

Sampler and Synthesizer for a low-cost Vector Network Analyzer

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Diplom-Ingenieurs

unter der Leitung von

Univ.-Ass. Dipl.-Ing. Dr. Holger Arthaber
und

Univ.-Prof. Dipl.-Ing. Dr. Gottfried Magerl

am

Institut für Elektrische Mess- und Schaltungstechnik (E354)

eingereicht an der Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

von

Paul-Jürgen Wagner

Matrikelnummer 9825610

1150 Wien, Geibelgasse 5/29

Wien, im Jänner 2007

*Dedicated to the loving memory of my grandfather,
who passed away on March 30, 2004.*

(This page unintentionally left blank.)

ABSTRACT

Vector network analyzers are the High-End of electronic measurement equipment, both in complexity and price. This project aims at designing basic building blocks for a small, PC-controlled two-port network analyzer suitable for everyday-measurements up to 3 GHz. The analyzer should be designed with standard devices to make construction easy and affordable.

This thesis describes two main modules of such a network analyzer's RF section: A *sampler*, used to convert the RF signals into lower frequency signals suitable for digital signal processing, and a *synthesizer* to provide the necessary excitation of the device under test.

In contrast to traditional sampler designs, the one presented here uses a differential pair to provide the sample pulses. This eliminates the need for transmission line balancing structures, which request for expensive high-permittivity substrates. The sampler's 3 dB-bandwidth is 2.5 GHz, its 6 dB-bandwidth is 4 GHz. The dynamic range is 60 dB.

The synthesizer module generates signals between 2 GHz and 3 GHz with an output power between -3.7 dBm and -7.7 dBm. The worst-case phase noise at a frequency offset of 1 kHz is -80 dBc/Hz. The module features an additional fixed 1 GHz output, which facilitates the downconversion of the main synthesizer's signal to extend the frequency range.

KURZFASSUNG

Vektorielle Netzwerkanalysatoren gehören zu den leistungsfähigsten, aber auch teuersten Messgeräten der elektronischen Messtechnik. Ziel dieser Arbeit ist, grundlegende Funktionsblöcke für einen einfachen, PC-gesteuerten Zweitor-Netzwerkanalysator zu entwickeln, der für Messaufgaben der täglichen Praxis bis 3 GHz geeignet ist. Der Analysator soll, um den Aufbau einfach und kostengünstig zu halten, ausschließlich mit Standardbauteilen aufgebaut sein.

Diese Arbeit beschreibt zwei Hauptkomponenten des Hochfrequenzteils eines solchen Netzwerkanalysators: Einen *Sampler*, der die Hochfrequenzsignale in niederfrequente, der digitalen Signalverarbeitung zugängliche Signale umsetzt, und einen *Synthesizer*, der die zur Anregung des Messobjektes notwendigen Signale erzeugt.

Im Gegensatz zu herkömmlichen Ausführungen wird beim hier entwickelten Sampler zur Erzeugung der Abtastimpulse ein Differenzverstärker verwendet. Diese Variante vermeidet die sonst zur Symmetrierung üblichen Leitungsstrukturen, die nur auf teuren Substraten mit hoher Dielektrizitätszahl vernünftig realisiert werden können. Die 3 dB-Bandbreite des Samplers ist 2.5 GHz, die 6 dB-Bandbreite 4 GHz. Der Dynamikbereich umfasst 60 dB.

Der Synthesizer erzeugt Signale zwischen 2 GHz und 3 GHz mit einer Ausgangsleistung im Bereich zwischen -3.7 dBm und -7.7 dBm. Das Phasenrauschen beträgt schlechtestenfalls -80 dBc/Hz bei einer Offsetfrequenz von 1 kHz. Das Modul hat einen zusätzlichen 1 GHz-Ausgang, mit dessen Hilfe der Frequenzbereich des Hauptsynthesizers durch Mischung erweitert werden kann.

Contents

Introduction	≡	vii
1 Network Analysis	≡	1
1.1 Scattering Parameters		1
1.2 Scalar Network Analysis		3
1.3 Vectorial Network Analysis		4
2 Sampler	≡	9
2.1 Theory of Operation		9
2.2 Circuits		10
2.3 Frequency Response		13
2.4 Noise		16
2.5 Pulse Generation		18
2.6 Results		22
2.7 Improvements		27
3 Synthesizer	≡	29
3.1 RF Oscillators—Theory and Practice		31
3.2 Phase-Locked Loop		40
3.3 Results		41
3.4 Improvements and Prospects		46
A Annotated Schematics	≡	48
A.1 Sampler		48
A.2 Synthesizer		51
B Board Layouts	≡	54
B.1 Sampler		54
B.2 Synthesizer		56
Bibliography	≡	58

Introduction

Vector Network Analyzers (VNAs) have been around since the 1960s, being the RF engineer's instrument of choice to characterize linear (and, to some extent, nonlinear) circuits. While in the early days they occupied a full 19-inch-rack, they have shrunk to desktop size in the meantime. On the other hand, the price of commercially available VNAs still reaches the 100.000 Euro-range without any troubles, basically because these instruments employ features often unnecessary in daily business.

This project tries to fill a gap: To provide a simple-as-possible two-port VNA suitable for measuring the bulk of today's RF components, which operate below 2.5 GHz. This includes components for broadcasting-, GSM900-, GSM1800-, and UMTS-equipment, all devices operating at the ISM-bands (excluding the highest ISM-band at 5.8 MHz), and the currently most popular WLAN-standard IEEE 802.11b/g. The VNA should be controlled by a Personal Computer, eliminating the need for a separate display unit. Standard parts available to everyone (not only business customers) at a random electronics distributor should be used exclusively, making it possible for the VNA to be built by a sufficiently skilled amateur. In some sense this project is also meant to be a case study about how much performance can be achieved in an RF circuit by using more-or-less standard devices.

NOTATION

\sim	if the right hand side is a variable expression, means ‘proportional to’, e.g. $P \sim V^2$; if the right hand side is a constant, means ‘in the order of magnitude of’
\approx	approximately equal, equal after neglecting second order terms
\doteq	rounded to the next meaningful standard value
\equiv	identically equal to
$:=$	defined to be
\parallel	reciprocal sum, $x \parallel y := (x^{-1} + y^{-1})^{-1}$; defined to bind stronger than multiplication, but weaker than raising to a power
\hat{x}	maximum value of x (in the current context)
\check{x}	minimum value of x (in the current context)
\bar{x}	typical (or average) value of x (in the current context)
$\lfloor x \rfloor$	floor of x ; integer that satisfies $\lfloor x \rfloor \leq x < \lfloor x + 1 \rfloor$
$\lceil x \rceil$	ceiling of x ; integer that satisfies $\lceil x - 1 \rceil < x \leq \lceil x \rceil$
$x^{(n)}$	n -th Fourier coefficient of $x(t)$
$\mathcal{F}x(t)$	Fourier transform of $x(t)$
$\arg x$	complex argument; $\arg x = x/ x $
$O(\cdot)$	Landau’s big-O; $g(x) = O(f(x)) \Rightarrow \exists x_0, c \forall x > x_0: g(x) \leq c f(x) $
$\text{si } x$	sine cardinal (sinc); $\text{si } x := (\sin x)/x$
$\sigma(t)$	unit step function; $\sigma(t) := 1$ if $t \geq 0$, $\sigma(t) := 0$ if $t < 0$
\mathbf{x}	column matrix (aka ‘vector’)
\mathbf{X}	matrix
\mathbf{X}^\top	\mathbf{X} transposed
\mathbf{X}^{-1}	inverse of \mathbf{X}
k_B	Boltzmann’s constant, $k_B = 1.380\,650\,5 \times 10^{-23}$ J/K
Q_e	elementary charge, $Q_e = 1.602\,176\,53 \times 10^{-19}$ C
V_T	thermal voltage, $V_T = k_B T / Q_e$
ADC	Analog- to Digital Converter
DUT	Device Under Test
VNA	Vector Network Analyzer

ACKNOWLEDGEMENTS

While in theory circuits always work, in practice they often do not. While in theory my plans have always been perfect, in practice they often did not work out. But I got to know a lot of interesting people that have influenced the way I am thinking, and it is time to say ‘Thank you’ to them.

I would like to thank the staff of the Institut für elektrische Meß- und Schaltungstechnik, especially Markus Mayer, Holger Arthaber, Christian Schuberth and Mike Gadringer, for their support, both technical and social; Tibor Grasser for believing in my skills; the colleagues I had the pleasure to work with at Agilent Technologies Österreich — you have taught me lessons no university can teach.

Special thanks go to my dear friends and colleagues Bernhard Sogl and Ayse Adalan, for their encouragement and for joining me at Colosseum concerts. Bernhard has to be thanked also for providing me with the quote in Chapter 3.

Very special thanks go to my boyfriend Christoph, who endured my bad moods, talked me out of doubts and proofread this text.

But the biggest ‘Thank You’ goes to my parents and grandparents. Not only have they provided the financial support for my journey so far, but also more important they have laid the foundations to everything I have achieved in the past and to everything I may achieve in the future.

CHAPTER 1

Network Analysis

*The need to know
and the chance to try
the hope to find
the other side of the sky.*
— Colosseum, “The Other Side of the Sky”

At low frequencies, the customary concept of *voltage* and *current* is perfectly suitable to fully describe the electrical properties of a certain point in an electrical circuit at a certain instant of time. But if frequency rises to the point where the physical extent of the circuit has the same order of magnitude as the associated wavelength $\lambda = c/f$, the nature of wave propagation must be taken into account. Voltage and current can not be defined unambiguously (in case of TEM-capable structures) any more; sometimes they can not even be defined in any way, e.g. with waveguides. They are replaced with *incident* and *reflected wave*.

1.1 SCATTERING PARAMETERS

Consider a linear n -port. The n voltages and currents are collected into column matrices $\mathbf{V} = (V_1 \ V_2 \ \dots \ V_n)^\top$ and $\mathbf{I} = (I_1 \ I_2 \ \dots \ I_n)^\top$. Now, at each port an arbitrary impedance $Z_{0,i}$ called the *reference impedance* is chosen. In general these impedances are complex, but for most applications real reference impedances are sufficient. Moreover, based on different physical interpretations, different definitions for the wave variables are in use.[#] If the reference impedances are real, however, these differences vanish. Moreover, complex reference impedances are of little practical importance. Therefore, in the rest of

[#]Detailed discussions on the rather complicated issues involved with scattering matrices normalized to complex impedances can be found in [Chen], [Pauli], [Anritsu AN246], [Wohlers], [Carlstein], [Zhu], [Frickey], and [Marks].

the thesis purely real reference impedances are assumed. With the matrix

$$\mathbf{W} = \begin{pmatrix} \sqrt{Z_{0,1}} & 0 & \cdots & 0 \\ 0 & \sqrt{Z_{0,2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \sqrt{Z_{0,n}} \end{pmatrix}$$

the columns \mathbf{V} and \mathbf{I} are transformed into the column of incident waves \mathbf{a} and the column of reflected waves \mathbf{b} by the ‘coordinate transform’

$$\mathbf{a} = \frac{\mathbf{W}^{-1}\mathbf{V} + \mathbf{W}\mathbf{I}}{2} \quad \text{and} \quad \mathbf{b} = \frac{\mathbf{W}^{-1}\mathbf{V} - \mathbf{W}\mathbf{I}}{2} .$$

Since the n -port considered is linear, a_i and b_i are related to each other in a linear fashion, just as V_i and I_i are. These relations are expressed as

$$\mathbf{b} = \mathbf{S}\mathbf{a} ,$$

where \mathbf{S} is a complex $n \times n$ matrix called the *scattering matrix*.

Determining the elements of \mathbf{S} is straight forward: Terminate $n - 1$ ports with their respective reference impedances, and excite the remaining port. Assuming port number l is excited, $a_i \equiv 0$ for all $i \neq l$. The equation for b_k reduces to $b_k = S_{kl}a_l$, and by measuring b_k and a_l we can calculate

$$S_{kl} = \left. \frac{b_k}{a_l} \right|_{a_i \equiv 0 \ \forall i \neq l} .$$

The scattering matrix characterizes an n -port just as the impedance-, admittance-, or the various hybrid matrices do. But there is an important difference: In contrast to the other matrices mentioned, the scattering matrix *always exists*. The reason for this is twofold: First, even if, due to a short or open circuit, voltage or current at a certain point becomes zero, the quotient of incident and reflected wave stays bounded. And second, the procedure to get S_{kl} described above terminates any port with some finite impedance, opposed to the procedures used to determine the other matrices, which terminate with short or open circuits. These ‘singular terminations’ potentially result in singular matrices. Furthermore, at high frequencies short and open circuit are cumbersome idealizations: In order to exclude any parasitic inductance, a possible short would need to have zero length, and in the light of the fact that free space has a wave impedance of about 377Ω , the concept of an open circuit becomes more than questionable. The scattering parameters’ definition incorporating finite termination impedances is therefore ideally suited for high frequencies. For a rigorous proof of the existence of scattering matrices for passive networks cf. [Carlin].

The theory of scattering parameter measurements only deals with two-ports, because the results can easily be applied to n -ports with $n > 2$ by terminating $n - 2$ ports with matched loads and measuring the remaining two ports. Then two other ports are chosen to be the ‘active’ ones, and measure-

ment is started again. After running through all possible combinations the complete $n \times n$ -scattering matrix is determined.

1.2 SCALAR NETWORK ANALYSIS

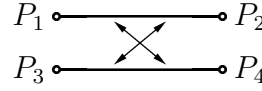
If the scattering matrix of an n -port is to be measured, a device that separates the incident and reflected wave on a transmission line is required. Among such devices are *directional couplers*. They come in different flavours, depending on the type (waveguide, microstrip, etc.) of line used. A directional coupler transmits a fraction of the wave incident on the *input port* (port 1) to the *coupled port* (port 4); the remaining portion arrives at the *transmitted port* (port 2), whereas in an ideal world no signal reaches the *isolated port* (port 3). The coupler is usually specified by the quantities

$$\text{Coupling factor} \quad C = \frac{P_1}{P_4} ,$$

$$\text{Insertion loss} \quad L = \frac{P_1}{P_2} \approx \frac{C}{C-1} ,$$

$$\text{Isolation} \quad I = \frac{P_1}{P_3} ,$$

$$\text{Directivity} \quad D = \frac{P_4}{P_3} = \frac{I}{C} .$$



Being only interested in the magnitudes of the scattering parameters, it should be sufficient to measure the magnitudes of the waves. This can be done by an arrangement as shown in Figure 1.1. If the directional coupler and the source are ideal, i.e. $I \rightarrow \infty$ and $\rho_s \rightarrow \infty$,

$$P_x = |a|^2 \frac{C-1}{C^2} |\rho_x|^2 .$$

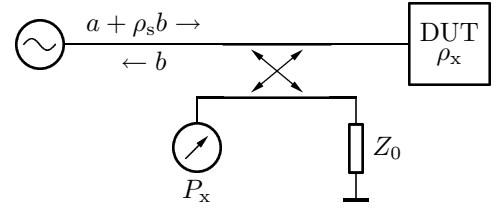


Figure 1.1 Scalar measurement of ρ_x .

The unknown source power $|a|^2$ can be eliminated if the setup is calibrated using a reference element of known reflection coefficient. In general such a calibration standard will be a short or an open, yielding the power meter readout P_c . The unknown reflection coefficient then is

$$|\rho_x|^2 = \frac{P_x}{P_c} .$$

Things get complicated if the non-idealities are considered. Only the magnitudes of the directional coupler's specifications are known, but the waves are added vectorially. By introducing the complex directional coupler quantities γ , λ , and ι , which have the magnitudes $|\gamma|^2 = C$, $|\lambda|^2 = L$, and $|\iota|^2 = I$, but

unknown phases, the power meter readout is

$$P_x = |a|^2 \frac{C-1}{C^2} \left| \rho_x + \underbrace{\frac{\lambda\gamma}{\iota} + \frac{\rho_s\gamma}{\lambda^2}\rho_x + \frac{\rho_s}{\lambda^2}\rho_x^2}_{=:\Delta\rho_x} \right|^2.$$

The measurement error $\Delta\rho_x$ consists of a constant term due to the finite isolation (directivity), a term quadratic in ρ_x caused by the source mismatch and a term linear in ρ_x . The last term has its reason in the fact that when calibrating the measurement setup, the same errors were present (cf. [HPWorkshop]). Since no phase information is available, it remains unknown in which way the error terms add up. The ‘true’ value of $|\rho_x|$ must be assumed somewhere in the interval $[|\rho_x - \Delta\rho_x|, |\rho_x + \Delta\rho_x|]$. The measurement accuracy can be enhanced by using a levelled source, effectively zeroing ρ_s , and by performing two calibrations, one with a short and one with an open, yielding two calibration curves where the errors have maximum magnitude but different signs. The error caused by the finite directivity, however, remains.

A way to get rid of the systematic errors introduced by the non-ideal measurement system is to include all possible sources of errors in a mathematical model and use a number of calibration measurements to get enough data to solve the model for the unknown reflection or transmission coefficient(s). An example of such a model is the *sixport-method* (cf. [Schiek]), shown in Figure 1.2. It uses four scalar instruments (power meters) to measure the magnitude and phase of a reflection coefficient. The systematic errors of the instrument are subsumed in the six-port, whose scattering matrix \mathbf{K} must (at least partially) be determined through calibration. Clearly, the requirement of four independent measurements to calculate one complex, i.e. two real numbers, does not leave too good prospects on measuring a complete two-port. But the idea to represent the whole system by a (linear) equation system is the base for vectorial network analysis.

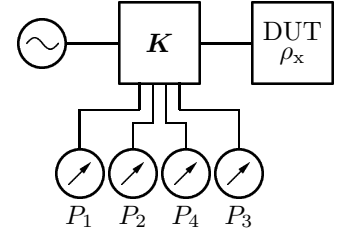


Figure 1.2 Sixport measurement setup to determine magnitude and phase of ρ_x by a calibrated arrangement with four power meters.

1.3 VECTORIAL NETWORK ANALYSIS

Measuring not only the magnitudes, but also the phases of the waves does not only add substantial information about the DUT; it furthermore makes it possible to account for all measurement errors due to imperfect matching or finite directivity, and (at least theoretically) provides a perfect correction of these errors by measuring one (complex) quantity for each quantity seeked.

Figure 1.3 shows a four-receiver VNA, which is the most general setup to measure a two-port. The switches are always operated in opposite directions, i.e. if port 1 of the DUT is excited, port 2 is matched, and vice versa. Since this operation in theory always leaves one incident wave zero, actually three receivers are sufficient: Two that measure the reflected waves from the DUT on both sides, and one that measures the source power at the point where it is split. However, this requires switches with perfect isolation; using the four-receiver method, imperfections of the switches are included in the error model.

The error model does not require that the directional couplers are in fact directional. What is needed to calculate \mathbf{S} are eight vectorial measurements (four measurements in the two switch states, each) that are *linearly independent*. This allows for replacing the directional couplers with resistive π -networks, enabling lower measurement frequencies that could not be handled by directional couplers, since their coupling efficiency degrades when their physical dimensions get small compared to the wavelength.

To determine both magnitude and phase of a signal, two approaches are possible. Figure 1.4 shows a *quadrature demodulator* (IQ-demodulator). The RF signal is split and separately multiplied with two sinewaves that have the same frequency as the input signal but are 90° out of phase. The resulting signals are filtered through narrowband lowpass filters to pick out the DC levels. These DC levels are proportional to the real and imaginary parts of the phasor associated with the input signal. The second method is to sample and digitize the signal and calculate the magnitude and phase by numerically fitting a sinewave to the sequence of samples.

Actually, both methods share a common ground: They both aim at calculating the complex Fourier coefficient of the input signal's fundamental. But whereas the first method directly applies the formulae for the real Fourier coefficients (the lowpass filters carry out the integration), the second disguises as a curve fitting problem. Nevertheless, the least squares fit of a sinewave to a series of equidistant samples is indeed the fundamental Fourier series component.

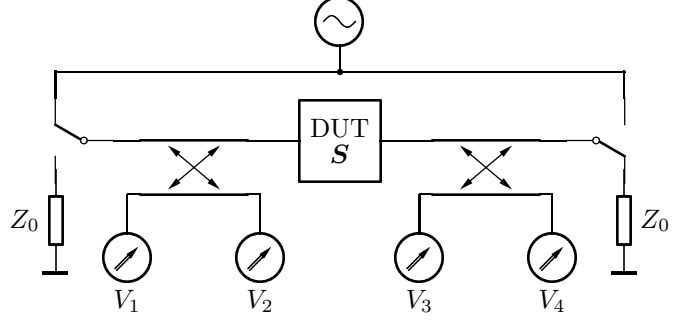


Figure 1.3 Vector Network Analyzer with four measurement receivers for complete measurement of the DUT's scattering matrix \mathbf{S} . The double-arrows indicate that the signal (voltage) is measured as a phasor, i.e. magnitude and phase are determined.

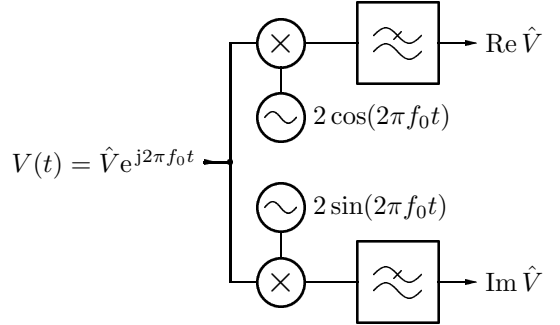


Figure 1.4 Quadrature demodulator used to determine real and imaginary part of the input signal's phasor.

