

Universidade Estadual de Campinas

Faculdade de Engenharia Elétrica e de Computação

Paulo José Martins Tavares

Circuito Integrado PLL para Banda GSM de 900MHz

Campinas 2003

Paulo José Martins Tavares

Circuito Integrado PLL para Banda GSM de 900MHz

Dissertação apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica, na área de Eletrônica, Microeletrônica e Optoeletrônica

Orientador: Prof. Dr. Furio Damiani

ESTE TRABALHO CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO PAULO JOSÉ MARTINS TAVARES, E ORIENTADA PELO PROF. DR. FÚRIO DAMIANI

Campinas 2003

Universidade Estadual de Campinas Biblioteca da Área de Engenharia e Arquitetura Luciana Pietrosanto Milla - CRB 8/8129

Tavares, Paulo José Martins, 1971-Circuito integrado PLL para banda GSM de 900 MHz / Paulo José Martins Tavares. – Campinas, SP : [s.n.], 2003.
Orientador: Furio Damiani. Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.
1. Radiofrequência. 2. Circuitos integrados. 3. Circuito de retenção de fase.
4. Sintetizadores de frequência. 5. Teoria da modulação. I. Damiani, Furio, 1943-2016. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Integrated PLL circuit for 900-MHz GSM band Palavras-chave em inglês: Radio frequency Integrated circuits Phase-locked loops Frequency synthesizers Modulation theory Área de concentração: Eletrônica, Microeletrônica e Optoeletrônica **Titulação:** Mestre em Engenharia Elétrica Banca examinadora: Furio Damiani [Orientador] Jacobus Willibrordus Swart Marcelo Arturo Jara Pérez Saulo Finco Data de defesa: 27-08-2003 Programa de Pós-Graduação: Engenharia Elétrica

Identificação e informações acadêmicas do(a) aluno(a) - ORCID do autor: https://orcid.org/0000-0003-3875-031X

⁻ Currículo Lattes do autor: http://lattes.cnpq.br/2437833868458768

COMISSÃO JULGADORA - DISSERTAÇÃO DE MESTRADO

Prof. Dr. Furio Damiani (Presidente) Prof. Dr. Jacobus Willibrordus Swart Dr. Marcelo Arturo Jara Pérez Dr. Saulo Finco

Candidato: Paulo José Martins Tavares RA: 994974 Data da Defesa: 27 de Agosto de 2003 Título da Tese: Circuito Integrado PLL para Banda GSM de 900MHz

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e de Computação.

Agradecimentos

Agradeço ao Centro de Pesquisas Renato Archer (CenPRA) por ter podido usar as suas instalações e equipamentos para o desenvolvimento deste projecto, e ao CNPQ e à FAPESP pelo apoio financeiro que possibilitou a concretização deste projecto.

Agradeço em especial ao Dr. Saulo Finco, chefe da Divisão de Concepção de Sistemas de Hardware, do CenPRA, pelo apoio e amizade que ele e a sua família me ofereceram desde que cheguei ao Brasil.

Agradeço ao Prof. Dr. Furio Damiani por ter aceitado orientar-me e por ter tido flexibilidade e tolerância para comigo, especialmente na fase de definição do tema de pesquisa.

Vim de Lisboa, onde nasci e cresci, para Campinas, sem conhecer praticamente ninguém aqui. Tem sido uma experiência de vida com os seus altos e baixos, mas que me completou. Devo isso em particular a algumas pessoas cuja amizade e carinho nunca vou esquecer. Pessoas especiais, cada uma do seu jeito: Ana Cláudia Faustino, Eloine Ferreira, Flávia Dzimba e Márcia Amaduci.

Durante o meu mestrado ocorreu a morte de pessoas que também ficarão na minha lembrança: a Cecília Maria Ridolfo foi uma amiga que só tinha para mim palavras de incentivo e elogio, mais até do que eu poderia merecer; a professora Maria Inês Simas, que orientou o meu projecto de iniciação científica após a licenciatura, sempre me apoiou e deu crédito ao meu trabalho. Também a sua filha Diana foi uma pessoa marcante, cuja vida foi lamentavelmente encurtada.

Agradeço aos meus amigos de Portugal, pela amizade que me proporcionaram e continuam a me proporcionar: João Pedro Pereira, Luis Rodrigues e Miguel Prado. Para além de grandes amigos, o Luís e o Miguel foram também, nos anos finais da licenciatura, os melhores colegas de trabalho que eu poderia ter tido.

A Érika Luciano tem sido nestes últimos meses uma companhia muito especial, de cujo convívio a escrita desta tese frequentemente me privou. Espero poder reparar isso e retribuir o carinho, paciência e encorajamento que recebi.

Agradeço ao meus pais por me terem proporcionado educação, mesmo com sacrifícios, e por me continuarem a ajudar. Tenho passado pouco tempo com eles, nos anos mais recentes, mas hei-de fazer por melhorar isso.

Agradeço à minha irmã por me ter ensinado as primeiras letras, e por ter sido sempre amiga. Agradeço-lhe também mais um sobrinho que há-de nascer, e há-de ser tão trabalhoso mas também tão maravilhoso quanto a Mónica, o Sérgio e a Joana.

Muito obrigado a todos vocês.

Resumo

Foi feita uma resenha da teoria de sistemas de rádio-freqüência, com ênfase em arquiteturas de transmissão/recepção, técnicas de modulação, demodulação e partilha de canal. No âmbito deste estudo, criou-se uma biblioteca de blocos funcionais que permite simular as principais técnicas de modulação e demodulação digital no software Matlab/Simulink.

Foi projetado um circuito integrado (CI) para síntese de freqüência baseada numa malha de controle de fase (PLL). Usou-se uma tecnologia BiCMOS de 0,8 μ m, com dispositivos HBT de SiGe. O PLL inclui um VCO ressonante LC, completamente integrado, capaz de oscilar a 915MHz ± 20%. Inclui também dois divisores de freqüência: um pré-divisor em lógica ECL, para alta-freqüência, e um divisor programável em lógica CMOS. O conjunto dos divisores de freqüência pode ser utilizado para formar um PLL de divisão inteira ou de divisão fracionária. O CI contém ainda um bloco comparador de freqüência e fase, baseado em máquina de estados, e um circuito bomba de carga, que aciona um filtro de malha externo. A escolha dos canais de freqüência baseou-se nas especificações do sistema GSM da banda de 900MHz. São apresentados os resultados de simulação e o *layout* de máscaras do CI.

Palavras-Chave: rádiofrequência, rádio-frequência, circuitos integrados, circuito de retenção de fase, sintetizadores de frequência, teoria da modulação.

Abstract

The theory of radio-frequency systems was reviewed, with emphasis on transceiver architectures, modulation, demodulation, and multiple access techniques. Within this study, a library of functional blocks was created, allowing to simulate the main digital modulation and demodulation techniques using the Matlab/Simulink software environment.

An integrated circuit (IC) for frequency synthesis based on a phase-locked loop (PLL) was designed, using a 0.8 μ m BiCMOS technology with SiGe HBT devices. The PLL includes a fully-integrated LC-tank resonant VCO, capable of oscillating at 915MHz ± 20%. Two frequency dividers are also included: a pre-scaler in ECL logic, operating at high-frequency, and a programmable divider using CMOS logic. These dividers can be configured for both integer-N and fractional-N PLL architectures. The IC also contains a phase-frequency detector, based on a simple finite state machine, and a charge pump circuit, which will drive an external loop filter. The choice of frequency channels was based on the specifications of the GSM system for the 900MHz band. Simulation results and IC physical mask layout are presented.

Key-Words: radio frequency, integrated circuits, phase-locked loops, frequency synthesizers, modulation theory.

Sumário

1	Inti	rodução
	1.1	Contexto
	1.2	Motivação e Objectivos
	1.3	Organização deste Trabalho
2	Vis	são Geral dos
		Sistemas de RF17
	2.1	Introdução
		2.1.1 Evolução Histórica
		2.1.2 Utilização do Espectro
		Legislação Brasileira
	2.2	Modulação
		2.2.1 Introdução
		2.2.2 Modulação Analógica
		AM — Modulação em Amplitude
		FM/ PM — Modulação em Frequência ou Fase
		Filtragem de Pré-ênfase e Desênfase
		2.2.3 Modulação Digital
		ASK. FSK e PSK
		Sinais Digitais de M Níveis ("M-ários")
		Funções de Base
		Detecção Óntima
		Detecção Coerente e Não-coerente
		Especificação de Largura de Banda
		Modulação Digital Binária
		Modulação em Quadratura
	22	
	2.3	2.2.1. Cadificação
		$2.5.2 \text{Acesso ao Canal} \qquad \dots \qquad $
		FDMA (Frequency-Division Multiple Access) $\ldots \ldots \ldots$
		IDMA (<i>Time-Division Multiple Access</i>)
		CDMA (Code-Division Multiple Access)
		2.3.4 Técnicas de Espalhamento Espectral
		Fundamentos
		Terminologia
		O Espalhamento Espectral e as Bandas ISM
	2.4	Arquitecturas
		2.4.1 Receptores Heteródinos
		Topologia de Dupla IF
		2.4.2 Receptores Homódinos
		Selecção de Canal
		Offsets DC
		Diferenças nas Componentes de I/Q
		Distorção de Segunda Ordem
		\mathbf{Ru} ído 1/f

	Radiação do LO	52
	2.4.3 Desmodulador de Quadratura	52
	2.4.4 Receptores de Rejeição de Imagem	52
	Arquitectura de Hartley	52
	Arquitectura de Weaver	54
	Receptores de IF Digital	55
	Receptores de Sub-Amostragem	55
	2.4.5 Transmissores	56
	2.4.6 Escolha da Frequência do LO	57
	2.4.7 Conclusões	57
2.5	Componentes	58
2.0	2.5.1 Problemas Gerais a Considerar	58
	Não-Linearidade e Seus Efeitos	58
	Interferência Inter-Símbolos	60
	A dantação de Impedância	61
	252 Amplificadores de Baixo Buído	62
	2.5.2 Amplificadores de Baixo Kuldo	62
		02 (2
	$Exemplo \dots \dots$	03
	Classificação de Amplificadores de Potencia	63
	2.5.4 Misturador	64
	2.5.5 Oscilador Local	64
	Ruido de Fase em Osciladores	65
	2.5.6 Amplificador de IF	65
	2.5.7 Filtro de IF	66
	2.5.8 Circuito de RSSI	66
	2.5.9 Limitador	66
	2.5.10 Data Slicer	67
2.6	Tecnologias	67
	2.6.1 Tecnologias Orientadas para RF	67
	2.6.2 Tecnologia CMOS	68
	2.6.3 Integração de Componentes Passivos	69
	Indutâncias	69
	Integração de Capacidade	70
	2.6.4 Modelamento e Simulação	70
2.7	Aplicações	71
	2.7.1 Áreas de Aplicação	71
	Telemetria	72
	Controlo de Acesso	73
	272 Exemplos de Padrões e Sistemas	74
	AMPS (Advanced Mobile Phone Service)	74
	NADC (North American Digital Collular)	74
	GSM (Global System for Mobile communications)	75
	CDMA de Quelcomm	75
	DECT (Digital European Cordless Tolenhone)	76
	HomoDE	76
	Divetanth	76
	Bluetootn	/0
2.0	GPS (Global Positioning System).	//
2.8		//
	2.8.1 Expectativas lecnologicas	77
	"Renascimento" do Projecto Analógico	78
	2.8.2 Expectativas de Mercado	79
_		-
Pro	ojecto do Circuito PLL	80
3.1	Descrição Geral	80
	3.1.1 Características Funcionais	80
	3.1.2 Princípio de Funcionamento	81
	3.1.3 Principais Aplicações	82
	3.1.4 Tecnologia de Fabrico Utilizada	83

	3.2	Oscilador Controlado por Tensão	83
		3.2.1 Introdução	83
		3.2.2 Osciladores Ressonantes	84
	3.3	Divisores de Frequência	88
		3.3.1 Canais de Frequência do Sistema GSM	89
		3.3.2 Arquitecturas de PLL e Divisão de Frequencia	89
		Arquitectura de Divisão Inteira	89
		3.3.3 Arquitectura de Divisao Fraccionaria	91
	2 1	S.S.4 Circuitos Divisores de Frequencia	92
	3.4 3.5	Bomba de Carga e Filtro de Realimentação	95
	5.5		91
4	Res	sultados de Simulação	98
-	4.1	VCO e Pré-divisor	98
	4.2	Divisores de Frequência Programáveis	100
	4.3	Comparador de Fase e Bomba de Carga	100
5	Lay	<i>yout</i> Físico do Circuito	104
	5.1	Introdução	104
	5.2	Oscilador Controlado por Tensão	105
	5.3	Divisores de Frequência	106
	5.4	Comparador de Fase e Bomba de Carga	109
	5.5	Layout Global	109
6	Cor	nclusões	115
	6.1	Sobre a Evolução da Area	115
	6.2	Sobre este Projecto	116
	6.3	Trabalho Futuro	116
D	oforô	anaing	117
R	eferê	encias	117
R(A	eferê Sim	èncias nulacões de Sistema em Matlab	117 120
R¢ A	eferê Sim A.1	encias nulações de Sistema em Matlab Introducão	117120120
R¢ A	eferê Sim A.1 A.2	encias nulações de Sistema em Matlab Introdução	 117 120 120 120
R¢ A	eferê Sim A.1 A.2 A.3	Incias Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca	 117 120 120 120 122
R(A	Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Introdução Implementação do Componentes da Biblioteca Introdução A.3.1 Geração, Conversão e Selecção de Sinais	 117 120 120 120 122 122
Ro A	eferê Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Introdução Implementação do Componentes da Biblioteca Introdução A.3.1 Geração, Conversão e Selecção de Sinais Introdução Onda Quadrada Introducta	<pre>117 120 120 120 122 122 122</pre>
Ro A	eferê Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Introduca Implementação do Componentes da Biblioteca Introduca A.3.1 Geração, Conversão e Selecção de Sinais Introduca Onda Quadrada Sequência de Bits	<pre>117 120 120 120 122 122 122 122 122</pre>
Ro A	Sim A.1 A.2 A.3	Pulações de Sistema em Matlab Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Bits Aleatórios.	<pre>117 120 120 120 122 122 122 122 122 122</pre>
R	eferê Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Introdução Implementação do Componentes da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Implementação Onda Quadrada Implementação de Bits Bits Aleatórios Implementação	<pre>117 120 120 120 122 122 122 122 122 122 122</pre>
Ro A	Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca Implementação, Conversão e Selecção de Sinais Implementação, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Bits Aleatórios. Bits Aleatórios. Conversor Série-Paralelo Conversor Paralelo-Série	<pre>117 120 120 120 122 122 122 122 122 122 122</pre>
R	Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca Implementação do Componentes da Biblioteca Implementação, Conversão e Selecção de Sinais A.3.1 Geração, Conversão e Selecção de Sinais Implementação do Componentes da Biblioteca Bits Aleatórios Implementação de Sinais Conversor Série-Paralelo Implementação Conversor Paralelo-Série Implementação Selectores 2 para 1 e 3 para 1 Implementação	<pre>117 120 120 120 122 122 122 122 122 122 123 123</pre>
R	Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca Implementação do Componentes da Biblioteca Implementação, Conversão e Selecção de Sinais A.3.1 Geração, Conversão e Selecção de Sinais Implementação Onda Quadrada Implementa de Bits Sequência de Bits Implementação Conversor Série-Paralelo Implementação Conversor Paralelo-Série Implementação A.3.2 Moduladores Implementação	<pre>117 120 120 120 120 122 122 122 122 122 123 123 124</pre>
R	Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca Implementação do Componentes da Biblioteca Implementação, Conversão e Selecção de Sinais A.3.1 Geração, Conversão e Selecção de Sinais Implementação do Componentes da Biblioteca Sequência de Bits Implementação de Sinais Bits Aleatórios Implementação de Bits Conversor Série-Paralelo Implementação de Sinais Conversor Paralelo-Série Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Implitude Modulada	<pre>117 120 120 120 120 122 122 122 122 122 123 123 124 124</pre>
R	Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Sequência de Bits Bits Aleatórios Conversor Série-Paralelo Conversor Série Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Amplitude Modulada Frequência Modulada Frequência Modulada	<pre>117 120 120 120 120 122 122 122 122 122 123 123 124 124 124</pre>
A	Sim A.1 A.2 A.3	mulações de Sistema em Matlab Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Bits Aleatórios. Conversor Série-Paralelo Conversor Paralelo-Série Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Amplitude Modulada BFSK	117 120 120 120 122 122 122 122 122 122 123 123 124
A	eferê Sim A.1 A.2 A.3	Introdução	<pre>117 120 120 120 120 122 122 122 122 122 123 123 124 124 124 124 124</pre>
A	eferê Sim A.1 A.2 A.3	Pulações de Sistema em Matlab Introdução Organização da Biblioteca Organização do Componentes da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Sequência de Bits Bits Aleatórios Conversor Série-Paralelo Conversor Paralelo-Série Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Amplitude Modulada Frequência Modulada BFSK BPSK OPSK	<pre>117 120 120 120 120 122 122 122 122 122 123 123 124 124 124 124 124 125 125</pre>
A	eferê Sim A.1 A.2 A.3	Pulações de Sistema em Matlab Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Bits Aleatórios. Conversor Série-Paralelo Conversor Paralelo-Série Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Amplitude Modulada Frequência Modulada BFSK QPSK OQPSK p/4-OPSK	117 120 120 120 122 122 122 122 122 122 123 123 124 124 124 124 124 125 125
A	eferê Sim A.1 A.2 A.3	Pulações de Sistema em Matlab Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Bits Aleatórios. Conversor Série-Paralelo Conversor Paralelo-Série Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Amplitude Modulada Frequência Modulada BFSK QPSK QPSK QPSK QPSK Matta Matta	117 120 120 122 122 122 122 122 122 123 123 124 124 124 124 124 125 125 125
A	eferê Sim A.1 A.2 A.3	Introdução	<pre>117 120 120 120 120 122 122 122 122 122 123 123 124 124 124 124 124 125 125 125 126 126</pre>
A	eferê Sim A.1 A.2 A.3	initasi nulações de Sistema em Matlab Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Bits Aleatórios Conversor Série-Paralelo Conversor Paralelo-Série Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Amplitude Modulada Frequência Modulada BFSK BPSK QPSK QPSK QPSK MSK. A.3.3 Desmoduladores Annplitude Modulada	<pre>117 120 120 120 120 122 122 122 122 122 122</pre>
A	Sim A.1 A.2 A.3	initasi nulações de Sistema em Matlab Introdução Organização da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Bits Aleatórios. Conversor Série-Paralelo Conversor Paralelo-Série Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Amplitude Modulada Frequência Modulada BFSK QPSK QPSK MSK. A.3.3 Desmoduladores Amplitude Modulada Desmoduladores Amplitude Modulada	<pre>117 120 120 120 120 122 122 122 122 122 122</pre>
Ra	Sim A.1 A.2 A.3	Introdução	<pre>117 120 120 120 120 122 122 122 122 122 122</pre>
A	eferê Sim A.1 A.2 A.3	Introdução Introdução Organização da Biblioteca Organização da Biblioteca Implementação do Componentes da Biblioteca Implementação A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada A.3.1 Geração, Conversão e Selecção de Sinais Sequência de Bits Bits Aleatórios Sequência de Bits Conversor Série-Paralelo Conversor Paralelo-Série Conversor Paralelo-Série Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Amplitude Modulada Frequência Modulada Frequência Modulada PSK OQPSK QPSK OQPSK p/4-QPSK MSK. A.3.3 Desmoduladores Amplitude Modulada A.3.3 Desmoduladores Amplitude Modulada BFSK BSK MSK_ BSK A.3.3 Desmoduladores Amplitude Modulada Amplitude Modulada Desmodulador Coerente Amplitude Modulada BFSK BFSK BSM BFSK BSM Asis Desmoduladores BSMOULAGOR Amplitude Modulada BSMOULAGOR Amplitude Modulada BSMOULAGOR	<pre>117 120 120 120 120 122 122 122 122 122 123 123 124 124 124 124 124 125 125 125 126 126 126 126 127 127</pre>
A	eferê Sim A.1 A.2 A.3	introdução Organização da Biblioteca Implementação do Componentes da Biblioteca A.3.1 Geração, Conversão e Selecção de Sinais Onda Quadrada Sequência de Bits Bits Aleatórios. Conversor Série-Paralelo Conversor Série-Paralelo Conversor Série-Paralelo Selectores 2 para 1 e 3 para 1 A.3.2 Moduladores Amplitude Modulada Frequência Modulada PSK QPSK p/4-QPSK MSK A.3.3 Desmoduladores Amplitude Modulada Desmodulador Coerente Amplitude Modulada BFSK BPSK Desmodulador Coerente Amplitude Modulada	<pre>117 120 120 120 120 122 122 122 122 122 122</pre>

	A.4	QPSK 1 OQPSK 1 p/4-QPSK 1 MSK. 1 A.3.4 Diversos 1 Canal com AWGN (Ruído Aditivo Gaussiano Branco) 1 Integral com <i>Reset</i> Periódico 1 Cálculo das Componentes em Fase e Quadratura 1 A.3.5 Diversos do Simulink 1 Demonstração da Biblioteca 1 A.4.1 Amplitude Modulada 1 A.4.2 Frequência Modulada 1 A.4.3 BFSK 1 A.4.4 BPSK 1	128 128 129 129 129 129 131 131 131 132 132 135 137 143
B	Tra	Insferência de Potência	.51
С	And	otações de Teoria dos Sinais 1	52
	C.1	Análise de Fourier	152
	C.2	Amostragem	153
	C.3	Conversão AD-DA, Pré- e Pós-Filtragem	154
	C.4	Sinais Complexos	155
	C.5	Decomposição Harmónica	157
	C.6	O Algoritmo da FFT	163

Capítulo 1 Introdução

1.1 Contexto

São evidentes a grande evolução e o grande sucesso comercial que as telecomunicações móveis tiveram nos últimos anos. A dimensão desse sucesso da comunicação "sem fios", ou mais especificamente, da comunicação por rádio-frequência¹ (RF), generalizou a expectativa e a apetência por novas aplicações de RF para o grande mercado, aplicações essas que se tornem tão acessíveis e disseminadas quanto são hoje os telefones móveis.

Redes de computadores sem fios, redes de automação industrial ou residencial via rádio, transmissão de dados entre computadores pessoais e equipamentos periféricos como câmaras de vídeo ou máquinas fotográficas digitais, são exemplos de produtos já existentes, embora com diferentes graus de penetração no mercado. Numerosos factores influenciaram o sucesso desses produtos: a solução técnica encontrada, o custo dos equipamentos, as necessidades do público, o *marketing*, etc.

Mas a evolução que globalmente se verificou, e as expectativas que se criaram, tiveram também como resultado, para os especialistas, um "renascimento" da pesquisa e desenvolvimento em RF, que nas duas ou três décadas anteriores se encontravam em "regime estacionário" [1].

Outra tendência que se verifica, e que já antes tinha tocado muitas outras áreas da electrónica e muitos outros domínios de aplicação, é a criação de produtos portáteis, com baixo consumo de energia, utilização de sinais digitais e de um microprocessamento sofisticado para controle de operação e tratamento de sinal. Estas características requerem a utilização cres-

¹O autor é português, e escreveu usando o seu "sotaque" natural (neste caso expresso em ortografia e gramática), agradecendo aos leitores brasileiros a sua amável tolerância face a consoantes mudas e a outras pequenas "esquisitices". Agradece também a tolerância de quaisquer leitores lusófonos, perante inconsistências de vocabulário que, sendo ainda predominantemente lusitano, já não é inteiramente de lugar nenhum. Espera-se contudo que poucas destas diferenças de vocabulário coloquem reais dificuldades — e onde for necessário, haverá uma breve nota de rodapé.

cente de circuitos integrados (CIs).

Os requisitos referidos, quando aplicados aos sistemas de RF, implicam, do ponto de vista técnico e de formação dos projectistas, uma síntese de duas "tradições" de projecto: a do projecto de RF usando componentes discretos, e a do projecto de CIs de baixa frequência [1]. Só recentemente surgiram livros introdutórios que procuram efectuar essa síntese (como por exemplo [1], [2] e [3]) e ferramentas tentando integrar os dois domínios de análise (veja-se em [4] uma apresentação sobre a variedade de ferramentas "tradicionais" utilizadas ao longo do fluxo de projecto de um CI transmissor-receptor completo).

1.2 Motivação e Objectivos

A motivação para elaborar uma tese em microelectrónica para RF resultou, por um lado, do desejo de ficar capacitado a participar do crescimento desta disciplina e, por outro lado, do fascínio exercido pela "magia" e mistério das comunicações sem fio.

Realizar um projecto nesta área, que se encontra na intersecção entre várias disciplinas tradicionais, exige o estudo de uma quantidade apreciável de tópicos, uns de carácter mais teórico, como:

- Arquitecturas de sistemas de RF (heterodinação ou conversão directa, esquemas de modulação analógica ou digital, técnicas de acesso a canal, etc.);
- Conceitos teóricos para análise de sistemas e componentes de RF (análise de ruído, efeitos de distorção e intermodulação, etc.);
- Conhecimento detalhado de diversos blocos funcionais (osciladores, moduladores ou desmoduladores, amplificadores, misturadores, antenas);

juntamente com outros de carácter mais prático:

- Ferramentas de software para análise e simulação de sistemas;
- Ferramentas de software para análise e simulação de circuitos analógicos e digitais;
- Características da tecnologia específica a ser usada no fabrico do circuito integrado.
- Ferramentas de software para *layout* e verificação física de circuitos integrados;
- Projecto de placas de circuito impresso para teste de protótipos de RF (considerando

blindagem de sinais de alta frequência, adaptação de impedâncias, criação de trilhas com impedâncias de valor pré-definido, etc);

• Equipamentos de teste e medida para sinais de alta frequência e análise espectral;

O objectivo principal deste trabalho de mestrado foi a *aprendizagem dos conceitos fundamentais* da área. Naturalmente que nem todos os tópicos referidos acima poderiam ser abarcados em detalhe, mas pretendia-se adquirir os conceitos fundamentais que permitissem realizar um primeiro projecto e efectuar estudos futuros mais aprofundados.

Um outro objectivo foi a *aprendizagem de ferramentas de análise, simulação e projecto* que são actualmente utilizadas na indústria microelectrónica, para ficar capacitado a implementar um projecto passível de fabricação numa tecnologia actual.

Outro objectivo ainda, foi a utilização da capacitação teórica e prática adquirida para projectar e enviar para prototipagem um bloco funcional significativo de um micro-circuito de RF.

O bloco funcional projectado foi um oscilador local, pelo facto de ser algo necessário para qualquer topologia de sistema, quer se trate de um transmissor ou de um receptor.

Existem numerosos tipos de oscilador local, de maior ou menor grau de sofisticação, de acordo com a aplicação pretendida. Optou-se por implementar um oscilador local baseado numa malha de controle de fase (PLL — *phase locked loop*) programável digitalmente. É um circuito razoavelmente sofisticado que, pelo facto de poder gerar uma frequência seleccionável de entre um conjunto de valores pré-estabelecidos, também é designado como "sintetizador de frequência". Ambos os termos serão utilizados ao longo deste texto². Sendo o PLL um circuito misto, i.e. com sub-blocos analógicos e digitais, permite também adquirir experiência prática nos dois fluxos de projecto, analógico e digital.

Na escolha inicial da frequência de operação, levou-se em conta a existência de bandas de frequência de uso livre para aplicações industriais, científicas e médicas (bandas ISM). Estas têm um conjunto de valores típicos mas, em função da legislação de cada país, apenas algumas estarão efectivamente atribuídas. No Brasil, essas frequências foram determinadas pela Agência Nacional de Telecomunicações, a Anatel [5]. Optou-se por trabalhar na

²Houve dois motivos para o autor preferir usar "PLL" no título: apesar de ser uma abreviatura, é um termo mais conhecido do que "sintetizador de frequência", e a escrita é a mesma na ortografía brasileira *e* portuguesa...

banda dos 902-928 MHz.

Finalmente, a escolha exacta do conjunto de valores de frequência a sintetizar e respectivo espaçamento entre canais foi motivada pelas especificações de canal da telefonia móvel de padrão GSM.

1.3 Organização deste Trabalho

O trabalho no projecto desta tese, com os objectivos anteriormente referidos, começou pelo estudo de arquitecturas de sistemas de RF e dos principais conceitos teóricos associados. O capítulo 2 deste texto constitui uma resenha do estudo efectuado.

Uma das "ramificações" desse estudo, em particular sobre tipos de modulação analógica e digital, foi o desenvolvimento de uma biblioteca de blocos de simulação no software *Simulink*. A descrição dessa biblioteca e de resultados com ela obtidos encontra-se no anexo A.

Em seguida avançou-se para o projecto de um sintetizador de frequência baseado em PLL. A descrição funcional do circuito é feita no capítulo 3. Começa por uma secção que apresenta em termos gerais as principais topologias para síntese de frequência, e a topologia escolhida neste projecto. Seguem-se secções associadas a cada um dos principais sub-blocos funcionais do PLL. No entanto, estas secções discutem não só os aspectos específicos de cada um desses blocos, mas também a forma como os mesmos afectam as características globais do sintetizador de frequência.

A familiarização com as ferramentas de projecto de CIs consumiu tempo significativo. Ao longo dessa aprendizagem, o autor escreveu um texto tutorial de 75 páginas [6], que poderá servir como introdução às ferramentas e ao *design-kit* utilizado. Esse material, embora seja uma "ramificação" adicional do estudo efectuado, não se presta a inclusão num texto académico, mas pode ser disponibilizado a eventuais interessados.

Os principais resultados de simulação obtidos são apresentados no capítulo 4, e os *layouts* de máscaras do circuito integrado, no capítulo 5. As conclusões gerais ocupam o capítulo 6.

Capítulo 2 Visão Geral dos Sistemas de RF

2.1 Introdução

2.1.1 Evolução Histórica

As comunicações sem fio percorreram um longo caminho, desde as primeiras experiências de Marconi, em 1895, com a transmissão a poucas dezenas de metros de sinais impulsivos, obtidos pela descarga de bobines. Logo foi sentida a necessidade de minimizar interferências, tanto de origem natural como artificial. Já em 1900, Marconi patenteou um circuito receptor com filtros passa-banda, para esse efeito.

As primeiras transmissões transatlânticas ocorreram em 1901. Eram, nos meios científicos da época, consideradas impossíveis, por saber-se da curvatura da terra e ignorar-se a existência da ionosfera [7]. Nesses primeiros tempos tratava-se efectivamente de *telegrafia* sem fios (TSF), e a grande aplicação que surgiu, para além das já existentes com a telegrafia convencional, foi a comunicação com navios em alto mar.

Durante a década seguinte tornar-se-ia evidente a superioridade dos transmissores baseados em sinais de onda contínua, e teriam lugar as primeiras transmissões de sinais de *áudio* via RF [2].

A transmissão de sinais de áudio possibilitou o surgimento de emissoras de rádio com programas de entretenimento (e inevitáveis intervalos publicitários) e a disseminação de receptores de rádio domésticos. Também equipamentos completos, para comunicação a grandes distâncias, atingiram alguma popularidade entre rádio-amadores entusiastas. Poucas décadas mais tarde, chegaria a televisão, que não "matou" o rádio, mas acabou ocupando o lugar de maior destaque nas comunicações de grande massa.

2.1 Introdução

A electrónica cresceu juntamente com as telecomunicações, e a invenção de válvulas e transístores foi motivada pela necessidade de amplificar os sinais enfraquecidos pela transmissão a grandes distâncias.

Os transístores revolucionaram a electrónica, pelo seu baixo custo e por permitirem níveis de miniaturização anteriormente impensáveis. O passo seguinte foi o circuito integrado, que com escalas de integração cada vez maiores, fez o custo de cada transístor tender para zero. Mas o mesmo não aconteceu com a complexidade do projecto.

Os telefones celulares que hoje em dia são tão comuns contêm milhares de transístores, e embora apenas uma pequena porção deles opere na faixa de RF, é essa secção do sistema que constitui tradicionalmente o "gargalo" do projecto [2] (sec. 1.2). Isso porque se trata de uma área multidisciplinar, o que a torna mais difícil.

Tradicionalmente, um projecto de RF de grandes dimensões divide-se em:

- Esquema de modulação e o processamento de banda base, projectados por especialistas em teoria da comunicação.
- 2. Arquitectura de transmissão-recepção, projectada por especialistas em RF.
- 3. Implementação dos blocos funcionais, feita por projectistas de circuitos integrados.

