Università degli Studi di Cagliari

DOTTORATO DI RICERCA

In Ingegneria Elettronica ed Informatica

Ciclo XXI

Settore/i scientifico disciplinari di afferenza

ING-INF/01 – Elettronica

Titolo Tesi

New design methodologies for Microwave Oscillators based on Negative Impedances. Study and Development of the Solution Space Concept.

Presentata da:………… Ing. Monni Marco

Coordinatore Dottorato: Prof. Alessandro Giua

Relatore:……………….. Prof. Ing. Giovanni Martines

Esame finale a.a 2007/2008

Eccoci arrivati all’unica pagina in Italiano di questo lavoro….
Beh mi sembra doveroso ringraziare persone che hanno reso questo cammino di lavoro un pò meno faticoso a molto spesso più allegro.
In primo luogo mi sembra doveroso ringraziare il mio ex prof di tesi Massimo Barbaro e il mio Tutor Giovanni Martines per avermi dato la possibilità di partecipare a questo interessante ambito di ricerca. Poi vorrei ringraziare tutti i colleghi del DIEE per l’allegria e la compagnia di questi 3 anni, Andrea A , Andrea, GianNicola , Paolo, Roberto, Gianmarco, Gianni, Francesca, Alessandra Daniela, Sabrina…Grazie per le pause caffè (macchinetta mi mancherai tanto)..per i pranzi al poetto e i panini caddotzi negli assolati venerdì primaverili quando il cervello andava in malora…A Roberto in particolare per avermi fatto conoscere il panino con EQUINA e CAPRESE che è entrato ormai nel mio menù… Vorrei ringraziare anche i tesisti che mi hanno sopportato e con i quali sono stati raggiunti dei notevoli traguardi. GRAZIE Sandro, Francesco e M. Luisa.
Poi vorrei ringraziare la mia chitarra e i miei maestri Massimo e Riccardo per avermi fatto diventare un “drago di ordine superiore”, nonché i miei amici delle giovani Marmitte per gli sfoghi e i bei momenti passati insieme. Un saluto alla mia rock band!! Ciao Rubiera!!!
Vorrei ringraziare inoltre i miei genitori per avermi supportato come dei veri genitori in tutto questo periodo di vita e il mio fratello Gianluca per tutti i bei momenti che passiamo insieme, vi voglio bene.
Una zampa al mio cagnaccio Milo per tutto l’affetto che mi da.
Il dottorato è stato bello soprattutto perché mi ha dato la possibilità di conoscere una persona veramente fantastica della quale mi sono innamorato…Ora stiamo insieme e nonostante i momenti difficili traiamo forza l’uno dall’altra e stiamo veramente bene …Grazie Lolla per tutto quello che fai per me!

INDEX

CHAPER1

- The nature of waves ……………………………………………………………………………1
- The Sinusoidal waves and the Fourier Transformation- RF allocation chart……………..2-6
- Modulations………………………………………………………………………………....6-13
- Microwave Circuit elements; Lumped or Distributed?...14-15
- The Smith Chart, the reflection coefficient, the bilinear transformation and the characteristic impedance……………………………………………………………………………15-23
- The resonance in the Smith Chart and the Instability…………………………………...23-24
- References……………………………………………………………………………..24

CHAPTER2

- The GaAs Technology…………………………………………………………………..25-26
- The GaAs material properties…………………………………………………………...26-30
- The varactor as junction characteristic of GaAs…………………………………......28-38
- Metal Semiconductor –GaAs-Field-Effect Transistor…………………………………..38-46
- Biasing transistor Networks and operations mode for MW transistors…………………46-55
- Characteristics of passive GaAs components………………………………………......55-79
- The GaAs and CMOS processes review………………………………………………80-109
- References……………………………………………………………………………109-110

CHAPTER3

- Oscillators General Theory and Classification…………………………………………….111
- The Harmonic Balance Simulator…………………………………………………113-117
- Oscillator analysis using probes……………………………………………………117-118
- Single transistor Oscillator analysis using reflection coefficients of the device and resonant load: The Start-Up……………………………………………………………………118-134
- The active network small signal models and their combination with resonant circuits134-138
- The State of the art in the design of single transistor negative impedance oscillators..138-145
- A different way to consider oscillators: design examples……………………………145-156
- References……………………………………………………………………………157

CHAPTER4

- The Phase noise……………………………………………………………………158-159
- General Considerations……………………………………………………………159-161
- Phase noise of ideal oscillator…………………………………………………………...161-164
- A Linear Time Varying Phase Noise Theory…………………………………………164-171
- Circuit Examples……………………………………………………………………...171-175
- Conclusion………………………………………………………………………175-178
- References……………………………………………………………………………178

CHAPTER5

- An application example of State of The Art in Microwave Negative Impedance Oscillators Design.……………………………………………………………..…………………179-190
- The Solution Space Approach: the basic idea and generation………………………..190-196
- The PFOM and the decrypting of Solution Space information (The new design method)……………………………………………………………………………196-199
- An applicative Example: the design of a 38GHz VCO in GaAs 0.2μm technology……199-205
- CONCLUSION…………………………………………………………………205-206
- References…………………………………………………………………………206

CHAPTER6

- The oscillator tool……………………………………………………………207-211

APPENDIX A

- The analysis of resonant series circuits………………………………………………212-215

APPENDIX B

- The loaded Quality factor…………………………………………………………216-221

APPENDIX C

- Synthesis procedure for microwave resonant circuits……………………………….222-228

In questo lavoro le figure sono numerate in ordine acescente per capitolo.
Ogni capitolo ha una numerazione delle figure che inizia da 1 a seguire.
Anche le formule sono numerate nel medesimo modo.
Laddove in diversi capitoli si fa riferimento a diverse formule e figure provenienti da diversi capitoli esse sono richiamate facendo riferimento al capitolo di provenienza ex…looking the figure number 2 of charter one, cosi come per le formule. Questo per una questione di praticita e consecutzio logico nella numerazione dato che in questa tesi troverete degli argomenti di diversa provenienza.

Buon Lavoro

Marco Monni

The nature of the waves

Radio transmission is a young science. It had beginnings by the theoretical works of Mawell in the nineteenth century and by the experiment of Hertz, but there are a lot of people that given their contribute including the development of devices which could detect the presence of radio waves. The question of who first transmitted radio signals is attributed to Guglielmo Marconi who was born in Italy in 1874, and his first British patent was taken out in 1896 and covered the use of speak transmitter. By the 1921, the thermo ionic transmitter tube became available and made possible to design a transmitter to operate on a range of frequencies. The power output available increased with the development of electronic tubes which could, increasingly, handle higher powers with the aid of air or liquid cooling systems. Over the years and stimulated by the need of the First and Second World Wars, radio transmission has become an established technology which is taken for granted and which, among other things, provides for the broadcasting to our homes of entertainment, news and information of every kind in both the radio and television spheres. The most recent development, resulting in the domestic satellite dish antenna, brings the quasi optical nature of microwaves to the notice of the consumer.

The use of semiconductor devices (transistors) has become common-place and as a result the mass and volume of electronic products for a given function is much less than that of their earlier counterparts which used electronic tubes. However, in the high power transmitter field electronic tubes are still the mainstay of transmitters. These use very high voltages, depending on power output. Normally the range of voltage is around 40kV for high power equipments. In last years the technology of tubes improve through the applications of new nanotechnologies and the discover of cool cathodes.

Semiconductor devices are being to used in trasmitters of more modest power and also in spaced array radar equipments and do not need high voltages. Semiconductors are obviously a considerable role in transmitter drives, audio circuits, control systems, and thanks to the possibility to implement in small areas an high number of transistors it is possible to produce systems on chip (S.O.C) which contains more elements integrated as mixed-signal circuits namely: low noise and power amplifiers mixer oscillators, data converters and digital signal processors for a whole processing of radio signals.

Whit the vast increase of terrestrial and satellite broadcasting and communications, and enormous number of mobile phones now in use, homes, work and recreational places are irradiated by a vast number of electromagnetic signals. Many are intended to operate receiving equipment, most of which are at very low levels because the high sensitivity of receiver does not necessitate large signals. Mobile phones do however communicate both was and thus incorporate transmitters and receivers. As usage increases there is the pressure for the use of more frequencies such that governments now sell licences to use parts of the RF spectrum.

The word ‘wireless’ largely passed out of use many years ago. Radio is now the more general term in use, though strangely enough in domestic use, mainly being interpreted as meaning sound broadcasting with the term ‘television’ or ‘TV’ to describe television pictures and broadcasting. There are many words used to describe forms of radio system including satellite communications, radar, microwave links, mobile telephones etc.

Despite the profusion of terms in use to describe the transmission of intelligence by electromagnetic waves, the nature of these waves is basically the same, the variable being the way in which the intelligence (signal) is added. It is therefore convenient to refer to these electromagnetic waves as ‘radio waves’ and the frequencies of the waves as ‘radio frequencies’.

The Sinusoidal waves and the Fourier Transformation- RF allocation chart.

As we will see in other chapters an important class of signal in RF design it that of sinusoidal signals of both voltage and current. A generic sinusoidal waveform can be depicted as showed in the figure below:

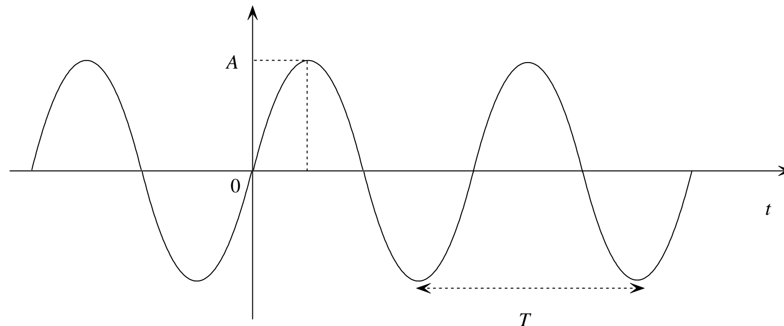


Figure 1: A sinusoidal waveform

An electronic sinusoidal waveform can be represent the trade parameterized in time of a generic electrical signal and its nature can be of both voltage and current.
In the general form we assume that the signal has the nature of a voltage and we will express them as

$$v(t) = A\sin(\omega_0 t + \phi) \qquad (1)$$

where A is the amplitude and ω_0 the radiant frequency. The phase ϕ represents the fact that the peak can be shifted from the time origin. Equation (1) implies that v(t) repeat itself for all time. The repetition of the period is calculated as the distance between two peaks and indicated by means of T_0. The reciprocal of the period is the frequency and it is defined as (2) and measured in Hz.

$$f_0 = \frac{1}{T_0} = \frac{\omega_0}{2\pi} \qquad (2)$$

Obviously, no real signals goes on forever, but Eq.(1) could be a reasonable model for a sinusoidal waveform that lasts a long time compared to the period. The AC steady-state circuit analysis depends upon the assumption of an external sinusoid waveform usually represented by a complex exponential phasor.

Usually a generic sinusoidal signal can be represented by a complex exponential quantity or a phasor by using the Eulero's theorem. If $\theta=\omega_0 t+\phi$ a generic signal can be expressed as a phasor for example:

$$v(t) = A\cos(\omega_0 t + \phi) = A\left[\mathrm{Re}(e^{j\theta}) + j\,\mathrm{Im}(e^{j\theta})\right] \qquad (3)$$

The representation of an sinusoidal signal in the frequency domain is achieved through the Fourier's transformation defined as:

$$F(\omega) = \int_{t=-\infty}^{+\infty} e^{-j\omega t} v(t)\,dt \qquad (4)$$

To compute the Fourier transform of a signal it is indispensable to introduce the Dirac impulse function

$$\int_{t=-\infty}^{+\infty} e^{-j\omega t} dt = \delta(t)$$

Where δ(t) is represented in the time domain by a infinitely tight impulse.
Therefore if we are going to compute the Fourier transform of a generic signal $v(t) = V_a \sin(\omega_0 t)$ by using the Eulero's formula we have:

$$v(t) = V_a \sin(\omega_0 t) = V_a \frac{e^{j\alpha} - e^{-j\alpha}}{2} \Rightarrow$$

$$V(\omega) = V_a \int_{t=-\infty}^{t=+\infty} e^{-j\omega} \sin(\omega_0 t) dt = -j\frac{V_a}{2} \int_{t=-\infty}^{t=+\infty} e^{-j2\pi f t}\left(e^{j2\pi f_0 t} - e^{-j2\pi f_0 t}\right) dt = -j\frac{V_a}{2} \int_{t=-\infty}^{t=+\infty} e^{-j(2\pi f - 2\pi f_0)t} dt + j\frac{V_a}{2} \int_{t=-\infty}^{t=+\infty} e^{-j(2\pi f + 2\pi f_0)t} dt$$

$$-j\frac{V_a}{2}\delta(f - f_0) + j\frac{V_a}{2}\delta(f + f_0)$$

That can be sketched as:

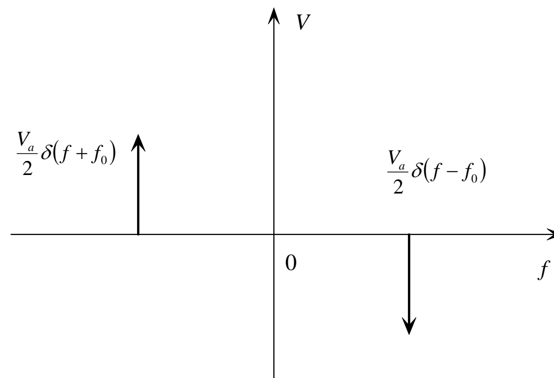


Figure 2

Fig(2) present two impulse for a sinusoidal signal. The negative impulse does not have physically means and the positive impulse have the amplitude inversed, that derives from mathematical properties of the Fourier transform.
Physically a sinusoidal in the time domain is presented by a delta impulse in the frequency domain (the amplitude must be pick up as positive).
In this thesis I will treats new methodologies to design microwave oscillators, an oscillator is an autonomus circuit, namely a circuit that does not contain any signal source and does not treat any signal source but only DC-BIAS sources and convert them in a sinusoidal signal ad a definite amplitude A_0 and frequency f_0.
Before to introduce this techniques it is very important to classify the frequencies ranges in the RF bandwidths and to introduce the physical mechanism of wave propagation in the space. By the physic it is note that a current flowing in a conductor gives rise to a magnetic field around it. When such a current is varying, it gives rise to a similarly changing electric field. Similarly a changing electric field will give rise to a magnetic field. Unchanging fields of either kind will not result in the production of the other kind of the field. With changing field and electric field are thus inextricably linked. Hence alternating currents and voltages do, by definition, involve time-varying fields.

It is easy to imagine that from any source of such fields some energy may be unintentionally released (transmitted) into the free space, causing interference with receivers or other equipment, without necessarily understanding the phenomenon. This is because such interference has been

experienced by most people in they every lives. Perhaps the most common example is the motor car ignition system which can also prove to be a rudimentary example of the spark transmitter.

In the case of radio transmitters, however, the whole intention is to transmit RF energy into free space and the antenna used to do so is specifically designed to achieve this objective. The antenna are transducer of a electrical signals in electromagnetic signals and viceversa. The Electric field (E) at any point is defined as the force acting on a unit positive charge at that point. The magnitude of the electric field is expressed in volt per metre (Vm^{-1}). The magnetic fields at a point is also a force and is defined as the force which could act on an isolated north pole at that point. The classic demonstration of this is that the earth's magnetic field exerts a force on a compass needle, to the great blessing of navigators. The ampere is defines on the basis of the magnetic force exerted when a current flows in a conductor and magnetic fields strength is measured in ampere per metre (Am^{-1}).

Being forces, both quantities are vector quantities having magnitude and direction. The Ohm's law equations for power when the voltage and current are in phase (plane wave conditions) can be used in an analogous way and with the same phase qualification to calculate power density.
Since wave are propagated into the free space it is necessary to introduce the concept of frère space impedance defined as:

$$Z_0 = \sqrt{\frac{\varepsilon_0}{\mu_0}} = \sqrt{\frac{4\pi \cdot 10^{-7}}{8.854 \cdot 19^{-12}}} = 376.7\Omega \approx 377\Omega \qquad (5)$$

Where ε_0 and μ_0 are respectively the permeability and the permittivity of free space. Electromagnetic waves that propagated in free space have the electric and magnetic fields perpendicular to each others and to the direction of propagation represented in the figure below:

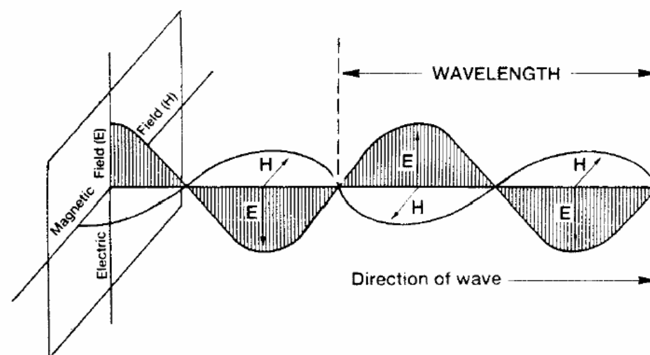


Figure 3

The plane of polarization of a wave is, by convention, that of the electric field, for example the polarization in the figure 3 is vertical. This convention has the advantage that for a vertical polarization the antenna will also vertical. Of the diagram is rotated until the electric field is horizontal then the wave polarization is horizontal. Apart from linear polarization, other forms such as circulator or elliptical polarisation are also used for specific purposes. There is another approach to RF radiation where by the concept of particles (photons) is used to describe the radiated signal. However, for the purposes of this work , the wave concept seems to serve the purpose best and is generally so used.
The main characteristics of a wave are frequency and wavelength (λ). The frequency are already defined by relationship (2), instead to define the wavelength we need to introduce the velocity of propagation for the wave.

The velocity of propagation of all electromagnetic waves are indicate by means of (c) and it is a constant in a given homogeneous medium. In the free space it is equal to the light velocity defined as $2.997925 \cdot 10^8$ ms^{-1}. If the wave travel in a different medium for which change the dielectric constant ε_r then the velocity are reduced by the factor $\varepsilon_r^{(1/2)}$.

The velocity is defined as

$$c = f\lambda \qquad (6)$$

Depends by frequency there are a classification of waves, and the RF space is divided in frequency bands used for specifics functions. Actually this is a most confused topic and the frequency allocation bandwidth are not well standardised but depends by laws of the governments in different countries.
A generally accepted allocation band comprise the frequencies in the range [300Hz-300GHz] and is summarized in the following table:

Frequency	Band Code	Band Description
300 Hz-3 KHz	ELF	Extra Low Frequency
3 KHz-30 KHz	VLF	Very Low Frequency
30 KHz-300 KHz	LF	Low Frequency
300 KHz-3 MHz	MF	Medium Frequency
3 MHz-30 MHz	HF	High Frequency
30 MHz-300 MHz	VHF	Very High Frequency
300 MHz- 3 GHz	UHF	Ultra High Frequency
3 GHz-30 GHz	SHF	Super High Frequency
30 GHz- 300 GHz	EHF	Extra High Frequency

In microwave engineering field the bands between [1-40] GHz are divided in sub-band and usually indicated with a letter

Frequency- GHz	Band Letter
1-2	L
2-4	S
4-8	C
8-12.5	X
12-5-18	Ku
18-26.5	K
26.5-40	Ka

The frequency bandwidth presented here are only one small portion of the electromagnetic spectrum. A sketch that represent all parts is presented in the figure below

IONISING	ULTRAVIOLET			VISIBLE LIGHT	INFRARED	
Ionising Radiations	UVC	UVB	UVA	Visible	Infrared	Radio Frequencies

WAVELENGTH 100 nm 280 nm 315 nm 400 nm 780 nm 1 mm
300 GHz

Figure 4

A very good explanation for allocation chart of the U.S are reported in the following allocation chart
www.ntia.doc.gov/osmhome/allochrt.pdf

UNITED STATES FREQUENCY ALLOCATIONS

THE RADIO SPECTRUM

Modulations

When a wave of a given frequency is radiated continuously, i.e. a continuous series of sinusoidal waves, no intelligence is conveyed and the signal is called a carrier. This mode of transmission is known as continuous wave (CW). Nothing can be heard unless there is a local oscillator to beat with carrier and produce a note at the different frequency. This is referred to as heterodyning. If the carrier is switched on and off in accordance with some kind of code, e.g. morse code, then this intelligence can be interpreted. More generally, for broadcasting the intelligence may be speech, music, and television pictures. Other professional work include voice and data transmission by a variety of methods, radar transmitters transmit RF signals in a series of pulses and so on. Modulation and coding are operations performed at the transmitter to achieve efficient reliable information transmission. Modulation involves two waveforms: a modulating signal that represent the message and a carrier wave generated by the local oscillator. The modulation translate the information that is contained at a lower frequency at an high frequency necessary to transmission

and it is a reversible process , so the message can be retrieved by the complementary process of demodulation. The first purpose of modulation in a communication system is to generate a modulated signal suited to the characteristics of the transmission channel. Actually there are several practical benefits and applications of modulation briefly discussed below.

a) Efficient Transmission: Signal transmission over appreciable distance always involves a travelling electromagnetic wave, with or without a guiding medium. The efficiency of any particular transmission method depends upon the frequency of the signal being transmitted. By exploiting the frequency-translation property of a CW modulation, message information can be impressed on a carrier whose frequency has been selected for the desired transmission period. An efficient transmission requires antennas which dimensions are almost 1/10 of the wavelength. Unmodulated transmission of an audio signal containing frequencies below 100Hz would require antennas of 300Km long. Modulate transmission at 100MHz allows to reduce antenna size about one meter.
b) Another benefit it consist in allow to the designer to place a signal in some frequency places that have an hardware limitations, and allows also multiplexing process in order to combining several signals for simultaneous transmission on one channel. For example Frequency division multiplexing (FDM) uses CW modulations to put each signal on a different carrier frequency, and bank of filters to separate the signals at destination. Time division multiplexing (TDM) uses pulse modulation to put samples of different signals in non overlapping slots.

A very exhaustive discussion about different techniques for modulation can be founded in [1]. For our purpose it is important only understand the modulations by a practical point of view in order to reach the design constraints for oscillator circuits. To do this we are going to show some examples of analog and digital modulations achievable through the use of ADS simulator.

AM MODULATION

The first example of modulation is the amplitude modulation. In the AM the amplitude of a carrier is varied by using a modulating signal of definite amplitude and frequency. The basic frequency of modulating signal must be smallest than of the carrier. Mathematically we can write

$$v_m(t) = V_m \cos\omega_m t$$
$$v_p(t) = V_p \cos\omega_p t$$
$$v_{AM} = \left(V_p + V_m \cos\omega_m t\right)\cdot \cos\omega_p t =$$
$$V_p \cos\omega_p t + V_m \cos\omega_m t \cos\omega_p t = V_p \cos\omega_p t + \frac{V_m}{2}\left[\cos\left(\omega_m t \omega_p t -\right) + \cos\left(\omega_m t + \omega_p t\right)\right]$$

The ratio between the amplitude of signals is defined as modulation index and indicated with m then we can write

$$m = \frac{V_m}{V_p} \rightarrow V_m = mV_p \Rightarrow$$
$$V_{AM} = V_p \cos\omega_p t + mV_p \frac{1}{2}\left[\cos\left(\omega_p t - \omega_m t\right) + \cos\left(\omega_p t + \omega_m t\right)\right] =$$
$$= V_p \cos\omega_p t + mV_p \frac{1}{2}\cos\left(\omega_p t - \omega_m t\right) + mV_p \frac{1}{2}\cos\left(\omega_p t + \omega_m t\right)$$

In others words as a sum of three sinusoidal signals. In the frequency domain the spectra is achieved in easy way by means of the Fourier transform function

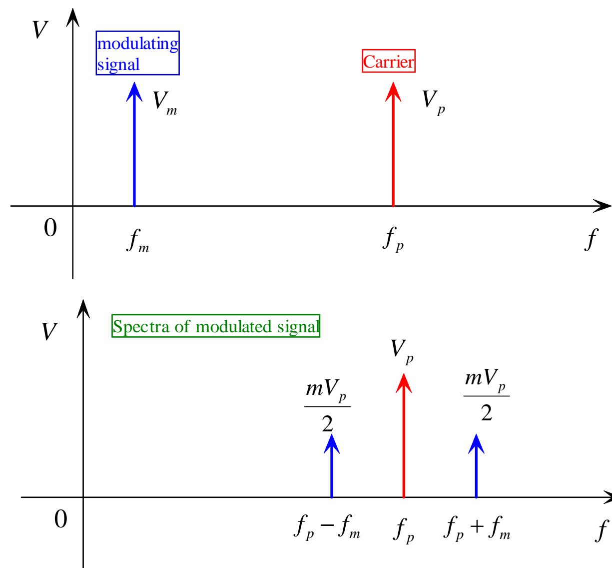


Figure 5

In ADS there are a library that contain modulation systems for example it is possible use an amplitude modulator and two time domain signal in order to simulate the signal at the output of the modulator

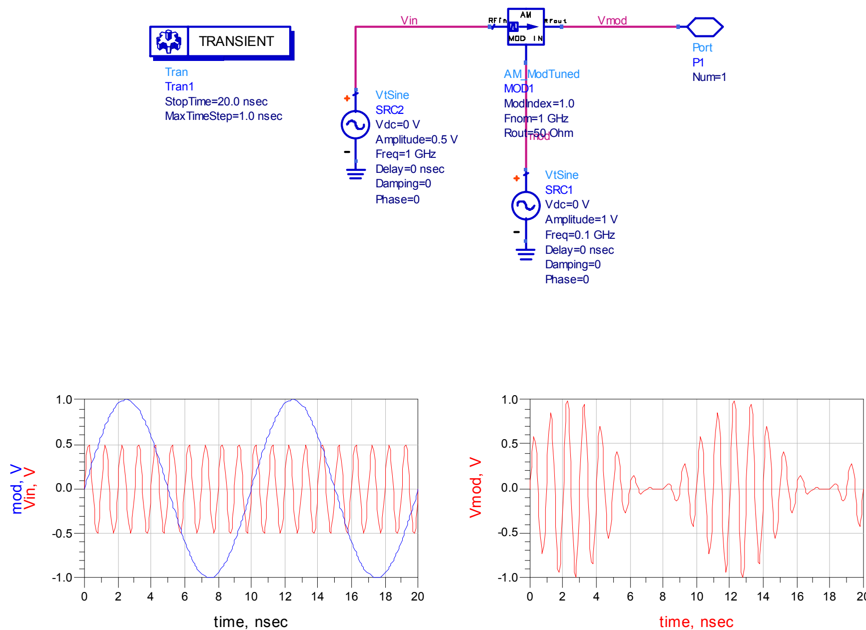


Figure 6

By using also the demodulator it is possible to recover the information contained in the modulated signals that are transmitted by operation implemented in modulation. Usually the carrier are necessary to transmit the signal whereas the modulating signal is what contain information and modulated signal will be transmitted by antenna. Also demodulation operation the signal at the output will have the following form

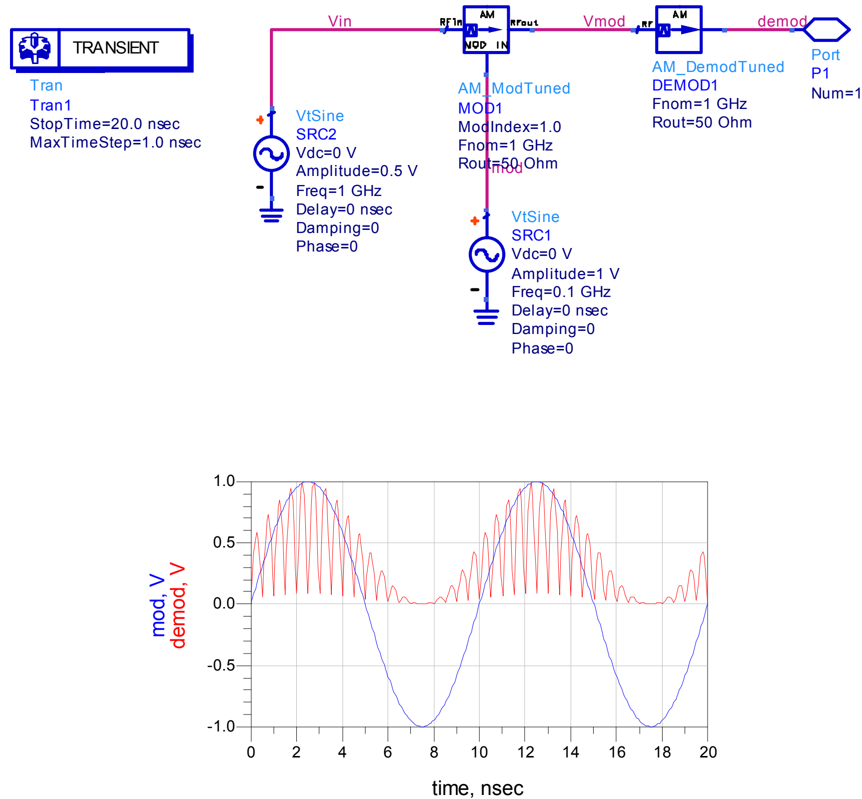


Figure 7

FM MODULATION

A more common way to modulate a signal is that of frequency modulation. The FM is better than AM especially because it has a greater robustness compared to noise.
Mathematically the FM modulation can be treated as follow:

$$v_m(t) = V_m \cos\omega_m t$$
$$v_p(t) = V_p \cos\omega_p t$$
$$\omega_m >> \omega_p$$

The first thing it consist in calculate the instant pulsation

$$\omega_{FM} = \omega_p + K_F V_m \cos(\omega_m t) = 2\pi\left(f_p + \frac{K_F V_m}{2\pi}\cos\omega_m t\right) = 2\pi\left(f_p + \Delta f \cos\omega_m t\right)$$

$$\Delta_f = \frac{K_F V_m}{2\pi}; \omega(t) = \frac{d\varphi(t)}{dt}; d\varphi(t) = \omega(t)dt$$

$$\varphi(t) = \int\left(\omega_p + K_F V_m \cos\omega_m t\right)dt = \omega_p t + \frac{K_F V_m}{\omega_{m,}}\sin\omega_m t$$

$$V_{FM}(t) = V_p \cos\left(\omega_p t + \frac{K_F V_m}{\omega_m}\sin\omega_m t\right) = V_p \cos(\omega_p t + m\sin\omega_m t)$$

$$m = \frac{K_F V_m}{\omega_m} = \frac{K_F V_m}{2\pi f_m} = \frac{\Delta f}{f_m}$$

$$V_{FM}(t) = V_p \cos(\omega_p t + m\sin\omega_m t)$$

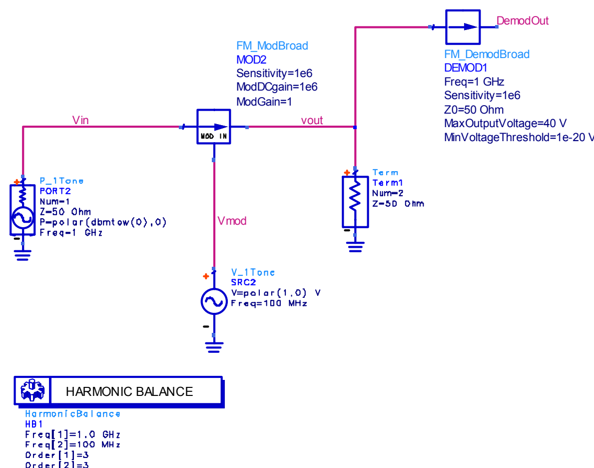
By using the Bessel Series it is possible to show as a FM modulated signal is represented by the sum of a infinite number of sinusoids

$$V_{FM}(t) = V_p J_0 m \sin(\omega_m t) + V_p J_1 m\left[\sin(\omega_p + \omega_m)t - \sin(\omega_p - \omega_m)t\right] + V_p J_2 m\left[\sin(\omega_p + 2\omega_m)t - \sin(\omega_p - 2\omega_m)t\right] + ..$$

The bandwidth of a FM modulate signal is defined as the frequencies with amplitude greater than 1% of non modulated signal. In order to estimate the bandwidth is possible to use the Carson formula

$$B = 2(\Delta f + f_m \max)$$

Where Δf is the offset compared to the carrier and f_m max is the maximum modulating frequency.
In order to simulate a simple FM modulation system in ADS we need to use the HB simulator setting two different frequencies for the modulating and for the carrier signals, or Envelope simulations. For more information of FM modulation tools look the Agilent Knowledge centre.
An example is showed below



It is difficult to see the trade of the modulated signal in the time domain because the simulator used is the HB in the frequency domain. Intuitively when the modulating signal is greater than 0 the modulated signal has a great number of zero crossing while when is less than 0 the modulated signal has a minor frequency.

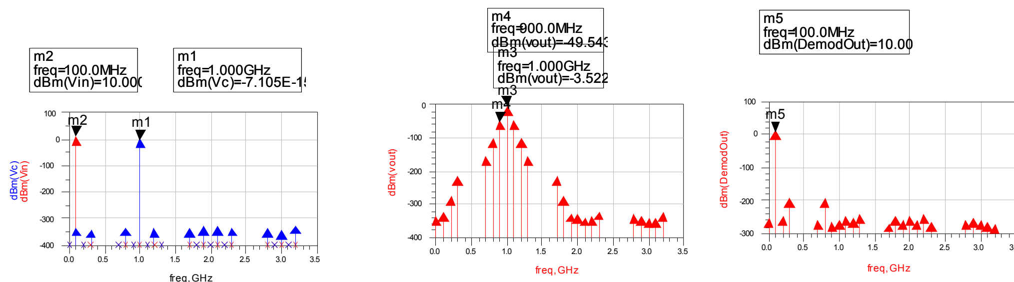
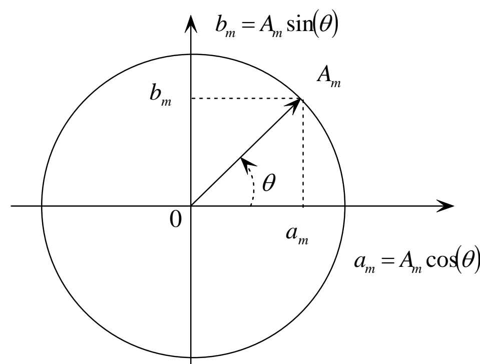


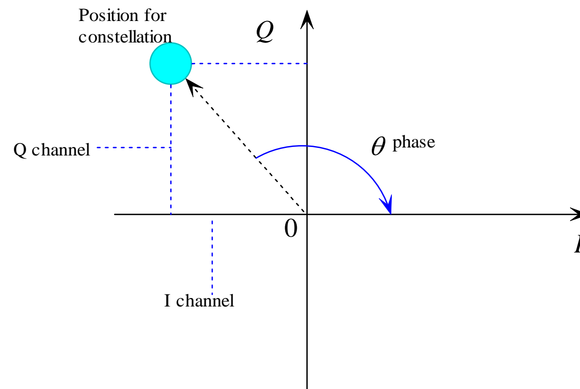
Figure 8

IQ MODULATION

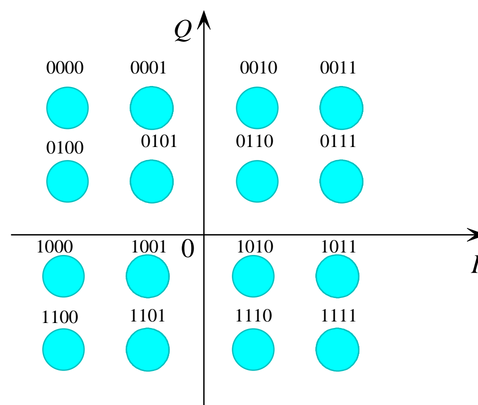
For modern wireless systems is used a IQ modulation called also phase and quadrature modulation. This kind of modulation is based on the fact that you can distinguish two different signals called X1(t) and X2(t) after a multiplication respectively for a cosine function (phase modulation) and for a sine function (quadrature modulation) ands sum the resulting signals. The terms modulation in phase and quadrature are referred to the fact that the sin and cos functions have a phase difference of 90°. The frequency f_0 of cosinusoidal modulating signal is called phase carrier and the frequency of sinusoidal modulating signal is called quadrature carrier. The IQ modulation is used to define a QAM known as quadrature amplitude modulation that is explained in the figure below. The IQ modulation is an example of numerical modulation and then follow a digitalization of the analogic signal.



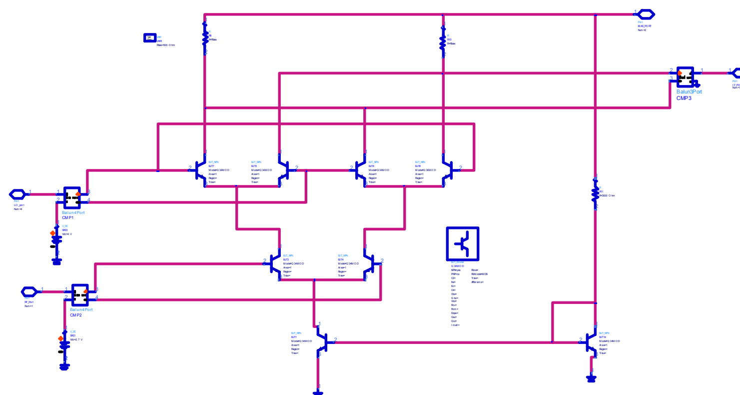
The two carriers have a difference of phase of 90°. The PM and PSK can be considered as particular cases of QAM in which the amplitude of modulating signals is constant, and only the phase varies. The same is for FM and FSK that are particular cases of phase modulation.
The QAM for example generate a constellation of values that represents in binary code the value transmitted.

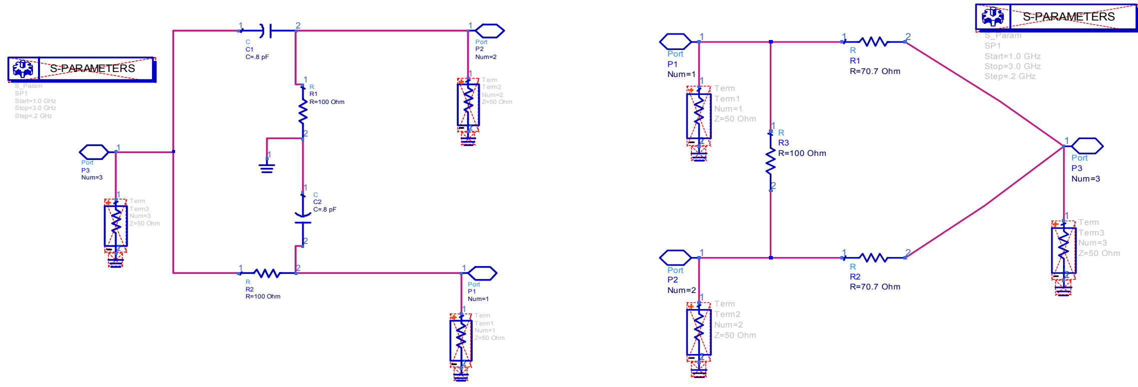


For example a 16 QAM modulation generate a constellation as that represented in the following figure



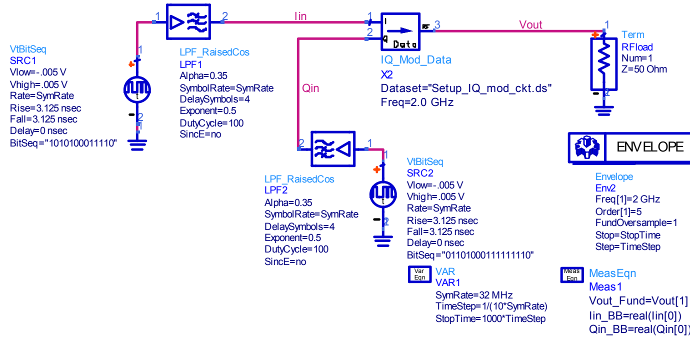
Obviously the noise degenerate the symbols. The distance between two symbols is known as Hamming's distance. An example of behavioural model for IQ modulation system can be obtained from ADS, by using clock generators that represent the converted analog to digital signals, ideals low pass filters and IQ modulator systems. At circuit detail the IQ modulator is constructed by a Gilbert cell, phase shift and Wilkinson power dividers. The schematic for these three sub-circuits are showed below





behavioral_level_QAM.dsn

behavioral_level_QAM.dsn is a schematic for a behavioral level Circuit Envelope simulation with input I_Q waveforms for a 16 QAM modulation. IQ_mod_ckt.dsn is modeled behaviorally with IQ_Mod_Data along with the dataset Setup_IQ_mod_ckt.ds (NOT circuit_level_QAM.ds). The result is in the dataset behavioral_level_QAM.ds.



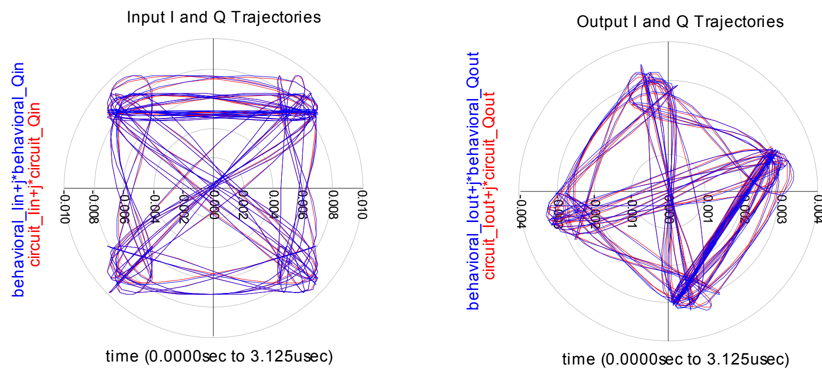
behavioral_level_QAM.dds

behavioral_level_QAM.dds compares the accuracy of the behavioral model versus the circuit level simulation for a QAM modulation schemel.

This page compares trajectory plots of the circuit and behavioral models, for both the input and output I and Q waveforms.

A comparison of the waveform plots is on the "waveforms" page.

Equations used to define the input and output I and Q signals are on the "equations" page.



Circuit level stopwatch time: 24 sec
Behavioral level stopwatch time: 5 sec
(HP 785/B2000 running HPUX 10.20)

Microwaves Circuit elements Lumped or distributed?

At frequencies where the wavelength is several orders of magnitude larger than the greatest dimensions of the circuit or system being examined, conventional circuits elements such capacitors, inductors, resistors, electron tubes or transistors are the basic building blocks for the information transmitting, receiving, and processing circuits used. The description of such circuits may be adequately carried out in terms of loop currents and nodes voltages without consideration of propagation effects. The time delay between cause and effect at different points in these circuits is so small compared with the period of the applied signal as to be negligible. It might be noted here that an electromagnetic wave propagates a distance of one wavelength in a time interval equal to one period of a sinusoidal time-varying applied signal. As a consequence, when the distance involved are short compared with a wavelength the time delay is not significant. If the wavelength has the same dimensions of the circuit then propagations effects can no longer be ignored. In most textbook it is easy to find that "for low frequencies you can treat the circuit by neglecting the propagation effects and at microwave frequencies it is better to study the transistor considering the propagative effects". Actually the bug of this claim is due to the fact of not consider the technology. In the past when for the microwave circuits only the GaAs are used to create microwave circuits and technologies as the CMOS were forbidden to high frequencies this claim may be true. Today most of RF front-ends are built in Si technology and the improvement of technology has allowed to reduce in a significant way the dimensions of the circuits. Also technologies in GaAs has been improved and scaled down ion dimensions so not all transmission lines used in design are treat as distributed lines, but also as metal connection between various elements. In conclusion in modern microwave design the distinction between lumped or distributed element cannot be done a priori, but it is possible to implement each element as a lumped or distributed depends by technology and characteristic values for the single element. For example with moderns GaAs at 24GHz an inductor can be implemented both lumped or distributed ways depends also by the value, and an LC circuit can be implemented through only a λ/4 line or a smaller line loaded with a lumped capacitance. Of course the propagative effects are most useful in MIC design in which the circuit dimensions are bigger. For example at 1.5GHz in fr-4 substrates in which the ε_r=4.5 a λ/4 may be of 20mm whereas in MMIC design these dimensions does not have sense. To make the reader able to understand the limits of lumped or distributed it is possible to look the following example in which are reported some components and their dimensions.

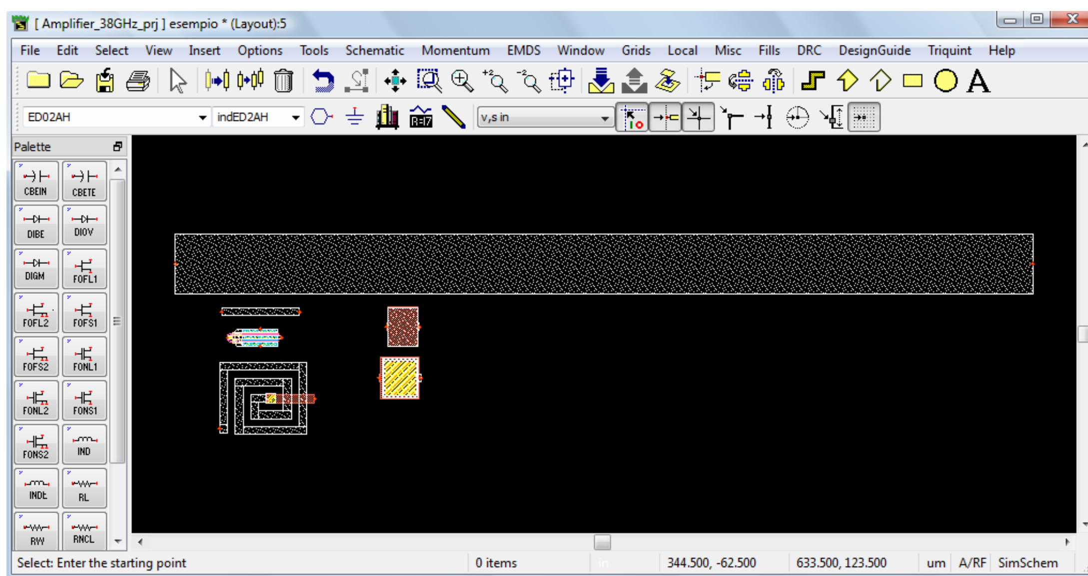


Figure 9

In this example are drown a transistor PHEMT with W=45μm and 2 fingers which channel is 0.2μm, two capacitors implemented in different way which dimensions are 1pF(yellow) and 100fF brown, an inductor of 500pH and two lines the smallest which W=10μm and L=100 μm that correspond to a line with Z_0=94Ω and 7° of electrical length and the bigger that have W=77μm and L=1103 μm and correspond to a characteristic impedance of 50Ω and a λ/4=90° of electrical length. It is possible to show as for a modern GaAs technology the lumped elements do not have the main sense in MMIC therefore it is most significant to consider both approaches in the design. In MIC design instead the dimensions of the circuit are bigger compared to MMIC and the distributed approach make sense.
For example if we wont design an amplifier at 4GHz by using a standard alumina substrate and a nec NE76184A commercial general purple transistor which datasheet can be downloaded from the link:
http://www.datasheetcatalog.com/datasheets_pdf/N/E/7/6/NE76184A.shtml

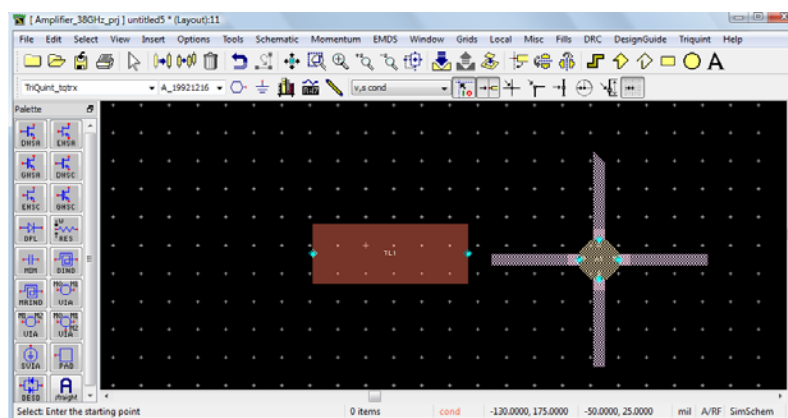
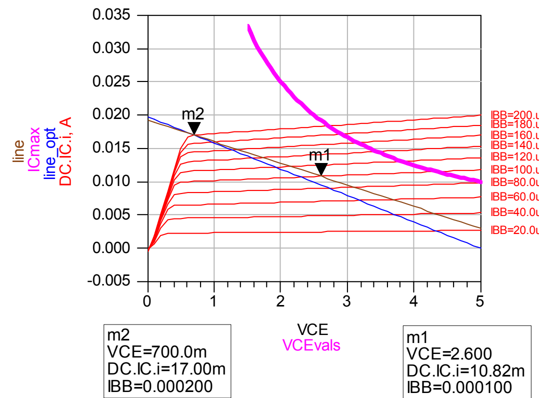


Figure 10

As results from the previously figure the transistor is much large compared with the wavelength of signal. Note the dimensions of a λ/4 transmission line of impedance 50Ω which dimensions are W=106mil and L=277mil. By definition a mil is a thousandth of a inch 1mil=0.0254mm =25.4μm.

The Smith Chart, the reflection coefficient, the bilinear transformation and the characteristic impedance.

In every case when you design at microwave frequency a most important chart in which you can represent the behaviour o the circuit is the Smith Chart that can be used to represent the impedance parameterized in frequency or the reflection coefficient.
The definition of impedance is a bit complicated because it is possible only when linear elements are studied. There are a lot of definition of impedances and in the first approximation it is possible to define the impedance for a non linear element also by looking the characteristic I/V. Intuitively a non linear impedance is a device in which at a linear variation of tension dos not correspond a linear variation of current. An example of non linear impedance is for example the transcharacteristic of a transistor that formally it is called transfer characteristic because at a linear variation of voltage does not correspond a linear variation of the current for the drain.



Generally an impedance needs to a linear behaviour, or is defined under steady-state analysis by the AC theory of circuits and under this hypothesis it is possible to represent a general uniport circuit as composed by a dissipative and reactive parts. A general one pot network can be represented as a impedance and then as Z=R+jX or admittance Y=G+jB where Z and Y indicate respectively the impedance or admittance, whereas R and G the dissipative parts called resistance and conductance and X and B the reactive parts called respectively reactance and susceptance. By point of view of physic an generic element represented by means of a impedance or admittance does not dissipate only energy but stored energy in the imaginary part.

The ratio of the energy stored and loss per cycle indicate the quality factor for the one port circuit. This definition is used for circuits one port that is characterized by a definite behaviour or inductive or capacitive.

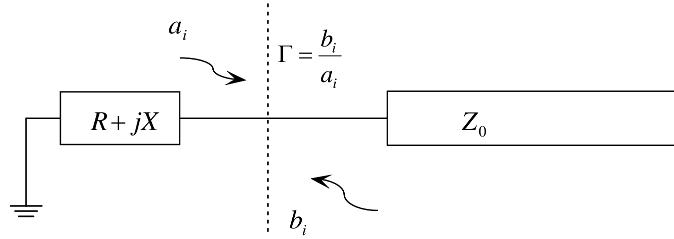
$$Q = \frac{X}{R};or;\frac{B}{G} \qquad (7)$$

To achieve the equations that allow to draw a chart which contain information about nature of devices represented it is indispensable to define the reflection coefficient.

In the previous section we have define the travelling waves in a specific medium, and we have told that the medium has a proper characteristic impedance that allows to define the propagation properties of electromagnetic field. If the travelling wave meet another medium characterized by a different value of impedance then a part of the signal coming back in the negative direction and only a portion continues the travel in the positive direction. Then it is possible to define a reflection coefficient that measures the portion of signal transmitted compared to the reflected portion. There reflection coefficient for two mediums that have two distinct values of characteristic impedance Z_1 and Z_2 is defined as

$$\Gamma = \frac{Z_1 - Z_2}{Z_1 + Z_2} \qquad (8)$$

In MIC or MMIC circuits the lines (microstrips or transmission lines, etc) the characteristic impedance assumes a specific value (50Ω,75Ω,100Ω) and then the characteristic impedances of the medium is indicated by Z_0 . When in the simple case the wave pass by a one port circuit to a transmission line which length is indefinitely so to be represented by a one port circuit (load) of Z_0 value then the reflection coefficient can be expressed as:



$$\Gamma = \frac{b_i}{a_i} = \frac{Z - Z_0}{Z + Z_0} = \frac{1}{Z_0}\frac{z-1}{z+1} \qquad (9)$$

Relationship (9) is the canonical form of a reflection coefficient normalized that can be usually employed in the design.
There a bit of confusion about reflection coefficients that derive from the difficult to distinguish when you can consider your system as a distributed or lumped. The question is dimension of my system is small compared to the wavelength then I must not use the reflection coefficients? The answer is that probably you don't need to use them, but you can. If there an advantage of reflection coefficients approach is that you can use them independently by frequency range of your circuit operate. Also if the wavelength is much bigger of dimension of your circuit you can consider the one port circuit by point of view of Γ considering them connected by means of a indefinitely length line of a specific characteristic impedance for example. Nota that in the design it is impossible to know exactly if you have a perfectly lumped or distributed circuit. But you can measure the reflection coefficient at every frequency and the value of Γ can be used to derive some important insight in for the design of a specific network that can be also implemented by means of a lumped elements.
Relationship (9) can be represented some circumferences in the polar chart that can be extracted in the following way: in first we consider the normalized reflection coefficient expression (*Γ*) and the Z_0 as a starting-point and reverse them to found the expression of the impedance in function of the reflection coefficient

$$\Gamma = \frac{z-1}{z+1} \rightarrow \Gamma(z+1) = z-1 \rightarrow \Gamma z + \Gamma = z - 1 \rightarrow -\Gamma z + z = 1 + \Gamma \leftarrow z = \frac{1+\Gamma}{1-\Gamma}$$

$$z = r + jx = \frac{1+\Gamma}{1-\Gamma} = \frac{1+\mathrm{Re}(\Gamma) + j\,\mathrm{Im}(\Gamma)}{1-\mathrm{Re}(\Gamma) - j\,\mathrm{Im}(\Gamma)} \qquad (10)$$

Then by separating real and imaginary part after rationalizing operation we can obtain two circumferences forms respectively for real and imaginary parts:

$$r + jx = \frac{1+\mathrm{Re}(\Gamma) + j\,\mathrm{Im}(\Gamma)}{1-\mathrm{Re}(\Gamma) - j\,\mathrm{Im}(\Gamma)} \cdot \frac{1-\mathrm{Re}(\Gamma) + j\,\mathrm{Im}(\Gamma)}{1-\mathrm{Re}(\Gamma) + j\,\mathrm{Im}(\Gamma)} =$$

$$= \frac{1+\mathrm{Re}(\Gamma) + j\,\mathrm{Im}(\Gamma)}{1-\mathrm{Re}(\Gamma) - j\,\mathrm{Im}(\Gamma)} \cdot \frac{1-\mathrm{Re}(\Gamma) + j\,\mathrm{Im}(\Gamma)}{1-\mathrm{Re}(\Gamma) + j\,\mathrm{Im}(\Gamma)} =$$

$$= \frac{1-\mathrm{Re}(\Gamma) + j\,\mathrm{Im}(\Gamma) + \mathrm{Re}(\Gamma) - \mathrm{Re}^2(\Gamma) + j\,\mathrm{Im}(\Gamma)\mathrm{Re}(\Gamma) + j\,\mathrm{Im}(\Gamma) - j\,\mathrm{Re}(\Gamma)\mathrm{Im}(\Gamma) - \mathrm{Im}^2(\Gamma)}{(1-\mathrm{Re}(\Gamma))^2 + \mathrm{Im}^2(\Gamma)}$$

$$\frac{1-\mathrm{Re}^2(\Gamma) - \mathrm{Im}^2(\Gamma) + 2j2\,\mathrm{Im}(\Gamma)}{(1-\mathrm{Re}(\Gamma))^2 + \mathrm{Im}^2(\Gamma)} = r + jx$$

By separating the real and imaginary parts we can achieve the equations for two circumferences in terms of Γ.

$$r = \frac{1-\mathrm{Re}^2(\Gamma)-\mathrm{Im}^2(\Gamma)}{(1-\mathrm{Re}(\Gamma))^2+\mathrm{Im}^2(\Gamma)} \rightarrow \mathrm{Re}^2(\Gamma)+\mathrm{Im}^2(\Gamma)-2\frac{r}{r+1}\mathrm{Re}(\Gamma)-\frac{r-1}{r+1}=0 \qquad (11)$$

$$x = \frac{2\,\mathrm{Im}(\Gamma)}{(1-\mathrm{Re}(\Gamma))^2+\mathrm{Im}^2(\Gamma)} \rightarrow \mathrm{Re}^2(\Gamma)+\mathrm{Im}^2(\Gamma)-2\,\mathrm{Re}(\Gamma)-\frac{2\,\mathrm{Im}(\Gamma)}{x}+1=0 \qquad (12)$$

Relationship (11) describes a circumference equation which centre (r/(r+1),0) and radius (1/(r+1)) whereas the equation (12) describes a circumference equation of centre (1,1/x) and radius 1/|x|. By keeping the real part constant in (11) and imaginary part constant in (12) we can obtain the smith Chart of impedances. At the same way it is possible to develop the Smith admittance chart and plot the circumferences obtained by keeping the conductance and susceptance at a fixed value.

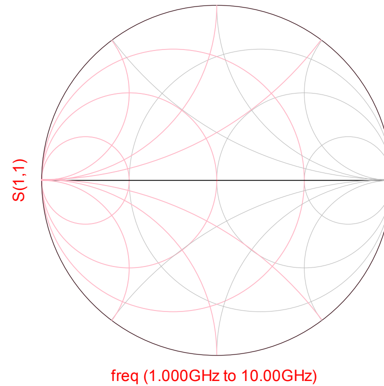


Figure 11: Impedance and admittance chart.

With The S parameters method it is possible to measure the reflection coefficient for a generic bipole and then to extract some properties for the passive or actives networks that are very interesting in the design. We start to examine the properties that can be derived from the impedances chart.

Some Interesting properties in the design for the Smith chart and S parameters.

NB all properties refer to a specific value for the characteristic impedance Z_0 value. The Z_0 is not only a fixed random value and cannot be chance arbitrarily but must be chosen with particular care in order to represent correctly the system under studio. For a typical communication system the characteristic impedance of a line is about 50Ω the and this value is common for load as well. A microstrip for MMIC technologies has a characteristic impedance value about 100Ω, depends by the width of the line and by the dielectric constant value but approximately this values is characteristic for lines with a width of 10µm. As the width of the line increase the characteristic impedance value decreases.

For a generic bipole the real part can be greater or smaller compared to the characteristic impedance value. The smaller values are represented in the left half-plane whereas the greater in the right half-plane. For the reactive parts depends by the nature. In general when the reflection coefficient is in

the upper half plane the bipole has an inductive nature, in the lower half plane a capacitive nature otherwise. A purely real part is represented in the real axis and a purely imaginary part in the unitary circles. Physically a real part not zero represent the loss in the reactive element. As will be discuss in next chapter the loss of an inductor or capacitor can be represented in first approximation by a series resistor. The impedances are often considered normalized, for real parts for example the ratio between the real part and characteristic impedance value must be considered. Obviously a ratio greater than one represent a real part bipole which resistance are greater that the characteristic impedance value whereas a ratio smaller than one means a resistive value smaller than Z_0.

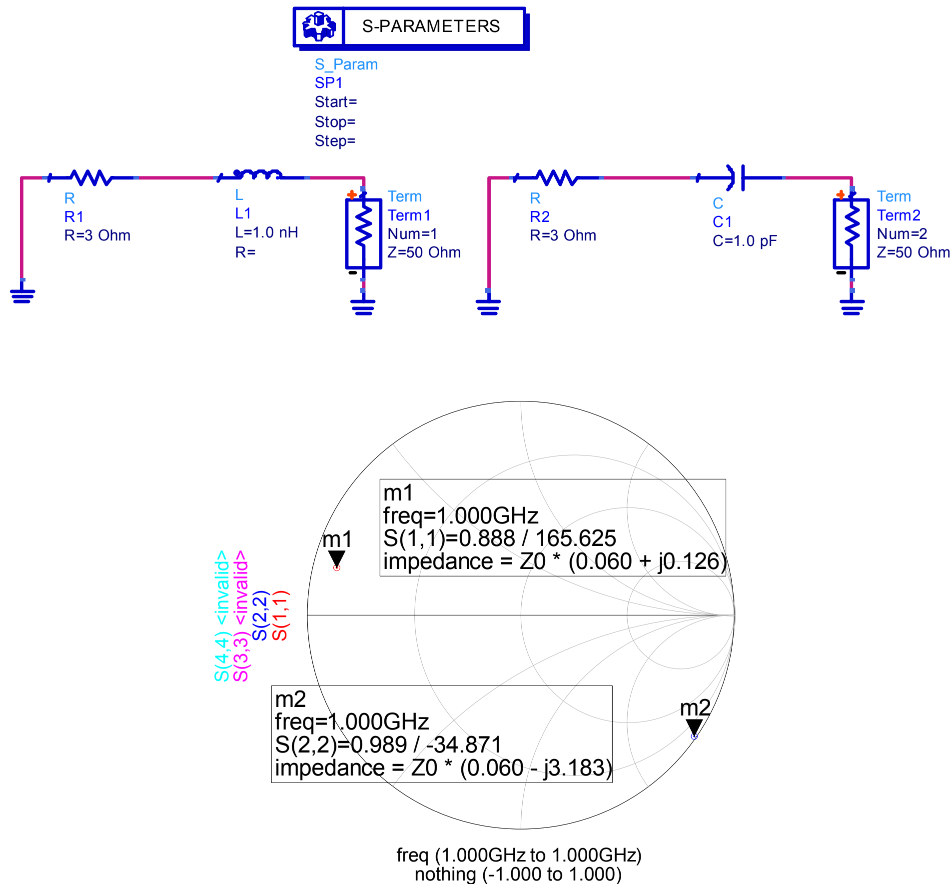


Figure 12: Representation of reactive elements with loss in the Smith impedance chart.

By this representation derives that a positive reactive value in the smith chart represent an inductive nature whereas a negative value represent a capacitive nature in the Smith impedances chart.
As the real part increase the reflection coefficients stay in a smaller circumference values, and if the real parts values exceed the characteristic impedance value the reflection coefficients that represent this circuits are inside the real constant circumference that pass for the zero point. For this properties it is possible to divide the Smith impedance chart in four areas. Inside the chart it is possible to individuate the areas represented in the following figure

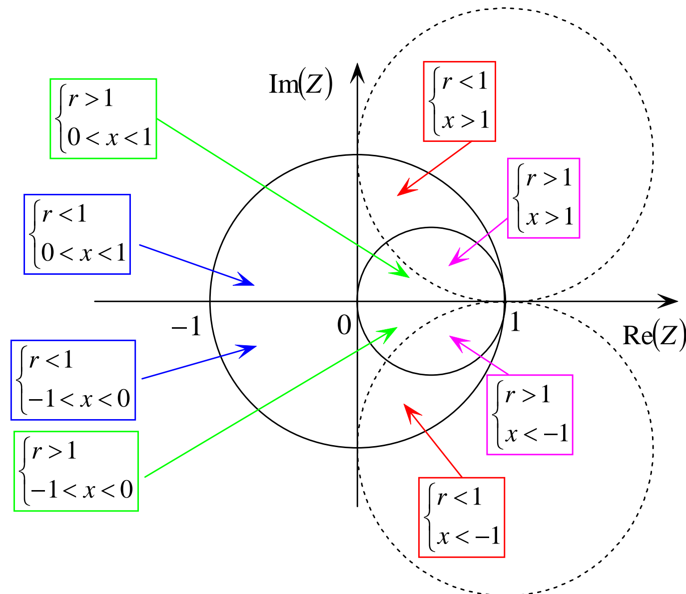


Figure 13: Different areas for the internal Smith chart of impedances.

For microwave active circuits it is important also the area outside the unitary circle that represent the devices characterized by a negative resistance value. In fact an active device does not dissipate but provide energy at a passive load that can contain or not also reactive components.

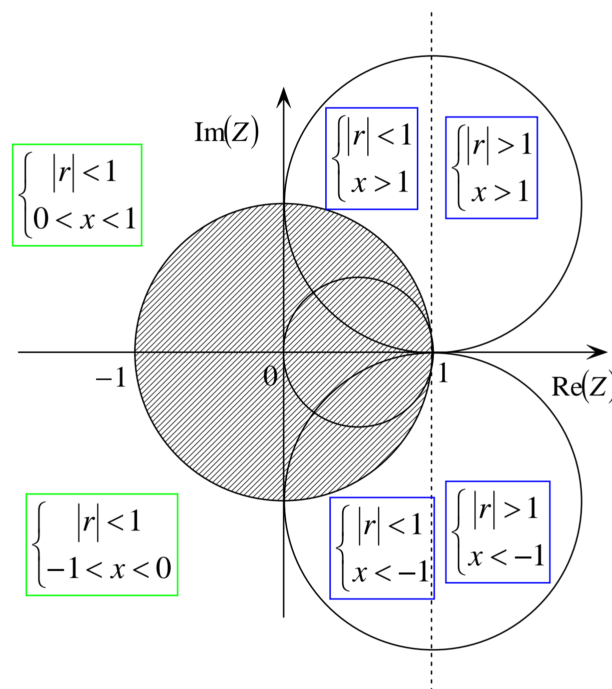


Figure 14:Different areas for external Smith impedance chart.

Actually an active device is a nonlinear elements because the impedance showed does not depends only by frequency but by the amplitude as well. A nonlinear device can not be studied as a linear device because the impedance for him is not well defined. In this case it is possible to talk about of I-V characteristic like the transistors, otherwise to consider the small-signal S-parameters showed by these devices. The Small signal S-parameters can be used in order to design low noise amplifiers or active components that work at small signal levels, as for example the intermediate stage of an amplifier in order to maximize the gain but in case of design of power amplifiers mixers and oscillators, and in general for every component that represent a great nonlinear behaviour can be

used with care. There are a lot of references that use the **LARGE-SIGNAL-S_PARAMETERS[6]** without to consider as by using these , it is impossible to represent the start-up for the oscillator. In this thesis as the small signal s parameters can be successfully used in order to design an oscillator will be showed.

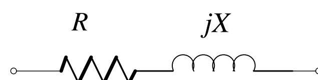
In most practical real circuits a network that contain only one reactive element does not exist but the networks contains in general both elements inductive and capacitive and often happens that a single network has an inductive behaviour for some frequencies and a capacitive for other. In these cases the designer must take care at the resonant frequencies that are represented by frequencies for which the reflection coefficient parameterized in frequency crosses the real axis of the Smith Chart. In fact the frequencies for which the energy stored in the inductive element and in capacitive is equal. The resonant circuit are very important in the design of MMIC and MIC systems and an accurate study will be present in a specific chapter. Both passive and active circuits can have resonant frequencies. When an active device has a proper resonance frequency it is an oscillator that represent a key point for moderns communication system and also the topic of this Ph.D. thesis.

The admittance chart the duality of the representation.

In order to give a unique representation in function of device properties we can consider a first definition of quality factor Q that can be defined as the ratio between imaginary and real part of impedances for a PASSIVE element (13). In this way the definition represent the ability of circuit to dissipate energy. In general a reactive element must have a Q higher as much as possible in order to provide good performances in the circuit.

$$Q = \frac{x}{r} = \frac{b}{g} \qquad (13)$$

Relation (13) can be used to define the quality factor for a passive circuit, where *x* and *b* represent respectively the reactance and susceptance and *r* and *g* the resistance and the conductance. Now we can represent the polar chart by considering the properties for the circuits represented in the different areas as function of quality factor *Q* defined as (13). For practical applications a good value of quality factor for passive devices is about 20.
The admittances Chart has got every property previously discussed, the main difference it consist in an inversion of sign for reactive parts. In fact whereas in the impedance chart a positive value of reactive impedance part represent an inductive behaviour in that of admittances it represent a capacitive nature because the “susceptance is equal to – reactance”. Obviously in this case the devices is not represented by considering a characteristic impedance value, rather the characteristic admittance $Y_0=1/Z_0$. Consequently the real properties will be mirrored compared to the impedance chart. In order to clarify these assertion we are going to consider as a inductive and capacitive bipole can be represented in the admittance chart.

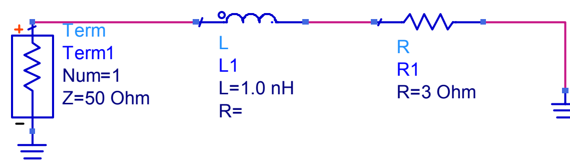


$$Z = R + jX \Rightarrow z = \frac{Z}{Z_0} = \frac{R}{Z_0} + j\frac{X}{Z_0}$$

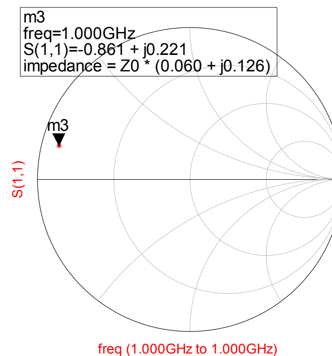
Therefore a positive value of reactance is represented in the positive half plane of Smith impedance chart. In this case if the X value is greater than R an high Q inductor are represented as showed in (Fig 13). If the same circuit is represented in the admittance chart then the series inductance is represented by a negative susceptance value.

$$Y = \frac{1}{R + jX} \rightarrow Y = \frac{R}{R^2 + X^2} - j\frac{X}{R^2 + X^2} \Rightarrow y = \frac{Y}{Y_0} = \frac{\frac{R}{R^2 + X^2}}{Y_0} - j\frac{\frac{X}{R^2 + X^2}}{Y_0}$$

In admittances chart this value will be in the upper half plane. Then we can conclude that in the admittance Smith chart a reactive element in the upper half plane is negative whereas in the impedances Smith chart would be positive.
To proof this we can build a numerical example.
Let us consider the following simple circuit at the frequency of 1GHz.



Through the common S parameter commercial simulators we can measure the value of reflection coefficient for the uniport circuit.



The CAD convert also the Γ value in impedance value and represent him by considering Z_0=50Ω.
If we consider the admittance value we found

$$y = \frac{1}{z} = \frac{1}{0.06 + j0.126} = 3.07 - j6.05$$

The reflection coefficient is represented in the upper half plane of the admittance value but at this point correspond a negative reactive value.
Commercial CAD as ADS provide only the value computed by considering the characteristic impedance value then compute only the Z value. Physically a bipole can be considered always as a impedance but we must consider the duality of representation for many applications. This point is often neglected in many textbooks and it may cause often a error source for results interpretation in analysis and design of microwaves circuits.
Putting together the impedance and admittance charts the polar chart can be obtained. In the polar chart the detailed information of the circuit are hidden and must be investigated in a way that depend by the circuit which are represented. The polar chart is depicted in function of the Γ value

and then can contain the information in both impedances or admittances terms. In practical cases from the reflection coefficient it is possible to achieve the Z values independently from the element of circuits are constituted. In a general complex circuit it is always possible to reduce the representation at a reflection value composed by a real and a imaginary part and from the *Γ* value it is possible to obtain easily the impedance or admittance representing the network. The designer must take care to the correct means of the representation in the different charts.

The resonance in the Smith chart and the instability.

The resonant frequency for an active or passive circuit must be separately studied because represent a complex and very important topic in microwave circuit design. At this point it is important to present a way in order to recognize the presence of a resonant frequency by the trend of reflection coefficient parameterized in frequency, when it is plotted in the Smith chart. The simplest way to recognize a resonance frequency is look the frequency trend of *Γ*. In general it may have a complex trend because a network characterized in terms of Z can show a inductive behaviour for some frequencies and a capacitive behaviour for others. To do this it may cross the real axis. When this happens, it means that the energy stored in the positive reactive part is equal that stored in the negative and then the *Γ* does not has a imaginary part. At this frequency the circuit has a resonant frequency. Mathematically you can see easily in fact:

$$\Gamma = \frac{z-1}{z+1} = \frac{r+jx-1}{r+jx+1} = \frac{(r-1)-jx}{(r+1)+jx} = \frac{(r-1)-jx}{(r+1)+jx}\cdot\frac{(r+1)-jx}{(r+1)-jx} = \frac{r^2-1-jxr+jx-jxr-jx-x^2}{(r+1)+x^2} \Rightarrow$$

$$\mathrm{Re}(\Gamma)+j\,\mathrm{Im}(\Gamma) = \frac{r^2-1-x^2}{(r+1)+x^2} - j\frac{2xr}{(r+1)+x^2}$$

The imaginary part of *Γ* is zero if and only if the imaginary part of impedance is zero.

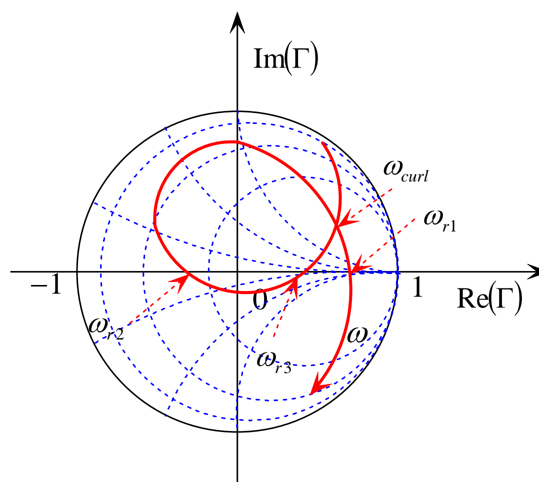


Figure 15:a general trend of reflection coefficient parameterized in frequency and represented impedance Smith chart

Actually this topic is very confused. The origin of the confusion may be generated by the shape of the reflection coefficient trend that for some frequencies may present curls, in fact some professors present the curls as possible resonance frequencies, whit special emphasis when they talk about oscillators . The proof of this should be done by the fact that changing the characteristic impedance value the point in which there is the curl (ω_{curl}) may cross the real axis. Actually change the

characteristic impedance value has not a serious means because, as discussed before, the Z_0 values are characteristics and depends only by the physical properties of the frequency. Moreover as exposed in Boyles[2-6] the presence of a curl may be meet the start up condition for a negative resistance oscillator at a specified frequency. In this paper Boyles show as a very employed criteria for check the start-up of a negative resistance oscillator is to plot the reflection coefficient trend parameterized in frequency in the polar plot. If this encircle in clockwise sense the 1+j0 point (Nyquist Criterion) then the system at start up is instable and then can oscillate. In our work I'm going to show as the start-up conditions expressed only in terms of Γ do not have a serious means but it must consider these expressed in terms of impedances as well. Anyway the instability condition that may be checked with Nyquist is necessary but not sufficient. For this purpose it is important to consider the trade of the reflection coefficient and then the curls of the reflection coefficient product used for this test. As will show in next chapter in order to verify the instability through the Nyquist criterion the reflection coefficient product is used. Obviously the possibility to encircle the 1+j0 point with growing frequency depends by the trade of reflection coefficient for the active and the passive networks and by the way to combine these trends. In particular cases like these showed in Fig 15 may happens that if for some frequencies before the curl, the product is unable to encircle the 1+j0 point, this may be encircles as the frequency increase as the torward change after the curl.

References

[1] A.Bruce Carlson-Paul B.Crilly Janet C. Routledge "Communication Systems"Mc Graw-Hill N.York 2002

[2] Robert Jackson "Criteria for the Oneset of Oscillation in microwave circuits" IEEE transaction on microwave theories and techniques VOL 40 N 03 March1992

[3] John W. Boyles*, "The* Oscillator *as a Reflection Amplifier: an Intuitive Approach to.* Oscillator *Design",* Microwave *Journal, pp. 83-98, June 1986 .*

[4] K. Kurokawa, "Some Basic Characteristic of Broad Band Negative Resistance Oscillator Circuits", Bell Sys Tech. J., ,pp 1937-1955, July-August1969.

[5] Michael Odyniec "RF and Micorwave Oscillator Design" Boston London Artech House Publishers 2002.

[6] D.A. Warren- J.M.Golio- W.L. Seely, "Large and Small Signal Oscillator Analysis", Microwave Journal, pp 228-246 May 1989

Others REFERENCES

- *R.E Collin, D.G. Dudley"Foundations for Microwave Engineering",McGraw-Hill, 1992.*
- *HP-Application Note A008, "Microwave Oscillator Design".*
- *Peter Russer, "Si and SiGe Millimeter-Wave Integrated Circuits", IEEE Trans. Microwave Theories and Tech., vol. MTT-46, n° 5, pag. 590-603, Maggio 1998.*
- *Stephen A. Mass , "Non linear Microwave circuit", IEEE Press, 1997.*
- *www.ntia.doc.gov/osmhome/allochrt.pdf*
- *http://www.datasheetcatalog.com/datasheets pdf/N/E/7/6/NE76184A.shtml*

2.1-The GaAs Technology

This chapter gives to reader an overview of why gallium-arsenide (GaAs) is often used in the design and fabrication of the MMICs and present a brief summary of the development of the monolithic microwave integrated circuit.
Perhaps the primary benefit of GaAs comes from its electron-dynamic properties. In equivalent doped n-type GaAs and silicon, the effective mass of the electric charge carriers in GaAs is far less than in silicon. This means that the electrons in GaAs are accelerated to higher velocities and therefore transverse the transistor channel in less time. This improvement in electron mobility is the fundamental property that enables higher frequencies of operation and faster switching speeds.
While the principal reason for making transistors out of GaAs is greater speed in performance, which is realized either as a higher maximum frequency of operation of higher logic switching speeds, the physical and chemical properties of GaAs make its use in transistor fabrication difficult. Most of early development in solid-state electronic devices centred on silicon and germanium based materials because on the relative ease with which the material could be processed. Silicon and germanium are elementary semiconductor materials, whereas GaAs is a binary compound. This is the root fact that cause many technical obstacles in the use of GaAs. Other properties not in GaAs's favour for early solid state device development included a lower thermal conductivity and a higher coefficient of thermal expansion that silicon and germanium. However as new market applications demanded higher performances that could be achieved only with superior electron dynamics of GaAs, these obstacles have been overcome.

The markets that drove the breakthroughs in material-growth and device –fabrication techniques of GaAs semiconductors were the defence and space industries. These industries required systems with higher frequency circuit for radars, secure communications and sensors. Many federal agencies put in place programs to develop GaAs devices as primary products in their systems. The maturity of GaAs led to the emergence of new commercial markets, such as wireless local area networks (WLANs), personal communication systems (PCSs), direct broadcast satellite (DBS) transmission and reception by the consumer, global positioning systems (GPSs), and global cellular communication. These commercial markets required the insertion of GaAs technology to meet system performances not attainable (at least since at 10 years ago) with silicon and germanium. Some of the advance achieved with GaAs technology include the use of higher frequencies to avoid spectrum crowding, new digital transmission techniques that require linear amplifiers at higher RF power levels, and lower voltage/lower current amplifier to maximize the operating and stability times of equipment that had to be powered by batteries. In some instances, GaAs is the system "enabler", without which there would be no product or service to sell. Although these emerging markets offer the advanced services and products to the consumer, several limitation to their acceptance over silicon-based systems exist. One drawback is the failure mechanism and reliability of silicon are better understood than those of GaAs. Another drawback is the cost and availability of GaAs when compared to silicon. The use of silicon in lower frequency analog circuit and in very large scale integration (VLSI) technology has developed proven practices and strong production base for semiconductor industry. This manufacturing maturity equates to a lower cost for silicon-based rather than GaAs-based technology. However, then the cost to manufacture is compared to performance, the value added to the system by the GaAs technology in most cases more than pays for the increased fabrication cost. As the WLAN, PCS, DBS, GPS, and cellular markets grow, the cost to manufacture GaAs will decrease. And the issue of using GaAs than silicon will hinge on the ability of GaAs to satisfy the technical needs to the marketplace.

Actually the silicon also improved himself continuously and gain also performances compared to GaAs. At the state of the art the battle is open and the two technologies both are courting the

market. The GaAs is more used in defence and aerospace whereas the CMOS is used in customer services as cellular phones.
In the following part of this chapter will be present the active and passive elements and present some examples of their applications with CAD.

2.2-The GaAs material properties

GaAs is a III-V compound semiconductor composed of the element gallium (Ga) from III column and the element arsenic (As) from column V of the periodic table of the elements. GaAs was first created by Goldschmidt and reported in 1929, but the first reported electron properties of III-V compounds as semiconductors did not appears until 1952. The GaAs crystal is composed of two sublattices, each face centred cubic (fcc) and offset with respect to each other by half the diagonal of the fcc cube. This crystal configuration is known as cubic sphalerite or zinc blend. The following shows a unit cube for GaAs and the table (Table 1) summarize a listing of some electrical properties.

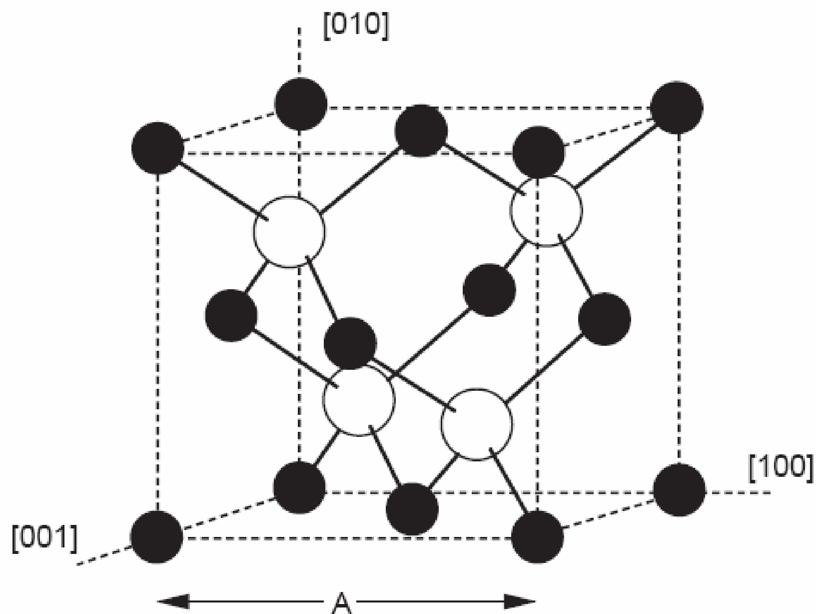


Figure 1: Unit cube of GaAs crystal lattice

As a result of the law of quantum mechanics, electrons in isolated atoms can have only certain discrete energy values. As these isolated atoms are brought together to form a crystal, the electrons become restricted not to single energy levels, but rather to ranges of allowed energies, or bands called the valence and conduction bands. These two bands are separated by an energy band cap, which is a very important characteristic of the semiconductor material. At zero Kelvin, all the electrons are confined to the valence band ant the material is a perfect insulator. Above zero Kelvin, some electrons have sufficient thermal energy to make a transition to the conduction band where they are fee to move and conduct current through the crystal. The probability of an electron having enough energy to make the transition is given by the Fermi distribution function. The Fermi level is the energy level at which the probability function is equal to one half. For pure semiconductors, the Fermi level is approximately in the centre of the gap. The amount of energy required by an electron to move from the valence band to the conduction band (energy band gap) depend on the temperature, the semiconductor material, and the material's purity and doping profile. For undoped

GaAs, the energy band gap at room temperature is 1.42eV. The energy band diagram is usually referenced to a potential called vacuum potential . GaAs is a direct band gap semiconductor, which means that the minimum of the conduction band is directly over the maximum of the valence band. Transitions between the valance band and the conduction band require only a change in energy, and no change in momentum, unlike indirect band-gap semiconductors such as silicon (Si). This property makes GaAs a very useful material for the manufacture of emitting diodes and semiconductor lasers, since a photon is emitted when an electron changes energy levels from the conduction band to the valance band.

Property	Parameter
Crystal structure	Zinc blende
Lattice constant	5.65 Å
Density	5.32 g/cm^3
Atomic density	4.5×10^{22} $atoms/cm^3$
Molecular weight	144.64
Bulk modulus	7.55×10^{11} dyn/cm^2
Sheer modulus	3.26×10^{11} dyn/cm^2
Coefficient of thermal expansion	5.8×10^{-6} K^{-1}
Specific heat	0.327 J/g-K
Lattice thermal conductivity	0.55 W/cm-°C
Dielectric constant	12.85
Band gap	1.42 eV
Threshold field	3.3 kV/cm
Peak drift velocity	2.1×10^7 cm/s
Electron mobility (undoped)	8500 cm^2/V-s
Hole mobility (undoped)	400 cm^2/V-s
Melting point	1238°C

Table 1 : Electrical and physical GaAs properties

Alternately an incident can excite an electron from the valence band to the conduction band , allowing the GaAs to be used in photo detector.