The quadrature demodulator has the virtue of delivering its output as DC signal, which can be digitized with high resolution at low speed. On the other hand, this strength is one of its major weaknesses: DC offsets of both amplifiers and the ADC affect the measurement. The sampling and digitizing with the second technique must be done at high speed, more than two samples need to be acquired in every period according to Nyquist's Theorem. Presently, analog to digital conversion in the two-digit gigahertz range is not feasible, but the application of the heterodyne principle (see below) can shift the frequency the ADC is required to handle down to some kilohertz. The additional signal processing involved with the second method does not really hurt; further processing of the measured data requires a good deal of number crunching, anyway.

Manoeuvring in the Frequency Domain

As indicated, direct processing of the RF signals can be difficult, if not impossible. Luckily, the property

$$\mathcal{F}(x(t)e^{j2\pi f_0 t})(f) = \mathcal{F}(x(t))(f - f_0)$$

shows that multiplying a signal with a harmonic oscillation leaves the magnitude and phase of the signal intact, while allowing for an arbitrary shift in frequency. Since our world is restricted to real-valued quantities, the equation above has to be revised. Using the identity $\cos x = (e^{jx} + e^{-jx})/2$ leads to

$$\mathcal{F}(x(t) \cos(2\pi f_0 t))(f) = \frac{1}{2} \left(\mathcal{F}(x(t))(f - f_0) + \mathcal{F}(x(t))(f + f_0) \right).$$

Multiplying a signal with a cosine (aka *mixing*) results in two images of the original signal's spectrum. Usually only one is of interest, the other one is filtered out. If the mixer is followed by some bandpass with bandwidth B and center frequency f_1 , the system acts as a bandpass with bandwidth B and center frequency $f_0 + f_1$. This *heterodyne principle* allows to select the spectral components of interest by adjusting the oscillator producing f_0 . Furthermore, all subsequent signal processing is done at a constant frequency band around f_1 and can therefore be designed narrowband. Unfortunately, every real signal's spectrum has also a hermitian component at negative frequencies, resulting in a frequency band of width B and center at $f_0 - f_1$ also being translated to the output. To ensure uniqueness this *image band* must be filtered out before the mixer, what is easier when f_1 is not too small. Therefore, a heterodyne receiver is often made up from several frequency conversion stages. Image rejection is easiest when both f_0 and f_1 are above the input band. In this case, actually the image band is used for conversion, whereas the band $f_0 + f_1$ is located beyond any potential input signal. The drawback is that for a given maximum input frequency, a system that works at even higher frequencies must be constructed. For a VNA, the image band issue is not a great one, anyway. Even if the DUT is nonlinear, i.e. producing harmonics, none of them is mixed down if $f_0 + f_1$

equals the excitation frequency, because then $f_0 - f_1$ is below f_0 , whereas all harmonics are well above.

An ideal mixer is a *bilinear* device. If the local oscillator's signal contains harmonics, these harmonics conduct separate downconversion processes,

$$\mathcal{F}\left(x(t) \sum_n s_n e^{j2\pi n f_0 t}\right)(f) = \sum_n s_n \mathcal{F}(x(t))(f - n f_0) .$$

These do not interfere with the main downconversion process if the DUT is linear. If it is nonlinear, unwanted superposition of spectra is avoided as long as $k(f_0 \pm f_1) \neq n f_0 \pm f_1$ for all the arbitrary combinations of integers n and k , except $n = k = 1$ of course. On the other hand, the side effect of multiple mixing can be exploited by making the principal conversion with $n > 1$. This is called *harmonic mixing*. Note that for a Fourier series to converge, the coefficients must fall at least with $O(1/n)$, which means that the higher the harmonic the mixing is done with, the lower the mixer's conversion gain. This effect can be mitigated to a certain extent by using Fourier series that, at least up to some harmonic index, decay as slow as possible, thus 'broadening' the oscillator's spectrum. More spread in the frequency domain (more high-frequency components) translates into a more compact, 'localized' shape in the time domain. The most extreme signal in this respect is a Dirac pulse train, which has the Fourier coefficients[#] $s_n = f_0$. But multiplying a signal with a Dirac pulse is the ideal model of a *sampler*, and downconversion with a sampler actually means making use of the otherwise undesired *aliasing*.

Mixer-based vs. Sampler-based Network Analyzers

When designing a VNA, it must be decided whether conventional mixers (working on the fundamental of the local oscillator) or samplers (harmonic mixers) are to be employed. Using samplers has the advantage that the local oscillators in the downconverters can operate at much lower frequencies than the measurement frequency, and that the oscillator's required frequency span is considerably lower. Whereas in a mixer-based VNA the mixers must be fed from an oscillator capable of essentially the same frequency range as the VNA's measurement range, thus effectively doubling the costly broadband oscillator, the oscillator in a sampler only needs to cover a frequency ratio of two. If a higher or lower sampling frequency is requested, there is always another harmonic which can be used to convert with a sampling frequency within the stated range. The drawback of a sampler-VNA is that samplers have a higher noise figure than fundamental mixers, reducing the system's dynamic range.

[#]These Fourier coefficients actually do not obey the $O(1/n)$ rule, but the Dirac pulse is not a function in the common sense that a series may converge to.

The first automated Vector Network Analyzer was the Hewlett Packard 8510. It is a sampler-based analyzer with four receivers. The control unit with the baseband signal processing, the test set with the directional couplers and the samplers and the source are distinct instruments that are mounted together in a rack. The system covers a frequency range from 45 MHz up to 26.5 GHz, 50 GHz, or 110 GHz, depending on the available budget. The sampling frequency is in the range of 50 to 300 MHz, the first IF at 20 MHz, the second at 100 kHz. At the time of introduction of the 8510, swept sources were quite popular. They are programmed with a start- and a stop frequency and exercise a continuous frequency ramp instead of stepping through like synthesized sources. The problem with sweepers is that since the frequency ramp control signal is generated by analog circuitry, the frequency accuracy is not too good. Moreover, the continuous sweep does not allow the receivers to operate on a constant input frequency, requiring the sampling frequency to rise synchronously with the measurement frequency. The 8510 solves this by controlling the sampling signal generator with a ‘dual phase-locked loop’. The first one, responsible for course adjustments, brings the sampler close to the measurement frequency before the ramp starts. When the first loop locked, the sweep is initiated, and the second loop tracks the sweeper. The sampler for the incident wave at port 1 (a_1) is used as part of the phase detector of the phase-locked loop.

A more recent generation of VNAs is the PNA series from Agilent. Being the successor of the discontinued 8510, they have reached the 110 GHz-range by now. In contrast to the older instrument, they are one-box solutions using a step synthesizer to generate the excitation and mixers instead of samplers to increase the dynamic range. The lower noise with mixers also has an impact on the measurement speed: Allowing for the same noise level in the measurement, a lower instrument noise allows a wider resolution bandwidth, in turn enabling higher sweep speeds. The PNA series also uses four measurement receivers.

CHAPTER 2

Sampler

*Time as an independent variable,
given at low cost by Nature,
has the advantage of nearly,
if not actually, infinite resolution.*
— George A. Philbrick

As described in the previous chapter, every VNA needs at least three vectorial receivers. In this low-cost VNA, these receivers are built with samplers, which translate the RF measurement signal to baseband using a harmonic of the sampling frequency.

2.1 THEORY OF OPERATION

With a sampler, any sampling frequency f_s satisfying

$$f_i = nf_s + f_o$$

with integer n can be used to convert signal components at frequency f_i down to f_o . Since all signals involved are real valued and therefore possess hermitian spectra, the equation also holds if some or all frequencies are negative. Of course, reversing all signs does not add significant information, as well as choosing f_s to be negative, which is absorbed by a negative n . So we arbitrarily choose $f_i > 0$. If then $f_i > f_o$, which is the case if the sampler is actually downconverting, only $n > 0$ is possible, and apart from the previous equation also

$$f_i = nf_s - f_o$$

is a valid scenario.

The principle freedom in choosing f_s is limited by the aforementioned fact that harmonics with higher indices in general have lower power. To yield the highest possible conversion gain (or, rather say, the lowest possible conversion loss) for given input and output frequencies, n must be chosen as small as possible. If the sampler is capable of sampling frequencies $\check{f}_s \leq f_s \leq \hat{f}_s$,

sampling at harmonic

$$n = \left\lceil \frac{f_i - f_o}{\hat{f}_s} \right\rceil$$

with

$$f_s = \frac{f_i - f_o}{n} = \frac{f_i - f_o}{\left\lceil \frac{f_i - f_o}{\hat{f}_s} \right\rceil}$$

results in the best conversion gain possible for the given f_i . Note that in most cases the second solution $f_i = n f_s - f_o$ yields the same n ; in some cases, however, an n that is higher by 1.

But not only the conversion gain calls for low harmonic numbers and high sampling frequencies. When a signal at f_i is downconverted by harmonic n , the noise in the vicinity of f_i is downconverted, too. Moreover, all the other harmonics of f_s are present as well, and even if at the frequencies $k f_s \pm f_o$ with integer $k \neq n$ there are no signals that interfere with the one at $n f_s + f_o = f_i$, the inevitable noise present at these frequencies is translated to f_o . The input port of the sampler is somehow bandlimited, and the additional noise power is directly proportional to the number of harmonics of f_s that fit into the samplers input bandwidth. Therefore, a sampling frequency as large as possible also minimizes the sampler's noise figure.

The required range for f_s may be easily derived by assuming that some frequency f_i is sampled at a certain harmonic n of the maximum sampling frequency \hat{f}_s . The same f_i could also be sampled with the next higher harmonic $n + 1$ using a lower sampling frequency f'_s . Therefore we have $f_i = n \hat{f}_s + f_o$, and on the other hand $f_i = (n + 1) f'_s + f_o$; equating these two yields

$$\frac{\hat{f}_s}{f'_s} = \frac{n + 1}{n} \leq 2 .$$

A tuning range of one octave for f_s is indeed sufficient, at least as long as $f_i > \hat{f}_s/2$. In the range $f_o < f_i \leq \hat{f}_s/2$ only $n = 1$ is possible, and f_s has to be lowered beyond $\hat{f}_s/2$. This unfortunately means that the number of sampling pulses per time interval gets very small, drastically lowering the sampling signal's average power and therefore the sampler's conversion gain.

In the remainder of this chapter, $f_o \ll f_s < f_i$ is assumed.

2.2 CIRCUITS

In the simplest theory, sampling a signal is modelled by multiplying it with a Dirac pulse train

$$\sum_{n=-\infty}^{\infty} \delta(t - n T_s) = f_s \sum_{n=-\infty}^{\infty} e^{j 2 \pi n f_s t} , \quad f_s = 1/T_s .$$

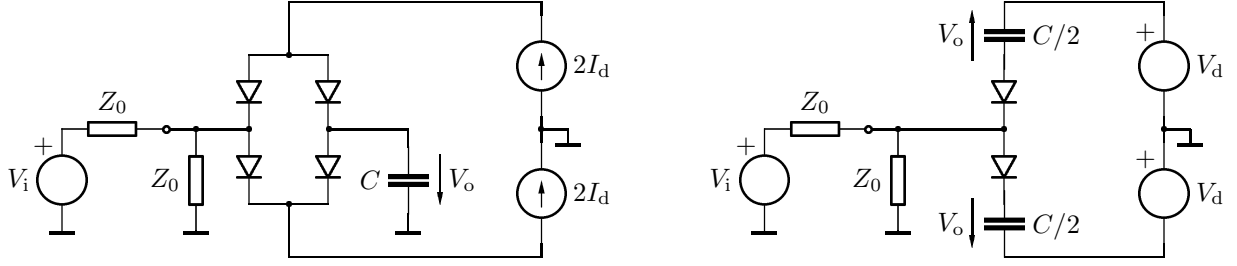


Figure 2.1 Diode switching gate circuits (cf. [Kampe]). Left: Basic circuit, where the sampling pulses must be provided from current sources. Right: Splitting the capacitor and moving it in series with the diode gate allows the usage of voltage sources.

This shows that the amplitude of the sampling pulse's harmonics, and therefore the samplers gain, depends on the sampling frequency. A remedy for this is not only to build a sampler, but a *sample and hold* circuit. The pulse response of a hold circuit is a rectangular pulse of duration T_s , corresponding to the transfer function $T_s e^{-j\pi f T_s} \text{si}(\pi f T_s)$. As long as the output frequency is small compared to f_s , $\text{si}(\pi f T_s) \approx 1$ and the hold circuit cancels the samplers dependence on f_s .

The sampling itself is usually done with a diode switching gate, rarely with FETs (cf. [Hafdallah]). The left part of Figure 2.1 shows a classic diode switching gate. The current sources drive the diodes into forward bias, where they allow the hold capacitor C to be charged to $V_i/2$. If the diodes have the same IV -characteristic, the current distribution is fully symmetric, and the sampling pulses do not interfere with the voltage to be sampled. One problem with the circuit is the need for *current* pulses: The pulse sources' impedances are in parallel with C ; if the sources have low impedance, they short-circuit the input signal. The second problem is the fact that with respect to the input signal, two diodes are in series in each branch. Their cumulative small-signal resistance slows the charging of C .

Both issues are solved by dividing the hold capacitor and moving the two equal parts in series with the pulse sources. Now, only one diode is in series with each hold capacitor, and low-impedance pulse sources present no problem because the hold capacitor is in series. The two diodes on the right of the diode gate do not have a function any more and are left out. The only drawback of this circuit is that the sampling pulse is carried by the hold capacitors now, imposing a lower bound on their value. Furthermore, while in the first variant the output was fully isolated from the sampling pulses, the sampling pulses now appear in full strength at the hold capacitors, but fortunately with opposite signs. By adding the voltages from the two capacitors, isolation can be reestablished. Since in practice sampling- and output frequency are well spaced apart, filtering out the sampling pulses at the sampler's output is easily possible, and only intermodulation in the subsequent amplifier is a potential issue.

The leakage of the sampling signal to the input port is a different story. Termed *sampler kick-out*, it can only be minimized by maintaining high symmetry, i.e. identical diodes and equal sampling pulses in both branches. Kick-out may have an adverse effect through intermodulation in preceding stages of the signal path, and can be suppressed by adding an amplifier to the sampler input. This not only isolates the sampler from the remaining RF path but also improves the sampler's input match.

Besides the traditional *voltage sampler* described here, there is a dual circuit called *current sampler*: It uses a hold inductor instead of a hold capacitor, the switching diodes are in parallel with the inductor and the input, and are normally 'on'. This circuit, as well as a circuit with a series- and a shunt-switch, which maintains a constant input impedance, are presented in [Kampe].

The main challenge when building a sampler is the generation of sampling pulses. This process can be split into three stages: Pulse generation, pulse shaping, and symmetrizing. In the first stage, a rectangular pulse with a rise time as short as possible is generated. For this purpose *step-recovery diodes* (SRDs) are often used. These are pn- or pin-diodes with a special doping profile which, when driven from forward- into reverse bias, first continue to carry their forward current until all carriers are cleared from the junction. At this point the diode's current immediately drops to zero, causing a sharp edge in the diode's voltage. Through this, pulses with a rise time of ~ 100 ps and a repetition rate of some ten megahertz may be generated from sinewaves. For a deeper discussion of SRDs cf. [Jungmeister]; a sampler using an SRD is described in [Han05]. Unfortunately, SRDs are nonstandard devices, and couldn't be used within this project.

If the rise times of SRDs are not sufficient, as it is the case in extremely high bandwidth applications, *nonlinear transmission lines* (NLTLs) are employed. Basically an NLTL is a homogeneous transmission line that is periodically shunted with Schottky diodes or varactors along its length. Despite these discontinuities the line still appears homogeneous if the distance between two diodes is short compared to the wavelength of the highest frequency on the line. If a pulse propagates along the line, the pulse amplitude modulates the diodes' capacitances, locally modifying the line's group velocity. Thereby it is possible to compress one edge of a pulse, usually the falling one. Fall times of 480 fs corresponding to a sampler bandwidth of ~ 1000 GHz have been reported [Kahrs]. Nonlinear transmission lines are only reasonable when built with high-speed GaAs diodes on a thick- or thin-film substrate, or when monolithically integrated. Fortunately, the requirements in this project are not that high that an NLTL is necessary. For a more complete description of nonlinear transmission lines, see for example [Rodwell] and [Afshari].

The second step, pulse shaping, aims at shortening the pulses. This is most easily done by a shorted transmission line. Different line types are used, mainly because this second step is often combined with the third step,

symmetrization. For this purpose, structures such as microstrip- to coplanar-waveguide-transitions (cf. [Lee01a]), microstrip- to slotline-transitions (cf. [Madani]), or special balun structures such as the *Marchand balun* (cf. [Devlin]) are in use. Unfortunately, they are difficult to manufacture in a low-cost environment, or request for high-permittivity substrates; for instance, slotlines only have meaningful dimensions if built on a substrate with $\epsilon_r \sim 10$.

In order to design a sampling circuit within these restrictions, an approach similar to [Devlin], albeit in a somewhat different context and configuration, is used: A differential amplifier acts as an active balun. The pulses are generated in a driving stage, sharpened and symmetrized in the differential amplifier, and the pulse shaping can then be done by two shorted microstrip lines, making the whole circuit uniplanar.

In the following two sections, general properties of diode sampling gates are derived. The matter of sampling pulse generation is picked up again in Section 2.5.

2.3 FREQUENCY RESPONSE

In order to analyze the sampler's frequency response, the sampling diodes are modelled by a time dependent conductance $g_d(t) = I_d(t)/V_T$. Since (at least in theory) the circuit is balanced with respect to the sampling pulse, the two branches of the sampling gate are combined into one; the two branches' voltages are added, therefore the single branch equivalent circuit yields only $V_o/2$. The input resistance of the adder/amplifier is modelled by R_l .

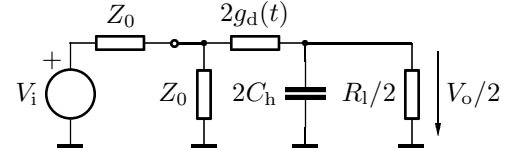


Figure 2.2 Equivalent circuit of the sampler to calculate the frequency response.

The circuit has associated the differential equation

$$V_i R_l g_d = \left(1 + R_l C_h \frac{d}{dt}\right) V_o + g_d Z_0 \left(1 + R_l/Z_0 + R_l C_h \frac{d}{dt}\right) V_o .$$

To find the output voltage under steady state excitation,

$$V_i = \hat{V}_i e^{j2\pi f_i t} ,$$

a harmonic ansatz is made. Components at arbitrary combinations of both the input- and the sampling frequency are expected to be found in V_o :

$$V_o = \sum_{k_i, k_s = -\infty}^{\infty} V_o^{(k_i, k_s)} e^{j2\pi(k_i f_i + k_s f_s)t} .$$

The diode's small signal conductance is represented as Fourier series in f_s , where the maximum conductance (when the diode current has its peak) is factored out for convenience,

$$g_d = \hat{g}_d \sum_{l=-\infty}^{\infty} g_d^{(l)} e^{j2\pi l f_s t} .$$

Inserting these into the differential equation and using the abbreviation

$$f_1 = \frac{1}{2\pi R_l C_h}$$

we get

$$\begin{aligned} \hat{V}_i R_l \hat{g}_d \sum_{l=-\infty}^{\infty} g_d^{(l)} e^{j2\pi(f_i + lf_s)t} = \\ \sum_{k_i, k_s=-\infty}^{\infty} \left(1 + j(k_i f_i + k_s f_s)/f_1\right) V_o^{(k_i, k_s)} e^{j2\pi(k_i f_i + k_s f_s)t} \\ + \hat{g}_d Z_0 \sum_{k_i, k_s, l=-\infty}^{\infty} \left(1 + R_l/Z_0 + j(k_i f_i + k_s f_s)/f_1\right) g_d^{(l)} V_o^{(k_i, k_s)} e^{j2\pi(k_i f_i + (k_s + l)f_s)t} \end{aligned}$$

The components at distinct frequencies can be separated. Assuming f_i and f_s incommensurate, $V_o^{(k_i, k_s)} \equiv 0$ for all $k_i \neq 1$, because on the left side only terms at the fundamental of f_i are present. For $k_i = 1$ and by using the sampling equation $f_i = n f_s + f_o$, the last formula can be rewritten as

$$\begin{aligned} \hat{V}_i R_l \hat{g}_d \sum_{l=-\infty}^{\infty} g_d^{(l-n)} e^{j2\pi(f_o + lf_s)t} = \\ \sum_{k=-\infty}^{\infty} \left(1 + j(f_o + k f_s)/f_1\right) V_o^{(1, k-n)} e^{j2\pi(f_o + k f_s)t} \\ + \hat{g}_d Z_0 \sum_{k, l=-\infty}^{\infty} \left(1 + R_l/Z_0 + j(f_o + (k-l)f_s)/f_1\right) g_d^{(l)} V_o^{(1, k-l-n)} e^{j2\pi(f_o + k f_s)t} . \end{aligned}$$

Taking the individual spectral components apart,

$$\begin{aligned} \hat{V}_i R_l \hat{g}_d g_d^{(k-n)} = \left(1 + j(f_o + k f_s)/f_1\right) V_o^{(1, k-n)} \\ + \hat{g}_d Z_0 \sum_{l=-\infty}^{\infty} \left(1 + R_l/Z_0 + j(f_o + (k-l)f_s)/f_1\right) g_d^{(l)} V_o^{(1, k-l-n)} . \end{aligned}$$

We are after $V_o^{(1, -n)}$, which is the output amplitude at f_o . By truncating the infinite series, a linear equation system can be formulated which can be solved for $V_o^{(1, -n)}$. However, if the lowpass formed by C_h and R_l has a sufficiently small cutoff frequency, $f_s \gg f_1$ and the terms $V_o^{(1, k-n)}$ with $k \neq 0$ can be neglected. Therefore, the sampler's frequency response is

$$H_s(f_i) = \frac{V_o^{(1, -n)}}{\hat{V}_i} = \frac{R_l \hat{g}_d g_d^{(-n)}}{1 + Z_0 \hat{g}_d g_d^{(0)} (1 + R_l/Z_0) + j(1 + Z_0 \hat{g}_d g_d^{(0)}) f_o/f_1} .$$

Note that f_i does not appear explicitly in the equation for H_s , but it is implicitly present, since both n and f_s (and therefore the $g_d^{(l)}$ themselves) depend on f_i . Although it does not look so at first glance, H_s is indeed hermitian: Changing $f_i \rightarrow -f_i$, the sampling equation requires both $n \rightarrow -n$ and $f_o \rightarrow -f_o$. But $g_d(t)$ is real-valued, and therefore $g_d^{(n)} = (g_d^{(-n)})^*$.