Este trabalho de mestrado irá incidir sobretudo no segundo e terceiro itens da lista (*arquitectura* e *implementação*), mas este capítulo procura dar uma visão geral da área.

2.1.2 Utilização do Espectro

Logo nos primórdios da história da TSF, tornou-se evidente a necessidade de regulamentar as frequências de transmissão. Essa regulamentação foi feita levando em conta interesses comerciais, estatais ou militares. Com os progressos técnicos, regiões adicionais do espectro foram-se tornando utilizáveis, em maior ou menor grau, e hoje em dia a regulamentação cobre toda a gama de radiação electromagnética, levando também em consideração os problemas técnicos particulares de cada faixa de frequência. Em [8], por exemplo, apresenta-se uma tabela com a utilização, nos EUA, da faixa de frequência entre os 2,15 e os 40 GHz, e são referidos alguns desses problemas, tais como:

• Nas frequências abaixo de 7GHz, o fading (interferência destrutiva devido a múltiplos

2.1 Introdução

percursos de propagação de um mesmo sinal) domina. Acima de 15GHz domina a atenuação devida à chuva. Entre 7 e 15GHz é necessário considerar ambos os factores.

- A chuva provoca maior atenuação se as ondas de transmissão tiverem polarização horizontal. Existem tabelas que classificam as diferentes regiões do mundo pela pluviosidade, e que são utilizadas pela indústria de telecomunicações.
- Existe também outro tipo de *fading*, chamado *multi-path* ou selectivo, que é devido a variações do índice de refracção da atmosfera. Este tipo de *fading* é combatido usando técnicas de diversidade espacial (utilização simultânea de várias antenas) ou diversidade em frequência (utilização simultânea de várias frequências).

Em [8] afirma-se ainda que, para frequências de giga-hertz, o mercado encontra-se fragmentado, pois a arquitectura do sistema depende muito da frequência utilizada e da potência de transmissão. Em geral, quanto mais elevada a frequência, mais complexo o sistema.

Recentemente surgiu um conjunto de técnicas de modulação e partilha de canal conhecidas como *espalhamento espectral* (cf. secção 2.3.4), que motivaram alterações à atribuição de frequências. Por exemplo, a faixa de frequência ISM (*Industrial, Scientific and Medical applications*) começou por ser atribuída aos fornos de micro-ondas (que são a fonte de interferência mais forte [9]) e aos aparelhos de diatermia por ondas curtas (para fisioterapia). Até alguns anos atrás ninguém pensava em utilizar essa faixa de frequência simplesmente porque a interferência era muito forte e não eram conhecidas técnicas de minorá-la. Com o advento das técnicas de espalhamento espectral isso tornou-se possível, e a utilização da faixa ISM foi aberta a novas aplicações. Esse facto despertou bastante interesse por parte dos fabricante de equipamento electrónico e de telecomunicações, para aplicações como WLANs (*Wireless Local Area Networks*).

Em [8] é fornecido ainda um conjunto interessante de referências relativas a regulamentação de espectro, entre as quais a do *site* www.tr.com (Telecommunications Research International) que reúne informação sobre leilões de espectro ao redor do mundo.

2.1.2.1 Legislação Brasileira

No Brasil, a agência reguladora é a Anatel (www.anatel.gov.br). Reproduzem-se abaixo dois trechos da legislação da Anatel relativos à utilização das faixas ISM [5]. Os parágrafos S5.13 e S5.280, referidos nos trechos, não se encontram em [5], mas podem também ser

consultados no site.

S5.138 - As faixas abaixo listadas estão destinadas a aplicações industriais, científicas e médicas (ISM). A utilização destas faixas de freqüências para aplicações ISM está sujeita a autorização especial concedida pela administração interessada, em acordo com as outras administrações cujos serviços de radiocomunicações possam ser afetados. Ao aplicarem esta disposição, as administrações levarão devidamente em conta as Recomendações mais recentes da UIT-R sobre o assunto:

6765–6795kHz	(freqüência central 6780kHz)
433,05–434,79MHz	(freqüência central 433,92MHz) na Região 1
	exceto nos países mencionados no Nº S5.280
61–61,5GHz	(freqüência central 61,25GHz)
122–123 GHz	(freqüência central 122,5GHz)
244–246GHz	(freqüência central 245 GHz)

S5.150 - As faixas abaixo listadas estão também destinadas às aplicações industriais, científicas e médicas (ISM). Os serviços de radiocomunicações operando nestas faixas devem aceitar a interferência prejudicial que possa resultar destas aplicações. Os equipamentos ISM operando nestas faixas estão sujeitos às disposições do Nº S5.13.

13553–13567kHz	(freqüência central 13560kHz)
26957–27283 kHz	(freqüência central 27120kHz)
40,66–40,70MHz	(freqüência central 40,68MHz)
902–928 MHz	na Região 2 (freqüência central 915MHz)
2400–2500 MHz	(freqüência central 2450MHz)
5725–5875 MHz	(freqüência central 5800MHz)
24–24,25GHz	(freqüência central 24,125GHz)

2.2 Modulação

2.2.1 Introdução

Porquê modular? As principais razões são:

- Para optimizar a partilha do espectro e cumprir a regulamentação vigente.
- Na transmissão por rádio, porque a antena deve ser uma fracção significativa do comprimento de onda, para ter um bom ganho.

- Na transmissão por cabo, porque se obtém um melhor isolamento em altas frequências.
- Porque em alguns casos isso permite contornar "não-idealidades" do canal.

Ao sinal original, com espectro concentrado em torno de $\omega = 0$, chama-se "sinal de banda-base", enquanto que o sinal modulado, cujo espectro fica concentrado em torno de uma frequência portadora ω_c , é designado por "sinal passa-banda"¹.

O processo inverso da modulação chama-se "desmodulação" ou "detecção".

Um sinal modulado pode sempre ser descrito pela expressão

$$x(t) = a(t) \cos \left[\omega c \ t + \theta(t) \right].$$

É comum chamar-se *excess phase* e *total phase* a $\theta(t)$ e a $\omega_c t + \theta(t)$, respectivamente. Como seria de esperar, também se usam os termos *excess frequency* e *total frequency* para $\theta'(t) \in \omega_c + \theta'(t)$.

Ao conjunto formado pelos circuitos modulador e desmodulador chama-se "modem", mas no contexto da análise de sistemas de RF, o termo também pode designar uma *técnica* de modulação e desmodulação.

Os principais critérios para avaliar o desempenho de um modem são:

- Tolerância ao ruído.
- Eficiência espectral.
- Eficiência energética, ou power efficiency.

A eficiência energética está relacionada com a escolha do tipo de amplificador de potência (PA, do inglês *power amplifier*) e um facto a ter em consideração é que PAs nãolineares podem atingir maiores eficiências.

Para análises comparativas entre modems, é comum considerar que o canal de transmissão é afectado por AWGN (*Additive Gaussian White Noise*, ou "ruído aditivo branco e gaussiano").

¹Em inglês pode chamar-se tanto *bandpass* como *passband*.

2.2.2 Modulação Analógica

2.2.2.1 AM — Modulação em Amplitude

Para AM (do inglês Amplitude Modulation), tem-se

$$x_{\rm AM}(t) = A_{\rm c} \left[1 + m \, x_{\rm bb}(t) \right] \cos \omega_{\rm c} t \,. \tag{1}$$

onde *m* é o índice de modulação.

A modulação é obtida com um circuito multiplicador. O sinal modulado irá ocupar o dobro da largura de banda do sinal de banda base.

A desmodulação pode ser feita de uma das seguintes formas:

- 1. Multiplicação por $\cos \omega_c t$, seguida de filtragem passa-baixo.
- 2. Detector de pico com carga resistiva (se a envolvente não passar por zero).

O segundo método é o mais simples e o mais popular.

A relação sinal-ruído (SNR, de Signal to Noise Ratio) é dada por [2]:

$$\text{SNR}_{\text{out}} = \frac{A_{\text{c}}^2 m^2 \overline{x_{\text{BB}}^2(t)}}{2N_0 B}$$

Num livro de 1990 [10], na pág. 1, afirma-se que a modulação em amplitude ainda é a mais utilizada e que tanto as formas *single-sideband* (SSB) como *double-sideband* (DSB) são comuns. Já em [2] (uma obra de 1998) é dito que a aplicação de AM é limitada a estações emissoras de rádio e ao som televisivo, e que as principais desvantagens do AM são a susceptibilidade ao ruído e a necessidade de um PA altamente linear.

A equação (1) corresponde ao tipo mais simples de modulação AM, que é o DSBAM (*Double SideBand Amplitude Modulation*), e pode ser reescrita como

$$x_{\rm AM} = A_c \cos \omega_c t + mA_c/2 \cos (\omega_c - \omega_{\rm BB})t + mA_c/2 \cos (\omega_c + \omega_{\rm BB})t.$$

Ou seja, de cada componente de frequência em banda-base resultam duas frequências em torno de ω_c , e por isso se fala de "banda lateral dupla". Em [10] o autor gasta algum tempo com cálculos relativos ao valor eficaz (ou RMS, do inglês *root mean square*) de uma onda modulada e com uma representação fasorial deste tipo de modulação. Uma das conclusões mais importantes desses cálculos de valor eficaz é que uma grande parte da potência de transmissão é utilizada para enviar a portadora, enquanto que apenas as bandas laterais são aproveitadas para transmitir informação. Em contrapartida, isso permite usar um circuito de desmodulação por detecção de envolvente, cuja implementação é muito simples.

Se o coeficiente de modulação *m* aumentar muito, a eficiência energética é maior mas o circuito de desmodulação terá de ser mais complexo. Também é possível eliminar completamente a transmissão da portadora. Esse tipo de modulação é designado por DSBSC(AM) (*Double SideBand Supressed Carrier*). Na desmodulação DSBSC é preciso reintroduzir a portadora com frequência *e* fase correctas, o que requer circuitaria complexa e faz com que o sistema DSBSC quase não seja usado [10], a não ser para transmitir o sinal de cor em televisão ou para transmitir informação estéreo em emissões de áudio VHF.

Outra possibilidade é suprimir não só a portadora mas também uma das banda laterais. Daí resulta a modulação conhecida como SSBSC(AM) (*Single SideBand Supressed Carrier*). Existem dois métodos para eliminar uma das banda laterais:

- Filtragem
- Circuito de desfasagem (ver fig. 2.1)



Figura 2.1 Circuito de modulação SSBSC por desfasagem.

No caso da filtragem, esta terá de ser feita numa frequência intermédia (cf. secção 2.4.1), para que possa ser selectiva o suficiente.

No caso do circuito por desfasagem, tem-se $x_m = \sin \omega_m t$ e $x_c = \sin \omega_c t$, mas por

simplicidade vamos escrever $x_m = \sin a$ e $x_c = \sin b$. Assim, à saída do multiplicador inferior tem-se

$$\sin a \sin b = \left[\cos(a-b) - \cos(a+b)\right]/2$$

enquanto que à saída do multiplicador superior, com as desfasagens, se tem

$$\sin(a + 90^{\circ}) \times \sin(b + 90^{\circ})$$

$$= \left[\cos(a + 90^{\circ} - b - 90^{\circ}) - \cos(a + 90^{\circ} + b + 90^{\circ})\right]/2$$

$$= \left[\cos(a - b) - \cos(a + b + 180^{\circ})\right]/2$$

A soma das saídas dos moduladores é simplesmente

$$\cos(a-b) = \cos(\omega_c - \omega_m)t.$$

Este método tem ainda a vantagem de permitir comutar facilmente entre a transmissão da banda lateral superior ou inferior.

De forma geral, a modulação SSBSC tem as seguintes vantagens sobre DSBAM:

- A largura de canal necessária é reduzida a metade.
- A relação sinal-ruído é melhor (ver detalhes em [10]).
- A eficiência energética é maior.

Na modulação DSB, se a portadora sofrer variações devidas a desvanecimento, isso provocará uma distorção considerável sobre o sinal. Se a portadora for reintroduzida no receptor, como acontece na SSBSC, esse problema é eliminado².

2.2.2.2 FM/ PM — Modulação em Frequência ou Fase

Segundo Razavi [2], O conceito de PM (*phase modulation*) e FM (*frequency mod-ulation*) é importante não só para modems mas também para a análise de circuitos como osciladores e sintetizadores de frequência.

Num sinal sinusoidal existem basicamente três variáveis cujos valores podem ser modulados: amplitude, frequência e fase. Os processos de modulação em frequência e fase são

²Se bem que exista o problema da fase da portadora reintroduzida (cf. [10]).

muito semelhantes e muitas vezes são analisados conjuntamente sob a designação de modulação de ângulo [10].

Estes dois tipos de modulação são descritos por

$$x_{\rm PM}(t) = A_{\rm c} \cos[\omega_{\rm c} t + m x_{\rm bb}(t)],$$

$$x_{\rm FM}(t) = A_{\rm c} \cos \left[\omega_{\rm c} t + m \int_{-\infty}^{t} x_{\rm BB}(\hat{t}) d\hat{t} \right].$$

Note-se que não é possível distinguir entre sinais PM e FM se os sinais de bandabase não forem conhecidos.

A modulação FM pode ser feita, por exemplo, usando um oscilador LC em que a capacidade é a de um díodo varistor cuja polarização inversa varia com $x_{bb}(t)$.

A desmodulação pode ser feita, por exemplo, através de um filtro passa-alto, que faz com que o sinal FM fique também modulado em amplitude. Em seguida aplica-se uma desmodulação em amplitude. Ao usar esta abordagem, há que ter em conta que variações espúrias na amplitude do sinal FM poderão introduzir ruído no sinal desmodulado. Por essa razão, há que interpor um estágio limitador entre o sinal FM e o filtro passa-alto.

O espectro do sinal FM é um pouco mais difícil de obter do que no caso AM. Se o excesso de fase for pequeno, ou seja,

$$\left| m \int_{-\infty}^{t} x_{\rm BB}(\hat{t}) d\hat{t} \right| \ll 1 \, \text{rad} \,, \tag{2}$$

diz-se que se trata de FM de banda estreita. Nesse caso [2],

$$x_{\rm FM}(t) \approx A_{\rm c} \cos \omega_{\rm c} t - A_{\rm c} m \sin \omega_{\rm c} t \int_{-\infty}^{t} x_{\rm BB}(\hat{t}) d\hat{t}$$

Pode dar a impressão que, ao limitar a amplitude de $\theta(t)$, à medida que $\omega_c t \rightarrow \infty$, a influência de $\theta(t)$ sobre o sinal modulado será cada vez menor, a ponto de se ficar apenas com uma modulação em fase, e não em frequência. Mas isso é uma impressão imprecisa³, pois uma função como o seno ou o cosseno é igualmente sensível a uma variação em torno de *x* ou em

³Na verdade é mais que imprecisa, é incorrecta, mas "impressão imprecisa" soa tão melhor...

torno de $x + 2000\pi$.

Se o sinal modulante for um sinal aleatório, o espectro de FM de banda estreita será ligeiramente mais estreito que o espectro de banda-base, devido à atenuação de 1/s resultante da integração.

Já se o sinal modulante for $x_{BB}(t) = A_m \cos \omega_m t$, então

$$x_{\rm FM}(t) \approx A_{\rm c} \cos \omega_{\rm c} t + A_{\rm c} A_{\rm m} \frac{m}{\omega_{\rm m}} \sin \omega_{\rm c} t \sin \omega_{\rm m} t$$
$$= A_{\rm c} \cos \omega_{\rm c} t + \frac{A_{\rm c} A_{\rm m} m}{2\omega_{\rm m}} \cos (\omega_{\rm c} - \omega_{\rm m}) t - \frac{A_{\rm c} A_{\rm m} m}{2\omega_{\rm m}} \cos (\omega_{\rm c} + \omega_{\rm m}) t$$

Assim, o espectro consistirá de impulsos em $\pm \omega_c$ e de bandas laterais (*sidebands*) em $\pm (\omega_c \pm \omega_m)$.

Tenha-se em atenção os seguintes factos:

- FM de banda estreita $\Rightarrow mA_{\rm m}/\omega_{\rm m} \ll 1 \, {\rm rad}$.
- Se ω_m aumenta, as bandas laterais ficam menores $\left(\frac{A_m A_c m}{2\omega_m} \rightarrow 0\right)$.
- A máxima variação instantânea de frequência (*frequency deviation*) é dada por |θ'(t)|_{max}
 = m A_m.

Note-se que a máxima variação instantânea de frequência $(m A_m)$ é diferente da distância entre as bandas laterais $(2\omega_m)$, e que para FM de banda estreita se tem $\frac{mA_m}{\omega_m} \ll 1$ e portanto $mA_m \ll \omega_m$, ou seja, a variação instantânea é muito menor que a largura de banda ocupada.

De acordo com [2], o conceito de FM de banda estreita, embora útil, tem pouca aplicação prática devido ao seu baixo valor de SNR.

Ainda de acordo com [2], se $x_{bb}(t) = A_m \cos \omega_m t$ mas não aplicarmos a restrição de FM de banda estreita (2), é possível expandir o sinal modulado como

$$x_{\rm FM}(t) = A_{\rm c} \sum_{n = -\infty}^{\infty} J_n \left(\frac{mA_{\rm m}}{\omega_{\rm m}}\right) \cos\left[(\omega_{\rm c} + n\omega_{\rm m})t\right]$$
(3)

2.2 Modulação

onde J_n é a função de Bessel de ordem n do primeiro tipo.

À medida que o valor de mA_m/ω_m aumenta, vão surgindo componentes significativas em $\omega_c \pm n\omega_m$.

Casos mais gerais requerem um tratamento mais pormenorizado, mas um resultado conhecido como Regra de Carson diz-nos que

$$B_{\rm FM} \approx 2 \left(\frac{mA_{\rm m}}{\omega_{\rm m}} + 1 \right) B_{\rm BB} ,$$
⁽⁴⁾

em que $B_{\rm BB}$ é a largura de banda do sinal de banda-base.

Relativamente à eficiência energética da modulação FM, é possível utilizar um PA não-linear (de alta eficiência), pois pelo facto de a informação num sinal FM estar contida nas passagens por zero, isso não corrompe o sinal de banda-base.

2.2.2.3 Filtragem de Pré-ênfase e Desênfase

Como foi visto na pág. 26, o aumento da frequência de modulação acarreta uma diminuição de amplitude nas bandas laterais. Para compensar esse efeito, o sinal de banda-base pode passar no emissor por um filtro de pré-ênfase, que irá amplificar as componentes de alta-frequência do sinal. No receptor será aplicado o filtro inverso, após a desmodulação.

De acordo com [2], se o sinal modulante tiver um espectro passa-baixo, a largura de banda do sinal FM não irá aumentar significativamente.

Sem aplicar esta técnica, tem-se

$$\frac{\text{SNR}_{\text{out}}}{\text{SNR}_{\text{in}}} = 6\beta^2(\beta+1)\frac{\overline{x_{\text{BB}}^2(t)}}{V_p^2},$$
(5)

enquanto que, com ela,

$$\frac{\text{SNR}_{\text{out}}}{\text{SNR}_{\text{in}}} = 2\beta^2(\beta+1)\left(\frac{B}{f_1}\right)^2 \frac{\overline{x_{\text{BB}}^2(t)}}{V_p^2} , \qquad (6)$$

onde $V_{\rm pk}$ é o valor de pico de $x_{\rm bb}(t)$, $\beta = \frac{mA_{\rm m}}{\omega_{\rm m}}$, *B* é a largura de banda do sinal de banda-base e f_1 a frequência de corte do filtro de desênfase. De acordo com [2], (6) permite

tipicamente um ganho de 10 a 15 dB relativamente a (5).

2.2.3 Modulação Digital

2.2.3.1 ASK, FSK e PSK

Os tipos de modulação citados na secção 2.2.2 são utilizados para transmissão de sinais analógicos. Para cada tipo referido de modulação analógica existe uma modulação digital "análoga"...

AM	ASK (amplitude shift keying)
PM	PSK (phase shift keying)
FM	FSK (frequency shift keying)

A modulação ASK, embora simples, é uma má solução pois é extremamente susceptível a interferência [11]. Por apresentarem uma menor sensibilidade ao ruído de amplitude, PSK e FSK são mais utilizados do que ASK [2]. A modulação FSK, em particular, tem ainda a vantagem de ser de implementação relativamente simples [11].

A tolerância ao ruído, em modulação digital, é frequentemente expressa como taxa de erro por bit, ou BER (*bit error rate*), definida como o número de bits recebidos incorrectamente, a dividir pelo número de bits recebidos.

2.2.3.1.1 Exemplo de Dimensionamento de Canais para FSK

Não há uma expressão rigorosa para a largura de banda ocupada por um sinal modulado em FSK, mas podem usar-se algumas expressões aproximadas.

O desvio máximo, Δf , é a separação entre a frequência nominal da portadora e a frequência correspondente a um "zero" ou "um". Em [11] sugere-se escolher esse valor em função da taxa de transmissão f_{TX} e do erro de frequência por imprecisões do oscilador, f_{ϵ} :

$$\Delta f = 2f_{\varepsilon} + f_{\mathrm{TX}}$$

É dado um exemplo: para um sistema que irá operar com uma taxa de transmissão de 57.6 kbit/s, em torno dos 900MHz, e que utilize um cristal de 10MHz com uma precisão de ±10 ppm, tem-se $f_{\epsilon} = 900$ MHz × 10 · 10⁻⁶ = 9kHz, e portanto

$$\Delta f = 2 \times 9 \text{ kHz} + 57.6 \text{ kHz} = 75.6 \text{ kHz}.$$

Em seguida pode aplicar-se a Regra de Carson para estimar a largura de banda ocupada.

Uma primeira estimativa equipara o sinal a modular a uma sinusóide (não uma onda quadrada), considerando tolerável uma perda de 3dB (o que irá aumentar a BER):

$$B_{\rm FSK} = 2 \Delta f + f_{\rm TX}$$

Uma segunda estimativa, mais realista, pretende recuperar até à terceira harmónica do sinal, com a contrapartida de ocupar mais espectro e assim ser também mais sensível a interferências.

$$B_{\rm FSK} = 2 \Delta f + 3 f_{\rm TX}$$

Aplicando a segunda estimativa ao exemplo anterior, o valor resultante é de $B_{FSK} = 2 \times 75.6$ kHz + 3 × 57.6 kHz = 324 kHz (cerca de $\frac{1}{80}$ da banda ISM a ser utilizada).

Em função desses valores, em [11] é apresentado um exemplo de atribuição de canais dentro da banda espectral a ser utilizada. Essa atribuição tem os seguintes parâmetros (cf. tabela 2.1 e fig. 2.2):

Largura de banda RF por canal:	330kHz
Número de canais:	5
Separação entre canais:	500kHz
Distância transmissão-recepção	21,9MHz

Tabela 2.1 Exemplo de atribuição de canais.



Figura 2.2 Exemplo de atribuição de canais.

2.2.3.2 Sinais Digitais de M Níveis ("M-ários")

Por vezes é vantajoso converter um fluxo de dados binários num sinal de M níveis, o que pode permitir um melhor aproveitamento da largura de banda disponível [2].

Por exemplo, se os bits forem agrupados dois a dois e os grupos assim formados forem colocados à entrada dum conversor DA de dois bits, o receptor precisará de uma resolução de amplitude maior, mas em contrapartida o sinal de quatro níveis terá menos de metade das transições do original. Assim, no exemplo anterior, a taxa de símbolos seria metade da taxa de bits original.

Para distinguir do caso binário, a cada nível transmitido chama-se "símbolo".

2.2.3.3 Funções de Base

Um sinal binário modulado em FSK pode ser representado por

$$x_{\text{FSK}}(t) = \begin{cases} A_{\text{c}} \cos(\omega_0 t) & \text{se } b_n = 0\\ A_{\text{c}} \cos(\omega_1 t) & \text{se } b_n = 1 \end{cases}$$

Uma outra notação possível é

$$x_{\text{FSK}}(t) = \alpha_0 \phi_0(t) + \alpha_1 \phi_1(t),$$

o que também pode ser escrito como um produto interno,

$$x_{\text{FSK}}(t) = [\alpha_0 \alpha_1] \cdot [\phi_0(t) \phi_1(t)],$$

onde $[\alpha_0 \alpha_1] = [0 A_c]$ ou $[A_c 0]$.

De forma mais geral, pode escrever-se a forma modulada de um símbolo como uma combinação linear de N funções de base ortogonais, ou seja, (usando várias notações possíveis)

$$x(t) = \alpha_1 \phi_1(t) + \alpha_2 \phi_2(t) + \dots + \alpha_N \phi_N(t)$$
(7)

$$=\sum_{i=1}^{N}\alpha_{i}\phi_{i}(t)$$
(8)

2.2 Modulação

$$= \overrightarrow{\alpha} \cdot \overrightarrow{\phi}(t) \,. \tag{9}$$

As funções de base serão ortogonais se, e só se,

$$\int_{0}^{T_{s}} \phi_{i}(t)\phi_{j}(t)dt = 0 \quad \text{para } i \neq j , \qquad (10)$$

onde T_S é a duração de um símbolo.

Este formalismo permite também representar uma técnica de modulação por uma "constelação" de símbolos, ao assinalar o valor de cada coeficiente α no respectivo eixo.

Por exemplo, um sinal binário em FSK pode ser representado por um de dois pontos num sistema de dois eixos, $\alpha_0 \in \alpha_1$. Por outro lado, um sinal binário em ASK pode ser representado por dois pontos sobre um único eixo.

A representação em constelação de símbolos facilita a visualização do efeito do ruído sobre os símbolos e da classificação no receptor.

2.2.3.4 Detecção Óptima

Admitamos que queremos detectar um impulso à saída do desmodulador. Como minimizar a influência do ruído? Pode demonstrar-se que se um pulso p(t) estiver corrompido por *ruído aditivo branco*, existe um filtro óptimo, na medida em que maximiza o SNR no instante de amostragem. A esse filtro chama-se filtro adaptado (*matched filter*), e deve ter uma resposta impulsiva dada por $h(t) = p^* (T_b - t)$, onde p^* denota o complexo conjugado de p. Mas em muitos casos a função p é real e portanto $h(t) = p (T_b - t)$.

Se à entrada do filtro também estiver presente ruído branco com densidade de potência N_0 / 2, então a melhor relação sinal-ruído possível será SNR_{max} = 2 E_p / N_0 , onde E_p é a energia do sinal,

$$E_p = \int_{-\infty}^{\infty} |p(t)|^2 dt \, .$$

Note-se que o SNR_{max} não depende nem da energia do pulso nem da sua forma ou largura de banda.

2.2.3.5 Detecção Coerente e Não-coerente

Recorrendo ao formalismo das funções de base e ao princípio de detecção óptima, podemos compor um receptor para uma modulação que utilize N funções de base, se no receptor fizermos o sinal modulado x(t) passar por N ramos, e em cada um deles for calculada a correlação entre o sinal e uma das funções de base. Depois disso, o sinal à saída de cada ramo será o coeficiente relativo a essa função de base.

Matematicamente, isso corresponde a dizer que, sendo o sinal modulado dado por (7), então em cada "ramo" *i* do receptor vamos calcular

$$\int_{0}^{T_{s}} x(t)\phi_{i}(t)dt = \int_{0}^{T_{s}} \left(\sum_{j=1}^{N} \alpha_{j}\phi_{j}(t)\right)\phi_{i}(t)dt$$
$$= \sum_{i=1}^{N} \alpha_{j}\int_{0}^{T_{s}}\phi_{j}(t)\phi_{i}(t)dt$$

Pela condição de ortogonalidade das funções de base, (10),

$$\int_0^{T_s} x(t)\phi_i(t)dt = \alpha_i \int_0^{T_s} \phi_i(t)\phi_i(t)dt$$

O valor da autocorrelação de $\phi_i(t)$ pode ser unitário ou não, mas em todo o caso será um valor conhecido. Com os valores de α_i assim obtidos no receptor, um bloco de codificação pode depois tomar a decisão sobre o símbolo correspondente.

Um problema prático desta abordagem é a necessidade de coerência entre $\phi_i(t)$ gerada no transmissor e $\phi_i(t)$ gerada no receptor, ou seja, não pode haver desfasagem entre elas. Por exemplo, se no receptor tivermos $\phi_i(t) = \cos \omega_i t$ e no receptor $\phi_i(t) = \cos (\omega_i t + \theta)$,

$$\int_{0}^{T_{s}} x(t)\phi_{i}(t)dt = \alpha_{i}\int_{0}^{T_{s}} \cos(\omega_{i}t + \theta)\cos(\omega_{i}t)dt$$
(11)

$$= \frac{\alpha_i}{2} \int_0^{T_s} \cos(2\omega_i t) \cos\theta \, dt \tag{12}$$

$$=\frac{\alpha_i T_S}{2} \cos\theta, \qquad (13)$$

o que significa que o valor de α_i obtido no receptor irá variar com θ , podendo ficar nulo.

Também existem detectores não-coerentes, assim chamados por não gerarem uma função de base coerente com a original. Em vez disso, se as funções de base forem sinusóides de diferentes frequências (como em FSK), cada ramo do receptor terá um filtro passa-banda sintonizado para uma das funções de base, seguido de um detector de envolvente. De acordo com [2], receptores coerentes têm melhor BER, mas são menos utilizados devido à sua maior complexidade.

2.2.3.6 Especificação de Largura de Banda

O critério mais comum para medir a largura de banda, em modulação digital, é a largura da faixa de espectro que contém 99% da potência do sinal.

Outra forma comum de indicar a largura de banda de um sinal é pela indicação da fracção de potência num canal adjacente, ACP (*adjacent signal power*). Por exemplo, se dissermos que num canal de 30kHz o sinal apresenta uma ACP de -50dB, isso significa que 10 $\log_{10} \frac{P_{AC}}{P_{C}} = -50$, ou seja, $P_{AC} = P_{C} \cdot 10^{-5}$, em que P_{C} é a potência do sinal contida no canal principal de 30kHz e P_{AC} a potência do sinal contida num canal adjacente da mesma largura.

2.2.3.7 Modulação Digital Binária

Se a operação de codificação após os correladores for linear, é possível simplificar a arquitectura do receptor.

Por exemplo, se na codificação o critério de classificação for $x(t) \phi_1(t) - x(t) \phi_2(t)$, pode ser utilizado um único filtro que calcula $\int_0^{T_s} x[\phi_1(t) - \phi_2(t)] dt$, ou seja, um filtro adaptado ao sinal $\phi_1(t) - \phi_2(t)$, apesar de este sinal nunca ser intencionalmente fornecido ao receptor. Isto permite não só simplificar a arquitectura, mas também simplificar a análise do desempenho. Neste exemplo, SNR_{max} = $2 E_d / N_0$, onde $E_d = \int_{-\infty}^{\infty} [\phi_1(t) - \phi_2(t)]^2 dt$.

Em [2] demonstra-se que a probabilidade de erro para modulação binária é

$$P_e = Q\left(\sqrt{\frac{E_d}{2N_0}}\right),$$

onde Q é a função de erro, E_d é a energia do sinal $\phi_1 - \phi_2$ e N_0 a densidade espec-

tral de potência do ruído à entrada do receptor. Este resultado pressupõe ruído AWGN e um filtro adaptado, mas é válido para todos os esquemas de modulação binária. No entanto, para fazer uma comparação, P_e será escrita em função da energia média por bit E_d , que de um modem para outro tem uma relação diferente com E_d .

Para BPSK tem-se $E_b = A_c 2 T_b / 2$, pelo que [2]

$$P_{e, \text{ BPSK}} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right).$$

Note-se que é possível diminuir P_e , aumentando a amplitude do sinal ou baixando o seu período (e consequentemente a taxa de transmissão).

Quanto ao espectro usado por um sinal BPSK, a densidade espectral de potência é

$$S_{x}(\omega) = \{ |P(\omega + \omega_{c})|^{2} + |P(\omega - \omega_{c})|^{2} \} / T_{b},$$
(14)

onde P é a densidade espectral da função de base utilizada.

Para BFSK, $E_d = A_c^2 T_b$, e portanto [2]

$$P_{e, \text{ BFSK}} = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

Note-se que para um obter um mesmo valor de *Pe*, a modulação BFSK requer o dobro da energia média por bit, em comparação com a BPSK. Isso pode ser entendido intuitivamente comparando as duas constelações de sinais. Para amplitudes de sinal idênticas, os sinais em BFSK ficam mais próximos. Por esta razão, diz-se que a modulação BPSK tem um "ganho" de 3 dB sobre a modulação BFSK. Mas apesar disso a modulação BFSK é mais popular, sobretudo em aplicações que não requerem uma taxa de transmissão tão elevada, pois permite uma implementação mais simples (com receptores não-coerentes) e uma maior eficiência energética.