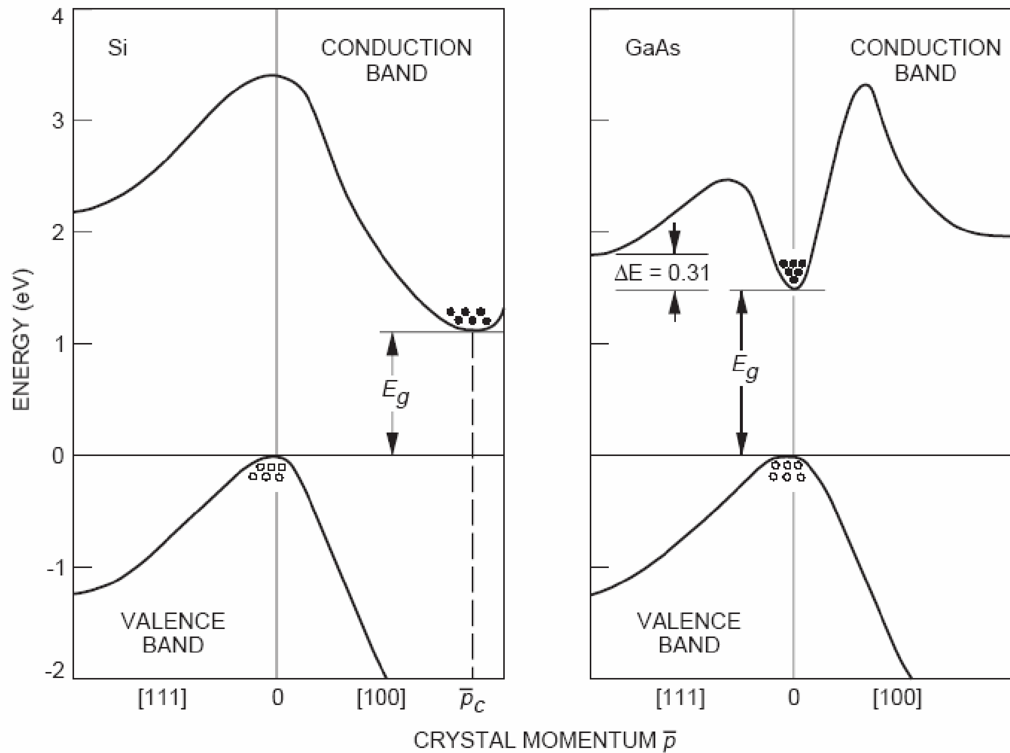


Figure 2 Energy band structure of Si and GaAs.

For more properties and studies you can check the [1]

2.3-The varactor as junction characteristic of GaAs.

In order to apply the GaAs technology at the design of VCO it is important to study the junction properties. A schematic of a metal-semiconductor junction formed on a type GaAs substrate with an external bias supply connected to the metal. Although the schematic is simple, it is also an accurate representation of the junction. To understand the junction dynamics, it is necessary to examine the energy-band diagram of the junction. It helps to first study the energy band diagram for a metal and n-type semiconductor separated from each other such that neither material is influenced by the other. Normally a finite number of electrons exist in the conduction band of the semiconductor, and the number of these free electrons is dependent on the temperature and doping concentration on purity of the material. Likewise, there are a number of free electrons in the metal, and the number of free electrons is dependent on the metal and temperature. The work function represent the energy required to remove an electron from the Fermi level of the metal to a vacuum potential. Most of the metals commonly used in GaAs circuits and devices have work function between 4 and 5.5eV.

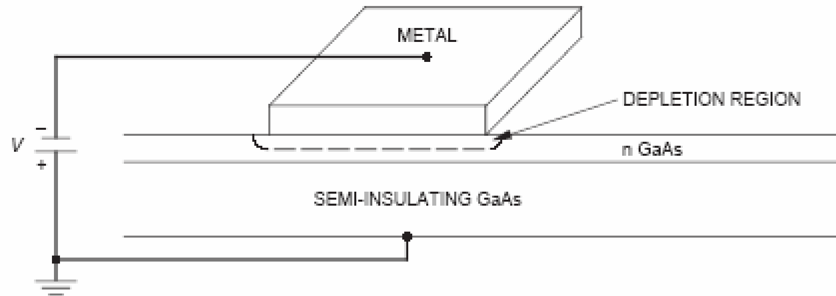


Figure 3-6. Schematic and cross section of metal–GaAs junction.

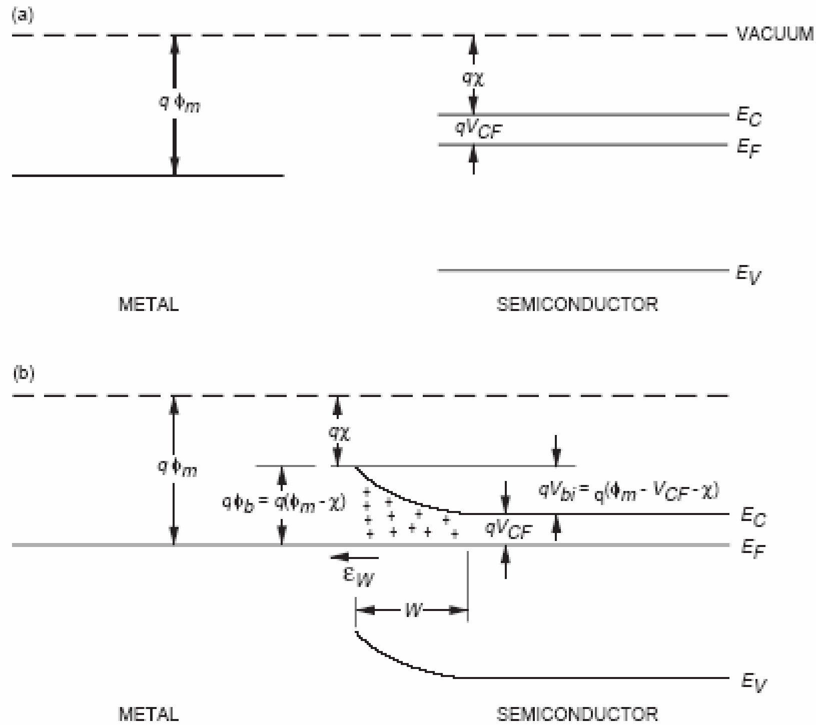


Figure 3- Energy band diagram of metal semiconductor (a) separate from each other and (b) in intimate contact.

If the semiconductor Fermi level is greater than the metal Fermi level, $\chi+V_{CF}<\phi_m$,as is shown, then the metal and semiconductor are put in intimate contact, electrons will diffuse from semiconductor to the metal. As electrons are depleted from semiconductor, a net positive charge is created in semiconductor at the junction. This positive charge will exert a force on the electrons that opposes the diffusion current. Equilibrium is established when these two forces are equal. Note that semiconductor energy bands bend in response to the forces just described. It is within this region, called the depletion region, that all of the junction's electrical properties are established. The amount of band bending is called the built-in potential, V_{bi} . For an electron to cross from the semiconductor to the metal, it must overcome V_{bi} , whereas an electron moving from the metal to the semiconductor must overcome the barrier of potential ϕ_b . To a first approximation, the barrier heioght is independent of the semiconductor, properties, whereas V_{bi} is dependent on the doping level. If an external potential is applied across the junction, the added electric field will disturb the equilibrium conditions. Consider first a positive external potential represented in the Figure below

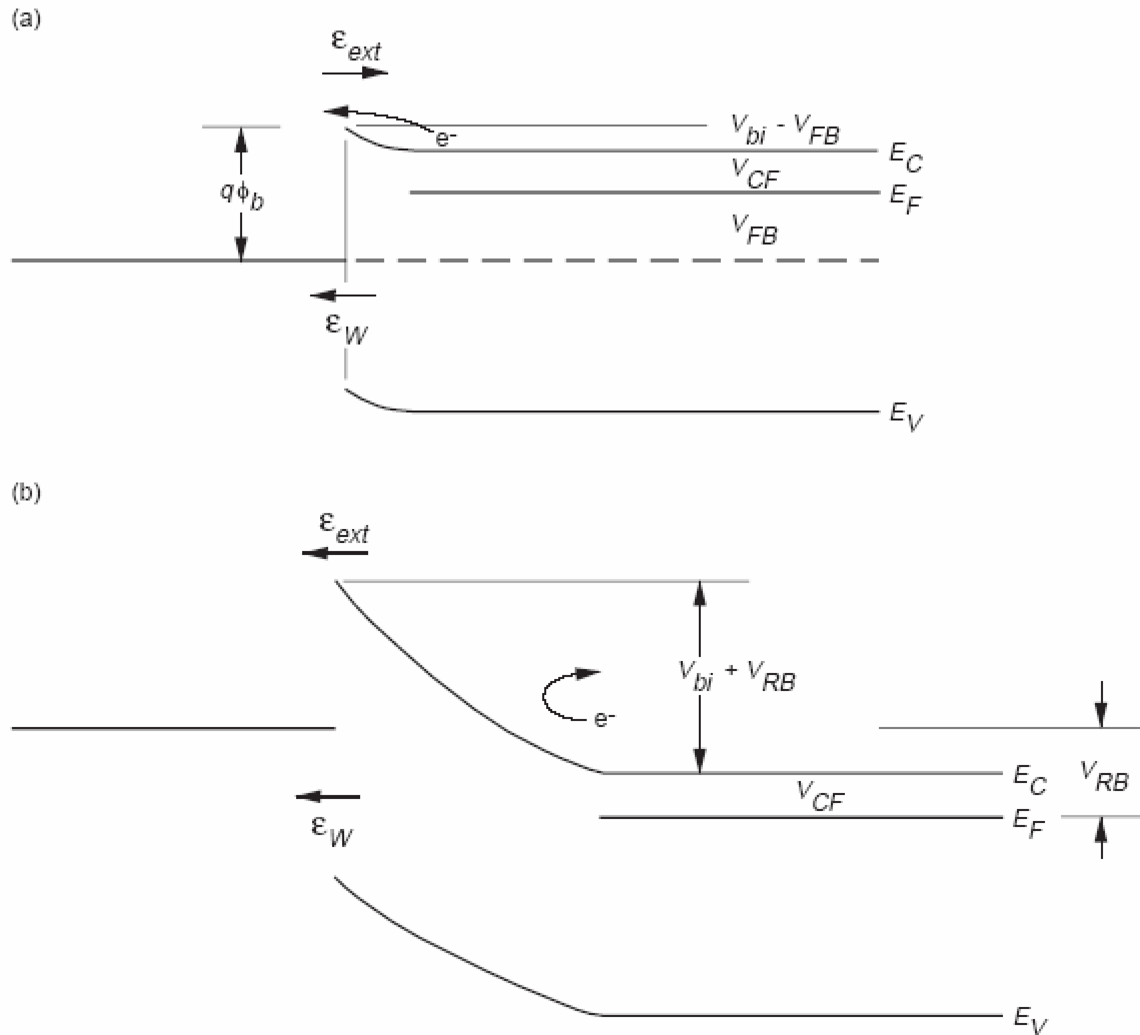


Figure 4- Energy band diagram of metal-semiconductor junction number (a) forward bias and (b) reverse bias.

This will create an electric field across the junction that is opposite to the electric field caused by the depleted GaAs atoms. The result is that the diffusion current will not be sufficiently opposed, and current will flow across the junction. Note the reduction in the barrier from the semiconductor to the metal, but not for electrons flowing from the metal to the semiconductor. If a negative voltage is applied to the metal, the external field will reinforce the electric field caused by the depleted carriers, increase the band bending at the junction, and prevent the diffusion current from flowing.
The preceding description assumed ideal material conditions. Specifically, it was assumed that the semiconductor lattice structure was uniform and perfect, even at the surface of the material. In practical cases, this is not possible. The atoms on the exposed surface do not have the require neighboring atom to complete all the covalent bonds. Therefore, these surface atoms may either give up an electron and becomes a positively charged donor ion, or accept an electron and become a negatively charged acceptor ion. Surface states and their associated charge cause the energy bands of the semiconductor to band even before the metal is introduced, as shown in Figure 5. Furthermore when the metal is brought into contact with the semiconductor, the surface states may be able to accommodate all of the charge movement required to equalize the free electrons between the two materials. When this occurs, the barrier potential is no longer dependent on the metal work function. Also, no additional band bending of the semiconductor occurs because of the metal-semiconductor contact. In other words, the junction characteristics are not dependent on the metal interface. Surface states can create severe reliability problems for GaAs devices since they are generally planar devices that use only the upper few thousand angstroms of the substrate. Therefore,

besides altering the built-in voltage of the contact, surface states may also provide leakage paths for current.

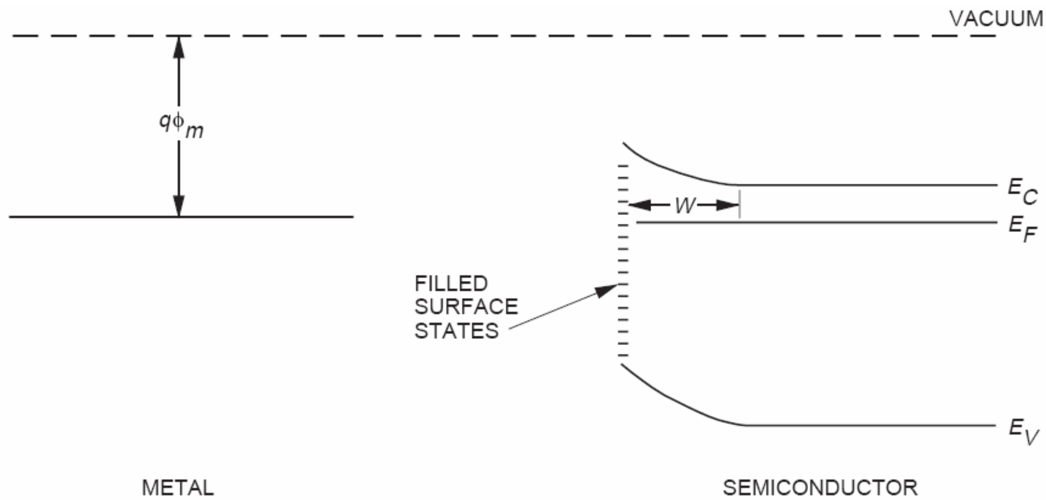


Figure 5- Energy band diagram of metal and semiconductor separate from each other when semiconductor surface states exist.

Now that the critical parameters have been introduced, their dependence on the semiconductor and metal properties can be examined. First, consider the depletion width. Upper abrupt barrier approximations, which are valid for junctions between metals and semiconductors, the width is given by:

$$W = \sqrt{\frac{2\varepsilon_r \varepsilon_0}{qN_d}\left(V_{bi} - V - \frac{kT}{q}\right)} \qquad (1)$$

Where N_d is the donor doping concentration, k is Boltzmann's constant, and q is the charge of an electron. The term *kT/q,* often referred to as V_T, is approximately 0.026V at room temperature whereas V_{bi} is approximately qV. From equation (1) it is seen that the depletion width is smaller for highly doped semiconductors, and that the depletion width varies inversely with the applied bias. Based on the preceding discussion relating to Figure it is noted that a positive bias increases current flow and decreases the depletion width. The opposite occurs for a negative bias.

Depletion widths can be quite large. As an example, consider two GaAs substrates at room temperature with an aluminium contact. Let the first have a typical MESFET channel of $N_d=10^{17}/cm^2$ and second have a typical ohmic contact of $N_d=10^{19}/cm^2$. With no external bias supplied the depletion widths for these two samples are approximately 0.048µm and 0.006µm respectively. Although there appear to be very small quantities, it will become apparent throughout the rest of this chapter that these depletion widths are in fact large compared to the dimensions required for microwave circuits. Another critical parameter is the electric field across the depletion region. The concern is that the maximum electric field that occurs at the metal-semiconductor interface must be kept smaller than the breakdown field of GaAs, approximately $4X10^5$ V/cm. If $E_m>4X10^5$V/cm, electrons have enough kinetic energy to create electron/hole pairs during electron/atom collisions at a faster rate than the free charges can recombine. These new electrons also are accelerated by the electric field and create more electron/hole pairs. This runway process is called avalanche breakdown. The result of avalanche breakdown is often a catastrophic junction failure. The maximum electric field is given by:

$$E_m = \frac{qN_d}{\varepsilon_r \varepsilon_0} W = \sqrt{\frac{2qN_d}{\varepsilon_r \varepsilon_0}\left(V_{bi} - V - \frac{kT}{q}\right)} \quad (2)$$

The field is stronger when large doping concentrations are used or if a large reverse bias is applied across the junction.
The charge storage in the depletion region also creates a capacitance across the junction, which is given by

$$C = \frac{\varepsilon_r \varepsilon_0}{W} = \sqrt{\frac{q\varepsilon_r \varepsilon_0 N_d}{2\left(V_{bi} - V - \frac{kT}{q}\right)}} \quad (3)$$

Note two things about equation 3. First the capacitance is a function of the applied voltage. Therefore, the junction behaves as a voltage-controlled capacitance. It is this feature of the junction that is exploited in varactor diodes, which are commonly used in phase shifters and voltage controlled oscillators (VCOs)[5]. The second thing to note is that the capacitance is dependent on the doping concentration. Therefore, by varying the doping profile across the junction, the capacitance-voltage curve can be varied. Alternatively, if the doping concentration is altered during the life of the diode the capacitance will change, and frequency shift in the VCO or phase change from phase shifter will occur.
Although an understanding of the depletion width and its associated capacitance are critical for the gate design of a field effect transistor, it is the current flow through the junction that the circuit designer is ultimately concerned with. In general, current flow through the junction is due to several mechanisms. It is necessary to examine only two of these for the purpose of this text. The first is the transport of electrons over the potential barrier, usually called thermionic emission. Thermionic emission current assumes that only electrons with energies greater than the energy of the potential barrier add to the current flow. Several methods of analysis have been proposed to determine the current density, and although each uses different assumptions and boundary conditions, they all result in an equation of the form:

$$J = J_0^* e^{-q\phi_b/kT} \cdot \left[e^{qV/kT} - 1\right] \quad (4)$$

J_0 increases with the doping concentration, N_d, and temperature. Note that J is exponentially dependent on the barrier potential, temperature, and the applied voltage. It is thing strong dependence on the applied voltage that makes the junction a good rectifier. Furthermore, the dependence on temperature makes this current mechanism dominant at higher temperatures. When Schottky diodes are characterized, the measured current does not fit Equation 4 exactly but rather

$$J = J_0^* e^{-q\phi_b/kT} \cdot \left[e^{nqV/kT} - 1\right] \quad (5)$$

Where n is called ideality factor. An ideal diode would have n=1 but for actual diodes, n>1. A change in ideality factor over the life of the diode is an indication that the metal-semiconductor interface is changing.
The second current mechanism that needs to be described is due to quantum mechanical tunnelling through the potential barrier. Recall from quantum mechanics that the position of a particle is not absolute, but described by a distribution function. Therefore, although the majority of electrons will be confined by a potential barrier, there is a probability that some of the electrons will exist in the region of the potential barrier. Furthermore, if the potential barrier is thin enough, there is a

probability that some electrons will travel through the barrier. This current component is referred to as tunnelling current.

The practical implementation of a planar diode is showed in the figure below

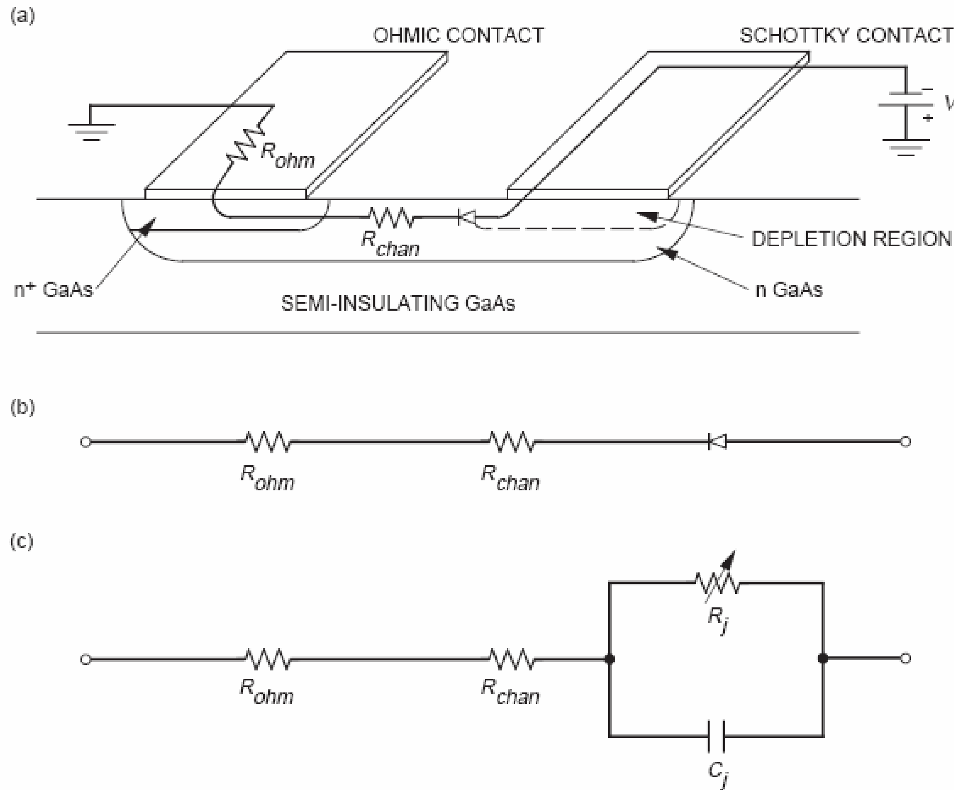


Figure 6- GaAs planar diode: (a) schematic, (b) simple equivalent circuit, (c) equivalent circuit for a planar Schottky diode.

The diode is fabricated either on a molecular-beam-epitaxy- (MBE) grown n layer or by ion implantation of an n region in the semi-insulating GaAs substrate. This is followed by the deposition of an ohmic contact metal, normally AuGe, and an ohmic contact alloying bake. Lastly, the Schottky contact metal is deposited. A simple equivalent circuit for the diode is showed in Figure 6 (b), R_{ohm} refers to the ohmic contact junction resistance and R_{chan} refers to the resistance between the two metal contacts. Although both resistances are parasitic and ideally would be eliminated, practical limitations do not permit this [6]. The diode electrical specifications will normally determine the doping concentration of the n region. Therefore, the ohmic contact resistance cannot be altered unless an n^+ region is formed upon which the ohmic contact can be made. R_{chan} can be reduced if the distance between the two contacts is reduced. Modern lithography permits the contacts to be separated by as little as 0.2μm, although the contact separation is typically on order of 1 μm. Unfortunately, the electric field between the two contacts increases as the spacing is reduced. If the electric field is increased too much because of the RF power or the dc bias, metal shorts may develop between the contacts leading to device failure. Therefore, limitations on the power handling capability of the diode are normally imposed.
Lastly, consider the diode itself. It has already been shown that the depletion region creates a capacitance called tha junction capacitance, C_j. In addition, there is also a junction resistance, R_j, which is in parallel to C_j. R_j accounts for the current flow through the depletion region and can be derived from equation (4)

$$R_j = \frac{nkT}{qJA} \qquad (6)$$

where A is the diode area, therefore the equivalent circuit for a planar GaAs diode is showed in Figure 6 (c). Notice that R_j is shown as a variable resistance due to its dependence on J, which in turn is dependent on the applied voltage. The important figure of merit for Schottky diodes is the forward current cut-off frequency:

$$f_c = \frac{1}{2\pi R_F C_F} \qquad (7)$$

where R_F is the total series resistance and C_F is the junction capacitance a slight forward bias. Schottky diodes have been fabricated with cut-off frequencies greater than 1THz. In general, a diode can be used at frequencies less than $f_c/10$. Therefore, it is desirable to have a small R_j and C_j as well as a small R_{ohm} and R_{chan}. To minimize the R_j, the diode area must be increased, but to minimize C_j, the diode area must be decreased. Furthermore, a parasitic capacitance due to the fringing fields along the edges of the Schottky contact exists. This parasitic capacitance is proportional to the diode periphery and the number of corners on the contact. Since the ratio of the contact periphery to area increases as the area of the contacts decreases, it is not practical to reduce C_j solely by decreasing area. When the diode is used as varicap (voltage controlled capacitance) then it is necessary to build a bias network in order to reverse biasing the diode and to increase or decrease the capacitance in present in the depletion region. The diode will be reverse biased if the doped negative part o the junction is connected to a positive voltage source and the positive doped part is connected to ground. To realize the bias network in RF and microwave frequencies field it is very important to separate the RF paths to the direct sources. Then components as DC-FEED (usually inductor of very large size) and DC-BLOCK (implemented through very large capacitances) are employed. The "very large" term must be intended as compared to other sizes present in the circuit. For example @24GHz values of inductances and capacitances usually employed in order to implement respectively the DC-FEED (that realize a path for the bias voltage) and DC-Block (a path for the signals) are respectively about 1nH and 1pF. A simple network to reverse biasing a diode is showed in the figure below.

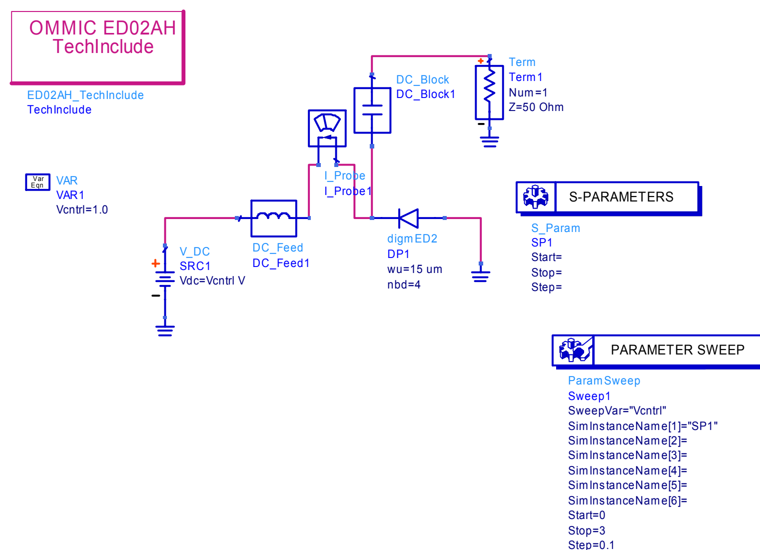


Figure 7- Reverse bias network of a GaAs diode

In the figure is possible to note the presence of a DC-FEED to connect the negative port of the diode to the VDC, a current probe present only for measure the current a DC-BLOCK used to connect the port an external circuit whereas the positive part of the junction is connected to ground. Through S-parameters simulators it is possible to study the behaviour of the diode for a fixed frequency by varying the DC voltage. Representing on the Smith impedance chart the value of the reflection coefficient extracted by the circuit it is possible to know the capacitance that the diode have for a definite value of bias voltage. This capacitance value can be changed for example by changing the dimensions of the diode as length or number of finger. To show how the results of the S-parameters analysis are represented for the circuit of Figure 7 in the Smith chart you can check the Figure 8.

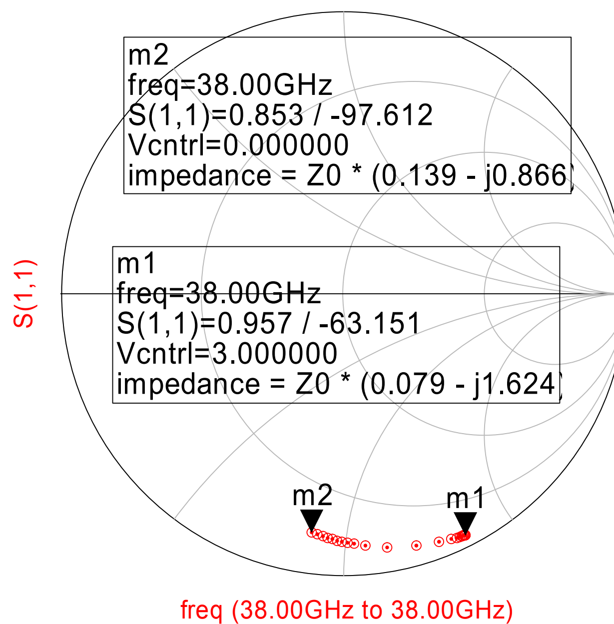


Figure 8- S-parameters analysis of circuit showed in Figure 7

It is easy to compute the capacitance value showed by the diode as varying the DC-voltage. As example we can compute the capacitance equivalent showed for *Vcontrl=3V* (position of marker m1)

$$Z = 50(0.139 - j \cdot 0.866) = 6.95 - 43.3i$$

to compute the capacitance only the reactive part must be considered; then we have:

$$-jX = \frac{1}{j\omega C} = -j\frac{1}{\omega C} \Rightarrow X = \frac{1}{\omega C} \Rightarrow C = \frac{1}{2\pi 38^9 \cdot 43.3} \approx 96.7\, fF$$

This is a normal value for GaAs diodes for high frequency circuits. Obviously the value of capacitance depends also by frequency range and dimensions. As the Vcntrl decrease the

capacitance obviously decrease because the electric field that exist across the junction is more weak decrease and its value for the ***0V*** is about ***51fF***.
In many VCOs applications it is important to measure the capacitance parameterizing the test for large signal levels. For example if the varicap is employed in a Voltage Controlled Oscillator In fact, an high level output power, can be responsible of a bias diode inversion. For the VCO this situation must be forbidden because would represent a non linear tuning range for the VCO. Using a varicap in a VCO allows only to achieve small frequency tuning ranges. To clarify this assertion we can look the Figure A Large signal parametric simulation show as the diode enter in the linear condition if the signal power is greater of ***8dBm*** at *Vcntrl=0V*.

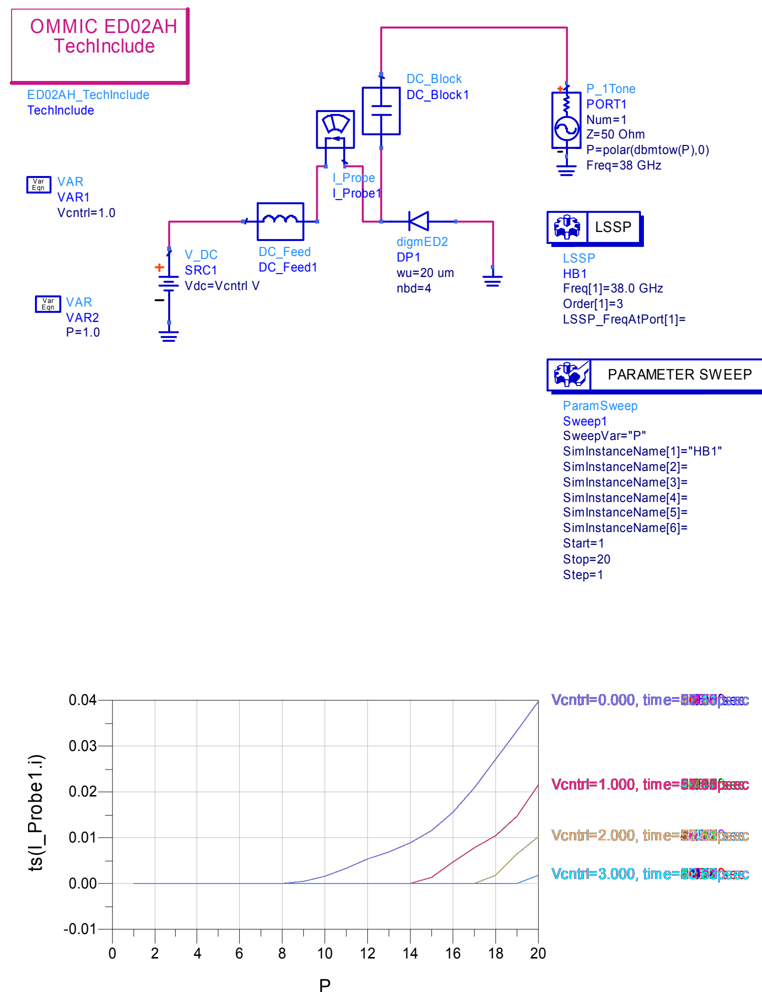


Figure 9- Large Scattering Simulation Parameters for the varicap

To overcome this problem the diode should be sized to avoid the direct biasing for any tuning voltage, when the power of signal is greater than 10dBm.

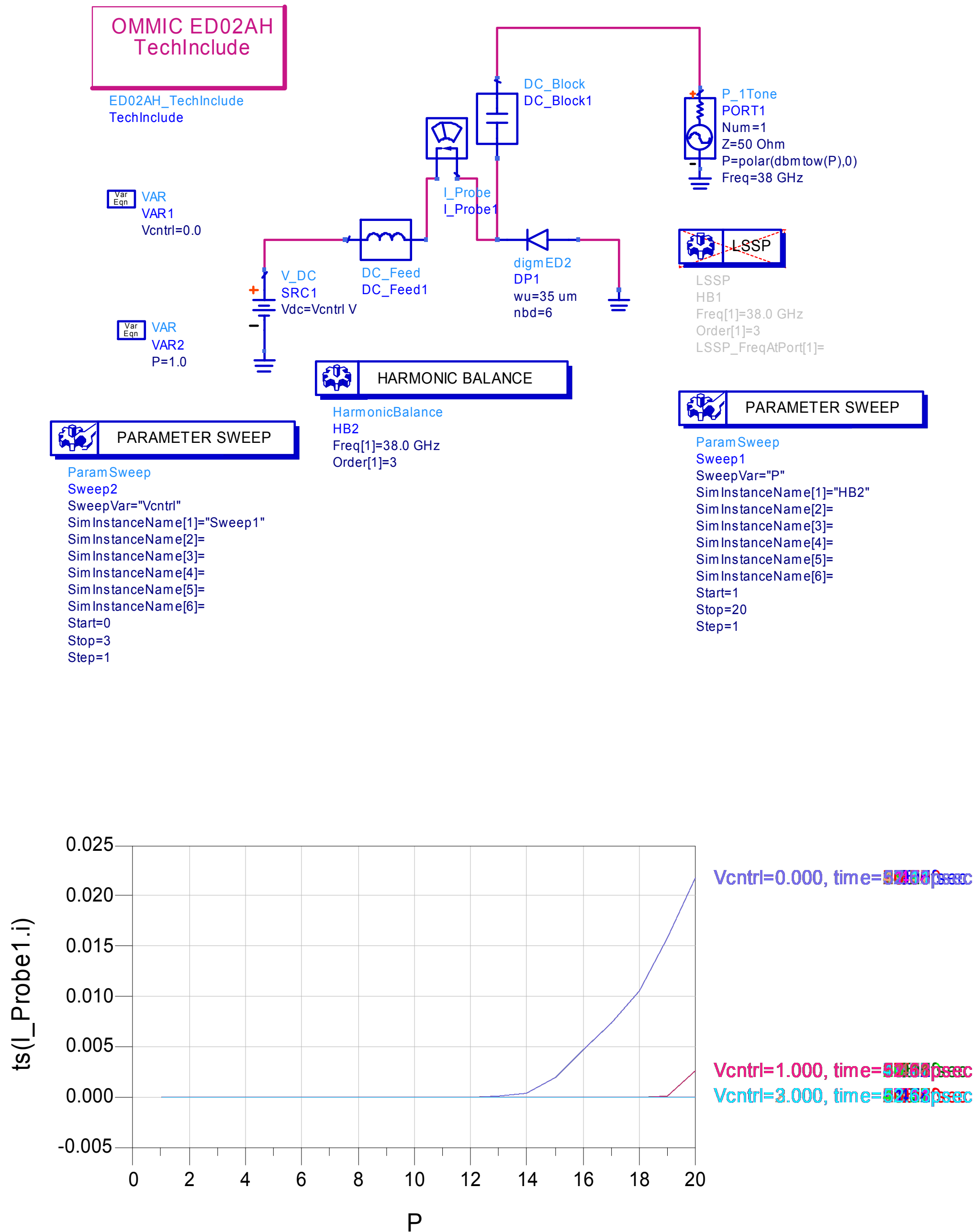


Figure 10- LSSP simulation with HB

The Figure 10 show how the diode with den sizes does not enter in conduction for the power levels of interests, and how the Harmonic Balance simulator can be employed at the same manner of the Large Signal S Parameters simulator in commercial CADs as ADS in order to perform the diode behaviour. The diode used to study and present in practical manner the behaviour of a generic GaAs diode is the ***digmED0 0.2 µm*** of the ***OMMIC ED02AH*** library provided at our university to research purposes. For more information about foundry check the link
http://www.ommic.com/

RELIABILITY

Reducing the diode area has been discussed as a method to reducing C_j. Besides the disadvantages of having a small diode area already discussed, there are other disadvantages. First, the fringing fields around the periphery of the diode will be greater and will lead to increased leackage current. Second, the fringing fields can be larger those predicted, especially around the corners of the contact. Therefore, the reverse breakdown voltage will be smaller. Moreover the current density must be increased as the area is decreased to maintain reasonable current through the device. If the current density is increased too much, failures due to electromigration may occur. Finally, the increased current density and the reduced junction area may cause the junction temperature to increase. Since GaAs is a relatively poor thermal conductor, thermal-related failure mechanism may increase as well. To get around these problems, it is better to maximize the area and to minimize C_j by decreasing N_d. Note that reducing N_d requires n^+ regions for the ohmic contacts and increase the risk of ionic contamination failures. The critical points to remember about the junction as it relates to device reliability are:

- The sensitivity of its electrical characteristics to the semiconductor doping concentration
- The interface barrier potential.
- The junction temperature.

Small changes in any of these parameters can greatly change the junction impedance and therefore the current that flows through the junction. While the circuit designer can control the junction temperature through proper packaging and heat sinking, unfortunately the barrier potential and doping concentration may change unpredictably over the life of the junction-especially at higher operating temperatures or if metal- semiconductor interactions occur.

2.4-Metal Semiconductor –GaAs-Field-Effect Transistors

GaAs metal-semiconductor field-effect transistor (MESFETs) are the most commonly used and important active devices in microwave circuits. In fact, until the late 1980s, almost all microwave integrated circuits used GaAs MESFETs. Although more complicated devices with better performances dor some applications have been introduced, the MESFET is still the dominant active device for power amplifies and switching circuits in the microwave spectrum.

The basic MESFET is shown schematically in Figure 3. The base material which the transistor is fabricated is a semi-insulating GaAs substrate. A buffer layer is epitaxially grown over the semi-insulating substrate to isolate defects in the substrate for the transistor. The channel or the conducting layer is a thin, lightly doped (n) conducting layer of semiconducting material epitaxially grown over the buffer layer. Since the electron mobility is approximately 20 times greater than the hole mobility for GaAs the conducting channel is always n type for microwave transistors. Finally, a highly doped (n^+) layer is grown on the surface to aid in the fabrication of low-resistance ohmic contacts to the transistor- This layer etched away in the channel region. Alternatively, ion implantation may be used to create the n channel and the highly doped ohmic contact regions directly in the semi-insulating substrate. Two ohmic contacts, the source and drain, are fabricated on the highly doped layer to provide access to the external circuit. Between the two ohmic contacts, a rectifying on Schottky contact is fabricated. Tipically, the ohmic contacts are Au-Ge based and the Schottky contact is Ti-Pt-Au.

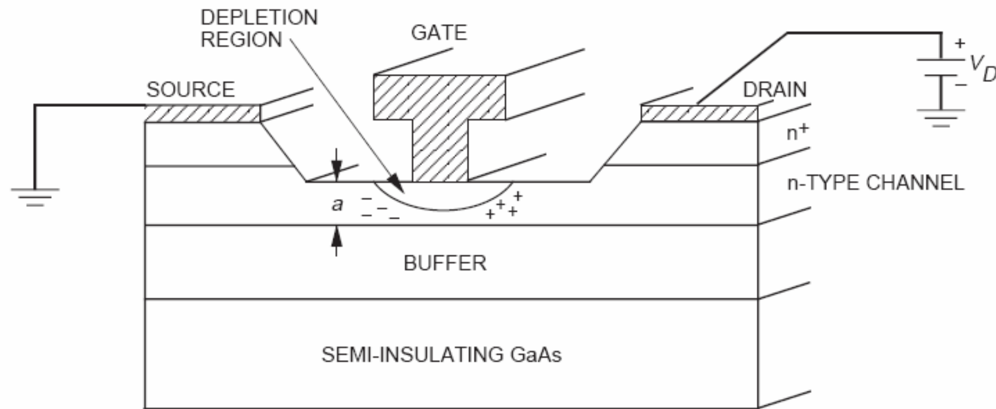


Figure 11 Schematic and cross section of a MESFET

The basic operation of the MESFET is basically understood by first considering the I-V characteristics of the device without the gate contact, as shown in Figure. If a small voltage is applied between the source and drain, a current will flow between the two contacts. As showed in the Figure below the voltage is increased, the current increases linearly with an associated resistance that is the sum of the two ohmic resistances, R_S and R_D, and the channel resistance R_{DS}.

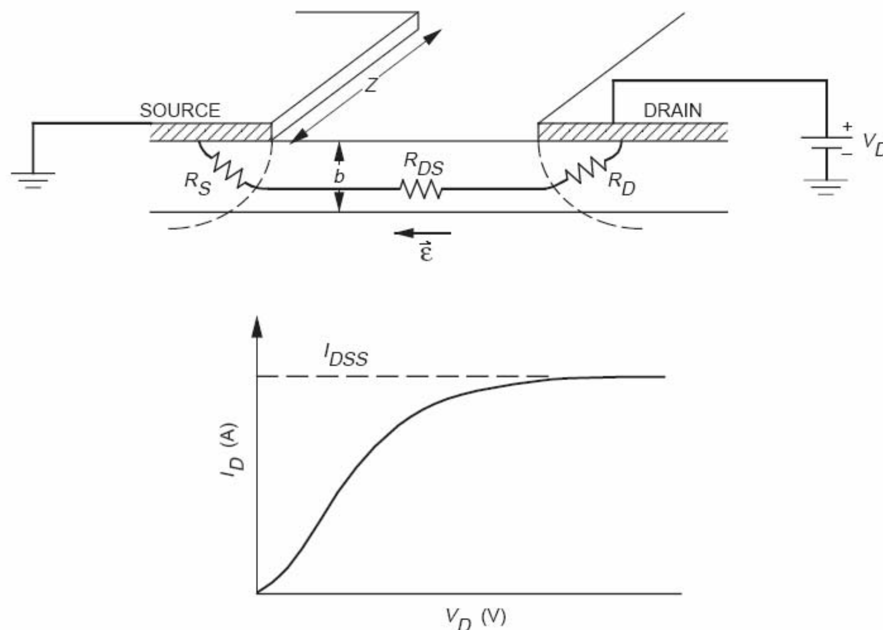


Figure 12- Schematic and I-V characteristic for an ungated MESFET

The current can be expressed as the relationship (8):

$$I_D = \frac{V_D}{R_D + R_S + R_{DS}} \qquad (8)$$

If the voltage is increased further, the applied electric field will become greater than the electric field required for saturation of electron velocity, as show in Figure 4. Under large bias conditions, an alternative expression for ID is usefull; this expression relates the current directly to the channel parameters:

$$I_D = Q(x)v(x) = Zb(x)qn(x)v(x) \qquad (9)$$

The expression omits the parasitic resistances R_D and R_S . The parameters in equation (2) are *Z*, the width of the channel; *b(x)*, the effective channel depth; q, the electron charge; *n(x)*, the electron density; and *v(x)*, the electron velocity, which is related to the electric field across the channel. Note that if *v(x)* saturates, I_D will saturate. This saturation current is called I_{DSS}.
Now consider the effect of the fate electrode placed over the channel but without any gate bias, $V_G=0$. A depletion region formed under the gate electrode reduces the effective channel depth, *b(x)*and therefore increases the resistance to current flow under the gate. The depletion region is dependent on the voltage drop across Schottky junction. Since the current flowing through the channel is equivalent to a current flow through a distributed resistor, there is a large voltage drop across the drain end of the channel than at the source end. This result in the depletion region depth being greater on the drain side of the channel. The nonuniform channel depth has two effects on the device operation. First, there is an accumulation of electrons on the source side and a depletion of electrons on the drain side of the depletion region. This dipole of charge creates a feedback capacitance between the drain and the channel; this capacitance is typically called C_{DC}. The second effect is that the electric field due to the dipole adds to the applied electric field causing the saturation conditions to occur at a lower V_D. By applying a bias to the gate junction, the depletion depth and therefore the resistance of the current flow between the source and drain and the saturation current can be controlled. If a large enough negative gate bias is applied, the depletion region depth will equal the channel depth, or the channel will be pinched off. This gate bias is called the pinch-off voltage and is given by:

$$V_P = \left(\frac{qN_d}{2\varepsilon_r \varepsilon_0} \right) a^2 \qquad (10)$$

Under pinch off conditions, the drain current drops to a very small value. Therefore, the transistor can act as a voltage-controlled resistor or a switch. The most important feature of MESFET is that they may be used to increase the power level of a microwave signal, or they provide gain. Because the drain current can be made to vary greatly by introducing small variations in the gate potential, the MESFET can be modelled as a voltage-controlled current source. The transconductance of the MESFET is defined as

$$g_m = - \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{DS}=const.} \qquad (11)$$

Using short-channel approximations, it can be shown that the transconductance may be written as

$$g_m = \frac{I_S}{2V_P} \left(\frac{I_S}{I_S - I_D} \right) \qquad (12)$$

where I_S is the maximum current that can flow if the channel were fully undepleted under saturate velocity conditions. This is the same as the saturation current discussed for the device without the gate electrode shown previously. Since I_S is proportional to the channel depth, a, and V_P is proportional to the square of the channel depth, g_m is inversely proportional to the channel depth. In addition, from equation (8) we note that for large I_S and g_m the parasitic resistance R_S and R_D must

be minimized. The most commonly used figures of merit for microwave transistor are the gain bandwidth product, the maximum frequency of oscillation, f_{max}, and the frequency where the unilateral power gain of the device is equal to one, f_t. Consider first the parameter f_t. If short gate length approximation are made, f_t, can be related to the transit time of the electrons through the channel ,*t*, by the expression

$$f_t = \frac{1}{2\pi\tau} = \frac{v_{sat}}{2\pi L} \qquad (13)$$

Since v_{sat} is approximately $6X10^{10}$µm/s for GaAs with doping levels typically used in the channel, the gate length must be less than 1 µm for ft to be greater than 10 GHz. The parameter f_{max} may be approximated by

$$f_{\max} = \frac{f_t}{2}\sqrt{\frac{R_{DS}}{R_G}} \qquad (14)$$

where R_G is the gate resistance. From the above two expressions for f_t and f_{max} , it is apparent that the gate length should be made as small as possible. Both the limits of fabrication and the need to keep the electric field under the channel less than the critical field strength required for avalanche breakdowns set the lower limit on *L* at approximately 0.1µm. For the gate to have effective control of the channel current, the gate length *L* must belarger than the channel depth, *a*, or $L/a>1$. This requires that the carrier concentration in the channel be as high as possible to maintain high current.

RELIABILITY

The small feature sizes described above may create reliability problems in microwave GaAS MESFETs. The small cross section of the gate electrode results in a high current density, especially for power transistors, which leads to electromigration failures. To reduce the gate resistance, gold is typically used over the gate refractory metal. Dince gold creates deep-level traps in GaAs, which effectively reduce the carrier concentration and therefore the current of the device, barrier metals such as platinum must be used. IN addition, because the channel depth is so small, any diffusion of gate metals into the GaAs creates large changes in the current that flows through the channel and decreases the pingh-off voltage. The small distance between the gate and drain electrodes also creates high electric fields, which may create an avalanche generation of electrons. These “hot electrons” may then become trapped in the surface states of the GaAs or in the passivation material that is commonly deposited over the device. The reliability problems that occur greatly depend on the device technology, as well as its application. In small-signal applications, the degradation ot the ohmic contacts or interdiffusion of the gate metals with the GaAs in the channel leads to shifts in I_D,g_m and V_P. Althrough power MESFETs also suffer from parametric degradation, catastrophic failures are more common. However, advances in device technology and operation within safe limits have decreased the incidence of burnout. For power amplifiers, the MESFET must be designed for maximum power output. This is equivalent to requiring a large drain-to-source voltage and a large drain current. To maximize I_D , a large carrier concentration or a large gate width is required; note that the channel depth may bot be increased since that would degrade the frequency range of the device. The carrier concentration may not be increased without defrading the gate-to-drain breakdown voltages, which must be maximized to maximize V_{DS}. Therefore, the only alternative is to increase the gate width, Z. Unfortunately, in microwave circuit design, long line lengths do not appear as lumped elements with an uniform potential along the length, but rather as distributed transmission lines with potential nulls occurring every half wavelength. The general rule

of thumb is that a line should be less than one tenth of a wavelength long to be considered a lumped element. For a GaAs,this is equivalent to

$$Z < \frac{11.3}{f} mm \qquad (15)$$

Where f is the frequency in GHz. Therefore, at X-band, 8 to 12 GHz, the maximum gate width that may be used is approximately 1 mm. If greater current is required, multiple gate fingers may be used in a parallel connection. This parallel connection of gate gingers in a tightly packed region increases the localized temperature of the circuit. Dince GaAs is poor thermal conductor, power transistors will typically operate at least 10 deg above the carrier temperature. This increased device temperature, with the higher fields and current used in power MESFETs, often leads to catastrophic failures.

HEMTs and PHEMTs

GaAs-bsed high electron mobility transistors (HEMTs) and pseudomorphic HEMT or PHEMTs are rapidally replacing conventional MESFET technology in military and commericl application requiring low noise figures and high gain, particularly at millimetre-wave frequencies. The application of PHEMTs for high-efficiency power amplification is gaining popularity. Other commonly used names for HEMTs include MODFET (modulation doped FET), TEGFET (two-dimensional electron gas ET) and SDHT(selectively doped heterojunction transistor). Since HEMTs and PHEMTs are field-effect transistors, the basic principle of operation are very similar those od the MESFETs described previously. The main difference between HEMTs and MESFETs is the epitaxial layer structure. In the HEMT structure, compositionally different layers are grown in order to optimize and to extend the performance of the FET. For III-V semiconductor using a GaAs substrate, the common materials used are $Al_XGa1_{-l-X}As$ and GaAs. There are some difference in layers from heterojunctions since each layer has different bad-gap. Structures grown with the same lattice constant but different band gaps are simply referred to as lattice-matched HEMTs. Those structures grown with slightly different lattice constants are called pseudomorphic HEMTs or PHEMTs. The structure of a basic AlGaAs/GaAs HEMT is showed below.

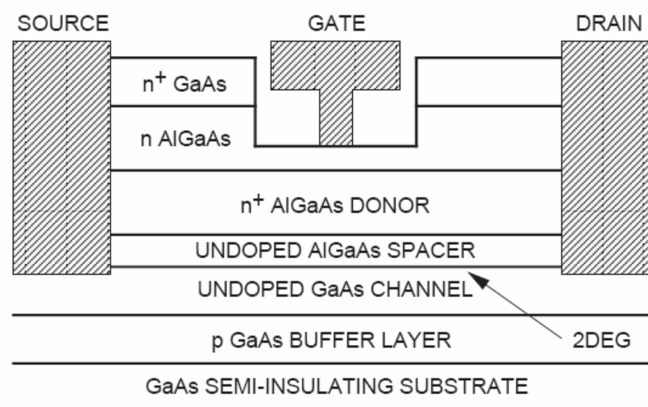


Figure 13- Basic AlGaAs/GaAs HEMT.

The structured showed before is a basic HEMT structure. Most of the structures used today are variants of this, having been optimized for performance and applications. For instance, many PHEMTs used for power applications will incorporate two silicon pulses, the second one below th channel, to increase the total charge available. Fore more information refears to [2]. Under operation, HEMTs and MESFETs are biased similarly. When a negative gate bias is applied to the

HEMT device, the Schottky layer becomes depleted. As the gate is biased further, the 2DEG becomes depleted. This results in the modulation od the channel (2DEG) by a negatively applied gate bias where gain and amplification occur until the channel is pinched off. The transconductance is given by

$$g_m = \left(\varepsilon v_{sat} W_g\right)/d \tag{16}$$

where ε is permittivity of substrate v_{sat} the saturated velocity, W_g the unit gate width and d the distance from the gate to 2DEG. Under high-electric-field conditions the HEMT shows an higher saturated velocity over the MESFET. Since the conduction of electrons from the source to the drain takes place in a channel that is well confined, g_m will remain very high at low drain current, the distance d will increase because the edge of the depletion region enters the tail of the doping profile. This results in a compression of the g_m. The higher mobility of the HET results in lower parasitic drain and source resistances. As a result $f_t=g_m/(2\pi C_{gs})$ and f_{max} are increased from the MESFET case for a given gate length leading to a lower noise figure and higher gain.

RELIABILITY

The reliability of HEMTs and PHEMTs is affected by the epitaxial structure, device fabrication, and device geometry. To take advantage of the high g_m , HEMTs and PHEMTs, rely on small geometries for optimum performance. Parallelinfg the MESFET, HEMTs, PHEMTs suffer the same electromigration and metal interdiffusion reliability problems associated with the ohmic and gate metallizations under the device operation. In addition, hot electron traps resulting from the generation of avalanche electrons are a problem for the HEMts and PHEMTs. The hot electrons cause a degradation in the current and in the gain and power under microwave drive as they become trapped in the passivation or the substrate. For power devices, catastrophic failures or burnout can also be an issue due to the high channel temperatures resulting from the large currents required for high power.

HETEROJUNCTION BIPOLAR TRANSISTORS

AlGaAs/GaAs heterojunction bipolar transistors (HBTs) are used for digital and analog microwave applications with frequencies as high as Ku band. HBTs can provide faster switching speeds than silicon bipolar transistors mainly because of the reduced base resistance and collector to substrate capacitance. HBT processing requires less demanding lithography than GaAsFETs, therefore, HBTs can cost less to fabricate and can provide improved lithographic yeld. This technology can also provide higher breackdown voltages and easier broad-band impedance matching than GaAs FETs.
In comparision with Si bipolar junction transistors (BJTs), HBTs show better performance in terms of emitter injection efficiency, base resistance, base-emitter capacitance, and cutoff frequency. They also offer good linearity, low phase noise and high power-added efficiency.

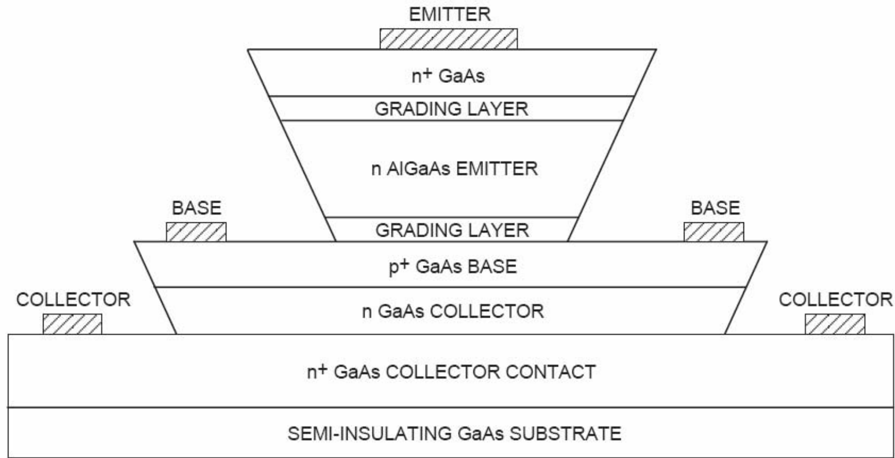


Figure 14- Cross section of an example HBT.

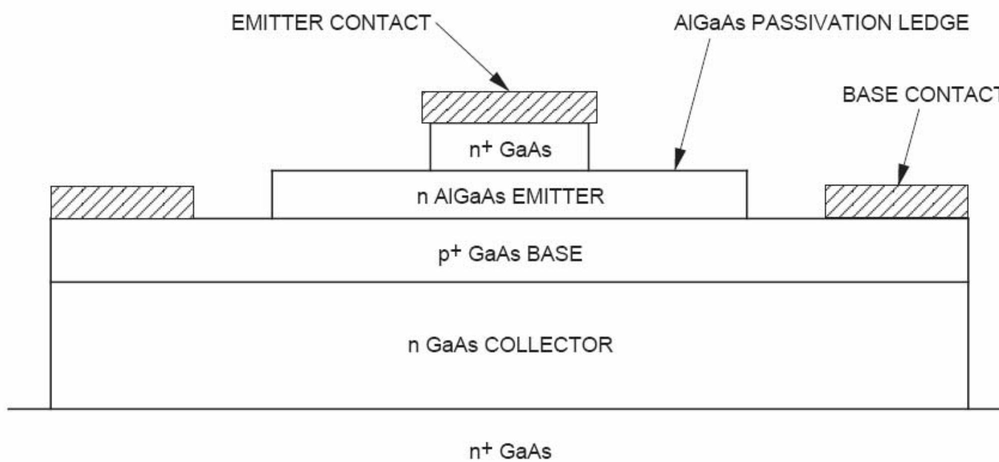


Figure 15 An HBT cross section showing a thin ledge of AlGaAs/GaAs.

The operating principle is the same like a BJT. The potential barriers for hole injection (ΔV_p) and electron injection (ΔV_n) in a graded E-B junction differ by the band-gap difference (ΔE_g) between the AlGaAs emitter and the GaAs base. Therefore, we have

$$q = \left(\Delta V_p - \Delta V_n\right) = \Delta E_g$$
$$\left(\Delta E_g\right) = E_g\left(AlGaAs\right) - E_g\left(GaAs\right) \qquad (17)$$

Now is possible to write the expression for the injected current from the emitter into the base (I_n) and the undesired hole injection current from the bas into the emitter (I_p). These current will be expressed by using the Boltzman approximation,

$$I_n = qAN_B\left(D_n/W\right)e^{\left(-q\Delta V_n/kT\right)}$$
$$I_p = qAN_B\left(D_p/L_p\right)e^{\left(-q\Delta V_p/kT\right)} \qquad (18)$$
$$\Rightarrow \frac{I_n}{I_p} = \left(D_n/D_p\right)\left(L_p/W\right)\left(N_E/N_B\right)e^{\Delta E_g/kT}$$

where T is temperature, A the emitter-base junction area D_n the electron diffusivity in the base; W the base width, N_E the emitter doping concentration; D_p the hole diffusivity in the emitter; L_p the hole diffusion in length emitter.
The typical I-V characteristic of an HBT is showed in the Figure 16.

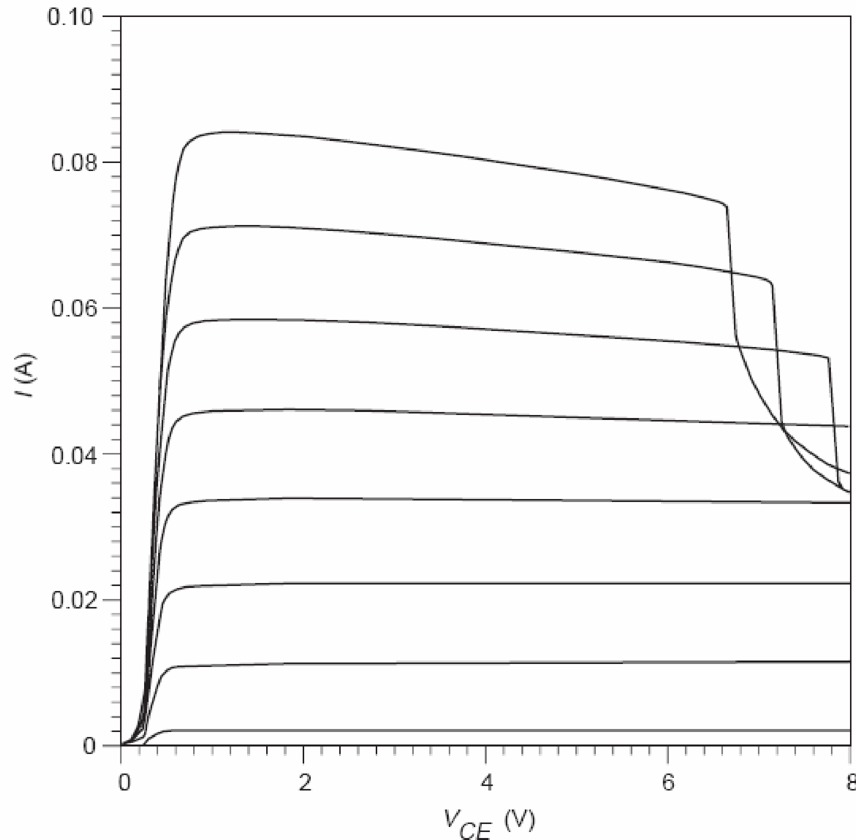


Figure 16- Typical I-V characteristic of a power HBT with multi finger design under collector current collapse.

RELIABILITY

As with other semiconductor devices, there are a number of mechanisms by which heterojunction bipolar transistor can fail The degradation mechanisms that have been reported in heterojunction bipolar transistors include the following:

- Decrease in current gain and increase in base-emitter voltage at high emitter currents.
- Increases in contact resistance caused by degradation of the interface between the emitter ohmic contact metallization and the emitter semiconductor. In AlGaAs contact layer is helpful in solving this problem
- Gettering of crystalline defects at the emitter-base heterojunction
- Decreasein current gain and increase in base-emitter voltage for a specified collector current caused by oxidation of the emitter mesa surface in the region of emitter base junction.

Specialized epitaxy growth for Be-doped-base HBTs, strain-relaxed base layer for C-coped base HBTs, the use oif an IgGaAs emitter cap layer, and emitter ledge passivation are some techniques used to alleviate some observed degradation mechanism.

At the state of the art there are three technologies used in order to produce GaAs transistor based

1) Pseudomorphic GaAs HEMTs
2)lattice matched and pseudomorphic Inp HEMTs

3)metamorphic GaAs HEMTs

The technological details are well recovered in literature. An excellent textbook is the [3]
In this context is more important to consider a comparison of bipolar and field effect transistor to look at their practical applications. The results of this research are summarized in the following tables:

Technology	Generic Requirement	Implication
Bipolar	Small base and emitter capacitance	Small emitter geometry
	Large current gain	Thin,highly doped base
	Small base resistance	Heterojunction technology
	Low parasitic resistance and capacitance	High mobility and velocity materials
FET	Small gate capacitance	Small footprint T-gate structure
	Large transconductance	Device channel close to surface
	Small gate resistance	Heterojunction technology
	Low parasitic resistance and capacitance	High mobility and velocity materials

Technology	Minimum feature size	f_T	f_{max}
Si bipolar	0.5 µm	50 GHz	50GHz
SiGe HBT	0.8 µm	130 GHz	160GHz
GaAs HBT	1.0 µm	180 GHz	280GHz
InP HBT	1.0 µm	228 GHz	270GHz
GaAs MESFET	0.2 µm	80 GHz	120GHz
GaAs pHEMT	0.12 µm	120 GHz	200GHz
InP HEMT	0.12 µm	250 GHz	350GHz
GaAs mHEMT	0.12 µm	225 GHz	250 GHz

It is clear that III-V devices outperform silicon technologies, with InP-based HBT and short gate-length HEMT processes offering the highest frequency operation, in keeping with generic requirements summarised in Table

2.5-Biasing transistor Networks and operations mode for MW transistors.

As well known by basic electronics transistors may be used either as amplifiers and switches. There exist many different ways to bias a transistor related to different class of operations [7]. For amplifiers as interstage gain and Low Noise Amplifiers the most commonly used class is the A class, for output stages (buffers) it is possible to use the B class or A/B class. Others class as C, D , E , F… can be used to employ the transistor as power amplifier. Actually the class of operation is a confused topic. This confusion is generated by the different definitions of "class of operation"; in fact, with this term, , a variety of different subjects are indicated. A first definition is given by considering the duty cycle shape of the collector or drain current for a device.

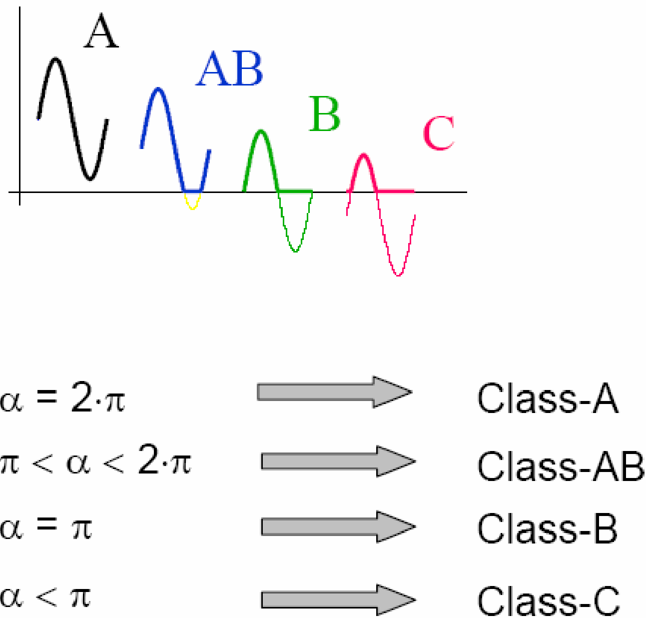


Figure 17- A first operational class definition by considering the duty cycle of the current.

Another definition can be done by considering the selected quiescent bias point

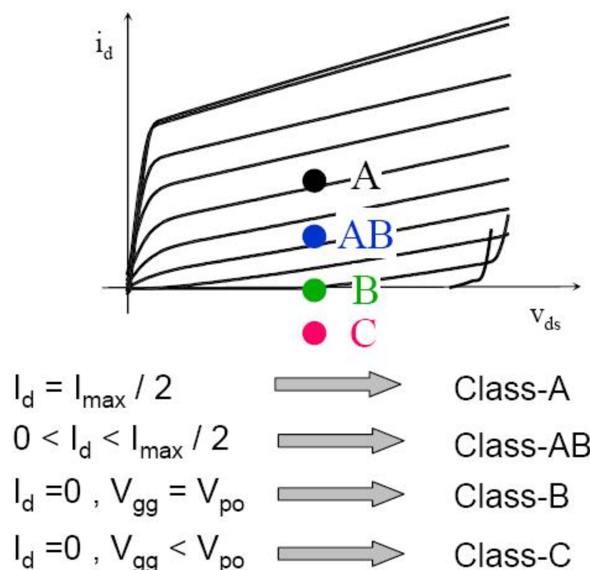


Figure 18- Classo f operations related to the position of quiescent bias point.

Since the duty cycle depends on the quiescent bias point, on the drive level and on the output termination, therefore the two definitions are not equivalent. Besides the term class C can be also indicated as an operating mode in which the device is forces to operate as a switch in subclasses E and D or an harmonic input/output tuning strategy for example when it charges with tunable loads or in class F. In this situation the active device acts as a controlled current source (as an amplifier) and is loaded by suitable terminations at harmonic frequencies. A given device may be biased in a given biasing class and may adopt an harmonic tuning strategy: for instance ,a class AB-class F amplifier stage. The design strategies in general for power amplifiers that represent a particular class in RF and MW circuit design, become mandatory to obtain highest performances from the available devices as examples lets considers; the switched mode operating condition that can be indicated as Class E the harmonic tuning approaches as (Class F, HM,….) and the architectural solutions as (Doherty, EER, ET….) amplifiers.

In general for the design of LOW NOISE amplifiers and Oscillators the class of operation may be considered as A class. It is important to consider then for a given transistor the DC bias characteristics achievable from a given bias network. The commercial CAD systems as ADS provide a series of tools to evaluate the DC performances for a class A BJT or HEMT.
Evaluate the DC properties is the first step to chose the configuration and the bias network.
We will show hoe work some of these tests for a PHEMT, same results can be achieved by considering a HBT from others foundries or commercial transistors used in MIC design.

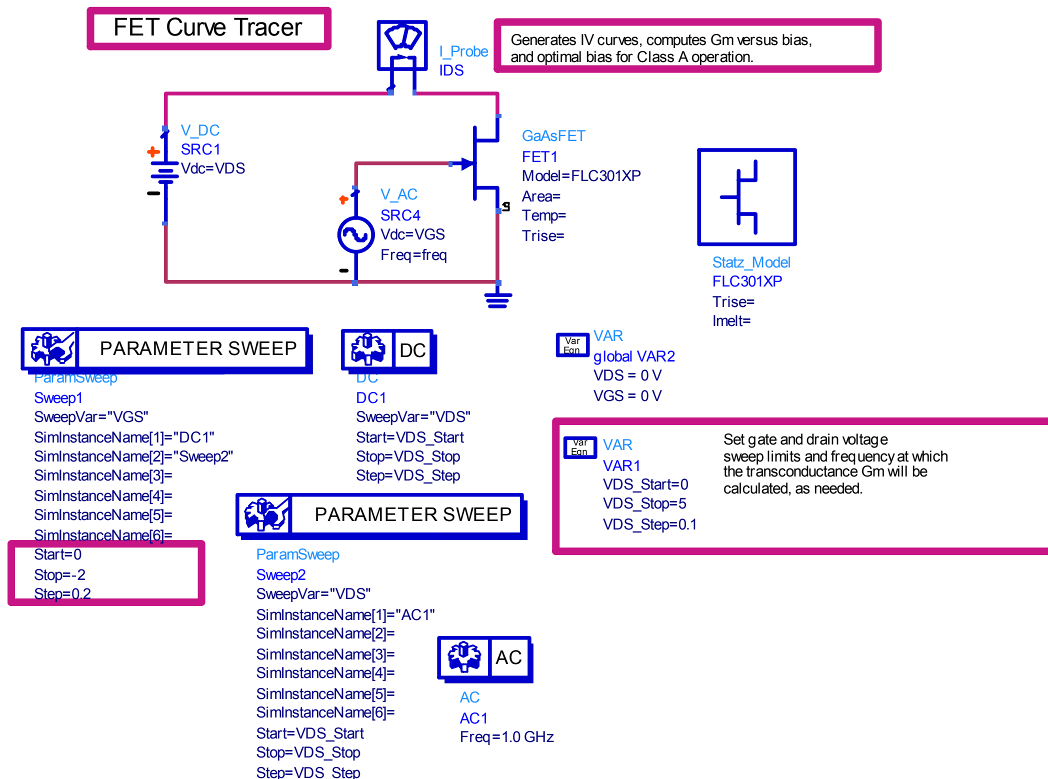


Figure 19- FET curve tracer schematic with ADS.

The figure shoe how to test a FET transistor in order to trace with CAD its transcharacteristic I-V curve. The FET and the reference of its models are placed in the schematic. Therefore through a signal and a time constant voltage sources, and a current probe, it is possible to execute the parameterized DC analysis. To achieve a good estimation it is important to set correctly the parameters of simulation that would be given by the constructor, as for example the bias voltage and the range of V_{GS} for which the transistor is given as operating. The V_{DS} must be changed from 0 and V_{DD} where V_{DD} is the maximum supply voltage. For modern MMIC the supply decrease continuously and became more difficult to manage bias and signals of good amplitude because. The Obviously the I-V characteristic for the device depends by V_{GS} value as well then the whole set of I-V curves is called transcharacteristic of the device. The V_{GS} values must be chosen carefully. For a MESFET the V_{GS} is negative then the voltage in the gate is zero or minus than in the source. The results of the DC analysis is showed in Figure 20

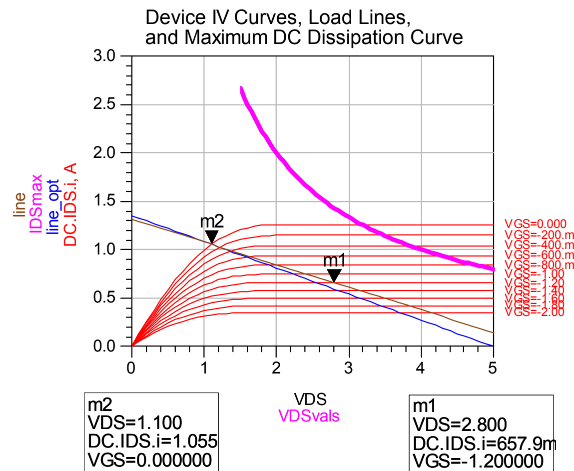


Figure 20- DC analysis of MESFET

The transistor presented in the example is an ideal model. For real models as that present in ED02AH library the curves are not as straight but for the modulation channel length also the saturated part of the curves will be actually skew.
Once the I-V curves are sketched it is possible to determinate the optimal Class A bias point that will be represented by the *VDS* and *VGS* values for which the current at the drain will be exactly half of the maximum possible. At this point the conversion efficiency for the transistor will be maximum and its value will be about 25%.

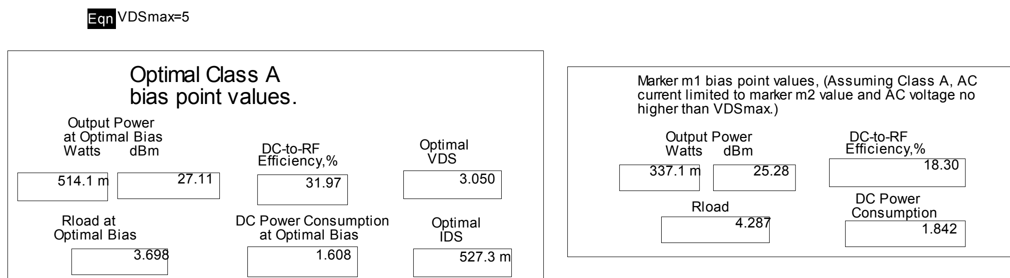


Figure 21- Power calculus for bias point selected by marker m1.

Moving the marker m1 it is possible to compute some different parameters for the devices as the Rload namely the total resistance of bias for the transistor , the current, the DC-RF conversion and DC power consumption for the chosen couple of *VDS* and *VGS* values. The AC simulation may be used in order to compute the transconductance value for the selected bias.

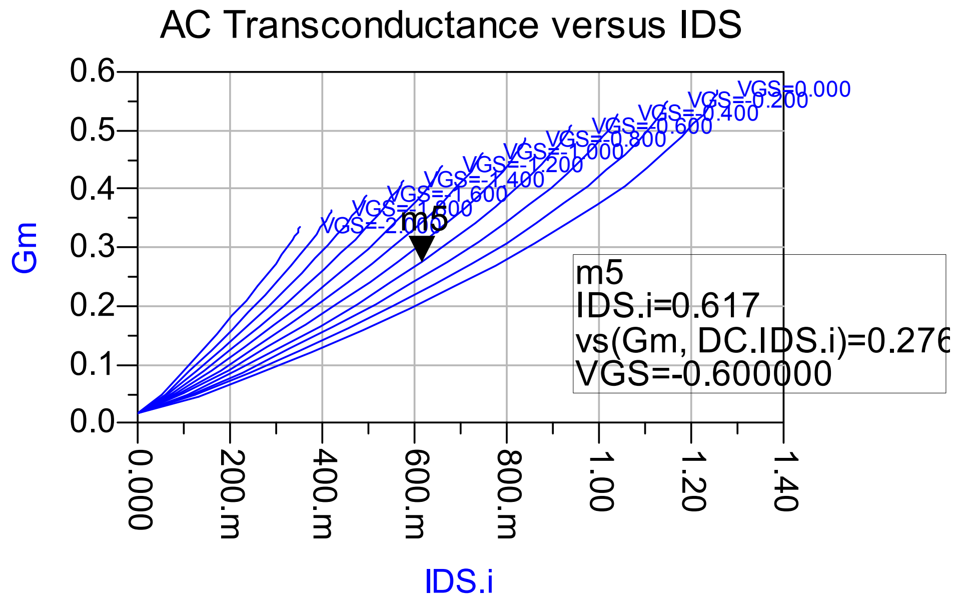
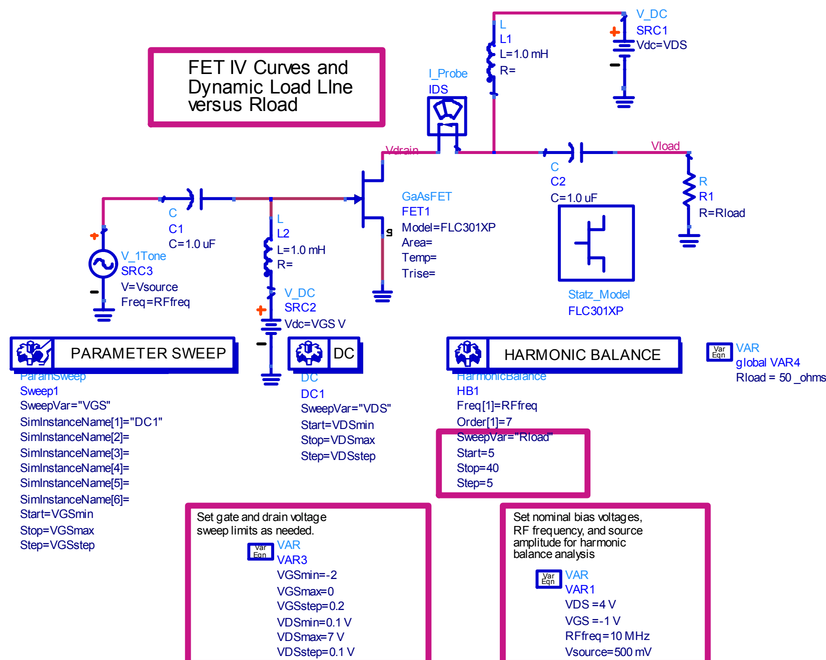


Figure 22- AC simulation of MESFET to compute the value of gm for different bias points.

Once the bias point has been chosen another important test it consist in measuring the output power versus load resistance. Obviously to execute this test it is necessary to know the level of the signal that the transistor must process. An example of this simulation is showed in the following figure



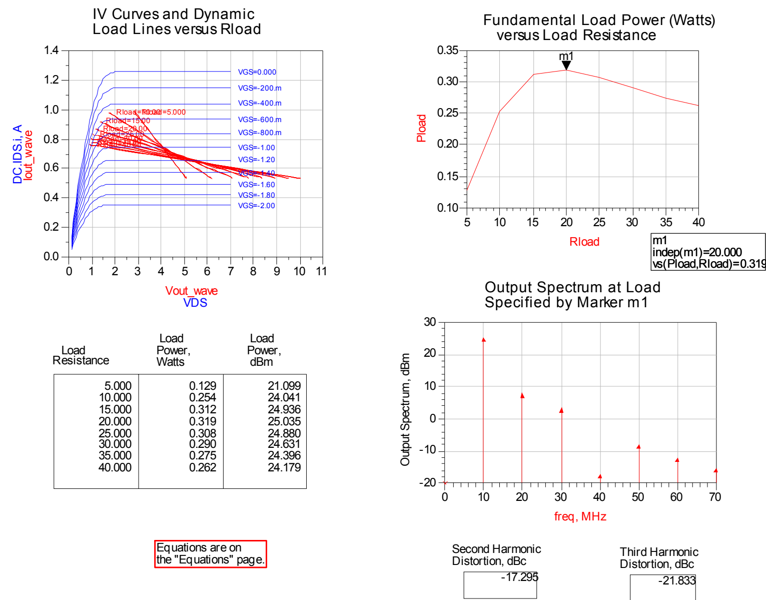
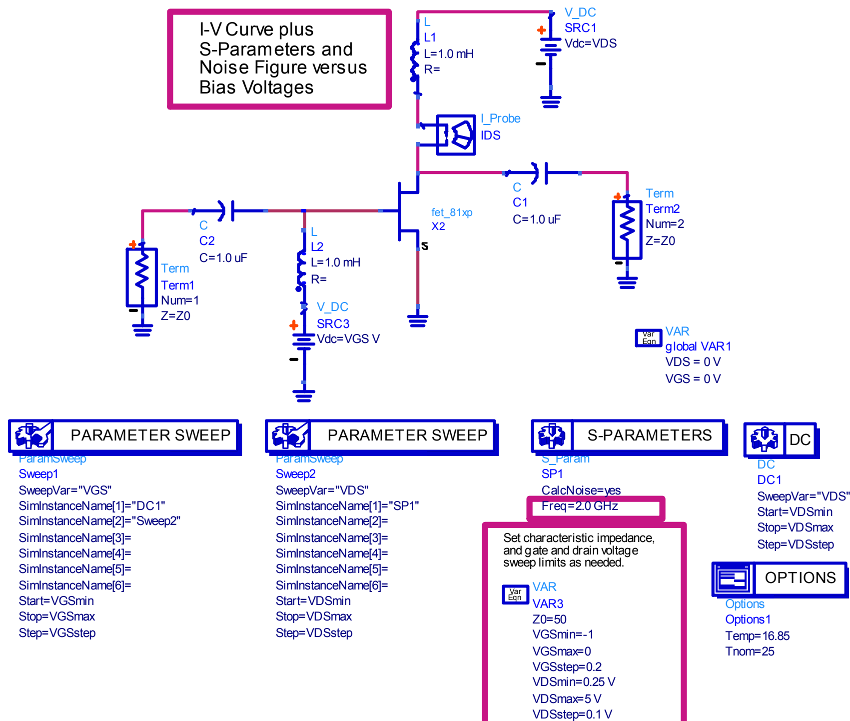
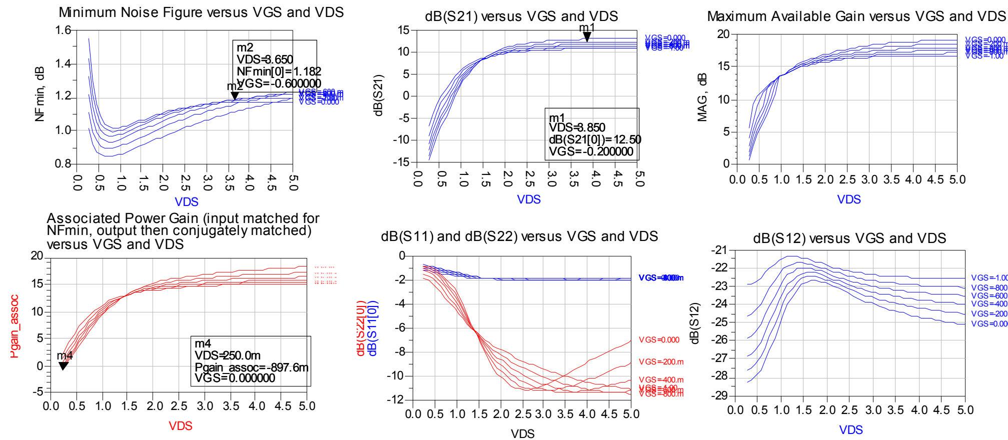


Figure 23- DC power loads.

Others test can be done used the S-parameters simulators in order to discover which loadswould be better to maximum gain and to minimum noise figure.





Set step sizes and number of circles, here.

Eqn num_NFcircles=3
Eqn NFstep_size=2
Eqn num_GAcircles=3
Eqn GAstep_size=1
Eqn num_GPcircles=3
Eqn GPstep_size=1

Stability Factor, K: 0.563

Source Stable Region: Outside

Zsource, Source Gamma — DUT* — Zload, Load Gamma

*DUT= Device Under Test (simulated circuitor device)

Available Gain & Noise Circles, Source Stability Circle, Source Gamma, Corresponding Load Gamma, (Black Dot)

GammaS

Move markers GammaS and GammaL to select arbitrary source and load reflection coefficients. The impedances, power gains, and noise figures below will be updated. The transducer power gains are invalid if the markers are moved into the unstable regions (usually inside the stability circles.)

Move marker mBiasPt to desired bias point. Smith Chart and data below will be updated.

IDS

mBiasPt

VDS

mBiasPt
VDS=2.550
IDS=275.6m
VGS=-0.200000

RF Frequency: 2.000 GHz

System Impedance, Z0: 50.000

Power Gain Circles, Load Stability Circle, Load Gamma, GammaL, Corresponding Source Gamma, (Black Dot)

GammaL

Load Stable Region: Outside

See "NF, SP, Gains at all Bias Pts." page.

See "Matching at 1 Bias Point" page for optimal source and load impedances to match for gain or noise figure. Equations for this page are on the "Equations2" page.