In order to be able to confront the model presented with reality, an assumption about the sampling pulse must be made. It would be asking for too much to tempt calculating $g_d(t)$ exactly, since the processes involved are highly nonlinear. Not only would the dynamics of the driving stage require to be included, but also the huge exemplar variations of the diode parameters and their temperature dependence make it utterly impossible to predict the precise waveforms. Measuring the voltage across the diodes on an assembled sampling circuit is also pretty difficult, because the whole path from the probe tip to the scope must be as homogeneous as possible to avoid reflections and dispersion of the pulse while presenting a negligible load to the circuit. On the other hand, the precise shape of the diode's conductance is of minor importance to the sampler's frequency response as long as the approximation used matches the actual pulse's profile by and large. The decision whether the pulse is modelled as rectangular, trapezoidal, or Gaussian has only little impact on the model (through a slightly different decline of the Fourier coefficients), as long as peak value and pulse duration are properly chosen.

In order to avoid too complicated calculations but to use a pulse function that is sufficiently smooth, $g_d(t)$ is modelled as a raised cosine,

$$g_d(t) = \begin{cases} \hat{g}_d \frac{1}{2} (1 - \cos(\pi t / \tau_s)) & 0 \leq t < 2\tau_s \\ 0 & 2\tau_s \leq t < T_s \end{cases} ;$$

its Fourier coefficients are

$$g_d^{(l)} = \frac{1}{T_s} \int_0^{T_s} \frac{g_d(t)}{\hat{g}_d} e^{-j2\pi l f_s t} dt = e^{-j2\pi l f_s \tau_s} \text{si}(2\pi l f_s \tau_s) \frac{\tau_s f_s}{1 - (2l f_s \tau_s)^2} .$$

The waveform and the Fourier coefficients are depicted in Figure 2.2, as well as the resulting frequency response (solid traces). Also shown in the figure is the spectrum for a rectangular pulse. Up to the first null it does not significantly differ from the raised cosine spectrum; above, the sampler is not usable anymore.

The model can also be used to show the detrimental effects of sampling pulse echos. Since both the diodes and the pulse generating stages (either diodes or transistors) are nonlinear, one can not expect that the equivalent source- and load impedances can be perfectly matched to the transmission lines carrying the pulses and taking part in the pulse forming process. Also device variations and temperature effects, especially in the diode's IV -curve, cause imperfect match. Therefore, the main sampling pulse will be followed by smaller, delayed replica. In the model above, the effects of a second pulse with normalized peak value α_2 and delay τ_2 ,

$$g'_d(t) = g_d(t) + \alpha_2 g_d(t - \tau_2) ,$$

can be seen by inserting the new Fourier coefficients

$$g_d'^{(l)} = \left(1 + \alpha_2 e^{-j2\pi l f_s \tau_2}\right) g_d^{(l)} .$$

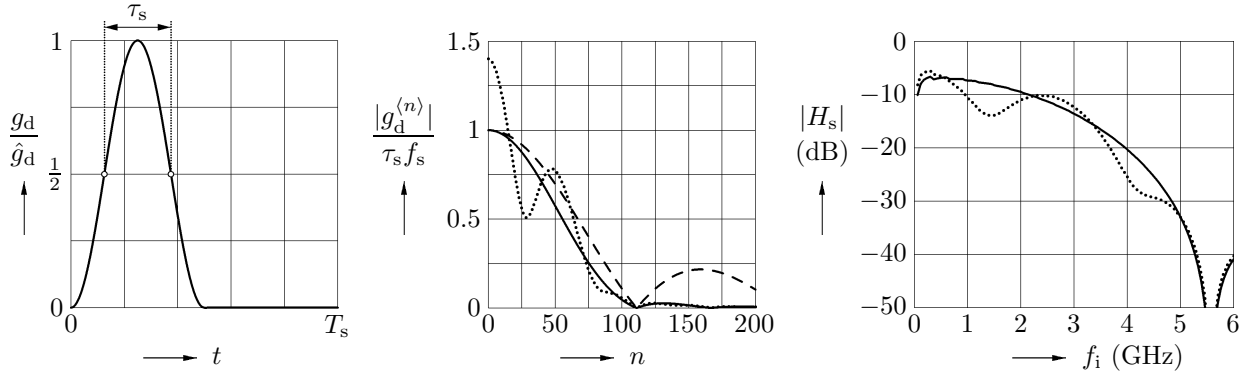


Figure 2.3 Left: Shape of the sampling diode's conductance; Center: Its spectrum (solid) and the spectrum of the conductance with echo pulse (dotted) compared to the spectrum of a rectangular pulse of width τ_s (dashed), all three drawn for $f_s = \hat{f}_s$; Right: The sampler's frequency response; solid for a single pulse, dotted for a pulse with echo. The parameters were $R_1 = 1 \text{ k}\Omega$, $C_h = 10 \text{ pF}$, $Z_0 = 50 \Omega$, $\hat{I}_d = 3 \text{ mA}$, $\hat{f}_s = 50 \text{ MHz}$, $f_o = 10 \text{ MHz}$, $\tau_s = 180 \text{ ps}$, $\tau_2 = 4\tau_s$ and $\alpha_2 = 0.4$.

Depending on whether the transmission line's impedance is too large or too small to match the impedance level of the remaining circuit, the echo will have positive or negative amplitude. Echos with negative amplitudes, however, can be disregarded, because they just drive the diodes into heavier reverse bias; on the other hand, they themselves will have echos, this time with positive amplitudes. So, in any case, mismatch will cause trailing pulses, either with $\tau_2 = 2\tau_1$ in the case of a too high line impedance, or with $\tau_2 = 4\tau_1$ in the case of a too low line impedance; τ_1 is the propagation delay of the line, which is connected to the sampling pulse width by $\tau_s = 2\tau_1$. The dotted curves in Figure 2.3 show the effects of an echo pulse on the spectrum and on the frequency response. The small bumps at low input frequencies are caused by the variations in $g_d^{(l)}$ with f_s . They get steeper at higher input frequencies because there the variations in f_s are smaller.

Further sampler analyses are provided in [Raleigh] and [Williams].

2.4 NOISE

To estimate the noise performance of the sampler, the equivalent circuit in Figure 2.4 is used. Only thermal and shot noise are considered, where for the shot noise component an ideal (rectangular) current pulse with amplitude \hat{I}_d through the sampling diodes is assumed. The noise in the adder's input resistance R_1 is calculated separately, so it is assumed noise-free. Since the sampler's output signal is bandpass-filtered with center frequency f_o and bandwidth B in the subsequent signal processing stage, noise is considered within a bandwidth B .

The noise at the sampler's output due to thermal noise in the source resistance and the match resistor in the sampler can simply be calculated using the transfer characteristic H_s from the last section. The squared noise

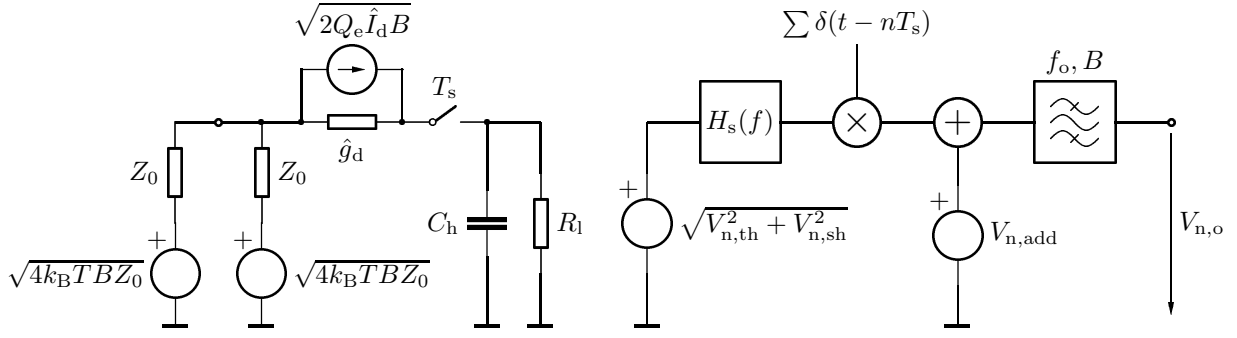


Figure 2.4 Left: Equivalent circuit of a single sampler branch with regard to noise. Right: Block diagram of the entire sampler's noise model; the bandpass at the output is not part of the sampler itself, but of the following signal processing stage.

voltage referred to the sampler input is

$$V_{n,th}^2 = 8k_B T Z_0 B .$$

The treatment of the shot noise is difficult, since it is not known to what extent one diode also downconverts the noise of the other diode. In order to calculate the distribution of the noise currents between the two branches, their individual impedances must be known. But the diodes are driven into heavy large signal operation by the sampling pulses, so these impedances are difficult to define. If the branches are assumed independent, the shot noise current source can be transformed into a voltage source with $\sqrt{2Q_e\hat{I}_dB}/(2\hat{g}_d)$ at the input (where also the thermal noise sources are situated). The factor 1/2 is included because the sampler model assumes identical voltages in both branches; adding these results in doubling the voltage found in one branch (see Figure 2.2), which must be compensated for. If the branches are assumed tightly coupled, one diode also mixes down the noise of the other, and the equivalent input voltage noise source has the value $\sqrt{2Q_e\hat{I}_dB}/\hat{g}_d$ (now both branches contribute the same noise). However, until now only one diode was taken into account. The other branches' diode also contributes independent noise, so the resulting noise powers have to be doubled. For a worst-case estimate ('tight coupling'), the squared noise voltage due to shot noise is

$$V_{n,sh}^2 = 4Q_e\hat{I}_dB/\hat{g}_d^2 = 4k_B^2 T^2 B/(\hat{I}_d Q_e) .$$

For the adder, an opamp with noise densities $v_{n,op}$ and $i_{n,op}$ in an inverting adder circuit is assumed. The circuit has unity gain, so all resistors are equal to R_l . The adder contributes the squared noise voltage

$$V_{n,add}^2 = 12k_B T B R_l + R_l^2 B i_{n,op}^2 + 9B v_{n,op}^2 .$$

The sampling operation causes an in theory infinitely number of harmonics of f_s , where noise is present but no signal, to be sampled down. The contributions are weighted by the sampler's frequency response at the particular

harmonic. Adding all parts together yields

$$V_{n,o}^2 = V_{n,\text{add}}^2 + \sum_{n=-\infty}^{\infty} (V_{n,\text{th}}^2 + V_{n,\text{sh}}^2) |H_s(nf_s + f_o)|^2 ;$$

no summing occurs over the adder's noise, since it is placed after the sampling diodes and contributes only 'baseband noise'.

Since $H_s(f)$ describes a 'real' system (in both senses), $|H_s(-f)| = |H_s^*(f)| = |H_s(f)|$. By using $f_o \ll f_s$ the approximations

$$\begin{aligned} |H_s(nf_s + f_o)| &\approx |H_s(nf_s)| \quad \text{and} \\ |H_s(-nf_s + f_o)| &= |H_s(nf_s - f_o)| \approx |H_s(nf_s)| \end{aligned}$$

are valid. With some intermediate input frequency \bar{f}_i the *noise-effective number of harmonics*

$$N_{\text{eff}} := \sum_{n=1}^{\infty} \frac{|H_s(nf_s)|^2}{|H_s(\bar{f}_i)|^2}$$

can be defined. The noise at the output then becomes

$$V_{n,o}^2 \approx V_{n,\text{add}}^2 + (V_{n,\text{th}}^2 + V_{n,\text{sh}}^2) 2N_{\text{eff}} |H_s(\bar{f}_i)|^2 .$$

Note that in the equation above the sum term for $n = 0$ was dropped; since $f_o \ll f_s$, the contribution of $|H_s(f_o)|$ is indeed negligible.

It should be stressed that N_{eff} , and therefore the sampler's noise figure, heavily depend on f_s . Assuming $|H_s(f)|$ constant up to some frequency \hat{f}_i and immediately dropping to negligible magnitude above,

$$N_{\text{eff}} = \lfloor \hat{f}_i / f_s \rfloor \approx \hat{f}_i / f_s .$$

The noise contributed by the adder can be reduced by decreasing R_1 . But the reduction is limited by the conversion gain: The smaller R_1 , the faster the hold capacitor is discharged, causing H_s to drop.

2.5 PULSE GENERATION

As indicated in Section 2.2, a differential pair is used to generate the sharp edges required to drive the sampling diodes. It is driven by a Schmitt trigger gate, used to get definite pulse times from a sinusoidal input signal. The gate's output pulses have a rise time of $1 \dots 2$ ns. C_1 differentiates the signal for a fast transistor turn-on, R_1 provides a DC path for the base currents and clears the base-emitter junction charge. The right half of the pair is identical to the left, maintaining the symmetry of the circuit. To put the

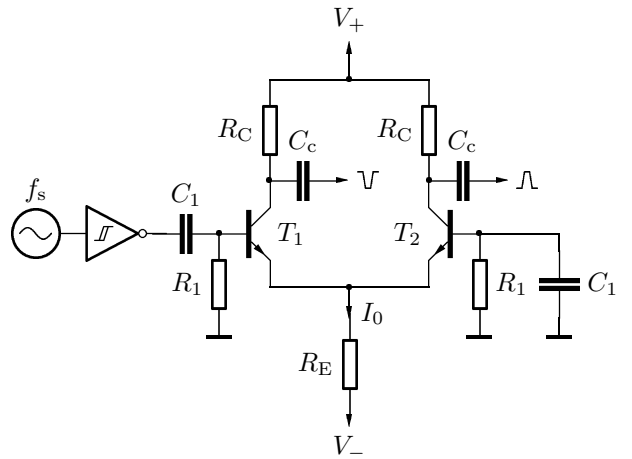


Figure 2.5 Sampling pulse generation.

differential pair into a defined state, a small positive bias voltage is applied to the base of the right transistor in the real circuit. The two symmetrical pulses are coupled out via capacitors at the collectors of the two transistors. The quiescent current, which flows completely through either the left or right transistor, is

$$I_0 = \frac{V_- - V_{BE}}{R_E}.$$

The transistor type used is the BFP450, a SiGe heterojunction bipolar transistor with a transit frequency of 24 GHz and a maximum collector current of 100 mA. As usual with HBTs the collector-emitter breakdown voltage is quite low (4.5 V); therefore, $V_+ = 3.3$ V was chosen. V_- was set to -5.0 V, since it was expected that this would be one of the standard power supply voltages in the VNA.

In order too see if the circuit is capable of delivering signal edges sharp enough, and—equally important—if it can be made stable, i.e. not self-oscillating, a small test circuit with $I_0 = 10$ mA was simulated and built. It turned out that resistors of approximately $10\ \Omega$ in series with the transistor's bases were necessary to stabilize the circuit and damp ringing. The rise time of the collector currents obtained by this circuit was about 200 ps.

To choose the impedance of the transmission line stubs, the equivalent circuit in Figure 2.6 was used. The transistor is represented by a current source; all parasitics are neglected, otherwise it would not be possible to give an analytically solvable model.[#] The coupling capacitor C_c is assumed to be so large that it can be essentially left out; how large it actually has to be for this to be valid will be calculated later. All parasitics in the diode are neglected for the same reasons as with the transistor.

The diodes used are Schottky types from the BAT15 series. Since switching speed is an issue, Schottky diodes are customary in sampler circuits. The diodes are modelled by the simple Shockley relation for forward bias

$$I_d = I_s e^{\frac{V_d}{mV_T}},$$

where I_s and m are fit parameters. They were determined by measuring four diodes (two devices with two junctions each) and applying a least-squares fit algorithm, illustrated in Figure 2.7.

At $t = 0$ the transistor switches on, $I_C = I_0\sigma(t)$. For $t = 0^+$ a wave with voltage \hat{V}_d and current amplitude \hat{V}_d/Z_w starts to propagate towards the end

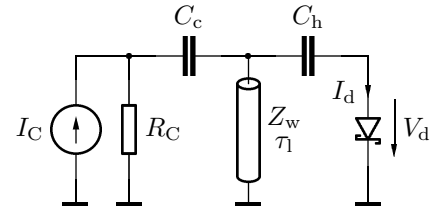


Figure 2.6 Equivalent circuit of one sampler branch with respect to its sampling pulse.

[#]In fact, an Early conductance could be included in the model, since it can easily be combined with R_C . On the other hand, especially with high-frequency heterojunction transistors, the approximation of the sloped output characteristics of the transistor with a constant conductance is not very accurate, especially under the present dynamic large-signal conditions.

of the transmission line. Kirchhoff's current law gives

$$I_0 = \frac{\hat{V}_d}{R_C} + \frac{\hat{V}_d}{Z_w} + \hat{I}_d, \quad t = 0^+.$$

The wave is reflected at the line's shorted end at $t = \tau_l$, and a wave with voltage $-\hat{V}_d$ and current \hat{V}_d/Z_w travels to the line's beginning, where it is superimposed on the circuit currents to yield the primed quantities

$$I_0 = \frac{V'_d}{R_C} + 2\frac{\hat{V}_d}{Z_w} + I'_d, \quad t = (2\tau_l)^+.$$

The diode is to be closed at this point, so $V'_d = 0$ and $I'_d = 0$ must be fulfilled. From this, the last equation gives

$$\hat{V}_d = \frac{I_0 Z_w}{2}$$

whereas the first equation results in

$$\hat{I}_d = \frac{I_0}{2}(1 - Z_w/R_C).$$

Inserting one into another, the quite intuitive formula

$$Z_w = R_C \parallel \frac{\hat{V}_d}{\hat{I}_d}$$

can be derived: The initial current wave \hat{V}_d/Z_w on the transmission line must, after reflection, completely cancel the currents in the collector resistor *and* the diode. To achieve this, the wave impedance must be equal to R_C in parallel with the 'diode resistance' \hat{V}_d/\hat{I}_d .

Presently, there are four unknown quantities (I_0 , \hat{I}_d , Z_w , and R_C , since \hat{V}_d and \hat{I}_d are related by the diode's IV -curve), but only two equations, leaving us with two degrees of freedom. To narrow it down a little bit more, two additional boundary conditions can be formulated: First, the collector resistance R_C should be not so large that the presently conducting transistor of the differential pair is driven into saturation. With the base-emitter voltage V_{BE} , the differential pair's positive supply voltage V_+ , and the transistor's saturation voltage V_{CEsat} the condition $V_{CE} \geq V_{CEsat}$ gives

$$R_C \leq \frac{V_+ + V_{BE} - V_{CEsat}}{I_0}.$$

The second condition comes from experimental results with the differential amplifier. I observed that for load resistances smaller than about $30\,\Omega$ the

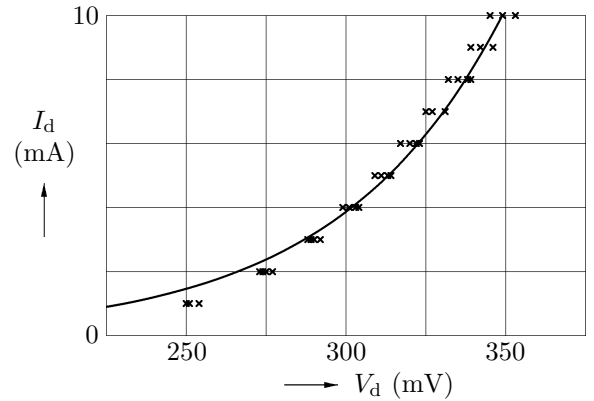


Figure 2.7 Measured static diode characteristic. The dots mark measured points, the trace is the least-squares fit. The parameters are $I_s = 11.3\,\mu\text{A}$ and $m = 1.99$ ($T = 300\,\text{K}$).

differential amplifier is unstable, even with high ($\sim 60 \Omega$) resistances in series with the base. In this context, ‘load resistance’ refers to the overall resistance the transistor sees at its collector, all ‘real’ load resistors in parallel with R_C . For the circuit above, this equivalent load resistance is easily calculated: Since at $t = 0^+$ the current into the transmission line on one hand and the combined current through R_C and the diode on the other must be equal,

$$R_{\text{eq}} = \frac{Z_w}{2} \geq \check{R}_{\text{eq}} = 30 \Omega .$$

Since the peak diode current \hat{I}_d directly affects the sampler’s frequency response (cf. Section 2.3), it is natural to select it as an independent variable; one other quantity can be chosen freely. For Z_w and I_0 the constraints

$$\begin{cases} Z_w \geq 2\check{R}_{\text{eq}} \\ Z_w \leq \frac{\hat{V}_d}{\hat{I}_d} \left(1 - \frac{2\hat{V}_d}{V_+ + V_{\text{BE}} - V_{\text{CEsat}}} \right) \\ I_0 \geq 2\hat{I}_d \left(1 - \frac{2\hat{V}_d}{V_+ + V_{\text{BE}} - V_{\text{CEsat}}} \right)^{-1} \\ I_0 \leq \frac{\hat{V}_d}{\check{R}_{\text{eq}}} \end{cases}$$

drawn in Figure 2.8 must be obeyed.

