

SIMON HAYKIN | MICHAEL MOHER

Introdução aos
**SISTEMAS DE
COMUNICAÇÃO**
| SEGUNDA EDIÇÃO



Introdução aos

SISTEMAS DE COMUNICAÇÃO



H419i Haykin, Simon.

Introdução aos sistemas de comunicação [recurso eletrônico]
/ Simon Haykin, Michael Moher ; tradução Gustavo Guimarães
Parma. – Dados eletrônicos. – 2. ed. – Porto Alegre : Bookman,
2008.

Edited also as a printed book in 2008.
ISBN 978-85-7780-329-3

1. Telecomunicações. I. Moher, Michael. II. Título.

CDU 621.39

SIMON HAYKIN | MICHAEL MOHER

McMaster University, Hamilton,
Ontario, Canada

Space-Time DSP, Ottawa,
Ontario, Canada

Introdução aos
**SISTEMAS DE
COMUNICAÇÃO**
| SEGUNDA EDIÇÃO

Tradução:

Gustavo Guimarães Parma
Doutor em Engenharia Elétrica pela UFMG

Consultoria, supervisão e revisão técnica desta edição:

Antonio Pertence Júnior
Engenheiro Eletrônico e de Telecomunicações
Especialista em Processamento de Sinais (Ryerson University – Canadá)
Professor de Telecomunicações da FUMEC/MG
Professor Titular da Faculdade de Sabará/MG
Membro da Sociedade Brasileira de Eletromagnetismo (SBmag)

Versão impressa
desta obra: 2008



2008

Obra originalmente publicada sob o título
An Introduction to Digital and Analog Communications, 2nd. Edition
ISBN 978-0-471-43222-7

Copyright © 2007 John Wiley & Sons, Inc. All rights reserved.
This translation published under license.

Capa: *Gustavo Demarchi, arte sobre capa orginal*

Preparação de original: *Rachel Garcia Valdez*

Supervisão editorial: *Denise Weber Nowaczyk*

Editoração eletrônica: *Techbooks*

Reservados todos os direitos de publicação, em língua portuguesa, à
ARTMED® EDITORA S.A.
(BOOKMAN® COMPANHIA EDITORA é uma divisão da ARTMED® EDITORA S. A.)
Av. Jerônimo de Ornelas, 670 - Santana
90040-340 Porto Alegre RS
Fone (51) 3027-7000 Fax (51) 3027-7070

É proibida a duplicação ou reprodução deste volume, no todo ou em parte,
sob quaisquer formas ou por quaisquer meios (eletrônico, mecânico, gravação,
fotocópia, distribuição na Web e outros), sem permissão expressa da Editora.

SÃO PAULO
Av. Angélica, 1.091 - Higienópolis
01227-100 São Paulo SP
Fone (11) 3665-1100 Fax (11) 3667-1333

SAC 0800 703-3444

IMPRESSO NO BRASIL
PRINTED IN BRAZIL

Para os pioneiros da comunicação do século 20, os quais,
por meio de suas teorias matemáticas e dispositivos geniais,
transformaram nosso planeta em uma vila global.

Agradecimentos

Os autores gostariam de expressar sua profunda gratidão a:

- Lily Jiang, anteriormente na MacMaster University, por sua ajuda na execução de vários dos experimentos de computador incluídos no texto.
- Wei Zhang, por toda ajuda, correções e melhorias que ela fez ao texto.

Também agradecem ao Dr. Stewart Crozier e Dr. Paul Guinand, ambos do Centro de Pesquisa em Comunicações, Ottawa, pela ajuda em diferentes partes do livro.

Eles estão em débito com Catherine Fields Shultz, Editora Sênior de Aquisições e Gerente de Produtos (Engenharia e Ciência da Computação) da John Wiley and Sons, Bill Zobrist anteriormente da Wiley e Lisa Wojcik, Editora Sênior de Produção da Wiley, pela orientação e dedicação na produção deste livro.

Por fim, mas não por menos, eles são gratos a Lola Brooks, MacMaster University, por seu trabalho na preparação do manuscrito e das questões relacionadas com o livro.

Uma disciplina introdutória de comunicações analógicas e digitais é fundamental para o programa de graduação em engenharia elétrica. Em geral, é oferecida no meio do curso e considera-se que o estudante já tenha cursado disciplinas de cálculo, eletrônica, sinais e sistema e, possivelmente, teoria de probabilidade.

Tendo em mente a natureza introdutória dessa disciplina, um livro-texto recomendado deve ser de fácil leitura, preciso e conter exemplos, problemas e experimentos de computador em abundância. Essas características são necessárias para facilitar o aprendizado dos fundamentos de sistemas de comunicação em um nível introdutório e de uma maneira eficaz. Este livro foi escrito com esses objetivos em mente.

Dada a natureza matemática da teoria de comunicação, é muito fácil o leitor esquecer o lado prático dos sistemas de comunicação. Neste livro, fizemos um esforço especial para não cair nesta armadilha abordando os assuntos de forma ordenada, tentando manter sempre o tratamento matemático em um nível de fácil entendimento e mostrando a relevância prática da teoria sempre que apropriado.

FILOSOFIA DA ESTRUTURA DO LIVRO

Para facilitar e reforçar a aprendizagem, o formato e a diagramação do livro foram estruturados para:

- Motivar a leitura do livro e aprender com ele.
- Enfatizar conceitos básicos de uma perspectiva de “sistemas”, fazendo isto de forma ordenada.
- Ilustrar a aplicação da teoria pertinente por meio de exemplos e experimentos de computador, sempre que apropriado.
- Fornecer problemas de desenvolvimento logo após a discussão de conceitos fundamentais para ajudar o leitor a verificar e reforçar os conceitos sob discussão.
- Fornecer problemas adicionais ao fim de cada capítulo, alguns de natureza avançada, para estender a teoria estudada no texto.

ORGANIZAÇÃO DO LIVRO

1. *Motivação.* Antes de se envolver com o estudo de comunicações analógicas e digitais, é imperativo que o leitor seja motivado a utilizá-lo e a aprender com ele. Por isso, o Capítulo 1 começa com um histórico de sistemas de comunicação e aplicações importantes do assunto.
2. *Teoria da modulação.* As comunicações digitais têm sobrepujado as comunicações analógicas como forma dominante de comunicação. Apesar das duas formas de comunicação trabalharem maneira diferente, a teoria da modulação é básica a ambas. Além disso, é mais fácil entender esse importante assunto analisando primeiro seus conceitos fundamentais aplicados a comunicações analógicas e, então, seguir para as comunicações digitais. Também, a modulação em amplitude é mais simples de ser apresentada do que a modulação em ângulo. Outro ponto relevante é o fato de que para entender a teoria da modulação, é importante que a teoria de Fourier seja dominada primeiro. Com esses pontos em mente, os Capítulos 2 a 7 estão organizados da seguinte maneira:

- ▶ Capítulo 2 é dedicado à revisão da representação de Fourier para sinais e sistemas.
 - ▶ Capítulos 3 e 4 são dedicados às comunicações analógicas, com o Capítulo 3 mostrando a modulação em amplitude e o Capítulo 4 apresentando a modulação em ângulo.
 - ▶ Capítulo 5 sobre modulação de pulso cobre os conceitos pertinentes à transição de comunicações analógicas para digitais.
 - ▶ Capítulos 6 e 7 são dedicados às comunicações digitais, com o Capítulo 6 cobrindo a transmissão de dados banda base e o Capítulo 7 cobrindo a transmissão de dados banda passante.
3. *Teoria da probabilidade e Detecção de sinal.* Tal como a análise de Fourier é fundamental para a teoria da modulação, a teoria da probabilidade é fundamental para a detecção de sinal e cálculo de performance do receptor na presença de ruído aditivo. Como a teoria da probabilidade não é crítica para o entendimento da modulação, propositadamente atrasamos a revisão da teoria da probabilidade, sinais aleatórios e ruídos para o Capítulo 8. Então, com um bom entendimento da teoria da modulação aplicada a comunicações analógicas e digitais e com os relevantes conceitos da teoria da probabilidade e modelos probabilísticos, o estágio necessário é revisar receptores de comunicação analógica e digital, como resumido por:
- ▶ Capítulo 9 discute o ruído em comunicações analógicas.
 - ▶ Capítulo 10 discute o ruído em comunicações digitais. Como as comunicações analógicas e digitais operam de forma diferente, é natural ver algumas das diferenças no tratamento dos efeitos do ruído nestes dois capítulos.
4. *Ruído.* O estudo introdutório de comunicações analógicas e digitais é completado no Capítulo 11. Este capítulo ilustra os papéis da modulação e ruído em sistemas de comunicação analisando quatro coisas:
- ▶ Primeiro, as fontes físicas do ruído são descritas, principalmente o ruído térmico e o ruído shot.
 - ▶ Segundo, as métricas de figura de ruído e temperatura de ruído são introduzidas.
 - ▶ Terceiro, é explicado como a propagação afeta a potência do sinal em comunicações sem fio, via satélite e terrestre.
 - ▶ Finalmente, mostramos como a potência do sinal e os cálculos de ruído podem ser combinados para fornecer uma estimativa da razão sinal/ruído, a figura de mérito fundamental em sistemas de comunicação.
5. *Exemplos de temas.* Para ressaltar aplicações práticas importantes da teoria de comunicações, exemplos de temas são incluídos sempre que apropriado. Os exemplos são retirados tanto das comunicações analógicas quanto digitais.
6. *Apêndices.* Para fornecer material de apoio ao texto, oito apêndices estão incluídos no final do livro e cobrem o seguinte:
- ▶ Razões de potência e decibel
 - ▶ Série de Fourier
 - ▶ Funções de Bessel
 - ▶ A função Q e sua relação com a função erro
 - ▶ Desigualdade de Schwarz
 - ▶ Tabelas matemáticas
 - ▶ Scripts MATLAB para experimentos de computador dos problemas dos Capítulos 7 a 10
 - ▶ Respostas dos problemas de desenvolvimento

7. *Notas de rodapé.* Incluídas ao longo do livro, ajudam o leitor interessado, fornecendo referências selecionadas para o aprendizado avançado.
8. *Material auxiliar.* O livro é, essencialmente, autocontido. Um glossário de símbolos e bibliografia é apresentado ao final do livro. O professor que adotar o livro terá um *manual de soluções* detalhado para todos os problemas.*

COMO UTILIZAR O LIVRO

O livro pode ser utilizado de diversas formas para um curso introdutório sobre comunicações analógicas e digitais, dependendo do conhecimento prévio dos estudantes e do interesse e das responsabilidades do professor. A seguir são mostrados dois modelos:

Modelo de curso A: Curso completo em dois semestres

- (A.1) O curso do primeiro semestre sobre teoria da modulação estuda os Capítulos 2 a 7.
- (A.2) O curso do segundo semestre sobre ruído em sistemas de comunicação estuda os Capítulos 8 a 11.

Modelo de curso B: Cursos de dois semestres, um sobre analógico e outro sobre digital

- (B.1) O primeiro curso, sobre comunicação analógica, começa como o material de revisão do Capítulo 2 sobre análise de Fourier, seguido pelo Capítulo 3 sobre modulação em amplitude e Capítulo 4 sobre modulação em ângulo. Ele, então, segue com uma revisão de partes relevantes do Capítulo 8, sobre ruído e, finalmente, termina com o Capítulo 9 sobre ruído em comunicações analógicas.
- (B.2) O segundo curso sobre comunicações digitais começa com o Capítulo 5 sobre modulação de pulso, seguido pelo Capítulo 6 sobre transmissão de dados banda base e Capítulo 7 sobre técnicas de modulação digital. Ele, então, segue com uma revisão de aspectos relevantes sobre a teoria da probabilidade no Capítulo 8 e, finalmente, termina com o Capítulo 10 sobre ruído em comunicações digitais.

Simon Haykin
Ancaster, Ontário, Canadá

Michael Moher
Ottawa, Ontário, Canadá

* Os professores interessados neste manual devem acessar a Área do Professor no site da Editora, www.bookman.com.br.

Capítulo 1 Introdução 19

- 1.1 Conhecimento histórico 19
- 1.2 Aplicações 22
- 1.3 Recursos primários e condições operacionais 30
- 1.4 A base das teorias de sistemas de comunicação 32
- 1.5 Notas de conclusão 34

Capítulo 2 Representação de Fourier de Sinais e Sistemas 35

- 2.1 A transformada de Fourier 36
 - 2.2 Propriedades da transformada de Fourier 42
 - 2.3 A relação inversa entre tempo e freqüência 56
 - 2.4 Função Delta de Dirac 58
 - 2.5 Transformada de Fourier de sinais periódicos 66
 - 2.6 Transmissão de sinais através de sistemas lineares:
a convolução revista 69
 - 2.7 Filtros passa-baixa ideais 76
 - 2.8 Correlação e densidade espectral: sinais de energia 86
 - 2.9 Densidade espectral de potência 95
 - 2.10 Determinação numérica da transformada de Fourier 97
 - 2.11 Tema de exemplo: par trançado para telefonia 105
 - 2.12 Resumo e discussão 106
- Problemas adicionais* 107
- Problemas avançados* 114

Capítulo 3 Modulação em Amplitude 116

- 3.1 Modulação em amplitude 117
- 3.2 Virtudes, limitações e modificações na modulação em amplitude 128
- 3.3 Modulação de faixa lateral dupla – portadora suprimida 129
- 3.4 Receptor Costas 135

3.5	Multiplexação por portadora em quadratura	137
3.6	Modulação por faixa lateral única	137
3.7	Modulação por faixa lateral vestigial	145
3.8	Representação banda base de ondas moduladas e filtros faixa passante	151
3.9	Exemplos de temas	156
3.10	Resumo e discussão	161
<i>Problemas adicionais</i>		162
<i>Problemas avançados</i>		164

Capítulo 4 Modulação em Ângulo 166

4.1	Definições básicas	166
4.2	Propriedades de ondas moduladas em ângulo	168
4.3	Relação entre ondas PM e FM	173
4.4	Modulação em freqüência de faixa estreita	174
4.5	Modulação em freqüência de faixa larga	177
4.6	Largura de faixa de transmissão de ondas FM	183
4.7	Geração de ondas FM	185
4.8	Demodulação de sinais FM	187
4.9	Exemplo de tema: multiplexação de FM estéreo	195
4.10	Resumo e discussão	197
<i>Problemas adicionais</i>		198
<i>Problemas avançados</i>		200

Capítulo 5 Modulação de Pulso: Transição de Comunicações Analógicas para Digitais

203

5.1	Processo de amostragem	204
5.2	Modulação por amplitude de pulso	210
5.3	Modulação por posição de pulso	214
5.4	Completando a transição de analógico para digital	216
5.5	Processo de quantização	217
5.6	Modulação por codificação de pulso	218

5.7	Modulação delta	224
5.8	Modulação por codificação diferencial de pulso	228
5.9	Código de linha	231
5.10	Exemplos de temas	232
5.11	Resumo e discussão	237
<i>Problemas adicionais</i> 238		
<i>Problemas avançados</i> 240		

Capítulo 6 Transmissão de Dados Banda Base 243

6.1	Transmissão banda base de dados digitais	243
6.2	Problema de interferência intersimbólica	245
6.3	Canal de Nyquist	246
6.4	Espectro do pulso de cosseno levantado	249
6.5	Transmissão banda base de dados M -ário	256
6.6	Padrão de olho	257
6.7	Experimento de computador: diagramas de olho para sistemas binários e quaternários	260
6.8	Exemplo de tema: equalização	263
6.9	Resumo e discussão	267
<i>Problemas adicionais</i> 267		
<i>Problemas avançados</i> 270		

Capítulo 7 Técnicas de Modulação Passa-Faixa 273

7.1	Algumas preliminares	273
7.2	Chaveamento binário de amplitude	276
7.3	Chaveamento de fase	280
7.4	Chaveamento em freqüência	289
7.5	Resumo dos três esquemas de sinalização binária	300
7.6	Esquemas de modulação digital não coerentes	300
7.7	Esquemas de Modulação Digital M -ário	305
7.8	Mapeamento de formas de onda digitalmente moduladas em constelações de pontos de sinais	309

7.9	Exemplo de temas	312
7.10	Resumo e discussão	318
<i>Problemas adicionais</i> 319		
<i>Problemas avançados</i> 321		
<i>Experimentos de computador</i> 322		

Capítulo 8 Sinais e Ruídos Aleatórios**324**

8.1	Probabilidade e variáveis aleatórias	325
8.2	Esperança	336
8.3	Transformação de variáveis aleatórias	339
8.4	Variáveis aleatórias Gaussianas	341
8.5	Teorema do limite central	344
8.6	Processos aleatórios	345
8.7	Correlação de processos aleatórios	348
8.8	Espectro de sinais aleatórios	353
8.9	Processos Gaussianos	357
8.10	Ruído branco	359
8.11	Ruído de faixa estreita	362
8.12	Resumo e discussão	367
<i>Problemas adicionais</i> 368		
<i>Problemas avançados</i> 371		
<i>Experimentos de computador</i> 373		

Capítulo 9 Ruído em Comunicações Analógicas**374**

9.1	Ruído em sistemas de comunicação	374
9.2	Razões sinal/ruído	376
9.3	Estruturas de receptor passa-faixa	379
9.4	Ruído em receptores lineares utilizando detecção coerente	380
9.5	Ruído em receptores AM utilizando a detecção de envelope	383
9.6	Ruído em receptores SSB	386
9.7	Detecção de modulação em freqüência (FM)	388

9.8	Pré-ênfase e de-ênfase FM	396
9.9	Resumo e discussão	398
<i>Problemas adicionais</i> 399		
<i>Problemas avançados</i> 400		
<i>Experimentos de computador</i> 401		

Capítulo 10 Ruído em Comunicações Digitais 402

10.1	Taxa de erro de bit	403
10.2	Detecção de um único pulso com ruído	404
10.3	Detecção ótima de PAM binário com ruído	407
10.4	Detecção ótima de BPSK	413
10.5	Detecção de QPSK e QAM com ruído	416
10.6	Detecção ótima de FSK binário	421
10.7	Detecção diferencial com ruído	424
10.8	Resumo de performance digital	425
10.9	Detecção e correção de erro	430
10.10	Resumo e discussão	441
<i>Problemas adicionais</i> 442		
<i>Problemas avançados</i> 442		
<i>Experimentos de computador</i> 444		

Capítulo 11 Cálculos de Sistemas e Ruídos 445

11.1	Ruído elétrico	445
11.2	Figura de ruído	450
11.3	Temperatura equivalente de ruído	451
11.4	Conexão em série de circuitos de duas portas	452
11.5	Cálculos de conexão em espaço livre	454
11.6	Rádio móvel terrestre	459
11.7	Resumo e discussão	464
<i>Problemas adicionais</i> 464		
<i>Problemas avançados</i> 465		

APÊNDICE 1 RAZÕES DE POTÊNCIA E DECIBEL 467

APÊNDICE 2 SÉRIE DE FOURIER 468

APÊNDICE 3 FUNÇÕES DE BESSEL 475

APÊNDICE 4 A FUNÇÃO Q E SUA RELAÇÃO COM A FUNÇÃO ERRO 478

APÊNDICE 5 DESIGUALDADE DE SCHWARZ 480

APÊNDICE 6 TABELAS MATEMÁTICAS 482

APÊNDICE 7 SCRIPTS MATLAB PARA EXPERIMENTOS DE COMPUTADOR
DOS PROBLEMAS DOS CAPÍTULOS 7 A 10 487

APÊNDICE 8 RESPOSTAS DOS PROBLEMAS 495

GLOSSÁRIO 501

BIBLIOGRAFIA 503

ÍNDICE 505

INTRODUÇÃO

“Para conhecer a ciência, é necessário conhecer sua história”
— Auguste Comte (1798-1857)

1.1 CONHECIMENTO HISTÓRICO

Com este comentário de Auguste Comte em mente, começamos este estudo introdutório aos sistemas de comunicação com um balanço histórico desta disciplina que nos rodeia diariamente de uma forma ou outra.¹ Cada subseção desta seção focaliza alguns eventos importantes e relacionados com a evolução histórica da comunicação.

Telégrafo

O telégrafo foi aperfeiçoadado por Samuel Morse, um pintor. Com as palavras “*What hath God wrought**”, transmitido pelo telégrafo elétrico de Morse entre Washington, D.C., e Baltimore, Maryland, em 1844, uma forma completamente revolucionária de comunicação a longa distância, em tempo real, foi iniciada. O telégrafo, idealmente adequado para chaveamento manual, é o precursor das comunicações digitais. Especificamente, o Código Morse é um código de *tamanho variável* usando um alfabeto de quatro símbolos: um ponto, um traço, a letra espaço e a palavra espaço. Seqüências curtas representam letras freqüentes, enquanto que seqüências longas representam letras não usuais.

Rádio

Em 1864, James Clerk Maxwell formulou a teoria eletromagnética da luz e previu a existência de ondas de rádio. O conjunto de equações básicas definidas por ele recebeu seu nome. A existência de ondas de rádio foi confirmada experimentalmente por Heirich Hertz em 1887. Em 1894, Oliver Lodge demonstrou a comunicação sem fio através de uma distância relativamente curta (aproximadamente 137 metros). Após estes eventos, em 12 de dezembro de 1901, Guglielmo Marconi recebeu um sinal de rádio em Signal Hill em Newfoudland. O sinal de rádio foi gerado em Cornwall, Inglaterra, 2736km através do Atlântico. O caminho estava, portanto, aberto a uma enorme ampliação das comunicações. Em 1906, Reginald Fessenden, um acadêmico autodidata, fez história ao ser o primeiro radialista em uma transmissão de rádio.

Em 1918, Edwin H. Armstrong inventou o receptor de rádio super-heteródino. Até hoje, quase todos os receptores de rádio são deste tipo. Em 1933, Armstrong demonstrou outro conceito revolucionário – chamado de esquema de modulação, o qual ele denominou *modulação em freqüência* (FM). O artigo de Armstrong apresentando o rádio FM foi publicado em 1936.

¹ Este conhecimento histórico é uma adaptação do livro de Haykin (2001).

* N. de T.: Em português, “Eis o que Deus realizou”.

Telefone

Em 1875, o telefone foi inventado por Alexander Graham Bell, um professor de surdos. O telefone torna a transmissão em tempo real da fala através da codificação elétrica e replicação do som uma realidade prática. A primeira versão do telefone era fraca e crua, permitindo que pessoas conversassem apenas por curtas distâncias. Quando o serviço de telefonia estava com apenas alguns anos surgiu o interesse em desenvolvê-lo. Notavelmente, em 1897, A. B. Strowger, um agente funerário da cidade de Kansas, Missouren, desenvolveu uma chave automática de comutação passo-a-passo, a qual recebeu seu nome. De todas as chaves eletromecânicas projetadas ao longo dos anos, a chave de Strowger é a mais popular e amplamente utilizada.

Componentes eletrônicos

Em 1904, John Ambrose Fleming inventou o diodo a tubo de vácuo, o que permitiu a invenção do triodo a tubo de vácuo por Lee de Forest em 1906. A descoberta do triodo foi o instrumental necessário para o desenvolvimento da telefonia transcontinental em 1913 e sinalizou o nascer das comunicações de voz sem fio. De fato, até a invenção e o aperfeiçoamento do transistor, o triodo era o dispositivo supremo para o projeto de amplificadores eletrônicos.

O transistor foi inventado em 1948 por Walter H. Brattain, John Bardeen e William Shockley nos laboratórios Bell. O primeiro circuito integrado (CI) de silício foi produzido por Robert Noyce em 1958. Estes marcos de inovações em dispositivos de estado sólido e circuitos integrados resultaram no desenvolvimento de circuitos VLSI (*Very Large Scale Integrated*) e microprocessadores em um único chip, e com eles a natureza do processamento de sinais e a indústria da telecomunicação mudaram para sempre.

Televisão

O primeiro sistema televisivo totalmente eletrônico foi demonstrado por Philo T. Farnsworth em 1928, e, após, Vladimir K. Zworykin, em 1929. Por volta de 1939, a *British Broadcasting Corporation* (BBC) estava transmitindo televisão em base comercial.

Comunicação digital

Em 1928, Harry Nyquist publicou um artigo clássico sobre a teoria da transmissão de sinal em telegrafia. Em particular, Nyquist desenvolveu um critério para a recepção correta de sinais telegráficos transmitidos em canais dispersivos na ausência de ruído. Grande parte do trabalho de Nyquist foi aplicado, posteriormente, na transmissão digital de dados em canais dispersivos.

Em 1937, Alex Reeve inventou a modulação por codificação de pulso (PCM*) para a codificação digital de sinais de fala. A técnica foi desenvolvida durante a Segunda Guerra Mundial para permitir a criptografia de sinais de fala. De fato, um sistema completo de 24 canais foi utilizado em campo pelo exército dos Estados Unidos no final da guerra. Entretanto, o PCM teve que esperar pela descoberta do transistor e consequente desenvolvimento de circuitos de alta escala de integração para ser explorado comercialmente.

A invenção do transistor em 1948 acelerou a aplicação de componentes eletrônicos no chaveamento e comunicações digitais. A motivação era melhorar a confiabilidade, aumentar a capacidade e reduzir os custos. A primeira chamada através de um sistema de programa armazenado foi realizada em março de 1958, nos Laboratórios Bell, e o primeiro serviço de telefonia comercial com chaveamento digital começou em Morris, Illinois, em junho de 1960. O primeiro sistema de transmissão de portadora T-1 foi instalado em 1962 nos laboratórios Bell.

* N. de T.: Pulse-code Modulation.

Em 1943, D. O. North projetou o filtro casado para a detecção ótima de um sinal conhecido, com ruído branco adicionado. Um resultado similar foi obtido independentemente em 1946 por J. H. Van Vleck e D. Middleton, os quais cunharam o termo *filtro casado*.

Em 1948, a fase teórica das comunicações digitais foi apresentada por Claude Shannon em um artigo intitulado *A Mathematical Theory of Communication*^{*}. O artigo de Shannon foi recebido com aclamação imediata e entusiasmada. Talvez seja esta resposta que encorajou Shannon a enfatizar o título de seu trabalho, alterando-o para *The Mathematical Theory of Communication* quando ele foi editado novamente, um ano depois, em um livro de co-autoria com Warren Weaver. É digno de nota que anteriormente à publicação do artigo clássico de Shannon em 1948, acreditava-se que aumentar a taxa de transmissão de informação em um canal aumentaria a probabilidade de erro. A comunidade da teoria da comunicação surpreendeu-se quando Shannon provou que isto não era verdade, desde que a taxa de transmissão estivesse abaixo da capacidade do canal.

Redes de computadores

Durante o período de 1943 a 1946, o primeiro computador digital eletrônico, chamado de ENIAC, foi construído na Escola Moore de Engenharia Elétrica da Universidade da Pensilvânia, sob direção técnica de J. Presper Eckert Jr. e John W. Mauchly. Entretanto, as contribuições de John von Neumann estavam dentre as primeiras e mais fundamentais para a teoria, projeto e aplicação de computadores digitais, reportando ao primeiro rascunho de um relatório escrito em 1945. Computadores e terminais iniciaram a comunicação em longa distância entre eles no começo da década de 1950. Os canais utilizados eram inicialmente canais de telefone de voz operando em baixas velocidades (300 a 1200 bps). Vários fatores contribuíram para o aumento dramático das taxas de transmissão, ressaltando a idéia de *equalização adaptativa*, cujo pioneiro foi Robert Lucky em 1965 além de eficientes técnicas de modulação inseridas por G. Ungerboeck em 1982. Outra idéia amplamente implementada em comunicações de computadores foi a chamada *requisição automática de repetição* (ARQ^{**}). O método ARQ foi originalmente proposto por H. C. A. van Duuren durante a Segunda Guerra Mundial e publicada em 1946. Ela era utilizada para melhorar a rádio-telefonia na transmissão de telex em longas distâncias.

De 1950 a 1970 vários estudos foram realizados sobre as redes de computadores. Entretanto, o estudo mais significativo em termos de impacto em comunicação de computadores foi a *Advanced Research Projects Agency Network* (ARPANET), colocada em serviço em 1971. O desenvolvimento da ARPANET foi financiado pela Agência de Pesquisa de Projetos Avançados do Departamento de Defesa dos EUA. O trabalho pioneiro em chaveamento de pacotes foi feito na ARPANET. Em 1985, a ARPANET foi renomeada para *Internet*. O ponto de virada na evolução da Internet ocorreu em 1990 quando Tim Berners-Lee propôs um software de interface de hipermídia para a Internet, o qual ele chamou de *World Wide Web*. Em apenas dois anos, a Web foi da não existência para a popularidade mundial, culminando em sua comercialização em 1994. Podemos explicar o crescimento explosivo da Internet através dos seguintes argumentos:

- ▶ Antes da Web explodir para a existência, os ingredientes para a sua criação já estavam dispostos. Em particular, graças ao VLSI, computadores pessoais (PCs) já haviam se tornado comuns nas casas ao redor do mundo e eram cada vez mais equipados com modems para a interconectividade ao mundo externo.
- ▶ Por aproximadamente duas décadas, a Internet crescia continuamente (apesar de estar confinada a uma comunidade ainda restrita de usuários).

* N. de T.: A teoria matemática da comunicação.

** N. de T.: *Automatic repeat-request*.

- Padrões para a descrição e transferência de documentos, linguagem de formatação de hipertextos (HTML) e protocolos de transferência de hipertexto (HTTP) já haviam sido adotados.

Portanto, tudo o que era necessário para a criação da Web já estava em seus devidos lugares, exceto por dois ingredientes críticos: uma interface simples com o usuário e um conceito brilhante de serviço.

Comunicações via satélite

Em 1955, John R. Pierce propôs a utilização de satélites para comunicações. Esta proposta foi precedida, entretanto, por um artigo de Arthur C. Clark, publicado em 1945, também propondo a idéia de utilizar um satélite orbitando a terra como ponto de repetição para a comunicação entre duas estações terrestres. Em 1957, a União Soviética lançou o Sputnik I, o qual transmitiu sinais de telemetria por 21 dias. Logo após houve o lançamento do Explorer 1, pelos Estados Unidos, em 1958, o qual transmitiu sinais de telemetria por cerca de 5 meses. Um grande passo experimental para a tecnologia de comunicação via satélite ocorreu com o lançamento do Telstar I, de Cabo Canaveral, em 10 de julho de 1962. O satélite Telstar foi construído pelos Laboratórios Bell, o qual adquiriu um considerável conhecimento do pioneiro trabalho de Pierce. O satélite foi capaz de retransmitir programas de TV através do Atlântico, possível apenas com o uso de receptores *maser* e grandes antenas.

Comunicações ópticas

A utilização de meios ópticos (por exemplo, fumaça e sinais de fogo) para a transmissão de informação data de tempos pré-históricos. Entretanto, nenhum grande marco nas comunicações ópticas surgiu até 1966, quando K. C. Kao e G. A. Hockham do *Standard Telephone Laboratories*, Reino Unido, propuseram a utilização de uma fibra de vidro recoberta como dielétrico de um guia de onda. O laser havia sido inventado e desenvolvido em 1959 e 1960. Kao e Hockham mostraram que (1) a atenuação em uma fibra óptica era devido a impurezas no vidro e (2) que a perda intrínseca, determinada pela dispersão de Rayleigh é muito baixa. De fato, eles previram que uma perda de 20 dB/Km poderia ser obtida. Esta memorável previsão feita quando a perda de potência em uma fibra de vidro era da ordem de 1000 dB/Km, teve que ser demonstrada posteriormente. Atualmente, perdas de transmissão tão baixas quanto 0,1 dB/Km são obtidas.

Os avanços espetaculares na microeletrônica, computadores digitais e sistemas ópticos, que testemunhamos e que continuarão a acontecer no futuro, são todos responsáveis por dramáticas mudanças no ambiente das telecomunicações. Muitas destas mudanças já estão funcionando e muitas outras ocorrerão ao longo do tempo.

1.2 APLICAÇÕES

O conhecimento histórico da Seção 1.1 permeia muitas das aplicações dos sistemas de comunicação, alguns dos quais são exemplificados pelo telégrafo que já deixou de funcionar há muito tempo, enquanto outras são exemplificadas pela Internet, de origem recente. No texto a seguir, iremos focar em rádio, redes de comunicação (exemplificados pelo telefone) e Internet, os quais dominam a forma pela qual nos comunicamos de uma dentre duas formas básicas, resumidas a seguir:

- *Broadcasting*: envolve a utilização de um único e potente transmissor e vários receptores que são relativamente baratos. Nesta classe de sistemas de comunicação, os sinais contendo a informação fluem para apenas uma direção, do transmissor para cada um dos receptores que estiverem na área de cobertura.
- Comunicação ponto-a-ponto: o processo de comunicação ocorre em um canal entre um único transmissor e um único receptor. Nesta segunda classe de sistemas de co-

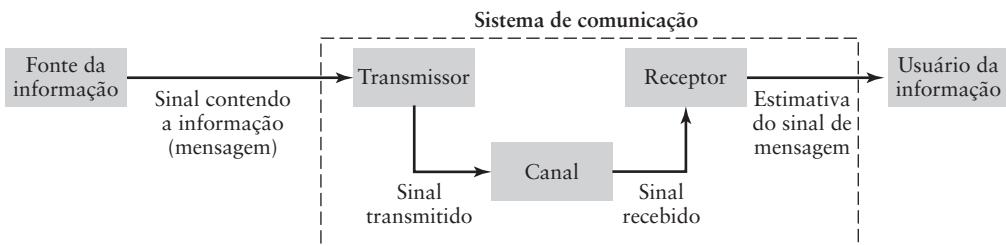


Figura 1.1 Elementos de um sistema de comunicação.

municação, existe, geralmente, um fluxo bidirecional de sinais contendo informação, requisitando, na realidade, o uso de um transmissor e um receptor (ou seja, um transreceptor) em cada ponta do canal.

O diagrama em blocos da Fig. 1.1 apresenta os componentes básicos de um sistema de comunicação. O *transmissor*, em alguma posição do espaço, converte o *sinal de mensagem* produzido pela *fonte de informação* em uma forma adequada para a transmissão no canal. O *canal*, por sua vez, transporta o sinal de mensagem e a entrega para o receptor em alguma outra localização no espaço. Entretanto, ao longo da transmissão no canal, o sinal é *distorcido* devido a imperfeições do canal. Além disso, ruído e sinais de interferência (originado por outras fontes) são adicionados à saída do canal, o que resulta na recepção de uma versão corrompida do sinal transmitido. O *receptor* possui a tarefa de operar no sinal recebido, de forma a produzir uma *estimativa* do sinal original da mensagem para o *usuário da informação*. Dizemos uma “estimativa” devido ao inevitável desvio, mesmo que pequeno, da saída do receptor quando comparado à entrada do transmissor, sendo este desvio atribuído às imperfeições do canal, aos ruídos e às interferências.

■ Rádio

Falando em termos genéricos, o rádio implementa tanto uma forma de broadcasting quanto uma comunicação ponto-a-ponto, dependendo de como ele é utilizado.

Tanto o rádio AM quanto o FM são comuns a todos nós. (AM representando modulação em amplitude e FM, modulação em freqüência). Os dois são construídos de forma integrada em uma única unidade e os encontramos em qualquer casa e instalado na maioria dos carros. Através do rádio escutamos as notícias locais, nacionais e internacionais, comentários, música e previsão do tempo, os quais são transmitidos por estações de broadcasting que operam em nossa vizinhança. Tradicionalmente, o rádio AM e FM são construídos utilizando circuitos eletrônicos analógicos. Entretanto, graças aos constantes melhoramentos e redução de custo de componentes eletrônicos digitais, o rádio digital (tanto AM quanto FM) já está em uso.

O rádio transmite a voz através de sinais elétricos. A *televisão*, a qual opera com os mesmos princípios da teoria de comunicação e eletromagnéticos, também transmite imagens visuais por sinais elétricos. Um sinal de voz é naturalmente definido como uma *função unidimensional do tempo*, o qual, portanto, resulta em operações simples de processamento de sinais. Por outro lado, uma imagem com movimento é uma *função bidimensional no tempo*, e requer uma atenção especial. Especificamente, cada imagem em um instante particular de tempo é vista como um quadro subdividido em um número de pequenos quadrados, chamados de *elementos de imagem* ou *pixels*. Quanto maior o número de pixels utilizado para representar a imagem, melhor a sua resolução. Varrendo os pixels em uma seqüência ordenada, a informação contida na imagem é convertida para um sinal elétrico cuja amplitude é proporcional ao nível de brilho dos pixels individuais. O sinal elétrico gerado na saída do circuito de varredura é o *sinal de vídeo* a ser transmitido. A geração do sinal de vídeo é o resultado de um *processo de mapeamento* bem definido, conhecido pelo receptor. Logo, dado um sinal de vídeo, o receptor é capaz de reconstruir a imagem

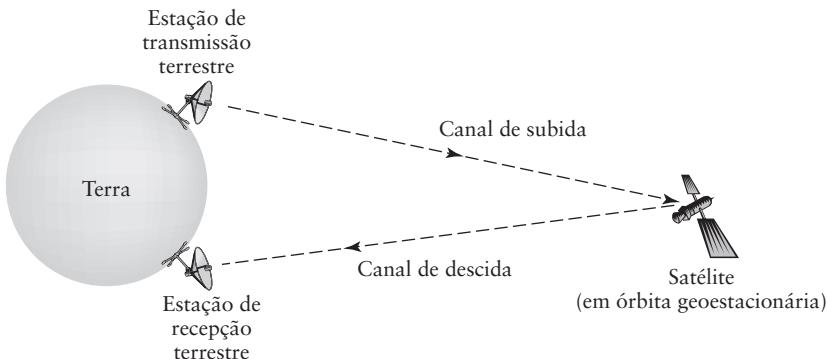


Figura 1.2 Sistema de comunicação via satélite.

original. Tal como no rádio digital, a televisão também foi beneficiada pelos avanços espetaculares dos componentes digitais. Estes avanços, juntamente com a aplicação de técnicas avançadas de processamento de sinal e à demanda dos consumidores, motivaram o desenvolvimento da chamada *televisão de alta definição* (HDTV), a qual apresenta uma melhora significativa na qualidade das imagens reconstruídas pelo receptor.

Vamos, agora, voltar nossas atenções para outro cenário de comunicação ponto-a-ponto. O rádio permeia nossas vidas cotidianas de forma significativa por dois caminhos distintos: a comunicação via satélite e comunicações sem fio. As *comunicações via satélite*, orquestradas por um satélite em órbita geoestacionária, é baseada em uma propagação de rádio por visada (ou linha) direta, para a operação de um canal de subida (*uplink*) e um canal de descida (*downlink*). O canal de subida conecta o terminal terrestre a um transponder (isto é, um circuito eletrônico) a bordo do satélite, enquanto que o canal de descida conecta o transponder a outro terminal terrestre. Portanto, o sinal contendo a informação é transmitido do terminal terrestre para o satélite pelo canal de subida, é amplificado pelo transponder e, então, retransmitido do satélite pelo canal de descida para outro terminal terrestre, como ilustrado na Fig. 1.2. Desta forma, o sistema de comunicação via satélite oferece uma possibilidade única: *cobertura global*.

De forma genérica, a *comunicação sem fio* opera de maneira similar a comunicação via satélite, no sentido de que ela também envolve um canal de descida e um de subida. O canal de descida é responsável pela transmissão direta de rádio de uma *estaçao base* para o usuário móvel. O canal de subida é responsável pelo caminho inverso, do usuário móvel para a sua estação base. Ao contrário da comunicação via satélite, a operação de comunicação sem fio é dominada pelo *fenômeno multi-caminho*, devido a reflexões do sinal transmitido por objetos (por exemplo, prédios, árvores, etc.) que estão no caminho da propagação. Este fenômeno tende a degradar a performance do receptor, o que torna o projeto do receptor um desafio. Em qualquer evento, a comunicação sem fio oferece uma capacidade própria única: *mobilidade*. Além disso, através do uso do conceito celular, o sistema de comunicação sem fio é capaz de reutilizar o espectro de rádio em uma grande área quantas vezes for possível. Dentro de uma célula, os recursos disponíveis para a comunicação podem ser compartilhados pelos usuários móveis operando dentro da célula.

■ Redes de comunicação

O computador foi, originalmente, projetado como uma máquina para trabalhar isoladamente, executando cálculos numéricos. Entretanto, dada a habilidade natural de um computador executar funções lógicas, reconheceu-se logo que o computador é idealmente adequado para o projeto de redes de comunicação. Como ilustrado na Fig. 1.3, uma rede de comunicação é constituída pela interconexão de um número de roteadores compostos por processadores inteligentes (por exemplo, microprocessadores). O objetivo primário destes

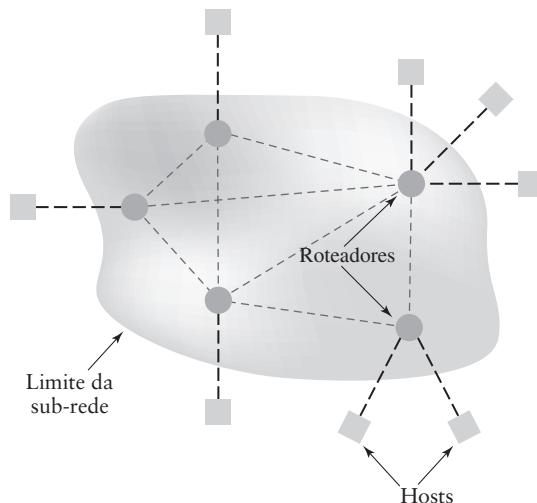


Figura 1.3 Rede de comunicação.

processadores é rotear voz ou dados através da rede, por isso o nome “roteador”. Cada roteador possui um ou mais *hosts* associados a ele. Por *host* nos referimos a dispositivos que comunicam entre si. O propósito da rede é possibilitar a entrega ou troca de voz, vídeo ou dados entre os hosts, o que é possível devido à utilização de chaveamento digital. Existem duas formas principais de chaveamento: chaveamento por circuito e chaveamento por pacote.

No chaveamento por circuito, caminhos dedicados para a comunicação são estabelecidos para a transmissão de mensagens entre dois ou mais terminais, chamados de estações. O caminho de comunicação, ou circuito, é constituído por uma seqüência de canais conectados da fonte ao destino. Por exemplo, os canais podem ser constituídos por slots de tempo (tal como em sistemas multiplexados por divisão no tempo), no qual um canal comum é disponibilizado para múltiplos usuários. O ponto importante é que, uma vez estipulado, o circuito permanece ininterrupto durante toda a transmissão. O chaveamento por circuito é geralmente controlado por um mecanismo de controle hierárquico centralizado, com o conhecimento de toda a organização da rede. Para estabelecer uma conexão com chaveamento por circuito, um caminho disponível pela rede de telefonia é reservado e, então, dedicado ao uso exclusivo de dois usuários que querem comunicar entre eles. Em particular, um sinal de requisição de chamada propaga todo o caminho até o destino, no qual ele é aceito antes que a comunicação possa começar. Então, a rede é efetivamente transparente aos usuários, o que significa que durante todo o tempo de conexão, os recursos alocados ao circuito essencialmente “pertencem” aos dois usuários. Este estado de conexão permanece até que o circuito seja desconectado.

O chaveamento por circuito é adequado para redes de telefonia, nas quais a transmissão de voz constitui o corpo principal de tráfego na rede. Dizemos isso porque a voz aumenta o fluxo do tráfego e conversações tendem a ser longas (2 minutos na média), comparado com o tempo necessário para reservar o circuito (cerca de 0,1 a 0,5 segundos).

No chaveamento por pacote², por outro lado, o compartilhamento dos recursos da rede é feito com base na demanda. Logo, o chaveamento por pacote possui uma vantagem sobre o chaveamento por circuito no sentido de que quando um canal possui tráfego a ser

² O chaveamento por pacote foi inventado por P. Baran em 1964 para satisfazer uma necessidade da defesa nacional dos EUA. A necessidade original era construir uma rede distribuída com diferentes níveis de conexões redundantes, a qual é robusta no sentido de que a rede pode resistir à destruição de vários nós devido a um plano de ataque, com os nós sobreviventes sendo capazes de manter a intercomunicação transportando informações comuns e de controle; veja Baran (1990).

enviado, ele tende a ser mais totalmente utilizado. Ao contrário de sinais de voz, dados tendem a ocorrer na forma de rajadas ocasionais.

O princípio da rede de chaveamento por pacote é o armazenamento e envio. Especificamente, em uma rede por chaveamento de pacote, qualquer mensagem maior do que um tamanho especificado é subdividida antes da transmissão, em segmentos que não sejam maior do que um tamanho especificado. Os seguimentos formados são chamados de pacotes. Após o transporte dos pacotes ao longo de diferentes caminhos na rede, a mensagem original é remontada no destino, pacote a pacote. A rede pode, portanto, ser vista como uma piscina de recursos de rede (isto é, largura de faixa do canal, buffers e processadores de chaveamento), com os recursos dinamicamente compartilhados com a comunidade de hosts que querem comunicar competindo. Este compartilhamento dinâmico dos recursos da rede é um contraste direto com a rede chaveada por circuito, na qual os recursos são dedicados a um par de hosts durante todo o período em que eles estiverem comunicando.

■ Redes de dados

Uma rede de comunicação na qual os hosts são todos compostos de computadores e terminais é geralmente chamada de *rede de dados*. O projeto de tal rede é realizado analisando a rede em termos de uma arquitetura de camadas, a qual é considerada como uma hierarquia de camadas sobrepostas. Uma camada se refere a um processo ou dispositivo dentro de um sistema computacional, projetado para executar uma função específica. Naturalmente, os projetistas de uma camada terão o conhecimento dos detalhes e operação internos. Ao nível de sistema, entretanto, um usuário enxerga a camada em questão simplesmente como uma “caixa preta”, a qual é descrita em termos de entradas, saídas e a função de relação entre as saídas e entradas. Em uma arquitetura de camadas, cada camada trata a próxima camada de nível mais baixo como uma ou mais caixas pretas com alguma especificação funcional a ser utilizada pela camada de nível mais alto. Desta forma, o problema de comunicação altamente complexo em redes de computadores é resolvido como um conjunto gerenciável de funções interconectadas bem conhecidas. Esta é a linha de pensamento que resultou no desenvolvimento do modelo de referência de interconexão de sistemas abertos (OSI*)³. O termo “aberto” se refere à capacidade de quaisquer dois sistemas serem interconectados, desde que eles estejam em conformidade com o modelo de referência e seus padrões associados.

No modelo de referência OSI, as comunicações e as funções relacionadas à conexão são organizadas por uma série de camadas com interface bem definidas. Cada camada é construída sobre sua predecessora. Em particular, cada camada executa um subconjunto relacionado de funções primitivas e confia que a próxima camada mais baixa irá executar funções primitivas adicionais. Além disso, cada camada oferece certos serviços à próxima camada de nível mais alto e protege aquela camada dos detalhes de implementação daqueles serviços. Entre cada par de camadas existe uma interface que define os serviços oferecidos pela camada inferior à camada superior.

Como ilustrado na Fig. 1.4, o modelo OSI é composto por sete camadas. A figura também inclui a descrição das funções das camadas individuais do modelo. A camada k do sistema A , digamos, comunica com a camada R de algum outro sistema B de acordo com um conjunto de regras e convenções, as quais, coletivamente, constituem o protocolo da camada k , onde $k = 1, 2, \dots, 7$. (O termo “protocolo” foi emprestado do uso comum que descreve a convenção social de comportamento entre seres humanos.) As entidades que constituem as camadas correspondentes em sistemas diferentes são chamadas de *processos de ponta*. Em outras palavras, a comunicação entre o sistema A e o sistema B é alcançada

* N. de T.: *Open systems interconnection*.

³ O modelo de referência OSI foi desenvolvido por um subcomitê da ISO (International Organization for Standardization) em 1977. Para a discussão dos princípios envolvidos na definição das sete camadas originais do modelo OSI e descrição das camadas, veja Tannenbaum (1996).

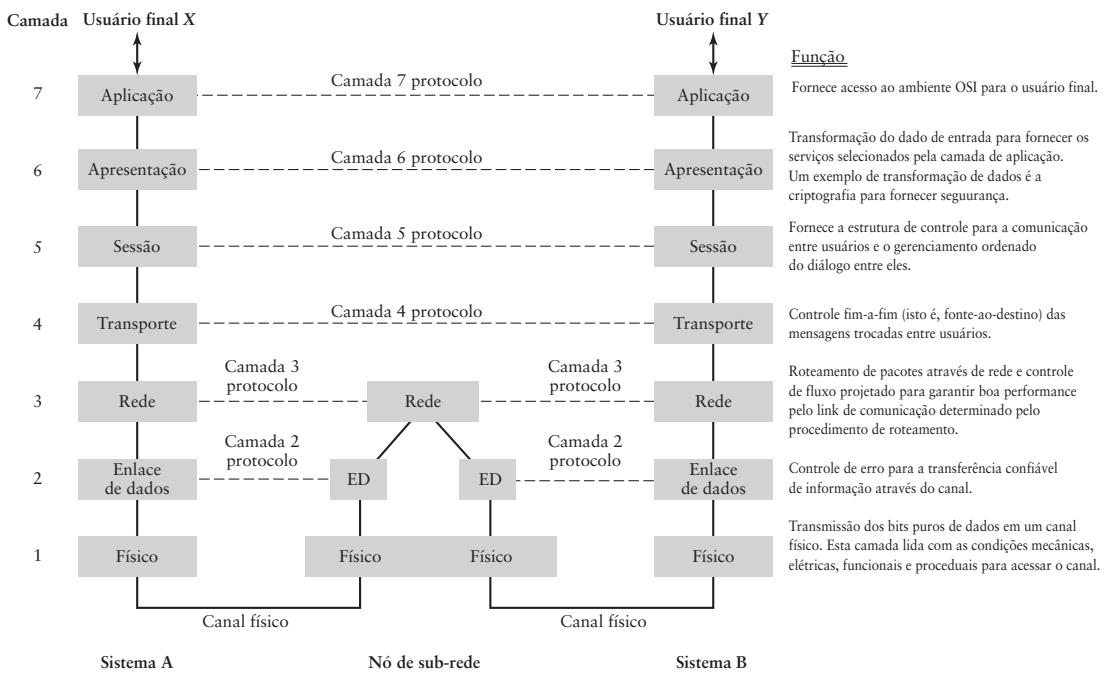


Figura 1.4 Modelo OSI. O acrônimo ED no meio da figura representa *Enlace de dados*.

tendo processos de ponta nos dois sistemas comunicando via protocolos. A conexão física entre os processos de ponta existe apenas na camada 1, chamada de *camada física*. As camadas restantes, de 2 a 7 estão em *comunicação virtual* com seus pontos distantes. Cada uma destas últimas seis camadas troca informação de dados e controle com suas camadas vizinhas (superior e inferior) através de interfaces camada-camada. Na Fig. 1.4, a comunicação física é mostrada por linhas sólidas e comunicações virtuais por linhas tracejadas.

■ Internet⁴

A discussão de redes de dados apresentada resulta na *Internet*. No paradigma da Internet, a tecnologia de base da rede é desacoplada das aplicações pela adoção de uma definição abstrata de rede de serviço. Em termos mais específicos, podemos dizer que:

- ▶ As aplicações são executadas independentemente da tecnologia adotada para a construção da rede.
- ▶ Pelo mesmo motivo, a tecnologia da rede é capaz de evoluir sem afetar as aplicações.

A aplicação da Internet descrita na Fig. 1.5 possui três blocos funcionais: hosts, sub-redes e roteadores. Os hosts constituem nós na rede, nos quais os dados são originados ou para onde eles são entregues. Os roteadores constituem nós intermediários utilizados para cruzar os limites das sub-redes. Dentro de uma sub-rede, todos os hosts pertencentes a ela trocam dados diretamente. Veja, por exemplo, as sub-redes 1 e 3 da Fig. 1.5. Em termos elementares, a operação interna de uma sub-rede é organizada de duas formas diferentes (Tanenbaum, 1996):

⁴ Para uma história fascinante da Internet, sua evolução histórica a partir da ARPANET e padrões internacionais, veja Abbate (2000). Para artigos de fácil leitura na internet, veja a edição especial da *IEEE Communications Magazine* (2002). Os artigos apresentados são escritos pelos pioneiros no desenvolvimento da Internet.

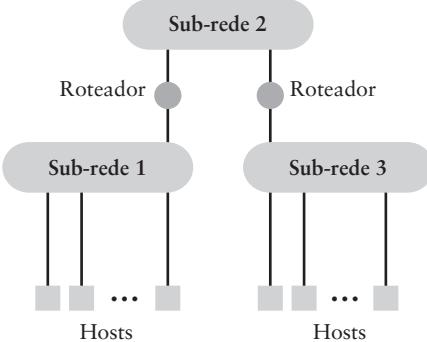
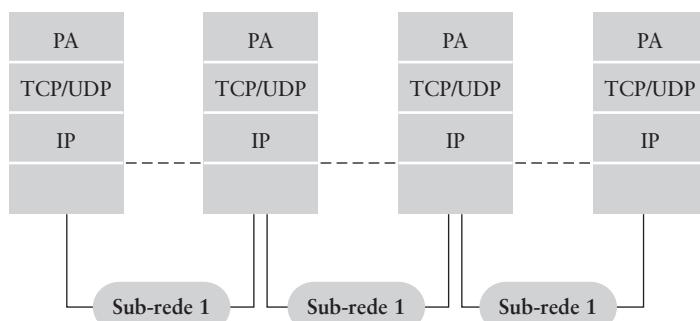


Figura 1.5 A conexão de uma rede de sub-redes.

1. Maneira conectada, na qual as conexões são chamadas de circuitos virtuais, em analogia com os circuitos físicos em um sistema de telefonia.
2. Maneira não conectada, na qual pacotes independentes são chamados de datagramas, em analogia com os telegramas.

Tal como outras redes de dados, a Internet possui um conjunto de camadas de protocolos. Em particular, a troca de dados entre os hosts e os roteadores é realizada através do *Internet Protocol* (IP), como ilustrado na Fig. 1.6. O IP é um protocolo universal que reside na camada de rede (isto é, a camada 3 do modelo de referência OSI). Ele é simples, definindo um plano de endereçamento com capacidade intrínseca de transportar dados na forma de pacotes de um nó para outro. Ao cruzar os limites de uma sub-rede, os roteadores tomam decisões de como os pacotes endereçados a um destino específico devem ser roteados. Isto é realizado com base em tabelas de roteamento desenvolvidas através da utilização de protocolos específicos para a troca de informação com outros roteadores. O resultado final da utilização de um conjunto de camadas de protocolos é o fornecimento do melhor esforço de serviço. Ou seja, a Internet oferece a capacidade de entregar cada pacote de dados, mas não existem garantias do tempo de trânsito necessário para a entrega ou mesmo se os pacotes serão entregues ao destino pretendido.

A Internet evoluiu para um sistema mundial, colocando os computadores no coração das mídias de comunicação, alterando profundamente nosso dia-a-dia em casa e no trabalho. Podemos mandar uma mensagem de e-mail de um host no Brasil para outro host no Japão, no outro lado do globo, com a mensagem chegando em seu destino em questão de segundos. Isto é ainda mais surpreendente porque os pacotes que constituem a mensagem muito provavelmente passaram por caminhos diferentes enquanto eram transportados ao redor do mundo.



PA: Protocolo de aplicação

TCP: Transmission control protocol

UDP: User datagram protocol

IP: Internet protocol

Figura 1.6 Ilustração da arquitetura de rede da Internet.

Outra aplicação que demonstra o poder surpreendente da Internet é a sua utilização para surfar na Web. Por exemplo, podemos utilizar uma ferramenta de busca para identificar as referências pertencentes a um assunto particular de interesse. Uma tarefa que antes demandava horas e algumas vezes dias de pesquisa em livros e jornais na biblioteca agora é realizada em questão de segundos!

Para utilizarmos totalmente o poder de computação da Internet de um host localizado em uma região remota, precisamos de um modem (isto é, um modulador-demodulador) de banda larga para possibilitar um canal de conexão rápido entre o host e sua sub-rede. Pelo termo “rápido”, entendemos velocidade de operação na ordem de megabits por segundo ou mais. Um dispositivo que satisfaz esta condição é o chamado DSL (*Digital Subscriber Line*). O que torna o DSL surpreendente é o fato dele operar em um canal de banda larga linear com resposta em frequência arbitrária. Tal canal é exemplificado por um canal ordinário de telefonia, construído usando par trançado para a transmissão do sinal. Um par trançado é constituído por dois condutores sólidos de cobre, cada um encapsulado por uma camada de PVC (*polyvinyl chloride*). Pares trançados são geralmente utilizados em cabos, com cada cabo constituído por vários pares trançados próximos uns dos outros. Do ponto de vista de transmissão de sinal, o DSL satisfaz a condição desafiadora descrita anteriormente pelo bem conhecido princípio de *dividir e conquistar*. Especificamente, o canal de banda larga dado é aproximado por um conjunto de canais próximos, cada um deles podem, então, ser acomodado de forma relativamente direta.

Um último comentário é necessário. Normalmente, o acesso à Internet é estabelecido via hosts na forma de terminais de computadores (isto é, servidores). O acesso é expandido usando dispositivos de mão que atuam como hosts, os quais comunicam com as sub-redes da Internet via canais sem fio. Portanto, pela adição da mobilidade através do uso de comunicação sem fio ao poder de computação da Internet, temos uma nova mídia de comunicação com enormes possibilidades práticas.

■ Integração do telefone com a Internet

Um dos importantes desafios relacionados com a indústria das telecomunicações é a transmissão de voz sobre IP (VoIP), o que tornaria possível a integração dos serviços de telefonia com as aplicações baseadas na Internet com rápido crescimento. O desafio é mais difícil porque o IP é projetado para acomodar a troca de dados entre hosts e roteadores, dificultando a implementação de suporte à qualidade de serviço para VoIP. A qualidade de serviço (QoS – Quality of Service) é medida em termos de dois parâmetros:

- ▶ Razão de perda de pacote, definida como o número de pacotes perdidos no transporte ao longo da rede para o número total de pacotes enviados na rede.
- ▶ Atraso de conexão, definida como o tempo necessário para um pacote de uma conexão particular ponto-a-ponto ser transmitido na rede.

Testes subjetivos executados em VoIP mostraram que para fornecer o serviço de telefonia a nível de voz, a perda de pacotes deve ser mantida abaixo de 1%, e o atraso de conexão em uma direção pode acumular até 160 ns sem uma degradação significativa na qualidade. Redes de VoIP bem gerenciadas e bem projetadas, satisfazendo estas condições, estão sendo desenvolvidas. Entretanto, a questão de controle inicial de eco permanece um desafio.⁵ O eco inicial se refere ao eco observado no começo de uma ligação, na primeira palavra ou par de palavras fora da boca do usuário. O eco aparece devido a uma falta de casamento de impedância em algum lugar na rede, no qual o sinal incidente é refletido de volta à fonte.

Olhando para o futuro, podemos fazer as seguintes considerações sobre a telefonia na Internet:

⁵ Os limites das grandezas de QoS mencionadas aqui foram retiradas de um artigo de Janes, Chen e Garrison (2004), publicado em uma edição especial da IEEE Communications Magazine, dedicada a VoIP e qualidade de serviço.

1. VoIP irá substituir os PABXs (*private branch exchanges*) e outras centrais de escritório. PABXs são unidades de chaveamento remotas que possuem seus próprios controles independentes.⁶
2. VoIP já está apresentando sucesso com chamadas de longa distância, mas isto ocorre principalmente devido ao excesso de capacidade que agora está disponível em redes de longa distância. Se a carga nestas redes aumentar, os atrasos também aumentarão e um serviço em tempo real, tal como VoIP, será degradado. Por outro lado, se os provedores de serviços longos continuarem a adicionar capacidade, de tal forma que a carga sempre será pequena e o tempo de resposta rápido, então, garantindo a qualidade de serviço, a telefonia VoIP se tornará o principal meio de comunicação e amplamente distribuído.

■ Armazenamento de dados

Quando consideramos aplicações importantes dos princípios da comunicação digital, é natural pensarmos em termos de sistemas de comunicação por broadcasting ou ponto-a-ponto. Entretanto, os mesmos princípios também são aplicados ao armazenamento digital de sinais de áudio e vídeo, exemplificados pelo *compact disc* (CD) e *digital versatile disc* (DVD). DVDs são refinamentos de CDs no sentido de sua capacidade de armazenamento (na ordem de dezenas de gigabytes) ser algumas ordens de grandeza maior do que a de CDs, além deles poderem entregar dados em uma taxa muito mais alta.

O domínio digital é preferido com relação ao domínio analógico para o armazenamento de áudio e vídeo pelas seguintes razões:

- (i) A qualidade de um sinal de áudio/vídeo digitalizado, medido em termos da resposta em frequência, linearidade e ruído é determinada pelo processo de conversão digital para analógico (DAC), um parâmetro que está sob o controle do projetista.
- (ii) Uma vez que o sinal de áudio/vídeo tenha sido digitalizado, podemos utilizar técnicas bem conhecidas e poderosas de codificação para a compressão de dados, reduzindo a largura de faixa e codificando o controle de erro fornecendo proteção contra a possibilidade de erros ao longo do armazenamento.
- (iii) Para a maioria das aplicações práticas, o armazenamento digital de sinais de áudio e vídeo não é degradado ao longo do tempo.
- (iv) Melhorias contínuas na fabricação de circuitos integrados utilizados para fabricar CDs e DVDs garantem custo sempre decrescente destes dispositivos de armazenamento.

Com a ajuda de poderosas técnicas de codificação inseridas nos projetos, DVDs podem armazenar horas de conteúdo áudio-visual de alta qualidade, os quais, por sua vez, tornam-o idealmente adequado para aplicações de multimídia interativas.

1.3 RECURSOS PRIMÁRIOS E CONDIÇÕES OPERACIONAIS

Os sistemas de comunicação descritos na Seção 1.2 abrangem várias áreas. De qualquer forma, dentro de seus próprios caminhos, os sistemas são projetados para fornecer uma utilização eficiente de dois recursos primários de comunicação:

- Potência transmitida, a qual é definida como a potência média do sinal transmitido.
- Largura de faixa do canal, a qual é definida como o comprimento da faixa passante do canal.

Dependendo de como estes dois recursos são considerados como fatores limitadores, podemos classificar os canais de comunicação como:

⁶ PABXs são discutidos em McDonald (1990).

- (i) Canal limitado em potência, no qual a potência transmitida é o prêmio buscado. Exemplos de tais canais incluem:
- ▶ Canais sem fio, nos quais é desejável manter a potência transmitida o mais baixo possível para prolongar o tempo de uso da bateria dos dispositivos.
 - ▶ Canais de satélite, nos quais a potência disponível a bordo do transponder do satélite é limitada, a qual, por sua vez, necessita manter a potência transmitida no canal de descida em um nível baixo.
 - ▶ Conexão de espaço profundo, no qual a potência disponível a bordo de uma sonda explorando o espaço é extremamente limitada, o que, novamente, requer que a potência média dos sinais de informação enviados da sonda para a estação em terra seja mantida a mais baixa possível.
- (ii) Canais limitados em faixa, nos quais a largura de faixa do canal é o premio buscado. Exemplos desta segunda categoria de canais de comunicação incluem:
- ▶ Canais de telefonia, nos quais, em um ambiente multiusuário, a condição é minimizar a faixa de freqüência alocada para a transmissão de cada sinal de voz, enquanto a qualidade de serviço é garantida para cada usuário.
 - ▶ Canais de televisão, nos quais a largura de faixa do canal é limitada por agências de regulamento e a qualidade de recepção é garantida usando uma potência de transmissão suficientemente alta.

Outro ponto importante a se ter em mente é a presença inevitável de ruído na entrada do receptor de um sistema de comunicação. De forma genérica, ruído se refere a sinais indesejados que tendem a perturbar a qualidade do sinal recebido em um sistema de comunicação. A fonte do ruído pode ser interna ou externa ao sistema. Um exemplo de ruído interno é o constante ruído de canal produzido pela agitação térmica de elétrons na frente do amplificador do receptor. Exemplos de ruído externos incluem ruído atmosférico e interferência devido a sinais transmitidos pertencentes a outros usuários.

Uma forma quantitativa de contabilizar o efeito benéfico da potência transmitida em relação ao efeito degradante do ruído (isto é, avaliar a qualidade do sinal recebido) é realizada em termos da relação sinal/ruído (RSR), a qual é um parâmetro sem dimensão. Em particular, a RSR na entrada do receptor é formalmente definida como a razão da potência média do sinal recebido (isto é, a saída do canal) pela potência média do sinal medido na entrada do receptor. A prática comum é expressar a RSR em decibel (dB), a qual é definida como 10 vezes o logaritmo (base 10) da relação de potência.⁷ Por exemplo, RSRs de 10, 100 e 1000 são 10, 20 e 30 dBs, respectivamente.

A luz desta discussão, agora é aparente que, considerando-se a análise de performance, existem apenas dois parâmetros de projeto do sistema: relação sinal/ruído e largura de faixa do canal. Dito em termos mais concretos:

O projeto de um sistema de comunicação é resumido em uma relação entre a relação sinal/ruído e a largura de faixa do canal.

Portanto, podemos melhorar a performance do sistema através de uma dentre duas alternativas de estratégia de projeto, dependendo das restrições do sistema:

1. A relação sinal/ruído é aumentada para acomodar a limitação imposta pela largura de faixa do canal.
2. A largura de faixa do canal é aumentada para acomodar a limitação imposta pela relação sinal/ruído.

Com estas duas possíveis abordagens de projeto, podemos determinar ordinariamente que a estratégia 1 é mais simples de ser implementada do que a estratégia 2, porque o aumento da relação sinal/ruído pode ser realizado simplesmente com o aumento da po-

⁷ Para uma discussão sobre decibel, veja o Apêndice 1.

tência transmitida. Por outro lado, para explorar o aumento da largura de faixa do canal, precisamos aumentar a largura de faixa do sinal transmitido, o que, por sua vez, requer o aumento da complexidade tanto do transmissor quanto do receptor.

1.4 A BASE DAS TEORIAS DE SISTEMAS DE COMUNICAÇÃO

O estudo de sistemas de comunicação é desafiador não somente em termos técnicos, mas também em termos teóricos. Nesta seção iremos ressaltar quatro teorias, cada uma delas é essencial para o entendimento de um aspecto específico de sistemas de comunicação.⁸

■ Teoria da modulação

Modulação é uma operação de processamento de sinal básica à transmissão de sinais de informação em um canal de comunicação, seja no contexto de comunicação digital ou analógica. Esta operação é realizada alterando algum parâmetro da onda portadora de acordo com o sinal contendo a informação (mensagem). A portadora pode assumir uma de duas formas básicas, dependendo da aplicação em interesse:

- ▶ Portadora senoidal, cuja amplitude, fase ou freqüência é o parâmetro escolhido para ser modificado pelo sinal de informação.
- ▶ Seqüência periódica de pulsos, cuja amplitude, largura ou posição é o parâmetro escolhido para ser modificado pelo sinal de informação.