O espectro ocupado por um sinal BFSK é difícil de calcular, mas pode ser estimado usando a regra de Carson, que segundo [2] também pode ser expressa como $B_T = 2 (\Delta f + 1 / T_b)$. Se a diferença entre frequências $f_1 e f_2$ for a mínima, $\Delta f = 1 / (2 T_b) e$ portanto $B_T \approx 3 / T_b$. T_b .

2.2.3.8 Modulação em Quadratura

Neste tipo de modulação, cada par de bits de banda base é convertido num único símbolo, sendo usada uma única portadora do tipo

$$x(t) = b_m A_c \cos \omega_c t - b_{m+1} A_c \sin \omega_c t, \qquad (15)$$

onde os valores dos coeficientes b_m e b_{m+1} indicam os bits a ser codificados no símbolo. Desta forma, a taxa de transmissão de símbolos é metade da taxa de transmissão de bits, o que contribui para a popularidade desta modulação.

Existem vários tipos e sub-tipos de modulação em quadratura:

- QPSK (quadrature phase shift keying)
 - OQPSK (offset QPSK)
 - $\pi/4$ -QPSK
- MSK (minimum phase shift keying)
 - GMSK (gaussian MSK)

Cada um desses tipos será em seguida descrito com mais pormenor.

2.2.3.8.1 QPSK (Quadrature Phase Shift Keying)

Neste caso, os coeficientes b_m e b_{m+1} valem ±1, conforme o valor do bit associado. Assim, a constelação de sinais será um quadrado, em que os sinais possíveis estarão nos vértices desse quadrado.

Uma forma equivalente de descrever o sinal é pela expressão

$$x(t) = \sqrt{2} A_c \cos(\omega_c t + k \pi/4), k = 1, 3, 5, 7$$

A desmodulação coerente de um sinal em QPSK é conceptualmente simples, com o sinal passando por dois trajectos onde é correlacionado com cos $\omega_c t$ e sin $\omega_c t$, respectivamente. O par de bits assim obtido é então aplicado a um bloco conversor paralelo-série.

Se compararmos a QPSK com a BPSK, conclui-se que com uma mesma energia de sinal transmitido, a distância entre símbolos é menor, para além de haver mais formas de errar, uma vez que passámos de dois símbolos para quatro símbolos. No entanto, como a duração de cada símbolo foi duplicada, a energia do símbolo recebido é maior. Pode demonstrar-se que o BER resultante é aproximadamente igual ao da modulação BPSK.

Já a eficiência espectral é melhor, pois o espectro de um sinal em QPSK é semelhante à soma de dois espectros BPSK, portanto semelhante a (14), mas com metade da amplitude, pois o valor de T_b é duplicado.

Uma desvantagem da QPSK são as possíveis mudanças de fase de 180°, indesejáveis se a forma de onda tiver de ser filtrada ou processada por um amplificador não-linear [2].

2.2.3.8.2 OQPSK (Offset QPSK)

Esta modulação é uma variante da QPSK que elimina as mudanças de fase de 180°. Isso é feito introduzindo em (15) um desfasamento entre as formas de onda de b_m e b_{m+1} . Desse modo as mudanças de fase serão apenas de ±90°, mas perde-se a vantagem das transições ocorrerem a intervalos de 2 T_b . No entanto, em [2] diz-se que o BER e o espectro obtidos são idênticos aos da QPSK.

Ainda segundo [2], uma desvantagem importante da OQPSK é a impossibilidade de usar codificação diferencial, uma técnica que permite o recurso a detectores não-coerentes, os mais populares devido à sua simplicidade.

2.2.3.8.3 π/4-QPSK

Esta modulação é conceptualmente equivalente a utilizar dois moduladores QPSK desfasados de 45°, sendo o sinal modulado enviado ora de um, ora de outro modulador. Desta forma reduz-se a variação máxima de fase a $3\pi/4$.

Segundo [2], o BER e a eficiência espectral são semelhantes aos da QPSK.

2.2.3.8.4 QPSK e Pulsos Não-Rectangulares

Embora nas secções anteriores tenham sido utilizados pulsos rectangulares, é preferível utilizar uma forma de pulso com espectro limitado. O pulso de cosseno levantado (16) tem a vantagem de minorar a interferência com canais adjacentes e de ao mesmo tempo minorar a interferência inter-simbólica.
Em [2] refere-se a possibilidade de decompor o filtro de cosseno levantado em duas "metades", uma no emissor e outra no receptor. Dessa forma a secção utilizada no receptor serve também como filtro adaptado.

2.2.3.8.5 MSK (Minimum Phase Shift Keying)

A modulação MSK pode ser encarada como uma variante da modulação OQPSK, em que não são usados pulsos rectangulares mas sim arcadas de sinusóide, positivas ou negativas conforme o bit for "um" ou "zero".

2.3 Protocolos

Neste contexto, os protocolos englobam os protocolos utilizados na codificação da sequência de dados e os utilizados no acesso ao canal.

2.3.1 Codificação

A codificação diz respeito à forma como os valores lógicos são traduzidos em valores ou transições de sinal. As principais preocupações são com a frequência máxima ou mínima que o sinal resultante terá, e também com o seu valor médio.

Em [11] são mencionadas como exemplos a codificação de Manchester, a codificação AMI, e o leitor é referido para [12] e para a demais literatura. Note-se que a codificação pode ser implementada em hardware ou software, dependendo do seu tipo e da taxa de transmissão pretendida.

2.3.2 Acesso ao Canal

O acesso ao canal, também denominado *duplexing*, diz respeito à forma como um dado equipamento pode transmitir e receber sinais simultaneamente [2].

Eis os possíveis tipos de acesso ao canal:

Simplex, que consiste em comunicação uni-direccional.

- *Half-duplex*, que permite a cada utilizador do canal ora ser emissor, ora ser receptor (como nos *walkie-talkies* tradicionais).
- *Full-duplex*, que permite emitir e receber simultaneamente (como num telefone, onde ambos os interlocutores podem falar e ouvir simultaneamente).

2.3 Protocolos

Se no caso *half-duplex* os dois circuitos utilizarem o mesmo canal de frequência alternadamente, então trata-se de TDD (*time-division duplexing*). Note-se que essa alternância pode ser rápida o suficiente para se tornar imperceptível em aplicações como transmissão de voz (como por exemplo em telefones celulares do sistema TDMA).

Uma arquitectura *half-duplex* é bastante interessante, pois o acesso TDD dispensa a implementação de um filtro de *duplex*, permite usar apenas um oscilador, e elimina que o circuito receptor sofre interferência do circuito transmissor [11], [9] (mas havendo vários emissores numa região, cada receptor estará sujeito à interferência dos outros emissores).

Se os circuitos de recepção e transmissão puderem funcionar em simultâneo, usando canais de frequências distintas, trata-se de FDD (*frequency-division duplexing*). Nesse caso o par de filtros do *front-end* do sistema será combinado num único filtro de duplex. Devido a limitações deste filtro, haverá alguma interferência entre o transmissor e o receptor do mesmo aparelho, ou haverá perdas significativas nos sinais filtrados.

2.3.3 Técnicas de Múltiplo Acesso

Muitos autores, entre os quais Razavi [2], usam uma analogia de pessoas a tentar falar numa festa barulhenta, como introdução ao tema. Mas essa analogia é um tanto grosseira, pois numa festa normal há demasiados mecanismos que permitem melhorar a comunicação: a proximidade entre interlocutores, a capacidade de distinguir entre timbres de voz e até a leitura de lábios...

Mas imagine-se um grupo de pessoas com vozes idênticas (talvez um aniversário comemorado entre irmãs gémeas sêxtuplas: Maria Antónia, Maria Beatriz, Maria Cecília, Maria Dolores, Maria Eulália e Maria Fernanda...) Por alguma razão elas estão de olhos vendados (não podem correlacionar o que ouvem com a direcção da voz ou com o movimento dos lábios), e Maria A. chama Maria B., Maria C. chama Maria D. e Maria E. chama Maria F., começando então cada uma das Marias A., C. e E. a transmitir à irmã que chamou uma sequência de informações do tipo:

"Sim, não, não, sim, não, sim, sim..."

É uma tarefa impossível! Têm de concordar num protocolo de comunicação, por exemplo:

- Maria A. pode pedir às irmãs C. e E. que esperem pela vez delas para falar, ou que se revezem a cada palavra (TDMA).
- Ou podem tentar usar tons de voz diferentes por exemplo contralto, soprano e mezzosoprano (FDMA), mas se não conseguirem fazer isso (se considerarmos um único canal disponível), essa técnica não ajuda.
- Ou ainda, podem usar termos diferentes: Maria A. e B. usam "positivo" e "negativo", Maria C. e D. usam "certo" e "errado" e Maria E. e F. usam "concordo" e "discordo" (CDMA). Dessa forma as irmãs receptoras identificarão muito mais facilmente a mensagem, apesar de as irmãs emissoras terem de falar palavras mais extensas e assim até aumentar o "burburinho"⁴.

Mas avancemos para uma descrição mais técnica do tema...

2.3.3.1 FDMA (Frequency-Division Multiple Access)

O espectro disponível é dividido em canais e a cada utilizador é atribuído um canal para emissão e outro para recepção.

2.3.3.2 TDMA (*Time-Division Multiple Access*)

Cada canal é partilhado por vários utilizadores, que utilizam o canal durante curtos intervalos de tempo, de forma rotativa. No caso de comunicação por voz, o sinal é digitalizado e compactado, para poder ser enviado dentro do intervalo de tempo disponível.

2.3.3.3 CDMA (Code-Division Multiple Access)

Embora em [2] seja usada a analogia da "festa", para introduzir a ideia do CDMA, parece-me mais simples considerá-la uma nova aplicação de funções de base ortogonais, só que numa fase de *codificação*, anterior à modulação propriamente dita. Essa codificação consiste em representar cada bit original por uma sequência de bits pseudo-aleatória, diferente para cada utilizador do canal. É como se cada utilizador do canal usasse um vocabulário diferente para comunicar com a central...

Para distinguir mais claramente entre um bit do sinal de banda-base e um bit do código, a estes último chama-se *chip*. Assim, se forem utilizados códigos de comprimento 64,

⁴Normalmente é feita a analogia das línguas diferentes, mas essa analogia não permite referir estes dois factores, o uso de palavras mais extensas e o maior ruído de fundo.

a taxa de transmissão de chips (*chip rate*) é 64 vezes mais elevada que a taxa de bits (*bit rate*). A sequência de chips é diferente para cada utilizador do canal, e vai ser o código (a "língua" própria) com que o utilizador se comunica com a central.

Os códigos de Walsh, ou códigos de Hadamard, são as linhas não-nulas de uma matriz *WN* gerada pela equação recursiva

$$W_1 = 0$$
$$W_{2N} = \begin{bmatrix} W_N & W_N \\ W_N & \overline{W_N} \end{bmatrix}$$

Por exemplo, o padrão CDMA IS-95 utiliza códigos de Walsh [13] de comprimento 64.

Estes códigos, para além da ortogonalidade, têm outra importante característica: pelo facto de serem uma sequência de bits pseudo-aleatória, têm um espectro relativamente largo. A multiplicação dos sinais no domínio do tempo faz com que o espectro resultante seja a convolução dos espectros originais, e neste caso a multiplicação pelos códigos de Walsh provoca o *espalhamento do espectro* do sinal de banda base.

Este espalhamento espectral pode parecer contra-producente, diminuindo a eficiência espectral, tal como a entendíamos, mas desta forma podem ser sobrepostos os espectros espalhados de vários utilizadores, pelo que a capacidade não é menor que a do FDMA ou TDMA.

Uma vantagem do CDMA é o facto de não impor um limite rígido ao número de utilizadores num canal. À medida que o número de utilizadores aumenta, a consequência é que cada receptor no canal sente um aumento proporcional no nível de ruído.

Por outro lado, esta técnica tem um requisito adicional, o de que todos os utilizadores transmitam com potências aproximadamente iguais, caso contrários os sinais mais fortes irão bloquear os mais fracos. Nos sistemas de telefonia celular esse controlo é feito pela estação rádio-base. Esse controlo de potência traz um benefício adicional para os terminais móveis, a redução do consumo de energia, pois não irão transmitir com uma potência desnecessariamente alta. O CDMA tal como foi descrito acima também é designado por DS-CDMA (*direct-sequence* CDMA) ou DS-SS (*direct-sequence spread spectrum*). Surgiu mais recentemente uma técnica chamada *frequency hopping* (FH-SS ou FH-CDMA), que consiste em emitir numa sequência pseudo-aleatória de frequências portadoras.

2.3.4 Técnicas de Espalhamento Espectral

A ideia do CDMA, introduzida como uma das técnicas de múltiplo acesso, é na verdade muito mais do que isso. Faz parte de um conjunto de técnicas sofisticadas, que apesar da sua maior complexidade de implementação, permitem fazer face a um grande número de problemas [14]:

- Endereçamento selectivo das transmissões
- Multiplexagem por divisão de código para múltiplo acesso a um canal
- Densidade espectral de potência reduzida, para ocultar sinais
- Protecção contra a intercepção de mensagens
- Medição de distâncias (ranging) com alta resolução
- Rejeição de interferências acidentais ou intencionais (anti-jamming)

As ideias base das técnicas de espalhamento espectral não são novas. Há mesmo quem considere que a primeira patente de espalhamento espectral foi registada pela actriz de Hollywood Hedy Lamarr, de origem austríaca, juntamente com o seu marido. Mas só avanços recentes na microelectrónica, como o surgimento de circuitos VLSI de alta velocidade e baixo custo, tornaram possível a aplicação prática desses princípios, e foram as aplicações militares as primeiras a efectivamente utilizar estas técnicas.

2.3.4.1 Fundamentos

De acordo com [14], a definição mais básica de sistema de espalhamento espectral seria um sistema de comunicação em que o sinal modulado ocuparia uma largura de banda muito maior que o sinal de banda base. De acordo com esta definição, um sinal de FM obtido com um índice de modulação elevado seria um caso de espalhamento espectral. O ganho de processamento seria, segundo [14],

$$3\beta^2 \left(\frac{S}{N}\right)_{info}$$

onde $\beta = \Delta f_{\text{carrier}} / f_{\text{modulation}}$ é o *deviation ratio*, e $\left(\frac{S}{N}\right)_{\text{info}}$ é a relação sinal-ruído na banda base.

No entanto, a definição mais aceite para espalhamento espectral acrescentaria à anterior a utilização de um outro sinal para realizar o espalhamento do sinal transmitido.

Existem três principais tipos de espalhamento espectral:

- Sequência directa, em que a portadora é modulada por um código binário cuja taxa de transmissão é muito mais elevada que a largura de banda do sinal de banda base.
- *Frequency hopping*, em que o valor da frequência da portadora é alterado de acordo com uma sequência de código pseudo-aleatória, pré-estabelecida.
- **Modulação por** *chirp*, em que são transmitidos pulsos e, para cada pulso, há um varrimento da frequência de oscilação.

Podem ainda considerar-se outras variantes, como o *time hopping*, em que os instantes de transmissão, feita com baixo *duty-cycle*, são variados também de acordo com uma sequência de código, ou até mesmo combinações de *time hopping* e *frequency hopping*, em que se varia simultaneamente o instante de transmissão e a frequência de portadora utilizada.

Uma relação fundamental da Teoria da Comunicação é a da capacidade de um canal, descoberta por Claude Shannon:

$$C = W \log_2 \left(1 + \frac{S}{N} \right),$$

onde C é a capacidade do canal em bit/segundo, W é a largura de banda em hertz, N é a potência do ruído e S a potência do sinal.

Se na expressão anterior fizermos a conversão para logaritmo natural e a expansão do logaritmo em série, obtém-se a aproximação

$$C \approx W1,44 \frac{S}{N} \approx W \frac{S}{N}$$

2.3 Protocolos

Esta relação diz-nos que a capacidade de um canal pode ser aumentada melhorando a relação sinal-ruído ou aumentando a largura de banda utilizada.

Por exemplo, se quisermos transmitir num canal a 3 kbit/s e nesse canal a potência do ruído for 100 vezes superior à do sinal transmitido, será necessário usar uma largura de banda de $2,08 \cdot 10^5$ Hz.

Quando à forma de "introduzir" o sinal de banda base no sinal a transmitir por espalhamento espectral, existem duas possibilidades:

A mais comum consiste em somar a informação (que estará sob forma digital) ao código de espalhamento espectral, antes de usar este código na modulação de espalhamento. Neste caso, o espalhamento pode ser tanto por sequência directa como por *frequency hopping*.

Outra possibilidade é efectuar a modulação pelo sinal de banda base antes do espalhamento. Os tipos de modulação mais comuns são os de modulação de ângulo (fase ou frequência), pois normalmente pretende-se um sinal espalhado com um envelope de potência constante.

2.3.4.2 Terminologia

"Ganho de processamento", de forma geral, indica o ganho em SNR obtido por um bloco funcional, por exemplo um amplificador de sinal. Se à entrada de um amplificador tivermos um SNR de 10dB e à saída um SNR de 16dB, o ganho de processamento será de 6dB.

No contexto de sistemas de espalhamento espectral, o termo pode também simplesmente significar a razão entre a largura de banda da informação e a largura de banda utilizada na transmissão. Este significado pode ser relacionado com o outro por intermédio do teorema anteriormente apresentado, demonstrado por Shannon.

A largura de banda (LB) de sinais transmitidos com espalhamento espectral de sequência directa (DS-SS) será a LB de 3dB do espectro do sinal transmitido, e irá portanto depender do tipo de modulação (bifásica, quadrifásica ou de fase mínima).

Para sistemas de *frequency hopping* (FH-SS), a LB será dada pelo número de canais usados multiplicado pela largura de cada canal.

O chip rate, ou "taxa de chips" será a frequência de relógio do código, no caso de

DS-SS, ou o hop-rate (taxa de comutação de canal), no caso de sistemas FH-SS.

2.3.4.3 O Espalhamento Espectral e as Bandas ISM

Existem várias bandas ISM, entre as quais [15]:

- 902–928 MHz
- 2400–2483.5MHz
- 5725–5850MHz

No entanto, a legislação impõe um limite à potência dos emissores [11]:

- 0dBm para transmissão contínua
- 30 dBm para transmissão com espalhamento espectral

Sistemas com potência radiada entre 0,75µW e um máximo de 1,0W devem implementar algum tipo de espalhamento espectral, ou por *frequency hopping* (FH-SS) ou por sequência directa (DS-SS), para reduzir a densidade espectral de potência (SPD, do inglês *spectral power density*) [15].

Os canais em que é feito o *frequency hopping* devem ter largura inferior a 500kHz, e devem ser utilizados pelo menos 50 canais a cada 20 segundos (o que corresponde a uma permanência de 400ms por canal). Recomenda-se⁵ que, para sequência directa, esta tenha pelo menos 127 bits.

Em ambos os tipos de espalhamento espectral o critério é o mesmo: não ultrapassar os 8dBm/3kHz. Por exemplo, a limitação do *frequency hopping* a canais de 500kHz com uma potência máxima de 1 W corresponde a ter

$$SPD_{dBm} = 10\log_{10} \frac{1 \text{ W} / 500 \text{ kHz}}{1 \text{ mW} / 3 \text{ kHz}} = 7,78$$

A melhoria de SNR obtida por aplicação de espalhamento espectral é designada por ganho de processamento, que deve ser pelo menos 10dB⁶. O ganho de processamento dependerá do código utilizado ou taxa de chips (caso do DS-SS) ou do número de canais e

⁵Mas o autor não deixa claro quem recomenda...

⁶O autor usa um "deve" como se fosse algo objecto de regulamentação, mas acho mais provável que seja apenas um critério de projectista para verificar o sucesso da implementação do espalhamento espectral.

hopping rate (caso do FH-SS).

2.4 Arquitecturas

Uma escolha importante, em termos de arquitectura de sistema, é entre utilizar um sistema homódino (também chamado de conversão directa) ou um sistema heteródino (que também pode ser chamado de *up-conversion*, no caso do transmissor, ou *down-conversion*, no caso do receptor). Outra questão correlacionada é se irá ser aplicado algum tipo de codificação aos dados a transmitir, o que poderá determinar a frequência mínima contida no sinal de banda base.

A conversão directa (fig. 2.3) consiste em modular directamente a frequência portadora com o sinal de banda-base. Segundo [11], é "extremamente fácil" de implementar e não gera produtos de modulação indesejados. Neste caso o oscilador local (LO) oscila na frequência portadora. Em contrapartida, é difícil gerar com precisão frequências acima de algumas centenas de megahertz, pelo que se torna necessário usar um oscilador estável (e.g. cristal) e



Figura 2.3 Princípio da conversão directa.

um PLL. O inconveniente do PLL é a possibilidade de *tracking-out*, situação em que o próprio PLL tenta cancelar os desvios de frequência decorrentes da modulação. Segundo [11], esse é um problema "particularmente frustrante".

No caso da *up-conversion*, a modulação é feita numa frequência intermédia (IF). O sinal modulado em IF é depois misturado (multiplicado) com o sinal gerado pelo oscilador local, LO. As frequências de IF e LO são escolhidas de forma a que a frequência de um dos





Figura 2.4 Princípio da up-conversion.

Note-se que se, por exemplo, a portadora pretendida estiver em LO + IF, na frequência LO – IF será criado um produto indesejável da modulação, que deverá ser filtrado. E quanto menor for IF, mais fácil será modular, mas mais difícil será filtrar esse produto indesejável (que se encontra a uma distância de 2 IF, no espectro).

Embora a *up-conversion* resulte num circuito mais complexo, tem uma vantagem: uma IF baixa o suficiente pode ser gerada com precisão mesmo sem usar um PLL. E não havendo PLL, não haverá *tracking-out* e não será necessário garantir uma frequência mínima para o sinal a transmitir, através de uma codificação do sinal.

Em [2] são referidas as seguintes arquitecturas:

- Receptores
 - heteródinos
 - homódinos
 - de rejeição de imagem
 - de IF digital
 - de sub-amostragem

- Transmissores
 - de conversão directa
 - de dois estágios

2.4.1 Receptores Heteródinos

A arquitectura genérica de um receptor heteródino é a apresentada na fig. 2.5.



Figura 2.5 Parte de um receptor heteródino.

Os vários blocos desempenham as seguintes funções:

- Selecção da banda de frequências de interesse (pode também ser parte de um sistema de FDD).
- Amplificador de baixo ruído (LNA, do inglês *low-noise amplifier*), necessário porque o misturador introduz tipicamente bastante ruído.
- Filtro de rejeição de frequência-imagem (necessário se a filtragem passa-banda anterior não for suficiente).
- 4. Misturador de *down-conversion*, para que a desmodulação e selecção de canal possam ser feitas mais facilmente, a uma frequência mais baixa.
- 5. Filtro passa-baixo para selecção do canal.

Os principais benefícios de uma arquitectura heteródina são

- é mais fácil filtrar ou amplificar a frequências baixas e fixas.
- é mais fácil sintonizar variando apenas o LO.
- é mais fácil evitar realimentações parasitas entre diferentes estágios de ganho.

2.4 Arquitecturas

Uma desvantagem da arquitectura heteródina é que o filtro de rejeição de imagem tem normalmente de ser implementado com componentes passivos discretos. Isto requer ainda que a impedância de saída do LNA seja de 50 Ω , o que dificulta o projecto desse amplificador.

Ao escolher o valor de IF há que ter em conta o seguinte compromisso:

- Um valor de IF elevado facilita a rejeição da frequência-imagem.
- Um valor de IF menor facilita a selecção de canal por filtragem e amplificação.

Ver a este respeito [2], fig. 5.10. Outro factor a considerar na escolha da IF é a disponibilidade de componentes comerciais. Algumas frequências são historicamente mais utilizadas, pelo que é mais fácil encontrar filtros SAW ou de cristal para elas: 10,7MHz, 71MHz, etc.

A interferência provocada por uma frequência-imagem pode surgir não apenas por filtragem insuficiente, mas também por distorção resultante de não-linearidades. Um exemplo desse problema é a interferência de "meia IF". Por exemplo, a frequência $(\omega_{RF} + \omega_{LO})/2$ pode também acabar interferindo com ω_{IF} . Se em algum ponto do receptor houver distorção de segunda ordem, teremos uma componente em $\omega_{RF} + \omega_{LO}$. Se o oscilador local tiver uma segunda harmónica significativa, teremos então uma *down-conversion* indesejável para |2LO - (RF + LO)| = |LO - RF| = IF. E existem mais possibilidades, claro (em [2], pág. 126,é dado ainda outro exemplo). Para suprimir essas possibilidades há que:

- Minimizar a distorção de segunda ordem
- Assegurar um *duty-cycle* de 50% no LO
- E se ainda for necessário, fazer com que o filtro de rejeição de imagem atenue o suficiente também para (ω_{RF} + ω_{LO}) / 2.

2.4.1.1 Topologia de Dupla IF

É possível estender o conceito de receptor heteródino a múltiplas *down-conversions*, cada uma com o seu filtro e amplificador. Desta forma o projecto de cada um desses filtros e amplificadores pode ficar facilitado. O funcionamento de tal sistema é bem ilustrado em [2], fig. 5.12. Ainda segundo [2], a maioria dos receptores de FM actuais usam dois estágios de *down-conversion*, daí o nome "dupla IF". O autor tece ainda algumas considerações a respeito do dimensionamento de ganho e linearidade para cada um dos elementos da cascata, de forma a optimizar o resultado global. E de acordo com ele, é necessário um número apreciável de iterações, alternando entre os níveis de arquitectura e de circuito, até chegar a uma implementação satisfatória.

2.4.2 Receptores Homódinos

Nestes receptores, também chamados de conversão-directa ou zero-IF, a rádio-frequência é convertida directamente para banda-base (LO = RF).

No caso da modulação AM, o sinal assim obtido já é o sinal de banda-base. No caso de modulação em frequência ou fase, é preciso ainda separar as componentes em fase e em quadratura.

Com esta arquitectura não se coloca o problema de frequência-imagem, pelo que não é preciso usar um filtro de rejeição de imagem, o LNA não precisa de ter uma impedância de saída de 50 Ω e os filtros subsequentes podem ser filtros passa-baixo e passa-banda mais facilmente integráveis. Claro que, em contrapartida, surgem outros problemas [2]:

- Selecção de canal
- Offsets DC
- Diferenças nas componentes de I/Q
- Distorção de segunda ordem
- Ruído 1/*f*
- Radiação do LO

Estes problemas serão descritos com mais detalhe nos parágrafos seguintes.

2.4.2.1 Selecção de Canal

É mais difícil fazer uma boa selecção de canal usando um filtro passa-baixo activo do que usando um filtro implementado com componentes passivos discretos.

Qualquer que seja a disposição do filtro relativamente ao amplificador, normalmente o que vier primeiro na "cascata" de blocos deve obedecer a requisitos mais exigentes.

2.4.2.2 Offsets DC

Como na arquitectura homódina o LO e o sinal que atravessa o LNA têm a mesma frequência, o acoplamento entre o LO e a entrada do LNA pode mais facilmente provocar uma "auto-mistura", que introduzirá uma componente DC à saída do misturador. A amplitude dessa componente DC poderá até ser maior que a do sinal de banda base, chegando a saturar os circuitos subsequentes, o que impediria uma amplificação correcta do sinal de banda-base.

É necessário cancelar os *offsets*, de alguma forma. A solução mais óbvia, de simplesmente aplicar um filtro passa-alto, não é aceitável. Segundo [2], num canal de 200kHz, a simples filtragem da banda 0–20 Hz é suficiente para elevar o BER para valores superiores a 10^{-3} , e não conseguiria cancelar variações mais rápidas no *offset*. Existem duas abordagens mais adequadas:

- Os dados a transmitir podem ser codificados de forma a não ter componente DC (em contrapartida a largura do canal aumenta ou a taxa de transmissão diminui), o que permite filtrar passa-alto com uma frequência de corte mais elevada.
- Se a recepção de sinal não for contínua (se e.g. for utilizado TDD), o *offset* pode ser armazenado num condensador⁷, durante o período de inactividade do receptor, e durante a recepção o cancelamento pode permitir a passagem de praticamente todas as frequências.

2.4.2.3 Diferenças nas Componentes de I/Q

Em [2] diz-se que para modulação de fase ou frequência, um receptor homódino tem de separar as componentes em fase e quadratura (I e Q). Se houver diferença entre os ganhos ou fases de I e Q, o sinal desmodulado será afectado. Os efeitos são bem ilustrados em [2] (fig.s 5.20–21).

A arquitectura heteródina é menos sensível a esse problema pois as componentes I/ Q são separadas a uma frequência mais baixa, e por isso os trajectos percorridos pelas componentes são menos afectados por parasitas. Para além disso, a maior parte da amplificação já terá sido aplicada antes da separação I/Q, pelo que são necessários menos estágios da ganho depois.

⁷Br.: capacitor.

2.4.2.4 Distorção de Segunda Ordem

Se por exemplo o LNA apresentar uma não-linearidade do tipo $y(t) = \alpha_1 x(t) + \alpha_2 x_2(t)$, isso pode provocar interferência sobre o sinal de banda-base através de vários mecanismos:

- Se houver uma frequência próxima do canal então o sinal recebido pode ser da forma x(t)
 = A₁ cosω₁ t + A₂ cosω₂ t, e y(t) conterá um termo da forma α₂ A₁ A₂ cos (ω₁ ω₂) t. Como todos os misturadores terão inevitavelmente algum *feedthrough* da entrada de RF para a saída, uma versão deste sinal de baixa frequência, atenuado talvez por 30–40dB, irá ser sobreposto ao sinal de banda-base.
- Se o sinal recebido apresentar variações na amplitude da forma x_{in}(t) = (A + ε cosω_mt) × (a cosω_C t + b sinω_C t), devidas a filtragem no transmissor ou perturbações na propagação, a distorção de segunda ordem introduzirá um termo da forma (a² + b²) A ε cosω_m t. Ou seja, a distorção de segunda ordem "desmodula" componentes AM. Se esse termo de frequência ω_m atravessar o misturador com atenuação finita, irá corromper o sinal de interesse.

A distorção de segunda ordem também pode ser quantificada através de um parâmetro IIP₂, definido de forma semelhante ao IIP₃. Segundo [2], a utilização de um LNA e misturador diferencial permitiria reduzir a distorção de segunda ordem. Só que a conversão de/ para a forma diferencial iria introduzir perdas de vários decibéis. Para além disso, um circuito LNA diferencial teria um maior consumo de energia.

2.4.2.5 Ruído 1/f

A amplificação no LNA e no misturador é tipicamente de cerca de 30dB, do que resulta um nível de sinal da ordem das dezenas de microvolt. Por isso, o ruído dos estágios seguintes ainda é crítico. Numa implementação CMOS o ruído 1/f irá afectar o sinal, o que será especialmente inconveniente se o espectro do sinal tiver componente DC. Por esta razão é desejável ter um ganho mais elevado na parte RF, utilizando por exemplo misturadores activos em vez de passivos.

Se for utilizada uma codificação de sinal que elimine a componente DC do sinal, também é possível simplesmente filtrar o ruído.

2.4.2.6 Radiação do LO

Se houver um acoplamento parasita do LO para a antena, o sinal do LO será radiado e poderá constituir uma interferência para outros receptores na mesma banda de frequências.

2.4.3 Desmodulador de Quadratura

A maior parte dos CIs utilizam um desmodulador de quadratura, pois os componentes dependentes da frequência (o indutor, o condensador e a resistência de amortecimento) são externos ao CI [15].

A parte do desmodulador que é implementada internamente é o multiplicador de quatro quadrantes, que pode ser integrado de forma económica e é facilmente ligado aos componentes externos.

O detector de quadratura opera da forma indicada na fig. 2.6.



Figura 2.6 Detector de quadratura.

2.4.4 Receptores de Rejeição de Imagem

Existem duas principais arquitecturas de rejeição de imagem...

2.4.4.1 Arquitectura de Hartley

Esta arquitectura é descrita pela fig. 2.7.