After choosing the currents, lower bounds for the capacitors C_c and C_h can be given for the above considerations to stay valid. The hold capacitor carries the diode current \hat{I}_d for a period of $\tau_s = 2\tau_l$. The charge $\hat{I}_d\tau_s$ brought onto the capacitor causes the diode voltage to drop by

$$\Delta V_d = \frac{\hat{I}_d\tau_s}{C_h} .$$

This voltage drop should be small compared to V_d for the diode current to remain approximately constant over the sampling time duration, ensuring a good cancellation of current pulses after $t = \tau_s$. On the other hand, if C_h is made too large the conversion gain of the sampler drops because of the lowpass C_h forms with R_l .

A similar consideration holds for C_c , but on a slower timescale: After $2\tau_l$, the diode current \hat{I}_d is cancelled, but the transmission line must still be fed with I_0 . C_c must therefore be much larger than C_h .

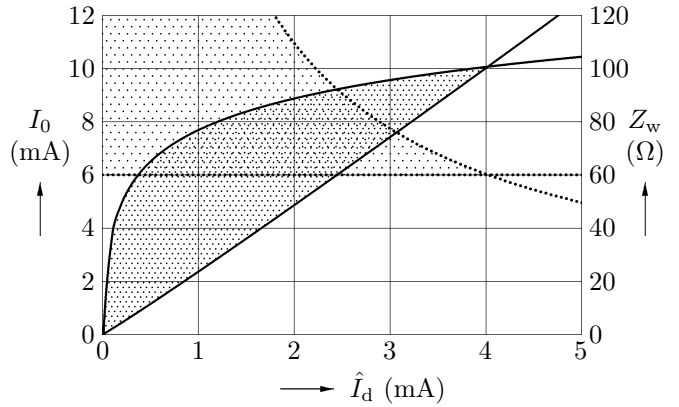


Figure 2.8 Valid ranges for the current pulse’s amplitude I_0 (solid lines, left axis) and the line impedance Z_w (dotted lines, right axis). Parameters used are $V_+ = 3.3 \text{ V}$, $V_{\text{BE}} = 0.7 \text{ V}$, $V_{\text{CEsat}} = 1 \text{ V}$, and $\check{R}_{\text{eq}} = 30 \Omega$.

2.6 RESULTS

With the sampler three design cycles were carried out. The first version was unusable due to a major design flaw: The short circuit stubs and the sampling diodes were designed with much too low impedance, violating the lower bound on the transistor's equivalent load resistance. This caused severe ringing, and in turn a really weird frequency response.

The second version, otherwise completely identical to the first, used stubs with $Z_w = 58.5 \Omega$, a differential amplifier quiescent current $I_0 = 9.5 \text{ mA}$ and a peak diode current $I_d = 3.5 \text{ mA}$. Initially the stubs were made to a propagation time $\tau_1 = 250 \text{ ps}$, and were subsequently shortened while observing the frequency response of the whole sampler. The adder was built around an AD8054A, which was found to have a not particularly good noise performance ($v_n = 16 \text{ nV}/\sqrt{\text{Hz}}$ and $i_n = 900 \text{ fA}/\sqrt{\text{Hz}}$), resulting in -125 dBm/Hz of noise power added by the amplifier compared to approximately -140 dBm/Hz which the sampler would produce *if it had no conversion loss*. In addition, to eventually aid in speeding up the shut down of the sampling diodes, the possibility for reverse biasing was added (voltage from $100 \text{ k}\Omega$ potentiometers connected to the appropriate supply voltage and ground, fed to the circuit where the output voltages from the hold capacitors are taken by $1 \text{ M}\Omega$ resistors). It was found that this was of no use, probably because of its high equivalent source resistance. Lowering the resistance on the other hand is expected to have a deteriorating effect on the frequency response by contributing to the load presented to the sampling capacitors. Extensive tests (by permutating hold capacitor-, coupling capacitor-, collector resistance-, quiescent current-, base resistance-, and input- RC -circuit-values and attempting to measure the sampling pulse) led to the conclusion that the ringing and strong overshoot degrade the performance of this sampler. These effects could be reduced by changing the base resistors from 10Ω to 68Ω , which slows down the switching process, and by changing the collector resistors from the theoretical 220Ω to 82Ω , questioning the validity of the considerations regarding transmission line impedance.

In the third version, the reverse-bias circuitry was omitted, and I tried to improve on the pulse shape by increasing the equivalent load resistance of the differential pair. This version was manufactured in two flavors, one with $Z_w = 67 \Omega$, $I_d = 3 \text{ mA}$, and $I_0 = 8.6 \text{ mA}$, and one with $Z_w = 100 \Omega$, $I_d = 1.2 \text{ mA}$, and $I_0 = 4.8 \text{ mA}$. The latter version, though, did not perform better than the former, showing the same frequency response and dynamic range, with a conversion loss increased by 2 dB . Both had a slightly flatter frequency response than the second version described in the previous paragraph, at the cost of an even higher conversion loss. Apart from that, the third generation uses an AD8651 instead of the AD8054, which has exceptionally low noise ($v_n = 5 \text{ nV}/\sqrt{\text{Hz}}$ and $i_n = 4 \text{ fA}/\sqrt{\text{Hz}}$). Together with lowering the impedance level of the opamp's feedback by a factor of 100, the noise contributed by the

amplifier drops to -142 dBm/Hz. The stubs are now 90 ps long, resulting in a sampling pulse duration $\tau_s = 180$ ps. This value was found in the experiments with the second version.

The sampler works up to a sampling frequency of 310 MHz, although with $f_s > 50$ MHz the frequency response starts to develop considerable ripple. This is due to the effect of sampling pulse echoes, as discussed in Section 2.3. The sampling signal must have at least 10 dBm; the measurements presented below were all taken with $P_s = 13$ dBm and $\hat{f}_s = 50$ MHz.

Frequency Response

The frequency response was obtained by varying the input frequency in steps of 10 MHz and measuring the output power with a spectrum analyzer in zero-span mode. To increase accuracy by lowering measurement noise, the resolution bandwidth was set to 10 Hz, which made it necessary to cheat at certain input frequencies: $f_i = 120$ MHz for example would require $n = 3$ and $f_s = 36.6$ MHz. Assuming a resolution of the sampling signal source of 1 Hz, setting $f'_s = 36\,666\,667$ Hz does the job. However, the deviation resulting from rounding f_s is multiplied with n , causing problems at higher harmonics: at $f_i = 1140$ MHz, $n = 23$ and $f_s = 49\,130\,434.78 \dots$ Hz; rounding gives $f'_s = 49\,130\,435$ Hz. The input signal is therefore converted to $9\,999\,995$ Hz, which is at the shoulder of the spectrum analyzer's resolution filter. The input source was therefore adjusted to $nf'_s + f_o$.

The sampler's conversion loss was much greater than predicted from the theory. This can be explained by assuming a much lower diode current than originally used in the design. If the model from Section 2.3 is evaluated with $\hat{I}_d = 55$ μ A and an echo amplitude of $a_2 = 0.1$, it comes close to the measurements.

The tiny diode current may be due to parasitic inductances in the diodes and capacitors, or to the diode's junction and diffusion capacitances. Another explanation is that \hat{I}_d is in fact only a means of modelling g_d . The actual diode current may have the assumed value, but the diode's small signal conductance may be influenced by the parasitics, and these influences can only enter the

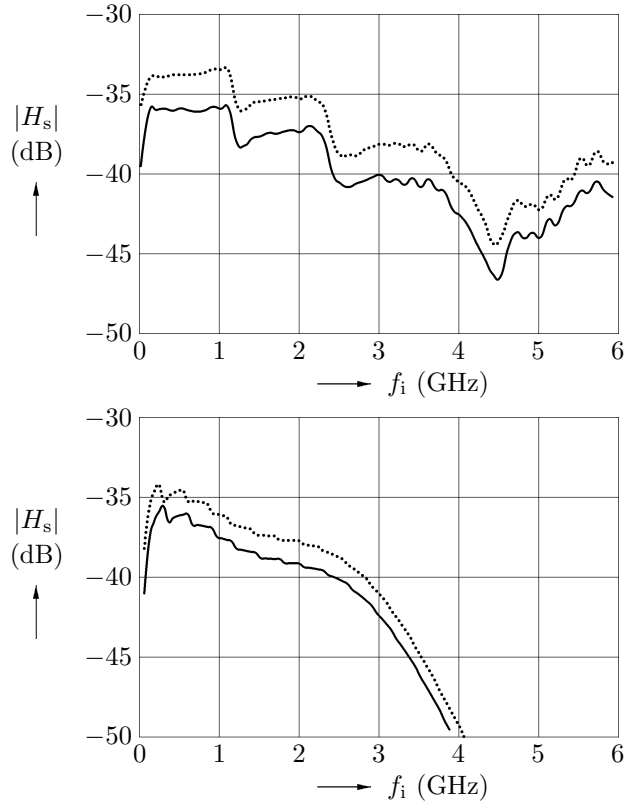


Figure 2.9 Measured frequency response (above) and theoretical response with $\hat{I}_d = 55$ μ A and $a_2 = 0.1$ (below); $f_o = 10$ MHz (solid trace) or $f_o = 1$ MHz (dotted trace) in both graphs.

model via \hat{I}_d . The figure also shows that lowering f_o decreases the conversion loss, an effect that could not be observed in Figure 2.3.

The bad conversion of the sampler is alleviated by the fact that it in principle does not influence the measurement of the VNA as long as the same bad conversion is used both in calibration and in the actual measurement process. In practice, however, a high conversion loss impairs the dynamic range of the network analyzer, because the noise floor keeps constant. Improving the conversion by adding a baseband amplifier or changing the feedback of the adder to result in a gain greater unity does no good, because the noise is amplified, too—even more since a good deal of noise comes from the opamp adder itself.

Dynamic Range

Figure 2.10 shows a sweep of the sampler input power. Such sweeps were done for different input frequencies; apart from the regular dependence of sampler gain shown in the previous figure, the traces are equal.

An interesting observation was made: At 10 MHz a coherent carrier is observable at the spectrum analyzer when a cable is connected to its input, even if the cable is open at its other end. For this reason, when decreasing the input power the dotted trace in Figure 2.10 flattens far earlier than the solid trace, which was measured at 1 MHz. The analyzers noise floor at this measurement was -144 dBm, so no separate noise figure measurement was necessary.

The noise floor of the sampler is slightly below -130 dBm, measured with $B_r = 10$ Hz, which matches the prediction of -142 dBm/Hz. Compression, however, starts quite early at a 1 dB-compression point $P_{1\text{dB}} = -20$ dBm. If allowing for a tolerance of 1 dB also at the lower end, we are left with a dynamic range of only $-80 \dots -20$ dBm.

I believe that the high conversion loss and the early compression have a common cause: If the sampling signal at the diodes is weaker than expected, a lower RF level is sufficient to cause considerable intermodulation, putting the sampler into compression.

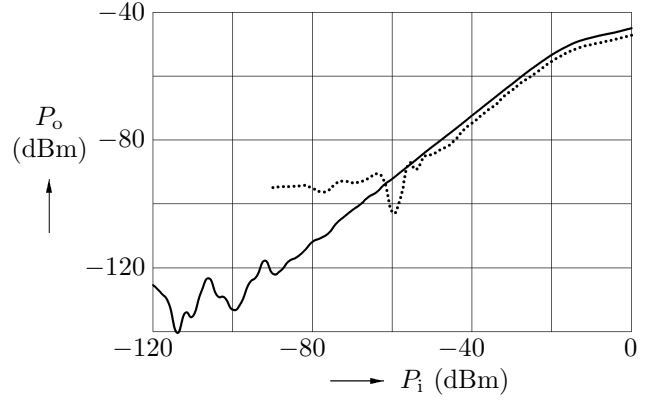


Figure 2.10 Power sweep; $f_i = 1$ GHz, $f_o = 10$ MHz (dotted trace) or $f_o = 1$ MHz (solid trace), resolution bandwidth 10 Hz.

Return Loss

The input reflection coefficient was measured using a VNA on a ‘cold’ sampler, i.e. a sampler with the power supply connected, but no sampling signal applied. If the sampler is set into operation, the reflection coefficient does not change except for a few discrete frequencies, where the sampler kick-out disturbs the VNA. The incident power was -20 dBm, a lower power level resulted in the same picture.

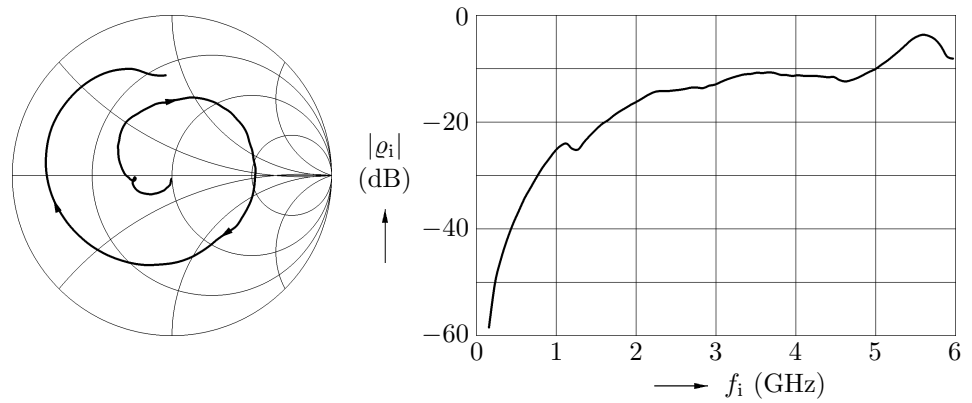


Figure 2.11 Input reflection coefficient at the sampler’s RF input.

Kick-out

The sampler kick-out was measured by connecting a spectrum analyzer to the RF input while in operation. The strong kick-out and the bad input match indicate that an isolating amplifier is indeed necessary when including the sampler in a measurement setup.

Of course the kick-out spectrum consists of many distinct carriers at multiples of f_s . For the diagram, the peaks were connected to mimic a continuous spectrum.

Kick-out is caused by imbalances, and while I tried to build a highly symmetric circuit, there are still quite a few sources of asymmetry. First of all, the differential amplifier is driven asymmetric. One transistor is directly driven on its base, whereas the second is driven on its emitter by a signal that had to propagate through the first one. Another thing to be considered is the fact that one transistor conducts the quiescent current I_0 most of the time, while the other transistor conducts only for very short periods. The difference in average dissipated power results in a different chip temperature, adding to the imbalance of the circuit. The next measurement tries to investigate more on the symmetry of the circuit.

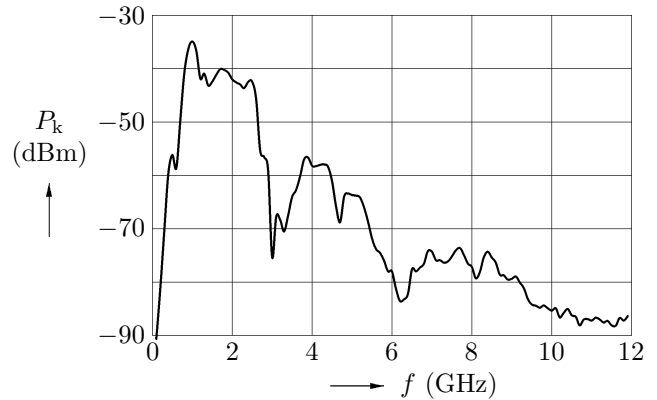


Figure 2.12 Sampler kick-out at $f_s = 50$ MHz.

Symmetry

To test the symmetry, the individual contributions of the sampler branches to the output signal were measured by breaking one of the adder's inputs just before the inverting input and connecting the respective resistor to ground to maintain all impedance levels (actually, the right side of either the capacitor C_{20} or C_{21} in the schematic in Figure A.1 was grounded).

The upper solid trace shows the signal contributed by the 'direct' diode, i.e. the diode connected to the branch of the differential amplifier driven by the Schmitt trigger; the lower solid trace the signal obtained by the 'indirect' diode. The dotted trace shows the summed output for comparison. Apart from the fact that most of the output signal is contributed from the 'direct' sampler branch, demonstrating bad balance between the sampling pulse amplitudes, the sum shows a dip at 4.47 GHz, which the two individual portions do not show. This is probably caused by the time skew between the two sampling pulses, which causes destructive addition of the two sampler branches' signals at that particular frequency. This cancellation happens when the time skew is equal to half the cycle time of the input signal, so the time skew is approximately

$$\Delta t_s = \frac{1}{2 \cdot 4.47 \text{ GHz}} = 112 \text{ ps} .$$

Sampling Pulse

Finally, I tried to record the sampling pulses themselves by soldering two 910Ω resistors (case 0402) to the beginning of the microstrip stubs, right where C_{15} and C_{16} (C_{14} and C_{17} for the other branch) meet, cf. figure-fig:schemsamplerfinal. Two coaxial 50Ω lines with SMA connectors at the other ends were soldered to the resistors, forming a 19.2 : 1 probe. The signals were picked up by a 20 GHz sampling scope, where the different line lengths were electronically compensated.

The solid trace shows the 'direct' branch's voltage (at C_{15}), the dotted the 'indirect' branch's (C_{14}). They differ greatly in ampli-

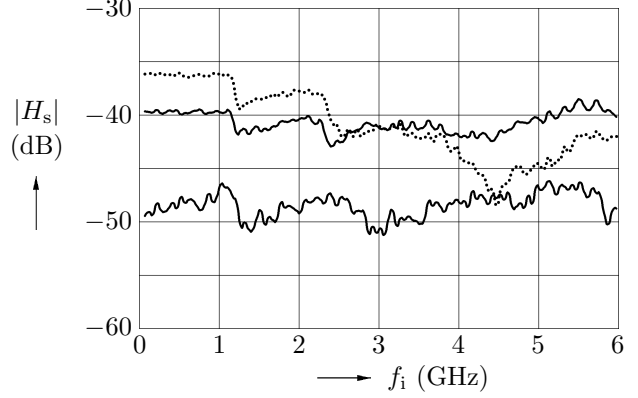


Figure 2.13 Individual contributions of the sampler branches (solid) to the output signal (dotted).

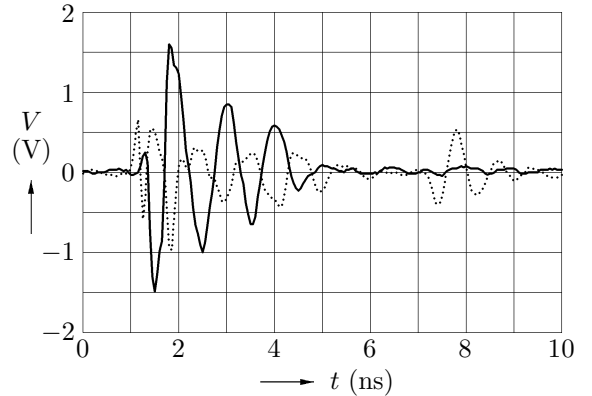


Figure 2.14 Sampling pulse of the 'direct' diode (solid) and the 'indirect' diode (dotted).

tude, explaining the largely differing contributions of the sampler branches above. The voltages also show severe ringing, and the quite surprising fact that the edge of the indirect branch comes *before* the direct one. While the latter can only be explained by a faulty line length compensation, I suspect the former is more of an artefact of the less than ideal signal pickup; e.g. it was not possible to solder the SMA-cables directly to the resistors, instead a short (a few millimeters) thin wire had to be used, adding series inductance. The waveforms in Figure 2.14 should therefore be taken with a grain of salt. However, if they resembled reality, the fact that in the direct branch voltages far above the Schottky diode's forward voltage occur would indicate the presence of considerable parasitic series inductances. Impeding a fast current turn-on, those inductances may be the cause for the high conversion loss of the sampler.

2.7 IMPROVEMENTS

Some ideas for improving the sampler's performance are discussed at last. They came to my mind while experimenting with the samplers, but lack of time did not allow to implement them.

Faster devices

The BFP450 has a faster ($f_t = 30$ GHz instead of 24 GHz) cousin, the BFP540F, which has been available through small volume distributors only recently. The transistor also has a smaller case (TSFP-4 instead of SOT343), lowering the parasitic inductances. The use of the BFP540F would allow for sharper edges of the sampling pulses, enabling shorter pulse widths while keeping the pulse amplitude constant. This in turn leads to a higher cut-off frequency. On the other hand a faster transistor has a higher tendency to self-oscillation. Building the circuit with the BFP540F may introduce heavy troubles with stability.

The Schottky diodes may be replaced by the variant BAT15-099LRH. Employing the same chip with the same electrical characteristics, the smaller case (TSLP-4-7 instead of SOT143 with the BAT15-099) incorporates smaller inductance.