Noise Figure (dB) with Source Impedance at marker GammaS	Source Impedance at marker GammaS	Optimal load impedance for power transfer when source impedance at marker GammaS is presented to input	Transducer Power Gain, dB when these source and load impedances are used
1.175	16.125 + j11.197	21.363 + j25.182	16.305

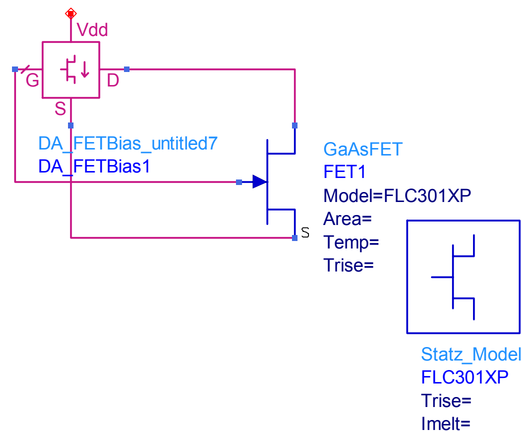
NFmin, dB	Source Impedance, Zopt, for Minimum NF	Optimal load impedance for power transfer when source impedance is Zopt	Transducer Power Gain, dB when these source and load impedances are used
1.139	19.717 + j14.807	21.229 + j20.794	15.790

Noise Figure (dB) with Zsource (only valid with K>1)	Simultaneous Conjugate Matching (only valid if stability factor K > 1) Zsource	Zload	Maximum Available Gain, dB (Maximum Stable Gain, S21/S12) if K<1
1000	50.000	50.000	17.463

Noise Figure (dB) with this optimal source impedance (at right)	Optimal source impedance for power transfer when load impedance at marker GammaL is presented to output	Load Impedance at marker GammaL	Transducer Power Gain, dB when these source and load impedances are used
2.010	5.475 + j9.182	123.450 + j91.013	13.752

Figure 24- Test for noise figure and to automatize the chose of matching input and output networks in function of the bias point.

The bias network requires actives elements, necessary to provide a constant voltage at some terminals of the transistor. For this purpose the inductances or RF-choke are applied. The RF chokes can implemented with a large inductance values or with λ/4 stubs used for example in MIC implementation. With referring of the schematic of Figure 24 we can describe a special feature of the RF-choke. Since they are active elements their property is to store energy. Then the Drain can overcome the limits imposed by the supply voltage without make off the transistor. A consequence of this is that in order to maximize the power for the transistor it is possible to choose the bias point in way that the $I_D=(I_{Dmax})/2$ and $V_{DS}=V_{DD}$. This situation can be sometimes used to choose a bias point in oscillator design. Once the bias point is selected it is necessary to implement the bias network. The bias network can be passive or active depends by situations and by the designer's purposes. Commercial CADs as ADS provide some tools to design automatically the network. As an example if the designer would be implement a possible network that fixes the bias point at the m1 point for the transistor represented in Figures 20 it is possible to employ a powerful tool called "Transistor bias tool" then put into the apposite spaces the values achieved from the previously analysis end chose the bias network between all possible.



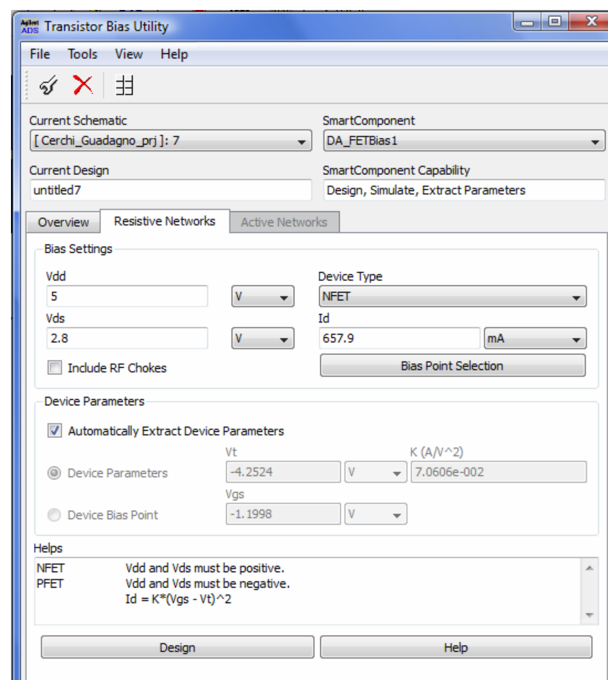


Figure 25

After click on design the tools propose a series of networks that can be used to achieve the desired point and the designer must chose one of these. The bias designed is not in general a good network but must be adapted to reduce the power consumption an modified to add the RF-chokes and DC-block. The starting and final lumped ideals networks with the DC simulation are represented in the figure below.

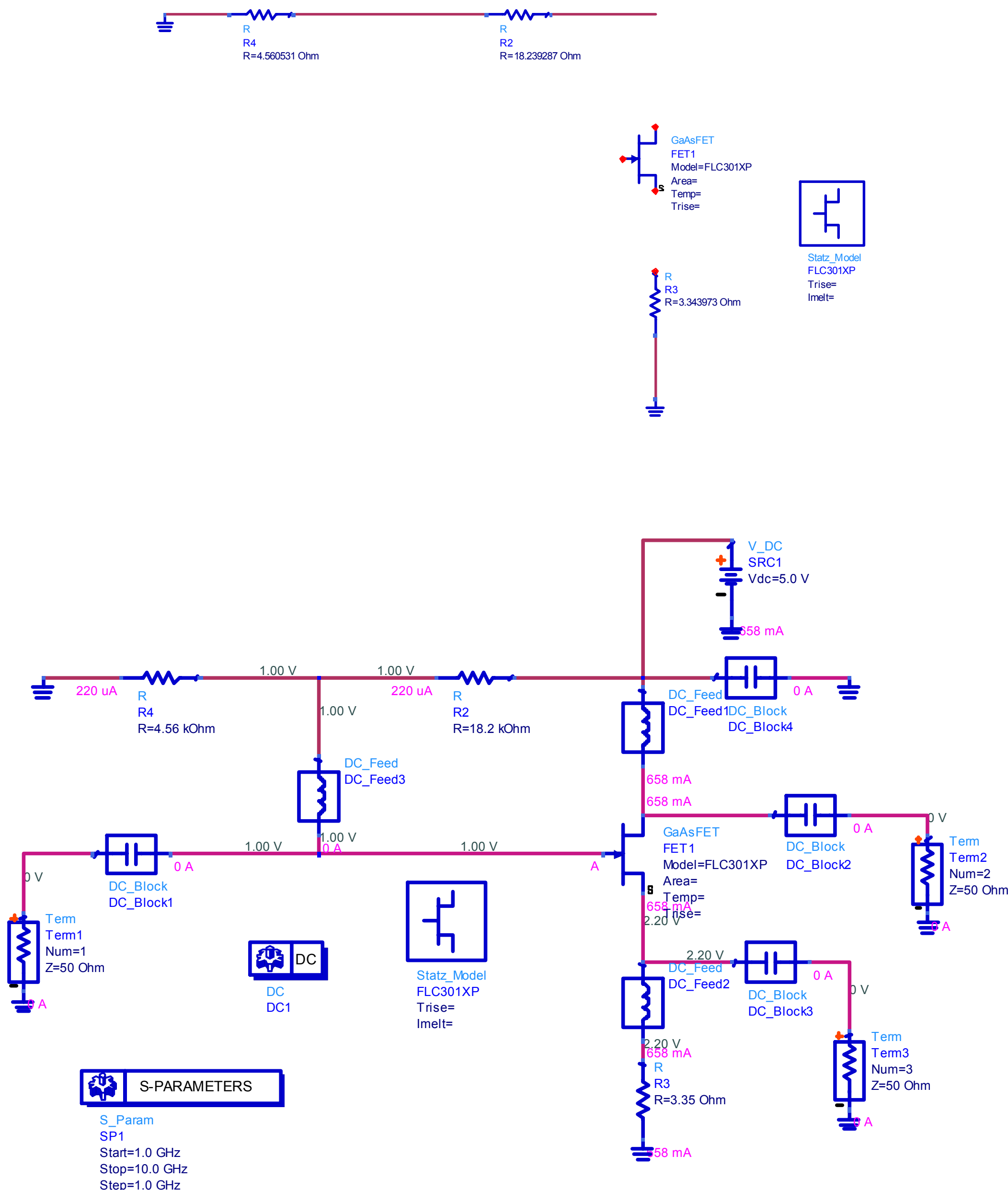


Figure 26- Bias Network ideal with DC simulations DC-FEED and DC-BLOCK

Figure 26 show a possible real bias network for the transistor. The DC-FEEDs and DC-Block in the real circuit would be replaced with real inductor or capacitances which sizing would be determined through the classical try and error procedure. The DC-BLOCK near to the DC source is useful to block some signals that may enter inside the chip. The Terms element are used to determine the S parameter of S-Parameters simulations at various ports. Once the DC-FEEDs and DC-BLOCKs are replaced by proper elements (also implemented in microstrip solutions) the elements would be connected through microstrip lines. The next step consists in realize the layout of the circuit. With

this transistor is no possible to generate the layout because for this purpose we need of a physical transistor model. But these steps can be followed in the same way when we are going to design a real circuit. In Next chapters we will present some layout examples for real circuits.
Same circuits and tools can be used to perform the analysis of HBT or BJT, for every technology and models.
Another tool that can be used to study the bias the transistor is the active tool. The active bias networks are used in CMOS technology and to perform biasing for differential amplifiers, or oscillator. The design principle is the well known current mirror biasing. There are many references for active bias networks as for example [4].
In this context is more important to present a list of default libraries of ADS for active transistors that can be used in MIC design.
The libraries that contain transistors with layout are listed below:
I modelli di transistor per cui si dispone di Layout sono riportati nella libreria:

High frequency diodes
Microwave Transistor Library
RF Transistor Library

In each library there are a lot of transistors of different topologies. Some of these are listed in the table, but some are obsolete and it is impossible to find the Datasheet. Ph refer to pseudomorphic hemt, Pb to hbt and mf to power mosfets.

Ph_nec_NE33284A_19921223	Ok
Ph_sny_2SK676_19921217	Ok
Ph_fuj_FH35LG_19921222	Obsolete
Ph_syn_SGH5002F_19931224	Ok
Ph_hp_ATF35176_19921221	Ok
Pb_nec_NE68030_19961001	Obsolete
Pb_mot_MRF942_19911009	Ok
Pb_sms_BFQ72_19911009	Ok
Pb_sms_PZT43_19961209	Obsolete
Pb_mot_MRF2396_19911003	Obsolete
Mf_phi_UF2805B_19930106	Ok
Mf_ply_F2001_19930106	Obsolete
Mf_phi_LF_199303106	Obsolete
Mf_mot_MRF134_19930106	Ok
Mf_mot_MRF136_19930106	Ok

The designer can require design kits to companies as for example ED02AH, Triquint, jazz semiconductors etc.

2.6-Characteristics of passive GaAs components

The passive elements are critical for the performance and reliability of the circuit, and determine the circuit’s bandwidth, the center frequency and other electrical characteristics. Besides connecting the various active elements together, passive elements are used to set the bias point for the circuit and impedance match the active devices to themselves and the input and output connection for the MMIC. Passive elements are composed of lumped elements such as resistors, capacitors and inductors and distributed elements such as transmission lines. In general, distributed elements are

physically large enough that transmission line characteristics play a significant role in their function. Distributed elements have inductive, capacitive and resistive aspects, all of which are taken into account by the transmission line analysis. The rule of thumb is that an element must be considered as a distributed element if it has dimensions greater than $\lambda/10$, where λ is the wavelength. Lumped elements on the other hand are small enough that transmission line effects do not play a significant role in their function. Nevertheless, even lumped elements are not purely inductive, resistive or capacitive, but have aspects of all three due to parasitic. Because of these parasitic effects, lumped elements will resonate at some frequency. Failure to account for this complexity may lead to significant errors in circuit design. CAD programs such ADS, MWO that are used in MMIC design contains models for most passive elements. Unfortunately, circuit designs do not always meet the modelled electrical performance goals because of unacconted-for electromagnetic coupling effects and limitations of the models themselves. The ability to fabricate MMICs and realize the advantage of the size, ruggedness reproducibility, and cost require that methods exist do fabricate the required passive components. Economic considerations require the MMIC chip size to be minimized as much as possible to maximize the number of chips on wafer. Unfortunately, the passive components, especially inductive elements, tend to be large. In addition the elements must be separated from each other to minimize electromagnetic coupling between the elements. The result is that the active elements occupy a small portion of the MMIC area. Lastly, all passive components must be fabricated on the semi-insulating GaAs substrate. The following is a briefly description of passive elements, methods o fabrication.

A- Resistors

Resistors are used for feedback circuits, setting the bias point of active devices, isolation, and terminations in power combiners and couplers. Two type of resistors are used in MMIC fabrication: thin films of lossy metals and lightly doped GaAs active layers. Fig 27 shows schematics of each of these two types of resistors. The resistance for both types of resistors may be given by:

$$R = \frac{\rho_s L}{A} \qquad (19)$$

Where ρ_s is the sheet resistivity, $A = W \cdot t$ is the cross-sectional of the resistor, t is the resistor thickness, W is the width of the resistor and L is the length. The efficiency of the resistor as determined by the resistance per unit length is a function of ρ_s . For metal thin film resistors, ρ_s is a function of the metal. For GaAs based resistors, ρ_s is a function of doping concentration. Metal thin.film resistors are used for accurate, low-resistance applications. They are usually fabricated for TaN and NiCr, Ti, Ta, Ta_2N and AuGeNi alloys have also been used. Some of the advantages of thin-film resistors are a low Temperature Coefficient of Resistance (TCR), tight tolerances, small parasitics, and low sheet resistivity. The major disadvantage of thin film resistors is the added processing steps required to fabricate them, although thermal dissipation difficulties and electromigration failures are also a concern. Short resistors or those made from materials with a large resistivity have fewer parasitics, but they have an higher thermal load to dissipate. Then resistors must dissipate large amounts of power, they can have the highest temperature on the MMIC and limit the MMIC reliability. Sidegating on the flow of current around the perimeter of the resistor is usually eliminated by depositing the resistor on top of an insulating film such as Si_3N_4 .Electromigration failures result from large current densities that can flow through the thin metal films. Tantalum resistors have exhibited this problem for thin, 0.006mA/µm of line width Lastly, NiCr resistors are susceptible to degradation due to moisture. Therefore, these resistors must be passivated. GaAs-based resistors are implemented by the use of FET channel and ohmic contacts that are alteady available in the MMIC fabrication process. The total resistance of these elements is the sum of the resistance of the GaAs channel itself and the two ohmic contacts. The advantage of

GaAs –based resistors is the wide range of resistivities available through changes in the doping level. GaAs-based resistors have several potential problems, however: current saturation, Gunn domain formation and an high TCR. Above a critical field, the current in GaAs will saturate and the device loses its linearity. In practice, this is not a severe limitation since the length of the resistors is usually sufficient to prevent the electric field from reaching its critical value. Gunn domain formation, the initiation of microwave oscillations due to an applied static electric field, also occurs only if large electric field are present. A more serious problem is the large positive TCR +3000ppm°/C. This can result in significant resistance changes over temperature. Fortunately, modelling tecniques can be used to determine the resistance change that is tolerable for a given circuit and application. The issues of current handling capacity, thermal dissipation, and distributed effects also play a role in the design and operation of GaAs base resistors. The resistor, especially one in the DC source or drain circuit, must be able to handle the current passing through it without reaching saturation. GaAs is a relatively poor thermal conductor and will not be able to remove heat rapidly enough if too much power is dissipated in too small an area. Physically large resistors on the other hand will become distributed elements and act as lossy transmission lines. In general, GaAs-based resistors may be thought of as a gateless FET. If the resistor is properly designed to operate below the current handling limit and the critical electric field limit, thermal heating should not be a problem.

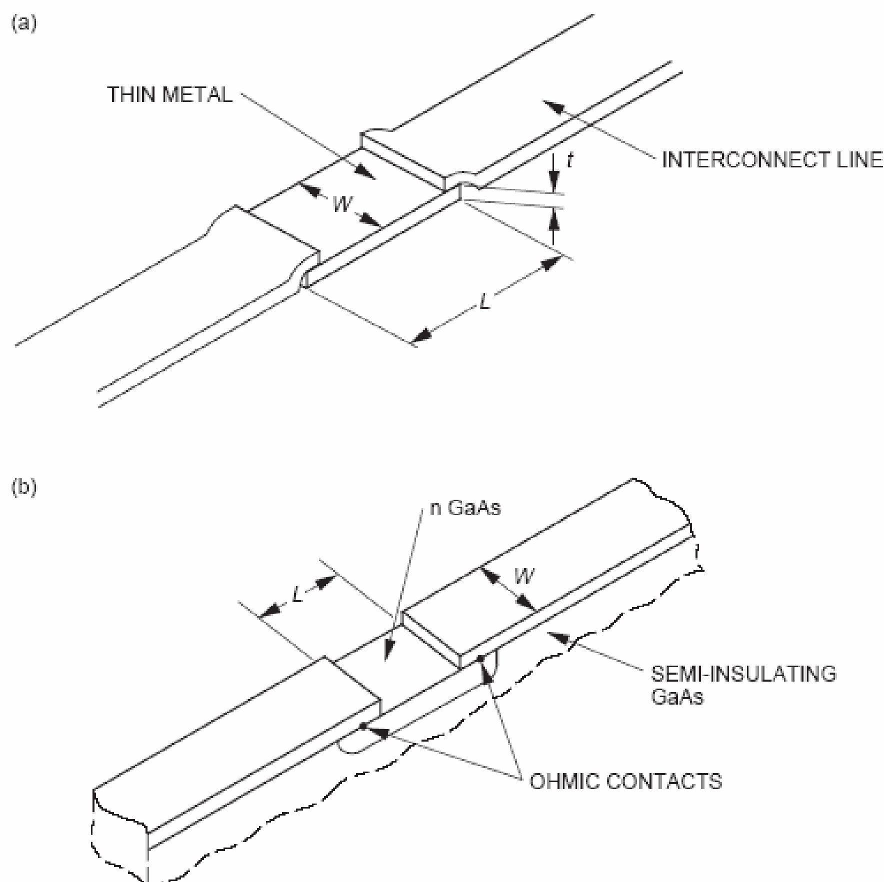


Figure 27- Two resistors type in MMIC fabrication: (a) thin film and (b) GaAs-based resistor incorporating an GaAs channel ohmic contacts.

In the case of semiconductor resistors, a linear current voltage relationship only exists for electric field strengths of less than $1KVcm^{-1}$. To avoid non-linearity, the resistor length should be chosen so

that the field across the structure is less than $1KVcm^{-1}$, and thus a resistor that is required to drop 2V should be at least 20µm long. The choice of resistor structure (thin film or semiconductor) will depend on the desired value. Thin metal films typically have a sheet resistance in the range 10-100Ω/square and thus are useful for relatively low values of resistor up to a few hundred ohms. Larger resistor values (e.g for biasing circuitry) can be realized using layers in the substrate's vertical architecture. Typical sheet resistances of such semiconductors layers are in the range of 200-600Ω/square and so are well suited to realizing resistances up to a few KΩs. A simple RF equivalent model for a resistor is showed in the figure below

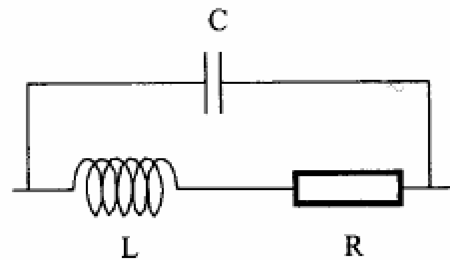


Figure 28- Lumped equivalent circuito f a resistor

As show in figure a resistor does not has only a resistive part but also an inductive and capacitive part. It is possible by applying the Laplace's transform to compute the expression for its equivalent impedance parameterized in frequency, and then also admittance or reflection coefficient expression.

$$
\begin{aligned}
Z(\omega) &= (R + j\omega L) // \left(\frac{-j}{\omega C} \right) \Rightarrow \\
R(\omega) + jX(\omega) &= \frac{L\omega^2 CR - R + \omega^2 LCR}{(\omega CR)^2 + (1-\omega^2 LC)^2} + j\frac{\omega^3 L^2 C - \omega CR^2 - L\omega}{(\omega CR)^2 + (1-\omega^2 LC)^2}
\end{aligned}
\tag{20}
$$

Both real and imaginary parts expressions are more complicated and can be better simulated on CAD. For frequency values that make zero the imaginary part of impedance and can be better studied with CAD simulators.

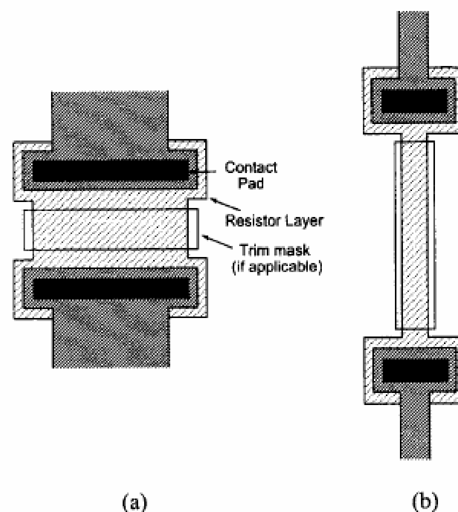


Figure 29- Resistor examples : (a) small value (50Ω) and (b) large value (3000Ω)

Capacitors

Capacitance may be used in MMIC circuits in any of four basic configurations: an open-circuit transmission line, coupled lines or interdigitated capacitors, Shottky diodes, and metal-insulator-metal (MIM) capacitors. Coupled lined and open-circuit transmission lines can be used to provide fairly low capacitance values. For these two capacitor types, the capacitance is dependent on the electrical length of the transmission lines. Therefore, the capacitance is highly frequency dependent. The advantage of these capacitors is that they are easy to fabricate since they require only a single metal layer. The most popular type of capacitors for MMICs has become the MIM capacitor because of the high capacitance per unit area can be obtained. Therefore, smaller and less costly circuits are possible. A schematic of an MIM capacitor is shown in Figure 29. This thin-film capacitor is composed of two metal plates separated by a dielectric material. Typically, the dielectric materials overlaps the first metal layer and the upper metal layer has a smaller area than the lower. This configuration helps to minimize fringing fields to ground and shorts between the upper and lower capacitor plates. Although the air bridge shown in Figure 29 is not required, it is often included for further minimize parasitic capacitance. The dielectric is typically silicon nitride (Si_3N_4 0.1 to 0.4 μm thick) since it is already used in MMIC fabrication process for circuit incapsulation although SiO_2 and Ta_2O_5 are also used. Since the dielectric layer is substantially thinner than the substrate thickness, MIM capacitors exhibit significant fringing effects, which are function of the perimeter. Careful experimentation to determine the magnitude of this effects for specific process parameters, such as dielectric type and thickness, is essential for any stable process. Test capacitors should also be included on the wafer for in-process verification. The yield of MIM capacitors on a wafer plays a major role in determining the total yield for the wafer. One pinhole that ruins one capacitor also ruins the entire chip. The problem can be illustrated by considering a complex MMIC chip with ten capacitor yield of 95%. This case would produce a chip yield of 60% on capacitor defects alone. The major yield limiting factor for MIM capacitors is shorts caused by pinholes in the dielectric or sharp points on the metal plates. Pinholes are very difficult to eliminate completely, but they can be minimized by good cleaning and deposition processes. Again, trade-offs require an engineering judgment based on the experimental and the realized yields for a particular process. In addition to pinholes, the circuit design must assure that the electric field across the capacitor does not exceed the dielectric breakdown field. These MIM that consist of a metal-insulator-metal "sandwich" with the most common insulators being silicon nitride, silicon dioxide, and polyimide. Silicon nitride is popular since it has a fairly ε_r and can also be used for passivatinf the exposed GaAs in the active devices. IN processes which use polyimide as the spacer dielectric for spiral inductors, the polyimide can also be used to realise small value capacitors for applications such matching networks and filter inter-resonator coupling, and they are significantly smaller than interdigital capacitors. The type of connection used from the capacitor to the res of the circuit depends on whether air-bridges or a polyimide-based two metal level process are used. In a two metal level process the main microstrip circuitry is normally on the upper metal, and so the lower plate has to be brought up to the top metal through a via-hole in the dielectric, as shown in Figure 31-a. Alternatively , with air-bridges the upper plate is connected to the rest of the circuit as shown in Figure 31-b. This is preferable to the method shown in Figure 31-c because the sharp edge along the interconnection would make a short-circuit failure more likely and reduce the breakdown voltage rating of the capacitor. Since the width of air-bridges is limited by the need during fabrication to remove the temporary photoresist support, often several air-bridges would be put in parallel in order to achieve a low inductance connection, rather than using a single wide air-bridge. The typical equivalent model of a monolithic capacitor is showed in Figure 32.

The equivalent circuit s much complicated to study without a CAD. In general the parasitic inductances and capacitances should be minimized in order to improve the yield.

Another way to build a GaAs-based capacitors is through the interdigital technique.

These consists of a number of interleaved microstrip fingers coupled together as shown in Figure 32.

The maximum value of an interdigital capacitor is limited by its physical size, and its maximum usable operating frequency is limited by the distributed nature of the fingers. They certainly cannot be used for values above 1pF and even 0.5 pF, interdigital capacitors will measure approximately 400X400 μm^2 . Nevertheless, since interdigital capacitors do not use a dielectric film, their capacitance tolerance is very good and is limited only by the accuracy of the metal pattern definition. Hence they are ideal as tuning, coupling and matching elements, where small capacitors values are required and precise values are necessary.

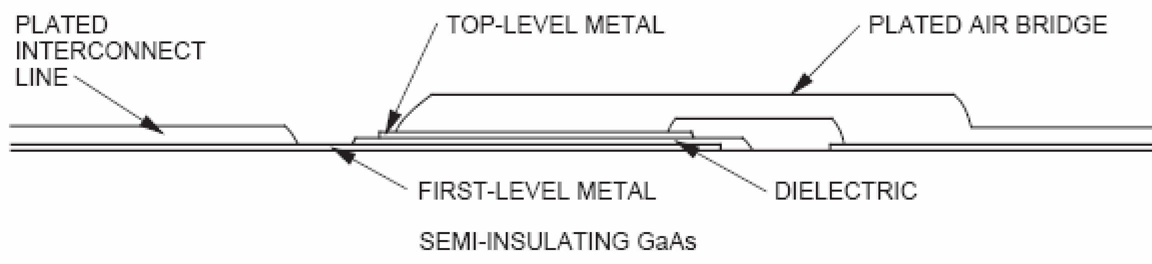


Figure 30- MIM capacitor using an air bridge for top-level interconnect.

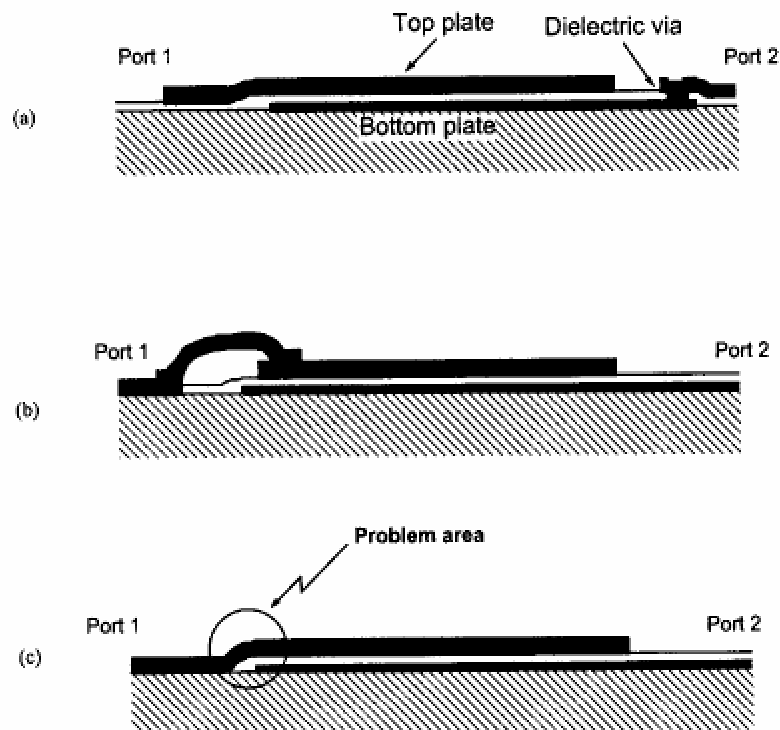


Figure 31- Overlay capacitors: (a) using a dielectric via (b) with an air-bridge and (c) without an air bridge or spacer dielectric.

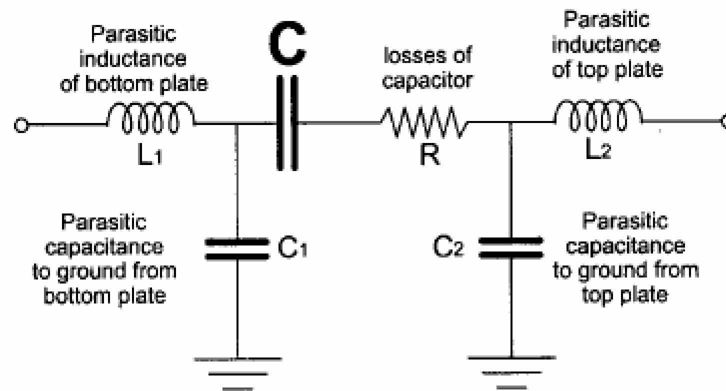


Figure 32- Capacitor model

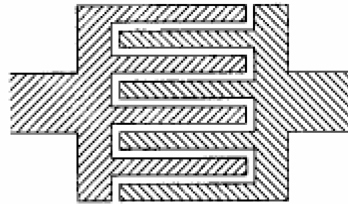


Figure 33- Interdigital capacitor

Inductors

Inductors are necessary elements in MMICs where they function as tuning elements and RF chockes in DC bias circuits. Inductors are one of the easiest passive elements to fabricate. As a distribute elements they are realized by a section of high-impedance transmission line. These inductors are limited to inductance values below 2 to 3 nH because of the high losses associated with the long length of high-impedance transmission lines. Lumped element inductors can be used to provide inductance up to 20nH. Depending on the inductance requires, MMICs inductors can be realised either as straight narrow tracks (ribbon inductors),as sinfle loop inductors or as multi-turn spiral inductors. A microstrip ribbon inductor and its equivalent circuit are shown in Figure 34.
For short lengths ($<\lambda/4$), the inductance and shunt end capacitances can be calculated from the following well known equations:

$$L = \frac{Z_0}{2\pi f}\sin\left(\frac{2\pi l}{\lambda}\right) \qquad (21)$$

and

$$C = \frac{1}{2\pi f Z_0}\tan\left(\frac{\pi l}{\lambda}\right) \qquad (22)$$

A narrow track with high Z_0 is needed to achieve high inductance with low parasitic capacitance. However, in practice the choice of track width is determined by fabrication limits, by the DC current carrying capacity and by the high resistance of very narrow tracks.

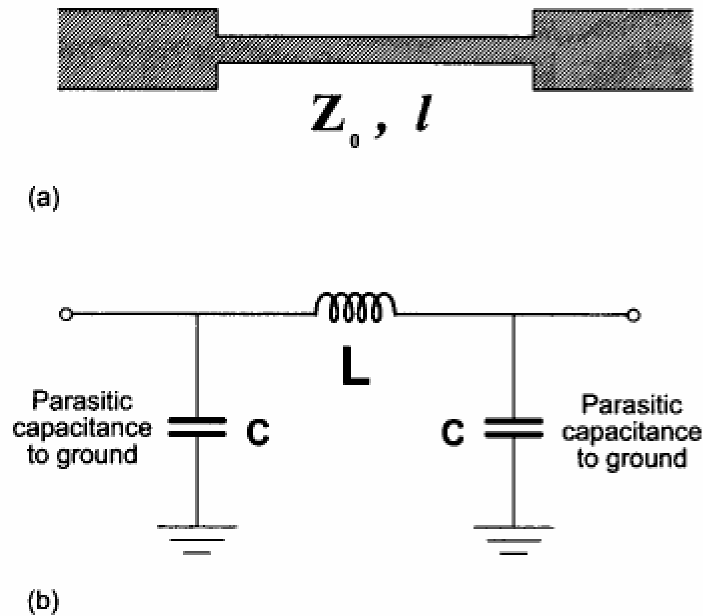


Figure 34- Ribbon inductor (a): microstrip layout and (b) equivalent circuit.

The track length is limited simply by the need to ensure a realistic economical chip size. The ribbon inductor is thus limited to values of less than 1nH, but is a relatively 'pure' inductor with low parasitics. Hence it is often used in distributed amplifiers where very large bandwidths are required. Another very simply inductor is the loop inductor like the one shown in Figure 35 were used extensively the pioneering days of MMICs. This probably because the processing did not a first offer air-bridges for spiral inductors, and because the limited experience of coupled lines on GaAs made designers reluctant to use meandered ribbon inductors. It is fair to say that in recent years loop inductors have been used very little because of their inefficient use of chip area. A number of text give useful design equations for loop inductors [3]

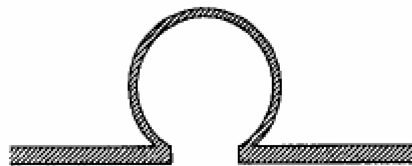


Figure 35- Layout of a single loop inductor

Most useful are the spiral inductor whose values can be chosen in the range of values from a few hundred of pH to a few nH. To understand the advantage of spiral inductor it is possible to compare the two different architectures showed in Figure 36 in which it is possible to note that for the Figure 36-(a) the current in adjacent tracks is flowing in opposite directions, whereasfor the spiral inductor Figure 36-(b) the current in adjacent tracks is in the same direction. The resulting mutual inductance yields a significant increase in the spiral inductor's overall inductance. IN the meandered case, however, the inductance is reduced and the overall inductance is significantly less than that of a straight track equal to the unwound length. Since the DC resistance is the same, the spiral inductor thus has higher Q. The drawback of the spiral is that the need to connect the centre turn back to the outside circuit dictates that either air-brige crossovers or dielectric spaced underpassed must be used. There are a number of different solutions to this same fundamental connection problem, and these are illusrated in Figure 37. Note that square inductors are shown; rectangular ones could be used if layout constraints made them necessary, and (approximately)

rounds ones would have slightly superior performance at the cost of greatly increased layout complexity and a less convenient shape for integration with other components. However, these alternative spiral inductors would have to be characterised first.

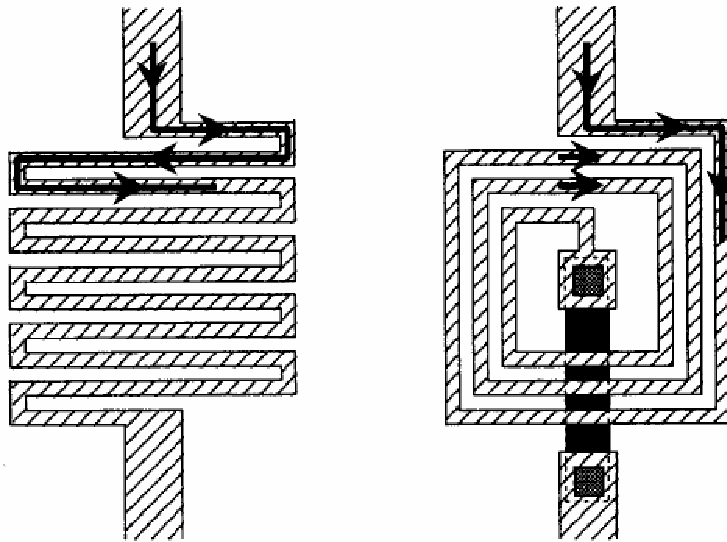


Figure 36- Large value inductance using (a) meandered track and (b) spiral.

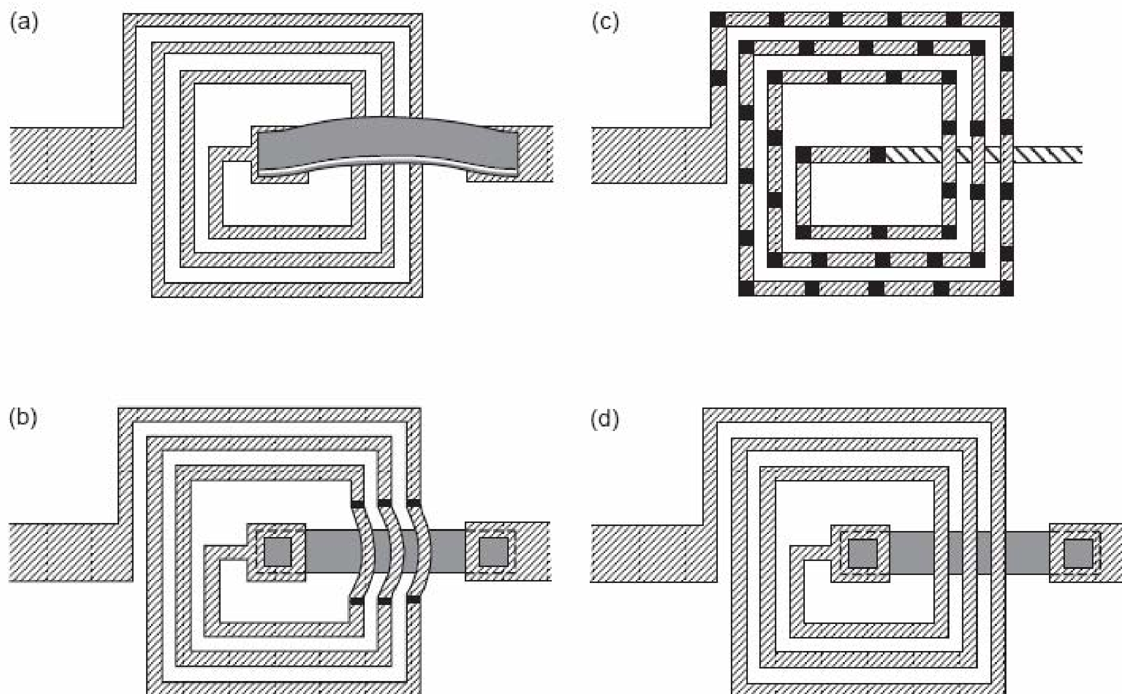


Figure 37- Spiral inductors; (a) single air-bridge, (b) air-bridge over an underpass, (c)formed entirely of air-bridges and (d) using two metal levels for underpass.

Note that square inductors are shown; rectangular ones could be used if layout constraints made them necessary, and (approximately) round ones would have slightly superior performances at the cost of greatly increased layout complexity and a less convenient shape for integration with other components. However, these alternative inductors would have to be characterised first.

Inductor type (a) has a single air-bridge span from the centre of the spiral to the outside. Whilst this type has been used successfully, it has the major drawback that the span of the ait-bridge is perilously large, especially when the inductor has many turns. The maximum span of an air-bridge is limited by fabrication technology and reliability issues. It is difficult to ease the problem by

having landing posts between the turns because adjacent tracks are very close. Hence, the more commonly accepted air-bridge solution is the type (b). In this type of spiral inductor the connection to the centre turns stays at the substrate level, and each turn of the inductor jumps over it using a short air bridge. This type thus has more ait-bridges, but they are each of minimum length so that reliability is improved. The third type of spiral inductor (c) actually forms the entire spiral out of air-bridges, which jump from landing post to landing post. Tha advantage of this structure is that since the track spends much of the time off the substrate surface, the parasitic capacitance between turns and ground is reduced, giving a higher useful maximum operating frequency. The alternative to using air bridges is to employ a two-level metal process with a spacer dielectric. Polyimide is often chosen as it has a low ε_r and because thick films can be deposited very quickly. In this type of inductors (d) , the turns are normally on the top metal layer above the dielectric so that the parasitic capacitance is reduced, and so that the turns can be plated up for less series resistance. The connection between the upper metal and the underpass is made with a via in the dielectric. These via should not be confused with the through-substrates used for grounding components. On standard silicon substrates the spiral inductors has parasitic which are so severe that the inductor is virtually unless. Thus, there has been a great deal of activity to make improved inductors on silicon, using either multilayer techniques or micromachining.
When one spiral inductor does not have enough performances it is possible to implement stacked spiral inductors. Stacked spiral inductors sonsist of a pair of inter-wound spirals placed on separate metal layers. The major advantage is that the turns are much more tightly packed than normal potholithography and metal patterning would allow for a single layer spiral, and since they are separated vertically to some extent there is less capacitance between adjacent turns than there would normal be with such small gaps. However, since the dielectric spacer can only planarise the circuit to certain extent, the lower metal thickness is limited to 1 μm or less, and this means that the stacked spiral has a high series resistance. The inductance per unit area of stacked spiral inductors is very high, but theur resonant frequency is much lower and so they are limited to frequencies below 5 GHz or so. They often find application as bias chokes, where high inductance is required, and in this application the high series resistance can actually be beneficial to the amplifier's stability, matching and bandwidth. The basic inductor equivalent circuit consists of a prime inductance along with its associated series resistance, inter-turn and crossover feedback capacitance, and some capacitance to ground. This lumped element model is shown in Figure 38.

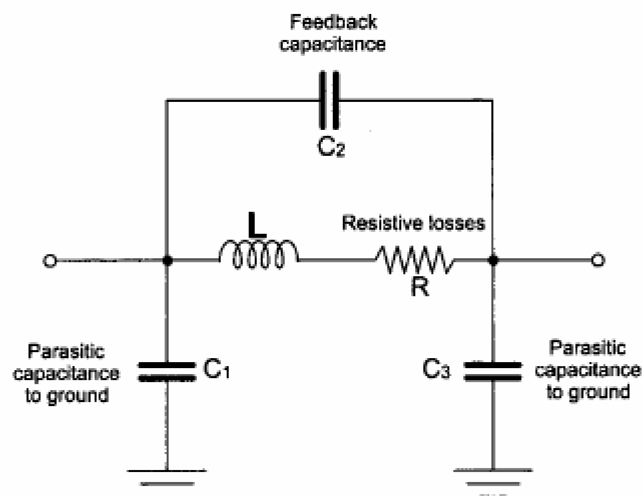


Figure 38- Inductor model

For silicon substrates it is necessary to add shunt resistance elements to ground, which are coupled via capacitors if there is an added dielectric layer or oxide. Direct calculations of prime inductance can be made using the classical formulas given by Grover [3], with appropriate correction factors. It

is more common, however, for the equivalent circuit data to be found for a number of different inductor geometries, and to use empirical curve-fitted expressions for the values of the equivalent circuit elements. Any equivalent circuit using lumped elements can only accurately model the component when its dimensions are less than approximately λ/20. In practice, such lumped equivalent circuits are accurate to frequencies above the self-resonant frequency of the inductor, beyond which one could argue that the inductor is no longer useful anyway. However, in a circuit such a mixer, an inductor may be used for IF matching at 1 GHz in a circuit which has a 20 GHz RF input. In this case it is essential that the inductor is modelled properly at both the IF and RF frequency. In order to achieve an accurate model at such high frequencies the distributed nature of the inductor must be considered. In essence, this means treating the inductor is an interconnection of multiple coupled microstrip lines. For a square inductor it surprisingly effective to consider each side of the inductor as a separate coupled-line problem, with special treatment for the corners. A number of quite rigorous approaches have been developed for the modelling of spiral inductors. The coupling between adjacent spiral inductor has received far lees theoretical treatment, but some very useful practical investigations have been made. In recent years, numerical electromagnetic field solutions have been reported as well. This type of modelling can be used in conjunction with measured inductor data to develop inductor models which are accurate and computationally efficient for incorporation into standard CAD packages. This approach allows more flexibility to the designer compared with the approach where a large number of standard inductor geometries are characterised by the foundry. Nevertheless, the foundry's own models should always be used in preference to those from any CAD package, unless that particular foundry has evaluated the alternative models for its own inductors.

Via-holes and grounding

The through-substrate via-hole, shown schematically in Figure 39-a is an essential component for all but the most simple microwave circuits as it enables a low inductance ground to be established within the circuit. Without via-holes, devices would have to be placed very near the edge of the chip and grounded with either wrap-around metallisation or bond-wires, as illustrated in Figure 39-b and c. These methods of grounding impose severe restrictions on the topology of the circuit. However, even when via-holes are used there are several pitfalls regarding grounding to be avoided at the design layout stage. The most important principle to remember is that a mass of metal can never be assumed to be a ground bar. The grounding connection must always be considered as a transmission line element in its own right and modelled accordingly. Figure 40 show a classic example of a circuit layout based on the ground bar principle that may not work in practive. In Figure 40-a the FET of a two stage amplifier share a common grounding pad, which is grounded with a via-hole. As the equivalent circuit shows, this grounding method provides a direct feedback path that is almost certain to cause instability. The solution in this case is to provide each FET whit its own ground pad, as show in Figure 40-b. The source inductance is not removed, but there is no longer a feedback path. A second example of inappropriate use of ground bar is shown in Figure 41. The ground bar is used as a neat common ground for the DC bias decoupling capacitors in a two-stage amplifier. However, the distributed nature of the ground bar means that none of the decoupling capacitors are truly grounded, and so a dangerous feedback path is again created. In many cases, when two components share a single via-hole ground, even the inductance of the via-hole itself is enough to allow the two components to interact undesirably. This can normally be detected by proper modelling. For MMIC design in the millimetre-wave range, even more subtle grounding effects have become apparent, such as the mutual coupling between via-holes placed in close proximity. In this situation, the inductance is much higher than expected as the current crowds together down one side of each via-hole. Hence, because grounding is so critical in millimetre-wave circuits, even the through-substrate via-hole ground may not offer low enough inductance. CPW

offers a very attractive solution to this problem, and is becoming increasing popular as the design tools improve.

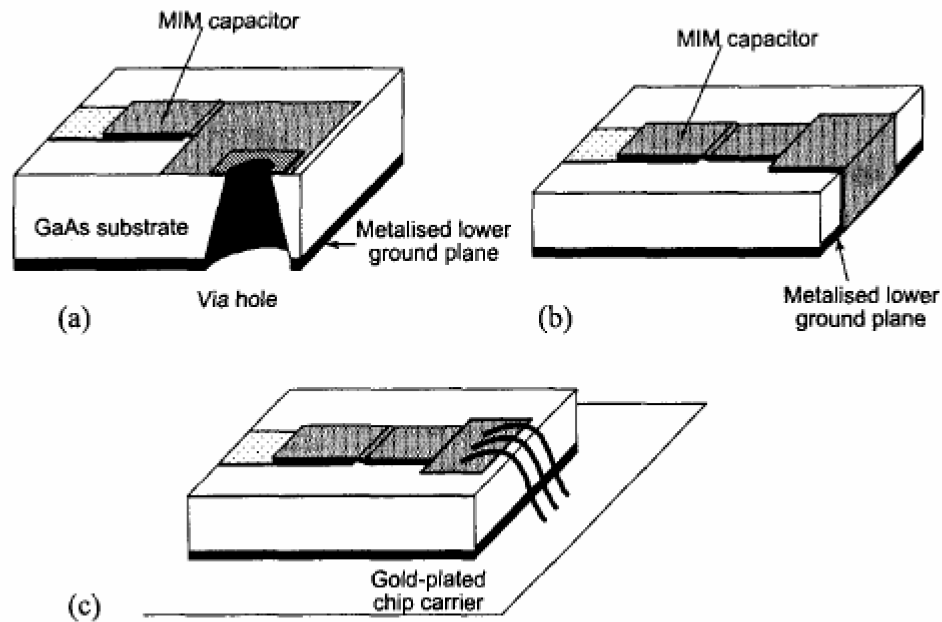


Figure 39

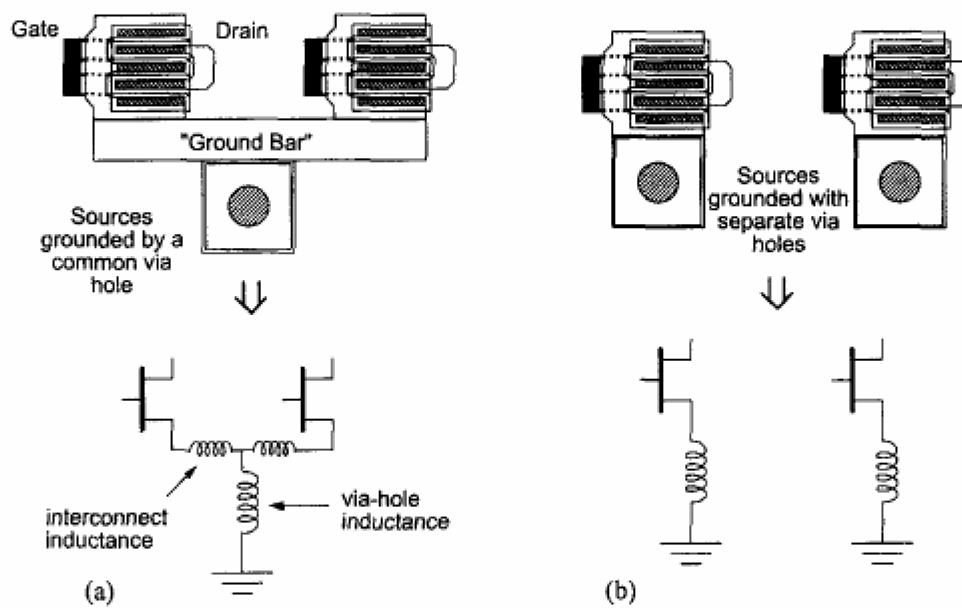


Figure 40

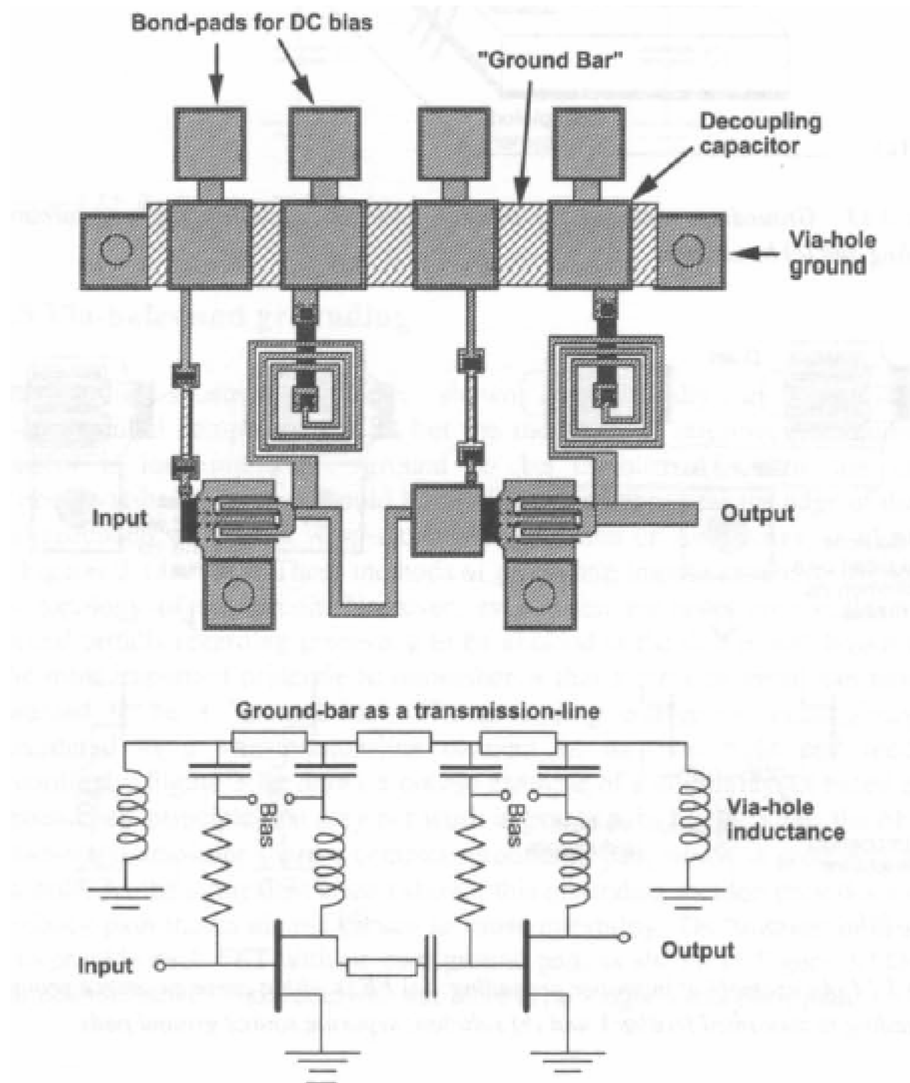


Figure 41

Microstrip components.

For many years microwave CAD packages have offered a wide range of models for microstrip lines and discontinuities. However, most of these have been developed principally for MIC circuits on substrates such as alumina and Duroid. MMICs have much smaller features than MICs, and it is important to ensure that the standard models are not used outside their range of validity.
The models are further limited because MMICs often have extra dielectric passivation and space layers. Many CAD packages have addressed these limitations and now provide models that are suitable for MMIC microstrip elements. However, the problem remains that on MMICs the components are packed more tightly and the coupling between elements must be accounted for. As a result it may necessary to employ field-based simulation for some parts of circuit layout. In this paragraph I would be some example of application on MIC circuit of microstrip elements to design some circuits.

The first example is the design of a 1.5GHz centre bandwidth frequency hybrid circuit for power division performed through ADS and the electromagnetic simulators
For the design of couplers circuit there are a lot of references in literature. The general structure of a circular coupler it consist in realize a square structures formed by λ/4 sides of same and different width.

In ADS it is possible to select the t-lines microstrip palette choose the substrate modeller and put them in the schematic. Save the substrate as FR-4.

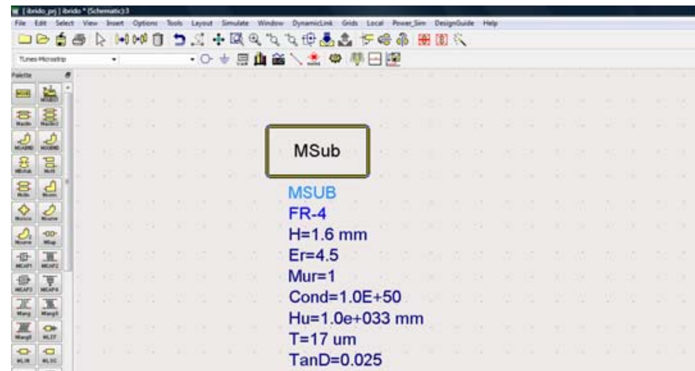


Figure 42

Now you can use linecalc tool to compute the lengths and width for your circuit.
Then press tools and start linecalc. In this tool you can define the same parameters of substrate and save them.

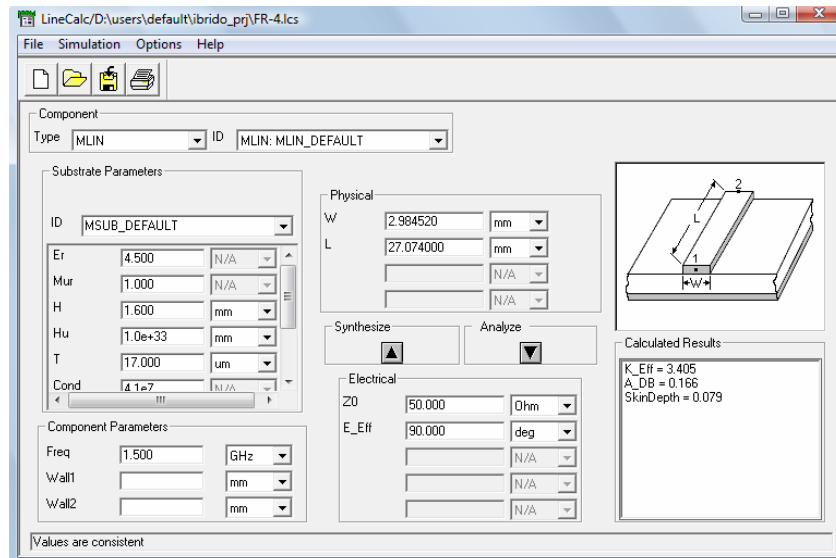


Figure 43

After the modelled was done you can insert the desired value of characteristic impedance Z_0 and electric equivalent wavelength E-Eff then press the button sinthesize and width and length are calculate automatically. Repeat the design steps with linecalc but in order to generate a transmission line with characteristic impedance of

$$\frac{Z_0}{\sqrt{2}} = \frac{50}{\sqrt{2}} \approx 35.35\Omega$$

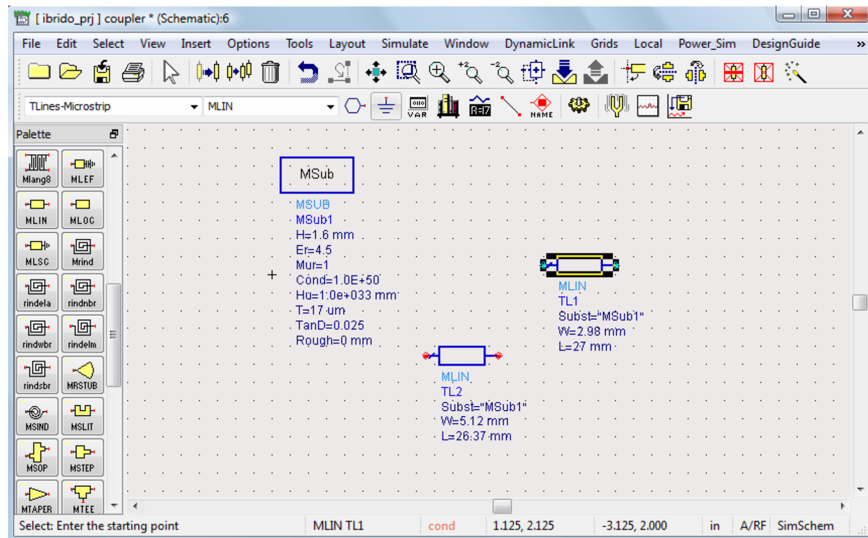


Figure 39

Now you can generate the layout and work at layout system to achieve the complete design. Press Layout then generate update and click on ok, the following window appears.

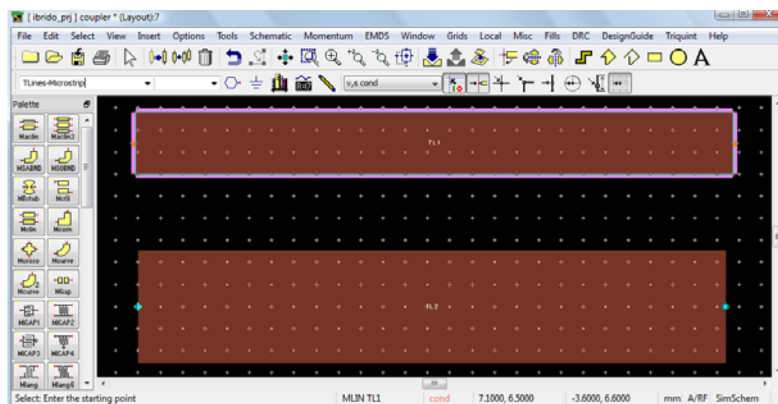


Figure 45

Now you can add, remove or modify components to achieve the circuit.
Add the MTE component to match and complete the design selecting for terminals one and three the same dimensions of the line parts. At the end you can click on schematic to update the schematic from layout and verify the connections. After you can add the ports by schematic and recreate the layout. These design steps are showed in the figures below:

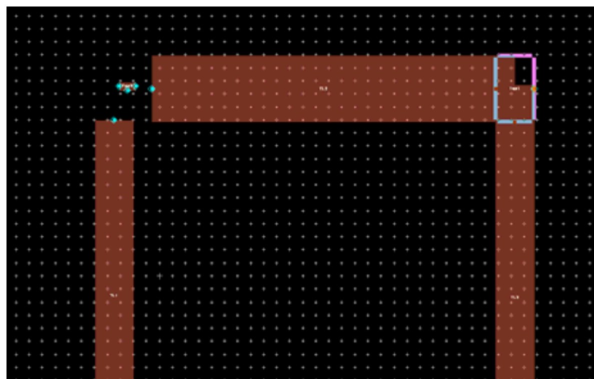


Figure 40

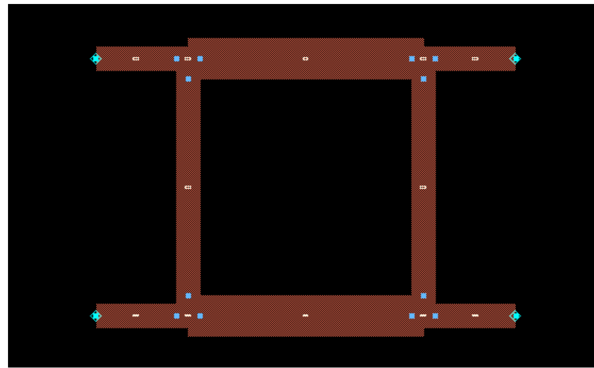


Figure 41

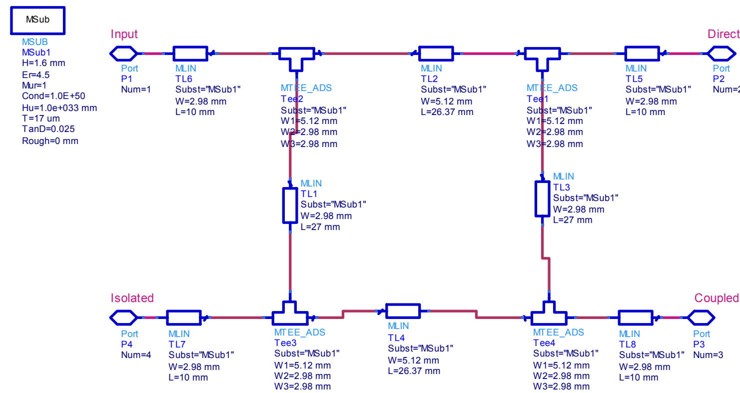


Figure 42

Then the last figure that show the layout view obtained by the schematic

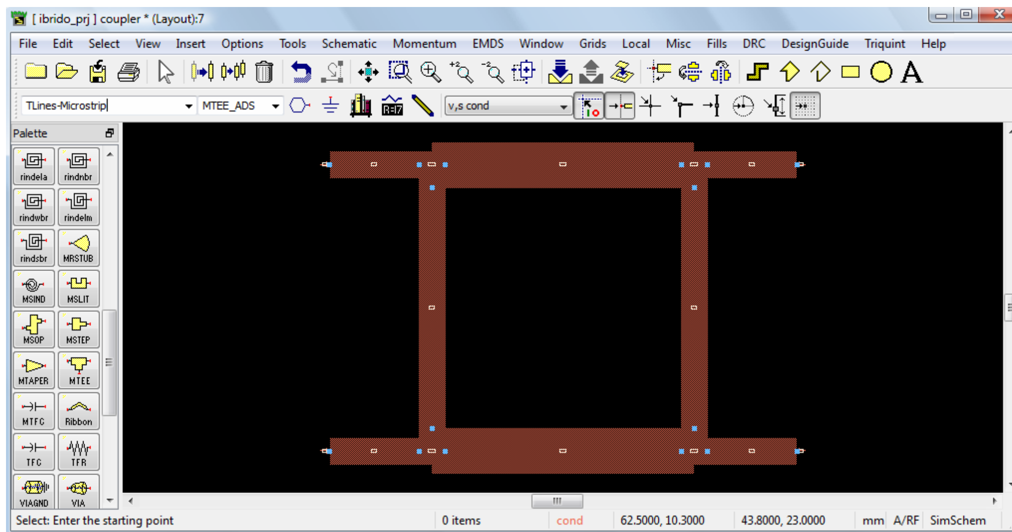


Figure 43

For simulations you can at first verify the S-parameters simulation or put a kind of time source at the input port and see how the signal behaves at others ports.
To simulate the circuit for example by using momentum you must import the substrate from the schematic as show in the figure below

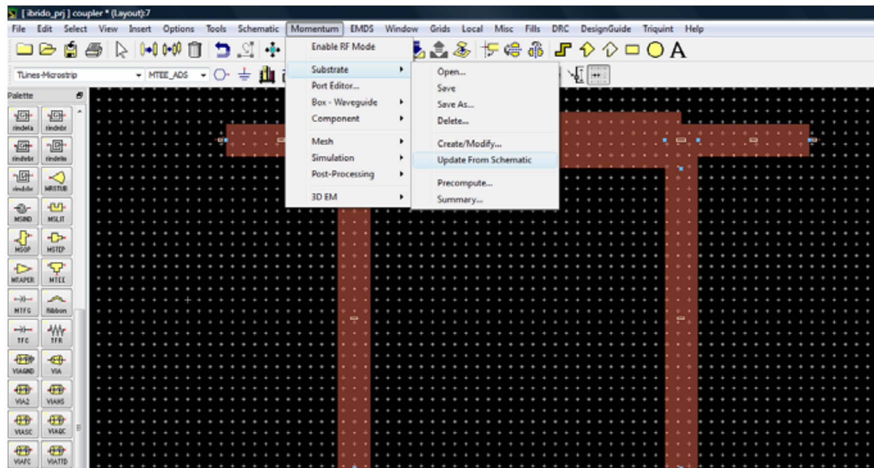


Figure 44

You can also to save the data for the 3D view. And click on simulation voice of the menu to define the simulation parameters. At least click on simulate

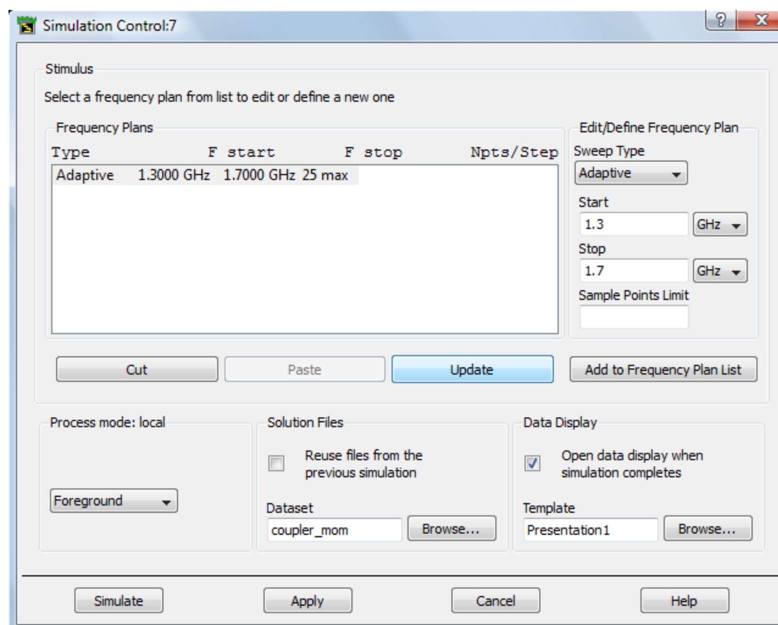


Figure 51

After the simulation you can visualize the results and compare to of s parameters simulations.

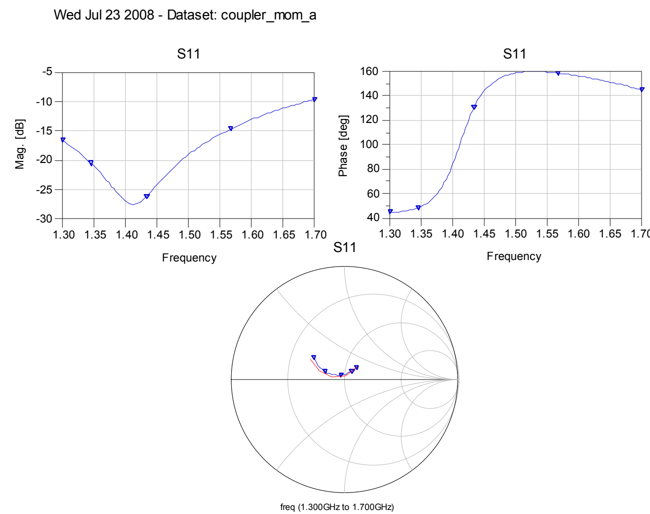


Figure 52

The red line is the trade of S11 parameterized in frequency of S-parameters simulations whereas the blue line represent the results of Momentum simulation. The results can be visualized using the post processing visualization by means of which it is possible to simulate the current view for all ports

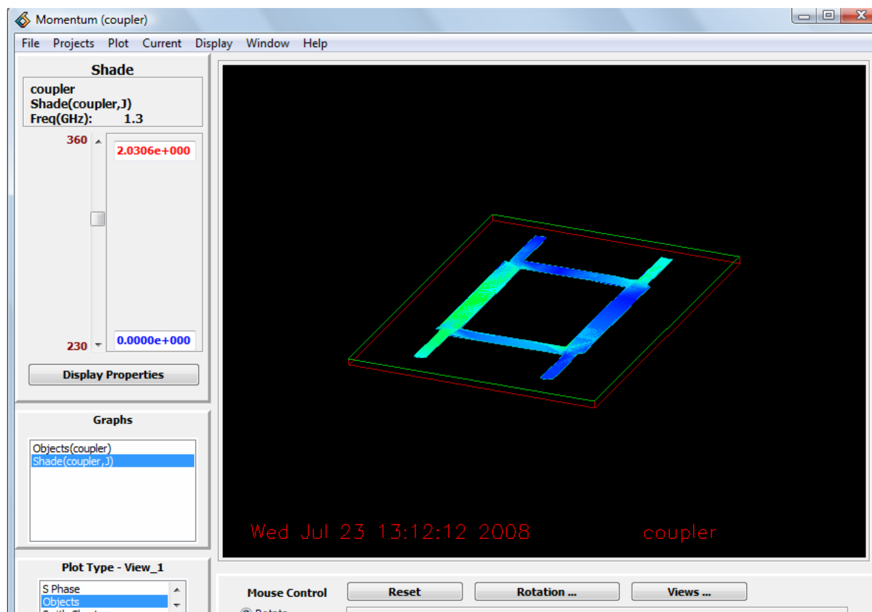


Figure 53

At least we will use EMDS to achieve a 3D electromagnetic simulation of the circuit.

The second example is about the realization of low-pass filter with cut frequency of 4GHz.

Filter constraints

Z0=50Ω ; N=3; f=4GHz equal ripple a 3 dB

From equiripple table in [1] and N=3 there are:

g_1=3.3487
g_2=0.7117
g_3=3.3487

The first design step is to model the filter by using a passive lumped component. The values for both possible configuration can be computed from following relationships:

$$L1 = \frac{g_1 Z_0}{2 \cdot \pi \cdot f} = \frac{3.3487 \cdot 50}{2 \cdot \pi \cdot 4^9} = 6.66nH \ ;$$

$$C2 = \frac{g_2}{2 \cdot \pi \cdot fZ_0} = \frac{0.7117}{2 \cdot \pi \cdot 4^9 \cdot 50} = 566 fH \ ;$$

$$L3 = \frac{g_3 Z_0}{2 \cdot \pi \cdot f} = \frac{3.3487 \cdot 50}{2 \cdot \pi \cdot 4^9} = 6.66nH$$

Or

$$C1 = \frac{g_1}{2 \cdot \pi \cdot fZ_0} = \frac{3.3487}{2 \cdot \pi \cdot 4^9 \cdot 50} = 2.66 pH$$

$$L2 = \frac{g_2 Z_0}{2 \cdot \pi \cdot f} = \frac{0.7117}{2 \cdot \pi \cdot 4^9 \cdot 50} = 1.4159nH$$

$$C3 = \frac{g_3}{2 \cdot \pi \cdot fZ_0} = \frac{3.3487}{2 \cdot \pi \cdot 4^9 \cdot 50} = 2.66 pH$$

The frequency transmission response (S_{21}) can be simulated through a S-parameter Simulation tool as showed in the following figure:

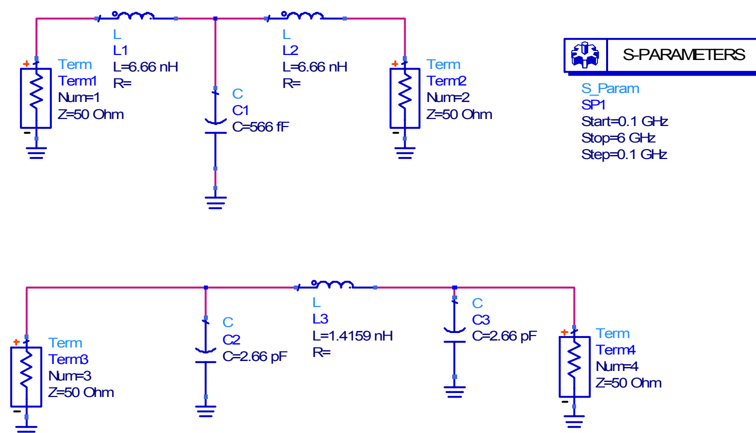


Figure 45

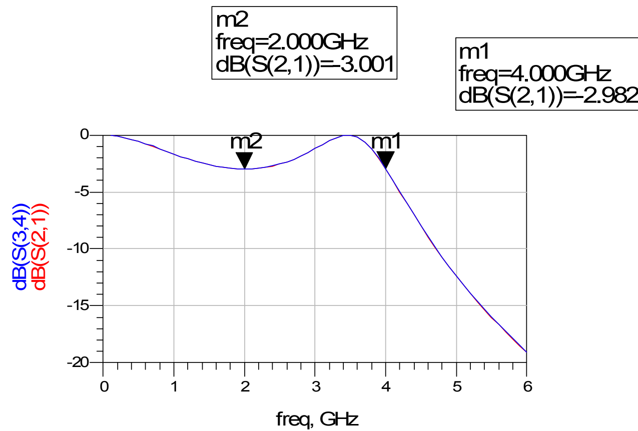


Figure 55

In order too design the real filter we must in transmission lines components we must compute the equivalences between lines capacitors and insuctors as well.

In order to compute the size of transmission lines we can consider the inductances and capacitances values to find the Z_0 and thus the W for each transmission line.

For example; for the first circuiti t is possible to write for the inductor

$$L = g_1 = 3.487 \rightarrow Z_O = L = 3.3487$$

Instead for the capacitor

$$C = g_1 = 0.7117 \rightarrow Z_O = \frac{1}{C} = 1.405$$

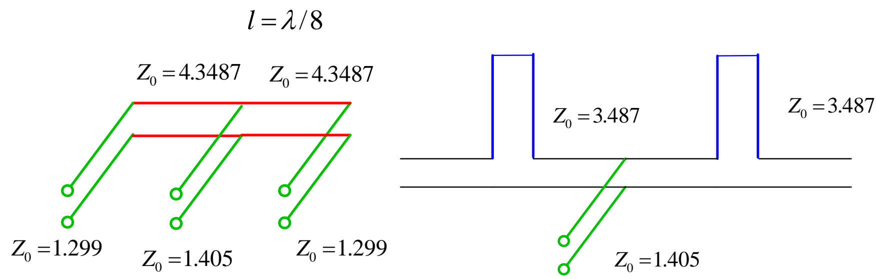
In the second case we have

$$L = g_1 = 0.7117 \rightarrow Z_O = L = 0.7117$$

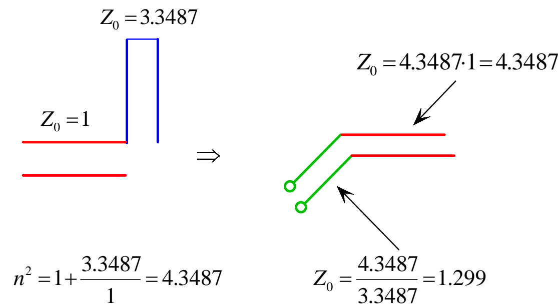
$$C = g_1 = 3.487 \rightarrow Z_O = \frac{1}{C} = 0.286$$

First possible circuit L-C-L.

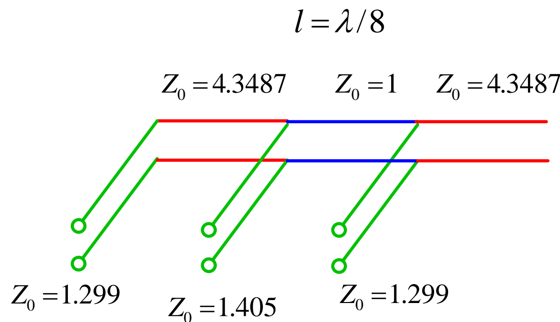
Substituting at lumped elements the circuit in form of transmission lines the two following schematics are depicted:



Now it is possible to apply the Kuroda identity to substitute the short circuit series stub with a parallel open circuit stub and a series transmission line. To do this a unitary impedance characteristic lines is add in series in order to not alter the behaviour of the circuit.



Therefore the equivalent circuit in transmission lines term is



The electrical length is the same for each component and this solution is also called conform.

The corresponding ideal circuit is found imposing the electrical length equal to l=λ/8 at the design frequency therefore we have electrical length of 45° while le weights will computed through the characteristic impedances of components after a Z_0=50Ω multiplication.

The ideal lines circuit is showed below

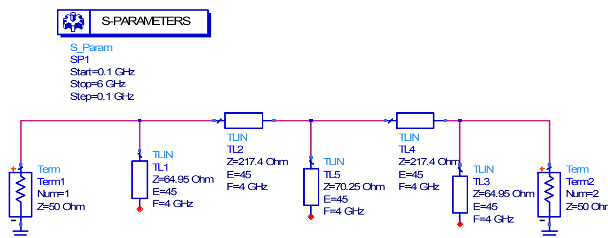


Figure 56

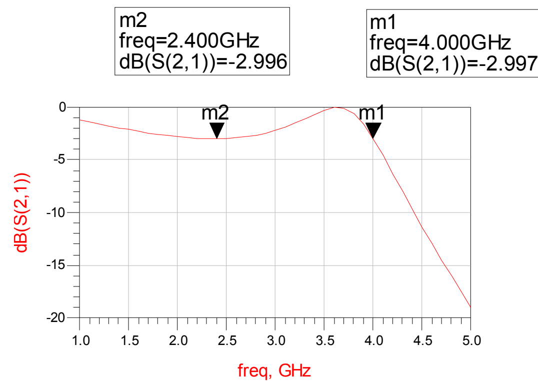


Figure 57

Now it is necessary to implement the circuit with real elements.

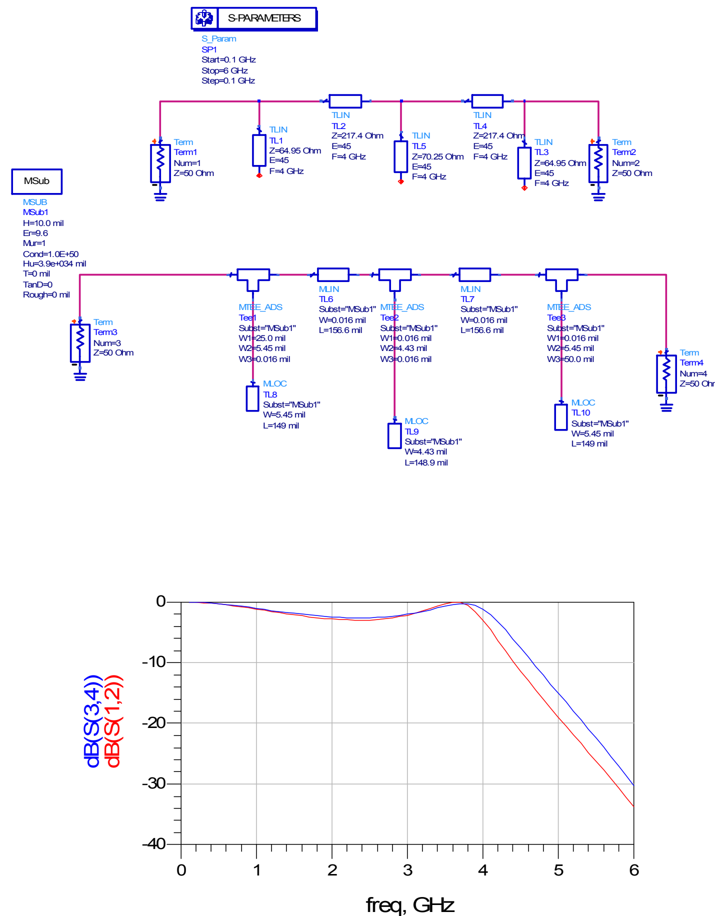


Figure 58

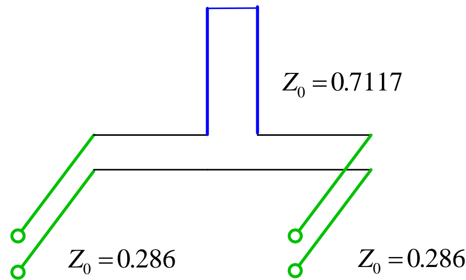
Although the simulation results are good it is difficult, for this foundry, to be able to have a value of impedance so high. In fact in order to achieve the outline corresponding to the transformation of Kuroda should make extremely thin lines.

Second possible circuit -C-L-C.

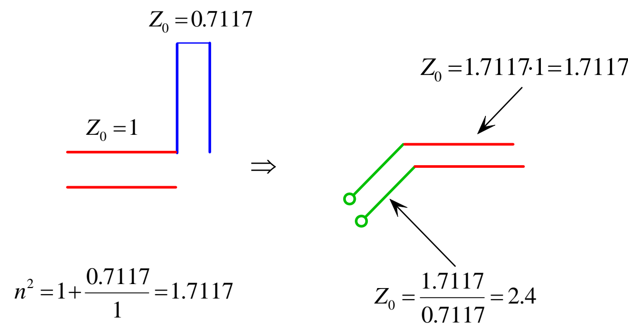
$$L = g_1 = 0.7117 \rightarrow Z_O = L = 0.7117$$

$$C = g_1 = 3.487 \rightarrow Z_O = \frac{1}{C} = 0.286$$

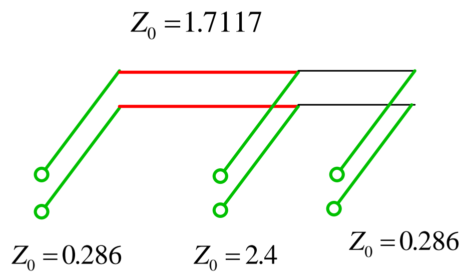
The equivalent circuit in terms of transmission lines for the second case should be the following:



By applying the Kuroda transformation we have:



Therefore the equivalent circuit becames



And his implementation in ideal lines:

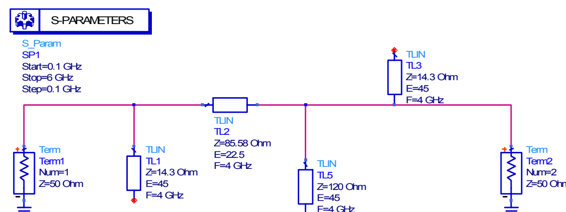
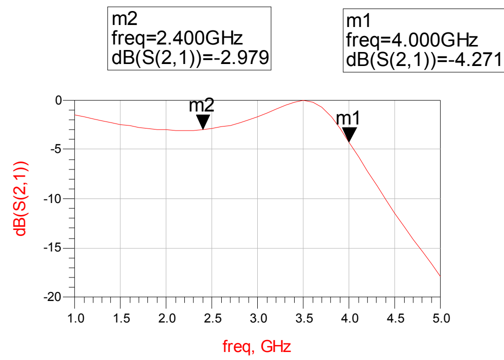
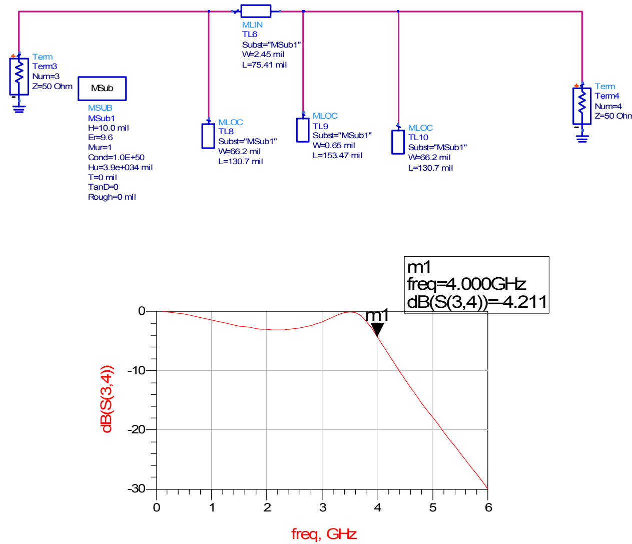


Figure 59

Note that the length of central lines has been reduced by a factor 2, for the presence of double stub.



