMOS, 1GHz VCO/Prescaler-Set for Very Low Power Applications"*, Proc. IEEE Custom Integrated Circuits Conference, pp. 649-652, San Diego, CA, USA, May 16-19, 1999

[9] M. Helfenstein, A. Muralt, G. Fischer, P. Zbinden, D. Pfaff, F. Frey and G.S. Moschyt, *"SC and SI filters in baseband applications: A comparison"*,

Proc. IEEE Int. Symposium on Circuits and Systems, Hong Kong, pp. 297-300, June 1997

Curriculum Vitae

Dirk Pfaff was born on April 4, 1970, in Zurich, Switzerland. He is a citizen of Liestal, Canton Baselland. Dirk Pfaff completed secondary school in Baden in 1986. He obtained Matura (Baccalaureate) of mathematical/scientific type from Kantonschule Baden in 1990. From 1990 he studied electrical engineering at the Swiss Federal Institute of Technology (ETH), Zurich, where he received the Dipl. Ing. degree in 1995. He joined Microswiss, Grenchen, Switzerland, in 1995 as a research and development engineer. There, he was contributing to several industrial IC design projects, mainly in the area of low power, low voltage, analog and mixed signal circuit design. Since 1997 he is with the Integrated Systems Laboratory of the Swiss Federal Institute of Technology in Zurich, where he was working towards the Ph.D degree in the field of analog and radio frequency integrated circuit design. His scientific interest is on CMOS frequency synthesis and other high frequency circuits for telecommunication systems. During this period, he was involved in several industrial research and development projects for Swiss and international customers.