Even the capacitors are candidates for improvement: Using case type 0402 instead of 0603 where possible (capacitance values up to 27 pF with ATC's 600L series) also should result in an enhanced frequency response of the whole sampler.

Two-stage design

By introducing an additional driver stage, the two duties of the differential amplifier are separated. The inner transistors produce two symmetrical pulses, the outer transistors drive the sampling diodes. The reduced load to the differential amplifier is expected to improve the pulse shape and the symmetry of the two pulses, especially with respect to their time skew.

Positive feedback

As in a Schmitt trigger, adding feedback from one transistor's collector to the other's base and vice versa accelerates the transition process from one state of the differential amplifier to the other. It may be even possible to omit the dedicated CMOS Schmitt trigger gate and subsume its function into the differential amplifier. On the other hand, being blissful about the circuit not self-oscillating in its current configuration, some positive feedback may be an audacious attempt.

CHAPTER 3

Synthesizer

Zwaa Hertz
— Anonymous

A VNA is a broadband measurement system, and therefore calls for a broadband signal source. Commercial instruments employ *YIG-Oscillators*, where a small monocrystalline pellet made of yttrium-iron-garnet is excited by a magnetic RF field, causing the precession of the electron spins. If the field's frequency agrees with the natural precession frequency in the crystal, a resonance effect is observed. The natural precession frequency can be influenced by applying a static magnetic field perpendicular to the RF field, turning the YIG arrangement into a tunable bandpass filter with an exceptional quality factor ($\sim 10^4$) and very wide tuning range (a few hundred megahertz up to some ten gigahertz). When used as the frequency determining device in an oscillator, the YIG resonator's high quality factor ensures low phase noise.

Unfortunately, YIG pellets are rather special devices, and I was forced to fall back on a more ordinary oscillator design using a varactor diode. These devices exploit the voltage-dependent junction capacitance of a reverse-biased pn-junction. The maximum capacitance variation (between zero bias and just before reverse breakdown) is typically $\hat{C}/\check{C} \approx 10$. Since in a simple resonator $f_0 \sim 1/\sqrt{C}$, in theory an oscillator with a varactor may be tuned over a relative range of $\hat{f}/\check{f} \approx 3$. However, the rest of the oscillator circuit also includes frequency dependent reactances, some wanted and some parasitic, restricting the practically achievable tuning range to $\hat{f}/\check{f} \approx 2$ (cf. [Alpha APN1006]).[#]

[#]This can be argued using Foster's Reactance Theorem: By lumping together all devices in the oscillator apart from the varactor into an impedance $Z_r = R_r + jX_r$, which must have $X_r > 0$, the inequality $dX_r/df \geq X_r/f$ states that the varactor's counterpart changes its reactance at least as fast with frequency as a pure inductor. For a particular change in the varactor's reactance, the necessary change in frequency to reach equilibrium again is therefore smaller or equal the change in frequency the circuit would show if Z_r were a pure inductor. More precisely, at resonance the equation $2\pi f_0 C_v X_r(f_0) = 1$ holds. Taking the derivative with respect to C_v and using $dX_r/dC_v = (dX_r/df_0)(df_0/dC_v)$ yields the relative change rate

$$\frac{df_0/f_0}{dC_v/C_v} = -\frac{1}{1 + \frac{dX_r/X_r}{df_0/f_0}}.$$

For this reason, it is not possible to span a broad frequency range with only one single varactor-tuned oscillator.

Two remedies for this situation are frequency multiplication and frequency conversion. While in a frequency multiplier a nonlinear device generates harmonics of a single signal, in a mixer one frequency is transposed by another signal's frequency. Frequency multiplication leaves the relative variation \hat{f}/\check{f} constant, whereas upconversion reduces it and downconversion raises it. In a mixer, the phase noise of the output signal is the sum of the phase noise of the two input signals, whereas a frequency multiplier also multiplies the phase noise.

A simple, yet very effective approach to generate frequencies in the range $\check{f} < f < \hat{f}$ is to mix a signal within $\check{f} + \hat{f} < f_1 < 2\hat{f}$ with \hat{f} and take the lower sideband. The oscillator for f_1 must accomplish $\hat{f}_1/\check{f}_1 = 2/(1 + \check{f}/\hat{f}) < 2$, and the upper sideband after the mixer has its lowest component at $2\hat{f}$, which makes filtering out easily. The lowest attainable frequency is mainly determined by the mixer's output characteristics. The drawback of this approach is that signals with twice the required output frequency are present in the circuit. This may necessitate a transition to a more sophisticated (and more expensive) technology. For precisely this reason, I took another approach.

The variable-frequency source in Figure 3.1 generates frequencies $\check{f}_v \leq f_v \leq \hat{f}_v$, while the fixed source oscillates at f_f . Through a nonlinear device a number of harmonics of f_f are generated, and a particular harmonic $n f_f$ is picked by the first tunable bandpass. The two frequencies are mixed, and either the upper or the lower sideband is selected by the bandpass at the output. This arrangement clearly covers a contiguous frequency

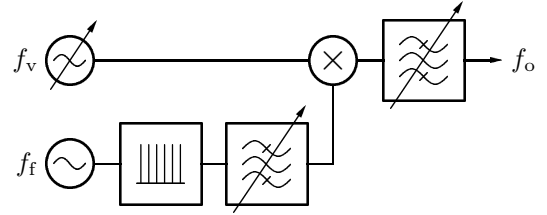


Figure 3.1 Block diagram of the broadband source.

band if $\hat{f}_v - \check{f}_v \geq f_f$; but in fact, by making clever use of both sidebands after the mixer, $\hat{f}_v - \check{f}_v \geq f_f/2$ is sufficient. For example, the first approach to the source used $1 \text{ GHz} \leq f_v \leq 1.5 \text{ GHz}$ and $f_f = 1 \text{ GHz}$. By generating at least three harmonics of f_f , the required frequency range can be covered by putting together the subbands as shown in Table 3.1 on the left.[#]

Later I changed the setup to $2 \text{ GHz} \leq f_v \leq 3 \text{ GHz}$ and $f_f = 1 \text{ GHz}$, because it incorporates the same relative variation of f_v but splits the range of f_o into only three subbands, reducing both the number of harmonics and

For a simple inductor $(dX_r/X_r)/(df_0/f_0) \equiv 1$, whereas for a more complicated arrangement Foster's Reactance Theorem gives $(dX_r/X_r)/(df_0/f_0) \geq 1$, effectively reducing the relative tuning sensitivity. Since this happens over the whole tuning range, the overall range is reduced, too.

[#]This is one of the rare instances where the existence of negative frequencies comes in handy: Using a lower sideband is nothing else than using an upper sideband, but with a negative frequency at one of the two mixer inputs. Since, taking negative frequencies into account, f_v covers a segment of length $2(\hat{f}_v - \check{f}_v)$ on the frequency axis, the harmonics of f_f must only be spaced apart that far.

subband	generated by		subband	generated by	
	n	sideband		n	sideband
0.0 ... 0.5 GHz	1	lower	0 ... 1 GHz	2	lower
0.5 ... 1.0 GHz	2	lower	1 ... 2 GHz	1	lower
1.0 ... 1.5 GHz	$(f_v \text{ used directly})$		2 ... 3 GHz	$(f_v \text{ used directly})$	
1.5 ... 2.0 GHz	3	lower			
2.0 ... 2.5 GHz	1	upper			
2.5 ... 3.0 GHz	4	lower			
(3.0 ... 3.5 GHz	2	upper)			

Table 3.1 Possible splitting of the frequency range 0 ... 3 GHz with a variable frequency oscillator between 1 GHz and 1.5 GHz (left) or 2 GHz and 3 GHz (right). The fixed oscillator has $f_f = 1$ GHz in both cases, n is the index of the harmonic used. The last line in the left table shows a subband not necessary to fulfill the specifications, but available virtually free of charge.

the number of bandpass filters. Only lower sidebands are used with the new configuration, and the upper sidebands are spaced farther apart, easing the requirements on the bandpasses' skirt factors.

To ensure accurate control over the frequency, frequency stability and low phase noise, the two sources that generate f_v and f_f are phase-locked loops. This chapter describes their design. In the last section, a few considerations on the other blocks are presented.

3.1 RF OSCILLATORS — THEORY AND PRACTICE

Rumour has it that circuits initially termed ‘amplifiers’ often turn out as good oscillators. Sometimes, however, circuits do not oscillate, even when designed with the intention to do so. This section tries to describe the problems encountered with the design of the oscillators for the synthesizer module.

Two and a Half Views on the Same Thing

Oscillators are autonomous systems. They oscillate on their own, requiring only a power supply as ‘input’. However, they were switched on some time in the past, where all voltages and currents in the circuit were zero. After switch-on, if the oscillator was designed properly, all voltages and currents start to oscillate with steadily rising amplitudes. Let V_i and I_i denote the complex voltage and current amplitudes in the i -th branch of the circuit,

$$V_i = \hat{V}_i e^{st} , \quad I_i = \hat{I}_i e^{st} , \quad s = \alpha + j\omega , \quad \alpha \geq 0 .$$

Assuming that every branch of the circuit consists either of a resistor, an inductor, or a capacitor with the constitutive equations $V_i = R_i I_i$, $V_i = s L_i I_i$, and $I_i = s C_i V_i$ respectively, the principle of conservation of complex power gives

$$\sum_i V_i I_i^* = \sum_{\text{resistors}} R_i |I_i|^2 + \sum_{\text{inductors}} s L_i |I_i|^2 + \sum_{\text{capacitors}} s^* C_i |V_i|^2 = 0 ,$$

or by splitting in real and imaginary part

$$\sum R_i |I_i|^2 + \alpha \left(\sum L_i |I_i|^2 + \sum C_i |V_i|^2 \right) = 0$$

and

$$\omega \left(\sum L_i |I_i|^2 - \sum C_i |V_i|^2 \right) = 0 .$$

If the circuit is to start up oscillation, $\alpha > 0$; but since $|I_i| > 0$ and $|V_i| > 0$, this is only possible if at least for one branch $R_i < 0$. If $R_i \equiv 0$ it follows that $\alpha = 0$, and the circuit is in the state of steady oscillation. The frequency of oscillation is controlled by the equation incorporating the imaginary parts. Note that the assumption of ‘branch impedances’ is not really intuitive if devices with more than two terminals are present in the circuit. Usually these are represented with equivalent circuits incorporating controlled sources. Suppose branch j consists of a current source controlled by the voltage in branch k , together with a source resistance of R_j in parallel; branch k is assumed to contain a resistor, too (think of a transistor). Then $V_j = R_j(I_j - g_m V_k) = R_j(I_j - g_m R_k I_k)$, and the power balance gives

$$\sum V_i I_i^* = \sum R_i |I_i|^2 + \sum s L_i |I_i|^2 + \sum s^* C_i |V_i|^2 - R_j R_k g_m I_k I_j^* = 0 .$$

Taking the real part,

$$\sum R_i |I_i|^2 + \alpha \left(\sum L_i |I_i|^2 + \sum C_i |V_i|^2 \right) - R_j R_k g_m |I_k| |I_j| \cos(\arg I_k - \arg I_j) = 0 .$$

Now the minus-sign and the cosine-term make nonzero current amplitudes possible. Thus, every oscillator requires an active device, either a two-terminal device with a negative resistance, i.e. a device whose IV -curve has a region with negative slope (e.g. tunnel diodes, Gunn diodes, or gas discharge lamps), or a device that exhibits gain and can be equivalently described by a controlled source.

From these two possibilities two different views on oscillators arise: The *negative resistance view* and the *feedback view*. Both aim at simplifying the design procedure, which would be quite complicated if all branch currents were to be considered. In the negative resistance view, the active device, its biasing circuitry, and the load resistance are subsumed into a single impedance Z_a with a negative real part. Those parts of the oscillator responsible for frequency determination through their mostly reactive behaviour remain; they are collected into a resonator impedance Z_r . The characteristic equations from the last paragraph are now

$$\operatorname{Re} Z_r + \operatorname{Re} Z_a \leq 0 \quad \text{and} \quad \operatorname{Im} Z_r + \operatorname{Im} Z_a = 0 .$$

With amplifying devices, the feedback view is more intuitive: The active device, together with bias circuitry and load is regarded as a two-port with transfer characteristic G , the rest of the oscillator is represented by a feedback

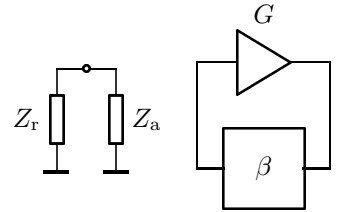


Figure 3.2 Negative resistance view (left) and feedback view (right) on an oscillator.

two-port β . The condition for oscillation is

$$|G| |\beta| \geq 1 \quad \text{and} \quad (\arg G + \arg \beta) \bmod (2\pi) = 0 ,$$

often referred to as the *Barkhausen criterion*. Finally, a reformulation of the negative resistance criterion simply transforms the impedances into reflection coefficients,

$$|\rho_r| |\rho_a| \geq 1 \quad \text{and} \quad \arg \rho_r + \arg \rho_a = 0 .$$

Particularly the first equation is quite demonstrative: The power reflected hence and forth between the two devices increases by $|\rho_r \rho_a|^2 \geq 1$ at each turn-around.

The negative resistance viewpoint is appropriate for oscillators using amplifying devices, too, because amplifiers with positive feedback can exhibit input impedances with negative real part.

In all the criteria stated, oscillation build-up requires inequality, whereas equality occurs for steady oscillation. Clearly, somehow a transition from build-up to steady state must take place. Responsible for this transition is the active device, more precisely its nonlinearities. Both the negative resistance and the gain decline with increasing amplitude, because through nonlinear effects more power is translated into harmonics; this power is missing for the fundamental.

Unfortunately, all of the three criteria stated are problematic. To start with, the phase criteria of the reflection coefficient formulation is only valid for the steady state: For two impedances whose imaginary parts cancel and whose sum of the real parts is negative, the phases of the reflection coefficients in general *do not cancel* (cf. [Chramiec]). Moreover, the requirement on the magnitudes is not necessary, as pointed out by [Jackson]: His example of $Z_r = 60 \Omega$ and $Z_a = -70 \Omega$ shows that an oscillating circuit may have $|\rho_r| |\rho_a| = 0.45 < 1$. The feedback equations are in principle correct, but they are merely *necessary* for oscillation, but not *sufficient*, as pointed out by [Nguyen]. Instead the Nyquist criterion should be used: The *complete* Nyquist plot (i.e. $G\beta$ drawn in the complex plane also for negative frequencies) must encircle the point $+1$ in clockwise direction when f is swept from $-\infty$ to ∞ .[#] The negative resistance oscillation conditions suffer from the same deficiency. The application of the Nyquist criterion to negative resistance oscillators is explained in [Jackson]. A particularly lucid explanation of negative resistance oscillators and the stability issues involved can be found in [Kurokawa73].

[#]It should be noted that in connection with control systems, the Nyquist criterion uses the encirclement of the point -1 in the complex plane. This is because control systems normally employ negative feedback, and therefore the fed back quantity is *subtracted* at the controller, while oscillators use positive feedback. Moreover, the sense of *stability* is reversed: A stable oscillator *oscillates*, i.e. it has poles in the right half plane; a stable control system *does not oscillate*, i.e. it has no pole in the right half plane.

The feedback criterion is often misused in another respect: Dividing the circuit into two blocks and only dealing with their transfer characteristics may lead to the conclusion that the two blocks are independent. In general, this is *not* the case. Instead, they present a load to each other. To make things still more complicated, the input impedance of one block, which is the load impedance of the other, depends on the load impedance of the first block, which is the input impedance of the second; the input impedance of the second, in turn, is a function of the load impedance of the second, at which point we are chasing our own tail. Calculating the two blocks' open circuit voltage transfer may work at low frequencies, where ideal isolation can be achieved through amplifiers with infinite input- and zero output impedance. At high frequencies this clearly does not work. Using scattering parameters (or scattering transmission parameters) makes for a better approach, since their definition assumes some finite source- and load impedances; but in general, the impedances in the closed loop are different from the reference impedance over which the scattering parameters were defined, and hence any predictions about the oscillator are unsound. This issue is addressed by [Randall], who proposes a method using eigenvalues based on scattering parameters. They also provide a quite intuitive interpretation of the Nyquist criterion: Having $|G\beta| > 1$ and $\arg(G\beta) = 0$ leads to amplitudes increasing with time by e^{st} with $s = \alpha + j\omega$ only if $\alpha > 0$ and t actually *increases*. The latter is the case if the system has a positive group delay, $-\mathrm{d}\arg(G\beta)/\mathrm{d}f > 0$. But $|G\beta| > 1$ and $\arg(G\beta) = 0$ mean that the Nyquist plot crosses the real axis *at the right of the point +1*; since $\mathrm{d}\arg(G\beta)/\mathrm{d}f < 0$, the crossing is downwards, a movement that results in a clockwise encirclement of the point +1, indicating oscillation. However, if the phase of $G\beta$ has a positive slope, the real axis is crossed upwards, and +1 is encircled counter-clockwise; on the other hand, in the case of a positive slope the group delay is negative, corresponding to a regression in time. In this case a positive α yields decaying amplitudes.

At microwave frequencies, the negative resistance interpretation is preferred over the feedback view, because the active devices' parasitics often provide sufficient feedback to make dedicated feedback networks obsolete (cf. the introductory sentence to this section). At very high frequencies, two-‘terminal’ devices with negative sloped characteristics may be the only available active devices at all.

Oscillator Simulation

The autonomous nature of oscillators and the inherent necessity to include non-linear effects make oscillator simulation a cumbersome task. Most simulation modes implemented in today's circuit simulators are ‘response-simulations’: The circuit is excited by some idealized signal and the response to that signal is taken as ‘output’. But oscillators do not have an input port. With the feedback view presented above, the problem of the missing input may be tack-

led by opening the loop. This, however, leads to wrong simulations, since it substantially changes the circuit, as described above. Therefore, the negative resistance view is preferred.

Simulating an oscillator with *AC simulation* may give some clues on the circuit's ability to oscillate. However, AC simulation has a number of serious drawbacks: First, the simulation mode uses small-signal equivalent circuits and thus neglects all nonlinear effects. But the inclusion of these is not only necessary to predict the oscillation amplitude (output power), but also to accurately simulate the oscillation frequency.

Transient analysis on the other hand includes all nonlinear effects (at least, all that were included in the models). Moreover, it has the virtue that it is the only simulation that from its principle can deal with autonomous systems. Time domain simulation yields accurate information on the oscillation frequency, the output power, and (through fast-fourier transform) the harmonic content of the output signal. The downside is that this simulation mode is computationally over-expensive: The oscillation build-up phase, which is of no relevance to the user, must be simulated. Since most simulators use the results of the bias point calculation as initial conditions for a transient simulation, a kick-off is required in order to get the oscillator going. The most convenient method to do this is to introduce a short pulse on the circuit's power supply.

The third simulation mode discussed may be regarded as middle ground between AC- and transient simulation. *Harmonic balance* simulation assumes the circuit in a state of stationary harmonic excitation, like AC simulation. But the currents and voltages are not limited to their fundamentals. Rather, a determinate number of harmonics is admitted, allowing nonlinear device equations to be used. The only catch with harmonic balance is that it is a response simulation, posing again the hassle of finding a way to excite the circuit without exciting it. Most harmonic balance simulators (at least Agilent's ADS and AWR's Microwave Office) solve this by providing a special *oscillator probe*, which triggers a special mode within the simulator: The probe is essentially a voltage source delivering a fundamental of a certain amplitude, frequency, and phase, and a predefined number of harmonics whose amplitudes and phases are variable, too. The source has a finite source resistance. In principle the probe may be connected to any node in the circuit; in practice, the node where the circuit would be split into 'active part' and 'resonator' (Z_a and Z_r from above) is used. At the start of the simulation, the frequency of the source as well as the amplitudes and phases of the fundamental and all harmonics are set to some initial guesses; then a regular harmonic balance simulation is carried out. If the oscillator probe's signal is not identical to the signal the oscillator circuit had exhibited without the probe, the probe has to source or sink current. Then, the probe's parameters are varied and the next simulation step is performed. With enough crossing of fingers, the process converges, identifiable by a vanishing current into or from the source. The voltages and currents in

the circuit at the last simulation step are those of the steady oscillating state.

Unfortunately, the best simulation is as worse as the models it uses. The first test oscillator I built was preceded by heavy simulation work, and, for various reasons, failed horribly. Only then I realized that the vendor's model for the varactor I picked is a good example for a particularly bad model.

Being used solely in reverse bias, the conductance of a varactor is negligible, as well as the diffusion capacitance. The equivalent circuit therefore consists of the junction capacitance

$$C_j = \frac{C_{j0}}{(1 - V_d/V_j)^m},$$

where the zero-bias capacitance C_{j0} , the junction built-in potential V_j and the grading coefficient m are model parameters and V_d is the diode voltage (counted negatively in reverse bias), and the spreading resistance R_s . Since the doping profiles of varactors are optimized for large tuning ranges, the simple law presented above does not reproduce the actual $C_j(V_d)$ very well (cf. [Alpha APN1004]). On the other hand, the equation is implemented in SPICE, therefore forming a de-facto standard. To enhance the model's accuracy, sometimes a fictitious capacitor is added in parallel to the varactor. Its value is treated as a fit parameter. Note that for good approximation the models often use $m > 1$, what is physically impossible; some SPICE variants limit $m \leq 0.9$. The varactor model is augmented by the capacitors and inductors that represent the case parasitics (bond inductances and pad capacitances).