Independentemente da abordagem particular utilizada para executar o processo de modulação, os pontos importantes da teoria da modulação que precisam ser apresentados são:

- ▶ Descrição no domínio do tempo do sinal modulado.
- ▶ Descrição no domínio da freqüência do sinal modulado.
- ▶ Detecção do sinal de informação original e determinação do efeito do ruído no receptor.

■ Análise de Fourier

A transformada de Fourier é uma operação matemática linear que transforma uma descrição no domínio do tempo de um sinal em uma descrição no domínio da freqüência, sem perda de informação, o que significa que o sinal original pode ser totalmente recuperado da descrição no domínio da freqüência. Entretanto, para um sinal poder ser transformado por Fourier, certas condições precisam ser satisfeitas. Felizmente, estas condições são satisfeitas pelo tipo de sinais encontrados no estudo dos sistemas de comunicação.

A análise de Fourier fornece a base matemática para a determinação das seguintes características:

- ▶ Descrição no domínio da freqüência de um sinal modulado, incluindo sua largura de faixa de transmissão.
- ▶ Transmissão de um sinal através de um sistema linear exemplificado por um canal de comunicação ou um filtro (seletivo em freqüência).
- ▶ Correlação (isto é, similaridade) entre um par de sinais.

A determinação destas características se torna ainda mais importante devido a um algoritmo chamado de transformada rápida de Fourier (FFT), o qual fornece um método eficiente para o cálculo da transformada de Fourier.

⁸ Uma outra teoria – chamada de Teoria da Informação – é básica ao estudo de sistemas de comunicação. Esta teoria não foi incluída porque ela é altamente matemática e, portanto, de natureza avançada, o que a torna inapropriada para um livro introdutório.

■ Teoria de detecção

Dado um sinal recebido, o qual é perturbado por um ruído aditivo do canal, uma das tarefas que o projetista do receptor precisa implementar é como determinar o sinal de informação original de maneira realizável. O problema de detecção de sinal é complicado por duas questões:

- ▶ Presença do ruído.
- ▶ Fatores tais como o desconhecimento do deslocamento de fase introduzido pela portadora devido à transmissão de sinais modulados senoidalmente no canal.

Lidar com estas características em comunicação analógica é radicalmente diferente de lidar com elas em comunicações digitais. Em comunicações analógicas, a abordagem normal é focalizar na relação sinal/ruído de saída, além de cálculos relacionados. Em comunicações digitais, por outro lado, o problema de detecção de sinal é visto como um teste de hipótese. Por exemplo, no caso específico de uma transmissão binária de dados, dado que o símbolo binário 1 é transmitido, qual é a probabilidade deste símbolo ser detectado corretamente e como a probabilidade é afetada alterando-se a relação sinal/ruído recebida na entrada do receptor?

Portanto, ao se trabalhar com a teoria de detecção, lidamos com as seguintes questões em comunicações analógicas:

- ▶ A figura de mérito para avaliar a performance do ruído em uma estratégia de modulação específica.
 - ▶ O fenômeno de limiar que aparece quando a relação sinal/ruído transmitida cai abaixo de um valor crítico.
 - ▶ Comparação de performance de uma estratégia de modulação em relação a outra.
- Em comunicações digitais, por outro lado, olhamos para:
- ▶ A probabilidade média de um erro de símbolo na saída do receptor.
 - ▶ O problema de lidar com fatores não-controláveis.
 - ▶ Comparação de um esquema de modulação digital com relação a outro.

■ Teoria da probabilidade e processos aleatórios

A partir da rápida discussão apresentada anteriormente sobre o papel da teoria de detecção no estudo de sistemas de comunicação, fica evidente que precisamos desenvolver um bom conhecimento sobre:

- ▶ Teoria da probabilidade para a descrição do comportamento de eventos ocorrendo aleatoriamente, em termos matemáticos.
- ▶ Caracterização estatística de sinais e ruídos aleatórios.

Ao contrário de um sinal determinístico, um sinal aleatório é um sinal no qual existe alguma incerteza antes dele ocorrer. Devido a esta incerteza, um sinal aleatório pode ser visto como pertencente a um grupo de sinais, com cada sinal do grupo tendo uma forma de onda diferente dos outros. Além disso, cada sinal dentro do grupo possui certa probabilidade de acontecer. O grupo de sinais é chamado de processo aleatório ou processo estocástico. Exemplos de processos aleatórios incluem:

- ▶ Ruído elétrico gerado na frente do amplificador de um receptor de rádio ou televisão.
- ▶ Sinal de voz produzido por um locutor masculino ou feminino.
- ▶ Sinal de vídeo transmitido pela antena de uma estação de TV.

Ao lidarmos com a teoria da probabilidade, sinais e ruídos aleatórios, lidamos com as seguintes características:

- ▶ Conceitos básicos da teoria da probabilidade e modelos probabilísticos.

- Descrição estatística de um processo aleatório em termos do grupo, além de médias temporais.
- Análise matemática de processamento de sinais aleatórios.

1.5 NOTAS DE CONCLUSÃO

Neste capítulo, fizemos um balanço histórico e apresentamos aplicações de comunicações além de uma breve apresentação das teorias básicas de sistemas de comunicação. Além disso, apresentamos os seguintes pontos para suportar nosso ponto de vista de que a disciplina é tanto altamente desafiadora quanto extremamente excitante:

- (i) Sistemas de comunicação englobam várias e diversas aplicações: rádio, televisão, comunicação sem fio, comunicação via satélite, comunicação em espaço profundo, telefonia, redes de dados e Internet, dentre outras.
- (ii) A comunicação digital tem se estabelecido como a forma dominante de comunicação. Muito do progresso que temos testemunhado no avanço dos sistemas de comunicação podem ser rastreados até certas teorias e tecnologias, resumidas a seguir:
 - Idéias matemáticas abstratas que são altamente relevantes para o conhecimento profundo de sinais de informação e sua transmissão através de mídias físicas.
 - Algoritmos de processamento digital de sinal para cálculo eficiente do espectro, correlação e filtragem de sinais.
 - Desenvolvimento de software e novas arquiteturas para o projeto de microprocessadores.
 - Avanços espetaculares na física de dispositivos de estado sólido e fabricação de chips VLSI.
- (iii) O estudo de sistemas de comunicação é uma disciplina dinâmica, em contínua evolução explorando novas inovações tecnológicas e outras disciplinas, e respondendo a novas necessidades sociais.
- (iv) Por fim, sistemas de comunicação permeiam nossa vida diária, tanto em casa como no trabalho, e nossa vida seria muito mais difícil sem a grande disponibilidade de dispositivos de comunicação que não damos o devido valor.

O restante do livro, dividido em dez capítulos, fornece um tratamento introdutório em sistemas de comunicação analógicos e digitais. O livro deve preparar o leitor para aprofundar seu conhecimento de uma disciplina que é melhor descrita como sendo quase sem limite de escopo. Este é especialmente o caso na tendência de unificação de redes cabeadas e sem fio, para acomodar a integração de transmissão de voz, vídeo e dados.

REPRESENTAÇÃO DE FOURIER DE SINAIS E SISTEMAS

Em termos matemáticos, um sinal é ordinariamente descrito como uma função do tempo, a qual é como nós geralmente vemos o sinal quando sua forma de onda é mostrada em um osciloscópio. Entretanto, como mostrado no Capítulo 1, da perspectiva de sistemas de comunicação, é importante que conheçamos o *conteúdo de freqüência* do sinal em estudo. A ferramenta matemática que relaciona a descrição no domínio da freqüência de um sinal com a descrição no seu domínio do tempo é a *Transformada de Fourier*. Existe, de fato, diversas versões da transformada de Fourier disponíveis. Neste capítulo iremos manter nossas discussões em, principalmente, duas versões:

- ▶ A transformada contínua de Fourier, ou, simplesmente, transformada de Fourier (TF), a qual trabalha com funções contínuas tanto no tempo quanto na freqüência.
- ▶ A transformada discreta de Fourier, ou TDF, a qual trabalha com dados discretos tanto no domínio do tempo quanto no domínio da freqüência.

Muito do material apresentado neste capítulo focaliza a transformada de Fourier, pois a motivação primária deste capítulo é determinar o conteúdo de freqüência de um sinal contínuo no tempo ou, então, calcular o que acontece com este conteúdo de freqüência quando o sinal passa através de um sistema linear invariante no tempo (LIT). Em contraste, a transformada discreta de Fourier, discutida no final do capítulo, tem sua importância própria quando precisamos calcular o conteúdo de freqüência de um sinal em um computador digital ou, então, calcular o que acontece com o sinal quando ele é processado por um dispositivo digital, tal como acontece em uma comunicação digital.

O extensivo material apresentado neste capítulo apresenta as seguintes lições:

- ▶ *Lição 1: A transformada de Fourier de um sinal especifica as amplitudes complexas das componentes que constituem a descrição no domínio da freqüência ou o conteúdo espectral do sinal. A transformada inversa de Fourier recupera, unicamente, o sinal, dada a sua descrição no domínio da freqüência.*
- ▶ *Lição 2: A transformada de Fourier possui várias propriedades importantes, as quais, individual ou coletivamente fornecem pistas valiosas sobre a relação entre o sinal definido no domínio do tempo e sua descrição no domínio da freqüência.*
- ▶ *Lição 3: Um sinal pode ser estritamente limitado no domínio do tempo ou no domínio da freqüência, mas não em ambos.*
- ▶ *Lição 4: A largura de faixa é um parâmetro importante na descrição do conteúdo espectral de um sinal e da resposta em freqüência de um filtro invariante no tempo.*
- ▶ *Lição 5: Um algoritmo amplamente utilizado, chamado de transformada rápida de Fourier, é uma poderosa ferramenta para o cálculo da transforma discreta de Fourier. Ele é a ferramenta matemática para cálculos digitais envolvendo a transformada de Fourier.*

2.1 A TRANSFORMADA DE FOURIER¹

■ Definições

Seja $g(t)$ um sinal não periódico determinístico, expresso como alguma função do tempo t . Pela definição, a transformada de Fourier do sinal $g(t)$ é dada pela integral

$$G(f) = \int_{-\infty}^{\infty} g(t) \exp(-j2\pi ft) dt \quad (2.1)$$

Na qual $j = \sqrt{-1}$ e a variável f representa *freqüência*. A função exponencial $\exp(-j2\pi ft)$ é referida como o núcleo da fórmula que define a transformada de Fourier. Dada a transformada de Fourier $G(f)$, o sinal original $g(t)$ é recuperado exatamente usando a fórmula da transformada inversa de Fourier:

$$g(t) = \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft) df \quad (2.2)$$

Na qual a exponencial $\exp(j2\pi ft)$ é o núcleo da fórmula que define a transformada inversa de Fourier. Os dois núcleos das Eqs. (2.1) e (2.2) são, portanto, o complexo conjugado um do outro.

Observe, também, que nas Eqs. (2.1) e (2.2), utilizamos letra minúscula para representar a função no tempo e letra maiúscula para representar a função correspondente na freqüência. Dizemos que as funções $g(t)$ e $G(f)$ constituem um par da transformada de Fourier. No Apêndice 2, obtemos as definições da transformada de Fourier e sua inversa, começando da Série de Fourier de uma forma de onda periódica.

Referimos à Eq. (2.1) como uma equação de análise. Dado o comportamento no domínio do tempo de um sistema, somos capazes de analisar o comportamento do sistema no domínio da freqüência. A vantagem básica de transformar um comportamento no domínio do tempo para o domínio da freqüência é que a solução em senóides infinitas apresenta o comportamento como o da superposição de efeitos de regime permanente. Para sistemas cujo comportamento no domínio do tempo é descrito por equações lineares diferenciais, as soluções separadas do regime permanente geralmente são simples de serem analisadas, tanto em termos teóricos quanto em termos práticos.

De forma semelhante, chamamos a Eq. (2.2) de *equação de síntese*. Dada a superposição de efeitos em regime permanente no domínio da freqüência, podemos reconstruir o comportamento original no domínio do tempo do sistema sem perda de qualquer informação. As equações de análise e síntese trabalhando lado a lado como mostrado na Fig. 2.1, enriquecem a representação de sinais e sistemas tornando possível a visualização da representação em dois domínios interativos: o domínio do tempo e o domínio da freqüência.

Para a transformada de Fourier de um sinal $g(t)$ existir é suficiente, mas não necessário, que $g(t)$ satisfaça três condições conhecidas coletivamente como *condições de Dirichlet*:

1. A função $g(t)$ é de valor único, com um número finito de máximos e mínimos em qualquer intervalo de tempo finito.

¹ Joseph Fourier estudou o fluxo de calor no início do século 19. O entendimento do fluxo de calor era um problema com significado prático e científico naquele tempo e envolvia a solução de uma equação diferencial parcial chamada de equação de calor. Fourier desenvolveu uma técnica para resolver equações diferenciais parciais baseada na consideração de que a solução era a soma ponderada de senóides harmonicamente relacionadas com coeficientes desconhecidos, o que, agora, chamamos de Série de Fourier. O trabalho inicial de Fourier sobre a condição de calor foi submetido como um artigo para a Academia de Ciências de Paris em 1807 e rejeitado após a revisão de Lagrange, Laplace e Legendre. Fourier persistiu no desenvolvimento de suas idéias apesar de ser criticado pela falta de rigor por seus contemporâneos. Em 1822, ele publicou um livro contendo grande parte de seu trabalho, *Theorie Analytique de la Chaleur*, o qual, agora, é considerado um clássico da matemática.

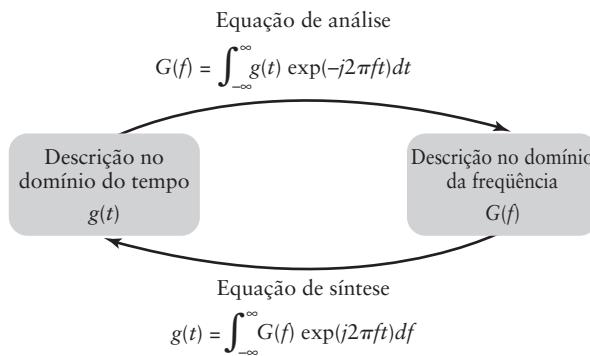


Figura 2.1 Rascunho da ação entre as equações de síntese e análise intrínsecas à transformada de Fourier.

2. A função $g(t)$ possui um número finito de descontinuidades em qualquer intervalo de tempo finito.
3. A função $g(t)$ é absolutamente integrável, isto é,

$$\int_{-\infty}^{\infty} |g(t)| dt < \infty$$

Podemos ignorar com segurança a questão de existência da transformada de Fourier de uma função no tempo $g(t)$ quando ela é uma descrição precisa de um sinal realizável (isto é, um sinal de voz ou sinal de vídeo). Em outras palavras, a realização física é uma condição suficiente para a existência da transformada de Fourier. Para a realização física de um sinal $g(t)$, a energia do sinal definida por $\int_{-\infty}^{\infty} |g(t)|^2 dt$ deve satisfazer a condição

$$\int_{-\infty}^{\infty} |g(t)|^2 dt < \infty$$

Tal sinal é referido como um sinal de tipo energia, ou simplesmente um sinal de energia. O que estamos dizendo, portanto, é que *todo sinal de energia possui transformada de Fourier*.

■ Representações

As fórmulas para a transformada de Fourier e sua inversa, apresentadas nas Eqs. (2.1) e (2.2), são escritas em termos de duas variáveis: *tempo t*, medido em segundos (s) e *freqüência f* medida em hertz (Hz). A freqüência f é relacionada com a freqüência angular ω por

$$\omega = 2\pi f$$

A qual é medida em *radianos por segundo* (rad/s). Podemos simplificar as expressões para os expoentes dos integrandos das Eqs. (2.1) e (2.2) usando ω no lugar de f . Entretanto, a utilização de f no lugar de ω é preferida por duas razões. Primeiro, a utilização da freqüência resulta em simetria matemática das Eqs. (2.1) e (2.2), uma com relação a outra, de forma natural. Segundo, os conteúdos espectrais de sinais de comunicação (isto é, sinais de voz e vídeo) são, geralmente, expressos em hertz.

Uma notação reduzida conveniente para a representação das relações de transformação das Eqs. (2.1) e (2.2) é obtida escrevendo

$$G(f) = F[g(t)] \quad (2.3)$$

e

$$g(t) = F^{-1}[G(f)] \quad (2.4)$$

Na qual $F[\cdot]$ e $F^{-1}[\cdot]$ fazem o papel de *operadores lineares*. Outra forma de notação conveniente para o par de transformada de Fourier, representada por $g(t)$ e $G(f)$ é

$$g(t) \iff G(f) \quad (2.5)$$

As notações apresentadas nas Eqs. (2.3) a (2.5) são utilizadas no texto, quando apropriado.

■ Espectro contínuo

Utilizando a operação de transformação de Fourier, um sinal de pulso $g(t)$ com energia finita é expresso como a soma contínua de funções exponenciais com freqüência no intervalo de $-\infty$ a ∞ . A componente de amplitude da freqüência f é proporcional a $G(f)$, na qual $G(f)$ é a transformada de Fourier de $g(t)$. Especificamente, para qualquer freqüência f , a função exponencial ($j2\pi ft$) é ponderada pelo fator $G(f) df$, o qual é a contribuição de $G(f)$ em um intervalo infinitesimal df centrado na freqüência f . Portanto, podemos expressar a função $g(t)$ em termos da soma contínua destas componentes infinitesimais, como mostrado pela integral

$$g(t) = \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft) df$$

Reafirmando o que foi mencionado previamente, a transformada de Fourier fornece uma ferramenta para resolver um dado sinal $g(t)$ em suas componentes exponenciais complexas ocupando todo o intervalo de freqüência de $-\infty$ a ∞ . Em particular, a transformada de Fourier $G(f)$ do sinal define a representação no domínio da freqüência do sinal ao definir as amplitudes complexas das várias componentes de freqüência do sinal. Podemos definir, de forma equivalente, o sinal em termos de sua representação no domínio do tempo especificando a função $g(t)$ em cada instante de tempo t . O sinal é unicamente definido por qualquer das representações.

Em geral, a transformada de Fourier $G(f)$ é uma função complexa da freqüência f , tal que podemos expressá-la na forma

$$G(f) = |G(f)| \exp[j\theta(f)] \quad (2.6)$$

Na qual $|G(f)|$ é chamado de *espectro contínuo de amplitude* de $g(t)$ e $\theta(f)$ é chamado de *espectro contínuo de fase* de $g(t)$. Neste caso o espectro é referido como *espectro contínuo* porque tanto a amplitude quanto a fase de $G(f)$ são unicamente definidos para todas as freqüências.

Para o caso especial de uma função $g(t)$ de valor real, temos:

$$G(-f) = G^*(f)$$

Na qual o asterisco denota conjugado complexo. Portanto, temos que se $g(t)$ é uma função de valor real do tempo t , então,

$$|G(-f)| = |G(f)|$$

e

$$\theta(-f) = -\theta(f)$$

De acordo, podemos fazer as seguintes afirmativas para o espectro de um sinal de valor real:

1. O espectro de amplitude de um sinal é uma função par da freqüência. Isto é, o espectro de amplitude é simétrico com relação à origem $f = 0$.

2. O espectro de fase do sinal é uma função ímpar da freqüência. Isto é, o espectro de amplitude é anti-simétrico com relação à origem $f = 0$.

Estas duas afirmativas são resumidas dizendo que o espectro de um sinal de valor real exibe uma *simetria conjugada*.

EXEMPLO 2.1 Pulso retangular

Considere uma função caixa, ou um *pulso retangular* de duração T e amplitude A , como mostrado na Fig. 2.2(a). Para definirmos este pulso matematicamente de forma conveniente, utilizamos a notação

$$\text{ret}(t) = \begin{cases} 1, & -\frac{1}{2} \leq t \leq \frac{1}{2} \\ 0, & t < -\frac{1}{2} \text{ ou } t > \frac{1}{2} \end{cases} \quad (2.7)$$

A qual define uma *função retangular* de amplitude e duração unitárias, centrada em $t = 0$. Então, em termos desta função padrão, podemos expressar o pulso retangular da Fig. 2.2(a) simplesmente como

$$g(t) = A \text{ ret}\left(\frac{t}{T}\right)$$

A transformada de Fourier do pulso retangular $g(t)$ é, então, dada por

$$\begin{aligned} G(f) &= \int_{-T/2}^{T/2} A \exp(-j2\pi ft) dt \\ &= AT \left(\frac{\sin(\pi f T)}{\pi f T} \right) \end{aligned} \quad (2.8)$$

Para simplificar a notação do resultado anterior e de posteriores, iremos introduzir outra função padrão, chamada de função *sinc*, definida por

$$\text{sinc}(\lambda) = \frac{\sin(\pi\lambda)}{\pi\lambda} \quad (2.9)$$

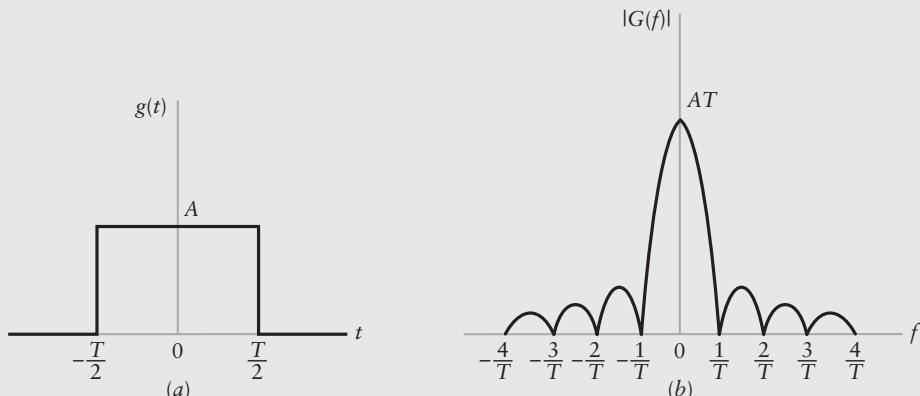


Figura 2.2 (a) Pulso retangular. (b) Espectro de amplitude.

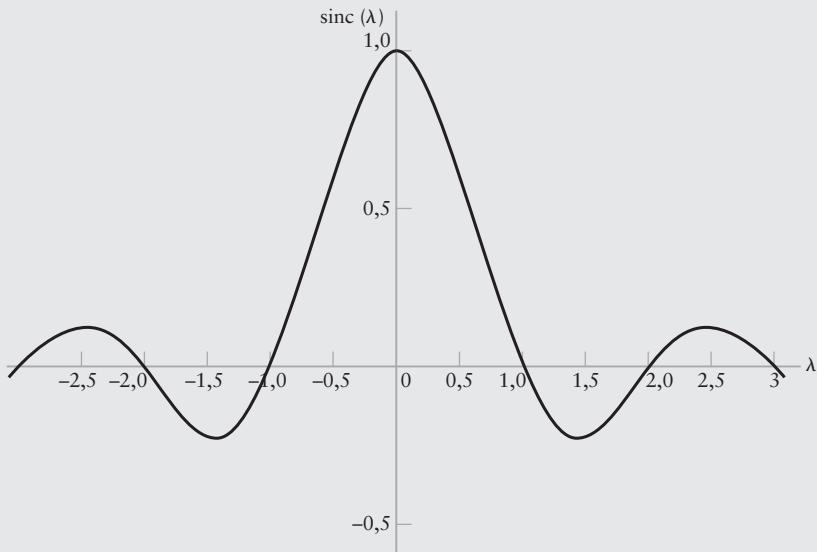


Figura 2.3 A função sinc.

Na qual λ é a variável independente. A função sinc possui um papel importante na teoria das comunicações. Como mostrado na Fig. 2.3, ela possui um valor máximo unitário em $\lambda = 0$ e aproxima-se de zero quando λ tende ao infinito, oscilando entre valores positivos e negativos. Ela vale zero para $\lambda = \pm 1, \pm 2, \dots$, e assim por diante.

Portanto, em termos da função sinc, podemos reescrever a Eq. (2.8) como

$$A \operatorname{ret}\left(\frac{t}{T}\right) \iff AT \operatorname{sinc}(ft) \quad (2.10)$$

O espectro de amplitude $|G(f)|$ é mostrado na Fig. 2.2(b). O primeiro cruzamento com zero do espectro ocorre em $f = \pm 1/T$. Quando a duração T do pulso diminui, seu primeiro cruzamento aumenta em freqüência. Por outro lado, quando a duração T do pulso aumenta, o primeiro cruzamento se move em direção à origem.

Este exemplo mostra que a relação entre as descrições no domínio do tempo e o domínio da freqüência de um sinal é inversa. Ou seja, um pulso estreito no tempo possui uma significante descrição em freqüência em uma ampla faixa de freqüências, e vice-versa. Teremos mais para concluir sobre a relação inversa entre tempo e freqüência na Seção 2.3.

Observe também que, neste exemplo, a transformada de Fourier $G(f)$ é uma função de valor real e simétrica em freqüência f . Esta é uma consequência direta do fato do pulso retangular $g(t)$ mostrado na Fig. 2.2(a) ser uma função simétrica no tempo t .

EXEMPLO 2.2 Pulso exponencial

Um pulso exponencial decrescente truncado é mostrado na Fig. 2.4(a). Definimos matematicamente este pulso de forma conveniente usando a função degrau unitário:

$$u(t) = \begin{cases} 1, & t > 0 \\ \frac{1}{2}, & t = 0 \\ 0, & t < 0 \end{cases} \quad (2.11)$$

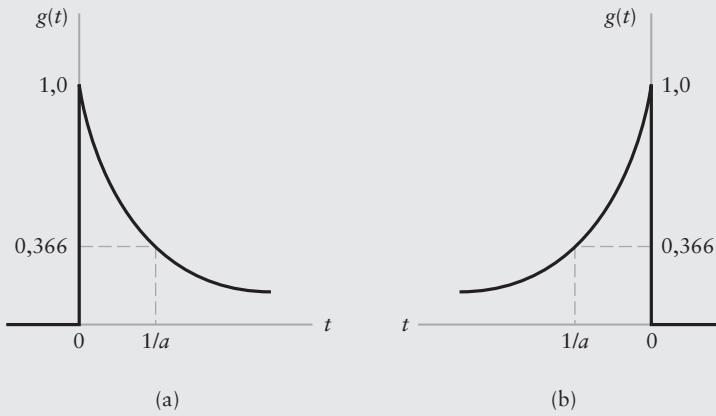


Figura 2.4 (a) Pulso exponencial decrescente. (b) Pulso exponencial crescente.

Podemos, então, descrever o pulso exponencial decrescente da Fig. 2.4(a) por

$$g(t) = \exp(-at)u(t)$$

Percebendo que \$g(t)\$ é zero para \$t < 0\$, a transformada de Fourier desse pulso é

$$\begin{aligned} G(f) &= \int_0^\infty \exp(-at) \exp(-j2\pi ft) dt \\ &= \int_0^\infty \exp[-t(a + j2\pi f)] dt \\ &= \frac{1}{a + j2\pi f} \end{aligned}$$

O par transformada de Fourier do pulso exponencial decrescente da Fig. 2.4(a) é, portanto,

$$\exp(-at)u(t) \iff \frac{1}{a + j2\pi f} \quad (2.12)$$

Um pulso exponencial crescente truncado é mostrado na Fig. 2.4(b), o qual é definido por

$$g(t) = \exp(at)u(-t)$$

Observe que \$u(-t)\$ é igual a um para \$t < 0\$, metade para \$t = 0\$ e zero para \$t > 0\$. Com \$g(t)\$ igual a zero para \$t > 0\$, a transformada de Fourier deste pulso é

$$G(f) = \int_{-\infty}^0 \exp(at) \exp(-j2\pi ft) dt$$

Substituindo \$t\$ por \$-t\$, podemos escrever que

$$\begin{aligned} G(f) &= \int_0^\infty \exp[-t(a - j2\pi f)] dt \\ &= \frac{1}{a - j2\pi f} \end{aligned}$$

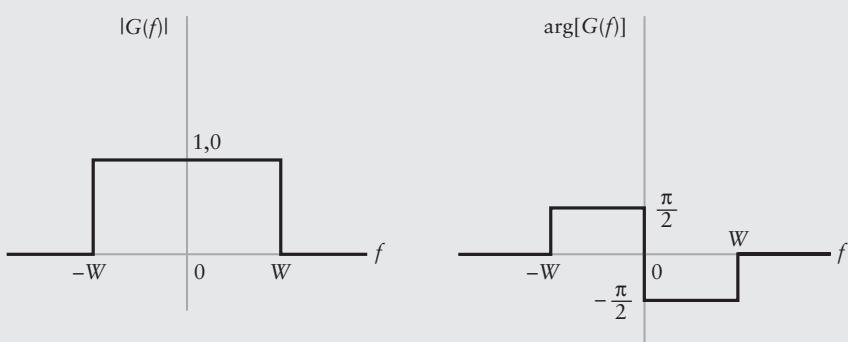


Figura 2.5 Função em freqüência $g(f)$ para o Problema 2.2.

O par transformada de Fourier para o pulso exponencial crescente da Fig. 2.4(b) é, portanto,

$$\exp(-at)u(-t) \iff \frac{1}{a - j2\pi f} \quad (2.13)$$

Os pulsos exponenciais decrescente ou crescente da Fig. 2.4 são ambos funções assimétricas do tempo t . Suas transformadas de Fourier são, portanto, funções de valor complexo, como podemos facilmente ver pelo fato dos pulsos exponenciais, decrescente ou crescente, truncados possuírem o mesmo espectro de amplitude, mas o espectro de fase de um é o negativo do espectro de fase do outro.

- **PROBLEMA 2.1** Calcule a transformada de Fourier da forma de onda senoidal amortecida $g(t) = \exp(-t) \sin(2\pi f_c t)u(t)$, na qual $u(t)$ é a função degrau unitário. ◀
- **PROBLEMA 2.2** Determine a transformada inversa de Fourier da função em freqüência $G(f)$ definida pelo espectro de amplitude e fase mostrado na Fig. 2.5. ◀

2.2 PROPRIEDADES DA TRANSFORMADA DE FOURIER

É útil termos algumas pistas sobre a relação entre uma função no tempo $g(t)$ e sua transformada de Fourier $G(f)$, além de como as várias operações na função $g(t)$ afetam a transformada $G(f)$. Esta tarefa pode ser realizada examinando certas propriedades da transformada de Fourier. Nesta seção iremos descrever quatorze propriedades, as quais iremos provar uma a uma. Estas propriedades são resumidas na Tabela A8.1 do Apêndice 8 no final do livro.

- **PROPRIEDADE 1 – LINEARIDADE (SUPERPOSIÇÃO)** Seja $g_1(t) \iff G_1(f)$ e $g_2(t) \iff G_2(f)$.

Então, para toda constante c_1 e c_2 , temos

$$c_1 g_1(t) + c_2 g_2(t) \iff c_1 G_1(f) + c_2 G_2(f) \quad (2.14)$$

A prova desta propriedade segue simplesmente da linearidade das integrais definindo $G(f)$ e $g(t)$.

A propriedade 1 nos permite determinar a transformada de Fourier $G(f)$ de uma função $g(t)$ que seja a combinação linear de duas outras funções $g_1(t)$ e $g_2(t)$ cujas transformadas de Fourier $G_1(f)$ e $G_2(f)$ sejam conhecidas, como ilustrado pelo exemplo a seguir.

EXEMPLO 2.3 Combinação de pulsos exponenciais

Considere o pulso exponencial duplo definido por (veja Fig. 2.6(a)),

$$g(t) = \begin{cases} \exp(-at), & t > 0 \\ 1, & t = 0 \\ \exp(at), & t < 0 \end{cases}$$

$$= \exp(-a|t|) \quad (2.15)$$

Este pulso pode ser visto como a soma de um pulso exponencial decrescente truncado e um pulso exponencial crescente truncado. Portanto, usando a propriedade da linearidade e os pares transformada de Fourier das Eqs. (2.12) e (2.13), determinamos que a transformada de Fourier do pulso exponencial duplo da Fig. 2.6(a) é

$$G(f) = \frac{1}{a + j2\pi f} + \frac{1}{a - j2\pi f}$$

$$= \frac{2a}{a^2 + (2\pi f)^2}$$

Portanto, temos o seguinte par transformada de Fourier para o pulso exponencial duplo da Fig. 2.6(a):

$$\exp(-a|t|) \iff \frac{2a}{a^2 + (2\pi f)^2} \quad (2.16)$$

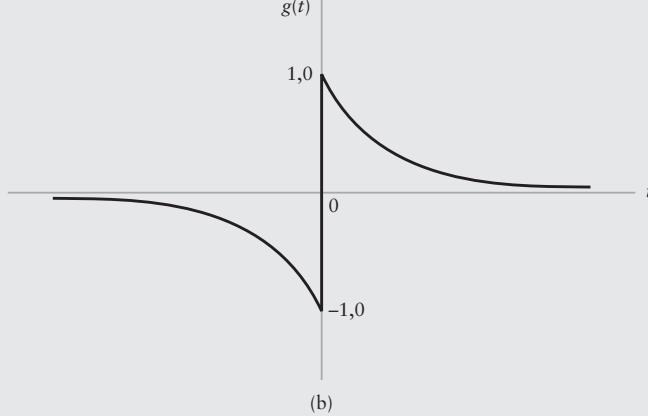
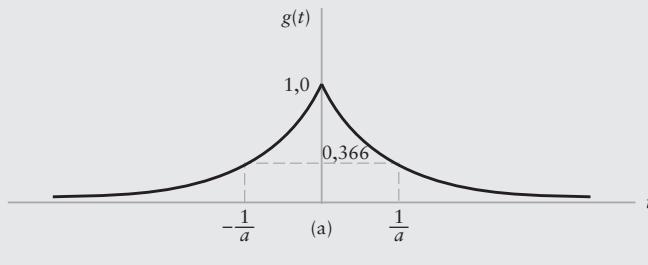


Figura 2.6 (a) Pulso exponencial duplo (simétrico). (b) Outro pulso exponencial duplo (simetria ímpar).

Observe que devido à simetria no domínio do tempo, mostrada na Fig. 2.6(a), o espectro é real e simétrico. Esta é uma propriedade geral de tais pares transformada de Fourier.

Outra combinação interessante é a diferença entre um pulso exponencial decrescente truncado e um pulso exponencial crescente truncado, como mostrado na Fig. 2.6(b). Neste caso temos,

$$g(t) = \begin{cases} \exp(-at), & t > 0 \\ 0, & t = 0 \\ -\exp(at), & t < 0 \end{cases} \quad (2.17)$$

Podemos formular uma notação compacta para este sinal composto usando a *função sinal*, igual a +1 para tempo positivo e -1 para tempo negativo, como mostrado por

$$\text{sgn}(t) = \begin{cases} +1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases} \quad (2.18)$$

A função sinal é mostrada na Fig. 2.7. Portanto, podemos reformular o sinal composto $g(t)$, definido na Eq. (2.17) por

$$g(t) = \exp(-a|t|) \text{sgn}(t)$$

Logo, aplicando a propriedade da linearidade da transformada de Fourier, podemos facilmente determinar, usando as Eqs. (2.12) e (2.13), que a transformada de Fourier do sinal $g(t)$ é dada por

$$\begin{aligned} \mathbf{F}[\exp(-a|t|) \text{sgn}(t)] &= \frac{1}{a + j2\pi f} - \frac{1}{a - j2\pi f} \\ &= \frac{-j4\pi f}{a^2 + (2\pi f)^2} \end{aligned}$$

Desta forma, temos o seguinte par transformada de Fourier

$$\exp(-a|t|) \text{sgn}(t) \iff \frac{-j4\pi f}{a^2 + (2\pi f)^2} \quad (2.19)$$

Em contraste com o par transformada de Fourier da Eq. (2.16), a transformada de Fourier da Eq. (2.19) é ímpar e puramente imaginária. Esta é uma propriedade geral de pares transformada de Fourier que se aplica a funções no tempo com simetria ímpar, as quais satisfazem a condição $g(-t) = -g(t)$, como mostrado na Fig. 2.6(b). Tais funções no tempo possuem funções ímpares e puramente imaginárias com transformadas de Fourier.

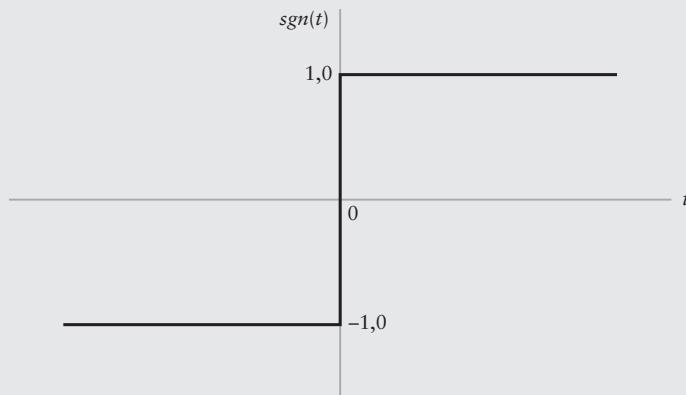


Figura 2.7 Função sinal.

- **PROPRIEDADE 2 – DILATAÇÃO** Seja $g(t) \iff G(f)$. Então, a propriedade da dilatação ou propriedade da similaridade afirma que

$$g(at) \iff \frac{1}{|a|} G\left(\frac{f}{a}\right) \quad (2.20)$$

na qual o fator de dilatação, representado por a , é um número real.

Para provar esta propriedade, note que

$$\mathcal{F}[g(at)] = \int_{-\infty}^{\infty} g(at) \exp(-j2\pi ft) dt$$

Seja $\tau = at$. Neste caso temos duas possibilidades, dependente se o fator de dilatação a é positivo ou negativo. Se $a > 0$, então

$$\begin{aligned} \mathcal{F}[g(at)] &= \frac{1}{a} \int_{-\infty}^{\infty} g(\tau) \exp\left[-j2\pi\left(\frac{f}{a}\right)\tau\right] d\tau \\ &= \frac{1}{a} G\left(\frac{f}{a}\right) \end{aligned}$$

Por outro lado, se $a < 0$, os limites de integração são trocados de tal forma que teremos o fator de multiplicação $-(1/a)$ ou, de forma equivalente, $1/|a|$, completando a prova da Eq. (2.20).

Observe que o fator de dilatação a e $1/a$ utilizado nas funções no tempo e na freqüência são recíprocos. Em particular, a função $g(at)$ representa $g(t)$ comprimido no tempo pelo fator a , enquanto que a função $G(f/a)$ representa $G(f)$ expandido em freqüência pelo mesmo fator a , assumindo que $0 < a < 1$. Logo, a regra da dilatação afirma que a compressão da função $g(t)$ no tempo é equivalente à expansão de sua transformada de Fourier $G(f)$ no domínio da freqüência pelo mesmo fator, ou vice-versa.

Para o caso especial quando $a = -1$, a regra de dilatação da Eq. (2.20) é reduzida à *propriedade da reflexão*, a qual afirma que se $g(t) \iff G(f)$ então,

$$g(-t) \iff G(-f) \quad (2.21)$$

Referindo à Fig. 2.4, vimos que o pulso exponencial crescente mostrado na parte (b) da figura é a reflexão do pulso exponencial decrescente mostrado na parte (a) com relação ao eixo vertical. Logo, aplicando a regra da reflexão na Eq. (2.12) que pertence ao pulso exponencial decrescente, facilmente vemos que a transformada de Fourier do pulso exponencial crescente é $1/(a - j2\pi f)$, a qual é exatamente o que obtivemos na Eq. (2.13).

- **PROPRIEDADE 3 – REGRA DA CONJUGAÇÃO** Seja $g(t) \iff G(f)$. Então, para uma função de valor complexo no tempo $g(t)$, temos

$$g^*(t) \iff G^*(-f) \quad (2.22)$$

na qual o asterisco representa a operação de conjugado complexo.

Para provar esta propriedade, sabemos da transformada inversa de Fourier que

$$g(t) = \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft) df$$

Calculando o conjugado complexo dos dois lados temos,

$$g^*(t) = \int_{-\infty}^{\infty} G^*(-f) \exp(-j2\pi ft) df$$

A seguir, substituindo f por $-f$, obtemos

$$\begin{aligned} g^*(t) &= - \int_{-\infty}^{-\infty} G^*(-f) \exp(j2\pi ft) df \\ &= \int_{-\infty}^{\infty} G^*(-f) \exp(j2\pi ft) df \end{aligned}$$

ou seja, $g^*(t)$ é a transformada inversa de Fourier de $G^*(-f)$, o qual é o resultado desejado.

Como um corolário da regra de conjugação da Eq. (2.22), podemos afirmar que se $g(t) \rightleftharpoons G(f)$, então

$$g^*(-t) \rightleftharpoons G^*(f) \quad (2.23)$$

Este resultado é obtido diretamente da Eq. (2.22) pela aplicação da regra da reflexão descrita na Eq. (2.21).

► **PROPRIEDADE 4 – DUALIDADE** Se $g(t) \rightleftharpoons G(f)$, então

$$G(t) \rightleftharpoons g(-f) \quad (2.24)$$

Esta propriedade é obtida da relação que define a transformada inversa de Fourier da Eq. (2.21) substituindo, primeiro, t por $-t$ e, então, escrevendo-a na forma

$$g(-t) = \int_{-\infty}^{\infty} G(f) \exp(-j2\pi ft) df$$

Finalmente, trocando t com f (isto é, substituindo t por f no lado esquerdo da equação e f por t no lado direito), temos

$$g(-f) = \int_{-\infty}^{\infty} G(t) \exp(-j2\pi ft) dt$$

a qual é a parte expandida da Eq. (2.24) ao se ir do domínio do tempo para o domínio da freqüência.

EXEMPLO 2.4 Pulso sinc

Considere um sinal $g(t)$ na forma da função sinc, mostrado por

$$g(t) = A \operatorname{sinc}(2Wt)$$

Para calcularmos a transformada de Fourier desta função, aplicamos as propriedades da dualidade e dilatação ao par transformada de Fourier da Eq. (2.10). Então, percebendo que a função retangular é uma função par no tempo, obtemos o resultado

$$A \operatorname{sinc}(2Wt) \rightleftharpoons \frac{A}{2W} \operatorname{ret}\left(\frac{f}{2W}\right) \quad (2.25)$$

ilustrada na Fig. 2.8. Vemos, portanto, que a transformada de Fourier do pulso sinc é zero para $|f| > W$. Note, também, que o pulso sinc é, por ele mesmo, apenas assintoticamente limitado no tempo, no sentido de que ele tende a zero quando o tempo t tende ao infinito. É esta característica assintótica que torna a função sinc um sinal de energia e, portanto, passível de sofrer uma transformação de Fourier.

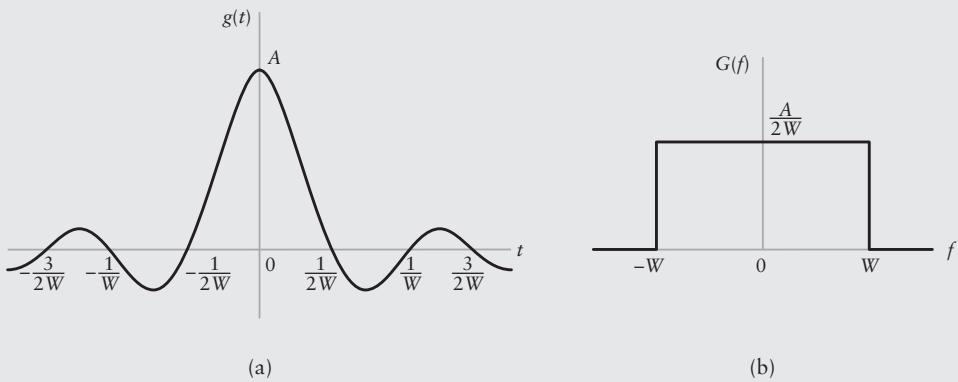


Figura 2.8 (a) Pulso sinc $g(t)$. (b) Transformada de Fourier $G(f)$.

► **PROPRIEDADE 5 – DESLOCAMENTO NO TEMPO** Se $g(t) \iff G(f)$, então

$$g(t - t_0) \iff G(f) \exp(-j2\pi f t_0) \quad (2.26)$$

onde t_0 é um deslocamento real no tempo constante.

Para provar esta teoria, calculamos a transformada de Fourier de $g(t - t_0)$ e, então, fazemos $\tau = (t - t_0)$ ou, de forma equivalente, $t = \tau + t_0$. Obtemos, portanto,

$$\begin{aligned} F[g(t - t_0)] &= \exp(-j2\pi f t_0) \int_{-\infty}^{\infty} g(\tau) \exp(-j2\pi f \tau) d\tau \\ &= \exp(-j2\pi f t_0) G(f) \end{aligned}$$

A propriedade de deslocamento no tempo afirma que se uma função $g(t)$ é deslocada ao longo do eixo de tempo por um valor t_0 , o efeito é equivalente a multiplicar sua transformada de Fourier $G(f)$ pelo fator $\exp(-j2\pi f t_0)$. Isto significa que a amplitude $G(f)$ permanece inalterada pelo deslocamento no tempo, mas sua fase é alterada pelo fator linear $-2\pi f t_0$, o qual varia linearmente com a freqüência f .

► **PROPRIEDADE 6 – DESLOCAMENTO NA FREQÜÊNCIA** Se $g(t) \iff G(f)$, então

$$\exp(j2\pi f_c t)g(t) \iff G(f - f_c) \quad (2.27)$$

onde f_c é uma freqüência real constante.

Esta propriedade é obtida do fato de que

$$\begin{aligned} F[\exp(j2\pi f_c t)g(t)] &= \int_{-\infty}^{\infty} g(t) \exp[-j2\pi t(f - f_c)] dt \\ &= G(f - f_c) \end{aligned}$$

ou seja, a multiplicação de uma função $g(t)$ pelo fator $\exp(j2\pi f_c t)$ é equivalente a deslocar sua transformada de Fourier $G(f)$ ao longo do eixo das freqüências por um fator f_c . Esta propriedade é um caso especial do *teorema da modulação*, discutido posteriormente na Propriedade 11. Basicamente, um deslocamento na faixa de freqüências de um sinal é realizado utilizando o processo de modulação. Note a dualidade entre as operações de deslocamento no tempo e na freqüência descritas pelas Eqs. (2.26) e (2.27).

EXEMPLO 2.5 Pulso de rádio freqüência (RF)

Considere o sinal de pulso $g(t)$ mostrado na Fig. 2.9(a), o qual é constituído de uma forma de onda senoidal com amplitude unitária e freqüência f_c , estendendo em duração de $t = -T/2$ até $t = T/2$. Este sinal é algumas vezes referenciado como um pulso de RF quando a freqüência f_c cai na faixa de rádio freqüência. O sinal $g(t)$ da Fig. 2.9(a) pode ser matematicamente expresso por:

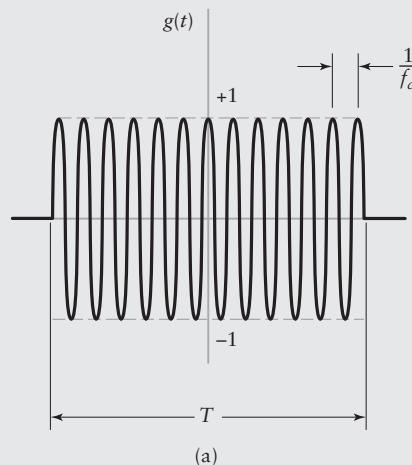
$$g(t) = \text{ret}\left(\frac{t}{T}\right) \cos(2\pi f_c t) \quad (2.28)$$

Para determinarmos a transformada de Fourier do sinal RF, primeiro utilizamos a fórmula de Euler para escrever

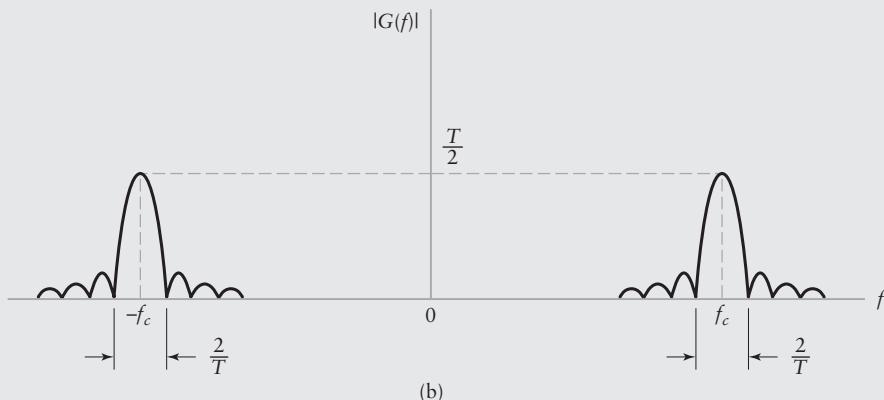
$$\cos(2\pi f_c t) = \frac{1}{2}[\exp(j2\pi f_c t) + \exp(-j2\pi f_c t)]$$

Portanto, aplicando a propriedade de deslocamento em freqüência ao par transformada de Fourier da Eq. (2.10) e, então, invocando a propriedade da linearidade da transformada de Fourier, obtemos o resultado desejado

$$\text{ret}\left(\frac{t}{T}\right) \cos(2\pi f_c t) \iff \frac{T}{2} \{\text{sinc}[T(f - f_c)] + \text{sinc}[T(f + f_c)]\}$$



(a)



(b)

Figura 2.9 (a) Pulso RF de amplitude unitária e duração T . (b) Espectro de amplitude.

No caso especial de $f_c T \gg 1$ – ou seja, a freqüência f_c é grande quando comparada com o recíproco da duração T do pulso – podemos utilizar o resultado aproximado

$$G(f) \approx \begin{cases} \frac{T}{2} \operatorname{sinc}[T(f - f_c)], & f > 0 \\ 0, & f = 0 \\ \frac{T}{2} \operatorname{sinc}[T(f + f_c)], & f < 0 \end{cases} \quad (2.30)$$

Com a condição $f_c T \gg 1$, o espectro de amplitude do pulso RF é mostrado na Fig. 2.9(b). Este diagrama, em relação à Fig. 2.2(b) ilustra, claramente, a propriedade de deslocamento em freqüência da transformada de Fourier.

► **PROPRIEDADE 7 – ÁREA SOB $g(t)$** Se $g(t) \iff G(f)$, então

$$\int_{-\infty}^{\infty} g(t) dt = G(0) \quad (2.31)$$

ou seja, a área sob a função $g(t)$ é igual ao valor de sua transformada de Fourier $G(f)$ em $f = 0$.

Este resultado é obtido fazendo, simplesmente, $f = 0$ na Eq. (2.1), a qual define a transformada de Fourier de $g(t)$.

► **PROBLEMA 2.3** Suponha que $g(t)$ é uma função de valor real com transformada de Fourier $G(f)$ de valor complexo. Explique como a regra da Eq. (2.31) pode ser satisfeita para este tipo de sinal. ◀

► **PROPRIEDADE 8 – ÁREA SOBRE $G(f)$** Se $g(t) \iff G(f)$, então

$$g(0) = \int_{-\infty}^{\infty} G(f) df \quad (2.32)$$

ou seja, o valor da função $g(t)$ em $t = 0$ é igual à área sobre sua transformada de Fourier $G(f)$.

► **PROBLEMA 2.4** Explique como a regra da Eq. (2.32) pode ser satisfeita pelo sinal $g(t)$ descrito no Problema 2.3. ◀

► **PROPRIEDADE 9 – DIFERENCIACÃO NO DOMÍNIO DO TEMPO** Seja $g(t) \iff G(f)$ e assuma que a primeira derivada de $g(t)$ com respeito ao tempo t possui transformada de Fourier. Então,

$$\frac{d}{dt} g(t) \iff j2\pi f G(f) \quad (2.33)$$

ou seja, a diferenciação de uma função no tempo $g(t)$ possui o efeito de multiplicar sua transformada de Fourier $G(f)$ pelo fator puramente imaginário $j2\pi f$.

Este resultado é obtido em dois passos. No primeiro passo, obtemos a derivada primeira dos dois lados da integral da Eq. (2.2), a qual define a transformada inversa de Fourier de $G(f)$. No segundo passo, trocamos as operações de integração e diferenciação.

Podemos generalizar a Eq. (2.33) para derivadas de ordem superior da função no tempo $g(t)$ utilizando:

$$\frac{d^n}{dt^n} g(t) \iff (j2\pi f)^n G(f) \quad (2.34)$$

a qual inclui a Eq. (2.33) como caso especial. A Eq. (2.34) assume que a transformada de Fourier da derivada de ordem superior de $g(t)$ existe.

EXEMPLO 2.6 Pulso Gaussiano unitário

Normalmente, um sinal pulso $g(t)$ e sua transformada de Fourier $G(f)$ possuem formas matemáticas diferentes. Esta observação é ilustrada pelos pares transformada de Fourier estudados nos Exemplos 2.1 a 2.5. Neste exemplo, iremos considerar uma exceção a esta regra. Em particular, utilizamos a propriedade da diferenciação da transformada de Fourier para obter a forma particular do sinal de pulso que possui a mesma forma matemática que sua própria transformada de Fourier.

Seja $g(t)$ a representação de um sinal de pulso expresso como uma função do tempo t e $G(f)$ sua transformada de Fourier. Diferenciando a fórmula da transformada de Fourier da Eq. (2.1) com relação à freqüência f , obtemos

$$-j2\pi t g(t) \iff \frac{d}{df} G(f)$$

ou, de forma equivalente,

$$2\pi t g(t) \iff j \frac{d}{df} G(f) \quad (2.35)$$

Suponha, agora, que iremos impor a seguinte condição aos lados esquerdos das Eqs. (2.33) e (2.35):

$$\frac{d}{dt} g(t) = -2\pi t g(t) \quad (2.36)$$

Então, de forma correspondente, temos que os lados direitos destas duas equações devem (após o cancelamento do fator comum j) satisfazer a condição

$$\frac{d}{df} G(f) = -2\pi f G(f) \quad (2.37)$$

As Eqs. (2.36) e (2.37) mostram que um sinal de pulso $g(t)$ e sua transformada de Fourier $G(f)$ possuem exatamente a mesma forma matemática. Em outras palavras, desde que o sinal de pulso $g(t)$ satisfaça a equação diferencial (2.36), então $G(f) = g(f)$, na qual $g(f)$ é obtida de $g(t)$ substituindo t por f . Resolvendo a Eq. (2.36) para $g(t)$, obtemos,

$$g(t) = \exp(-\pi t^2) \quad (2.38)$$

O pulso definido pela Eq. (2.38) é chamado de *pulso Gaussiano*, nome derivado da similaridade com a função densidade de probabilidade Gaussiana da teoria da probabilidade (veja o Capítulo 8). Ele é mostrado na Fig. 2.10. Aplicando a Eq. (2.31), determinaremos que a área sob este pulso unitário Gaussiano é dada por

$$\int_{-\infty}^{\infty} \exp(-\pi t^2) dt = 1 \quad (2.39)$$

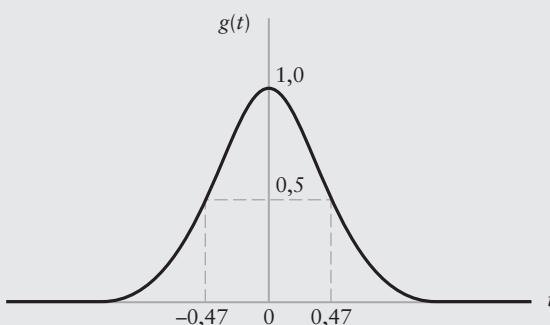


Figura 2.10 Pulso Gaussiano.

Quando a ordenada central e a área sob a curva de um pulso são ambos unitários, tal como nas Eqs. (2.38) e (2.39), dizemos que o pulso Gaussiano é um pulso unitário. Concluímos, portanto, que o pulso Gaussiano unitário é sua própria transformada de Fourier, como mostrado por:

$$\exp(-\pi t^2) \iff \exp(-\pi f^2) \quad (2.40)$$

► **PROPRIEDADE 10 – INTEGRAÇÃO NO DOMÍNIO DO TEMPO** Seja $g(t) \iff G(f)$, então, desde que $G(0) = 0$, temos

$$\int_{-\infty}^t g(\tau) d\tau \iff \frac{1}{j2\pi f} G(f) \quad (2.41)$$

ou seja, a integração de uma função no tempo $g(t)$ possui o efeito de dividir sua transformada de Fourier $G(f)$ pelo fator $j2\pi f$, desde que $G(0)$ seja zero.

Esta propriedade é verificada expressando $g(t)$ por

$$g(t) = \frac{d}{dt} \left[\int_{-\infty}^t g(\tau) d\tau \right]$$

e, então, aplicando a propriedade da diferenciação no tempo da transformada de Fourier para obtermos

$$G(f) = (j2\pi f) \left\{ F \left[\int_{-\infty}^t g(\tau) d\tau \right] \right\}$$

da qual a Eq. (2.41) é obtida diretamente.

É uma tarefa direta generalizar a Eq. (2.41) para múltiplas integrações. Entretanto, a notação se torna desajeitada.

A Eq. (2.41) assume que $G(0)$ – ou seja, a área sob $g(t)$ – é zero. O caso mais geral de $G(0) \neq 0$ é mostrado na Seção 2.4.

EXEMPLO 2.7 Pulso triangular

Considere o pulso anti-simétrico $g_1(t)$ mostrado na Fig. 2.11(a). Integrando este pulso com relação ao tempo, obtemos o pulso triangular $g_2(t)$ mostrado na Fig. 2.11(b). Observamos que o pulso anti-simétrico $g_1(t)$ é constituído por dois pulsos retangulares, um de amplitude A , definido no intervalo $-T \leq t \leq 0$ e outro de amplitude $-A$, definido no intervalo $0 \leq t \leq T$. Aplicando a propriedade de deslocamento no tempo da transformada de Fourier à Eq. (2.10), determinamos que as transformadas de Fourier destes dois pulsos retangulares são iguais a $AT \operatorname{sinc}(fT) \exp(j\pi fT)$ e $-AT \operatorname{sinc}(fT) \exp(-j\pi fT)$, respectivamente. Logo, invocando a propriedade da linearidade da transformada de Fourier, determinamos que a transformada de Fourier $G_1(f)$ do pulso anti-simétrico $g_1(t)$ da Fig. 2.11(a) é dada por

$$\begin{aligned} G_1(f) &= AT \operatorname{sinc}(fT) [\exp(j\pi fT) - \exp(-j\pi fT)] \\ &= 2jAT \operatorname{sinc}(fT) \operatorname{sen}(\pi fT) \end{aligned}$$

Notamos, ainda, da Eq. (2.42) que $G_1(0)$ é zero. Logo, usando as Eqs. (2.41) e (2.42), determinamos que a transformada de Fourier $G_2(f)$ do pulso triangular $g_2(t)$ da Fig. 2.11(b) é dada por

$$\begin{aligned} G_2(f) &= \frac{1}{j2\pi f} G_1(f) \\ &= AT \frac{\operatorname{sen}(\pi fT)}{\pi f} \operatorname{sinc}(fT) \\ &= AT^2 \operatorname{sinc}^2(fT) \end{aligned} \quad (2.43)$$

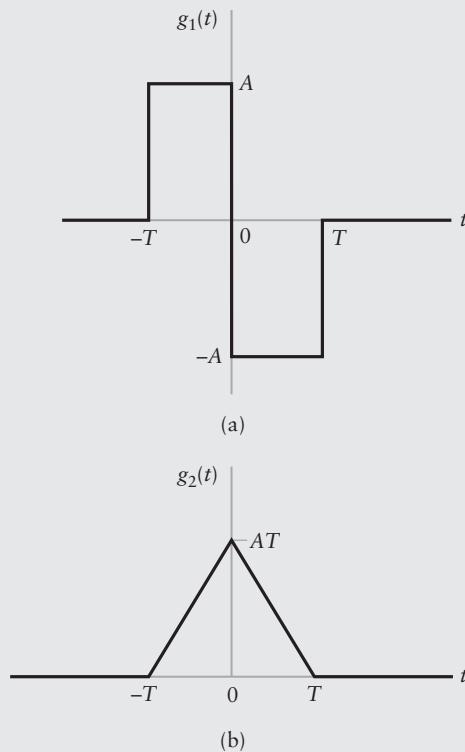


Figura 2.11 (a) Pulso anti-simétrico $g_1(t)$. (b) Pulso triangular $g_2(t)$ obtido pela integração de $g_1(t)$ com respeito ao tempo t .

Note que o pulso anti-simétrico da Fig. 2.11(a) é real com simetria ímpar e sua transformada de Fourier é, portanto, ímpar e puramente imaginária, enquanto que o pulso triangular da Fig. 2.11(b) é real e com simetria par e sua transformada de Fourier é, portanto, simétrica (simetria par) e puramente real.

EXEMPLO 2.8 Partes real e imaginária de uma função no tempo

Até agora, neste capítulo, discutimos a representação de Fourier de vários sinais, alguns puramente reais, outros puramente imaginários outros ainda sendo valores complexos com partes reais e imaginárias. É um momento oportuno, portanto, que neste estágio na análise de Fourier de sinais, utilizemos este exemplo para desenvolver algumas fórmulas para sinais complexos e seus espectros.

Expressando uma função $g(t)$ de valor complexo em termos de suas partes real e imaginária, podemos escrever

$$g(t) = \operatorname{Re}[g(t)] + j\operatorname{Im}[g(t)] \quad (2.44)$$

Na qual **Re** representa “a parte real de” e **Im** representa “a parte imaginária de”. O conjugado complexo de $g(t)$ é definido por

$$g^*(t) = \operatorname{Re}[g(t)] - j\operatorname{Im}[g(t)] \quad (2.45)$$

Somando as Eqs. (2.44) e (2.45), obtemos

$$\operatorname{Re}[g(t)] = \frac{1}{2}[g(t) + g^*(t)] \quad (2.46)$$

E subtraindo-as, temos

$$\operatorname{Im}[g(t)] = \frac{1}{2}[g(t) - g^*(t)] \quad (2.47)$$

Portanto, aplicando a regra da conjugação da Eq. (2.22), obtemos os seguintes pares transformada de Fourier:

$$\left. \begin{aligned} \operatorname{Re}[g(t)] &\iff \frac{1}{2}[G(f) + G^*(-f)] \\ \operatorname{Im}[g(t)] &\iff \frac{1}{2}[G(f) - G^*(-f)] \end{aligned} \right\} \quad (2.48)$$

Da segunda linha da Eq. (2.48) é evidente que no caso de uma função no tempo $g(t)$ de valor real temos $G(f) = G^*(-f)$, ou seja, a transformada de Fourier $G(f)$ exibe uma simetria de conjugado, confirmado o resultado já afirmado na Seção 2.2

► **PROPRIEDADE 11 – TEOREMA DA MODULAÇÃO** Seja $g_1(t) \iff G_1(f)$ e $g_2(t) \iff G_2(f)$. Então,

$$g_1(t)g_2(t) \iff \int_{-\infty}^{\infty} G_1(\lambda)G_2(f - \lambda) d\lambda \quad (2.49)$$

Para provarmos esta propriedade, inicialmente representamos a transformada de Fourier do produto $g_1(t)g_2(t)$ por $G_{12}(f)$, escrevendo

$$g_1(t)g_2(t) \iff G_{12}(f)$$

na qual

$$G_{12}(f) = \int_{-\infty}^{\infty} g_1(t)g_2(t) \exp(-j2\pi ft) dt$$

Para $g_2(t)$, substituímos, a seguir, a transformada inversa de Fourier

$$g_2(t) = \int_{-\infty}^{\infty} G_2(f') \exp(j2\pi f' t) df'$$

na integral definindo $G_{12}(f)$, obtendo

$$G_{12}(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(t)G_2(f') \exp[-j2\pi(f - f')t] df' dt$$

Definindo $\lambda = f - f'$, eliminamos a variável f' e trocamos a ordem de integração, obtendo (após o rearranjo dos termos)

$$G_{12}(f) = \int_{-\infty}^{\infty} G_2(f - \lambda) \left[\int_{-\infty}^{\infty} g_1(t) \exp(-j2\pi\lambda t) dt \right] d\lambda$$

assumindo que f é fixo. A integral interna (dentro dos colchetes) é reconhecida simplesmente com $G_1(\lambda)$ e, portanto, podemos escrever

$$G_{12}(f) = \int_{-\infty}^{\infty} G_1(\lambda)G_2(f - \lambda) d\lambda$$

a qual é o resultado desejado. Esta integral é conhecida como a *integral de convolução*, expressa no domínio da freqüência, e a função $G_{12}(f)$ é chamada de a *convolução* de $G_1(f)$ e $G_2(f)$. Concluímos que a *multiplicação de dois sinais no domínio do tempo é transformada na convolução de suas transformadas de Fourier no domínio da freqüência*. Esta propriedade também é chamada de *teorema da modulação*. Falaremos mais sobre as implicações práticas desta propriedade nos capítulos seguintes.

Na discussão da convolução, a notação abreviada a seguir é freqüentemente utilizada:

$$G_{12}(f) = G_1(f) \star G_2(f)$$

De acordo, podemos reformular a Eq. (2.49) na seguinte forma simbólica

$$g_1(t)g_2(t) \rightleftharpoons G_1(f) \star G_2(f) \quad (2.50)$$

Na qual o símbolo \star representa a convolução. Note que a convolução é comutativa, ou seja,

$$G_1(f) \star G_2(f) = G_2(f) \star G_1(f)$$

a qual é obtida diretamente da Eq. (2.50).

- **PROPRIEDADE 12 – TEOREMA DA CONVOLUÇÃO** Seja $g_1(t) \rightleftharpoons G_1(f)$ e $g_2(t) \rightleftharpoons G_2(f)$. Então,

$$\int_{-\infty}^{\infty} g_1(\tau)g_2(t - \tau) d\tau \rightleftharpoons G_1(f)G_2(f) \quad (2.51)$$

A Eq. (2.51) é obtida diretamente da combinação da Propriedade 4 (dualidade) com a Propriedade 11 (modulação). Podemos, portanto, afirmar que *a convolução de dois sinais no domínio do tempo é transformado na multiplicação de suas transformadas de Fourier no domínio da freqüência*. Esta propriedade é chamada de *teorema da convolução*. Ela nos permite trocar a operação da convolução no domínio do tempo pela multiplicação das duas transformadas de Fourier, uma operação é que ordinariamente mais simples de ser manipulada. Temos mais a dizer sobre a convolução posteriormente no capítulo, quando a questão de filtragem é discutida.