By using the real elements we have:



A simple layout for the circuit is showed in the figure below.

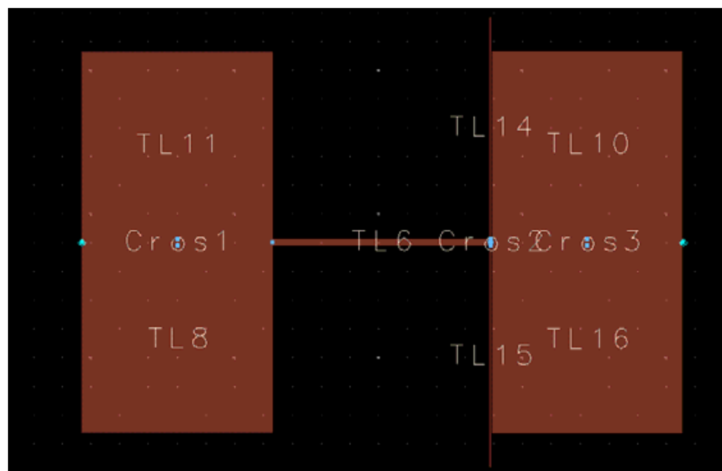


Figure 60-layout of the Low pass filter.

After the layout operation ADS allows to extract easily the equivalent circuit and re-simulate them.

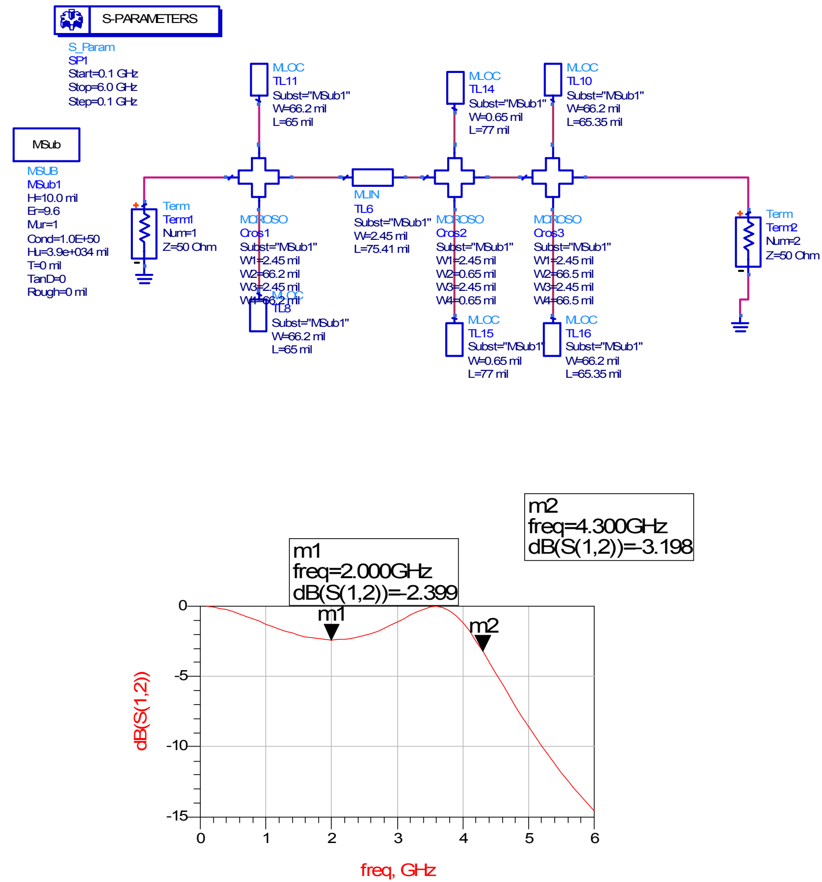


Figure 61- post layout schematic and simulation of the circuit.

From this point it is possible through the classical try and error procedure to optimize the design constraints by changing for example the width of the second line and re-simulating the design since the desired performances are obtained.

2.7-The GaAs and CMOS processes review.

In order to complete our brief description of GaAs technology we can give a review of the processes.
The starting wafers must be selected based on the specific process requirements. Low-noise processes require a different set of starting material characteristics than power processes, and each manufacturer has a defined set of wafer characteristics based on the selected process. A typical ;ESFET process flow is showed in Figures 62 and 63 with optional steps shown for clarity and completeness of flow.

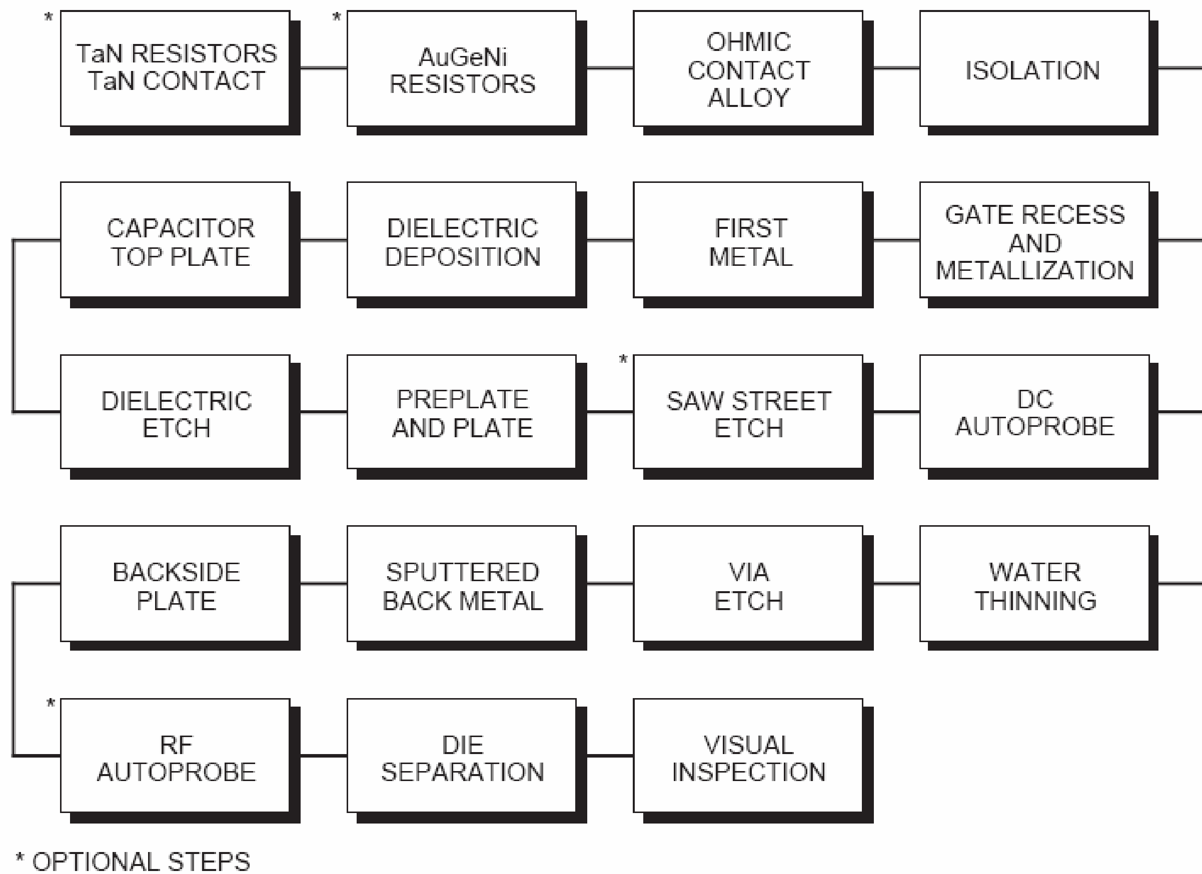


Figure 62- Basic sequence of process steps

The first step normally involves the fabrication of the thin-film resistors. The AuGeNi resistor metal is evaporated, and the TaN resistor metal is sputtered and followed by a TaN contact-metal evaporation step. AuGeNi is normally used for designing low-values resistors, while TaN is used for medium-values resistors. An ohmic-contact deposition step normally follows with an alloy steo, which results in a low-resistance ohmic contact to the active GaAs and also serves to stabilize the metal-film resistors.

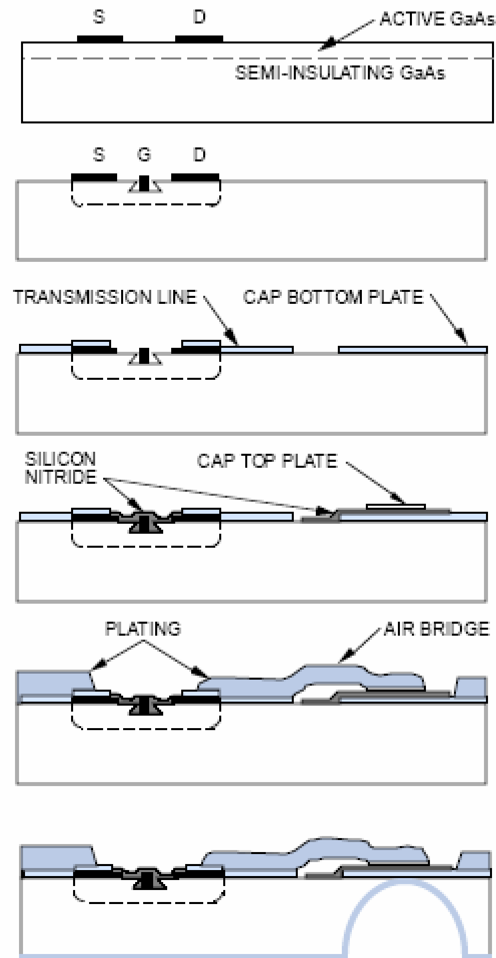


Figure 63- Basic process steps for MESFETs.

An ion implant such as boron is used to deactivate the conduction GaAs layer and form isolation patterns where desired. A direct-write e-beam can then be used to pattern the gate and gate recess in the active areas. The first metal layer is normally an evaporated metal layer, which contacts the semi-insulating GaAs and forms the first-level interconnect. Dielectric deposition of silicon nitride is used to protect circuit elements and provide a dielectric for capacitors. Capacitor top-plate metal is then deposited on top of the silicon-nitride dielectric. A pattern step is implemented to open the contact and define the bottom plate of the capacitor. Figure 64 shows a cross section of thin-film resistors and ohmic contacts on GaAs.

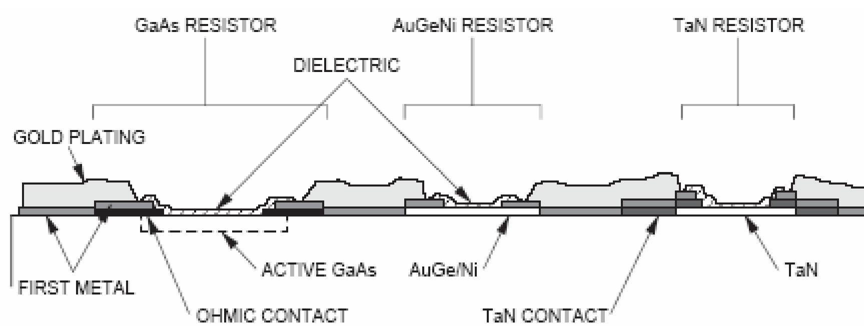


Figure 64

Plating is used to deposit thick layers of gold to construct air bridges, low-loss transmission lines, high-current-carrying lines, bonds pads, resistor, contacts, and evaporated metal step coverage. Two resist patterns are required to define the plated gold layer; the preplate and plate patterns work

together to define the electroplated gold-metal level. The preplate layer defines the area where the plating contacts underlying metal. The preplate level defines the horizontal extent of the plated region. The preplate resist pattern is deposited derectly on the front side of the wafer. Openings in the preplate resist are exposed and developed to define those areas where the plated metal will contact the underlying metal layer. The underlying metal is usually first metal, but plated metal can also contact capacitor top plates and other conducting layers. After the preplate pattern is formed, a thin layer-to-metal, which serves to carry the electroplating current. On top of the preplate metal, a second resist pattern is formed to define the horizontal extent of any plated geometry, whether it is part of an air bridge or in contact with underlying metal. The preplate metal is removed, along with photoresist, in all unplated areas. The preplate metal remains underneath all plated areas. To allow for electrical connections between the frontside metal and the backside ground plane, via holes are formed and plated with gold. The size of the via hole depends on the substrate thickness; a circular pattern having 50-to 60 µm diameter is normally used on 100 µm thick substrates. The plated gold layer also serves as the contact layer for die attach and a thermal path to a substrate. In processing, the via diameter at the frontside contact may vary from 12 to 160 µm thivk substrates. At backside, the via may be 2 to 3 times larger than the frontside pattern. The following figure shows a via hole and the process-dependent parameters.

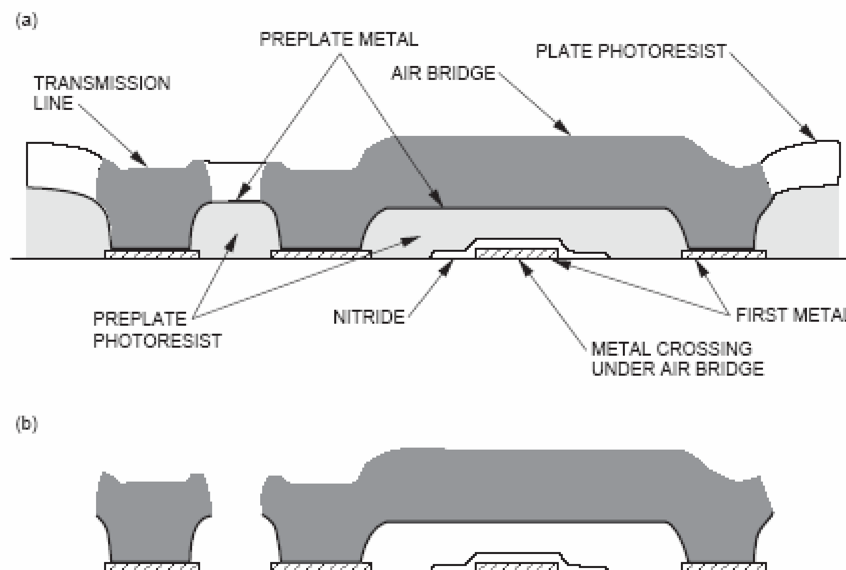


Figure 3-32. The air-bridge process: (a) with plate and preplate photoresist patterns and (b) after resist is removed.

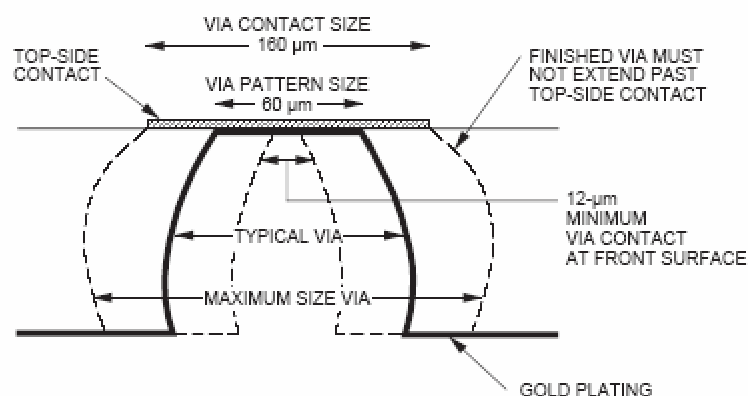


Figure 64 via hole eith process-dependent parameters.

Reactive Ion Etch (RIE) is normally used to open the backside via holes. A layer of sputtered metal is then deposited over the entire backside of the wafer. Gold plating, approximately 6 µm thick, is

then added to the sputtered metal for die-attach capability. The last step in the process is to physically separate the devices on the wafer. This is done by either scribing and breacking apart the devices or by using areas called saw streets, which are stripts void of plated metal and are outer boundaries of the individual MMICs. Final visual inspection is normally used in conjunction with dc probe data to select acceptable device.

The typical HEMT/PHEMT PROCESS FLOW

The starting materials for HEMT-based devices require specific and stringent parameter control. Device manufacturers normally specify the applicable parameters that affect their process and are suitable to the processing flow. After the usual wafer cleaning and inspection, an epitaxial layer must be grown to provide the required material characteristics necessary for HEMT and PHEMT devices. The process starts with a GaAs buffer layer epitaxially grown to isolate defects drom the substrate and provide a smooth foundation for further growth of the active layer of the transistor. Although differences exist in the fabrication of HEMT and PHEMT devices, the general approach and processing flow remain essentially the same. The following brief description of the general processing flow is depicted in Figure 65.

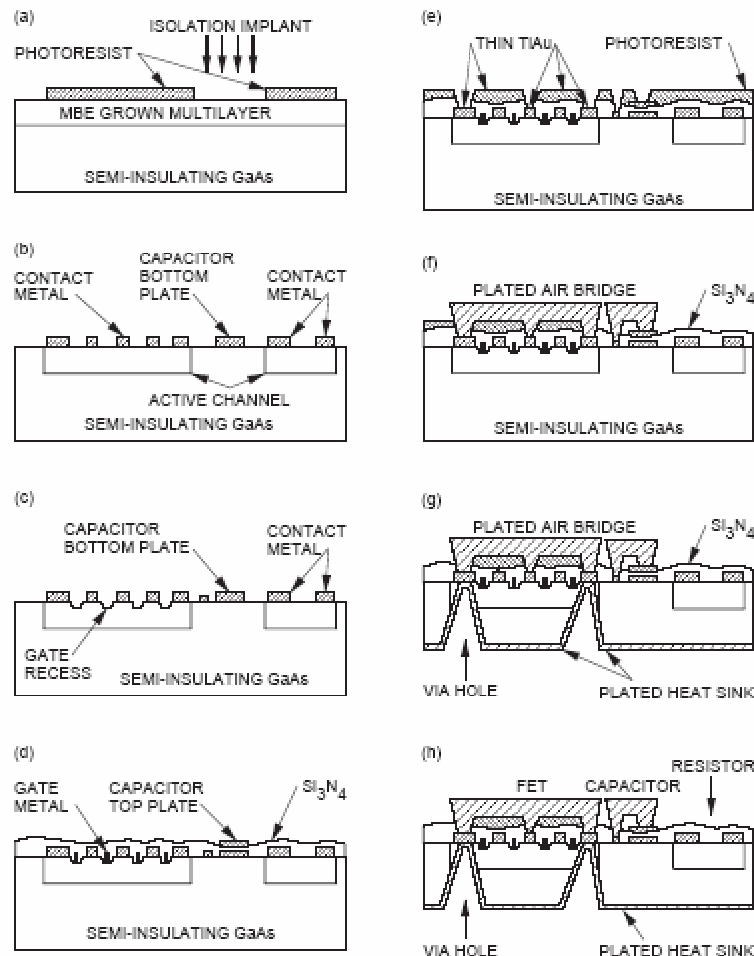


Figure 65 Typical HEMT/PHEMT process flow: (a) active canne definition and isolation implant, (b) ohmic-metal formation, (c) gate-recess formation, (d) gate-metal formation and nitride deposition, (e) source and contact each, (f) air-bridge formation, (g) via-hole formation and backside processing, (h) completed typical MMIC structure.

The typical HBT Process Flow.

The device fabrication sequence basically consists of etching steps to reveal the various layers in the structure and fabricating electrical contacts to each layer. Finally devices are isolated and interconnections are made within each device as well as between devices. The following general steps as shown in Figures 66 and 67.

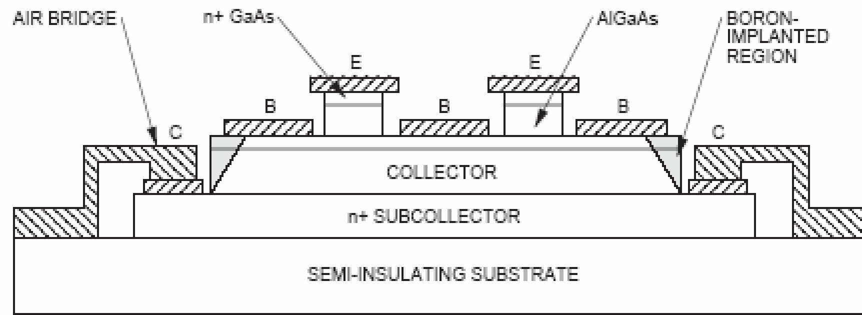


Figure 3-36. Self-aligned HBT cross-sectional view. (Courtesy of Artech House.)

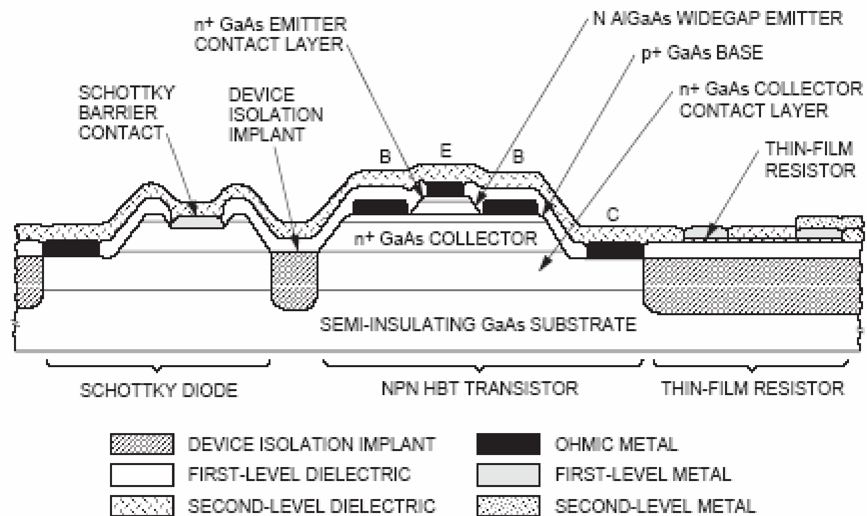
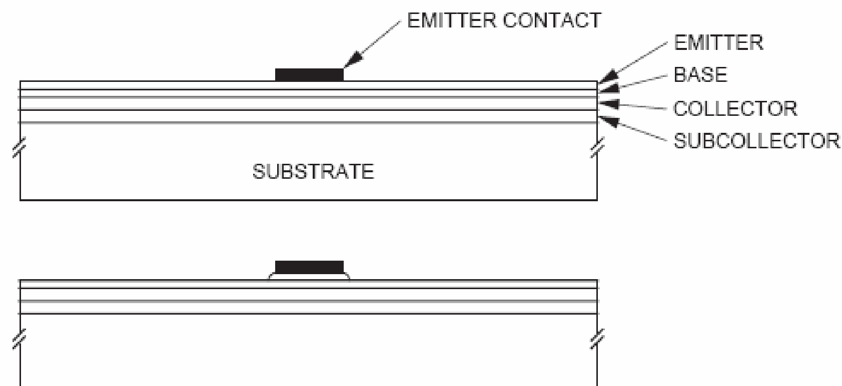


Figure 66 Schematic cross section of the self-aligned HBT IC structure with integrated NPN transistor, Schottky diode, thin film resistor, and metal-insulator (SiN)-metal capacitor(not show). (Courtesy of Artech House.)



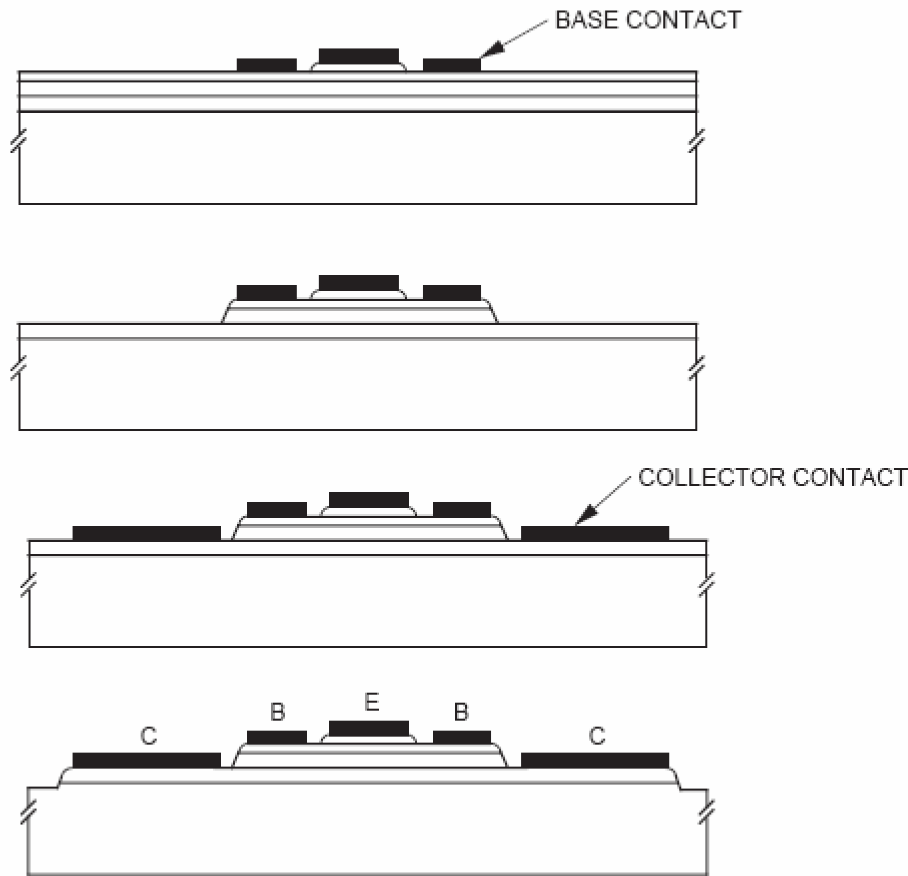


Figure 67 A typical HBT process flow.

2.8 The CMOS technology in RF design

The CMOS technology is used in the design of microwave and RF circuit. Today this technology is very important and its size is small compare its the frequency, therefore the design by using CMOS is as lumped components. In this paragraph we will describe the realization of capacitors, inductors, and resistors, transistors and a review of CMOS technology process as well.
The CMOS technology is very known and for this reason the following description is more technical compared with previously done on GaAs-based devices. Commercial software for microwave as ADS does not have special foundries to simulate devices at layout level but only to describe components which equations derive by experimental simulations. Our university does not have any license or design kit in CMOS for ADS, therefore the examples done in this thesis by using the CMOS will be done only by using ideal devices. We will describe the general process and technical issue for a generic CMOS process.

The CMOS process

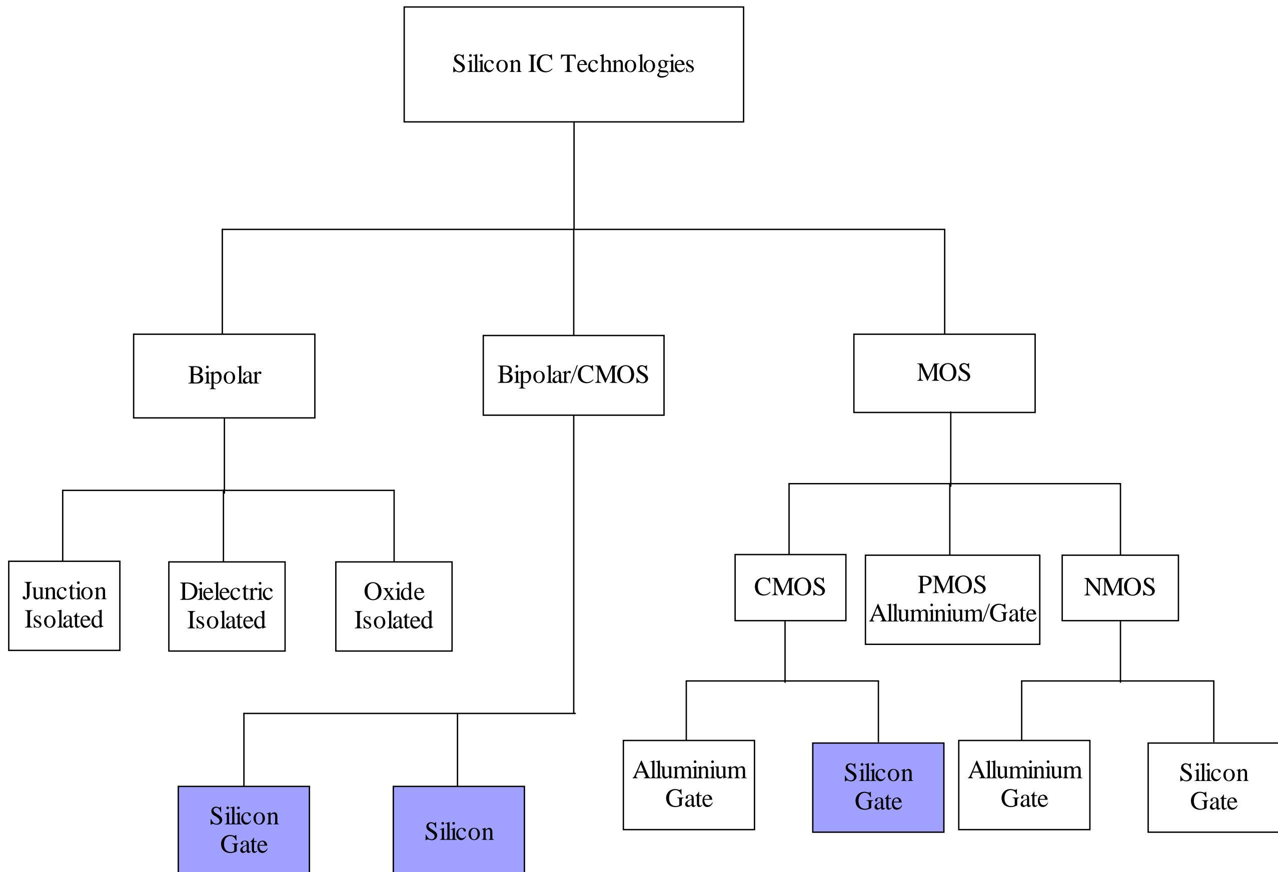


Figura 67 Classification of different Silicon technologies

As previously told silicon technologies presents a lot of advantages compare to GaAs technologies about cost integration capabilities. Between the Silicon technologies the most popular and used the CMOS technology. The BJT is also used but for different application. The comparison between BJT and CMOS is summarised in the following Table:

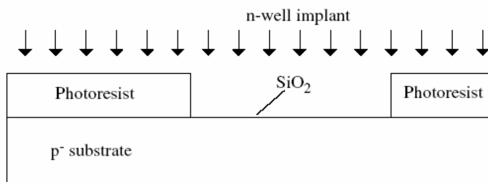
Feature	BJT	MOSFET
Cutoff Frequency f_T	100GHz	50GHz (0.25µm)
Noise (Thermal about the same	Less 1/f	More 1/f
DC Range of operation	9 decades of exponential current versus v_{BE}	2-3 decades of square law behavior
Small signal Output Resistance	Slightly larger	Smaller for short channel
Switch implementation	Poor	Good
Capacito Implementation	Voltage dependent	Reasonably good

Therefore:
Almnost every comparison favors the BJT, however a similar comparison from a digital viewpoint would come un on the side of CMOS. A large-volume technology will be driven by digital demands, CMOS is an obvious result as the technology of availability.

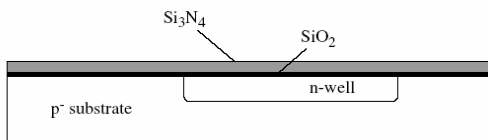
Moreover the potential for technology improvement for CMOS is greater than for BJT, and performances improvements in general increases decreasing the channel length. In addition a deep-n well that can be utilized to reduce substrate noise coupling and a MOS varactor that can serve in VCOs. Finally at least 6 levels of metal that can form many useful structures such as inductors, capacitors, and transmission lines.
The principals technological steps for CMOS and BICMOS technologies are showed in the following figures flow.

FOR CMOS:

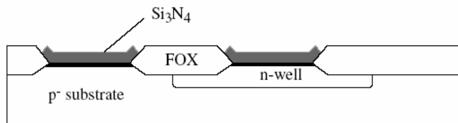
Step 1 - Implantation and diffusion of the n-wells



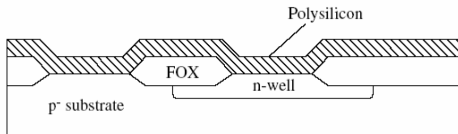
Step 2 - Growth of thin oxide and deposition of silicon nitride



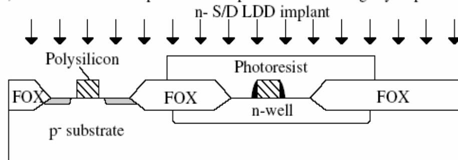
Step 5.) Growth of the thick field oxide (LOCOS - *loc*alized *o*xidation of *s*ilicon)



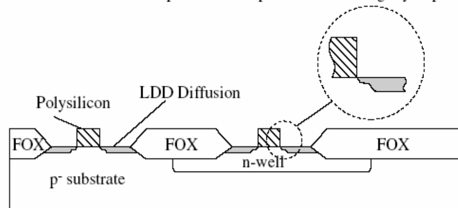
Step 6.) Growth of the gate thin oxide and deposition of polysilicon. The thresholds can be shifted by an implantation before the deposition of polysilicon.



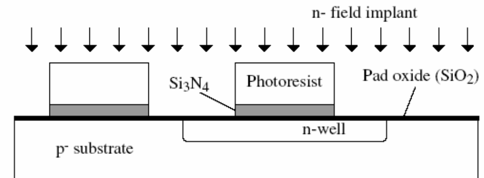
Step 9.) Remove sidewall spacers and implant the NMOS lightly doped source/drains



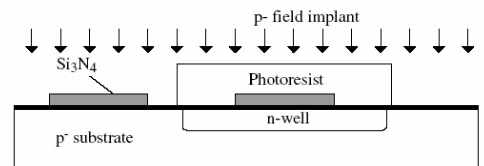
Step 10.) Implant the PMOS source/drains and contacts to the p⁻ substrate (not shown), remove the sidewall spacers and implant the PMOS lightly doped source/drains



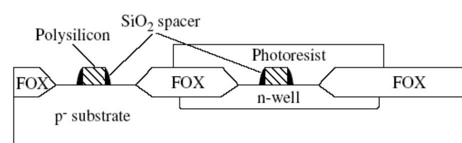
Step 3.) Implantation of the n-type field channel stop



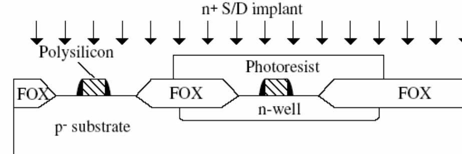
Step 4.) Implantation of the p-type field channel stop



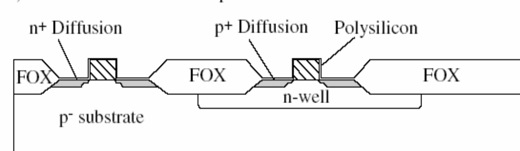
Step 7.) Removal of polysilicon and formation of the sidewall spacers



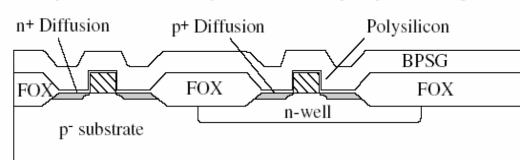
Step 8.) Implantation of NMOS source and drain and contact to n-well (not shown)



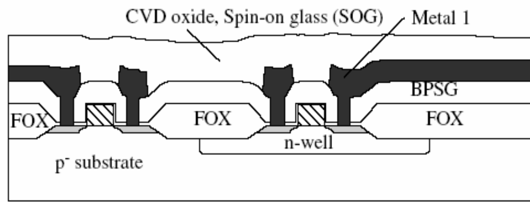
Step 11.) Anneal to activate the implanted ions



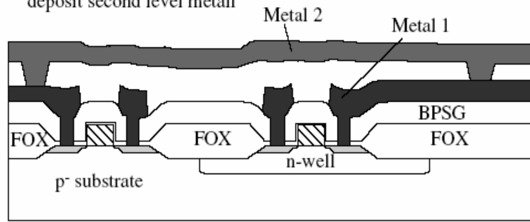
Step 12.) Deposit a thick oxide layer (BPSG - borophosphosilicate glass)



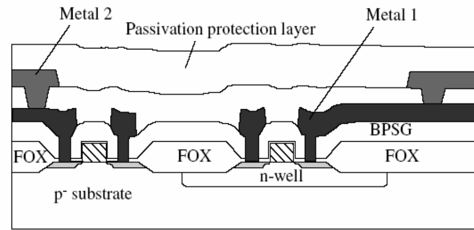
Step 13.) Open contacts, deposit first level metal and etch unwanted metal

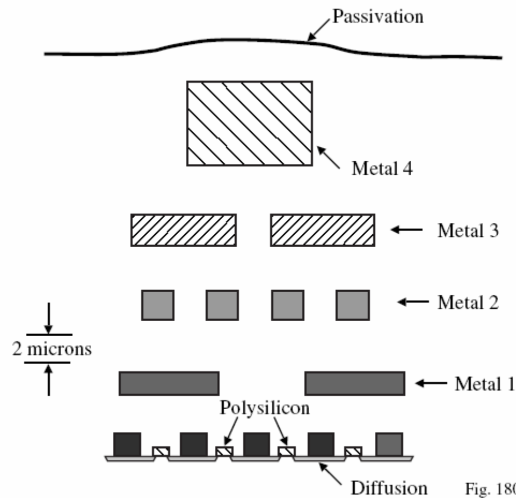


Step 14.) Deposit another interlayer dielectric (CVD SiO_2), open contacts, deposit second level metall



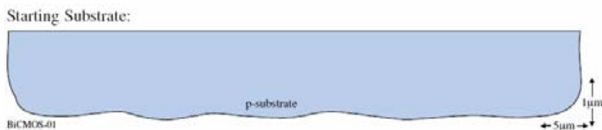
Step 15.) Etch unwanted metal and deposit a passivation layer and open over bonding pads



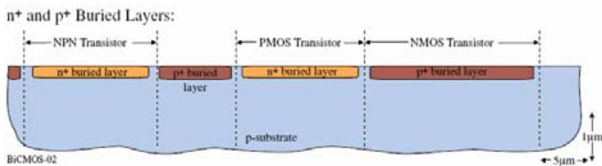


FOR BICMOS:

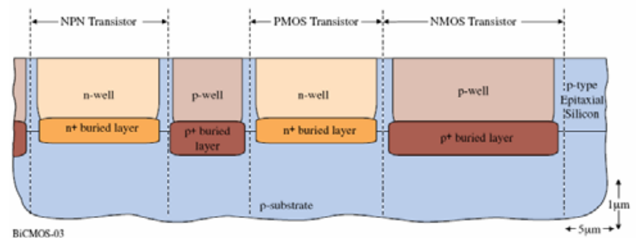
n+ and p+ Buried Layers



n+ and p+ Buried Layers:



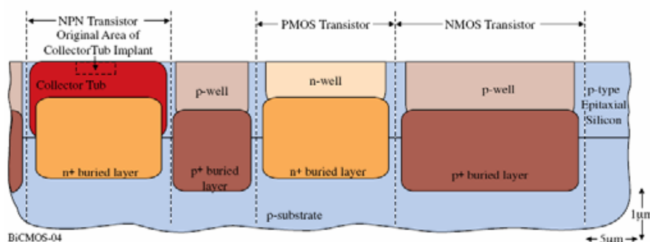
Epitaxial Growth



Comment:

- As the epi layer grows vertically, it assumes the doping level of the substrate beneath it.
- In addition, the high temperature of the epitaxial process causes the buried layers to diffuse upward and downward.

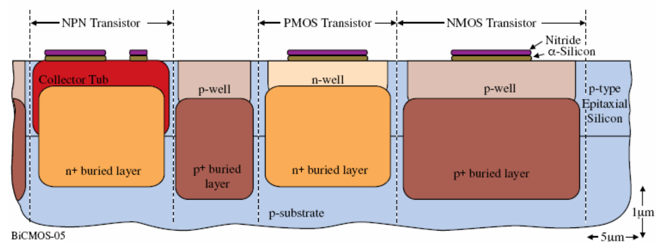
Collector Tub



Comment:

- The collector area is developed by an initial implant followed by a drive-in diffusion to form the collector tub.

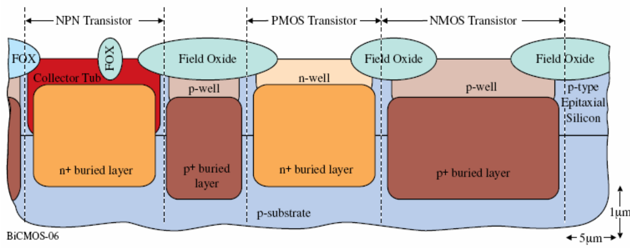
Active Area Definition



Comment:

- The silicon nitride is use to impede the growth of the thick oxide which allows contact to the substrate
- α-silicon is used for stress relief and to minimize the bird's beak encroachment

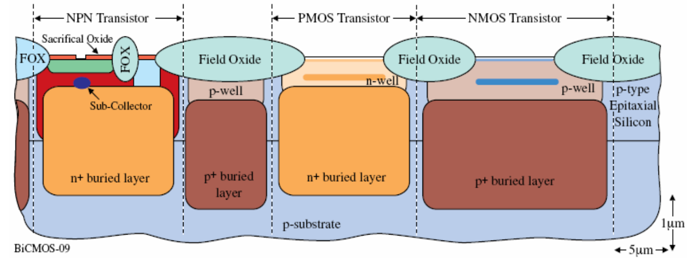
Field Oxide



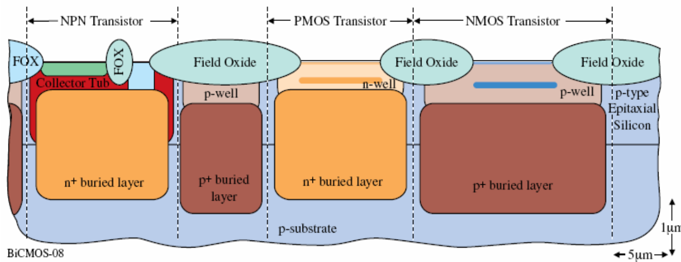
Comments:

- The field oxide is used to isolate surface structures (i.e. metal) from the substrate

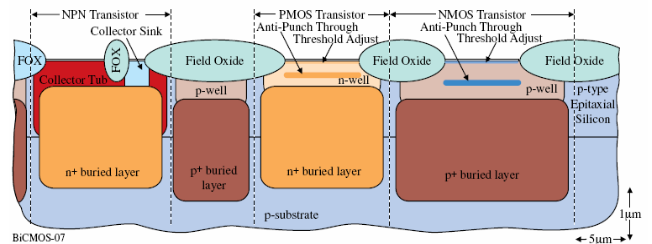
Definition of the Emitter Window and Sub-Collector Implant



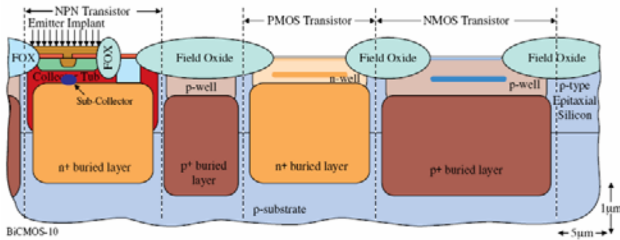
Base Definition



Collector Sink and n-Well and p-Well Definitions



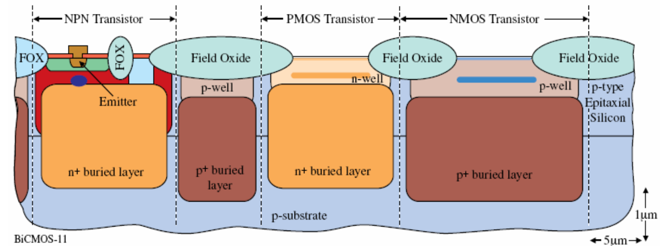
Emitter Implant



Comments:

- The polysilicon above the base is implanted with n-type carriers

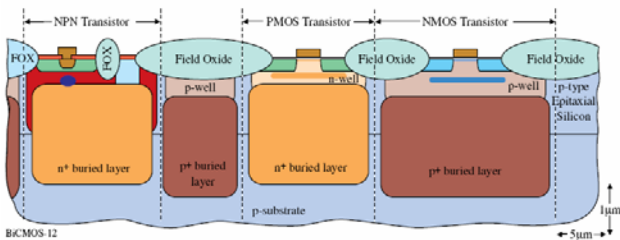
Emitter Diffusion



Comments:

- The polysilicon not over the emitter window is removed and the n-type carriers diffuse into the base forming the emitter

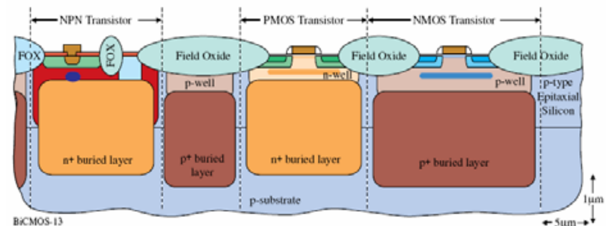
Formation of the MOS Gates and LD Drains/Sources



Comments:

- The surface of the region where the MOSFETs are to be built is cleared and a thin gate oxide is deposited with a polysilicon layer on top of the thin oxide
- The polysilicon is removed over the source and drain areas
- A light source/drain diffusion is done for the NMOS and PMOS (separately)

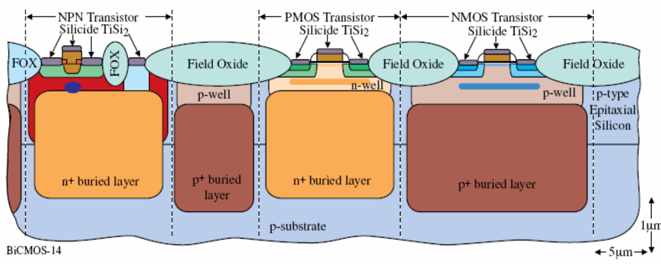
Heavily Doped Source/Drain



Comments:

- The sidewall spacers prevent the heavy source/drain doping from being near the channel of the MOSFET

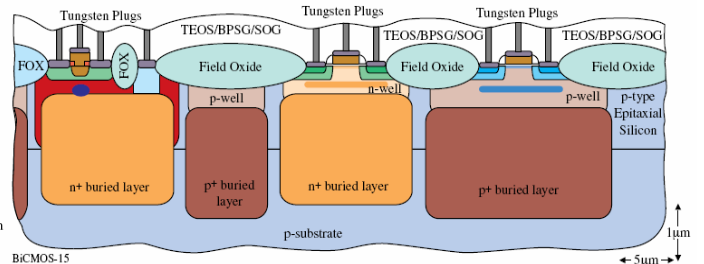
Siliciding



Comments:

- Siliciding is used to reduce the resistance of the polysilicon and to provide ohmic contacts to the base, emitter, collector, sources and drains

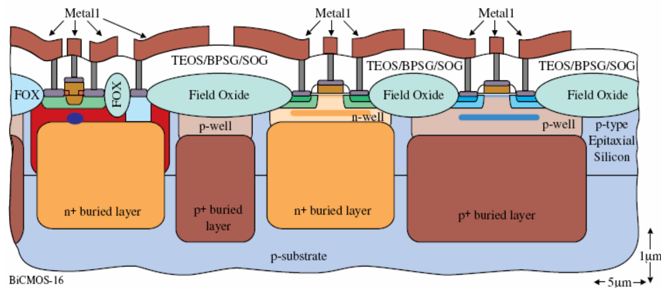
Contacts



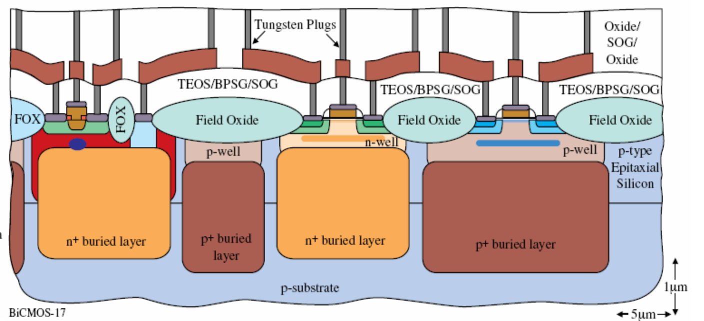
Comments:

- A dielectric is deposited over the entire wafer
- One of the purposes of the dielectric is to smooth out the surface
- Tungsten plugs are used to make electrical contact between the transistors and metal1

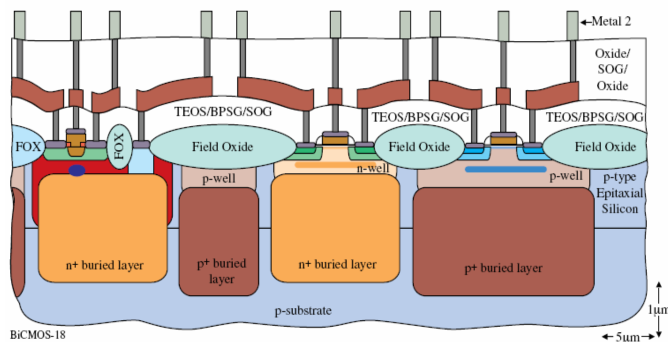
Metal1



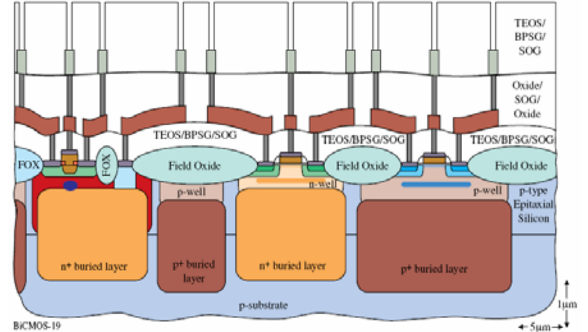
Metal1-Metal2 Vias



Metal2



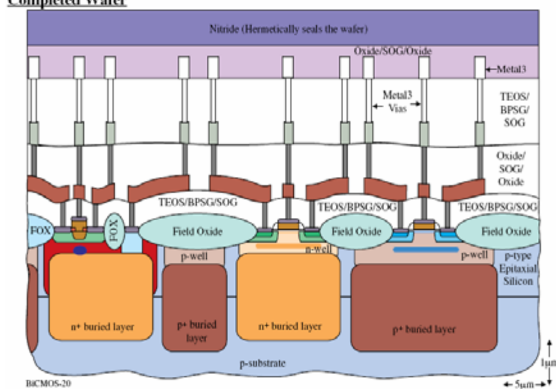
Metal2-Metal3 Vias



Comments:

- The metal2-metal3 vias will be filled with metal3 as opposed to tungsten plugs

Completed Wafer



In microelectronic exist two different kinds of accuracy. The absolute accuracy has in general a poor values and requires to increase the cost in order to improve. Fortunately it is easy to obtain a relative very good accuracy that allows to keep low costs.

Resistors

There are two techniques to implement resistors in CMOS technology the Diffusion through it is possible to obtain values about 10-100Ω/square with relative accuracy of 2% or ion implanted which values are about 500-2000Ω/square with same relative accuracy.

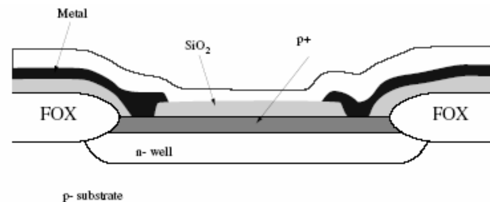


Figure 67 – general Source Drain resistor

Another common way is to implement the resistor through a polysilicon which values are about 30-100Ω/square for unshielded and 100-500Ω/Square for shielded.

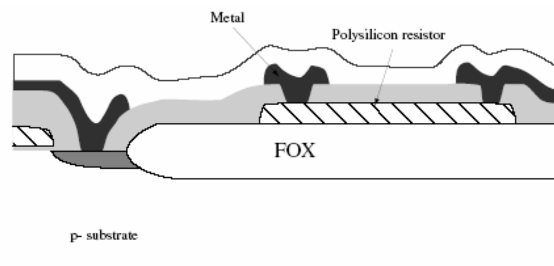


Figure 68 – general polysilicon resistor

Another way is through a n-well but to obtain resistor of 1000-5000Ω/square as showed in Figure 69.

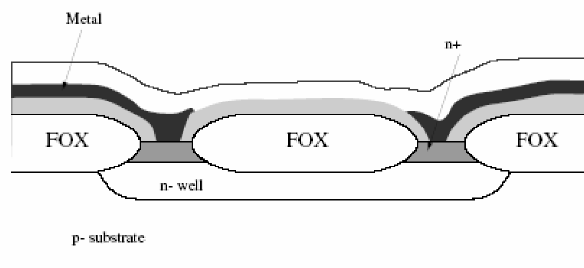


Figure 69 – n-well resistor

Capacitors and varactors.

The techniques to achieve the capacitors in CMOS technologies are different for example :
a)pn junction capacitors (as GaAs diodes)
b)Standard MOS capacitors
c)Accunulation mode MOS capacitors
d)Poly poly capacitors
e)Metal metal capacitors

Obviously the values depends by the specific technology. It is important for our purpose focusing the attention on pn junction capacitors and how to simulate and measure the capacitor by using a MOS transistor in diode connection.
The use of CMOS technology in ADS is not ready to designer. It is possible to implement some models on ADS for the technology and sing the standard models of SPICE to simulate the CMOS transistors.
As example it is possible to download by our website (www.diee.unica.it/miclab/miclab.html) the generic0.25 lib for use CMOS technology on ADS. This file contains a description and equations on the standard 0.25µm CMOS technology.

A simple test to change the capacitor is showed in the figure below

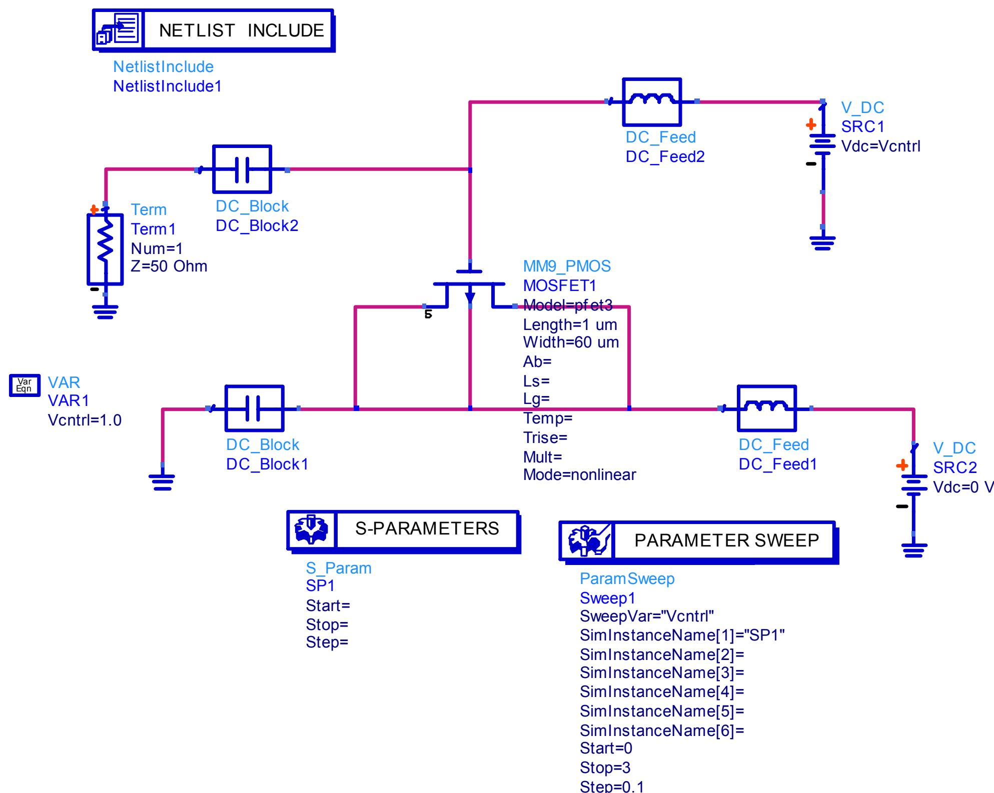


Figure 70 – a Simple test to use a PMOS transistor as varactor

A transistor is connected into diode configuration. The Drain and the gate must be connected through a DC-FEED to DC biasing. The Source must be grounded through a DC-Block and the gate Show through a DC-Block the capacitance seen for a fixed frequency by varying the Vcontrol.
The result is showed in the Figure 71. The capacitance value can be computed following the same procedure showed previously. Also taking into account that the presence of a signal with a large amplitude level can invert the values of the Gate compared with Drain and bring the transistor in conductivity. Obviously these simulations make sense only with physical models extracted by layout or provided by foundries services.

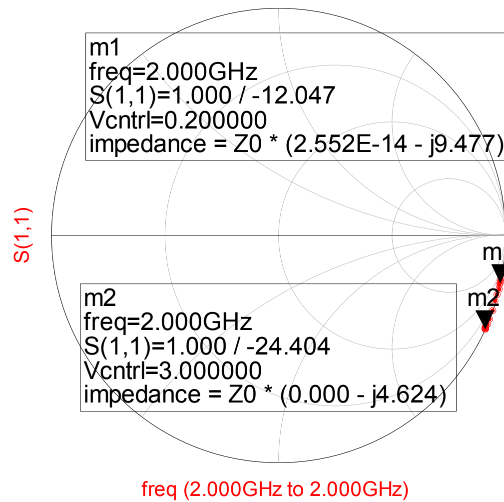


Figure 71 – The PMOS measured capacitance

The most great advantage of CMOS technology compared with GaAs is the possibility to use a PMOS and a NMOS transistor to implement a very good switch. In GaAs in fact the mobility positive charges is 20 times less than those for negative charges, then it is not convenient to realise positive channel devices It is possible also implement switches in GaAs technology but are very complex devices used to attenuate the digital or analog signals, where the attenuating function can be in a number of cascade transistor that can be swithed or not to realise the desired value of attenuation. In CMOS the difference betweens the two mobility is only 3 times, then to implement a very good switch it is sufficient to use a PMOS transistor 3 times large than NMOS. In addition the CMOS process allows to realise excellent switches with only two small transistors.

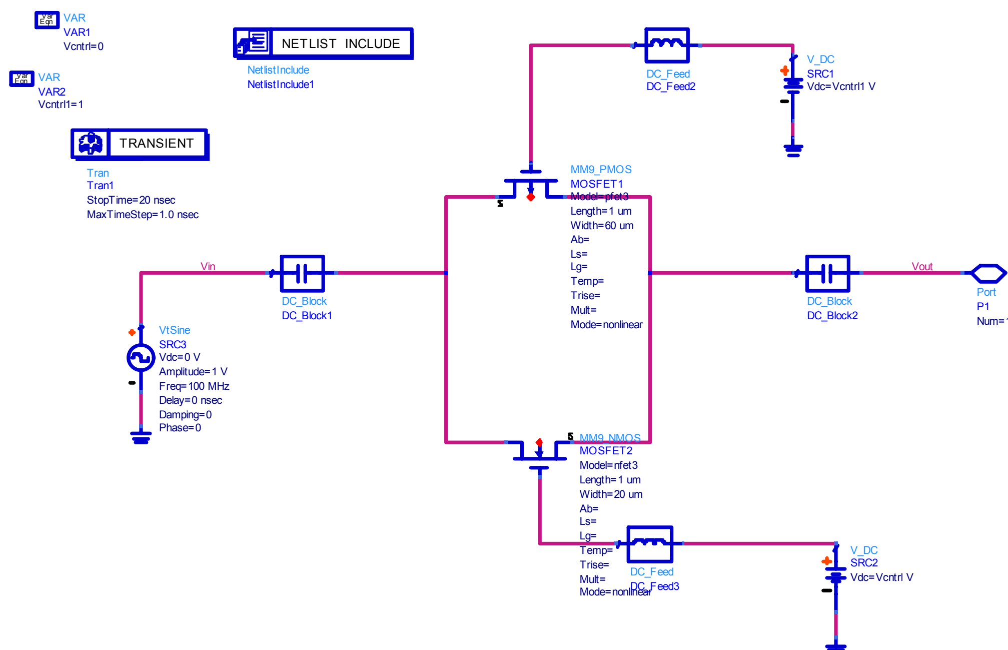


Figure 72 a transmission gate in CMOS

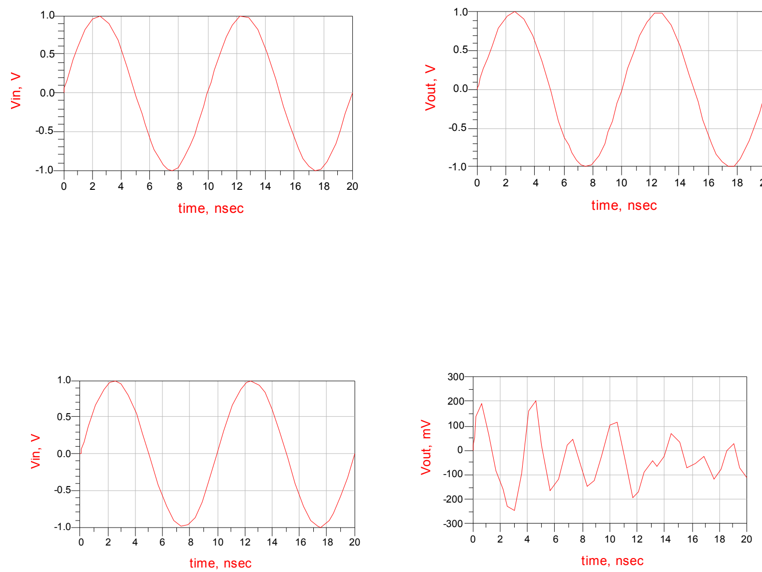


Figure 73 switch operation if at gate of PMOS is applyed 0V ue and at that of NMOS the VDD the transmission gate is open and allow to signal to pass; if VPMOS=3V and VNMOS=0V the switch is off.

The possibility to implement very good switch allow to create large varactors to guarantee a large tuning range. The disadvantage is generation of spark of power during the commutation power and this can get worse the phase noise performances. To minimize these effects in VCO that employ CMOS switches tuning is often used a sigma delta modulator that spread the phase noise of commutation in the frequency range.

A simple example of ideal switch capacitor it is represented in the figure:

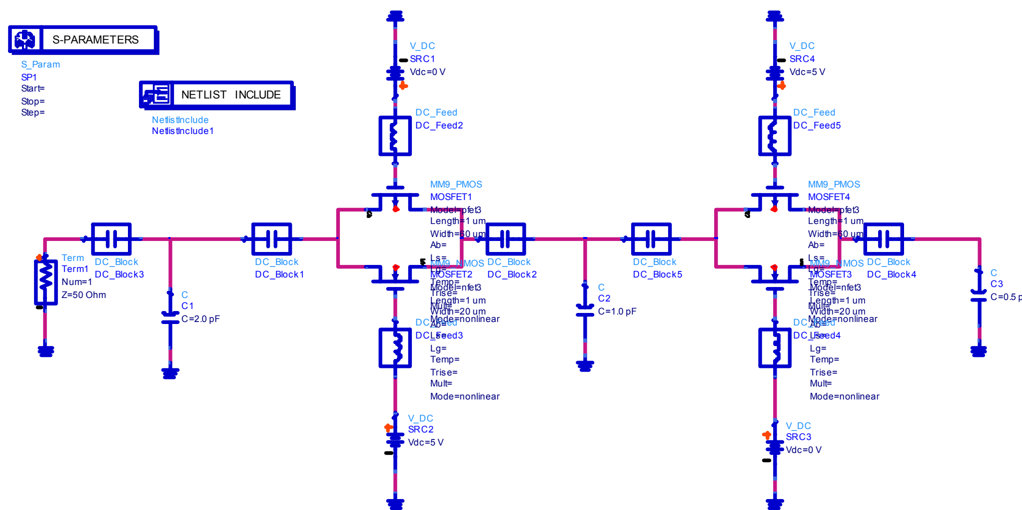


Figure 74 Switched Varactor

Depends on the values at the transmission gate ports the capacitance value must changed in weight way. The capacitors can be substituted by the circuit showed in Figure 70 to obtain a continuous change of the capacitance between a set and another selected by the switches.

Inductors

The way to present the inductors in CMOS technology is the same of the GaAs inductor based.
IN this section we will focus the attention to the design of an inductor. There are some university tools useful for this purpose as ASITIC (http://rfic.eecs.berkeley.edu/~niknejad/asitic.html). This software under windows necessitate of an emulator linux software as cygwin (http://www.cygwin.com/). The information about this software are not well explained therefore it is better to spent some words to helps the reader to set the correct parameters.
The first thing is to download cygwin and to run it. There are many tutorial but never explain that to operate correctly with cygwin, during the install procedure you must expand the X11 (last package of the list) and select all components one by one, after installing you can double click on cygwin and digit startx to make start the console. The program asitic_cygwin must be put in the /bin directory of the cygwin and then in the console you can write asitic_cygwin. At the start the program will demand you the tek file. Asitic work with some teck CMOS files. You can download a sample tek file from the website (http://rfic.eecs.berkeley.edu/~niknejad/doc-05-26-02/sample1.tek), then at the start you must digit the path of the sample1.tek for example d:/asitic/cygwin/sample1.tek then the program will be start.
Example of use of asitic

ASITIC> sq name=a len=175 w=10 s=1 n=5 metal=m3 exit=m2

ASITIC> sq= name=b len=150 w=15 s=2 n=4 metal=m1 exit=m2 xorg=300 yorg=300

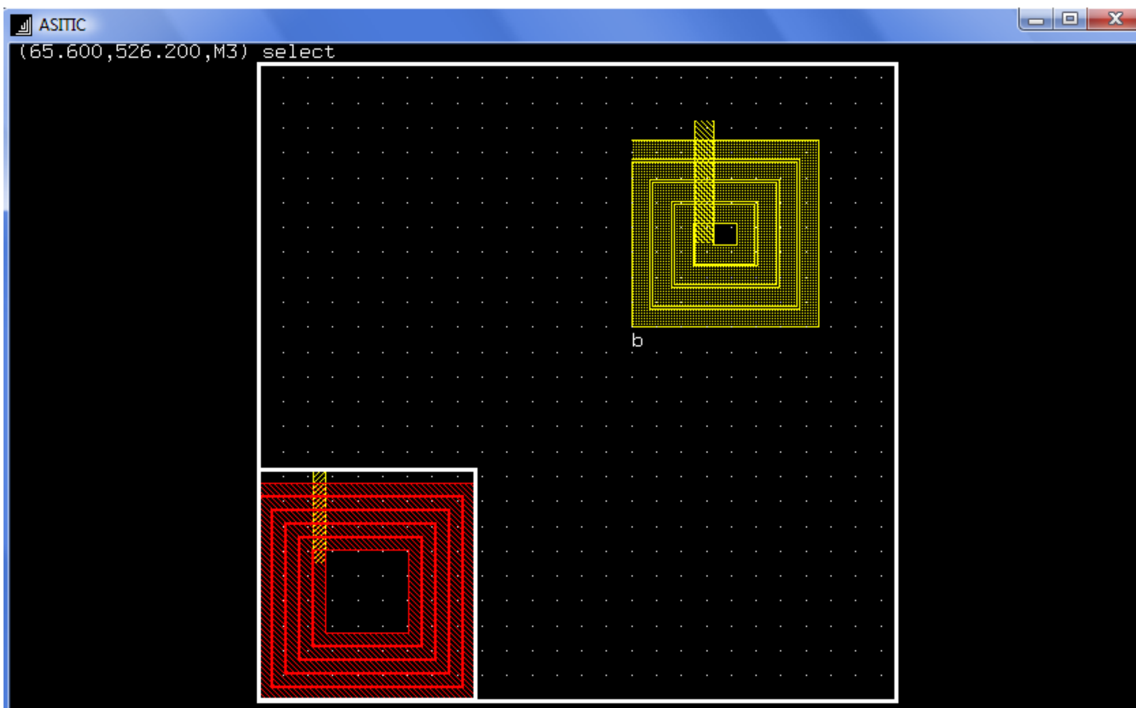


Figure 75 examples of inductors with asitic

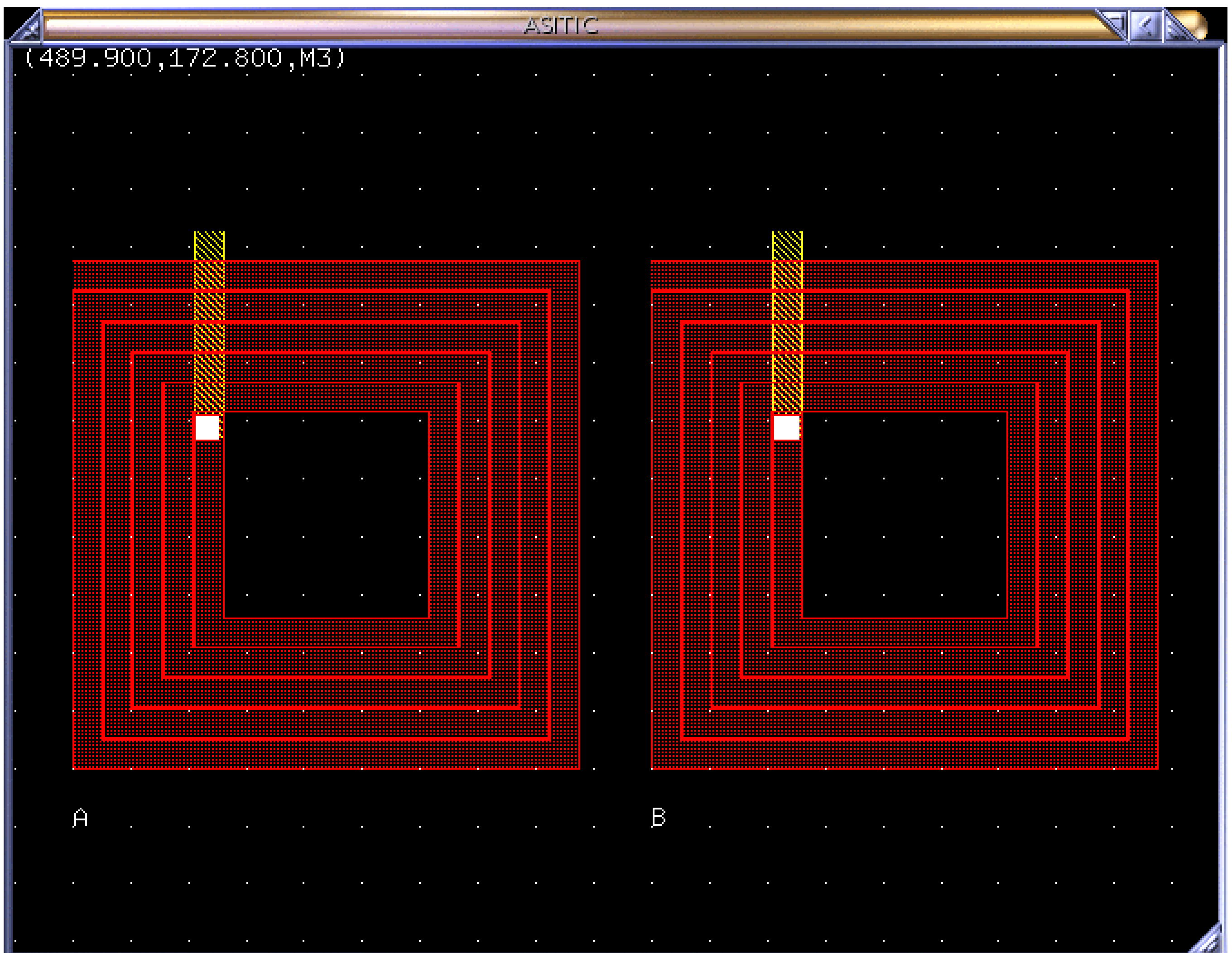
A very good tutorial is available on (http://rfic.eecs.berkeley.edu/~niknejad/doc-05-26-02/sample.html), nb to delete the inductance tape (del a..).
Asitic may be used also to compute the properties of CMOS metals layers for example it is possible to edit a simple square of metal that may be analyzed with the command pix (name) freq (GHz). A part of this tutorial is reported in this document for a fast consulting

Analysis of an Inductor Pair

So far we have only looked at a single device. Let's create a replica of this spiral and look at the interaction between two spirals.

```
ASITIC>  del solid
ASITIC>  cp a b
ASITIC>  mv a -100 0
ASITIC>  mv b 100 0
ASITIC>  friend a b
ASITIC>  mv a -25 0
```

You should now see both spirals in the layout.



By making the spirals *friends* we can move them both simultaneously. Let look at the magnetic coupling between these devices

```
ASITIC>  k a b

Coupling coefficient of A and B: k =   -0.02748 and M =   -0.11355
(nH).
```

```
ASITIC>  k2 2 a b

lambda = 37500.00, delta =  1.95

maxL = 1875.00, maxT =  1.56, maxW =  1.56
Generating inductance matrix (252x252)..
Inverting matrix......
Ind Timing:  tot =  4673, setup =    20, fill =  2117
           invert =  2526, reduce =    33, eddy =    00
L(A,A) =    4.03648 nH R(A,A) = 6.120
L(A,B) =   -0.11181 nH R(A,B) = -0.074
L(B,B) =    4.03624 nH R(B,B) = 6.127
```

The first command is strictly the DC coupling factor. When we compute the coupling at 2 GHz, we get similar results from the partial indutance matrix for these two devices. Each winding has a certain self inductance and resistance. The resistance is much higher than the DC resistance due to skin and proximity effects, in other words due to eddy currents in the metallization. The coupling term Z(A,B) is of course equal to Z(B,A) since the devices are passive. The imaginary part of the coupling is of course the mutual inductance but the real part is due to the change in the distributed current flow when one device is placed next to the other. In fact, if we were to connect these devices in series, then this term would lead to a reduction in the total series R by a small amount.

How well are these devices isolated from one another? In other words, if we ground one device and compute the impedance to ground from the other device, what do we get?

```
ASITIC>  cap a 2 b

lambda = 37500.00, delta =  1.95

maxL = 1875.00, maxT =  1.56, maxW =  1.56
Performing Analysis at 2.00 GHz
Generating capacitance matrix (233x233)..
At 2.000 GHz:
Total Capacitance = 281.395 (fF)
Total Resistance = 240.863.
```

Again this capacitance is lossy due to the substrate losses. Let's treat these two inductors as two windings of a transformer. What's the equivalent circuit?

```
ASITIC>  calctrans a b 2

lambda = 37500.00, delta =  1.95

maxL = 1875.00, maxT =  1.56, maxW =  1.56
Performing Analysis at 2.00 GHz
Generating capacitance matrix (233x233)...
Generating inductance matrix (315x315)..
Inverting matrix......
Ind Timing:  tot =  8065, setup =   -10, fill =  3714
           invert =  4357, reduce =    28, eddy =    00
Narrowband Model at f=2.00 GHz:
L1=  4.12  R1=  6.46  L2=  4.09 R2=  6.25 M=-0.117 (k=-0.0286) Re(Z12) = -0.0852
```

This result is very similar to what we would expect since the coupling is small. The magnetic coupling term is one order of magnitude larger than the substrate coupling term (the real part). Let's now join these two devices in series. First, create a wire to physically connect them:

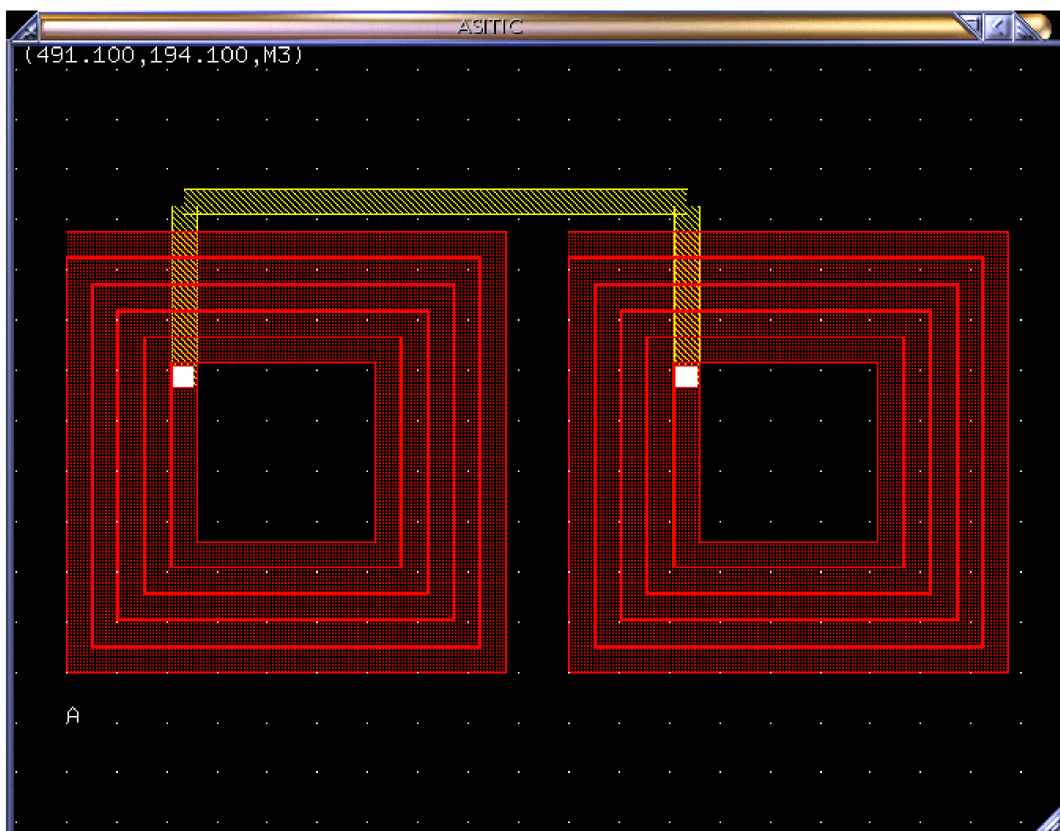
```
ASITIC>  wire name=c len=200 w=10 metal=m2 xorg=120 yorg=380
```

In practice you can move the wire to the proper location by selecting it with the mouse and moving it by dragging the structure to the appropriate location. The `snap` command controls the snap to grid size. Observe that the grid size and snap size are independent parameters. The coordinates of the mouse should be continuously displayed in the upper left corner of the window.

Since we are now about to create a user defined structure (as opposed to an internally synthesized structure), we have to exercise caution. Inside *ASTIIC* each device is a series interconnection of *super* segments. Each *super* segment consists of an arbitrary number of segments connected in shunt. To see this, use the `psegs` command (print segments). Now, in order to join spirals A and B in series, we have to make sure that the segments are in correct order with current flowing in the correct direction. We first must thus `flip` spiral B and change the phase since current now enters the inner port and exits from the outer port. The follwoing commands show this

```
ASITIC>  showdir

ASITIC>  flip b

ASITIC>  flipphase b

ASITIC>  join a c b

ASITIC>  ind a

Inductance of A = 8.80189 (nH).
```

The `showdir` command shows the phase of each segment. The total DC inductance is as we expect 2*(L+M). Here is the final layout:



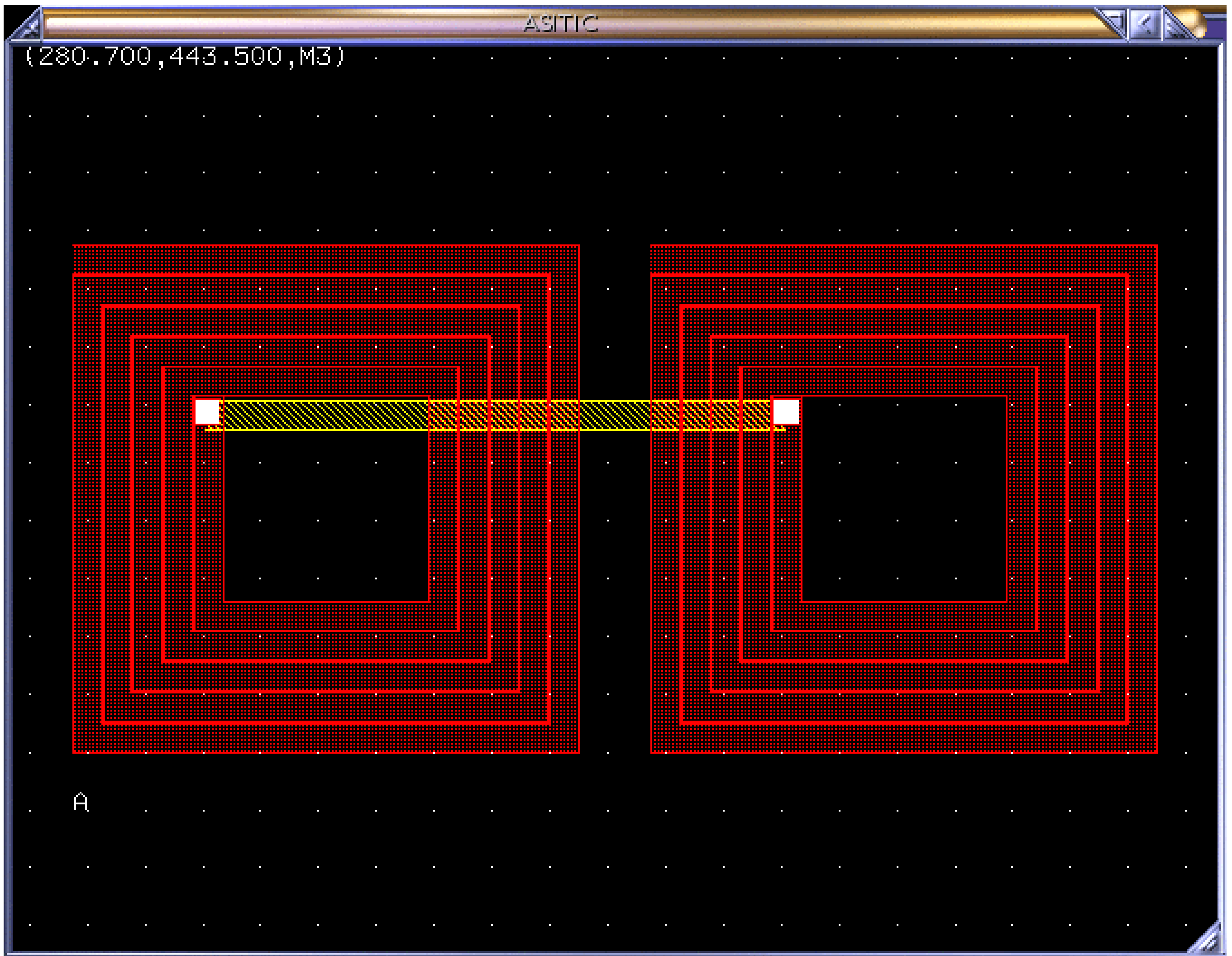
The high frequency performance of this pair is disappointing

```
ASITIC>  pix a 2

lambda = 37500.00, delta =  1.95

maxL = 1875.00, maxT =  1.56, maxW =  1.56
Performing Analysis at 2.00 GHz
Generating capacitance matrix (215x215)...
Generating inductance matrix (258x258)..
Inverting matrix......
Ind Timing:  tot =  3048, setup =    24, fill =  1294
           invert =  1720, reduce =    06, eddy =    00
Calc Times (ms):  total =  5684, cap =  2595, ind =  3059, node =    29
Pi Model at f=2.00 GHz:  Q = 4.58, 4.42, 6.82
L  =  8.93 nH  R  =  8.66
Cs1=   195 fF  Rs1=   653
Cs2=   203 fF  Rs2=   616     f_res =  3.82GHz
```

The capacitance is now doubled as expected. But notice that the Q of the device is reduced substantially due to increased substrate parasitics. In absence of substrate parasitics Q should be the same since we doubled both the series inductance and resistance. On the other hand, if we drive the structure differentially, then the Q degradation is tolerable (the third Q number of 6.82 is the differential Q).