Figure 3.4 shows measured data and the performance of the model, both provided by the varactor's vendor. The predicted series resonance frequency due to the parasitic inductances is about 50 % higher than the one measured, and the losses in the model are too low by a factor of 20. Starting to distrust the vendor, I measured the varactor myself. The results were comparable to the measured data from the vendor. In an attempt to create a reliable oscillator simulation, I fitted my own model to the measured data. This model uses the chip parameters from the vendor (of which most important are $V_j = 6.22$ V, $C_{j0} = 9.12$ pF and $m = 2.42$) but different case parasitics and an additional resistor. The fitting was done using Microwave Office's optimizer.

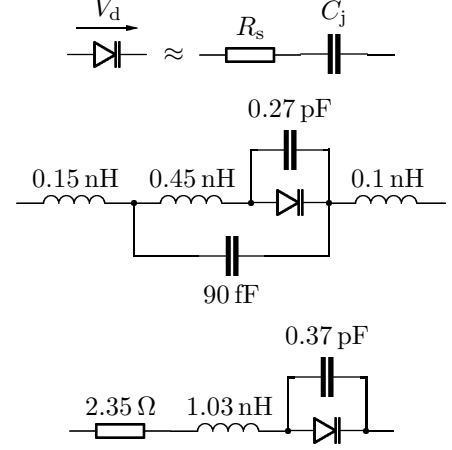


Figure 3.3 Simplified chip model (top), device model from the vendor (middle) and device model fitted to the measured data (bottom).

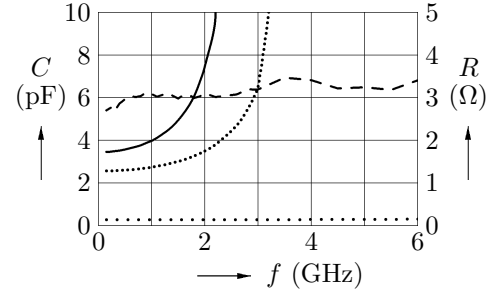


Figure 3.4 Capacitance (measured solid, model dotted) and resistance (measured dashed, model wide dotted) of an RC series equivalent circuit of the BB857 at $V_d = -5$ V.

Since in the sampler described in the previous chapter the BFP450 is already used, and this transistor has a transit frequency far higher than necessary to build an oscillator for up to 3 GHz, the device was tried in the oscillator, too. With this transistor (like with many small power bipolar RF transistors), adding an emitter resistor is sufficient to generate some negative resistance at the base; hence the grounded collector configuration suggested itself. From the various possible resonator configurations the Colpitts type was chosen, because it needs only one inductor. Notwithstanding the transistor's intrinsic affinity to oscillation, both of the two 'Colpitts'-capacitors were added in order to ensure enough negative resistance over the whole frequency range where oscillation was desired.

Besides the DC tuning voltage V_t , also a part of the RF voltage from the oscillation drops at the varactors. To minimize the effects that may stem from this unwanted superposition, it is customary to connect two varactors in series: They divide the RF voltage equally between both of them while maintaining the capacitance variation \hat{C}/\bar{C} of a single varactor. The two resistors R_d feed the tuning voltage to the varactors; the ground connection for the left varactor

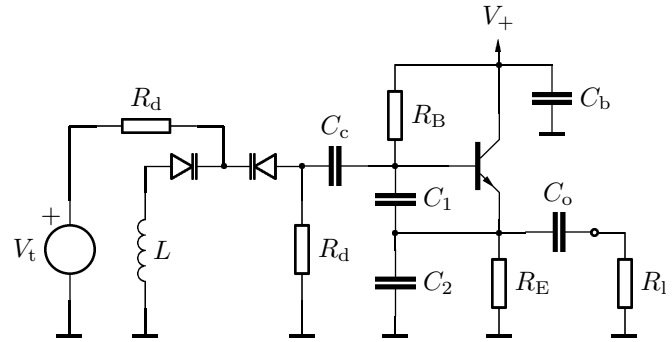


Figure 3.5 Colpitts oscillator with bipolar transistor in common collector configuration.

is provided through the inductor. Resistors instead of RF chokes were used to minimize the chance of spurious resonances, cf. [Alpha APN1006]; besides, resistors are cheaper. Since varactors are used in reverse bias, R_d may be made very large. But as R_d is increased, I supposed that the growing thermal noise voltage of the resistors would increase the oscillator's phase noise. In fact, within certain bounds, the opposite is true: Increasing R_d enhances the resonator's quality factor, lowering the phase noise. The free-running frequency of the circuit was very unstable. Therefore optimization with respect to phase noise had to be deferred until the oscillator was included in the phase-locked loop. Connected in series with the varactors, the coupling capacitor C_c must be made much larger than the zero-bias capacitance of the varactors, otherwise their capacitance variation (and hence the frequency range) will be narrowed unnecessarily. In fact, the capacitor's parasitic series inductance can easily be compensated for by making the resonator's inductor smaller. C_b is actually a cascade of capacitors in a geometric progression, starting at 10 pF and going down in steps of 100 : 1 to provide a small-impedance RF path from the positive supply to ground from DC to beyond the desired oscillation frequency. The quiescent current of the transistor was set to 10 mA, because there the transistor has its lowest noise figure. Since the maximum admissible collector

emitter voltage is only 4.5 V, V_+ was chosen to be 3.3 V.

The first oscillator was designed to yield a frequency range of 1...1.5 GHz, because at that time the band splitting on the left side of Table 3.1 was used. Simulations showed that with $L = 12$ nH and a tuning voltage range of $V_t = 0 \dots 20$ V the oscillator should be able to cover a frequency range of 950...2420 MHz; nonlinear simulation in the time domain was used. It is interesting to note that the linear (small signal) AC simulation predicted a frequency range of 1150...2290 MHz. The harmonic balance simulation with an oscillator probe did not converge, and therefore was given up. The real oscillator, however, did not start to oscillate until $V_t = 2$ V at 850 MHz, and ceased operation at $V_t = 12$ V with an oscillation frequency of 1340 MHz.

At that point, I discovered the faulty varactor model and replaced it. But also passive devices include noteworthy parasitics, most important series inductances. Measurements on three capacitors of case type 0603 with values 4.7 pF, 47 pF, and 1 nF showed that the first resonance could well be matched for all three by adding a series inductor of 0.5 nH. A 330 Ω resistor (case 0603) was measured and modelled with 0.44 nH series inductance; in addition, an 83 fF capacitor was added in parallel with the intrinsic resistor. The same model but with 3 nH series inductance was used for 3.3 k Ω resistors. The most difficult passive devices to model are inductors, since e.g. the losses due to the skin effect are proportional to \sqrt{f} . The model used in the simulations was tailored for the planned oscillation frequency, i.e. it used fixed parasitics more-or-less appropriate for the whole frequency range the circuit operated in. The only remaining part is the transistor. Its vendor-supplied model does not match the measured data from the vendor, either. The differences, though, are by far less drastical than with the varactor. Moreover, I did not succeed in finding a set of case parasitics that matched the measurements better than the one available.

After including the parasitics, the linear simulation's frequency range shifted to 1050...1905 MHz. In complete contrast to the measurement, the nonlinear simulation predicted oscillation only from $V_t = 0 \dots 2$ V (frequency range 850...950 MHz). At this point I decided to abandon simulation and to proceed with cut-and-try. Step-by-step disassembling the circuit revealed that at 1360 MHz some resonance effect occurred. The cause for this resonance can be seen in Figure 3.6, which shows a detail of the oscillator's board layout. The capacitor C_{b2} is one of the DC bypass cascade's and shares the via to the ground plane with R_E and C_2 . It had a value of 1 nF and hence was inductive above 190 MHz. But the via also has a parasitic inductance, and at 1360 MHz the structure exhibited a series resonance, shorting the transistor and impeding oscillation.

The next oscillator had a separate via for each component, but still did not reach the maximum frequency of 1500 MHz. The reason is as subtle as with the first oscillator: Each individual capacitor in the DC bypass cascade

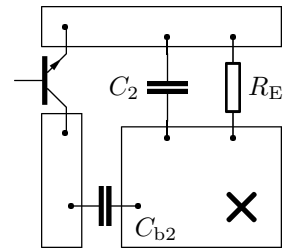


Figure 3.6 Detail of the first oscillator's board; the cross marks a via.

has a series resonance frequency determined by its capacitance and its parasitic inductance. But if a number of series-resonant circuits is connected in parallel, Foster's Reactance Theorem states that a parallel resonance must occur in between every two series resonances, i.e. there are frequencies for which the whole DC bypass cascade is essentially without effect! At these frequencies, the collector is not AC-grounded any longer, resulting in a considerable loss in transistor gain.

Figure 3.7 shows that there are frequency intervals in which a single 100 nF capacitor, which is purely inductive at these frequencies, is actually a better DC bypass than a parallel cascade. One remedy is to include a low-value resistor in series with each individual capacitor. The other is to use only a single DC bypass capacitor that has its series resonance in the middle of the frequency band the circuit operates. For the common collector oscillator circuit, still a better alternative suggested itself: Using a negative power supply, which includes connecting the resonator's inductor directly to the collector.

But it was not until I changed the BFP450 to a BFG520W, which has a *lower* transit frequency, that the oscillator showed a satisfactory frequency range. Having moved to the three-subband configuration (cf. Table 3.1) meanwhile, the BFP450 was unusable since it could not be made to oscillate above 2.4 GHz, not even when used in a common base circuit. The final oscillator uses the BFG520W at a collector current of 6 mA (where this transistor has its minimum noise figure), a negative supply voltage (-5 V, because the new transistor has a higher breakdown voltage), only one DC bypass capacitor (8.2 pF proved to be the most appropriate value for a frequency range of 2...3 GHz), and a narrow PCB track of about 0.25 mm width and 9 mm length instead of the lumped inductor. The PCB inductor is used because it enables the final adjustment of the frequency range to compensate for device tolerances. In this setup the frequency range of 2...3 GHz can be covered with a tuning voltage of 2...18 V. At lower tuning voltages the RF voltage drives the varactor diodes into forward conduction, observable as a quite sudden broadening of the oscillator's noise shape. When used in the phase-locked loop, however, other effects supposedly connected with varactor RF modulation were encountered (see below).

The oscillator for the fixed channel at 1 GHz was derived from the circuit above by changing the DC bypass capacitor to 27 pF, adapting the feedback capacitors C_1 and C_2 and of course increasing the resonator inductance to roughly the six-fold value by using a meander-like PCB trace.

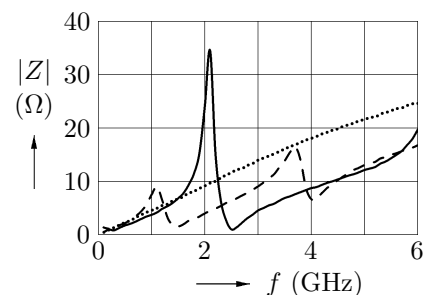


Figure 3.7 Impedance of a cascade of 100 nF, 1 nF, 10 pF, and 1 pF (solid); the same cascade, but with 1.5 Ω-resistors in series with each capacitor (dashed); and a single 100 nF capacitor (dotted).

3.2 PHASE-LOCKED LOOP

The PLL chip chosen was the ADF4252 from Analog Devices. It accommodates a fractional- N and an integer- N PLL. Presumably targeted at wireless communication devices, where the fractional part serves the first, the integer part the second local oscillator, the device seemed ideally suited to serve both the variable and fixed oscillators in this application.

Variable channel

The fractional- N part of the ADF4252 comprises a phase/frequency discriminator usable up to an input frequency of $\hat{f}_{\text{PFD1}} = 30$ MHz and a fractional divider including a prescaler with adjustable divide ratio of $P_1 = 4$ or $P_1 = 8$. For the fractional divider to operate properly, the prescaler's output frequency must be less than $\hat{f}_{\text{pre1}} = 375$ MHz. The fractional divider realizes a division ratio of $N = I + F/M$, where $31 \leq I \leq 255$ when $P_1 = 4$; otherwise ($P_1 = 8$), $91 \leq I \leq 255$. Independent of P_1 , the fractional part is restricted by $2 \leq M \leq 4095$ and $0 \leq F \leq M$. The externally applied reference frequency may be doubled ($D_1 = 1$) and divided by $1 \leq R_1 \leq 15$. Thus, the synthesizer's output frequency is

$$f_v = f_{\text{ref}} \frac{1 + D_1}{R_1} \left(I + \frac{F}{M} \right) .$$

The variable oscillator should have an output frequency range of $2 \text{ GHz} \leq f_v \leq 3 \text{ GHz}$, with a frequency resolution as fine as possible. Since $\hat{f}_v / \hat{f}_{\text{pre1}} = 8$, the prescaler must be programmed with $P_1 = 8$. From the range of possible division ratios (note $\hat{N} = 91$) it follows that

$$f_{\text{PFD1}} \leq \check{f}_v / \check{N} = 22.0 \text{ MHz} \quad \text{and} \quad f_{\text{PFD1}} \geq \hat{f}_v / \hat{N} = 11.7 \text{ MHz} .$$

If the standard system reference frequency of 10 MHz should be used, the only possible choice is $D_1 = R_1 = 1$, resulting in $f_{\text{PFD1}} = 20$ MHz and a maximum frequency resolution of $\Delta f_v = f_{\text{PFD1}} \Delta N = f_{\text{PFD1}} / \hat{M} = 4.88$ kHz. Practically, M will be set to 4000, resulting in $\Delta f_v = 5$ kHz.

Since the varactors need a tuning voltage of up to 20 V to achieve the broad frequency range, an active loop filter is used. The filter itself, shown in Figure 3.8, is a second order lowpass filter suggested by Analog Devices' design program ADsimPLL. Unfortunately, ADsimPLL is not smart enough to include the opamp's finite gain-bandwidth product in its considerations. The filter's frequency response is

$$H(f) = \frac{V_t}{I_p} = H_0(f) H'(f)$$

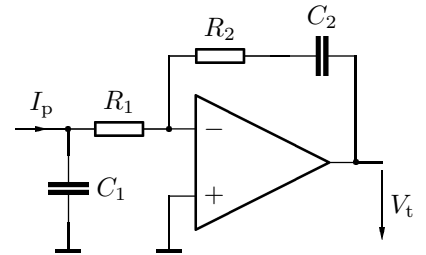


Figure 3.8 Active loop filter for the variable channel of the PLL. I_p is the charge pump current from the PLL chip, V_t the tuning voltage for the varactors.

with the ideal transfer characteristic

$$H_0(f) = \frac{R_2}{jf/f_2} \frac{1 + jf/f_2}{1 + jf/f_1}, \quad f_i = \frac{1}{2\pi R_i C_i},$$

and the correction term

$$H'(f) = \frac{1}{1 + \frac{1}{v_d} \left(1 + \frac{C_1}{C_2} \frac{1 + jf/f_2}{1 + jf/f_1} \right)}$$

where $v_d = v_0/(1 + jv_0f/f_t)$ is the opamp's open loop gain. The initially selected AD8671 provides very low noise, but the low gain bandwidth product of $f_t = 10$ MHz has an impact on the filter transfer function despite the high open-loop gain of $\bar{v}_0 = 6 \times 10^6$, cf. Figure 3.9. As an alternative choice, the AD817 (with $\bar{f}_t = 50$ MHz and $\bar{v}_0 = 6 \times 10^3$) was also used to rule out the opamp as the reason for the encountered PLL malfunction when higher loop bandwidths were employed (see below).

Fixed channel

The phase/frequency discriminator of the ADF4252's integer- N part accepts input frequencies of up to $\hat{f}_{\text{PFD2}} = 55$ MHz. The dual modulus prescaler can be programmed with $P_2 = 8, 16, 32$, or 64 , and must have an output frequency of less than $\hat{f}_{\text{pre2}} = 150$ MHz. The PLL's output frequency is $f_{\text{PFD2}}(BP_2 + A)$, where the prescaler control registers A and B allow $3 \leq B \leq 4095$ and $0 \leq A \leq 63$. The reference path has its own doubler bit D_2 and a prescaler with $1 \leq R_2 \leq 32767$.

For the required frequency of $f_f = 1$ GHz, $P_2 \geq f_f/\hat{f}_{\text{pre2}} \approx 7$. To have the spurious frequencies away from the carrier as far as possible, $D_2 = R_2 = 1$, leaving the settings $P_2 = 8, B = 6, A = 2$, or $P'_2 = 16, B' = 3, A' = 2$.

The loop filter is a passive second order lowpass with a pole at $s = 0$.

3.3 RESULTS

Though the power supply lines were separated as well as possible and shielding structures were placed on the PCB to minimize interaction between the two oscillators, considerable crosstalk occurred. While normally this is just a matter of spectral purity, an interesting behaviour can be observed when the variable oscillator is set to the low or high end of its tuning range. There, its own oscillation frequency is close to the first or second harmonic of the fixed channel's frequency, and both spectra show strong distortion, possibly due to synchronization or related phenomena observable in nonlinear systems. I believe that the coupling is mainly caused by the PCB inductors, which produce a lot of stray field. Using separate PLL modules for the two sources and fitting

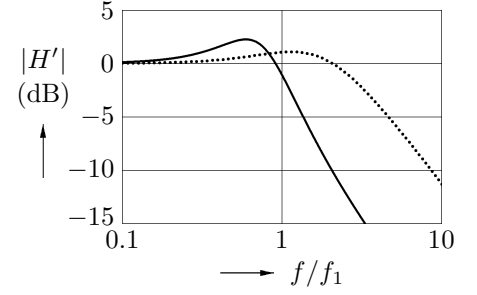


Figure 3.9 Loop filter frequency response error due to a finite gain-bandwidth product, normalized to the filter's second pole frequency f_1 for the AD8671 (solid) and the AD817 (dotted). Parameters were $R_1 = 130 \Omega$, $C_1 = 1$ nF, $R_2 = 1.8$ k Ω , and $C_2 = 470$ pF, resulting in a filter for a loop bandwidth of 500 kHz.

them into separate metal cabinets is mandatory for good performance.

The second important observation made when measuring the module was that the circuit is very susceptible to power line hum. The PLL's evaluation kit from Analog Devices circumvents this by using a battery for power supply, what is not a particularly practical solution. Additional 1000 μF electrolyte capacitors at the power supply connectors lowered the power line spurs by about 10 dB to approximately -35 dBc . Adding linear voltage regulators and incorporating of the module into a metal housing are expected to improve the spectral purity.

Since within the loop bandwidth the output signal's phase noise is reduced to the PLL's phase noise, I tried to make the loop filters as broadband as possible. In fact, this also ensures fast locking, which directly affects the VNA's measurement speed. With charge pump phase/frequency discriminators, the loop bandwidth should be less than $f_{\text{PFD}}/5$ (cf. [Analog AN30]) in order to ensure the stability of the loop; in the case of the variable channel, this means a maximum loop bandwidth of 4 MHz. I started the design with a loop bandwidth

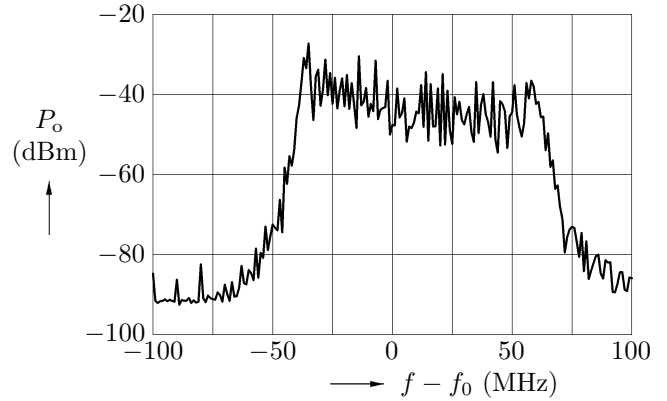


Figure 3.10 ‘Brick’ output spectrum, centered to the actually preset output frequency.

of 100 kHz and tuned the oscillator's PCB coil such that the frequency range of 2...3 GHz was reached with a tuning voltage of 2...18 V. Then I tried to increase the loop bandwidth, and at the lower end of the frequency range the output started to show a ‘brick’-like spectrum like the one shown in Figure 3.10. The spectrum actually consists of a large number of distinct carriers spaced a few hundred kilohertz apart, with slowly varying amplitudes. The brick spectrum did not occur at frequencies $f_v > 2.4\text{ GHz}$, or when the PLL's doubler-bit was cleared (effectively reducing f_{PFD} to 10 MHz, but also the maximum output frequency \hat{f}_v to 2.56 GHz through the restriction on the fractional divider's I -register). The brick spectrum also showed up with the 100 kHz-loop when the resistors providing the DC ground for the varactors (R_d in Figure 3.5) were increased from their initial value of 3.3 k Ω . Moreover, even at those frequencies where the synthesizer worked (only a single carrier showed up), the increase of the loop bandwidth *increased* the overall phase noise, because the oscillators noise skirt grew wider, shifting the point where the noise skirt and loop filter characteristic intersect not only to higher offset frequencies, but also to higher powers. This was most critical in the case of fractional divider ratios, less for integer multiples of the PFD frequency. Changing the opamp type to ensure smaller deviations from the ideal loop filter frequency response (cf. Figure 3.9) did not help.