Usando a notação abreviada de convolução, podemos reescrever a Eq. (2.51) na forma mais simples dada por

$$g_1(t) \star g_2(t) \rightleftharpoons G_1(f)G_2(f) \quad (2.52)$$

Note que as Propriedades 11 e 12, descritas pelas Eqs. (2.49) e (2.51), respectivamente, são o dual uma da outra.

- **PROBLEMA 2.5** Desenvolva os passos detalhados para mostrar que os teoremas da modulação e convolução são, de fato,iais um do outro. ◀
- **PROPRIEDADE 13 – TEOREMA DA CORRELAÇÃO** Seja $g_1(t) \rightleftharpoons G_1(f)$ e $g_2(t) \rightleftharpoons G_2(f)$. Então, assumindo que $g_1(t)$ e $g_2(t)$ são de valor complexo,

$$\int_{-\infty}^{\infty} g_1(t)g_2^*(t - \tau) dt \rightleftharpoons G_1(f)G_2^*(f) \quad (2.53)$$

Na qual $G_2^*(f)$ é o complexo conjugado de $G_2(f)$ e τ é uma variável temporal necessária na definição da transformada inversa de Fourier do produto $G_1(f)G_2^*(f)$.

Para provarmos a Eq. (2.53), começamos reformulando a integral de convolução com os papéis das variáveis temporais t e τ trocados. Neste caso simplesmente reescrevemos a Eq. (2.51) como

$$\int_{-\infty}^{\infty} g_1(t)g_2(\tau - t) dt \rightleftharpoons G_1(f)G_2(f) \quad (2.54)$$

Como já mostrado na definição da Propriedade 13, a transformada inversa de Fourier do termo de produto $G_1(f)G_2(f)$ possui τ como sua variável de tempo. Ou seja, $\exp(j2\pi f\tau)$ está em seu núcleo. Com a fórmula da Eq. (2.54) em mãos, a Eq. (2.53) é obtida diretamente da combinação da regra de reflexão (caso especial da propriedade da dilatação) e da regra de conjugação.

A integral do lado esquerdo da Eq. (2.53) define uma medida de *similaridade* que pode existir entre um par de sinais de valor complexo. Esta medida é chamada de *correlação*, sobre a qual teremos mais a dizer posteriormente no capítulo.

► **PROBLEMA 2.6** Desenvolva os passos detalhados envolvidos na obtenção da Eq. (2.53), começando pela Eq. (2.51). ◀

► **PROBLEMA 2.7** Prove as seguintes propriedades do processo de convolução:

(a) propriedade comutativa: $g_1(t) \star g_2(t) = g_2(t) \star g_1(t)$

(b) propriedade associativa: $g_1(t) \star [g_2(t) \star g_3(t)] = [g_1(t) \star g_2(t)] \star g_3(t)$

(c) propriedade distributiva: $g_1(t) \star [g_2(t) + g_3(t)] = g_1(t) \star g_2(t) + g_1(t) \star g_3(t)$ ◀

► **PROPRIEDADE 14 – TEOREMA DE RAYLEIGH DA ENERGIA** Seja $g(t) \iff G(f)$. Então,

$$\int_{-\infty}^{\infty} |g(t)|^2 dt = \int_{-\infty}^{\infty} |G(f)|^2 df \quad (2.55)$$

Para provar a Eq. (2.55), fazemos $g_1(t) = g_2(t) = g(t)$ na Eq. (2.53), e nesse caso o teorema da correlação se reduz a

$$\int_{-\infty}^{\infty} g(t)g^*(t - \tau) dt \iff G(f)G^*(f) = |G(f)|^2$$

Na forma expandida, podemos escrever

$$\int_{-\infty}^{\infty} g(t)g^*(t - \tau) dt = \int_{-\infty}^{\infty} |G(f)|^2 \exp(j2\pi f\tau) df \quad (2.56)$$

Finalmente, fazendo $\tau = 0$ na Eq. (2.56), observamos que $g(t)g^*(t) = |g(t)|^2$ e obtemos o resultado desejado.

A Eq. (2.55), conhecida como *teorema de Rayleigh da energia*, afirma que a energia total do sinal transformado em Fourier é igual a área total sob a curva do quadrado do espectro de amplitude do sinal. A determinação da energia é geralmente simplificada utilizando o teorema de Rayleigh da energia, como mostrado no exemplo a seguir.

EXEMPLO 2.9 Pulso sinc (continuação)

Considere, novamente, o pulso sinc dado por $A \operatorname{sinc}(2Wt)$. A energia do pulso é igual à

$$E = A^2 \int_{-\infty}^{\infty} \operatorname{sinc}^2(2Wt) dt$$

A integral do lado direito desta equação é de difícil determinação. Entretanto, observamos do Exemplo 2.4 que a transformada de Fourier do pulso sinc dado por $A \operatorname{sinc}(2Wt)$ é igual a $(A/2W) \operatorname{rect}(f/2W)$. Logo aplicando o teorema de Rayleigh da energia a este problema, obtemos facilmente o resultado desejado:

$$\begin{aligned} E &= \left(\frac{A}{2W}\right)^2 \int_{-\infty}^{\infty} \operatorname{rect}^2\left(\frac{f}{2W}\right) df \\ &= \left(\frac{A}{2W}\right)^2 \int_{-W}^{W} df \\ &= \frac{A^2}{2W} \end{aligned} \quad (2.57)$$

Este exemplo mostra claramente a utilidade do teorema de Rayleigh da energia.

► **PROBLEMA 2.8** Considere a função de pulso sinc(t) e mostre que $\int_{-\infty}^{\infty} \text{sinc}^2(t) dt = 1$. ◀

2.3 A RELAÇÃO INVERSA ENTRE TEMPO E FREQUÊNCIA

As propriedades da transformada de Fourier discutidas na Seção 2.2 mostraram que as descrições no domínio do tempo e no domínio da freqüência de um sinal são inversamente relacionadas uma com a outra. Em particular, podemos fazer duas importantes afirmações:

1. Se a descrição no domínio do tempo de um sinal é alterada, a descrição no domínio da freqüência do sinal é alterada de maneira inversa, e vice-versa. Esta relação inversa impede especificações arbitrárias do sinal nos dois domínios. Em outras palavras, *podemos especificar uma função arbitrária no tempo ou um espectro arbitrário, mas não podemos especificar ambos ao mesmo tempo.*
2. Se um sinal é estritamente limitado em freqüência, a descrição no domínio do tempo do sinal será infinita, mesmo que sua amplitude assuma valores progressivamente menores. Dizemos que um sinal é *estritamente limitado em freqüência* ou *estritamente limitado em faixa* se sua transformada de Fourier for exatamente zero fora de uma faixa finita de freqüências. O pulso sinc é um exemplo de um sinal estritamente limitado em faixa, como ilustrado na Fig. 2.8. Esta figura também mostra que o pulso sinc é somente assintoticamente limitado no tempo. De forma inversa, se um sinal é estritamente limitado no tempo (isto é, o sinal é exatamente zero fora de um limite finito de tempo), então o espectro do sinal é infinito em extensão. Este comportamento é exemplificado tanto pelo pulso retangular (descrito na Fig. 2.2) quanto pelo pulso triangular (descrito na Fig. 2.11(b)). Desta forma, podemos dizer que *um sinal não pode ser estritamente limitado no tempo e na freqüência.*

■ Largura de faixa

A *largura de faixa* de um sinal fornece uma medida da *extensão do conteúdo espectral significativo de um sinal para freqüências positivas*. Quando o sinal é estritamente limitado em faixa, a largura de faixa é bem definida. Por exemplo, o pulso sinc descrito na Fig. 2.8(a) possui largura de faixa igual a W . Entretanto, quando o sinal não é estritamente limitado em faixa, o que geralmente ocorre, temos dificuldade em definir sua largura de faixa. A dificuldade aparece porque a palavra “significativo”, associada ao conteúdo espectral de um sinal, é matematicamente imprecisa. Conseqüentemente, não existe definição universal aceita para largura de faixa.

De qualquer forma, existem algumas definições geralmente aceitas para largura de faixa. Nesta seção, iremos considerar três destas definições. A formulação de cada definição depende se o sinal é passa-baixa ou passa-faixa. Um sinal é dito ser passa-baixa se o conteúdo espectral significativo for centrado na origem, em $f = 0$. Um sinal é dito ser passa-faixa se seu conteúdo espectral significativo estiver centrado em $\pm f_c$, onde f_c é uma freqüência constante.

Quando o espectro do sinal é simétrico com um lóbulo principal limitado por *nulos* bem definidos (isto é, freqüências nas quais o espectro é zero), podemos utilizar o lóbulo principal como base para definição da largura de faixa. A razão para isto é que o lóbulo espectral principal contém a porção significativa da energia do sinal. Se um sinal é passa-baixa, a largura de faixa é definida como metade da largura total do lóbulo espectral principal, pois apenas metade do lóbulo estará dentro da região de freqüências positivas. Por exemplo, um pulso retangular de duração T segundos possui lóbulo espectral principal com largura total de $(2/T)$ hertz, centrado na origem, como mostrado na Fig. 2.2(b). Portanto, podemos definir a largura de faixa deste pulso retangular como $(1/T)$ hertz. Por outro lado, se o sinal é passa-faixa com lóbulo espectral principal centrado em $\pm f_c$ na qual f_c é grande,

a largura de faixa é definida como a largura do lóbulo principal para freqüências positivas. Esta definição de largura de faixa é chamada de *largura de faixa de nulo para nulo*. Por exemplo, um pulso RF de duração T segundos e freqüência f_c possui lóbulos espectral principais com largura $(2/T)$ hertz centrados em $\pm f_c$, como mostrado na Fig. 2.9(b). Logo, podemos definir a largura de faixa de nulo para nulo deste pulso RF como sendo $(2/T)$ hertz. Com base nas definições apresentadas aqui, podemos afirmar que o deslocamento espectral do conteúdo de um sinal passa-baixa por uma freqüência suficientemente grande possui o efeito de dobrar a largura de faixa do sinal. Tal translação de freqüência é obtida utilizando o processo de modulação, o qual será discutido em detalhes no Capítulo 3.

Outra definição popular de largura de faixa é a *largura de faixa de 3dB*. Especificamente, se o sinal é passa-baixa, a largura de faixa de 3dB é definida como a separação entre a freqüência zero, na qual a amplitude do espectro é mantida em seu valor de pico e a *freqüência positiva* na qual a amplitude do sinal cai para $1/\sqrt{2}$ de seu valor de pico. Por exemplo, os pulsos exponenciais decrescente e crescente definidos na Fig. 2.4 possuem largura de faixa de 3dB de $(a/2\pi)$ hertz. Se, por outro lado, o sinal é passa-faixa, centrado em $\pm f_c$, a largura de faixa de 3dB é definida como a separação (ao longo do eixo de freqüências positivas) entre as duas freqüências na qual o espectro de amplitude do sinal cai para $1/\sqrt{2}$ do valor de pico em f_c . A largura de faixa de 3dB possui a vantagem de poder ser obtida diretamente do gráfico de espectro de amplitude. Entretanto, ela possui a desvantagem de poder resultar em uma análise errônea se o espectro de amplitude decrescer muito lentamente.

Outra medida de largura de faixa de um sinal é a *largura de faixa de raiz média quadrática (rms – root mean-square)*, ou largura de faixa eficaz, definida como a raiz quadrática do segundo momento de uma forma adequadamente normalizada do quadrado do espectro de amplitude, de tal forma que o segundo momento pode ser calculado na origem. Para a forma normalizada do quadrado do espectro de amplitude, utilizamos a função não negativa $|G(f)|^2 / \int_{-\infty}^{\infty} |G(f)|^2 df$, na qual o denominador resulta na correta normalização

no sentido do valor integrado desta razão sobre todo o eixo de freqüências é unitário. Podemos, portanto, definir formalmente a largura de faixa rms de um sinal passa-baixa $g(t)$ com transformada de Fourier $G(f)$ como sendo dada por:

$$W_{\text{rms}} = \left(\frac{\int_{-\infty}^{\infty} f^2 |G(f)|^2 df}{\int_{-\infty}^{\infty} |G(f)|^2 df} \right)^{1/2} \quad (2.58)$$

Uma característica atraente da largura de faixa rms, W_{rms} , é que ela por si só possibilita a determinação matemática mais fácil do que as duas outras definições de largura de faixa, apesar dela não ser facilmente medida em laboratório.

■ Produto tempo-largura de faixa

Para qualquer família de sinais de pulso que diferem por um fator de escala no tempo, o produto da duração do sinal por sua largura de faixa é constante, mostrado por

$$(\text{duração}) \times (\text{largura de faixa}) = \text{constante}$$

O produto é chamado de *produto tempo-largura de faixa* ou *produto largura de faixa-duração*. O fato do produto tempo-largura de faixa ser constante é outra manifestação da relação inversa que existe entre as descrições no domínio do tempo e no domínio da freqüência do sinal. Em particular, se a duração de um sinal de pulso é diminuída comprimindo-se a escala de tempo por um fator a , digamos, a escala de freqüência do espectro do sinal e, portanto, a

largura de faixa do sinal, é expandida pelo mesmo fator a , em virtude da Propriedade 2 (dilatação), e o produto tempo-largura de faixa do sinal será, portanto, mantido constante. Por exemplo, um pulso retangular de duração T segundos possui largura de faixa (definida com base na parte de freqüências positivas do lóbulo principal) igual a $(1/T)$ hertz, fazendo com que o produto tempo-largura de faixa do pulso seja unitário. O ponto importante a ser notado é que qualquer que seja a definição de largura de faixa utilizada para o sinal, o produto tempo-largura de faixa será mantido constante para uma certa classe de sinais de pulso. A escolha de uma definição particular para largura de faixa simplesmente altera o valor da constante.

Para sermos mais específicos, considere a largura de faixa rms definida na Eq. (2.58). A definição correspondente de *duração rms* para o sinal $g(t)$ é

$$T_{\text{rms}} = \left(\frac{\int_{-\infty}^{\infty} t^2 |g(t)|^2 dt}{\int_{-\infty}^{\infty} |g(t)|^2 dt} \right)^{1/2} \quad (2.59)$$

na qual assume-se que o sinal $g(t)$ esteja centrado na origem. Pode ser mostrado que usando as definições de rms das Eqs. (2.58) e (2.59), o produto tempo-largura de faixa possui a seguinte forma:

$$T_{\text{rms}} W_{\text{rms}} \geq \frac{1}{4\pi} \quad (2.60)$$

na qual a constante é $(1/4\pi)$. Também pode ser mostrado que o pulso Gaussiano satisfaz esta condição com o sinal de igualdade. Para detalhes sobre estes cálculos, o leitor deve se referir ao Problema 2.51.

2.4 FUNÇÃO DELTA DE DIRAC

Falando estritamente, a teoria da transformada de Fourier, como descrita nas Seções 2.2 e 2.3, é aplicável apenas a funções no tempo que satisfazem as condições de Dirichlet. Tais funções incluem sinais de energia – ou seja, sinais para os quais a condição

$$\int_{-\infty}^{\infty} |g(t)|^2 dt < \infty$$

é válida. Entretanto, pode ser extremamente desejável estender a teoria por dois motivos:

1. Para combinar a teoria de séries de Fourier e Transformada de Fourier em uma forma unificada, de tal forma que as séries de Fourier possam ser tratadas como um caso especial da transformada de Fourier (uma revisão de séries de Fourier é apresentada no Apêndice 2).
2. Para expandir a aplicabilidade da transformada de Fourier para incluir sinais de potência, ou seja, sinais para os quais a condição

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{T} |g(t)|^2 dt < \infty$$

é válida.

Estes dois objetivos se encontram através do uso adequado da *função Delta de Dirac* ou *impulso unitário*.

A função Delta de Dirac, representada por $\delta(t)$, é definida como tendo amplitude zero em todo local menos em $t = 0$, no qual ela é infinitamente grande, de tal forma que sua área sob a curva é unitária. Especificamente, $\delta(t)$ satisfaz o seguinte par de relações

$$\delta(t) = 0, \quad t \neq 0 \quad (2.61)$$

e

$$\int_{-\infty}^{\infty} \delta(t) dt = 1 \quad (2.62)$$

Uma implicação deste par de relações é que a função delta $\delta(t)$ deve ser uma função par do tempo t .

Para a função delta ter significado, entretanto, ela precisa aparecer como um fator no integrando de uma integral com relação ao tempo e apenas quando o outro fator do integrando é uma função contínua no tempo. Seja $g(t)$ esta função, considere o produto de $g(t)$ e a função delta deslocada no tempo $\delta(t - t_0)$. À luz das duas equações de definição (2.61) e (2.62), podemos expressar a integral do produto $g(t)\delta(t - t_0)$ com relação ao tempo t como mostrado a seguir:

$$\int_{-\infty}^{\infty} g(t)\delta(t - t_0) dt = g(t_0) \quad (2.63)$$

A operação indicada no lado esquerdo desta equação desloca o valor $g(t_0)$ da função $g(t)$ para o tempo $t = t_0$, onde $-\infty < t < \infty$. Portanto, a Eq. (2.63) é chamada de *propriedade de peneiramento* da função delta. Esta propriedade é, algumas vezes, utilizada como a equação de definição da função delta. De fato, ela incorpora as Eqs. (2.61) e (2.62) em uma única relação.

Observando que a função delta $\delta(t)$ é uma função par de t , podemos reescrever a Eq. (2.63) de forma a enfatizar a semelhança com a integral de convolução, mostrada por

$$\int_{-\infty}^{\infty} g(\tau)\delta(t - \tau) d\tau = g(t) \quad (2.64)$$

ou, usando a notação de convolução,

$$g(t) \star \delta(t) = g(t)$$

Em palavras, a convolução de qualquer função temporal $g(t)$ com a função delta $\delta(t)$ resulta na função completamente inalterada. Referimos a esta afirmativa como a *propriedade da replicação* da função delta.

Pela definição, a transformada de Fourier da função delta é dada por

$$F[\delta(t)] = \int_{-\infty}^{\infty} \delta(t) \exp(-j2\pi ft) dt$$

Logo, utilizando a propriedade de deslocamento da função delta e observando que $\exp(-j2\pi ft)$ é igual a 1 (um) para $t = 0$, obtemos,

$$F[\delta(t)] = 1$$

Portanto, temos o par transformada de Fourier para a função Delta de Dirac:

$$\delta(t) \iff 1 \quad (2.65)$$

Esta relação afirma que o espectro da função delta $\delta(t)$ se estende uniformemente sobre todo o intervalo de freqüência, como mostrado na Fig. 2.12.

É importante observar que o par transformada de Fourier da Eq. (2.65) existe apenas em sentido limitado. O ponto é que nenhuma função no sentido ordinário possui as duas propriedades das Eqs. (2.61) e (2.62) ou, de forma equivalente, a propriedade de deslocamento da Eq. (2.63). Entretanto, podemos imaginar uma seqüência de funções que possuem, progressivamente, picos mais altos e mais estreitos para $t = 0$, com área sob a curva

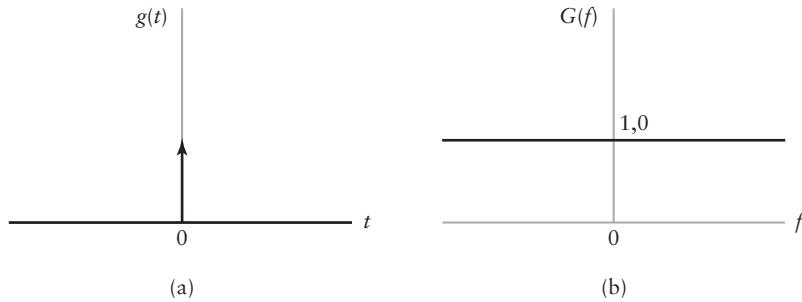


Figura 2.12 (a) A função delta de Dirac $\delta(t)$. (b) Espectro de $\delta(t)$.

permanecendo igual à unidade, enquanto que o valor da função tende para zero em todo ponto exceto $t = 0$, no qual ele tende ao infinito. Ou seja, podemos ver a função delta como *a forma limite de um pulso de área unitária e duração tendendo a zero*. O tipo do pulso utilizado não é importante.

Em sentido rigoroso, a função delta de Dirac pertence a uma classe especial de funções chamadas de *funções generalizadas* ou de *distribuições*. De fato, em algumas situações seu uso requer um cuidado considerável. De qualquer forma, um belo aspecto da função delta de Dirac está, precisamente, no fato de que, um tratamento intuitivo da função, considerando as questões descritas aqui, geralmente resulta na resposta correta.

EXEMPLO 2.10 A função delta como uma forma limite do pulso Gaussiano

Considere o pulso Gaussiano de área unitária, definido por

$$g(t) = \frac{1}{\tau} \exp\left(-\frac{\pi t^2}{\tau^2}\right) \quad (2.66)$$

na qual τ é um parâmetro variável. A função Gaussiana $g(t)$ possui duas propriedades úteis: (1) suas derivadas são todas contínuas e (2) ela desaparece mais rapidamente do que qualquer potência de t . A função delta $\delta(t)$ é obtida determinando-se o limite $\tau \rightarrow 0$. O pulso Gaussiano, então, se torna infinitamente estreito em duração e infinitamente grande em amplitude, mantendo a sua área finita e fixa em 1. A Fig. 2.13(a) ilustra a seqüência de tais pulsos quando o parâmetro τ é diminuído.

O pulso Gaussiano $g(t)$, definido aqui, é o mesmo pulso Gaussiano unitário $\exp(-\pi t^2)$ obtido no Exemplo 2.6, exceto pelo fato de que, agora, ele é escalonado no tempo pelo fator τ e escalonado em amplitude pelo fator $1/\tau$. Portanto, aplicando as propriedades da linearidade e dilatação da transformada de Fourier ao par transformada de Fourier da Eq. (2.40), observamos que a transformada de Fourier do pulso Gaussiano $g(t)$ definido na Eq. (2.66) também é Gaussiana, mostrado por

$$G(f) = \exp(-\pi\tau^2f^2)$$

A Fig. 2.13(b) ilustra o efeito da variação do parâmetro τ no espectro do pulso Gaussiano $g(t)$. Portanto, fazendo $\tau = 0$, obtemos, como esperado, que a transformada de Fourier da função delta é unitária.

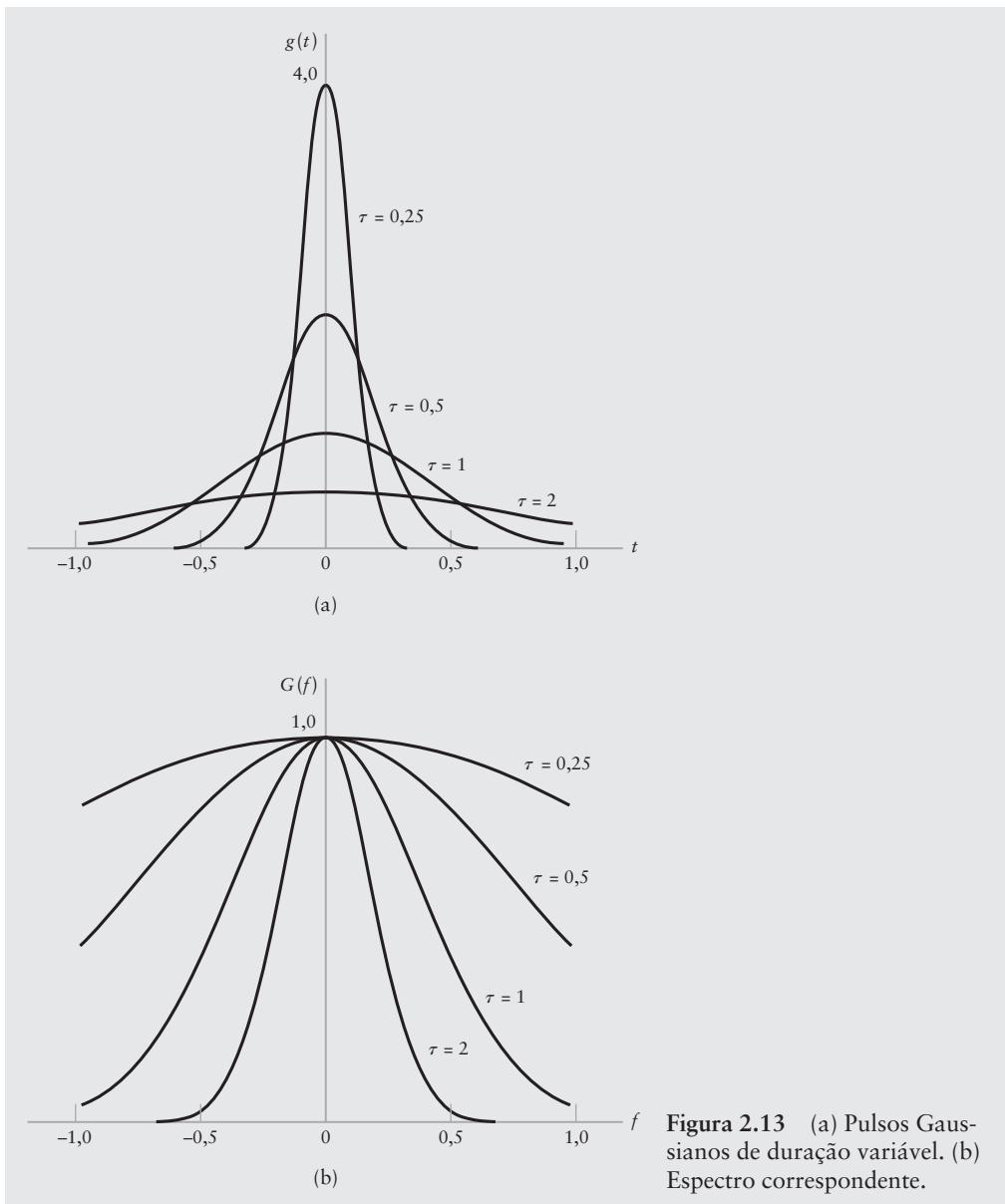


Figura 2.13 (a) Pulses Gaus-sianos de duração variável, (b) Espectro correspondente.

■ Aplicações da função delta

1. Sinal CC

Aplicando-se a propriedade da dualidade ao par da Eq. (2.65) e observando que a função delta é uma função par, obtemos,

$$1 \iff \delta(f) \quad (2.67)$$

A Eq. (2.67) afirma que um sinal CC é transformado no domínio da freqüência em uma função delta $\delta(t)$ ocorrendo na freqüência zero, como mostrado na Fig. 2.14. Obviamente este resultado satisfaz intuitivamente.

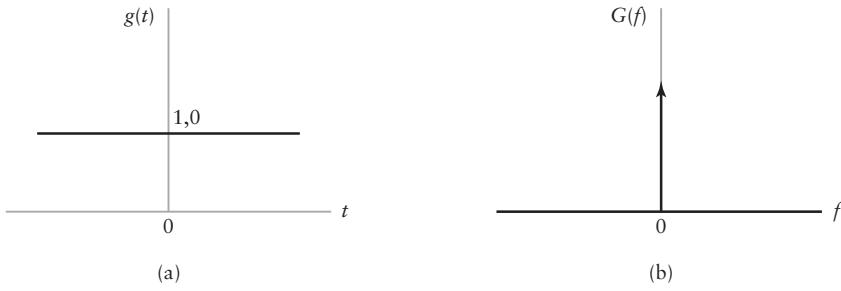


Figura 2.14 (a) Sinal CC. (b) Espectro.

Invocando a definição da transformada de Fourier, rapidamente deduzimos da Eq. (2.67) a relação útil

$$\int_{-\infty}^{\infty} \exp(-j2\pi ft) dt = \delta(f)$$

Reconhecendo que a função delta $\delta(f)$ é de valor real, podemos simplificar esta relação para

$$\int_{-\infty}^{\infty} \cos(2\pi ft) dt = \delta(f) \quad (2.68)$$

A qual provê outra definição para a função delta, apesar de estar no domínio da freqüência.

2. Função exponencial complexa

A seguir, aplicando a propriedade de deslocamento em freqüência à Eq. (2.67), obtemos o par transformada de Fourier

$$\exp(j2\pi f_c t) \iff \delta(f - f_c) \quad (2.69)$$

Para a função exponencial complexa de freqüência f_c . A Eq. (2.69) afirma que a função exponencial complexa $\exp(j2\pi f_c t)$ é transformada no domínio da freqüência em uma função delta $\delta(f - f_c)$ ocorrendo em $f = f_c$.

3. Funções senoidais

Considere o seguinte problema de se determinar a transformada de Fourier da função cosseno $\cos(2\pi f_c t)$. Primeiro utilizamos a fórmula de *Euler* para escrever

$$\cos(2\pi f_c t) = \frac{1}{2}[\exp(j2\pi f_c t) + \exp(-j2\pi f_c t)] \quad (2.70)$$

Portanto, utilizando a Eq. (2.69), verificamos que a função cosseno $\cos(2\pi f_c t)$ é representada pelo par transformada de Fourier

$$\cos(2\pi f_c t) \iff \frac{1}{2}[\delta(f - f_c) + \delta(f + f_c)] \quad (2.71)$$

Em outras palavras, o espectro da função cosseno $\cos(2\pi f_c t)$ consiste de um par de funções delta ocorrendo em $f = \pm f_c$. Cada um ponderado por um fator $1/2$, como mostrado na Fig. 2.15.

Similarmente, podemos mostrar que a função seno $\sin(2\pi f_c t)$ é representada pelo par transformada de Fourier

$$\sin(2\pi f_c t) \iff \frac{1}{2j}[\delta(f - f_c) - \delta(f + f_c)] \quad (2.72)$$

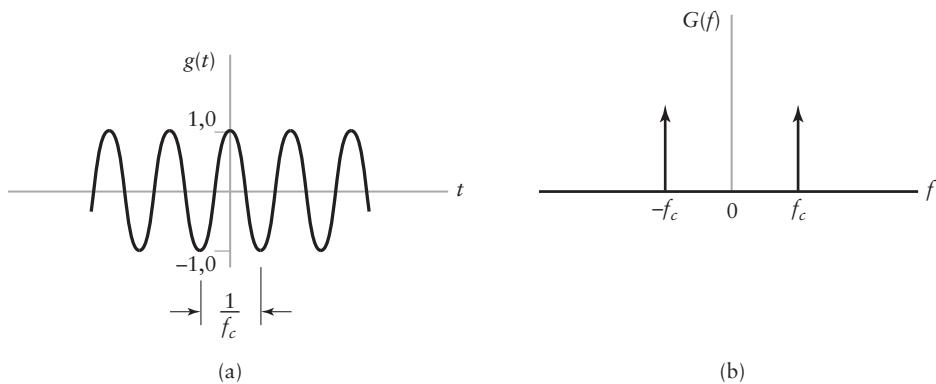


Figura 2.15 (a) Função cosseno. (b) Espectro.

a qual é ilustra na Fig. 2.16.

► **PROBLEMA 2.9** Determine a transformada de Fourier dos sinais quadráticos senoidais:

- (i) $g(t) = \cos^2(2\pi f_c t)$
- (ii) $g(t) = \sin^2(2\pi f_c t)$

4. Função sinal

A função sinal $\text{sgn}(t)$ é igual a $+1$ para valores positivos de tempo e -1 para valores negativos de tempo, como mostrado pela curva sólida da Fig. 2.17(a). A função sinal foi definida previamente na Eq. (2.18). Esta definição é reproduzida a seguir por conveniência:

$$\text{sgn}(t) = \begin{cases} +1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases}$$

A função sinal não satisfaz as condições de Dirichlet e, portanto, não possui transformada de Fourier. Entretanto, podemos definir a transformada de Fourier para a função sinal analisando-a como a forma limite do pulso exponencial dobrado com simetria ímpar.

$$g(t) = \begin{cases} \exp(-at), & t > 0 \\ 0, & t = 0 \\ -\exp(at), & t < 0 \end{cases} \quad (2.73)$$

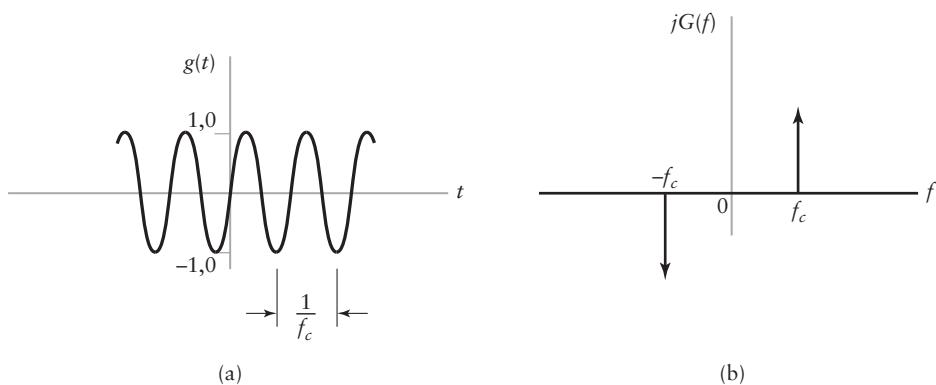


Figura 2.16 (a) Função seno. (b) Espectro.

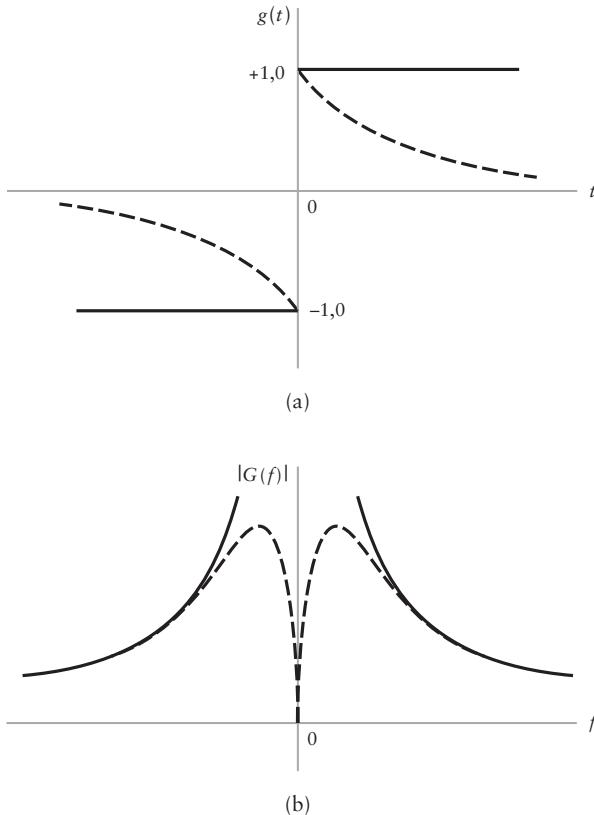


Figura 2.17 (a) Função sinal (curva contínua) e pulso exponencial dobrado (curva tracejada). (b) Espectro de amplitude da função sinal (curva contínua) e do pulso exponencial dobrado (curva tracejada).

quando o parâmetro a tende a zero. O sinal $g(t)$, mostrado em tracejado na Fig. 2.17(a), satisfaz as condições de Dirichlet. Sua transformada de Fourier foi obtida no Exemplo 2.3, sendo o resultado dado por [veja a Eq. (2.19)]

$$G(f) = \frac{-j4\pi f}{a^2 + (2\pi f)^2}$$

O espectro de amplitude de $|G(f)|$ é mostrado na curva tracejada da Fig. 2.17(b). No limite, quando a tende a zero, temos

$$\begin{aligned} F[\text{sgn}(t)] &= \lim_{a \rightarrow 0} \frac{-4j\pi f}{a^2 + (2\pi f)^2} \\ &= \frac{1}{j\pi f} \end{aligned}$$

ou seja,

$$\text{sgn}(t) \iff \frac{1}{j\pi f} \quad (2.74)$$

O espectro de amplitude da função sinal é mostrado na curva contínua da Fig. 2.17(b). Vemos que, para valores pequenos de a , a aproximação é muito boa, exceto próximo da origem do eixo de freqüências. Na origem, o espectro da função de aproximação $g(t)$ é zero para $a > 0$, enquanto que o espectro da função sinal tende ao infinito.

5. Função degrau unitário

A função degrau unitário $u(t)$ é igual a +1 para tempos positivos e zero para tempos negativos. Previamente definido pela Eq. (2.11), sua equação é reproduzida aqui por conveniência:

$$u(t) = \begin{cases} 1, & t > 0 \\ \frac{1}{2}, & t = 0 \\ 0, & t < 0 \end{cases}$$

A forma de onda da função degrau unitário é mostrada na Fig. 2.18(a). A partir da equação de definição e da função sinal, ou das formas de onda das Figs. 2.17(a) e 2.18(a), vemos que a função degrau unitário e a função sinal são relacionadas por:

$$u(t) = \frac{1}{2}[\operatorname{sgn}(t) + 1] \quad (2.75)$$

Logo, utilizando a propriedade da linearidade da transformada de Fourier e os pares transformada de Fourier das Eqs. (2.67) e (2.75), determinamos que a função degrau unitário é representada pelo par transformada de Fourier

$$u(t) \iff \frac{1}{j2\pi f} + \frac{1}{2}\delta(f) \quad (2.76)$$

Isto significa que o espectro da função degrau unitário contém a função delta ponderada pelo fator $\frac{1}{2}$ ocorrendo na freqüência zero, como mostrado na Fig. 2.18(b).

6. Integração no domínio do tempo (revista)

A relação da Eq. (2.41) descreve o efeito da integração na transformada de Fourier de um sinal $g(t)$, assumindo que $G(0)$ é zero. Iremos, agora, considerar um caso mais geral, sem esta restrição.

Seja

$$y(t) = \int_{-\infty}^t g(\tau) d\tau \quad (2.77)$$

O sinal integrado $y(t)$ pode ser visto como a convolução do sinal original $g(t)$ e da função degrau unitário $u(t)$, mostrado por

$$y(t) = \int_{-\infty}^{\infty} g(\tau)u(t - \tau) d\tau$$

na qual a função degrau unitário deslocada no tempo $u(t - \tau)$ é definida por

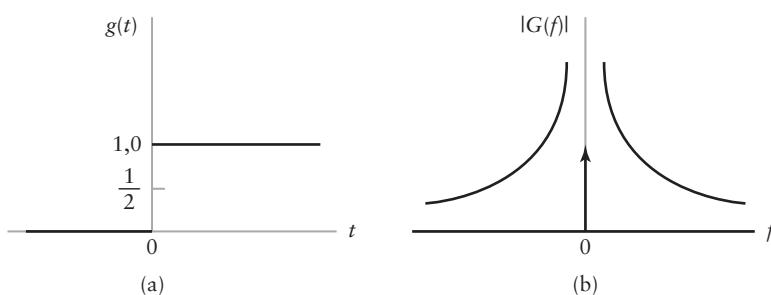


Figura 2.18 (a) Função degrau unitário. (b) Espectro de amplitude.

$$u(t - \tau) = \begin{cases} 1, & \tau < t \\ \frac{1}{2}, & \tau = t \\ 0, & \tau > t \end{cases}$$

Sabendo que a convolução no domínio do tempo é transformada em uma multiplicação do domínio da freqüência, de acordo com a propriedade 12, e utilizando o par transformada de Fourier da Eq. (2.76) para a função degrau unitário $u(t)$, determinamos a transformada de Fourier de $y(t)$ como sendo

$$Y(f) = G(f) \left[\frac{1}{j2\pi f} + \frac{1}{2} \delta(f) \right] \quad (2.78)$$

na qual $G(f)$ é a transformada de Fourier de $g(t)$. De acordo com a propriedade de deslocamento de uma função delta desenvolvida no domínio da freqüência, temos

$$G(f)\delta(f) = G(0)\delta(f)$$

Logo, podemos reescrever a Eq. (2.78) de forma equivalente:

$$Y(f) = \frac{1}{j2\pi f} G(f) + \frac{1}{2} G(0)\delta(f)$$

Em geral, o efeito da integração do sinal $g(t)$ é, portanto, descrito pelo par transformada de Fourier

$$\int_{-\infty}^t g(\tau) d\tau \iff \frac{1}{j2\pi f} G(f) + \frac{1}{2} G(0)\delta(f) \quad (2.79)$$

Este é o resultado desejado, o qual inclui a Eq. (2.41) como caso particular (ou seja, $G(0) = 0$).

► **PROBLEMA 2.10** Considere a função

$$g(t) = \delta\left(t + \frac{1}{2}\right) - \delta\left(t - \frac{1}{2}\right)$$

a qual é constituída da diferença entre duas funções delta em $t = \pm 1/2$. A integração de $g(t)$ com relação ao tempo t resulta na função retangular unitária $\text{ret}(t)$. Utilizando a Eq. (2.79) mostre que

$$\text{ret}(t) \iff \text{sinc}(f)$$

é uma forma especial da Eq. (2.10). ◀

2.5 TRANSFORMADA DE FOURIER DE SINAIS PERIÓDICOS

É bem conhecido que utilizando a série de Fourier, um sinal periódico pode ser representado pela soma de exponenciais complexas. (O Apêndice 2 apresenta uma revisão sobre série de Fourier.) Além disso, em sentido limitado, a transformada de Fourier pode ser definida por exponenciais complexas, como demonstrado nas Eqs. (2.69), (2.71) e (2.72). Portanto, parece ser razoável representar um sinal periódico em termos da transformada de Fourier, desde que seja permitido incluir as funções delta nesta transformada.

Considere o sinal periódico $g_{T_0}(t)$, no qual o subscrito T_0 representa o período do sinal. Sabemos que $g_{T_0}(t)$ pode ser representado em termos da série de Fourier de exponenciais complexas (veja o Apêndice 2)

$$g_{T_0}(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_0 t) \quad (2.80)$$

na qual c_n é o coeficiente complexo de Fourier, definido por

$$c_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_{T_0}(t) \exp(-j2\pi n f_0 t) dt \quad (2.81)$$

e f_0 a freqüência fundamental definida como o recíproco do período T_0 , ou seja,

$$f_0 = \frac{1}{T_0} \quad (2.82)$$

Seja $g(t)$ uma função do tipo pulso, a qual é igual a $g_{T_0}(t)$ em um período e zero no restante. Ou seja,

$$g(t) = \begin{cases} g_{T_0}(t), & -\frac{T_0}{2} \leq t \leq \frac{T_0}{2} \\ 0, & \text{caso contrário} \end{cases} \quad (2.83)$$

O sinal periódico $g_{T_0}(t)$ pode, agora, ser descrito em termos da função $g(t)$ como o somatório infinito

$$g_{T_0}(t) = \sum_{m=-\infty}^{\infty} g(t - mT_0) \quad (2.84)$$

Baseado nesta representação, podemos analisar $g(t)$ como uma *função de geração*, na qual ela gera o sinal periódico $g_{T_0}(t)$. Sendo um pulso com alguma energia finita, a função $g(t)$ possui transformada de Fourier. De acordo com as Eqs. (2.82) e (2.83), podemos escrever a fórmula do coeficiente complexo de Fourier c_n por:

$$\begin{aligned} c_n &= f_0 \int_{-\infty}^{\infty} g(t) \exp(-j2\pi n f_0 t) dt \\ &= f_0 G(n f_0) \end{aligned} \quad (2.85)$$

na qual $G(n f_0)$ é a transformada de Fourier de $g(t)$, calculada na freqüência $f = n f_0$. Podemos, portanto, reescrever a fórmula da Eq. (2.80) para a reconstrução do sinal periódico $g_{T_0}(t)$ como sendo

$$g_{T_0}(t) = f_0 \sum_{n=-\infty}^{\infty} G(n f_0) \exp(j2\pi n f_0 t) \quad (2.86)$$

Portanto, eliminando-se $g_{T_0}(t)$ entre as Eqs. (2.84) e (2.86), podemos escrever

$$\sum_{m=-\infty}^{\infty} g(t - mT_0) = f_0 \sum_{n=-\infty}^{\infty} G(n f_0) \exp(j2\pi n f_0 t) \quad (2.87)$$

a qual define uma forma da *fórmula da soma de Poisson*.

Finalmente, utilizando a Eq. (2.69), a qual define a transformada de Fourier de uma função exponencial complexa, na Eq. (2.87), deduzimos o par transformada de Fourier:

$$\sum_{m=-\infty}^{\infty} g(t - mT_0) \iff f_0 \sum_{n=-\infty}^{\infty} G(n f_0) \delta(f - n f_0) \quad (2.88)$$

para o sinal periódico $g_{T_0}(t)$ cuja freqüência fundamental é $f_0 = (1/T_0)$. A Eq. (2.88) simplesmente afirma que a transformada de Fourier de um sinal periódico é constituída de funções delta ocorrendo em múltiplos inteiros da freqüência fundamental f_0 , incluindo a origem, e que cada função delta é ponderada por um fator igual ao valor correspondente de $G(n f_0)$.

De fato, esta relação apenas prova um método de mostrar o conteúdo de freqüência de um sinal periódico $g_{T_0}(t)$.

É interessante observar que a função tipo pulso $g(t)$, constituindo de um período do sinal periódico $g_{T_0}(t)$, possui um espectro contínuo definido por $G(f)$. Por outro lado, o sinal periódico $g_{T_0}(t)$ possui um espectro discreto. Em palavras, podemos, portanto, resumir a transformação embutida na Eq. (2.88) por:

A periodicidade no domínio do tempo possui o efeito de alterar o espectro de um sinal do tipo pulso em uma forma discreta definida para múltiplos inteiros da freqüência fundamental e vice-versa.

EXEMPLO 2.11 Função de amostragem ideal

Uma função de amostragem ideal, ou também chamada de *Dirac Comb*, é constituída por uma seqüência infinita de funções delta uniformemente espaçadas, como mostrado na Fig. 2.19(a). Representamos esta forma de onda por:

$$\delta_{T_0}(t) = \sum_{m=-\infty}^{\infty} \delta(t - mT_0) \quad (2.89)$$

observe que a função de geração $g(t)$ para a função de amostragem ideal $\delta_{T_0}(t)$ consiste simplesmente da função delta $\delta(t)$. Portanto, temos $G(f) = 1$ e

$$G(nf_0) = 1 \text{ para todo } n$$

Portanto, a utilização da Eq. (2.88) resulta em um novo resultado

$$\sum_{m=-\infty}^{\infty} \delta(t - mT_0) \iff f_0 \sum_{n=-\infty}^{\infty} \delta(f - nf_0) \quad (2.90)$$

A Eq. (2.90) afirma que a transformada de Fourier de trem periódico de funções delta, espaçado por T_0 segundos, consiste de outro conjunto de funções delta ponderado pelo fator $f_0 = (1/T_0)$ e regularmente espaçado por f_0 Hz ao longo do eixo de freqüências, como mostrado na Fig. 2.19(b). No caso especial de $T_0 = 1$, o trem periódico de funções delta é, tal como o pulso Gaussiano, a própria transformada de Fourier.

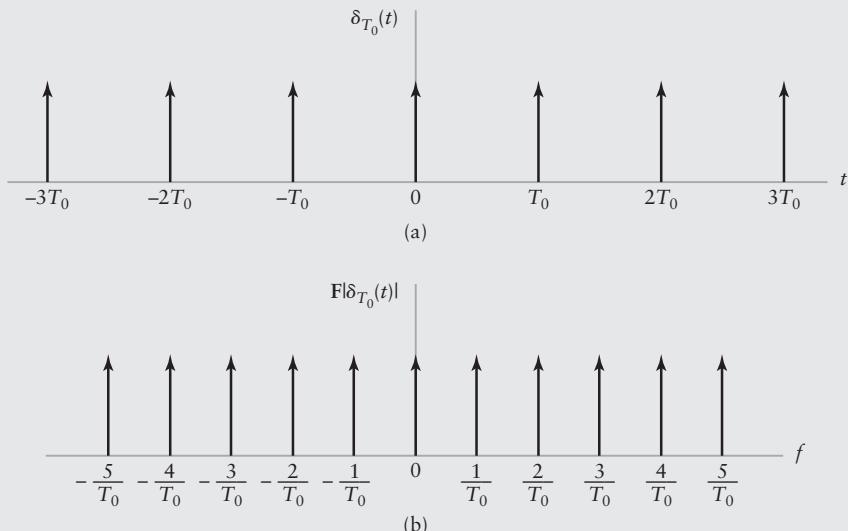


Figura 2.19 (a) Função de amostragem ideal. (b) Espectro.

Aplicando a transformada inversa de Fourier ao lado direito da Eq. (2.90), obtemos a relação

$$\sum_{m=-\infty}^{\infty} \delta(t - mT_0) = f_0 \sum_{n=-\infty}^{\infty} \exp(j2\pi n f_0 t) \quad (2.91)$$

Por outro lado, aplicando a transformada de Fourier ao lado esquerdo da Eq. (2.90), obtemos a relação dual

$$T_0 \sum_{m=-\infty}^{\infty} \exp(j2\pi m f T_0) = \sum_{n=-\infty}^{\infty} \delta(f - n f_0) \quad (2.92)$$

na qual utilizamos a relação da Eq. (2.82) para escrevê-la na forma $T_0 = 1/f_0$. As Eqs. (2.91) e (2.92) são duais uma da outra, com as funções delta mostradas no domínio do tempo na Eq. (2.91) e no domínio da freqüência na Eq. (2.92).

- **PROBLEMA 2.11** Utilizando a fórmula de Euler $x = 1/2[\exp(jx) + \exp(-jx)]$, reformule as Eqs. (2.91) e (2.92) em termos de funções cossenoidais. ◀

2.6 TRANSMISSÃO DE SINAIS ATRAVÉS DE SISTEMAS LINEARES: A CONVOLUÇÃO REVISTA

Com a teoria da transformada de Fourier apresentada nas seções anteriores, estamos prontos para voltarmos nossa atenção ao estudo de uma classe especial de sistemas conhecidos como *lineares*. Um sistema se refere a qualquer dispositivo ou fenômeno físico que produz um sinal de saída em resposta a um sinal de entrada. É comum referirmos ao sinal de entrada por *excitação* e ao sinal de saída como a *resposta*. Em um sistema *linear*, o princípio da *superposição* é válido. Ou seja, a *resposta de um sistema linear a um número de excitações aplicadas simultaneamente é igual à soma das respostas do sistema quando cada excitação é aplicada individualmente*. Exemplos importantes de sistemas lineares incluem *filtros* e *canais de comunicação* operando em suas regiões lineares. Um filtro se refere a um dispositivo seletivo em freqüência utilizado para limitar o espectro do sinal a alguma faixa de freqüências. Um canal se refere a uma mídia física que conecta o transmissor e o receptor de um sistema de comunicação. Queremos determinar os efeitos da transmissão de sinais através de filtros lineares e canais de comunicação. Esta determinação pode ser realizada de duas formas, dependendo de descrição adotada para o filtro ou canal. Ou seja, podemos utilizar características no domínio do tempo ou domínio da freqüência, como descrito a seguir.

■ Resposta no tempo

No domínio do tempo, um sistema linear é descrito em termos de sua *resposta ao impulso*, a qual é definida como a *resposta do sistema (com condições iniciais nulas) ao impulso unitário ou a função delta $\delta(t)$ aplicada à entrada do sistema*. Se o sistema for invariante no tempo, então esta propriedade implica em que um impulso unitário deslocado no tempo na entrada do sistema produz uma resposta ao impulso na saída, deslocada exatamente pelo mesmo valor. Em outras palavras, a forma da resposta ao impulso de um sistema linear invariante no tempo é a mesma, independente de quando o impulso é aplicado ao sistema. Portanto, assumindo que o impulso unitário ou a função delta é aplicada no tempo $t = 0$, podemos representar a resposta ao impulso do sistema linear invariante no tempo por $h(t)$. Seja este sistema sujeito a uma excitação arbitrária $x(t)$, como na Fig. 2.20(a). Para determinarmos a resposta $y(t)$ do sistema, começamos primeiro aproximando $x(t)$ por uma

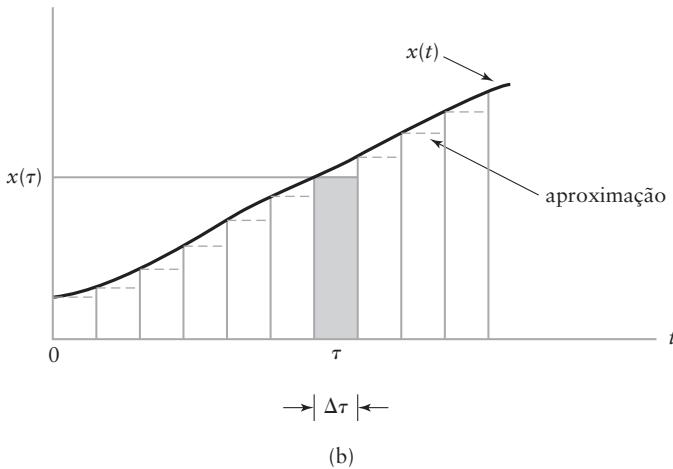
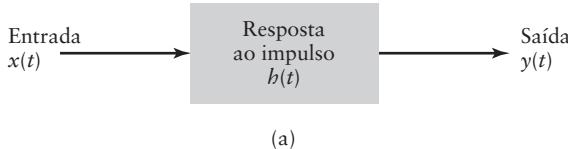


Figura 2.20 (a) Sistema linear com entrada $x(t)$ e saída $y(t)$. (b) Aproximação em degrau da entrada $x(t)$.

função em degrau composta por pulsos retangulares estreitos, cada um com duração $\Delta\tau$, como mostrado na Fig. 2.20(b). Claramente, a aproximação se torna melhor para $\Delta\tau$ pequeno. Quando $\Delta\tau$ se aproxima de zero, cada pulso se aproxima, no limite, à função delta ponderada pelo fator igual a altura do pulso vezes $\Delta\tau$. Considere um pulso típico, mostrado sombreado na Fig. 2.20(b), o qual ocorre em $t = \tau$. Este pulso possui área igual a $x(\tau)\Delta\tau$. Pela definição, a resposta do sistema ao impulso unitário, ou função delta $\delta(t)$ ocorrendo em $t = 0$ é $h(t)$. Temos, portanto, que a resposta do sistema a função delta, ponderada pelo fator $x(\tau)\Delta\tau$ e ocorrendo em $t = \tau$ deve ser $x(\tau)h(t - \tau)\Delta\tau$. Para determinarmos a resposta $y(t)$ em algum tempo t , aplicamos o princípio da superposição. Portanto, somando as várias respostas infinitesimais devido aos vários pulsos de entrada, obtemos, no limite, quando $\Delta\tau$ tende a zero,

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau) d\tau \quad (2.93)$$

Esta relação é chamada de *integral de convolução*.

Na Eq. (2.93), três escalas de tempo diferentes estão envolvidas: o tempo de excitação τ , o tempo de resposta t e o tempo de memória do sistema ($t - \tau$). Esta relação é a base da análise no domínio do tempo de sistemas lineares invariantes no tempo. Ela afirma que *o valor atual da resposta de um sistema linear invariante no tempo é a integral ponderada sobre a história passada do sinal de entrada ponderado de acordo com a resposta ao impulso do sistema*. Portanto, a resposta ao impulso funciona como *função de memória* do sistema.

Na Eq. (2.93), a excitação $x(t)$ é convoluída com a resposta ao impulso $h(t)$ para produzir a resposta $y(t)$. Como a convolução é comutativa, também podemos escrever

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t - \tau) d\tau \quad (2.94)$$

na qual $h(t)$ é convoluído com $x(t)$.

EXEMPLO 2.12 Filtro de linha de atraso

Considere um filtro linear invariante no tempo com resposta ao impulso $h(t)$. Fazemos duas considerações:

1. Casualidade, o que significa que a resposta ao impulso $h(t)$ é zero para $t < 0$.
2. Suporte finito, o que significa que a resposta ao impulso do filtro possui alguma duração finita T_f , de tal forma que podemos escrever $h(t) = 0$ para $t \geq T_f$.

Com estas duas considerações, podemos expressar a saída $y(t)$ do filtro, produzida em resposta à entrada $x(t)$, por

$$y(t) = \int_0^{T_f} h(\tau)x(t - \tau) d\tau \quad (2.95)$$

Seja a entrada $x(t)$, a resposta ao impulso $h(t)$ e a saída $y(t)$ uniformemente amostrados com uma taxa de $(1/\Delta\tau)$ amostras por segundo, podendo escrever

$$t = n \Delta\tau$$

e

$$\tau = k \Delta\tau$$

nas quais k e n são inteiros e $\Delta\tau$ é o período de amostragem. Assumindo que $\Delta\tau$ é suficientemente pequeno para que o produto $h(\tau)x(t - \tau)$ permaneça constante para $k \Delta\tau \leq \tau \leq (k + 1) \Delta\tau$ para todos os valores de k e τ , podemos aproximar a Eq. (2.95) pela soma de convolução mostrada por

$$y(n \Delta\tau) = \sum_{k=0}^{N-1} h(k \Delta\tau)x(n \Delta\tau - k \Delta\tau) \Delta\tau$$

na qual $N \Delta\tau = T_f$. Definindo o peso

$$w_k = h(k \Delta\tau) \Delta\tau, \quad k = 0, 1, \dots, N - 1$$

podemos reescrever a fórmula para $y(n\Delta\tau)$ como

$$y(n \Delta\tau) = \sum_{k=0}^{N-1} w_k x(n \Delta\tau - k \Delta\tau)$$

A Eq. (2.97) pode ser implementada utilizando a estrutura mostrada na Fig. 2.21, a qual é constituída de um conjunto de elementos de atraso (cada um produzindo um atraso de $\Delta\tau$ segundos), um conjunto de multiplicadores conectados à linha de atraso, um conjunto correspondente de pesos fornecidos aos multiplicadores e um somador para implementar a soma dos multiplicadores. Esta estrutura é conhecida como *filtro de linha de atraso* ou *filtro transversal*. Note que na Fig. 2.21 o espaçamento ou incremento básico do atraso é igual ao período de amostragem da seqüência de entrada $\{x(n\Delta\tau)\}$.

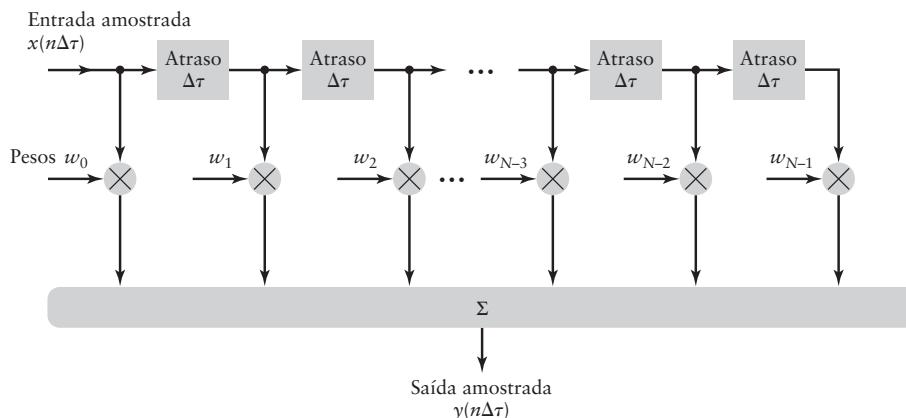


Figura 2.21 Filtro de linha de atraso.

■ Causalidade e estabilidade

Um sistema é dito ser *causal* se ele não responder antes da excitação ser aplicada. Para um sistema linear invariante no tempo ser causal, sua resposta $h(t)$ ao impulso deve desaparecer para tempos negativos, como afirmado no Exemplo 2.12. Ou seja, podemos afirmar formalmente que a condição necessária e suficiente para que um sistema linear invariante no tempo seja causal é

$$h(t) = 0, \quad t < 0 \quad (2.98)$$

Claramente, para um sistema operando em *tempo real* ser fisicamente realizável, ele deve ser causal. Entretanto, existem diversas aplicações nas quais o sinal a ser processado está disponível apenas em sua forma armazenada. Nestas situações, o sistema pode ser não causal e, mesmo assim, fisicamente realizável.

O sistema é dito ser estável se o sinal de saída é limitado para todo sinal de entrada limitado. Referimos a esta restrição como critério de estabilidade entrada limitada-saída limitada (BIBO*), a qual é bem adequada para a análise de sistema linear invariante no tempo. Seja o sinal de entrada $x(t)$ limitado, mostrado por:

$$|x(t)| < M \text{ para todo } t$$

na qual M é um número real positivo finito. Determinando o valor absoluto dos dois lados da Eq. (2.94), temos

$$|y(t)| = \left| \int_{-\infty}^{\infty} h(\tau)x(t - \tau) d\tau \right| \quad (2.99)$$

A seguir percebemos que o valor absoluto da integral é limitado pela integral do valor absoluto do integrando, mostrado por

$$\begin{aligned} \left| \int_{-\infty}^{\infty} h(\tau)x(t - \tau) d\tau \right| &\leq \int_{-\infty}^{\infty} |h(\tau)x(t - \tau)| d\tau \\ &= M \int_{-\infty}^{\infty} |h(\tau)| d\tau \end{aligned}$$

Logo, substituindo esta desigualdade na Eq. (2.99), obtemos o importante resultado

$$|y(t)| \leq M \int_{-\infty}^{\infty} |h(\tau)| d\tau$$

Temos, portanto, que para um sistema linear invariante no tempo ser estável, a resposta $h(t)$ ao impulso deve ser absolutamente integrável. Ou seja, a condição necessária e suficiente para estabilidade BIBO de um sistema linear invariante no tempo é descrita por

$$\int_{-\infty}^{\infty} |h(t)| dt < \infty \quad (2.100)$$

na qual $h(t)$ é a resposta ao impulso do sistema.

■ Resposta em freqüência

Considere, a seguir, um sistema linear invariante no tempo com resposta $h(t)$ ao impulso, a qual é obtida da entrada exponencial complexa de amplitude unitária e freqüência f , ou seja,

$$x(t) = \exp(j2\pi ft) \quad (2.101)$$

* N. de R. T.: Do inglês, *bounded input-bounded output*.

Usando as Eqs. (2.101) na Eq. (2.94), a resposta do sistema é obtida como

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau) \exp[j2\pi f(t - \tau)] d\tau \\ &= \exp(j2\pi ft) \int_{-\infty}^{\infty} h(\tau) \exp(-j2\pi f\tau) d\tau \end{aligned} \quad (2.102)$$

Define-se a *função de transferência* ou *resposta em freqüência* do sistema como a transformada de Fourier de sua resposta ao impulso, mostrado por

$$H(f) = \int_{-\infty}^{\infty} h(t) \exp(-j2\pi ft) dt \quad (2.103)$$

Os termos função de transferência e resposta em freqüência são utilizados um no lugar do outro. A integral na última linha da Eq. (2.102) é a mesma da Eq. (2.103), exceto pelo fato de que τ é utilizado no lugar de t . Logo, podemos reescrever a Eq. (2.102) na forma,

$$y(t) = H(f) \exp(j2\pi ft) \quad (2.104)$$

A Eq. (2.104) afirma que a resposta de um sistema linear invariante no tempo a uma função exponencial complexa com freqüência f é a mesma função exponencial complexa multiplicada pelo coeficiente constante $H(f)$.

A Eq. (2.103) é uma das definições de função de transferência $H(f)$. Uma definição alternativa de função de transferência pode ser deduzida dividindo-se a Eq. (2.104) pela (2.101), obtendo

$$H(f) = \frac{y(t)}{x(t)} \Big|_{x(t) = \exp(j2\pi ft)} \quad (2.105)$$

Considere, a seguir, um sinal arbitrário $x(t)$ aplicado ao sistema. O sinal $x(t)$ pode ser descrito em termos de sua transformada inversa de Fourier por

$$x(t) = \int_{-\infty}^{\infty} X(f) \exp(j2\pi ft) df \quad (2.106)$$

De forma equivalente, podemos expressar $x(t)$ como o limite de

$$x(t) = \lim_{\substack{\Delta f \rightarrow 0 \\ f=k}} \sum_{k=-\infty}^{\infty} X(f) \exp(j2\pi ft) \Delta f \quad (2.107)$$

ou seja, o sinal de entrada $x(t)$ pode ser visto como a superposição de exponenciais complexas com amplitude incremental. Como o sistema é linear, a resposta a esta superposição de entradas exponenciais complexas é dada por

$$\begin{aligned} y(t) &= \lim_{\substack{\Delta f \rightarrow 0 \\ f=k}} \sum_{k=-\infty}^{\infty} H(f)X(f) \exp(j2\pi ft) \Delta f \\ &= \int_{-\infty}^{\infty} H(f)X(f) \exp(j2\pi ft) df \end{aligned} \quad (2.108)$$

A transformada de Fourier do sinal de saída $y(t)$ é, portanto, facilmente obtida por

$$Y(f) = H(f)X(f) \quad (2.109)$$

De acordo com a Eq. (2.109), um sistema linear invariante no tempo pode, portanto, ser descrito de forma simples no domínio da freqüência observando que a transformada de

Fourier da saída é igual ao produto da resposta em freqüência do sistema pela transformada de Fourier da entrada.

Obviamente, poderíamos ter deduzido o resultado da Eq. (2.109) diretamente observando que:

1. A resposta $y(t)$ de um sistema linear invariante no tempo com resposta $h(t)$ ao impulso a uma entrada arbitrária $x(t)$ é obtida convoluindo-se $x(t)$ com $h(t)$, de acordo com a Eq. (2.93).
2. A convolução de um par de funções temporais é transformada na multiplicação de suas transformadas de Fourier.

A obtenção alternativa da Eq. (2.109) mostrada anteriormente é primariamente apresentada para desenvolver o entendimento de por que a representação de Fourier de uma função temporal como uma superposição de exponenciais complexas é tão importante na análise do comportamento de sistemas lineares invariantes no tempo.

A resposta em freqüência $H(f)$ é uma propriedade característica de um sistema linear invariante no tempo. Ela é, em geral, uma grandeza complexa e, portanto, podemos expressá-la na forma

$$H(f) = |H(f)| \exp[j\beta(f)] \quad (2.110)$$

Na qual $|H(f)|$ é chamada de *resposta em amplitude* ou *resposta em magnitude* e $\beta(f)$ é a *fase* ou *resposta de fase*. No caso especial de um sistema linear com resposta $h(t)$ ao impulso de valor real, a resposta em freqüência $H(f)$ exibe simetria conjugada, o que significa que

$$|H(f)| = |H(-f)|$$

e

$$\beta(f) = -\beta(-f)$$

ou seja, a resposta em amplitude $|H(f)|$ de um sistema linear com resposta ao impulso de valor real é uma função par da freqüência, enquanto que a fase $\beta(f)$ é uma função ímpar da freqüência.