Notice that we could also join the spirals by simply joining the inner ports directly. To make this modification we'll have to split the devices up, delete the extra segments, and rejoin the spirals:

```
ASITIC>  split a 0 b
ASITIC>  split b 2 c
ASITIC>  split b 0 d
ASITIC>  del d
ASITIC>  split a -1 e
ASITIC>  del e
ASITIC>  mv b 0 -65
ASITIC>  who

List of Spirals:

C
B
A

ASITIC>  join a b c

ASITIC>  ind a

Inductance of A = 8.57191 (nH).
```



The `split` command with a "0" argument splits a device down the middle creating a new device. With a non-zero argument $i > 0$, the device is cut after the i'th segment. With $i < 0$, the couting starts from the end. Hence the `split a -1 e` command removes the last segment from the spiral.

As expected, the partial inductance is reduced since we have reduced the electrical path from input to output, or equivalently the mangetic flux of the device.

The following commands calculate the equivalent *pi* circuit of the new device with and without a shield.

```
ASITIC>  pix a 2



lambda = 37500.00, delta =  1.95

maxL = 1875.00, maxT =  1.56, maxW =  1.56
Performing Analysis at 2.00 GHz
Generating capacitance matrix (205x205)...
Generating inductance matrix (246x246)..
Inverting matrix......
Ind Timing:  tot =  3964, setup =    02, fill =  1994
           invert =  1963, reduce =    04, eddy =    00
```

```
Calc Times (ms):  total =  6084, cap =  2099, ind =  3983, node =    02
Pi Model at f=2.00 GHz:  Q = 4.90, 4.77, 7.36
L  =  8.64 nH   R  =  7.67
Cs1=   187 fF   Rs1=   648
Cs2=   193 fF   Rs2=   624      f_res =  3.96GHz

ASITIC>  wire name=sh len=450 wid=200 metal=msub

ASITIC>  mv sh 37 -13

ASITIC>  geom sh


Wire  has the following geometry:
L = 450.00, W = 200.00, Metal = MSUB
Total length =  450.00 (um), Total Area = 90000.00 (um^2)
Located at (37.00,187.00) with 1 segments.

ASITIC>  pix a 2 sh

lambda = 37500.00, delta =  1.95

maxL = 1875.00, maxT =  1.56, maxW =  1.56
Performing Analysis at 2.00 GHz
Generating capacitance matrix (435x435)...
Generating inductance matrix (246x246)..
Inverting matrix......
Ind Timing:  tot =  3156, setup =    21, fill =  1306
             invert =  1823, reduce =    03, eddy =    00
Calc Times (ms):  total = 16351, cap = 13176, ind =  3147, node =    26
Pi Model at f=2.00 GHz:  Q = 6.39, 6.35, 7.73
L  =  7.65 nH   R  =  10.6
Cs1=   232 fF   Rs1=  2.73
Cs2=   241 fF   Rs2=  1.61      f_res =  3.77GHz
```

The above discussion applies to non-square structures. For instance, the `spiral` command creates a polygon structure while the `symsq` command creates symmetric inductors. See the creation reference sections for more details.

Multi-Layer Stacked Inductors

Note: For the purpose of this discussion, please locate the technology file `sample2.tek`. This technology file resembles a modern multi-layer IC process.

ASITIC has several built in commands that generate inductors on multiple metal layers. The command `sqsh` creates a spiral inductor identical in layout to the `sq` command with the exception that multiple metal layers are put in parallel to lower the resistance of the device. Let's first try a two layer structure:

```
ASITIC>  sqsh name=a2 len=150 w=8 s=1 n=3.75 metal=m3 exit=m2 xorg=200 yorg=200
cbegincend exit90
```

This structure resides on metal m3 and m2 in shunt and the exit occurs through layer m1. To see this, type `psegs` to see a list of segments:

```
ASITIC>  psegs a2
.
.
.
Segment  14:  (   119.0,    31.0,    45.0)-(    69.0,    31.0,    45.0) on
```

```
        Shunt:  (  119.0,   31.0,   47.0)-(   69.0,   31.0,   47.0) on
Segment  15:  (   65.8,   31.0,   47.0)-(   72.2,   31.0,   45.0) on
Segment  16:  (   69.0,   31.0,   48.0)-(   69.0,   -8.0,   48.0) on
```

Notice that ASITIC reports that each segment before the last two are "super" segments as they consist of two metal layers strapped together. Even though you can't see this (there are no vias), rest assured that this is the case. Let's compare the low-frequency impedance of this structure to an identical single layer structure:

```
ASITIC>  hide a2
ASITIC>  sq name=a1 len=150 w=8 s=1 n=3.75 metal=m3 exit=m2 xorg=200 yorg=200
cbegin cend exit90

ASITIC>  indmat a1 .1

lambda = 750000.00, delta = 10.07
maxL = 37500.00, maxT =  8.05, maxW =  8.05
Generating inductance matrix (16x16)..
Inverting matrix......
L =   2.4777073e-09            R =      8.50000000
ASITIC>  indmat a2 .1

lambda = 750000.00, delta = 10.07
maxL = 37500.00, maxT =  8.05, maxW =  8.05
Generating inductance matrix (31x31)..
Inverting matrix......
L =   2.3760413e-09            R =      4.35750000
```

As expected, the low frequency resistance dropped by about a factor of 2 and the inductance dropped slightly since the two stacked windings are strongly coupled. Let's compare the high frequency behavior:

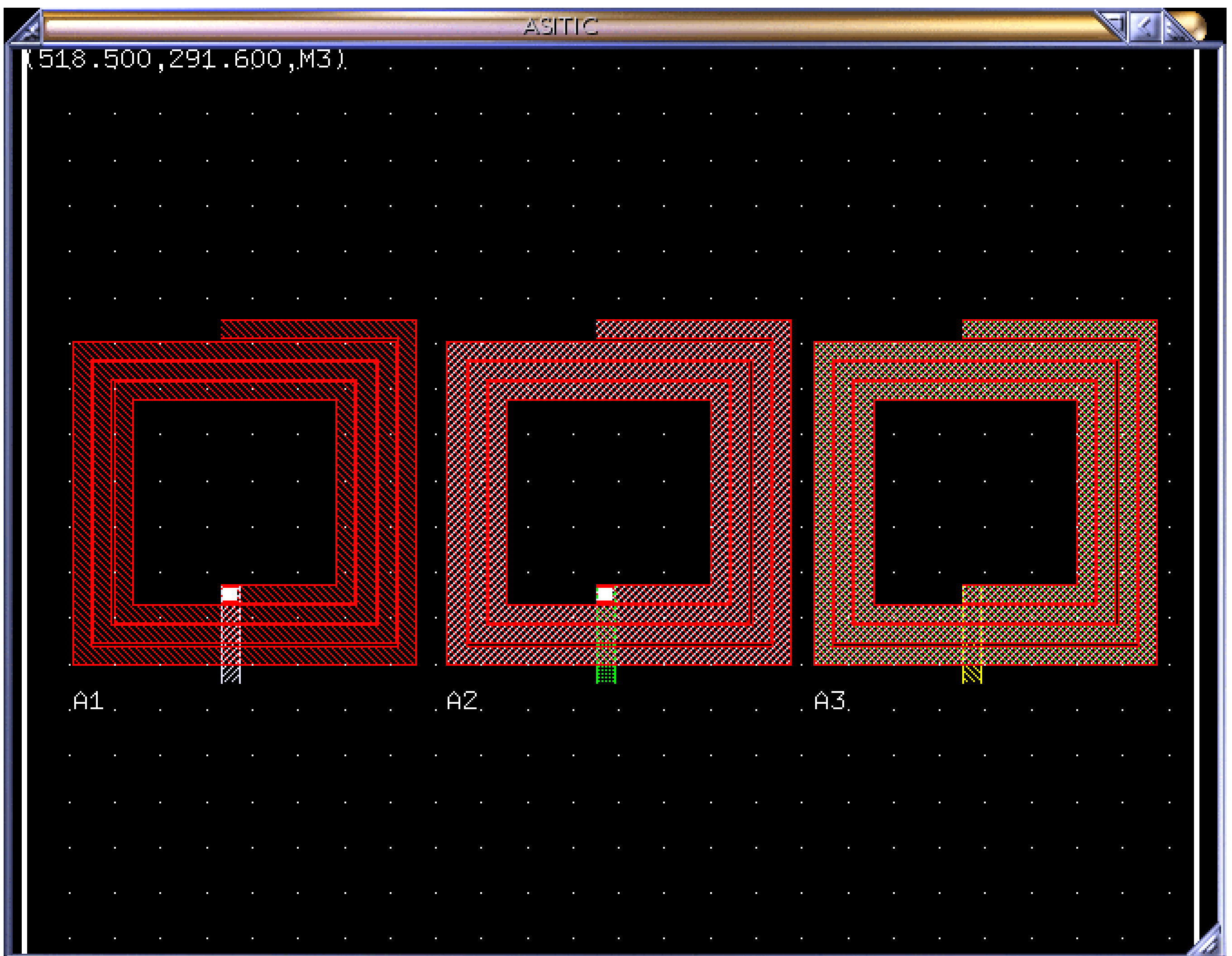
```
ASITIC>  pix a1 3

lambda = 25000.00, delta =  1.84
maxL = 1250.00, maxT =  1.47, maxW =  1.47
Performing Analysis at 3.00 GHz
Generating capacitance matrix (64x64)...
Generating inductance matrix (80x80)..
Inverting matrix......
Pi Model at f=3.00 GHz:  Q = 4.62, 4.66, 4.95
L  =  2.47 nH   R  =  8.77
Cs1=  50.6 fF   Rs1=   633
Cs2=    47 fF   Rs2=   704      f_res = 14.25GHz


ASITIC>  pix a2 3

lambda = 25000.00, delta =  1.84
maxL = 1250.00, maxT =  1.47, maxW =  1.47
Performing Analysis at 3.00 GHz
Generating capacitance matrix (124x124)...
Generating inductance matrix (155x155)..
Inverting matrix......
Pi Model at f=3.00 GHz:  Q = 7.02, 7.10, 7.89
L  =  2.31 nH   R  =   4.9
Cs1=  63.2 fF   Rs1=   639
Cs2=  59.9 fF   Rs2=   702      f_res = 13.16GHz
```

Also as expected, the stacked structure has higher capacitance as it lies closer to the substrate. Also at 3 GHz the quality factor of the two layer structure is better due to the lower winding loss. How about a three layer structure?

```
ASITIC>  timing
ASITIC>  hide a1
```

```
ASITIC>  sqsh name=a3 len=150 w=8 s=1 n=3.75 metal=m3 exit=m1 xorg=200 yorg=200
cbegin cend exit90

ASITIC>  pix a3 3

lambda = 25000.00, delta =  1.84
maxL = 1250.00, maxT =  1.47, maxW =  1.47
Performing Analysis at 3.00 GHz
Generating capacitance matrix (184x184)...
Generating inductance matrix (230x230)..
Inverting matrix......
Ind Timing:  tot =  3006, setup =    32, fill =  2444
             invert =   495, reduce =    32, eddy =    00
Calc Times (ms):  total =  5508, cap =  2502, ind =  3000, node =    05

Pi Model at f=3.00 GHz:  Q = 7.33, 7.34, 8.59
L  =  2.21 nH   R  =  4.04
Cs1=  89.2 fF   Rs1=   665
Cs2=  89.7 fF   Rs2=   683      f_res = 11.34GHz
```

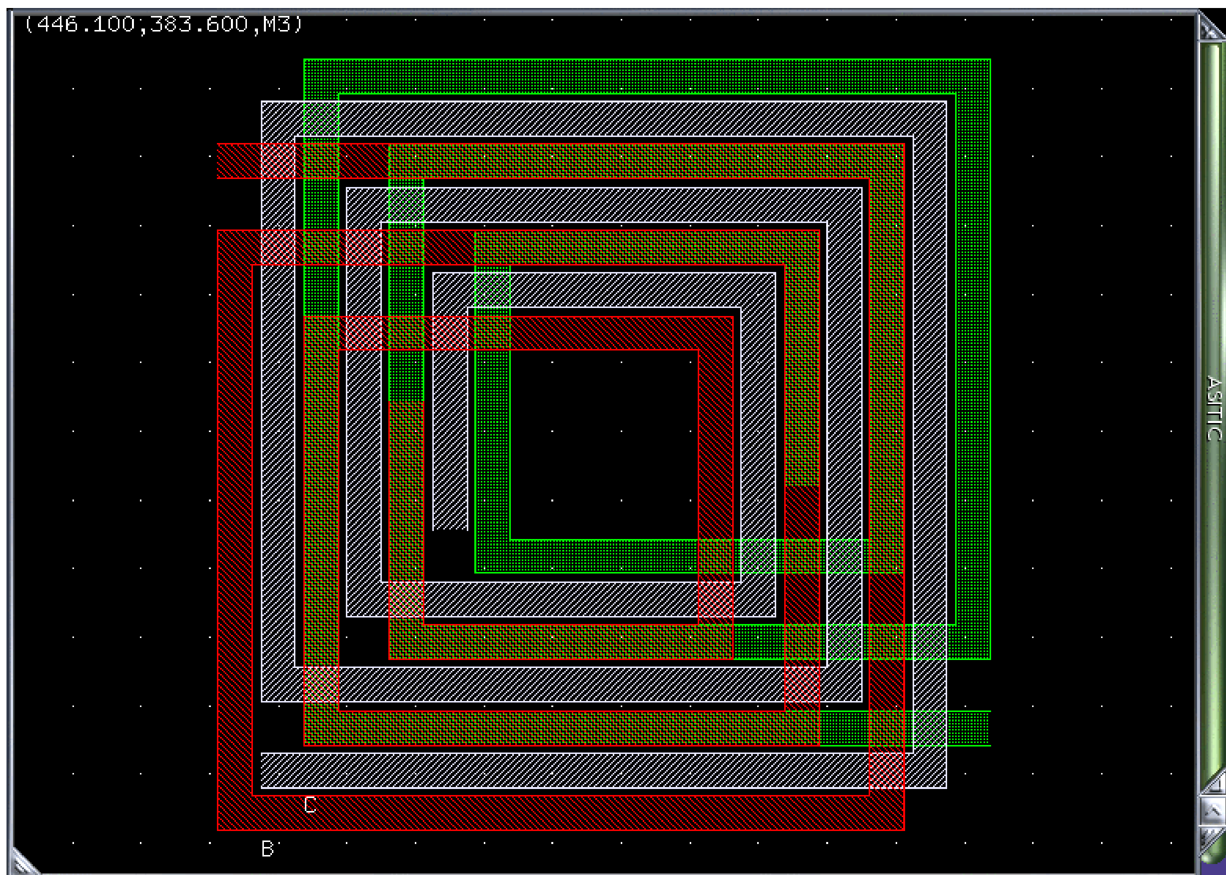
We're getting close to a point of of diminishing returns. While the series lesses drop, the increased substrate losses due to the close proximity of the substrate begin to limit the improvement in Q. The following figure shows the three devices side-by-side for a comparison.



While the `sqsh` structures tend to reduce loss at approximately constant inductance, the `sqmm` series connected devices increase inductance almost quadratically with the number of turns with only a linear increase in resistance. Let's see this

```
ASITIC>  del a1 a2 a3
ASITIC>  sq name=s1 len=200 w=10 s=1 n=4 metal=m3 exit=m2 xorg=200 yorg=200
exit90
ASITIC>  sqmm name=s2 len=200 w=10 s=1 n=4 metal=m3 exit=m2 xorg=200 yorg=200
exit90
```

The figure below shows how the multi-layer series device is wound. Observe that the successive winding on layers below run parrallel to the top windings in order to reinforce the magnetic field. Simply stacking two windings in series would result in a reduction of inductance.




```
ASITIC>  indmat s1 .1

lambda = 750000.00, delta = 10.07
maxL = 37500.00, maxT =  8.05, maxW =  8.05
Generating inductance matrix (17x17)..
Inverting matrix......
Ind Timing:  tot =    24, setup =    00, fill =    19
           invert =    02, reduce =    01, eddy =    00
L =   4.0726003e-09             R =     10.25600000
ASITIC>  indmat s2 .1


lambda = 750000.00, delta = 10.07
maxL = 37500.00, maxT =  8.05, maxW =  8.05
Generating inductance matrix (32x32)..
Inverting matrix......
Ind Timing:  tot =    92, setup =    00, fill =    76
           invert =    10, reduce =    03, eddy =    00
```

```
L =   1.4433833e-08             R =     20.08800000
```

The low frequency results are close to our expectation. How about the high frequency resistance?

```
ASITIC>  indmat s1 3


lambda = 25000.00, delta =  1.84

maxL = 1250.00, maxT =  1.47, maxW =  1.47
Generating inductance matrix (119x119)..
Inverting matrix......
Ind Timing:  tot =  1314, setup =    02, fill =  1160
           invert =   126, reduce =    22, eddy =    00
L =   4.0283518e-09             R =     11.61582699
ASITIC>  indmat s2 3


lambda = 25000.00, delta =  1.84

maxL = 1250.00, maxT =  1.47, maxW =  1.47
Generating inductance matrix (224x224)..
Inverting matrix......
Ind Timing:  tot =  3285, setup =    31, fill =  2039
           invert =  1207, reduce =    03, eddy =    00
L =    1.517041e-08             R =     29.20356681
```

Notice that the AC resistance of the stacked device is much higher (45% increase) whereas the single layer stucture shows a less pronounced increase in loss (13%). Another issue with a multi-layer structure is the large inter-winding capacitance which lowers the frequency of self-resonance

```
ASITIC>  pix s1 3

lambda = 25000.00, delta =  1.84
maxL = 1250.00, maxT =  1.47, maxW =  1.47
Performing Analysis at 3.00 GHz
Generating capacitance matrix (85x85)...
Generating inductance matrix (119x119)..
Inverting matrix......
Ind Timing:  tot =  1477, setup =    11, fill =  1167
           invert =   269, reduce =    26, eddy =    00
Calc Times (ms):  total =  2032, cap =   539, ind =  1487, node =    06
Pi Model at f=3.00 GHz:  Q = 5.04, 5.16, 6.24
L  =  4.11 nH   R  =   9.8
Cs1=  80.2 fF   Rs1=   614
Cs2=  75.5 fF   Rs2=   677      f_res =  8.77GHz
ASITIC>  pix s2 3

lambda = 25000.00, delta =  1.84
maxL = 1250.00, maxT =  1.47, maxW =  1.47
Performing Analysis at 3.00 GHz
Generating capacitance matrix (160x160)...
Generating inductance matrix (224x224)..
Inverting matrix......
Ind Timing:  tot =  5379, setup =    10, fill =  3759
           invert =  1522, reduce =    85, eddy =    00
Calc Times (ms):  total =  8759, cap =  3338, ind =  5391, node =    28
Pi Model at f=3.00 GHz:  Q = 3.06, 1.66, 3.52
C  =  75.3 fF   R  =  75.1
Cs1=    47 fF   Rs1=1.23e+03
Cs2=   148 fF   Rs2=   441
```

Even at 3 GHz, the device has experienced self-resonance and looks like a capacitor. Whereas the single layer device is still a healthy inductor. Let's check at a lower frequency (1 GHz)

```
ASITIC>  pix s1 1
```

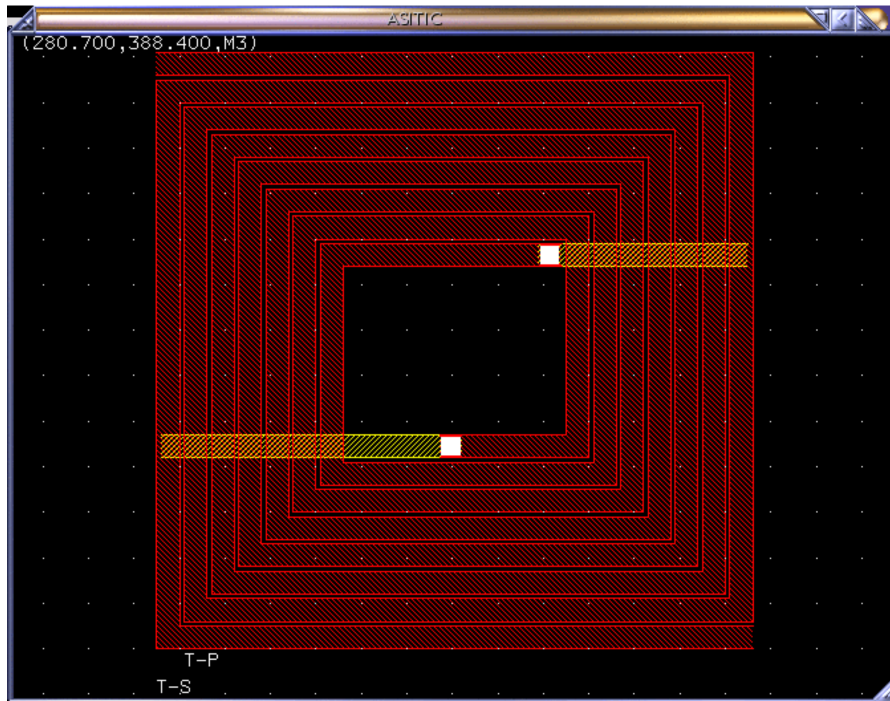
```
lambda = 75000.00, delta =  3.18
maxL = 3750.00, maxT =  2.55, maxW =  2.55
Performing Analysis at 1.00 GHz
Generating capacitance matrix (68x68)...
Generating inductance matrix (68x68)..
Inverting matrix......
Ind Timing:  tot =   101, setup =    01, fill =    78
           invert =    14, reduce =    07, eddy =    00
Calc Times (ms):  total =   282, cap =   174, ind =   102, node =    05
Pi Model at f=1.00 GHz:  Q = 2.40, 2.40, 2.44
L  =  4.06 nH   R  =  10.3
Cs1=  94.8 fF   Rs1=   642
Cs2=  88.2 fF   Rs2=   690      f_res =  8.11GHz
ASITIC>  pix s2 1

lambda = 75000.00, delta =  3.18
maxL = 3750.00, maxT =  2.55, maxW =  2.55
Performing Analysis at 1.00 GHz
Generating capacitance matrix (128x128)...
Generating inductance matrix (128x128)..
Inverting matrix......
Ind Timing:  tot =   807, setup =    12, fill =   703
           invert =    83, reduce =    06, eddy =    00
Calc Times (ms):  total =  2488, cap =  1663, ind =   803, node =    21
Pi Model at f=1.00 GHz:  Q = 3.64, 3.04, 3.75
L  =  17.5 nH   R  =  26.9
Cs1=  65.7 fF   Rs1=1.19e+03
Cs2=   173 fF   Rs2=   465      f_res =  4.69GHz
```

Now things look better at 1 GHz. Note that the multi-layer structure has a self-shiedling property. In other words, the bottom coil acts as a substrate shield and thus the substrate losses of the top shield are reduced.

Transformers and Baluns

Planar transformers can be created with the `trans` command:

```
ASITIC>  trans name=t len=250 w=10 s=2 n=3.75
ASITIC>  mv t-p 120 120
```



ASITIC concatenates a "-P" and "-S" to the name you specify in order to designate the primary and secondary windings. In the above case we wound the primary and secondary similarly (with equal turns and width). Let's create an asymmetric transformer as follows

```
ASITIC>  hide t-p t-s
ASITIC>  mv x-p 120 120
ASITIC>  trans name=x len=250 w=10 s=2 np=4.5 ns=2
```

You'll notice that *ASITIC* made the primary and secondary windings "friends". Thus when you move one, the other moves as well. Otherwise the windings are two distict independent spiral.

You can find the inductance matrix of the device at high frequency with the `k2` command.

```
ASITIC>  k2 1.2 t-s t-p

lambda = 62500.00, delta =  2.52
maxL = 3125.00, maxT =  2.01, maxW =  2.01
Generating inductance matrix (160x160)..
Inverting matrix......
L(T-S,T-S) =    3.34111 nH      R(T-S,T-S) = 5.135
L(T-S,T-P) =    2.66867 nH      R(T-S,T-P) = 0.300
L(T-P,T-P) =    3.34187 nH      R(T-P,T-P) = 5.404
ASITIC>  k2 1.2 x-s x-p

lambda = 62500.00, delta =  2.52
maxL = 3125.00, maxT =  2.01, maxW =  2.01
Generating inductance matrix (140x140)..
Inverting matrix......
L(X-S,X-S) =    1.83162 nH      R(X-S,X-S) = 3.480
L(X-S,X-P) =    1.83813 nH      R(X-S,X-P) = 0.152
L(X-P,X-P) =    3.67132 nH      R(X-P,X-P) = 5.598
```

The `calctrans` command computes the high-frequency inductance and capacitive behavior of the device at a particular frequency:

```
ASITIC>  calctrans t-s t-p 2

lambda = 37500.00, delta =  1.95
```

```
maxL = 1875.00, maxT =  1.56, maxW =  1.56
Performing Analysis at 2.00 GHz
Generating capacitance matrix (160x160)...
Generating inductance matrix (192x192)..
Inverting matrix......
Narrowband Model at f=2.00 GHz:
L1=   3.4  R1=  5.16  L2=   3.4 R2=  5.44 M=  2.73 (k= 0.805) Re(Z12) =  0.818
```

It's important to realize that this is not a circuit model for the transformer but simply a translation of the 2-port z-parameters into a particular circuit representation. In other words, R1 + i*w*L1 = z11, and M = imag(z12)/(2*pi*freq), and so on. To actually design a broadband model for the spiral, the `transs` command should be used to generate *s*-parameters over a wide frequency range. See the section broading modeling for more details.

If a large turns ratio transformer is desired for a particular application more exotic structures can be synthesized by hand. For instance, multi-layer spirals are very easy to generate and analyze:

```
ASITIC>  del t-p t-s x-p x-s
ASITIC>  sq name=pri len=200 w=8 s=8 n=5 metal=m3 exit=m1
ASITIC>  sq name=sec len=200 w=4 s=4 n=10 metal=m2 exit=m1
ASITIC>  rot sec 90
ASITIC>  k2 2 pri sec

lambda = 37500.00, delta =  1.95
maxL = 1875.00, maxT =  1.56, maxW =  1.56
Generating inductance matrix (228x228)..
Inverting matrix......
L(PRI,PRI) =    3.69316 nH      R(PRI,PRI) = 6.410
L(PRI,SEC) =    6.18359 nH      R(PRI,SEC) = 1.071
L(SEC,SEC) =   13.55803 nH      R(SEC,SEC) = 103.351
```

Notice that a large secondary to primary inductance ratio is obtained. Even larger ratios can be obtained by mixing a multi-layer spiral with a single layer spiral. For instance the primary can have two metal layers in series and the secondary can have two metal layers in shunt. You get the idea.

To realize a balun, we can tap the center of the secondary and create a three-port device with respect to a common ground. The "inductive" center, though, does not naturally coincide with the geometric center and this implies asymmetric capacitance and resistance on the secondary windings. For fully-differential circutis this is undesirable. The `balun` command solves this problem by creating a symmetric structure:

```
ASITIC>  balun name=b len=200 w1=10 s=1 n=5 metal=m3 metal2=m2 xorg=200 yorg=200
ASITIC>  mv b-s -50 -50
ASITIC>  ind b-s

Inductance of B-S = 0.96325 (nH).
ASITIC>  ind b-p

Inductance of B-P = 1.76683 (nH).
ASITIC>  split b-p 0 b-p2

ASITIC>  ind b-p

Inductance of B-P = 0.646206 (nH).
ASITIC>  ind b-p2

Inductance of B-P2 = 0.646206 (nH).
ASITIC>  res b-p

Resistance of B-P = 1.576453 (Ohms).
```

```
ASITIC>  res b-p2

Resistance of B-P2 = 1.576453 (Ohms).
ASITIC>  k b-s b-p

Coupling coefficient of B-S and B-P: k =    0.67907 and M =    0.53576 (nH).
ASITIC>  k b-s b-p2

Coupling coefficient of B-S and B-P2: k =    0.67925 and M =    0.53590 (nH).
```

NB the syntax of some commands can change in your machine and there are some mistakes in this edit. The correct syntax could be the following:

Type:(sq, wire) name=(name) len=(length) w=(width) n=(number of turns) s=(space between turns) metal=(msub, m2 m3) exit(msub,m2,m3), xorg=(horizontal pos) yorg(vertical pos).

When you analyze the structure following the example previously describe you must write the commands parts in this order; could help.

pix name frequency.

REFERENCES

[1] Sammy Kayali- George Ponchak Roland Shaw "GaAs MMIC Reliability Assurance Guideline for Space Applications. JPL Pubblication 96-25" 1996

[2]S. M. Sze, Semiconductor devices: physics and technology, John Wiley & Sons, 2002

[3]I.D. Robertson and S.Lucyszyn "RFIC and MMIC design and technology"

[4]BRHZAD RAZAVI "RF microelectronics"Prentice Hall 1998

[5] Jwo-Shiun Sun, Design And Analysis of Microwave Varactor-Tuned Oscillators", Microwave Journal, pag.302-310, Maggio 1999.

[6] S. Chen – S. Tadayon, U-Band MMIC HBT DRO", IEEE Microwave and Guided Letters, vol. 4, n°2, pag . 50-52, Febbraio 1994.

[7]I. Bahal, P. Barthia "Microwave solid state circuit design" Ed. Jhon Wiley & Sons, 1987.

Others References

- *Paul Gray, Paul Hurst, Steve Lewis and Robert mayer, Analysis and Design of Analog Integrated Circuits,Fourth Editon, Jhon Wiley and Sons, Inc 2001*
- *P.E.Allen and D.R.Holberg, CMOS Analog Integrated Circuit Design-Second Edition, Oxford University Press,2002*
- *B.Razavi, Design of Integrated Circuits for Optical Communications, MCGraw-Hill, 2003*
- *METZ, E.P.A., MILLER,R.C, andMAZELSKY, R, "A technique for pulling crystals of volatile material", Journal pf Applied Physics, 1962*
- *TURNER, J. A "History of the GaAs FET at Caswell(1964-1985)" IEE Colloquium on modelling, design and application on MMICs, London 1994 digest no. 1994/092 pp 1/1-1/3*
- *MEHAL, E. W., and WACKER, R. W.: "GaAs Integrate microwave Circuits", IEEE of Solid StateCircuits, 1968 SC-3(2) pp 117-123*
- *STRID, E.W,and GLEESON, K.R "A DC-12GHz monolithic GaAs FET distributed amplifier" IEEE Transaction on Mictowave Theories and Techniques, 1982 MTT-30(7) pp 969-975*

- *ARNOLD R.G. and PEDDER D.J "Microwave characterization of microstrip lines and spiral inductors" in MCM-D technology, IEEE Transaction on Components, Hybrids and Manifacturing Technology, 15(6) pp1038-1045*
- *WOOD, J.SABO, C LONDON A., and BARNWELL P: "Processing of photoimageable thick film materials", Proceeding of SPIE – the International Society for Optical Engineering, 1999, 3582, pp783-8*
- *LUCYSZYN, S. et al: "Therahertz multi-chip Modules(T-MCM) technology fot the 21st Century?", IEE Colloquium on Multi-Chip Modules and RFIC's London 1998,pp 6/1-8.*
- *BLONDY, P., BROWN, A.R., CROS, D., and REBEIZ, G.M "Micromachining for Therahertz Applications", IEEE Transactions on Microwave theories and techniques, 1998 46(11), pp 1821-1831*
- www.diee.unica.it/miclab/miclab.html

3 OSCILLATORS GENERAL THEORY AND CLASSIFICATION

Oscillators are the key building of integral transceivers. An oscillator classification can be based on one of the basic properties of the oscillator as frequency, tuning range, noise performances pr functionality (single or multiphase outputs). The first possible classification group together oscillators which frequency can be changed by electric control and oscillator which frequency is fixed. These oscillators are summarized in the Fig.1

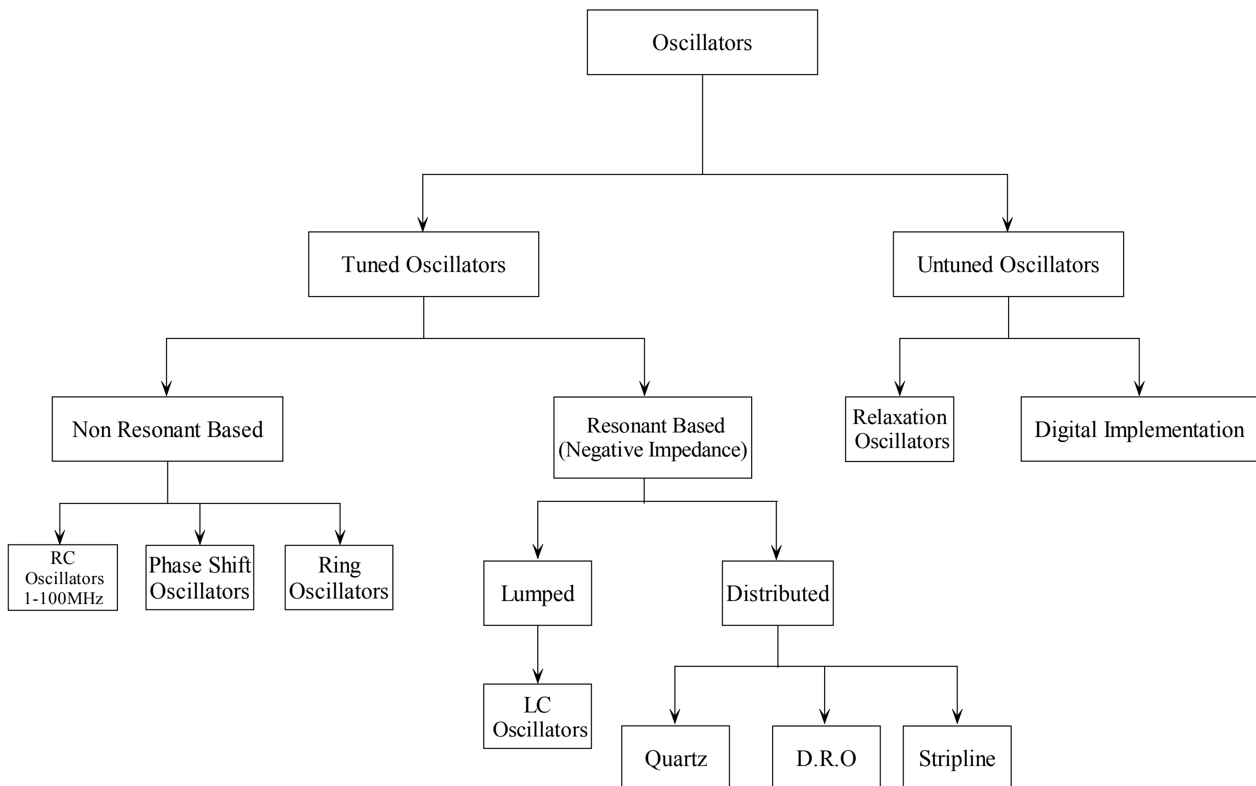


Figure 1- First classification for oscillators

Belong to interest of this work the Tuned Oscillators.
For all categories there is the possibility to produce a sinusoidal waveform or a square waveform, depends by the implementation way.
In particular the LC oscillator can be implemented by following two different architectures composed by a single transistor or two transistor in a cross coupled configuration.
The topic of this thesis focuses the attention on single transistor oscillators, whereas the ***most popular structure*** is that of ***cross coupled oscillators*** that are employed in modern wireless transceivers.

3.1 The Harmonic Balance Simulator.

Oscillators are nonlinear elements, and their behaviour can be correctly formal studied only through the non-linear techniques and differential equations in the time domain. This because the oscillator is a time varying non linear circuit. However these techniques are very difficult to employ without a specific CAD and today does not make sense as well, in addition would not give any insight in the design. If you must analyze an oscillator use the phase portrait method and the non-linear techniques give you best performance but if you must design an oscillator it is impossible or very

very difficult to use these methods and employing them is time consuming and cannot give insight at the designer when the design does not give the correct performances.
The history of circuit simulation for RF design has meandered back and forth between improvements to device models and improvement to the circuit simulators themselves. In the early 1970s, SPICE was pioneered and prompted a range of device modelling efforts, including the Gummel-Poon model for the bipolar transistor. Simple MESFET models were also introduced in the early 1980s, but the limitations of SPICE soon rendered further improvements to modellinf accuracy wasted. It was not until the late 1980s and the commercialization of harmonic balance simulators for PCs that device modelling efforts once again accelerated, particularly for microwave monolithic integrated circuit and GaAs technology. The 1990s saw device manufacturers properly characterizing their devices and introducing device libraries, while a number of improved device models once again appeared. Now, in the early 2000s, PC simulation techniques have matured to the point where nonlinear characterization of diverse phenomenon such as oscillation, phase noise, and high-order distortion products can be performed with relative ease. Perhaps the later years of this will again see the modelers playing catch-up, particularly for compound semiconductors devices. Without describe in details the harmonic balance method and the history of RF simulators (for more details you can check the Practical RF circuits of Besser). All simulations in this work use the Harmonic Balance method. The assumption a priori for the harmonic balance is that in the circuit there are only periodic or quasi-periodic waveform. For example amplifiers and mixers are invevitably measured, at least initially, using continuous-wave excitation, which is periodic by its very nature. Oscillators are even self-excited with a periodic waveform. Thus we will focus exclusively in the harmonic balance simulations.
The harmonic balance is an algorithm that employs the space variables to analyze the circuit splitting the linear and nonlinear parts of the circuit.
In order to expose how the harmonic balance simulator works we can consider an arbitrary single-transistor circuit deembedded into its linear and nonlinear parts. This enables us to define a clear split between those parts of the circuit and analyzed in the frequency domain and those in time domain. The linear circuit contains any matching networks, the bias network, device parasitic, sources and so on. The nonlinear circuit contains only those elements within the device model whose value is a function of voltage and current. For instance, in the case of a MESFET, the nonlinear components might consist of the capacitance C_{GS} between the intrinsic gate and source, and C_{DS} between the intrinsic drain and source in parallel with the drain current source. These elements are all modelled by equations in which the current through them is a function of the applied voltages at the intrinsic device terminals, perhaps with a transit time delay included, as well as possibly the derivatives of these voltages. The inclusion of a time delay means that the model need not necessarily be quasi-static. These voltages, $v_1(t)$ and $v_2(t)$ in the figure, are taken as the state variables of the system. In the more general case of multiple transistors, the numbers of nodes and branches joining the linear and nonlinear subnetworks will extend to some higher node number N rather than 2 as shown. This is a convenient representation, because a model of the form $C=C(v)$, perhaps derived from the semiconductor physics of the device, or from empirical observation a number of bias points, is assumed to imply $C[t]=C(v[t])$. The current is then $i(t)=C(v(t))dv(t)/dt$. More generally, the nonlinear currents in the branches joining the linear and nonlinear subnetworks can be modelled by equations on the form

$$i_j(t) = N(v_1(t), v_2(t), \ldots\ldots, v_N(t)) \quad J = 1 \ldots N \tag{1}$$

Where J is the branch number in consideration. We allow differentiation and integration in the equation to account for currents in nonlinear capacitors and inductors. N is any general nonlinear function of the state variables, typical functions are power series, exponential, and special functions as Bessel functions.

In the case of Figure 1, the linear circuit can be modelled as a four-port network, in which two ports connect the linear and nonlinear subnetworks, and the other two are for applied bias and RF voltages. More generally, there will be *N+M* ports, where *M* is the number of ports at which external sources are added. Such a liner network can be analyzed at each of the *K* harmonic components present in the circuit. The relationship between the linear applied voltages and the resulting currents at those ports is then an *N+M* port admittance matrix at each frequency. Generally then, we can calculate an augmented admittance matrix at each frequency $k\omega_0$ where

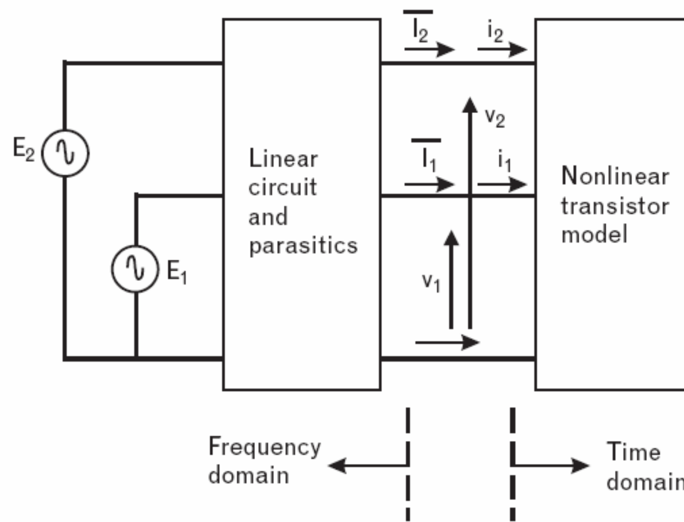


Figure 2- A simple transistor circuit deembedded into its linear and nonlinear parts.

$$\begin{pmatrix} \bar{I}_1(k\omega_0) \\ \cdots \\ \bar{I}_N(k\omega_0) \end{pmatrix} = \begin{pmatrix} Y_{11}(k\omega_0) & \cdots & Y_{1N+M}(k\omega_0) \\ \cdots & \cdots & \cdots \\ Y_{N1}(k\omega_0) & \cdots & Y_{NN+M}(k\omega_0) \end{pmatrix} \begin{pmatrix} V_1(k\omega_0) \\ \cdots \\ V_N(k\omega_0) \\ E_1(k\omega_0) \\ \cdots \\ E_M(k\omega_0) \end{pmatrix} k = 0,1 \ldots k \qquad (2)$$

and the matrix is augmented from the normal square *NXN* admittance matrix to account for the additional *M* ports where external voltages are applied. Most linear simulators already calculate the admittance matrix of a circuit in this form, although we should note in passing that the dc case usually requires special attention since the disappearance of inductors and capacitors at dc can result in singular matrices with many zero and infinite elements. The matrix in (2) is also usually required to be the definite admittance matrix, meaning that the ports connecting the linear and nonlinear subnetworks will not always be defined with one terminal as ground. Returning now to Figure 2 for simplicity, we can describe the general process step for the principles of harmonic balance. These are illustrated in the figure below. For simplicity let us assume that there are up to four harmonics of the fundamental frequency present in the circuit, generated by either an applied voltage or from harmonics created through distortion in the device. Establish initial guesses for the frequency

components of the state variables V_1 and V_2. If we let capitalized variables refer to phasor or frequency-domain quantities, then we establish initial guesses for

$$\begin{aligned} &V_1(0), V_1(\omega_0), V_1(2\omega_0), V_1(3\omega_0), V_1(4\omega_0) \\ &V_2(0), V_2(\omega_0), V_1(2\omega_0), V_2(3\omega_0), V_2(4\omega_0) \end{aligned} \quad (3)$$

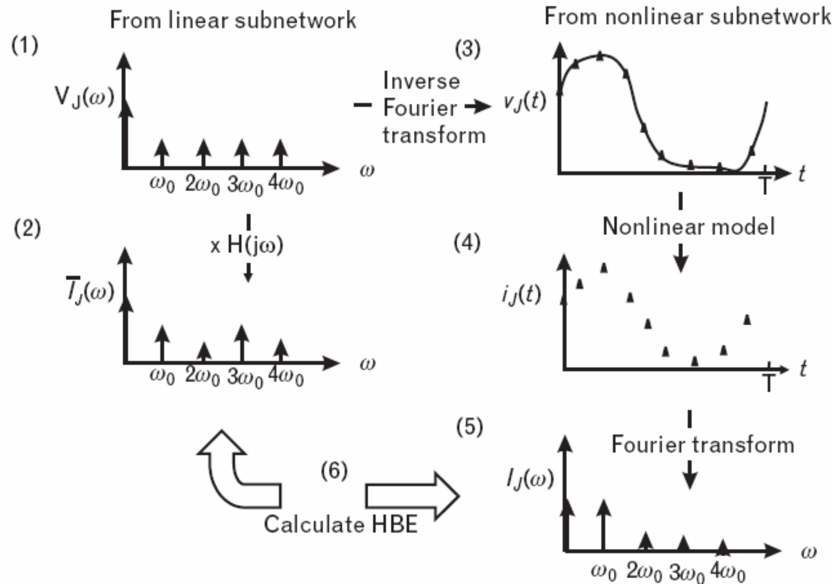


Figure 3- The harmonic balance steps process for the ciruit in Figure 2

These initial guesses should ideally correspond to the expected steady-state value of the variables. In practice, the dc values probably correspond to the expected bias conditions, while the RF value can be initially set to zero since they are not usually known a priori.
Referar now to the linear subnetwork only. Use the values above for the state variable together with the known applied source voltages and the definite admittance matrix to calculate the corresponding values of phasor current that flow into the linear subnetwork, that is,

$$\begin{aligned} &\bar{I}_1(0), \bar{I}_1(\omega_0), \bar{I}_1(2\omega_0), \bar{I}_1(3\omega_0), \bar{I}_1(4\omega_0) \\ &\bar{I}_2(0), \bar{I}_2(\omega_0), \bar{I}_1(2\omega_0), \bar{I}_2(3\omega_0), \bar{I}_2(4\omega_0) \end{aligned} \quad (3)$$

The overbar is simply used to indicate the current flowing into the linear network. Using an expression of the form

$$v_J(t) = \mathrm{Re} \sum_k V_J(k\omega_0) e^{jk\omega_0 t} \quad (4)$$

we can calculate the time-domain waveform corresponding to the two state variables v_1 and v_2 if required the derivatives and integrals of the state variables can also be calculated directly from (4) by differentiation or integration. Since the waveforms are periodic, we need to evaluate only at the Nyquist rate over one period. In our example with four harmonics, we calculate nine time samples of

Now referring only to the nonlinear subnetwork, substitute the state variables from v_1 and v_2 within one period T that is

$$\begin{gathered} v_1(T/9), v_1(2T/9)...v_1(T) \\ v_2(T/9), v_2(2T/9)...v_2(T) \end{gathered} \tag{5}$$

Now referring only to the nonlinear subnetwork, substitute the state variables from (5) and their derivatives or integrals into the nonlinear model, to yield values of nonlinear current $i_1(t)$ and $i_2(t)$that flow at the same time instants, that is,

$$\begin{gathered} i_1(T/9), i_1(2T/9)...i_1(T) \\ i_2(T/9), i_2(2T/9)...i_2(T) \end{gathered} \tag{6}$$

Using a discrete Fourier transform, extract the frequency content of the time samples of current in (5). Since there are nine time samples, we can extract a dc component and four harmonics:

$$\begin{gathered} I_1(0), I_1(\omega_0), I_1(2\omega_0), I_1(3\omega_0), I_1(4\omega_0) \\ I_2(0), I_2(\omega_0), I_2(2\omega_0), I_2(3\omega_0), I_2(4\omega_0) \end{gathered} \tag{7}$$

Now putting the linear and subnetworks together, Kirchoff's current law is applied at each branch and requires that

$$I_J(\omega) = \bar{I}_J(\omega) \quad J = 1,2 \tag{8}$$

At all frequency components. We cann calculate an error function comparing the components of current flowing into the two sub-networks as

$$HBE = \sum_{k=0}^{K} \left| I_1(k\omega_0) - \bar{I}_1(k\omega_0) \right|^2 + \sum_{k=0}^{K} \left| I_2(k\omega_0) - \bar{I}_2(k\omega_0) \right|^2 < \varepsilon \tag{9}$$

If we have reached a solution, then the current components at each branch and at every frequency component will be equal and opposite, and the harmonic balance error (HBE) will be zero. The method is called harmonic balance because the harmonics in the linear and nonlinear "sides" must be balanced each other out.

We return to step 1 and adjust the values of the state variables. The process steps above are successively continued until $HBE<\varepsilon$ and the procedure is said to have converged. ε is typically of the order of 10^{-6} or smaller.

Once convergence is obtained, the solution to the state variables has been determined and the (2) can be used to find the branch currents. The linear subnetwork has therefore been solved. Quantities such as the distortion power, dc power, and gain can all be founf through solution of the relevant currents and voltages within the linear subnetwork.

The harmonic balance procedure is thus an iterative procedure. Like all iterative procedures, there is no guarantee of convergence and even with today's "fail-proff" simulators certain circuits will have increasing value of HBE on successive iteration steps. The values of the state variables are usually adjusted using a quasi-Newton approach, in which slight adjustments are made in turn to each of the components of the state variables in (3), and the sensitivity of each of the resulting harmonic currents at each branch to that change can be calculated. In most cases, this involves the creation of

what is known as the Jacobian, which is a sensitivity matrix, and inverting that matrix. Since there are *(K+1)* unknown harmonic components at N ports, as the number of the devices or harmonics increases, the time to find a solution increase quite rapidly. Convergence can sometimes be achieved by adjusting some of the default parameters that control the harmonic balance engine. Increasing the number of harmonics to reduce the level of aliasing, decreasing the step size in the state variables between iterations, and sweeping the input power level from small-signal up to the desired large-signal level can all help approach the desired solution incrementally. The largest improvements in convergence, however, have come through mathematical tricks used within the simulators themselves. For instance using the logarithm of the base voltage as a state variable for the bipolar transistor, rather than the base voltage itself, can help improve the convergence since the base and the collector current then vary linearly with that state variable rather exponentially. As can be seen from close inspection of the steps above, the algorithm can be applied generally to any circuit whose driving function, and thus response is periodic, and whose nonlinearity may be modelled as a time domain expression of the chosen state variables. Most RF amplifiers, mixers, attenuators, and filters fall into this class of circuits, as well as many systems. In the case of mixers, where both RF and LO signal provide two (usually) nonharmonically related input fundamental frequencies, a two-dimensional Fourier transform is required to support all possible linear combinations of these frequencies that are created within the circuit, including the IF. The same is required in the case of an amplifier to simulate its third-order intermodulation response. Most simulators also allow for a third fundamental frequency input, which is required for two-tone determination of the mixer RF response in order to simulate its third-order intermodulation performances. For more information about intermodulation amplifiers and RF systems you can check the **[3].**

Once this point we have discussed the harmonic balance method for non autonomus circuits, with applied input signals. These input signals force the device into linear and nonlinear regimes that can be analyzed at known excitations and frequency.

In the case of autonomus circuits as oscillators , there is no applied RF signal and the frequency is initially indeterminate. Yet the solution is indeed periodic, thus should still be amenable to the harmonic balance approach. However, without a known frequency or excitation level, how can the state variables be driven to a steady-state value? There are a number of different solutions to this problem.

3.2 Oscillator analysis using probes

A probe is a voltage source with series impedance, or a current source with shunt impedance, that is inserted into a circuit in order to drive the circuit in a forced regime. Ideally, it is attached to a node between the oscillating device and its resonant load. The probe is defined by its amplitude and fundamental frequency, and is assumed to have zero phase. Standard harmonic balance can then be used to analyze the circuit when the probe is inserted at a convenient point to drive the circuit, since it force excitation. In order for the probe not perturb the steady-state solution of the circuit, the series impedance of the voltage probe is set to be infinite at all frequencies except the fundamental where it is set to zero. Then the voltage amplitude and frequency of the probe are adjusted so that at steady state the ratio of the probe current to its voltage equals zero at the fundamental. (The dual is true for the current probe).Imposing this constraint on the converged solution implies that the probe can be removed from the circuit without affecting the result. The steady-state conditions will occur at that point where the circuit sustains its own excitation equal to the probe voltage and frequency.

Introducing the probe has now introduced two additional state variables (the fundamental frequency probe voltage and the fundamental frequency itself) into the harmonic balance system of equations. Both are initially unknown and need to be assigned initial values. However, the number of state-space equations that provide the boundary conditions has also increased by two as well, since the real and imaginary parts of the ratio of probe current to voltage must equal zero at the solution

point. Thus, the system of equations remains square and can be solved using the same algorithm discussed earlier.
In actual implementation , the prove voltage and frequency are assigned initial values and an inner harmonic balance loop is first solved for the other state variables. A second outer optimization loop is then used to adjust the probe voltage and frequency until the ratio of probe current to voltage equals zero. The initial value for the probe voltage is sometimes found by a separate search for that value of voltage that forces the loop gain to be one, which indicate starting point is in the vicinity of oscillation. The nesting of two optimization loops is more demanding in terms of computer time but simpler to implement since its requires no modification to the harmonic balance engine for nonautonomous circuits. It also allows other variables to be associated with the probe and optimized as part of the outer optimization loop, such as a tuning voltage or a component value that yields a desired oscillation frequency. The probe type of analysis can also be extended to stability analysis of autonomous circuits.
Software CADs that employ the probe in order to test the oscillation is Ansoft®.
Tests based on probes are powerful and practical but does not give information or design insight to the designer when the test fails.

3.3Single transistor Oscillator analysis using reflection coefficients of the device and resonant load: The Start-Up.

These limits can be overcome if the oscillation condition are formulated as a stability constraint on the reflection coefficients, that allows the designer to understand in which way modify the design in order to meet the start-up. When this criteria is used it is possible to design an oscillator that meet the start-up automatically without use any probe to force the system. Simply the system has the start-up at the desired frequency and will evolves in the steady-state automatically.
Define a criteria for the analysis of oscillators by using the reflection coefficient presents a lot of problems, that for a lot of time and by many authors and designers have been trivialized without a good understanding bringing the scholars to make a great debate on this topic. This debate is well reported in literature but in the past no author has formalized the argument and dissolved the doubts in a consistent manner, rather has been defined an erroneous relationship that people know as the rule of thumb, that in many cases may fails. This condition is so well rooted in the background designers that the CAD for oscillators design of regarded as the golden rule.
In addition the state of the art in microwave oscillator design seems to have decreed the victory of CMOS technology and differential configurations hat allow to have ***I-Q signals*** (cross coupled configuration) very important in modern transceivers. For these class of oscillators the condition based on reflection coefficients may be bypassed because to guarantee the start-up are sufficient some arrangements that can be established without the use of the coefficients of reflection.
The reflection coefficients are used mostly for the design of planar circuits in GaAs. Obviously in this technology also it is possible to design an oscillator with cross coupled configuration, but often together to usage of single-pole antennas and for radars (military applications and weapons design, radios, aircraft etc…) the single transistor oscillators are favourites.
Then when will talk about this condition (***main topic of this thesis***) we will refers to the old schools of oscillator design that is ***no used in consumer electronics***.
The stability condition born by the observation that every oscillator circuit can be always sketched as showed in figure below.

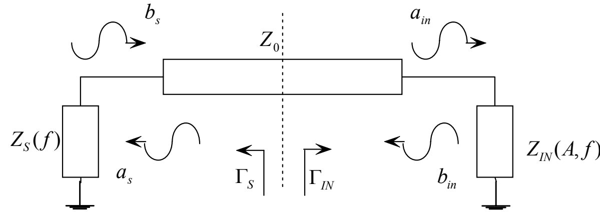


Figure 4- Basic schematic of a oscillator

These oscillators can be composed as a linear network for which the reflection coefficient is a function only of the frequency and an active non-linear network for which the reflection coefficient is a function of amplitude of the generated signal and frequency.
When a reflection coefficient is defines is defined also the associated impedance or admittance.

$$Z = \frac{1+\Gamma}{1-\Gamma} \quad Y = \frac{1-\Gamma}{1+\Gamma} \tag{10}$$

But the oscillator is a non-linear circuit therefore it is not possible to give a formal definition of impedance that is defined only for linear circuits. Actually there are a lot of definition of impedance and the impedance of an oscillator must be seen as characteristic I-V that physically represent the impedance of the circuit but is not in agreement with the formal definition. Elsewhere the reflection coefficient represent the system at the initial point of work (start-up) and making the assumptions that under the start-up the signal is small and does not excite the non linearities of the circuit the amplitude dependence can be neglected. In addition to use the S-parameters and the information of Γ in terms of impedances and admittances the Kurokawa assumption is adopted.
The paper of Kurokawa is more complex and focuses the attention on expressing stable build-up and start-up condition for the oscillator.

As showed by Kurokawa, a steady state free-running oscillation, with amplitude A_0 and frequency f_0, exists, if the nonlinear operating point defined by:

$$\begin{cases} R_{IN}(A_0, f_0) + R_S(f_0) = 0 \\ X_{IN}(A_0, f_0) + X_S(f_0) = 0 \end{cases} \tag{11}$$

is stable. In other words, if and only if the relationship

$$\left(\frac{\partial R_{IN}(A,f)}{\partial A}\right)\Bigg|_{A=A_0} \frac{dX_S}{df}\Bigg|_{f=f_0} - \left(\frac{\partial X_{IN}(A,f)}{\partial A}\right)\Bigg|_{A=A_0} \frac{dR_S}{df}\Bigg|_{f=f_0} > 0 \tag{12}$$

where $Z_{IN}(A,f)=R_{IN}(A,f)+jX_{IN}(A,f)$ represents the active network and $Z_S(f)$ the passive, is verified. The proof was given assuming multiple-resonant circuits in steady state almost sinusoidal oscillation.

Rohde et al. [1], have claimed that, under the hypothesis of high Q system, the steady-state oscillation conditions (1) are guaranteed if the relationship (2) has meet, by imposing operating conditions on the network which describes the small-signal operation of the system. In fact at the start-up the amplitude does not excites the nonlinearities of the active network that can then be represented by the input impedance $Z_{IN}(f)=R_{IN}(f)+jX_{IN}(f)$. Therefore the fundamental question of oscillator design lies in ensuring the start-up of the linearized circuit at a frequency f_X near to the steady state frequency f_0. The conditions for the oscillation start-up and build-up are consequently expressed in terms of impedance by:

$$\begin{cases} X_{IN}(f_X) + X_S(f_X) = 0 \\ R_{IN}(f_X) + R_S(f_X) < 0 \\ \dfrac{\partial}{\partial f}\left[X_{IN}(f) + X_S(f)\right]_{f=f_X} > 0 \end{cases} \qquad (13)$$

These results were achieved by studying the systems that can be represented as a series connection of one-port circuits. For the duals systems, namely those that can be represented as a parallel connection of one-port circuits, the start-up conditions must be given in terms of admittances by

$$\begin{cases} B_{IN}(f_X) + B_S(f_X) = 0 \\ G_{IN}(f_X) + G_S(f_X) < 0 \\ \dfrac{\partial}{\partial f}\left[B_{IN}(f) + B_S(f)\right]_{f=f_X} > 0 \end{cases} \qquad (14)$$

In both relationships sets, the inequality on the real parts must be limited. Indeed a unstable focus in the phase portrait, becomes a unstable node if the correspondent natural frequencies of the circuit are two real positive (Fig. 5), while only near a unstable focus the qualitative behavior of the system correspond to a growing sinusoidal oscillation.

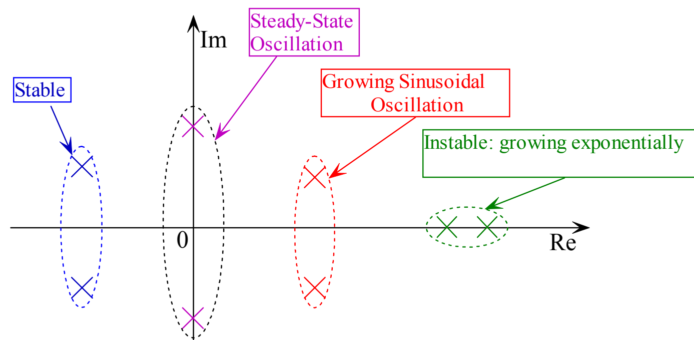


Figure 5- Qualitative behavior of the natural frequencies

As an example let's consider the simplest RLC circuit showed in Fig. 3,

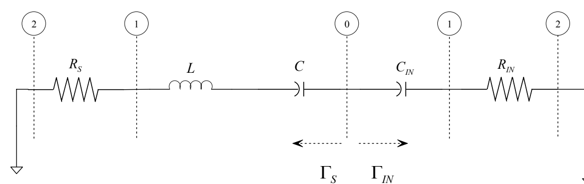


Figure 6 – The simplest circuit able to satisfy the start-up and build-up conditions.

where the natural frequencies can be easily derived from the transfer function that describes the circuit excited by voltage series generator. If the negative resistance exceeds a limit value, the complex pole pair still stays in the Gauss positive half-plane but become two real distinct. By expressing the equivalent capacitance of the circuit with C_{eq} the boundary negative resistance value is given by:

$$\left|R_{IN}(f_X) + R_S(f_X)\right| < 2\sqrt{\frac{L}{C_{eq}}} \qquad (15)$$

This work aims at giving the start-up and build-up conditions expressed by (13) or (14), only in terms of power waves, incident (a_s,, a_{in}) and reflected (b_{in}, b_s), that is of reflection coefficients ($\Gamma_{IN}=b_{in}/a_{in}$ and $\Gamma_S=b_s/a_s$) and characteristic impedance (Z_0). In literature, the most common condition in reflection coefficients terms is:

$$\begin{cases} \left|\Gamma_S \Gamma_{IN}\right| > 1 \\ \angle\Gamma_S + \angle\Gamma_{IN} = 0 \end{cases} \tag{16}$$

that is intuitively attractive but normally not true because, as show, it can predicts the oscillator start-up only if the measuring section is selected so as to include all reactive elements in one side.

Some authors have already made known that relationship (16) cannot predict the instability or stability in all cases. Indeed the product $\Gamma_S(s)\Gamma_{IN}(s)$ is the open-loop transfer function of a generic feedback system, expressed in terms of power wave. So the stability of the closed loop system can be determined by the Nyquist stability criterion, namely through the analysis of the polar plot of $\Gamma_S(j\omega)\Gamma_{IN}(j\omega)$ for $-\infty<\omega<+\infty$ (Fig.7). Unfortunately the Nyquist method does not give any information about oscillation frequency.

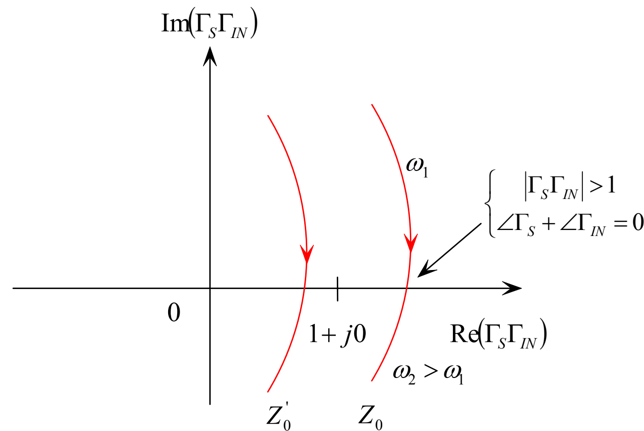


Figure 7 – $\Gamma_S\Gamma_{IN}$ behaviour versus frequency. For some characteristic impedance values, the function encircles in clockwise sense the $1+j0$ point.

The authors have evidenced that the product value also depends on Z_0 and its frequency behavior can looks like that of for the same system. Moreover, that the start-up conditions at f_X, (13) or (14), cannot be never represented by relationship (16), has been showed.

Based on this analysis they determine the geometrical meaning of the reflection coefficient product at the measuring section, a sufficient condition in terms of Γ able to ensure the oscillation without the need to use other validation tools, selection criteria to set Z_0 values to verify the tests effectiveness employed on CAD. At last they propose a simple way to calculate the resonant system loaded quality factor Q_L, at the start-up, and an example design for a low phase noise oscillator.

Some of these results have been presented in the previous paper [2]. However for the sake of completeness, they are included in this paper also.

The major problem for translation of start-up conditions in reflection coefficient terms is that a single value of Γ can be read as an impedance or admittance value simultaneously, in fact by definition we have:

$$z_{IN} = \frac{Z_{IN}}{Z_0} = \frac{1+\Gamma_{IN}}{1-\Gamma_{IN}} \qquad y_{IN} = Z_0 Y_{IN} = \frac{1-\Gamma_{IN}}{1+\Gamma_{IN}} \tag{17}$$

We show that by fixing a Γ_S or Γ_{IN} value, the set of corresponding impedances and admittances values that satisfy the startup conditions is represented by an arc of circumference on the plane of $\Gamma_S\Gamma_{IN}$ product. The equations are developed in terms of impedances then referring to the first two equations in (13) and results are extended to the admittances dual case as well.
Let's consider when the Γ_S value is fixed. In order to get the Γ_{IN} values that meet the startup condition, the second term of relationship (17) can be multiplied and divided by Γ_S. Then by separating real and imaginary parts, we have:

$$r_{IN} + jx_{IN} = \frac{\mathrm{Re}(\Gamma_S) + j\,\mathrm{Im}(\Gamma_S) + \mathrm{Re}(\Gamma_S\Gamma_{IN}) + j\,\mathrm{Im}(\Gamma_S\Gamma_{IN})}{\mathrm{Re}(\Gamma_S) + j\,\mathrm{Im}(\Gamma_S) - \mathrm{Re}(\Gamma_{IN}) - j\,\mathrm{Im}(\Gamma_S\Gamma_{IN})} \quad (18)$$

The expression for the imaginary parts is obtained after rationalization of the previous expression

$$x_{IN} = \frac{2[\mathrm{Re}(\Gamma_S) + \mathrm{Im}(\Gamma_S\Gamma_{IN}) - \mathrm{Re}(\Gamma_S\Gamma_{IN})\mathrm{Im}(\Gamma_S)]}{[\mathrm{Re}(\Gamma_S) - \mathrm{Re}(\Gamma_{IN})]^2 + [\mathrm{Im}(\Gamma_S) - \mathrm{Im}(\Gamma_S\Gamma_{IN})]^2} \quad (19)$$

Now imposing $x_{IN} = -x_S$, on relationship (19) this can be rewritten as

$$\begin{aligned}&\mathrm{Re}^2(\Gamma_S\Gamma_{IN}) + \mathrm{Im}^2(\Gamma_S\Gamma_{IN}) - 2\left[\mathrm{Re}(\Gamma_S) + \frac{\mathrm{Im}(\Gamma_S)}{x_S}\right]\mathrm{Re}(\Gamma_S\Gamma_{IN}) \\ &-2\left[\mathrm{Im}(\Gamma_S) - \frac{\mathrm{Re}(\Gamma_S)}{x_S}\right]\mathrm{Im}(\Gamma_S\Gamma_{IN}) + |\Gamma_S|^2 = 0\end{aligned} \quad (20)$$

The (20) represents a circumference equation on the $\Gamma_S\Gamma_{IN}$ Gauss plane, whose centre and radius are given by:

$$\begin{cases} C\Gamma_S = \left[\mathrm{Re}(\Gamma_S) + \dfrac{\mathrm{Im}(\Gamma_S)}{x_S}; \mathrm{Im}(\Gamma_S) - \dfrac{\mathrm{Re}(\Gamma_S)}{x_S}\right] \\ \rho\Gamma_S = \dfrac{|\Gamma_S|}{|x_S|} \end{cases} \quad (21)$$

Note that this circumference represents all possible values that the $\Gamma_S \Gamma_{IN}$ product can assume for the fixed value of Γ_S, once the constraint for the imaginary part at the impedance represented by Γ_{IN} having been imposed. In fact for the selected Γ_S , the correspondent Γ_{IN} values stay on the x=-x_S circle of the Smith chart. Then, in order to draw the product circumference, r_{IN} must assume values between [+∞;-∞]. In this circumference, the instability arc is composed by r_{IN} values that meet relationship

$$r_{IN} + r_S < 0 \quad (22)$$

It is evident that for $\Gamma_{IN}=1/\Gamma_S$ the (22) is equal to zero and $\Gamma_S \Gamma_{IN}=1$. For $\Gamma_{IN} = 1$ (open circuit) $r_{IN}=\pm\infty$ and $\Gamma_S\Gamma_{IN}= \Gamma_S$. These two values identify two arcs of circumference (Fig. 11) but relationship (22) is represented from the one which does not crosses the real axis. In fact, it turns out that in this cross point $\Gamma_{IN} = \Gamma_S^*$ which corresponds to $\Gamma_S\Gamma_{IN} = |\Gamma_S|^2$ and $r_{IN} + r_S > 0$.

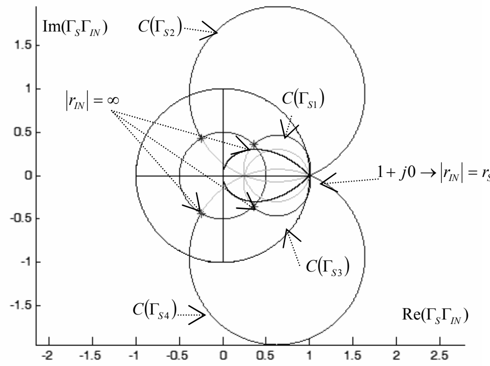


Figure 8- $\Gamma_S\Gamma_{IN}$ Circumferences versus r_{IN} drawn for the following Γ_S values: Γ_{S1}=(0.5;45°) ; Γ_{S2}=(0.5;120°) ; Γ_{S3}=(0.5; -45°); Γ_{S4}=(0.5;-120°).

From Fig. 8 it is clear that the relationship (22) can be satisfied for $|\Gamma_S\Gamma_{IN}|<>1$; this shows that relationship (16) does not represent the start-up conditions. It is possible to define the set of Γ_S values for which the circumference (20) has its center on the real axis and results completely enclosed in the unitary circle. This locus called "*Drop Contour*", is a curve on the Γ_S plane, (Fig. 9) defined by

$$\mathrm{Im}(C\Gamma_S) = 0 \Rightarrow \begin{cases} |\Gamma_S| = \dfrac{1-|\sin(\varphi_S)|}{\cos(\varphi_S)} \\ -\dfrac{\pi}{2} < \varphi_S < \dfrac{\pi}{2} \end{cases} \qquad (23)$$

Same expressions can also achieve in admittances terms (14), by replacing the reactance x_S with susceptance b_S in equations (20) and (21). Fig. 7 shows that, in this case $g_S = \pm\infty$ (short circuit) correspond to $\Gamma_{IN} = -1$ and $\Gamma_S\Gamma_{IN} = -\Gamma_S$. In this way we can conclude that for a fixed Γ_S value there are two different sets of Γ_{IN} values that satisfy respectively (3) and (4).

In order to draw the $\Gamma_S\,\Gamma_{IN}$ circumference, for a selected Γ_{IN} value, r_S (or g_S) varies in the range [+∞;-∞]. The equations are written in terms of x_{IN} or b_{IN} like in (10) and (11). For the new circumference, the arc that satisfies the relationship (12), is the one that intercepts the real axis (Fig. 8).

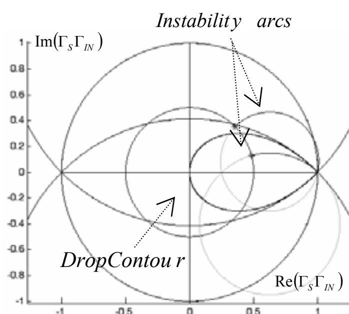


Figure 9 - $\Gamma_S\Gamma_{IN}$ instability arcs versus r_{IN}. When the Γ_S value belongs to the Drop Contour, the instability arc is all enclosed in the unitary circle.

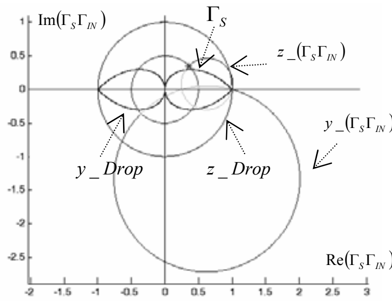


Figure 10 - $\Gamma_S\Gamma_{IN}$ circumferences for admittance ($y_\Gamma_S\Gamma_{IN}$) and impedance ($z_\Gamma_S\Gamma_{IN}$) corresponding to fixed Γ_S value

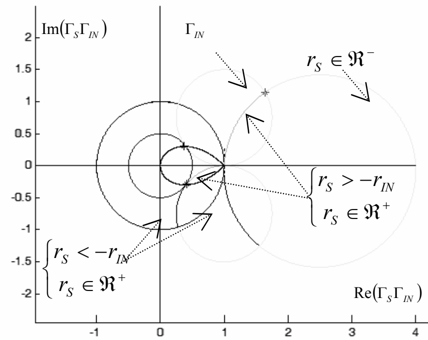


Figure 11 - $\Gamma_S\,\Gamma_{IN}$ circumference parameterized in r_S for a fixed Γ_{IN} value. The circumference is split in three arcs. The instability arc (dark grey) include the Γ_S values that satisfy relationship (22)

In this arc, only the points representing Γ_S values with $r_S > 0$ are able to meet relationship (22). The others Γ_S values ensuring the condition $x_S = -x_{IN}$ but with $r_S < 0$ lie on the light arcs.

To ensure a growing sinusoidal response, the R_{IN} or G_{IN} value must be limited in order to avoid that complex pole pair becomes real (Fig.5). Then the R_{IN} value should not be chosen much bigger than R_S (Fig. 5); therefore the Γ_{IN} should be set in the vicinity of the point for which $\Gamma_S\Gamma_{IN}=1$. This result is in agreement to the most popular rule of thumb presented by some textbooks.

For two Γ_S values in opposite sides compared to Drop Contour, the instability arcs are enclosed either completely or partially into the unitary circle (Fig. 9). This means that test based on Fig. 7 may fail for Γ_S values that fall into the Drop.

Since Z_S can be represented on the Smith chart as a Z_0 function, then corresponding Γ_S values, describe the loci that meet the following condition:

$$z_S = \frac{X_S}{R_S} = \frac{x_S}{r_S} = \pm const. \qquad (24)$$

These loci are actually two circumferences that contain the +1 and -1 points and have the center on the imaginary axis. Very important are those with radius $\rho=\sqrt{2}$ named "*Eye Contour*" being these tangent to the Drop Contours (Fig. 9), that have to coincide with the well known contour curves $Q=1$. The loci interceptions allow to define an area "Eye Area", that here emphasizes the Γ_S values lying outside the Drop for any Z_0 value.

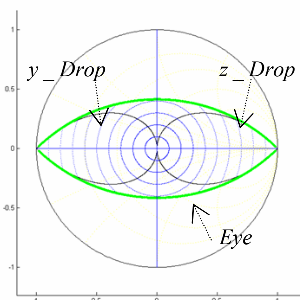
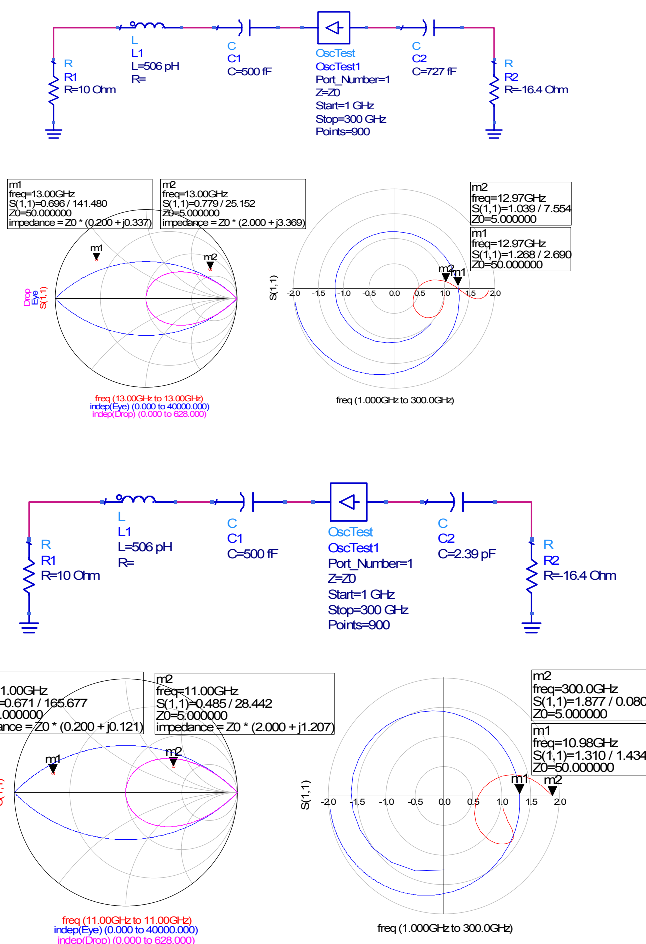


Figure 12- The thin lines represent all the Γ_S values that can belong to the Drop varying Z_0. The large contour is the Eye . The black lines are the Drops for impedances and admittances.

Then it is possible to express only in terms of reflection coefficients, and for a given frequency f_X , a sufficient condition for the startup of negative resistance oscillators.

$$
\begin{array}{lll}
\text{if} & \Gamma_S \notin \text{Eye Area} & \\
\text{then} & \left|\Gamma_S \Gamma_{IN}\right| > 1 & (25) \\
\text{and} & x_S(f_X) = -x_{IN}(f_X)\ (Z) \text{ or} & \forall\ Z_0 \\
 & b_S(f_X) = -b_{IN}(f_X)\ (Y) &
\end{array}
$$

For simple circuits like the one showed in Fig. 3 it is possible to simulate through a CAD system the failure for common oscillation test, when the Γ_S belongs in the areas previously described. In fact, if the Γ_S does not fall at the startup frequency into the area delimited by Eye Contour, common test succeed for any Z_0 values (Fig. 10), whereas the test becomes inaccurate for Z_0 values that bring the Γ_S into the Drop.



In conclusion we can affirm that relation based on reflection coefficient products does not express in anyway the start-up. The misleading of this condition is given by the fact that it is impossible to meet at the start-up the relationship $\Gamma_S \Gamma_{IN}=1$ because this means that the equality of resistive parts and the inverse for the imaginary parts. The condition presented by Rhode (13 and 14) must also be limited because if the negative resistance presents a too high value then the complex conjugate pair with positive real part degenerate into a purely real part pole pair and the system will continue to be unstable but the start-up conditions will be forbidden.

In order to use the start-up in reflection coefficient terms you can use the reflection coefficients products circles parameterized for a fixed R_{IN} and by varying R_S as represented in Fig. 11, this exclude the possible loosing of the start-up conditions due to the add of a too large value of negative resistance that can determine the degeneration of the complex conjugate pair in two distinct real poles. As an example you can use the tools developed in matlab product, that require the R_{IN} value and the resonance kind and for these values plot all the Γ_S values that ensure the generation of a complex conjugate pole pair.

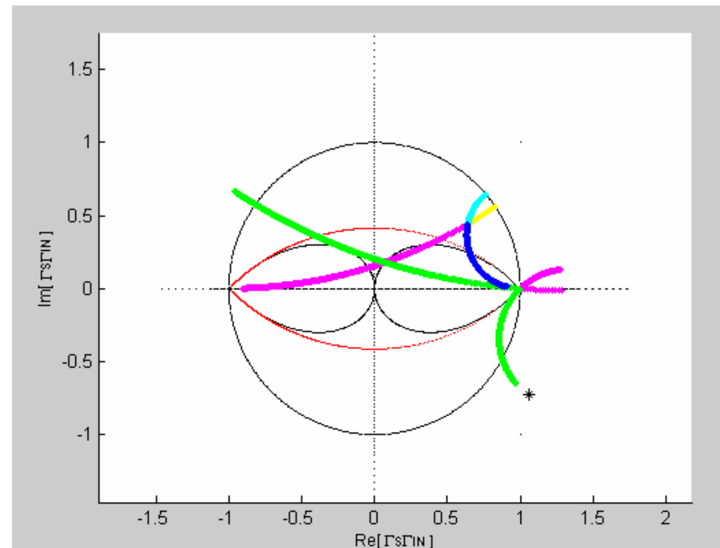


Figure 13- Reflection coefficients product trade for a fixed GIN and by varying RS. The magenta-yellow arc represent the instability arc(yellow) for the parallel resonator whereas the blu-cryian that for the series resonant circuit.

In every cases if you design by following this criteria then the product $\Gamma_S\Gamma_{IN}$ parameterized in frequency guarantee the effectiveness of the instability test based on Nyquist criterion.
It is possible to explain some examples in order to clarify the criterion exposed.

The topic is the following: the Nyquist criterion is the only way to test the instability of the oscillator considered as a linear system at small signal (under start-up conditions). The failure of the Nyquist criterion assert the impossibility to oscillate for the system (assert stability and the HB does not start) whereas the effectiveness does not guarantee the start-up of the oscillators, as previously showed that depends also by the value of characteristic impedance Z_0.
In almost all cases when the system generate a complex conjugate poles pair the Nyhquist criterion give good result if applied at the section of reflection coefficient product whereas it is necessary to change the section and consider all the system by a only side.
Let us present some examples to clarify the different situations in both series and parallel cases.
Lets consider a series *RLC* resonant circuit with resonance frequency of 10 GHz. (In the paragraph of resonant circuit design we will present a procedure to synthesize the resonant circuits starting by the desired resonance frequency and by the knowledge of the value of a reflection coefficient for another frequency).

Instead for these cases we will consider the classics relationship for the design of an RLC circuit.
The starting point are the resonant frequency and the Q.

$$\omega_r = \frac{1}{\sqrt{LC}}\ ;\ Q = \frac{2\pi f L}{R} = \frac{1}{\omega RC}\ .$$

The value of Q for the resonant circuit depends of both by the inductance and capacitance.
In general a good value of Q for modern RF processes is about 20.

The first step may be to calculate the value of the inductance starting by the Q. As described in the previous chapter the inductance is a complex circuit constituted by a parasite resistance in series. The value of R influence the value of the inductance.
The design depends by the technology and process.
Assuming the value of Q it is possible to determinate the values of R and L at a give frequency. The pair of values must be chosen in agreement with physical meanings.

$$Q = \frac{2\pi f L}{R} = QR = 2\pi f L \Rightarrow \frac{R}{L} = \frac{2\pi f}{Q}$$

Then you can proceed iteratively in order to find a value for R and a value for L hat have a correct means.
For the example we can find by imposing R=5Ω a value of L=1.59nH and these are reasonable values.
The values of the capacitance can be computed through

$$\omega_r = \frac{1}{\sqrt{LC}} \Rightarrow C = \frac{1}{L\omega_r^2} = 157\,fF$$

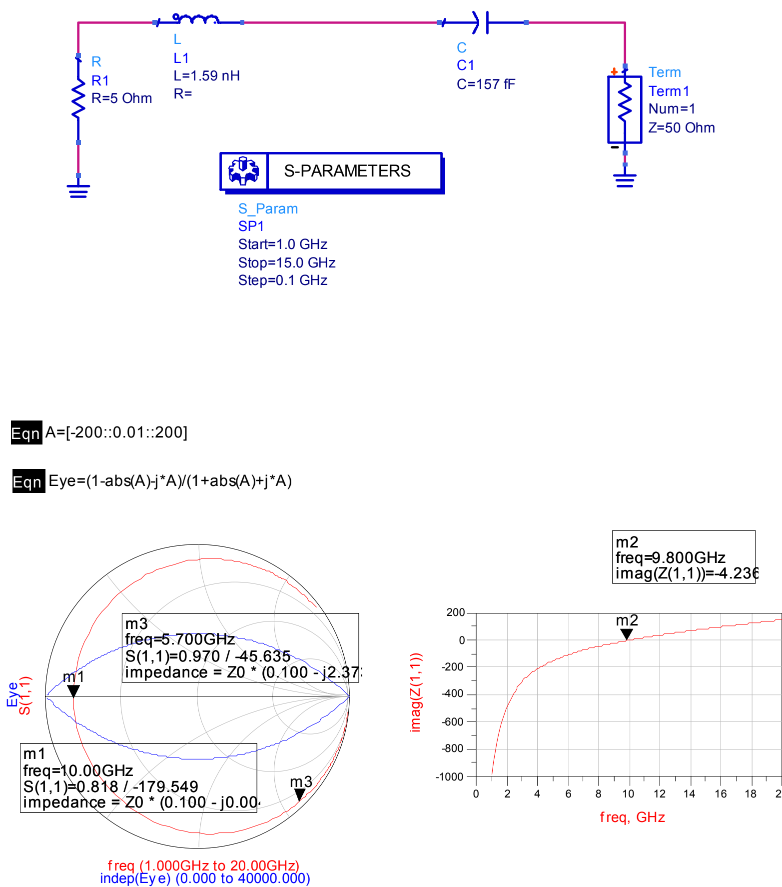


Figure 14 – A Simple RLC circuit with Q=20 and the simulations for the reflection coefficietn seen at the input port and for the imaginary part of impedance.