The only sensible explanation for the brick spectra is that self-modulation of the varactors causes subharmonic oscillations: The RF voltage superimposed

on the varactors in addition to the tuning voltage modulates their capacitance. Self-modulation gets more serious the lower the tuning voltage and the higher the RF voltage is, explaining the fact that brick spectra only appear at output frequencies at the lower end (low tuning voltage). Increasing the DC resistors reduces the losses in the resonator. The higher quality factor, observable at the output as a rather drastical (≈ 10 dB) lowering of the phase noise, results in a larger RF amplitude, which explains another of the observations described in the previous paragraph. Knowing too little about nonlinear phenomena like multiple oscillations, the dependence on the loop bandwidth can only be speculated about. Maybe the multiple oscillations need a special stimulus only provided when higher frequency components are present in the tuning voltage. This would also explain why these effects were not observed when the oscillator was tested alone, using a DC lab source to provide the tuning voltage.

The final remedy for these problems was to deliberately shift the oscillator's frequency range by adjusting the PCB coil such that the output frequency of $\tilde{f}_v = 2$ GHz is reached with a tuning voltage as high as possible. This is of course limited by the requirement of reaching \hat{f}_v within the opamp's output swing (which could not be increased much because of both the admissible supply voltage of the opamp and the maximum reverse voltage of the varactors). The DC resistors for the varactors were set to $33\text{ k}\Omega$; higher values do not improve phase noise any more. The loop bandwidth must be limited to 200 kHz ; in fact, measurements showed that reducing it to 50 kHz does not necessarily increase phase noise.

With the fixed channel's oscillator, the initial design's loop bandwidth of 2 MHz presented none of the above problems. Although, increasing the DC path resistors also produces the same brick spectrum. Nevertheless, in order to reduce phase noise, the oscillator's tuning voltage sensitivity is lowered by a smaller coupling capacitor (C_c in Figure 3.5). In theory, this should also mitigate the self-modulation effects, since the RF voltage is divided between the varactors and the coupling capacitor. The smaller the capacitor, the larger the fraction of the voltage that drops at it, leaving room for an increase in the resonator's quality factor. In practice, unfortunately, this was not the case.

With the final circuit, tuning ranges were determined to be 1838 MHz to 3041 MHz (tuning voltage 2.2 V to 29.2 V) for the variable channel and 910 MHz to 1130 MHz (tuning voltage 0.5 V to 4.7 V) for the fixed channel. Note that for reasons discussed above, at the lower ends of the given intervals the spectra have very poor quality.

Output Power

The output power was measured using a spectrum analyzer in zero-span mode with a resolution bandwidth of 1 kHz .

In addition, the figure displays the tuning voltage. Shifting the tuning voltage required for \hat{f}_v from the (theoretically possible) 2 V to 5 V results in an over-proportional increase in the tuning voltage required for \hat{f}_v , because the CV -characteristic of a varactor causes the tuning voltage sensitivity to decline for raising tuning voltages. A lower tuning voltage sensitivity, however, improves the phase noise performance, adding to the effect of increased resonator quality factor. The output power for the fixed channel is -12.1 dBm.

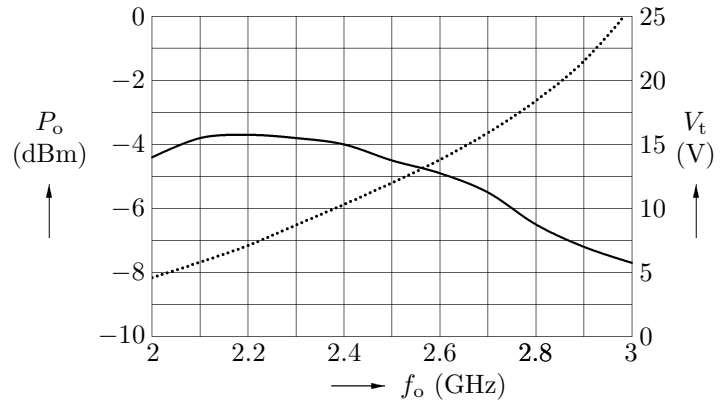


Figure 3.11 Output power (solid, left scale) and tuning voltage (dotted, right scale) of the synthesizer's variable channel.

Harmonic Content

The first harmonic at the output is quite strong. An output filter will be necessary in almost every application.

The harmonics for the fixed channel are 20.4 dBc, 26.6 dBc, 30.9 dBc, and 44.9 dBc for the first to the fourth harmonic, respectively. Thanks to Mr. Murphy, the first harmonic of this oscillator, where harmonics are favoured, is quite low.

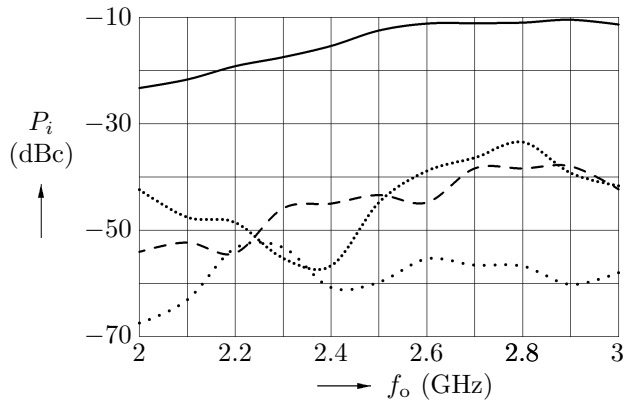


Figure 3.12 Power of the harmonics (first harmonic solid, second narrow dotted, third dashed, fourth wide dotted) at the synthesizer's variable channel.

Phase Noise

The phase noise was measured with a spectrum analyzer in zero-span mode, using a resolution bandwidth of 10 Hz and a sweep time of 1 s. An average of ten sweeps was obtained, and the 601 points of the average trace were again averaged to yield the phase noise figure at this particular frequency.

Regarding the offset frequencies where phase noise was measured, a little cheating was necessary: Due to the always present power line hum at multiples of 100 Hz, the given measurement frequencies were actually shifted 15 Hz upwards. At multiples of f_r , the measured value was recorded as a spurious and a second measurement with a frequency shift of 33.333 kHz was taken.

The phase noise of the variable channel was measured at different center frequencies and with different loop bandwidths. The solid trace shows the spectrum at 2103.005 MHz, which was the worst case with a loop bandwidth of 200 kHz. With this loop bandwidth, the phase noise at fractional divider ratios is much stronger than with integer ratios. Reducing the loop bandwidth to 50 kHz, the differences between the phase noise at fractional and integer ratios vanish; the dotted trace shows the worst-case phase noise for 50 kHz bandwidth, which was measured at 2102 MHz. Virtually the same phase noise performance can be achieved with the 200 kHz-loop by setting the charge pump current to its minimum (0.625 mA in this configuration).

The peaks of the phase noise curves at 400 Hz and 800 Hz are supposedly caused by power supply ripple. The shape of the phase noise curves agrees with [Hajimiri], including the peak near cut-off.

Lock Time

The lock time was only determined for the variable oscillator part of the module, since the fixed part (hopefully) locks at power-up and never unlocks.

The measurement setup employs the AD8302 from Analog Devices, a gain/phase detector usable up to 2700 MHz with an output bandwidth of 30 MHz. The synthesizer's signal is fed into one input of the phase detector, the other input is supplied with a signal of frequency f_2 from a lab RF source. The PLL is locked to some frequency f_1 and programmed to exercise a jump to a frequency f_2 . The programming through the serial interface lines *DAT*, *CLK*, and *LE* is finished at the last falling edge of the *LE* signal. The output of the phase detector is recorded on a digital storage scope triggered by the *LE* signal. When the PLL's instantaneous frequency approaches f_2 within ± 30 MHz, the phase detector's output starts to display a signal whose instantaneous frequency is the difference of the synthesizer's instantaneous frequency and f_2 . When the detector output

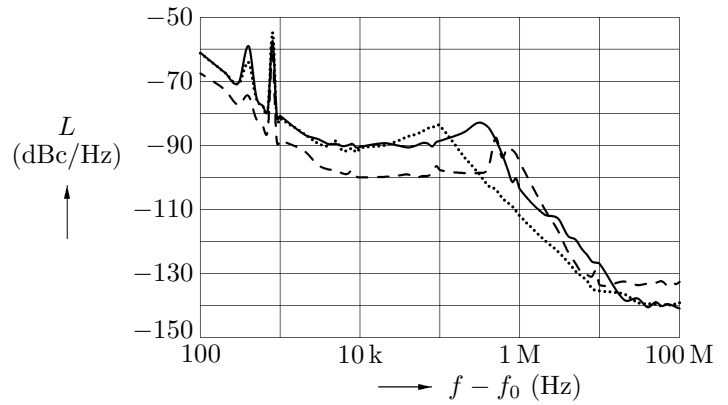


Figure 3.13 Worst-case phase noise of the variable channel (solid for a loop bandwidth of 200 kHz, dotted for 50 kHz) and of the fixed channel (dashed).

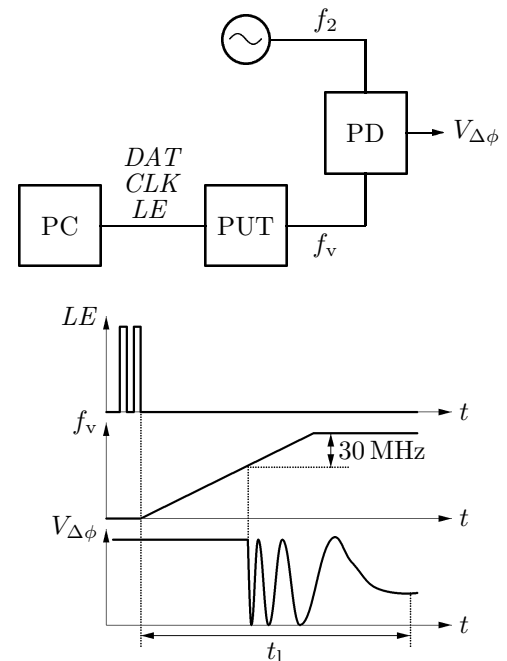


Figure 3.14 Measurement setup for the PLL's lock time t_l (timing diagram not to scale). PUT ... PLL Under Test, PD ... Phase Detector.

signal settles, the PLL has acquired lock. Note that the voltage at the phase detector output after locking in general does not get zero but settles at a non-zero value, since the RF source for f_2 and the source from which the PLL reference is drawn will have a nonzero phase shift.

In theory, large frequency steps cause the PLL to lose lock and regain it after the slip of some cycles, whereas small frequency steps should be followed without any cycle slips. However, when measuring the lock time it turned out that the loop unlocks even at the smallest jumps (5 kHz). The lock time for a jump from 2 GHz to 2.5 GHz was 35 μ s for a PLL with loop bandwidth 200 kHz.

3.4 IMPROVEMENTS AND PROSPECTS

Although time did not permit to develop the whole broadband source, an idea for a possible improvement of the oscillators and a few considerations regarding the function blocks necessary to complete the circuit in Figure 3.1 are discussed at last.

Oscillator

To overcome the troubles with varactor capacitance self-modulation at low tuning voltages, a quadruple configuration may be used. It distributes the RF voltage over four diodes instead of two, effectively cutting in half the lowest allowed tuning voltage. Similar to the two-diode configuration, the circuit quarters the resulting capacitance while maintaining the relative variability \hat{C}/\check{C} of a single varactor. With the BB857, however, $\check{C} \approx 0.5$ pF; dividing this by four enters an order where stray capacitances on the PCB might have to be considered, which restricts the effective capacitance variability. In addition, the four diodes accumulate more parasitic series inductance. Both issues may negatively affect the achievable frequency range.

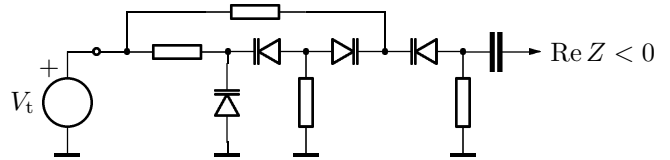


Figure 3.15 Alternative circuit for lowering self-modulation effects.

Mixer and Bandpasses

The nonlinear device to generate the harmonics of f_f may be either a Schottky-diode, a varactor, or an additional transistor amplifier stage optimized for the second harmonic. Depending on the tolerable frustration level, some variations in the feedback and/or bias point of the fixed oscillator may yield a circuit which produces enough harmonic content to make an additional nonlinearity obsolete.

Finding a mixer suitable for the broadband source is a real challenge, since it must be able to handle comparatively large frequency ranges on all ports.

Furthermore, good isolation from both input ports to the output is vital for high harmonic purity of the output signal, disqualifying diode ring mixers. Being left with integrated Gilbert cell mixers, most of them are dedicated up- or downconverters with a limited frequency range on one input- and the output port. The only suitable types at the time of design of this project were Linear Technology's LT5521 and LT5560. When using the LT5521 (or an equivalent circuit which operates as overdriven Gilbert mixer), it is advisable to use the variable oscillator as LO, and feed the fixed oscillator to the IF input. The former is internally wide-band matched, and the external matching network required for the latter is easier to design when only two discrete frequencies have to be considered instead of a rather broad frequency range. Moreover, in this configuration the inherent harmonics of the LO due to the overdrive are already damped by the mixer's input and output frequency range restrictions, whereas multiples of 1 GHz had to be specially handled if the inputs were exchanged.

The tunable bandpass filters are most easily implemented as a filter bank. While for the bandpass that selects the harmonics of f_f high stopband attenuation is the most important parameter, whereas the exact cutoff frequencies are secondary, the difficulty with the output filters are their broad bandwidths. The output filters are probably best realized as stepped impedance- or hairpin transmission line Cauer filters. For the harmonic selection filters a structure similar to a 90° microstrip hybrid coupler with appropriate taps switched by pin-diodes may perform the harmonic selection and matching of the mixer simultaneously.

APPENDIX A

Annotated Schematics

*Wires connecting are indicated by a heavy black dot;
wires crossing, but not connecting, have no dot
(don't use a little half-circular 'jog';
it went out in the 1950s).*

— Paul Horowitz and Winfield Hill, “The Art Of Electronics”

This chapter contains the final schematics of the sampler and the synthesizer module described in the main part, together with remarks on the choice of device values.

Resistors are unbranded 1 % (E24), 1/10 W (case 0603) or 1/16 W (case 0402) types in general. Capacitors ≤ 100 pF are series 600S-types from ATC (American Technical Ceramics), all others are unbranded types.

A.1 SAMPLER

The capacitors C_1 through C_{10} and C_{22} through C_{24} are the usual DC bypasses. R_1 is the load for the sampling frequency source. The sine is coupled through C_{11} ; R_2 and R_3 shift the DC level to 1 V. Since the Schmitt trigger's input thresholds have large tolerances ($V_{\text{in,low}} = 0.8 \dots 1.5$ V, $V_{\text{in,high}} = 1.1 \dots 2.0$ V), it may be necessary to adapt the divider. The Schmitt trigger itself was chosen from the LVC-family because of its high output current capabilities (± 50 mA).

The differential amplifier is discussed in depth in Section 2.5. The values of C_{12} , C_{13} , R_4 , R_5 , R_8 , and R_{10} were determined empirically. R_9 provides a slightly positive (160 mV) bias to the differential pair; since $R_9 \gg R_{10}$, $R_9 \parallel R_{10} \approx R_4$. C_{14} and C_{15} are uncritical as long as they follow the boundaries in Section 2.5. They should, however, have a low series inductance, thus too large values are prohibited. Measurements showed that for values up to about 1 nF, all case-0603 capacitors have an equal series inductance of roughly 0.5 nH.

The choice of the quiescent current follows Section 2.5. Setting $\hat{I}_d = 3$ mA, $\hat{V}_d = 289$ mV. If $Z_w = 70 \Omega$ were chosen, $R_c = 256 \Omega$; to gain a little elbow-room to compensate for tolerances, I set $R_6 = R_7 \doteq 220 \Omega$, requiring $Z_w = 67 \Omega$ and $I_0 = 8.6$ mA. $R_{18} = (V_+ - V_{\text{BE}})/I_0 = 492 \Omega \doteq 510 \Omega$. When measuring the sampler, these values proved to be optimal.

Initially C_{16} and C_{17} were 15 pF; they were optimized empirically.

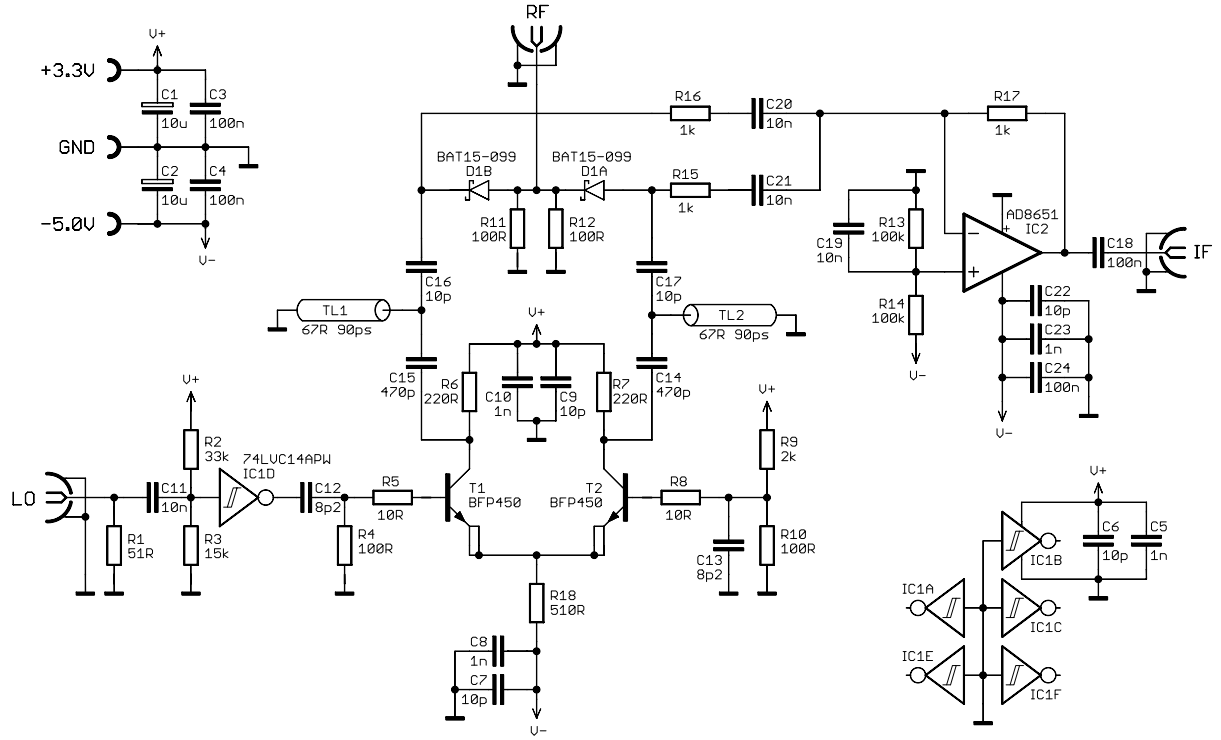


Figure A.1 Schematic of the final sampler circuit.

The RF input's match resistor is split (R_{11} and R_{12}), since 50Ω resistors are not standard. Also the two resistors provide a visually more appealing layout.

The resistors $R_{15} = R_{16} = R_{17}$ are a compromise between least noise (small values) and least conversion loss (large values). Since conversion loss is high anyway, it seems sensible to enhance the dynamic range by lowering the noise floor as far as possible. For that reason, the AD8651 was chosen. It has a maximum supply voltage of 5.5 V, enabling single supply, which requires biasing of the non-inverting input through R_{13} and R_{14} . C_{20} and C_{21} block this bias; they introduce an unwanted high-pass with a cut-off frequency $\approx 16\text{ kHz}$. C_{19} is used to damp the voltage divider's thermal noise. It has a quite interesting property: R_{13} and R_{14} produce a total noise voltage of $\sqrt{2k_B T B R_{13}}$; together with C_{19} they form a first-order lowpass with $f_c = 1/(\pi R_{13} C_{19})$. If for the sampler's output frequency $f_o \gg f_c$ is true, the noise voltage at the non-inverting input is $V_n = \sqrt{2k_B T B R_{13}} |1/(1 + j f_o/f_c)| \approx \sqrt{2k_B T B R_{13}} |1/(j f_o/f_c)| = \sqrt{2k_B T B} / (f_o \pi C_{19} \sqrt{R_{13}})$. So, while increasing the resistance increases the noise voltage, the cut-off frequency is lowered, resulting in a higher attenuation at a fixed frequency. While the former effect follows a square-root law, the latter goes linearly with the resistance. R_{13} and R_{14} should therefore be chosen high (which is no problem since the AD8651 is a CMOS device with small input currents), and $C_{19} \gg 1/(\pi f_o R_{13})$. Every increase in C_{19} above that value decreases the noise contribution in a linear fashion.

Power supply

	Voltage	Current
V_+	3.3 V	18 mA
V_-	−5.0 V	22 mA

Adjustments

To adapt to device variations, varying R_6 and R_7 may improve the frequency response flatness. The frequency range may be extended by slightly altering C_{12} and R_4 (and C_{13} and R_{10} , respectively). Depending on the Schmitt trigger threshold voltages, changing the ratio of R_2 and R_3 may be necessary.

shipped with it. The *CLK*, *DATA*, and *LE* lines connect to the appropriate pins on the controlling PC's parallel (printer port) interface.

R_{13} selects the maximum charge pump current (4.375 mA), to facilitate fast locking.