Em algumas aplicações, é preferível trabalhar com o logaritmo de $H(f)$, expresso na forma polar, ao invés de diretamente com $H(f)$. Definindo o logaritmo natural

$$\ln H(f) = \alpha(f) + j\beta(f) \quad (2.111)$$

no qual

$$\alpha(f) = \ln|H(f)| \quad (2.112)$$

A função $\alpha(f)$ é uma definição de *ganho* do sistema. Ele é medido em *neper*, enquanto que a fase $\beta(f)$ é medida em radianos. A Eq. (2.111) indica que o ganho $\alpha(f)$ e fase $\beta(f)$ são as partes real e imaginária do logaritmo natural da resposta em freqüência $H(f)$, respectivamente. O ganho também pode ser expresso em *decibel* (dB), utilizando a definição

$$\alpha'(f) = 20 \log_{10}|H(f)| \quad (2.113)$$

As duas funções de ganho $\alpha(f)$ e $\alpha'(f)$ são relacionadas por

$$\alpha'(f) = 8,69\alpha(f) \quad (2.114)$$

ou seja, 1 neper é igual a 8,69 dB.

A partir da discussão apresentada na Seção 2.3, observamos que a largura de faixa do sistema é especificada pela resposta em amplitude constante. A largura de faixa de um sistema passa-baixa é, portanto, definida como a freqüência na qual a resposta em amplitude $|H(f)|$ é $1/\sqrt{2}$ vezes seu valor na freqüência zero ou, de forma equivalente, a freqüência na qual o ganho $\alpha'(f)$ cai para 3 dB abaixo de seu valor na freqüência zero, como ilustrado

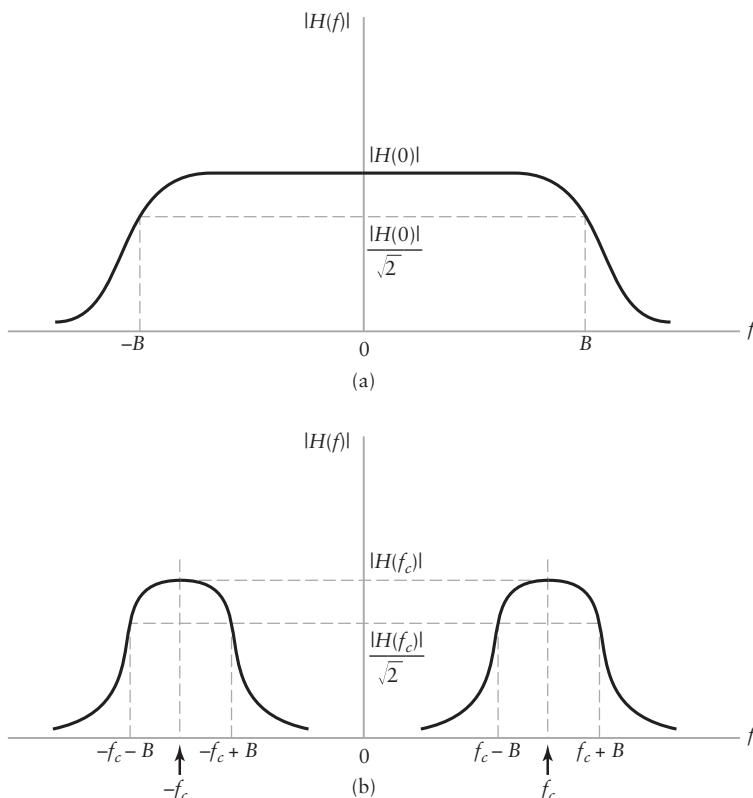


Figura 2.22 Exemplo da definição de largura de faixa do sistema.
(a) Sistema passa-baixa. (b) Sistema passa-alta.

na Fig. 2.22(a). De forma análoga, a largura de faixa de um sistema passa-faixa é definida como a faixa de freqüências nas quais a resposta de amplitude permanece $1/\sqrt{2}$ vezes seu valor na freqüência de meia faixa, como ilustrado na Fig. 2.22(b).

■ Critério de Paley-Wiener

A condição necessária e suficiente para que a função $\alpha(f)$ seja o ganho de um filtro causal é a convergência da integral

$$\int_{-\infty}^{\infty} \left(\frac{|\alpha(f)|}{1 + f^2} \right) df < \infty \quad (2.115)$$

Esta condição é conhecida como critério de Paley-Wiener. Ela afirma que, desde que o ganho $\alpha(f)$ satisfaça a condição da Eq. (2.115), então podemos associar a este ganho uma fase $\beta(f)$ adequada tal que o filtro resultante seja uma resposta ao impulso causal, ou seja, zero para valores negativos de tempo. Em outras palavras, o critério de Paley-Wiener é o equivalente no domínio da freqüência da condição de causalidade. Um sistema com característica de ganho realizable pode possuir atenuação infinita [isto é, $\alpha(f) = -\infty$] para um conjunto discreto de freqüências, mas ele não pode possuir atenuação infinita para uma faixa de freqüências ou, caso contrário, o critério de Paley-Wiener será violado.

► **PROBLEMAS 2.12** Discuta as seguintes questões, citando exemplo para as respostas:

- (a) É possível um sistema linear invariante no tempo ser causal, mas instável?
- (b) É possível tal sistema ser não causal, mas estável? ◀

- **PROBLEMA 2.13** A resposta ao impulso de um sistema linear é definido pela função Gaussiana

$$h(t) = \exp\left(-\frac{t^2}{2\tau^2}\right)$$

na qual τ é um parâmetro ajustável que define a duração do pulso. Determine a resposta em freqüência do sistema.

- **PROBLEMA 2.14** Um filtro de atraso de linha é constituído por N pesos, nos quais N é ímpar. Ele é simétrico com relação à linha central, ou seja, os pesos satisfazem a condição

$$w_n = w_{N-1-n}, \quad 0 \leq n \leq N - 1$$

- (a) Determine a resposta em amplitude do filtro.
- (b) Mostre que este filtro possui resposta de fase linear. Qual é a implicação dessa propriedade?
- (c) Qual é o atraso de tempo produzido pelo filtro?

2.7 FILTROS PASSA-BAIXA IDEALS

Como mencionado anteriormente, um *filtro* é um sistema seletivo em freqüência utilizado para limitar o espectro de um sinal em alguma faixa de freqüências especificadas. Sua resposta em freqüência é caracterizada por uma *faixa passante* e uma *faixa filtrada*. As freqüências dentro da faixa passante são transmitidas sem (ou com pouca) distorção, enquanto que aquelas na faixa filtrada são rejeitadas. O filtro pode ser *passa-baixa*, *passa-alta*, *passa-faixa* ou *pára-faixa*, dependendo se ele transmite freqüências baixa, alta, intermediária ou se ele rejeita freqüências intermediárias, respectivamente. Já vimos exemplos de sistemas passa-baixa e passa-alta na Fig. 2.22.

Filtros, de uma forma ou outra, representam um importante bloco funcional na construção de sistemas de comunicação. Neste livro estaremos voltados para o uso de filtros passa-alta, passa-baixa e passa-faixa.

Nesta seção estudaremos a resposta no tempo de um *filtro passa-baixa ideal*, o qual transmite, sem qualquer distorção, todas as freqüências dentro da faixa passante e rejeita completamente todas as freqüências dentro da faixa filtrada, como ilustrado na Fig. 2.23. De acordo com esta figura, a resposta em freqüência de um filtro passa-baixa ideal satisfaz duas condições necessárias:

1. A resposta em amplitude do filtro é uma constante dentro da faixa passante $-B \leq f \leq B$. (A constante na Fig. 2.23 é feita igual a um por conveniência da representação.)
2. A resposta em fase varia linearmente com a freqüência dentro da faixa passante do filtro. (Fora da faixa passante, a resposta em freqüência assume valores arbitrários.)

Em termos matemáticos, a função de transferência de um filtro passa-baixa ideal é, portanto, definida por

$$H(f) = \begin{cases} \exp(-j2\pi ft_0), & -B \leq f \leq B \\ 0, & |f| > B \end{cases} \quad (2.116)$$

O parâmetro B define a *largura de faixa* do filtro. O filtro passa-baixa ideal é, obviamente, não causal, pois ele viola o critério de Paley-Wiener. Esta observação também pode ser confirmada examinando-se a resposta $h(t)$ ao impulso. Portanto, calculando-se a transformada inversa de Fourier da Eq. (2.116), obtemos

$$h(t) = \int_{-B}^B \exp[j2\pi f(t - t_0)] df \quad (2.117)$$

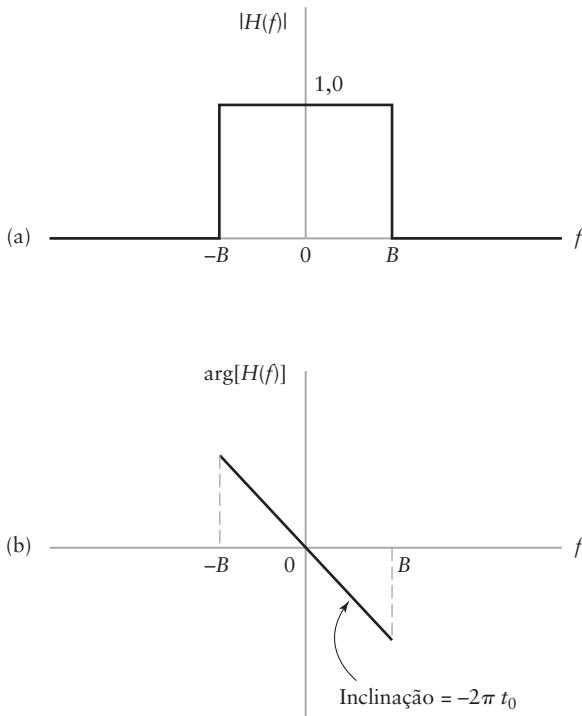


Figura 2.23 Resposta em freqüência de um filtro passa-baixa ideal. (a) Resposta em amplitude. (b) Resposta em fase. Fora da faixa $-B \leq f \leq B$ a resposta em fase assume uma forma arbitrária (não mostrada na Figura).

na qual os limites de integração foram reduzidos para a faixa de freqüências dentro da qual $H(f)$ não desaparece. A Eq. (2.117) é facilmente integrada, resultando em

$$\begin{aligned} h(t) &= \frac{\sin[2\pi B(t - t_0)]}{\pi(t - t_0)} \\ &= 2B \operatorname{sinc}[2B(t - t_0)] \end{aligned} \quad (2.118)$$

A resposta ao impulso possui uma amplitude de pico de $2B$ centrada no tempo t_0 , como mostrado na Fig. 2.24 para $t_0 = 1/B$. A duração do lóbulo principal da resposta ao impulso é $1/B$ e o tempo necessário do zero no começo do lóbulo principal até o valor de pico é $1/2B$. Podemos observar na Fig. 2.24 que, para qualquer valor finito de t_0 , existe

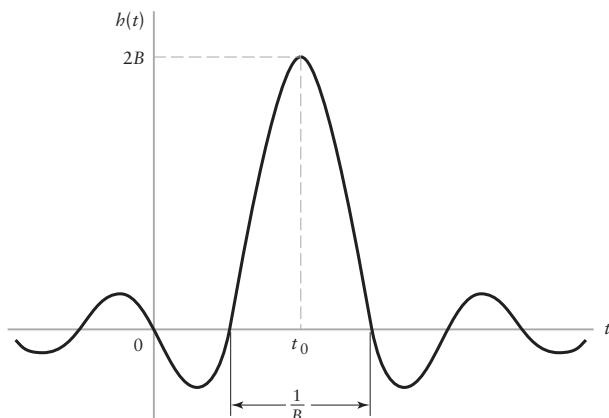


Figura 2.24 Resposta ao impulso de um filtro passa-baixa ideal.

alguma resposta do filtro antes do tempo $t = 0$ no qual o impulso unitário é aplicado à entrada. Esta observação confirma que o filtro passa-baixa ideal é não causal. Note, entretanto, que podemos sempre fazer com que o atraso t_0 seja grande o suficiente para que a condição

$$|\text{sinc}[2B(t - t_0)]| \ll 1, \quad \text{para } t < 0$$

seja satisfeita. Fazendo isso, estaremos aptos a construir um filtro causal que se aproxima de um filtro passa-baixa ideal, com uma aproximação cada vez melhor a medida que reduzimos o atraso t_0 .

■ Resposta ao pulso de filtros passa-baixa ideais

Considere o pulso retangular $x(t)$ com amplitude unitária e duração T , o qual é aplicado a um filtro passa-baixa ideal com largura de faixa B . O problema é determinar a resposta $y(t)$ do filtro.

A resposta $h(t)$ ao impulso do filtro é definida pela Eq. (2.118). Claramente, o atraso t_0 não possui efeito na forma da resposta $y(t)$ do filtro. Sem perda de generalidade, podemos, portanto, simplificar a exposição fazendo $t_0 = 0$. Neste caso, a resposta ao impulso da Eq. (2.118) se reduz para

$$h(t) = 2B \text{sinc}(2Bt) \quad (2.119)$$

Com a entrada $x(t) = 1$ para $-(T/2) \leq t \leq (T/2)$, a resposta resultante do filtro é dada pela integral de convolução

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} x(\tau)h(t - \tau) d\tau \\ &= 2B \int_{-T/2}^{T/2} \text{sinc}[2B(t - \tau)] d\tau \\ &= 2B \int_{-T/2}^{T/2} \left(\frac{\text{sen}[2\pi B(t - \tau)]}{2\pi B(t - \tau)} \right) d\tau \end{aligned} \quad (2.120)$$

Definindo uma nova variável dimensional

$$\lambda = 2\pi B(t - \tau)$$

e, alterando a variável de integração de τ para λ , podemos escrever a Eq. (2.120) como

$$\begin{aligned} y(t) &= \frac{1}{\pi} \int_{2\pi B(t-T/2)}^{2\pi B(t+T/2)} \left(\frac{\text{sen } \lambda}{\lambda} \right) d\lambda \\ &= \frac{1}{\pi} \left[\int_0^{2\pi B(t+T/2)} \left(\frac{\text{sen } \lambda}{\lambda} \right) d\lambda - \int_0^{2\pi B(t-T/2)} \left(\frac{\text{sen } \lambda}{\lambda} \right) d\lambda \right] \\ &= \frac{1}{\pi} \{ \text{Si}[2\pi B(t + T/2)] - \text{Si}[2\pi B(t - T/2)] \} \end{aligned} \quad (2.121)$$

Na Eq. (2.121), introduzimos uma nova expressão chamada de *integral de seno*, a qual é definida por

$$\text{Si}(u) = \int_0^u \frac{\text{sen } x}{x} dx \quad (2.122)$$

Infelizmente, a integral de seno $\text{Si}(u)$ não pode ser calculada analiticamente, em termos de funções elementares. Entretanto, ela pode ser integrada como uma série de potênc-

cias, a qual, por sua vez, resulta no gráfico mostrado na Fig. 2.25. A partir desta figura, podemos fazer três observações:

1. A integral de seno $\text{Si}(u)$ é uma função oscilatória de u , possuindo simetria ímpar com relação a origem $u = 0$.
2. Ela possui máximos e mínimos para múltiplos de π .
3. Ela se aproxima do valor limite ($\pi/2$) para grandes valores positivos de u .

Na Fig. 2.25 vemos que a integral de seno $\text{Si}(u)$ oscila com freqüência $1/2\pi$. Portanto, a resposta $y(t)$ do filtro também irá oscilar com freqüência igual à freqüência de corte (isto é, largura de faixa) B do filtro passa-baixa, como indicado na Fig. 2.26. O valor máximo de $\text{Si}(u)$ ocorre em $u_{\text{máx}} = \pi$, sendo igual a

$$1,8519 = (1,179) \times \left(\frac{\pi}{2} \right)$$

Também podemos mostrar que a resposta $y(t)$ do filtro possui máximo e mínimo em

$$t_{\text{máx}} = \pm \frac{T}{2} \pm \frac{1}{2B}$$

com

$$\begin{aligned} y(t_{\text{máx}}) &= \frac{1}{\pi} [\text{Si}(\pi) - \text{Si}(\pi - 2\pi BT)] \\ &= \frac{1}{\pi} [\text{Si}(\pi) + \text{Si}(2\pi BT - \pi)] \end{aligned}$$

logo, na segunda linha, utilizamos a propriedade da simetria ímpar da integral de seno. Seja

$$\text{Si}(2\pi BT - \pi) = \frac{\pi}{2}(1 \pm \Delta)$$

na qual Δ é o valor absoluto do desvio no valor de $\text{Si}(2\pi BT - \pi)$ descrito como uma fração do valor final $+\pi/2$. Portanto, percebendo que

$$\text{Si}(\pi) = (1,179)(\pi/2)$$

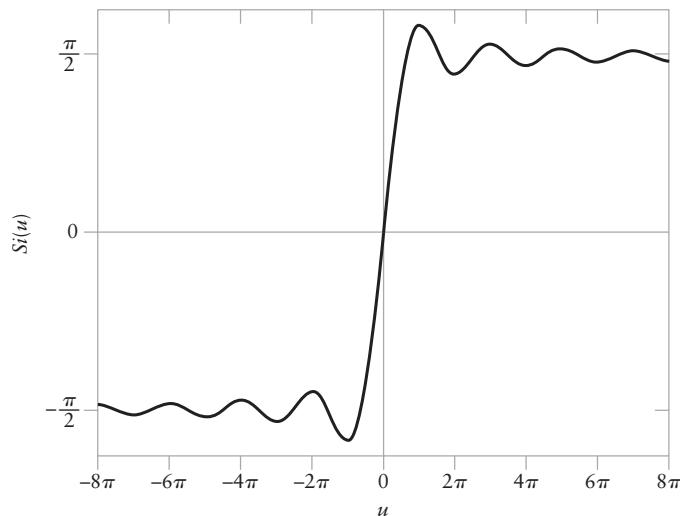


Figura 2.25 A integral de seno $\text{Si}(u)$.

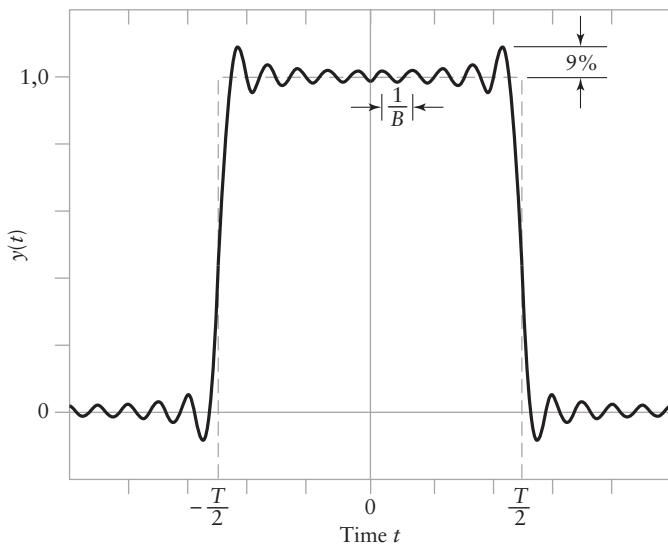


Figura 2.26 Resposta de um filtro passa-baixa ideal a um pulso quadrado.

podemos redefinir $y(t_{\max})$ para

$$\begin{aligned} y(t_{\max}) &= \frac{1}{2}(1,179 + 1 \pm \Delta) \\ &\approx 1,09 \pm \frac{1}{2}\Delta \end{aligned} \quad (2.123)$$

Para um produto tempo-largura de faixa limitado $BT \gg 1$, o desvio fracional Δ possui valor muito pequeno e, neste caso, podemos fazer duas observações importantes a partir da Eq. (2.123):

1. O sobressinal percentual da resposta do filtro é aproximadamente 9%.
2. O sobressinal é praticamente independente da largura de faixa B do filtro.

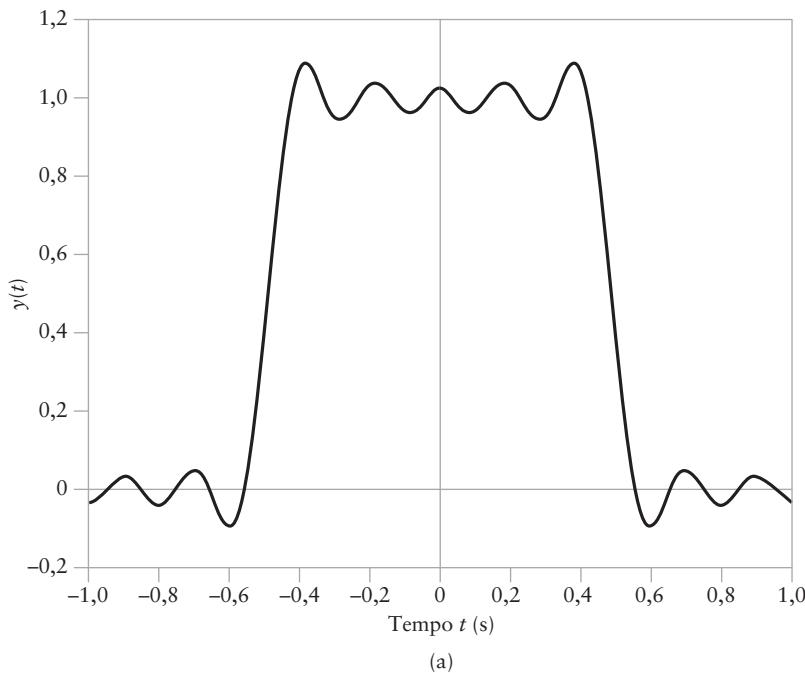
O fenômeno básico destas duas observações é chamado de *fenômeno de Gibbs*. A Fig. 2.26 mostra a natureza oscilatória da resposta do filtro e o sobressinal de 9% caracterizando a resposta, assumindo que $BT \gg 1$.

A Figura 2.27, ocupando as páginas 81 e 82, mostra a resposta do filtro para quatro produtos tempo-largura de faixa: $BT = 5, 10, 20$ e 100 , assumindo que a duração T do pulso é 1 segundo. A Tabela 2.1 mostra as freqüências correspondentes de oscilação e o sobressinal para estes produtos tempo-largura de faixa, confirmando as observações 1 e 2.

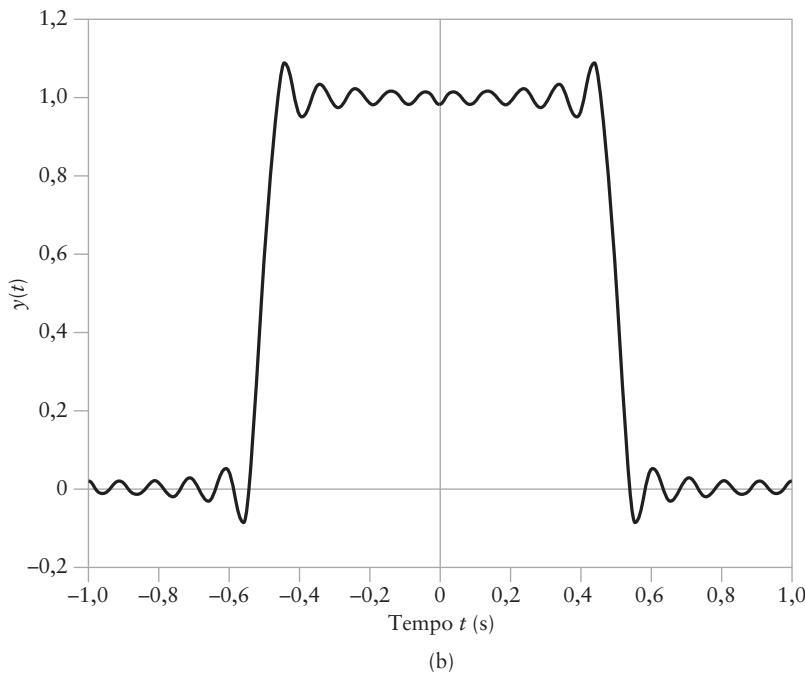
TABELA 2.1 Freqüência de oscilação e percentual de sobressinal para produto tempo-largura de faixa variando

BT	Freqüência de oscilação	Percentual de sobressinal
5	5 Hz	9,11
10	10 Hz	8,98
20	20 Hz	8,99
100	100 Hz	9,63

A Fig. 2.28, ocupando as páginas 83 e 84, mostra a resposta do filtro para entradas periódicas quadradas com freqüências fundamentais diferentes: $f_0 = 0,1; 0,25; 0,5$ e 1 Hz

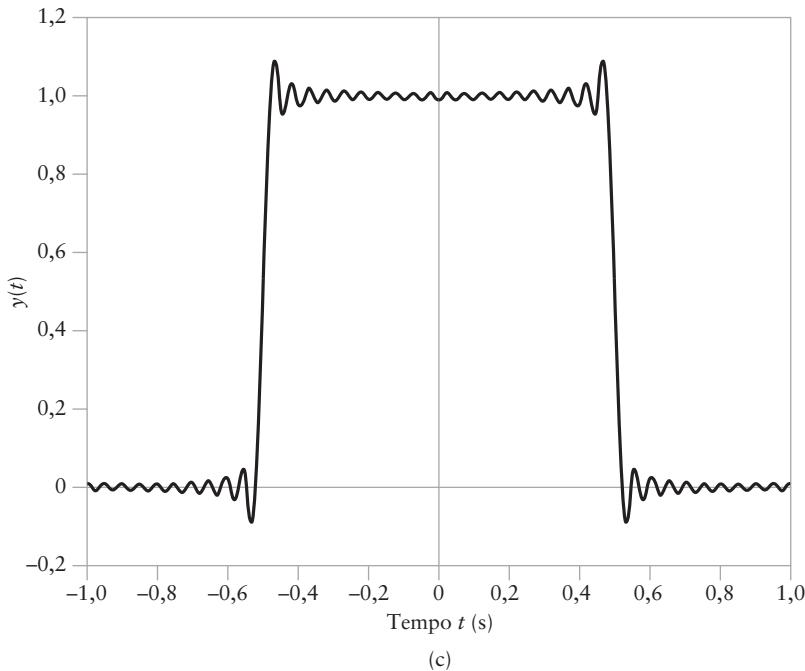


(a)

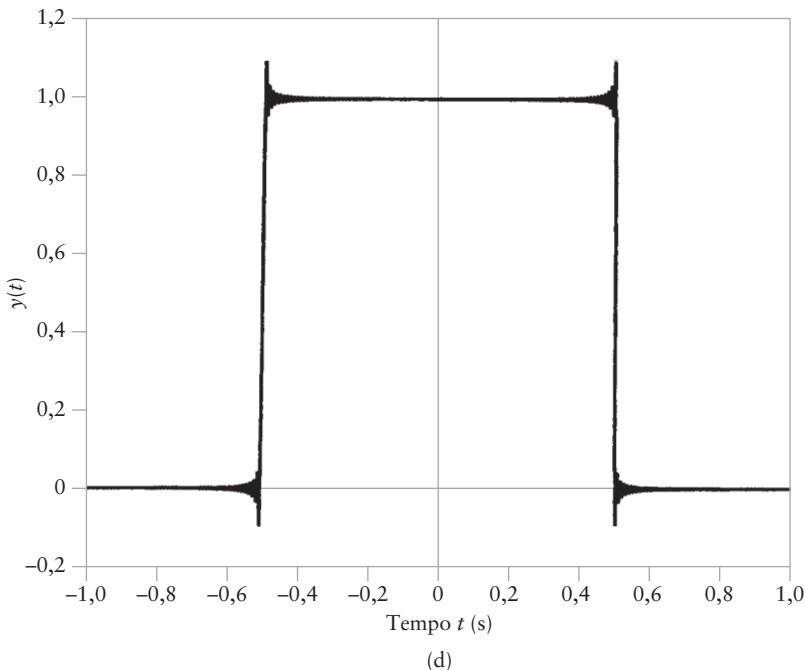


(b)

Figura 2.27 Resposta ao pulso de um filtro passa-baixa ideal para um pulso de duração $T = 1\text{s}$ e variando o produto tempo-largura de faixa (BT). (a) $BT = 5$, (b) $BT = 10$.

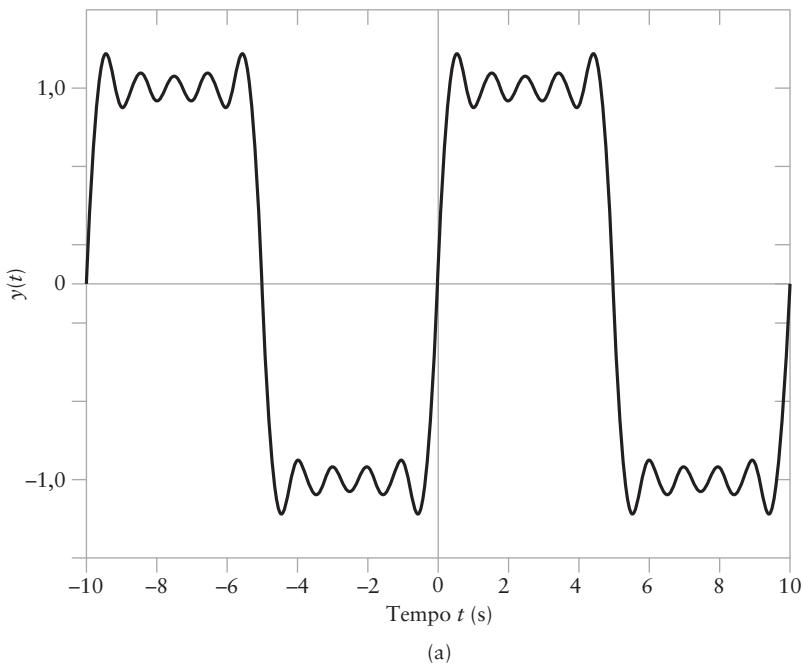


(c)

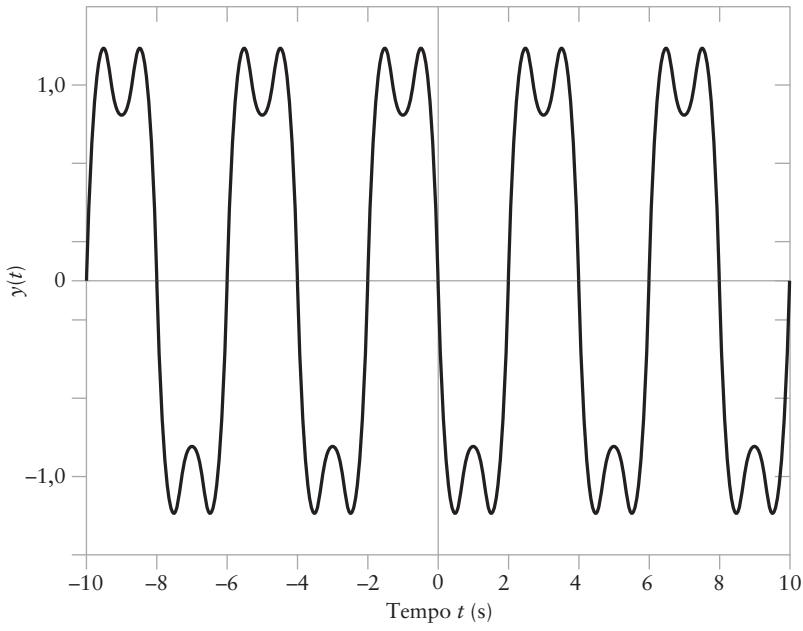


(d)

Figura 2.27 (continuação) (c) $BT = 20$. (d) $BT = 100$.

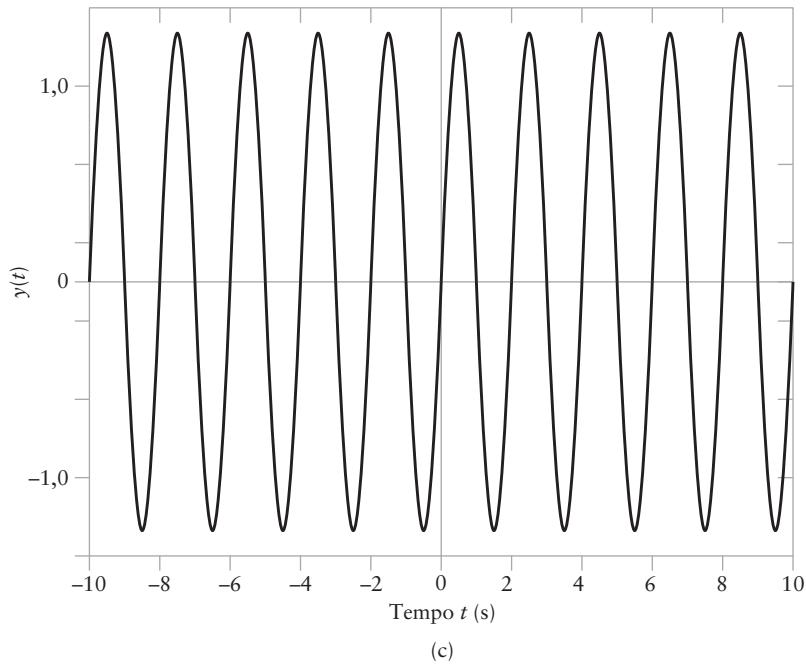


(a)

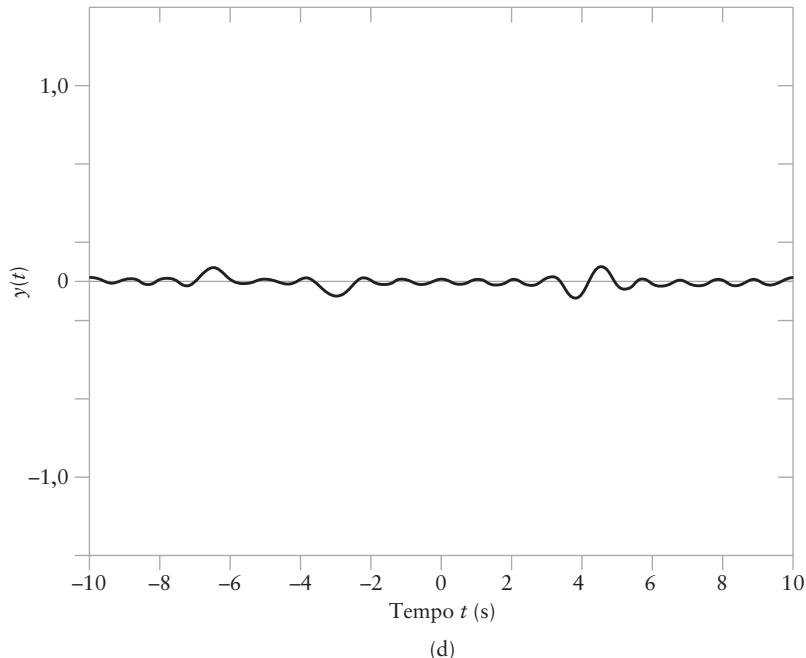


(b)

Figura 2.28 Resposta de um filtro passa-baixa ideal a uma forma de onda quadrada com frequência f_0 variando. (a) $f_0 = 0,1\text{Hz}$. (b) $f_0 = 0,25\text{Hz}$.



(c)



(d)

Figura 2.28 (continuação) (c) $f_0 = 0,5\text{Hz}$. (d) $f_0 = 1\text{Hz}$.

e com largura de faixa do filtro passa-baixa limitada em $B = 1\text{Hz}$. A partir da Fig. 2.28 podemos fazer as seguintes observações:

- ▶ Para $f_0 = 0,1\text{Hz}$, correspondendo a um produto tempo-largura de faixa $BT = 5$, o filtro distorce um pouco o pulso quadrado de entrada, mas a forma da entrada ainda é evidente na saída do filtro. Ao contrário da entrada, a saída do filtro possui tempos de subida e descida não nulos que são inversamente proporcionais à largura de faixa do filtro. Além disso, a saída exibe oscilações tanto antes da borda de subida quanto após a borda de descida.
- ▶ Quando a freqüência fundamental f_0 da forma de onda quadrada de entrada aumenta, o filtro passa-baixa corta mais as componentes de alta freqüência da entrada. Portanto, quando $f_0 = 0,25\text{Hz}$, correspondendo a $BT = 2$, apenas a freqüência fundamental e a componente de primeira harmônica passam pelo filtro. Os tempos de subida e descida da saída são significativos quando comparados com a duração do pulso de entrada. Quando $f_0 = 0,5\text{Hz}$, correspondendo a $BT = 1$, apenas a componente de freqüência fundamental da forma de onda quadrada de entrada é preservada pelo filtro, resultando em uma saída essencialmente senoidal.
- ▶ Quando a freqüência fundamental da forma de onda de entrada é aumentada ainda mais para o alto valor $f_0 = 1\text{Hz}$, que corresponde ao produto tempo-largura de faixa $BT = 0,5$, a componente CC se torna a saída dominante, e a forma da onda quadrada de entrada é completamente destruída pelo filtro.

A partir destes resultados, podemos escrever uma importante conclusão: *Quando utilizamos um filtro passa-baixa ideal, devemos utilizar um produto tempo-largura de faixa $BT \geq 1$ para garantir que a forma de onda da entrada do filtro seja reconhecida na saída resultante.* Um valor de BT maior do que a unidade tende a reduzir o tempo de subida e o tempo de descida da resposta ao pulso do filtro.

■ Aproximação de filtros passa-baixa ideais

Um filtro pode ser caracterizado pela especificação de sua resposta $h(t)$ ao impulso ou, de forma equivalente, de sua função de transferência $H(f)$. Entretanto, a aplicação de um filtro geralmente envolve a separação dos sinais com base em seu espectro (isto é, conteúdos de freqüência). Isto, por sua vez, significa que o projeto de filtros geralmente ocorre no domínio da freqüência. Existem, basicamente, dois passos necessários ao projeto de um filtro:

1. A *aproximação* de uma resposta em freqüência desejada (isto é, resposta em amplitude, resposta em fase ou ambas) por uma função de transferência realizável.
2. A *realização* da função de transferência aproximada por um dispositivo físico.

Para uma aproximação de uma função de transferência $H(f)$ ser fisicamente realizável, ela deve representar um sistema estável. A estabilidade é definida com base no critério entrada limitada-saída limitada, descrito pela Eq. (2.110), a qual envolve na resposta $h(t)$ ao impulso. Para especificar a condição correspondente de estabilidade em termos da função de transferência, a abordagem tradicional é substituir $j2\pi f$ por s e reescrever a função de transferência em termos de s . A nova variável s pode possuir parte real e parte imaginária. Desta forma nos referimos a s como *freqüência complexa*. Seja $H'(s)$ a representação da função de transferência do sistema, definida na forma descrita aqui. Ordinariamente, a função de transferência aproximada $H'(s)$ é uma função racional, a qual, portanto, pode ser descrita em uma forma fatorada

$$\begin{aligned} H'(s) &= H(f)|_{j2\pi f=s} \\ &= K \frac{(s - z_1)(s - z_2) \cdots (s - z_m)}{(s - p_1)(s - p_2) \cdots (s - p_n)} \end{aligned}$$

na qual K é chamado de fator de escala; z_1, z_2, \dots, z_m são chamados de *zeros* da função de transferência e p_1, p_2, \dots, p_n são chamados de *pólos*. Para uma função de transferência

passa-baixa, o número m de zeros é menor do que o número n de pólos. Se o sistema for causal, então a condição de entrada limitada-saída limitada para a estabilidade do sistema é satisfeita restringindo-se todos os pólos da função de transferência $H'(s)$ a permanecerem no lado esquerdo do plano s , ou seja,

$$\operatorname{Re}([p_i]) < 0, \quad \text{para todo } i$$

Note que a condição para a estabilidade envolve apenas os pólos da função de transferência $H'(s)$. Os zeros podem estar em qualquer local do plano s . Dois tipos de sistemas podem ser distinguidos, dependendo da localização dos m zeros no plano s :

- ▶ *Sistemas de fase mínima*, caracterizados por uma função de transferência cujos pólos e zeros estejam todos dentro do lado esquerdo do plano s .
- ▶ *Sistemas de fase não mínima*, cuja função de transferência pode possuir zeros no eixo imaginário e no lado direito do plano s .

Sistemas de fase mínima distinguem deles mesmos pela propriedade de que a resposta de fase desta classe de sistema linear invariante no tempo é unicamente relacionada com a resposta do ganho.

No caso de filtros passa-baixa, onde a principal restrição é aproximar a resposta em amplitude ideal mostrada na Fig. 2.23, podemos mencionar duas famílias de filtros: Filtros de Butterworth e Filtros de Chebyshev. Os dois possuem todos os seus zeros em $s = \infty$. Em um filtro de Butterworth, os pólos da função de transferência $H'(s)$ estão em um círculo com o centro na origem e raio igual a $2\pi B$, na qual B é a largura de faixa de 3dB do filtro. Em um filtro Chebyshev, por outro lado, os pólos estão em uma elipse. Nos dois casos, obviamente, os pólos são confinados ao lado esquerdo do plano s .

Voltando nossa atenção à questão de realização física do filtro, vemos que existem duas opções básicas para esta realização. Uma analógica e outra digital:

- ▶ *Filtros analógicos*, construídos usando (a) indutores e capacitores ou (b) capacitores, resistores e amplificadores operacionais. A vantagem dos filtros analógicos é a simplicidade de implementação.
- ▶ *Filtros digitais*, para os quais os sinais são amostrados no tempo e suas amplitudes quantizadas. Estes filtros são construídos usando hardware digital, por isso o nome. Uma característica importante de um filtro digital é que ele é programável, oferecendo, portanto, um alto grau de flexibilidade no projeto. De fato, existe um compromisso entre complexidade e flexibilidade.

2.8 CORRELAÇÃO E DENSIDADE ESPECTRAL: SINAIS DE ENERGIA

Nesta seção, iremos continuar a caracterizar sinais e sistemas considerando a classe de sinais de energia e, portanto, focalizando na noção de *energia*. (A caracterização de sinais e sistemas é completada na Seção 2.9, na qual iremos considerar outra classe de sinais, chamados de sinais de potência.) Em particular, apresentamos um novo parâmetro chamado de *densidade espectral*, a qual é definida como o quadrado do espectro de amplitude de um sinal de interesse. Vemos que a densidade espectral é a transformada de Fourier da função de correlação, a qual foi mostrada na Propriedade 13 na Seção 2.2.

■ Função de autocorrelação

Assuma um sinal de energia $x(t)$ que, para o propósito de generalidade, é considerado como um valor complexo. Segundo o material apresentado no teorema da correlação (Propriedade 13), na Seção 2.2, definimos formalmente a *função de autocorrelação* do sinal de energia $x(t)$ para o *atraso* τ como

$$R_x(\tau) = \int_{-\infty}^{\infty} x(t)x^*(t - \tau) dt \quad (2.124)$$

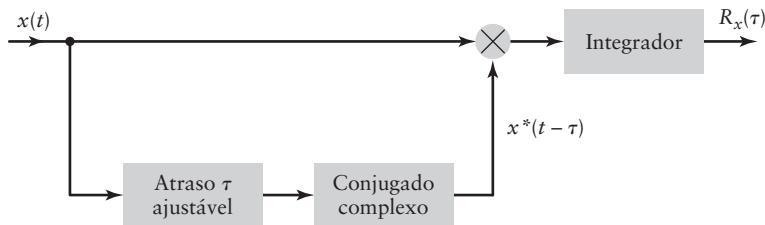


Figura 2.29 Esquema para a medição da função de autocorrelação $R_x(\tau)$ de um sinal de energia $x(t)$ para um atraso τ .

De acordo com esta fórmula, a função de autocorrelação $R_x(\tau)$ fornece uma medida da similaridade entre o sinal $x(t)$ e sua versão atrasada $x(t - \tau)$. E, como tal, pode ser medida usando o arranjo mostrado na Fig. 2.29. O tempo de atraso τ possui o papel de *varrer* ou de ser uma *variável de busca*. Note que $R_x(\tau)$ é um valor complexo se $x(t)$ for de valor complexo.

Da Eq. (2.124), podemos facilmente ver que o valor da função de autocorrelação $R_x(\tau)$ para $\tau = 0$ é igual a energia do sinal $x(t)$, ou seja,

$$R_x(0) = \int_{-\infty}^{\infty} |x(t)|^2 dt$$

■ Densidade espectral de energia

O teorema de Rayleigh para energia, discutido na Propriedade 14, Seção 2.2, é importante porque ele não somente fornece um método útil para o cálculo da energia de um sinal de pulso, mas também ressalta o quadrado do espectro de amplitude como a distribuição de energia do sinal medida no domínio da freqüência. É à luz deste teorema que definimos, formalmente, a *densidade espectral de energia* ou *espectro de densidade de energia* de um sinal de energia $x(t)$, dada por

$$\psi_x(f) = |X(f)|^2 \quad (2.125)$$

na qual $|X(f)|$ é o espectro de amplitude de $x(t)$. Claramente, a densidade espectral de energia $\psi_x(f)$ é uma quantidade de valor real não negativa para todo f , mesmo que o sinal $x(t)$ seja, ele próprio, de valor complexo.

■ Relações de Wiener-Khitchine para sinais de energia

Referindo ao teorema da correlação, descrito na Eq. (2.53), seja $g_1(t) = g_2(t) = x(t)$, no qual $x(t)$ é um sinal de energia e, portanto, possui transformada de Fourier. Sob esta condição, o lado esquerdo resultante da Eq. (2.53) define a função de autocorrelação $R_x(\tau)$ do sinal $x(t)$. No domínio do tempo, de forma correspondente, temos $G_1(f) = G_2(f) = X(f)$ e, neste caso, o lado direito da Eq. (2.53) define a densidade de energia espectral $\psi_x(f)$. Com esta base, podemos, portanto, afirmar que *dado um sinal de energia $x(t)$, a função de autocorrelação $R_x(\tau)$ e a densidade espectral de energia $\psi_x(f)$ formam um par transformada de Fourier*. Especificamente, temos o par de relações:

$$\psi_x(f) = \int_{-\infty}^{\infty} R_x(\tau) \exp(-j2\pi f\tau) d\tau \quad (2.126)$$

e

$$R_x(\tau) = \int_{-\infty}^{\infty} \psi_x(f) \exp(j2\pi f\tau) df \quad (2.127)$$

Note, entretanto, que a transformação de Fourier da Eq. (2.126) é realizada com relação à variável de atraso τ . O par de Equações (2.126) e (2.127) constitui as *relações de Wiener-Khitchine para sinais de energia*.

Das Eqs. (2.126) e (2.127), podemos facilmente deduzir as duas propriedades:

1. Fazendo $f = 0$ na Eq. (2.126), temos

$$\int_{-\infty}^{\infty} R_x(\tau) d\tau = \psi_x(0)$$

a qual afirma que *a área total sob a curva de uma função de autocorrelação de valor complexo de um sinal de energia de valor complexo é igual à energia espectral $\psi_x(0)$ de valor real na freqüência zero*.

2. Fazendo $\tau = 0$ na Eq. (2.127), temos

$$\int_{-\infty}^{\infty} \psi_x(f) df = R_x(0)$$

a qual afirma que *a área total sob a curva de uma densidade espectral de energia de valor real de um sinal de energia é igual a energia total do sinal*. Este segundo resultado é meramente outra forma de descrever o teorema de Rayleigh para energia.

EXEMPLO 2.13 Função de autocorrelação do pulso sinc

Do Exemplo 2.4, a transformada de Fourier do pulso sinc

$$x(t) = A \operatorname{sinc}(2Wt)$$

é dada por

$$X(f) = \frac{A}{2W} \operatorname{ret}\left(\frac{f}{2W}\right)$$

Como a função retangular $\operatorname{ret}(f/2W)$ não é afetada pelo quadrado, a densidade espectral de energia de $x(t)$ é, portanto,

$$\psi_x(f) = \left(\frac{A}{2W}\right)^2 \operatorname{ret}\left(\frac{f}{2W}\right)$$

Determinando a transformada inversa de Fourier de $\psi_x(f)$, obtemos a função de autocorrelação do pulso sinc $A \operatorname{sinc}(2Wt)$, dada por

$$R_x(\tau) = \frac{A^2}{2W} \operatorname{sinc}(2W\tau) \quad (2.128)$$

a qual é uma forma de onda similar ao próprio pulso sinc, traçada em função do atraso τ .

Este exemplo mostra que algumas vezes é mais fácil utilizar um procedimento indireto baseado na densidade espectral de energia para determinarmos a função de autocorrelação de um sinal de energia do que utilizarmos diretamente a fórmula para a função de autocorrelação.

Efeito da filtragem na densidade espectral de energia

Suponha, agora, que o sinal de energia $x(t)$ é passado através de um sistema linear invariante no tempo com função de transferência $H(f)$, resultando no sinal de saída $y(t)$, como

mostrado na Fig. 2.20(a). Então, de acordo com a Eq. (2.109), a transformada de Fourier da saída $y(t)$ é relacionada com a transformada de Fourier da entrada $x(t)$ por

$$Y(f) = H(f)X(f)$$

Calculando o quadrado da amplitude dos dois lados desta equação, obtemos

$$\psi_y(f) = |H(f)|^2\psi_x(f) \quad (2.129)$$

na qual, pela definição, $\psi_x(f) = |X(f)|^2$ e $\psi_y(f) = |Y(f)|^2$. A Eq. (2.129) afirma que *quando um sinal de energia é transmitido através de um filtro linear invariante no tempo, a densidade espectral de energia da saída resultante é igual à densidade espectral de energia da entrada multiplicada pelo quadrado da resposta em amplitude do filtro*. A simplicidade desta afirmação enfatiza a importância da densidade espectral como parâmetro para caracterização da distribuição de energia em um sinal no domínio da freqüência.

Além disso, com base nas equações de Wiener-Khintchine (2.126) e (2.127) e da relação da Eq. (2.129), podemos descrever um *método indireto* para a determinação do efeito da filtragem linear invariante no tempo na função de autocorrelação de um sinal de energia:

1. Determine as transformadas de Fourier de $x(t)$ e $h(t)$, obtendo $X(f)$ e $H(f)$, respectivamente.
2. Utilize a Eq. (2.129) para determinar a densidade espectral de energia, $\psi_y(f)$, da saída $y(t)$.
3. Determine $R_y(\tau)$ aplicando a transformada inversa de Fourier a $\psi_y(f)$ obtida no ponto 2.

EXEMPLO 2.14 Energia da versão filtrada passa-baixa do pulso retangular

Um pulso retangular com amplitude e duração unitária é passado através de um filtro passa-baixa ideal com largura de faixa B , como ilustrado na Fig. 2.30(a). A parte (b) desta figura mostra a forma de onda do pulso retangular. A resposta em amplitude do filtro é definida por (veja a Fig. 2.30(c))

$$|H(f)| = \begin{cases} 1, & -B \leq f \leq B \\ 0, & \text{caso contrário} \end{cases}$$

O pulso retangular de entrada do filtro possui energia unitária. Queremos determinar o efeito da variação de B na energia da saída do filtro.

Começamos com o par transformada de Fourier:

$$\text{ret}(t) \iff \text{sinc}(f)$$

a qual representa a versão normalizada do par transformada de Fourier dado pela Eq. (2.10). Logo, com a entrada do filtro definida por

$$x(t) = \text{ret}(t)$$

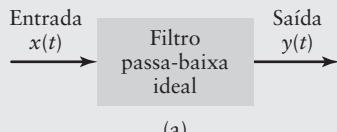
sua transformada de Fourier é igual a

$$X(f) = \text{sinc}(f)$$

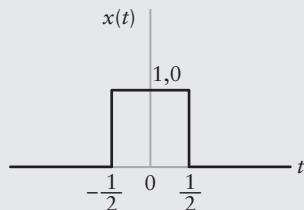
A densidade espectral de energia da entrada do filtro é, portanto,

$$\begin{aligned} \psi_x(f) &= |X(f)|^2 \\ &= \text{sinc}^2(f) \end{aligned} \quad (2.130)$$

Esta densidade espectral de energia normalizada é apresentada na Fig. 2.31.



(a)



(b)

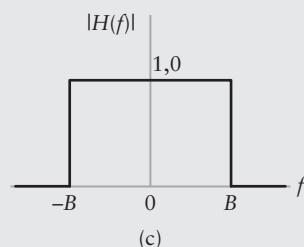


Figura 2.30 (a) Filtragem passa-baixa ideal.
 (b) Entrada do filtro.
 (c) Resposta em amplitude do filtro.

Para calcularmos a densidade espectral de energia $\psi_x(f)$ da saída $y(t)$ do filtro, utilizamos a Eq. (2.129), obtendo,

$$\begin{aligned}\psi_y(f) &= |H(f)|^2 \psi_x(f) \\ &= \begin{cases} \psi_x(f), & -B \leq f \leq B \\ 0, & \text{caso contrário} \end{cases}\end{aligned}\quad (2.131)$$

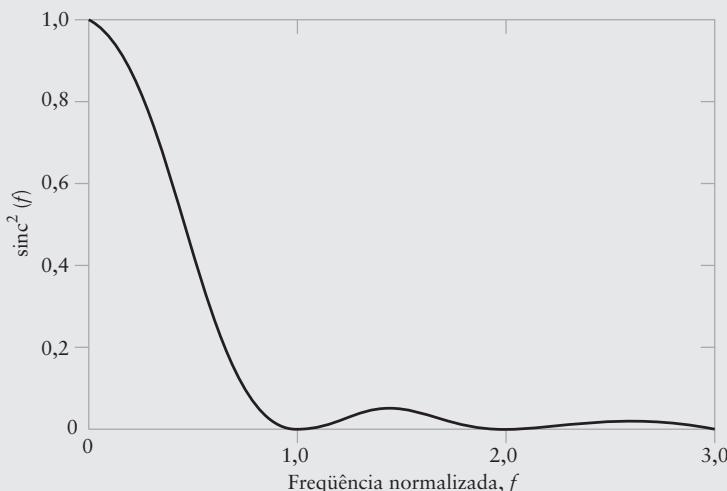


Figura 2.31 Densidade espectral de energia da entrada $x(t)$ do filtro.
 Apenas os valores para freqüência positiva são mostrados na figura.

A energia da saída do filtro é, portanto,

$$\begin{aligned}
 E_y &= \int_{-\infty}^{\infty} \psi_y(f) df \\
 &= \int_{-B}^{B} \psi_x(f) df \\
 &= 2 \int_0^B \psi_x(f) df \\
 &= 2 \int_0^B \operatorname{sinc}^2(f) df
 \end{aligned} \tag{2.132}$$

Como a entrada do filtro possui energia unitária, podemos analisar o resultado dado pela Eq. (2.132) como a *relação de energia da saída do filtro pela energia da entrada do filtro*, para o caso geral, como um pulso retangular de amplitude e duração arbitrários, processado por um filtro passa-baixa ideal com largura de faixa B . Desta forma, para o caso geral, podemos escrever

$$\begin{aligned}
 \rho &= \frac{\text{Energia da saída do filtro}}{\text{Energia da entrada do filtro}} \\
 &= 2 \int_0^B \operatorname{sinc}^2(f) df
 \end{aligned} \tag{2.133}$$

De acordo com a Fig. 2.30(b), o pulso retangular aplicado à entrada do filtro possui duração unitária. Logo, a variável f na Eq. (2.133) representa a *freqüência normalizada*. A Eq. (2.133) é traçada na Fig. 2.32. Esta figura mostra que cerca de 90% da energia de um pulso retangular reside dentro do lóbulo espectral principal do pulso.

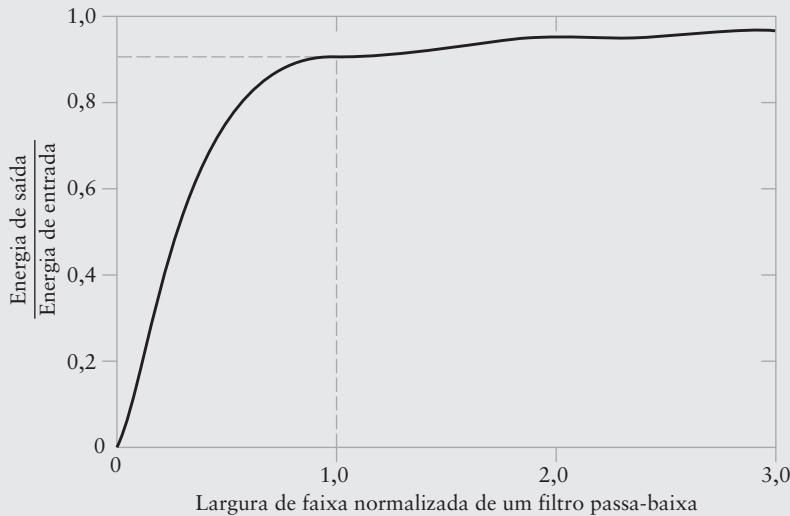


Figura 2.32 Razão de energia de saída por energia de entrada *versus* largura de faixa normalizada.

■ Interpretação da densidade espectral de energia

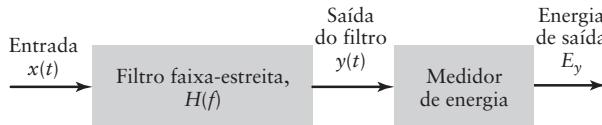
A Eq. (2.129) é importante, pois ela não apenas relaciona a densidade espectral de energia de saída de um sistema linear invariante no tempo com a densidade espectral de energia de entrada, mas também fornece a base para a interpretação física do conceito de densidade

espectral de energia. Sendo mais específico, considere o arranjo mostrado na Fig. 2.33(a), no qual um sinal $x(t)$ de energia passa através de um filtro de faixa estreita seguido por um *medidor de energia*. A Fig. 2.33(b) mostra a resposta em amplitude idealizada do filtro. Ou seja, o filtro é um filtro passa-faixa cuja resposta em amplitude é definida por

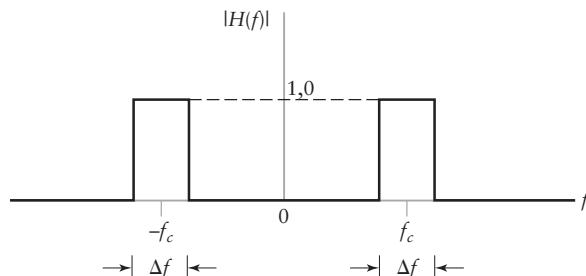
$$|H(f)| = \begin{cases} 1, & f_c - \frac{\Delta f}{2} \leq |f| \leq f_c + \frac{\Delta f}{2} \\ 0, & \text{caso contrário} \end{cases} \quad (2.134)$$

Assumimos que a *largura de faixa do filtro* Δf é pequena o suficiente para que a resposta em amplitude do sinal $x(t)$ de entrada seja essencialmente plana no intervalo de freqüência coberto pela faixa passante do filtro. Desta forma podemos expressar o espectro de amplitude da saída do filtro pela fórmula aproximada

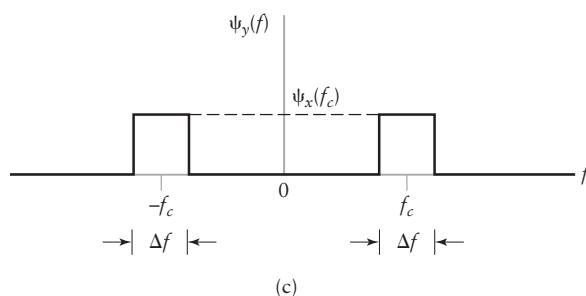
$$\begin{aligned} |Y(f)| &= |H(f)||X(f)| \\ &\approx \begin{cases} |X(f_c)|, & f_c - \frac{\Delta f}{2} \leq |f| \leq f_c + \frac{\Delta f}{2} \\ 0, & \text{caso contrário} \end{cases} \end{aligned} \quad (2.135)$$



(a)



(b)



(c)

Figura 2.33 (a) Diagrama em blocos do sistema para a medição da densidade espectral de energia. (b) Resposta em amplitude idealizada do filtro. (c) Densidade espectral de energia da saída do filtro.

Assim sendo, a densidade espectral de energia, $\psi_y(f)$, da saída $y(t)$ do filtro é aproximadamente relacionada com a densidade espectral de energia $\psi_x(f)$ da entrada do filtro $x(t)$ por:

$$\psi_y(f) \approx \begin{cases} \psi_x(f_c), & f_c - \frac{\Delta f}{2} \leq |f| \leq f_c + \frac{\Delta f}{2} \\ 0, & \text{caso contrário} \end{cases} \quad (2.136)$$

Esta relação é ilustrada na Fig. 2.33(c), a qual mostra que apenas as componentes de freqüência do sinal $x(t)$ que estão dentro da faixa passante do filtro passa-faixa ideal atingem a saída. Utilizando o teorema de Rayleigh para energia, a energia da saída $y(t)$ do filtro é dada por

$$\begin{aligned} E_y &= \int_{-\infty}^{\infty} \psi_y(f) df \\ &= 2 \int_0^{\infty} \psi_y(f) df \end{aligned}$$

À luz da Eq. (2.136), podemos aproximar E_y por

$$E_y \approx 2\psi_x(f_c) \Delta f \quad (2.137)$$

A multiplicação pelo fator 2 é devida a contribuição da freqüência positiva e negativa. Podemos reescrever a Eq. (2.137) na forma

$$\psi_x(f_c) \approx \frac{E_y}{2\Delta f} \quad (2.138)$$

A Eq. (2.138) afirma que a densidade espectral de energia da entrada do filtro em alguma freqüência f_c é igual a energia da saída do filtro dividida por $2\Delta f$, na qual Δf é a largura de faixa do filtro centrada em f_c . Podemos, portanto, interpretar a densidade espectral de energia de um sinal de energia para qualquer freqüência f como *a energia por unidade de largura de faixa, para a qual contribui as componentes de freqüência do sinal ao redor da freqüência f*.

O arranjo mostrado no diagrama em blocos da Fig. 2.33(a) fornece, portanto, a base para a medição da densidade espectral de energia de um sinal de energia. Especificamente, utilizando um filtro passa-faixa variável para varrer a faixa de freqüências de interesse e determinando a energia da saída do filtro para cada freqüência dentro desta faixa, podemos obter o gráfico da densidade espectral de energia em função da freqüência. Note, entretanto, que para a fórmula da Eq. (2.138) ser válida e, portanto, para que o arranjo da Fig. 2.33(a) funcione, a largura de faixa Δf deve permanecer fixa enquanto variamos f_c .

■ Correlação cruzada de sinais de energia

A função de autocorrelação fornece uma medida de similaridade entre um sinal e sua própria versão atrasada no tempo. De forma similar, podemos utilizar a *função de correlação cruzada* como uma medida de similaridade entre um sinal e a versão atrasada no tempo de um segundo sinal. Seja $x(t)$ e $y(t)$ um par de sinais de energia de valor complexo. A correlação cruzada deste par de sinais é definida por

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y^*(t - \tau) dt \quad (2.139)$$

Vimos que se os dois sinais $x(t)$ e $y(t)$ forem de alguma forma similar, então a função de correlação cruzada $R_{xy}(\tau)$ será finita em alguma faixa de τ , fornecendo, pois, uma medi-

da quantitativa da similaridade, ou coerência entre os sinais. Os sinais de energia $x(t)$ e $y(t)$ são ditos ortogonais sobre todo intervalo de tempo se $R_{xy}(0)$ for zero, ou seja, se

$$\int_{-\infty}^{\infty} x(t)y^*(t) dt = 0 \quad (2.140)$$

A Eq. (2.139) define um possível valor para a função de correlação cruzada para um valor específico da variável de atraso τ . Podemos definir uma segunda função de correlação cruzada para o sinal de energia $x(t)$ e $y(t)$, dada por

$$R_{yx}(\tau) = \int_{-\infty}^{\infty} y(t)x^*(t - \tau) dt \quad (2.141)$$

A partir das definições das funções de correlação cruzada $R_{xy}(\tau)$ e $R_{yx}(\tau)$, obtemos a relação fundamental

$$R_{xy}(\tau) = R_{yx}^*(-\tau) \quad (2.142)$$

A Eq. (2.142) indica que, ao contrário da convolução, a correlação não é, em geral, comutativa. Ou seja, $R_{xy}(\tau) \neq R_{yx}(\tau)$.

Para caracterizar o comportamento da correlação cruzada de sinais de energia no domínio da freqüência, introduzimos a notação de *densidade espectral cruzada*. Especificamente, dado um par de sinais de energia de valor complexo $x(t)$ e $y(t)$, definimos suas densidades espectrais cruzadas, representadas por $\psi_{xy}(f)$ e $\psi_{yx}(f)$, como as transformadas de Fourier das respectivas funções de correlação cruzada $R_{xy}(\tau)$ e $R_{yx}(\tau)$, como mostrado por

$$\psi_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) \exp(-j2\pi f\tau) d\tau \quad (2.143)$$

e

$$\psi_{yx}(f) = \int_{-\infty}^{\infty} R_{yx}(\tau) \exp(-j2\pi f\tau) d\tau \quad (2.144)$$

De acordo com o teorema da correlação (isto é, Propriedade 13 da Seção 2.2), temos, portanto,

$$\psi_{xy}(f) = X(f)Y^*(f) \quad (2.145)$$

e

$$\psi_{yx}(f) = Y(f)X^*(f) \quad (2.146)$$

A partir destes pares de relação, vemos rapidamente duas propriedades da densidade espectral cruzada.

1. Ao contrário da densidade espectral de energia, a densidade espectral cruzada é, geralmente, um valor complexo.
2. $\psi_{xy}(f) = \psi_{yx}^*(f)$ da qual temos, no caso geral, $\psi_{xy}(f) \neq \psi_{yx}(f)$.

► **PROBLEMA 2.15** Obtenha a relação da Eq. (2.142) entre duas funções de correlação cruzada $R_{xy}(t)$ e $R_{yx}(t)$. ◀

► **PROBLEMA 2.16** Considere o pulso exponencial decrescente

$$g(t) = \begin{cases} \exp(-at), & t > 0 \\ 1, & t = 0 \\ 0, & t < 0 \end{cases}$$

Determine a densidade espectral de energia do pulso $g(t)$. ◀

► **PROBLEMA 2.17** Repita o Problema 2.16 para o pulso exponencial dobrado

$$g(t) = \begin{cases} \exp(-at), & t > 0 \\ 1, & t = 0 \\ \exp(at), & t < 0 \end{cases}$$



2.9 DENSIDADE ESPECTRAL DE POTÊNCIA

Nesta seção, iremos expandir a importante noção de densidade espectral para incluir a classe de sinais de potência. A *potência média* de um sinal $x(t)$ é definida por

$$P = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |x(t)|^2 dt \quad (2.147)$$

O sinal $x(t)$ é dito ser um *sinal de potência* se a condição

$$P < \infty$$

for válida. Exemplos de sinais de potência incluem sinais periódicos e ruído. Iremos considerar sinais periódicos nesta seção. (O ruído será considerado no Capítulo 8.)