To test the effectiveness of the start-up conditions it is possible to build a series of circuits by using for the transistor a model constituted by an RC circuit with negative real part. In order do not

remove the Kurokawa assumptions to build an oscillator with robustness conditions for the build-up the resonance must be considered outside the active circuit. Theoretically it is possible to divide the resonance between the active circuit and passive network without considering a passive resonant network by simply resonating the imaginary part of active circuit for a given frequency with a proper element ,i.e. if the transistor have a capacitive behaviour in order to achieve resonance it is sufficient to add a inductance. This simple way to design the oscillator does not ensure good reliability performance because the transistor change its impedance with the time, in fact the core is non linear and time varying and this cause a change in the frequency of the generated signal without ensuring a good control.
Actually the resonant circuit is a dominant part and must be controlled by voltage, and the active source must only by ensuring to resonant circuit the storage of the loosing energy (resistance loss).
As showed by Kurokawa a robust start-up oscillation condition is guarantee if the active part does not have own resonance frequencies in the range of frequency of interest. Therefore the two functions for active and passive circuit must be well defined.
In order to guarantee this assumption the start-up of the oscillator must be near to the resonance frequency of resonant circuit unloaded.
We can examine for the circuit showed in figure (14) different RC circuits that give the start-up at different frequencies and test the instability of the circuit with the reflection coefficient product. In order to test the effectiveness of reflection coefficients product with different Z_0 values we can execute the same test for different Z_0 values especially for start-up conditions into the Eye contour.
If we look the current example we can implement the linear system as:

$f_r=10.20GHz$
$Z_S=0.100+j\cdot 0.050$
$Z_{IN}=-0.03-j\cdot 0.05$

The active circuit can be implemented through a series $R_{IN}=-15\Omega$ and a capacitance in series computed through

$$C = \frac{1}{\omega X Z_0} = 6.24\,pF$$

In this case the value of start-up frequency represent a Γ_S value into the Eye

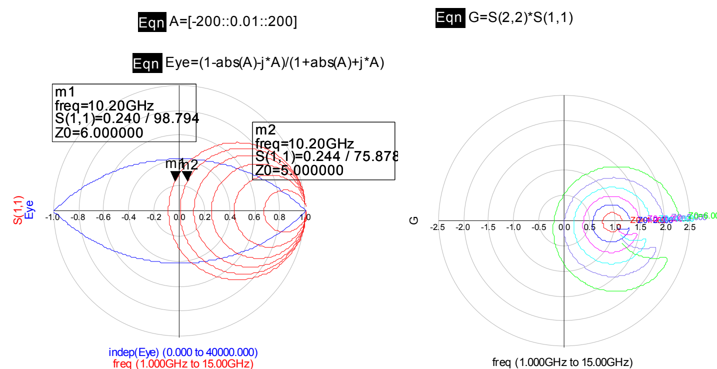


Figure 15

If we repeat the system for a frequency value outside the Drop for example $f_r=10.20GHz$ which correspond a $C=873fF$ the following results are achieved

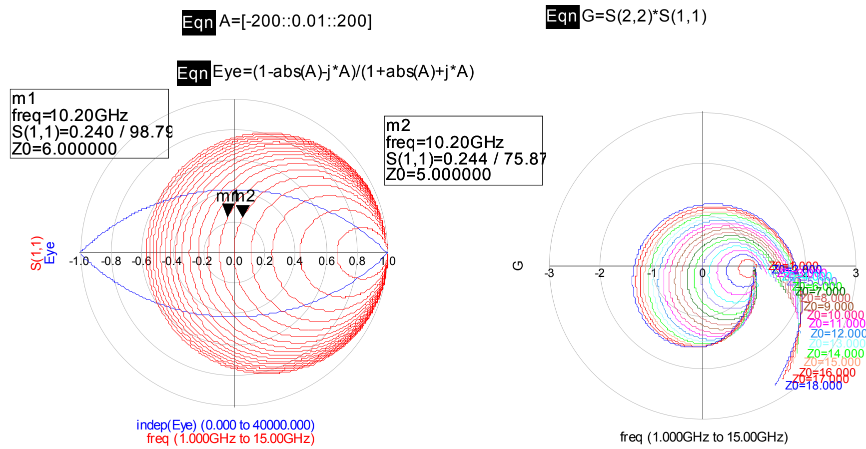


Figure 16

The test always converge then we can conclude that the Eye and the Drop represent the reflection coefficients with a different compression for real parts, and this show how the results presented in the paper are correct.

If we change section and consider the trade of the Γ parameterized in the frequency at a section that comprise all the system , by choosing a frequency into the Eye the test for the instability fails for values of Z_0 into the Drop.

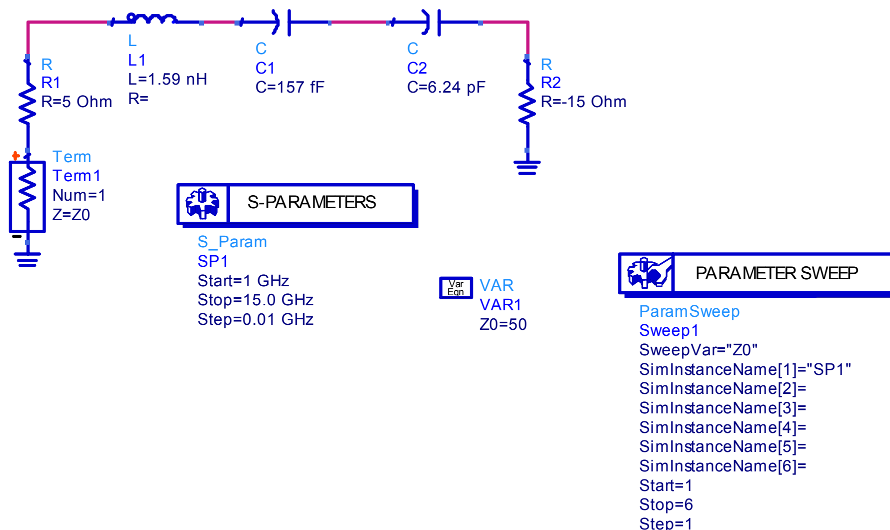


Figure 17

For values of Z_0 outside the Drop the test fail if the $-R_{IN}+R_S+Z_0\leq 0$ whereas the test converge. This result is not contemplated in the paper but must be taken into account.

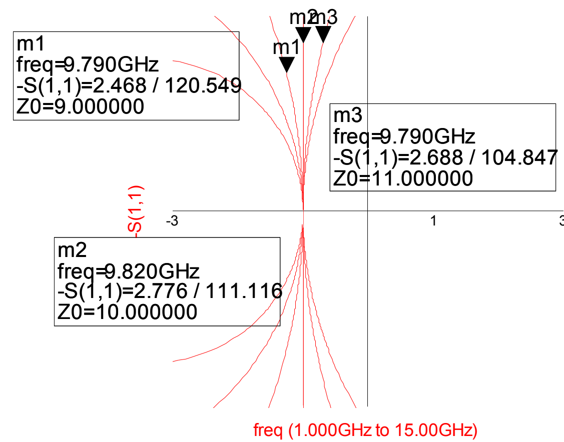
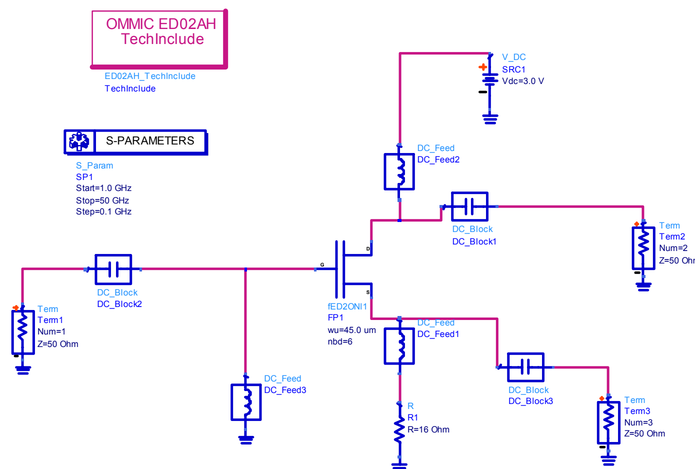


Figure 18

As a general result the Z_0 value does not must be chosen arbitrarily but by considering the values achievable with technology which are limited by dimension (50-100Ω) for modern GaAs technology processes. Note that this result (Fig.18) keep validity ever for the test also if the system has a start-up at a frequency for which the Γ_S is outside the Drop.
In general is very useless to model the oscillator splitting the system in two parts active circuit and resonant network by using for the transistor a significant model.
In order to be useless in microwave transistor oscillator design the transistor model must trace the behaviour of a small signal model for the reflection coefficient seen at the resonant port for a black box system constituted by a transistor, the bias network, and the loads networks used to generate the negative resistance.
For example we can consider the reflection coefficient seen at different ports for a simple biased transistor loaded with two terminals Z_0 values. In oscillator design a transistor can be used in both configurations common source and common gate the second when can be used is preferred to increase the frequency oscillation and to overcome the theoretical maximum oscillation frequency of $f_T/2$.
For an oscillator the transistor must be set to create a negative resistance by side of resonant circuit, and for this purpose it must be loaded with passive networks. The behaviour of the system is different if we look the gate or the source.



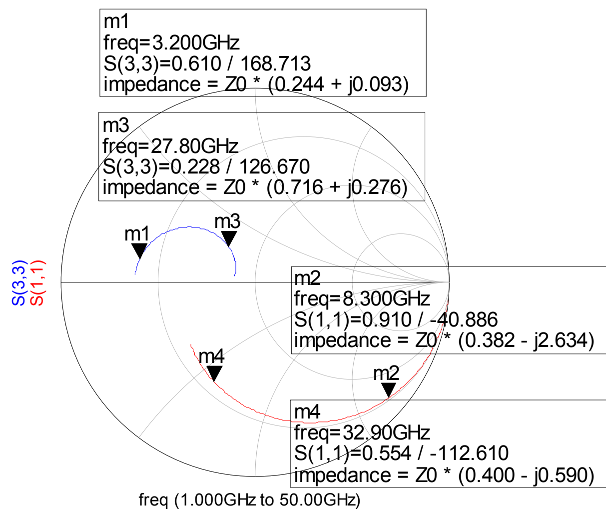
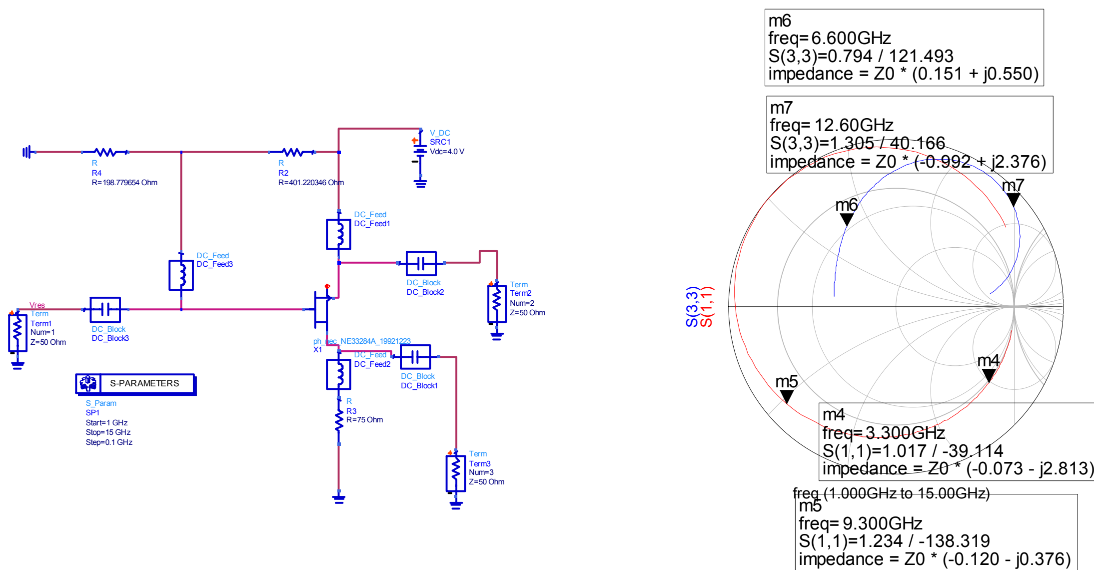


Figure 19 – Gate and source behaviour of a biased transistor.

This behaviour keep with different transistors and different technologies because every transistor can be represented as the same principal model even for the packaged models of some transistors there are also frequency ranges in which the present a negative resistance value without the need of the addition of passive networks as for example the following:



When the transistor is loaded with passive networks in two ports the behaviour frequency trend of Γ keep the same shape that may be distorted to bring the system outside or inside the unitary circle. The function of the system loaded is to give at the reflection coefficient a shape that does not varies with the operating conditions i.e. temperature etc…

When the transistor is loaded in the Γ may appear some curls that indicate some locals change of behaviour for example the trend of reflection coefficient may results inverted for example “grow with frequency for a small portion of frequency range” whereas before was decreasing with

frequency. The curls may play a role in the start-up CAD tests based on Nyquist criterion. In fact the Nyquist stability criterion assert that a linear system is unstable if the open loop gain ($\Gamma_S\Gamma_{IN}$ in the case of reflection coefficients used for microwave) encircle the 1+j0 point in clockwise sense.
To explain the importance of curls in the failure of stability Nyquist criterion we can look the following situations:

Case a) Γ_{IN} looked by gate with series resonant circuit and $|\Gamma_{IN}|>1$

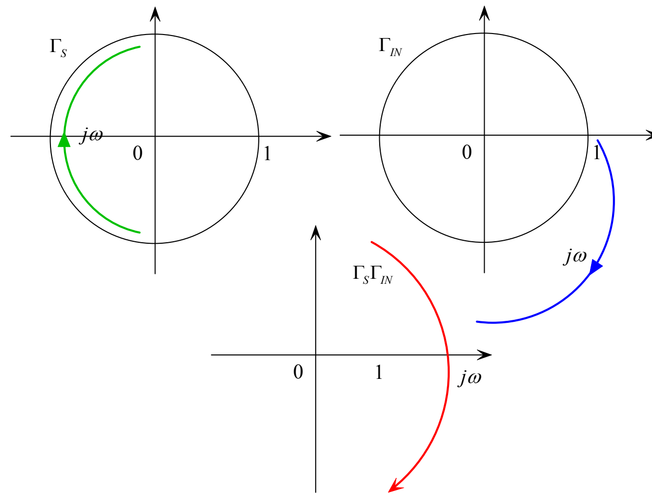


Figure 20 – Composition for reflection coefficients product in the frequency for resonant circuit and active circuit

In this case the reflection coefficient for series resonant circuit has a clockwise direction with the frequency and that of active circuit has a clockwise sense as well. Therefore does not exist the possibility to failure for the Nyquist instability criterion.

Case b) Γ_{IN} looked by source with series resonant circuit and $|\Gamma_{IN}|>1$

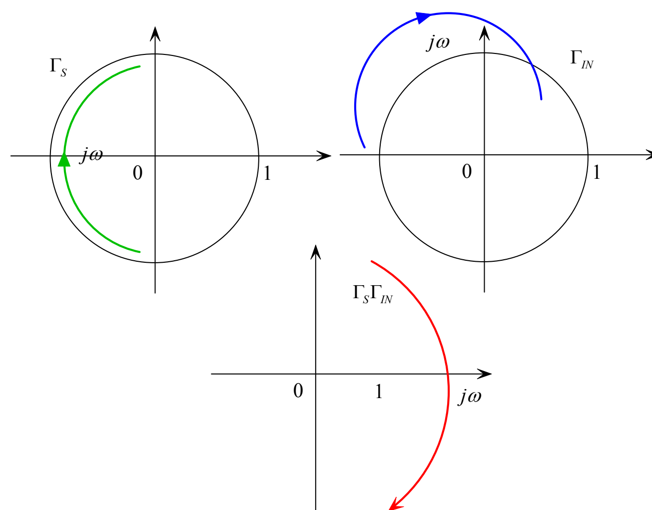


Figure 21

Case C) Γ_{IN} looked by gate with parallel resonant circuit and $|\Gamma_{IN}|>1$

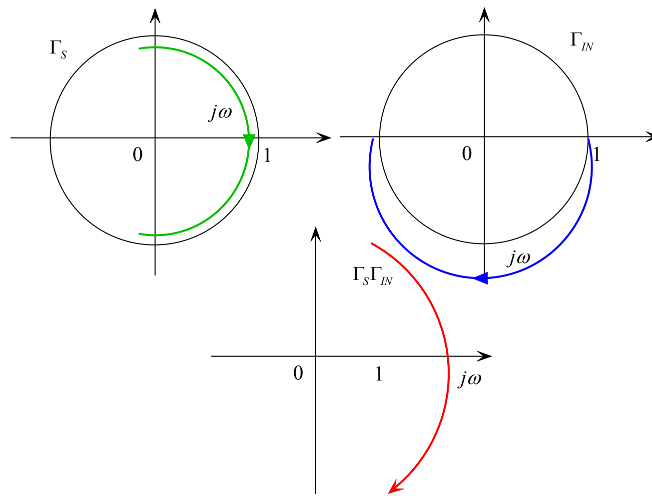


Figure 22

Case d) Γ_{IN} looked by source with parallel resonant circuit and $|\Gamma_{IN}|>1$

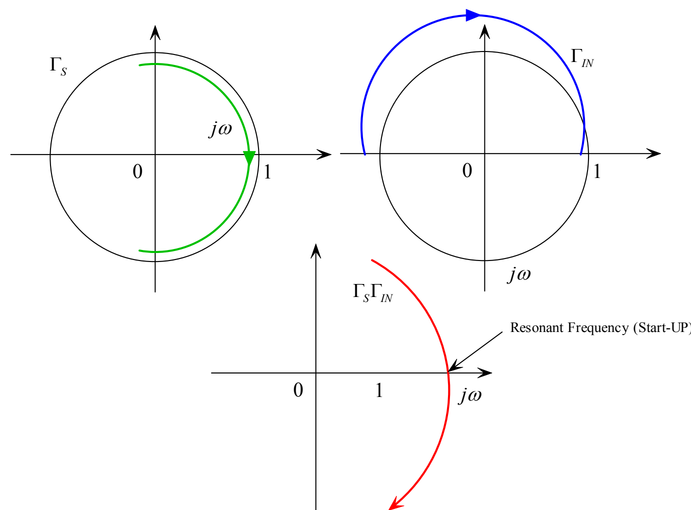


Figure 23

IN general we can conclude that in simplest cases there is not possible that the instability Nyquist based criterion may fails but it is possible that presence of curls may invert its trade with frequency if the Γ_{IN} will present a curl

Case d) Γ_{IN} with curl

In this case the reflection coefficient of resonant load has a clockwise trend with the frequency whereas the reflection coefficient for the Γ_{IN} presents a curl. In this case the curl indicate an inversion of frequency trend for the reflection coefficient from the clockwise sense to counter-clockwise sense and also a greater velocity of variation for the reflection coefficient. In this case the combination of reflection coefficient depends by velocity of variation of both Γ_S and Γ_{IN} and it is possible that the global product trend assumes the shape showed in the figure below. In this case the system may results stable.

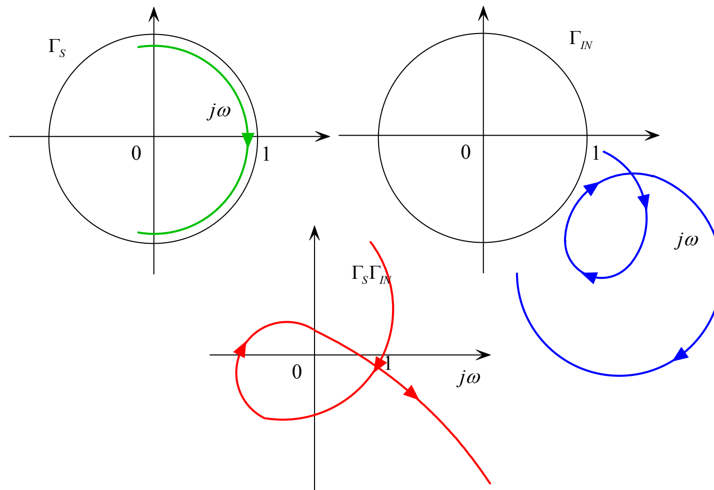


Figure 24 – Reflection coefficient product parameterized with frequency with curls . The curl may appear when a load resonant network is attached at the transistor

3.4 The active network small signal models and their combination with resonant circuits

From this study appear that whereas the passive network can be always modelled as a series or parallel RLC that active can be modelled as a series RC connection if you look the reflection coefficient by the Gate port.

We will show that for these systems the build up condition is always verified.

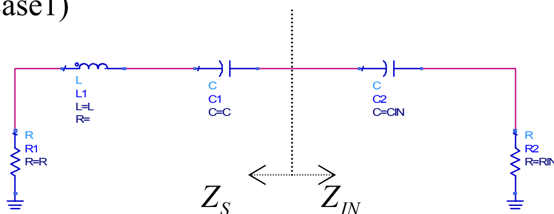
$$\frac{\partial}{\partial f}\left[X_{IN}(f)+X_S(f)\right]_{f=f_X}>0 \text{ or } \frac{\partial}{\partial f}\left[B_{IN}(f)+B_S(f)\right]_{f=f_X}>0$$

(These conditions means that the loaded system core+ resonator keep the resonance condition)

are always verified.

Let us examine the case of ***/R***

Case1)



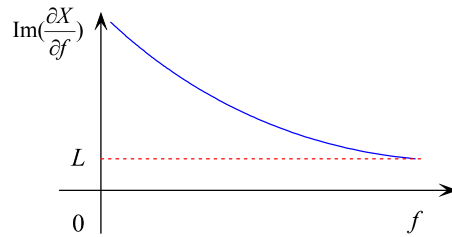
From resonant circuit side

$$Z_S=R+j\left(\omega L-\frac{1}{\omega C}\right)$$

$$X_S=\left(\omega L-\frac{1}{\omega C}\right)=\left(\frac{\omega^2 LC-1}{\omega C}\right) \qquad \lim_{\omega\to 0^+}\left(\frac{\omega^2 LC^2+C}{\omega^2 C^2}\right)=+\infty$$

$$\frac{\partial X_S}{\partial f}=\frac{(2\omega LC)\omega C-(\omega^2 LC-1)C}{\omega^2 C^2} \qquad \lim_{\omega\to +\infty}\left(\frac{\omega^2 LC^2+C}{\omega^2 C^2}\right)=L$$

$$=\frac{2\omega^2 LC^2-\omega^2 LC^2+C}{\omega^2 C^2}=\frac{\omega^2 LC^2+C}{\omega^2 C^2}$$

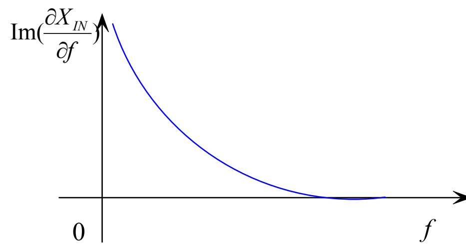


From active port side:

$$Z_{IN} = R_{IN} - j\frac{1}{\omega C_{IN}} \qquad \lim_{\omega \to 0^+} \frac{1}{\omega^2 C_{IN}} = +\infty$$

$$X_{IN} = -\frac{1}{\omega C_{IN}} \qquad \lim_{\omega \to +\infty} \frac{1}{\omega^2 C_{IN}} = 0^+$$

$$\frac{\partial X_{IN}}{\partial f} = \frac{C_{IN}}{\omega^2 C_{IN}^{\ 2}} = \frac{1}{\omega^2 C_{IN}}$$



Then

$$\frac{\partial}{\partial f}[X_S + X_{IN}] > 0$$
$$\forall f \in \Re^+$$

If $|R_{IN}|$ <>Z_0 it is possible to observe that the reflection coefficient have a different behaviour but the same means

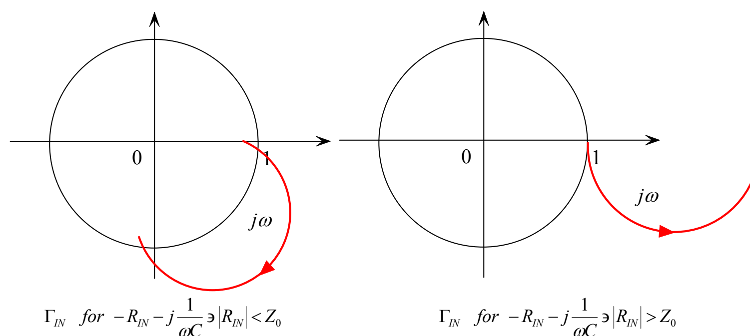


Figure 25 –Γ_{IN} trade versus frequency fora an active circuit for $|R_{IN}|$<>Z_0

The value of R_{IN} compared on Z_0 depends by the natures of the loads networks used. For example there are some loads network that may emphasise the first others the second.
In every cases the reflection coefficient representation is only a convenient way to represent microwave quantities but can bring the designer in misleading interpretations because the reflection coefficient can represent two distinct systems (impedance or admittance) with the same value and at

single frequency is meaningless. By varying the frequency the reflection coefficient can give some information and suggest insight for the design but only if the physical model has been decided at the beginning of considerations.

For these considerations
In order to exclude the possibility for the core to oscillate or to have a resonant frequency without the passive load circuit in the frequency range of interest when the resonant circuit is used at the gate port and the active system is modelled through a negative resistance with a series capacitance. Note that this modelling impose that the Γ_{IN} must be selected into the fourth quadrant of the polar chart. It is easy to show that for the behaviour that the transistor assumes in this area the core will be resonance less whereas if the Γ_{IN} will be chosen in others areas a free running oscillation may occur easily.

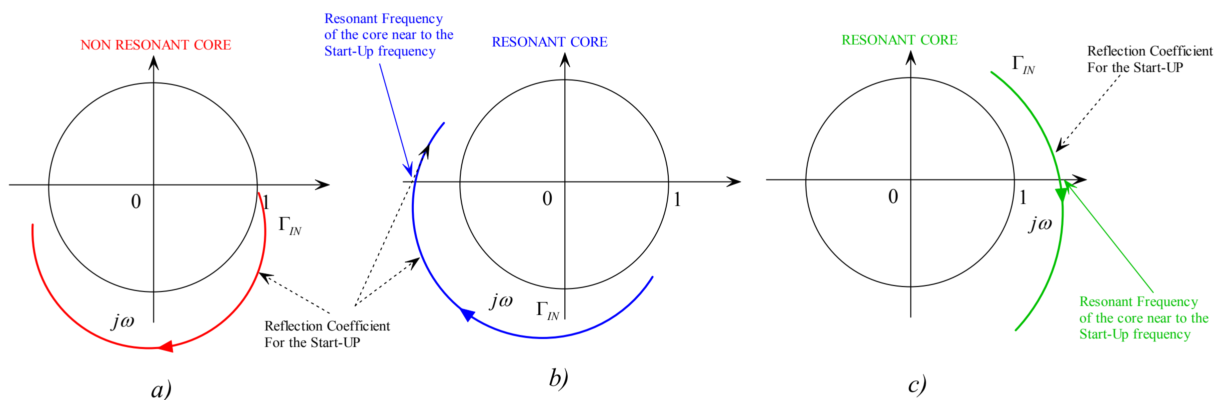


Figure 26 – The four cases of possible evolution of the Γ_{IN} for $|R_{IN}|<Z_0$ *a)* the GIN is chosen for frequency values that exclude the possibility of resonance frequencies *b) c)* in others three areas the GIN is chosen near to a possible resonant frequency.

By this study appears that the aforementioned architecture sketched in general in the following Figure suffers to these general limits:

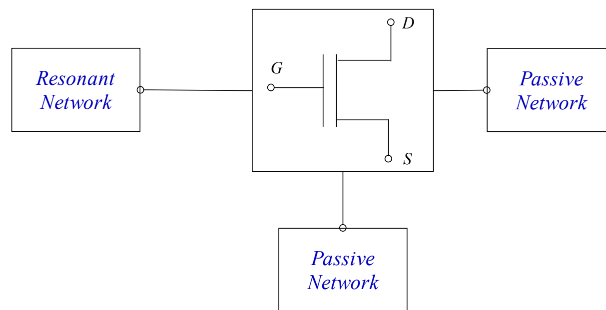


Figure 27 – General sketch for a negative resistance oscillator.

The Γ_{IN} will belongs to the fourth quadrant of the Smith chart if the resonator will be connected to the gate port whereas it will belongs to the third quadrant it is look by the source.
Then the two most popular configuration can be summarized in the following figure.

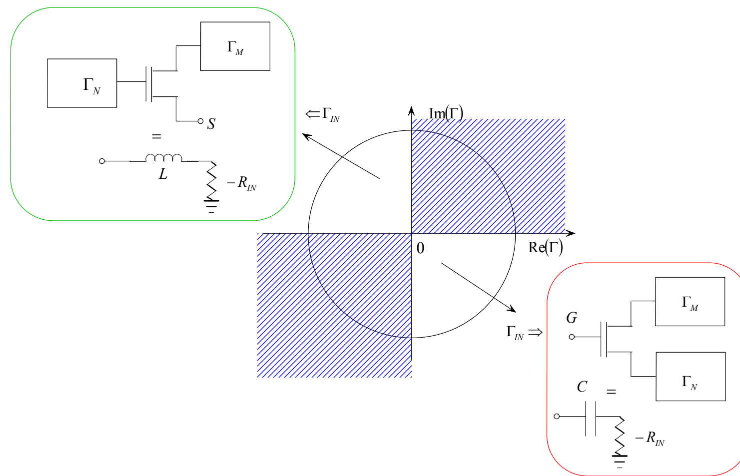


Figure 28 – Equivalent models for negative resistance circuit looking from Gate and Source.

If the Γ_{IN} belongs to this areas then the possibility of self oscillation are reduce in general but in some cases self resonance frequencies may appear depends by the trade of the reflection coefficient for the physical circuit.
This facts limits obviously the resonance frequencies range for creating a self resonant circuit.
For example if we wont build an oscillator connecting the resonant circuit in the gate port then we can model the transistor as a RC circuit and in general case the trend of Γ_{IN} parameterized in frequency will looks like that of fig.25(a). The addition of passive networks consent to distort the shape of the Γ_{IN} frequency trend to obtain a negative impedance.

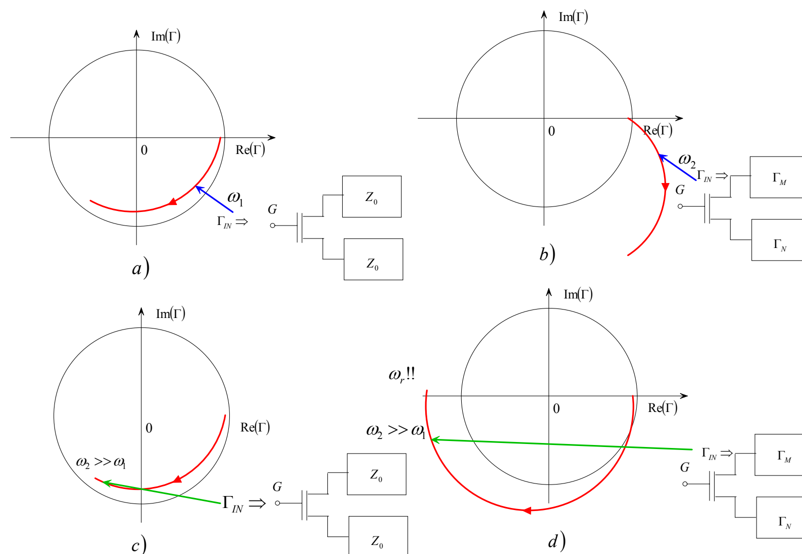


Figure 29 – Possible different situations for a negative resistance oscillator with a resonant circuit in the gate port.

Looking the previously figure if the transistor loaded with Z_0 does not show a negative resistance value (Figures a,c) at low frequency values the addition of loads (Γ_M and Γ_N)can distort the Γ_{IN} as showed in figure b) in this case we are sure that does not exist self resonance frequencies for the core near to the Start-Up frequency. If we design the passive networks so that the core shows a negative resistance value at ω_2 frequency then it is more probable the presence of a resonance frequency for the core. The core in the case of figure (d) will have a proper resonance frequency before the start-up frequency therefore the action of frequency determining circuit represented by a Γ_S will be ineffective. In the case of (d) it is possible in some cases to modify the Γ_{IN} trend in order to modify its trade and make it less of one, before the crossing point of X axis. It is important to note that this solution is very difficult to practice because the passive networks must be designed to

satisfy others constraints about transmission properties for example low pass network or high Q reactive element.
It is important to note also that the configuration with resonance network into the source port has the same behaviour but in opposite side. Obviously for a single transistor oscillator in independent way by the Γ_{IN} frequency trend the maximum oscillating frequency will be limited by the maximum oscillating transistor frequency that is about $f_T/2$.
Another very important disadvantage it consist in the great difficult to obtain a differential oscillator signal and (I/Q) signals used in modern communication systems. Moreover it is important to define in different manner the functions associated to two oscillator parts. These disadvantages make this class of oscillators antiquate and obsolete.
These oscillators make sense in the custom circuit design, focusing attention in MIC circuits and systems in which you can build the circuit by using a commercial substrate which microstrips may be modelled with proper instruments for metal cutting and the transistor may be bolt on in a specific point of the circuit. For complex transceiver structures the most common oscillator architecture preferred by state of the art is the cross-coupled also taking into account the advantages possible by the easy architecture on silicon substrate and the better phase noise performances showed in many MTT papers (es Andreani et all).

3.5 The State of the Art in the design of a single transistor negative impedance oscillators.

There are many books and references on state of the art on design of single transistor oscillator. The methodology cab be summarized as follow:
1) The first step it consist in chose the technology and the transistor properties, the class of operation and the bias network include DC-FEED and DC-BLOCK to foresee through a non linear CAD software which current it is necessary to provide the desired output power at 50Ω load. This step can be overcomed if you decide to put at the output a buffer stage, but this worsens the phase noise performances. As the transistor and its bias network are designed then the system can be studied as a two port network, by connecting the common port to ground and others ports to characteristic impedance Z_0 able to represent the characteristic impedance of lines to connect output and input.

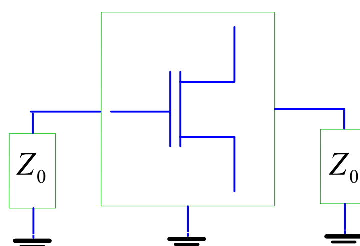


Figure 30 – a two port active network.

In this case it is easy to extract the S-matrix (scattering) that give information about the signal propagation of the device.

The scattering matrix is defined as

$[S]=\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$ where S_{11} indicate the measure of portion of signal that are reflected at port 1; S_{12} measure the signal quantity that arrive at port 2 and come back at port 1; S_{21} measure the signal that arrive at port 2 when is injected to port 1 and S_{22} have the same means to S_{11} but looking the system by the output port. The trade of $S_{12}S_{21}$ can be considered as small signal gain of the active network, or as transmission frequency behaviour if the network is passive.

Based on S parameters it is possible to give some definition useful for the design of active circuits by following the classic theory.

The first parameter is the Rollet Stability factor or K factor that give information for the instability or stability of the active circuit at a give frequency. If the active network must be employed as an amplifier the K factor must be smaller than one whereas when $K>1$ the system is potentially unstable or conditionally unstable and its stability depends on values of Γ representing the passive networks that you can connect at input or output port.

When $K<-1$ the system is unconditionally unstable.

The K factor is defined as:

$$K=\frac{1-|S_{11}|^2-|S_{22}|^2+|\Delta|^2}{2|S_{12}S_{21}|} \quad (26)$$

$$\Delta = S_{11}S_{22}-S_{12}S_{21}$$

The relationships that lies the reflection coefficient at one port when the other is loaded with a passive networks are respectively:

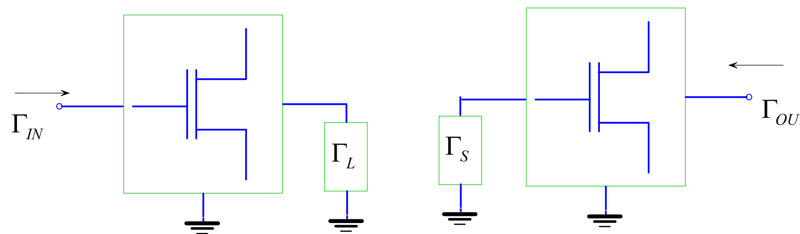


Figure 31

$$\Gamma_{IN}=S_{11}+\frac{S_{12}S_{21}\Gamma_L}{1-S_{22}\Gamma_L} \quad (27)$$

$$\Gamma_{OUT}=S_{22}+\frac{S_{12}S_{21}\Gamma_S}{1-S_{11}\Gamma_S} \quad (28)$$

Depends by loads the reflection coefficient seen at one port can assume a magnitude greatest or smallest than one. When $|\Gamma_{IN}|>1$ or $|\Gamma_{OUT}|>1$ the system is said to be potentially unstable, in these situations there is the presence of a negative resistance and then if the condition on resonance on reactances is satisfied the system can produce a complex conjugate poles pair in the right half-plane and then can oscillate. It this possible then to define the stability circles as circles that separate the value of Γ_S that make $|\Gamma_{IN}|<>1$ and the same for the Γ_L and $|\Gamma_{OUT}|$. To obtain a mathematical form of these circles it is possible to resolve the relation (27) or (28) by imposing $|\Gamma_{IN}|=1$.

$$1=S_{11}+\frac{S_{12}S_{21}\Gamma_L}{1-S_{22}\Gamma_L} \Rightarrow$$

It easy to extract the center and radius of circumference that can be obtained from that equation

$$C_{\Gamma_L} = \frac{\left(S_{11} - \Delta S_{22}^*\right)^*}{\left|S_{11}\right|^2 - \left|\Delta\right|^2} \quad , r_{\Gamma_L} = \frac{\left|S_{21} \cdot S_{12}\right|}{\left|S_{11}\right|^2 - \left|\Delta\right|^2} \quad (29)$$

At the same way the well known stability circle for Γ_S can be obtained. The discussion about the stability circles can be found in every microwave book.
At this time we can consider only that the instability circles can be founded only if the K-factor is less than one because this situation depict the possibility for the circuit to be conditionally unstable.
The state-of art procedure for SINGLE TRANSISTOR OSCILLATOR design can be depicted in the following figure.

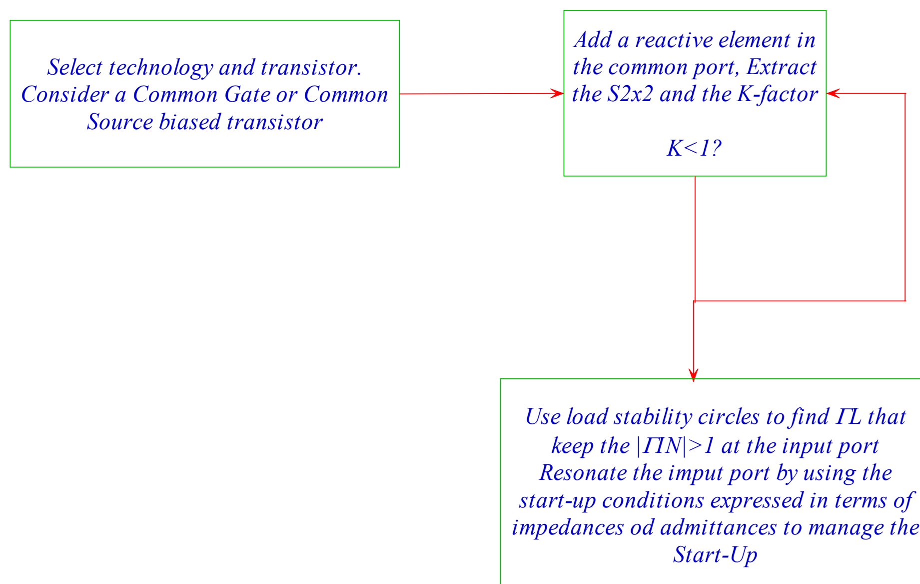


Figure 32

As shown in figure the schematics composed by few simple steps, but a lot of non linear simulations and large signal S-parameters simulations can be contemporary used to accomplish the design.
A good example of how can be complex this procedure can be found in [3-4]
This procedure suffers of the limits imposed by the impossibility to manage the output power because it is impossible to separate the Start-Up obtained through the choice of Γ_L by the manage of power made through the transmission behaviour of Γ_L network. We can consider an example performed through the ADS simulator.
Let's consider a PHEMT transistor biased by only a single resistor in the source. The drain can be connected directly to VDD because the Dc-FEED (RF-choke) must store the energy that allows to transistor to overcome the VDD limits. In these conditions the PHEMT work in C-class.
It is also possible to use the transistor in others operation class as A or B, it depends by position of bias point chosen by transcharacteristic point. As example we can consider the transistor ATF35176 from HP (transistor in package) and a network that allow to bias this transistor with a 10mA Drain-Source current

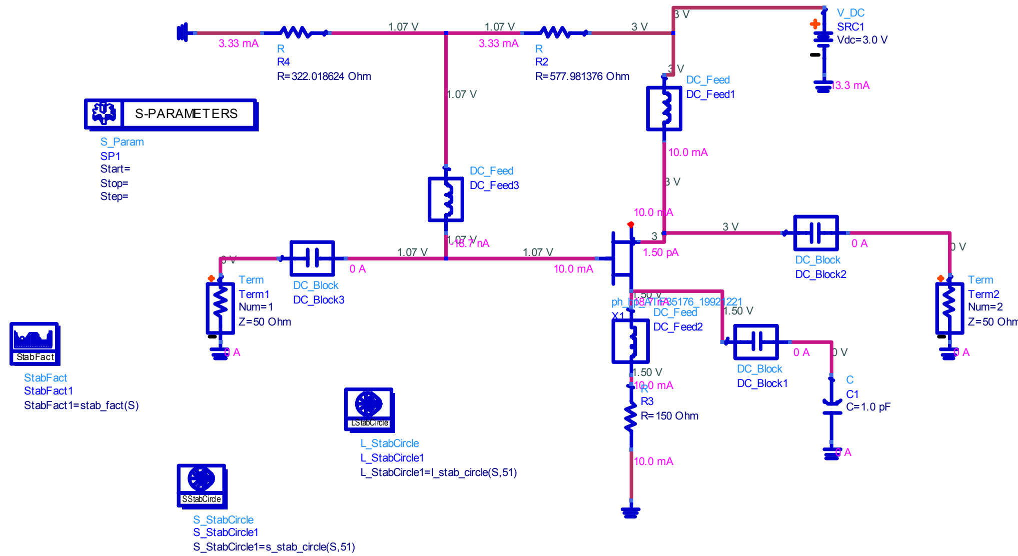


Figure 33

With S-parameter simulator we can extract the S-parameter @4GHz in order to obtain a 4GHz oscillator. It is impossible to say a priori which kind of passive element must be added to common port in order to destabilize the transistor, it may be both capacitive or inductive. You can try add at first time an inductive element. In this case an inductive reaction tends to stabilize the transistor, whereas a capacitive element make the K- factor smaller than one.
The result of S-parameter simulation are represented in the following table:

freq	S(1,1)	S(1,2)	S(2,1)	S(2,2)	StabFact1
4.000 GHz	1.093 - j0.749	0.107 + j0.110	-2.333 - j0.590	0.672 - j0.838	-0.831

The K- factor is smaller than one then the circuit is potentially unstable.
The next step is to determine the load network that allows the Γ_{IN} at resonant port greater than one. This may be done by plotting the stability circle to the load port in order to individuate the loads that can give the desired value of Γ_{IN}. The load applied to the output in order to achieve the desired Γ_{IN} does not must have reactive properties but must be resistive (near as much as possible to Z_0).
The stability circle for the load is:

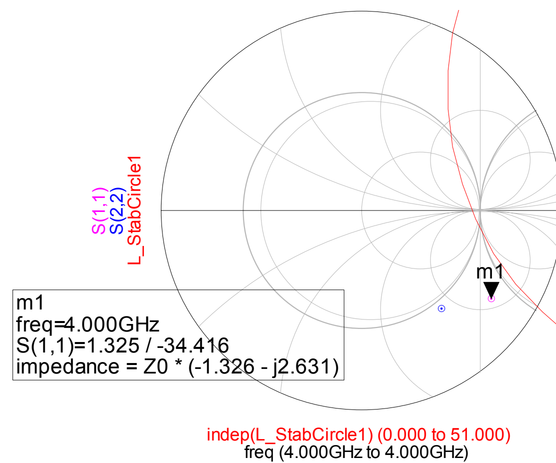


Figure 34

The Γ_{IN} has a magnitude greater than one and the circle does not contain the origin then the points expressed in terms of Γ outside the interception between he unitary circle and the load stability circle keep the Γ_{IN} with a magnitude greater than one. It is better to chose a network for which it is possible an implementation with a load pass transmission network then for example we can add a network which G is represented in the following figure

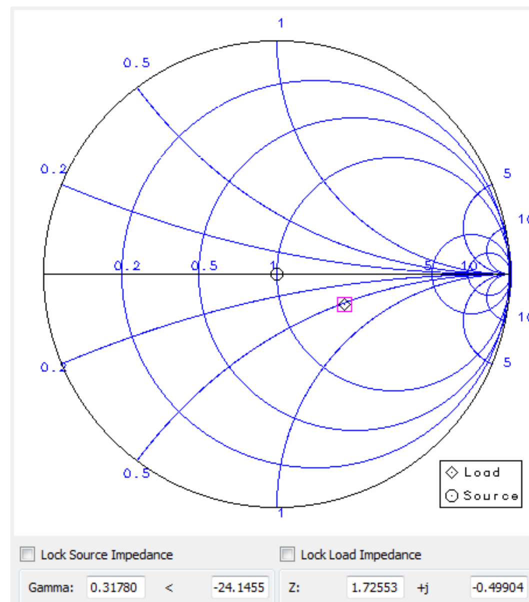


Figure 35

Then it is necessary to implement a low pass network that have a cut-off frequency greater of 4GHz, and presents a Γ which impedance is $Z=86.2765\ -j{\cdot}24.9520$. A generic circuit that can implement this reflection coefficient values having contemporary a low pass behaviour can be performed through a series inductance L and a shunt capacitance C.

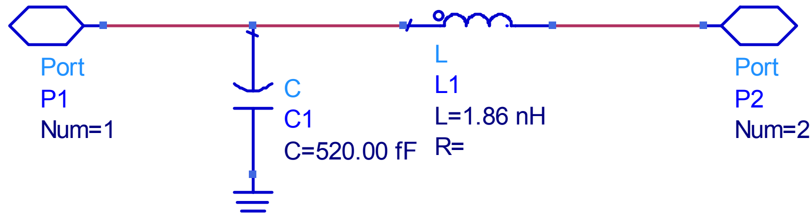


Figure 36

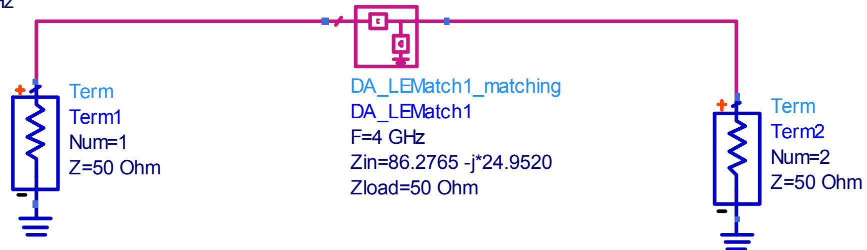


Figure 37

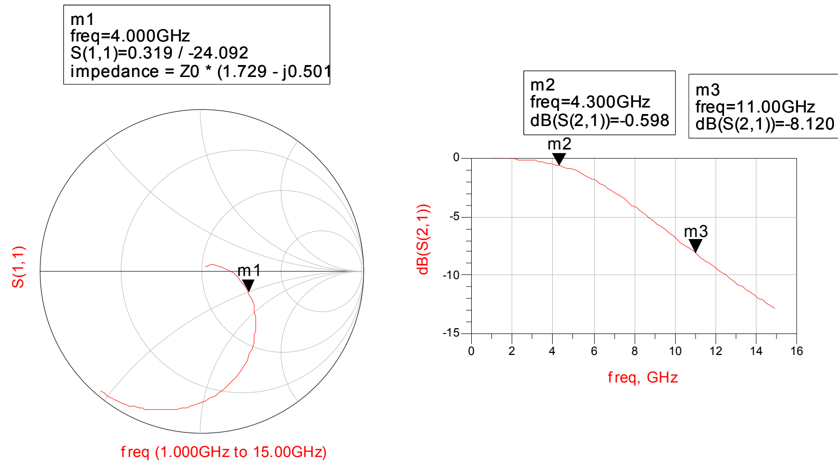


Figure 38

Then the core circuit and its simulation in the frequency domain is sketched below

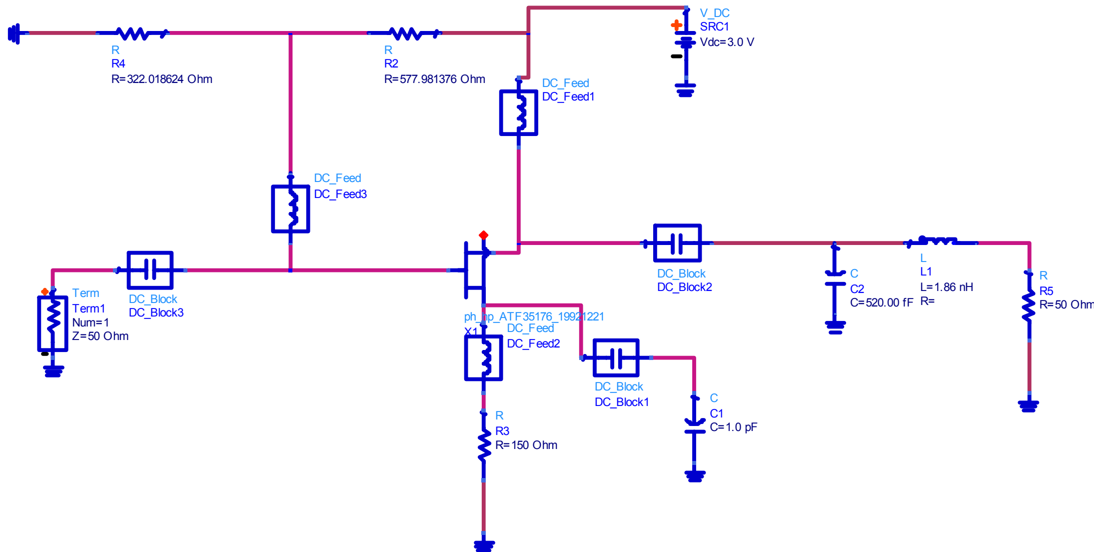


Figure 39- Oscillator Core.

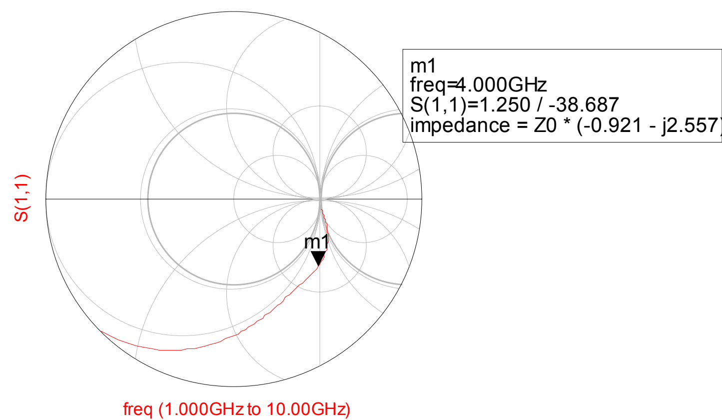


Figure 40 – Frequency trend of Γ_{IN} for the Core.

Now you must apply the Start-Up conditions in terms of impedances or admittances in order to design a frequency determining that allows to the system to generate a couple of complex conjugated poles pair.

In this case the impedance showed by the core is $Z_{IN}=R_{IN}+jX_{IN}$ but from position of Γ_{IN} showed by the figure implies that the resonance circuit that better satisfies the resonance condition for the system is a parallel resonant circuit then it will be better to use the Start-Up conditions in admittance terms then in order to obtain the Start-UP with $Y_{IN}=-0.1247 + 0.3462i$ you should use a parallel resonant circuit with admittance like this: $Y_{IN}=+0.0998 - 0.3462i$ where the positive conductance has been considered 20% less in absolute value compared to the negative reactance.

Then in order to complete the design of the oscillator it will be necessary to implement the resonant circuit, but in order to evaluate the performance it is possible also use a not frequency determining circuit but only a passive circuit constituted as well also by only one element. In this case the cirtuit may be

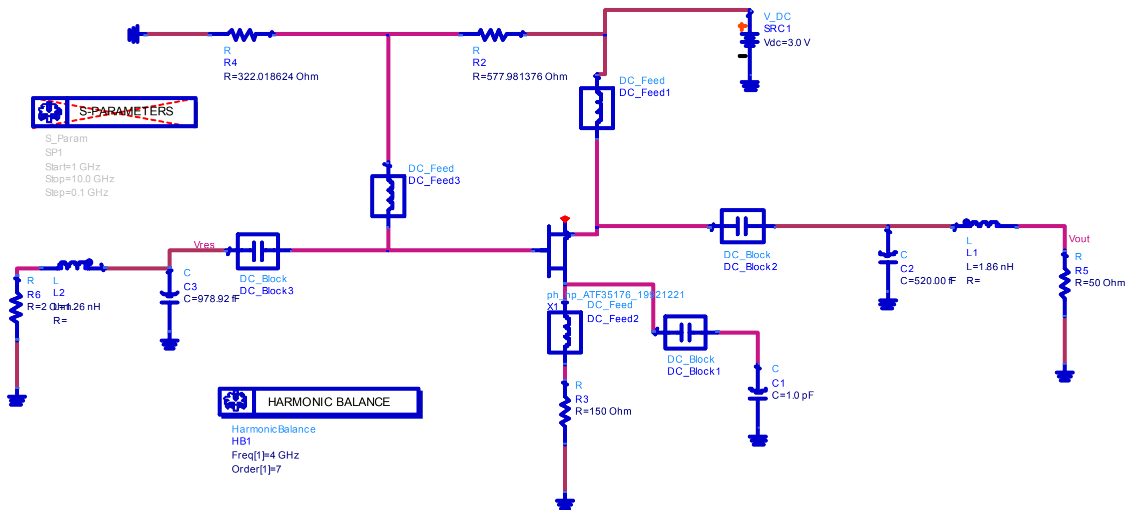


Figure 41

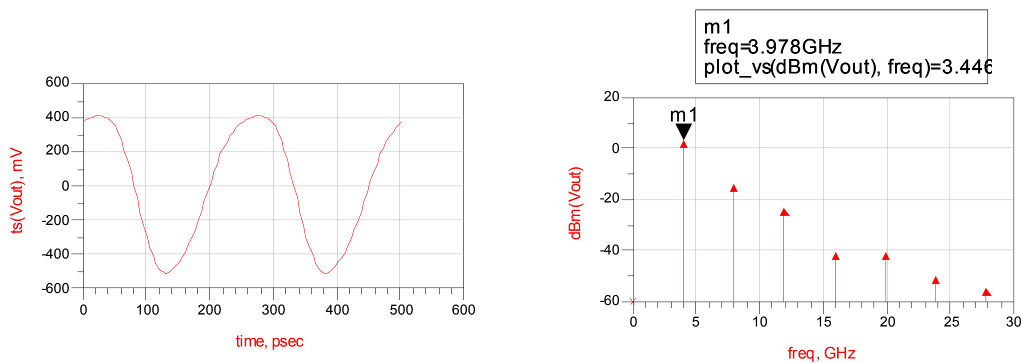


Figure 42

The output power cannot be estimated in any mode in the design and others performances as the loaded quality factor responsible of phase noise reduction for example cannot be estimated.

By changing the reaction port or only the reactive element or also only the output network can give better results. The problem is that improvements are only achieved through the experience of the designer and there are not a well defined procedure that can help him to obtain better performances.

In this thesis a way to overcome these limits and a new design procedure has been defined. This procedure will be discussed in the next chapter, whereas we can complete the discussion of performances of oscillator in first giving some examples of different architectures and by focusing the attention on the most important parameter to measure the goodness of an oscillator the “phase noise”.

It is important also to note that the possibility to obtain an oscillator with a positive reaction exists since the gain of the transistors ($S_{12}S_{21}$) is greater than one. Change configuration in order to pass from a common drain or source is different to common gate common drain configurations because the ($S_{12}S_{21}$) may change and in general if a configuration consent to obtain a different configuration may not allow an oscillator. If the gain is greater than one remember that the possibility of oscillation is limited by the cut off frequency of the transistor.

3.6 A different way to consider oscillators: design examples

Although there are not limits to apply the reflection coefficient in the design of microwave integrate circuits, (you can define the reflection coefficient at any frequency and for any application), there is another method to design oscillators, considering the classical methodology based on Barkhausen criterion. The Barkhausen criterion is the most popular for the design of oscillator and the mistakes on Start-Up for negative resistance oscillator are derived directly from an erroneous application of this criterion.
If we consider a feedback amplifier we can give the following description:

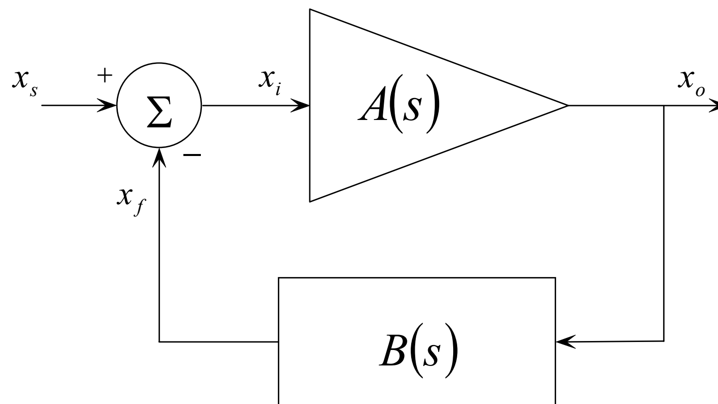


Figure 43 – General feedback architecture

The equations that define the system behaviour are:

$$x_o = A(s)x_i\,;\, x_f = B(s)x_o = A(s)B(s)x_i\,;\, x_s = x_i - x_f \Rightarrow \frac{x_o}{A(s)} - B(s)x_o = x_o\left(\frac{1}{A(s)} - B(s)\right)$$

$$\frac{x_s}{x_o} = \left(\frac{1}{A(s)} - B(s)\right) = \left(\frac{1 - A(s)B(s)}{A(s)}\right) \Rightarrow \frac{x_o}{x_s} = \left(\frac{A(s)}{1 - A(s)B(s)}\right)$$

The open loop gain is defined as *A(s)B(s)* and when *A(s)B(s)=1* the system can oscillate. The frequency for which this happens can be found imposing $s=j\omega$.

$$\frac{x_o}{x_s} = \left(\frac{A(j\omega)}{1 - A(j\omega)B(j\omega)}\right)$$

The system will be in a stationary state oscillation condition if the open loop gain expressed by $A(j\omega)B(j\omega)=1$. Since $A(j\omega)$ and $B(j\omega)$ are complex quantities their product will be one only if:

$$A(j\omega)B(j\omega) = 1 \Leftrightarrow \begin{cases} |A(j\omega)B(j\omega)| = 1 \\ \angle A(j\omega)B(j\omega) = 0 \end{cases}$$

That represent the Barkhausen criterion for oscillation. This situation correspond to a system that have a complex conjugate poles pair in the imaginary axix, and at this condition correspond also a steady state oscillation condition. The oscillation condition does not give any information about Start-Up. It is necessary that the system has a complex conjugate pole pair in the positive Gauss

half-plane to achieve the Start-Up as explained before. Then in a feedback system it is possible to affirm that for the Start-Up it is necessary to guarantee

$$\begin{cases} |A(j\omega)B(j\omega)| > 1 \\ \angle A(j\omega)B(j\omega) = 0 \end{cases} \qquad (27)$$

And it is easy to note as the condition is the same of that expressed in terms of reflection coefficient terms. The problem of difference is that in this case the condition is true but between A and B are not modelled as impedances or admittances. The kind of reaction is different and whereas for a feedback system relationship (27) it the only way to express the possibility of system to generate a positive complex conjugate pole pairs, in a reflection system the correspondent relationship is not true in every cases, but only when the condition reflect the Start-Up in impedances or admittances terms, but it is necessary to note that in these cases also the condition express only a instability properties but does not take into account the possibility of degeneration of complex poles pair into a purely real coupled of poles.
In every cases there are a lot of oscillators for which this condition may be applied.

a)- The most simple case is constituted by a ring oscillator.

A ring oscillator is an oscillator architecture for which a number of identical amplifiers stages is configured in a cascade with feedback in order to obtain a growing oscillating response. The simplest cases is constituted by inverter stages as showed by this example.

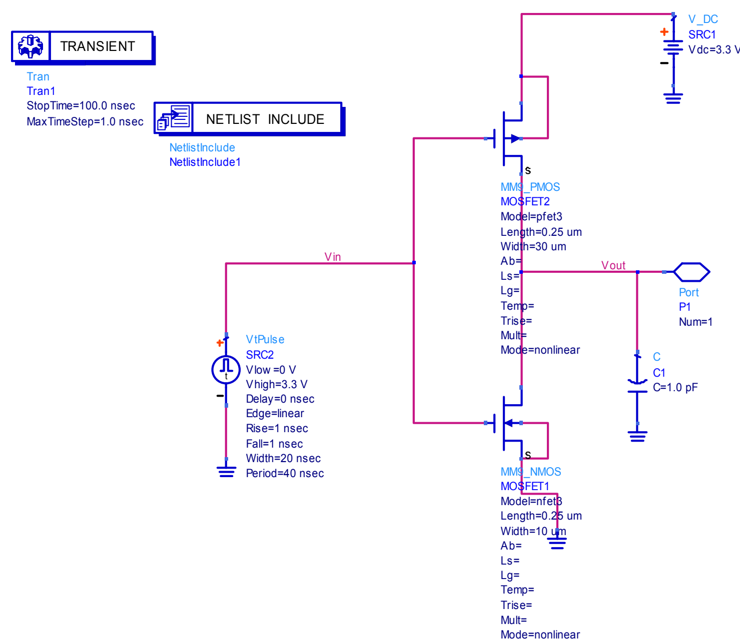


Figure 44 – a simple inverter circuit constituted by a NMOS and a PMOS.

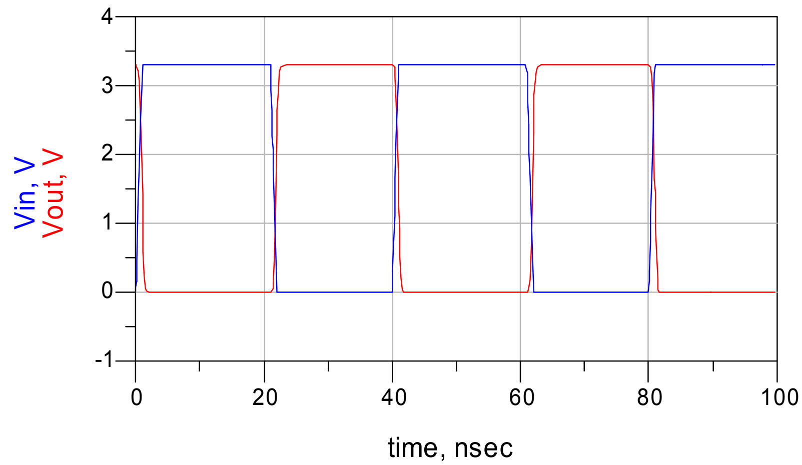


Figure 45 – transient simulation of inverter showed in Fig. 44.

For these kind of oscillator the frequency can be computed by knowing the delay time of the digital port.
For a single CMOS inverter the delay time is computed after considering the charge and discharge time of the capacitor.
The rise time ca be expressed as

$$\tau_R = \frac{4C}{K_p \left(\frac{W}{L} \right)_P V_{DD}}$$

where C is the load capacitance K_p is the intrinsic conductance of the p-MOS and depend by $\mu n, C_{OX}$.
The fall time has the same expression but for the N-MOS transistor

$$\tau_F = \frac{4C}{K_N \left(\frac{W}{L} \right)_N V_{DD}}$$

The frequency of oscillation is simply defined as

$$f_{OSC} = \frac{1}{N(\tau_R + \tau_F)} \rightarrow = \frac{1}{N\left(\dfrac{4C}{K_P\left(\dfrac{W}{L}\right)_P V_{DD}} + \dfrac{4C}{K_N\left(\dfrac{W}{L}\right)_N V_{DD}}\right)} = \frac{1}{N\left(\dfrac{4C\left(K_P\left(\dfrac{W}{L}\right)_P + K_N\left(\dfrac{W}{L}\right)_N\right)}{K_N K_P\left(\dfrac{W}{L}\right)_P\left(\dfrac{W}{L}\right)_N V_{DD}}\right)}$$

$$= \frac{K_N K_P\left(\dfrac{W}{L}\right)_P\left(\dfrac{W}{L}\right)_N V_{DD}}{N\left(4C\left(K_P\left(\dfrac{W}{L}\right)_P + K_N\left(\dfrac{W}{L}\right)_N\right)\right)}$$

From the file .lib it is possible to estimate the parameters for transistors N and P

$\mu_N \approx 112^{-6}$ and $\mu_P \approx 23^{-6}$

Then employing the previous relationship will be found, neglecting the parasite capacitances of transistors

$$f_{OSC} = \frac{1}{N(\tau_R + \tau_F)} \approx 281 MHz$$

Results obtained with simulators (Transient and HB) obviously are better accurate.
The next step consists in use the hierarchical design topology to built the full schematic of the oscillator
Then you can construct the full schematic for the oscillator

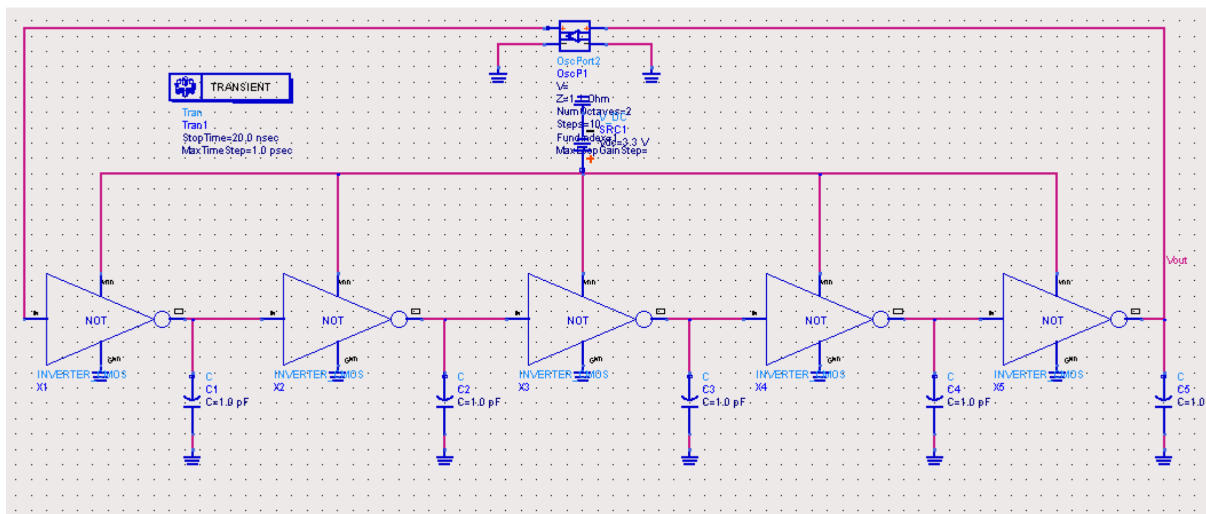


Figure 46

.

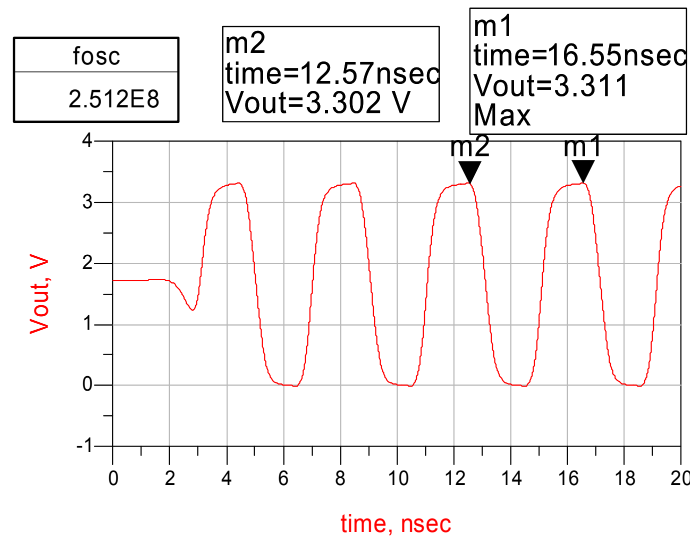


Figure 47 – Transient simulation of the oscillator

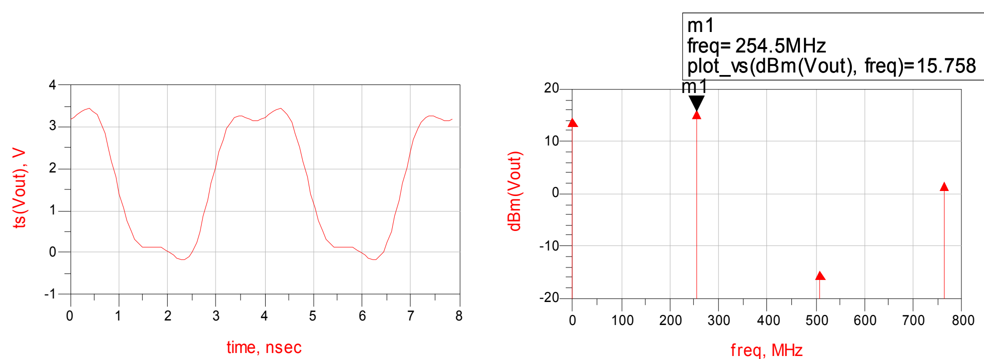


Figure 48 – HB simulation of the oscillator

Since the frequency is relatively low it is possible to analyze this circuit with both Transient and HB simulators. The Osc Port2 taken from the Simulation-HB library it is necessary to introduce initial conditions different from zero and to see the start-up. The output can be see plotting the *Vout* versus Time and show how the waveform for the circuit is a square wave. The output frequency can be computed through the equation the equation editor by defining the Eqn fosc . The results show a oscillator frequency of ***251MHz***

b)- The cross coupled oscillator.

A variation of a electronic oscillator can be built by using a cascade of single tuned amplifiers. In this case the ring oscillator can produce a sinusoidal waveform. For this circuital topology only two stages are necessary to produce a growing sinusoidal response. This is the base for most popular VCO architecture used in modern wireless transceiver, the cross-coupled architecture.

There are many reasons for which this architecture is preferred some of these are; great power of signal produced compared with DC-power dissipation, in fact the cross-coupled works in small time intervals as a bistable element, small chip are consumption, it is possible to implement with great success this architecture with CMOS technology that are cheapest compared with GaAs, moreover there not are problem for the Start-Up because the bistable circuit with reactive feedback does not suffers for complex conditions that are able to generate the Start-Up in terms of negative impedance system, and more important it has been demonstrated under both theoretical [5-6-7] and practical [8-9] points of view that this architecture allows to achieve better performances compared with the single transistor architecture. Another greater advantage consists in the possibility to implement quadrature (1-Q)VCO in very easy way, that are indispensable devices in modern wireless systems. Others advantages are derived by the possibility to use the silicon compared wit gallium arsenide technology. In fact the silicon give the possibility to use both PMOS and NMOS circuits and then to implement very good switches, that allows to obtain a wide range of tuning for the circuits. Moreover the tuning for circuits that employ the silicon is much faster compared with tuning achieved with GaAs as explained in [1], and this property is very required in modern systems. Finally the symmetrical architectures given by the differential configuration allows to obtain greater robustness compared with constructive tolerances and better performances for the chip and greater integration possibility with complex systems. If you read this thesis and you wont continue my work I suggest to you to get great importance to these class of oscillator because there are the state of the art of oscillator circuitry. The most common cross-coupled circuits are summarized in the figure below.

PMOS LC Oscillator   NMOS LC Oscillator   Improved NMOS LC Oscillator

It is important to note that the cross coupled structures is obtained considering a ring between the following circuits

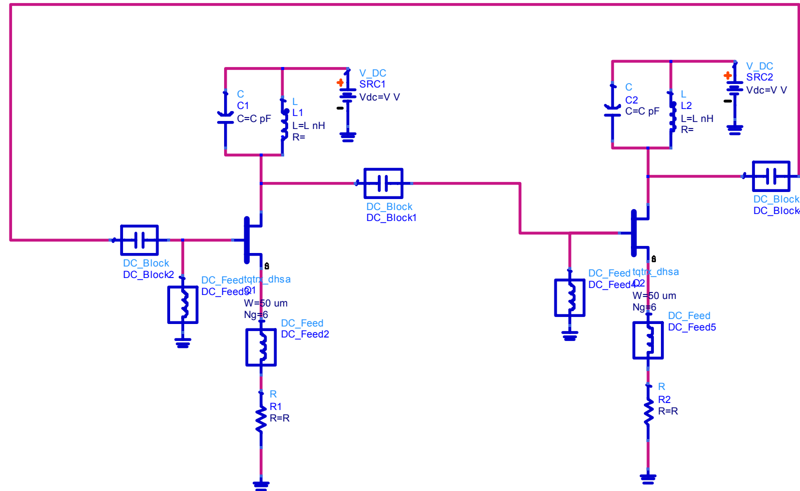
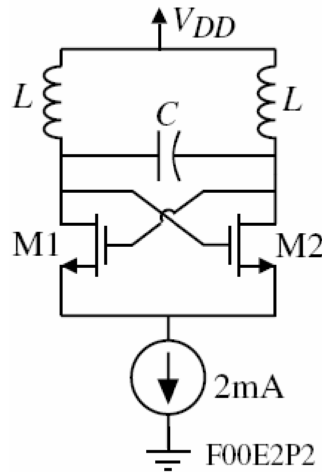


Figure 49 – Basic schematic of an LC oscillator.

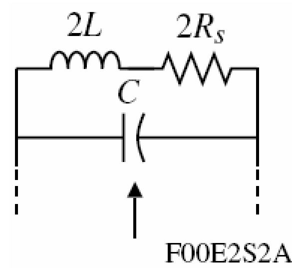
In this case the two LC couples can be considered as two L and a C. then the circuits showed in the previous figure are generated.
The oscillation condition is expressed by:

$$H(s)=\left[\frac{\frac{g_m}{C}s}{s^2+\frac{s}{RC}+\frac{1}{LC}}\right]^2 \Rightarrow H(j\omega)=\left[\frac{\frac{g_m}{C}j\omega}{-\omega^2+\frac{j\omega}{RC}+\frac{1}{LC}}\right]^2 = 1+j0 \Rightarrow \omega_{osc}^2=\frac{1}{LC} \;\&\; \frac{g_m}{C}=1 \quad (27)$$

Also the cross coupled circuits can be seen as a negative resistance oscillators and obviously the negative resistance must be compensate the losses in the LC circuit to Start-UP. The cross coupled transistor circuit is an example of negative resistance circuit but the mechanism is different from that of negative resistance oscillator in which you use the resonant circuit as reflection system, in this case in-fact you are using the LC as transmission system!.
For the synthesis you can consider the cross coupled as a negative impedance system and simulating them you can note as its behaviour is that of a parallel RL circuit with an high value of R then you can employ easily the conditions expressed in the previous section in terms of impedances.
We can consider as example the following problem:
An oscillator is showed in the figure below: The value for both inductance are L=5nH and C=2.5pF. If the *Q* for both inductors is 5 it is possible to determine the ratio W/L that allows the circuit to oscillate.



The equivalent circuit seen by side of negative resistance circuit is



The oscillation frequency is known through the formula:

$$\omega_{osc} = \frac{1}{\sqrt{2LC}} = \frac{1}{\sqrt{2 \cdot 5^{-9} 2.5^{-12}}} = 2\pi^9 \Rightarrow f_{osc} = 1GHz$$

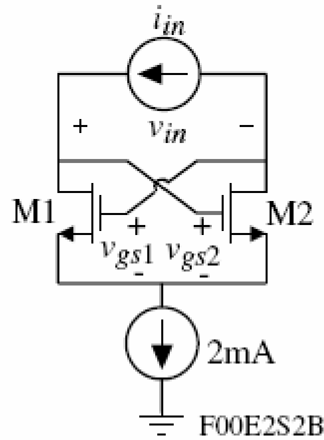
Therefore the series resistance for single inductor is known by the *Q*

$$R_S = \frac{\omega L}{Q} = \frac{2\pi \cdot 1^9 \cdot 5^{-9}}{5} = 6.2832\Omega$$

The series impedance seen by active circuit s Z=2Rs+2L and may be converted in an admittance to use the well known start-Up conditions:

$$Z_S = 2R_S + 2L \Rightarrow Y_S = \frac{1}{2R_S + 2L} = \frac{0.5}{R_S + j\omega L} = \frac{0.5}{R_S + j\omega L} \cdot \frac{R_S - j\omega L}{R_S - j\omega L} = \frac{0.5(R_S - j\omega L)}{\left(R_S^2 + \omega^2 L^2\right)} = 0.0031 - j0.0153$$

the reciprocal of the conductance is the parallel resistance *Rp=322.5Ω.*
The negative resistance of the active circuit must be slightly greater compared with *Rp.*
Rneg=-324.
Then we can design the transistor pair in order to provide this value of negative resistance.



$$i_{in} = gm_1 v_{gs1} = -g_{m2} v_{gs2}$$

$$R_{in} = \frac{v_{in}}{i_{in}} = \frac{v_{gs2} - v_{gs1}}{i_{in}} = \frac{-1}{g_{m1}} - \frac{-1}{g_{m2}} = \frac{-2}{g_m}$$

If the current is assumed of *2mA* splits between *M1* and *M2* for the negative resistance calculation we can write, after considering that the K_n factor $\mu n C_{ox}$ for the CMOS technology is about 110^{-6}

$$g_m = g_{m1} = g_{m2} = \sqrt{2 \cdot 2^{-3} \cdot 110^{-6}(W/L)} \Rightarrow \frac{\sqrt{(W/L)}}{1508} = 2\frac{1}{3242}$$

$$\sqrt{(W/L)} = 2\frac{1508}{324} \Rightarrow \frac{W}{L} = 85$$

It is easy to implement this solution with ADS, for example you can look the following circuit

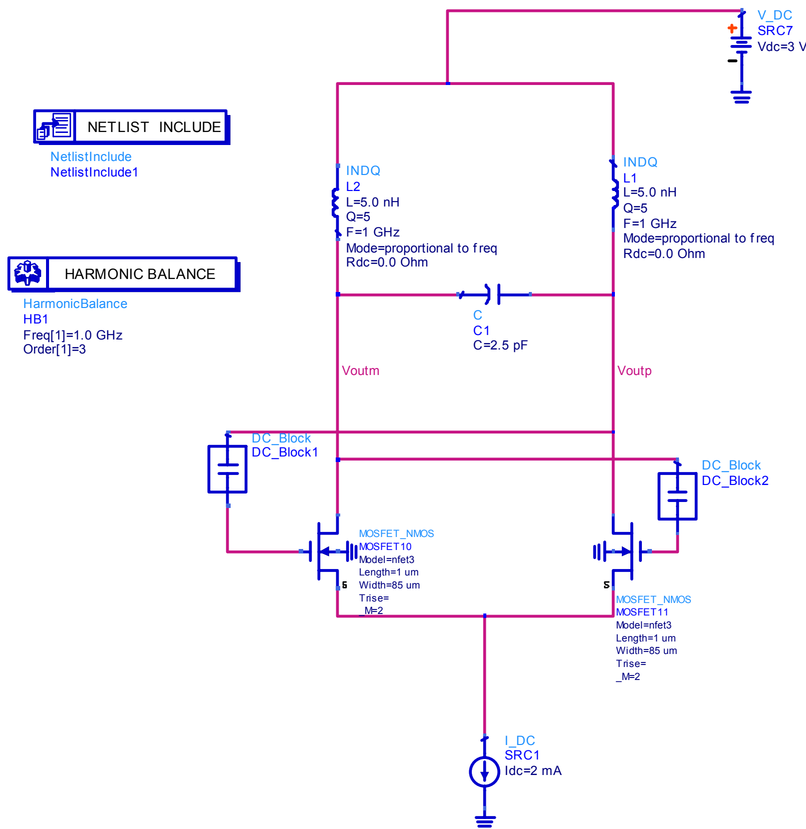


Figure 50 – Cross-Coupled oscillator implementation

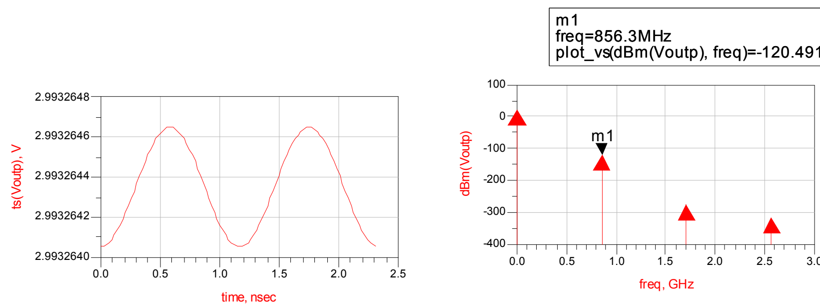


Figure 51 – HB simulation

As you can see the circuit works in agreement with theoretical calculus, obviously the poor quality factor of inductors is responsible to a very small signal generated in HB simulation. To improve the performances you can change for example the Quality factor of inductors, and then change the capacitance value to change the frequency.
ES: the frequency of *1 GHz* and a better output power is obtained for inductors *Q=20* and *C=1.95pF*.

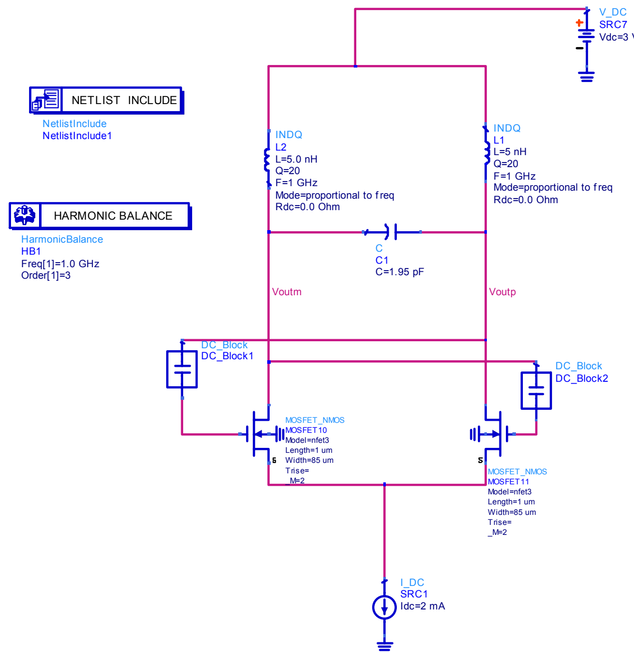


Figure 52 – Cross coupled optimized

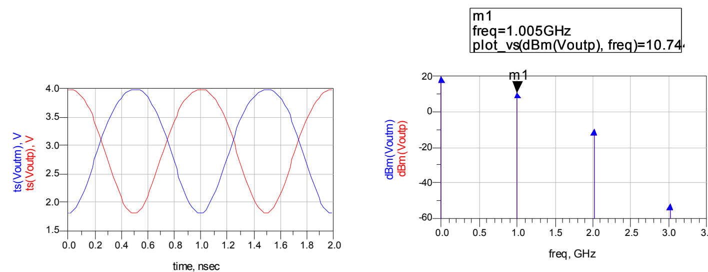


Figure 53 – HB simulations.

In order to pick-up the output signal in some cases the design of a buffer stage is needed. By changing the capacitance value it is possible to implement a VCO. It is important to note as good performances can be obtained with very small DC-bias current and then reducing the power consumption of the device that works in B-class. The only limits of this architecture is given by the speed switching for the cross coupled that imposes a superior limits to frequency, but with continuous technology improvements these limits are going to be overcomed. The continuous component can be cut by using a series capacitance in the output necessary to coupling the signal with the buffer stage. When the frequency is relatively low it is possible to use another simulation procedure and the Spice simulator.

Spice is based on time domain equations and provide very good results for the system.

The application of time domain simulation is limited by the frequency range. In modern CAD simulators the upper limit of time circuit simulation is about 10GHz. As the frequency increases the matrix used to describe the system becomes too large and the analysis it not performed with good results.