Se admitirmos que à entrada do receptor temos

$$x(t) = A_{\rm RF} \cos \omega_{\rm RF} t + A_{IM} \cos \omega_{IM} t,$$

então no ramo superior tem-se

fica



Figura 2.7 Arquitectura de Hartley.

$$x(t)\sin\omega_{\rm LO}t = \frac{A_{\rm RF}}{2}\sin(\omega_{\rm LO} + \omega_{\rm RF})t + \frac{A_{\rm RF}}{2}\sin(\omega_{\rm LO} - \omega_{\rm RF})t + \frac{A_{\rm IM}}{2}\sin(\omega_{\rm LO} + \omega_{\rm IM})t + \frac{A_{\rm IM}}{2}\sin(\omega_{\rm LO} - \omega_{\rm IM})t$$

Depois do filtro passa-baixo (e admitindo injecção low-side a título de exemplo)

$$\frac{A_{\rm RF}}{2}\sin(\omega_{\rm LO} - \omega_{\rm RF})t + \frac{A_{\rm IM}}{2}\sin(\omega_{\rm LO} - \omega_{\rm IM})t$$
$$= -\frac{A_{\rm RF}}{2}\sin(\omega_{\rm LO} - \omega_{\rm RF})t + \frac{A_{\rm IM}}{2}\sin(\omega_{\rm LO} - \omega_{\rm IM})t$$

e depois da desfasagem de 90°(que transforma sin $x \text{ em} -\cos x$),

$$\frac{A_{\rm RF}}{2}\cos(\omega_{\rm LO}-\omega_{\rm RF})t - \frac{A_{\rm IM}}{2}\cos(\omega_{\rm LO}-\omega_{\rm IM})t.$$

Já no ramo inferior tem-se

$$x(t)\cos\omega_{\rm LO}t = \frac{A_{\rm RF}}{2}\cos(\omega_{\rm LO} + \omega_{\rm RF})t + \frac{A_{\rm RF}}{2}\cos(\omega_{\rm LO} - \omega_{\rm RF})t + \frac{A_{\rm IM}}{2}\cos(\omega_{\rm LO} + \omega_{\rm IM})t + \frac{A_{\rm IM}}{2}\cos(\omega_{\rm LO} - \omega_{\rm IM})t$$

,

e depois do filtro passa-baixo

2.4 Arquitecturas

$$\frac{A_{\rm RF}}{2}\cos(\omega_{\rm LO}-\omega_{\rm RF})t + \frac{A_{\rm IM}}{2}\cos(\omega_{\rm LO}-\omega_{\rm IM})t.$$

Assim, à saída do somador cancelam-se os termos na frequência $\omega_{LO} - \omega_{IM}$, ou seja, conseguiu-se cancelar a frequência-imagem.

O funcionamento também pode ser explicado de forma gráfica, como poderá ser visto em [2], na fig. 5.26.

Segundo [2], é frequente o bloco de 90° de desfasagem ser substituído por dois blocos introduzindo, cada um, 45° de desfasagem.

A principal desvantagem desta arquitectura é a sensibilidade aos diferentes ganhos e desfasagens introduzidos em cada um dos percursos do sinal. Razavi [2] faz uma análise quantitativa desses efeitos e define o IIR (*Image Rejection Ratio*). Valores típicos para esta rejeição de imagem seriam 30 a 40dB. Um valor total, incluindo a rejeição devida a outros estágios de filtragem, seria de 60 a 70dB. No entanto, é dito que isso exige uma semelhança ou "emparelhamento" de ganhos muito maior de que para arquitecturas homódinas.

A integração da arquitectura de Hartley coloca ainda os seguintes problemas: uma vez que os filtros passa-baixo integrados não irão apresentar uma supressão de interferências tão boa, nos canais vizinhos, a linearidade do somador é crítica. Para além disso, as perdas e o ruído no estágio de desfasagem são geralmente importantes.

2.4.4.2 Arquitectura de Weaver

Esta arquitectura é descrita pela fig. 2.8. O princípio de funcionamento é semelhante ao da arquitectura de Hartley. A diferença é que o bloco de desfasagem é substituído pela multiplicação por sinusóides desfasadas, com uma segunda frequência. O funcionamento no domínio da frequência está ilustrado em [2] (fig.s 5.29–30).

A segunda frequência pode, ou não, ser escolhida de forma a que o espectro do sinal de saída fique já em torno do zero. Se isso não acontecer, coloca-se o problema de uma frequência-imagem secundária! A eliminação dessa imagem secundária obriga a utilizar filtros passa-banda em vez de filtros passa-baixo.

De acordo com [2], esta arquitectura evita alguns dos problemas associados às arquitecturas homódinas. A distância entre ω_1 e ω_{RF} permite que o acoplamento entre LO e a



Figura 2.8 Arquitectura de Weaver.

entrada seja minorado e as componentes DC devidas a auto-mistura sejam filtradas. Também é possível introduzir ganho suficiente na IF e no segundo conjunto de misturadores de forma a tornar o ruído 1/f desprezável na banda-base.

O problema introduzido por diferenças de fase, já existente na arquitectura de Hartley, mantém-se.

2.4.4.3 Receptores de IF Digital

Numa arquitectura que use duas ou mais IFs é possível efectuar as operações de baixa frequência, como a segunda mistura e filtragem, sobre um sinal já digitalizado. Isto evitaria o problemas das diferenças entre componentes I/Q do sinal.

Só que para um valor de IF de 50–200MHz, seria necessário um conversor AD com uma taxa de aquisição de 100–400MHz e uma resolução de mais de 14 bits. Essas especificações são actualmente inalcançáveis.

É possível reduzir a taxa de amostragem necessária a metade fazendo a *down-conversion* por sub-amostragem. Se a frequência de amostragem f_S for ligeiramente inferior a f_{IF} , o espectro do sinal será transladado para f_{IF} - f_S . Segundo [2], esta técnica é utilizada em equipamentos fixos que precisem de receber e processar um grande número de canais em simultâneo, mas ainda não é viável para equipamentos móveis, de baixo consumo.

2.4.4.4 Receptores de Sub-Amostragem

Se um sinal de RF de banda estreita for amostrado a uma frequência muito mais baixa, o espectro do sinal será transferido para uma banda mais baixa. Isso permite simplificar o projecto do LO e do seu sintetizador. O circuito de amostragem também será mais simples de projectar do que um circuito de mistura. A grande desvantagem desta arquitectura é que o ruído de entrada sofrerá *aliasing* e por isso a sua densidade espectral de potência será multiplicada por 2m, em que *m* é a razão entre a frequência RF e a frequência de amostragem. Outro problema é a amplificação do ruído de fase do relógio de amostragem ser por m^2 .

2.4.5 Transmissores

Segundo [2], existem menos arquitecturas de transmissão que de recepção porque no transmissor os problemas de ruído, rejeição de interferências e selectividade são menos marcados.

Os estágios principais de um transmissor são a modulação, a *upconversion* e a amplificação de potência. Pode não haver um estágio separado de *upconversion*, e nesse caso fala-se de uma arquitectura de conversão directa.

Num transmissor FM ou FSK, poderia pensar-se que o sinal a transmitir seria simplesmente aplicado a um VCO. Mas normalmente existirá uma etapa de condicionamento de sinal para filtrar o sinal de banda-base (e assim minimizar a largura de banda ocupada e eventualmente a interferência intersimbólica) e para adaptar automaticamente o ganho, de forma a compensar variações nas características do VCO devidas ao fabrico. O próprio oscilador precisará de uma malha de realimentação para controlar as variações com o tempo e com a temperatura.

O condicionamento do sinal de banda base pode ser feito digitalmente, como exemplificado em [2] (fig.s 5.37–38).

Nos tipos de modulação que envolvem separação da componente em fase e em quadratura do sinal, também se colocam os problemas de simetria entre os dois percursos das componentes I/Q.

Já o problema do ruído no misturador é menos sério, pois o sinal de banda base é produzido no transmissor e por isso é suficientemente forte.

O principal problema das arquitecturas de conversão directa é o facto de o espectro do LO ser corrompido pelo sinal de saída do transmissor. Existe uma variante da arquitectura de conversão directa em que são utilizados dois VCOs com frequências ω_1 e ω_2 , cujas saídas são multiplicadas de forma a aproveitar a componente do produto em $\omega_1 + \omega_2$. Se as frequências originais estiverem suficientemente afastadas da frequência de saída, a interferência é minimizada.

Se houver duas ou mais etapas de *upconversion*, esse problema também é resolvido, com a vantagem adicional de permitir modulação de quadratura a frequências mais baixas e portanto com melhor simetria. Um inconveniente é que o filtro passa-banda da segunda *upconversion* pode requerer uma implementação discreta e com componentes relativamente caros.

2.4.6 Escolha da Frequência do LO

Em [15] é apresentado um exemplo de *down-conversion* de um sinal de 144,45 MHz para uma IF de 10,7 MHz (um dos valores mais comuns). Há duas escolhas possíveis para a frequência de LO:

- 155,15 MHz (seria *high-side injection*, ou *supradyne*, pois LO > 144,45 MHz)
- 133,75 MHz (seria low-side injection, ou infradyne, pois LO < 144,45 MHz)

A escolha dependerá sobretudo da ocupação do espectro adjacente. Se, por exemplo, utilizarmos um LO de 155,15MHz, estaremos a fazer *down-conversion* do nosso sinal de 144,45MHz para os 10,7MHz, como pretendido, mas também estaremos a fazer *down-conversion* de qualquer outro sinal ou ruído que se encontre nos 165,85MHz, "para cima" dos mesmos 10,7MHz (ver fig. 2.9, e comparar com fig. 2.4 da pág. 46).

Por isso se diz que teremos uma frequência-imagem de 165,85MHz. Já se o LO fosse de 133,75MHz, a frequência-imagem seria de 123,05MHz.

O ideal seria escolher a frequência de LO de forma a que a frequência-imagem correspondesse a uma região "tranquila" do espectro. Mas isso nem sempre é praticável, pelo que é necessário aplicar uma forte filtragem antes do misturador, para atenuar a imagem.

2.4.7 Conclusões

Como medir o desempenho dos transceptores? O critério mais simples é a distância máxima de operação. Esta distância dependerá da potência entregue à antena e da sensibilidade do receptor (principalmente do ruído no LNA e do tipo de modulação) [2].



Figura 2.9 Down-conversion com injecção low-side (a) e high-side (b).

2.5 Componentes

Descrevem-se nesta secção os principais blocos funcionais que compõem os emissores e transmissores descritos na secção 2.1.

2.5.1 Problemas Gerais a Considerar

2.5.1.1 Não-Linearidade e Seus Efeitos

A propósito de classificar um sistema como linear ou não-linear, e variante ou invariante no tempo, [2] (pág. 12) chama a atenção para a necessidade de explicitar qual ou quais as entradas de interesse. Em [16] também são referidos alguns outros efeitos de distorção.

A resposta de um sistema não-linear, sem dinâmica (mas eventualmente variante no tempo), pode ser aproximada por

$$y(t) = \alpha_1 x(t) + \alpha_2 x_2(t) + \alpha_3 x_3(t)$$
.

2.5.1.1.1 Distorção Harmónica

Uma consequência imediata é o surgimento de harmónicas da frequência fundamental. Efectuando algumas aproximações, a *n*-ésima harmónica cresce na proporção de A^n .

2.5.1.1.2 Compressão de Ganho

Na maioria dos circuitos reais, a saída tem uma saturação. Isto ocorre se $\alpha_3 < 0$. Uma forma de quantificar este efeito é pelo ponto de compressão de 1 dB, definido como o valor de amplitude A para o qual o ganho diminui 1 dB. De acordo com as aproximações acima,

$$A_{1\rm dB} = \sqrt{0.145 \left| \frac{\alpha_1}{\alpha_3} \right|}$$

O valor dá-nos a noção de gama de tensões admissíveis na entrada do circuito. Valores típicos para amplificadores de *front-end* serão -20 a -25 dBm (63.2 a 35.6mV pico a pico para uma impedância de 50 Ω).

2.5.1.1.3 Perda de Sensibilidade e Bloqueio

Trata-se de um "efeito secundário" da compressão de ganho, quando um circuito com saturação recebe um sinal de pequena amplitude e um outro sinal de interferência de grande amplitude. Este último provoca uma redução do ganho que torna o circuito "insensível" a sinais de pequena intensidade.

Tipicamente pretende-se que um receptor RF resista a bloqueio por sinais até 60– 70 dB mais intensos que o sinal desejado.

2.5.1.1.4 Modulação Cruzada

Este efeito ocorre quando a modulação (ou simples variação aleatória de amplitude) presente num sinal de interferência provoca modulação também na amplitude do sinal de interesse.

2.5.1.1.5 Intermodulação

Quando dois sinais de diferentes frequências são aplicados a um sistema nãolinear, à saída surgirão componentes que não são harmónicas das frequências de entrada, resultando da multiplicação dos dois sinais.

Para frequências de entrada $\omega_1 \in \omega_2$, os produtos de terceira ordem terão frequências $2\omega_1 - \omega_2 \in 2\omega_2 - \omega_1$. Assim, este efeito é mais visível quando os valores de $\omega_1 \in \omega_2$ são próximos. Frequentemente este efeito é quantificado em dBc, em que o "c" indica um valor relativo à portadora (*carrier*, em inglês).

A corrupção de sinais devida à presença de duas interferências próximas é tão comum e crítica que foi definida uma forma de quantificar esta característica, o *third intercept point*, ou IP₃. Também se definem as amplitudes de entrada e saída correspondentes ao IP₃ como IIP₃ e OIP₃, respectivamente.

2.5.1.1.6 Estágios em Série

Como num sistema de RF os sinais são processados por vários estágios ligados em série, é útil poder calcular o IP₃ do conjunto a partir do IP₃ e do ganho de cada estágio.

Em [2] (sec. 2.1.2) afirma-se que é interessante a possibilidade de obter um valor de IP₃ arbitrariamente alto dependendo dos coeficientes da não-linearidade de cada estágio. É óbvio que isso significa que a não-linearidade de um bloco pode cancelar a do outro, se as funções de transferência forem uma a inversa da outra. Mas na prática será difícil aproveitar esse facto.

Após uma análise aproximada, o autor conclui que

$$\frac{1}{A_{\mathrm{IP}_3}^2} \approx \frac{1}{A_{\mathrm{IP}_{3,1}}^2} + \frac{\alpha_1^2}{A_{\mathrm{IP}_{3,2}}^2} + \frac{\alpha_1^2 \beta_1^2}{A_{\mathrm{IP}_{3,3}}^2} + \dots ,$$

ou seja, a não-linearidade dos últimos estágios torna-se cada vez mais crítica.

2.5.1.2 Interferência Inter-Símbolos

Mesmo os sistemas lineares e invariantes no tempo (SLITs) introduzem uma certa distorção num sinal, em virtude da sua largura de banda limitada. Um exemplo disso é quando se pretende descodificar uma sequência aleatória de "zeros" e "uns" que tenha passado por um filtro passa-baixo. Se os "uns" forem representados por pulsos rectangulares e os zeros "pela" ausência de pulso, a discriminação entre um símbolo e outro é dificultada pelo facto de a resposta do filtro ao pulso rectangular apresentar um decaimento prolongado.

Há duas abordagens para minorar esse problema: a equalização, que consiste num filtro colocado no receptor que tenta compensar o efeito do filtro do emissor (ou do canal de transmissão), e a utilização de sinais de Nyquist (*Nyquist signaling* ou *pulse-shaping*).

A ideia dos sinais de Nyquist é simples: trata-se de utilizar impulsos com forma não rectangular, que tenham um espectro limitado na frequência, de forma a não serem distorcidos pela limitação em largura de banda da transmissão, e que ao mesmo tempo tenham valor nulo noutros instantes de amostragem.

Um impulso do tipo sinc(*t*) pode satisfazer esses requisitos. Mas a função sinc teria algumas desvantagens: o filtro seria difícil de projectar, e o sistema ficaria vulnerável a erros no instante de amostragem. Uma generalização da função sinc é um impulso com espectro de "cosseno levantado" (*raised cosine*), ou seja, um espectro plano em que os flancos têm forma de arcada de cosseno. O sinal correspondente é dado por

$$p(t) = \frac{\sin \pi t/T_S}{\pi t/T_S} \cdot \frac{\cos \pi \alpha t/T_S}{1 - 4\alpha^2 t^2/T_S^2},$$
(16)

onde o parâmetro α (0 < α < 1) permite ajustar as características do pulso.

2.5.1.3 Adaptação de Impedância

Em [2] (pág. 50) é explicada a necessidade utilizar circuitos de adaptação de impedância⁸. A ideia geral consiste em fazer a fonte de sinal "ver" uma carga de um valor que drene maior potência. Em RF isso é normalmente feito com componentes passivos, pois componentes activos não permitiriam obter o mesmo desempenho.

Uma forma de adaptar impedâncias seria pelo uso de transformadores. Se um transformador ideal tiver uma razão de número de espiras de m, pode reduzir uma impedância por m^2 . No entanto, transformadores de alta-frequência apresentam perdas, acoplamento capacitivo entre primário e secundário e até mesmo ressonâncias indesejadas. Tudo isso complica o projecto e requer uma modelação cuidadosa. Por essa razão, frequentemente se utilizam circuitos baseados em condensadores e bobines.

Uma resistência pode ser convertida num valor mais elevado utilizando um divisor capacitivo, que aumenta a resistência aparente por um factor de $(1 + C_P / C_1)^2$. O mesmo pode ser feito com um divisor indutivo, que introduzirá um factor de $(1 + L_1 / L_P)^2$.

Uma resistência pode também ser convertida num valor menor, utilizando um condensador e uma bobine funcionando perto da ressonância. A resistência é dividida por $(C_P \omega)^2$.

⁸Br.: *casamento* de impedância.

As transformações de circuitos apresentadas têm contudo limitações à sua validade: é necessário que o valor do factor de qualidade Q seja elevado e que a faixa de frequências de interesse seja restrita.

2.5.2 Amplificadores de Baixo Ruído

Um amplificador de baixo ruído (LNA, *Low-Noise Amplifier*) para RF pode consumir muitas horas de simulação e optimização, na tentativa de equilibrar requisitos de ruído e adaptação de impedância com minimização de intermodulação, ao mesmo tempo que se procura minimizar o consumo [17].

Outra possível fonte de dificuldade para o projecto do LNA é o facto de o ruído da corrente da porta, em implementações CMOS, estar parcialmente correlacionado com o ruído da corrente de dreno [18]. Mas o valor do factor de ruído tende a melhorar com a evolução da tecnologia.

Um outro efeito do encurtamento do canal é que a característica deixa de ser quadrática e passa a ser linear, o que acarreta uma melhor linearidade dos dispositivos de canal curto. Assim, dependendo do critério de comparação, o desempenho do MOSFET pode até ser bastante mais próximo do desempenho do BJT do que geralmente se considera.

Uma técnica que pode ser utilizada para ajustar a impedância de entrada do amplificador consiste em colocar um indutor em série com a porta do transístor.

Em [18], o autor refere também que existe um compromisso entre largura de canal do transístor e frequência de funcionamento. Para uma impedância de entrada de 50 Ω , o produto destes factores será de aproximadamente 500 μ m · GHz.

2.5.3 Amplificadores de Potência

De acordo com [2] (cap. 9), o projecto de amplificadores de potência (PAs, do inglês *Power Amplifiers*) permanece um problema difícil. Na prática são frequentemente necessárias várias tentativas para obter um bom circuito, pelo que normalmente se preferem implementações discretas ou híbridas.

2.5.3.1 Exemplo

Um transmissor deve fornecer 1W a uma antena de 50 Ω . Sabemos que $P = V_{\rm rms}^2 / R$, pelo que $V_{\rm rms} = \sqrt{PR}$ e portanto $V_{\rm rms} = \sqrt{50}$ e $V_{\rm pp} = 2\sqrt{2}V_{\rm rms} = 20$ V.

Portanto, a corrente de pico na carga deve ser $I_{\rm p} = \frac{V_{\rm pp}/2}{R} = 10/50 = 200 \,\mathrm{mA}$.

Para uma das topologias de amplificação mais simples, a de fonte comum, com a carga ligada entre a alimentação e o dreno do transístor MOS, é necessário que a alimentação tenha um valor superior a V_{pp} . Se a topologia for alterada pela introdução de um bobine (designada *radio frequency choke*, ou RFC), é possível reduzir a tensão de alimentação a metade. No entanto, o transístor continua a ser submetido a uma tensão de pico idêntica, e uma redução de V_{DD} a 10V não melhora muito as coisas, do ponto de vista da alimentação do sistema portátil.

Mas também pode ser utilizado um bloco de adaptação de impedância. Por exemplo, um transformador com uma razão de número de espiras de 1:4 permite reduzir a tensão pico-a-pico de 20 para 5V, mas a corrente de pico terá de aumentar de 200 para 800mA, a que se irá somar a corrente do RFC, pelo que o transístor terá de suportar 1,6A de corrente de pico!

Para uma corrente $i = A \sin 2\pi ft$, tem-se $\frac{di}{dt} = 2\pi fA \cos 2\pi ft$ e portanto o *slew-rate* máximo para f = 900 MHz será $2\pi \times 900 \cdot 10^6 \times 1,6 \approx 9 \cdot 10^9$ A/ns. Com tal valor de *slew-rate*, mesmo uma indutância parasita de 10pH causa uma queda de tensão de 90mV. Da mesma forma, uma resistência de algumas dezenas de mili-ohm em série com o transístor, a RFC ou a malha de adaptação de impedância pode provocar perdas significativas. Por estas razões, muitos detalhes de *layout* que podem normalmente ser ignorados noutros circuitos analógicos e de RF são essenciais em amplificadores de potência. Note-se ainda que devido às perdas no componente activo, no PA tal como no LNA, utiliza-se apenas um transístor.

Em [2] (pág. 300) é apresentada uma tabela com valores típicos para o desempenho de um PA de sistema portátil. Por exemplo, a potência de saída é de 20–30dBm, a eficiência de 30–60% e o ganho de 20–30dB.

2.5.3.2 Classificação de Amplificadores de Potência

Os PAs mais antigos (classes A, B e C) são projectados de forma a terem tensões e correntes de saída sinusoidais. Isso limita as eficiências que podem ser atingidas. Nas classes E e F utilizam-se circuitos de adaptação de impedância que tratam distintamente a frequência fundamental e as primeiras harmónicas, o que permite fazer com que o transístor não esteja

simultaneamente submetido a tensões e correntes elevadas. Dessa forma reduz-se a dissipação no transístor e aumenta-se a eficiência do amplificador. Na classe F, em particular, a amplificação é não-linear, o que permite atingir eficiências ainda maiores.

2.5.4 Misturador

O misturador deve multiplicar a portadora pelo sinal gerado pelo oscilador local (LO), usando um dispositivo não-linear (e.g. um díodo) [15]. Admitindo que esses sinais são do tipo $\cos \omega_{\text{LO}} t e A \cos \omega_{\text{RF}} t$, um multiplicador ideal deveria gerar termos nas frequências ω_{LO} + $\omega_{\text{RF}} e \omega_{\text{LO}} - \omega_{\text{RF}}$. Uma destas frequências será o ω_{IF} de interesse. Mas na realidade, todo um conjunto de outros termos será produzido, dependendo do tipo de circuito misturador utilizado.

O circuito misturador mais comum nos CIs actuais é o de configuração duplamente equilibrada. O termo "equilíbrio" descreve o grau de isolamento,

- da entrada LO para a entrada RF
- das entradas para a saída IF

Os circuitos duplamente equilibrados apresentam um melhor isolamento. Para além disso, todos os protestos excepto $\omega_{LO} + \omega_{RF}$ e $\omega_{LO} - \omega_{RF}$ são cancelados, o que facilita a filtragem da frequência IF resultante.

Os principais parâmetros do misturador são:

Ganho de conversão. Um ganho maior facilita o projecto do filtro passa-banda, pois o limitador subsequente irá impor uma amplitude mínima aceitável.

Impedância de entrada/saída. Deve estar adaptada aos blocos funcionais adjacentes.

Third Order Intercept. Trata-se de uma medida da supressão dos produtos de distorção. É normalmente medida em dBm, e quanto maior for, melhor.

2.5.5 Oscilador Local

Os principais parâmetros de um LO implementado no próprio CI são [15]:

- Estabilidade
- Poder de drive

Uma forma de melhorar a estabilidade térmica do oscilador é pela utilização de componentes com coeficientes térmicos complementares. Se for necessária uma grande estabilidade, é recomendável um cristal, que ainda permite um ajuste de frequência através da colocação de um indutor ajustável em série.

2.5.5.1 Ruído de Fase em Osciladores

Os dois principais tipos de imperfeição espectral em osciladores são a presença de ruído de banda larga e a presença de tons espúrios.

Durante muito tempo, a forma como o ruído de banda larga surgia num oscilador não foi bem entendida. Recentemente conseguiu-se uma melhor compreensão desse fenómeno através da aplicação de um modelo linear variante no tempo. A utilidade desse modelo pode ser explicada em termos qualitativos pelo facto de um impulso de ruído afectar mais a fase da oscilação se ocorrer perto da passagem por zero⁹.

O modelo variante no tempo aplicado a osciladores contribuiu também para solucionar o problema de translação na frequência do ruído 1/f, em que o ruído de baixa frequência característico dos MOSFETs surgia de alguma forma a frequências próximas da portadora. Esse efeito pode ser atenuado através duma topologia com dupla simetria, apresentada em [18], na fig. 16, com a qual os autores obtiveram um ruído de fase de -121 dBc/Hz.

2.5.6 Amplificador de IF

Este elemento é utilizado para amplificar a saída do misturador, fornecendo assim ao limitador um sinal com amplitude suficiente. Também deve fornecer corrente à entrada do circuito de RSSI (ver adiante). Os seus principais parâmetros são [15]:

- Impedância de entrada
- Ganho
- Impedância de saída

Historicamente, os dois valores de IF mais comuns são 455kHz e 10,7kHz. Filtros cerâmicos para estas frequências podem ser facilmente adaptados a amplificadores de IF comerciais usando resistências externas. É importante esta adaptação de impedâncias para que

⁹Este facto explica inclusive o melhor desempenho de um oscilador Collpits comparado com um oscilador em anel: o primeiro recebe energia nos picos de oscilação, o segundo nas passagens por zero.

os filtros cerâmicos cumpram as especificações.

2.5.7 Filtro de IF

Em [15] é apresentado um exemplo de componente comercial para filtragem de IF, o MC13156. É classificado, mais precisamente, como um circuito de *split IF*, pois pode ser aplicada filtragem ao sinal em dois pontos do sistema: depois do misturador e depois do amplificador de IF. Normalmente é mais fácil utilizar dois filtros do que apenas um, que teria de cumprir requisitos bastante mais rigorosos.

O autor indica também uma expressão para estimar a largura de banda (3dB) resultante de dois filtros,

$$LB_{TOT} = \frac{1}{\sqrt{\frac{1}{LB_1^2} + \frac{1}{LB_2^2}}},$$

e chama a atenção para a necessidade de calcular correctamente a largura de banda necessária para os diferentes estágios do sistema. Apresenta ainda um exemplo, para modulação GFSK.

2.5.8 Circuito de RSSI

O circuito de RSSI (do inglês *received signal strength indicator*) faz o que o nome sugere: indica a intensidade do sinal recebido. Essa intensidade é calculada pelo logaritmo da soma das correntes de saída do amplificador de IF e do limitador. Os principais parâmetros do RSSI são [15]:

- Gama dinâmica (medida em dB)
- Ganho (medido em μ V/dB ou μ A/dB)
- Impedância de saída

2.5.9 Limitador

O limitador remove as variações de amplitude que o sinal possa ter à saída do amplificador de IF. Isso é necessário porque normalmente o desmodulador também responde a variações de amplitude. Os principais parâmetros do limitador são [15]:

- Ganho de limitação
- Impedância de entrada

2.5.10 Data Slicer

Este bloco reconstrói os dados digitais a partir do sinal de banda base resultante da desmodulação. Normalmente o sinal de banda base terá sido fortemente filtrado nos estágios de IF do transmissor e receptor, com o objectivo de limitar o espectro utilizado e restringir o ruído.

Se não fosse aplicada essa filtragem, a transmissão de uma onda quadrada com degradação mínima das transições implicaria transmitir pelo menos até à sétima harmónica, o que multiplicaria por sete a largura de banda necessária (ver exemplo de cálculo em [15]).

Depois desta filtragem, o sinal digital de banda base tornou-se praticamente sinusoidal e não pode ser directamente usado num sistema digital.

O *data slicer* é basicamente um comparador, mas possui também um ajuste automático de limiar de comparação que acompanha o valor DC da sequência de dados recebida. As variações no valor DC são devidas à ocorrência de longas sequências de "zeros" ou "uns".

2.6 Tecnologias

A tecnologia de fabrico é determinante para a viabilidade técnica e económica de um sistema de RF. Em [20] é indicada uma referência interessante sobre a escolha de tecnologias, [19].

2.6.1 Tecnologias Orientadas para RF

As tecnologias mais importantes em RF são BiCMOS (Bipolar-CMOS) e AsGa (arsenieto de gálio) [2]. As tecnologias de AsGa tem as desvantagens de:

- alto custo
- baixo yield (rendimento de produção)
- elevado consumo,

mas continuam a ser importante, sobretudo para amplificadores de potência e inter-

ruptores do front-end. Têm as vantagens de:

- maior valor de $V_{breakdown} \times f_{cutoff}$
- substrato semi-isolante
- capacitores e indutores de alta qualidade

2.6.2 Tecnologia CMOS

Tradicionalmente, a tecnologia CMOS tem sido menos apropriada para aplicações de RF que outras tecnologias como a bipolar e a de AsGa. Por exemplo, os dispositivos bipolares apresentam valores de g_m/I superiores, e os dispositivos de AsGa têm maior mobilidade de portadores.

A tecnologia CMOS tornou-se recentemente mais viável (frequência de trânsito de dezenas de GHz para a geração de 0,35 µm) [2], mas tem ainda problemas de acoplamento pelo substrato, variação de parâmetros com a temperatura e o processo, e modelação insuficiente de dispositivos. Muitas vezes são apenas indicados valores típicos, sem que sejam dadas garantias relativamente à variação dos valores [17].

No entanto, o aproveitamento da tecnologia CMOS também para aplicações RF tem-se revelado cada vez mais "apetecível", por razões económicas. Afinal, é através dessa tecnologia que tem sido cumprida a previsão de desenvolvimento da Lei de Moore¹⁰. Só que até agora este desenvolvimento tem sido mantido sobretudo pelo mercado de aplicações digitais, cujos requisitos nem sempre coincidem com os de aplicações lineares. Um dos principais requisitos das aplicações digitais é a velocidade. Segundo [18], a geração de 0,18µm permite valores de f_T de mais de 50 GHz. Isso é desejável também para aplicações RF, embora em circuitos lineares se possa trabalhar apenas a uma fracção do valor de f_T . Já outros requisitos essenciais para aplicações RF, tais como baixo ruído, linearidade e disponibilidade de bons componentes passivos, não são cobertos por tecnologias optimizadas para aplicações digitais. Mas surgiram recentemente técnicas que tentam "contornar" essas limitações da tecnologia CMOS.

A tecnologia CMOS tem ainda a vantagem, sobre as bipolares, da a ausência de *shot-noise* [17].

¹⁰Segundo esta, a cada 18 meses a densidade de integração duplica e o preço baixa para metade.

2.6.3 Integração de Componentes Passivos

2.6.3.1 Indutâncias

Os indutores são componentes críticos para circuitos como osciladores. Já têm sido obtidas indutâncias da ordem de 1 nH/mm com *bond wires*¹¹ ou pinos de encapsulamento, mas esses elementos apresentam uma forte variabilidade com o fabrico [18].

Indutores planares têm características mais previsíveis e repetíveis, pois são definidos pelo processo litográfico. E recentemente foram obtidas, para um conjunto de geometrias, fórmulas simples e com uma precisão da ordem dos 2% a 3% [18]:

$$L = \frac{\mu n^2 d_{\text{avg}} c_1}{2} \left(\ln \left(\frac{c_2}{\rho} \right) + c_3 \rho + c_4 \rho^2 \right)$$

Na fórmula acima, μ é a permeabilidade, *n* o número de espiras, d_{avg} a média dos diâmetros interno e externo. ρ e os coeficientes $c_1...c_4$ são valores tabelados que dependem da geometria da bobine.

Indutores CMOS apresentam mecanismos dissipativos que reduzem o seu factor de qualidade Q a valores entre 5 e 10. Existe dissipação pelo substrato por dois mecanismos.

O primeiro deles, é o acoplamento capacitivo com o substrato. Este efeito pode ser reduzido pela introdução de uma camada de protecção electrostática. Esta camada é designada em inglês como *patterned grounded shield*, ou PGS. *Grounded* pois o material condutor (geralmente metal ou poli-silício) está ligado ao potencial de terra, e *patterned* pois é uma camada de material condutor descontínua, com um padrão de ranhuras que evitam perdas por correntes turbilhonares. A introdução de um PGS permite melhorar o valor de *Q* por um factor de dois.