Para desenvolver uma distribuição de potência no domínio da freqüência, precisamos conhecer a transformada de Fourier do sinal $x(t)$. Entretanto isto é um problema, pois sinais de potência possuem energia infinita e podem, portanto, não possuir transformada de Fourier. Para superar este problema, iremos considerar uma versão truncada do sinal $x(t)$. Em particular, definimos

$$\begin{aligned} x_T(t) &= x(t) \operatorname{ret}\left(\frac{t}{2T}\right) \\ &= \begin{cases} x(t), & -T \leq t \leq T \\ 0, & \text{caso contrário} \end{cases} \end{aligned} \quad (2.148)$$

Em quanto a duração T for finita, o sinal truncado $x_T(t)$ possui energia finita. Logo, $x_T(t)$ possui transformada de Fourier. Seja $X_T(f)$ a transformada de Fourier do sinal $x_T(t)$, ou seja,

$$x_T(t) \iff X_T(f)$$

Usando o sinal truncado $x_T(t)$, podemos reescrever a Eq. (2.147) para a potência média P em termos de $x_T(t)$ como

$$P = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-\infty}^{\infty} |x_T(t)|^2 dt \quad (2.149)$$

Como $x_T(t)$ possui energia finita, podemos utilizar o teorema de Rayleigh para energia para expressar a energia de $x_T(t)$ em termos de sua transformada de Fourier $X_T(f)$, usando

$$\int_{-\infty}^{\infty} |x_T(t)|^2 dt = \int_{-\infty}^{\infty} |X_T(f)|^2 df$$

na qual $|X_T(f)|$ é o espectro de amplitude de $x_T(t)$. Assim sendo, podemos reescrever a Eq. (2.149) na forma equivalente

$$P = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-\infty}^{\infty} |X_T(f)|^2 df \quad (2.150)$$

Quanto a duração T aumenta, a energia de $x_T(t)$ aumenta. Desta forma, a densidade espectral de energia $|X_T(f)|^2$ aumenta com T . De fato, quando T tende ao infinito, $|X_T(f)|^2$

também tenderá. Entretanto, para a potência média P ser finita, $|X_T(f)|^2$ deve tender ao infinito na mesma taxa que T . Esta condição garante a *convergência* da integral do lado direito da Eq. (2.150) no limite quando T tende ao infinito. A convergência, por sua vez, nos permite *alterar a ordem na qual a operação de limite e integração ocorrem na Eq. (2.150)*. Podemos reescrever esta equação como

$$P = \int_{-\infty}^{\infty} \left(\lim_{T \rightarrow \infty} \frac{1}{2T} |X_T(f)|^2 \right) df \quad (2.151)$$

Seja o integrando da Eq. (2.151) representado por

$$S_x(f) = \lim_{T \rightarrow \infty} \frac{1}{2T} |X_T(f)|^2 \quad (2.152)$$

A função $S_x(f)$ dependente da freqüência é chamada de *densidade espectral de potência* ou *espectro de potência* de um sinal de potência $x(t)$, e a quantidade $(|X_T(f)|^2/2T)$ é chamada de *periodograma* do sinal.

Da Eq. (2.152), facilmente vemos que a densidade espectral de potência é uma quantidade de valor real não negativa para todas as freqüências. Além disso, da Eq. (2.152), rapidamente vemos que

$$P = \int_{-\infty}^{\infty} S_x(f) df \quad (2.153)$$

A Eq. (2.153) afirma: *a área total sob a curva da densidade espectral de potência de um sinal de potência é igual a potência média do sinal*. A densidade espectral de potência de um sinal de potência possui, portanto, um papel similar à da densidade espectral de energia para um sinal de energia.

► **PROBLEMA 2.18** Em sentido implícito, a Eq. (2.153) engloba o *teorema de Parseval para potência*, o qual afirma que para um *sinal periódico* $x(t)$, temos

$$\frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 dt = \sum_{n=-\infty}^{\infty} |X(nf_0)|^2$$

no qual T é o período do sinal, f_0 é a freqüência fundamental e $X(nf_0)$ é a transformada de Fourier de $x(t)$ calculada na freqüência nf_0 . Prove este teorema. ◀

EXEMPLO 2.15 Onda modulada

Considere a onda modulada

$$x(t) = g(t) \cos(2\pi f_c t) \quad (2.154)$$

na qual $g(t)$ é um sinal de potência com faixa limitada a B hertz. Referimos a $x(t)$ como “onda modulada” no sentido de que a amplitude da “portadora” senoidal de freqüência f_c é variada linearmente com o sinal $g(t)$. (O assunto modulação é coberto em detalhes no Capítulo 3). Queremos determinar a densidade espectral de potência de $x(t)$ em termos da densidade espectral de potência de $g(t)$, dada que a freqüência f_c é maior do que a largura de faixa B .

Seja $g_T(t)$ a versão truncada de $g(t)$, definida de maneira similar a descrita pela Eq. (2.148). Desta forma, podemos expressar a versão truncada de $x(t)$ por

$$x_T(t) = g_T(t) \cos(2\pi f_c t) \quad (2.155)$$

Como

$$\cos(2\pi f_c t) = \frac{1}{2} [\exp(j2\pi f_c t) + \exp(-j2\pi f_c t)], \quad (2.156)$$

temos que, da propriedade de deslocamento em freqüência (isto é, Propriedade 6) da transformada de Fourier,

$$X_T(f) = \frac{1}{2}[G_T(f - f_c) + G_T(f + f_c)] \quad (2.157)$$

na qual $G_T(f)$ é a transformada de Fourier de $g_T(t)$.

Dado que $f_c > B$, observamos que $G_T(f - f_c)$ e $G_T(f + f_c)$ não apresentam espectro sobreposto e, portanto, seu produto é zero. Assim sendo, usando a Eq. (2.157) para calcular a amplitude quadrática de $X_T(f)$, temos

$$|X_T(f)|^2 = \frac{1}{4} [|G_T(f - f_c)|^2 + |G_T(f + f_c)|^2] \quad (2.158)$$

Finalmente, aplicando-se a definição da Eq. (2.152) para a densidade espectral de potência do sinal de potência $g(t)$ à Eq. (2.158), obtemos o resultado desejado:

$$S_x(f) = \frac{1}{4}[S_g(f - f_c) + S_g(f + f_c)] \quad (2.159)$$

Exceto pelo fator de escala de 1/4, a densidade espectral de potência da onda modulada $x(t)$ é igual a soma da densidade espectral de potência $S_g(f)$ deslocada para a direita por f_c e de $S_g(f)$ deslocada para a esquerda pelo mesmo fator f_c .

2.10 DETERMINAÇÃO NUMÉRICA DA TRANSFORMADA DE FOURIER

O material apresentado neste capítulo mostra, claramente, a importância da transformada de Fourier como ferramenta teórica para a representação de sinais determinísticos e sistema linear invariante no tempo. A importância da transformada de Fourier é maior ainda pelo fato de existir uma classe de algoritmos chamados de algoritmos de transformada de Fourier para a determinação numérica da transformada de Fourier de forma muito eficiente.

O algoritmo da transformada rápida de Fourier é derivado da transformada discreta de Fourier, na qual, como o próprio nome informa, tanto o tempo quanto a freqüência são representados na forma discreta. A transformada discreta de Fourier fornece uma *aproximação* da transformada de Fourier. Para representar adequadamente o conteúdo de informação do sinal original, precisamos tomar cuidado especial na execução de operações de amostragem envolvidas na definição da transformada discreta de Fourier. Uma análise detalhada do processo de amostragem será apresentada no Capítulo 5. Neste momento, é suficiente dizer que dado um sinal limitado em faixa, a taxa de amostragem deve ser maior do que duas vezes a maior componente de freqüência do sinal de entrada. Além disso, se as amostras forem uniformemente espaçadas por T_s segundos, o espectro do sinal se torna periódico, repetindo a cada $f_s = (1/T_s)$ Hz. Seja N o número de amostras de freqüência contidas no intervalo f_s . Logo, a *resolução de freqüência* envolvida na determinação numérica da transformada de Fourier é definida por

$$\Delta f = \frac{f_s}{N} = \frac{1}{NT_s} = \frac{1}{T} \quad (2.160)$$

na qual $T = NT_s$ é a duração total do sinal.

Considere, então, uma seqüência finita de dados $\{g_0, g_1, \dots, g_{N-1}\}$. Por simplicidade, iremos nos referir a esta seqüência como g_n , na qual o subscrito é o índice de tempo $n = 0, 1, \dots, N - 1$. Esta seqüência pode representar o resultado da amostragem de um sinal analógico $g(t)$ nos tempos $t = 0, T_s, \dots, (N - 1)T_s$, na qual T_s é o intervalo de amostragem. A ordem da seqüência de dados define o tempo de amostragem na qual g_0, g_1, \dots, g_{N-1}

representa as amostras de $g(t)$ tomadas nos tempos $0, T_s, \dots, (N - 1)T_s$, respectivamente. Portanto, temos

$$g_n = g(nT_s) \quad (2.161)$$

Definimos, formalmente, a transformada discreta de Fourier (TDF) da seqüência g_n por

$$G_k = \sum_{n=0}^{N-1} g_n \exp\left(-\frac{j2\pi}{N} kn\right), \quad k = 0, 1, \dots, N - 1 \quad (2.162)$$

A seqüência $\{G_0, G_1, \dots, G_{N-1}\}$ é chamada de seqüência transformada. Por simplicidade iremos nos referir a esta nova seqüência por G_k , na qual o subscrito é o índice de freqüência $k = 0, 1, \dots, N - 1$. De forma correspondente, definimos a transformada discreta inversa de Fourier (TDIF) de G_k por

$$g_n = \frac{1}{N} \sum_{k=0}^{N-1} G_k \exp\left(\frac{j2\pi}{N} kn\right), \quad n = 0, 1, \dots, N - 1 \quad (2.163)$$

A TDF e a TDIF formam um par transformada. Especificamente, dada a seqüência g_n , podemos utilizar a TDF para calcular a seqüência transformada G_k e dada a seqüência transformada G_k , podemos utilizar a TDIF para recuperar a seqüência de dados original g_n . A característica distinta da TDF é que para o somatório finito definido nas Eqs. (2.162) e (2.163) não existe a questão de convergência.

Quando discutimos a TDF (e algoritmos para a sua implementação), as palavras “amostra” e “ponto” são utilizadas com o mesmo significado, para nos referirmos a um valor da seqüência. Além disso, é prática comum se referir a seqüência de tamanho N como uma *seqüência de N pontos* e se referir à TDF de uma seqüência de tamanho N como uma TDF de N pontos.

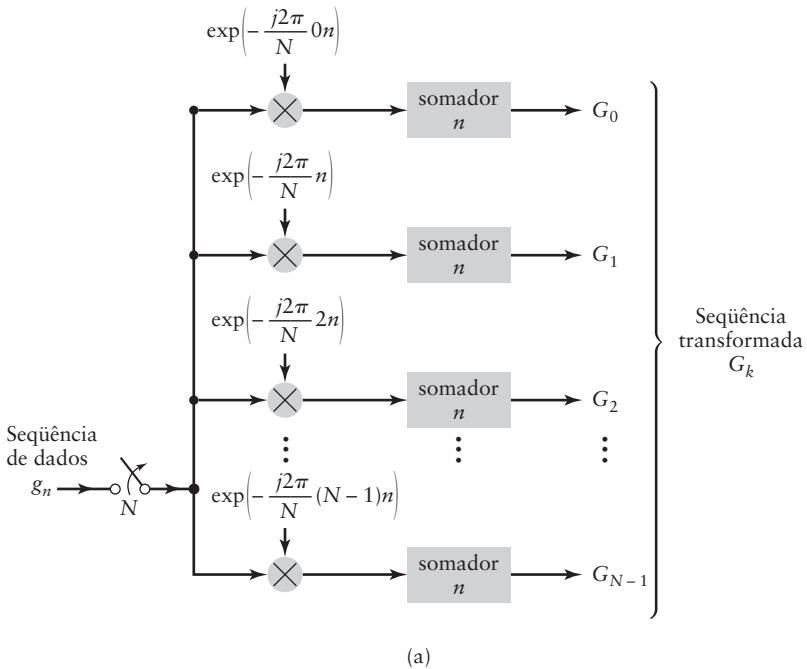
■ Interpretações da TDF e da TDIF

Podemos visualizar o processo da TDF, descrita pela Eq. (2.162), como uma coleção de N operações de média e heterodinagem complexa, como mostrado na Fig. 2.34(a). Na figura apresentada, a heterodinagem se refere à multiplicação da seqüência de dados g_n por uma exponencial complexa. Dizemos que a heterodinagem é complexa porque as amostras da seqüência de dados são multiplicadas por *seqüências exponenciais complexas*. Existe um total de N seqüências exponenciais complexas a serem consideradas, correspondendo ao índice de freqüência $k = 0, 1, \dots, N - 1$. Seus períodos foram selecionados de tal forma que cada seqüência exponencial complexa possua, precisamente, um número de ciclos inteiros no intervalo total de 0 a $N - 1$. A resposta de freqüência nula, correspondendo a $k = 0$, é a única exceção.

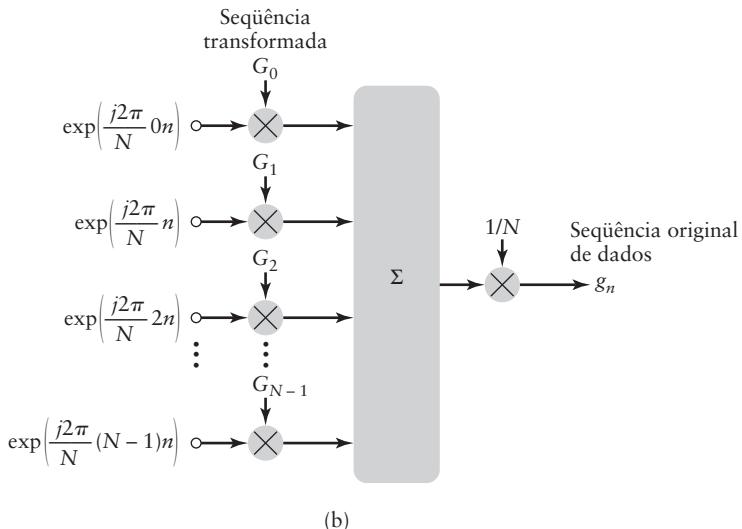
Para a interpretação do processo da TDIF, descrita na Eq. (2.163), podemos utilizar o esquema mostrado na Fig. 2.34(b), no qual temos uma coleção de N geradores de sinal complexos, cada um dos quais produz a *seqüência exponencial complexa*

$$\begin{aligned} \exp\left(\frac{j2\pi}{N} kn\right) &= \cos\left(\frac{2\pi}{N} kn\right) + j \sin\left(\frac{2\pi}{N} kn\right) \\ &= \left\{ \cos\left(\frac{2\pi}{N} kn\right), \sin\left(\frac{2\pi}{N} kn\right) \right\}, \quad k = 0, 1, \dots, N - 1 \end{aligned} \quad (2.164)$$

Portanto, na realidade, cada gerador de sinal complexo consiste de um par de geradores cuja saída é a seqüência cossoidal e senoidal de k ciclos por intervalo de observação. A saída de cada gerador de sinal complexo é ponderada pelo coeficiente complexo de Fourier G_k . Em cada índice de tempo n , uma saída é formada pela soma ponderada das saídas do gerador complexo.



(a)



(b)

Figura 2.34 Interpretações de (a) a TDF como método analisador da seqüência de dados g_n e (b) a TDIF como método sintetizador de g_n .

É importante ressaltar que, apesar da TDF e TDIF possuírem formulações matemáticas similares, como descrito nas Eqs. (2.162) e (2.163), suas interpretações, como mostrado nas Figs. 2.34(a) e 2.34(b), são completamente diferentes.

Além disso, a soma de sinas periódicos harmonicamente relacionados, como nas Figs. 2.34(a) e 2.34(b), sugere que as seqüências G_k e g_n devem ser periódicas. Mais ainda, os processadores mostrados nas Figs. 2.34(a) e 2.34(b) devem ser lineares, sugerindo que a TDF e TDIF são operações lineares. Esta importante propriedade também é óbvia a partir das equações de definição (2.162) e (2.163).

■ Algoritmos da Transformada Rápida de Fourier

Na transformada discreta de Fourier (TDF), tanto a entrada quanto a saída consistem de seqüências de números definidos por pontos uniformemente espaçados no tempo e freqüência, respectivamente. Esta característica torna a TDF ideal para a determinação numérica direta em um computador digital. Além disso, o cálculo pode ser implementado de forma eficiente utilizando uma classe de algoritmos chamados de *transformada rápida de Fourier* (FFT do original em inglês).² Um algoritmo se refere a uma receita que pode ser escrita na forma de um programa de computador.

Algoritmos de transformada rápida de Fourier são computacionalmente eficientes, porque eles utilizam um número de operações aritméticas muito reduzido quando comparado ao cálculo direto da TDF. Basicamente, um algoritmo de FFT obtém sua eficiência computacional seguindo uma estratégia de dividir-e-conquistar, na qual o cálculo original da TDF é decomposto sucessivamente em cálculos de TDF menores. Nesta seção, iremos descrever uma versão de um algoritmo popular de FFT, no qual o seu desenvolvimento é baseado nesta estratégia.

Para prosseguirmos com o desenvolvimento, iremos, primeiro, reescrever a Eq. (2.162), a qual define a TDF de g_n , na forma simplificada

$$G_k = \sum_{n=0}^{N-1} g_n W^{nk}, \quad k = 0, 1, \dots, N - 1 \quad (2.165)$$

na qual o novo coeficiente W é definido por

$$W = \exp\left(-j\frac{2\pi}{N}\right) \quad (2.166)$$

Utilizando esta definição, vemos que

$$\begin{aligned} W^N &= \exp(-j2\pi) = 1 \\ W^{N/2} &= \exp(-j\pi) = -1 \\ W^{(k+lN)(n+mN)} &= W^{kn}, \quad \text{para } m, l = 0, \pm 1, \pm 2, \dots \end{aligned}$$

ou seja, W^{kn} é periódico com período N . A periodicidade de W^{kn} é a característica chave no desenvolvimento de algoritmos de FFT.

Seja N , o número de pontos da seqüência de dados, *uma potência de dois inteiros*, como mostrado por

$$N = 2^L$$

na qual L é um inteiro. Como N é um inteiro par, $N/2$ é um inteiro e, portanto, podemos dividir a seqüência de dados em uma primeira e segunda metade de pontos. Portanto, podemos reescrever a Eq. (2.165) na forma equivalente,

$$\begin{aligned} G_k &= \sum_{n=0}^{(N/2)-1} g_n W^{nk} + \sum_{n=N/2}^{N-1} g_n W^{nk} \\ &= \sum_{n=0}^{(N/2)-1} g_n W^{nk} + \sum_{n=0}^{(N/2)-1} g_{n+N/2} W^{k(n+N/2)} \\ &= \sum_{n=0}^{(N/2)-1} (g_n + g_{n+N/2} W^{kN/2}) W^{kn}, \quad k = 0, 1, \dots, N - 1 \end{aligned} \quad (2.167)$$

² O algoritmo da transformada rápida de Fourier (FFT) possui uma longa história. Sua descoberta moderna (ou redescoberta para sermos mais precisos) é atribuída a Cooley e Tukey, em 1965. Veja o artigo de Cooley (1992) para detalhes.

Note que, na segunda linha da Eq. (2.167), mudamos o índice do segundo termo de somatório de tal forma que os dois temos cubram a mesma faixa. Como $W^{kN/2} = -1$, temos

$$W^{kN/2} = (-1)^k$$

Para o cálculo da Eq. (2.167) procedemos considerando dois casos, um correspondendo a valores pares de k e o outro correspondendo aos valores ímpares de k . Para o caso de k par, seja $k = 2l$, na qual $l = 0, 1, \dots, (N/2)$. Logo, definimos

$$x_n = g_n + g_{n+N/2} \quad (2.168)$$

Então, para k par, podemos colocar a Eq. (2.167) na nova forma

$$G_{2l} = \sum_{n=0}^{(N/2)-1} x_n (W^2)^{ln}, \quad l = 0, 1, \dots, \frac{N}{2} - 1 \quad (2.169)$$

A partir da definição de W dada pela Eq. (2.166), rapidamente vemos que

$$\begin{aligned} W^2 &= \exp\left(-\frac{j4\pi}{N}\right) \\ &= \exp\left(-\frac{j2\pi}{N/2}\right) \end{aligned}$$

Logo, percebemos que o somatório do lado direito da Eq. (2.169) é a seqüência de $(N/2)$ pontos da TDF da seqüência x_n .

Considere, a seguir, o caso restante de k ímpar, e seja

$$k = 2l + 1, \quad l = 0, 1, \dots, \frac{N}{2} - 1$$

Então, percebendo que para k ímpar, $W^{kN/2} = -1$, podemos definir

$$y_n = g_n - g_{n+N/2} \quad (2.170)$$

Logo, para o caso de k ímpar, podemos colocar a Eq. (2.167) na forma

$$\begin{aligned} G_{2l+1} &= \sum_{n=0}^{(N/2)-1} y_n W^{(2l+1)n} \\ &= \sum_{n=0}^{(N/2)-1} [y_n W^n] (W^2)^{ln}, \quad l = 0, 1, \dots, \frac{N}{2} - 1 \end{aligned} \quad (2.171)$$

Desta forma percebemos que o somatório do lado direito da Eq. (2.171) é a TDF de $(N/2)$ pontos da seqüência modificada $y_n W^n$. O coeficiente W^n multiplicando y_n é chamado de *fator de rotação*.

As Eqs. (2.169) e (2.171) mostram que amostras de valor par e ímpar da seqüência transformada G_k podem ser obtidas de TDF de $(N/2)$ pontos da seqüência x_n e $y_n W^n$, respectivamente. As seqüências x_n e y_n são relacionadas como a seqüência original de dados g_n pelas Eqs. (2.168) e (2.170), respectivamente. Portanto, o problema de cálculo de uma TDF de N pontos é reduzido para o cálculo de duas TDF de $(N/2)$ pontos. Este procedimento é repetido uma segunda vez, na qual $(N/2)$ pontos é decomposta em duas TDF de $(N/4)$ pontos. A decomposição (ou mais precisamente, o procedimento de dividir-e-conquistar) é continuada nesta forma até (após $L = \log_2 N$ estágios) que tenhamos atingindo o caso trivial TDF de ponto único.

A Fig. 2.35 ilustra os cálculos necessários para a aplicação das fórmulas das Eqs. (2.169) e (2.171) a uma seqüência de dados de 8 pontos. Ou seja, $N = 8$. Na construção das porções do lado esquerdo da figura, utilizamos a notação de grafo de fluxo de sinal.

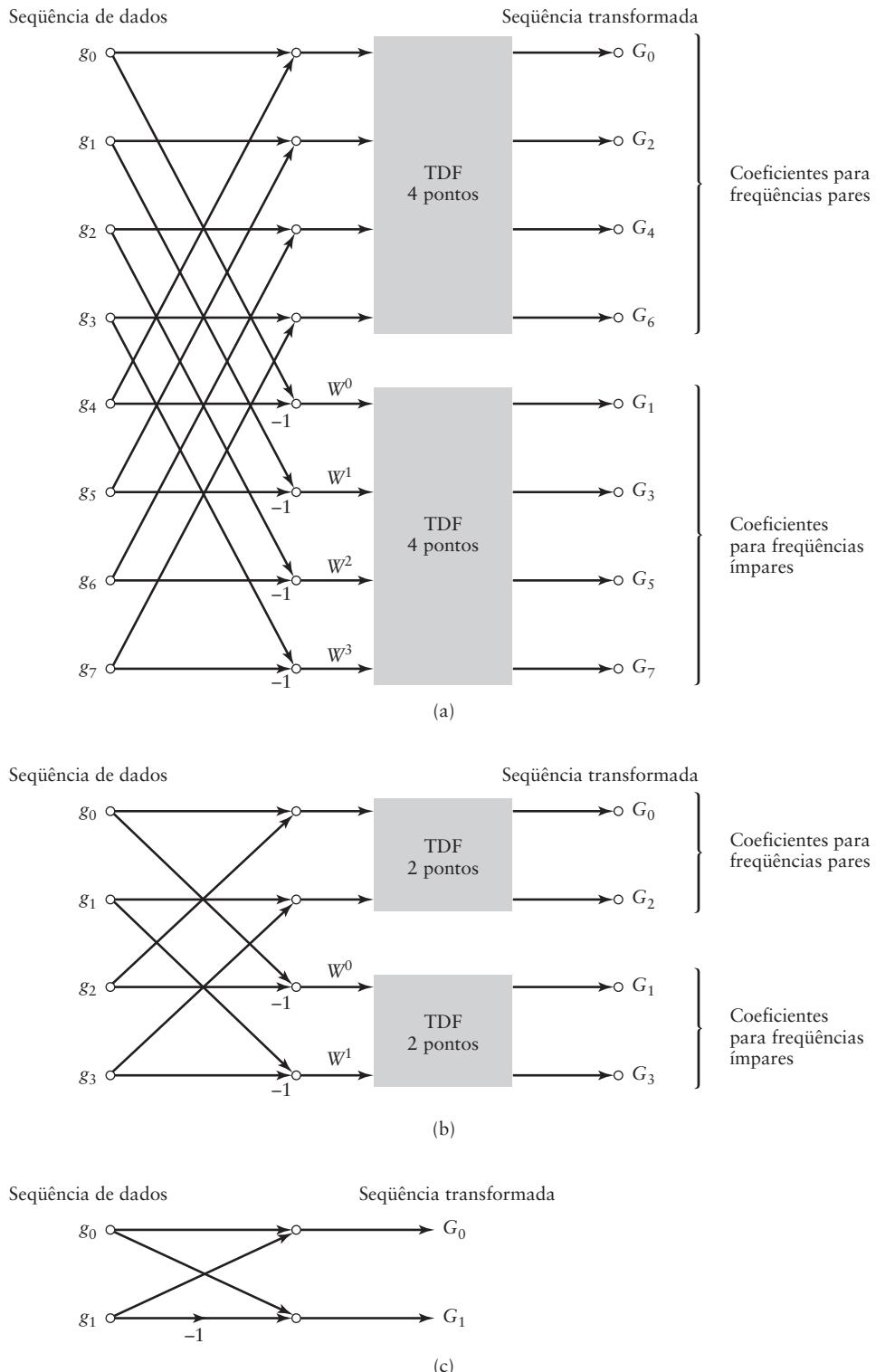


Figura 2.35 (a) Redução de uma TDF de 8 pontos em duas TDF de 4 pontos. (b) Redução de uma TDF de 4 pontos em duas TDF de 2 pontos. (c) Caso trivial da TDF de 2 pontos.

Um grafo de fluxo de sinal é constituído pela interconexão de nós e ramos. A direção da transmissão de sinal ao longo do ramo é indicada por uma seta. Um nó soma as saídas de todos os ramos de entrada. A convenção utilizada para o peso das transmissões nos ramos da Fig. 2.35 é explicada a seguir.

Quando nenhum coeficiente estiver indicado no ramo, a transmissão no ramo é considerada como sendo unitária. Para os outros ramos, a transmissão no ramo é indicada por -1 ou uma potência inteira de W , assinalada ao longo da seta do ramo.

Portanto, na Fig. 2.35(a), o cálculo de uma TDF de 8 pontos é reduzida a duas TDF de 4 pontos. O procedimento para a TDF de 8 pontos é simulado para simplificar o cálculo da TDF de 4 pontos. Este fato é ilustrado na Fig. 2.35(b), na qual o cálculo de uma TDF de 4 pontos é reduzida para duas TDF de 2 pontos. Finalmente, o cálculo da uma TDF de 2 pontos é mostrado na Fig. 2.35(c).

Combinando as idéias descritas na Fig. 2.35, obtemos o grafo de fluxo de sinal completo da Fig. 2.36 para o cálculo de uma TDF de 8 pontos. A estrutura repetitiva, chamada de *borboleta*, pode ser distinguida no algoritmo de FFT da Fig. 2.36. Uma borboleta possui duas entradas e duas saídas. Exemplos de borboletas (para os três estágios do algoritmo) são ilustradas pelas linhas em negrito da Fig. 2.36.

Para o caso geral de $N = 2^L$, o algoritmo requer $L = \log_2 N$ estágios de cálculo. Cada estágio requer $(N/2)$ borboletas. Cada borboleta envolve uma multiplicação complexa e duas somas complexas (para sermos precisos, uma adição e uma subtração). Desta forma,

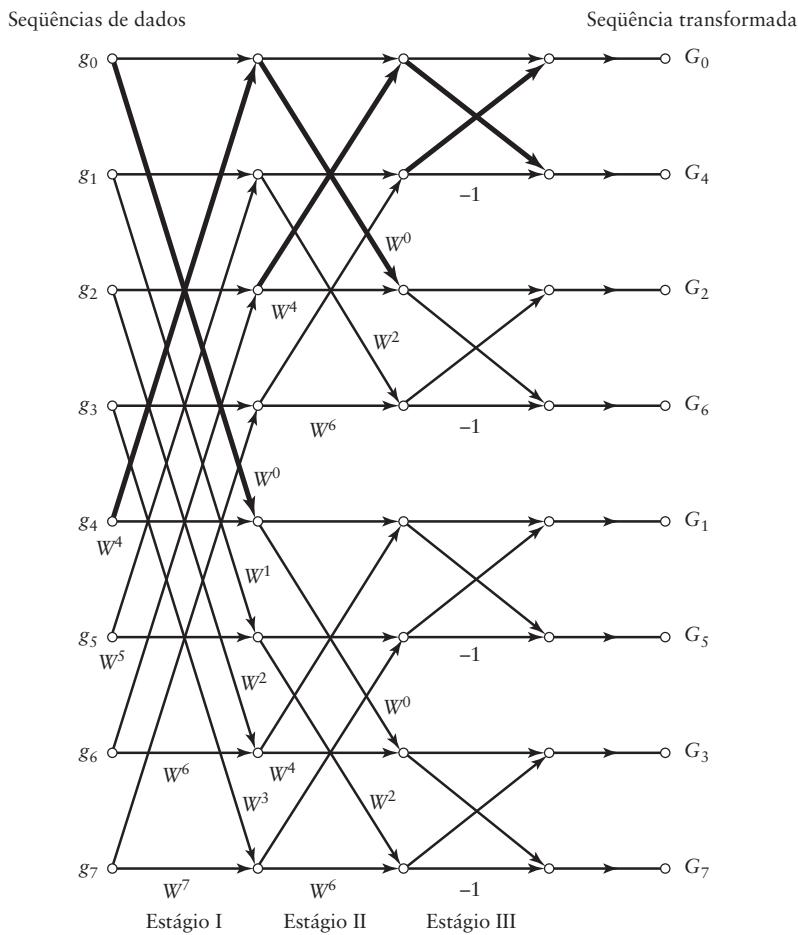


Figura 2.36 Algoritmo FFT de decimação em freqüência.

a estrutura do FFT descrita aqui requer $(N/2)\log_2 N$ multiplicações complexas e $N\log_2 N$ somas complexas. (Na realidade, o número de multiplicações estimado é pessimista, pois podemos omitir todos os fatores de rotação $W^0 = 1$ e $W^{N/2} = -1$, $W^{N/4} = -j$, $W^{3N/4} = j$.) Esta complexidade computacional é significativamente menor do que as N^2 multiplicações complexas e $N(N - 1)$ adições complexas necessárias para o cálculo direto da TDF. A redução computacional possibilitada pelo algoritmo de FFT se torna mais substancial quando aumentamos o tamanho N de dados.

Podemos estabelecer duas outras importantes características do algoritmo FFT examinando cuidadosamente o grafo de fluxo de sinal da Fig. 2.36.

1. Em cada estádio de cálculo, o novo conjunto de N números complexos resultantes dos cálculos pode ser armazenado nas mesmas posições de memória utilizadas para armazenar o conjunto anterior. Este tipo de computação é chamado de *computação in-place*.
2. As amostras da seqüência transformada G_k são armazenadas em uma ordem de reversão de bit. Para ilustrar o significado desta terminologia considere a Tabela 2.2, construída para o caso de $N = 8$. Na esquerda da tabela mostramos os oito possíveis valores do índice de freqüência k (em sua ordem natural) e suas representações binárias utilizando três bits. Na direita da tabela mostramos as representações binárias com bit revertido e índices correspondentes. Observe que os índices de bit revertido na coluna mais à direita da Tabela 2.2 aparecem na mesma ordem que os índices da saída do algoritmo de FFT da Fig. 2.36.

TABELA 2.2 Ilustrando a reversão de bit

Índice de freqüência, k	Representação binária	Representação binária com bit revertido	Índice com bit revertido
0	000	000	0
1	001	100	4
2	010	010	2
3	011	110	6
4	1000	001	1
5	101	101	5
6	110	011	3
7	111	111	7

O algoritmo de FFT mostrado na Fig. 2.36 é chamado de *algoritmo de decimação em freqüência*, pois a seqüência transformada (freqüência) G_k é dividida sucessivamente em subseqüências menores. Em outro algoritmo de FFT popular, chamado de algoritmo de *decimação no tempo*, a seqüência de dados (tempo) g_n é dividida sucessivamente em subseqüências menores. Eles diferem um do outro em duas questões. Primeiro, para a decimação em freqüência, a entrada é a ordem natural, enquanto que a saída é a ordem com bit revertido. O inverso é válido para a decimação no tempo. Segundo, a borboleta para a decimação no tempo é um pouco diferente da utilizada na decimação em freqüência. O leitor é convidado a obter os detalhes do algoritmo de decimação no tempo utilizando a estratégia de dividir-e-conquistar que resultou no desenvolvimento do algoritmo descrito na Fig. 2.36. Veja o Problema 2.50.

Cálculo da TDIF

A TDIF da seqüência transformada G_k é definida pela Eq. (2.163). Podemos reescrever esta equação em termos do parâmetro complexo W por

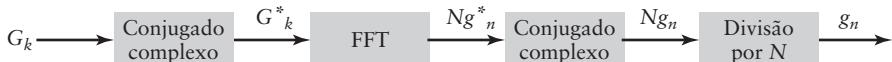


Figura 2.37 Utilização do algoritmo de FFT para o cálculo da TDIF.

$$g_n = \frac{1}{N} \sum_{k=0}^{N-1} G_k W^{-kn}, \quad n = 0, 1, \dots, N - 1 \quad (2.172)$$

Obtendo o conjugado complexo da Eq. (2.172), multiplicando por N e percebendo da definição da Eq. (2.166) que $\bar{W}^* = W^{-1}$, temos

$$Ng_n^* = \sum_{k=0}^{N-1} G_k^* W^{kn}, \quad 0, 1, \dots, N - 1 \quad (2.173)$$

O lado direito da Eq. (2.173) é reconhecido como sendo a TDF de N pontos do conjugado complexo da seqüência G_k^* . Ou seja, a Eq. (2.173) sugere que podemos calcular a seqüência g_n desejada usando o esquema mostrado na Fig. 2.37, baseado no algoritmo de FFT de N pontos. Portanto, o mesmo algoritmo da FFT pode ser utilizado, essencialmente, para efetuar o cálculo tanto da TDIF quanto da TDF.

2.11 EXEMPLO DE TEMA: PAR TRANÇADO PARA TELEFONIA

O meio fundamental de transmissão para a conexão de casas às centrais telefônicas para chaveamento é o *par trançado*. O par trançado é, geralmente, um par de fios de cobre com uma capa de polietileno. Se o cobre possui um diâmetro de 0,4 mm, este tamanho de cabo é referenciado como #26 na *American Wire Gauge*, ou, simplesmente 26 AWG. Um par trançado é um exemplo de uma *linha de transmissão*.

Uma linha de transmissão é constituída por dois condutores, cada um com sua própria resistência e indutância. Como os dois condutores estão, geralmente próximos, existe também um efeito capacitivo entre os dois, além de uma potencial condutância através do material utilizado para isolar os dois cabos. Uma linha de transmissão construída é geralmente representada pelo *círcuito concentrado* mostrado na Fig. 2.38. Apesar das impedâncias serem mostradas como elementos discretos na Fig. 2.38, é mais correto considerá-las como distribuídas ao longo do tamanho da linha de transmissão.

Dependendo dos valores dos elementos do circuito da Fig. 2.38, é claro que a linha de transmissão irá ter um efeito de distorção no sinal transmitido. Além disso, como a impedância total aumenta com o tamanho da linha, o mesmo acontecerá com a resposta em frequência da linha de transmissão.

Na Fig. 2.39 mostramos a resposta típica de um par trançado com comprimentos de 2 a 8 quilômetros. Existem diversas observações a serem feitas a partir da figura:

- ▶ Pares trançados correm diretamente do escritório central de telefonia para a casa com um par dedicado a cada linha telefônica. Conseqüentemente, as linhas de transmissão são relativamente longas.
- ▶ Os resultados da Fig. 2.39 consideram um cabo contínuo. Na prática, podem haver diversas seções separadas de cabo, diferentes dimensões ao longo de diferentes partes

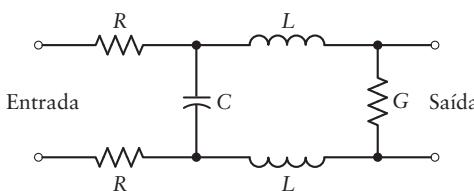


Figura 2.38 Modelo de circuito concentrado de uma linha de transmissão.

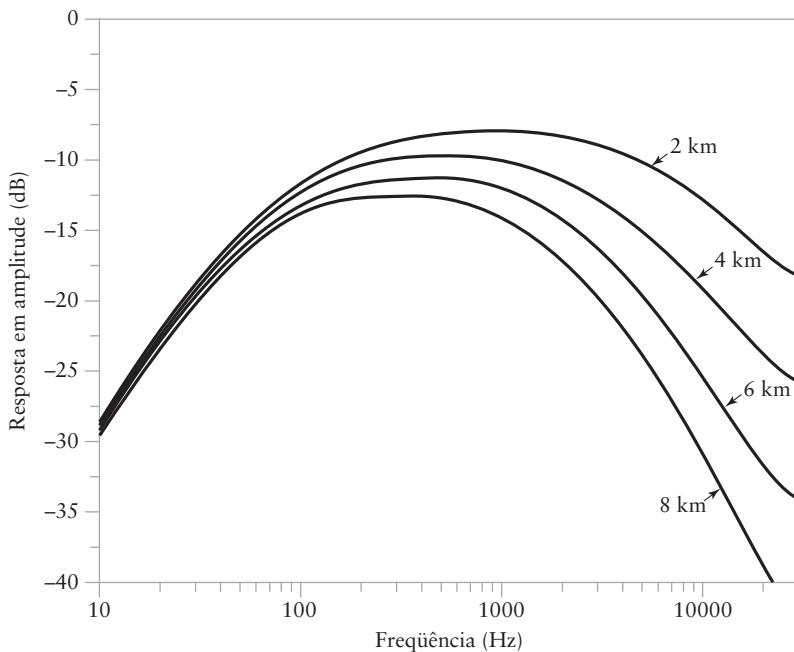


Figura 2.39 Resposta em freqüência típica de uma linha de transmissão utilizando par trançado de 26 AWG, para diferentes comprimentos com impedâncias da fonte e carga de $(600\Omega + 2\mu F)$.

do cabeamento e assim por diante. Estas descontinuidades no meio de transmissão trarão outros efeitos na resposta de freqüência do cabo.

- Vemos que, para um cabo de 2 km, a resposta em freqüência é bem plana na faixa de voz de 200 a 3100 Hz, para comunicação telefônica. Entretanto, para um cabo de 8 km, a resposta em freqüência começa a cair em, aproximadamente, 1 kHz.
- A resposta em freqüência desvia da freqüência zero porque existe uma conexão capacitiva na carga e na fonte. Esta conexão capacitiva é implementada para permitir a transmissão de potência CC ao longo do cabo para alimentar o aparelho telefônico.

A análise da resposta em freqüência de cabos longos indica que ela pode ser melhorada adicionando-se alguma carga reativa. Por esta razão, geralmente escutamos o termo linhas carregadas que incluem indutores em intervalos regulares (tipicamente 66 milihenry (mH) a cada dois quilômetros, aproximadamente). A carga melhora a resposta em freqüência do circuito na faixa correspondente aos sinais de voz sem a necessidade de potência adicional. A desvantagem das linhas carregadas, entretanto, é a performance degradada para altas freqüências. Serviços tais como DSL (*Digital Subscriber Line*) (discutidos no Capítulo 7), que requerem uma resposta de alta freqüência dos pares trançados, não trabalham bem em linhas telefônicas carregadas.

Na maior parte no restante do livro iremos assumir, geralmente, que o meio não afeta a transmissão, exceto, possivelmente, através da adição de ruído ao sinal. Na prática, o meio pode afetar o sinal de diversas formas, como ilustrado no tema de exemplo descrito nesta seção.

2.12 RESUMO E DISCUSSÃO

Neste capítulo descrevemos a transformada de Fourier como a ferramenta fundamental para relacionarmos as descrições de um sinal determinístico no domínio do tempo e no domínio da freqüência. A transformada de Fourier inclui a série exponencial de Fourier como caso especial, desde que possamos utilizar a função delta de Dirac.

Uma relação inversa existe entre as descrições de um sinal no domínio do tempo e no domínio da freqüência. Sempre que uma operação é executada na forma de onda do sinal no domínio do tempo, uma modificação correspondente é aplicada ao espectro do sinal no domínio da freqüência. Uma consequência importante desta relação inversa é o fato de que um produto tempo-largura de faixa de um sinal de energia é constante. As definições de duração de sinal e largura de faixa simplesmente afetam o valor da constante.

Uma operação de processamento de sinal importante geralmente encontrada em sistemas de comunicação é a filtragem linear. Esta operação envolve a convolução do sinal de entrada com a resposta ao impulso do filtro ou, de forma equivalente, a multiplicação da transformada de Fourier do sinal de entrada pela função de transferência (isto é, a transformada de Fourier da resposta ao impulso) do filtro. Note, entretanto, que o material apresentado no capítulo de filtragem linear assume que o filtro é invariante no tempo (isto é, a forma da resposta ao impulso do filtro é invariante com relação a quando o pulso unitário, ou a função delta, é aplicado ao filtro).

Outra importante operação no processamento de sinais encontrada em sistemas de comunicação é a correlação. Esta operação pode fornecer uma medida de similaridade entre o sinal e sua versão atrasada, no caso de falarmos da função de autocorrelação. Quando a medida de similaridade envolve um par de sinais diferentes, entretanto, falamos da função de correlação cruzada. A transformada de Fourier da função de autocorrelação é chamada de densidade espectral. A transformada de Fourier da função de correlação cruzada é chamada de densidade espectral cruzada. Discussões sobre a correlação e densidade espectral apresentadas no capítulo são restrinidas a sinais de energia e sinais de potência, exemplificados por sinais do tipo pulso e sinais periódicos, respectivamente. O tratamento do ruído (outra realização de sinal de potência) será apresentado no Capítulo 8.

A parte final do capítulo focalizou a transformada discreta de Fourier e seus cálculos numéricos. Basicamente, a transformada discreta de Fourier é obtida da transformada de Fourier padrão amostrando uniformemente tanto o sinal de entrada quanto o espectro de saída. O algoritmo da transformada rápida de Fourier fornece um meio prático para a implementação eficiente da transformada discreta de Fourier em computadores digitais, tornando o algoritmo da transformada rápida de Fourier em uma ferramenta computacional poderosa na análise espectral e filtragem linear.

Problemas adicionais

- 2.19 (a) Determine a transformada de Fourier do pulso de meio cosseno mostrado na Fig. 2.40(a).

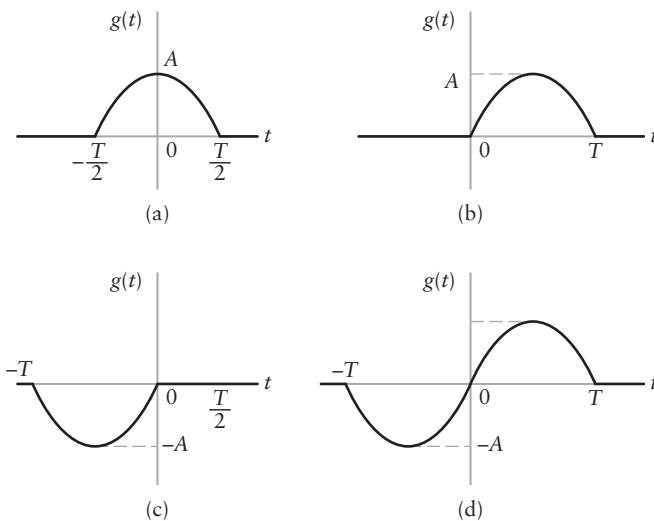


Figura 2.40
Problema 2.19

- (b) Aplique a propriedade de deslocamento no tempo ao resultado obtido na parte (a) para calcular o espectro do pulso de meio seno mostrado na Fig. 2.40(b).
- (c) Qual é o espectro do pulso de meio seno com duração igual a aT ?
- (d) Qual é o espectro do pulso de meio seno negativo mostrado na Fig. 2.40(c)?
- (e) Determine o espectro do pulso de meio seno mostrado na Fig. 2.40(d).
- 2.20 Qualquer função $g(t)$ pode ser dividida de forma não ambígua em uma parte par e uma parte ímpar, como mostrado por

$$g(t) = g_e(t) + g_o(t)$$

A parte par é definida por

$$g_e(t) = \frac{1}{2}[g(t) + g(-t)]$$

e a parte ímpar é definida por

$$g_o(t) = \frac{1}{2}[g(t) - g(-t)]$$

- (a) Calcule as partes par e ímpar do pulso retangular definido por

$$g(t) = A \operatorname{ret}\left(\frac{t}{T} - \frac{1}{2}\right)$$

- (b) Quais são as transformadas de Fourier destas duas partes do pulso?

- 2.21 A expressão a seguir pode ser vista como uma representação aproximada de um pulso com tempo de subida finito.

$$g(t) = \frac{1}{\tau} \int_{t-\tau}^{t+\tau} \exp\left(-\frac{\pi u^2}{\tau^2}\right) du$$

na qual assume-se que $T \gg \tau$. Determine a transformada de Fourier de $g(t)$. O que acontecerá a esta transformada se τ se tornar zero?

- 2.22 A transformada de Fourier de um sinal $g(t)$ é representada por $G(f)$. Prove as seguintes propriedades da transformada de Fourier:

- (a) Se um sinal real $g(t)$ é uma função par do tempo t , a transformada de Fourier $G(f)$ é puramente real. Se um sinal real $g(t)$ é uma função ímpar do tempo t , a transformada de Fourier $G(f)$ é puramente imaginária.

(b) $t^n g(t) \iff \left(\frac{j}{2\pi}\right)^n G^{(n)}(f)$, na qual $G^{(n)}(f)$ é a n -ésima $G(f)$ com respeito a f .

(c) $\int_{-\infty}^{\infty} t^n g(t) dt = \left(\frac{j}{2\pi}\right)^n G^{(n)}(0)$

(d) $\int_{-\infty}^{\infty} g_1(t) g_2^*(t) dt = \int_{-\infty}^{\infty} G_1(f) G_2^*(f) df$

- 2.23 A transformada de Fourier $G(f)$ de um sinal $g(t)$ é limitada pelas três seguintes desigualdades:

(a) $|G(f)| \leq \int_{-\infty}^{\infty} |g(t)| dt$

(b) $|j2\pi f G(f)| \leq \int_{-\infty}^{\infty} \left| \frac{dg(t)}{dt} \right| dt$

(c) $|(j2\pi f)^2 G(f)| \leq \int_{-\infty}^{\infty} \left| \frac{d^2 g(t)}{dt^2} \right| dt$

Considera-se, ainda, que a primeira e a segunda derivada de $g(t)$ existe.

Construa estes três limites para o pulso triangular mostrado na Fig. 2.41 e compare seu resultado com o espectro de amplitude do pulso regular.

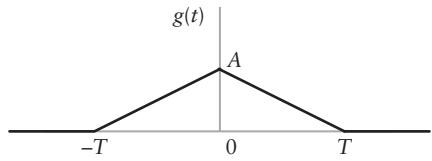


Figura 2.41
Problema 2.23

- 2.24** Considere a convolução de dois sinais $g_1(t)$ e $g_2(t)$. Mostre que

$$(a) \frac{d}{dt} [g_1(t) \star g_2(t)] = \left[\frac{d}{dt} g_1(t) \right] \star g_2(t)$$

$$(b) \int_{-\infty}^t [g_1(\tau) \star g_2(\tau)] d\tau = \left[\int_{-\infty}^t g_1(\tau) d\tau \right] \star g_2(t)$$

- 2.25** Um sinal $x(t)$ de energia finita é aplicado a um dispositivo de lei quadrática cuja saída $y(t)$ é definida por

$$y(t) = x^2(t)$$

O espectro de $x(t)$ é limitado ao intervalo de freqüência $-W \leq f \leq W$. Logo, mostre que o espectro de $y(t)$ é limitado a $-2W \leq f \leq 2W$. Dica: Expresse $y(t)$ como $x(t)$ multiplicado por ele mesmo.

- 2.26** Calcule a transformada de Fourier da função delta considerando-a como forma limite de (a) pulso retangular de área unitária e (b) pulso sinc com área unitária.

- 2.27** A transformada de Fourier $G(f)$ de um sinal $g(t)$ é definida por

$$G(f) = \begin{cases} 1, & f > 0 \\ \frac{1}{2}, & f = 0 \\ 0, & f < 0 \end{cases}$$

Determine o sinal $g(t)$.

- 2.28** Considere uma função $g(t)$ do tipo pulso constituída por um pequeno número de segmentos retos de linha. Suponha que esta função é diferenciável duas vezes com respeito ao tempo t de forma a gerar uma seqüência de funções delta ponderadas, como mostrado por

$$\frac{d^2g(t)}{dt^2} = \sum_i k_i \delta(t - t_i)$$

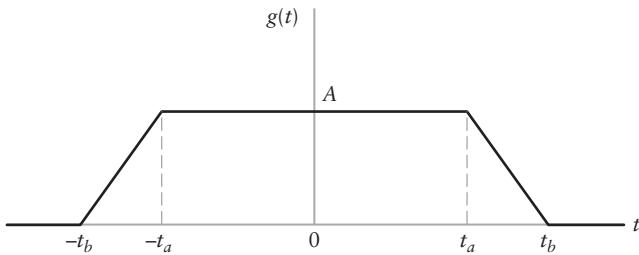
Na qual k_i estão relacionados a inclinação dos segmentos de linha reta.

- (a) Dados os valores de k_i e t_i , mostre que a transformada de Fourier de $g(t)$ é dada por

$$G(f) = -\frac{1}{4\pi^2 f^2} \sum_i k_i \exp(-j2\pi f t_i)$$

- (b) Usando este procedimento, mostre que a transformada de Fourier do pulso trapezoidal mostrado na Fig. 2.42 é dado por

$$G(f) = \frac{A}{\pi^2 f^2 (t_b - t_a)} \sin[\pi f(t_b - t_a)] \sin[\pi f(t_b + t_a)]$$

Figura 2.42
Problema 2.28

- 2.29 O pulso retangular de amplitude A e duração $2t_a$ pode ser visto como o caso limite do pulso trapezoidal mostrado na Fig. 2.42 quando t_b se aproxima de t_a .
- Começando do resultado dado na parte (b) do Problema 2.28, mostre que quando t_b se aproxima de t_a , este resultado se aproxima da função sinc.
 - Reconcilie o resultado obtido na parte (a) com o par transformada de Fourier da Eq. (2.10).
- 2.30 Sejam $x(t)$ e $y(t)$ sinais de entrada e saída de um filtro linear invariante no tempo. Utilizando o teorema de Rayleigh para energia, mostre que se o filtro for estável e o sinal $x(t)$ possuir energia finita, então o sinal de saída $y(t)$ também terá energia finita. Ou seja, se

$$\int_{-\infty}^{\infty} |x(t)|^2 dt < \infty$$

então

$$\int_{-\infty}^{\infty} |y(t)|^2 dt < \infty$$

- 2.31 (a) Determine a resposta em amplitude total da conexão em série mostrada na Fig. 2.43, constituída por N estágios idênticos, cada um com constante de tempo RC igual a τ_0 .
- (b) Mostre que quando N tende ao infinito, a resposta em amplitude da conexão em série se aproxima da função gaussiana $\exp\left(-\frac{1}{2}f^2T^2\right)$, na qual, para cada valor de N , a constante de tempo τ_0 é selecionada de tal forma que a condição

$$\tau_0^2 = \frac{T^2}{4\pi^2 N}$$

é satisfeita.

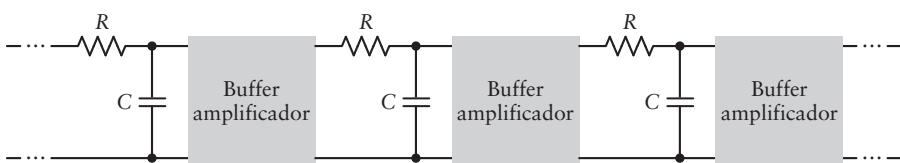


Figura 2.43 Problema 2.31

- 2.32 Suponha que, para um dado sinal $x(t)$, o valor integrado do sinal em um intervalo T é necessário, como mostrado por

$$y(t) = \int_{t-T}^t x(\tau) d\tau$$

- (a) Mostre que $y(t)$ pode ser obtido pela transmissão do sinal $x(t)$ através de um filtro cuja função de transferência é dada por

$$H(f) = T \operatorname{sinc}(fT) \exp(-j\pi fT)$$

- (b) Uma aproximação adequada desta função de transferência é obtida utilizando um filtro passa-baixa com largura de faixa igual a $1/T$, resposta em amplitude da faixa passante igual a T e atraso $T/2$. Assumindo este filtro passa-baixa como ideal, determine a saída do filtro no tempo $t = T$ devido a uma função degrau unitário aplicada ao filtro em $t = 0$ e compare o resultado com a saída correspondente do integrador ideal. Note que $\text{Si}(\pi) = 1,85$ e $\text{Si}(\infty) = \pi/2$.

- 2.33 Mostre que os dois pulsos diferentes definidos nas partes (a) e (b) da Fig. 2.44 possuem a mesma densidade espectral de energia:

$$\Psi_g(f) = \frac{4A^2T^2 \cos^2(\pi Tf)}{\pi^2(4T^2f^2 - 1)^2}$$

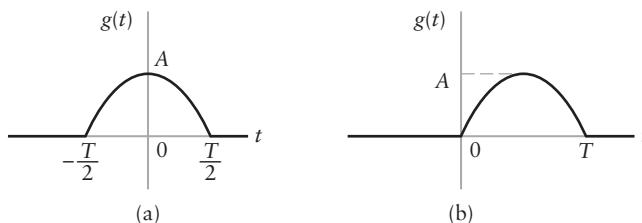


Figura 2.44 Problema 2.33

- 2.34 Determine e rascunhe as funções de autocorrelação dos seguintes pulsos exponenciais:

(a) $g(t) = \exp(-at)u(t)$

(b) $g(t) = \exp(-a|t|)$

(c) $g(t) = \exp(-at)u(t) - \exp(at)u(-t)$

na qual $u(t)$ é a função degrau unitário e $u(-t)$ é sua versão revertida no tempo.

- 2.35 Determine e rascunhe a função de autocorrelação do pulso Gaussiano definido por

$$g(t) = \frac{1}{t_0} \exp\left(-\frac{\pi t^2}{t_0^2}\right)$$

- 2.36 A transformada de Fourier de um sinal é definida por $\text{sinc}(f)$. Mostre que a função de autocorrelação deste sinal possui forma triangular.

- 2.37 Especifique dois sinais diferentes de pulso que possuem exatamente a mesma função de autocorrelação.

- 2.38 Considere o sinal senoidal $g(t)$ definido por

$$g(t) = A_0 + A_1 \cos(2\pi f_1 t + \theta_1) + A_2 \cos(2\pi f_2 t + \theta_2)$$

- (a) Determine a função de autocorrelação $R_g(\tau)$ deste sinal.

- (b) Qual é o valor de $R_g(0)$?

- (c) Alguma informação sobre $g(t)$ foi perdida na obtenção da função de autocorrelação? Explique.

- 2.39 Determine a função de autocorrelação do pulso triplo mostrado na Fig. 2.45.

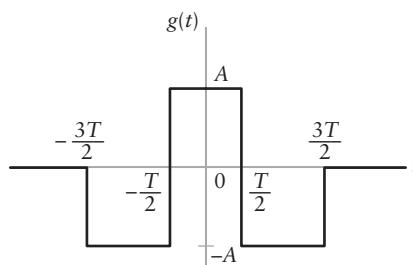


Figura 2.45 Problema 2.39

- 2.40 Seja $G(f)$ a transformada de Fourier de um sinal de energia de valor real $g(t)$ e $R_g(\tau)$ a sua função de autocorrelação. Mostre que

$$\int_{-\infty}^{\infty} \left[\frac{dR_g(\tau)}{d\tau} \right] d\tau = 4\pi^2 \int_{-\infty}^{\infty} f^2 |G(f)|^4 df$$

- 2.41 Determine a função de correlação cruzada $R_{12}(\tau)$ do pulso retangular $g_1(t)$ e do pulso triplo $g_2(t)$ mostrado na Fig. 2.46 e mostre-a em um gráfico. Qual é $R_{21}(\tau)$? Estes dois sinais são ortogonais entre si? Por quê?

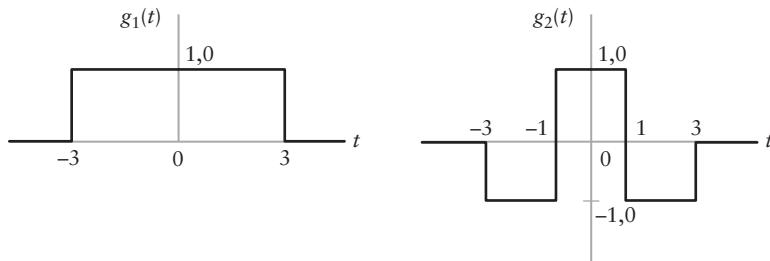


Figura 2.46 Problema 2.41

- 2.42 Considere dois sinais de energia $g_1(t)$ e $g_2(t)$. Estes dois sinais são, então, atrasados por um total igual a t_1 e t_2 segundos, respectivamente. Mostre que os atrasos de tempo são aditivos na convolução do par de sinais atrasados, enquanto que eles são subtrativos na correlação cruzada dos sinais.

- 2.43 (a) Um sinal de energia $x(t)$, sua transformada de Fourier $X(f)$, a função de autocorrelação $R_x(\tau)$ e a densidade espectral de energia $\Psi_x(f)$ estão todos relacionados, direta ou indiretamente. Construa um grafo de fluxo que mostre todas as possíveis relações diretas entre eles.
 (b) Se você tiver a descrição $X(f)$ no domínio do tempo, a função de autocorrelação $R_x(\tau)$ pode ser calculada a partir de $X(f)$. Descreva duas formas pelas quais este cálculo pode ser realizado.

- 2.44 Determine a função de autocorrelação de um sinal de potência $g(t)$ cuja densidade espectral de potência é mostrada na Fig. 2.47. Qual é o valor da função de autocorrelação na origem?

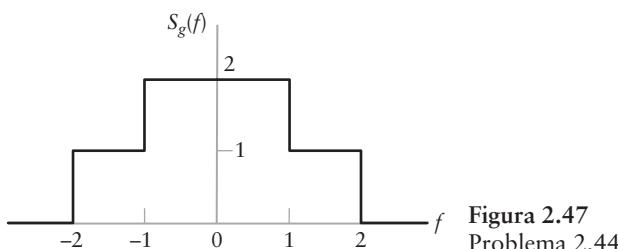


Figura 2.47
Problema 2.44

- 2.45 Considere a forma de onda quadrada $g(t)$ mostrada na Fig. 2.48. Determine a densidade espectral de potência e a função de autocorrelação desta onda quadrada. A forma de onda possui potência CC? Explique sua resposta.

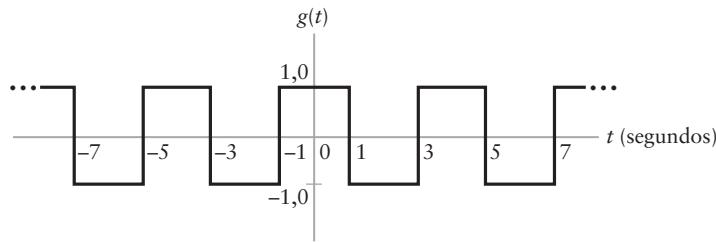


Figura 2.48 Problema 2.45

- 2.46** Considere dois sinais periódicos $g_{p1}(t)$ e $g_{p2}(t)$ que possuem as seguintes representações por série de Fourier:

$$g_{p1}(t) = \sum_{n=-\infty}^{\infty} c_{1,n} \exp\left(\frac{j2\pi nt}{T_0}\right)$$

e

$$g_{p2}(t) = \sum_{n=-\infty}^{\infty} c_{2,n} \exp\left(\frac{j2\pi nt}{T_0}\right)$$

Os dois sinais possuem um período comum igual a T_0 .

Usando a seguinte definição de função de correlação cruzada para um par de sinais periódicos,

$$R_{12}(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_{p1}(t) g_{p2}^*(t - \tau) dt$$

mostre que o par de sinais periódicos descritos satisfaz o par transformada de Fourier

$$R_{12}(\tau) \iff \sum_{n=-\infty}^{\infty} c_{1,n} c_{2,n}^* \delta\left(f - \frac{n}{T_0}\right)$$

- 2.47** Um sinal periódico $g_p(t)$ com período T_0 é representado pela série de Fourier complexa

$$g_p(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi nt/T_0)$$

na qual c_n são os coeficientes complexos de Fourier. A função de autocorrelação de $g_p(t)$ é definida por

$$R_{g_p}(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_p(t) g_p^*(t - \tau) dt$$

- (a) Considere a forma de onda senoidal

$$g_p(t) = A \cos(2\pi f_c t + \theta)$$

Determine a função de autocorrelação $R_{g_p}(\tau)$ e trace sua forma de onda.

- (b) Mostre que $R_{g_p}(0) = A^2/2$.

- 2.48** Repita as partes (a) e (b) do Problema 2.47 para a onda quadrada

$$g_p(t) = \begin{cases} A, & -\frac{T_0}{4} \leq t \leq \frac{T_0}{4} \\ 0, & \text{para o restante do período } T_0 \end{cases}$$

- 2.49 Determine a densidade espectral de potência de (a) a forma de onda senoidal do Problema 2.47 e (b) a onda quadrada do Problema 2.48.
- 2.50 Seguindo um procedimento similar ao descrito na Seção 2.10 que levou ao grafo de fluxo da Fig. 2.36 para o algoritmo de FFT de 8 pontos baseado na decimação em freqüência, faça o seguinte:
- Desenvolva o grafo de fluxo correspondente para um algoritmo de FFT de 8 pontos baseado na decimação no tempo.
 - Compare o grafo de fluxo obtido na parte (a) com o mostrado na Fig. 2.36, ressaltando as similaridades e diferenças entre estes dois métodos básicos para a obtenção do algoritmo de FFT.

Problemas avançados

- 2.51 (a) A *largura de faixa de raiz média quadrática (rms)* de um sinal passa-baixa $g(t)$ com energia finita é definido por

$$W_{\text{rms}} = \left[\frac{\int_{-\infty}^{\infty} f^2 |G(f)|^2 df}{\int_{-\infty}^{\infty} |G(f)|^2 df} \right]^{1/2}$$

Na qual $|G(f)|^2$ é a densidade espectral de energia do sinal. De forma correspondente, a *duração de raiz média quadrática* do sinal é definida por

$$T_{\text{rms}} = \left[\frac{\int_{-\infty}^{\infty} t^2 |g(t)|^2 dt}{\int_{-\infty}^{\infty} |g(t)|^2 dt} \right]^{1/2}$$

Usando estas definições, mostre que

$$T_{\text{rms}} W_{\text{rms}} \geq \frac{1}{4\pi}$$

Assuma que $|g(t)| \rightarrow 0$ mais rápido do que $1/\sqrt{|t|}$ quando $|t| \rightarrow \infty$.

- (b) Considere um pulso Gaussiano definido por

$$g(t) = \exp(-\pi t^2)$$

Mostre que, para este sinal, a igualdade

$$T_{\text{rms}} W_{\text{rms}} = \frac{1}{4\pi}$$

pode ser atingida.

Dica: Utilize a desigualdade de Schwarz (veja o Apêndice 5).

$$\left\{ \int_{-\infty}^{\infty} [g_1^*(t)g_2(t) + g_1(t)g_2^*(t)] dt \right\}^2 \leq 4 \int_{-\infty}^{\infty} |g_1(t)|^2 \int_{-\infty}^{\infty} |g_2(t)|^2 dt$$

na qual fazemos

$$g_1(t) = tg(t)$$

e

$$g_2(t) = \frac{dg(t)}{dt}$$

- 2.52 A *transformada de Hilbert* de um sinal $g(t)$ que possui transformada de Fourier é definida por

$$\hat{g}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{g(\tau)}{t - \tau} d\tau$$

De forma similar, a *transformada inversa de Hilbert* é definida por

$$g(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{g}(\tau)}{t - \tau} d\tau$$

Usando estas duas fórmulas, obtenha o seguinte conjunto de pares de transformada de Hilbert

$g(t)$	$\hat{g}(t)$
$\frac{\sin t}{t}$	$\frac{1 - \cos t}{t}$
$\operatorname{ret}(t)$	$-\frac{1}{\pi} \ln \left \left(t - \frac{1}{2} \right) / \left(t + \frac{1}{2} \right) \right $
$\delta(t)$	$\frac{1}{\pi t}$
$\frac{1}{1 + t^2}$	$\frac{t}{1 + t^2}$

- 2.53 Calcule a transformada inversa de Fourier $g(t)$ da função (a qual possui espectro apenas para freqüências não negativas).

$$G(f) = \begin{cases} \exp(-f), & f > 0 \\ \frac{1}{2}, & f = 0 \\ 0, & f < 0 \end{cases}$$

Logo, mostre que $g(t)$ é complexa, e que suas partes real e imaginária constituem um par transformada de Hilbert.

- 2.54 Um transformador de Hilbert pode ser visto como um dispositivo cuja função de transferência possui as seguintes características:

- (a) A resposta em amplitude é unitária para todas as freqüências positivas e negativas.
- (b) A resposta em fase é $+90^\circ$ para freqüências negativas e -90° para freqüências positivas. Começando com a definição da transformada de Hilbert dada no Problema 2.52, demonstre a descrição no domínio da freqüência embutida nas partes (a) e (b).
- (c) O transformador de Hilbert é fisicamente realizável? Justifique sua resposta.