For the circuit previously described the time domain analysis as we jet show in the previously chapter are:

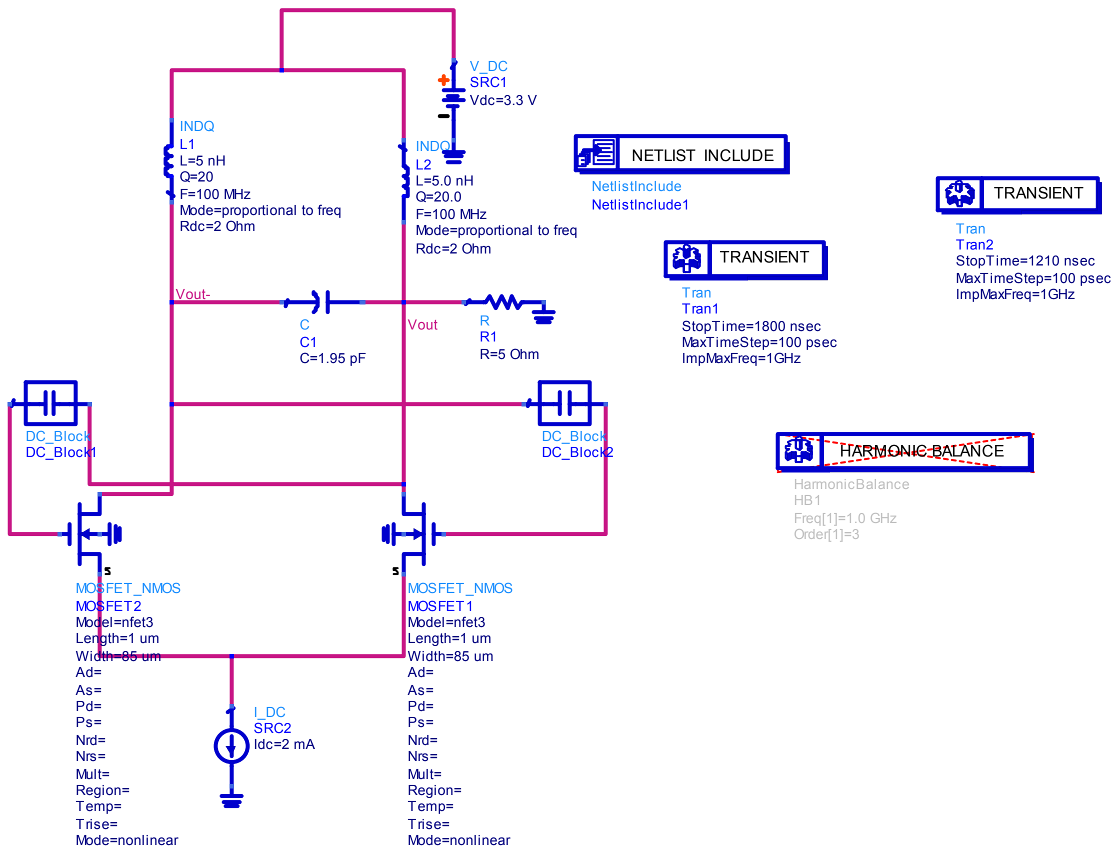


Figure 54

In this case we need to connect a series resistance of small value to pick out the signal from the oscillator. The value of this resistance will be small in order to not overcome the negative resistance value generated by the transistor pair and then the voltage waveform will have a small peak-to peak value. It is now possible to plot the trade of inductors current versus voltage across it and look the limit cycle of the system.

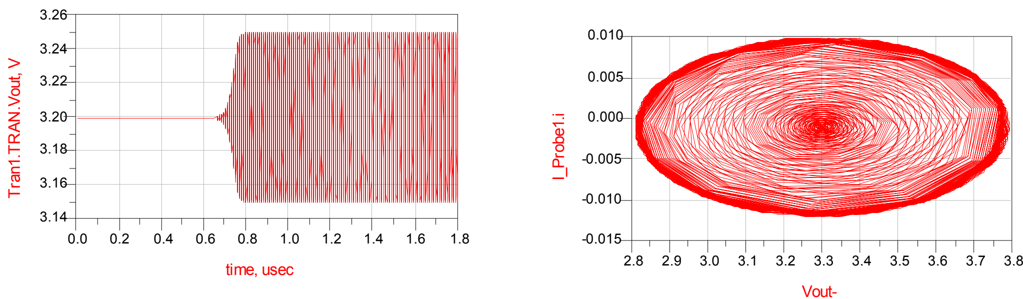


Figure 55

To do this measurement it is necessary to insert a current probe in the net of the inductor; then after the simulation is done the limit cycle is given by the I versus V inductor's plot.

REFERENCES

[1] Ulrich L. Rhode, Ajay K. Poddar, and Georg Böck, " General Theory of Oscillators," in The Design of Modern Microwave Oscillators for Wireless Applications , Canada: Jhon Wiley

[2] M.Monni, G.Martines,"A Novel Approach to Determine the Start-Up Conditions in Microwave Negative Impedance Oscillators", European Microwave Conference, in EuMW2007, Munich October 8-12, 2007.

[3] Less Besser Rowan Gilmore : "Practical RF circuit Design For modern Wireless Systems"Vol1: passive circuits and systems Artech House Boston London 2003

[4] Less Besser Rowan Gilmore : "Practical RF circuit Design For modern Wireless Systems Vol2: active circuits and systems" Artech House Boston London 2003

[5] T. Lee Ali Hajimiri "Oscillator Phase Noise: A Tutorial" IEEE Journal OF SOLID-STATE CIRCUITS VOL 35 NO3 MARCH 2000

[6] D.B. Leeson "A Simple model of feedback oscillator noise spectrum" in Proc. IEEE vol 54 Feb 1996

[7] A.Hajimiri and T.Lee "A general Theory of phase noise in electrical oscillators" IEEE J.Solid-State Circuits vol 33 Feb 1998

[8] P.Andreani "A Time Variant Analysis of the $1/f^2$ Phase Noise in CMOS Parallel LC-Tank Quadrature Oscillators" IEEE TRANSACTION ON CIRCUITS AND SYSTEMS VOL 53 August 2006

[9] P. Andreani Xiaoyan Wang "On the Phase Noise and Phase Error Performances of Multiphase LC CMOS VCOs"IEEE JOURNAL OF SOLID STATE CIRCUITS VOL 39 NO11 NOVEMBER 2004

4.1 The phase noise[1].

The phase noise is a crucial point in microwave design. It is a very confused topic and the better result of steady state analysis are published in the reference [1].
This chapter describe the phase noise as a review and a study of reference one with addition of my opinions and observations derived from my design experiences.
A List of publications about the phase noise can be found in the end of this chapter.

The theoretical and practical importance of oscillators has motivated the development of numerous treatment of phase noise. The sheer number of publications on this topic underscores the importance attached to it. At the same time, many of these disagree on rather fundamental points, and it may be argued that the abundance of such conflicting research quietly testifies to the inadequacies of many of those treatments. Complicating the search for a suitable theory is that noise in a circuit may undergo frequency translations before ultimately becoming oscillator phase noise. These translations are often attribute to the presence of obvious nonlinearities in practical oscillators. The simplest theories nevertheless simply ignore the nonlinearities altogether and frequently ignore the possibility of time variation as well. Such linear, time invariant (LTI) theories surprisingly manage to provide important qualitative design insights. However, these theories are understandably limited in their predictive power. Chief among the deficiencies of an LTI theory is that frequency translations are necessarily disallowed, begging the question of how the nearly symmetrical sidebands observed in practical oscillator arise. Despite this complication, and despite the obvious presence of nonlinearities necessary for amplitude stabilization, the noise-to-phase transfer function of oscillators nonetheless may be treated as linear. However, a quantitative understanding of the frequency translation process require abandonment of the principle of the time invariance implicitly assumed in most theories of phase noise. In addition to providing a quantitative reconciliation between theory and measurement, the time-varying phase noise model presented here identifies an important symmetry principle, which may be exploited to suppress the upconversion of *1/f* noise into close in phase noise. At the same time, it provides an explicit accommodation of cyclostationary effects, which are significant in many practical oscillators, and of amplitude-to-phase (AM-PM) conversion as well. These insight allow a reinterpretation of why certain topologies such Colpitts oscillators exhibit good performance. Perhaps more important, the theory informs design, suggesting novel optimization of well-known oscillators, as well as the invention of new circuit topologies. Tuned LC and ring oscillator circuit examples are presented to reinforce theoretical consideration developed. Simulations issues and the topic of amplitude noise are considered as well. We first revisit how one evaluates whether a system is linear or time invariant. Indeed, we find that we must even take care to define explicitly what is meant by the word “system”. We then identify some very general tradeoffs among key parameters, such as power dissipation, oscillation frequency, resonator Q and circuit noise power. These tradeoffs are first studied qualitatively in hypothetical ideal oscillator in which linearity of the noise to phase transfer function is assumed, allowing characterization y an impulse response. Althrough the assumption of linearity is defensible, we shall see that time invariance fails to hold even in this simple case. That is, oscillators are linear and time varying systems(LTV), where system is defined by the noise-to-phase transfer characteristic. Fortunately, complete characterization by an impulse response, depends only on linearity not time invariance. By studying the impulse response, we discover that periodic time variation leads to frequency translation of device noise to produce the phase noise spectra exhibited by real oscillators. In particular, the upconversion of *1/f* noise into close in phase noise is seen to depend on symmetry properties that are potentially controllable by the designer. Additionally, the same treatment easily subsumes the cyclostationarity of noise generators, and helps explain why class-C operation of active elements within an oscillator can be beneficial. Illustrative circuit examples reinforce key insights of the LTV model. In general circuit and device noise can perturb both the amplitude and phase of an oscillator’s output. Because amplitude

fluctuations are usually greatly attenuated as a result of the amplitude stabilization mechanism present in every practical oscillator, phase noise generally dominates, at least at frequencies not far removed from the carrier. Thus, even though it is possible to design oscillators in which amplitude noise is significant, we focus primarily on phase noise here. We show later that a simple modification of the theory allows the accommodation of output spectrum at frequencies well removed from the carrier.

4.2 General Considerations

Perhaps the simplest abstraction of an oscillator that still retains some connection to the real world is a combination of a lossy resonator and an energy restoration element. The latter precisely compensates for the tank loss to enable a constant-amplitude oscillation. To simplify matters, assume that the energy restored is noiseless. The tank resistance is therefore the only noiy element in this model.

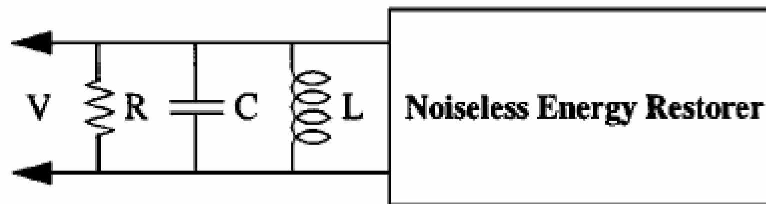


Figure 1- perfectly efficient oscillator

To gain some useful design insight, first compute the signal energy stored in the tank:

$$E^2 = \frac{1}{2} C V_{pk}^2 \qquad (1)$$

So that the mean square signal (carrier) voltage is:

$$\overline{V_{sig}^2} = \frac{E_{stored}}{C} \qquad (2)$$

Where we have assumed a sinusoidal waveform.
The total means square noise voltage is found by integrating the resistor's thermal noise density over the noise bandwidth of the RLC resonator.

$$\overline{V_n^2} = 4kTR \int_0^{\infty} \left| \frac{Z(f)}{R} \right|^2 df = 4kTR \frac{1}{4RC} = \frac{kT}{C} \qquad (3)$$

Combining 2 and 3 we obtain a noise-to-carrier ratio the reason for this upsie down ratio is simply one of convention

$$\frac{N}{C} = \frac{\overline{V_n^2}}{\overline{V_{sig}^2}} = \frac{kT}{Estored} \qquad (4)$$

Sensibly enough, one therefore needs to maximize the signal levels to minimize the noise-to-carrier ratio. We may bring power consumption and resonator Q explicitly into considreration by noting

that Q can be defined generally as proportional to the energy stored, divided by the energy dissipated:

$$Q = \omega_0 \frac{Estored}{Pdiss} \tag{5}$$

Therefore

$$\frac{N}{C} = \frac{\omega_0 kT}{QP_{diss}} \tag{6}$$

The power consumed by this model oscillator its simply equal to Pdiss the amount dissipated by the tank loss. The noise-to-carrier ratio is here inversely proportional to the oscillation frequency. This set of relationships still holds approximately for real oscillatorsa nd explains the near obsession of engineers with maximizing the resonator Q, for example. Other important design criteria become evident by coupling the foregoing with additional knowledge of practical oscillators. One is that oscillators generally operate in one of the two regimes that may be distinguished by their differing dependence of output amplitude on bias current so that one may write.

$$V_{sig} = I_{BIAS} R \tag{7}$$

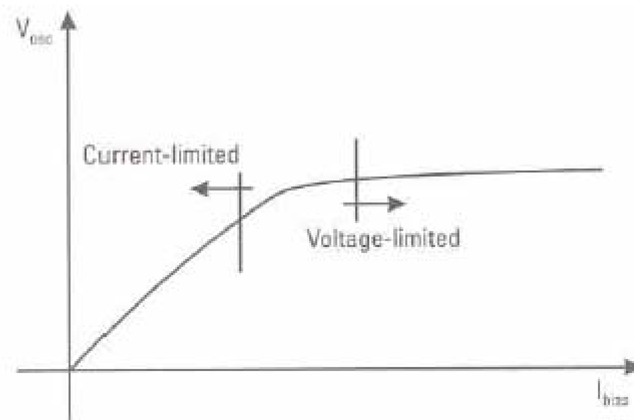


Figure 2- Oscillator operating regimes

Where R is constant of proportionality with the dimensions of resistance. This constant, in turn, is proportional to the equivalent parallel thank resistance so that

$$V_{sig} \propto I_{BIAS} R_{\tan k} \tag{8}$$

Implying that the carrier power may be expressed as

$$P_{sig} \propto \left(I_{BIAS} R_{\tan k}\right)^2 \tag{9}$$

The noise power has already been computed in terms of the tank capacitance as

$$\overline{V_n^2} = \frac{kT}{C} \tag{10}$$

But it may be expressed in terms of the tank inductance

$$\overline{V_n^2} = \frac{kT}{C} = \left(\frac{kT}{\frac{1}{\omega_0^2 L}} \right) = kT\omega_0^2 L \tag{11}$$

An alternative expression for the noise-to-carrier ratio in the current-limited regime is therefore

$$\frac{N}{C} \propto= \frac{kT\omega_0 L}{\left(I_{BIAS} R_{\tan k}\right)^2} \tag{12}$$

Assuming operation at a fixed supply voltage, a constraint output on power consumption implies an upper bound on the bias current. Of the remaining free parameters, then, only the tank inductance may be practically varied to minimize the *N/C* ratio. That is, optimization of such an oscillator corresponds to minimizing $L/(R_{tank})^2$. In many treatments, maximizing tank inductance is offered as a prescription for optimization. However, we see that a more valid objective is to minimize the $L/(R_{tank})^2$. Since the achieving this minimum is not always trivial. An additional consideration is that, below a certain minimum inductance, oscillation may case, Hence, the optimization prescription here presumes oscillation, and in a regime where the output amplitude is proportional to the bias current. To augment the qualitative insight of the foregoing analysis let us now determine the actual output spectrum of the ideal oscillator.

4.3 Phase Noise of Ideal Oscillator.

Assumes that the output of the circuit in Fig 1 is the voltage across the tank, as shown. By postulate, the only source of noise is the with thermal noise of the tank conductance which we represent as a current source across the tank with a mean-square spectral density of

$$\frac{\overline{i_n^2}}{\Delta f} \propto= 4kTG \tag{13}$$

This current noise becomes voltage noise when multiplied by the effective impedance facing the current source. In computing this impedance, however it is important to recognize that the energy restoration element must contruibute an average effective negative resistance that precisely cancels the positive resistance of the tank. Hence, the net result is that the effective impedance seen by the noise current source is simply that of a perfectly lossless LC network. For a relatively small offset frequency $\Delta\omega$ from the center frequency ω_0 the impedance of an LC tank may be approximated by

$$Z(\omega_0 + \Delta\omega) \approx j\frac{\omega_0 L}{\frac{2\Delta\omega}{\omega_0}} \tag{14}$$

We may rewrite the impedance in a more useful from by incorporating an expression for the unloaded tank Q:

$$Q = \frac{R}{\omega_0 L} = \frac{1}{\omega_0 G L} \tag{15}$$

Solving the (15) for L and substituting the (14) yield:

$$\left|Z(\omega_0 + \Delta\omega)\right| \approx \frac{1}{G}\frac{\omega_0}{2Q|\Delta\omega|} \qquad (16)$$

Thus we have treaded an explicit dependence on inductance for a dependence on Q and G.
Next, multiply the spectral density of the mean-square noise current by the squared magnitude of the tank impedance to obtain the spectral density of the mean-square noise voltage:

$$\frac{\overline{v_n^2}}{\Delta f} = \frac{\overline{i_n^2}}{\Delta f}|Z|^2 = 4kTR\left(\frac{\omega_0}{2Q\Delta\omega}\right)^2 \qquad (17)$$

The power spectra density of the output noise is frequency-dependent because of the filtering action of the tank, falling as the inverse-square of the offset frequency. This $1/f^2$ behaviour simply reflects the fact that the voltage frequency response of an RLC tank rolls off as $1/f$ to either side of the center frequency, and that power is proportional to the square of the voltage. Note also that an increase in tank Q reduces the noise density, when all other parameters are held constant, underscoring once again the value of increasing Q resonator. In our idealized LC model, thermal noise affects both amplitude and phase and (17) includes their combined effects. The equipartition theorem of the thermodynamics tells us that, in equilibrium, amplitude and phase noise power are equal. Therefore, the amplitude-limiting mechanism present in any practical oscillator suppress half the noise given by (17) It is traditional to normalize the mean-square noise voltage density to the mean square carrier voltage, and report the ratio in decibels, thereby explaining the upside down ratios presented previously. Performing this normalization yields the following equation for phase noise

$$L\{\Delta\omega\} = 10\log\left[\frac{2kT}{P_{sig}}\left(\frac{\omega_0}{2Q\Delta\omega}\right)^2\right] \qquad (18)$$

The units of phase noise are thus proportional to the log of the density. Specifically they are expressed as decibels below the carrier per hertz or dBc/Hz one might speak of a 2 GHz oscillator's phase noise as -110dBc/Hz @ 100KHz offset. It is important to note that the per Hz actually applies to the argument of the log not to the log itself; doubling the measurement bandwidth does nor double in the decibel quantity. As lacking in rigor DBc/Hz is it common usage. Equation (18) tell us that phase noise (at a given offset) improves as both the carrier power and Q increase, as predicted earlier. These dependencies make sense. Increasing the signal power improves the ratio simply because the thermal noise is fixed, while increase the Q improves the ratio quadratically because the tank impedance's fall off as $1/(Q\Delta\omega)$.
Because many simplifying assumptions have led us to this point, it should not be surprising that there are some significant differences between the spectrum predicted by (18) and what one typically measured in practice. For example although real spectra do possess a region where the observed density is proportional to $1/(\Delta\omega)^2$, the magnitude are typically quite a bit larger than predicted by (18) because there are additional important noise sources beside the tank loss. For example, any physical implementation of an energy stored will be noisy. Furthermore, measured spectra eventually flatten out for large frequency offsets, rather than continuing to drop quadratically. Such a floor may be due to the noise associated with active elements (such as buffers) placed between the tank and the outside world, or it can even reflect limitations in the measurement instrumentation itself. Even if the output were taken directly from the tank, any resistance in series with either the inductor or capacitor would impose a bound on the amount of filtering provided by the tank at large frequency offsets and thus ultimately produce noise floor. Finally there is almost

always a $1/(\Delta\omega)^3$ region at small offsets. A modification to (18) provides a means to account for these discrepancies

$$L\{\Delta\omega\} = 10\log\left[\frac{2FkT}{P_{sig}}\left\{1+\left(\frac{\omega_0}{2Q\Delta\omega}\right)^2\right\}\left(1+\frac{\Delta\omega_{1/f^3}}{|\Delta\omega|}\right)\right] \qquad (19)$$

These modifications, due to Leeson consist of a F factor to account for the increased noise $1/(\Delta\omega)^2$ region, an additive factor for unity (inside the braces) to account for the noise floor, and a multiplicative factor of unity (the term in the second set of parentheses) to provide a $1/|\Delta\omega|^3$ behaviour at sufficiently small offset frequencies. With these modifications the phase noise spectrum appears as in the figure below.

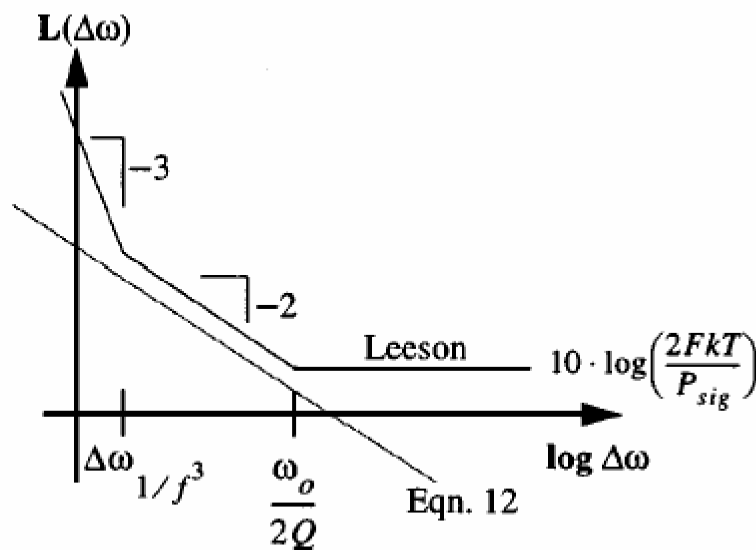


Figure 3

It is important to note that the factor F is an empirical fitting parameter and therefore must be limited from measurements, diminishing the predictive power of the phase noise equation. Furthermore, the model asserts that $\Delta\omega_{1/f}{}^3$, the boundary between the $1/(\Delta\omega)^2$ and a $1/|\Delta\omega|^3$ regions, is precisely equal to *1/f* corner of the device noise. However, measurements frequently show no such equality, and thus one must generally treat $\Delta\omega_{1/f}{}^3$ as an empirical fitting parameter as well. Also it is not clear what the corner frequency will be in the presence of more than one noise source, each with an individual *1/f* noise contribution (and generally differing *1/f* corner frequency). Finally, the frequency at which the noise flattens out is not always equal to half the resonator bandwidth $\omega_0/2Q$. **Both the ideal oscillator model and the Lesson models suggest that increasing resonator Q and signal power are ways to reduce phase noise**. The Leesson model traditionally introduces the F factor but without know precisely what depends it on, it is difficult to identify specific ways to reduce it. The same problem exists with $\Delta\omega_{1/f}{}^3$ as well. Finally, blind application of these models has periodically led to earnest but misguided attempts by some designers to use active circuits to boost Q. Sadly increase in Q through such means are necessarily accompanied by increases in F as well, generally preventing the anticipated improvement in phase noise. Note that a resonant active circuit is an oscillator then this attempt was done by using two oscillators. Again the lack of analytical expression for F can obscure this conclusion and one continues to encounter various doomed oscillator design based on the notion of active Q boosting.

That neither (18) nor (19) can make quantitative predictions about phase noise in an indication that at least more of the assumptions used in the derivations are invalid, despite their apparent resonableness. To develop a theory that does not possess the enumerated deficiencies, we need to revisit and perhaps revise these assumptions.

4.4 A Linear Time Varying Phase Noise Theory

The foregoing derivations have all assumed linearity and time invariance. Let'us reconsider each of these assumption in turn. Nonlinearity is clearly a fundamental property of all real oscillators, as it is necessary for amplitude limiting. Several phase noise theories have consequently attempted to explain certain observation as a consequence of nonlinear behaviour. One observation is that a single-frequency sinusoidal disturbance injected into an oscillator gives rise to two equal-amplitude sidebands, symmetrically disposed about the carrier. Since LTI system cannot perform frequency translation and non linear system can, nonlinear mixing has occasionally been proposed to explain phase noise. Unfortunately the amplitude of sidebands must then depend nonlinearly on the amplitude of the injected signal, and this dependency is not observed. One must conclude that memoryless nonlinearity cannot explain this discrepancies, despite initial attractiveness as the culprit. As we shall see momentarily, amplitude-control nonlinearities certainly do affect phase noise, but only incidentally, by controlling the detailed shape of the output waveform. An important insight is that disturbances are just that: perturbances superimposed on the main oscillation. They will always be much smaller in magnitude than the carrier in any oscillator worth designing or analyzing. Thus, if a certain amount of injected noise produces a certain amount of phase disturbance, we ought to expect doubling the injected noise to produce double the disturbance. Linearity would therefore appear to be a reasonable assumption as far as the noise to phase transfer function is concerned. It is therefore particularly important to keep in mind that when assessing linearity , it is essential to identify explicitly the input-output variables. Linear relationships may exist between certain variable pairs as the same time nonlinear ones exists between others. Linearization of the fundamentally nonlinear behaviour of the active devices. Indeed, we will perform a linearization around the steady-state solution, which automatically takes the effect of the device nonlinearity into account. There is therefore no contradiction here with the prior acknowledgment of nonlinear amplitude control.

We are left only with the assumption of time invariance to re-examined. In the previous derivations, we have extended time invariance to the noise sources themselves, meaning that the measures that characterize noise (e.g., spectral density) are time invariant (stationary). IN contrast with linearity, the assumption of time invariance is less obviously defensible. IN fact, it is surprisingly simple to demonstrate that oscillators are fundamentally time-varying systems. Recognizing this truth is the main key to developing a more accurate theory of phase noise. To show that time invariance fails to hold, consider explicitly how an impulse of current affects the waveform of the simplest resonant system a lossless LC tank (figure below)

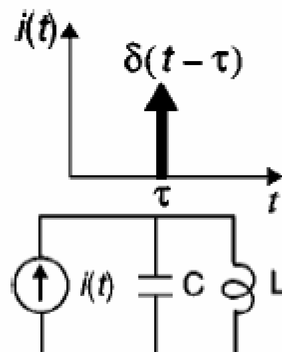


Figure 4- LC oscillator excited by current pulse

Assume that the system is oscillating with some constant amplitude until the impulse occurs, then consider how the system responds to an impulse injected at two different times as seen in Fig.4.
If the impulse happened to coincide with a voltage maximum (as in the left plot), the amplitude increases abruptly by an amount $\Delta V=\Delta Q/C$, but because the response to the impulse superposes exactly in phase with pre-existing oscillation, the timing of zero crossing does not change. On the other hand an impulse injected at some other time generally affects both the amplitude of oscillation and the timing of the zero crossings, as in the right plot. Interpreting the zero crossing timings as a measure of phase, we see that the amount of phase disturbance for a given injected impulse depends on then the injection occurs; time invariance thus fails to hold. An oscillator is therefore a linear, but (periodically) time-varying (LTV) system.
Because linearity remains a good assumption, the impulse response still completely characterizes the system, even with time variation thrown in. Nothing that an impulsive input produces a step change in phase, the impulse response may be written as

$$h_\phi(t,\tau)=\frac{\Gamma(\omega_0\tau)}{q_{\max}}u(t-\tau) \qquad (20)$$

Where $u(t)$ is the unit step function. Dividing by q_{max} , the maximum charge displacement across the capacitor, makes the function $\Gamma(x)$ through simulation, but there are also analytical methods (some approximate) that apply in special cases. IN any event, to develop a feel for typical shapes of the impulsive sensitivity function, consider two representative examples, first for a LC and a ring oscillator represented respectively in figure 5 a and b:

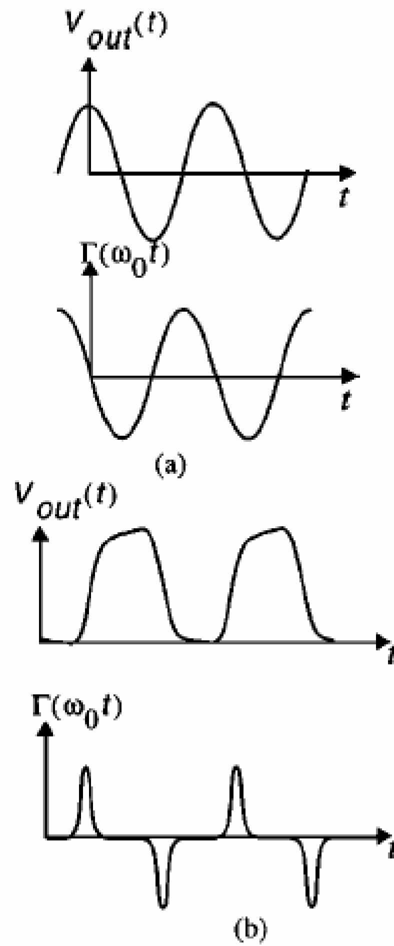


Figure 5 - Examples of ISF for a LC and a ring oscillator.

Once the ISF has been determined (by whatever means), we may compute the excess phase through use of the superposition integral. This computation is valid here since superposition is linked to linearity, not time invariance

$$\phi_\phi = \int_{-\infty}^{+\infty} h_\phi(t,\tau) i(\tau) d\tau = \frac{1}{q_{\max}} \int_{-\infty}^{t} \Gamma(\omega_0) i(\tau) d\tau \qquad (21)$$

This computation can be visualized with the help of the equivalent block diagram show in Fig.6.

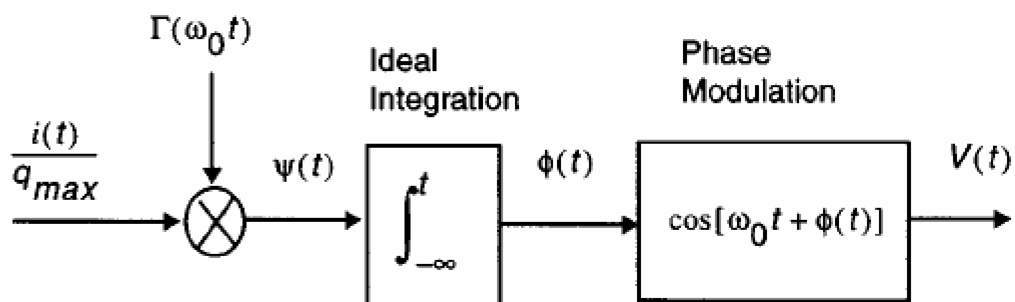


Figure 6- The equivalent block diagram of the process

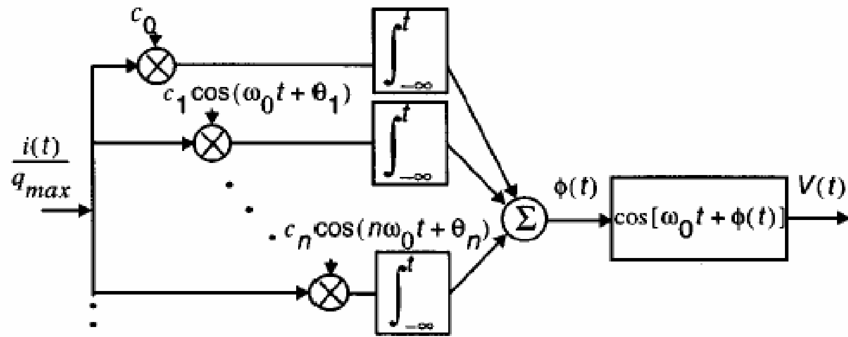


Figure 7- The equivalent system for the ISF decomposition

To cast this equation in a more practically useful form, note that the ISF is periodic and therefore expressible as a Fourier series.

$$\Gamma(\omega_0\tau)=\frac{c_0}{2}+\sum_{n=1}^{\infty}c_n\cos(n\omega_0\tau+\theta_n) \qquad (22)$$

Where the coefficients c_n are real and θ_n is the phase of the *nth* harmonics of the ISF. We will ignore θ_n in all that follows because we will be assuming that noise components are uncorrelated, so their relative phase is irrelevant.
The value of this decomposition is that, like many functions associated with physical phenomena, the series typically converges rapidly, so it is often well approximated by just the first few terms of the series. Substituting the Fourier expansion into (21) and exchanging summation and integration, one obtains

$$\phi_\phi=\frac{1}{q_{\max}}\left[\frac{c_0}{2}\int_{-\infty}^{t}i(\tau)d\tau+\sum_{n=1}^{\infty}c_n\int_{-\infty}^{t}i(\tau)\cos(n\omega_o\tau)d\tau\right] \qquad (23)$$

The corresponding sequence of mathematical operations is shown graphically in the left half of Fig.7. Note that the block diagram contains elements that are analogous to those of a super heterodyne receiver. The normalized noise current is a broad-band “RF” signal, whose Fourier components undergo simultaneous down conversion (multiplications) by local oscillator signals at all harmonics of the oscillation frequency. It is important to keep in mind that multiplication is a linear operation if one argument is held constant , as it is here. The relative contributions of these multiplications are determined by the Fourier coefficients of the ISF. Equation (23) thus allows us to compute the excess phase caused by an arbitrary noise current injected into the system, once the Fourier coefficients of the ISF have been determined. Earlier , we noted that signals (noise) injected into a non-linear system can exhibit qualitatively similar behaviour, as implied by the super heterodyne imagery invoked in the preceding paragraph. To demonstrate this property explicitly, consider injecting a sinusoidal current whose frequency is near an integer multiple m of the oscillation frequency, so that

$$i(t)=I_m\cos[(m\omega_0+\Delta\omega)t] \qquad (24)$$

Where $\Delta\omega \ll \omega_0$. Substituting (24) into (23) and noting that there is a negligible net contribution to the integral by terms other than when n=m, one obtains the following approximation

$$\phi(t) \approx \frac{I_m c_m \sin(\Delta\omega t)}{2q_{\max}\Delta_{\omega}} \qquad (25)$$

The spectrum of $\phi(t)$ therefore consists of two equal sidebands at $\pm\Delta\omega$, even though the injection occurs near some integer multiple of ω_0 . We see that we do not need to invoke nonlinearity to explain this frequency conversion (or "folding"). This observation is fundamental to understanding the evolution of noise in an oscillator. Unfortunately, we are not in quite done: (25) allows us to figure out the spectrum of $\phi(t)$, but we ultimately want to find the spectrum of the output voltage of the oscillator, which is not quite the same thing. The two quantities are linked through the actual output waveform. To illustrate what we mean by this linkage, consider a specific case where the output may be approximated as a sinusoid, so that $v_{out}(t)=\cos[\omega_0 t+\phi(\tau)]$.
This equation may be considered a phase to voltage converter; it takes phase as an input and produces from it the output voltage. This conversion is fundamentally nonlinear because it involves the phase modulation of a sinusoid.
Performing this phase-to-voltage conversion, and assuming "small" amplitude disturbances, we find the single-tone injection leading to (25) results in two equal-power sidebands symmetrically disposed about the carrier.

$$P_{SCB}(\Delta\omega) \approx 10 \cdot \log\left(\frac{I_m c_m}{4q_{\max}\Delta\omega}\right)^2 \qquad (26)$$

To distinguish this result from nonlinear mixing phenomena, note that the amplitude dependence is linear (the squaring operation simply reflects the fact that we are dealing with a power quantity here). This relationship can be, and has been, verified experimentally.
The foregoing result may be extended to the general case of white noise source

$$P_{SCB}(\Delta\omega) \approx 10 \cdot \log\left(\frac{\frac{\overline{i_n^2}}{\Delta f}\sum_{m=0}^{\infty} c_m^2}{4q_{\max}^2\Delta\omega^2}\right) \qquad (27)$$

Equation (26) implies both upward and downward frequency translations of noise into the noise near the carrier, as illustrated in Fig.8. This figure summarizes what the foregoing equations tell us: components of noise near integer multiples of the carrier frequency all fold into noise near the carrier itself. Noise near dc get unconverted weighted by coefficient c_0, so 1/f device noise ultimately becomes $1/f^3$ noise near the carrier; noise near the carrier stays here, weighted by c_1 and white noise near higher integer multiplies of the carrier undergoes down conversion, turning into noise in the $1/f^2$ region.
Note that $1/f^2$ shape results from the integration implied by the step to change in phase caused by an impulsive noise input. Since an integration (even time-varying one) gives a white voltage or current spectrum a 1/f character, the power spectral density will have a $1/f^2$ shape.

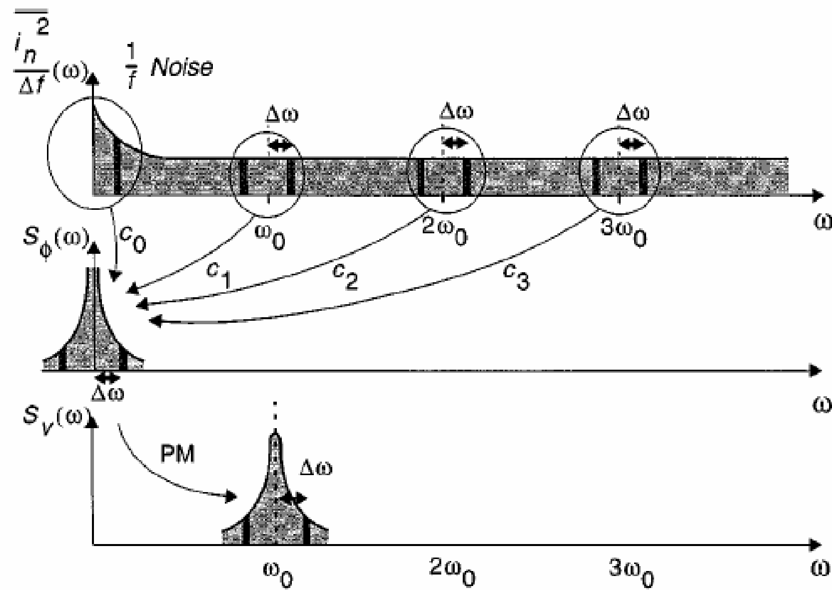


Figure 8- Evolution of circuit noise into phase noise.

It is clear form previously figure that in order to minimize the phase noise it is necessary to minimize the coefficients c_n of the ISF. To underscore this point qualitatively we can apply the Parseval's theorem on write:

$$\sum_{n=0}^{\infty} c_m^2 = \frac{1}{\pi}\int_0^{2\pi} |\Gamma(x)|^2 dx = 2\Gamma_{rms}^2 \qquad (28)$$

So that the spectrum in the $1/f^2$ region will be expressed as

$$L(\Delta\omega) = 10 \cdot \log\left(\frac{\frac{\overline{i_n^2}}{\Delta f}\Gamma_{rms}^2}{2q_{\max}^2 \Delta\omega^2} \right) \qquad (30)$$

Where Γ_{rms} is the rms value of the ISF. All other factors held equal, reducing Γ_{rms} will reduce the phase noise at all frequencies. Equation (30) is the rigorous equation for the $1/f^2$ region and is one key result of the LTV model. Note that no empirical curve-fitting parameter are presented in equation (30). Among other attributes (30) allows to study qualitatively the up conversion of 1/f noise into close in phase noise. Noise near the cattier is particularly important in communication systems with narrow channel spacing. In fact the allowable channel spacing are frequently constrained by the achievable phase noise. Unfortunately, it is not possible to predict close in phase noise correctly with LTI models.

This problem disappears in the LTV model is used. Specifically assume that the current noise behaves as follows in the 1/f region.

$$\overline{i_{n,1/f}^2} = \overline{i_n^2} \cdot \frac{\omega_{1/f}}{\Delta\omega} \qquad (31)$$

Where $\omega_{1/f}$ is the 1/f corner frequency. Using (30), we obtain the following for the noise in the $1/f^3$ region:

$$L(\Delta\omega) = 10 \cdot \log\left(\frac{\frac{\overline{i_n^2}}{\Delta f} c_0^2}{8q_{\max}^2 \Delta\omega^2} \cdot \frac{\omega_{1/f}}{\Delta\omega} \right) \qquad (32)$$

Which describes the phase noise in the $1/f^3$ region. The $1/f^3$ corner frequency is then

$$\Delta\omega_{1/f} = \omega_{1/f} \cdot \frac{c_0^2}{4\Gamma_{rms}^2} = \omega_{1/f} \cdot \left(\frac{\Gamma_{dc}}{\Gamma_{rms}} \right)^2 \qquad (33)$$

From which we see that the $1/f^3$ phase noise corner is not necessarily the same as the 1/f device circuit noise corner; it will generally be lower. IN fact since the Γ_{dc} is the dc value of the ISF, there is a possibility of reducing by large factors the $1/f^3$ phase-noise corner. The ISF is a function of the waveform, and hence potentially under the control of the designer, usually through adjustment of the rise-and fall-time symmetry. This result is not anticipated by LTI approaches, and is one of most powerful insights conferred by the LTV model. This result has particular significance for technologies such CMOS and GaAs MESFET's. A specific circuit example for how one may exploit this observation follows shortly.

One more extremely powerful insights concerns the influence of cyclostationary noise sources. As alluded to earlier, the noise sources in many oscillators cannot be well modelled as stationary. A typical example is the nominally white drain or collector noise current in a MOSFET. Noise currents are a function of bias currents, and the latter vary periodically with the oscillating waveform. The LTV model is able to accommodate a cyclostationary white noise source with ease, since such a source may be treated as the product of a stationary white noise source and a periodic function.

$$i_n(t) = i_{n0}(t) \cdot \alpha(\omega_0 t) \qquad (34)$$

Here , in0 is a stationary white noise source whose peak value is equal to that of the cyclostationary source, and a(x) is a periodic unity less function with a peak value of this unity. Substituting this into (22) allows us to treat cyclostationaty noise as a stationary noise source provided we define an effective ISF as follows:

$$\Gamma_{eff}(x) = \Gamma(x) \cdot \alpha(x) \qquad (35)$$

Thus, none of the foregoing conclusions changes as long as Γ_{eff} is used in all of the equations. Having identified the factors that influence the oscillator noise, we are now in a position to articulate the requirements that must be satisfied to make a good oscillator. First, in common with revelations of LTI models, both signal power and resonator Q should be maximized, all other factors held constant. In addition, note that an active devices is always necessary to compensate for

the tank loss, and that active devices always contribute noise. Note also that the ISF's tell us that there are sensitive and insensitive moments in an oscillation cycle. Of the infinitely many ways that an active element could return energy to the tank, this energy should be delivered all in once, where the ISF has its minimum value. In an ideal LC oscillator, therefore, the transistor would remain off almost all of the time, waking up periodically to deliver an impulse of current at the signal peak(s) of each circle. The extent to which real oscillators approximate this behaviour determines the quality of their phase-noise properties. Since an LTI theory treats all instants as equally important, such theories are unable to anticipate this important result. Last the best oscillators will possess the symmetry properties that lead to small Γ_{dc} for minimum up conversion of 1/f noise. In the following section we consider several circuit examples of how accomplish these ends in practice.

4.5 Circuit Examples

Let us examinee for first the Colpitts oscillator and its relevant waveforms, now that we have developed these insights.

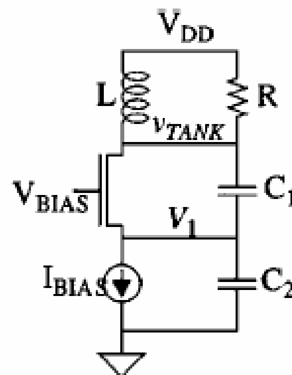


Figure 9

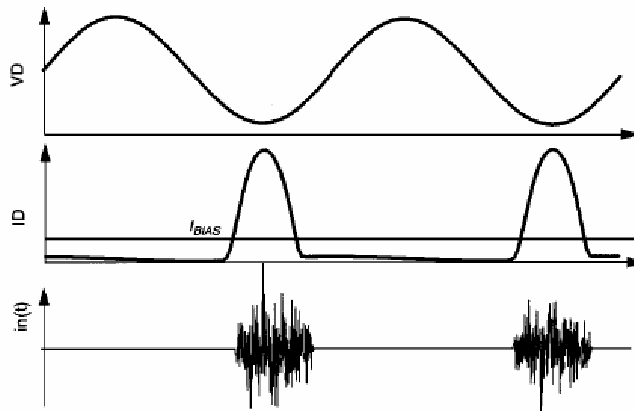


Figure 10- Approximate incremental tank voltage and drain current for Colpitts

Note that the drain current flows only during a short interval coincident with the most benign moments (the peaks of the tank voltage). Its corresponding excellent phase noise properties account for the popularity of this configuration. It has long been known that the best phase noise occur for a certain narrow range of tapping ratios (es 4:1 capacitance ratio), but before the LTV theory, no theoretical basis existed to explain a particular optimum. Both LTI and LTV models point out the value of maximizing signal amplitude. To evade supply-voltage or break down constraints, one may employ a tapped resonator to decouple resonator swings from device voltage limitations. A common configuration that does so is Clapp's modification to the Colpitt's oscillator.

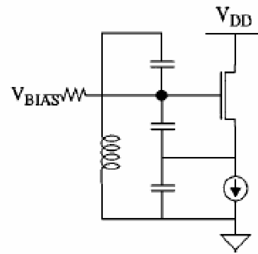


Figure 11- Clapp's Oscillator

Differential implementations of oscillators with tapped resonators have recently made an appearance in the literature. These kinds of oscillators become increasingly attractive as supply voltages scale downward, where conventional resonator connection s lead to VDD-constrained signal swings. Use of tapping allows signal energy to remain high level even with low supply voltages.

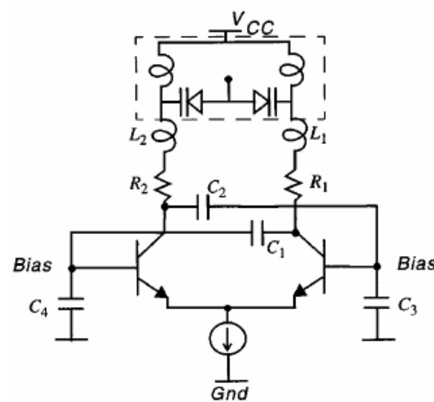


Figure 12- Simplified schematic for the VCO

Phase noise predictions using the LTV model are frequently more accurate for bipolar oscillators due to availability of better device noise models. In[10], impulse response modelling was used to optimize the noise performance of a differential bipolar voltage-controlled oscillator with an automatic amplitude control loop using an external resonator. A simplified schematic of this oscillator is shown in figure 12. A tapped resonator is used to increase the tank signal power P_{sig}. The optimum capacitive tapping ratio is calculated to be around 4.5 based on simulations that take the cyclostationarity of the noise into account. The effect of 1/f noise reduction is clearly seen, as a $1/f^3$ corner of 3KHz is both predicted and measured, in comparison with a device 1/f noise corner of 200KHz. The measured phase noise of -106dBc/Hz at 100KHz offset in the $1/f^2$ region is also an excellent agreement with the predicted value of -106.2dBc/Hz. The automatic amplitude control loop allows for independent optimization of the steady-state and start-up conditions in terms of phase noise.

As mentioned a key insight of the LTV theory concerns the importance of symmetry. A configuration that exploits this knowledge is the symmetrical negative resistance oscillator showed in figure 13

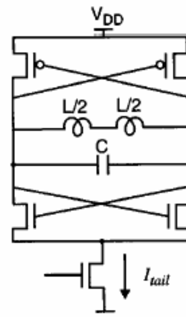


Figure 13- Symmetrical negative resistance oscillator

This configuration is hardly new, but an appreciation of its symmetry properties is. Here, it is the half circuit symmetry that is important, because noise in the two half circuits is only partially correlated at best. By selecting the relative widths of the PMOS and NMOS devices appropriately to minimize the dc value of the ISF (Γ_{dc}) for each half-circuit, one may minimize the up conversion of 1/f noise. Through exploitation of symmetry in this manner, the $1/f^3$ corner can be dropped to exceptrionally low values, even though device 1/f noise corners may be high (as is typically in the case for CMOS). Furthermore, the bridge-like arrangement of the transistor quad allows for generated signal swings, compounding the improvements in phase noise. As a result of all of these factors, a phase noise of -121 dBc/Hz at an offset of 600KHz at 1.8GHz has been obtained with low-Q on chip spiral inductors, on 6 mW of power consumption in a 0.25-mm CMOS technology. This result rivals what one may achieve with bipolar technologies. With a modest increase in power, the same oscillator's phase noise becomes compliant with specification for GSM1800.
The tapped resonators are the resonators implemented in microstrip technology for example by using two pieces of transmission line with a proper L and W to resonate at a proper frequency with an electromagnetic coupling between each others.
A tapped resonator can be well used in symmetrical negative resistance structures.

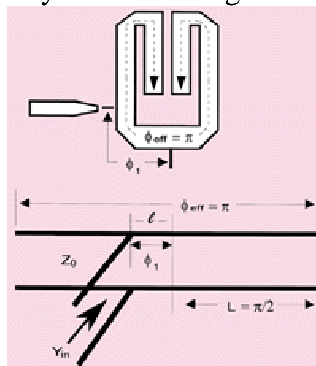


Figure 14- Schematic for a tapped resonator

Ring-oscillators

As an example of a circuit that does not well approximate ideal behaviour, consider a ring oscillator. First, the resonator Q is poor since the energy stored in the node capacitance is reset (discharged) every cycle. Hence, if the resonator of a Colpitts oscillator may be likened to a fine crystal wine glass, the resonator of a ring oscillator is mud. Next, energy is restored to the resonator during the edges (the worst possible times), rather than a the voltage maxima. These factors account for the well-known terrible phase noise performance for a ring oscillators. As a consequence, ring oscillators are found only if the most no critical applications, or inside wide band phase locked loops that clean un spectrum.
However there are certain aspect of ring oscillator that can be exploited to achieve better phase-noise performance in a mixed-mode integral circuit. Noise sources on different nodes of an

oscillator may be strongly correlated due to various reasons. Two examples of sources with strong correlation are substrate and supply noise, arising from current switching in other parts of the chip. The fluctuations on the supply and substrate will induce a similar perturbation on different stages of the ring oscillator. To understand the effect of this correlation, consider the special case of having identical noise sources on all the nodes of the ring oscillator, as shown in fig. 15.

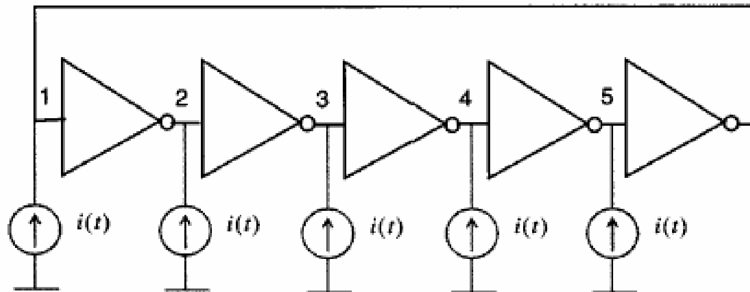


Figure 15- Five stage ring oscillator with identical noise sources on all nodes.

If all the inverters in the oscillator are the same, the ISF for different nodes will differ only in phase by multiplies of 2π/N , as shown in Fig 16.

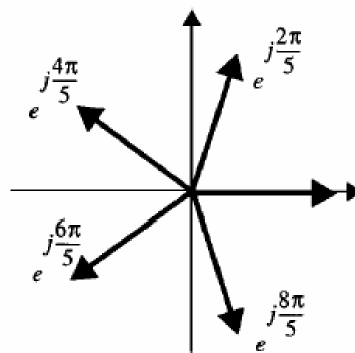


Figure 16- Phasors for noise contributions from each source

Therefore the total phase due to all the sources is given by

$$\phi(t) = \frac{1}{q_{\max}} \int_{-\infty}^{t} i(\tau) \left[\sum_{n=0}^{N-1} \Gamma\left(\omega_0 \tau + \frac{2\pi n}{N} \right) \right] d\tau \qquad (36)$$

Expanding the term in brackets in a Fourier series, it can be observed that is zero except at dc and multiplies of $N_{\omega 0}$.

$$\phi(t) = \frac{N}{q_{\max}} \sum_{n=0}^{\infty} c(nN) \int_{-\infty}^{t} i(\tau) \cos(nN\omega_0 \tau) d\tau \qquad (37)$$

Which means that for fully correlated sources, only noise in the vicinity of integer multiplies of $N_{\omega 0}$ affects the phase. Therefore, every effort should be made to maximize the correlations of noise arising from substrate and supply perturbations. This can be done by making the invert stages and the noise sources on each node as similar o each others as possible by proper layout and circuit design. For example, the layout should be kept symmetrical, and the inverter stages should be laid out close to each other so that the substrate noise appears as a common mode source.

This latter consideration is particularly important in the case of a lightly doped substrate, since such a substrate may no act as a single node. It is also important that the orientation of all stages be kept identical. The interconnecting wires between the stages must be identical, perhaps,, for example, by using dummy buffer stages as necessary. Use desired frequency will also be helpful because, as a practical matter, fever coefficients will then affect the phase noise. Last, as the low-frequency portion of the substrate and supply noise then dominates, one should exploit symmetry to minimize the Γ_{dc} .

Another common conundrum concerns the preferred topology for MOS ring oscillators, i.e, whether a single-ended or differential topology results in better jitter and phase noise performances for a given centre frequency f_0 and total power dissipation. Using the noise equations for transistors an approximate expression for the ISF, one may derive expressions for the phase noise of MOS differential and single-ended oscillators. Based on these expressions, the phase noise of a single ended inverter chain ring oscillator is found to be independent of the number of stages for a given power dissipation and frequency of operation. However for a differential ring oscillator, the phase noise (jitter) grow with the number of stages. Therefore, even a properly designed differential CMOS ring oscillator underperforms its single-ended counterpart, with a disparity that increases with the number of stages. The difference in the behaviour of these two types of oscillators with respect to the number of stages can be traced to the way they dissipate power. The dc current drawn from the supply is independent of the number and slope of the transitions in differential ring oscillators. IN contrast, inverter-chain ring oscillators dissipate power mainly on a per transition basis and therefore have better phase noise for a given power dissipation. However, a differential topology may still preferred in IC's with a large amount of digital circuitry because of the lower sensitivity to substrate and supply noise, as well as lower noise injection into other circuits on the same chip. The decision of which architecture to use should be based on both of these considerations. Yet another commonly debated question concerns the optimum number of inverter stages in a ring oscillator to achieve the best jitter and phase noise for a given f_0 and P. For single ended CMOS ring oscillators, the phase noise and jitter in the $1/f^2$ region are not strong functions of the number of stages. However if the symmetry criteria are not well satisfied, and or the process has large 1/f noise, a large N will reduce the jitter. In general the choice of the number of stages must be made on the basis of several design criteria, such as 1/f noise effects, the desired maximum frequency of oscillation, and the influence of external noise sources, such as supply and substrate noise, that may not scale with N.

The jitter and phase noise increase with an increasing number of stages (three or four) should be used to give the best performance. This recommendation holds even if the power dissipation is not a primary issue. It is not fair to argue that burning more power in a larger number of stages allows the achievement of better phase noise, since dissipating the same total power in a smaller number of stages with larger devices results in better jitter and phase noise, as long as possible to maximize the total charge swing.

CONCLUSION

The insight gained from the LTI phase noise models are simple and intuitively satisfying. One should maximize signal amplitude and resonator Q. In additional implicit insight is that the phase shifts around the loop generally must be arranged so that oscillation occurs at or very near the frequency of the resonator. This result is the same that derived by our approach that getting build on LTI approach

This way, there is a maximum attenuation by the resonator of off-centre spectral components.

Deeper insights provided by the LTV model are that the resonator energy should be restored impulsively at the ISF minimum, instead of evenly throughout a cycle, and that the dc value of the effective IDF should be made as close to zero as possible to suppress the up conversion of 1/f noise into close-in phase noise.

The theory also shows that the inferior broad-band noise performance of ring oscillators may be offset by their potentially superior ability to reject common-mode substrate and supply noise.
Exact analytical derivations of the ISF are usually not obtainable for any but the simplest oscillators. Various approximate methods are outlined in [3] and [4], but most generally accurate method is direct evaluation of the time-varying impulse response. In this direct method, an impulsive excitation perturbs the oscillator , and the steady state phase perturbation measured. The timing of the impulse with respect to the unperturbed oscillator's zero crossing is then incremental and the simulation repeated until the impulse has been walked through an entire cycle.
The impulse must have a small enough value to ensure that the assumption of linearity holds. Just as an amplifier step response cannot be valued properly with steps of arbitrary size, one must judiciously select the area of the impulse rather than blindly employing some fixed value. If one is unsure if the impulse chosen is sized properly, linearity may always be tested explicitly by scaling the size of impulse by some amount and verifying that the response scales by the same factor.
Last, some confusion persists about whether the LTV theory properly accommodates the phenomenon of amplitude-to-phase conversion that some oscillators exhibit. As long as linearity holds, the LTV theory provides the correct answer, provided that an exact ISF has been obtained. This is due to the fact that changes in the phase of oscillator due to an amplitude change appear in the impulse response if the oscillator. As noted in the preceding paragraphs the direct impulse response method is the most reliable one, as it makes no assumptions other than linearity. This reliability is in contrast with the approximate analytical approaches offered in [3].
While the close-in sidebands are dominated by phase noise, the far-out sidebands are greatly affected by amplitude noise. Unlike the induces excess phase, the excess amplitude A(t) due to the current impulse, decays with time. This decay is the direct result of the amplitude restoring mechanism always present in practical oscillators. The excess amplitude may decay very slowly or very quickly. Some circuits may even demonstrate an under damped second order amplitude response.

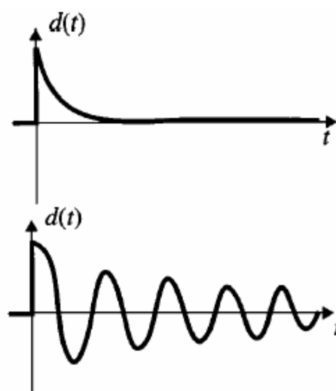


Figure 17: Over damped and under damped amplitude response

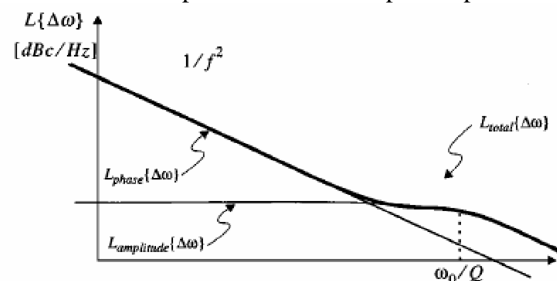


Figure 18 : phase amplitude and total output sideband power for the over damped response

In the context of the ideal LC oscillator a current impulse with an area Δq will induce an instantaneous change in the capacitor voltage, which in turn will result in a change in the oscillator

amplitude that depends on the instant of injection as show previously. The amplitude change is proportional to the instantaneous normalized voltage change ΔV/Vmax for small injected charge.

$$\Delta A = \Lambda(\omega_0 t)\frac{\Delta V}{V_{max}} = \Lambda(\omega_0 t)\frac{\Delta q}{q_{max}};\Delta_q << q_{swing} \qquad (38)$$

Where $\Lambda(\omega_0 t)$ is a periodic function that determinates the sensitivity of each point on the waveform to an impulse and is called the amplitude impulse sensitivity function. It is the amplitude counterpart of the phase impulse sensitivity function $\Gamma(\omega_0 t)$.
From a development similar to that of previously section, the amplitude impulse response can be written as:

$$h_A = \frac{\Lambda(\omega_0 t)}{q_{max}} d(t-\tau) \qquad (39)$$

Where d(t-τ) is a function that defines how the excess amplitude decays. Fig. 17 shows tow hypothetical examples of d(t) for a low Q oscillator with over damped response and high Q oscillator with under damped amplitude response. For most oscillators, the amplitude limiting system can be approximated as first or second orders.
The function d(t-τ) typically will thus be either a dying exponential or a damped sinusoid.
For a first order system

$$d(t-\tau) = e^{-\omega_0(t-\tau)/Q} \cdot u(t-\tau) \qquad (40)$$

Therefore the excess amplitude response to an arbitrary input current i(t) is given by the superposition integral

$$A(t) = \int_{-\infty}^{t} \frac{i(\tau)}{q_{max}} \Lambda(\omega_0 \tau) e^{-\omega_0(t-\tau)/Q} d\tau \qquad (41)$$

If i(t is a white noise source with power spectral density $\overline{i_n^2}/\Delta f$, the output power spectrum of the amplitude noise A(t can be shown to be

$$L_{amplitude}\{\Delta\omega\} = \frac{\Lambda_{rms}^2}{q_{max}^2} \cdot \frac{\overline{i_b^2}/\Delta f}{2 \cdot \left(\frac{\omega_0^2}{Q^2} + \Delta\omega^2\right)} \qquad (42)$$

Where Λ_{rms} is the rms value of $\Lambda(\omega_0 t)$. If L_{total} is measured, the sum of both L amplitude and L phase will be observed, and hence there will be a pedestal in the phase-noise spectrum at ω_0/Q, as shown in Fig.18. Also note that the significance of the amplitude response depends greatly on Λ_{rms}, which, in turn depends on the topology.

The result of the Leeson formula is a model for take into account the phase noise in microwave oscillator design. There are different reasons and different contribute in phase noise, but the result is only a fitting parameter. Hajimiri extend the theory and show as a particular architecture of oscillator based on cross coupled negative resistance active circuits and resonator used in

transmission but not in reflection has good performances in terms of phase noise. In fact for this particular architecture the phase noise depends also of resonator but the way in which the negative impedance circuit provide energy to compensate resonator loss improve the phase noise performances despite to resonant circuit with higher Q.
The impulse sensitivity function is applicable for impulsive resonator, in the cross couples the negative impedance circuit is impulsive, and a transistor leads signal for half way.
For a general class of oscillator in which the resonant circuit is used in reflection mode the LTV theory is not applicable, but often the transistor may results in A class or in C class. The only way to improve the phase noise performances is implement an high Q resonant circuit, and chose the start-up and then the oscillation frequency very close to the resonant frequency of tank circuit.
This result improve the Q for the oscillator that may be measured with the EYE method described in previously chapter.
The method developed in this thesis is not advantageous for phase noise and then is not very attractive for the design.
Our procedure has been developed only in order to solve the problem of the design for microwave reflection oscillators that has been misleading in more textbooks and hold paper.
Some authors as [2-3] have showed as the cross coupled oscillator has better performances in terms of phase noise compared with a Colpitts oscillator that works in C class.
Moreover the new improvement in CMOS technologies allows to achieve highest cut-off frequencies.

REFERENCES

[1] Michael Odyniec “RF and Microwave Oscillator Design” Boston London Artech House Publishers 2002.

[2] P.Andreani “A Time Variant Analysis of the 1/f^2 Phase Noise in CMOS Parallel LC-Tank Quadrature Oscillators” IEEE TRANSACTION ON CIRCUITS AND SYSTEMS VOL 53 August 2006

[3] P. Andreani Xiaoyan Wang “On the Phase Noise and Phase Error Performances of Multiphase LC CMOS VCOs”IEEE JOURNAL OF SOLID STATE CIRCUITS VOL 39 NO11 NOVEMBER 2004

- *T.H. Lee, The design of CMOS Radio Frequency Integrated Citcuits Cambridge Uk: Cambridge Univ. Press,1998*
- *D.B. Leeson, “A simple model of feedback oscillator noise spectrum”, inProc. IEEE, vol.54,Feb.1996,pp 329-330*
- *A.Hajimiri and T.Lee, “A general theory of phase noise in electrical oscillators”IEEE J. Solid-State Circuits, vol33, pp 179-194 Feb.1998*
- *Norwell, MA: Kluwer, The Design of Low-noise Oscillators. 1999*
- *A.Hajimiri and T.Lee “Design issues in CMOS differential LC oscillators, IEEE J.Solid-state Circuits, vol 34 pp 716-724, May 1999*
- *B Razavi, “A study of phase noise in CMOS oscillators”, IEEE J.Solid State Circuits vol 31, pp 331-343, Mar 1996*
- *F.X Kaetner, “Determination of the correlation spectrum of oscillators with low-noise”, IEEE Trans Microwave Theory Tech, vol 37,pp 90-101, Jan 1999*
- *T.I. Ahrens and T.H.Lee, “A 1.4Ghz,3mW CMOS LC lkow phase noise VCO using tapped bond wire inductance”, in Proc.ISLPED, Aug 1998, pp 16-19*
- *A.Hajimiri, S Liotyrakis and T.H.Lee, “Jitter and phase noise in ring oscillators”, IEEE J.Solid-State Circuits, vol 34, pp 790-804, June 1999*
- *A.Hajiniri T.H.Lee “Oscillator phase noise: A Tutorial”, IEEE Journal of Solid State circuits, vol 35, NO.3 MARCH 2000.*

5.1 An application example of State of The Art in Microwave Negative Impedance Oscillators Design.

Preface: this chapter collect the greatest part and results of my Ph.D. work at the university of Cagliari in the years 2005-2008.
The results exposed in this chapter are send in a paper “A New Method to Design Negative Impedance Oscillator Using The Solution Space Concept ” at the Microwave Theories and Techniques committee and are still under review.

In the design of microwave reflection oscillators the design procedure is not well standardized and depends strongly by experience of designer.
This procedure is convenient for single transistor oscillators, high frequencies and GaAs technologies, for low frequencies and others technologies the CMOS is better in terms of achieved performances.
The starting point of this approach consist in make unstable the transistor by introducing a reactive component in the common terminal. The concept of instability is developed y the classical theory of microwaves and consider the amplifier like a black boss model. By means of a small signal measurement and considering the amplifier plus a positive feedback like a two port network.

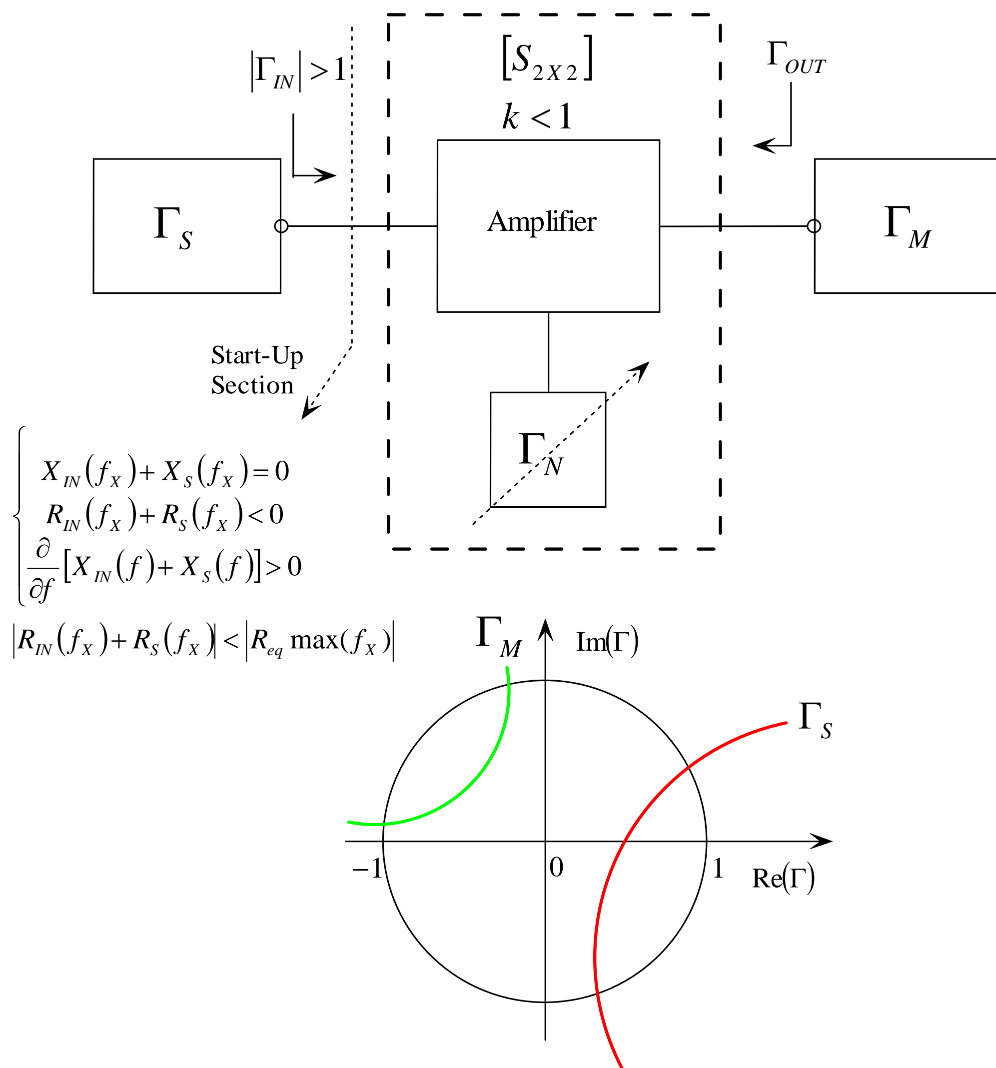


Figure 1

The matrix that represent the transistor is given by:

$$[S] = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

For a two port network and the positive feedback constituted by a reactive element introduced into the common terminal as Γ_N provide instability if the rollet stability factor (k) is smallest than one. The k factor is defined as

$$k = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2|S_{12}S_{21}|} \qquad (1)$$

Where $\Delta = S_{11}S_{22} - S_{12}S_{21}$ and the term at the denominator can be considered the gain of the active device.

A condition to achieve the instability through the insertion of a passive reactive element is that the active device must have a gain greater than zero.

When the instability is achieved it is possible to draw the stability circles that allows to view all possible values for passive loads that keep the transistor unstable.

The stability circles for the two ports equivalent circuits allows to separate all possible values that keep the potential instability condition namely all values that if implemented as a passive network at one terminal keep the reflection coefficient looking into the port greater or smaller than one.

An example could help to clarify this situation.

Lets consider for example the following circuit that show a biased transistor with its bias cirtuit including dc-feed and dc-block ideal elements.

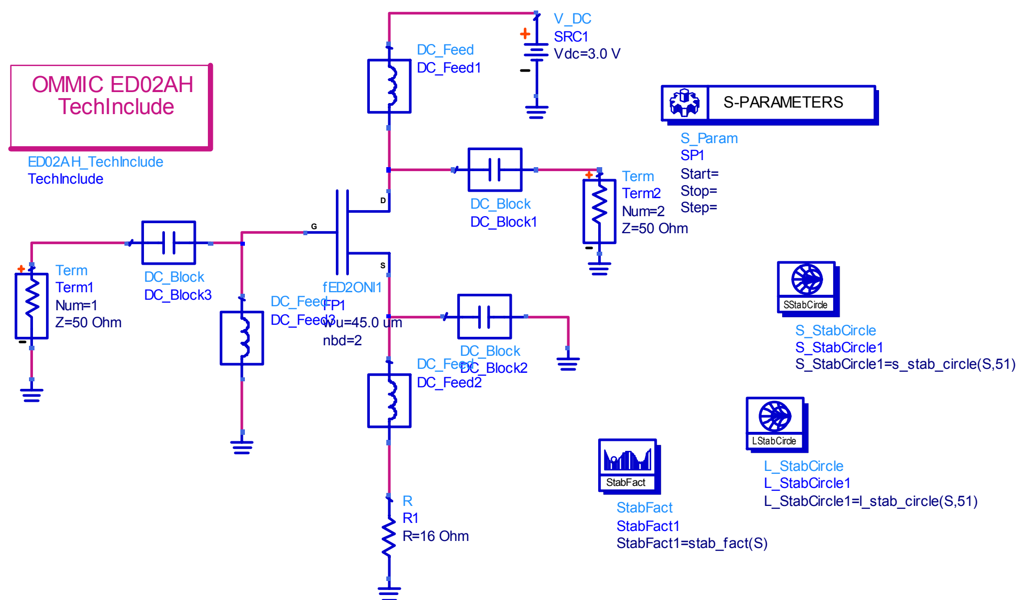


Figure 2

The simulation is reported in the following spot

freq	S(1,1)	S(1,2)	S(2,1)	S(2,2)	StabFact1
38.00 GHz	0.711 / -117.604	0.183 / 17.352	2.007 / 89.790	0.504 / -65.287	0.563

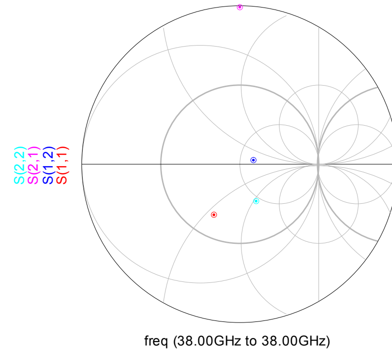


Figure 3

From simulation it is possible to see as the circuit is potentially unstable and the instability factor is small than one.
Once the S parameter for the two port network are known it is easy determinate the stability circles.
In the textbooks as “Collins” there is a well done description of instability circles and applications, therefore is more theoretic et does not has a good impact in the practice.
For the designer the stability circles are very helpful to determine the potentially instability and more of theoretical description are not usual in the practice.
We will explain the instability circles directly referring to the design of an oscillator so to give to the reader a more comprehensive description of the phenomena.
If we consider a general two port networks as showed in the following figure

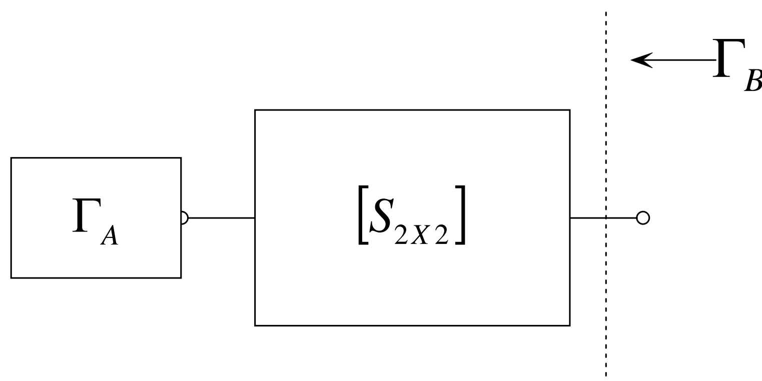


Figure 4

It is known that the reflection coefficient seen looking at the output port depends by the value of reflection coefficient used by the passive network, but also by the mgnitude of the reflection coefficient looked without any load (in the example the S_{22}).

$$\Gamma_B = S_{22} + \frac{S_{12}S_{21}\Gamma_A}{1 - S_{11}\Gamma_A} \qquad (2)$$

Where Γ_A is smallest than one because its represent a passive circuit reflection coefficient.
Γ_B must assume values smaller or greater than one in function of values assumed by S parameter and their combination with the Γ_A. Most important is the initial value of S_{22}.
The discussion of this topic can be found in any microwave textbook in this case we will derive the expression of the instability circles that are the limits region namely the corresponding Γ_A values for which Γ_B assumes the value one and the system is say to be in the limit stability condition.
Relationship (2) can be rewritten as

$$\Gamma_B = \frac{S_{22} - \Delta\Gamma_A}{1 - S_{11}\Gamma_A} \qquad (3)$$

Then it is possible to extract the equations for the centre and radius of the instability circle that are:

$$C_{\Gamma_A} = \frac{\left(S_{22} - \Delta S_{11}^*\right)^*}{|S_{22}|^2 - |\Delta|^2} \quad , \quad r_{\Gamma A} = \frac{|S_{21}S_{12}|}{|S_{22}|^2 - |\Delta|^2} \qquad (4)$$

Once the circle is drown it is possible to separate the Γ_A values that keep the system potentially unstable from these for which the system is stable.
The rule for separate the values is explained in this case for the system sketched in figure 4.If the instability circle intercept the unitary circle and the origin of the unitary circle fall inside the interception area and the S_{22} has a magnitude small than one then all points belonging at this area keep the system stable whereas the external points bring the system in the unstable region. Vice versa if in the same condition S_{22} was greater than one then all points belonging to interception area keep the system in the instability region.
Finally if the instability circle does not contain the origin of the unitary circle there are opposite considerations for the case in which $|S_{22}|>1$ the internal points keep instability vice versa in the opposite case. To explain the different situations is usefull to refears to the case of Fig2 and plotting the stability circle by looking from the term2.

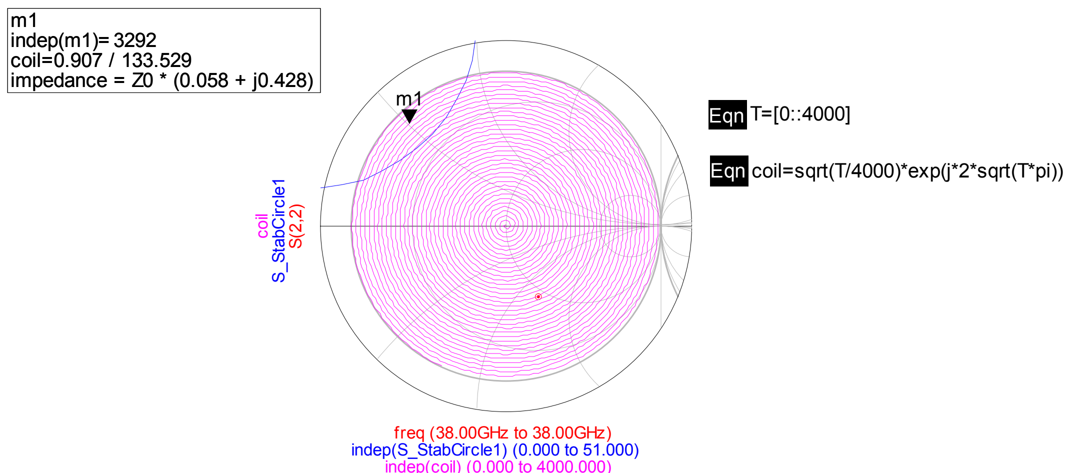


Figure 5

As show in figure 5 the coil curve is a numerical curve “a trick” that I’ve used in ADS to take into account all the possible reflection coefficient values in the Smith chart and through I can now chose a value by means of a marker.
In agreement with theory of stability if in the input port of my circuit I put a passive load network which has the reflection coefficient indicated by marker I would bring the S22 outside the unitary circle and then create a potential instability condition for the output port. By implementing the value through a simple passive network the circuit will becomes

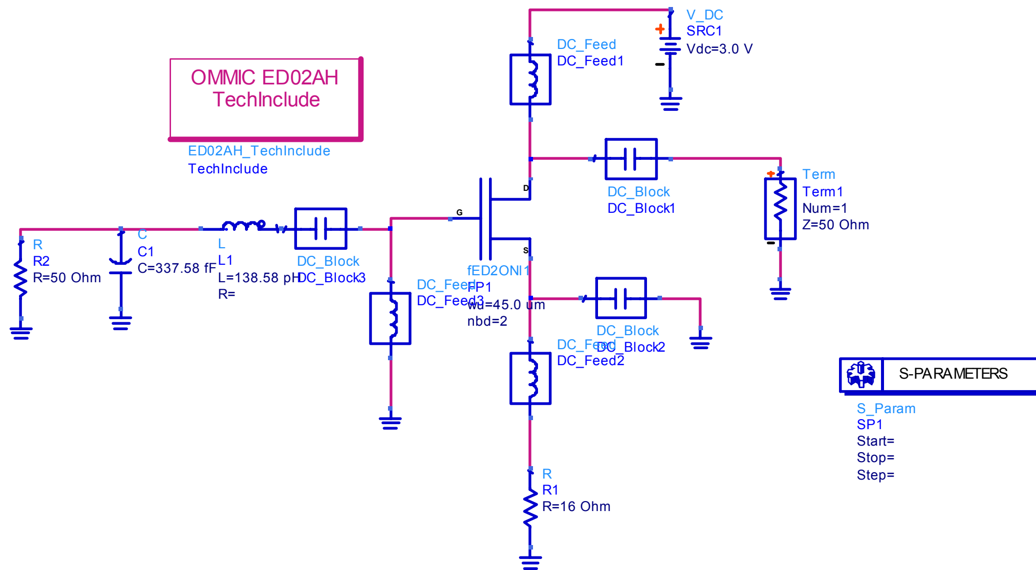


Figure 6

And the reflection coefficient looked by the output port will have a magnitude value greater than one

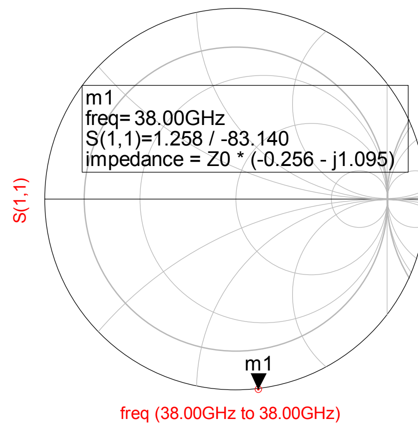


Figure 7

The whole set of possible combination are summarized in Figure 8 and obviously the same cases appear by opposite port.
This theory is used to design oscillators but is not very simple to apply.

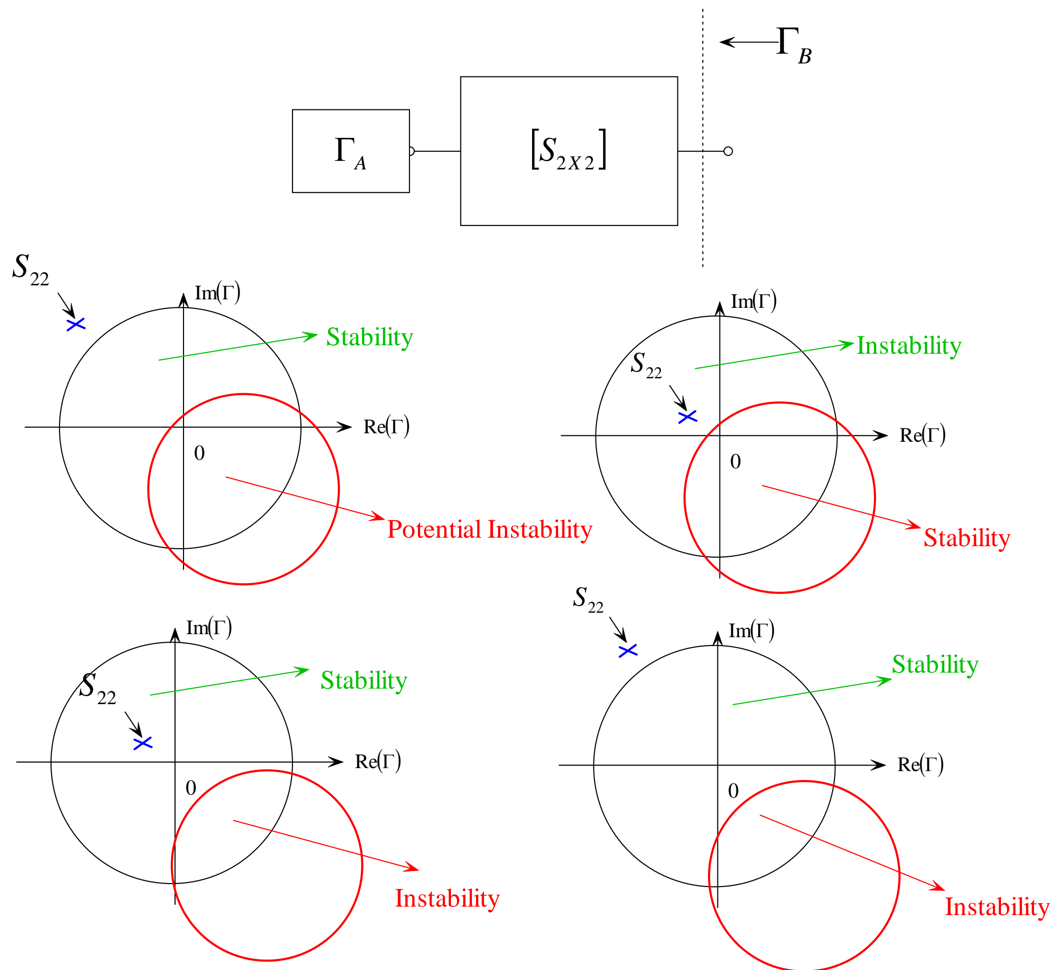


Figure 8

In fact there are different reason for which the design procedure of microwave reflection oscillator are not well explained by official theory.
The first reason is the possibility to loose the start-up coefficients and then the Kurokawa hypothesis. If the resonance property of the system is derived by composition of both resonant circuit and active devices then it is possible that the circuit does not meet high performances in terms of reliability because the active circuit may change its property during his work as for example change the S-parameter as increasing the temperature. Then the resonance properties (the pair of complex and conjugate poles) must be produced only by the resonant circuit while the active circuit must have the only function to restore the lose energy in the resonant circuit. The first effect for which it is difficult to apply this methodology is that the designer must satisfy at the same time the start-up conditions and the constraints derived form the output network in order to achieve a signal with good properties.
This is not easy because the output signal depends by nature of the output matching network but also by others factors discovered during my Ph.D. activity and then ever mentioned before.
My studies have put in evidence the existence of a Figure of Merit that allows to paint the possibility to find a solution in a multidimensional space called The Solution Spaces.
The mathematics passages that have evidenced the existence of this space are showed after.
Now we can show as the usual design procedure is time consuming and expensively in terms of work time and the design of a microwave reflection oscillator that is basically a very simple circuit can becomes very difficult.
Now we suppose to built a VCO starting by the circuit showed in figure 1. The circuit is unstable and does not need to further adjustments to guarantee the instability.