The variable frequency oscillator requires a comparably large tuning voltage range, which can not be achieved by the PLL itself; the necessary amplification is achieved with an active loop filter, built around IC_2 . The configuration shown results in a PLL loop bandwidth of 200 kHz. To minimize coupling of the RF signal into the control voltage through stray capacitances on the PCB, C_{15A} is added—the value is the same as C_{22} , the best possible approximation to a short circuit between 2 and 3 GHz. Since the charge pump prefers to have its output at half of its supply voltage, R_{24} and R_{25} form an appropriate voltage divider. The resistors' value is uncritical as long as the voltage drop across them due to the opamp's input bias current is negligible. They contribute noise to the loop, which is damped through C_{10} , at least for frequencies larger than $1/(2\pi R_{24} \parallel R_{25} C_{10}) = 1/(\pi R_{24} C_{10}) \approx 290$ Hz. The opamp initially selected was an Analog Devices AD8671, which was later changed to an AD817. The former has lower noise and supply current, but the latter has a larger gain-bandwidth product; it appears that both yield equal performance (see also Figure 3.9). The most critical parameter of the opamp in this application is the maximum supply voltage. Depending on the opamp's output voltage swing, increasing V_{+++} may be necessary. The maximum tuning voltage should, however, not exceed 30 V, since this is the maximum rating for the varactors.

A passive loop filter (C_{17} , C_{18} , and R_{26}) is sufficient for the fixed frequency oscillator; its bandwidth is 2 MHz.

The oscillators themselves are described in depth in Section 3.1. They are basically identical for both channels, and differ only in the DC bypass capacitor (as described above), the feedback capacitors, the coupling capacitor to the resonator and the resistors for the varactor's DC path. While in the variable-channel oscillator noise is reduced by larger resistors improving the resonator's quality factor, this approach leads to multiple oscillations in the fixed oscillator. There, a noise reduction is possible by lowering the tuning voltage sensitivity through a smaller coupling capacitor (C_{23}). The oscillators' output power is coupled out with C_{21} and C_{25} . R_{16} , R_{17} , and R_{18} (R_{19} , R_{20} , and R_{21}) form resistive 6 dB-dividers which split the RF signal between the modules' outputs and the PLL's RF inputs. These inputs are matched by R_{14} and R_{18} , C_{11} through C_{14} are DC blocks. These circuits are copied from the ADF4252 data sheet.

Power supply

	Voltage	Current
V_+	3.3 V	$\leq 12.6 \text{ mA}^\#$, $\leq 5.9 \text{ mA}^\P$
V_{++}	5.0 V	$\leq 2.0 \text{ mA}^\#$, $\leq 1.0 \text{ mA}^\P$
V_{+++}	30 V	$6.8 \text{ mA}^\#$, (n/a) ¶
V_-	-5.0 V	$5.6 \text{ mA}^\#^\P$

$^\#$ 2...3 GHz oscillator equipped; ¶ 1 GHz oscillator equipped

Adjustments

The oscillator's inductors are realized as tracks on the PCB, which must be shorted at the proper position. After assembling the module, a voltmeter is connected to the respective loop filter's output. A small piece of copper foil is soldered as short circuit to the PCB-coil so that the full tuning range of 2...3 GHz is covered with a tuning voltage of approximately 5...27 V for the variable oscillator. As described in Section 3.3, the tuning voltage for the lowest output frequency should be as high as possible (provided that the highest output frequency is still reached within the output voltage range of IC_2 , of course).

For the fixed oscillator, the tuning voltage for 1 GHz should be $V_{++}/2 = 2.5 \text{ V}$.

APPENDIX B

Board Layouts

*Tell me what you like,
I'll print it all.*
— Roger Hodgson, “My Magazine”

This chapter contains the board layouts for the modules from the main part. In general the passive devices have case size 0603, with the exception of the sampler's pulse forming stage, where size 0402 components were used wherever possible to minimize parasitic inductances.

B.1 SAMPLER

Since the quality of the pulse forming transmission lines is crucial to the operation of the sampler, the module is built on an RF substrate (Rogers RO4003C, 0.5 mm thick). The laminate's narrow tolerances on the permittivity, at least compared to FR4, and its low dielectric losses ensure low attenuation and low dispersion of the lines.

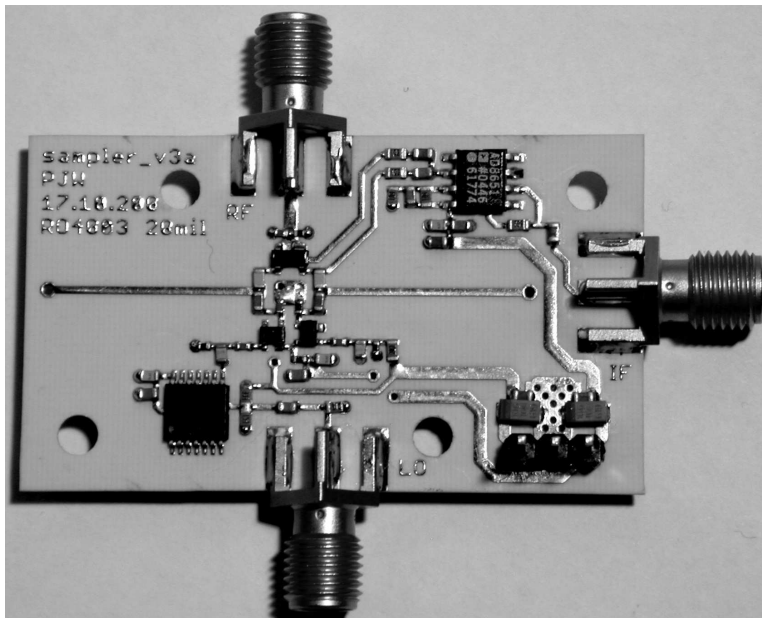


Figure B.1 Photograph of the sampler module.

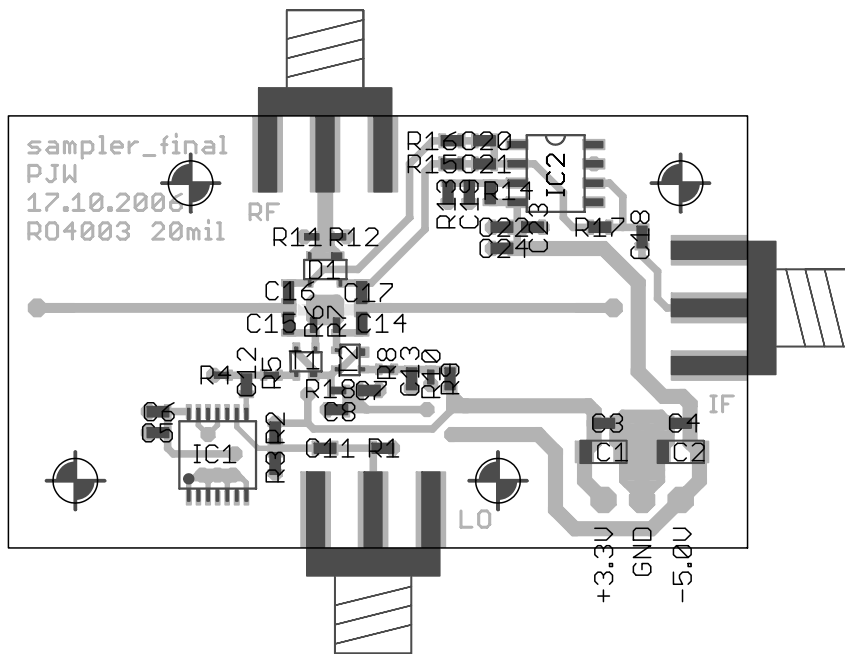


Figure B.2 Top side of the sampler PCB.

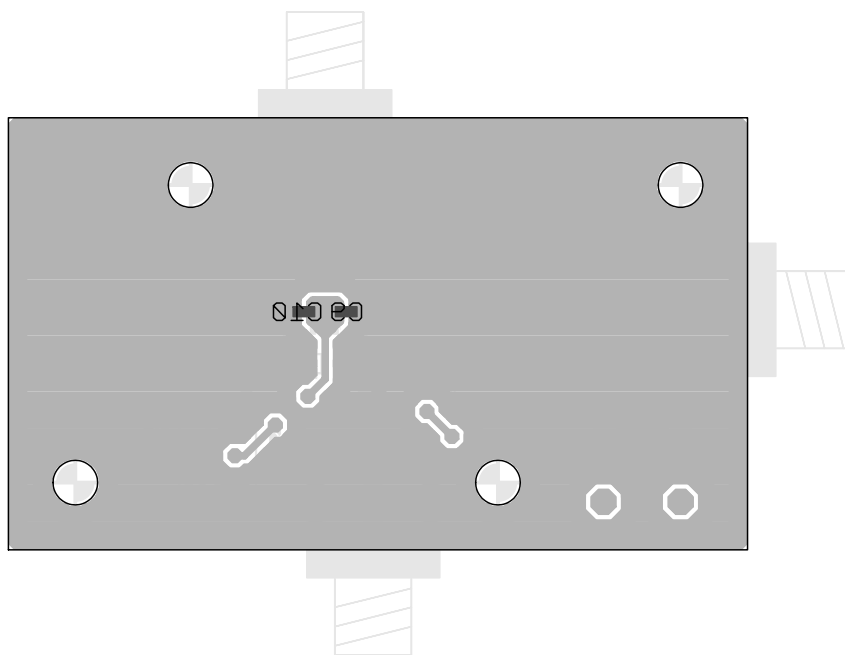


Figure B.3 Bottom side of the sampler PCB.

B.2 SYNTHESIZER

The VCO's inductors are formed by meander-like traces on the PCB. They are to be shorted to appropriate length, cf. Section A.2. The substrate is standard FR4 with a thickness of 1.5 mm. The module should be manufacturable with equal performance on any substrate, though.

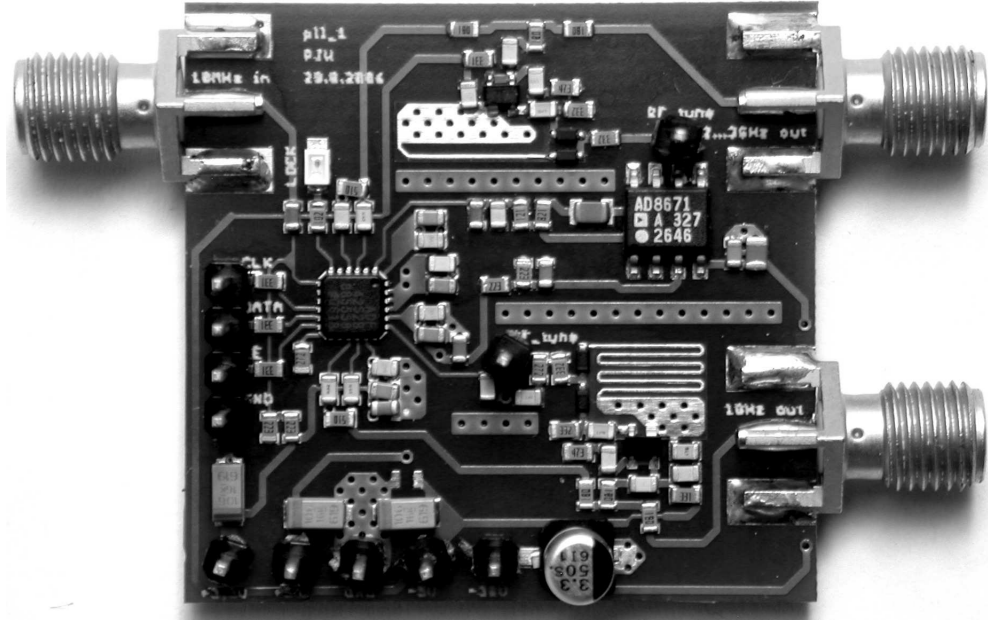


Figure B.4 Photograph of the synthesizer module with both oscillators equipped. Note that this picture actually shows the first and not the final version of the synthesizer, which only differ in some small details.

Bibliography

Ehsan Afshari and Ali Hajimiri: *Nonlinear Transmission Lines for Pulse Shaping in Silicon*. IEEE Journal of Solid-State Circuits, Vol. 40, No. 3, March 2005, pp. 744–752

Agoston Agoston, Steve Pepper, Robert Norton, John Ebner, and Kipp Schoen: *100 GHz Through-Line Sampler System with Sampling Rates in Excess of 10 GSa/s*. 2003 IEEE MTT-S International Microwave Symposium Digest, Vol. 3, June 2003, pp. 1519–1521

Alpha Industries, Inc. APN1004: *Varactor Spice Models for RF VCO Applications*. Application Note, 6/99A

Alpha Industries, Inc. APN1006: *A Colpitts VCO for Wideband (0.95–2.15 GHz) Set-Top TV Tuner Applications*. Application Note, 6/99A

Analog Devices AN30: *PLL Synthesizers*. Application Note, 2002

Anritsu AN246: *Arbitrary Impedance*. Application Note, Rev. B May 2002

Roland Best: *Theorie und Anwendungen des Phase-locked Loops*. AT-Verlag Aarau, 4. Auflage 1987

Herbert J. Carlin: *On the Existence of a Scattering Representation for Passive Networks*. IEEE Transactions on Circuit Theory, Vol. 14, Issue 4, December 1967, pp. 418–419

Joseph Carlstein: *A Physical Representation for Scattering Matrices Employing Complex Normalization*. Proceedings of the IEEE, Vol. 55, Issue 4, April 1967, pp. 553–554

Wai-Kai Chen and Qiang-Zhong Zha: *Factorizations and Complex Normalization of Scattering Matrices*. Proceedings of the 34th Midwest Symposium on Circuits and Systems, Vol. 2, May 1991, pp. 716–718

Jerzy Chramiec and Piotr S. Debicki: *Teaching the Principles of Microwave Oscillator Design*. 14th International Conference on Microwaves, Radar and Wireless Communications. Vol. 2, May 2002, pp. 429–432

L. M. Devlin, A. W. Dearn, G. A. Pearson, P. D. Beasley, and G. D. Morgan: *A Monolithic, 2 to 18 GHz Upconverter*. 2002 IEEE MTT-S International Microwave Symposium Digest, Vol. 1, June 2002, pp. 257–260

Dean A. Frickey: *Conversions Between S , Z , Y , h , $ABCD$, and T Parameters which are valid for Complex Source and Load Impedances*. IEEE Transactions on Microwave Theory and Techniques, Vol. 42, Issue 2, February 1994, pp. 205–211

Floyd M. Gardner: *Phaselock Techniques*. John Wiley & Sons, 2nd edition 1979

Hewlett Packard: *HP8510 Service Manual*.

Hewlett Packard: *Skalare Netzwerkanalyse im Mikrowellenbereich*. Workshop handout.

Habiba Hafdallah, Achour Ouslimani, Robert Adde, Guy Vernet and Paul Crozat: *Feasibility of Picosecond Electrical Sampling Using GaAs FET*. IEEE Transactions on Instrumentation and Measurement, Vol. 49, No. 1, February 2000, pp. 172–177

Habiba Hafdallah, Guy Vernet and Robert Adde: *2.4 GHz MESFET Sampler*. Electronics Letters, Vol. 24, No. 3, 4th February 1988, pp. 151–153

Ali Hajimiri: *Noise in Phase-Locked Loops*. Southwest Symposium on Mixed-Signal Design, February 2001, pp. 1–6

Jeongwoo Han and Cam Nguyen: *A New Ultra-Wideband, Ultra-Short Monocycle Pulse Generator with Reduced Ringing*. IEEE Microwave and Wireless Components Letters, Vol. 12, No. 6, June 2002, pp. 206–208

Jeongwoo Han and Cam Nguyen: *Coupled-Slotline-Hybrid Sampling Mixer Integrated with Step-Recovery-Diode Pulse Generator for UWB Applications*. IEEE Transactions on Microwave Theory and Techniques, Vol. 53, No. 6, June 2005, pp. 1875–1882

Robert W. Jackson: *Criteria for the Onset of Oscillation In Microwave Circuits*. IEEE Transactions on Microwave Theory and Techniques, Vol. 40, No. 3, March 1992, pp. 566–569

Hans G. Jungmeister and D. Schmidt: *High-Speed Pulse Circuits Obtained by Computer-Aided Nonlinear Analysis of Step-Recovery Diodes.* IEEE Journal of Solid-State Circuits, Vol. SC-5, No. 4, August 1970, pp. 135–141

Mark Kahrs: *50 Years of RF and Microwave Sampling.* IEEE Transactions on Microwave Theory and Techniques, Vol. 51, No. 6, June 2003, pp. 1787–1805

Jürgen Kampe: *Alternative Konzepte diodengeschalteter Abtastanordnungen als Empfangsteil für die HF- und Mikrowellen-Signalmeßtechnik.* PhD Thesis, Technische Universität Ilmenau, 1993

Kaneyuki Kurokawa: *Injection Locking of Microwave Solid-State Oscillators.* Proceedings of the IEEE, Vol. 61, No. 10, October 1973, pp. 1386–1410

Jeong S. Lee and Cam Nguyen: *A Low-Cost Uniplanar Sampling Down-Converter with Internal Local Oscillator, Pulse Generator, and IF Amplifier.* IEEE Transactions on Microwave Theory and Techniques, Vol. 49, No. 2, February 2001, pp. 390–392

Jeong S. Lee, Cam Nguyen and Tom Scullion: *New Uniplanar Subnanosecond Monocycle Pulse Generator and Transformer for Time-Domain Microwave Applications.* IEEE Transactions on Microwave Theory and Techniques, Vol. 49, No. 6, June 2001, pp. 1126–1129

K. Madani and C. S. Aitchison: *A 20 GHz Microwave Sampler.* IEEE Transactions on Microwave Theory and Techniques, Vol. 40, No. 10, October 1992, pp. 1960–1963

Roger B. Marks and Dylan F. Williams: *Comments on “Conversion Between S , Z , Y , h , $ABCD$, and T Parameters which are Valid for Complex Source and Load Impedances”.* IEEE Transactions on Microwave Theory and Techniques, Vol. 43, No. 4, April 1995, pp. 914–915

Nhat M. Nguyen and Robert G. Meyer: *Start-up and Frequency Stability in High-Frequency Oscillators.* IEEE Journal of Solid-State Circuits, Vol. 27, No 5, May 1992, pp. 810–820

Rainer Pauli: *On complex normalized scattering matrices.* IEEE Transactions on Circuits and Systems, Vol. 37, Issue 11, November 1990, pp. 1452–1454

Paul Penfield, Jr., Robert Spence, and Simon Duinker: *Tellegen's Theorem and Electrical Networks*. The M.I.T. Press, 1970

G. G. Raleigh and J. V. Bellantoni: *Determination of Schottky Diode Sampling Mixer Frequency Response From Diode Conductance Waveforms*. IEEE Transactions on Microwave Theory and Techniques, Vol. 41, No. 4, April 1993, pp. 723–726

Mitch Randall and Terry Hock: *General Oscillator Characterization Using Linear Open-Loop S-Parameters*. IEEE Transactions on Microwave Theory and Techniques, Vol. 49, No. 6, June 2001, pp. 1094–1100

Mark J. W. Rodwell, Masayuki Kamegawa, Ruai Yu, Michael Case, Eric Carman, and Kirk S. Giboney: *GaAs Nonlinear Transmission Lines for Picosecond Pulse Generation and Millimeter-Wave Sampling*. IEEE Transactions on Microwave Theory and Techniques, Vol. 39, No. 7, July 1991, pp. 1194–1204

Burkhard Schiek: *Grundlagen der Hochfrequenz-Messtechnik*. Springer-Verlag, 1999

Dylan F. Williams and Kate A. Remley: *Analytic Sampling-Circuit Model*. IEEE Transactions on Microwave Theory and Techniques, Vol. 49, No. 6, June 2001, pp. 1013–1019

Martin R. Wohlers: *Complex Normalization of Scattering Matrices and the Problem of Compatible Impedances*. IEEE Transactions on Circuit Theory, Vol. CT-12, No. 4, December 1965, pp. 528–535

Martin R. Wohlers: *On Scattering Marices Normalized to Active n -Ports at Real Frequencies*. IEEE Transactions on Circuit Theory, Vol. 16, Issue 2, May 1969, pp. 254–256

Yi-Sheng Zhu and Wai-Kai Chen: *On Complex-Normalized Reflection Coefficients of a Terminated Lossless Two-Port Network*. IEEE International Symposium on Circuits and Systems 1988, Vol. 2, June 1988, pp. 1187–1190

74LVC14A Datasheet, Philips Semiconductors, Rev. 04 February 2005

AD8051/AD8052/AD8054 Datasheet, Analog Devices, Rev. F 2004

AD817 Datasheet, Analog Devices, Rev. B 1995

AD8302 Datasheet, Analog Devices, Rev. A 2002

AD8651/AD8652 Datasheet, Analog Devices, Rev. B 2004

AD8671/AD8672/AD8674 Datasheet, Analog Devices, Rev. C 2005

ADF4252 Datasheet, Analog Devices, Rev. B 2003

ADF4252 Evaluation Board Datasheet, Analog Devices, Rev. O 05/02

ATC600L Series Datasheet, American Technical Ceramics, Rev. E 4/06

ATC600S Series Datasheet, American Technical Ceramics, Rev. J 10/05

BAT15... Datasheet, Infineon Technologies, 19. November 2004

BB837/BB857 Datasheet, Infineon Technologies, 7. November 2002

BFG520W Datasheet, Philips Semiconductors, 2. October 1998

BFP450 Datasheet, Infineon Technologies, 20. August 2001

BFP540F Datasheet, Infineon Technologies, 28. January 2004

RO4000 Series High Frequency Circuit Materials Datasheet, Rogers Corporation, 3. March 2003