Capítulo 3

MODULAÇÃO EM AMPLITUDE

A modulação é definida como *o processo pelo qual alguma característica da onda portadora é variada de acordo com o sinal contendo a informação*. A portadora é necessária para facilitar o transporte do sinal modulado ao longo de um canal passa-faixa, do transmissor ao receptor. Uma portadora geralmente utilizada é a onda senoidal, cuja fonte é fisicamente independente da fonte do sinal contendo a informação. Quando o sinal contendo a informação é do tipo analógico, falamos de uma *modulação de onda contínua*, um termo que ressalta a continuidade da onda modulada em função do tempo.

No contexto de comunicações, a motivação primária para a modulação é facilitar a transmissão de um sinal contendo informação em um canal de comunicação (por exemplo, um canal de rádio), com uma faixa passante determinada. Na modulação de onda contínua, isto é possível variando-se a amplitude ou o ângulo da forma de onda da portadora. Com esta base, podemos classificar a modulação de onda contínua em duas famílias genéricas: modulação em amplitude e modulação em ângulo. Estas duas famílias de modulação se distinguem pelo oferecimento de características espectrais completamente distintas e, portanto, benefícios práticos diferentes. A classificação é feita se a amplitude da portadora ou o ângulo de fase da portadora é variado de acordo com o sinal contendo a informação. A família de modulação em amplitude é estudada neste capítulo, seguido pelo estudo da modulação em ângulo no próximo capítulo.

No Capítulo 1, identificamos a complexidade dos sistemas e os dois recursos primários de comunicação – explicitamente, potência transmitida e largura de faixa do canal – como questões centrais envolvidas no projeto de um sistema de comunicação. Com estas questões em mente, neste capítulo iremos estudar quatro *estratégias de modulação lineares que constituem a família de modulação em amplitude*:

- ▶ Modulação em amplitude (AM)
- ▶ Faixa lateral dupla – portadora suprimida (DSB-SC)
- ▶ Faixa lateral única (SSB)
- ▶ Faixa lateral vestigial (VSB)

Estes quatro tipos de modulação diferem um do outro em virtude de suas características espectrais. Este estudo nos ensinará as seguintes lições:

- ▶ *Lição 1: A análise de Fourier é uma poderosa ferramenta matemática para o desenvolvimento da análise matemática e física das características espectrais das estratégias de modulação linear.*
- ▶ *Lição 2: A implementação de comunicações analógicas é significativamente simplificada pela utilização de AM, ao custo da potência transmitida e da largura de faixa do canal.*
- ▶ *Lição 3: A utilização da potência transmitida e da largura de faixa do canal é melhorada através de modificações bem definidas do conteúdo espectral da onda modulada em amplitude, ao custo de aumento na complexidade do sistema.*

Resumindo, podemos fazer a seguinte afirmação:

Não existe almoço grátis no projeto de um sistema de comunicação: Para cada ganho realizado, existe um preço a ser pago.

3.1 MODULAÇÃO EM AMPLITUDE

■ Teoria

Considere uma *onda portadora senoidal* $c(t)$ definida por

$$c(t) = A_c \cos(2\pi f_c t) \quad (3.1)$$

na qual A_c é a *amplitude da portadora* e f_c é a *freqüência da portadora*. O sinal contendo a informação, ou o *sinal de mensagem* é representado por $m(t)$. Os termos “sinal contendo a informação” e “sinal de mensagem” são utilizados como sinônimos ao longo do livro. Para simplificar a exposição sem afetar os resultados e conclusões obtidas, assumimos que a fase da portadora é zero na Eq. (3.1). A *modulação em amplitude*¹ é formalmente definida como o processo pelo qual o valor médio da amplitude da onda portadora $c(t)$ é variado linearmente com o sinal de mensagem $m(t)$. Um sinal modulado em amplitude (AM) pode, portanto, ser descrito como uma função do tempo usando:

$$s(t) = A_c[1 + k_a m(t)] \cos(2\pi f_c t) \quad (3.2)$$

na qual k_a é uma constante chamada de *sensibilidade de amplitude* do modulador responsável pela geração do sinal modulado $s(t)$. Tipicamente, a amplitude A_c da portadora e o sinal de mensagem $m(t)$ são medidos em volts, e, neste caso, a sensibilidade de amplitude k_a é medida em volt⁻¹. Observe que se o sinal de mensagem $m(t)$ for desligado, a portadora senoidal é mantida intacta.

A Fig. 3.1(a) mostra o sinal de mensagem $m(t)$ e as Figs. 3.1(b) e (c) mostram as ondas $s(t)$ AM correspondentes a dois valores de sensibilidade k_a , com uma amplitude da portadora $A_c = 1$ volt.

Na modulação em amplitude, a informação contida no sinal de mensagem $m(t)$ está somente no *envelope*, o qual é definido como a amplitude da onda modulada $s(t)$ – ou seja, $A_c|1 + k_a m(t)|$. Desta expressão, observamos que o envelope de $s(t)$ possui, essencialmente, a mesma forma do sinal de mensagem $m(t)$ desde que duas condições sejam satisfeitas:

1. A amplitude de $k_a m(t)$ é sempre menor do que a unidade, ou seja,

$$|k_a m(t)| < 1, \quad \text{para todo } t \quad (3.3)$$

Esta condição é ilustrada na Fig. 3.1(b). Ela garante que a função $1 + k_a m(t)$ é sempre positiva e, neste caso, podemos expressar o envelope da onda AM $s(t)$ da Eq. (3.2) simplesmente como $A_c[1 + k_a m(t)]$. Quando a sensibilidade k_a do modulador é grande o suficiente para garantir $|k_a m(t)| > 1$ para qualquer t , a onda portadora se torna sobremodulada, resultando em inversões de fase da portadora sempre que o fator $1 + k_a m(t)$ cruzar o zero. A onda modulada, então, exibe uma distorção do envelope, como mostrado na Fig. 3.1(c). Portanto, é aparente que evitando-se a sobremodulação, uma relação de um-par-um é mantida entre o envelope da onda AM e a onda modulante para todos os valores de tempo. O valor máximo absoluto de $k_a m(t)$ multiplicado por 100 é chamado de *percentual de modulação*.

2. A freqüência da portadora f_c é muito maior do que a maior componente de freqüência W do sinal de mensagem $m(t)$, ou seja,

$$f_c \gg W \quad (3.4)$$

Chamamos W de *largura de faixa da mensagem*. Se a condição da Eq. (3.4) não for satisfeita, um envelope não pode ser visualizado (e portanto não detectado) satisfatoriamente.

¹ Ao longo do livro, o termo “modulação em amplitude” ou AM, é utilizado para referir a uma forma particular de modulação na qual tanto a onda portadora quanto as duas faixas laterais estão presentes.

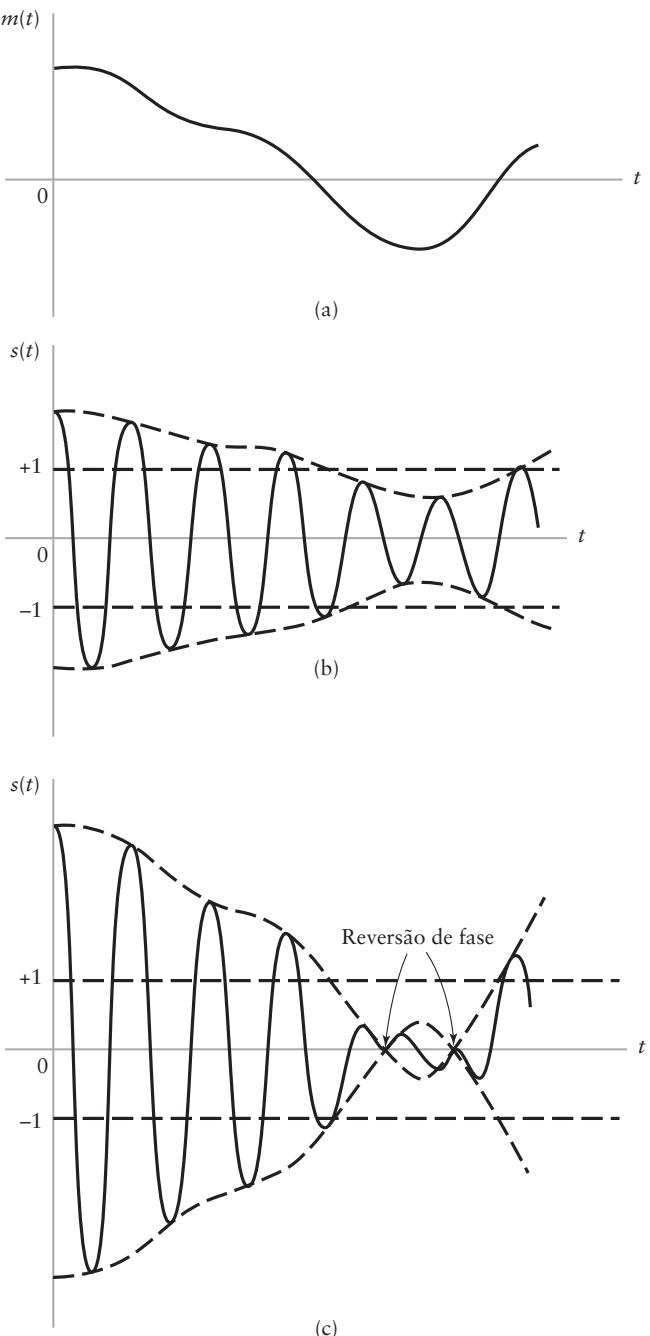


Figura 3.1 Ilustração do processo de modulação em amplitude. (a) Sinal de mensagem $m(t)$. (b) Onda AM para $k_a m(t) < 1$ para todo t . (c) Onda AM para $|k_a m(t)| > 1$ para algum tempo t .

Desde que as condições das Eqs. (3.3) e (3.4) sejam satisfeitas, a demodulação da onda AM é realizada utilizando um *detector de envelope*, o qual é definido como *um dispositivo cuja saída acompanha o envelope da onda AM que atua como sinal de entrada*. O processo de detecção de envelope será discutido mais adiante.

A próxima questão da nossa discussão é a descrição no domínio da freqüência da AM. Seja $m(t) \iff M(f)$, na qual a transformada de Fourier $M(f)$ é chamada de *espectro de mensagem*. A partir da Eq. (3.2), determinamos que a transformada de Fourier ou espectro da onda AM $s(t)$ é dada por

$$S(f) = \frac{A_c}{2} [\delta(f - f_c) + \delta(f + f_c)] + \frac{k_a A_c}{2} [M(f - f_c) + M(f + f_c)] \quad (3.5)$$

na qual utilizamos as relações:

$$\begin{aligned} \cos(2\pi f_c t) &= \frac{1}{2} [\exp(j2\pi f_c t) + \exp(-j2\pi f_c t)] \\ \exp(j2\pi f_c t) &\iff \delta(f - f_c) \end{aligned}$$

e

$$m(t) \exp(j2\pi f_c t) \iff M(f - f_c)$$

Seguindo a terminologia apresentada no Capítulo 2, a $\delta(t)$ representa a função delta de Dirac no domínio da freqüência.

Suponha que o sinal de mensagem $m(t)$ é limitado em faixa ao intervalo $-W \leq f \leq W$, como mostrado na Fig. 3.2(a). A forma do espectro mostrado nesta figura é utilizada somente com o propósito ilustrativo. Determinamos, da Eq. (3.5) que o espectro $S(f)$ da onda AM é como mostrado na Fig. 3.2(b) para o caso de $f_c > W$. Este espectro é constituído por duas funções delta ponderada pelo fator $A_c/2$ e ocorrendo em $\pm f_c$, e duas versões do espectro de mensagem transladasadas em freqüência por $\pm f_c$ e escalonadas em amplitude por $k_a A_c/2$. A partir do espectro da Fig. 3.2(b), podemos fazer três importantes observações:

1. Como resultado do processo de modulação, o espectro do sinal de mensagem $m(t)$ para freqüências negativas de $-W$ a 0 se torna completamente visível (isto é, mensurável) para freqüências positivas, desde que a freqüência da portadora satisfaça a

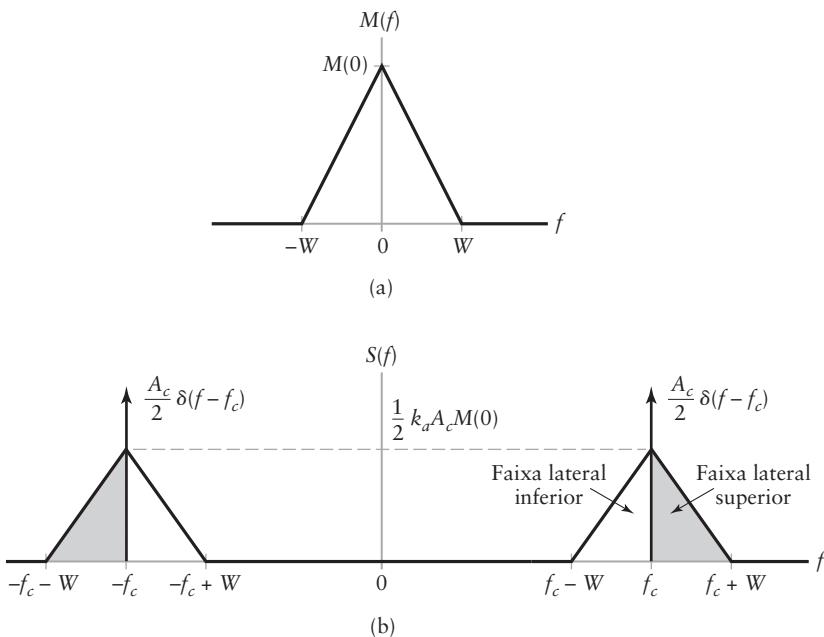


Figura 3.2 (a) Espectro do sinal de mensagem $m(t)$. (b) Espectro da onda AM $s(t)$.

condição $f_c > W$, aqui está a importância da idéia de freqüências “negativas”, a qual foi enfatizada no Capítulo 2.

2. Para freqüências positivas, a porção do espectro de uma onda AM acima da freqüência da portadora f_c é chamada de *faixa lateral superior*, enquanto que a porção simétrica abaixo de f_c é chamada de *faixa lateral inferior*. A condição $f_c > W$ garante que as faixas laterais não se sobreponham. Além disso, a faixa lateral superior, a faixa lateral inferior e a portadora são completamente representadas no espectro da Fig. 3.2(b), a onda modulada é chamada de AM, de acordo com a nota de rodapé da página 117.
3. Para freqüências positivas, a componente de freqüência mais alta da onda AM é igual a $f_c + W$, e a componente mais baixa de freqüência é $f_c - W$. A diferença entre estas duas freqüências define a *largura de faixa de transmissão* B_T da onda AM, a qual é exatamente duas vezes a largura de faixa W da mensagem. Ou seja,

$$B_T = 2W \quad (3.6)$$

EXEMPLO 3.1 Modulação de tom único

Considere a onda modulante $m(t)$ constituída por um único tom, ou componente de freqüência, ou seja,

$$m(t) = A_m \cos(2\pi f_m t)$$

na qual A_m é a amplitude da onda senoidal modulante e f_m é a sua freqüência (veja a Fig. 3.3(a)). A portadora senoidal possui amplitude A_c e freqüência f_c (veja a Fig. 3.3(b)). A onda AM correspondente é, portanto, dada por

$$s(t) = A_c[1 + \mu \cos(2\pi f_m t)] \cos(2\pi f_c t) \quad (3.7)$$

na qual

$$\mu = k_a A_m$$

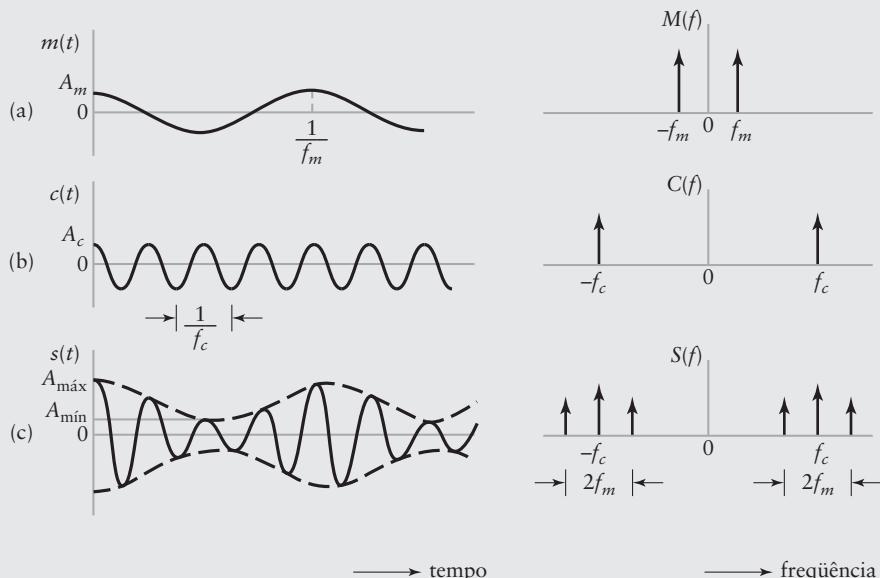


Figura 3.3 Ilustração das características da modulação em amplitude produzida por um sinal de único tom no domínio do tempo (à esquerda) e no domínio da freqüência (à direita). (a) Onda modulante. (b) Onda portadora. (c) Onda AM.

A constante adimensional μ é chamada de *fator de modulação*, ou *percentual de modulação*, quando é expressa numericamente como um percentual. Para evitar a distorção de envelope devido a sobremodulação, o fator de modulação μ deve ser mantido abaixo da unidade, como explicado anteriormente.

A Fig. 3.3(c) mostra um rascunho de $s(t)$ para μ menor do que a unidade. Sejam A_{\max} e A_{\min} os valores máximo e mínimo do envelope da onda modulada, respectivamente. Então, a partir da Eq. (3.7), temos que

$$\frac{A_{\max}}{A_{\min}} = \frac{A_c(1 + \mu)}{A_c(1 - \mu)}$$

Reorganizando esta equação, podemos expressar o fator de modulação por

$$\mu = \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}}$$

Expressando o produto dos dois cossenos da Eq. (3.7) como a soma de duas ondas senoidais, uma com freqüência $f_c + f_m$ e outra com freqüência $f_c - f_m$, temos

$$s(t) = A_c \cos(2\pi f_c t) + \frac{1}{2}\mu A_c \cos[2\pi(f_c + f_m)t] + \frac{1}{2}\mu A_c \cos[2\pi(f_c - f_m)t]$$

A transformada de Fourier de $s(t)$ é, portanto,

$$\begin{aligned} S(f) &= \frac{1}{2}A_c[\delta(f - f_c) + \delta(f + f_c)] \\ &\quad + \frac{1}{4}\mu A_c[\delta(f - f_c - f_m) + \delta(f + f_c + f_m)] \\ &\quad + \frac{1}{4}\mu A_c[\delta(f - f_c + f_m) + \delta(f + f_c - f_m)] \end{aligned}$$

Logo, o espectro da onda AM, para o caso especial de modulação senoidal, é constituído por funções delta em $\pm f_c$, $f_c \pm f_m$ e $-f_c \pm f_m$, como mostrado na Fig. 3.3c.

Na prática, a onda AM $s(t)$ é uma onda de tensão ou corrente. Em qualquer um dos casos, a potência média entregue em um resistor de 1 ohm por $s(t)$ é composta por três componentes:

$$\text{Potência da portadora} = \frac{1}{2}A_c^2$$

$$\text{Potência da freqüência lateral superior} = \frac{1}{8}\mu^2 A_c^2$$

$$\text{Potência da freqüência lateral inferior} = \frac{1}{8}\mu^2 A_c^2$$

Para uma carga resistiva diferente de 1 ohm, o que geralmente acontece na prática, as expressões para potência da portadora, potência da freqüência lateral superior e potência da freqüência lateral inferior são meramente escalonadas pelo fator $1/R$ ou R , dependendo se a onda modulada $s(t)$ é tensão ou corrente, respectivamente. Em qualquer caso, a relação da potência total das faixas laterais pela potência total na onda modulada é igual a $\mu^2/(2 + \mu^2)$, a qual depende somente do fator de modulação μ . Se $\mu = 1$ – ou seja, 100% de modulação é utilizado – a potência total das duas freqüências laterais da onda AM é somente um terço da potência total da onda modulada.

A Fig. 3.4 mostra o percentual da potência total das duas freqüências laterais e da portadora, traçados em função do percentual de modulação. Note que, quando o percentual de modulação é menor do que 20%, a potência em uma freqüência lateral é menor do que 1% da potência total da onda AM.

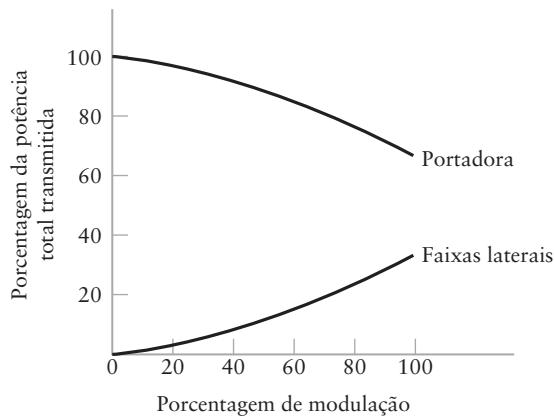


Figura 3.4 Variações da potência da portadora e da potência total das faixas laterais com o percentual de modulação na modulação em amplitude.

■ Experimento de computador: AM

Para o experimento AM, iremos estudar a modulação senoidal com base nos seguintes parâmetros:

$$\text{Amplitude da portadora, } A_c = 1$$

$$\text{Freqüência da portadora, } f_c = 0,4 \text{ Hz}$$

$$\text{Freqüência da modulante, } f_m = 0,05 \text{ Hz}$$

Queremos mostrar e analisar 10 ciclos completos da onda modulada, correspondendo a uma duração total de 200 segundos. Para executar este experimento em um computador digital, a onda modulada é amostrada a uma taxa $f_s = 10\text{Hz}$, obtendo um total de $200 \times f_s = 2000 \text{ Hz}$ pontos de dados. A faixa de freqüência ocupada pela onda modulada é $-5\text{Hz} \leq f \leq 5\text{Hz}$. Como a separação entre a freqüência da portadora e cada freqüência lateral é igual à freqüência modulante $f_m = 0,05 \text{ Hz}$, queremos ter uma resolução de freqüência $f_r = 0,005 \text{ Hz}$. Para alcançarmos esta resolução de freqüência, é recomendado que o número de amostras em freqüência satisfaça a condição:

$$M \geq \frac{f_s}{f_r} = \frac{10}{0,005} = 2000$$

Escolhemos, portanto, $M = 2000$. Para aproximar a transformada de Fourier da onda modulada, utilizamos um algoritmo de FFT de 200 pontos. O algoritmo de FFT foi descrito no Capítulo 2.

O único parâmetro variável em todo o experimento AM é o fator de modulação μ , com o qual três situações diferentes serão investigadas:

$$\mu = 0,5, \text{ correspondendo a submodulação}$$

$$\mu = 1,0, \text{ correspondendo a 100\% de modulação}$$

$$\mu = 2,0 \text{ correspondendo a sobremodulação}$$

Os resultados desta análise estão mostrados nas Figs. 3.5 a 3.7, sendo que os detalhes das figuras serão descritos a seguir.

1. Fator de modulação $\mu = 0,5$.

A Fig. 3.5(a) mostra 10 ciclos da onda AM, correspondendo a $\mu = 0,5$. O envelope da onda modulada é visto claramente, seguindo fielmente a onda senoidal modulante. Isto significa que podemos utilizar um detector de envelope para a demodulação.

A Fig. 3.5(b) mostra o espectro de amplitude da onda modulada. Na Fig. 3.5(c) fizemos uma ampliação da estrutura fina da onda modulada ao redor da freqüência da portadora. As duas figuras mostram, claramente, as relações exatas entre as duas

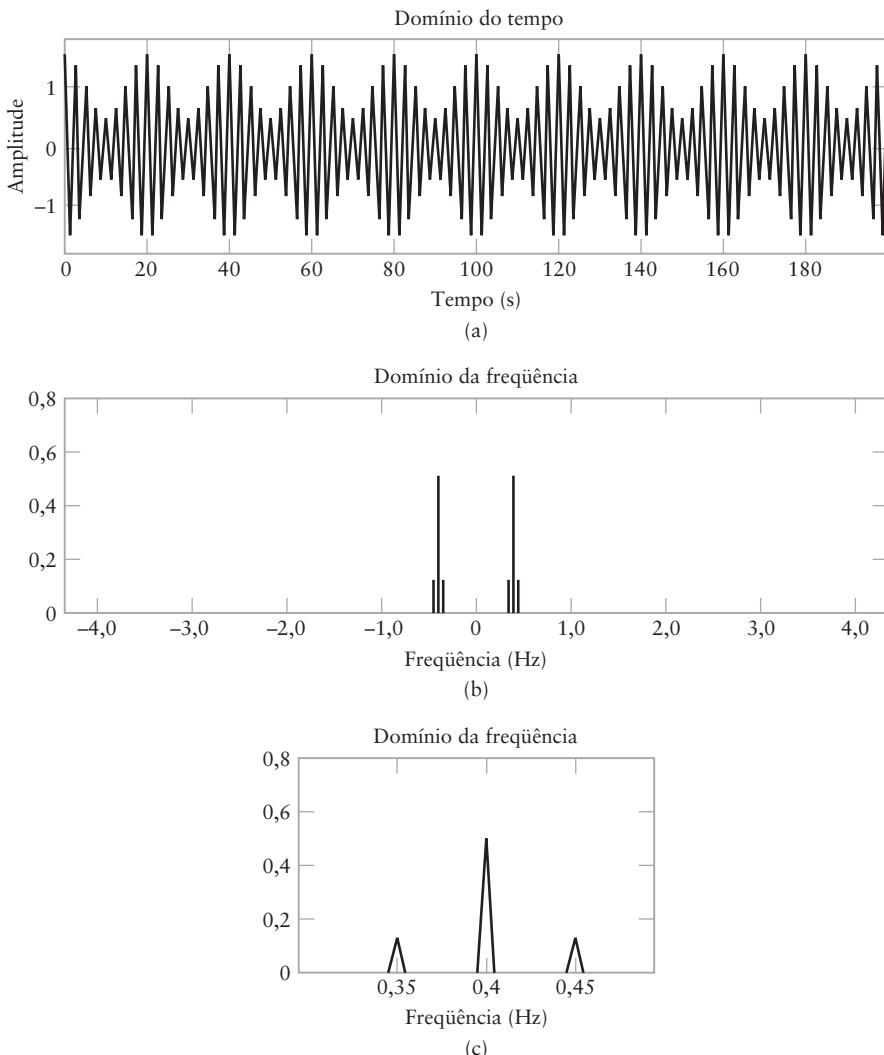


Figura 3.5 Modulação em amplitude com 50% de modulação: (a) onda AM, (b) espectro de amplitude da onda AM e (c) espectro expandido ao redor da freqüência da portadora.

freqüências laterais e a portadora, de acordo com a teoria de modulação em amplitude, resumidas em:

- A freqüência lateral inferior, a portadora e a freqüência lateral superior estão localizadas em $(f_c - f_m) = \pm 0,35\text{Hz}$, $f_c = 0,4\text{ Hz}$ e $(f_c + f_m) = \pm 0,45\text{ Hz}$.
- A amplitude das duas freqüências laterais é $(\mu/2) = 0,25$ vezes a amplitude da portadora.

2. Fator de modulação $\mu = 1,0$.

A Fig. 3.6(a) mostra 10 ciclos da onda AM com os mesmos parâmetros da Fig. 3.5(a), exceto pelo fato de que, agora, $\mu = 1,0$. Essa nova figura mostra que a onda modulada está, agora, no limite de sobremodulação.

O espectro de amplitude da onda modulada é mostrado na Fig. 3.6(b) e sua versão ampliada (ao redor da freqüência da portadora) mostrada na Fig. 3.6(c). Novamente, vemos que a estrutura básica do espectro em amplitude da onda modulada está em conformidade perfeita com a teoria de modulação em amplitude.

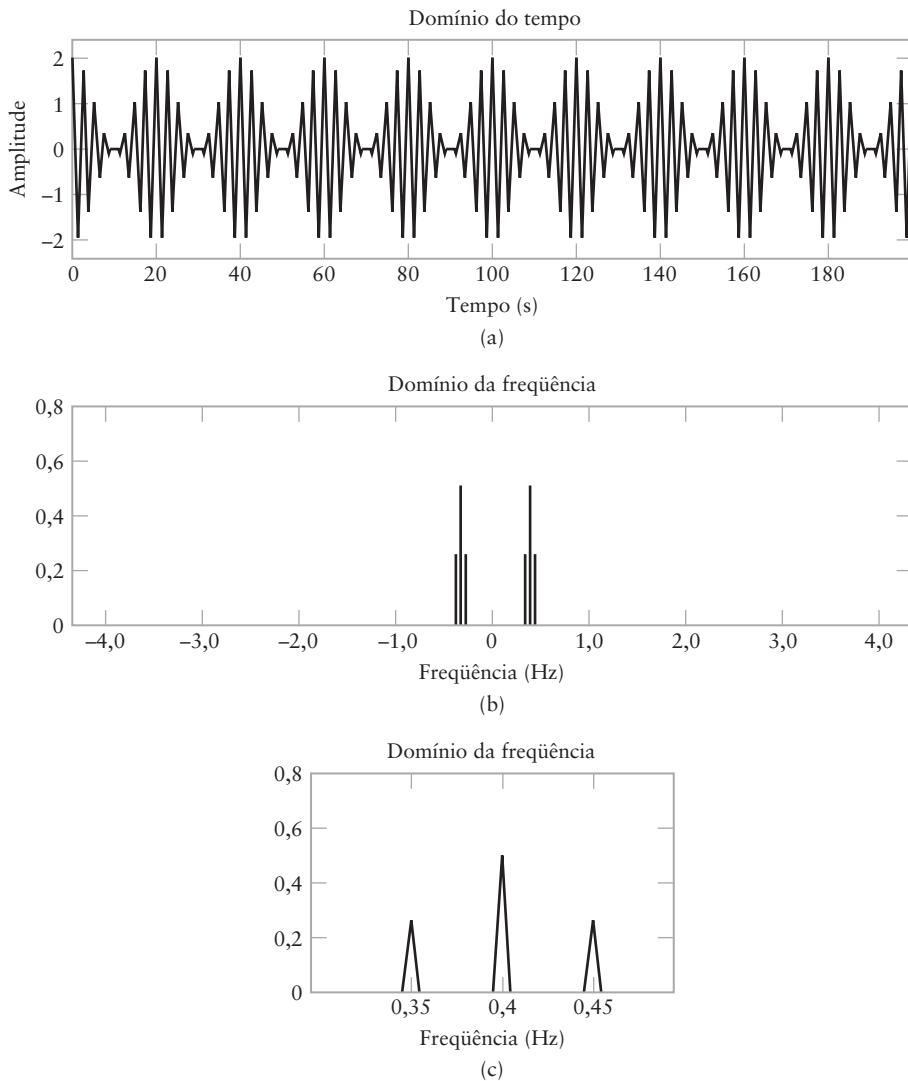


Figura 3.6 Modulação em amplitude com 100% de modulação: (a) onda AM, (b) espectro de amplitude da onda AM e (c) espectro expandido ao redor da freqüência da portadora.

3. Fator de modulação $\mu = 2,0$.

A Fig. 3.7(a) demonstra o efeito da sobremodulação utilizando o fator de modulação $\mu = 2,0$. Aqui vemos que não existe relação clara entre o envelope da onda sobre-modulada e a onda senoidal modulante. Como esperado, o resultado indica que um detector de envelope não irá funcionar para $\mu = 2,0$.

De qualquer forma, o conteúdo espectral da onda modulada mostrada nas Figs. 3.7(b) e 3.7(c) segue exatamente o que a teoria da modulação em amplitude prediz.

- **PROBLEMA 3.1** Para uma modulação de 100%, é possível que o envelope AM se torne zero para algum tempo t ? Justifique sua resposta. ◀
- **PROBLEMA 3.2** Para um caso particular de modulação AM utilizando onda modulante senoidal, o percentual de modulação é 20%. Calcule a potência média (a) na portadora e (b) em cada freqüência lateral. ◀

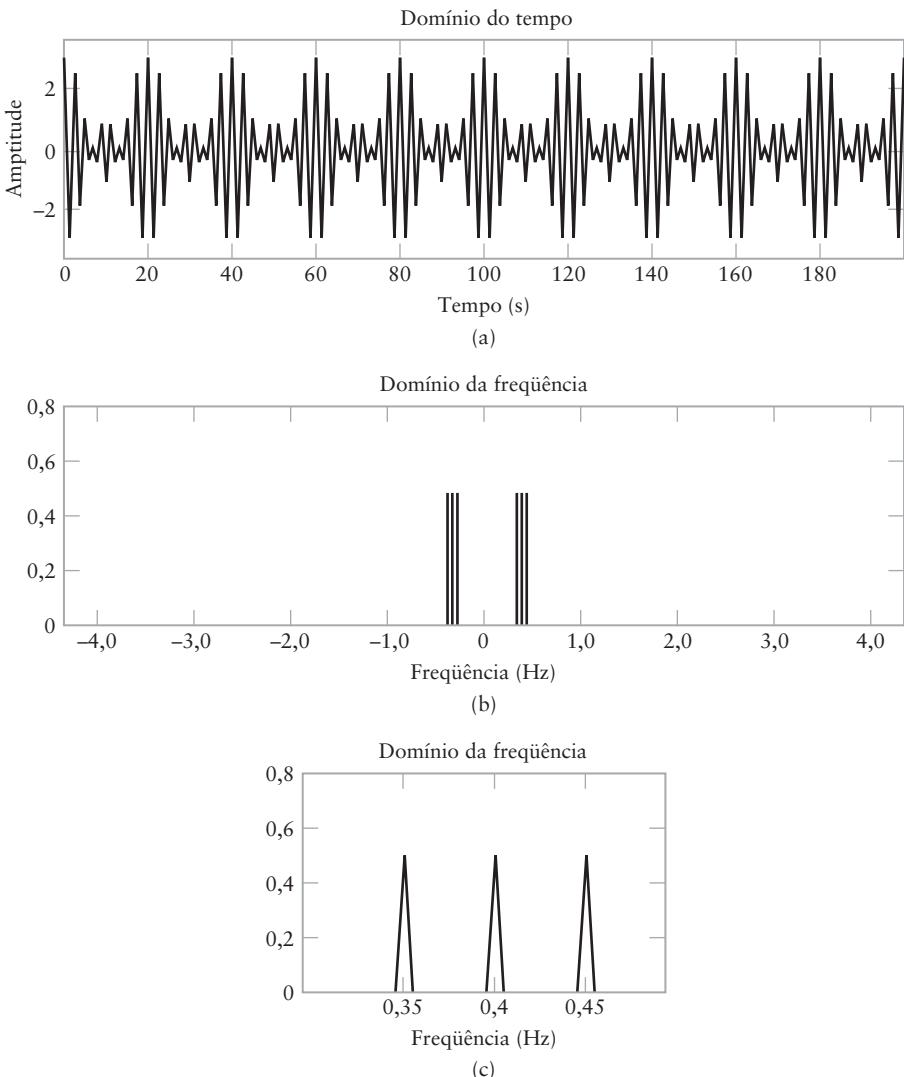


Figura 3.7 Modulação em amplitude com 200% de modulação: (a) onda AM, (b) espectro de amplitude da onda AM e (c) espectro expandido ao redor da freqüência da portadora.

- **PROBLEMA 3.3** No AM, diz-se que a sobreposição espectral ocorre se a faixa lateral inferior para freqüências positivas se sobrepor com sua imagem para freqüências negativas. Qual condição deve ser satisfeita pela onda modulada se quisermos evitar a sobreposição espectral? Assuma que o sinal de mensagem $m(t)$ é um sinal do tipo passa-baixa com largura de faixa W . ◀
- **PROBLEMA 3.4** Um modulador de lei quadrática para a geração de uma onda AM é baseado na utilização de um dispositivo não linear (por exemplo, um diodo). A Fig. 3.8 mostra a forma mais simples deste modulador. Ignorando os termos de alta ordem, a característica de entrada-saída da combinação diodo-carga desta figura é representada pela lei quadrática:

$$\nu_2(t) = \alpha_1 \nu_1(t) + \alpha_2 \nu_1^2(t)$$

na qual

$$\nu_1(t) = A_c \cos(2\pi f_c t) + m(t)$$

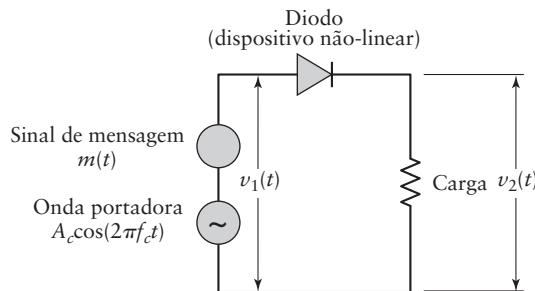


Figura 3.8 Circuito não linear utilizando um diodo.

é o sinal de entrada, $v_2(t)$ é o sinal de saída desenvolvido na carga resistiva, e a_1 e a_2 são constantes.

- Determine o conteúdo espectral do sinal de saída $v_2(t)$.
- Para extrair a forma de onda AM desejada de $v_2(t)$, precisamos de um filtro passa-faixa (não mostrado na Fig. 3.8). Determine as freqüências de corte do filtro necessário, assumindo que o sinal de mensagem é limitado a faixa $-W \leq f \leq W$.
- Para evitar a distorção espectral devida à presença de produtos de modulação não desejados em $v_2(t)$, a condição $W < f_c > 3W$ deve ser satisfeita. Certifique esta condição. ◀

■ Detecção de envelope

O modulador de lei quadrática apresentado no Problema 3.4 é um exemplo da simplicidade de implementação de um transmissor AM. A simplicidade de implementação do AM é ressaltada ainda mais quando consideramos a demodulação da onda AM, a qual é o inverso da modulação. Em particular, a demodulação de uma onda AM pode ser realizada através do simples e altamente eficaz circuito chamado *detector de envelope*², desde que duas condições práticas sejam satisfeitas:

- A onda AM é faixa estreita, o que significa que a freqüência da portadora é grande quando comparada com a largura de faixa da mensagem.
- O percentual de modulação da onda AM é menor do que 100%.

Um detector de envelope do tipo série é mostrado na Fig. 3.9(a), o qual é constituído por um diodo e um filtro resistor-capacitor (RC). A operação deste detector de envelope é descrita a seguir. No semi-ciclo positivo do sinal de entrada, o diodo estará diretamente polarizado e o capacitor C se carregará rapidamente até atingir o valor de pico do sinal de entrada. Quando o sinal de entrada diminui abaixo deste valor, o diodo ficará reversamente polarizado e o capacitor C começará a descarregar lentamente através do resistor de carga R_L . O processo de descarga continuará até o próximo semi-ciclo positivo. Quando o sinal de entrada se tornar maior do que a tensão no capacitor, o diodo conduzirá novamente e o processo se repetirá. Assumimos, aqui, que o diodo é ideal, apresentando uma resistência r_f ao fluxo de corrente na região e polarização direta e resistência infinita na região de polarização reversa. Assumimos, ainda, que a onda AM aplicada ao detector de envelope é fornecida por uma fonte de tensão com resistência interna R_s . A constante de tempo de carregamento $(r_f + R_s)C$ deve ser pequena se comparada com o período da portadora $1/f_c$, ou seja,

$$(r_f + R_s)C \ll \frac{1}{f_c}$$

² No prefácio, mostramos que a abordagem adotada neste livro é da perspectiva de sistemas. Na descrição em detalhes do detector de envelope deixamos claro que estamos fazendo uma exceção a esta abordagem. A razão para isso está no reconhecimento do fato de que o detector de envelope, devido a sua simplicidade, é utilizado em quase todo receptor AM comercial. De fato, a simplicidade da construção de transmissores e receptores AM é um atrativo comercial tão forte que, mesmo com o crescente domínio das comunicações digitais, a modulação em amplitude continuará a encontrar espaço prático de uma forma ou outra.

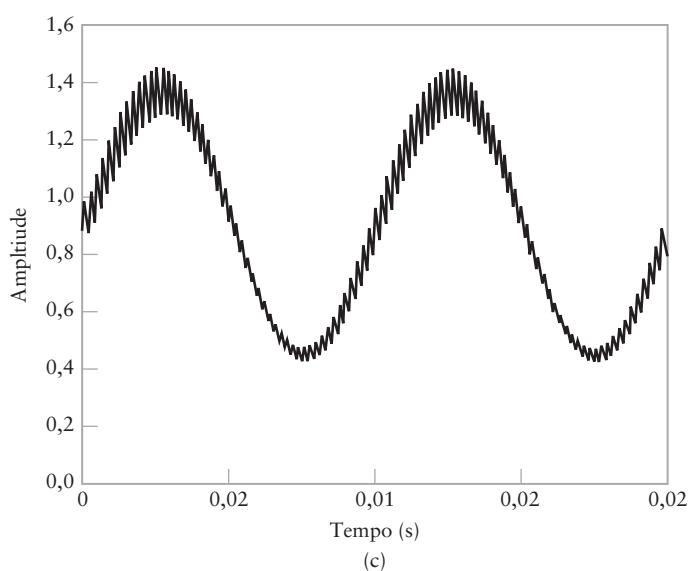
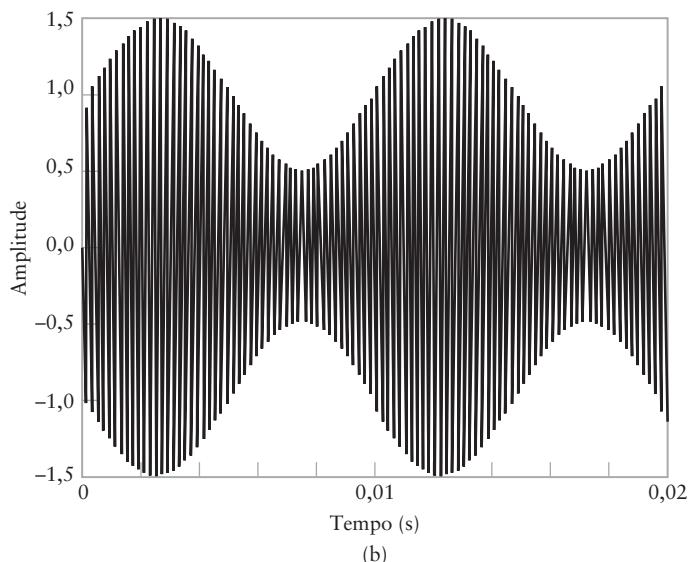
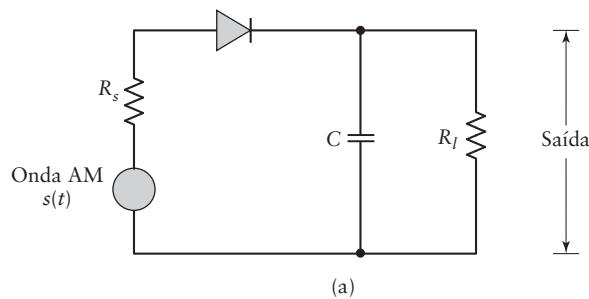


Figura 3.9 Detector de envelope. (a) Diagrama de circuito. (b) Onda AM de entrada. (c) Saída do detector de envelope.

de tal forma que o capacitor C se carrega rapidamente e, portanto, segue a tensão aplicada até o pico positivo quando o diodo estiver conduzindo. Por outro lado, a constante de tempo de descarga R_lC deve ser grande o suficiente para garantir que o capacitor se descarregue lentamente através do resistor de carga R_l entre os picos positivos da onda portadora, mas não tão grande a ponto da tensão do capacitor não diminuir à máxima taxa de variação da onda modulante, ou seja,

$$\frac{1}{f_c} \ll R_l C \ll \frac{1}{W}$$

na qual W é a largura de faixa da mensagem. O resultado é que a tensão do capacitor, ou a saída do detector, é aproximadamente a mesma do envelope da onda AM, como demonstrado a seguir.

■ Experimento de computador: detecção de envelope para AM senoidal

Considere a forma de onda AM senoidal da Fig. 3.9(b), assumindo 50% de modulação. A saída do detector de envelope é mostrada na Fig. 3.9(c). Esta última forma de onda é calculada assumindo que o diodo é ideal, possuindo uma resistência constante r_f quando polarizado diretamente e resistência infinita quando reversamente polarizado. Os valores numéricos utilizados no cálculo da Fig. 3.9(c) são mostrados a seguir:

Resistência da fonte	$R_s = 75\Omega$
Resistência direta	$r_f = 25\Omega$
Resistência de carga	$R_l = 10 k\Omega$
Capacitância	$C = 0,01\mu F$
Largura de faixa da mensagem	$W = 1 kHz$
Freqüência da portadora	$f_c = 20 kHz$

Observe que a saída do detector de envelope inclui um ripple de alta freqüência. Este ripple pode ser removido utilizando um filtro passa-baixa (não mostrado na Fig. 3.9(a)).

3.2 VIRTUDES, LIMITAÇÕES E MODIFICAÇÕES NA MODULAÇÃO EM AMPLITUDE

A modulação em amplitude é o método mais antigo de modulação. Como já ressaltado na Seção 3.1, sua grande virtude é a facilidade pela qual ela é gerada e revertida. O resultado final é que o sistema de modulação em amplitude é relativamente barato de ser construído.

Entretanto, do Capítulo 1, lembramos que a potência transmitida e a largura de faixa do canal são nossos dois recursos primários de comunicação e devem ser utilizados eficientemente. Neste contexto, observamos que a modulação em amplitude definida na Eq. (3.2) sofre de duas grandes limitações práticas:

1. *A modulação em amplitude desperdiça potência transmitida.* A onda portadora $c(t)$ é completamente independente do sinal $m(t)$ contendo a informação. A transmissão da onda portadora, portanto, representa um desperdício de potência, o que significa que na modulação em amplitude somente uma fração da potência total transmitida é realmente afetada por $m(t)$.
2. *A modulação em amplitude desperdiça a largura de faixa do canal.* As faixas laterais superior e inferior da onda AM são unicamente relacionadas uma com a outra em virtude de sua simetria com relação a freqüência da portadora. Logo, dado o espectro de amplitude e fase de uma das faixas, podemos determinar, unicamente, a outra. Isto significa que, enquanto a transmissão de informação for o foco principal,

somente uma faixa lateral é necessária e o canal de comunicação, portanto, precisa fornecer somente largura de faixa igual a do sinal de mensagem. A luz dessa observação, a modulação em amplitude desperdiça a largura de faixa do canal, pois ela requer uma largura de faixa de transmissão igual a duas vezes a largura de faixa da mensagem.

Para superar estas duas limitações do AM, devemos fazer certas modificações que resultam no aumento da complexidade do sistema do processo de modulação em amplitude. De fato, trocamos a complexidade do sistema pela melhor utilização dos recursos da comunicação. Começando com a modulação em amplitude, podemos distinguir três modificações na modulação:

1. *Modulação de faixa lateral dupla – portadora suprimida (DSB-SC*)*, na qual a onda transmitida é constituída somente das faixas laterais superior e inferior. A potência transmitida é salva através da supressão da onda portadora, mas o requisito da largura de faixa do canal é o mesmo do anterior (isto é, duas vezes a largura de faixa da mensagem).
2. *Modulação de faixa lateral única (SSB**)*, na qual a onda modulada é constituída somente da faixa lateral superior ou da faixa lateral inferior. A função essencial da modulação SSB é, portanto, transladar o espectro do sinal modulante (com ou sem inversão) para uma nova posição no domínio da freqüência. A modulação de faixa lateral única é particularmente adequada para a transmissão de sinais de voz em virtude do *gap de energia* que existe no espectro de sinais de voz entre zero e algumas centenas de hertz. A SSB é a forma ótima da modulação de onda contínua, no sentido de que ela requer o mínimo de potência transmitida e o mínimo de largura de faixa do canal. Sua principal desvantagem é o aumento na complexidade e aplicabilidade limitada.
3. *Modulação de faixa lateral vestigial (VSB***)*, na qual uma faixa lateral é transmitida quase que completamente, juntamente com somente um traço, ou um vestígio, da outra. A largura de faixa do canal é, portanto, maior do que a largura de faixa da mensagem por um total igual a largura da faixa lateral vestigial. Esta forma de modulação é bastante adequada para a transmissão de sinais de faixa larga, tais como sinais de televisão que contém componentes significantes em freqüências extremamente pequenas. Na transmissão de televisão aberta, uma portadora de tamanho ajustável é transmitida juntamente com a onda modulada, possibilitando a demodulação do sinal de entrada modulado por um detector de envelope no receptor, o que simplifica o projeto do receptor.

Na Seção 3.3 iremos discutir a modulação DSB-SC, seguida pela discussão das formas de modulação SSB e VSB nas seções subsequentes, nesta ordem.

3.3 MODULAÇÃO DE FAIXA LATERAL DUPLA – PORTADORA SUPRIMIDA

■ Teoria

Basicamente, a *modulação de faixa lateral dupla-portadora suprimida*, consiste do produto do sinal de mensagem $m(t)$ pela onda portadora $c(t)$ como mostrado pela equação

$$\begin{aligned} s(t) &= c(t)m(t) \\ &= A_c \cos(2\pi f_c t)m(t) \end{aligned} \tag{3.8}$$

* N. de T.: Do inglês, *double sideband-suppressed carrier*.

** N. de T.: Do inglês, *single sideband*.

*** N. de T.: Do inglês, *vestigial sideband*.

Desta forma, o dispositivo utilizado para gerar a onda DSB-SC é chamado de *modulador de produto*. Da Eq. (3.8) também notamos que, ao contrário da modulação em amplitude, a modulação DSB-SC é reduzida para zero quando o sinal de mensagem $m(t)$ é desligado.

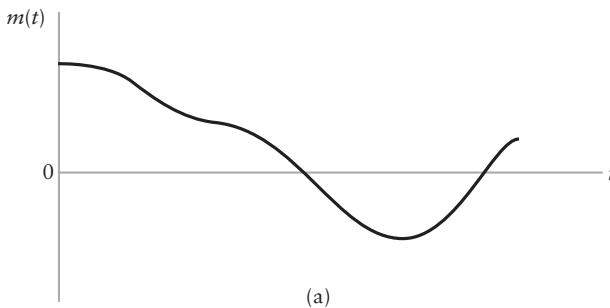
O mais importante, entretanto, está no fato do sinal modulado $s(t)$ sofrer uma *inversão de fase* sempre que o sinal de mensagem $m(t)$ cruzar o zero, como indicado na Fig. 3.10(b) para o sinal de mensagem $m(t)$ mostrado na parte (a) desta figura. O envelope de um sinal modulado DSB-SC é, portanto, diferente do sinal de mensagem, o que significa que a demodulação simples usando um detector de envelope não é uma opção viável para a modulação DSB-SC.

Da Eq. (3.8), a transformada de Fourier de $s(t)$ é obtida como

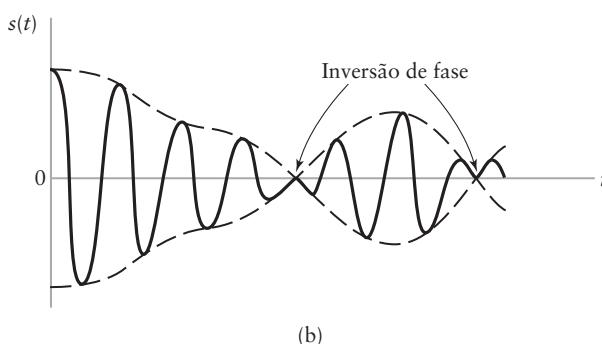
$$S(f) = \frac{1}{2}A_c[M(f - f_c) + M(f + f_c)] \quad (3.9)$$

na qual $m(t) \iff M(f)$. Quando o sinal de mensagem $m(t)$ é limitado ao intervalo $-W \leq f \leq W$, como mostrado na Fig. 3.11(a), verificamos que o espectro $S(f)$ da onda DSB-SC $s(t)$ é como mostrado na Fig. 3.11(b). Exceto pela mudança no fator de escala, o processo de modulação simplesmente translada o espectro do sinal de mensagem por f_c para a direita e por $-f_c$ para a esquerda. Obviamente, a largura de faixa de transmissão necessária pela modulação DSB-SC é a mesma da modulação em amplitude, ou seja, $2W$.

Resumindo, do ponto de vista da ocupação da largura de faixa, a DSB-SC não oferece vantagem frente ao AM. Sua única vantagem está no fato de salvar potência transmitida, o que já é suficientemente importante quando a potência transmitida estiver em prêmio.



(a)



(b)

Figura 3.10 (a) Sinal de mensagem $m(t)$. (b) Onda modulada DSB-SC $s(t)$.

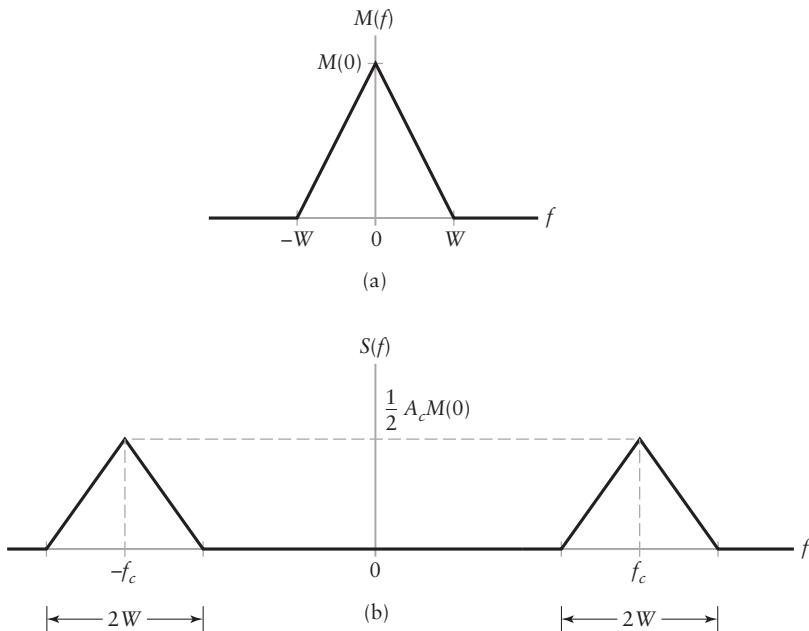


Figura 3.11 (a) Espectro do sinal de mensagem $m(t)$. (b) Espectro da onda modulada DSB-SC $s(t)$.

EXEMPLO 3.2 Espectro do DSB-SC senoidal

Considere a modulação DSB-SC utilizando uma onda modulante senoidal com amplitude A_m e freqüência f_m e operando com uma amplitude de portadora A_c e freqüência f_c . O espectro da mensagem é

$$M(f) = \frac{1}{2}A_m\delta(f - f_m) + \frac{1}{2}A_m\delta(f + f_m)$$

Invocando a Eq. (3.9), o espectro deslocado $\frac{1}{2}A_cM(f - f_c)$ define as duas freqüências laterais para freqüências positivas:

$$\frac{1}{4}A_cA_m\delta(f - (f_c + f_m)); \quad \frac{1}{4}A_cA_m\delta(f - (f_c - f_m))$$

O outro espectro deslocado da Eq. (3.9) $-\frac{1}{2}A_cM(f + f_c)$, define as duas freqüências laterais restantes para freqüências negativas:

$$\frac{1}{4}A_cA_m\delta(f + (f_c - f_m)); \quad \frac{1}{4}A_cA_m\delta(f + (f_c + f_m))$$

as quais são imagens das duas primeiras freqüências laterais com relação a origem, na ordem inversa.

- **PROBLEMA 3.5** Para a modulação DSB-SC senoidal considerada no Exemplo 3.2, qual é a potência média das freqüências laterais inferior e superior, expressas como percentual da potência média da onda modulada DSB-SC? ◀

■ Detecção coerente

Como o envelope da onda modulada DSB-SC $s(t)$ é diferente do sinal de mensagem $m(t)$, precisamos determinar outra forma de recuperar $m(t)$ de $s(t)$. Para isto, observamos que $\cos^2(2\pi f_c t)$ contém um tempo constante, como mostrado pela identidade trigonométrica

$$\cos^2(\theta) = \frac{1}{2} + \frac{1}{2} \cos(2\theta)$$

Considerando esta relação escrita para $\theta = 2\pi f_c t$, vemos da Eq. (3.8) que a recuperação do sinal de mensagem $m(t)$ pode ser feita multiplicando, primeiro, $s(t)$ por uma onda senoidal gerada localmente e, então, filtrando o produto com um filtro passa-baixa. Assume-se que o sinal oscilador local é exatamente coerente ou sincronizado, tanto em freqüência quanto em fase, com a onda portadora $c(t)$ utilizada no modulador de produto para gerar $s(t)$. Este método de demodulação é chamado de *detecção coerente* ou *demodulação síncrona*.

É interessante obter a detecção coerente como um caso especial de um processo de demodulação mais geral utilizando um sinal oscilador local com mesma freqüência, mas com diferença de fase ϕ arbitrária, medida com relação a onda portadora $c(t)$. Portanto, representando o sinal oscilador local por $A'_c \cos(2\pi f_c t + \phi)$ e usando a Eq. (3.8) para a onda DSB-SC $s(t)$, temos que a saída do modulador de produto da Fig. 3.12 é

$$\begin{aligned} v(t) &= A'_c \cos(2\pi f_c t + \phi) s(t) \\ &= A_c A'_c \cos(2\pi f_c t) \cos(2\pi f_c t + \phi) m(t) \\ &= \frac{1}{2} A_c A'_c \cos(4\pi f_c t + \phi) m(t) + \frac{1}{2} A_c A'_c \cos(\phi) m(t) \end{aligned} \quad (3.10)$$

na qual utilizamos a identidade trigonométrica

$$\cos(\theta_1) \cos(\theta_2) = \frac{1}{2} \cos(\theta_1 + \theta_2) + \frac{1}{2} \cos(\theta_1 - \theta_2)$$

O primeiro termo da Eq. (3.10) representa um novo sinal modulado DSB-SC com freqüência de portador $2f_c$, enquanto que o segundo termo é proporcional ao sinal de mensagem $m(t)$. Estes termos são melhor ilustrados no espectro $V(f)$ mostrado na Fig. 3.13, no qual assume-se que o sinal de mensagem é limitado ao intervalo $-W \leq f \leq W$. É, portanto, aparente que o primeiro termo da Eq. (3.10) é removido pelo filtro passa-baixa da Fig. 3.12, desde que a freqüência de corte deste filtro seja maior do que W mas menor do que $2f_c - W$. Isto é satisfeito escolhendo-se $f_c > W$. Na saída do filtro obtemos, então, um sinal dado por

$$v_o(t) = \frac{1}{2} A_c A'_c \cos(\phi) m(t)$$

O sinal demodulado $v_o(t)$ é, portanto, proporcional a $m(t)$ quando o erro de fase ϕ é constante. A amplitude do sinal demodulado é máxima quando $\phi = 0$ e mínimo (zero) quando $\phi = \pm\pi/2$. O sinal demodulado zero, o qual ocorre para $\phi = \pm\pi/2$, representa um *efeito nulo de quadratura*, o que é uma propriedade inherente da detecção coerente. Portanto, o erro de fase ϕ no oscilador local faz com que a saída do detector seja atenuada por um fator igual a $\cos \phi$. Enquanto o erro de fase ϕ for constante, a saída do detector fornece uma versão não distorcida do sinal de mensagem $m(t)$. Na prática, entretanto, geralmente temos o erro de fase ϕ variando aleatoriamente com o tempo, devido a variações aleatórias no canal de comunicação. O resultado é que, na saída do detector, o fator de multiplicação $\cos \phi$ irá

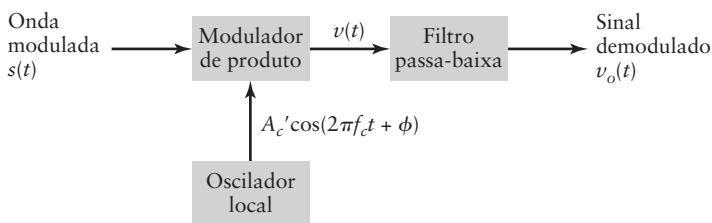


Figura 3.12 Diagrama em blocos do detector coerente, assumindo que o oscilador local está fora de fase por ϕ com respeito ao oscilador senoidal da portadora no transmissor.

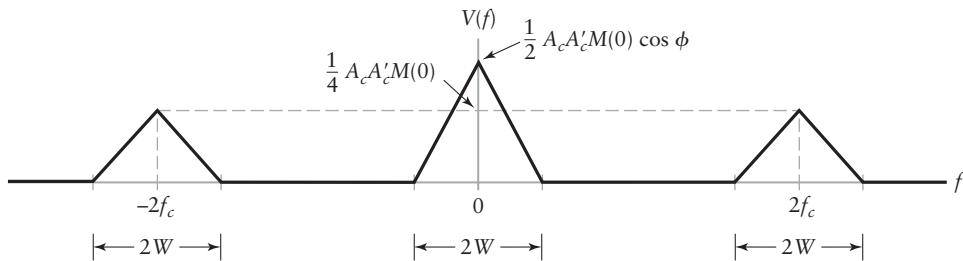


Figura 3.13 Ilustração do espectro da saída $v(t)$ do modulador de produto no detector coerente da Fig. 3.12, o qual é produzido em resposta a uma onda modulada DSB-SC na entrada do detector.

variar, também, aleatoriamente com o tempo, o que, obviamente, é indesejado. Portanto, deve ser feito algo no sistema para manter o oscilador local no receptor em *sincronismo*, tanto em freqüência quanto em fase, com a onda portadora utilizada para gerar o sinal modulado DSB-SC no transmissor. A complexidade resultante do sistema é o preço que deve ser pago pela supressão da onda portadora para salvar a potência transmitida.

- **PROBLEMA 3.6** A onda DSB-SC senoidalmente modulada do Exemplo 3.2 é aplicada ao modulador de produto usando uma senóide gerada localmente com amplitude unitária e em sincronismo com a portadora utilizada no modulador.
 - (a) Determine a saída do modulador de produto, representada por $v(t)$.
 - (b) Identifique os dois termos senoidais em $v(t)$ produzidos pela onda modulada DSB-SC para freqüências positivas e os dois termos senoidais restantes produzidos pela onda modulada DSB-SC para freqüências negativas.
- **PROBLEMA 3.7** O detector coerente para a demodulação de DSB-SC falha se houver sobreposição espectral no modulador. Explique a razão para esta falha.

■ Experimento de computador: DSB-SC

Para o estudo experimental da modulação DSB-SC, utilizaremos o mesmo esquema descrito na Seção 3.1, exceto pelas alterações trazidas pelo uso de DSB-SC no lugar do AM. Os resultados do experimento são descritos considerando dois pontos:

1. A Fig. 3.14(a) mostra 10 ciclos da onda DSB-SC produzida pela modulação de onda senoidal com freqüência de 0,05 Hz. Como esperado, o envelope do sinal modulado não possui relação clara com a onda senoidal modulante. Desta forma devemos utilizar a detecção coerente para a demodulação, a qual será discutida no ponto 2.

A Fig. 3.14(b) mostra o espectro de amplitude da onda modulada. Uma ampliação do espectro ao redor da freqüência da portadora (com freqüência 0,4 Hz) é mostrada na Fig. 3.14(c). Estas duas figuras mostram, claramente, que a portadora é, de fato, suprimida e que as duas faixas laterais superior e inferior são localizadas exatamente onde elas devem estar – ou seja, em $\pm 0,45$ Hz e $\pm 0,35$ Hz, respectivamente.

2. Para executar a detecção coerente, procedemos em dois estágios: (i) multiplicamos a onda modulada DSB-SC por uma réplica exata da portadora e (ii) passamos o produto em um filtro passa-baixa, como descrito na detecção coerente nesta seção. Com dois estágios necessários para o processo de detecção coerente, os resultados desta parte são mostrados como descrito a seguir:

- (i) A Fig. 3.15(a) mostra a forma de onda da saída do modulador de produto do detector coerente. O espectro de amplitude desta forma de onda é mostrado na Fig. 3.15(b), a qual mostra que a forma de onda é constituída pelas seguintes componentes:

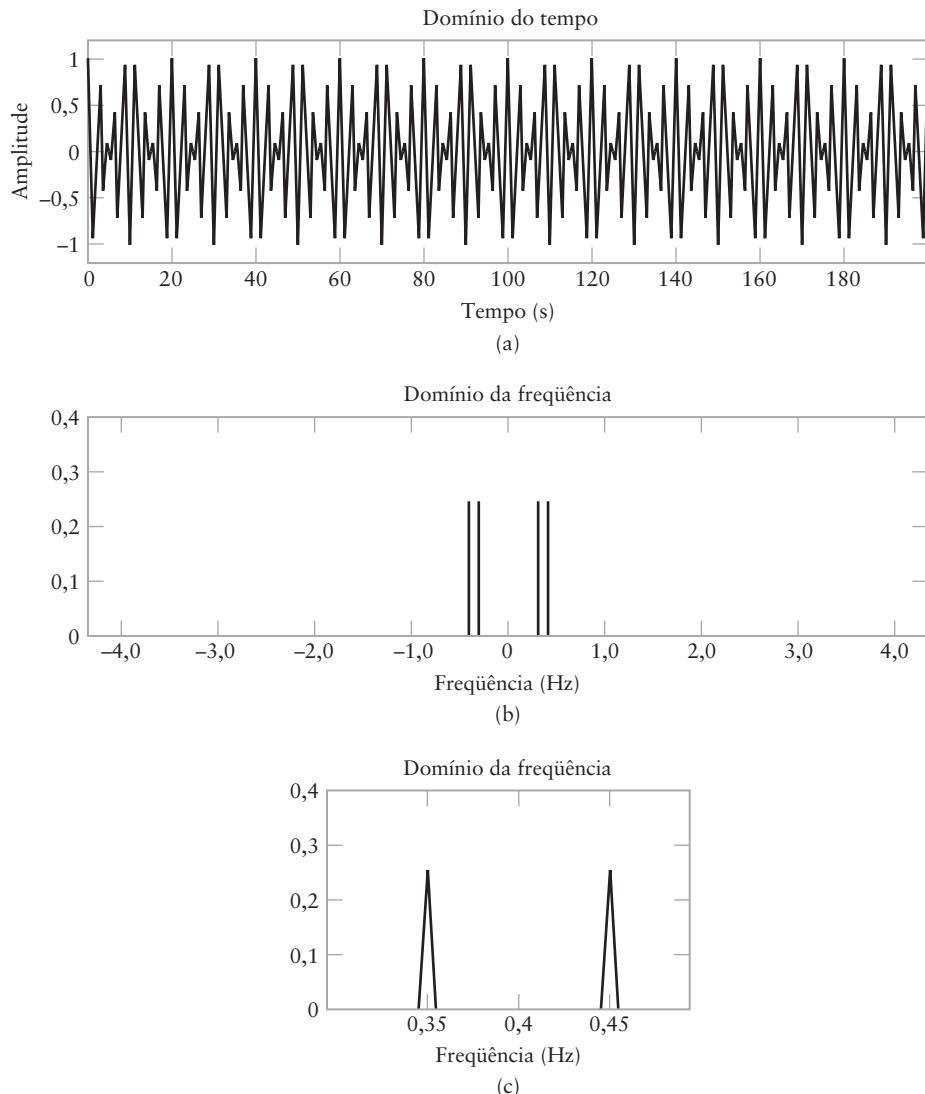


Figura 3.14 Modulação DSB-SC: (a) onda modulada DSB-SC, (b) espectro de amplitude da onda modulada e (c) ampliação do espectro ao redor da freqüência da portadora.