O segundo mecanismo torna-se importante para frequências mais elevadas, quando o efeito pelicular reduz a condução no substrato a uma camada mais fina e surgem perdas por correntes turbilhonares no substrato. Se a tecnologia de fabrico permitir vários níveis de metalização, podem ser implementados indutores verticais, que sofrerão menos deste problema mas em contrapartida apresentarão maior capacidade parasita entre espiras.

¹¹Fios de ligação entre o encapsulamento e a pastilha de silício.

2.6.3.2 Integração de Capacidade

Um dos problemas tradicionais de integrar condensadores lineares usando camadas condutoras paralelas é a inevitável capacidade parasita para o substrato. Outro problema é que a distância entre placas não foi diminuindo com a evolução das tecnologias, pois a maior capacidade assim obtida prejudicaria o desempenho dos circuitos digitais. Assim, os transístores "encolheram", e os condensadores não [18].

Surgiram recentemente capacitores integrados com geometrias que usam o fluxo eléctrico lateral, pois essas dimensões laterais *têm* diminuído. Esses "condensadores de fluxo lateral" (*lateral flux capacitors* ou LFCs) podem ter geometrias em pente ou geometrias fractais. As geometrias fractais têm menor auto-indutância e espera-se que em breve se atinjam valores de capacidade uma ordem de grandeza acima da de condensadores de placas paralelas com igual área. Outra medida de mérito é o produto *Q*-frequência, para o qual já se atingem valores de 100 a 200 GHz [18].

Outro componente essencial, o condensador variável, ou "varactor", pode utilizar a capacidade de junção entre uma difusão de fonte/dreno e o poço, mas como o poço tem normalmente uma resistência eléctrica elevada, é mais comum utilizar a capacidade de porta de um MOSFET. E com uma pequena modificação, a de usar difusões com o mesmo tipo de dopagem do poço, é possível usar o condensador com acumulação de portadores sob o óxido, o que reduz a resistência do canal e permite atingir valores de produto de *Q*-frequência da ordem dos 100GHz, e ainda melhorar esses valores a cada nova geração da tecnologia [18], [21].

2.6.4 Modelamento e Simulação

Outra dificuldade do projecto de CIs de RF é a falta de bons modelos, que calculem correctamente ruído e impedâncias de entrada para rádio-frequências. Isso deve-se sobretudo ao facto de os modelos tradicionais recorrerem à aproximação de quase-estacionaridade. A velocidade de trânsito dos portadores que atravessam o canal revela-se na forma de uma diferença de fase que aumenta com a frequência. Assim, a porta do MOSFET deixa de ser puramente capacitiva e começa a ser cada vez mais resistiva. Com a dissipação de potência que assim surge, surge também ruído térmico, que tem de ser considerado no projecto de um LNA.

Pesquisadores propuseram um modelo consistindo de uma condutância e uma fonte de corrente de ruído em paralelo com os terminais porta-fonte do transístor MOS. Os res-

pectivos valores seriam dados por

$$g_g = \frac{\omega^2 C_{gs}^2}{5g_{d0}}, \quad \frac{\overline{i_{ng}^2}}{\Delta f} = 4kT\delta g_g.$$

Consultar [18], pág. 1564, para mais detalhes.

Existem novos modelos de transístor que procuram melhorar a precisão a altas frequências [20]: VBIC, MEXTRAM e HICUM para BJTs, e Phillips MOS 9, BSIM3v3 e EKV para MOSFETs. Mas para além de imprecisões no modelo, o próprio simulador pode ter limitações que diminuem a sua utilidade para simulação em RF. Por exemplo, o SPICE "clássico" permite simular ruído sobre pequenos sinais, mas não sobre grandes sinais.

Em [2] (pág. 5) afirma-se que não-linearidade, variância no tempo e ruído requerem o estudo do espectro dos sinais, mas a análise AC do SPICE permite apenas modelos lineares e invariantes no tempo, pelo que é necessário fazer simulações no domínio do tempo e depois aplicar transformações para o domínio da frequência. O principal problema dessa abordagem é que, para discernir componentes de frequência muito próximas, torna-se necessário simular durante longos períodos de tempo. E se for injectado ruído aleatório no domínio do tempo, torna-se necessário calcular médias sobre o espectro. Outro problema é a necessidade de utilizar modelos de parâmetros distribuídos ou de parâmetros S, como por exemplo para filtros SAW.

Foi necessário desenvolver novos algoritmos de simulação optimizados para altas frequências e circuitos de modulação: *harmonic balance, shooting method*, análise de subespaço de Krylov, análise de regime estacionário periódico e quase-periódico, análise AC para ponto de funcionamento periodicamente variável, etc [20].

Os autores também referem sucintamente os problemas de ter um substrato comum a todo circuito.

2.7 Aplicações

2.7.1 Áreas de Aplicação

As aplicações de RF mais populares, para além da telefonia celular, são as seguintes (faixas de frequências citadas em [2], secção 1.3):

2.7 Aplicações

- Aparelhos telefónicos sem fio
- TV por satélite (10GHz)
- LANs de alta velocidade (900MHz e 2.4GHz)
- GPS (1.5 GHz)
- Terminais de dados ponto-a-ponto
- Telemetria
- Identificação e controlo de acesso (900MHz e 2.4GHz)

2.7.1.1 Telemetria

Em [22] é apresentada um exemplo de telemetria bastante curioso. Os autores começam por afirmar que o principal método de localização utilizado em pesquisa biológica de água doce¹² é a rádio-telemetria. É apresentado um caso de localização de trutas levado a cabo no estado de Montana, nos EUA. Biólogos capturam as trutas, anestesiam-nas, e implantam-lhes cirurgicamente um pequeno transmissor de rádio. Esse transmissor pesa entre 6 e 8,9 g (pouco menos de 2% da massa corporal do peixe), tem um tempo de vida de 237 a 344 dias, e transmite um sinal a cada cinco segundos, numa frequência de 150MHz.

O implante cirúrgico é colocado no ventre do peixe, e a antena (um filamento com cerca de dez centímetros de comprimento) fica no exterior do corpo do peixe, pendente como uma pequena cauda (ver fig. 2.10).



Figura 2.10 Implante cirúrgico de transmissor de RF numa truta (a antena, assinalada na figura, sai do corpo do peixe através de um orifício entre as barbatanas caudais).

¹²Não tão aplicável no mar, onde a água salgada, mais condutiva, atenua fortemente as ondas electromagnéticas.
2.7 Aplicações

Outro exemplo afim, mas ainda por concretizar, seria um transmissor consumível sob a forma de pílula para monitorização de dados fisiológicos [17].

Data do início da década de 1980 a disponibilização de partes do espectro para uso livre em aplicações de telemetria de baixa potência. Reproduzem-se na tabela 2.2 algumas das bandas que estão disponíveis num número significativo de países [23].

Frequência	Largura de Banda	Potência	Disponibilidade
49MHz	100MHz	100mW	Mundial
433 MHz	1,7 MHz	10mW	Europa e outros
868MHz	2MHz	25/500mW	Europa
2.4GHz	30MHz	100mW	Mundial

Tabela 2.2 Exemplo de atribuição de canais

Para além disso, [23] dá algumas indicações gerais sobre o grau de dificuldade em atingir determinadas especificações:

- 1. Alcance
 - 50m obtém-se facilmente com 1mW e uma antena simples.
 - 500m requer 10mW e uma antena eficiente e bem orientada.
 - 5km requer 100mW, uma boa antena e pelo menos 10m de altura do solo.
- 1. Velocidade de transmissão
 - 1 kbit/s é fácil de obter
 - 10kbit/s é aproximadamente o máximo para um sistema de banda estreita (canal de 25kHz).
 - 20kbit/s ou mais irá requerer um sistema de banda larga
 - 100kbit/s ou mais irá requerer um sistema projectado especificamente para o efeito

2.7.1.2 Controlo de Acesso

Em [24] são referidos numerosos CIs para utilização por fabricantes de aparelhos de controlo de acesso a automóveis. Já são bastante comuns chaves que permitem abrir as por-

tas de um carro, ou o portão de uma garagem, à distância. Essas "chaves" ou "controles remotos" baseiam-se em um de dois tipos de comunicação: infra-vermelhos ou RF. Os primeiros serão ligeiramente mais baratos mas em contrapartida requerem que haja uma linha de vista entre o emissor e o receptor. Já os sistemas que utilizam RF não têm essa limitação. A chave pode ser accionada sem mesmo tirá-la do bolso.

Os CIs apresentados em [24] utilizam as seguintes faixas de frequência:

- baixa-frequência (8–150kHz)
- rádio-frequência (13–2400 MHz)
- micro-ondas (5,8 e 24GHz)
- infra-vermelhos

Uma outra aplicação sugerida na mesma página para esses circuitos é na adaptação automática do carro ao condutor. Imagine-se um automóvel familiar que pertence a um casal, em que o marido e a esposa têm chaves diferentes. Ambas emitem o código que abre o automóvel, mas para além disso cada uma emite um código diferente que identifica o condutor. De acordo com esse código, o automóvel automaticamente ajusta a posição do assento do condutor, do volante e dos espelhos retrovisores.

2.7.2 Exemplos de Padrões e Sistemas

2.7.2.1 AMPS (Advanced Mobile Phone Service)

Origem: EUA, tendo sido um dos primeiros padrões de telefonia celular.

Modulação: FM analógico.

Acesso múltiplo: FDMA (canais de 30kHz, na faixa de 824–849MHz para móvel-base e na faixa de 869–894MHz para base-móvel).

Capacidade: 830 utilizadores ($\frac{894 - 869}{0.03} = \frac{849 - 824}{0.03} = \frac{25}{0.03} \approx 830$).

Duplex: FDD (a separação de 20MHz entre faixas permite baixas perdas).

2.7.2.2 NADC (North American Digital Cellular)

Origem: EUA, foi o primeiro padrão de telefonia celular digital.

Modulação: $\pi/4$ -DQPSK.

Acesso múltiplo: TDMA (mesmos canais de frequência que no sistema AMPS, para manter a compatibilidade¹³, mas com 6 utilizadores por canal).

Capacidade: aproximadamente 4980 utilizadores (830×6).

Taxa de transmissão: 48.6 kb/s por canal, cerca de 6.5 kb/s por utilizador.

Duplex: FDD (mas o TDMA também permite TDD¹⁴).

2.7.2.3 GSM (Global System for Mobile communications)

Origem: Europa. É hoje em dia o padrão de telefonia celular mais utilizado mundialmente.

Modulação: GMSK.

Acesso múltiplo: TDMA (canais de 200kHz, na faixa de 890–915MHz para móvel-base e na faixa de 935–960MHz para base-móvel, com 8 utilizadores por canal)¹⁵.

Capacidade: Cerca de 1000 utilizadores ($\frac{960-935}{0,2}8 = \frac{915-890}{0,2}8 = \frac{25}{0,2}8 = 125 \times 8 = 1000$).

Taxa de transmissão: 270 kb/s por canal, cerca de 24 kb/s por utilizador.

Duplex: FDD (mas o TDMA também permite TDD).

2.7.2.4 CDMA da Qualcomm

Origem: EUA, proposto pela empresa Qualcomm. Também conhecido pela sigla IS-95.

Modulação: OQPSK para móvel-base, QPSK para base-móvel¹⁶.

- Acesso múltiplo: DS-CDMA (mesmas *faixas* de frequência dos sistemas IS-54 e AMPS, mas com canais de 1.23 MHz partilhados por vários utilizadores).
- **Capacidade:** Aproximadamente 20 canais $\left(\frac{894-869}{1,23} = \frac{849-824}{1,23} = \frac{25}{1,23} \approx 20\right)$, com uma quantidade variável de utilizadores por canal¹⁷.
- **Taxa de transmissão:** 9.6 kb/s por utilizador (mas o codec utiliza uma taxa de transmissão variável, de 9.6, 4.8, 2.4 ou 1.2kb/s, para poupar energia e diminuir o ruído no canal).

¹³O padrão americano IS-54 (*interim standard* 54) define telefones celulares que podem funcionar com o sistema AMPS e com o NADC.

¹⁴Como a cada intervalo de 40ms o utilizador só emite durante 1/6 do tempo e também só recebe durante 1/6 do tempo, os intervalos de emissão e recepção não precisam de ficar sobrepostos, o que facilita o projecto do circuito transceptor.

¹⁵Posteriormente foram acrescentadas faixas em 1.8 GHz e 1.9 GHz. No Brasil é utilizada a faixa de 1.8 GHz.

¹⁶A modulação QPSK tem maior eficiência energética, algo particularmente importante para terminais móveis.

¹⁷O limite teórico (não referido em [2]) são 64 códigos disponíveis, donde resultam $20 \times 64 = 1280$ utilizadores.

Duplex: FDD (nada é dito em [2] sobre TDD).

2.7.2.5 DECT (Digital European Cordless Telephone)

Origem: Europa, orientado para aplicação em telefones sem fio.

Modulação: GFSK.

Acesso múltiplo: FDMA com canais de 1.73 MHz, na faixa de 1880–1900 MHz, e TDMA com 12 utilizadores por canal, que emitem durante 0.4 ms e recebem durante 0.4 ms, dentro de intervalos de 10 ms.

Capacidade: Approximadamente 132 utilizadores ($\frac{1900-1880}{1.73} \times 12 = 138$).

Taxa de transmissão: 1.15 Mb/s por canal, cerca de 38.8 kb/s por utilizador.

Duplex: TDD.

2.7.2.6 HomeRF

Em [25] é descrito o HomeRF, um sistema de rede RF doméstica. Pretende obter um alcance de 50m e utilizar uma frequência em torno dos 2,4GHz, para uma taxa de transmissão de 0,8–1,6Mbps.

Também é apresentada uma tabela comparativa entre diferentes sistemas— HomeRF, Bluetooth, IEEE 802.11 FH, HIPERLAN, HomePNA e IrDA (estas duas últimas redes utilizam cabo telefónico e luz infravermelha, e foram incluídas para comparação de custo/desempenho).

2.7.2.7 Bluetooth

O especificação do sistema Bluetooth começou a ser criada em Fevereiro de 1998 por cinco grandes empresas: Ericsson, Nokia, IBM, Toshiba e Intel, às quais, em Dezembro de 1999, se juntariam ainda a 3COM, Lucent, Microsoft e Motorola.

Eis um sumário das principais especificações do sistema [9].

- Modulação GFSK (como no padrão IEEE 802.11 para WLANs).
- Operação na faixa de frequência ISM de 2,4 a 2,4835 GHz.
- Envio de dados por pacotes, em *half-duplex*. Comutação entre emissão e recepção a cada 220 µs.

- Espalhamento espectral por *frequency-hopping* (frequência alterada a cada pacote enviado).
- Alcance de cerca de 10m.
- Arquitectura super-heteródina, com baixa IF (uma arquitectura homódina sofreria forte interferência dentro da sua faixa de frequência).

2.7.2.8 GPS (Global Positioning System)

O sistema GPS foi criado nos EUA para aplicações militares e emite sinais para uso civil em que a precisão é intencionalmente degradada, por razões de segurança militar [26]. No entanto, surgiu um grande mercado de aplicações civis, desde topografia a navegação aeronáutica, marítima e rodoviária, controlo de frotas, recreação, etc. Por isso é de esperar que a qualidade do sinal para aplicações civis venha a melhorar em futuras gerações do sistema, com a colocação de novos satélites em órbita.

Uma forma de melhorar a precisão da localização obtida é através do GPS diferencial, em que se compara o sinal recebido no local de interesse com o sinal recebido noutro local cujas coordenadas são conhecidas com precisão.

A Rússia possui também um sistema semelhante, o Glonass, mas que se encontra só parcialmente funcional. A União Europeia planeia lançar também um sistema de navegação por satélite, que se chamaria Galileo, e existem negociações entre a União Europeia e a Rússia para eventualmente construir esse futuro sistema aproveitando a infra-estrutura do Glonass.

2.8 Conclusões

2.8.1 Expectativas Tecnológicas

Durante muitos anos pensou-se que o receptor de rádio ideal seria um em que todo o espectro de sinais recebidos pela antena fosse digitalizado imediatamente, sendo a selecção de canal e desmodulação efectuadas digitalmente. Igualmente, o transmissor faria a síntese digital da portadora, sendo a antena alimentada pela saída do DAC. Esta abordagem teria a vantagem de permitir reprogramar o protocolo de transmissão por software.

Hoje em dia já é comum a digitalização ocorrer nas secções de IF (frequência intermédia), que trabalham na faixa de 10 a 300MHz [17]. A necessidade de um conversor de alta velocidade pode ser evitada utilizando uma IF muito baixa, ou deslocando o espectro RF directamente para a banda-base, usando técnicas de *zero-IF*. No entanto, o LNA e o misturador de abaixamento de frequência (*down-conversion*) do receptor continuarão a ser analógicos, enquanto que no transmissor o mesmo acontece para o *up-converter*, o amplificador de potência e o controlo de potência. Isto a não ser em casos excepcionais em que a potência do sinal recebido é elevada, a frequência da portadora baixa, e baixa também a potência do sinal transmitido.

Segundo [17], não é seguro que a aplicação do conceito de micro-sistema, ou SoC (*System on a Chip*) a um sistema de rádio venha a ser viável, por razões económicas, se os progressos na área de encapsulamento e montagem de baixo custo forem mais rápidos que os progressos na área de circuitos mistos, ou por razões técnicas. Por exemplo, a simples alteração da ordem dos pinos no encapsulamento, ou a forma como o chip é montado no encapsulamento, podem afectar o desempenho.

Outro exemplo em que a opção foi não integrar todas as partes do sistema foi em [25]. Os autores afirmaram que, embora seja possível integrar o *front-end* de RF (LNA, PA, chaves da antena e filtro passa-banda), preferiram separá-lo para fazer uma optimização deta-lhada e reduzir o consumo de energia.

Mas um sistema completamente integrado, mesmo que à partida não seja a solução óptima, do ponto de vista técnico ou económico, pode tornar possíveis novas aplicações. E essas aplicações poderão por sua vez torná-lo a alternativa mais atraente.

2.8.1.1 "Renascimento" do Projecto Analógico

Hoje em dia, apenas 5% dos CIs contêm componentes analógicos, mas espera-se, com aplicações como WLANs e Bluetooth, que dentro de cinco anos essa percentagem aumente para 70% [27].

A tendência actual seria de migrar de um fluxo de projecto *bottom-up* para um fluxo *top-down.*, com o surgimento de novas ferramentas de software de síntese analógica (embora ainda com funcionalidade limitada), por empresas como a Neolinear (Pittsburgh), a Barcelona Design (Sunnyvale, Califórnia), a Analog Design Automation (Otava, Canadá) e a Antrim Design (Scots Valley, Califórnia) [27]. Para além disso, é também sentida a necessidade de melhorar as técnicas de teste de circuitos analógicos. Uma solução considerada "ideal"

seria a integração de blocos de teste e instrumentação que fizessem o grosso das operações de teste *dentro* do CI.

2.8.2 Expectativas de Mercado

Nos EUA, companhias especializadas em RF e que tinham como principal cliente o sector militar, adaptaram o seu negócio para o mercado em expansão das telecomunicações civis [28]. Uma empresa citada como exemplo, a GEC Plessey Semiconductors, possuía alguns meios de fabrico próprios, como tecnologias CMOS de 0,7µm, tecnologia de silício sobre safira (SOS), módulos multi-chip, tecnologia bipolar para RF, etc.

A companhia afirma que o mercado de celulares analógicos se vai manter ainda por bastante tempo, porque a tecnologia digital é cara demais para os países sub-desenvolvidos.

O mercado que a GEC Plessey considera ter maior potencial de crescimento é o de WLANs funcionando na faixa de 2,4 a 2,483 MHz. Esta faixa, por ser de uso livre e já ser utilizada por diversas aplicações, como fornos de micro-ondas, requer cuidados especiais, para obter maior fiabilidade, como sejam a utilização de *frequency-hopping* e *spread-spectrum* [28].

Capítulo 3 Projecto do Circuito PLL

3.1 Descrição Geral

Um oscilador local, dependendo da aplicação em vista, pode ser mais ou menos complexo. Na sua forma mais simples, para transmissão ou recepção num canal pré-determinado, pode ser um oscilador de frequência fixa.

Osciladores de cristal de quartzo são muito precisos, mas não atingem frequências de RF, pelo que outros tipos de oscilador, operacionais a frequências mais elevadas, embora menos precisos, serão necessários. Estes osciladores de alta frequência, se forem de frequência fixa, necessitarão de um componente ajustável para corrigir desvios resultantes do processo de fabrico.

O oscilador local também pode ser, mais do que ajustável, variável, para fazer a sintonia entre diferentes canais de rádio.

Um oscilador que gere um sinal com uma frequência bem precisa, sendo o valor da frequência seleccionável de entre um conjunto relativamente vasto (algumas dezenas ou centenas de valores pré-definidos) é designado sintetizador de frequência.

Uma forma sofisticada de gerar uma frequência fixa com alta precisão consiste em utilizar um oscilador variável inserido numa malha de controle de fase (PLL), tendo como referência um oscilador de cristal de quartzo.

A malha de realimentação do PLL contém um bloco divisor de frequência, que normalmente é implementado com um contador digital. Se esse contador digital for programável, teremos então um sintetizador de frequência.

3.1.1 Características Funcionais

As principais características funcionais do sintetizador projectado são:

- Frequência de saída variando em torno dos 900MHz, em incrementos de 200kHz;
- Frequência seleccionável digitalmente, por meio de duas entradas de oito bits;
- Tensão de alimentação de 3.3 ou 5V.

3.1.2 Princípio de Funcionamento

Existem vários tipos de implementação possíveis para um sintetizador de frequência, nomeadamente [29]:

- Comutação entre um banco de osciladores de cristal.
- Divisão de frequência.
- Oscilador harmónico com banco de filtros.
- Multiplicação de frequência com misturadores.
- Malha de controle de fase (PLL), com divisão de frequência programável.

A última opção é de longe a mais popular para implementações microelectrónicas, tanto que em [2] é apresentada como sendo praticamente sinónimo de sintetizador de frequência.

O seu diagrama de blocos é apresentado na fig. 3.1.



Figura 3.1 Diagrama de blocos genérico de síntese de frequência por PLL.

O sinal de saída é gerado por um oscilador de Colpitts (um tipo de oscilador ressonante LC). Serão utilizados o indutor e o capacitor variável disponíveis na tecnologia escolhida, obtendo-se assim um *oscilador controlável por tensão* (VCO) e completamente integrado.

No entanto, para que se obtenha uma frequência de saída bem definida, face a

variações do processo de fabrico, é necessário aplicar ao VCO um controle por realimentação. A malha de controle de fase contém *divisores de frequência*, um *comparador de fase e frequência* (PFD), um *filtro de realimentação* e um *oscilador de referência* (não representado na fig. 3.1) que utiliza um cristal de quartzo externo ao circuito. Será feita uma apresentação mais detalhada destes blocos nas secções 3.2 a 3.5.

3.1.3 Principais Aplicações

As aplicações mais comuns para este tipo de circuito são a geração de sinais de relógio para microprocessadores e a sintonia de canal em emissores ou receptores de rádio-frequência (RF).

Sistemas digitais avançados precisam de múltiplos sinais de relógio, e sobretudo de sinais de relógio com o mínimo de *jitter* (ruído de fase) e *skew* (diferenças de fase entre diferentes pontos do circuito). Sinais de relógio gerados por PLL e distribuídos utilizando níveis de lógica ECL ou PECL (técnicas essas usadas neste projecto) estão a tornar-se a norma, em tais casos.

A utilização de sintetizadores de frequência para fazer a sintonia de um equipamento de rádio pode parecer óbvia num transmissor, mas nem tanto para um receptor. No entanto, é a técnica mais utilizada, desde a invenção do super-heteródino por Armstrong em 1917 (cf. uma óptima resenha histórica em [1]). A ideia mais óbvia para um receptor de rádio escolher diferentes canais seria utilizando um filtro passa-banda ajustável, ou um banco de filtros. Mas é mais eficaz usar um filtro fixo, depois de converter a frequência original para um valor mais baixo, deslocando-a por um valor ajustável.

Um sintetizador de frequência controlado digitalmente pode também ser utilizado para sistemas de RF que utilizem espalhamento espectral por *frequency hopping*.

No caso deste circuito, a gama de frequências e o espaçamento entre canais foi escolhido de forma a coincidir com os da telefonia móvel de padrão GSM ([13], [30]), embora, para um receptor super-heteródino, as do sintetizador e do canal possam diferir de um valor constante.

O PLL, por si só, também possui outras importantes aplicações para além da síntese de frequência, tais como a extracção do relógio de uma sequência de dados série¹, ou a desmodulação de sinais de frequência modulada (FM).

3.1.4 Tecnologia de Fabrico Utilizada

Existem numerosas tecnologias de fabrico para circuitos integrados. Neste projecto usou-se uma tecnologia BiCMOS (que conjuga transístores <u>bi</u>polares e circuitos <u>CMOS</u>), da companhia *Austria Micro Systems* (AMS). Outras características da tecnologia utilizada são:

- Dimensão característica de 0.8 µm;
- Transístores bipolares de hetero-junção silício-germânio (Si-Ge);
- Biblioteca de componentes caracterizados para alta-frequência, incluindo:
 - Indutores integrados;
 - Capacitores variáveis controlados por tensão;
 - Portas lógicas ECL (Emitter-Coupled Logic);

A biblioteca de componentes foi uma das principais motivações para escolher esta tecnologia para o circuito.

3.2 Oscilador Controlado por Tensão

3.2.1 Introdução

Um oscilador para o sistema em consideração deve possuir as seguintes características:

- Capaz de operar em alta frequência.
- Ter boa pureza espectral.
- Ser facilmente integrável.
- Ter frequência de operação ajustável.

Haveria diversos tipos de oscilador a considerar, e.g. [29]:

Cristal Excelente pureza espectral, mas baixa frequência e muito pouco ajustável.

Tanque LC Boa pureza espectral, frequência limitada por dispositivos activos, suficientemente ajustável.

¹Br.: "dados seriais".

Multi-vibrador Baixa pureza espectral, frequência limitada por dispositivos activos, facilmente ajustável.

Foi escolhido um circuito ressonante de tanque LC, com um indutor integrado. Em [29], por exemplo, a opção foi por um circuito multi-vibrador (oscilador de anel).

3.2.2 Osciladores Ressonantes

Um oscilador ressonante pode ser analisado como um circuito realimentado instável, cujas oscilações crescem em amplitude até que um outro mecanismo, não-linear, estabilize a amplitude das oscilações.

Essa perspectiva dá-nos uma forma de analisar o sistema, por meio de um formalismo típico da análise de sistemas dinâmicos, ilustrado na fig. 3.2.



Figura 3.2 Sistema genérico com realimentação.

A função de transferência do sistema representado é

$$\frac{Y(s)}{X(s)} = \frac{H(s)}{1 - H(s)}$$

Para que a resposta do sistema seja uma oscilação permanente, de amplitude constante, é necessário que a função de transferência tenha pólos puramente imaginários, i.e. $H(j\omega_0) = 1$, o que é equivalente ao bem conhecido critério de Barkhausen, segundo o qual o ganho de malha deverá unitário, e a fase de 180°.

Outra perspectiva que pode ser útil para análise é a de considerar um oscilador como dois circuitos de um porto ligados entre si. O circuito ressonante pode, para uma faixa de frequência suficientemente estreita, ser representado por um circuito equivalente RLC série ou paralelo. As oscilações perdem amplitude devido à energia dissipada no elemento resistivo. O bloco de ganho será equivalente a uma resistência negativa, ou seja, irá fornecer a energia dissipada na resistência do circuito ressonante.

A análise mais básica de circuitos osciladores frequentemente ignora as não-linearidades que determinam a amplitude que será obtida. Um tratamento mais completo, levando em conta também esses efeitos, pode ser encontrada em [1].

Para que o sistema tenha uma frequência de oscilação bem definida, deverá ter uma resposta em frequência do tipo passa-banda, e com uma banda tão estreita quanto possível. Isso normalmente é realizado decompondo o sistema H(s) em dois sub-blocos, um que fornece a selectividade na frequência, e outro que fornece o ganho. Nas implementações que iremos considerar, o primeiro bloco é implementado com um circuito ressonante do tipo tanque LC, e o segundo com um transístor, bipolar (BJT) ou de efeito de campo (FET). A maioria das topologia utilizadas em RF utiliza apenas um transístor como elemento de ganho.

Tradicionalmente isso era feito para reduzir o custo (quando eram utilizados componentes discretos) e para minimizar o ruído introduzido [2]. Para uma implementação monolítica, o custo dos transístores deixa de ser significativo, mas continua a ser importante minimizar o ruído.

Os osciladores ressonantes apresentam a forma básica da fig. 3.3 (onde é usado um BJT, mas poderia usar-se um FET com polarização apropriada). No entanto, normalmente não



Figura 3.3 Forma básica de oscilador ressonante (sem os detalhes de polarização do circuito).

é feita a ligação directa entre o tanque LC e o emissor (fonte) do transístor, pois a impedância vista ao terminal do emissor é baixa $(1 / g_m)$, e iria constituir uma forte carga para o tanque, dificultando o surgimento de oscilações. Por essa razão torna-se necessário efectuar uma transformação de impedância entre o tanque e o emissor do transístor. Existem várias formas de

solucionar este problema:

- Usando um transformador.
- Usando um divisor indutivo (oscilador de Hartley).
- Usando um divisor capacitivo (oscilador de Colpitts) (cf. fig. 3.4).
- Usando um segundo transístor como seguidor de emissor (cf. fig. 3.6).

Para uma implementação integrada, usa-se normalmente uma das duas últimas soluções. Como o transístor seguidor de emissor introduz ruído adicional, o oscilador de Colpitts é uma topologia frequentemente escolhida.



Figura 3.4 Oscilador de Colpitts.

Outra escolha frequente é a de utilizar topologias diferenciais. Este tipo de topologia tem diversas vantagens, entre as quais a de gerar menos ruído e ser menos sensível a ruído. Por este razão, permite também trabalhar com amplitudes de sinal menores e assim consumir menos energia. Todos estes factores são importantes de um modo geral, mas mais ainda em circuitos de alta-frequência.

A topologia finalmente escolhida para o VCO foi uma topologia diferencial, do tipo representado na fig. 3.5, que tem a vantagem de utilizar apenas um indutor.

Note-se que uma topologia deste tipo admite diversas variantes obtidas ligando de diferentes formas os capacitores e indutores, e mantendo um funcionamento similar. Um caso trivial é o da ligação dos capacitores não ao potencial de terra, mas a V_{DD} ou a um outro poten-

cial constante ou de variação muito lenta (como seria por exemplo o terminal de controle de um capacitor variável). Pode até mesmo deixar-se esse potencial flutuante (ficando a série dos dois capacitores em paralelo com o indutor) e, se o circuito for perfeitamente simétrico, esse mesmo potencial há-de manter-se constante. Mas ligar os capacitores ao potencial nulo pode ser vantajoso, porque permite eliminar alguns componentes parasitas.



Figura 3.5 Oscilador ressonante diferencial.

De um modo geral, ao considerar as variantes possíveis, há que ter em conta:

- a facilidade em polarizar
- as tensões absolutas a que são submetidos os componentes
- a possibilidade de eliminar componentes parasitas de forma útil



Figura 3.6 Oscilador ressonante com *buffer* seguidor de emissor, em dois desenhos equivalentes. O funcionamento de uma topologia diferencial pode ser entendido de diferentes

formas. Uma delas consiste em considerar que o par de transístores cruzados apresenta entre os terminais de dreno uma resistência incremental negativa, que irá compensar a resistência parasita do tanque LC que lhe é ligado. Outra forma é considerando esta topologia uma "evolução" da topologia de seguidor de emissor mostrada na fig. 3.6.

Finalmente, o circuito utilizado foi o representado na fig. 3.7.



Figura 3.7 Esquema eléctrico do VCO.

Note-se a substituição das fontes de corrente por espelhos e de parte das capacidades fixas por capacitores variáveis ligados a um potencial de polarização comum, assim como o acréscimo de um capacitor de 300 fF para compensar a diferença de capacidade parasita entre os terminais do indutor escolhido. Também foram acrescentados transístores de *buffer* com saídas em dreno aberto.

3.3 Divisores de Frequência

De um modo geral, pretende-se obter à saída do sintetizador de frequência um valor da forma $f_s = f_0 + k \cdot f_{can}$, onde f_0 é o valor mínimo, k o índice do canal e f_{can} o espaçamento entre canais.

O valor de k será determinado pela entrada digital do sintetizador.