We have show the instability circle by side of resonant circuit but not the circle by side of the output matching networks.
Once the circuit is unstable the instability circles are used in order to find a value for the output matching network that generate a possible reflection coefficient able to start an oscillator as presented by the fig.1.
The following figure show both instability circles for the transistor. The source is grounded through a DC-BLOCK in order to have potential instability therefore we are looking for an oscillator which output is picked-off by DRAIN terminal.

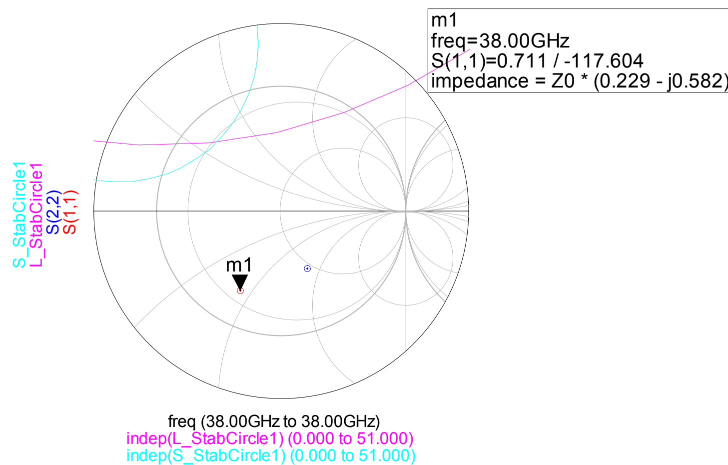


Figure 9

The first operation consist in generate at the gate a reflection coefficient greater than one in magnitude.
Then we must generate a passive network and applying this at the drain port.
In order to filter the signal the output port must be reliable with a low pass filter if it is possible.
In agreement with instability rules the value of Γ_{OUT} must be chosen inside the magenta arc.
The general design procedure does not impose criteria in order to chose a proper network but suggests to operate only under try and error. At same way it suggest to operate in order to obtain potential instability without take care at impose at first time the best reaction and without consider if the transistor must operate as a good oscillator when the output is picked-out by the DRAIN. Then we can choose a case value as for example that indicated in the figure

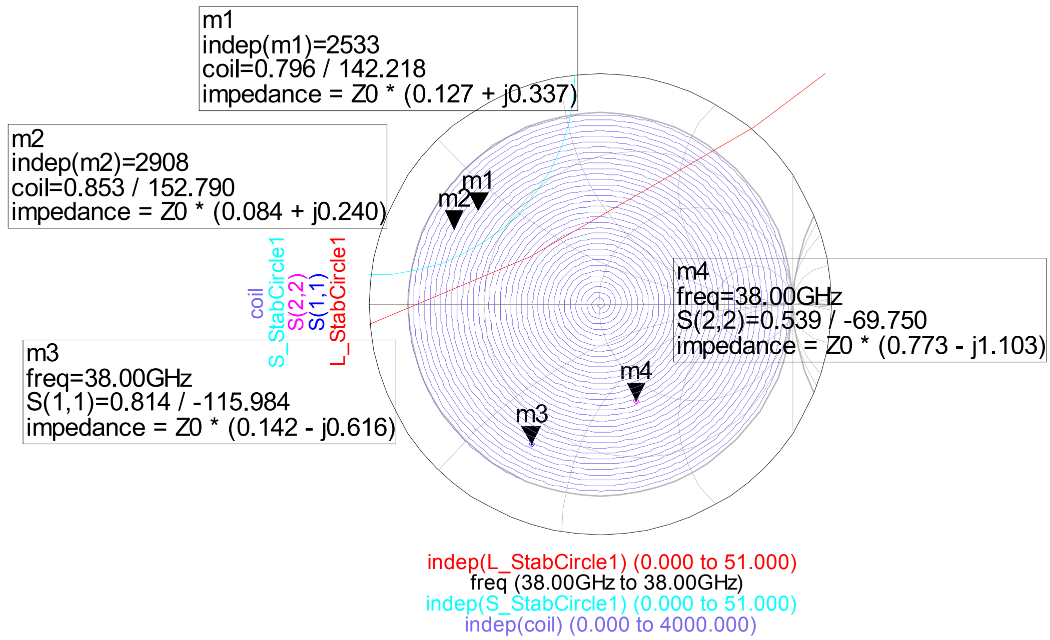


Figure 10

This value is feasible through the following networks and we choose that has a low pass behaviour namely the fort circuit with a series inductance and a shunt capacitance, in this way the higher harmonics are cutted. Moreover if I use a low pass network there is a greater possibility that the reflection coefficient does not meet the self resonating condition as explained in chapter 2.

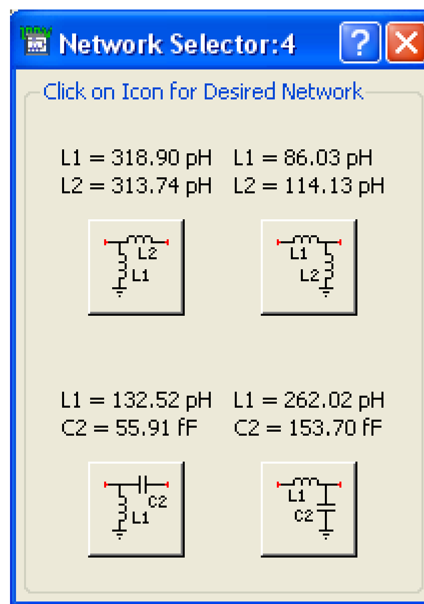
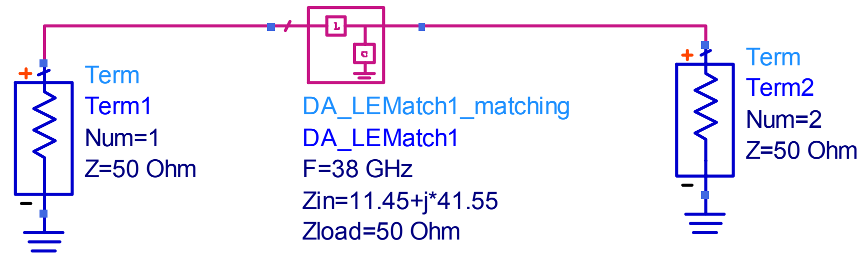


Figure 11

S-PARAMETERS

S_Param
SP1
Start=1.0 GHz
Stop=60 GHz
Step=0.1 GHz



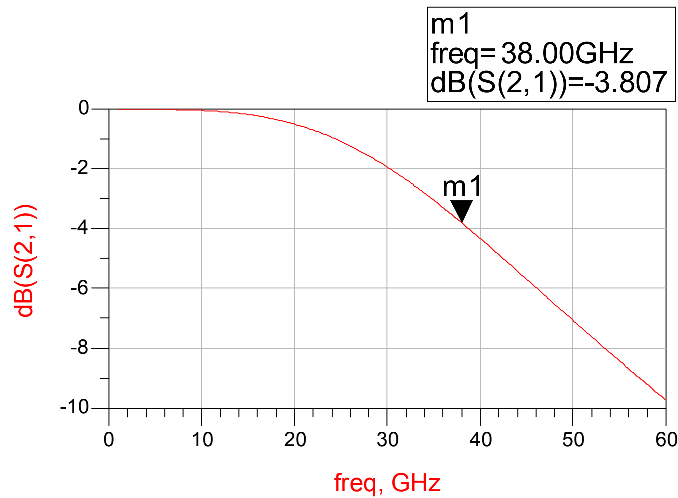
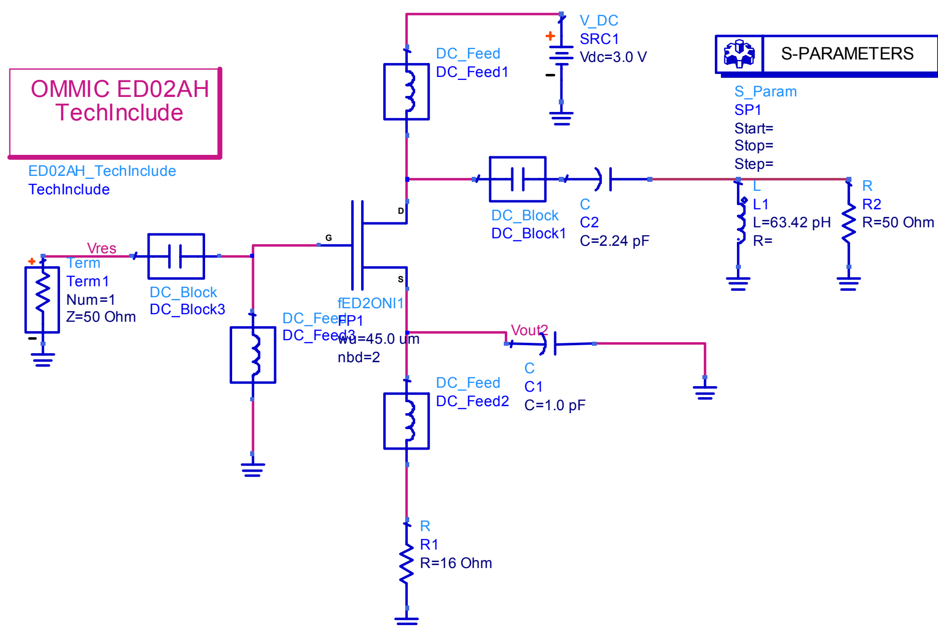
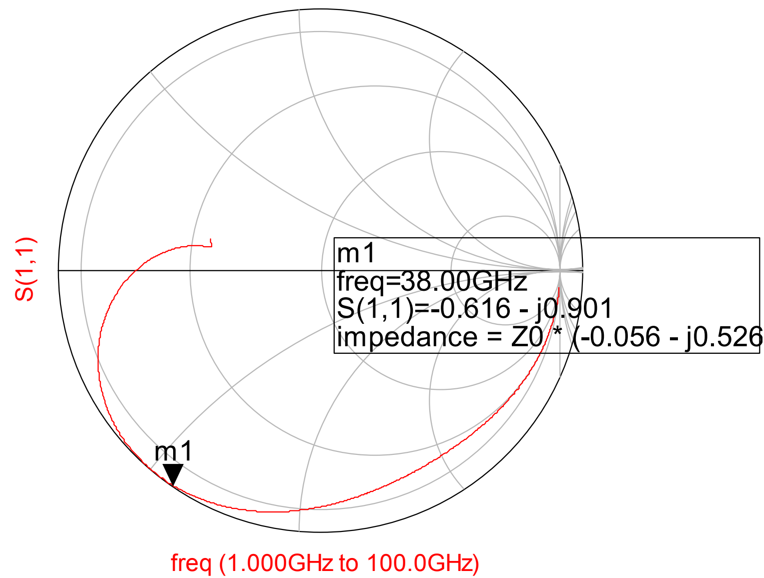


Figure 12

The circuit without resonator is showed in the figure below



Now we can apply the start-up conditions in terms of impedances or admittances in order to achieve an oscillation
For this case if my Z_{IN} is:



Then for the Γ_{IN} of my circuit I can use different Γ_S values for example by considering a series or a parallel resonant circuit. In this case I can consider a series resonant circuit because the resonator frequency is near to the star-up frequency and this is better for the phase noise in agreement with the LTI theory.

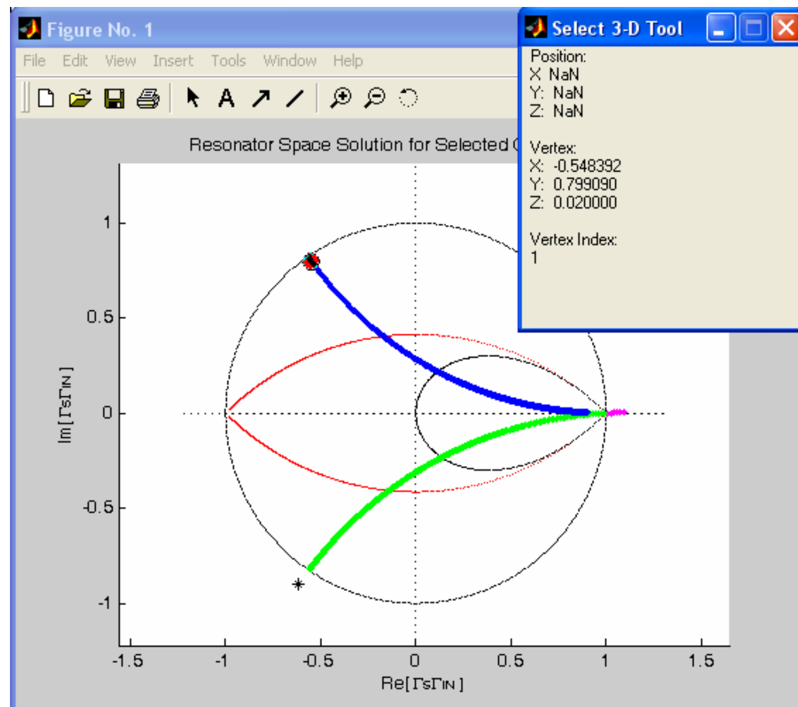


Figure 13

A possible value for the Γ_S may be Γ_S=-0.548+j0.799 with a series resonance type.
Then ZS=50(1+ΓS)/(1-ΓS)=1+j26.3199 that may be implemented through a simple L (jwL= j26.3199 at the frequency of 38GHz) inductor and a series resistance to test the oscillation. If the test produces a good result the frequency determining circuit must be substituted with a series resonant circuit.

The circuit used to test the oscillation is then

$$L = \frac{X}{2\pi f} = \frac{0.769 \cdot 50}{2\pi \cdot 38^9} = 161 pH \qquad (5)$$

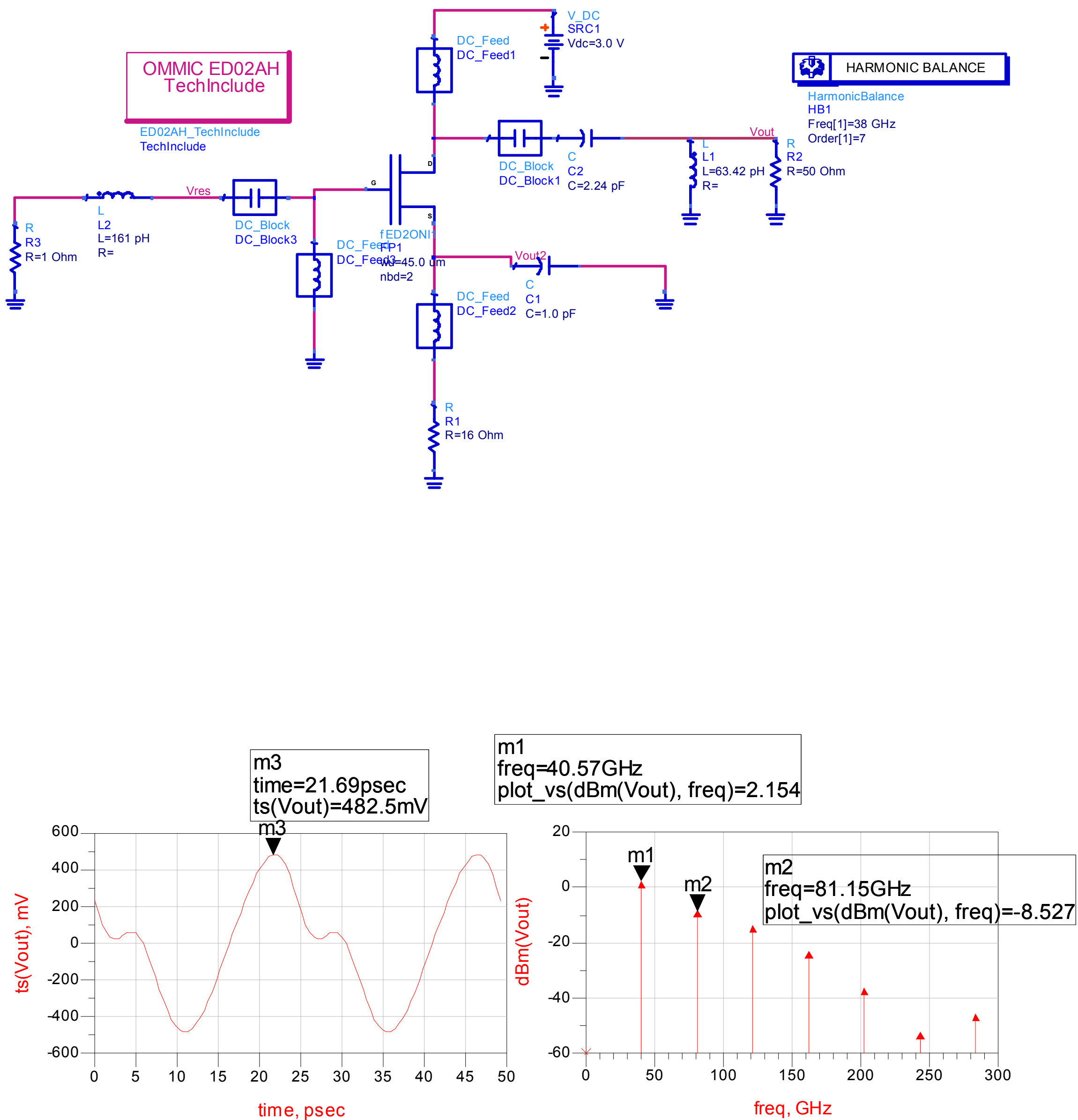


Figure 14

That is a result but not a good result because the output shape has a second harmonic very high, and the amplitude of the signal is very small only 2.1 dBm
To improve the performance of the oscillator now it is necessary to change the output network or to change the feedback and then change the resonant circuit. In addition it is necessary to use the non linear CAD design procedures in order to improve the performance of the oscillator. All these operation must be done without any guideline but the designer must move based on its experience and this operation is not ever feasible.
In order to overcome these limits I propose a new design methodology based on the concept of solution space.

5.2-THE SOLUTION SPACE APPROACH
The basic idea and the generation

The basic idea to Space Solution development it consist on considering the transistor as a three port network and to write the relationship between reflection coefficients representing the passive networks that can be connected at the outputs and the scattering matrix of the transistor

The Oscillator Space Solutions

In order to extract the equations that we allows to define the Space Solution generated by the three port networks we can consider the transistor as a three port network.

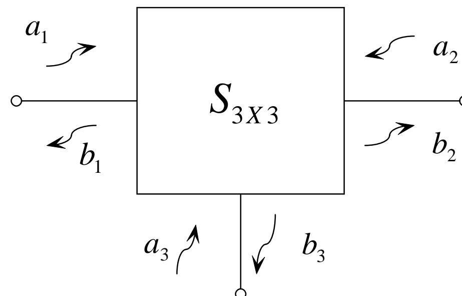


Figure 15

The reflection coefficients for a section can be wrote as the ratio between the reflected and incident waves.
Then we can define:

$$\begin{aligned}&\Gamma_S = \frac{a_1}{b_1}; \Gamma_{IN} = \frac{b_1}{a_1}; \Gamma_L = \frac{a_2}{b_2};\\&\Gamma_{OUT} = \frac{b_2}{a_2}; \Gamma_F = \frac{a_3}{b_3}; \Gamma_{OUTF} = \frac{b_3}{a_3}\end{aligned} \qquad (6)$$

Mathematically the reflection coefficient product of terms seen at the opposite sides of a section give one. Actually this does not represent a situation true in real case but only when the design of an oscillator is performed. The reflection coefficient seen looking into a specific port of active device when other ports are loaded can be greater or smaller than one than one. When this is greater indicate potential instability because in this situation exist the possibility for the product to assume the value one in others words this can indicate the possibility of oscillation because the system present a negative impedance at the considered port. Actually the reflection coefficient of active

device depends also by the scattering matrix and can be verified only for some value of reflection coefficient then can be used to separate the reflection coefficient values that give instability from those that keep the stability. For those last the it is impossible that the reflection coefficient product at the section would assume the unitary value.
The relationship that ties the reflection coefficient to the S-Parameter matrix can be found through previously showed equation and the following system

$$\begin{cases} b_1 = a_1 S_{11} + a_2 S_{12} + a_3 S_{13} \\ b_2 = a_1 S_{21} + a_2 S_{22} + a_3 S_{23} \\ b_3 = a_1 S_{31} + a_2 S_{32} + a_3 S_{33} \end{cases} \quad (7)$$

Depend by the way to solve the system it is possible to represent different physic situations and different kind of circuit implementation.

- ***Solution for Oscillator design***

In order to find the equations that allows to define the Space Solution for Oscillator design we can substitute the definition of Γ_F in the third equation of (2) and find the expression for a_3.

$$b_3 = \frac{a_3}{\Gamma_F}$$

$$\begin{cases} b_1 = a_1 S_{11} + a_2 S_{12} + a_3 S_{13} \\ b_2 = a_1 S_{21} + a_2 S_{22} + a_3 S_{23} \\ \dfrac{a_3}{\Gamma_F} = a_1 S_{31} + a_2 S_{32} + a_3 S_{33} \end{cases}$$

$$\frac{a_3}{\Gamma_F} = a_1 S_{31} + a_2 S_{32} + a_3 S_{33} \rightarrow a_3\left(\frac{1 - S_{33}\Gamma_F}{\Gamma_F}\right) =$$

$$a_1 S_{31} + a_2 S_{32} \rightarrow a_3 = \frac{a_1 S_{31}\Gamma_F + a_2 S_{32}\Gamma_F}{1 - S_{33}\Gamma_F}$$

Then the system can be written as follows

$$\begin{cases} b_1 = a_1 \dfrac{\left(S_{11} - S_{11}S_{33}\Gamma_F + S_{31}\Gamma_F\right)}{1 - S_{33}\Gamma_F} + a_2 \dfrac{\left(S_{12} - S_{12}S_{33}\Gamma_F + S_{32}\Gamma_F\right)}{1 - S_{33}\Gamma_F} \\ b_2 = a_1 \dfrac{\left(S_{21} - S_{21}S_{33}\Gamma_F + S_{31}\Gamma_F\right)}{1 - S_{33}\Gamma_F} + a_2 \dfrac{\left(S_{22} - S_{22}S_{33}\Gamma_F + S_{32}\Gamma_F\right)}{1 - S_{33}\Gamma_F} \end{cases}$$

$$\rightarrow \begin{cases} b_1 = a_1 \overline{S_{11}} + a_2 \overline{S_{12}} \\ b_2 = a_1 \overline{S_{21}} + a_2 \overline{S_{22}} \end{cases}$$

$$\Gamma_{IN} = \frac{b_1}{a_1} \rightarrow b_1 = \Gamma_{IN} a_1 = a_1 \overline{S_{11}} + a_2 \overline{S_{12}} \rightarrow a_1 = \frac{a_2 \overline{S_{12}}}{\left(\Gamma_{IN} - \overline{S_{11}}\right)}$$

$$b_2 = a_2 \frac{\overline{S_{12}}\,\overline{S_{11}} + \overline{S_{22}}\Gamma_{IN} - \overline{S_{22}}\,\overline{S_{11}}}{\left(\Gamma_{IN} - \overline{S_{11}}\right)}$$

$$-\Delta = \left(-\overline{S_{22}}\,\overline{S_{11}} + \overline{S_{12}}\,\overline{S_{11}}\right)$$

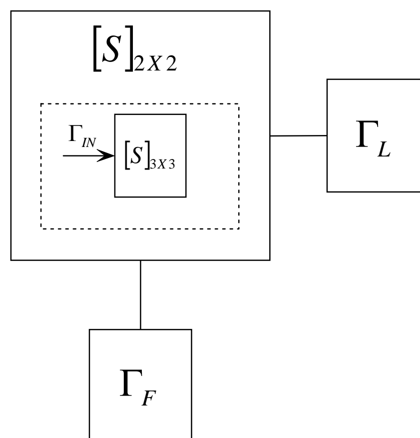
The expression for the reflection coefficient at the input port when other two ports are loaded with passive networks can be written as a function of S parameters Γ_F and Γ_L by classical form of bilinear transformation

$$\Gamma_{IN} = \frac{\left(\overline{S_{11}} - \Delta\Gamma_L\right)}{\left(1 - \overline{S_{22}}\Gamma_L\right)} \qquad (8)$$

Or in more explicit form as

$$\Gamma_{IN} = \left(S_{11} + \frac{S_{13}S_{31}\Gamma_F}{1 - S_{33}\Gamma_F}\right) + \frac{\left(S_{12} + \frac{S_{13}S_{32}\Gamma_F}{1 - S_{33}\Gamma_F}\right)\cdot\left(S_{21} + \frac{S_{23}S_{31}\Gamma_F}{1 - S_{33}\Gamma_F}\right)\cdot\Gamma_L}{1 - \left(S_{22} + \frac{S_{32}S_{23}\Gamma_F}{1 - S_{33}\Gamma_F}\right)\cdot\Gamma_L} \qquad (9)$$

From this relationship it is possible to map the spatial disposition of reflection coefficient seen at the input port for the system depicted in the figure below



The obtained map show for a transistor the capabilities to obtain a couple of loads that give the negative resistance at the input.
It is possible then choose a specific value in order to exclude the possibility of self resonance due to passive loads and make the choice in high density points area to obtain a greater robustness compared to constructive tolerances.

A Typical map is showed in the following figure in which the value for the design of an oscillator are outside the unitary circumference.

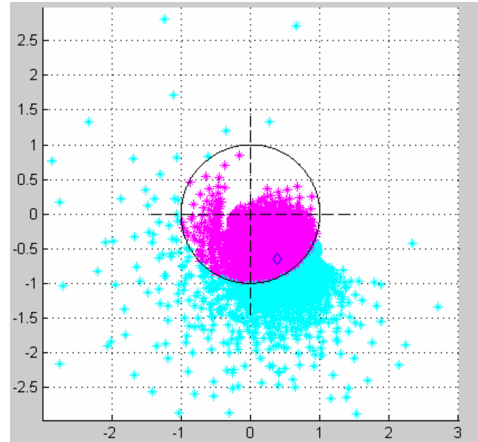


Figure 16

By choosing a specific value of Γ_{IN} it is possible to invert the equations and compute the S2x2 matrix for the system showed below

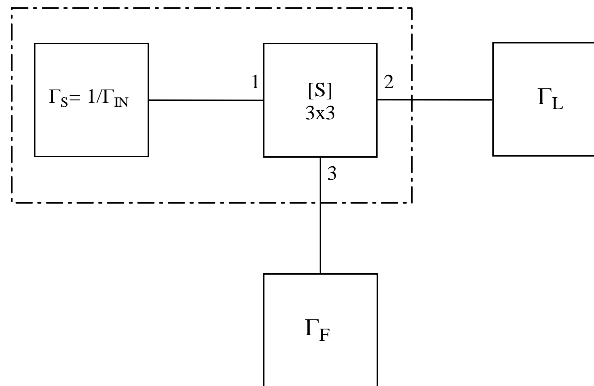


Figure 17

and then compute the new matrix starting from the S3x3 matrix of the transistor but imposing the constraints on $\Gamma_S=1/\Gamma_{IN}$ at the input port and extracting the new X matrix.
The X matrix does not exist because represent only a mathematical condition, but for the design of oscillators can give a starting point to study the transistor configuration.
Then starting from equation (2) and by expressing b_1 as a function of Γ_{IN} the system becomes:

$$b_1 = a_1\Gamma_{IN}$$

$$\begin{cases} a_1\Gamma_{IN1} = a_1S_{11} + a_2S_{12} + a_3S_{13} \\ b_2 = a_1S_{21} + a_2S_{22} + a_3S_{23} \\ b_3 = a_1S_{31} + a_2S_{32} + a_3S_{33} \end{cases}$$

$$a_1(\Gamma_{IN} - S_{11}) = a_2S_{12} + a_3S_{13} \rightarrow a_1 = \frac{a_2S_{12} + a_3S_{13}}{(\Gamma_{IN} - S_{11})}$$

Then the system can be written as follows

$$\begin{cases} b_2 = a_2\left(\dfrac{S_{12}S_{21} + S_{22}\Gamma_{IN} - S_{22}S_{11}}{\Gamma_{IN} - S_{11}}\right) + a_3\left(\dfrac{S_{13}S_{21} + S_{23}\Gamma_{IN} - S_{23}S_{11}}{\Gamma_{IN} - S_{11}}\right) \\ b_3 = a_2\left(\dfrac{S_{12}S_{31} + S_{32}\Gamma_{IN} - S_{32}S_{11}}{\Gamma_{IN} - S_{11}}\right) + a_3\left(\dfrac{S_{31}S_{13} + S_{33}\Gamma_{IN} - S_{33}S_{11}}{\Gamma_{IN} - S_{11}}\right) \end{cases}$$

$$\begin{cases} b_2 = a_2\overline{X_{11}} + a_3\overline{X_{12}} \\ b_3 = a_2\overline{X_{21}} + a_3\overline{X_{22}} \end{cases}$$

Then for the two ports equivalent circuit it is possible to write the equation for the centre and radius of stability circles. These circles give information about which couples of reflection coefficients for passive networks allow to bring the transistor in the desired position of the smith chart.

$$\begin{cases} C_F = \dfrac{(\overline{X_{22}} - \Delta\overline{X_{11}})^*}{\left|\overline{X_{22}}\right|^2 - |\Delta|^2} \\ R_F = \left|\dfrac{\overline{X_{12}}\,\overline{X_{21}}}{\left|\overline{X_{22}}\right|^2 - |\Delta|^2}\right| \end{cases} \begin{cases} C_L = \dfrac{(\overline{X_{11}} - \Delta\overline{X_{22}})^*}{\left|\overline{X_{11}}\right|^2 - |\Delta|^2} \\ R_L = \left|\dfrac{\overline{X_{12}}\,\overline{X_{21}}}{\left|\overline{X_{11}}\right|^2 - |\Delta|^2}\right| \end{cases} \quad (4)$$

The radius of circles is inversely proportional to the value of modulus of the reflection coefficient and indicate a potential greater negative resistance at the port. In the design of oscillators high level of negative resistance at the output indicate the terminal in which there is a grown of the output signal and then the terminal of interest is indicated for design the matching network.

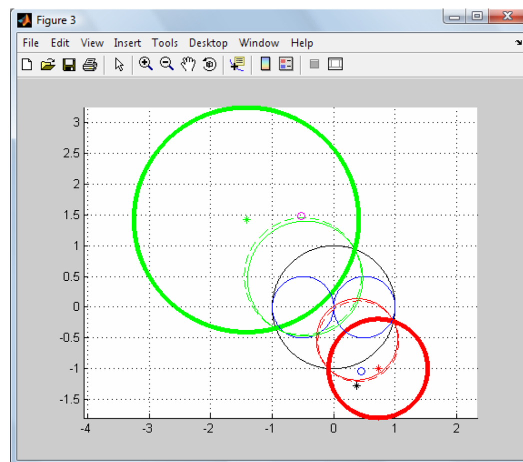


Figure 18

The green circle represent the instability circle for the load (Γ_L) network and then separate all Γ_L value that make the reflection coefficient seen looking into the third port greater than one from those that make them smaller than one. In reference to previously figure, the values for the instability will be inside or outside the stability circle if the circle contains the origin and the reflection coefficient looking into the unloaded port is greater or smaller than one. In the example showed in the figure the stability green circle does not contain the origin and the reflection coefficient Γ_{OUT} of the unloaded port is smallest than one. In this case the values of Γ_L that belong to the intersection between the unitary circle and the stability circle generate a reflection coefficient looking into the (Γ_F port) greater than one and then can be used to destabilize the circuit. The same

thing happens for the Γ_F network and the Γ_{OUT} reflection coefficient. By using the relationship defined by bilinear transformation that ties the reflection coefficients of the two port Scattering matrix to the reflection coefficient that representing the passive networks applied to the output ports it is possible to define a univocal relationship between the two areas and then determine a set of couples for Γ_F and Γ_L that allows to obtain the chosen input reflection coefficient Γ_{IN} .

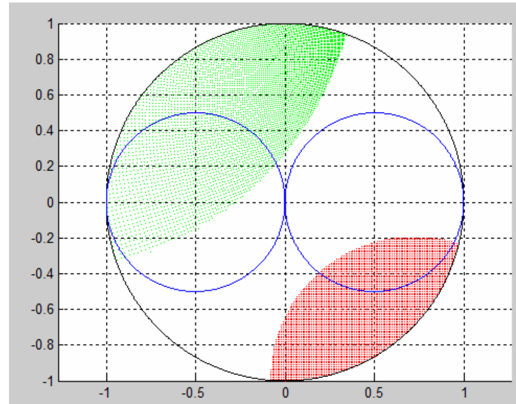
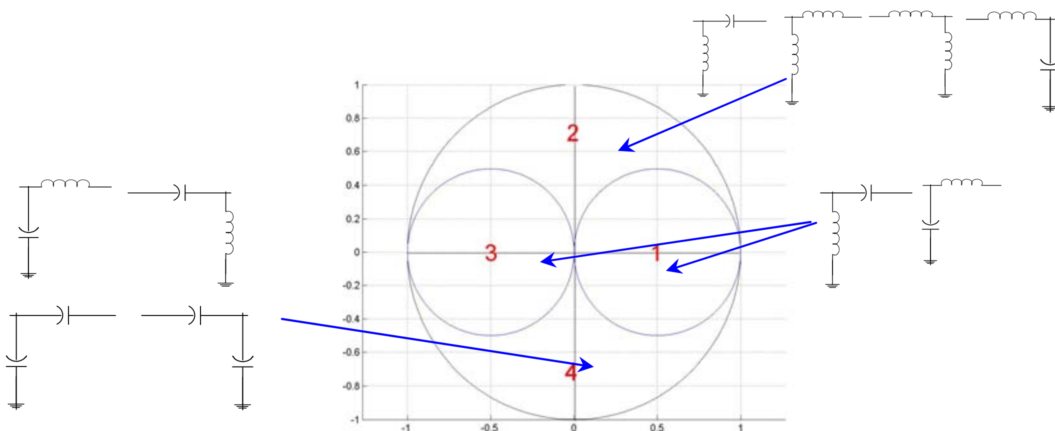


Figure 19

Where the achieved relationship to obtain the map can be written as:

$$\Gamma_L = \frac{1-\overline{X_{22}}\Gamma_F}{\overline{X_{11}}-\Delta\Gamma_F} \quad \Gamma_F = \frac{1-\overline{X_{11}}\Gamma_L}{\overline{X_{22}}-\Delta\Gamma_L} \qquad (9)$$

In this context the designer can choose the couples of values that allows to obtain the desired reflection coefficient at the input port. The blue circles inside the unitary circle have the function to separate in the Smith chart four different areas which correspond different kinds of networks and then help to chose the designer to find in few time the better solution for the case of study



This is only a first approximate step to design the oscillator and to estimate a first possible configuration of the active device. To complete the design the resonant circuit topology series-parallel can be chosen and designed to guarantee at the section the start-up and the build-up for the signal.

The choice can be done by means of the design chart that show for a given reflection coefficient how the resonant circuit can be chosen. For example let us consider the case of reflection coefficient
Γ_{IN} =0.8-j0.95 end computing the reflection coefficients of Γ_S extracted by the product that guarantee the start-up for the series and the parallel topologies.

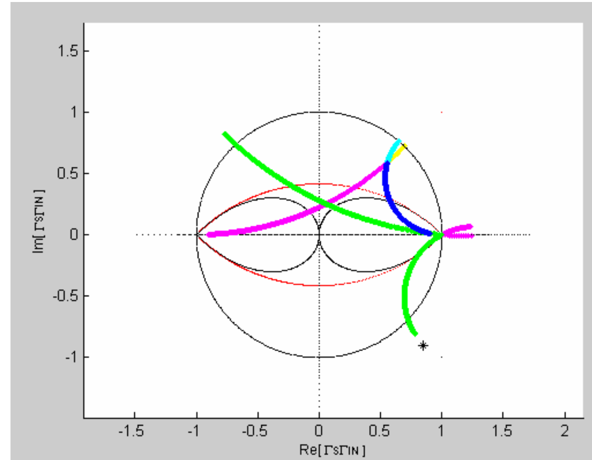


Figure 20

5.3-The PFOM and the decrypting of solution space information (The new design method).

The main feature of the solution space is the possibility to put in both numerical and graphical forms under the hands of the designers all the solutions and triplet values Γ_{IN} Γ_{OUT} Γ_F able to generate an oscillator for a given transistor. This allows to overcome the time consuming try and error methodology and to get in fast way the best solution achievable for the chosen technology.

This property is possible tanks to the discover of a way to ordinate the pairs of passive values in order to maximize the amplitude of the signal in a specific terminal.

In fact the possibility to develop a signal or in the Drain or in the Source terminal depends on some transistor characteristics that cannot correctly defined in a formal way.

The main discover of my works is that the transducer Gain that is considered as indefinite at the oscillation is not indefinite but plays a very important role in the design and must take into account in order to obtain good performance for the designed oscillator. For these reasons and by following the Kurokawa works in which is refer to a impedance without calling (Impedance) we refer to transducer gain but without calling transducer gain and define a Figure of Merit (FOM) that take into account the amplitude of generated signal for the circuit.

Since exist a lot of Γ_N and Γ_M pairs into the *Loads Space,* it is important to define design criteria that could help the designer to chose in fast way those that better satisfy the design aims. In the practice, it has been observed that from the course of some small-signal parameters in the *Loads Space* (at start-up condition), is possible to foresee the system behaviour in stationary oscillation condition.

For example, we derive here one figure-of-merit qualified for the output power aims. In fact, once the Γ_S value for ensuring the start-up is fixed, we define

$$\begin{cases} K_M = \dfrac{\left(1-|\Gamma_S|^2\right)|M_{21}|^2\left(1-|\Gamma_M|^2\right)}{\left|(1-M_{11}\Gamma_S)(1-M_{22}\Gamma_M)-M_{12}M_{21}\Gamma_S\Gamma_M\right|^2} \\ \text{where} \\ M = \begin{bmatrix} S_{11}+\dfrac{S_{13}S_{31}\Gamma_N}{1-S_{33}\Gamma_N} & S_{12}+\dfrac{S_{13}S_{32}\Gamma_N}{1-S_{33}\Gamma_N} \\ S_{21}+\dfrac{S_{23}S_{31}\Gamma_N}{1-S_{33}\Gamma_N} & S_{22}+\dfrac{S_{23}S_{32}\Gamma_N}{1-S_{33}\Gamma_N} \end{bmatrix} \end{cases} \quad (10)$$

and

$$\begin{cases} K_N = \dfrac{\left(1-|\Gamma_S|^2\right)|N_{21}|^2\left(1-|\Gamma_N|^2\right)}{\left|(1-N_{11}\Gamma_S)(1-N_{22}\Gamma_N)-N_{12}N_{21}\Gamma_S\Gamma_N\right|^2} \\ \text{where} \\ N = \begin{bmatrix} S_{11}+\dfrac{S_{12}S_{21}\Gamma_M}{1-S_{22}\Gamma_M} & S_{13}+\dfrac{S_{12}S_{23}\Gamma_M}{1-S_{22}\Gamma_M} \\ S_{31}+\dfrac{S_{21}S_{32}\Gamma_M}{1-S_{22}\Gamma_M} & S_{33}+\dfrac{S_{23}S_{32}\Gamma_M}{1-S_{22}\Gamma_M} \end{bmatrix} \end{cases} \quad (11)$$

and finally the figure-of-merit as

$$PFOM \equiv K_M - K_N \quad (12)$$

We have noted that the output power tends to a maximum in the *Loads Space*, when the $|PFOM|$ tends to a maximum as well. This condition generates a set of values $\{PFOM_i\}$ for which the greater output power will be present in Γ_M load if $PFOM_i > 0$ and vice versa in Γ_N load if $PFOM_i < 0$. Intuitively it is possible to assert that if, at the start-up, the system is in the occurrence sketched in Fig. 21 (in agreement with $PFOM_i > 0$), then the system state evolves conserving the power difference between the two loads.

We have noted also that the power will tends to the same value in both loads (Γ_M , Γ_N) in steady-state, if $|PFOM| \approx 0$. The output power will be much greater at a Γ_M , Γ_N pair for which $K_M \approx K_N$ assumes highest value. The inherent *PFOM* cases are summarized in Fig. 22.

The *PFOM* previously described does not take in to account the distortion of the generated signal which can be reduced, as example, by implementing the Γ value as low pass network.

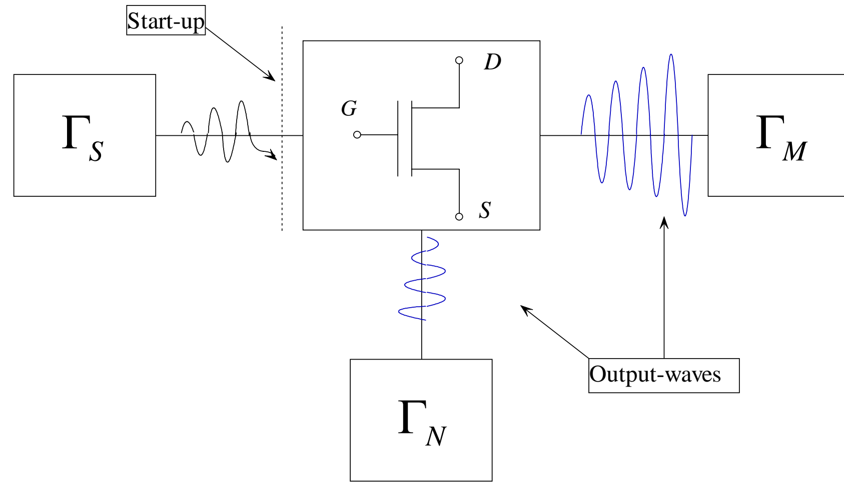


Figure 21

$$|PFOM| \rightarrow \max \quad and \quad K_M > K_N \Rightarrow \boxed{\text{maximum power on M load (the Drain in previous figure)}}$$

$$|PFOM| \rightarrow \max \quad and \quad K_M < K_N \Rightarrow \boxed{\text{maximum power on N load (the Source in previous figure)}}$$

$$PFOM \approx 0 \Rightarrow \boxed{\text{equal power on both M and N loads}}$$

$$PFOM \approx 0 \quad and \quad K_M \approx K_N \rightarrow \max \Rightarrow \boxed{\text{higher power on both M and N loads}}$$

Figure 22

These concepts allow to define a method to design the negative impedance oscillators taking into account the different design aims. This method is here presented as a step by step procedure oriented to design optimization of single-transistor oscillator at microwave.

The first step consists in the choose of the transistor and the bias point through which to reach the required performances. The frequency behaviour of the biased transistor is disclosed by the S_{3x3} matrix which can be acquired through a S-parameter simulator from the proper physical-mathematical model. By studying the S parameters trends, the designer decides on the start-up frequency (f_X) selection on own experience basis.

In the second step the *Negative Impedance Space* is generated by calculating the function (3) on a uniform grid into the unitary circle on both Γ_M and Γ_N plans, with a tool feasible by means of any technical computing software. If this *Space* is satisfactory then the Γ_{IN0} in a points high density area, is chosen, else the f_X or the bias conditions and/or the transistor are changed to obtain different one.

In the third step, after the Start-Up Space has been generated for the Γ_{IN0} value, remains determined the resonant circuit topology (resonance series or parallel). By using proper tools, the specifications that guide their synthesis (as resonance frequency f_r, merit factor Q_U, Γ_S at f_X) have been extract in order to minimize the phase noise as explained in [4], through the oscillator Q optimization.

The fourth step consists in generate the *Loads Space* for the selected Γ_{IN0} and Γ_S and in extract the *PFOM* information. Then the Γ_M and Γ_N pairs can be chosen on basis of output power aim. Finally the schematic of the oscillator with lumped ideal elements can be easily carried out. In order to

foresee the phase noise, the instability test presented in [4] can be used to measure the quality factor Q of the oscillator at the Start-Up. In the last step, by means of harmonic balance simulators the performances are checked. After the oscillator layout is generated and the performances are checked by a model near to the real circuit. Finally the optimization procedures are applied and the prototype is realized.

The workflow presented is summarized in Figure.

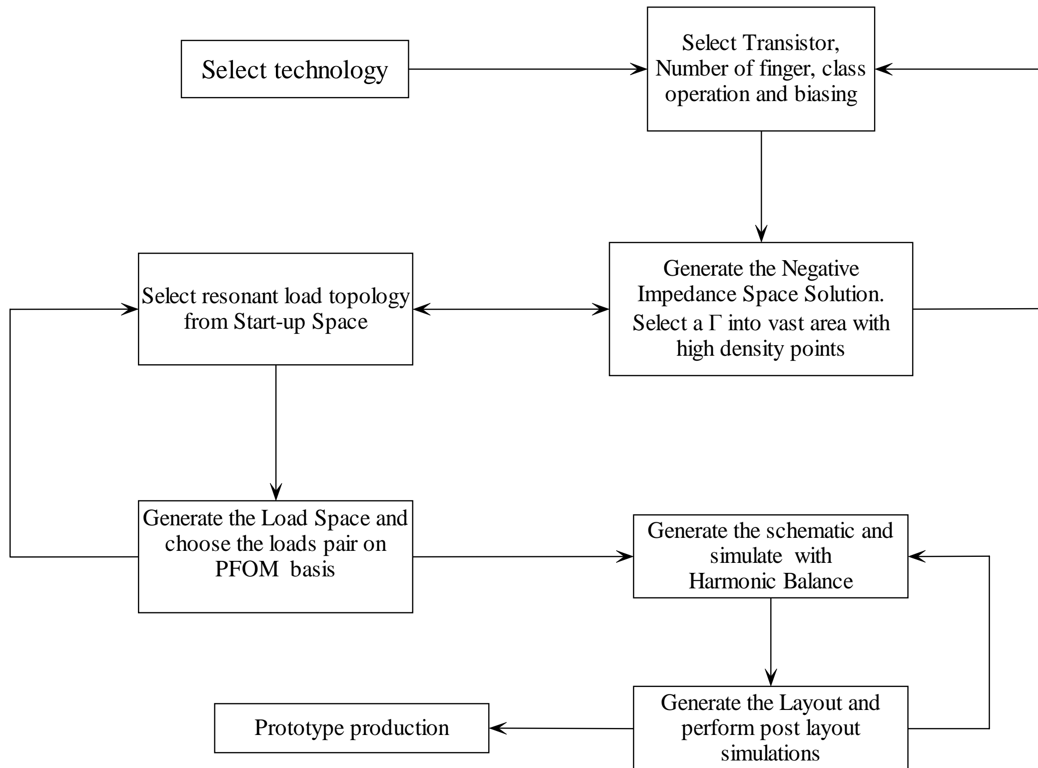
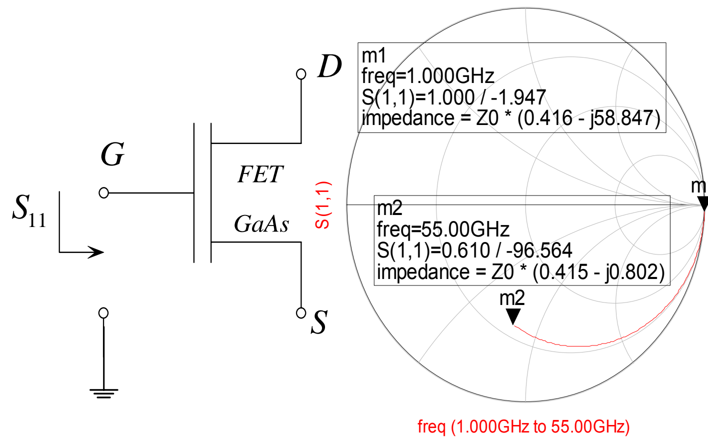


Figure 23

5.4-An applicative Example: the design of a 38GHz VCO in GaAs 0.2 μm technology

To illustrate how the Space Solution can be successfully employed in the design, a single transistor VCO with low phase noise, high conversion efficiency, output power of 10dBm almost constant in whole tuning range [38.0-39.5 GHz] is presented. The oscillator was performed in GaAs technology 0.2μm and the output power was obtained just through the choice of appropriate matching network, without employing any buffer stage. The starting point consists in the choice of an appropriate bias point for the transistor and in the estimation of output power with the CAD nonlinear tools. Next the small signal S_{3X3} parameters have been extracted in the frequency range of interest and the behaviour of the biased transistor has been studied .



Through the relationship the map of *Negative Impedance Space* (at transistor gate port), is drown in Fig.24 for $f_X = 38$ GHz .

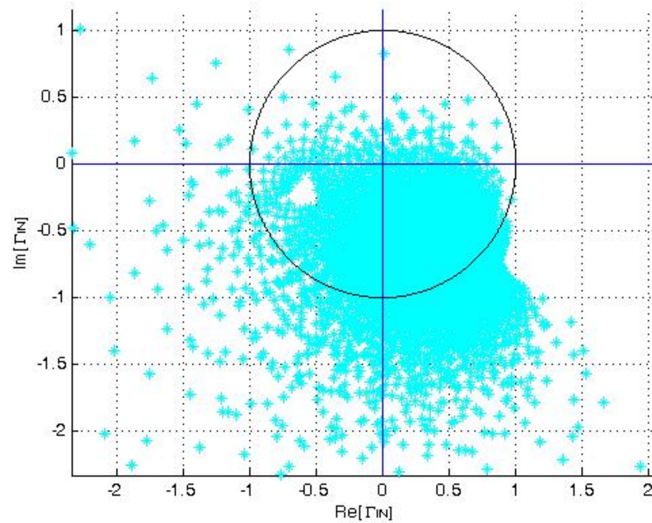


Figure 24

The Γ_{IN0} value is chosen to maximize the magnitude as possible inside the points high density area and to exclude the presence of resonance conditions near the f_X, namely a purely real value. Then the Solution Space equations allow to find all possible doable pairs that ensure the reliability of the selected Γ_{IN0} and the designer can choose those that for his design are significant.

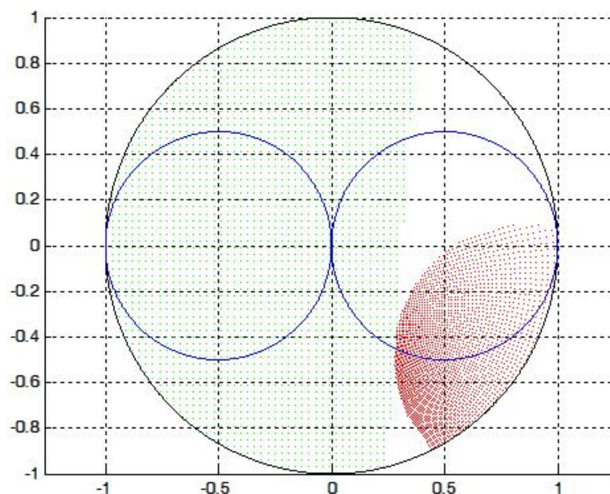


Figure 25

The output port can be chosen at this point by previous space once the solutions which kept limited *PFOM* around its maximum value, have been isolated. In this case the chosen output port is the transistor source because $K_N > K_M$ and the correspondent pair set is showed in Fig. 15.

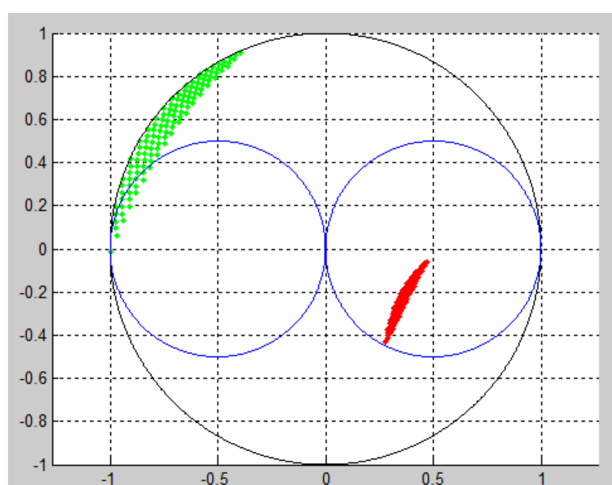
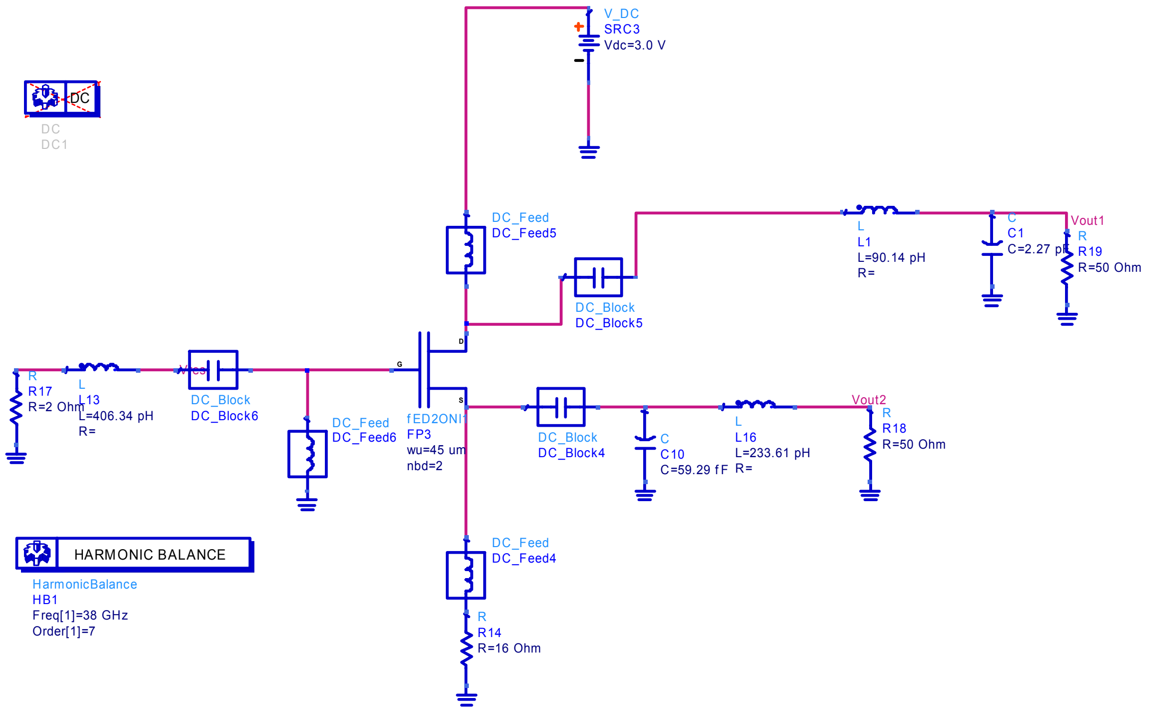


Figure 26

Next a simple lumped-elements circuit can be designed in order to test the performances. The following figures show respectively the schematic , the reflection coefficient seen at the gate, parameterized in frequency , and the estimation of the signals generated at source and drain port, through HB simulation .



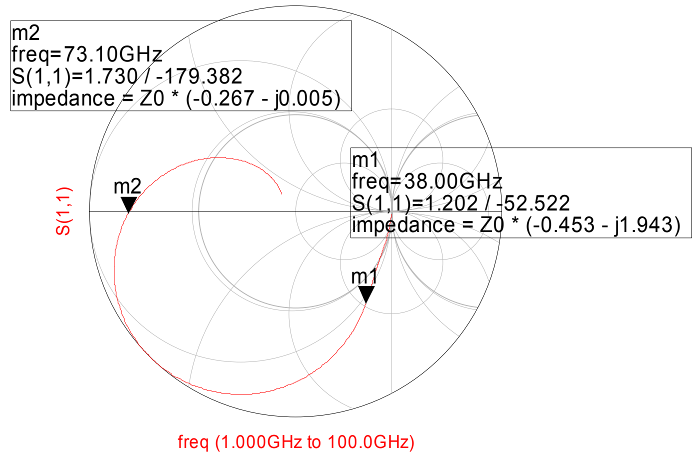


Figure 27

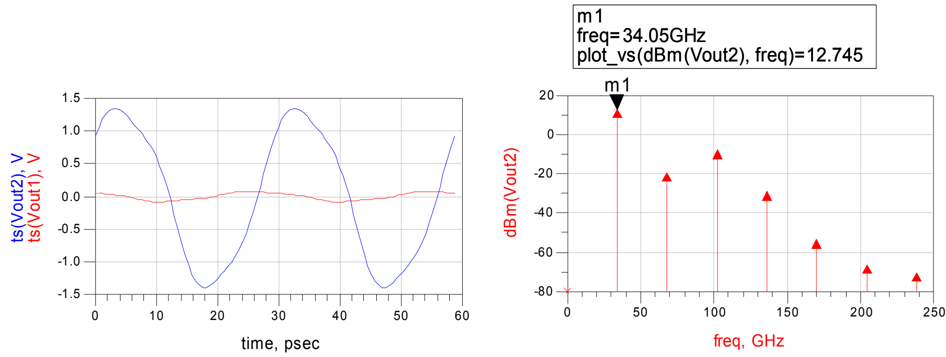


Figure 28

Finally the layout for the core and for the frequency determining circuit have been generated. In this case, the tuning network has been implemented in easy way through a varicap diode which capacitance variation in the voltage range 0÷3 V. The layout is shown in Fig 19 and the post-layout simulations are presented in Figure.

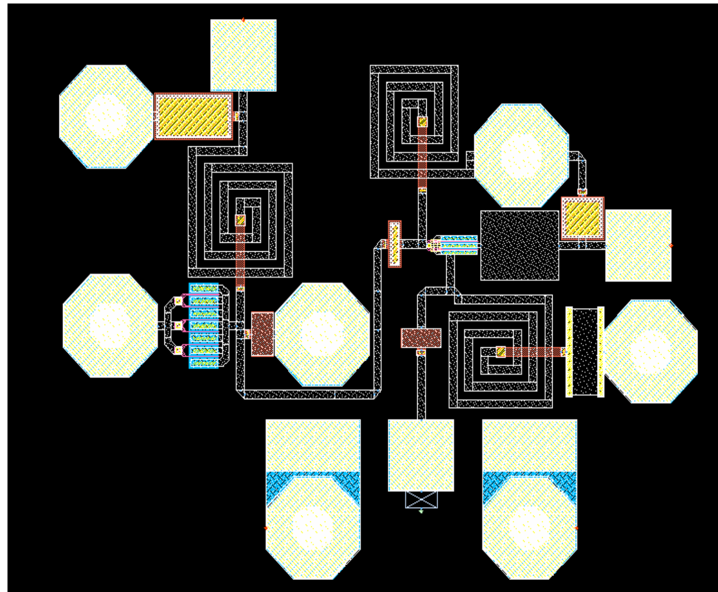


Figure 29

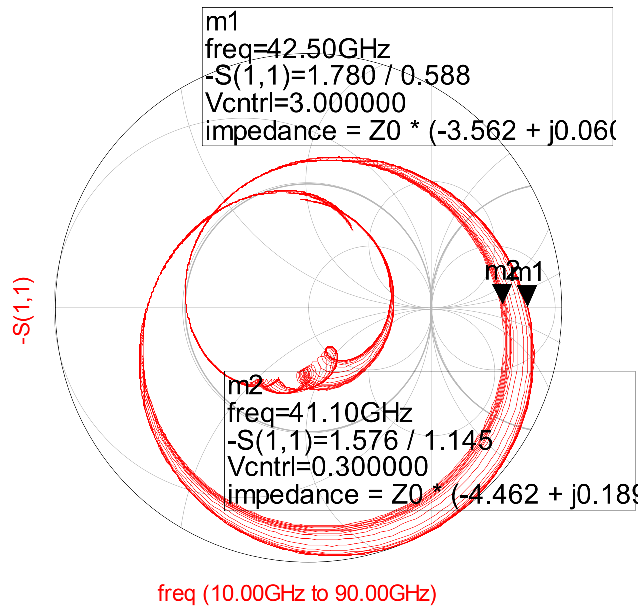


Figure 30

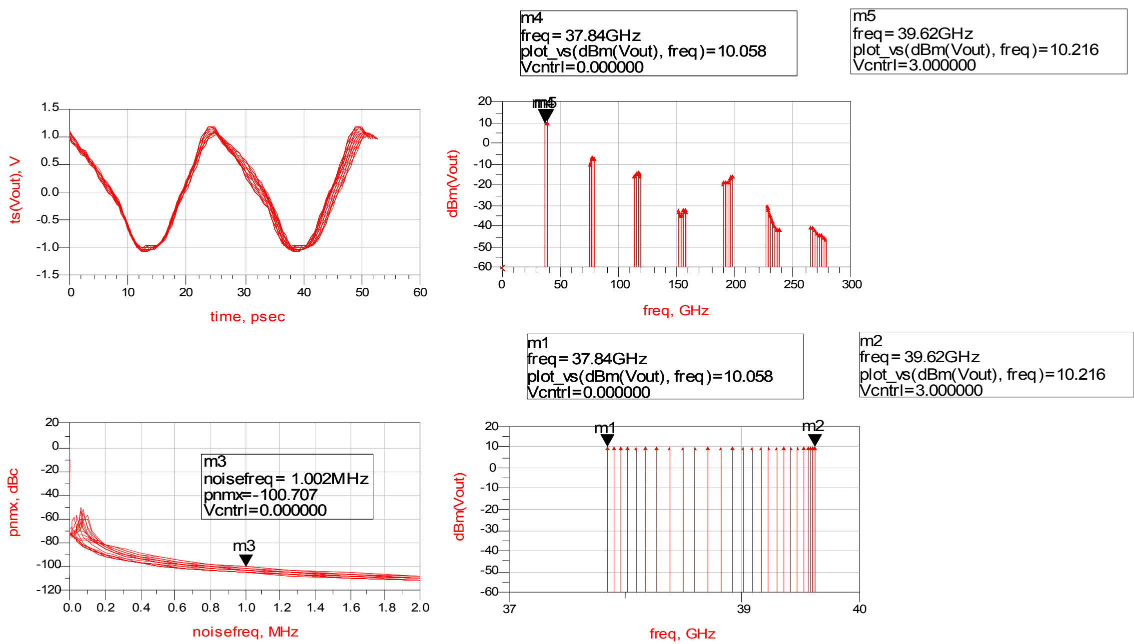


Figure 31

The FOM for the VCO circuit are computed through formulas presented in []:

$$FOM_1 = L(\Delta f_{offset}) - 20\log\left(\frac{f_0}{\Delta f_{offset}}\right) + 10\log\left(\frac{P_{DC}}{1mW}\right) \quad (11)$$

$$FOM_2 = FOM_1 + 20\log\left(\frac{f_0}{\Delta f_{TUNE}}\right) \quad (12)$$

and the performances are summarized in the following table:

Frequency Tuning Range	[37.84-39.62]GHz@[0-3]V
Output Power 50Ω	10dBm [±0.2dB]
Area	600X700 μm^2
PhaseNoise@1MHz	[-100;-112]dBc/Hz@[0-3]V
FOM1	-179.28
FOM2	-152.69
Development Time	<5 h

CONCLUSION.

In this work a new microwave negative impedance oscillator design methodology has been developed. This methodology presents a lot of significant advantages compared to the one considered to be the state of the art for this class of systems.
In first the study of the reflection coefficients and the modelling for the oscillator represented only through reflection coefficients for both resonant and negative impedance part leads to describe and classify the behaviours of possible combinations between the whole set of systems that it is possible to generate connecting these part each others. It was also dissolved any doubt about the Start- Up conditions by identifying the functions of each block and determining the specific behaviour for different parts; the resonant network, which must introduce the complex conjugate poles pair and the active network that must not introduce additional complex conjugate poles pairs in the frequency range of interest, but only to provides energy to the resonant circuit in order to ensure the oscillator start-up at desired frequency and keep it in steady stationary oscillation condition.
In fact; if the active device introduces additional complex conjugate poles pairs the Kurokawa's hypothesis on the domain on the frequency determining circuit (resonator) would no longer be verified and this may lead to a bad prediction of the reliability for the oscillator. Moreover would be more difficult to predict what will be the start-up frequency, because it will be determined by an arbitrary combination of different poles pairs. Furthermore, the oscillators built without this care would be characterized by a random operating condition because the active device and then poles pair that it introduce will be strongly dependents on the operating conditions of the active device i.e temperature.
These studies have led to determine robust Start-Up conditions and to solve in a definitive way the question on the accuracy of Start-Up expressed in terms of reflection coefficients and have allowed to define a true sufficient condition. Moreover criteria for the correct choice of value of the characteristic impedance Z_0 for employed in oscillation tests used in moderns CAD systems like ADS and Microwave office have been defined and employed with success.
The study and development of the s Solutions Space instead allowed to overcome the limits imposed by the state of the art design methodology which bases its success only on the designer

experience and try and error procedures. Instead to focuses the attention on a try and errors, a new design method has been developed, which allows since the first analysis to determine the best possible solutions at the first attempt; in terms of signal performance and start-up. The study of mathematical equations properties of systems that may be described through the trade in frequency of reflection coefficients, allowed to identify different areas with a greater density of points presences that indicates a greater robustness compared to constructive tolerances. The whole set of possible solutions showed in the maps, and generated by Solution Space equations, has also allowed to decrypt the properties hidden by different feasible configurations and to identify some figure of merit ,like the (PFOM), suitable to paint the generated maps in order to help the designer to isolate the solutions that provide the desired performance in terms of output power and signal cleaning, useful to improve the phase noise oscillator performances. They also can choose the port for the output feedback, by optimizing the layout according to the specific constraints of the design. Finally the developed tool allows to manage in an effective and easy way the design developed algorithms, by allowing the designer to get to the prototype of the optimal solution in a few clicks rather than many hours of trial work. The circuit prototype designed in the first phase, provides the maximum performances in terms of output power but still contains ideal and lumped elements. From this circuit it is easy to derive the physical circuit obtained by substituting real elements. The layout and post layout simulations keep the quality performances of ideal circuit. This good result indicate the good features of the design and a great robustness compared to constructive tolerances. The algorithms have been developed, only using the small signal S parameters that are most significant for the Start-Up, and allow you to bypass all of the analysis using the uncertain Large Signal S parameters models, by demonstrating that in order to achieve good performance in steady state behaviour is sufficient optimize the behaviour of circuit determined at small signal. They are proving to be particularly flexible and mouldable to define new methodologies for design of microwave circuits different from oscillators as LNA, Mixer and in general any system containing any number of microwave transistor, which are considered as 3-port networks whose are encoded through the plot of the reflection coefficient in the polar map of Smith.

REFERENCES

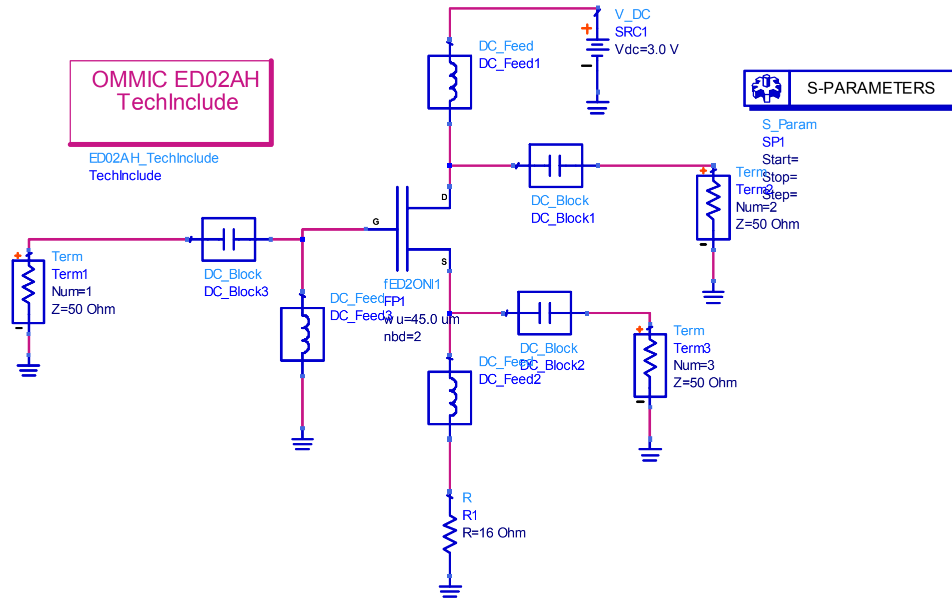
M.Monni G.Martines "A Novel Approach to Determine the Start-Up conditions in Microwave Negative Impedance Oscillator Design " EumC Munich 2007 Proceeding of the 37th European Microwave Conference.

M. Monni Giovanni Martines "A New Method to Design Microwave Negative Impedance Oscillators, Using the Solution Space Concept" Submitted to MTT

6.1-The oscillator Tool

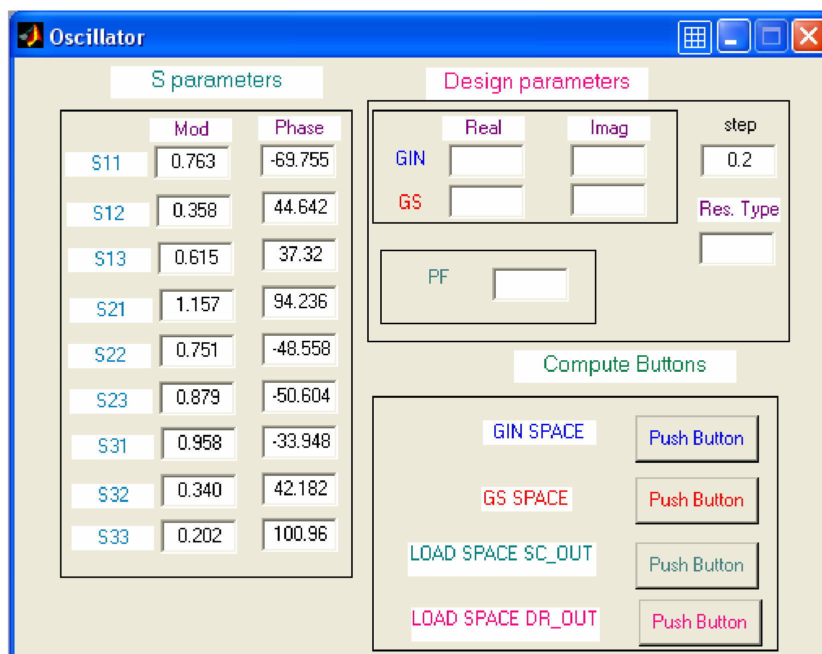
In order to make the work more boring I've developed a simple GUI in matlab ambience in which the designer can put in easy way the data and compute the different functions.
Lets we illustrate the use of this GUI.

At first we can chose a transistor with a bias network and to extract the S-parameters a desired frequency.

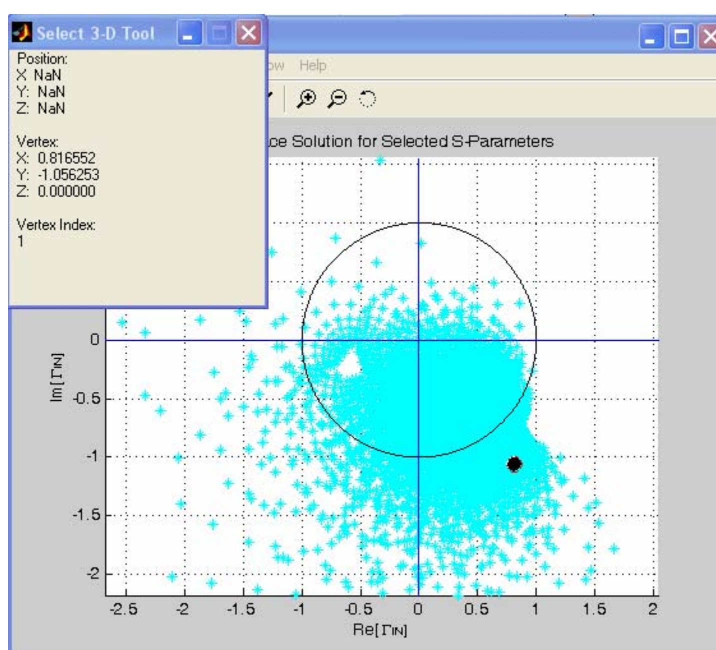


freq	S(1,1)	S(1,2)	S(1,3)	S(2,1)	S(2,2)	S(2,3)	S(3,1)	S(3,2)	S(3,3)
38.00 GHz	0.763 / -69.755	0.358 / 44.642	0.615 / 37.320	1.157 / 94.236	0.751 / -48.558	0.879 / -50.604	0.958 / -33.948	0.340 / 42.182	0.202 / 100.969

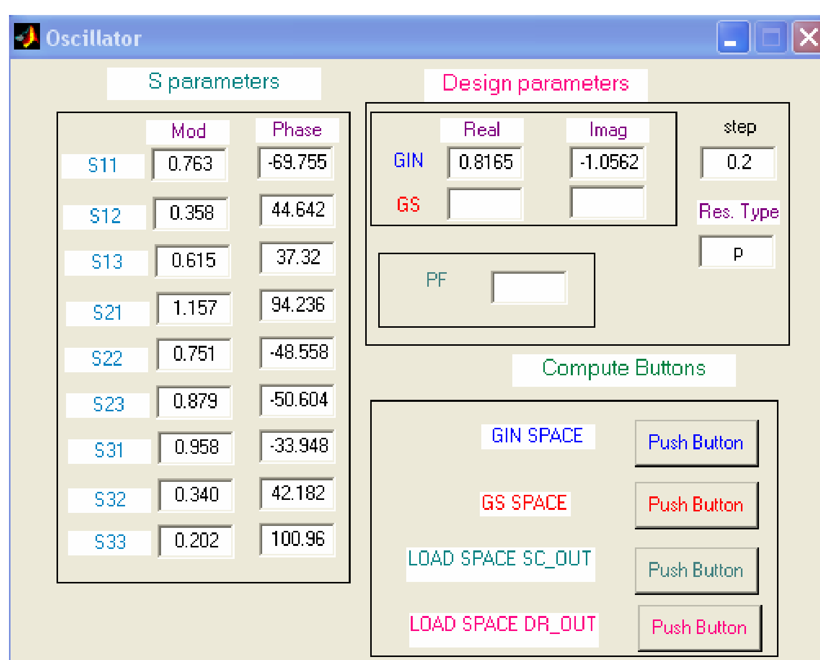
Then we can open matlab and digit oscillator to start-the GUI program.
IN the gui put the S-parameter values in module and phase and st-that indicate the step for the scan



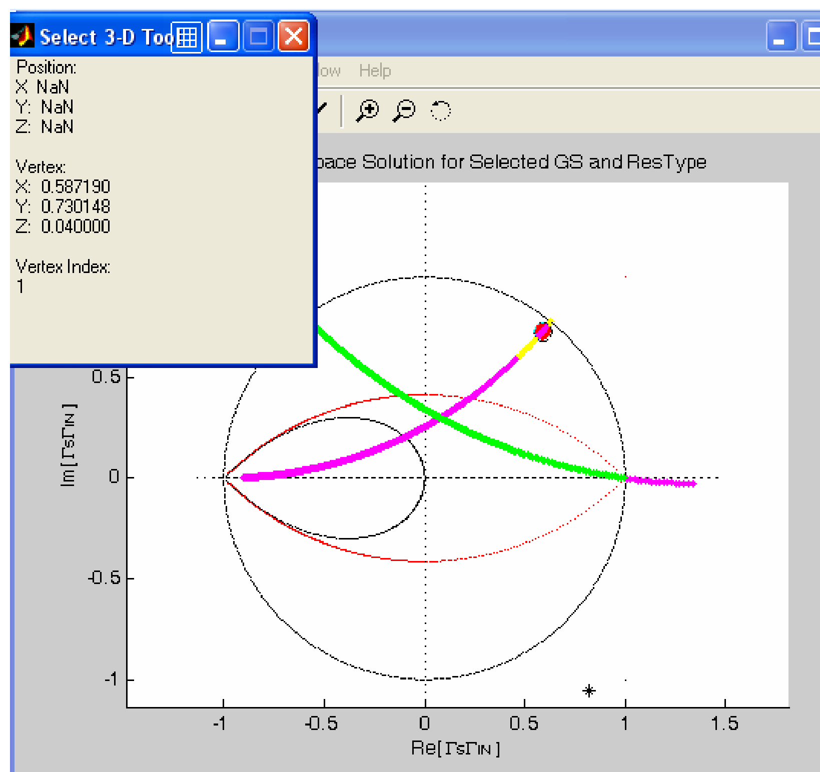
By pushing the GIN-SPACE push button the following window appear



Then you can chose the Γ_{IN} value by clicking with the mouse in a point. The corresponding value for Γ_{IN} in real and imaginary part will appear in the box called select 3-D Tool. If you click in a different point the value will be update automatically.
When a value is chosen you can put the value inside the box and select the kind of resonant series circuit in this case in agreement with the LTI theory the phase noise will be reduced if you chose a parallel circuit because it is easy to implement an high Q_U parallel resonant circuit that has a resonant frequency near to the start-up. If the RLC circuit is parallel and the Γ_{IN} has a capacitive behaviour then the start-up frequency will be greater than the resonant frequency since the unloaded transistor behaviour is the same that an RC series circuit.
Then you can push the GS SPACE BUTTON



And the following space appear



Then the possible Γ_S value to ensure the start-up are that comprises in the yellow arc.
In order to maximize the Quality factor for the system it is convenient to choose a point near as much as possible to the unitary circle.

Oscillator

S parameters

	Mod	Phase
S11	0.763	-69.755
S12	0.358	44.642
S13	0.615	37.32
S21	1.157	94.236
S22	0.751	-48.558
S23	0.879	-50.604
S31	0.958	-33.948
S32	0.340	42.182
S33	0.202	100.96

Design parameters

	Real	Imag
GIN	0.8165	-1.0562
GS	0.5871	0.7301

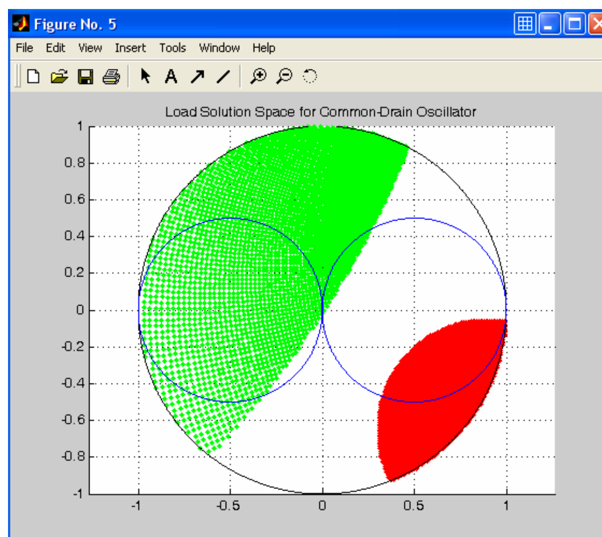
step 0.2
Res. Type p
PF 3

Compute Buttons

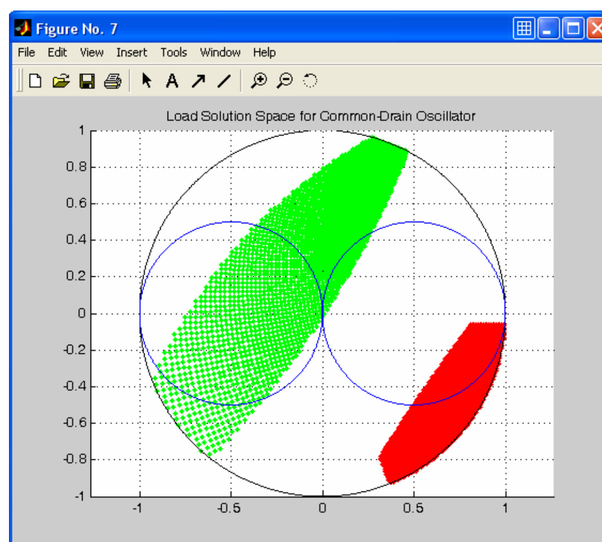
GIN SPACE Push Button
GS SPACE Push Button
LOAD SPACE SC_OUT Push Button
LOAD SPACE DR_OUT Push Button

Now we can put the Γ_S value in the box and chose a value for the PFOM in the range $-10<PFOM<10$. NB if we use a positive PFOM value then we are looking for solution in the SOURCE whereas a negative PFOM value indicates the solution that gives a greater output power in the DRAIN and hence for a common Source configuration.

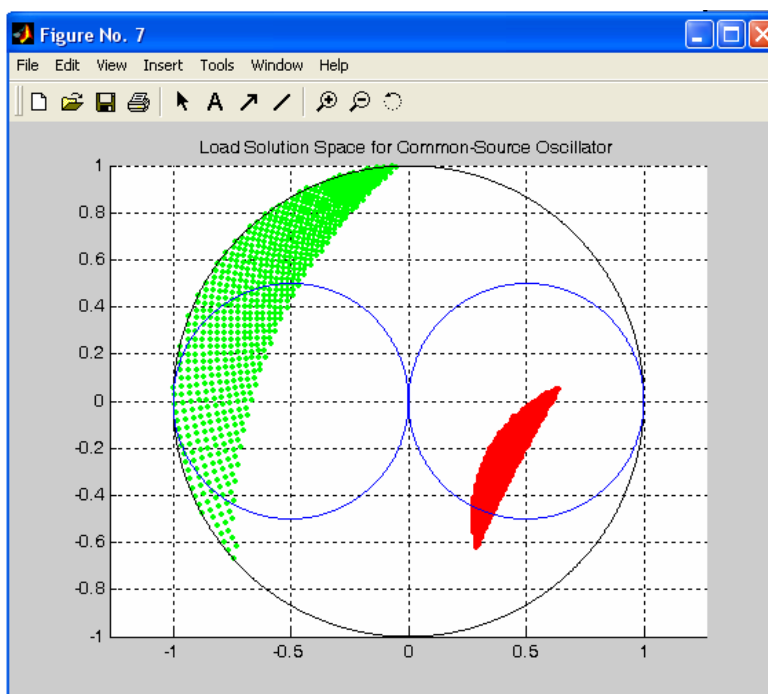
We don't know which value of PFOM is better at the beginning but by imposing a random value in the range guarantee a solution or for the common source of the common drain. If the value of PFOM is changed then change also the possibility to found a solution for the common drain or common source. In this case for minimum PFOM of 3 are founded solutions for the common drain and not for the common source.
By changing the st parameter it is possible to increase the scan step and then to find more solutions.



The high quantity of solution denotes the possibility to increase the PFOM In order to find only the solutions that give a great output power at the Drain terminal



Obviously the system generate a file .exl that can be studied in order to isolate the values with a specific PFOM for example these solutions can be used to increase the output power in the DRAIN and are traced with a PFOM of 0.8
The numerical file gives information about the solution and allows to find in a fast way the maximum PFOM achievable for the device.
If there not are solution it is possible to change in fast way the Γ_{IN} and the Γ_S and regenerate the load space in order to find the best solution as exposed in the previous paraghaph.
For example
The solutions that allows to maximize the output power at the source will be



For a PFOM minimum of 3.
As we note if we look solutions for the SOURCE output power the red point tend to be near to origin and the green point that represent the reactive components tend to be near to the unitary circle.
The dual case happens if we look solutions for the output in the DRAIN terminal.

Although the Space solution method is not innovative for the phase noise as to in agreement with hold school of LTI model, the Space solution and the oscillator tool overcome the limits of the hold method to design single transistor oscillator, and can be also used to design microwave amplifier with success.
Moreover y research activity has put in evidence how to do a desired oscillator with the output in a desired port.
The results of this method can be applied in every design problem that implies S-parameters.

Good luck.

Marco Monni