- ▶ Uma componente senoidal com freqüência 0,05 Hz, representando a onda senoidal modulante.
- ▶ Uma nova onda modulada DSB-SC com o dobro da freqüência da portadora, 0,8Hz, na verdade, as duas freqüências laterais desta onda modulada estão localizadas em 0,75 e 0,85 Hz, exatamente onde elas deveriam estar.
- (ii) A Fig. 3.15(c) mostra a forma de onda da saída final do detector coerente, a qual é o resultado da filtragem da saída do modulador de produto por um filtro passa-baixa. Exceto pelos efeitos transientes apresentados inicialmente pelo processo de detecção, a forma de onda é reconhecida como sendo a forma de onda senoidal modulante de freqüência 0,05 Hz. Este resultado é confirmado pelo espectro de amplitude mostrado na Fig. 3.15(d). A base na qual a componente de freqüência de 0,05 Hz está é devido aos efeitos transientes já mencionados.

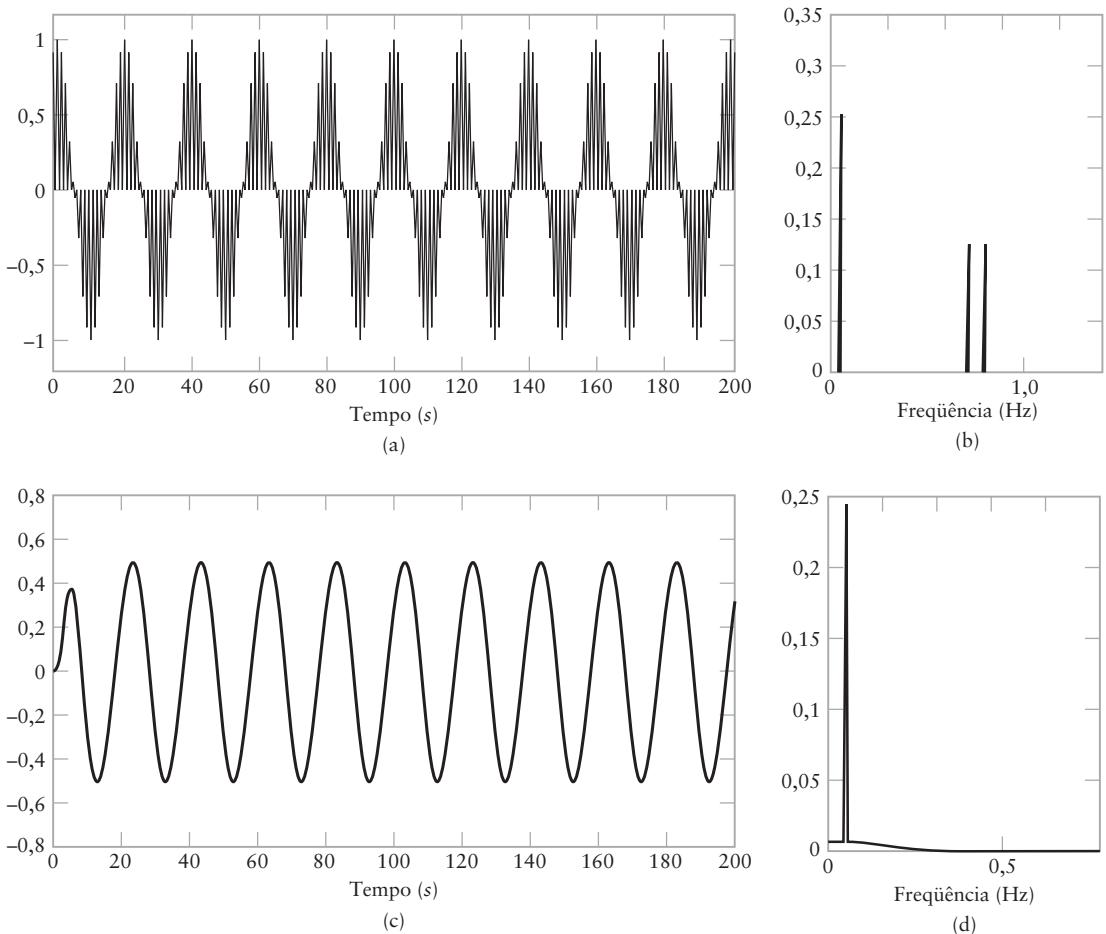


Figura 3.15 Detecção coerente da onda modulada DSB-SC: (a) Forma de onda do sinal produzido na saída do modulador de produto, (b) espectro de amplitude do sinal da parte (a), (c) forma de onda do filtro passa-baixa de saída e (d) espectro de amplitude do sinal da parte (c).

3.4 RECEPTOR COSTAS

A detecção coerente de uma onda modulada em DSB-SC requer que a portadora gerada localmente no receptor esteja sincronizada tanto em fase como em frequência com o oscilador responsável pela geração da portadora no transmissor. Essa é uma exigência extremamente importante, uma vez que a portadora é suprimida do sinal transmitido em DSB-SC. Um método de satisfazer esta condição é utilizar o *receptor Costas* mostrado na Fig. 3.16. Este receptor é constituído por dois detectores coerentes com o mesmo sinal de entrada – a onda DSB-SC $A_c \cos(2\pi f_c t)m(t)$, mas com dois sinais osciladores que estão em quadratura de fase entre eles. A frequência do oscilador local é ajustada para ser a mesma frequência f_c da onda portadora, a qual é considerada conhecida *a priori*. Esta consideração é razoável, pois o projetista do sistema possui acesso às especificações detalhadas tanto do transmissor quanto do receptor. O detector da parte superior é chamado de *detector coerente em fase*, ou *canal I*, e o detector da parte inferior é chamado de *detector coerente em quadratura*, ou *detector-Q*. Estes dois detectores são acoplados para formar um sistema de *realimentação negativa* projetado de forma a manter o oscilador local em sincronismo com a portadora.

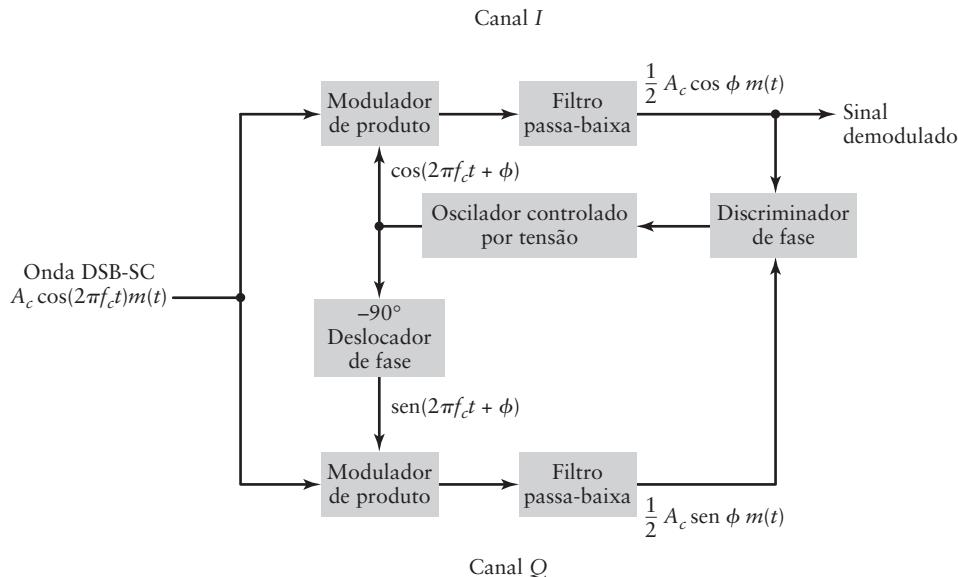


Figura 3.16 Receptor Costas para a demodulação de uma onda modulada DSB-SC.

Para compreender a operação deste receptor, suponha que o sinal do oscilador local possui mesma fase da onda portadora $A_c \cos(2\pi f_c t)$ utilizada para gerar o sinal DSB-SC de entrada. Com estas condições, verificamos que a saída do canal I contém o sinal demodulado $m(t)$ desejado, enquanto que a saída do canal Q é zero, devido ao efeito nulo da quadratura no canal Q . A seguir, suponha que a fase do oscilador local desvie de seu valor correto por um pequeno ângulo ϕ radianos. Da discussão de detecção coerente da Seção 3.3, sabemos que a saída do canal I é proporcional a $\cos \phi$ e $\cos \phi \approx 1$ para ϕ pequeno. Logo, a saída do canal I permanecerá, essencialmente, inalterada desde que ϕ seja pequeno. Mas agora existirá um sinal, mesmo que pequeno, na saída do canal Q , o qual é proporcional a $\sin \phi \approx \phi$ para ϕ pequeno. Esta saída do canal Q terá mesma polaridade da saída do canal I para uma direção do desvio de fase ϕ do oscilador local e polaridade oposta para a direção oposta de ϕ . Portanto, combinando as saídas dos canais I e Q em um *discriminador de fase* (o qual é constituído por um multiplicador seguido por uma unidade de média temporal), um sinal de controle cc, proporcional ao desvio de fase ϕ , é gerado. Com a realimentação negativa atuando no receptor Costas, o sinal de controle tende a corrigir automaticamente o erro de fase ϕ em um *oscilador controlado por tensão*.

Fica, então, aparente que o controle de fase no receptor Costas termina com o fim da modulação, o que significa que o atracamento de fase deve ser reestabelecido com o surgimento da modulação. Isso não é um problema sério, pois o processo de atracamento normalmente ocorre tão rapidamente que nenhuma distorção é percebida.

► **PROBLEMA 3.8** Como mencionado anteriormente, a discriminação de fase no receptor Costas da Fig. 3.16 é constituída por um multiplicador seguido por uma unidade de média temporal. Referindo a esta figura, faça o seguinte:

- Assumindo que o erro de fase ϕ é pequeno, comparado a um radiano, mostre que a saída $g(t)$ do componente multiplicador é aproximadamente $1/4\phi m^2(t)$.
- Além disso, passando $g(t)$ em uma unidade de média temporal definida por

$$\frac{1}{2T} \int_{-T}^T g(t) dt$$

na qual o intervalo de média $2T$ é grande o suficiente quando comparado com o recíproco da largura de faixa de $g(t)$, mostre que a saída do discriminador de fase é proporcional ao erro de

fase ϕ multiplicado pela componente CC (corrente contínua) de $m^2(t)$. A amplitude deste sinal (atuando como sinal de controle aplicado ao oscilador controlado por tensão da Fig. 3.16) irá, portanto, ter sempre o mesmo sinal algébrico que o erro de fase ϕ , como deveria ser. ◀

3.5 MULTIPLEXAÇÃO POR PORTADORA EM QUADRATURA

O efeito nulo de quadratura do detector coerente pode ter uma boa utilidade na construção do chamado *multiplexação por portadora em quadratura* ou *modulação de amplitude em quadratura* (QAM*). Este esquema permite que duas ondas DSB-SC (resultantes da aplicação de dois sinais de mensagem fisicamente *independentes*) ocupem a mesma largura de faixa do canal. Ainda assim ele permite a separação dos dois sinais de mensagem na saída do receptor. A multiplexação por portadora em quadratura é, portanto, um *sistema conservativo de largura de faixa*.

O diagrama em blocos do sistema é mostrado na Fig. 3.17. A parte de transmissão do sistema, mostrada na Fig. 3.17(a) envolve o uso de dois moduladores de produto separados que são supridos com duas ondas portadoras de mesma freqüência mas com uma diferença de fase de -90 graus. O sinal transmitido $s(t)$ consiste da soma da saída destes dois moduladores de produto, mostrado por

$$s(t) = A_c m_1(t) \cos(2\pi f_c t) + A_c m_2(t) \sin(2\pi f_c t) \quad (3.12)$$

na qual $m_1(t)$ e $m_2(t)$ representam os dois sinais de mensagem diferentes aplicados aos moduladores de produto. O sinal multiplexado $s(t)$ ocupa a largura de faixa do canal de $2W$, centrada na freqüência da portadora f_c , na qual W é a largura de faixa da mensagem, considerada comum aos dois sinais $m_1(t)$ e $m_2(t)$. De acordo com a Eq. (3.12), podemos imaginar $A_c m_1(t)$ como a componente em fase do sinal $s(t)$ passa-faixa multiplexado e $-A_c m_1(t)$ como sua componente em quadratura.

A parte do receptor do sistema é mostrada na Fig. 3.17(b). Especificamente, o sinal multiplexado $s(t)$ é aplicado simultaneamente a dois detectores coerentes separados, os quais são supridos por duas portadoras locais de mesma freqüência mas com diferença de fase de -90 graus. A saída do detector superior é $\frac{1}{2} A_c A'_c m_1(t)$, enquanto que a saída do detector inferior é $\frac{1}{2} A_c A'_c m_2(t)$. Para o sistema operar satisfatoriamente, é importante manter as relações de fase e freqüência corretas no oscilador utilizado para gerar as portadoras no transmissor e no oscilador local correspondente utilizado no receptor.

Para manter esta sincronização, podemos utilizar o receptor Costas, descrito na Seção 3.4. Outro método geralmente utilizado é enviar um *sinal piloto* fora da faixa passante do sinal modulado. Neste último método, o sinal piloto normalmente consiste de um tom senoidal de baixa potência cuja freqüência e fase são relacionadas com a portadora $c(t) = A_c(2\pi f_c t)$. No receptor, o sinal piloto é extraído através de um circuito adequadamente sintonizado e, então, transladado para a freqüência correta para ser utilizado no detector coerente.

- **PROBLEMA 3.9** Comprove que as saídas do receptor da Fig. 3.17(b) são as indicadas na figura, assumindo o sincronismo perfeito entre o receptor e o transmissor. ◀

3.6 MODULAÇÃO POR FAIXA LATERAL ÚNICA

Com a supressão da portadora, a modulação DSB-SC soluciona a principal limitação do AM, a qual era o desperdício da potência transmitida. Para resolver a outra principal limitação do AM, relacionada com a largura de faixa do canal, precisamos suprimir uma das duas faixas laterais da onda modulada por DSB-SC. Esta modificação na modulação

* N. de T.: Do inglês, *quadrature-amplitude modulation*.

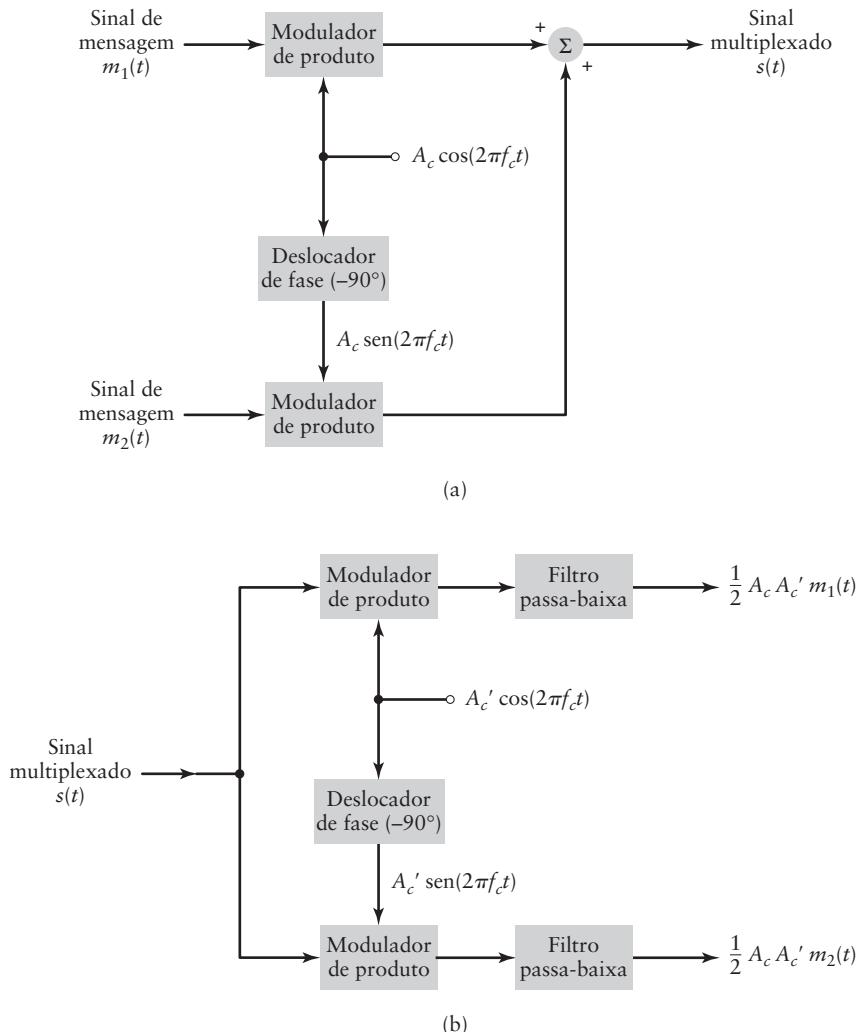


Figura 3.17 Sistema de multiplexação por portadora em quadratura: (a) Transmissor, (b) receptor.

DSB-SC é precisamente o que é feito na *modulação por faixa lateral única* (SSB). De fato, a modulação SSB utiliza apenas a faixa lateral inferior ou a faixa lateral superior para a transmissão de um sinal de mensagem em um canal de comunicação. Dependendo de qual faixa lateral for transmitida, falamos da modulação SSB *inferior* ou SSB *superior*.

■ Teoria

A análise rigorosa da teoria da modulação SSB aplicada a um sinal de mensagem arbitrário é muito difícil e, portanto, além do escopo deste livro. Para simplificar, iremos adotar uma abordagem diferente da utilizada na Seção 3.1 sobre AM e na Seção 3.3 sobre DSB-SC. Especificamente, começaremos a estudar a modulação SSB considerando, primeiro, o caso simples de uma *onda modulante senoidal*, e, então, iremos generalizar os resultados para um sinal modulante arbitrário, passo a passo.

Desta forma, considere um modulador DSB-SC que utiliza a onda modulante senoidal

$$m(t) = A_m \cos(2\pi f_m t)$$

Com a portadora $c(t) = A_c \cos(2\pi f_c t)$, a onda modulada DSB-SC é definida por

$$\begin{aligned} S_{\text{DSB}}(t) &= c(t)m(t) \\ &= A_c A_m \cos(2\pi f_c t) \cos(2\pi f_m t) \\ &= \frac{1}{2} A_c A_m \cos[2\pi(f_c + f_m)t] + \frac{1}{2} A_c A_m \cos[2\pi(f_c - f_m)t] \end{aligned} \quad (3.13)$$

a qual é caracterizada por duas freqüências laterais, uma em $f_c + f_m$ e outra em $f_c - f_m$. Suponha que queremos gerar uma onda modulada SSB senoidal que mantenha a faixa lateral superior em $f_c + f_m$. Então, suprimindo o segundo termo da Eq. (3.13), podemos expressar a onda SSB superior por

$$S_{\text{SSB superior}}(t) = \frac{1}{2} A_c A_m \cos[2\pi(f_c + f_m)t] \quad (3.14)$$

O termo em cosseno na Eq. (3.14) inclui a soma de dois ângulos $-2\pi f_c t$ e $2\pi f_m t$. Portanto, expandindo o termo em cosseno da Eq. (3.14) utilizando uma identidade trigonométrica bem conhecida, temos,

$$S_{\text{SSB superior}}(t) = \frac{1}{2} A_c A_m \cos(2\pi f_c t) \cos(2\pi f_m t) - \frac{1}{2} A_c A_m \sin(2\pi f_c t) \sin(2\pi f_m t) \quad (3.15)$$

Se, por outro lado, quisermos manter a freqüência lateral inferior em $f_c - f_m$ da onda modulada DSB-SC da Eq. (3.13), então teremos a onda modulada SSB inferior definida por

$$S_{\text{SSB inferior}}(t) = \frac{1}{2} A_c A_m \cos(2\pi f_c t) \cos(2\pi f_m t) + \frac{1}{2} A_c A_m \sin(2\pi f_c t) \sin(2\pi f_m t) \quad (3.16)$$

Examinando as Eqs. (3.15) e (3.16), vemos que elas diferem uma da outra no seguinte aspecto: o sinal menos da Eq. (3.15) é substituído pelo sinal mais na Eq. (3.16). Desta forma, podemos combinar estas duas equações e, portanto, definir a onda modulada SSB por:

$$S_{\text{SSB}}(t) = \frac{1}{2} A_c A_m \cos(2\pi f_c t) \cos(2\pi f_m t) \mp \frac{1}{2} A_c A_m \sin(2\pi f_c t) \sin(2\pi f_m t) \quad (3.17)$$

na qual o sinal mais se aplica ao SSB inferior e o sinal menos se aplica ao SSB superior.

Com a generalização da Eq. (3.17) como objetivo, podemos seguir em duas etapas. Na etapa 1 fazemos o sinal de mensagem ser periódico e no estágio 2 fazemos o sinal de mensagem ser não periódico. Considere, então, o *sinal de mensagem periódico* definido pela Série de Fourier

$$m(t) = \sum_n a_n \cos(2\pi f_n t) \quad (3.18)$$

o qual é constituído de uma mistura de ondas senoidais com freqüências harmonicamente relacionadas. Percebendo que a portadora $c(t)$ é comum a todas as componentes senoidais de $m(t)$, podemos inferir imediatamente da Eq. (3.17) a expressão

$$S_{\text{SSB}}(t) = \frac{1}{2} A_c \cos(2\pi f_c t) \sum_n a_n \cos(2\pi f_n t) \mp \frac{1}{2} A_c \sin(2\pi f_c t) \sum_n a_n \sin(2\pi f_n t) \quad (3.19)$$

como fórmula correspondente para a onda modulada SSB.

A seguir, vamos considerar outro sinal periódico definido pela série de Fourier

$$\hat{m}(t) = \sum_n a_n \sin(2\pi f_n t) \quad (3.20)$$

a qual é uma forma similar da Eq. (3.18), exceto pelo fato do termo $\cos(2\pi f_c t)$ ser substituído pelo termo $\sin(2\pi f_c t)$. Então, utilizando as definições das Eqs. (3.19) e (3.20), podemos reformular a onda modulada SSB da Eq. (3.17) por

$$S_{SSB}(t) = \frac{A_c}{2} m(t) \cos(2\pi f_c t) \mp \frac{A_c}{2} \hat{m}(t) \sin(2\pi f_c t) \quad (3.21)$$

Comparando a Eq. (3.20) com a Eq. (3.18), observamos que o sinal periódico $\hat{m}(t)$ pode ser obtido do sinal periódico modulante $m(t)$ simplesmente pelo deslocamento em fase de cada termo cosseno da Eq. (3.18) por -90° .

Em termos técnicos e práticos, a observação que acabamos de fazer é muito importante por duas razões:

1. Sabemos da análise de Fourier que *sob condições apropriadas, a representação por série de Fourier de um sinal periódico converge para a transformada de Fourier de um sinal não periódico.* Veja o Apêndice 2 para detalhes.
2. O sinal $\hat{m}(t)$ é a transformada de Hilbert do sinal $m(t)$. Basicamente, a transformada de Hilbert é um sistema cuja função de transferência é definida por

$$H(f) = -j \operatorname{sgn}(f) \quad (3.22)$$

Na qual $\operatorname{sgn}(f)$ é a função sinal. Para a definição da função sinal, veja a Seção 2.4. Em palavras, a transformação de Hilbert é um deslocador de fase de faixa larga, cuja resposta em freqüência é caracterizada por duas partes como mostrado a seguir (veja o Problema 2.52):

- A resposta em amplitude é unitária para todas as freqüências, tanto positivas quanto negativas.
- A resposta em fase é $+90^\circ$ para freqüências negativas e -90° para freqüências positivas.

Equipado analiticamente na forma descrita nos pontos 1 e 2, podemos finalmente generalizar a Eq. (3.21) para a fórmula de uma onda modulada por faixa lateral única produzida por um sinal de mensagem, independente se ele for periódico ou não. Especificamente, dado o sinal de mensagem $m(t)$ possível de ser transformado por Fourier, com sua transformada de Hilbert definida por $\hat{m}(t)$, a onda modulada SSB produzida por $m(t)$ é definida por

$$s(t) = \frac{A_c}{2} m(t) \cos(2\pi f_c t) \mp \frac{A_c}{2} \hat{m}(t) \sin(2\pi f_c t) \quad (3.23)$$

na qual $A_c \cos(2\pi f_c t)$ é a portadora, $A_c \sin(2\pi f_c t)$ é sua versão deslocada em fase por -90° , o sinal mais e menos são aplicados para SSB inferior e superior, respectivamente. Na Eq. (3.23), omitimos o uso do subscrito SSB em $s(t)$ por ficar implícito que esta equação se refere à modulação SSB em sua forma mais genérica.

- **PROBLEMA 3.10** Usando as Eqs. (3.22) e (3.23), mostre que, para freqüência positivas, os espectros dos dois tipos de ondas moduladas SSB são definidos por:

(a) Para SSB superior,

$$S(f) = \begin{cases} \frac{A_c}{2} M(f - f_c), & \text{para } f \geq f_c \\ 0, & \text{para } 0 < f \leq f_c \end{cases} \quad (3.24)$$

(b) Para SSB inferior

$$S(f) = \begin{cases} 0, & \text{para } f > f_c \\ \frac{A_c}{2} M(f - f_c), & \text{para } 0 < f \leq f_c \end{cases} \quad (3.25)$$

- **PROBLEMA 3.11** Mostre que se o sinal de mensagem $m(t)$ for passa-baixa, então a transformada de Hilbert $\hat{m}(t)$ também será passa-baixa com a mesma largura de faixa de $m(t)$. ◀

As duas fórmulas de espectro definidas nas partes (a) e (b) do Problema 3.10 são intuitivamente satisfatórias. Em particular, as duas estão de acordo com a figura mostrada nas partes (b) e (c) da Fig. 3.18, respectivamente. A Fig. 3.18(b) descreve a onda modulada SSB que manteve a faixa lateral superior, enquanto que a Fig. 3.18(c) descreve o outro tipo de modulação SSB, o qual manteve a faixa lateral inferior. Do ponto de vista prático, a única questão que distingue um tipo de modulação SSB do outro é a ocupação da largura de faixa.

■ Moduladores para SSB

Utilizando a teoria apresentada nesta seção, iremos desenvolver dois métodos para a geração de ondas moduladas SSB, descritas a seguir.

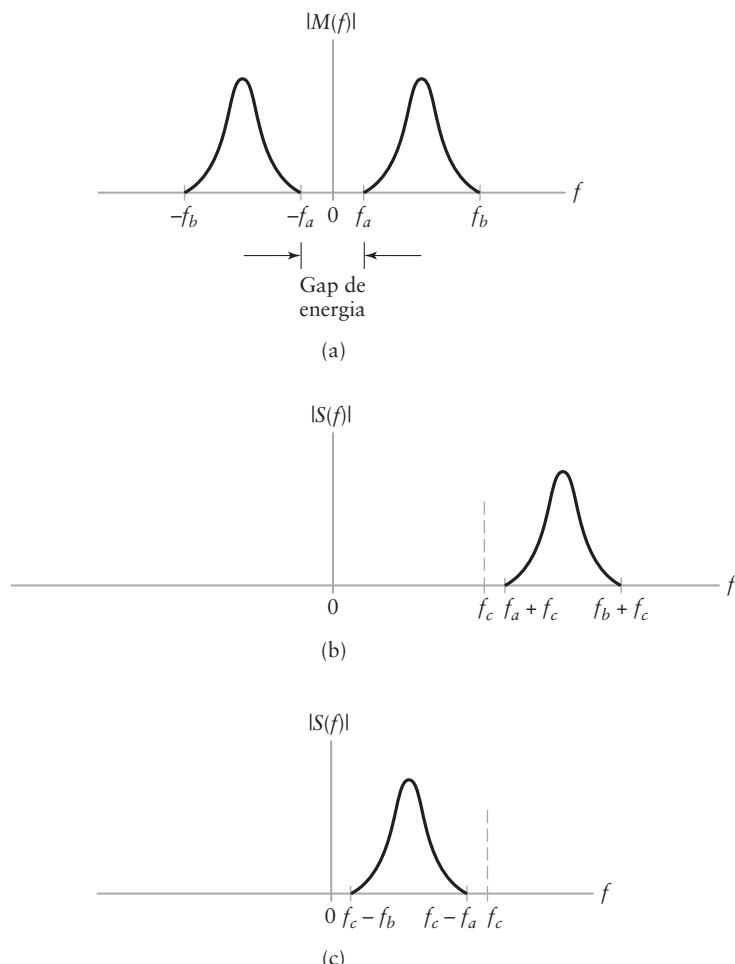


Figura 3.18 (a) Espectro do sinal de mensagem $m(t)$ com gap de energia centrado na freqüência zero. Espectro correspondente da onda modulada SSB usando (b) faixa lateral superior e (c) faixa lateral inferior. Nas partes (b) e (c), o espectro é mostrado apenas para freqüências positivas.

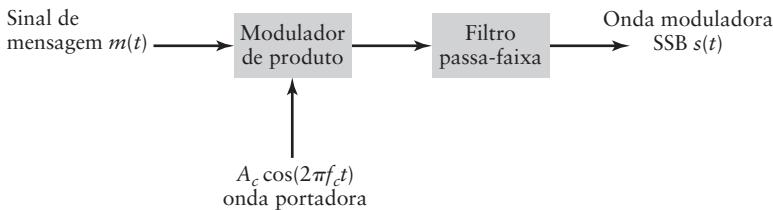


Figura 3.19 Esquema de discriminação em freqüência para geração da onda modulada SSB.

Método da discriminação em freqüência

Um método direto para a geração de SSB, chamado de *método de discriminação em freqüência* é mostrado na Fig. 3.19. Este discriminador é obtido diretamente das Eqs. (3.24) e (3.25) apresentadas no Problema 3.10. O modulador SSB da Fig. 3.19 é constituído por dois componentes: um modulador de produto seguido por um filtro passa-faixa. O modulador de produto produz a onda modulada DSB-SC com as faixas laterais superior e inferior. O filtro passa-faixa é projetado para transmitir uma destas duas faixas laterais, dependendo se a modulação SSB desejada é superior ou inferior. Para o projeto do filtro passa-faixa ser fisicamente realizável, deve existir uma certa separação entre as duas faixas laterais que seja suficientemente grande para acomodar a faixa de transição do filtro passa-faixa. Essa separação é equivalente a $2f_a$, na qual f_a é a menor componente de freqüência do sinal de mensagem, como mostrado na Fig. 3.18. Esta condição limita a aplicabilidade da modulação SSB a sinais de voz para os quais $f_a \approx 100$ Hz, mas exclui sinais de vídeo e dados de computadores, cujo conteúdo espectral chega quase à freqüência zero.

Método de discriminação em fase

O segundo método para geração de SSB, chamado de *método de discriminação em fase*, é mostrado na Fig. 3.20. Sua implementação é obtida da descrição no domínio do tempo da onda SSB definida na Eq. (3.23). Este segundo modulador SSB é constituído por dois caminhos paralelo, um chamado de *caminho em fase* e outro chamado de *caminho em quadratura*. Cada caminho envolve um modulador de produto. As ondas portadoras senoidais aplicadas aos dois moduladores de produto estão em quadratura de fase, o que é obtido pela simples utilização de um deslocador de fase de -90° como mostrado na Fig. 3.20. Entretanto, o único bloco funcional da Fig. 3.2 que merece atenção especial é o deslocador de fase de faixa larga, o qual é projetado para produzir a transformada de Hilbert $\hat{m}(t)$ em resposta ao sinal de mensagem de entrada $m(t)$. O papel do caminho em quadratura contendo o deslocador de fase de faixa larga é simplesmente interferir com o caminho em fase, de forma a eliminar potência em uma das duas faixas laterais, dependendo de qual modulação SSB, superior ou inferior, é desejada.

Os dois moduladores das Figs. 3.19 e 3.20 são claramente diferentes em estrutura. Em termos de desafio de projeto, o filtro passa-faixa no discriminador em freqüência da Fig. 3.19 permanece como bloco funcional que requer atenção especial. Por outro lado, no discriminador em fase da Fig. 3.20, é o deslocador de fase de faixa larga que requer atenção especial.

■ Detecção coerente de SSB

A demodulação de DSB-SC é complicada pela supressão da portadora no sinal transmitido. Para compensar a ausência da portadora no sinal recebido, o receptor utiliza a *detecção coerente*, a qual requer a sincronização de um oscilador local no receptor com o oscilador responsável pela geração da portadora no transmissor. O requisito de sincronização precisa ser tanto em fase quanto em freqüência. Apesar da portadora ser suprimida, a informação da fase e freqüência da portadora está implícita nas faixas laterais da onda modulada, a

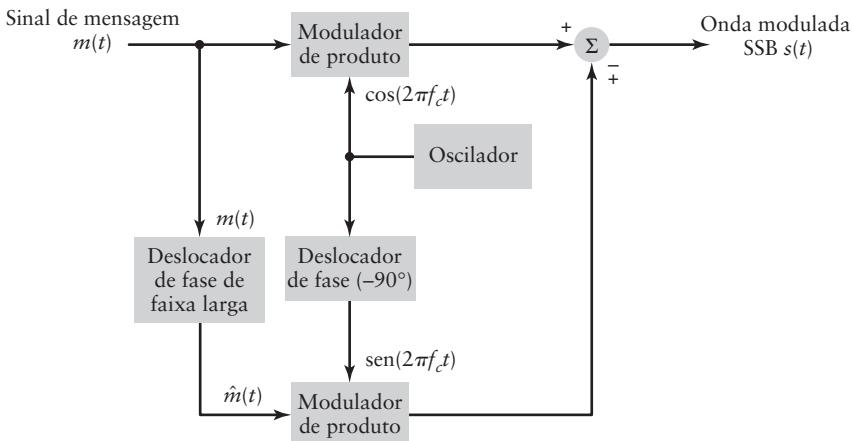


Figura 3.20 Método de discriminação em fase para a geração da onda modulada SSB.
Nota: O sinal mais no bloco somador representa a transmissão da faixa lateral inferior e o sinal menos representa a faixa lateral superior.

qual é explorada no receptor. Entretanto, a demodulação SSB é mais complicada ainda, pela supressão da faixa lateral superior ou inferior. Na realidade, entretanto, as duas faixas laterais compartilham uma importante propriedade: uma é imagem da outra com relação à portadora. Novamente, a detecção coerente aparece para salvar a demodulação SSB.

O detector coerente da Fig. 3.12 funciona bem tanto para a demodulação DSB-SC quanto para SSB. A única diferença entre estas duas aplicações é como a onda modulada $s(t)$ é definida.

- **PROBLEMA 3.12** Para o filtro passa-baixa da Fig. 3.12 (assumindo sincronismo perfeito) suprimir a onda SSB não desejada, a seguinte condição deve ser satisfeita

$$f_c > W, \quad f_c = \text{freqüência da portadora} \text{ e } W = \text{largura de faixa da mensagem}$$

Justifique esta condição. ◀

- **PROBLEMA 3.13** Começando da Eq. (3.23) para a onda modulada SSB, mostre que a saída produzida pelo detector coerente da Fig. 3.12, em resposta a esta onda modulada, é definida por

$$\nu_o(t) = \frac{A_c A'_c}{4} m(t)$$

Assuma erro de fase $\phi = 0$ na Fig. 3.12. ◀

■ Translação em freqüência

A operação básica realizada na modulação por faixa lateral única é, de fato, uma forma de *translação em freqüência*, motivo pelo qual a modulação por faixa lateral única é, algumas vezes, chamada de *mudança de freqüência, mixagem ou batimento*.

A idéia da modulação por faixa lateral única foi, portanto, já apresentada no contexto de um sinal de mensagem puro. Esta idéia pode ser generalizada para englobar a translação em freqüência como mostrado a seguir. Suponha que temos a onda modulada $s_1(t)$ cujo espectro está centrado em uma freqüência de portadora f_1 , cuja necessidade é transladá-lo pra cima ou para baixo na freqüência, de forma que a freqüência da portadora seja alterada de f_1 para um novo valor f_2 . Esta necessidade é satisfeita usando um *mixer*. Como mostrado na Fig. 3.21 um mixer é um bloco funcional constituído por um modulador de produto seguido por um filtro passa-faixa, tal como em um modulador SSB convencional,

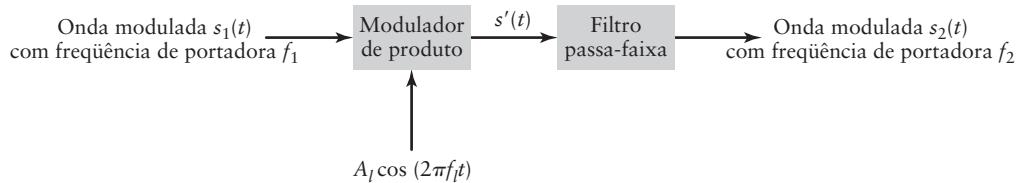


Figura 3.21 Diagrama em blocos do mixer.

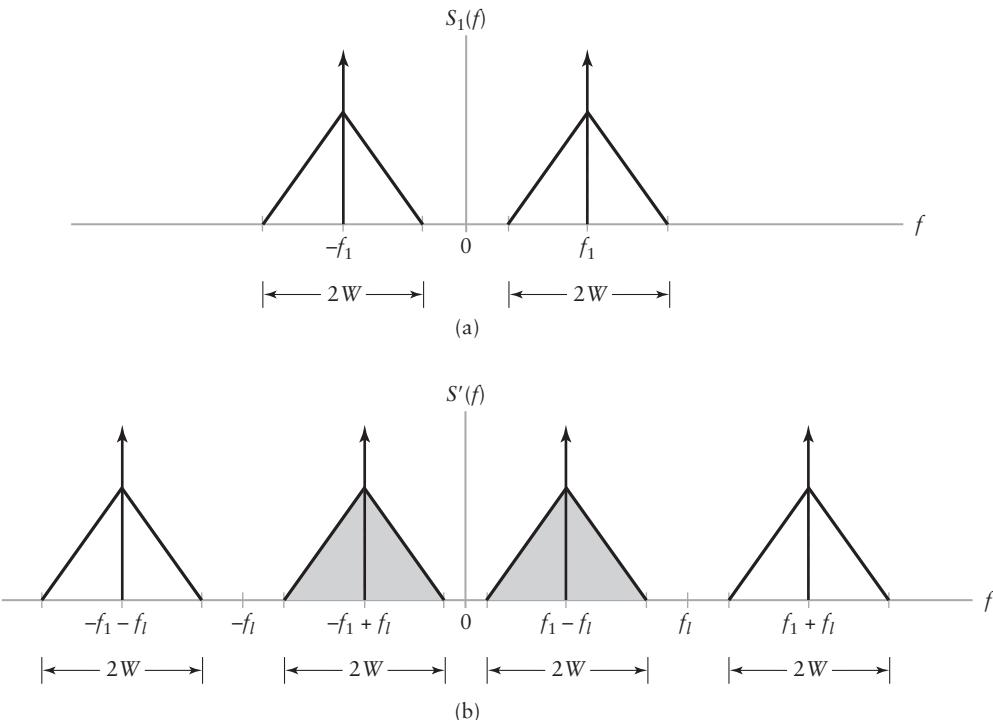
mas com uma importante diferença. O filtro passa-faixa, agora, é fácil de ser projetado, pelo motivo explicado a seguir.

Especificamente, para explicar a ação do mixer, considere a situação espectral mostrada na Fig. 3.22(a), na qual, para efeitos de ilustração, considera-se que a entrada $s_1(t)$ do mixer é uma onda com freqüência de portadora f_1 e largura de faixa $2W$. A Fig. 3.21(b) mostra o espectro $S'(f)$ do sinal resultante $s'(t)$ produzido na saída do modulador de produto da Fig. 3.21.

O sinal $s'(t)$ pode ser visto como a soma de duas componentes moduladas: uma componente representada pelo espectro sombreado da Fig. 3.22(b), e a outra componente representada pelo espectro não sombreado da figura. Dependendo se a freqüência da portadora f_1 deve ser transladada para cima ou para baixo, podemos identificar duas situações diferentes:

- (i) *Conversão para cima*. Nesta forma de mixagem, a freqüência da portadora transladada, representada por f_2 , é maior do que a freqüência da portadora de entrada f_1 . A freqüência f_l do oscilador local necessária é, portanto, definida por

$$f_2 = f_1 + f_l$$

Figura 3.22 (a) Espectro do sinal modulado $s_1(t)$ na entrada do mixer. (b) Espectro do sinal correspondente $s'(t)$ na saída do modulador de produto do mixer.

Resolvendo para f_1 , temos

$$f_l = f_2 - f_1$$

Nesta situação, a parte não sombreada do espectro da Fig. 3.22(b) define o sinal convertido para cima $s_2(t)$, e a parte sombreada deste espectro define a imagem do sinal associada com $s_2(t)$, a qual é removida pelo filtro passa-faixa da Fig. 3.21. Por motivos óbvios, o mixer, neste caso, é chamado de *conversor de freqüência para cima*.

- (ii) *Conversão para baixo.* Nesta segunda forma de mixagem, a freqüência transladada f_2 é menor do que a freqüência da portadora de entrada f_1 , mostrado por

$$f_2 = f_1 - f_l$$

A freqüência do oscilador local necessária é, portanto,

$$f_l = f_1 - f_2$$

A figura que temos neste caso é o inverso da pertencente à conversão para cima. Em particular, a parte sombreada do espectro da Fig. 3.22(b) define o sinal $s_2(t)$ convertido para baixo e a parte não sombreada deste espectro define a imagem associada ao sinal. Desta forma, este segundo mixer é chamado de *conversor de freqüência para baixo*. Note que, neste caso, a freqüência da portadora transladada f_2 precisa ser maior do que W (isto é, metade da largura de faixa do sinal modulado de entrada $s_2(t)$) para evitar a sobreposição de faixas laterais.

O propósito do filtro passa-faixa do mixer da Fig. 3.21 agora é claro: passar o sinal $s_2(t)$ e eliminar a imagem associada ao sinal. Este objetivo é realizado pelo alinhamento da freqüência central do filtro com a freqüência f_2 da portadora transladada e associar sua largura de faixa à do sinal modulado de entrada $s_1(t)$. Independentemente da conversão de freqüência para cima ou para baixo, a faixa de transição do filtro pode ocupar o espaço entre $f_1 - f_l + W$ a $f_1 + f_l + W$, ou seja, a largura possível da faixa de transição é $2(f_l + W)$, o que, de fato, requer que a freqüência do oscilador local f_l seja maior do que W . Além disso, para evitar a sobreposição espectral na conversão para baixo, precisamos, também, que $f_1 - f_l - W$ seja maior do que zero, ou seja, $f_l > f_1 - W$.

É importante notar que a mixagem é uma operação linear. Desta forma, a relação das faixas laterais da onda modulada de entrada com a portadora original existente na entrada do mixer é completamente preservada na saída do mixer.

3.7 MODULAÇÃO POR FAIXA LATERAL VESTIGIAL

■ Motivação

A modulação por faixa lateral única funciona satisfatoriamente bem para um sinal contendo a informação (por exemplo, sinal de voz) com um gap de energia centrado na freqüência zero. Entretanto, para a transmissão espectral eficiente de *sinais de faixa larga*, precisamos de outro método de modulação por dois motivos:

1. Normalmente, o espectro de sinais de faixa larga (exemplificados por sinais de vídeo de televisão e dados de computadores) contém baixas freqüências significantes, o que torna impraticável a modulação SSB.
2. As características espetrais de dados de faixa larga beneficiam o uso de DSB-SC. Entretanto, DSBSC requer uma largura de faixa de transmissão igual a duas vezes a largura de faixa da mensagem, o que viola a condição de conservação de largura de faixa.

Para sobrepujar estas duas limitações práticas, precisamos de um método de *compromisso* de modulação que esteja em algum lugar entre as características espetrais da SSB e

da DSB-SC. A modulação por faixa lateral *vestigial*, o esquema de modulação final a ser considerado nesta seção, é este esquema de compromisso.

A modulação por faixa lateral vestigial (VSB) se distingue da modulação SSB em dois pontos práticos:

1. Em vez de remover totalmente a faixa lateral, um traço ou vestígio da faixa lateral é transmitido. Por isso o nome “faixa lateral vestigial”.
2. Em vez de transmitir a outra faixa lateral completamente, a *quase* totalidade desta segunda faixa também é transmitida.

Desta forma, a largura de faixa de transmissão de um sinal modulado VSB é definida por

$$B_T = f_v + W$$

onde f_v é a largura de faixa vestigial e W é a largura de faixa da mensagem. Tipicamente, f_v é 25% de W , o que significa que a largura de faixa B_T da VSB está entre a largura de faixa W da SSB e a largura de faixa $2W$ da DSB-SC.

■ Filtro formador de faixa lateral

Para produzir a modulação VSB, podemos utilizar o modulador mostrado na Fig. 3.23, o qual é constituído por um modulador de produto seguido por um filtro passa-faixa. Para a modulação VSB, o filtro passa-faixa é chamado de *filtro formador de faixa lateral* (ou filtro vestigial). Assumindo que o vestígio da VSB está na faixa lateral inferior da onda modulada DSB-SC, a forma do espectro é definida pela função de transferência do filtro, representada por $H(f)$. A única condição que a formação da faixa lateral executada por $H(f)$ deve satisfazer é que o vestígio transmitido *compense* a porção espectral que falta na outra faixa lateral. Este requisito garante que a detecção coerente da onda modulada VSB recupere uma réplica do sinal de mensagem, exceto por um fator de escala.

Impondo esta condição ao processo de demodulação VSB, temos que o filtro formador de faixa lateral deve satisfazer a seguinte condição:

$$H(f + f_c) + H(f - f_c) = 1, \quad \text{para } -W \leq f \leq W \quad (3.26)$$

na qual f_c é a freqüência da portadora. O termo $H(f + f_c)$ é a parte positiva de freqüência da função de transferência passa-faixa $H(f)$ deslocada para a esquerda por f_c , e $H(f - f_c)$ é a parte negativa de freqüência de $H(f)$ deslocada para a direita por f_c . A prova da Eq. (3.26) lidando com um sinal de mensagem arbitrário possível de ser transformado de Fourier é apresentada posteriormente nesta seção, na detecção coerente de VSB.

Duas propriedades do filtro formador de faixa lateral seguem da Eq. (3.26):

1. A função de transferência do filtro formador de faixa lateral exibe simetria ímpar com relação à freqüência da portadora f_c . Para explicar esta propriedade, expressamos, primeiro, $H(f)$ como a diferença entre duas funções deslocadas no tempo, como mostrado a seguir:

$$H(f) = u(f - f_c) - H_v(f - f_c), \quad \text{para } f_c - f_v < |f| < f_c + W \quad (3.27)$$

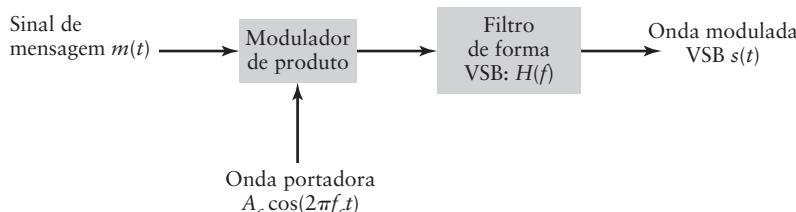


Figura 3.23 Modulador VSB utilizando discriminação em freqüência.

O primeiro termo $u(f - f_c)$ representa a versão deslocada no *tempo do degrau unitário na freqüência* $u(f)$, o qual é mostrado na Fig. 3.24(b). Ou seja,

$$u(f) = \begin{cases} 1, & \text{para } f > 0 \\ 0, & \text{para } f < 0 \end{cases} \quad (3.28)$$

O segundo termo $H_v(f - f_c)$ representa a versão deslocada na freqüência de uma nova *função de transferência passa-baixa*, $H_v(f)$, a qual, como mostrado na Fig. 3.24(c), é completamente determinada pelo vestígio da onda modulada $s(t)$. A relação definida pela Eq. (3.27) é obtida diretamente das três partes do exemplo da Fig. 3.24. O ponto importante da parte (c) da figura é que $H_v(f)$ satisfaz a propriedade de *simetria ímpar* com relação a freqüência zero, mostrado por

$$H_v(-f) = -H_v(f) \quad (3.29)$$

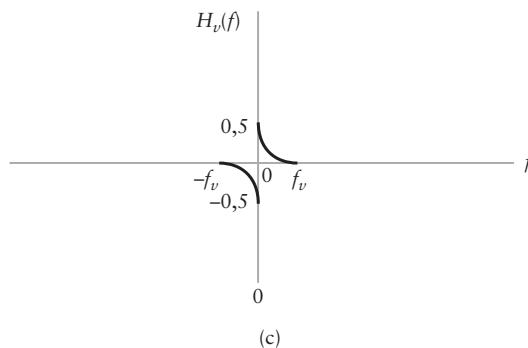
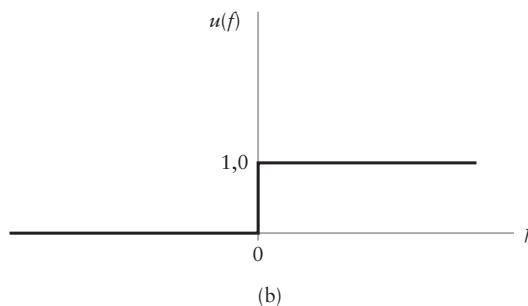
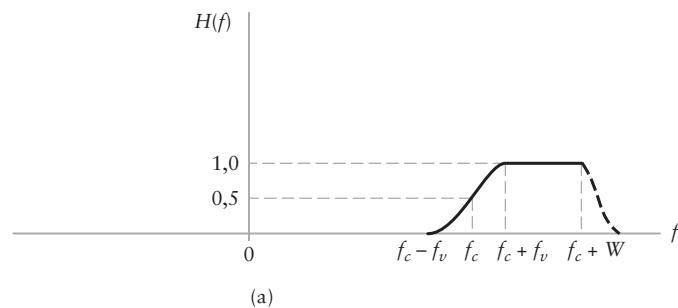


Figura 3.24 (a) Resposta em amplitude do filtro de formação de faixa lateral. Apenas a porção de freqüências positivas é mostrada, a parte tracejada da resposta em amplitude é arbitrária. (b) Função degrau unitário definida no domínio da freqüência. (c) Função de transferência passa-baixa $H_v(f)$.

É, portanto, neste sentido que a propriedade 1 é apresentada.

2. A função de transferência $H_v(f)$ deve satisfazer a condição da Eq. (3.26) somente para o intervalo de freqüência $-W \leq f \leq W$, na qual W é a largura de faixa da mensagem. A implicação prática desta segunda propriedade é que, para o caso do VSB mostrado na Fig. 3.24(a), a função de transferência do filtro formador de faixa lateral pode ter uma especificação arbitrária para $|f| > f_c + w$. É por esta razão que a parte do espectro acima de $f_c + W$ é mostrada pontilhada na Fig. 3.24(a).

EXEMPLO 3.3 VSB senoidal

Considere o exemplo simples de modulação VSB senoidal produzida pela onda modulante senoidal

$$m(t) = A_m \cos(2\pi f_m t)$$

e onda portadora

$$c(t) = A_c \cos(2\pi f_c t)$$

Seja a freqüência lateral superior em $f_c + f_m$ e sua imagem em $-(f_c + f_m)$ atenuada pelo fator k . Para satisfazer as condições da Eq. (3.26), a freqüência lateral inferior em $f_c - f_m$ e sua imagem $-(f_c - f_m)$ devem ser atenuadas pelo fator $(1 - k)$. O espectro VSB é, portanto,

$$\begin{aligned} S(f) = & \frac{1}{4} k A_c A_m [\delta(f - (f_c + f_m)) + \delta(f + (f_c + f_m))] \\ & + \frac{1}{4} (1 - k) A_c A_m [\delta(f - (f_c - f_m)) + \delta(f + (f_c - f_m))] \end{aligned}$$

Desta forma, a onda modulada VSB é definida por

$$\begin{aligned} s(t) = & \frac{1}{4} k A_c A_m [\exp(j2\pi(f_c + f_m)t) + \exp(-j2\pi(f_c + f_m)t)] \\ & + \frac{1}{4} (1 - k) A_c A_m [\exp(j2\pi(f_c - f_m)t) + \exp(-j2\pi(f_c - f_m)t)] \\ = & \frac{1}{2} k A_c A_m \cos(2\pi(f_c + f_m)t) + \frac{1}{2} (1 - k) A_c A_m \cos(2\pi(f_c - f_m)t) \end{aligned} \quad (3.30)$$

Usando identidades trigonométricas bem conhecidas para expandir os termos em cosseno $\cos(2\pi(f_c + f_m)t)$ e $\cos(2\pi(f_c - f_m)t)$, podemos reformular a Eq. (3.30) como a combinação linear de duas ondas senoidais moduladas DSB-SC.

$$\begin{aligned} s(t) = & \frac{1}{2} A_c A_m \cos(2\pi f_c t) \cos(2\pi f_m t) \\ & + \frac{1}{2} A_c A_m (1 - 2k) \sin((2\pi f_c t) \sin(2\pi f_m t)) \end{aligned} \quad (3.31)$$

na qual o primeiro termo do lado direito da equação está em fase com a componente de $s(t)$ e o segundo termo em quadratura.

Resumindo, dependendo de como o fator de atenuação k da Eq. (3.31) é definido no intervalo $(0, 1)$, podemos identificar todas as diferentes formas senoidais das ondas moduladas lineares estudadas nas Seções 3.3, 3.6 e 3.7, como mostrado a seguir:

1. $k = \frac{1}{2}$, para o qual $s(t)$ se reduz para SDSB-SC
2. $k = 0$, para o qual $s(t)$ se reduz para SSB inferior
 $k = 1$, para o qual $s(t)$ se reduz para SSB superior

3. $0 < k < \frac{1}{2}$, para o qual a versão atenuada da freqüência lateral superior define o vestígio de $s(t)$
 $\frac{1}{2} < k < 1$, para o qual a versão atenuada da freqüência lateral inferior define o vestígio de $s(t)$

■ Detecção coerente de VSB

Para a recuperação exata do sinal de mensagem $m(t)$ da onda modulada VSB $s(t)$, exceto para algum escalonamento em amplitude, podemos utilizar o detector coerente mostrado na Fig. 3.12. Tal como a demodulação de DSB-SC e SSB estudadas anteriormente, a demodulação de VSB consiste na multiplicação de $s(t)$ por uma senóide gerada localmente e, então, pela filtragem passa-baixa do produto resultante $v(t)$. Assume-se que a senóide local no detector coerente da Fig. 3.12 está em perfeito sincronismo com a portadora do modulador responsável pela geração da onda modulada VSB. Então, ajustando a fase ϕ da senóide local da Fig. 3.12 para zero, podemos expressar a transformada de Fourier do sinal produto

$$v(t) = A'_c s(t) \cos(2\pi f_c t)$$

como

$$V(f) = \frac{1}{2} A'_c [S(f - f_c) + S(f + f_c)] \quad (3.32)$$

na qual

$$s(t) \iff S(f)$$

A seguir, expressamos a transformada de Fourier da onda modulada VSB $s(t)$ por

$$S(f) = \frac{1}{2} A_c [M(f - f_c) + M(f + f_c)] H(f) \quad (3.33)$$

da qual, utilizando a Fig. 3.23 que mostra o modulador VSB: $M(f)$ é o espectro da mensagem e $H(f)$ é a função de transferência do filtro formador de faixa lateral. Deslocando o espectro VSG $S(f)$ para a direita, por f_c , temos

$$S(f - f_c) = \frac{1}{2} A_c [M(f - 2f_c) + M(f)] H(f - f_c) \quad (3.34)$$

e deslocando-o para a esquerda por f_c , temos

$$S(f + f_c) = \frac{1}{2} A_c [M(f) + M(f + 2f_c)] H(f + f_c) \quad (3.35)$$

Logo, substituindo as Eqs. (3.34) e (3.35) na Eq. (3.32) e combinando os termos, obtemos

$$\begin{aligned} V(f) &= \frac{1}{4} A_c A'_c M(f) [H(f - f_c) + H(f + f_c)] \\ &\quad + \frac{1}{4} A_c A'_c [M(f - 2f_c) H(f - f_c) + M(f + 2f_c) H(f + f_c)] \end{aligned}$$

a qual, considerando a condição imposta em $H(f)$ pela Eq. (3.26), se reduz para

$$\begin{aligned} V(f) &= \frac{1}{4} A_c A'_c M(f) \\ &\quad + \frac{1}{4} A_c A'_c [M(f - 2f_c) H(f - f_c) + M(f + 2f_c) H(f + f_c)] \end{aligned} \quad (3.36)$$

O primeiro termo do lado direito da Eq. (3.36) é a versão escalonada do espectro de mensagem $M(f)$. O segundo termo da Eq. (3.36) é a transformada de Fourier das compo-

nentes de alta freqüência, representando uma nova onda modulada VSB com freqüência de portadora $2f_c$. Desde que o filtro passa-baixa do detector coerente da Fig. 3.12 possua freqüência de corte um pouco maior do que a largura de faixa da mensagem, as componentes de alta freqüência de $v(t)$ são removidas pelo filtro. O sinal demodulado resultante é uma versão escalonada do sinal de mensagem $m(t)$ desejado.

- **PROBLEMA 3.14** Comprove a afirmativa de que as componentes de alta freqüência da Eq. (3.36) representam uma onda modulada VSB com freqüência de portadora $2f_c$. ◀

EXEMPLO 3.4 Detecção coerente de VSB senoidal

Lembre-se da Eq. (3.31) do Exemplo 3.3, que o sinal senoidal modulado VSB é definido por

$$\begin{aligned} s(t) &= \frac{1}{2} A_c A_m \cos(2\pi f_m t) \cos(2\pi f_c t) \\ &\quad + \frac{1}{2} A_c A_m (1 - 2k) \sin(2\pi f_m t) \sin(2\pi f_c t) \end{aligned}$$

A multiplicação de $s(t)$ por $A'_c \cos(2\pi f_c t)$, de acordo com a detecção coerente perfeita, resulta no sinal produto

$$\begin{aligned} v(t) &= A'_c s(t) \cos(2\pi f_c t) \\ &= \frac{1}{2} A_c A'_c A_m \cos(2\pi f_m t) \cos^2(2\pi f_c t) \\ &\quad + \frac{1}{2} A_c A'_c A_m (1 - 2k) \sin(2\pi f_m t) \sin(2\pi f_c t) \cos(2\pi f_c t) \end{aligned}$$

A seguir, utilizando identidades trigonométricas

$$\cos^2(2\pi f_c t) = \frac{1}{2}[1 + \cos(4\pi f_c t)]$$

e

$$\sin(2\pi f_c t) \cos(2\pi f_c t) = \frac{1}{2} \sin(4\pi f_c t)$$

podemos redefinir $v(t)$ por

$$\begin{aligned} v(t) &= \frac{1}{4} A_c A'_c A_m \cos(2\pi f_m t) \\ &\quad + \frac{1}{4} A_c A'_c A_m [\cos(2\pi f_m t) \cos(4\pi f_c t) + (1 - 2k) \sin(2\pi f_m t) \sin(4\pi f_c t)] \quad (3.37) \end{aligned}$$

O primeiro termo do lado direito da Eq. (3.37) é uma versão escalonada do sinal de mensagem $A_m \cos(2\pi f_m t)$. O segundo termo da equação é uma nova onda senoidal modulada VSB com freqüência de portadora $2f_c$, a qual representa as componentes de alta freqüência de $v(t)$. Este segundo termo é removido pela filtragem passa-baixa no detector da Fig. 3.12, desde que a freqüência de corte do filtro seja um pouco maior do que a freqüência da mensagem f_m .

EXEMPLO 3.5 Detecção de envelope de VSB mais portadora

A detecção coerente de VSB requer o sincronismo do receptor com o transmissor, o que aumenta a complexidade do sistema. Para simplificar o processo de demodulação, podemos adicionar, propositalmente, a portadora do sinal VSB (escalonada pelo fator k_a) antes da transmissão e, então, utilizar a detecção de envelope no receptor.³ Assumindo a modulação senoidal, o sinal “VSB mais portadora” é definido [veja a Eq. (3.31) do Exemplo 3.3] por

$$\begin{aligned}s_{\text{VSB+C}}(t) &= A_c \cos(2\pi f_c t) + k_a s(t), \quad k_a = \text{fator de sensibilidade a amplitude} \\ &= A_c \cos(2\pi f_c t) + \frac{k_a}{2} A_c A_m \cos(2\pi f_m t) \cos(2\pi f_c t) \\ &\quad + \frac{k_a}{2} A_c A_m (1 - 2k) \sin(2\pi f_m t) \sin(2\pi f_c t) \\ &= A_c \left[1 + \frac{k_a}{2} A_m \cos(2\pi f_m t) \right] \cos(2\pi f_c t) \\ &\quad + \frac{k_a}{2} A_c A_m (1 - 2k) \sin(2\pi f_m t) \sin(2\pi f_c t)\end{aligned}$$

O envelope de $s_{\text{VSB+C}}(t)$ é, portanto,

$$\begin{aligned}a(t) &= \left\{ A_c^2 \left[1 + \frac{k_a}{2} A_m \cos(2\pi f_m t) \right]^2 + A_c^2 \left[\frac{k_a}{2} A_m (1 - 2k) \sin(2\pi f_m t) \right]^2 \right\}^{1/2} \\ &= A_c \left[1 + \frac{k_a}{2} A_m \cos(2\pi f_m t) \right] \left\{ 1 + \left[\frac{\frac{k_a}{2} A_m (1 - 2k) \sin(2\pi f_m t)}{1 + \frac{k_a}{2} A_m \cos(2\pi f_m t)} \right]^2 \right\}^{1/2}\end{aligned}\quad (3.38)$$

A Eq. (3.38) mostra que a distorção na detecção de envelope efetuada no envelope $a(t)$ é contribuída pela componente em quadratura do sinal senoidal VSB. Esta distorção pode ser reduzida usando a combinação de dois métodos:

- O fator de sensibilidade k_a é reduzido, o que resulta na redução do percentual de modulação.
- O tamanho da faixa lateral vestigial é reduzido, resultando na redução do fator $(1 - 2k)$.

Os dois métodos são intuitivos quando analisamos o termo dentro do colchete da Eq. (3.38).

3.8 REPRESENTAÇÃO BANDA BASE DE ONDAS MODULADAS E FILTROS FAIXA PASSANTE

A partir da discussão de estratégias de modulação diferentes apresentadas neste capítulo, vemos que a onda modulada utilizando uma onda senoidal é, na realidade, um *sinal faixa passante centrado na freqüência da portadora*. Em função deste fato, a onda portadora está contida na estrutura da onda modulada. Em um sentido explícito, isso acontece quando a onda portadora é contida como uma componente separada no sinal transmitido. Quando a portadora é suprimida, sua presença é conhecida no receptor em

³ Outro procedimento utilizado para a detecção de uma onda modulada VSB é adicionar um piloto à onda modulada no transmissor. O piloto pode ser uma versão transladada em freqüência da portadora utilizada na geração da onda modulada, mas deve estar fora da faixa de freqüências ocupadas pela onda modulada. No receptor, o piloto é extraído pela filtragem passa-faixa e, então, transladado (para cima ou para baixo) para produzir uma réplica da portadora original. Com esta réplica da portadora disponível no receptor, a detecção coerente pode ser utilizada para recuperar o sinal de mensagem original. Um procedimento similar pode ser utilizado para a detecção coerente de ondas moduladas SSB.

um sentido implícito pelo posicionamento das faixas laterais do espectro transmitido ao redor da freqüência da portadora de uma forma ou de outra, dependendo do tipo da modulação utilizado.

Tipicamente, a freqüência da portadora é grande quando comparada com a largura de faixa da mensagem, o que torna o processamento da onda modulada em um computador digital uma tarefa difícil. Entretanto, da teoria de modulação apresentada neste capítulo, sabemos que todo o conteúdo de informação de um sinal de mensagem reside completamente nas faixas laterais da onda modulada. Desta forma, quando o objetivo é processar a onda modulada em um computador, o procedimento eficiente é processar a versão banda base da onda modulada, ao invés de processar diretamente a onda modulada. O termo “banda base” é utilizado para designar a *faixa de freqüências representando o sinal original, como entregue pela fonte de informação*.

■ Representação banda base de ondas moduladas

Considere, então, uma onda genérica, modulada linearmente, definida por

$$s(t) = s_I(t) \cos(2\pi f_c t) - s_Q(t) \sin(2\pi f_c t) \quad (3.39)$$

Seja

$$c(t) = \cos(2\pi f_c t)$$

a onda portadora com freqüência f_c e

$$\hat{c}(t) = \sin(2\pi f_c t)$$

a versão em *quadratura de fase* da portadora. Por simplicidade, sem perda de generalização, fizemos a amplitude da portadora igual a um. Expressamos, agora, a onda modulada na forma compacta

$$s(t) = s_I(t)c(t) - s_Q(t)\hat{c}(t) \quad (3.40)$$

O termo $s_I(t)$ é chamado de *componente em fase* da onda modulada $s(t)$, chamado desta forma porque ele é multiplicado pela portadora $c(t)$. Pelo mesmo motivo, o termo $s_Q(t)$ é chamado de *componente em quadratura de fase* ou simplesmente, *componente em quadratura* de $s(t)$, pois ele é multiplicado pela portadora em quadratura $\hat{c}(t)$. As portadoras $c(t)$ e $\hat{c}(t)$ são ortogonais entre si.

A Eq. (3.39) ou (3.40) é chamada de *representação canônica de ondas moduladas linearmente*. Mais importante ainda, esta representação inclui todos os membros da família de modulação em amplitude discutida neste capítulo, como mostrado na Tabela 3.1.