3.3.1 Canais de Frequência do Sistema GSM

De acordo com as especificações da telefonia móvel GSM [13], as bandas atribuídas ao sistema GSM são 890–915MHz para o *downlink* (comunicação do móvel para a base) e 935–960MHz para o *uplink* (comunicação da base para móvel). A banda em cada direcção está dividida em 124 canais, espaçados de 200kHz.

3.3.2 Arquitecturas de PLL e Divisão de Frequência

As arquitecturas de PLL podem também diferir no que diz respeito à forma como é feita a divisão de frequência.

3.3.2.1 Arquitectura de Divisão Inteira

Como já foi referido, a frequência de saída do VCO é dividida por um factor M e comparada com a frequência de referência. O caso mais simples é quando este factor M é um



Figura 3.8 PLL de divisão inteira.

número inteiro, e designa-se então o sistema por PLL de divisão inteira. Esse tipo de arquitectura está representado na fig. 3.8.

Uma consequência é imediata: a frequência de saída será sempre um múltiplo inteiro da frequência de referência, e portanto o espaçamento entre canais será necessariamente igual à frequência de referência. Desta consequência decorre outra:

A frequência de comparação de fase, que influencia a estabilidade e resposta a transientes da malha de controle do PLL, é determinada pelo oscilador de referência. Seria preferível que esta frequência fosse mais alta (intuitivamente, isso equivale a dizer que o controle será tão mais rigoroso quanto maior for a frequência com que é aplicado). Veremos uma alternativa na secção 3.3.3.



Figura 3.9 Divisor de frequência pulse-swallow.

Para fazer síntese de frequência, o divisor terá de ser ajustável. Uma topologia de divisor frequentemente utilizada para isso é a da fig. 3.9, cujo princípio de funcionamento é ilustrado na fig. 3.10.

Observe-se que para cada impulso na saída *s* foram necessários S(N + 1) + (P - S)N = NP + S pulsos na entrada, sendo portanto o factor de divisão de frequência M = NP + S.



Figura 3.10 Princípio de funcionamento do divisor de frequência pulse-swallow.

Apresenta-se na fig. 3.11 um exemplo de dimensionamento de frequência para um PLL de divisão inteira. Os valores dos canais de frequência são os correspondentes às especificações GSM para a banda de 900MHz. O sistema foi redesenhado de forma a evidenciar as duas entradas do bloco PFD.

Se admitirmos que a realimentação mantém os dois sinais de entrada do PFD com frequências iguais, à parte de um erro negligenciável, então a frequência de saída do VCO será, e.g.,

 $(NP + S) 200 \text{ kHz} = (32 \times 146 + S) 200 \text{ kHz} = 934.4 \text{ MHz} + S 0.2 \text{ MHz}$

Para S = 4, $f_{VCO} = 935.2$ MHz, enquanto que para S = 127, $f_{VCO} = 959.8$ MHz. Cobrem-se assim os canais pretendidos dentro da faixa dos 935–960 MHz.



Figura 3.11 Exemplo de dimensionamento de frequências para PLL de divisão inteira.

3.3.3 Arquitectura de Divisão Fraccionária

Se a comparação de frequências for efectuada pelo PFD *antes* dos blocos de divisão por P, como na fig. 3.12, o sistema continuará a funcionar da mesma forma, mas com a vantagem de se ter aumentado a frequência de amostragem da malha de controle e portanto ser possível obter uma melhor resposta dinâmica.



Figura 3.12 Exemplo de dimensionamento de frequências para PLL de divisão fraccionária.

Isto corresponde a transformar o bloco *pulse-swallow* num divisor de frequência com um factor de divisão de fraccionário. Se a saída do bloco *pulse-swallow* for tomada do ponto indicado na fig. 3.12, o número de pulsos à saída passará a ser simplesmente P, pelo que

o factor de divisão de frequência será

$$M = \frac{NP+S}{P} = N + \frac{S}{P} \,.$$

O exemplo anterior pode ser refeito como

$$\left(N + \frac{S}{P}\right)$$
 29.2 MHz = $\left(32 + \frac{S}{146}\right)$ 29.2 MHz = 934.4 MHz + S 0.2 MHz,

com o que se verifica a equivalência entre esta arquitectura e a anterior.

3.3.4 Circuitos Divisores de Frequência

Circuitos divisores de frequência são normalmente circuitos digitais que, por cada N impulsos aplicados à sua entrada, produzem um impulso na saída. Dessa forma, o sinal de saída apresenta pulsos com uma frequência N vezes menor que o sinal de entrada. Assim, um divisor de frequência é basicamente um contador.

Embora a definição anterior esteja essencialmente correcta, é incompleta para a aplicação em vista, que impõe requisitos adicionais:

- O circuito divisor deve ser capaz de operar em alta frequência, uma vez que o seu sinal de entrada será o sinal produzido pelo oscilador.
- O circuito divisor deve introduzir o mínimo de ruído de fase, ou seja, deve perturbar o mínimo possível a fase do sinal de entrada.

O último requisito pode também ser descrito de outra forma: é desejável que o atraso entre transições do sinal, da entrada para a saída, seja pequeno, mas é mais importante ainda que esse atraso seja tão constante quanto possível. Intuitivamente, é de esperar que isso seja mais fácil de obter com circuitos de topologia simples e de poucas portas lógicas, do que com circuito mais complexos.

Os circuitos contadores são circuitos digitais sequenciais, constituídos por flipflops e alguma lógica combinatória adicional. Podem ser classificados como:

Síncronos Quando todos os flip-flops têm a sua entrada de relógio ligada ao sinal de entrada.

Assíncronos Quando os flip-flop estão ligados em cascata, sendo o relógio de cada flip-flop dado pela saída do flip-flop a montante.

A fig. 3.13 é um exemplo de contador síncrono.



Figura 3.13 Exemplo de contador síncrono. Note-se como todos os flip-flops recebem o mesmo sinal nas suas entradas de relógio.

O sub-bloco "div 16" da fig. 3.14 é um exemplo de contador assíncrono (as portas lógicas servem apenas para detecção de fim de contagem).

No caso deste sistema, o sinal de entrada terá uma frequência de cerca de 915 MHz. A maioria das tecnologias CMOS está limitada a frequências de operação de aproximadamente 100 MHz, ou pouco mais.

Isso implica a utilização de lógica ECL, que nesta tecnologia permite, de acordo com o fabricante, atingir uma frequência de operação da ordem dos 2GHz. No entanto, este tipo de lógica tem as desvantagens de ocupar maior área e consumir mais energia que um circuito CMOS de lógica equivalente. Por essa razão, a lógica ECL será utilizada apenas onde estritamente necessária.

Assim, o bloco funcional de divisão de frequência que surge no esquema do PLL é repartido em dois blocos de implementação física:

Pré-divisor Usa lógica ECL. Recebe o sinal do VCO e reduz a frequência o suficiente para o estágio seguinte. Irá corresponder ao divisor 32/33.

Divisor programável Usa lógica CMOS. Irá conter os divisores por S e por P, programáveis.

O pré-divisor utilizado neste projecto tem a topologia do tipo representado na fig. 3.14, que permite comutar entre dois factores de divisão, 32/33. Isso é conseguido com dois sub-blocos, um divisor assíncrono de módulo 16 e um divisor síncrono de módulo 2/3. Quando se pretende uma divisão total por 33, são efectuadas 15 divisões por 2 e uma por 3,



Figura 3.14 Circuito lógico do pré-divisor 32/33.

caso contrário far-se-ão 16 divisões por 2.

É fácil de entender o funcionamento do sub-bloco div2/3. Se MC23 = 1, tem-se simplesmente $x_{n+1} = \overline{z}_n$, ou seja, tem-se o respectivo flip-flop D ligado como um *toggle*, dividindo a frequência por dois, sendo que o outro flip-flop apenas introduz um atraso na saída. Já se MC23 = 0, o funcionamento é mais claramente descrito pela tabela 3.1, de transição de estados.

x _n	y _n	$z_n = x_n y_n$	$x_{n+1} = \overline{z}_n$	$y_{n+1} = x_n$
0	0	0	1	0
1	0	0	1	1
1	1	1	0	1
0	1	0	1	0

Tabela 3.1 Transição de estados do sub-bloco divisor 2/3.

Vê-se que o sub-circuito gera no nó y uma sequência do tipo ...011011011011..., ou seja, para cada ciclo de relógio de entrada gera um pulso à saída, dividindo portanto a frequência de entrada por três. Ao implementar o circuito da fig. 3.14 com as portas ECL da tecnologia utilizada, o circuito que se obtém é o da fig. 3.15. Parece um circuito bastante diferente do original, mas isso deve-se apenas às seguintes razões:

- A lógica ECL utiliza um par de sinais eléctricos diferenciais, para cada sinal lógico.
- Não é necessário ter portas inversoras. A inversão lógica é conseguida trocando a ordem de ligação dos sinais diferenciais.
- Basta usar ou portas AND, ou portas OR, e aplicar as leis de Morgan para obter a porta complementar (observem-se as portas AND e OR do sub-bloco "div 2/3" da fig. 3.15).

3.4 Comparador de Fase

É necessário ainda um elemento que compare a fase do sinal realimentado com o sinal de referência. Isso pode ser feito de várias formas: multiplicadores, portas XOR ou máquinas de estado. As duas primeiras alternativas têm a desvantagem de, para grande diferenças de frequência, a saída ser o batimento das duas frequências, tendo portanto média nula.

O comparador de fase da fig. 3.16, que é basicamente uma máquina de estados implementada com dois flip-flops tipo D, não apresenta esse tipo de inconveniente. As saídas da comparador de fase irão controlar uma bomba de cargas, que mais não é que duas fontes de corrente que irão ora carregar, ora descarregar um condensador, que será o componente principal do filtro de realimentação. Se os sinais em A e B tiverem frequências semelhantes, mas A estiver adiantado em relação a B, haverá uma série regular de pulsos em QA, cuja duração será igual ao avanço de A sobre B. Assim, a tensão de saída irá subir tão mais rapidamente quanto maior for a diferença de fase.

Ou seja, se a referência estiver adiantada em relação ao VCO (ou ao VCO dividido por determinado factor), a tensão aplicada ao VCO irá subir para "acelerá-lo" ligeiramente, até "apanhar" a referência

Já se a frequência de A for muito maior que a de B, haverá uma sequência irregular de pulsos em QA e QB, mas a duração média dos pulsos em A será maior, o que fará também aumentar a tensão de realimentação. Esta resposta é desejável sobretudo para garantir um comportamento correcto do sistema nos instantes iniciais após ter sido ligado.



Figura 3.15 Circuito eléctrico do pré-divisor 32/33, usando lógica ECL.



Figura 3.16 Circuito lógico do comparador de fase e interligação à bomba de cargas.

3.5 Bomba de Carga e Filtro de Realimentação

As duas saídas do comparador da fase da fig. 3.7 accionam o circuito da bomba de carga, ou CP (*charge pump*), que tem uma implementação simples (cf. fig. 3.17). O CP irá car-



Figura 3.17 Bomba de carga.

regar ou descarregar um capacitor ou um filtro RC cuja tensão controla o VCO, fechando-se assim a malha de realimentação.

O filtro de realimentação irá determinar as características de estabilidade, rejeição de erro e resposta dinâmica do sintetizador.

Capítulo 4 Resultados de Simulação

4.1 VCO e Pré-divisor

O funcionamento do VCO está ilustrado na fig. 4.1. A frequência varia entre



Figura 4.1 Saída do VCO para alimentação de 3.3 V, V_{in} a passar de 0 para 3.3 V.

786MHz, para v_{in} nulo, e 1097MHz, para v_{in} = 3.3 V. Observa-se também uma variação na amplitude. As características tensão–frequência e tensão–amplitude são apresentadas na fig. 4.2.

Para demonstrar o funcionamento do pré-divisor iremos simulá-lo conjuntamente com o VCO. O circuito é o da fig. 4.3. A fonte de corrente ligada entre os terminais o1 e o2 do VCO serve apenas para fornecer alguns pulsos de corrente durante os primeiros momentos da simulação, de forma a acelerar o arranque do oscilador. Verificou-se que para uma alimentação de 3.3 V, o consumo do VCO é de 11.3 mA (excluindo *buffers* de saída). Com a alimentação de 5V, o circuito principal do VCO consome 19.5 mA. Mas esta diferença deve-se sobretudo ao facto de o circuito de polarização em corrente ser extremamente simples e não assegurar uma



Figura 4.2 Características V-A e V-f do VCO para alimentação de 3.3 V, V_{in} entre 0 e 3 V. corrente constante face a variações na tensão de alimentação. Já o pré-divisor ECL consome cerca de 30.35 mA, e não muda significativamente com a tensão de alimentação.

A fig. 4.4 mostra o funcionamento do pré-divisor, para MC = 1, nomeadamente o facto haver uma divisão por três a cada 15 divisões por dois, para MC = 1, no sinal de saída do sub-bloco div2/3.



Figura 4.3 Circuito de teste do pré-divisor, com acoplamento AC à saída do VCO.





4.2 Divisores de Frequência Programáveis

A fig. 4.5 contém o resultado de uma simulação digital demonstrando o funcionamento dos divisores programáveis., que se revela semelhante ao descrito na secção 3.1.2.

4.3 Comparador de Fase e Bomba de Carga

É interessante verificar as características do comparador de fase em conjunto com o *charge pump*. Para isso usou-se o circuito da fig. 4.6, onde os componentes VCOquad são modelos comportamentais de VCO, apenas para fornecer estímulos ao circuito em teste.

	155,461.813	,60 , 000	65,000	71,308.045 ns,
Group: A				
test. $P[7:0] = 'dx$	139			
test.S[7:0] = 'd x	65			
test.in = 0				
test.XTAL = 0				
test.top.frac = x				
test.top.ref = x				
test.top.int = x				
Group: swallow				
Group: PC				
test.top. $MC = x$				
test.med_VCO.Npuls = 'd x				
test.med_XTAL.Npuls = 'd x				
test.med_frac.Npuls = 'd x				
test.med_int.Npuls = 'd x	* 11	12	13	*
<pre>test.med_ref.Npuls = 'd x</pre>	11	12	13	14
test.med_VCO.freq = 0				
test.med_XTAL.freq = 0				
test.med_frac.freq = 0				
test.med_int.freq = 0				
test.med_ref.freq = 0				
st.top.Pdiv.Q[7:0] = 'h xx				
est.top.Pdiv.Q[7:0] = 'd x				
st.top.Pcnt.Q[7:0] = 'h xx				
est.top.Pcnt. $Q[7:0] = dx$				
$\underline{\text{st.top.Scnt.Q[7:0]}} = h xx$				
est tob.Scnt.0[7:0] = 'd x				

Figura 4.5 Resultados de simulação do divisor de frequência *pulse-swallow*.



Figura 4.6 Testbench do comparador de fase.





Figura 4.7 Resposta do PFD a diferenças de frequência.

incrementos, em MHz, sobre a frequência de base dos componentes VCOquad, ou seja, a frequência instantânea irá variar entre 15 e 30MHz, em ciclos de 1 µs de duração. Observa-se que o circuito é sensível não só à diferença de fase, mas também à diferença de frequência. Testar o PFD em conjunto com o CP faz com que a carga vista pelo PFD seja realista e permite calcular o consumo mais exactamente.

Verificou-se que, nesta simulação, o PFD consome em média 118µA e o charge

pump consome 178µa. O consumo do PFD é dependente da frequência dos sinais aplicados, e o do CP é aproximadamente proporcional à corrente que se pretende para a carga e descarga do capacitor de saída (neste caso 100µA).

Efectuou-se uma segunda simulação, que modelará melhor a situação em que a saída do PLL está próxima do valor de referência, ou seja, em que as frequências dos dois sinais de entrada são muito próximas, havendo apenas uma diferença de fase.



Figura 4.8 Resposta do PFD a diferenças de fase.

Os resultados são os da fig. 4.8, em que o consumo de corrente foi de $63.5 \,\mu\text{A}$ para o PFD e $165 \,\mu\text{A}$ para o CP.

Capítulo 5 Layout Físico do Circuito

5.1 Introdução

Este capítulo descreve os principais aspectos do trabalho de layout efectuado.

Nas figuras de *layout* apresentadas serão utilizadas legendas para mais facilmente identificar as camadas físicas utilizadas. A fig. 5.1 mostra essa legenda, com o nome abreviado das camadas físicas do processo e a designação completa:

NTUB	Poço n
ZINBUR	Camada profunda n+
DIFF	Difusão
₩ HBT	Camada de heterojunção de transístores bipolares
FIMP	Field implant
CAPLAY	Capacitor de precisão
Secol L	Colector de transístores bipolares
BPOLY	Camada de poli-silício da base de transístores bipolares
∭POLY1	1ª camada de poli-silício, ou poli-1
<u>∭</u> EMITT	Abertura de ligação ao emissor
NPLUS	Implante n+
PPLUS	Implante p+
POLY2	2ª camada de poli-silício, ou poli-2
HRES	Camada de alta resistividade
CONT	Contacto (entre metal-1 e camadas de difusão ou poli-silício)
MET1	1ª camada de metal, ou metal-1
VIA	Interligação entre as duas camadas de metal
MET2	2ª camada de metal, ou metal-2
∭PAD	Abertura na passivação

Figura 5.1 Máscaras correspondentes às camadas físicas da tecnologia de Si-Ge utilizada.

O facto de se dispor de um *design-kit* incluindo uma biblioteca de numerosas células básicas certamente poupou bastante trabalho, mas o processo de *layout* continua ainda assim sendo moroso. Aprender a conhecer e a utilizar apropriadamente os componentes da biblioteca requer o seu tempo. E mesmo depois de adquirido esse conhecimento preliminar, o trabalho de roteamento, que para este tipo de circuito é quase inteiramente manual, é minucioso e demorado. As verificações de regras físicas (DRC — *design rules check*) e, em especial, a verificação da conectividade do *layout* (LVS — *layout versus schematic*), com a subsequente localização e correcção dos erros consomem também bastante tempo, mas são etapas essenciais para garantir a correcção do *layout*.

5.2 Oscilador Controlado por Tensão

A fig. 5.2 apresenta o *layout* do VCO. A topologia escolhida é diferencial, e no *layout* procurou-se manter, na medida do possível, a simetria do circuito.

A assimetria que mais se repercute no desempenho do circuito é a do indutor. O indutor utilizado, da biblioteca da tecnologia, tem um plano de espiras em metal-1 e outro em metal-2, o que introduz um diferença significativa nas capacidades parasitas entre espira e substrato vistas aos terminais do indutor.

Sob os indutores há ainda um padrão gravado na camada profunda, para minimizar as perdas devidas a correntes induzidas no substrato.

Procurou-se usar extensivamente anéis de guarda em torno dos componentes, na maioria dos casos ligados ao potencial de terra, excepto no caso dos transístores de canal n, de polarização, cujos anéis de guarda foram ligados ao potencial de alimentação.

O capacitor variável fornecido na biblioteca tem a forma de uma matriz de capacitores elementares, e a capacidade nominal é ajustável pelo número destes elementos.

Ao experimentar as possibilidades de parametrização dos capacitores fixos da biblioteca, verificou-se que se obtinha uma diminuição significativa da resistência parasita, ao optar por contactos distribuídos sobre toda a superfície da placa superior do capacitor, em vez de contactos apenas ao longo da periferia.

Foram utilizados *pads* de RF para ligar o par de sinais de saída do VCO. Cada *pad* do par possui dois planos condutores, um inferior em metal-1, ligado ao potencial de terra, e outro superior em metal-2, ligado ao sinal de interesse. Isso permite reduzir a resistência parasita para o substrato [31], [32]. Esteticamente, a forma octogonal também quebra um pouco a monotonia...



Figura 5.2 Layout do VCO.

5.3 Divisores de Frequência

Cada célula lógica ECL possui um estágio de saída relativamente poderoso, e é aconselhável ligar o conjunto desses estágios de saída a uma alimentação separada da do "núcleo" lógico da célula. Assim, utilizam-se *pads* separados, vcc1 e vcc2, para cada uma das alimentações.

Pelo facto de a biblioteca desta tecnologia em particular ainda não incluir *pads* de alimentação específicos para a lógica ECL, a alimentação separada foi implementada usando dois pares de *pads* de alimentação analógicos.

Os *pads* de sinal possuem geralmente estruturas protectoras que requerem polarização, e por isso os *pads* são ligados entre si, ao longo da periferia do circuito, de forma a que os *pads* de alimentação polarizem estas estruturas¹. Quando se usam alimentações separadas, há que ter o cuidado de também interromper as interligações entre *pads*, usando células fornecidas para o efeito na biblioteca da tecnologia (designadas *power-cut cells*).

Também foram utilizados *pads* de RF para os sinais de entrada do circuito. Estes *pads* foram associados à alimentação de vcc1, que deverá ser menos ruidosa que vcc2.

Finalmente, os dois *pads* de potencial de terra de cada (vee1) foram ligados internamente. Veja-se o *layout* resultante na fig. 5.3.

Efectivamente, foram utilizadas alimentações separadas para cada um dos principais blocos funcionais do sintetizador de frequência. Isso com o intuito de minimizar as formas de propagação de ruído e maximizar a flexibilidade durante o teste dos protótipos do circuito, podendo assim ter-se blocos funcionais ligados ou desligados de forma selectiva.

Também no caso dos divisores de frequência programáveis, implementados em tecnologia CMOS, a alimentação é isolada da dos demais blocos do circuito. A excepção é o facto de existirem células de saída do bloco pré-divisor que efectuam a conversão de nível entre os dois tipos de lógica, e que precisam de receber os dois tipos de alimentação, ECL e CMOS.

As entradas e saídas de sinal do bloco de divisores programáveis utilizam *pads* digitais de entrada e saída, que para além das estruturas de protecção incorporam também *buf-fers*.

Os circuitos contadores, de oito bits, eram ainda pequenos para justificar a utiliza-

¹Estas estruturas contêm por vezes resistores em série, e quase sempre díodos ligados entre o *pad* de sinal e cada uma das trilhas de alimentação. Os díodos mantêm-se normalmente em polarização reversa e limitam a excursão dos sinais aplicados ao *pad*.





Figura 5.3 Layout do pré-divisor em lógica ECL.

ção de roteadores automáticos de circuitos digitais, mas já grandes o suficiente para justificar certos cuidados de organização no trabalho de *layout* manual.
Devido ao número já relativamente elevado de trilhas de sinal, estabeleceram-se barramentos de sinal. Adoptou-se também o método de definir ligações com trilhas verticais de um metal, e com trilhas horizontais de outro (e.g. metal-1 na horizontal e metal-2 na vertical).

Em cada célula da hierarquia colocaram-se pinos de alimentação e de sinal em metais diferentes. Por exemplo: alimentação com metal-1 na horizontal e sinal com metal-2 na vertical. Assim, as células mais elementares são dispostas em linhas, em cada uma das quais as células partilham o mesmo par de trilhas de alimentação.

É também aconselhável não deixar pinos diferentes de uma célula numa mesma vertical, ao fazer o *layout* de uma dada célula. Dessa forma será fácil, se necessário, deslocar a célula de uma linha para outra linha de células, mantendo as interligações já feitas ao barramento de sinais e sem provocar curto-circuitos acidentais no interior da célula.

O *layout* obtido é apresentado na fig. 5.4. Os oito *pads* à esquerda na figura correspondem a uma das duas palavras de oito bits usadas para programar os divisores. Os *pads* da segunda palavra podem ser vistos no *layout* global (fig. 5.8 ou 5.7).

5.4 Comparador de Fase e Bomba de Carga

Os layouts do comparador de fase e da bomba de carga encontram-se na fig. 5.5.

O par de *pads* Xin e Xout são especiais, na medida em que contêm um oscilador controlado por um cristal de quartzo a ser ligado aos *pads*. Este circuito oscilador é um componente da biblioteca e utiliza uma topologia simples de oscilador ressonante paralelo, em que o principal elemento activo é uma porta inversora CMOS com um resistor de alimentação.

O *layout* do PFD e da bomba de carga, propriamente ditos, também é simples e sem nada de inesperado.

5.5 Layout Global

Eis, finalmente, o *layout* global, apresentado nas fig.s 5.6–5.7 (dupla página 112–113, com detalhes do lado esquerdo e direito do circuito) e fig. 5.8 (vista geral).

O desejo de facilitar e flexibilizar os testes e a utilização de alimentações separadas implicou a utilização de um número relativamente elevado de *pads* (44). Isso impediu uma maior compactação do *layout* global.



Figura 5.4 Layout dos divisores programáveis em lógica CMOS.



Figura 5.5 *Layout* do comparador de fase e da bomba de carga.



Figura 5.6 Layout global do sintetizador de frequência — detalhe do lado esquerdo.



Figura 5.7 *Layout* global do sintetizador de frequência — detalhe do lado direito.



Figura 5.8 *Layout* global do circuito. As dimensões reais são de $2600 \,\mu\text{m} \times 1754 \,\mu\text{m}$.

Capítulo 6 Conclusões

6.1 Sobre a Evolução da Área

Do ponto de vista de mercado, a evolução tem sido de tal modo rápida que mesmo apenas durante o mestrado foi possível observar tendências e produtos ganharem ou perderem destaque. Um dos ganhadores é certamente o padrão IEEE 802 para *Wireless* LANS. Já conta com diversas revisões, e na mais recente a velocidade de transmissão foi quadruplicada. Já o padrão Bluetooth não conquistou nenhuma aplicação de óbvia necessidade nem tampouco ganhou destaque.

A esperada convergência entre PDAs e celulares ainda não aconteceu, e a "moda" das PDAs parece estar a passar, embora não a dos celulares.

No Brasil, a grande novidade da telefonia celular foi a entrada do GSM, na banda de 1800MHz. Para além disso, foi iniciada há pouco tempo, pelo governo brasileiro, o debate sobre a escolha do novo padrão de televisão digital.

Do ponto de vista técnico, as expectativas permanecem elevadas. A área de RF continua a ser vista como "quente", e surgiram também novas técnicas promissora, tais como *transponders* passivos baseados em dispositivos SAW (*surface acoustic wave*) [33].

A familiarização com as ferramentas de projecto exigiu bastante tempo, mesmo sem ter ficado completa e abrangente. Na indústria, a tendência é que os projectistas se especializem num determinado fluxo de projecto: circuitos analógicos, ou circuitos digitais, ou ainda FPGAs (*field programmable gate arrays*). E devido ao elevadíssimo preço das ferramentas de software, a maioria das empresas adquire apenas as estritamente necessárias. Essas são algumas das razões pelas quais os projectistas de circuitos mistos são relativamente raros na indústria, embora necessários. Pelo facto de o CenPRA ser um centro de pesquisa que coordena um serviço de difusão de projectos multi-usuário, dispõe de significativa variedade de ferramentas, permitindo adquirir uma valiosa experiência em projecto de circuitos mistos.

6.2 Sobre este Projecto

O desenvolvimento da biblioteca de simulação foi uma aprendizagem útil, mas seria bom cobrir mais tipos de modulação, incluindo técnicas de espalhamento espectral.

O CI descrito neste texto foi enviado há alguns meses para prototipagem, mas essa primeira rodada foi cancelada por falta de clientes. Isso é claramente uma desvantagem adicional de tecnologias mais caras e "exóticas" como a de SiGe.

Já tinha sido anteriormente afirmado que a motivação para a escolha da tecnologia não é baseada apenas nos méritos intrínsecos da mesma, mas também nas qualidades do *design-kit*. Também no espaço de tempo deste mestrado surgiram outras tecnologias da mesma *foundry*, como a CMOS de 0,35 µm, com modelos para RF tão completos quanto os da tecnologia BiCMOS.

A funcionalidade do CI projectado foi verificada por simulação. Os resultados foram encorajadores, mas requerem ainda uma confirmação por meio de testes eléctricos sobre o protótipo.

6.3 Trabalho Futuro

Após receber os protótipos do circuito, serão realizados testes experimentais para verificar o desempenho efectivamente obtido.

Pretende-se continuar o trabalho na área de microcircuitos de RF com o projecto de outros blocos funcionais, de forma a posteriormente poder-se implementar um sistema transmissor e/ou receptor completo, para transmissão de dados a curta distância e com taxas de transmissão moderadas.

Pretende-se aprofundar o estudo de técnicas de análise de circuitos de RF cuja necessidade ainda não foi plenamente sentida ou explorada no presente projecto. Entre estas técnicas, referem-se e.g. as de análise de ruído, adaptação de impedância e utilização de parâmetros de entrada-saída.

Pretende-se estudar o funcionamento e metodologia de projecto de novos dispositivos como *transponders* passivos SAW, nomeadamente para a criação de sensores sem fios.

Referências

- Thomas H. Lee, "The Design of CMOS Radio-Frequency Integrated Circuits", Cambridge University Press, 2001
- [2] Behzad Razavi, "RF Microelectronics", Prentice Hall, New Jersey, 1998
- [3] Derek K. Shaeffer, Thomas H. Lee, "The Design and Implementation of Low-Power CMOS Radio Receivers", Kluwer, Boston, 2001
- [4] Jacob Jude Rael et al., "Design Methodology Used in a Single-Chip CMOS 900MHz Spread-Spectrum Transceiver", Proceedings of the Design Automation Conference, São Francisco, Junho de 1998 Artigo disponível em www.sigda.org/Archives/ProceedingArchives/Dac/Dac98/papers/ 1998/dac98/pdffiles/03_1.pdf, transparências disponíveis em www.sigda.org/Archives/ ProceedingArchives/Dac/Dac98/papers/1998/dac98/slides/dp03_01.pdf
- [5] Anatel Gerência Geral de Certificação e Engenharia do Espectro da Superintendência de Radiofreqüência e Fiscalização, "Atribuição de Faixas de Freqüências no Brasil (Resolução 79 de 24/12/98", Janeiro de 2000 Disponível no *site* da Anatel, www.anatel.gov.br, na secção "Radiofreqüência"
- [6] Paulo Tavares, "Introdução às Ferramentas Cadence e ao *Design-Kit* AMS", relatório interno do CenPRA — Centro de Pesquisas Renato Archer, Rod. D. Pedro I (SP65) km 143.6, Amarais, 13082-120, Campinas-SP
- [7] Gian Carlo Corazza, "Marconi's History", Proceedings of the IEEE, Julho de 1998, 1307– 1311
- [8] Amitava Dutta-Roy, "Fixed Wireless Routes for Internet Access", IEEE Spectrum, Setembro de 1999, 61–69
- [9] Jaap C. Haartsen, Sven Mattisson, "Bluetooth—A New Low-Power Radio Interface Providing Short-Range Connectivity", Proceedings of the IEEE, Outubro de 2000, vol. 88, N° 10, 165–1661
- [10] D. C. Green, "Radio Systems Technology", Longman Scientific & Technical (co-publicado nos EUA por John Wiley & Sons, 1990, Longman House, Burnt Mill, Harlow, Essex CM20 2JE, Inglaterra
- [11] Paul Sofianos, "Practical Solutions for Medium Data Rate Wireless Communications— Application Note AN1691", Motorola Inc., Abril de 1996 Disponível no site da Motorola, em motorola.com/sps
- [12] Paul Sofianos, "A Full-Featured Wireless Interface for RS-232 Communications—Application Note AN1691", Motorola Inc. disponível no site da Motorola, em motorola.com/sps

- [13] Kaveh Pahlavan, Allen H. Levesque, "Wireless Information Networks", John Wiley & Sons, 1995
- [14] Robert C. Dixon, "Spread Spectrum Systems", Second ed., John Wiley & Sons, 1984
- [15] Albert Franceschino, "An IF Communication Circuit Tutorial—Application Note AN1539/D", Motorola Inc., disponível no site da Motorola, ou em Design-NET.com
- [16] Paul W. Tuinenga, "SPICE: A Guide to Circuit Simulation and Analysis Using PSpice", 3rd ed., Prentice-Hall Inc., Engelwood Cliffs, New Jersey, 1995
- [17] Barrie Gilbert, "Analog at Milepost 2000: A Personal Perspective", Proceedings of the IEEE, Março de 2001, vol. 89, N° 3, 289–304
- [18] Thomas H. Lee, S. Simon Wong, "CMOS RF Integrated Circuits at 5 GHz and Beyond", Proceedings of the IEEE, Outubro de 2000, vol. 88, Nº 10, 1560–1571
- [19]L. E. Larson, "Integrated Circuit Technology Options for RFIC's —Present Status and Future Directions", IEEE Journal of Solid-State Circuits, Março de 1998, vol. 33, 387– 399
- [20] Jim Lansford, Paramvir Bahl, "A Design System for RFIC: Challenges and Solutions", Proceedings of the IEEE, Outubro de 2000, vol. 88, Nº 10, 1613–1632
- [21] Theerachet Soorapanth et al., "Analysis and Optimization of Accumulation-Mode Varactor for RF ICs", Dig. Tech. Papers, 1998 VLSI Circuits Symposium, Junho de 1998
- [22] U.S. Fish and Wildlife Service, "Adopt-A-Trout Radio Telemetry", Setembro de 2001 Texto do "Adopt-A-Trout Website" (www.r6.fws.gov/pfw/MONTANA/adopt2.html)
- [23] Graham Sharples, "Radio Telemetry using RF Modules", Setembro de 2001 disponível em smub.st-e.ac.uk/biomedtelem98/sharples.html
- [24] "LF & RF Communications", Janeiro de 2002 www.austriamicrosystems.com/0automotive/lfrf/index.htm
- [25] Paolo Miliozzi et al., "The Design and Implementation of HomeRF: A Radio Frequency Wireless Networking Standard for the Connected Home", Proceedings of the IEEE, Outubro de 2000, vol. 88, Nº 10, 1662–1676
- [26] Elisabeth A. Bretz, "X Marks the Spot, Maibe", IEEE Spectrum, Abril de 2000, 26-36
- [27] Beth Martin, "Automation Comes to Analog", IEEE Spectrum, Junho de 2001, 70-75
- [28]Brian Kerridge, "High-Speed Bipolar Process Forms Bedrock for Wireless ICs", EDN Magazine, Agosto de 1994 Disponível em www.ednmag.com/ednmag/reg/1994/081894/17df4.htm
- [29] Jacob Jude Rael, "A 915MHz CMOS Frequency Synthesizer", Tese de M.Sc. da University of California (orientada pelo Prof. Asad A. Abidi), Los Angeles, 1995 Disponível em www.icsl.ucla.edu/aagroup/PDF files/jjr-thss.pdf
- [30] Khurram Waheed et al., "A Completely Integrated, Low Noise, Low Power CMOS Frequency Synthesizer for GSM Communications"
- [31] Ahmadreza Rofougaran et al., "A 1 GHz CMOS RF Front-End IC for a Direct-Conversion Wireless Receiver", IEEE Journal of Solid-State Circuits, vol. 31, Julho de 1996, 880–889
- [32] John T. Colvin et al., "Effects of Substrate Resistances on LNA Performance and a Bondpad Structure for Reducing the Effects in a Silicon Bipolar Technology", IEEE Journal of Solid-State Circuits, vol. 34, Setembro de 1999, 1339–1344

- [33]L. Reindl et al., "A Wireless AQP Pressure Sensor Using Chirped SAW Delay Lines Structures", IEEE Ultrasonics Symposium, 1998
- [34] Maurice Bellanger, "Digital Processing of Signals Theory and Practice", John Wiley & Sons, 1984
- [35] Leland B. Jackson, "Digital Filters and Signal Processing", Kluwer Academic Publishers, 1986

Anexo A Simulações de Sistema em Matlab

A.1 Introdução

Para melhor entender os principais tipos de modulação e o seu funcionamento com dados realistas, foi efectuado um conjunto de simulações com o Matlab 4.0 e o Simulink 1.2c.

Dessas simulações resultou um conjunto de blocos que acaba constituindo uma biblioteca de componentes para simulação de sistemas de comunicação, embora a um nível elementar, pois não era esse o objectivo principal deste trabalho de mestrado.

Biblioteca de blocos para comunicações v1.1 compatível com Mattab 4.0, Simulínk 1.2c, por Paulo Tavares, Out.-Nov. de 2001 Geração, conversão e selecção de sinais Modulação Des- modulação Diversos do Simulínk

A.2 Organização da Biblioteca

Figura A.1 O bloco "raiz" da biblioteca.

O Simulink permite construir um modelo de simulação sob a forma de um diagrama de blocos, havendo numerosos componentes disponíveis já à partida. À medida que um diagrama de blocos se vai tornando mais complexo, torna-se necessário organizá-lo de forma a conseguir-se gerir essa complexidade. Uma possibilidade consiste em agrupar vários componentes dentro de um único bloco, o que pode ser repetido agrupando vários blocos e/ou componentes num único bloco, criando-se assim uma descrição hierárquica de um sistema.

É ainda possível associar a um bloco um conjunto de parâmetros, que permitam tratar o conteúdo do bloco como uma "caixa preta", um simples componente de cuja constituição nos podemos abstrair. Na biblioteca que foi formada, os blocos foram organizados em 6 grupos, que podem ser vistos na fig. A.1 ("encapsulamentos") e na fig. A.2 (respectivos conteúdos).



(a) Blocos de geração, conversão e selecção de sinais.





(b) Modulação.

(c) Desmodulação.



(f) Demonstrações.

Figura A.2 Agrupamento dos blocos constituintes da biblioteca.

A biblioteca é constituída por 28 blocos originais e 8 sistemas de demonstração (cf. fig. A.2), construídos com os blocos originais e blocos provenientes do Simulink. Estes blocos de demonstração implementam os sistemas de modulação que se pretendia estudar. Foram ainda incluídos na biblioteca 3 blocos dos mais utilizados no Simulink.

A.3 Implementação do Componentes da Biblioteca

A.3.1 Geração, Conversão e Selecção de Sinais

A.3.1.1 Onda Quadrada

Nada de surpreendente, na implementação ou utilização deste bloco (fig. A.3). O utilizador pode escolher a frequência e os níveis de sinal.



Figura A.3 Gerador de onda quadrada.

A.3.1.2 Sequência de Bits

Este bloco permite, de forma cómoda, especificar uma sequência periódica de bits, através de um vector contendo zeros e uns (fig. A.4). O utilizador pode escolher os níveis de sinal correspondentes. Na implementação, o bloco de função (f(u)) contém a expressão rem(u[1], period), que fornece ao bloco seguinte uma "base de tempo" periódica.



Figura A.4 Gerador de sequência de bits.

A.3.1.3 Bits Aleatórios

Também nada de surpreendente, na implementação ou utilização deste bloco (fig. A.5). O utilizador pode escolher a duração de cada bit e os níveis de sinal.



Figura A.5 Gerador de sequência aleatória de bits.

A.3.1.4 Conversor Série-Paralelo

A implementação (fig. A.6) faz lembrar a utilização de funções de base nas arquitecturas BFSK, BPSK e similares, só que neste caso as funções de base são ondas quadradas sincronizadas com os bits pares e ímpares.

A integração evita *glitches* nas transições. Os níveis lógicos à saída serão os mesmos da entrada, actualizados após cada par de bits de entrada. O primeiro bit do par é a saída 1.



Figura A.6 Conversor série-paralelo.

A.3.1.5 Conversor Paralelo-Série

Também aqui a implementação (fig. A.7) faz lembrar um receptor coerente, e também aqui os níveis lógicos à saída serão os mesmos da entrada.



Figura A.7 Conversor paralelo-série.

A.3.1.6 Selectores 2 para 1 e 3 para 1

Nos blocos de multiplexação (fig. A.8 e fig. A.9) a última entrada permite escolher qual da outras entradas é colocada na saída.



Figura A.8 Selector 2 para 1.



Figura A.9 Selector 3 para 1.

A.3.2 Moduladores

A.3.2.1 Amplitude Modulada

Os blocos (fig. A.10) simplesmente calculam $A_c [1 + m u(t)] \sin \omega_c$. A segunda saída é a portadora, sem modulação.



Figura A.10 Modulador AM.

A.3.2.2 Frequência Modulada

Neste caso (fig. A.11) os cálculos são mais complexos, por isso recorreu-se a um bloco de função.



Figura A.11 Modulador FM.

A.3.2.3 BFSK

A implementação é óbvia (fig. A.12). O bloco Look Up Table serve apenas para adaptar os valores do zero e um lógico ao limiar de decisão.



Figura A.12 Modulador BFSK.

A.3.2.4 BPSK

(fig. A.13) Implementação semelhante à do bloco BFSK.



Figura A.13 Modulador BPSK.

A.3.2.5 QPSK

(fig. A.14) O par de blocos à saída do conversor S/P serve apenas para converter o nível do zero e um lógico para os valores -1 e 1.

Este bloco (e os demais da família QPSK) utiliza um conversor S/P, pelo que apresenta um atraso de 2 T_b , da entrada para a saída. Na minha implementação dos moduladores desta família, fiz a saída igual à soma das duas componentes, e não igual à diferença, ao contrário de alguns autores.



Figura A.14 Modulador QPSK.

A.3.2.6 OQPSK

(fig. A.15) A implementação é óbvia...



Figura A.15 Modulador OQPSK.

A.3.2.7 π/4-QPSK

(fig. A.16) A implementação também é óbvia...



Figura A.16 Modulador $\pi/4$ -QPSK.

A.3.2.8 MSK

(fig. A.17) Neste caso colocou-se o bloco de adaptação de nível antes do conversor S/P, pois o conversor coloca à saída sinais com níveis idênticos ao da entrada.



Figura A.17 Modulador MSK.

A.3.3 Desmoduladores

A.3.3.1 Amplitude Modulada—Desmodulador Coerente

(fig. A.18) O bloco utiliza um filtro de Chebychev cujos parâmetros são especificados pelo utilizador do bloco.

A.3.3.2 Amplitude Modulada—Detector de Pico

(fig. A.19) Este bloco reproduz a dinâmica de um díodo p-n e de um filtro RC. A



Figura A.18 Desmodulador AM coerente.

característica do díodo utilizada é $Is(e^{u[1]/(\eta Vt)} - 1)$, em que Is = 1 mA, $V_T = 25 \text{ mV}$, $\eta = 2$, sendo a corrente limitada a 10A, para evitar problemas de *overflow* nos cálculos. O valor de resistência usado é de $10k\Omega$ e a capacidade é calculada de forma a que a descarga do filtro RC tenha a rapidez especificada nos parâmetros do bloco.



Figura A.19 Desmodulador AM por detecção de pico.

A.3.3.3 Frequência Modulada

(fig. A.20) A desmodulação é feita calculando a derivada numérica do sinal de entrada, de forma a que a amplitude varie com a frequência, e depois efectuando uma detecção de pico, com o respectivo bloco de desmodulação AM.



Figura A.20 Desmodulador FM.

A.3.3.4 BFSK—Desmodulador Coerente

(fig. A.21) Implementação sem surpresas...

A.3.3.5 BFSK—Desmodulador Não-Coerente

(fig. A.22) São utilizados dois filtros passa-banda em torno das duas frequências usadas, e a largura de banda destes filtros deve ser especificada pelo utilizador.

A frequência de corte dos detectores de pico é 1/50 da menor frequência usada. Os blocos de atraso de $T_b/2$ (T_b é a duração de cada bit) fazem com que a decisão sobre o valor do bit de saída seja feita a meio do bit anterior.



Figura A.21 Desmodulador BFSK coerente



Figura A.22 Desmodulador BFSK não-coerente.

A.3.3.6 BPSK

(fig. A.23) Implementação previsível...



Figura A.23 Desmodulador BPSK.

A.3.3.7 QPSK

(fig. A.24) O atraso total da entrada para as saídas é de 3 T_b (2 T_b dos integradores e T_b de um conversor P/S).

A.3.3.8 OQPSK

(fig. A.25) Naturalmente, o atraso de T_b , no desmodulador, ficou no ramo oposto ao que se encontrava no modulador.



Figura A.24 Desmodulador QPSK.

O atraso total da entrada do modulador para as saídas do desmodulador é de 7 T_b .



Figura A.25 Desmodulador OQPSK.

A.3.3.9 π/4-QPSK

(fig. A.26) O atraso total da entrada para as saídas é de 3 T_b (2 T_b dos integradores e T_b de um conversor P/S).

A.3.3.10 MSK

(fig. A.27) O atraso total da entrada do modulador para as saídas do desmodulador é de 7 T_b .

A.3.4 Diversos

Grupo de blocos auxiliares, usados na construção dos restantes blocos.

A.3.4.1 Canal com AWGN (Ruído Aditivo Gaussiano Branco)

(fig. A.28) Para a simulação poder ser mais rápida, o ruído passa por um bloco de amostragem. Isso fá-lo ficar um pouco "menos branco", mas desde que a frequência de amostragem do ruído seja suficientemente maior que a dinâmica mais rápida do sistema, o ruído será "branco o suficiente".



Figura A.26 Desmodulador $\pi/4$ -QPSK.



Figura A.27 Desmodulador MSK.

Neste bloco é também aplicado um factor de ganho, calculado em função da variância pretendida para o ruído.



Figura A.28 Canal com AWGN (ruído aditivo gaussiano branco).

A.3.4.2 Integral com Reset Periódico

(fig. A.29) O bloco de função f (u) implementa a expressão rem (u[1]+offset, T), que permite gerar uma onda triangular com o período período e *offset* desejados. O bloco Memory e o bloco de soma permitem detectar o instante exacto da transição na onda triangular.



Figura A.29 Integral com reset periódico.

A.3.4.3 Cálculo das Componentes em Fase e Quadratura

(fig. A.30) A frequência deste cálculo pode ser ajustada de forma independente.

Valores razoáveis para o período de integração irão variar entre T_b (a duração de um bit) e $1/f_C$ (um período da portadora). No caso de modulações do tipo MSK, por exemplo, é interessante fazê-la várias vezes maior que a taxa de transmissão de bits, de forma a ver a variação contínua das componentes I e Q.



Figura A.30 Cálculo das componentes em fase e quadratura.

A.3.5 Diversos do Simulink

Este bloco não tem componentes originais, apenas agrupa alguns componentes do Simulink que são necessários com maior frequência, pelo que não serão aqui descritos.

A.4 Demonstração da Biblioteca

Nos diagramas seguintes, os blocos que terminam um sinal são normalmente blocos que indicam o nome de uma variável onde a sequência de valores temporais do sinal será armazenada.

A.4.1 Amplitude Modulada

O sistema da fig. A.31 testa simultaneamente os dois tipos de desmodulador implementados.



Figura A.31 Demonstrações—Modem AM.

Começou-se por testar o sistema com uma entrada sinusoidal, sem ruído no canal. Eis os programas utilizados, para inicializar a simulação...

```
1
      % entrada sinusoidal sem ruído
2
 3
      % Parâmetros do sinal de entrada
4
      i sel = 1;% sinal: 1=sinusóide; 2=onda quadrada; 3=bits aleatórios
 5
      fx = 15;% frequência da sinusóide ou onda quadrada
 6
     bps = 2*fx; % bit-rate é dobro da freq. de onda quad. corresp.
 7
      seed = 1;% inicialização do gerador de bits aleatórios
8
      xmin=-1;xmax=1;% níveis do "0" e "1" lógico
9
10
      % Parâmetros do pré-filtro
11
      fi sel = 1;% 1=nenhum; 2=Butterworth; 3=Chebyshev
12
      fc but i=100; ord but i=2;% butterworth de entrada
13
      fc che i=500; ord che i=6; aten che i=40;% chebyshev de entrada
14
```

```
15
      \% Parâmetros do modulador: x AM = Ac ( 1 + m x BB ) cos w c t
16
     Ac = 5;% Amplitude da portadora
17
     fc = 1e3;% frequência da portadora
18
     m = .5;% indice de modulação
19
20
     % Parâmetros do canal
21
     Nvar=.0;% variância do ruído
22
      %Nvar=.5;% variância do ruído
23
24
      % Parâmetros do desmodulador por detecção de pico
25
      fcut = 20;% freq. de corte do filtro RC
26
27
      % Parâmetros do desmodulador coerente
28
      fc che o=500; ord che o=6; aten che o=40;% filtro de chebyshev
29
30
      % Parâmetros do simulador
31
     tf=0.12; minstep=1e-4/fc; maxstep=5e-2/fc; tol=1e-3;
32
33
     % Parâmetros do pós-processamento
34
     N = 8192;% número de pontos de simulação armazenados
```

e para efectuar o pós-processamento...

```
1
      % entrada sinusoidal sem ruído
 2
 3
      mplot([1 1 2]',t,[x xrx xdem1 xdem2],...
 4
      ['sinal de banda base,'...
 5
       `sinal modulado,'...
 6
       `detector de pico,detector coerente']);
 7
 8
     bwlines;
 9
     wysiwyg;
10
     print -deps dsbamres1
```

Na fig. A.32 apresentam-se os resultados da simulação. Ambos os receptores apresentam um bom desempenho.

Na simulação seguinte a entrada é uma onda quadrada, que passou por um estágio de pré-filtragem (fig. A.38). Esse estágio tem por finalidade limitar a largura de banda do sinal a modular (e, por consequência, do sinal modulado). Admitiu-se que o canal era afectado por ruído branco gaussiano com variância de 0,5.

O código de inicialização sofreu as seguintes alterações, nas linhas 4, 11 e 21, respectivamente:



Figura A.32 Resultados-Modem AM.

```
i_sel = 2;% sinal: 1=sinusóide; 2=onda quadrada; 3=bits
aleatórios\lineno{11}
fi_sel = 2;% 1=nenhum; 2=Butterworth; 3=Chebyshev\lineno{21}
Nvar=.5;% variância do ruído
```

O código de pós-processamento sofreu alterações mais numerosas:

```
1
      % entrada sinusoidal sem ruído
 2
 3
      h=mplot([2 1 2]',t,[x xbb xrx xdem1 xdem2],...
 4
      ['sinal de banda base, sinal filtrado,'...
 5
       'sinal recebido (com ruído), '...
 6
       `detector de pico,detector coerente']);
 7
      subplot(3,1,2);set(gca,'ylim',[-15 15]);
 8
 9
      bwlines;
10
      subplot(3,1,3);hold on;
11
      set(plot(t,xrx),'color',.8*[1 1 1],'linestyle','-');
12
13
      wysiwyg;
14
      print -depsc dsbamres2
```

Os resultados encontram-se na fig. A.33. O receptor coerente apresentou um melhor desempenho.



Figura A.33 Resultados—Modem AM.

A.4.2 Frequência Modulada

Para testar o sistema de FM foram utilizados sinais e condições semelhantes às do caso anterior: primeiro uma entrada sinusoidal, sem ruído a afectar o sistema, e depois uma onda quadrada, na presença de ruído com variância de 0,1.



Figura A.34 Demonstrações—Modem FM.

Eis as listagens do primeiro caso, com o código de inicialização,

```
1 % arquivo: fm_i.mautor: paulo.tavares@ieee.org, nov 2001
2 % Inicialização dos parâmetros de simulação para FM.M (v1.0)
3
```

```
4
      % Parâmetros do sinal de entrada
 5
     i sel = 1;% sinal: 1=sinusóide; 2=onda quadrada; 3=bits aleatórios
6
     fx = 15;% frequência da sinusóide ou onda quadrada
7
     bps = 2*fx; % bit-rate é dobro da freq. de onda quad. corresp.
8
     xmin=-1;xmax=1;% níveis do "0" e "1" lógico
9
10
     % Parâmetros do pré-filtro
11
     fi sel = 1;% 1=nenhum; 2=Butterworth; 3=Chebyshev
12
     fc_but_i=100; ord_but_i=2;% butterworth de entrada
13
     fc_che_i=500; ord_che_i=6; aten_che_i=40;% chebyshev de entrada
14
15
     % Parâmetros do modulador:
16
     Ac = 5;% Amplitude da portadora
17
     fc = 1e3;% frequência da portadora
18
     df = fc/2; % desvio máximo de frequência
19
     m = 2*pi*df/max(abs([xmin xmax]));% indice de modulação
20
21
     % Parâmetros do canal
22
     Nvar = 0.0;% variância do ruído
23
     % (seria necessário uma saturação de amplitude)
24
25
     % Parâmetros do desmodulador por detecção de pico
26
     fcut = 20;% freq. de corte do filtro RC
27
28
     % Parâmetros do simulador
29
     tf=0.12;
30
     minstep=le-4/(fc+df); maxstep=5e-2/(fc+df); tol=le-3;
31
32
     % Parâmetros do pós-processamento
33
     N = 20e3; % número de pontos de simulação armazenados
```

e de pós-processamento:

```
1
      % entrada sinusoidal sem ruído
 2
 3
     mplot([1 1 1]',t,[x xrx xdem],...
 4
      ['sinal de banda base,'...
 5
      'sinal recebido (sem ruído), '...
 6
      `sinal desmodulado']);
 7
 8
     bwlines;
 9
     wysiwyg;
10
     print -depsc fmres1
```

Os resultados são os da fig. A.35.

Eis as listagens do segundo caso. Uma vez mais, o código de inicialização sofreu





poucas alterações, nas linhas 5, 11 e 22.

```
i_sel = 2;% sinal: 1=sinusóide; 2=onda quadrada; 3=bits
aleatórios\lineno{11}
fi_sel = 2;% 1=nenhum; 2=Butterworth; 3=Chebyshev\lineno{22}
Nvar = 0.1;% variância do ruído
```

e o código de pós-processamento sofreu alterações mais extensas:

```
% entrada sinusoidal sem ruído
 1
 2
 3
      mplot([2 1 1]',t,[x xbb xrx xdem],...
 4
      ['sinal de banda base, sinal filtrado,'...
 5
       'sinal recebido (com ruído),'...
 6
       `sinal desmodulado']);
 7
      subplot(3,1,2), set(gca, 'ylim', [-7 7])
 8
 9
      bwlines;
10
      wysiwyg;
11
      print -depsc fmres2
```

Os resultados são apresentados na fig. A.36 e fig. A.37

O receptor FM parece ser bastante sensível ao ruído, mas isso deve-se ao tipo de implementação utilizado, que se baseia num derivador para converter a variação de frequência numa variação de amplitude. Para agravar a situação, o bloco derivador não foi precedido nem de um bloco de limitação de amplitude, nem de um filtro passa-banda.

A.4.3 BFSK

A primeira simulação pretende apenas ilustrar o princípio de funcionamento do sis-



Figura A.37 Resultados com ruído-Modem FM.

Eis os arquivos de programa utilizados:

tema, que é bastante simples.



Figura A.38 Demonstrações—Geradores de sinal e pré-filtros dos modems AM e FM.



Figura A.39 Demonstrações—Modem BFSK.

```
1
      % arquivo: bfsk i.mautor: paulo.tavares@ieee.org, nov 2001
 2
      % Inicialização dos parâmetros de simulação para BFSK
 3
 4
      % Parâmetros do sinal de entrada
 5
      xmin=0;% xmin->f1
 6
      xmax=1;% xmax->f2
 7
      bps = 40;% bit-rate
 8
      Tb=1/bps;
 9
      i sel = 2;% 1=bits aleatórios; 2=sequência de bits
10
      seq = [1 0];% sequência = onda quadrada
11
12
      % Parâmetros da modulação:
      f1=1e3;f2=2e3;% frequências a utilizar
13
14
      Ac = 5;% Amplitude da portadora
15
16
      % Parâmetros do canal
17
      Nvar=0;% variância do ruído
18
19
      % Parâmetros do simulador
20
      tf=3.5*Tb; minstep=1e-3/max(f1,f2); maxstep=.1/max(f1,f2); tol=1e-3;
```

```
21
22
      % Parâmetros do pós-processamento
23
      N = 8192;% número de pontos de simulação armazenados
 1
      Tb=1/bps;
 2
      mplot(t/Tb,[x xrx xdemc],[...
 3
      'sinal de banda base,'...
 4
      'sinal recebido (sem ruído),'...
 5
      `detecção coerente']);
 6
 7
      bwlines;
 8
 9
      wysiwyg;
10
      print -deps bfskres1
```

Nos resultados (fig. A.40), a escala de tempo está graduada em duração de bit, e observa-se que o sinal desmodulação tem um atraso de um bit em relação ao sinal transmitido, mesmo ignorando atrasos de propagação no canal, que não foram incluídos nos modelos apresentados. Durante a duração do primeiro bit, a saída do receptor encontra-se num nível indefinido.



Figura A.40 Resultados-Modem BFSK.

Numa segunda simulação foi introduzido algum ruído:

```
1
      % arquivo: bfsk i.mautor: paulo.tavares@ieee.org, nov 2001
 2
      % Inicialização dos parâmetros de simulação para BFSK
 3
 4
      % Parâmetros do sinal de entrada
 5
      xmin=0;% xmin->f1
 6
      xmax=1;% xmax->f2
 7
      bps = 40;% bit-rate
 8
      Tb=1/bps;
 9
      i sel = 2;% 1=bits aleatórios; 2=sequência de bits
10
      seq = [1 0];% sequência = onda quadrada
11
12
      % Parâmetros da modulação:
13
      f1=1e3;f2=2e3;% frequências a utilizar
14
      Ac = 5; % Amplitude da portadora
15
16
      % Parâmetros do canal
17
      Nvar=395;% variância do ruído
18
19
      % Parâmetros do simulador
20
      tf=3.5*Tb; minstep=1e-3/max(f1,f2); maxstep=.1/max(f1,f2); tol=1e-3;
21
22
      % Parâmetros do pós-processamento
23
      N = 8192;% número de pontos de simulação armazenados
```

Eis o código de pós-processamento:

```
1
      Tb=1/bps;
 2
      mplot(t/Tb,[xrx xdemc xdemnc],[...
 3
      'sinal recebido (ruído com variância=395),'...
 4
      'detecção coerente,'...
 5
      'detecção não-coerente']);
 6
      subplot(3,1,1), set(gca, 'ylim', 50*[-1 1])
 7
 8
      bwlines;
 9
10
      wysiwyg;
11
      print -deps bfskres2
```

O sistema evidenciou robustez face ao ruído, uma vez que os primeiros erros de recepção começam a ocorrer para ruído com variância de cerca de 400 (cf. fig. A.41 e fig. A.42). As modificações ao código de inicialização e pós-processamento foram triviais.

A robustez face ao ruído deve-se ao facto de o receptor coerente integrar o sinal recebido, o que atenua o ruído. Também o receptor não-coerente parece mais robusto que o receptor FM anteriormente visto, pois o sinal de entrada sofre uma filtragem passa-banda.



Figura A.41 Resultados-Modem BFSK.

Observa-se que o receptor coerente é o primeiro a falhar (embora para tirar conclusões mais generalizáveis a simulação devesse ter sido mais prolongada).





A.4.4 BPSK



Figura A.43 Demonstrações—Modem BPSK.

Uma primeira simulação serve uma vez mais para ilustrar o princípio de funcionamento. Eis o respectivo código de inicialização,

```
1
      % arquivo: psk_i.mautor: paulo.tavares@ieee.org, nov 2001
 2
      % Inicialização dos parâmetros de simulação para a família de modems
 3
      % digitais por modulação de fase (BPSK, QPSK, OQPSK, pi/4-QPSK e MSK)
 4
 5
 6
      % Parâmetros do sinal de entrada
 7
      xmin=0;% xmin->phi1
 8
      xmax=1;% xmax->phi2
 9
      bps = 40;% bit-rate
10
      Tb = 1/bps;
11
      i sel = 2;% 1=bits aleatórios; 2=sequência de bits
12
      seq = [1 0];% sequência = onda quadrada
13
14
      % Parâmetros da modulação:
15
      fc = 1e3;% frequência da portadora
16
      Ac = 5;% Amplitude da portadora
17
18
      %Parâmetros do canal
19
      Nvar=0;% variância do ruído
20
21
      % Parâmetros do simulador
22
      tf=3.5/bps;% tf é modificado em cada sistema
23
      minstep=le-3/fc; maxstep=.1/fc; tol=le-3;
24
25
      % Parâmetros do pós-processamento
26
      N = 8192;% número de pontos de simulação armazenados
```

e de pós-processamento

```
1
     mplot(t/Tb,[x xrx xdem],[...
2
     'sinal de banda base,'...
3
     'sinal recebido (sem ruído), '...
4
     `sinal desmodulado']);
5
6
     bwlines;
7
8
     wysiwyg;
9
     print -deps bpskres1
```





Figura A.44 Resultados-Modem BPSK.

Os resultados tornam-se mais interessantes ao aplicar uma sequência de entrada mais prolongada e introduzir ruído no sistema (fig. A.45). O código de inicialização sofreu poucas modificações em relação ao caso anterior (apenas nas linhas 11, 19 e 22).

```
i_sel = 1;% 1=bits aleatórios; 2=sequência de bits\lineno{19}
Nvar=350;% variância do ruído\lineno{22}
tf=18.5/bps;% tf é modificado em cada sistema
```

O código de pós-processamento sofreu apenas alterações triviais.


Figura A.45 Resultados-Modem BPSK.

A.4.5 QPSK





Figura A.46 Demonstrações—Modem QPSK.

Vamos começar por analisar a relação entre o sinal de entrada, as componentes de fase e quadratura e o sinal desmodulado, sem considerar os efeitos do ruído. Isso é feito com o seguinte código:

(inicialização:)

```
1
      % arquivo: psk i.mautor: paulo.tavares@ieee.org, nov 2001
2
     % Inicialização dos parâmetros de simulação para a família de modems
3
     % digitais por modulação de fase (BPSK, QPSK, OQPSK, pi/4-QPSK e MSK)
4
5
6
     % Parâmetros do sinal de entrada
7
     xmin=0;% xmin->phi1
8
     xmax=1;% xmax->phi2
9
     bps = 40;% bit-rate
10
     i sel = 1;% 1=bits aleatórios; 2=sequência de bits
11
     seq = [1 0];% sequência = onda quadrada
12
     bitseed=7;% semente da sequência pseudo-aleatória
13
14
     % Parâmetros da modulação:
15
     fc = 1e3;% frequência da portadora
16
     Ac = 5; % Amplitude da portadora
17
18
     %Parâmetros do canal
19
     Nvar=0;% variância do ruído
20
21
     % Parâmetros do simulador
22
     tf=1;% tf é modificado em cada sistema
23
     minstep=le-3/fc; maxstep=.1/fc; tol=le-3;
24
25
     % Parâmetros do pós-processamento
26
     N = 8192;% número de pontos de simulação armazenados
        (pós-processamento:)
 1
     Tb=1/bps;
2
     set(0,'defaultaxesfontsize',10);
 3
4
     subplot(4,1,1),plot(t/Tb,x), ,title('sinal de banda base')
 5
     set(gca,'box','off','xtick',0:2:30,'xgrid','on')
6
     set(get(gca,'title'),'fontsize',10);
7
8
      subplot(4,1,2),plot(t/Tb,iq(:,1)),title('componente em fase')
9
     set(gca,'box','off','xlim',[0 35]+3,'xtick',(0:2:30)+3,'xgrid','on')
10
     set(get(gca,'title'),'fontsize',10);
11
12
     subplot(4,1,3),plot(t/Tb,iq(:,2)),title('componente em quadratura')
13
     set(gca,'box','off','xlim',[0 35]+3,'xtick',(0:2:30)+3,'xgrid','on')
14
      set(get(gca,'title'),'fontsize',10);
15
16
     subplot(4,1,4),plot(t/Tb,xdem), title('sinal desmodulado')
17
      set(gca,'box','off','xlim',[0 35]+5,'xtick',(0:2:30)+5,'xgrid','on')
```

```
18 set(get(gca,'title'),'fontsize',10);
```

```
19 % em condições normais o MPLOT faria o trabalho das 17 linhas
20 % anteriores, mas neste caso quis ter eixos diferentes em cada gráfico
21
22 bwlines;
23
24 wysiwyg;
25 print -deps qpskres1
```



Figura A.47 Resultados-Modem QPSK.

Neste tipo de modulação, os bits de entrada são divididos em pares. Aos bits ímpares irá corresponder uma das componentes, aos bits ímpares a outra componente. As componentes de fase e quadratura são obtidas com um atraso de 3 bits em relação ao sinal de entrada, e o sinal desmodulado com um atraso de 7 bits, mas para facilitar a comparação, os eixos das abcissas estão alinhados de acordo com os respectivos atrasos.

A relação entre as componentes de fase e quadratura e o sinal modulado só se torna clara se o sinal modulado for observado suficientemente "de perto". Para isso, o código de inicialização sofreu pequenas modificações (linhas 10, 11 e 22):

```
i_sel = 2;% 1=bits aleatórios; 2=sequência de bits\lineno{11}
seq = [1 1 0 1 1 0 0 0];\lineno{22}
tf=12.5/bps;
```

Eis o código de pré-processamento

```
1
      Tb=1/bps;
 2
      set(0,'defaultaxesfontsize',10);
 3
 4
      subplot(3,1,1),plot(t/Tb,xrx) ,title('sinal modulado')
 5
      set(gca,'box','off','xlim',[0 13],'xtick',0:2:13,'xgrid','on')
 6
      set(get(gca,'title'),'fontsize',10);
 7
 8
      subplot(3,1,2),plot(t/Tb,iq(:,1)),title('componente em fase')
 9
      set(gca,'box','off','xlim',[0 13],'xtick',(0:2:13),'xgrid','on')
10
      set(get(gca,'title'),'fontsize',10);
11
12
      subplot(3,1,3),plot(t/Tb,iq(:,2)),title('componente em quadratura')
13
      set(gca,'box','off','xlim',[0 13],'xtick',(0:2:13),'xgrid','on')
14
      set(get(gca,'title'),'fontsize',10);
15
16
      subplot(3,1,1)
17
      pos=get(gca,'pos');
18
      delete(gca);
19
      dx=.05*pos(3);% separação entre cada par de eixos
20
      width=(pos(3) - 3*dx)/4;
21
      h=[];
22
      for i=1:4;
23
         left=pos(1)+(i-1)*(dx+width);
24
         axes('position',[left pos(2) width pos(4)]);
25
         plot(t/Tb,xrx); title('sinal modulado');
26
         set(get(gca,'title'),'fontsize',9);
27
         set(gca,'xlim',i*2+[-.1 .1],'fontsize',9);
28
         if i~=1, set(gca,'yticklabels',''); end
29
      end
30
31
      bwlines;
32
33
      wysiwyg;
34
      print -deps qpskres2
```

O resultado é o da fig. A.48...

Ao observar a figura, há que não esquecer o atraso entre a transição no sinal modulado e a transição nas componentes $I \in Q$.

A fig. A.49 contém bastante informação. Mostra o sinal modulado afectado por um ruído de variância 100, e mostra como apesar disso as componentes I e Q são recuperadas sem grande erro (o sinal de saída não conteria erro absolutamente nenhum). Mostra ainda, no canto



Figura A.48 Resultados—Modem QPSK.



Figura A.49 Resultados—Modem QPSK.

superior direito, as componentes I-Q representadas sob a forma geralmente designada por constelação do sinal.

Do código utilizado, o de inicialização tem apenas uma alteração relativamente à segunda simulação:

```
Nvar=100;% variância do ruído
```

Eis o código de pós-processamento:

```
1
      Tb=1/bps;
 2
      set(0,'defaultaxesfontsize',10);
 3
 4
      subplot(3,1,2),plot(t/Tb,iq(:,1)),ylabel('fase');
 5
      set(gca,'box','off','xlim',[0 13],'xtick',(0:2:13),'xgrid','on')
 6
      set(get(gca,'title'),'fontsize',10);
 7
 8
      subplot(3,1,3),plot(t/Tb,iq(:,2)),ylabel('quadratura');
 9
      set(gca,'box','off','xlim',[0 13],'xtick',(0:2:13),'xgrid','on')
10
      set(get(gca,'title'),'fontsize',10);
11
12
      subplot(3,1,1)
13
      pos=get(gca,'pos');
14
      delete(gca);
15
      dx=.05*pos(3);% separação entre cada par de eixos
16
      width=(pos(3)-3*dx)/4;
17
     h=[];
18
      for i=1:4;
19
         left=pos(1)+(i-1)*(dx+width);
20
         axes('position',[left pos(2) width pos(4)]);
21
         plot(t/Tb,xrx); title('sinal modulado');
22
         set(get(gca,'title'),'fontsize',8);
23
         set(gca,'xlim',i*2+[-.1 .1],'fontsize',8);
24
         if i~=1, set(gca,'yticklabels',''); end
25
      end
26
27
      plot(iq(:,1),iq(:,2),'+'); title('constelação')
28
      v=version;
29
      if v(1) =='4'
30
         xform=eye(4); xform(1,1)=-1;
31
         set(gca,'xform',xform,'xdir','reverse')
32
      else
33
         set(gca,'yaxislocation','right')
34
      end
35
36
      wysiwyg;
37
      print -deps qpskres3
```

Anexo B Transferência de Potência

Consideremos um circuito constituído por uma fonte de tensão V, com uma resistência interna R_i , alimentando uma carga resistiva R. Vamos designar por P_V a potência fornecida pela fonte e por P_i e P_R a potência dissipada na resistência interna e na resistência de carga, respectivamente. A questão a que normalmente interessa responder é: como maximizar a potência transferida para a carga?

O valor de potência recebida pela carga é $P_R = \frac{V_L^2}{R}$, em que $V_L = \frac{R}{R_i + R}V$. Portanto,

$$P_R = \frac{R}{(R_i + R)^2} V^2$$

donde resulta que

$$\frac{\partial P_R}{\partial R} = \frac{(R_i + R)^2 - 2R(R_i + R)}{(R_i + R)^4} V^2 = \frac{R_i - R}{(R_i + R)^3} V^2,$$

$$\frac{\partial P_R}{\partial R_i} = -\frac{2R(R_i + R)}{(R_i + R)^4}V^2 = -\frac{2R}{(R_i + R)^3}V^2$$

Assim, conclui-se que para um valor de R_i fixo, o valor de R que recebe mais energia da fonte será $R = R_i$, e nesse caso $P_R = P_i = \frac{V^2}{4R_i}$, enquanto que $P_V = P_i + P_R = \frac{V^2}{2R_i}$.

Já se considerarmos que é R que tem um valor fixo, o máximo de P_R ocorre para R_i nulo, e é dado simplesmente por $P_R = P_V = \frac{V^2}{R}$.

É frequente encontrar textos onde não se distingue claramente entre $P_R e P_V e$ entre a possibilidade de variações no valor de *R* ou de R_i .

Anexo C Anotações de Teoria dos Sinais

C.1 Análise de Fourier

Uma forma de analisar sinais complexos consiste em decompô-los numa soma de sinais mais simples, que tenham propriedades matemáticas interessantes. A esses sinais "elementares" chama-se funções de base.

No caso da análise de Fourier, as funções de base são senos e cossenos ou, equivalentemente, exponenciais complexas:

$$e^{j 2\pi f t} = \cos(2\pi f t) + j \sin(2\pi f t).$$
(17)

Se o sinal em análise for periódico, com período *T*, o sinal pode ser expandido em série de Fourier como [34]

$$x(t) = \sum_{n = -\infty}^{\infty} C_n e^{j2\pi nt/T}.$$
(18)

Os coeficientes da série são dados por

$$C_n = \frac{1}{T} \int_0^T x(t) e^{-j2\pi nt/T} dt .$$
 (19)

Uma propriedade importante da série de Fourier é o Teorema de Bessel-Parseval, que mostra que a potência do sinal pode também ser calculada a partir dos seus coeficientes de Fourier:

$$\sum_{n = -\infty}^{\infty} |C_n|^2 = \frac{1}{T} \int_0^T |x(t)|^2 .$$
(20)

Ao analisar um sinal aperiódico, podemos considerar que se trata do caso limite de

fazer a frequência fundamental tender para zero, pelo que o somatório de (18) passa a ser um integral,

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi f t} df , \qquad (21)$$

e

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi f t} df , \qquad (22)$$

Como os intervalos de integração vão agora de $-\infty$ a ∞ , o sinal deve ser limitado não em potência, como no caso do sinal periódico, mas em energia. E a energia pode ser calculada tanto sobre o sinal original como sobre a sua transformada, segundo uma expressão semelhante a (20):

$$\int_{-\infty}^{\infty} |X(f)|^2 df = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad .$$
 (23)

C.2 Amostragem

Em [35], o autor começa por perguntar qual poderá ser o significado do espectro de uma sequência discreta de valores a(n). Para responder a essa pergunta, começa por construir um sinal contínuo "parecido" com o sinal discreto:

$$a^*(t) = \sum_{n = -\infty}^{\infty} a(n)\delta(t - nT).$$

Agora é possível calcular a transformada de Fourier desse sinal, e verifica-se que

$$a^*(j\omega) = \sum_{n = -\infty}^{\infty} a(n)e^{-j\omega nT} = A(z)|_{z = e^{j\omega T}} = A'(\omega) ,$$

em que A(z) é a transformada Z de a(n).

Outra abordagem consiste em considerar que a sequência de valores discretos a(n)é obtida por amostragem de um sinal contínuo $a_c(t)$, e que matematicamente a amostragem consistiria em multiplicar $a_c(t)$ por um "pente" de impulsos de Dirac i(t). Assim, no domínio da frequência teríamos a convolução de $A*(j\omega)$ e de $I(j\omega)$.

A transformada de Fourier do pente de impulsos também é um pente de impulsos:

$$I(j\omega) = \int_{-\infty}^{\infty} \left[\sum_{n = -\infty}^{\infty} \delta(t - nT) \right] e^{-j\omega t} dt$$
(24)

$$=\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-j\omega t}\delta(t-nT)dt$$
(25)

$$=\sum_{n=-\infty}^{\infty}e^{-j\omega nT}$$
(26)

Não é "completamente óbvio" que (26) seja um pente de impulsos, mas pode intuir-se que se $\omega T \neq 2\pi k$, para $k = 0, \pm 1, \pm 2,...$, estamos a somar vectores rodados em torno da origem que irão cancelar-se ao serem somados. Já se $\omega T = 2\pi k$, $e^{-j\omega nT} = 1$, para qualquer valor de *n*. Assim, teremos impulsos de Dirac localizados em frequências que sejam múltiplos da frequência de amostragem, ou seja,

$$I(j\omega) = \sum_{k=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi}{T}k\right).$$

Portanto, o espectro do sinal amostrado resulta da convolução com impulsos espaçados pela frequência de amostragem.

C.3 Conversão AD-DA, Pré- e Pós-Filtragem

Normalmente é utilizado um pré-filtro para evitar a ocorrência de *aliasing*. Para obter um determinado nível de atenuação de *aliasing* pode ser necessário aumentar a ordem do filtro, para que o decaimento fora da banda passante seja mais rápido, ou aumentar a frequência de amostragem, para que haja espaço (ou melhor dizendo, espectro) para ocorrer atenuação suficiente.

O conversor DA tem uma resposta impulsiva to tipo impulso rectangular, a que corresponde uma resposta em frequência do tipo $\frac{\sin f}{f}$. O pós-filtro terá não só de eliminar as réplicas posicionadas nos múltiplos de f_s mas também, em alguns casos, de compensar o efeito

da resposta em frequência do DA dentro da banda de interesse.

C.4 Sinais Complexos

Assim como os números *negativos* surgiram como soluções para equações do tipo x + 23 = 14, os números *imaginários* surgiram como soluções de equações do tipo $x^2 = -5$, e os números *complexos* como soluções de equações do tipo $(x + 1)^2 = -1$.

Considerando isso, todos deveriam ser igualmente fáceis ou difíceis de entender, mas tal não acontece. Talvez a designação infeliz de "imaginários" ou "complexos" contribua em parte para essa estranheza. É uma designação infeliz porque, afinal, *todos* os números são imaginários, na medida em que são fruto da imaginação humana. Todos são abstracções que podem ser usadas para modelar, com maior ou menor adequação, diferentes aspectos do nosso mundo físico.

Mas, em comparação com os números complexos, os números negativos parecem "quase naturais"¹, para quem teve alguma instrução matemática. Isso certamente acontece porque podem ser usados para representar diversos fenómenos razoavelmente familiares: temperaturas, dívidas, posições relativamente a uma referência, tensão eléctrica, etc.

Já com os números complexos, não acontece o mesmo. É preciso lidar com problemas de engenharia para lhes achar utilidade, e esses problemas são muito menos familiares para a grande maioria das pessoas.

E como surgem os números complexos na engenharia? Os casos mais elementares resultam do estudo de sistemas que apresentam variações sinusoidais. As sinusóides são tão frequentes porque constituem a solução de sistemas com movimentos rotativos, ou com oscilações (mesmo que amortecidas) de alguma grandeza. Exemplo disso são sistemas eléctricos de tensão alternada.

Ao efectuar cálculos com sinusóides, é preciso considerar a amplitude da sinusóide e a sua fase em relação a uma referência. Os números complexos constituem uma forma cómoda de armazenar essas duas informações numa única variável. Claro que, ao efectuar os cálculos, é preciso decompor cada variável complexa nas suas componentes "real" e "imaginária" (ou de amplitude e fase, no caso da representação polar), mas pelo menos as equações sim-

¹Ou, com mais rigor, *quase tão intuitivos* quanto os números *naturais*.

bólicas ficam mais compactas e a manipulação destas fica mais simples.

E como surgiu a relação entre trigonometria e números complexos? Surgiu quando alguém (Argand, para ser mais preciso) se lembrou de tratar a componente real e imaginária de uma quantidade complexa como se fossem coordenadas num plano. Ao aplicar as leis da trigonometria a esse plano, surgiu a equivalência entre dois tipos de representação de um número complexo, a representação cartesiana e a representação polar:

$$z = a + j b$$

$$z = r \cos \theta + j r \sin \theta$$
(27)

Vamos juntar agora dois "ingredientes" aos que já temos. O primeiro é o desenvolvimento de uma função em série de Taylor (ou de Mac-Laurin, quando o desenvolvimento é feito em torno do ponto x=0):

$$f(x) = f(0) + f'(0)\frac{x}{1} + f''(0)\frac{x^2}{2!} + \dots + f^{(n)}(0)\frac{x^n}{n!} + \dots$$

O segundo ingrediente é o génio matemático de Euler, que observou que

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots ,$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots ,$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots ,$$

e "ousou" introduzir um argumento complexo na exponencial! Observe-se que a generalização das operações aritméticas elementares aos números complexos é relativamente óbvia, mas a generalização dos argumentos de funções transcendentes a valores complexos não seria tanto. Qual seria o resultado dessa operação? O desenvolvimento em série de potências foi o caminho escolhido para justificar essa generalização, uma vez que a substituição da variável *x* por um valor complexo, em cada termo da série, não requer mais que a aplicação de operações aritméticas.

Dessa "ousadia" resultou portanto que

$$e^{jx} = 1 + jx + j^2 \frac{x^2}{2!} + j^3 \frac{x^3}{3!} + \dots$$
$$= 1 + jx - \frac{x^2}{2!} - j\frac{x^3}{3!} + \dots$$

de onde se obtém a celebrada equação de Euler:

$$e^{jx} = \cos x + j \sin x$$

Desta equação resultam outras igualmente úteis...

$$\cos x = \frac{e^{jx} + e^{-jx}}{2}, \quad \sin x = \frac{e^{jx} - e^{-jx}}{2j}.$$

C.5 Decomposição Harmónica

Os fundamentos da análise espectral são devidos sobretudo ao trabalho de Fourier, que propôs a decomposição de funções periódicas em séries de *harmónicas* (sinusóides cujas frequências seriam múltiplos inteiros da frequência da função, ou sinal, original).

Ao fazer essa decomposição, há várias escolhas possíveis em relação à escrita das equações de análise (processo de decompor o sinal original) e de síntese (processo de reconstruir o sinal original). Podem ser usadas séries de senos, de cossenos, apenas com frequências positivas ou também com frequências negativas...

Frequências negativas? Talvez esta aplicação dos números negativos pareça um pouco surpreendente à primeira vista, mas parecê-lo-á menos se nos lembrarmos que sin(-x) = -sin(x) e portanto uma sinusóide pode ser decomposta numa contribuição de uma frequência positiva e outra negativa, i.e.,

$$\sin(2\pi f t) = \frac{1}{2} \sin(2\pi f t) - \frac{1}{2} \sin(-2\pi f t)$$

Fechando o parêntesis das frequências negativas, e voltando à escolha da convenção usada nas equações de análise e síntese do sinal, uma outra escolha possível é fazer a decomposição em exponenciais complexas. Pode parecer a opção menos intuitiva, porque envolve um maior número de conceitos matemáticos mais avançados (exponenciais, números complexos e frequências negativas), mas tem várias vantagens: não é necessário fazer uma escolha arbitrária entre senos e cossenos², uma única variável complexa permite armazenar informação sobre amplitude e fase e permite até lidar com sinais de valor complexo.

São necessárias equações diferentes conforme o sinal x for contínuo ou discreto, e periódico ou aperiódico. A tabela abaixo resume a nomenclatura e as equações.



Tabela 6.1 Sumário de transformadas de Fourier.

A Transformada Rápida de Fourier (FFT, de *Fast Fourier Transform*) é uma implementação da DFT (*Discrete Fourier Transform*) que utiliza truques algorítmicos para acelerar o cálculo dos coeficientes.

Vamos verificar se os resultados da FFT implementada num pacote de software como o Matlab coincidem com os resultados obtidos por aplicação da fórmula da DFT da tabela 6.1. Para isso iremos definir dois sinais simples,

²Na verdade, essa escolha foi feita ao escrever a relação (27), e só tem consequências sobre a referência para a fase.

$$x_s(t) = \sin \omega t = \frac{e^{j\omega t} - e^{-j\omega t}}{2j}, \quad x_c(t) = \cos \omega t = \frac{e^{j\omega t} + e^{-j\omega t}}{2},$$

cujos coeficientes de DFT são facilmente obtidos por inspecção:

$$X_{s}[\omega] = \frac{1}{2j} = -\frac{1}{2}j = \frac{1}{2}e^{-j\pi/2}, \quad X_{s}[-\omega] = -\frac{1}{2j} = \frac{1}{2}j = \frac{1}{2}e^{j\pi/2}$$
(28)

$$X_{c}[\omega] = X_{c}[-\omega] = \frac{1}{2} = \frac{1}{2}e^{j0}.$$
 (29)

Iremos agora utilizar o seguinte programa de teste do Matlab:

```
1
      % Verificação do cálculo de DFT no Matlab
 2
 3
      % preparar definição de (co)sinusóide com período N
 4
      N=16; n=0:N-1; k=n;
 5
      wt=linspace(0,2*pi,N+1); wt=wt(1:N);
 6
 7
      xs=sin(wt); Xs=fft(xs); fase s = angle(Xs)/pi*180;
 8
      xc=cos(wt); Xc=fft(xc); fase c = angle(Xc)/pi*180;
9
10
      % Eliminar o valor de fase para amplitudes próximas de zero
11
      tol=1e-3;
12
      for i=1:length(Xs);
13
         if abs(Xs(i)) < tol, fase s(i) = NaN; end
14
         if abs(Xc(i)) < tol, fase c(i) = NaN; end
15
      end
16
17
      % visualizar resultados
18
      subplot(2,2,1), stem(k,fftshift(abs(Xs)))
19
      title('fft(sin(...))'),grid
20
      ylabel('abs(...)')
21
      set(gca,'xlim',[0 N-1],'xtick',0:4:N-1);
22
      subplot(2,2,3), stem(k,fftshift(fase s))
23
      ylabel('angle(...)')
24
      axis([0 N-1 -180 180]),grid
25
      set(gca,'xtick',0:4:N-1,'ytick',-180:90:180);
26
      subplot(2,2,2), stem(k,fftshift(abs(Xc)))
27
      title('fft(cos(...))'),grid
28
      set(gca,'xlim',[0 N-1],'xtick',0:4:N-1)
29
      subplot(2,2,4), stem(k,fftshift(fase c))
30
      axis([0 N-1 -180 180]),grid
31
      set(gca,'xtick',0:4:N-1,'ytick',-180:90:180);
```

Ao executar o programa anterior obtêm-se os resultados da fig. C.1. Esses resultados estão de acordo com os valores (28) e (29).



Figura C.1 Resultados de simulação.

De em modo geral, ao analisar um sinal por intermédio de uma FFT, há que ter dois tipos de cuidados básicos:

- Na amostragem do sinal de interesse,
- Na interpretação dos resultados, em especial...
 - da fase,
 - dos valores de frequência.

Esses cuidados foram observados no programa anterior. Por exemplo, ao amostrar o sinal, há que evitar um erro fácil de cometer: permitir que a última amostra seja a repetição da primeira. Isso é exemplificado na linha 5 do programa. Se essa linha tivesse sido substituída simplesmente por

```
wt=linspace(0,2*pi,N);
```

a fase ωt seria um vector de *N* elementos, a variar de 0 a 2π , pelo que as formas de onda em análise seriam tais que x[1] = x[N] (período N-1). Para que o período seja realmente

N, é necessário que

$$x[0] = x[N], x[1] = x[N+1], \dots,$$

como foi feito no programa original. Se isso não for feito e o sinal for um seno, essa amostra repetida tem um valor nulo. Nesse caso, o efeito é muito pequeno, pois é como se estivéssemos a analisar uma sinusóide de frequência ligeiramente menor que a pretendida e tivéssemos utilizado uma FFT maior do que o necessário (acrescentando amostras nulas) para obter um espectro com maior resolução (esta técnica será discutida adiante em maior detalhe). Se o sinal tiver uma amostra repetida de valor não nulo, como aconteceria com o cosseno, o efeito seria um pouco pior, pois de acordo com a FFT assim obtida, o nosso sinal tem média não nula. Isso pode ser observado na fig. C.2.



Figura C.2 Diferenças no resultado devidas a uma amostra repetida.

Um cuidado a ter na interpretação dos valores de fase é a eliminação dos valores de fase relativos a pontos do espectro de amplitude nula ou quase nula. Se a amplitude for nula, a fase é indefinida. Se omitirmos do programa as linhas 10–15, observa-se que os valores de fase apresentam variações sem significado (cf. fig. C.3).

Um outro cuidado, que facilita a interpretação dos valores de frequência, ou pelo menos a sua comparação com a representação tradicional do espectro (frequência nula no centro do eixo das abcissas, com valores de frequência negativos e positivos), consiste no desloca-



Figura C.3 Pontos onde a fase é indefinida apresentam valores aleatórios.

mento das metades esquerda e direita dos valores de resultado da FFT. No Matlab, esse deslocamento pode ser feito por intermédio da função fftshift.

O efeito da função fftshift e o espectro de sinais com diferentes frequências pode ser visualizado com o programa abaixo.

```
1
      % Comparação do espectro para sinais de diferentes frequências
 2
 3
      N=8; n=0:N-1; k=n;
 4
 5
      for nf=0:N/2;
 6
 7
        % preparar definição de (co)sinusóide com período N
 8
        wt =linspace(0,nf*2*pi,N+1); wt=wt(1:N);
 9
        % preparar definição do mesmo sinal com + amostras
10
        wti=linspace(wt(1),wt(N),10*N);
11
        ni =linspace(n (1), n (N), 10*N);
12
13
        x=cos(wt); xi=cos(wti);
14
        X=fft(x);
15
16
        subplot(3, N/2+1, 0*(N/2+1)+nf+1), stem(n, x)
17
      title(['f=' num2str(nf) '/N'])
18
      set(gca,'xlim',[0 N-1],'xtick',0:4:N-1)
19
      % sobrepôr sinal com amostragem mais fina
20
      hold on, plot(ni,xi,'--')
```

```
21
22
        k=0:N-1;
23
        subplot(3, N/2+1, 1*(N/2+1)+nf+1), stem(k, abs(X))
24
      if nf==0, ylabel('|X|'), end, grid
25
      set(gca,'ylim',[0 N+1],'ytick',0:4:8)
26
        set(gca,'xlim',[0 N-1],'xtick',0:4:N-1)
27
28
        X=fftshift(X);
29
        k=-4:N-1-4;
30
        subplot(3, N/2+1, 2*(N/2+1)+nf+1), stem(k, abs(X))
31
      if nf==0, ylabel(`fftshift(|X|)'), end, grid
32
      set(gca,'ylim',[0 N+1],'ytick',0:4:8)
33
        set(gca,'xlim',[-4 N-1-4],'xtick',-3:3:3)
34
35
      end
```

Os resultados são apresentados na fig. C.4.



Figura C.4 Sinais com diferentes frequências e respectivos espectros, com e sem reordenação das abcissas.

C.6 O Algoritmo da FFT

Conhecer o algoritmo da FFT não é de todo essencial para utilizá-lo com sucesso, mas pode ser interessante e certamente contribui para enfrentar com mais confiança os desafios técnicos da área.

A abordagem mais clara que encontrei do algoritmo foi a de [34], mas ainda assim

dá algum trabalho "interpolar" os passos da explicação. Por essa razão, aqui deixo registado esse trabalho.

A FFT é uma implementação da DFT, cuja fórmula é

$$X[k] = \sum_{n=0}^{N-1} x_n e^{-j2\pi kn/N} = \sum_{n=0}^{N-1} W_N^{kn} x_n ,$$

em que $W_N = e^{-j2\pi/N}$. Omitiu-se o factor de escala, que aliás pode ser colocado tanto na transformada directa como na inversa

A expressão anterior também pode ser escrita matricialmente:

$$\begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ \dots \\ X_{N-1} \end{bmatrix} = \begin{bmatrix} W_N^{0 \times 0} & W_N^{0 \times 1} & W_N^{0 \times 2} & \dots & W_N^{0 \times (N-1)} \\ W_N^{1 \times 0} & W_N^{1 \times 1} & W_N^{1 \times 2} & \dots & W_N^{1 \times (N-1)} \\ W_N^{2 \times 0} & W_N^{2 \times 1} & W_N^{2 \times 2} & \dots & W_N^{2 \times (N-1)} \\ \dots & \dots & \dots & \dots & \dots \\ W_N^{(N-1)0} & W_N^{(N-1)1} & W_N^{(N-1)2} & \dots & W_N^{(N-1)(N-1)} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \dots \\ x_{N-1} \end{bmatrix}$$

É fácil calcular o número de operações para o cálculo de uma DFT de N pontos. Se não contabilizarmos o cálculo dos coeficientes W_N^{kn} (pois são constantes e por isso podem ser armazenados em memória, para DFTs com o mesmo número de pontos), temos N^2 multiplicações e $N(N-1) \approx N^2$ adições. Na verdade, as operações referidas são sobre números complexos, pelo que deveríamos considerar que cada multiplicação complexa corresponde a quatro multiplicações reais e duas adições reais, enquanto que cada adição complexa corresponde a duas adições reais. Assim, tem-se um total de 4 N^2 somas e 4 N^2 produtos.

Para apresentar um exemplo concreto, vamos considerar que N = 8, e portanto

$$\begin{bmatrix} X_{0} \\ X_{1} \\ X_{2} \\ \dots \\ X_{7} \end{bmatrix} = \begin{bmatrix} W_{8}^{0} \ W_{8}^{0} \ W_{8}^{0} \ \dots \ W_{8}^{0} \\ W_{8}^{0} \ W_{8}^{1} \ W_{8}^{2} \ \dots \ W_{8}^{1} \\ W_{8}^{0} \ W_{8}^{2} \ W_{8}^{4} \ \dots \ W_{8}^{14} \\ \dots \ \dots \ \dots \ \dots \ \dots \\ W_{8}^{0} \ W_{8}^{7} \ W_{8}^{14} \ \dots \ W_{8}^{49} \end{bmatrix} \begin{bmatrix} x_{0} \\ x_{1} \\ x_{2} \\ \dots \\ x_{7} \end{bmatrix}$$

Vamos agora introduzir uma notação em que uma matriz de potências com a

mesma base é representada pela base elevada à matriz dos expoentes (pode soar complicado, mas observando a expressão seguinte torna-se óbvio) e o mesmo é feito relativamente aos índices de x e X.

$$\begin{array}{c} x \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 2 & 4 & 6 & 8 & 10 & 12 & 14 \\ 0 & 3 & 6 & 9 & 12 & 15 & 18 & 21 \\ 0 & 4 & 8 & 12 & 16 & 20 & 24 & 28 \\ 0 & 5 & 10 & 15 & 20 & 25 & 30 & 35 \\ 0 & 6 & 12 & 18 & 24 & 30 & 36 & 42 \\ 0 & 7 & 14 & 21 & 28 & 35 & 42 & 49 \end{bmatrix} \begin{array}{c} x \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{bmatrix}$$

Agora, iremos repartir o cálculo dos coeficientes de X em duas metades,

$${}^{X}\begin{bmatrix}0\\1\\2\\3\end{bmatrix} = \begin{bmatrix}0 & 0 & 0 & 0\\0 & 2 & 4 & 6\\0 & 4 & 8 & 12\\0 & 6 & 12 & 18\end{bmatrix} {}^{x}\begin{bmatrix}0\\2\\4\\6\end{bmatrix} + \begin{bmatrix}0 & 0 & 0 & 0\\1 & 3 & 5 & 7\\2 & 6 & 10 & 14\\3 & 9 & 15 & 21\end{bmatrix} {}^{x}\begin{bmatrix}1\\3\\5\\7\end{bmatrix}$$

$${}^{X}\begin{bmatrix}4\\5\\6\\7\end{bmatrix} = \begin{bmatrix}0 & 8 & 16 & 24\\0 & 10 & 20 & 30\\0 & 12 & 24 & 36\\0 & 14 & 28 & 42\end{bmatrix} {}^{x}\begin{bmatrix}0\\2\\4\\6\end{bmatrix} + \begin{bmatrix}4 & 12 & 20 & 28\\5 & 15 & 25 & 35\\6 & 18 & 30 & 42\\7 & 21 & 35 & 49\end{bmatrix} {}^{x}\begin{bmatrix}1\\3\\5\\7\end{bmatrix}$$

Sabemos que

e

$$\begin{bmatrix} x^{d_1} & 0 \\ 0 & x^{d_2} \end{bmatrix} \begin{bmatrix} x^{a_{11}} & x^{a_{12}} \\ x^{a_{21}} & x^{a_{22}} \end{bmatrix} = \begin{bmatrix} x^{d_1 + a_{11}} & x^{d_1 + a_{12}} \\ x^{d_2 + a_{21}} & x^{d_2 + a_{22}} \end{bmatrix}$$

resultado esse que pode ser representado na nossa notação por

$$\begin{bmatrix} x^{d_1} & 0 \\ 0 & x^{d_2} \end{bmatrix}_{x} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \left(\begin{bmatrix} d_1 & d_1 \\ d_2 & d_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right)$$

Relembrando que $W_N = e^{-j2\pi/N}$, verifica-se que

$$W_8^{24} = W_8^{16} = W_8^8 = W_8^0 = 1$$
 e $W_8^{28} = W_8^{20} = W_8^{12} = W_8^4 = -1$,

o que permite simplificar o cálculo da segunda metade dos coeficientes:

X		0 0	0	0	$\begin{bmatrix} x \\ 0 \end{bmatrix}$		([0	0	0	0		0 0	0	0	$)^{x}$	1
5	=	0 2	4	6	2	_		1	1	1	1	+	0 2	4	6		3
6		04	8	12	4			2	2	2	2		04	8	12		5
7	W_8	0 6	12	18	6	W_8		3	3	3	3		0 6	12	18)	7

Como $W_N^a = W_{N/2}^{a/2}$, podemos escrever

	0	0	0	0			0	0	0	0
	0	2	4	6	=		0	1	2	3
	0	4	8	12			0	2	4	6
W_8	0	6	12	18		W_8	0	3	6	9

do que se conclui que as metades da DFT de 8 pontos podem ser calculadas através de duas DFTs de 4 pontos. Este processo pode ser reiterado até atingirmos DFTs de um ponto, para as quais X = x.

Este processo de decomposição sucessiva é ilustrado na fig. C.5. Nessa figura os pontos negros indicam somas ou subtracções. A multiplicação de uma sub-FFT pela matriz diagonal pode ser feita simplesmente como um produto componente a componente entre dois

vectores.



Figura C.5 FFT de 8 pontos, com algumas das sub-FFTs assinaladas.

Para N pontos, quantas decomposições serão feitas? Os N pontos serão divididos por 2, até que N/2/2/.../2 = 1. Se d for o número de divisões ou decomposições por que passa uma amostra do sinal original, então

$$N/2^d = 1 \Leftrightarrow N = 2^d \Leftrightarrow d = \log_2 N$$

Nas expressões acima, para N pontos efectuam-se N/2 adições e N/2 subtracções (total N adições complexas) e uma multiplicação por matriz diagonal (N/2 multiplicações complexas). Isso corresponde a 3N adições reais e 2N multiplicações reais. O produto pelas matrizes de transformada não é efectuado, é substituído pela recursão para os níveis seguintes. Nesses restantes níveis, embora as amostras de sinal estejam repartidas, perfazem o mesmo número de amostras e portanto o mesmo número de operações, pelo que se tem um total de 3N $\log_2 N$ adições reais e $2N \log_2 N$ multiplicações reais.

algoritmo	multiplicações	adições
directo	4N2	4N2
FFT	2N log2 <i>N</i>	3N log2 <i>N</i>

Em suma,

Comparemos os números de multiplicações (a operação mais "pesada") da DFT

algoritmo	<i>N</i> =32	<i>N</i> =256	<i>N</i> =1024
directo	4096	262144	4.194.304
FFT	320	4096	20.480
ganho	12,8	64	204,8

directa e da FFT, para diferentes valores de N...

Os ganhos de tempo globais serão um pouco menores do que as estimativas acima, pois o algoritmo da FFT requer reordenamentos da sequência de valores, mas ainda assim serão ganhos apreciáveis.