A partir desta tabela, fica claro que o conteúdo de informação do sinal de mensagem $m(t)$ e a forma pela qual a estratégia de modulação é implementada são totalmente descritos pela componente em fase $s_I(t)$ tanto em AM quanto DSB-SC, ou pela combinação da componente em fase $s_I(t)$ e pela componente em quadratura $s_Q(t)$ em SSB e VSB. Além disso, a ortogonalidade de $s_I(t)$ e $s_Q(t)$ nos impele a introduzir um novo sinal chamado de *envelope complexo* da onda modulada $s(t)$, o qual é formalmente definido por

$$\tilde{s}(t) = s_I(t) + js_Q(t) \quad (3.41)$$

Esta definição é motivada pela forma na qual lidamos com números complexos. Em qualquer evento, o ponto importante a ser notado na Eq. (3.41) é o fato de que o envelope complexo $\tilde{s}(t)$ possui todo o conteúdo de informação de $s_I(t)$ e $s_Q(t)$. Note, entretanto, que o envelope complexo $\tilde{s}(t)$ é um sinal fictício, utilizado para simplificar as operações de processamento de sinal em sinais banda base, exemplificados pelas ondas moduladas baseadas em portadora senoidal.

TABELA 3.1 Formas diferentes de modulação linear como casos especiais da Eq. (3.39), assumindo amplitude unitária para a portadora

<i>Tipo de modulação</i>	<i>Componente em fase</i> $s_i(t)$	<i>Componente em quadratura</i> $S_Q(t)$	<i>Comentários</i>
AM	$1 + k_a m(t)$	0	k_a = sensibilidade à amplitude $m(t)$ = sinal de mensagem
DSB-SC	$m(t)$	0	
SSB:			
(a) Transmissão da faixa lateral superior	$\frac{1}{2}m(t)$	$\frac{1}{2}\hat{m}(t)$	$\hat{m}(t)$ = transformada de Hilbert de $m(t)$ (veja a parte (i) da nota de rodapé 4) ⁴
(b) Transmissão da faixa lateral inferior	$\frac{1}{2}m(t)$	$-\frac{1}{2}\hat{m}(t)$	
VSB:			
(a) Vestígio da faixa lateral inferior transmitido	$\frac{1}{2}m(t)$	$\frac{1}{2}m'(t)$	$m'(t)$ = resposta do filtro com função de transferência $H_Q(f)$ definido ao sinal de mensagem $m(t)$. $H_Q(f)$ é definido pela fórmula (veja parte (ii) da nota de rodapé 4)
(b) Vestígio da faixa lateral superior transmitido	$\frac{1}{2}m(t)$	$-\frac{1}{2}m'(t)$	$H_Q(f) = -j[H(f + f_c) - H(f - f_c)]$ na qual $H(f)$ é a função de transferência do filtro formador de faixa lateral VSB.

⁴ Dois comentários adicionais sobre a Tabela 3.1:

(i) Na modulação SSB, a transformada de Hilbert

$$\hat{m}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{m(\tau)}{t - \tau} d\tau$$

define a componente em quadratura da onda modulada $s(t)$, ou seja,

$$s_Q(t) = \hat{m}(t)$$

No domínio da freqüência, a transformada de Hilbert é descrita por

$$\hat{M}(f) = -j \operatorname{sgn}(f) M(f)$$

na qual

$$\operatorname{sgn}(f) = \begin{cases} 1, & \text{para } f > 0 \\ 0, & \text{para } f = 0 \\ -1, & \text{para } f < 0 \end{cases}$$

é a função sinal.

(ii) Na modulação VSB, a componente em quadratura $s_Q(t) = m'(t)$ é obtida passando o sinal de mensagem $m(t)$ em um filtro linear invariante no tempo cuja função de transferência é representada por $H_Q(t)$. $H_Q(t)$ é definido por,

$$H_Q(f) = -j[H(f + f_c) - H(f - f_c)]$$

na qual $H(f)$ é a função de transferência do filtro formador de faixa lateral. No limite, quando a faixa lateral vestigial f_v aproxima-se de zero, temos

$$\lim_{f_v \rightarrow 0} H_Q(f) = -j \operatorname{sgn}(f)$$

e, dessa forma, a VSB se reduz para a SSB, exatamente como deve ser.

De maneira correspondente à Eq. (3.41), podemos definir a *onda portadora complexa*

$$\begin{aligned}\tilde{c}(t) &= c(t) + j\hat{c}(t) \\ &= \cos(2\pi f_c t) + j \sin(2\pi f_c t) \\ &= \exp(j2\pi f_c t)\end{aligned}\quad (3.42)$$

Dessa forma, a onda modulada $s(t)$ é definida por

$$\begin{aligned}s(t) &= \operatorname{Re}[\tilde{s}(t)\tilde{c}(t)] \\ &= \operatorname{Re}[\tilde{s}(t) \exp(j2\pi f_c t)]\end{aligned}\quad (3.43)$$

Na qual $\operatorname{Re}[\cdot]$ extrai a parte real da grandeza complexa dentro dos colchetes.

Agora podemos ver a vantagem prática do envelope complexo $\tilde{s}(t)$ sobre a onda modulada de valor real $s(t)$:

1. A componente de mais alta freqüência de $s(t)$ pode ser tão grande quanto $f_c + W$, na qual f_c é a freqüência da portadora e W é a largura de faixa da mensagem.
2. Por outro lado, a componente de freqüência mais alta de $\tilde{s}(t)$ é consideravelmente menor, sendo limitada pela largura de faixa W da mensagem.

Mesmo assim, na utilização da Eq. (3.43) como a representação da onda modulada $s(t)$, não existe nenhuma perda.

Dada uma onda modulada arbitrariamente $s(t)$. Podemos obter a componente em fase $s_I(t)$ e a componente em quadratura $s_Q(t)$ utilizando o esquema mostrado na Fig. 3.25(a). De forma alternativa, dado o par em fase $s_I(t)$ e a componente em quadratura $s_Q(t)$, podemos gerar a onda modulada $s(t)$ utilizando o esquema complementar mostrado na Fig. 3.25(b). Por motivos óbvios, estes dois esquemas são respectivamente chamados de *analisador* e *sintetizador* de ondas moduladas.

- **PROBLEMA 3.15** A obtenção do sintetizador mostrado na Fig. 3.25(b) é feita diretamente da Eq. (3.39). Entretanto, a obtenção do analisador mostrado na Fig. 3.25(a) requer uma consideração mais detalhada. Dado que $f_c > W$ e as identidades trigonométricas:

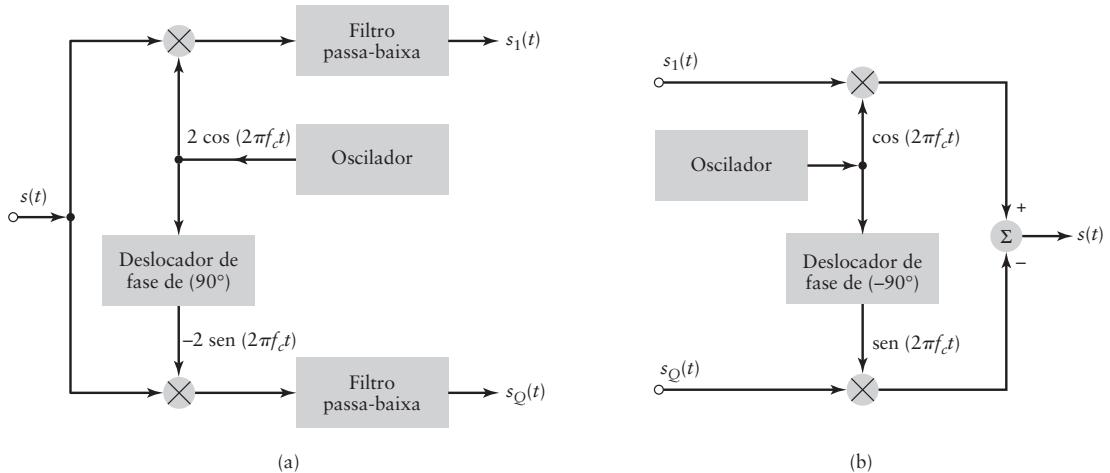


Figura 3.25 (a) Esquema para a obtenção das componentes em fase e em quadratura de um sinal modulado linearmente (isto é, passa-faixa). (b) Esquema para a reconstrução do sinal modulado de suas componentes em fase e em quadratura.

$$\cos^2(2\pi f_c t) = \frac{1}{2}[1 + \cos(4\pi f_c t)],$$

$$\sin^2(2\pi f_c t) = \frac{1}{2}[1 - \cos(4\pi f_c t)],$$

e

$$\sin(2\pi f_c t) \cos(2\pi f_c t) = \frac{1}{2} \sin(4\pi f_c t),$$

mostre que o analisador da Fig. 3.25(a) resulta em $s_I(t)$ e $s_Q(t)$ como suas duas saídas. ◀

■ Representação banda base de filtros passa-faixa

A representação banda base de um sinal faixa passante (exemplificado por uma onda modulada) desenvolvida nesta seção nos impele ao desenvolvimento de uma representação correspondente para filtro passa-faixa, incluindo canais de comunicação passa-faixa.

Para isto, considere o filtro passa-faixa linear cujo comportamento entrada/saída é definido pela função de transferência $H(f)$, a qual é limitada dentro das freqüências $\pm B$ da freqüência de meia faixa f_c . De fato, $2B$ define a largura de faixa do filtro. Suponha uma onda modulada $s(t)$ aplicada a este filtro, produzindo a saída $y(t)$, como mostrado na Fig. 3.26(a). Assumimos que a largura de faixa de transmissão da onda modulada é $2W$, centrada na freqüência de portadora f_c . Em outras palavras, o espectro da onda modulada e a resposta em freqüência do filtro passa-faixa são alinhados, com $B \leq W$. (A razão para ignorar o caso $B > W$ é que, em tal situação, a onda modulada $s(t)$ passa através do filtro completamente inalterada, o que, portanto, não possui importância prática.) Obviamente, podemos determinar o sinal de saída $y(t)$ calculando a transformada inversa de Fourier do produto $H(f)S(f)$. Um procedimento mais simples, entretanto, é utilizar uma transformação *passa-faixa para passa-baixa* (isto é, banda base),⁵ a qual elimina a freqüência da portadora f_c da análise. Especificamente, esta transformação é definida por

$$\tilde{H}(f - f_c) = 2H(f), \quad \text{para } f > 0 \quad (3.44)$$

A nova função de freqüência $\tilde{H}(f)$ é a função de transferência do filtro passa-baixa complexo, o que resulta na transformação definida na Eq. (3.44). O fator de escala 2 nesta equação é necessário para garantir que a transformação apresente o resultado exato quando calcularmos a saída $y(t)$.

De acordo com a Eq. (3.44), podemos determinar $\tilde{H}(f)$ pelo procedimento a seguir:

1. Dada a função de transferência $H(f)$ de um filtro passa-faixa, a qual é definida para freqüências positivas e negativas, mantenha a parte de $H(f)$ que corresponde às freqüências positivas. Seja $H_+(f)$ esta parte.

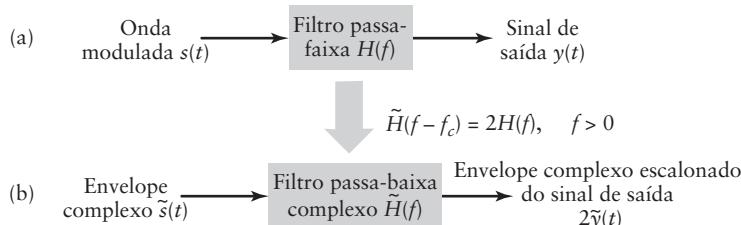


Figura 3.26 Filtro passa-faixa para a transformação complexa de sistemas passa-baixa.
(a) Configuração passa-faixa de valor real e (b) configuração correspondente passa-faixa de valor complexo.

⁵ Para a obtenção da transformação definida pela Eq. (3.44), veja Haykin (2000), p. 731.

2. Desloque $H_+(f)$ para a esquerda ao longo do eixo das freqüências por um total igual a f_c e escalone-o por um fator 2. O resultado obtido define $\tilde{H}(f)$ desejado.

Tendo determinado o filtro passa-baixa complexo caracterizado por $\tilde{H}(f)$, podemos, então, proceder com o próximo estágio do *processamento de sinal complexo*. Especificamente, a entrada deste filtro é o envelope complexo $\tilde{s}(t)$ da onda modulada $s(t)$. $\tilde{s}(t)$ é obtido de $s(t)$ de acordo com a Eq. (3.41). Então, aplicando $\tilde{s}(t)$ em $\tilde{H}(f)$ como mostrado na Fig. 3.26b, determinamos o envelope complexo $\tilde{y}(t)$ do sinal de saída $y(t)$. Finalmente, a saída real $y(t)$ é determinada pela fórmula

$$y(t) = \operatorname{Re}[\tilde{y}(t) \exp(j2\pi f_c t)] \quad (3.45)$$

a qual é simplesmente a Eq. (3.43) reescrita.

- **PROBLEMA 3.16** Começando do sistema passa-baixa complexo mostrado na Fig. 3.26(b), mostre que $y(t)$ obtido da Eq. (3.45) é idêntico a saída real $y(t)$ da Fig. 3.26(a). ◀

3.9 EXEMPLOS DE TEMAS

Nesta seção iremos descrever três exemplos de temas, os quais são desenvolvidos utilizando a teoria de modulação de onda contínua descrita nas seções anteriores deste capítulo. Os exemplos enfatizam aspectos operacionais de sistema de comunicação analógicos em vez de equações matemáticas ou detalhes de projetos.

■ Receptor super heterodino

Em um sistema de broadcasting, independente se ele for baseado em modulação em amplitude ou modulação em freqüência, o receptor não somente possui a tarefa de demodular o sinal modulado de entrada mas também de executar outras funções do sistema:

- *Sintonização da freqüência da portadora*, com o propósito de selecionar o sinal desejado (isto é, a estação desejada de rádio ou TV).
- *Filtragem*, necessária para separar o sinal desejado de outros sinais modulados que podem ser recebidos ao longo do caminho.
- *Amplificação*, necessária para compensar a perda de potência do sinal que ocorre ao longo da transmissão.

O *receptor super heterodino* é um tipo especial de receptor que implementa todas as três funções, particularmente as duas primeiras, de forma prática e elegante. Especificamente, ele sobrepõe a dificuldade de construir um filtro variável altamente seletivo em freqüência. De fato, praticamente todo receptor de rádio e TV é construído baseado no super heterodino.

Basicamente, o receptor é constituído de uma seção de rádio freqüência (RF), um mixer e um oscilador local, uma seção de freqüência intermediária (FI), um demodulador e um amplificador de potência. Parâmetros típicos de receptores de rádio AM comerciais são listados na Tabela 3.2. (A tabela também inclui parâmetros de freqüência correspondentes de receptores FM comerciais, para efeito de comparação. A teoria de modulação em freqüência (FM) é apresentada no Capítulo 4.) A Fig. 3.27 mostra o diagrama em blocos de um receptor super heterodino para modulação em amplitude utilizando um detector de envelope para a demodulação.

A onda modulada em amplitude de entrada é recebida pela antena do receptor e amplificada na seção RF, sintonizada na freqüência da portadora da onda de entrada. A combinação do mixer e do oscilador local (de freqüência ajustável) fornece a função *heteródina*, pela qual o sinal de entrada é convertido para uma *freqüência intermediária* previamente determinada, usualmente menor do que a freqüência da portadora de entrada. Essa translação de freqüência é obtida sem perturbar a relação das freqüências laterais com a portadora. O resultado é a produção de uma portadora de freqüência intermediária definida por

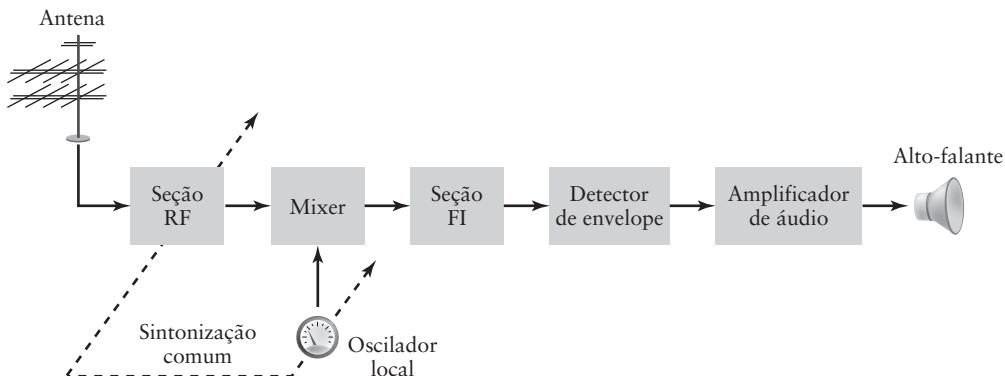


Figura 3.27 Elementos básicos de um rádio receptor AM do tipo superheterodino.

$$f_{FI} = f_{RF} - f_{LO} \quad (3.46)$$

na qual f_{LO} é a freqüência do oscilador local e f_{RF} é a freqüência da portadora do sinal RF de entrada. Referimos à f_{FI} como freqüência intermediária (FI) porque o sinal não está nem na freqüência de entrada original nem na freqüência final banda base. A combinação do mixer com o oscilador local algumas vezes é chamada de *primeiro detector* e, nesse caso, o demodulador (detector de envelope da Fig. 3.27) é chamado de *segundo detector*.

A seção FI consiste de um ou mais estágios de amplificação sintonizada, com a largura de faixa necessária para o tipo particular de sinal que o receptor deve trabalhar. Esta seção fornece grande parte da amplificação e seletividade do receptor. A saída da seção FI é aplicada ao demodulador, com o propósito de recuperar o sinal banda base. Se a detecção coerente for utilizada, então a fonte do sinal coerente deve ser fornecida no receptor. A operação final no receptor é a amplificação de potência do sinal de mensagem recuperado.

Em um receptor super heterodino, o mixer irá desenvolver uma saída de freqüência intermediária quando a freqüência do sinal de entrada for maior ou menor do que a freqüência do oscilador local, por um total igual à freqüência intermediária. Ou seja, existem duas freqüências de entrada $-|f_{LO} \pm f_{RF}|$, que resultarão na f_{LO} na saída do mixer. Isso introduz a possibilidade de recepção simultânea de dois sinais diferindo em freqüência por duas vezes a freqüência intermediária. Por exemplo, um receptor sintonizado para 1 MHz e tendo FI de 0,455 MHz é sujeito a uma interferência de imagem em 1,910 MHz. De fato, qualquer receptor com este valor de FI, quando sintonizado em qualquer estação, está sujeito a interferência de imagem para uma freqüência de 0,910 MHz maior do que a estação desejada. Como a função do mixer é produzir a diferença entre duas freqüências aplicadas, ele é incapaz de distinguir entre o sinal desejado e sua imagem enquanto ele produz a saída FI de qualquer uma delas. A única solução prática para a supressão da interferência de imagem é construir estágios altamente seletivos na seção RF (isto é, entre a antena e o mixer), para favorecer o sinal desejado e discriminar o sinal de imagem indesejado. A efetividade da supressão da imagem de sinal indesejada aumenta quando o número de estágios seletivos na seção de rádio freqüência aumenta e quando a relação do sinal intermediário de freqüência aumenta.

TABELA 3.2 Parâmetros de freqüência típicos de receptores de rádio AM e FM

	Rádio AM	Rádio FM
Faixa da portadora RF	0,535–1,605 MHz	88–108 MHz
Freqüência de meia-faixa da seção FI	0,455 MHz	10,7 MHz
Largura da faixa FI	10 kHz	200 kHz

■ Sinais de televisão

A modulação de faixa lateral vestigial, discutida na Seção 3.7, possui um importante papel na televisão comercial. Os detalhes exatos do formato de modulação utilizados para transmitir o sinal de vídeo que caracterizam o sistema de TV são influenciados por dois fatores:

1. O sinal de vídeo exibe uma grande largura de faixa e um pequeno conteúdo de baixa freqüência, o que sugere o uso de modulação de faixa lateral vestigial.
2. O circuito utilizado para a demodulação no receptor deve ser simples e, portanto, barato. Isto sugere o uso de detecção de envelope, o que requer a adição da portadora na onda modulada VSB.

Com relação ao ponto 1, entretanto, deve ser ressaltado que realmente existe o desejo básico de conservar a largura de faixa, na transmissão broadcasting de TV comercial, o sinal transmitido não é, exatamente, modulado em VSB. A razão é que os níveis de potência de transmissão são altos, o que significa que seria muito caro controlar rigidamente a filtragem das faixas laterais. Ao invés disso, um *filtro VSB* é inserido em cada receptor nos quais os níveis de potência são baixos. A performance total é a mesma de uma modulação de faixa lateral vestigial, exceto por alguma potência e largura de faixa desperdiçadas. Estas considerações são ilustradas na Fig. 3.28. Em particular, a Fig. 3.28(a) mostra o espectro idealizado de um sinal de TV transmitidos. A faixa lateral superior, 25% da faixa lateral inferior, e a portadora de vídeo são transmitidas. A resposta em freqüência do filtro VSB utilizado para a formação do espectro necessário no receptor é mostrada na Fig. 3.28(b).

A largura de faixa do canal utilizado para a transmissão broadcasting de TV na América do Norte é de 6 MHz, como indicado na Fig. 3.28(b). Esta largura de faixa de canal não acomoda somente os requisitos de largura de faixa do sinal de vídeo modulado VSB, mas também o sinal de som que modula uma portadora própria. Os valores apresentados no eixo de freqüências das Figs. 3.28(a) e 3.28(b) pertencem a um canal de TV específico. De acordo com a figura, a freqüência da portadora de vídeo está em 55,25 MHz e a freqüência da portadora de som está em 59,75 MHz. Note, entretanto, que o conteúdo de informação do sinal de TV está no espectro banda base, estendendo-se de 1,25 MHz abaixo da portadora de vídeo até 4,5 MHz acima dela.

Com relação ao ponto 2, a utilização da detecção de envelope (aplicada à onda modulada VSB acrescida pela portadora) produz uma *distorção da forma de onda* no sinal de vídeo recuperado na saída do detector. Como discutido no Exemplo 3.5, a distorção de forma de onda é produzida pela componente em quadratura da onda modulada VSB. Como ressaltado no exemplo, podemos reduzir a extensão da distorção da forma de onda reduzindo o percentual de modulação e minimizando a largura da faixa lateral vestigial.

■ Multiplexação por divisão em freqüência

Outra importante operação de processamento de sinal em comunicações analógicas é a multiplexação, pela qual um número de sinais independentes pode ser combinado em um sinal composto adequado para a transmissão em um canal comum. Freqüências de voz transmitidas em sistemas telefônicos, por exemplo, na faixa de 300 a 3100 Hz. Para transmitir um número destes sinais em um mesmo canal (isto é, cabo), os sinais devem ser mantidos separados, de tal forma que eles não interfiram uns com os outros e, portanto, possam ser separados no lado do receptor. Isto é feito separando os sinais na freqüência ou no tempo. A técnica de separação dos sinais na freqüência é chamada de *multiplexação por divisão na freqüência* (FDM, do termo em inglês), enquanto que a técnica de separação dos sinais no tempo é chamada de *multiplexação por divisão no tempo* (TDM, do termo em inglês). Nesta seção iremos discutir a FDM. A discussão da TDM é deixada para o Capítulo 5.

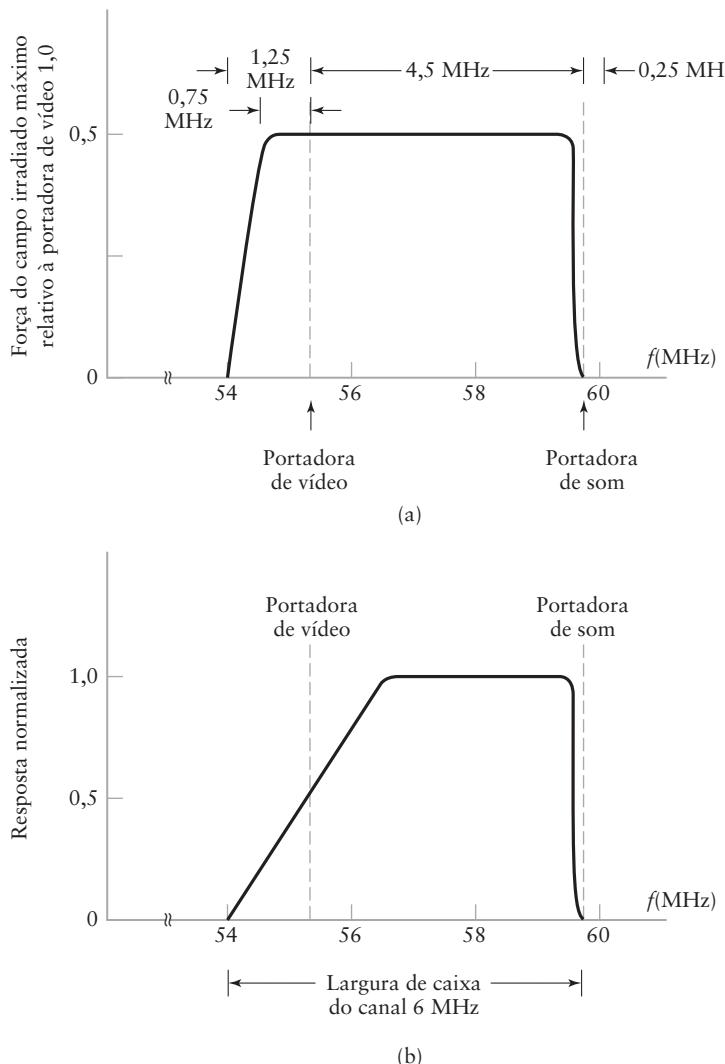


Figura 3.28 (a) Espectro em amplitude idealizado de um sinal de TV transmitido. (b) Resposta em amplitude de um filtro de formação VSB no receptor.

O diagrama em blocos do sistema FDM é mostrado na Fig. 3.29. Os sinais de mensagem de entrada são considerados como sendo do tipo passa-baixa, mas seus espectros não necessariamente possuem valores não nulos em todas as freqüências até a freqüência zero. Seguindo cada sinal de entrada, mostramos um filtro passa-baixa, o qual é projetado para remover componentes de alta freqüência que não contribuem significativamente para a representação do sinal, mas que são capazes de perturbar outros sinais de mensagem que partilham o canal comum. Estes filtros passa-baixa são omitidos somente se os sinais de entrada são, inicialmente, suficientemente limitados em faixa. Os sinais de entrada são aplicados a moduladores que deslocam as faixas de freqüência dos sinais, de forma a ocuparem intervalos de freqüência mutuamente exclusivos. As freqüências de portadora necessárias para executar estas translações de freqüência são obtidas de uma fornecedora de portadora. Para a modulação, podemos utilizar qualquer um dos métodos descritos nas seções anteriores deste capítulo. Entretanto, em telefonia, o método mais amplamente utilizado de modulação na multiplexação por divisão em freqüência é a modulação de faixa lateral úni-

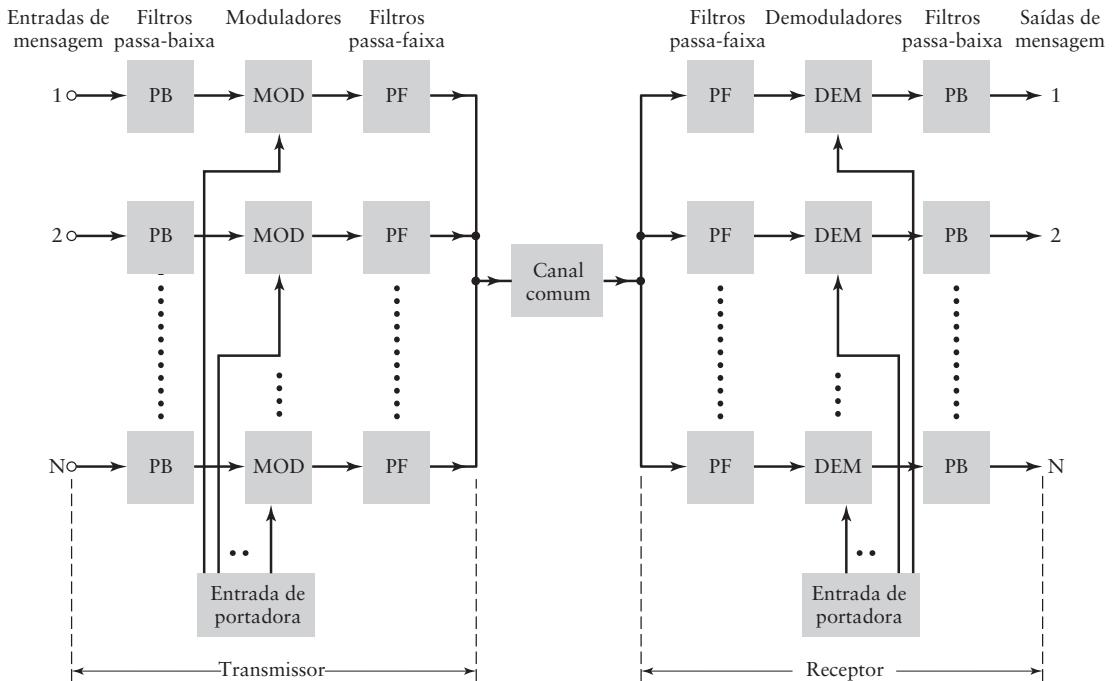


Figura 3.29 Diagrama em blocos de um sistema de multiplexação por divisão na freqüência (FDM).

ca, o que, no caso dos sinais de voz, requer uma largura de faixa aproximadamente igual à do sinal de voz original. Na prática, para cada sinal de voz é geralmente associado uma largura de faixa de 4 kHz. Os filtros passa-faixa seguindo os moduladores são utilizados para restringir a faixa de cada onda modulada para sua faixa especificada. Os resultados da saída dos filtros passa-faixa são combinados em paralelo para formar a entrada do canal comum. No terminal receptor, um banco de filtros passa-faixa, com suas entradas conectadas em paralelo, é utilizado para separar os sinais de mensagem com base na ocupação de freqüência. Note que o sistema FDM mostrado na Fig. 3.29 opera de forma unidirecional apenas. Para fornecer uma transmissão bidirecional, como no exemplo de telefonia, precisamos duplicar completamente o sistema de multiplexação, com os componentes conectados na ordem inversa e com as ondas de sinal indo da direita para a esquerda.

EXEMPLO 3.6 Passos de modulação em um sistema FGM de 60 canais

A implementação prática de um sistema FGM geralmente envolve vários passos de modulação e desmodulação, como ilustrado na Fig. 3.30. O primeiro passo de multiplexação combina 12 entradas de voz em um *grupo básico*, o qual é formado tendo a n -ésima entrada modulada por uma portadora com freqüência $f_c = 60 + 4n$ KHz, na qual $n = 1, 2, \dots, 12$. As faixas laterais inferiores são, então, selecionadas pela filtragem passa-faixa e combinadas para formar um grupo de 12 faixas laterais inferiores (uma para cada entrada de voz). Então, o grupo básico ocupa a faixa de freqüência de 60-108 kHz. O passo seguinte na hierarquia FDM envolve a combinação de cinco grupos básicos em um *supergrupo*. Isto é realizado usando o n -ésimo grupo para modular a freqüência de portadora $f_c = 372 + 48n$ kHz, na qual $n = 1, 2, \dots, 5$. Novamente, as faixas laterais são selecionadas pela filtragem e combinadas para formar o supergrupo na faixa de 312-552 kHz. Portanto, um supergrupo é projetado para acomodar 60 entradas de voz independentes. A razão para a formação do supergrupo desta forma é que filtros econômicos com as características desejadas são disponíveis somente para uma faixa limitada de freqüência. De forma similar, supergrupos são combinados em *grupos masters* e grupos masters são combinados em *grupos muito grandes*.

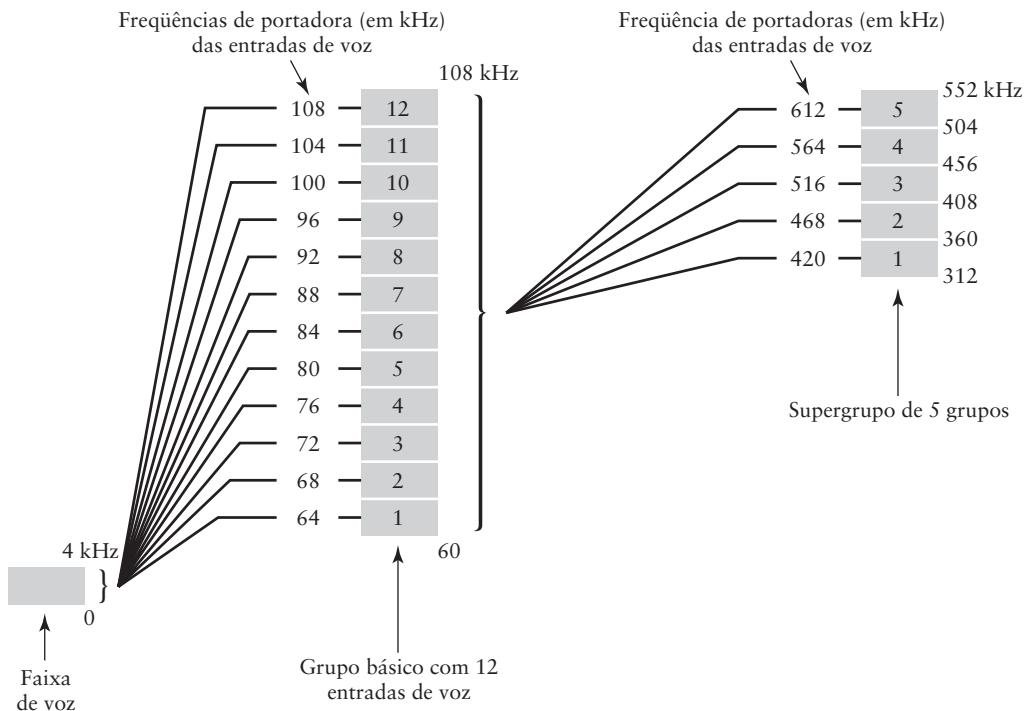


Figura 3.30 Ilustração dos passos de modulação em um sistema FDM.

3.10 RESUMO E DISCUSSÃO

Neste capítulo estudamos a família de amplitude em modulação, na qual a portadora é uma onda senoidal cuja amplitude é variada de acordo com o sinal de mensagem. O formato desta família de modulação analógica é generalizada pelo exemplo de onda modulada

$$s(t) = A_c m(t) \cos(2\pi f_c t) \quad (3.47)$$

na qual $m(t)$ é o sinal de mensagem e $A_c \cos(2\pi f_c t)$ é a portadora. A família de modulação em amplitude engloba quatro tipos de modulação de onda contínua, dependendo do conteúdo espectral da onda modulada. Os quatro tipos de modulação e seus méritos práticos são resumidos a seguir:

1. Modulação em amplitude (AM), na qual as faixas laterais superior e inferior são completamente transmitidas, acompanhadas pela onda portadora. A geração da onda AM pode ser realizada simplesmente utilizando um dispositivo não linear (por exemplo, um diodo) em um modulador de lei quadrática, por exemplo. Pelo mesmo motivo, a demodulação da onda AM é realizada de forma igualmente simples no receptor usando um detector de envelope, por exemplo. É por estas duas razões, geração simples e detecção simples, que a modulação em amplitude é amplamente utilizada na transmissão broadcasting de rádio AM, o que envolve um transmissor simples potente e vários receptores que são relativamente baratos de serem construídos.
2. Modulação por faixa lateral dupla – portadora suprimida (DSB-SC), definida pela Eq. (3.47), na qual apenas as faixas laterais superior e inferior são transmitidas. A supressão da onda portadora pela modulação DSB-SC requer menos potência do que o AM para transmitir o mesmo sinal de mensagem. Esta vantagem da modulação DSB-SC sobre a AM é, entretanto, paga pelo aumento da complexidade do receptor. A modulação DSB-SC é, portanto, adequada para a *comunicação ponto-*

a-ponto, envolvendo um transmissor e um receptor. Nesta forma de comunicação analógica, a potência transmitida é o prêmio e a utilização de um receptor completo é, pois, justificável.

3. Modulação por faixa lateral única (SSB), na qual apenas a faixa lateral superior ou inferior é transmitida. Ela é ótima no sentido de requerer o mínimo de potência transmitida e o mínimo de largura de faixa de canal para o envio de um sinal de mensagem de um ponto a outro. Entretanto, a implementação do transmissor SSB impõe várias restrições ao conteúdo espectral do sinal de mensagem de entrada. Especificamente, ela requer a presença de um gap de baixa freqüência ao redor da freqüência zero, o que, por exemplo, é satisfeito por sinais de voz em comunicações telefônicas.
4. Modulação de faixa lateral vestigial, na qual uma faixa lateral é transmitida “quase” na totalidade além de um vestígio da outra faixa lateral, de forma complementar. A modulação VSB requer um canal com largura de faixa intermediária entre o requerido pelos sistemas SSB e DSB-SC, e a economia na largura de faixa pode ser significativa se os sinais modulantes tiverem larguras de faixa grandes, tal como no caso de sinais de televisão e dados digitais de alta velocidade.

Um comentário final é necessário. Apesar do desenvolvimento da família de modulação em amplitude ter sido motivado pela sua relevância direta em comunicações analógicas, vários aspectos da teoria desse ramo de modulação são aplicáveis às comunicações digitais. Por exemplo, o sinal de mensagem da Eq. (3.47) para a onda modulada $s(t)$ é restrito aos níveis -1 ou $+1$, representando o binário “0” e “1”, respectivamente. Então temos a forma básica de modulação digital chamada de chaveamento binário de fase (BPSK*), a ser discutido no Capítulo 7.

Problemas adicionais

- 3.17 Ao longo deste capítulo focamos em

$$c(t) = A_c \cos(2\pi f_c t)$$

como onda portadora senoidal. Suponha que escolhemos

$$c(t) = A_c \sin(2\pi f_c t)$$

como onda portadora senoidal. Para sermos consistentes, suponha que definimos, também,

$$m(t) = A_m \sin(2\pi f_m t)$$

- (a) Calcule o espectro da nova definição de AM:

$$s(t) = A_c [1 + k_a m(t)] \sin(2\pi f_c t)$$

na qual k_a é a sensibilidade de amplitude.

- (b) Compare o resultado obtido na parte (a) com o estudado no Exemplo 3.1.
(c) Qual a diferença que a formulação neste problema resulta na formulação da teoria de modulação ilustrado no Exemplo 3.1?

- 3.18 Considere o sinal de mensagem

$$m(t) = 20 \cos(2\pi t) \text{ volts}$$

e onda portadora

$$c(t) = 50 \cos(100\pi t) \text{ volts}$$

- (a) Trace a onda AM resultante para 75% de modulação.
(b) Determine a potência desenvolvida em uma carga de 100 ohms devido a esta onda AM.

* N. de T.: Do inglês, *binary phase-shift-keying*.

3.19 Usando o sinal de mensagem

$$m(t) = \frac{t}{1 + t^2}$$

determine e trace a onda modulação para modulação em amplitude cujos percentuais de modulação sejam:

- (a) 50%
- (b) 100%
- (c) 125%

3.20 Suponha que um dispositivo não linear está disponível, para o qual a corrente de saída i_o e a tensão de entrada v_i estão relacionados por:

$$i_o = a_1 v_i + a_3 v_i^3$$

na qual a_1 e a_3 são constantes. Explique como este dispositivo pode ser utilizado para fornecer a modulação em amplitude. Este dispositivo também pode ser utilizado para a demodulação? Justifique sua resposta.

3.21 Considere a onda modulada DSB-SC obtida pela utilização da onda modulante senoidal

$$m(t) = A_m \cos(2\pi f_m t)$$

e onda portadora

$$c(t) = A_c \cos(2\pi f_c t + \phi)$$

O ângulo de fase ϕ , representando a diferença de fase entre $c(t)$ e $m(t)$ no tempo $t = 0$ é variável. Trace esta onda modulada para os seguintes valores de ϕ :

- (a) $\phi = 0$
- (b) $\phi = 45^\circ$
- (c) $\phi = 90^\circ$
- (d) $\phi = 135^\circ$

Comente seus resultados.

3.22 Dado o dispositivo não linear descrito no Problema 3.20, explique como ele pode ser utilizado para criar um modulador de produto.

3.23 Considere um sinal de mensagem $m(t)$ com o espectro mostrado na Fig. 3.31. A largura de faixa da mensagem é $W = 1\text{kHz}$. Este sinal é aplicado a um modulador de produto, juntamente com a onda portadora $A_c \cos(2\pi f_c t)$, produzindo a onda modulada DSB-SC $s(t)$. Esta onda modulada é, a seguir, aplicada a um detector coerente. Assumindo sincronismo perfeito entre as ondas portadoras no transmissor e detector, determine o espectro da saída do detector quando: (a) a freqüência da portadora é $f_c = 1,25\text{ kHz}$ e (b) a freqüência da portadora é $f_c = 0,75\text{ kHz}$. Qual é a menor freqüência da portadora para a qual cada componente da onda modulada $s(t)$ é unicamente determinada de $m(t)$?

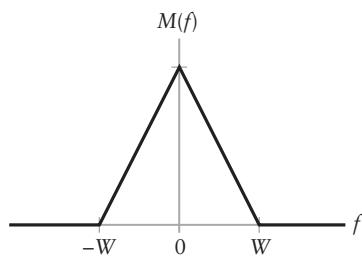


Figura 3.31 Problema 3.23.

3.24 Considere uma onda composta obtida pela adição de uma portadora não coerente $A_c \cos(2\pi f_c t + \phi)$ para a onda DSB-SC $\cos(2\pi f_c t)m(t)$. Esta onda composta é aplicada a um detector de envelope ideal. Determine a saída resultante do detector para

- (a) $\phi = 0$
- (b) $\phi \neq 0$ e $|m(t)| \ll A_c/2$

- 3.25 Uma onda DSB-SC é demodulada pela sua aplicação em um detector coerente.
- Calcule o efeito de um erro de freqüência Δf na freqüência portadora local do detector, medida com relação a freqüência da portadora do sinal DSB-SC de entrada.
 - Para o caso de uma onda modulante senoidal, mostre que, devido a este erro de freqüência, a onda demodulada exibe um batimento na freqüência do erro. Ilustre sua resposta com o rascunho desta onda demodulada. (Um batimento se refere a um sinal cuja freqüência é a diferença entre as freqüências de dois sinais de entrada.)
- 3.26 Considere um pulso de amplitude A e duração T . Este pulso é aplicado a um modulador SSB, produzindo a onda modulada $s(t)$. Determine o envelope de $s(t)$ e mostre que este envelope exibe picos no começo e fim do pulso.
- 3.27 (a) Considere um sinal de mensagem $m(t)$ contendo componentes de freqüência em 100, 200 e 400 Hz. Este sinal é aplicado a um modulador SSB juntamente com a portadora em 100 kHz, com apenas a faixa lateral superior transmitida. Em um detector coerente utilizado para recuperar $m(t)$, o oscilador local fornece a onda senoidal de freqüência 100,02 kHz. Determine as componentes de freqüência da saída do detector.
- (b) Repita sua análise assumindo que apenas a faixa lateral inferior é transmitida.
- 3.28 Ao longo deste capítulo, expressamos a onda portadora senoidal na forma

$$c(t) = A_c \cos(2\pi f_c t)$$

na qual A_c é a amplitude da portadora e f_c é a freqüência da portadora. No Capítulo 7, trabalhando com técnicas de modulação passa-faixa digital, iremos descobrir uma forma mais conveniente de expressar a portadora na forma

$$c(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_c t)$$

na qual T_b é a duração alocada para a transmissão do símbolo 1 ou símbolo 0. Determine o valor da amplitude A_c da portadora para a energia em $c(t)$ por símbolo ser unitária.

Problemas avançados

- 3.29 Para um diodo de junção p-n, a corrente i e a tensão v ao longo do diodo são relacionadas por

$$i = I_0 \left[\exp\left(-\frac{v}{V_T}\right) - 1 \right]$$

na qual I_0 é a corrente de saturação reversa e V_T é a tensão térmica definida por

$$V_T = \frac{kT}{e}$$

na qual k é a constante de Boltzmann em joules por grau Kelvin, T é a temperatura absoluta em graus kelvin e e é a carga de um elétron. À temperatura ambiente, $V_T = 0,026$ volts.

- Expanda i como uma série de potência em v , mantendo os termos até v^3 .
- Seja

$$v = 0,01 \cos(2\pi f_m t) + 0,01 \cos(2\pi f_c t) \text{ volts}$$

na qual $f_m = 1\text{kHz}$ e $f_c = 100\text{ kHz}$. Determine o espectro da corrente i resultante no diodo.

- Especifique o filtro passa-faixa necessário para extrair da corrente do diodo uma onda AM com freqüência de portadora f_c .
- Qual é o percentual de modulação desta onda AM?

- 3.30 Considere o sistema multiplexado por portadora em quadratura da Fig. 3.17. O sinal multiplexado $s(t)$ produzido na saída do transmissor da parte (a) desta figura é aplicado a um canal de comunicação com função de transferência $H(f)$. A saída deste canal é, por sua vez, aplicada à entrada do receptor da parte (b) da Fig. 3.17. Prove que a condição

$$H(f_c + f) = H^*(f_c - f), \quad \text{para } 0 \leq f \leq W$$

é necessária para a recuperação dos sinais de mensagem $m_1(t)$ e $m_2(t)$ na saída do receptor. f_c é a freqüência da portadora, W é a largura de faixa da mensagem. O asterisco em $H^*(f_c - f)$ representa o conjugado complexo.

Dica: Calcule o espectro das duas saídas do receptor.

- 3.31 (a) Seja $s_u(t)$ a onda SSB obtida pela transmissão da faixa lateral superior e $\hat{s}_u(t)$ sua transformada de Hilbert. Mostre que

$$m(t) = \frac{2}{A_c} [s_u(t) \cos(2\pi f_c t) + \hat{s}_u(t) \sin(2\pi f_c t)]$$

e

$$\hat{m}(t) = \frac{2}{A_c} [\hat{s}_u(t) \cos(2\pi f_c t) - s_u(t) \sin(2\pi f_c t)]$$

na qual $m(t)$ é o sinal de mensagem, $\hat{m}(t)$ é sua transformada de Hilbert, f_c é a freqüência da portadora e A_c é a amplitude da portadora.

- (b) Mostre que as equações correspondentes aos termos da onda SSB $s_l(t)$ obtida pela transmissão somente da faixa lateral inferior são

$$m(t) = \frac{2}{A_c} [s_l(t) \cos(2\pi f_c t) + \hat{s}_l(t) \sin(2\pi f_c t)]$$

e

$$\hat{m}(t) = \frac{2}{A_c} [s_l(t) \cos(2\pi f_c t) - \hat{s}_l(t) \sin(2\pi f_c t)]$$

- (c) Utilizando os resultados de (a) e (b), construa um diagrama de blocos de um receptor para demodulação da onda SSB.

Dica: A transformada de Hilbert é definida no Problema 2.52, veja a nota de rodapé 4 desse capítulo.

- 3.32 Neste problema, iremos continuar a discussão da modulação VSB para o caso quando o vestígio da faixa lateral inferior é transmitido. A Fig. 3.24 mostra a resposta em freqüência $H(f)$ do filtro formador de faixa lateral utilizado para gerar esta onda modulada. Em particular, queremos examinar a representação complexa deste filtro, representada por $\tilde{H}(f)$.

Sejam $H_I(f)$ e $H_Q(f)$ as componentes em fase e em quadratura de $\tilde{H}(f)$, respectivamente. Mostre que para todo o intervalo $-W \leq f \leq W$, temos

- (a) $H_I(f)$ representa todos os filtros passa-tudo, ou seja, a resposta em freqüência do filtro é constante como mostrado por

$$H(f) = 1, \quad \text{para } -W \leq f \leq W$$

onde W é a largura de faixa da mensagem.

- (b) $H_Q(f)$ representa um filtro passa-baixa com resposta em freqüência com simetria ímpar, descrito pelas três relações a seguir:

$$1. H_Q(-f) = -H_Q(f), \quad -W \leq f \leq W$$

$$2. H_Q(0) = 0$$

$$3. H_Q(f) = 1 \quad \text{para } f_v \leq f \leq W$$

onde f_v é a largura da faixa lateral vestigial.

Capítulo 4

MODULAÇÃO EM ÂNGULO

No capítulo anterior, investigamos o efeito de uma lenta variação na amplitude de uma onda portadora senoidal de acordo com o sinal contendo a informação, mantendo a freqüência da portadora fixa. Existe outra forma de modular uma onda portadora senoidal – chamada de *modulação em ângulo*, na qual o ângulo da onda portadora é variado de acordo com o sinal contendo a informação. Nesta segunda família de técnicas de modulação, a amplitude da onda portadora é mantida constante.

Uma característica importante da modulação em ângulo é que ela possibilita uma melhor discriminação contra ruído e interferência do que a modulação em amplitude. Como será mostrado no Capítulo 9, entretanto, esta melhoria é obtida ao custo do aumento da largura de faixa transmitida. Ou seja, a modulação em ângulo é um modo prático de trocar a largura de faixa do canal por melhoria com relação ao ruído. Tal compromisso não é possível na modulação em amplitude. Além disso, a melhoria com relação ao ruído na modulação em ângulo é obtida ao custo de aumento da complexidade do sistema tanto no transmissor quanto no receptor.

O material apresentado neste capítulo sobre modulação em ângulo irá nos ensinar três lições:

- *Lição 1: A modulação em ângulo é um processo não linear, o que comprova sua natureza sofisticada. No contexto de comunicações analógicas, esta propriedade distinta da modulação em ângulo possui duas implicações:*
 - Em termos analíticos, a análise espectral da modulação em ângulo é complicada.
 - Em termos práticos, a implementação da modulação em ângulo é laboriosa.
- *Lição 2: Enquanto a largura de faixa da transmissão de uma onda modulada em amplitude (ou suas variantes) possui tamanho limitado, a largura de faixa da transmissão de uma onda modulada em ângulo pode assumir um tamanho infinito, ao menos em teoria.*
- *Lição 3: Dado que a amplitude da onda portadora é mantida constante, intuitivamente esperamos que o ruído aditivo afete menos a performance da modulação em ângulo do que a modulação em amplitude.*

4.1 DEFINIÇÕES BÁSICAS

Seja $\theta_i(t)$ o ângulo de uma portadora senoidal modulada no tempo t . Considera-se que ele seja uma função do sinal contendo a informação ou sinal de mensagem. Expressamos a onda modulada em ângulo por

$$s(t) = A_c \cos[\theta_i(t)] \quad (4.1)$$

na qual A_c é a amplitude da portadora. Uma oscilação completa ocorre sempre que o ângulo $\theta_i(t)$ mudar por 2π radianos. Se $\theta_i(t)$ aumentar monotonicamente com o tempo, então a freqüência média em hertz, em um pequeno intervalo de t a $t + \Delta t$ é dada por

$$f_{\Delta t}(t) = \frac{\theta_i(t + \Delta t) - \theta_i(t)}{2\pi\Delta t}$$

Permitindo que o intervalo de tempo Δt tenda a zero, chegamos à seguinte definição para a *freqüência instantânea* de um sinal modulado em ângulo $s(t)$:

$$\begin{aligned}
f_i(t) &= \lim_{\Delta t \rightarrow 0} f_{\Delta t}(t) \\
&= \lim_{\Delta t \rightarrow 0} \left[\frac{\theta_t(t + \Delta t) - \theta_i(t)}{2\pi\Delta t} \right] \\
&= \frac{1}{2\pi} \frac{d\theta_i(t)}{dt}
\end{aligned} \tag{4.2}$$

na qual, na última linha, invocamos a definição para a derivada do ângulo $\theta_i(t)$ com relação ao tempo t .

Portanto, de acordo com a Eq. (4.1), podemos interpretar o sinal modulado em ângulo $s(t)$ como um fasor girante de tamanho A_c e ângulo $\theta_i(t)$. A velocidade angular deste fasor é $d\theta_i(t)/dt$, medida em radianos por segundo. No caso simples de uma portadora não modulada, o ângulo $\theta_i(t)$ é

$$\theta_i(t) = 2\pi f_c t + \phi_c, \quad \text{para } m(t) = 0$$

e o fasor correspondente gira com velocidade angular constante igual a $2\pi f_c$ radianos por segundo. A constante ϕ_c define o ângulo da portadora não modulada no tempo $t = 0$.

Existe um número infinito de formas pelas quais o ângulo $\theta_i(t)$ pode ser variado de alguma forma pelo sinal de mensagem. Entretanto, iremos considerar apenas dois métodos geralmente utilizados, a modulação em fase e a modulação em freqüência, definidos a seguir:

1. A modulação em fase (PM, do termo em inglês) é a forma de modulação em ângulo na qual o ângulo instantâneo $\theta_i(t)$ é variado linearmente com o sinal de mensagem $m(t)$, mostrado por

$$\theta_i(t) = 2\pi f_c t + k_p m(t) \tag{4.3}$$

O termo $2\pi f_c t$ representa o ângulo da portadora não modulada com a constante ϕ_c igual a zero por conveniência de representação. A constante k_p representa o fator de sensibilidade de fase do modulador, expresso em radianos por volt na condição de $m(t)$ ser uma forma de onda de tensão. A onda modulada em fase $s(t)$ é descrita no domínio do tempo por

$$s(t) = A_c \cos[2\pi f_c t + k_p m(t)] \tag{4.4}$$

2. A modulação em freqüência (FM) é a forma de modulação em ângulo na qual a freqüência instantânea $f_i(t)$ é variada linearmente com o sinal de mensagem $m(t)$, mostrado por

$$f_i(t) = f_c + k_f m(t) \tag{4.5}$$

O termo constante f_c representa a freqüência da portadora não modulada, a constante k_f representa o fator de sensibilidade de freqüência do modulador, expresso em hertz por volt com a condição de $m(t)$ ser uma forma de onda de tensão. Integrando a Eq. (4.5) com relação ao tempo e multiplicando o resultado por 2π , temos

$$\begin{aligned}
\theta_i(t) &= 2\pi \int_0^t f_i(\tau) d\tau \\
&= 2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau
\end{aligned} \tag{4.6}$$

na qual o segundo termo representa o aumento ou diminuição na fase instantânea $\theta_i(t)$ devido ao sinal de mensagem $m(t)$. A onda modulada em freqüência é, portanto,

$$s(t) = A_c \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right] \tag{4.7}$$

A Tabela 4.1 resume as definições básicas englobadas na geração de ondas moduladas em ângulo. Estas definições se aplicam a todos os sinais de mensagem, sejam eles analógicos ou digitais.

TABELA 4.1 Resumo das definições básicas na modulação em ângulo

	Modulação em fase	Modulação em freqüência	Comentários
Fase instantânea $\theta_i(t)$	$2\pi f_c t + k_p m(t)$	$2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau$	A_c : Amplitude da portadora f_c : Freqüência da portadora $m(t)$: Sinal de mensagem k_p : Fator de sensibilidade de fase k_f : Fator de sensibilidade de freqüência
Freqüência instantânea $f_i(t)$	$f_c + \frac{k_p}{2\pi} \frac{d}{dt} m(t)$	$f_c + k_f m(t)$	
Onda modulada $s(t)$	$A_c \cos[2\pi f_c t + k_p m(t)]$	$A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau\right]$	

4.2 PROPRIEDADES DE ONDAS MODULADAS EM ÂNGULO

Ondas moduladas em ângulo são caracterizadas por algumas propriedades importantes, as quais seguem das definições básicas resumidas na Tabela 4.1. De fato, são estas propriedades que fazem das ondas moduladas em ângulo uma família específica e distinta da família de ondas moduladas em amplitude, como mostrado na Fig. 4.1 para o exemplo de modulação senoidal. As Figs. 4.1(a) e 4.1(b) são a portadora senoidal e a onda modulante, respectivamente. As Figs. 4.1(c), 4.1(d) e 4.1(e) mostram as ondas modulada em amplitude (AM), modulada em fase (PM) e modulada em freqüência (FM), respectivamente.

- **PROPRIEDADE 1 – POTÊNCIA TRANSMITIDA CONSTANTE** Tanto da Eq. (4.4) quanto da Eq. (4.7), vemos facilmente que a amplitude das ondas PM e FM é mantida em um valor constante igual a amplitude da portadora A_c para todo tempo, independentemente dos fatores de sensibilidade k_p e k_f . Esta propriedade é bem demonstrada pela onda PM da Fig. 4.1(d) e pela onda FM da Fig. 4.1(e). Conseqüentemente, a potência média transmitida das ondas moduladas em ângulo é constante, mostrada por

$$P_{\text{máx.}} = \frac{1}{2} A_c^2 \quad (4.8)$$

na qual considera-se que o resistor de carga é 1 ohm.

- **PROPRIEDADE 2 – NÃO LINEARIDADE DO PROCESSO DE MODULAÇÃO** Outra característica distinta da modulação em ângulo é seu caráter não linear. Dizemos isso porque tanto as ondas PM quanto FM violam o princípio da superposição. Suponha, por exemplo, que o sinal de mensagem $m(t)$ é obtido por duas componentes diferentes $m_1(t)$ e $m_2(t)$, como mostrado por

$$m(t) = m_1(t) + m_2(t)$$

Seja $s(t)$, $s_1(t)$ e $s_2(t)$ as ondas PM produzidas por $m(t)$, $m_1(t)$ e $m_2(t)$, de acordo com a Eq. (4.4), respectivamente. Utilizando esta equação, podemos expressar estas ondas PM por:

$$\begin{aligned} s(t) &= A_c \cos[2\pi f_c t + k_p(m_1(t) + m_2(t))] \\ s_1(t) &= A_c \cos[2\pi f_c t + k_p m_1(t)] \end{aligned}$$

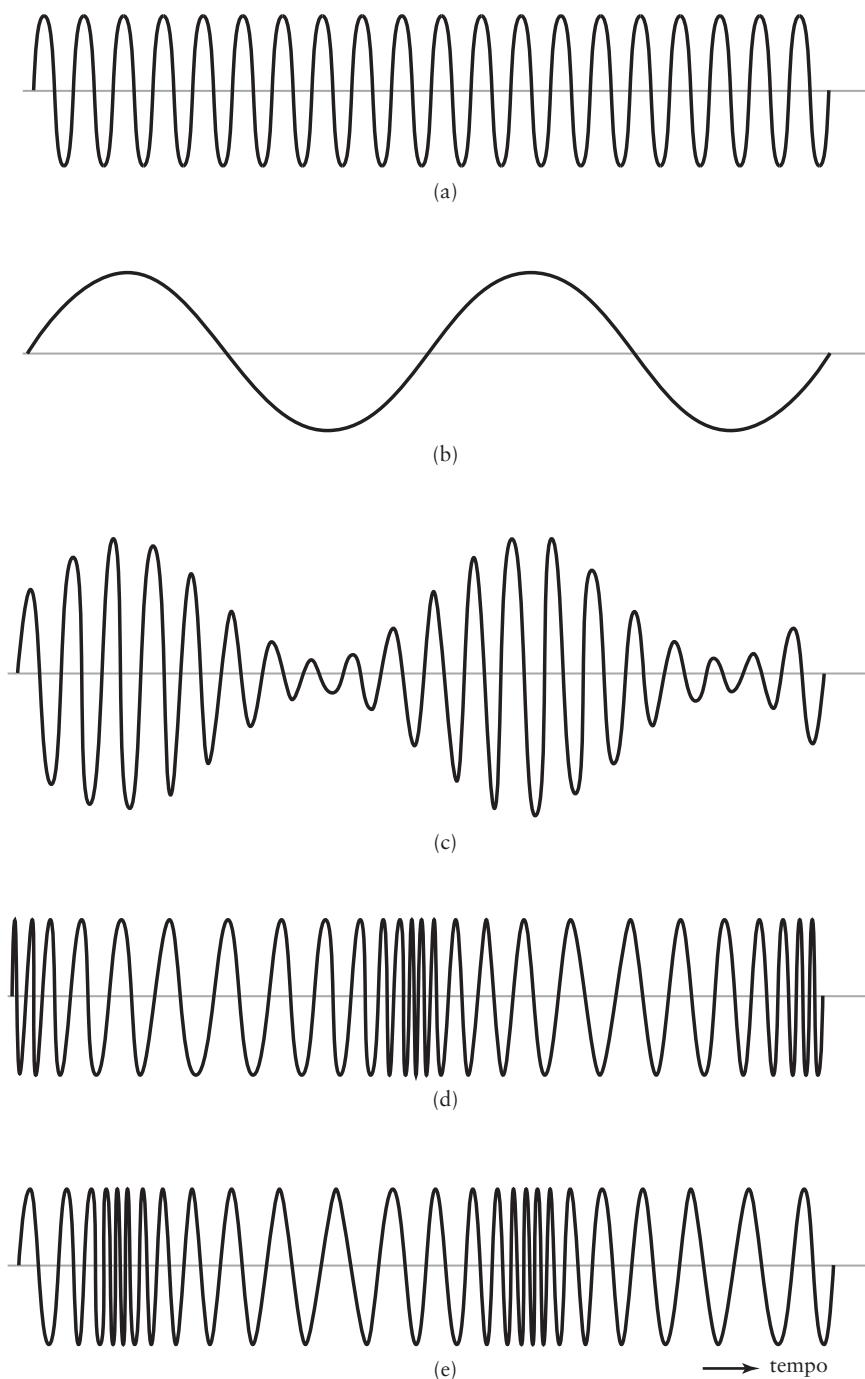


Figura 4.1 Ilustração de ondas AM, PM, FM produzidas por um tom único. (a) Onda portadora. (b) Sinal modulante senoidal. (c) Sinal modulado em amplitude. (d) Sinal modulado em fase. (e) Sinal modulado em freqüência.

e

$$s_2(t) = A_c \cos[2\pi f_c t + k_p m_2(t)]$$

A partir destas expressões, apesar do fato de $m(t) = m_1(t) + m_2(t)$, facilmente vemos que o princípio da superposição é violado, pois

$$s(t) \neq s_1(t) + s_2(t)$$

- **PROBLEMA 4.1** Utilizando a Eq. (4.7), mostre que as ondas FM também violam o princípio da superposição. ◀

O fato do processo de modulação em ângulo ser não linear complica a análise espectral e a análise de ruído de ondas PM e FM, comparado com a modulação em amplitude. Pelo mesmo motivo, o processo de modulação em ângulo possui benefícios próprios. Por exemplo, a modulação em freqüência oferece uma performance superior de ruído quando comparado com a modulação em amplitude, o que é atribuído ao caráter não linear da modulação em freqüência.

- **PROPRIEDADE 3 – IRREGULARIDADE NO CRUZAMENTO DE ZERO** Uma consequência de permitir que o ângulo instantâneo $\theta_i(t)$ se torne dependente do sinal de mensagem $m(t)$ como na Eq. (4.3) ou de sua integral $\int_0^t m(\tau) d\tau$ como na Eq. (4.6) é que, em geral, o cruzamento no zero de uma onda PM ou FM não possui regularidade perfeita em seu espaçamento ao longo da escala de tempo. Cruzamentos por zero são definidos como os instantes de tempo nos quais a forma de onda muda sua amplitude de um valor positivo para negativo ou vice-versa. De uma forma, a irregularidade no cruzamento de zero em ondas moduladas em ângulo também é atribuída ao caráter não linear do processo de modulação. Para ilustrar esta propriedade, podemos contrastar a onda PM da Fig. 4.1(d) e a onda FM da Fig. 4.1(c) com a onda AM correspondente.

Entretanto, podemos citar dois casos especiais nos quais a regularidade é mantida na modulação em ângulo:

1. O sinal de mensagem $m(t)$ aumenta ou diminui linearmente com o tempo t , neste caso, a freqüência instantânea $f_i(t)$ da onda PM muda da freqüência f_c da portadora não modulada para uma nova constante dependente da inclinação de $m(t)$.
2. O sinal de mensagem $m(t)$ é mantido em algum valor constante, positivo ou negativo, neste caso, a freqüência instantânea $f_i(t)$ da onda FM muda da freqüência f_c da portadora não modulada para uma nova constante dependente do valor constante de $m(t)$.

Em qualquer evento, é importante notar que na modulação em ângulo, o conteúdo da informação do sinal de mensagem $m(t)$ reside nos cruzamentos por zero da onda modulada. Esta afirmativa é válida desde que a freqüência f_c da portadora seja grande quando comparada com a maior componente de freqüência do sinal de mensagem $m(t)$.

- **PROPRIEDADE 4 – DIFICULDADE DE VISUALIZAÇÃO DA FORMA DE ONDA DA MENSAGEM** No AM, vemos a forma de onda da mensagem como o envelope da onda modulada, desde que o percentual de modulação seja menor do que 100%, como ilustrado na Fig. 4.1(c) para a modulação senoidal. Esta visualização não ocorre na modulação em ângulo, como ilustrado nas formas de onda correspondentes das Figs. 4.1(d) e 4.1(e), para PM e FM, respectivamente. Em geral, a dificuldade na visualização da forma de onda da mensagem em ondas moduladas em ângulo também é atribuída ao caráter não linear das ondas moduladas em ângulo.

- **PROPRIEDADE 5 – COMPROMISSO ENTRE O AUMENTO DA LARGURA DE FAIXA DE TRANSMISSÃO PARA A MELHORIA DA PERFORMANCE DO RUÍDO** Uma importante vantagem da modulação em ângulo sobre a modulação em amplitude é a realização de uma melhoria na performance do ruído. Esta vantagem é atribuída ao fato da transmissão de um sinal de mensagem pela modulação em ângulo de uma portadora senoidal ser menos sensível à presença de ruído aditivo do que

na transmissão por modulação em amplitude da portadora. A melhoria na performance do ruído é, entretanto, obtida ao preço do aumento correspondente no quesito de largura de faixa da transmissão da modulação em ângulo. Em outras palavras, a utilização da modulação em ângulo oferece a possibilidade de trocar o aumento da largura de faixa de transmissão por uma melhoria na performance de ruído. Tal compromisso não é possível na modulação em amplitude, pois a largura de faixa de transmissão de uma onda modulada em amplitude é mantida fixa entre a largura de faixa W da mensagem e $2W$, dependendo do tipo de modulação utilizado. O efeito do ruído na modulação em ângulo é discutido no Capítulo 9.

EXEMPLO 4.1 Cruzamentos por zero

Considere a onda modulante $m(t)$ que aumenta linearmente com o tempo t , começando em $t = 0$, como mostrado por

$$m(t) = \begin{cases} at, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

na qual a é o parâmetro de inclinação (veja a Fig. 4.2(a)). Desta forma, estudamos os cruzamentos por zero das ondas PM e FM produzidas por $m(t)$ para o seguinte conjunto de parâmetros:

$$f_c = \frac{1}{4} \text{ Hz}$$

$$a = 1 \text{ volt/s}$$

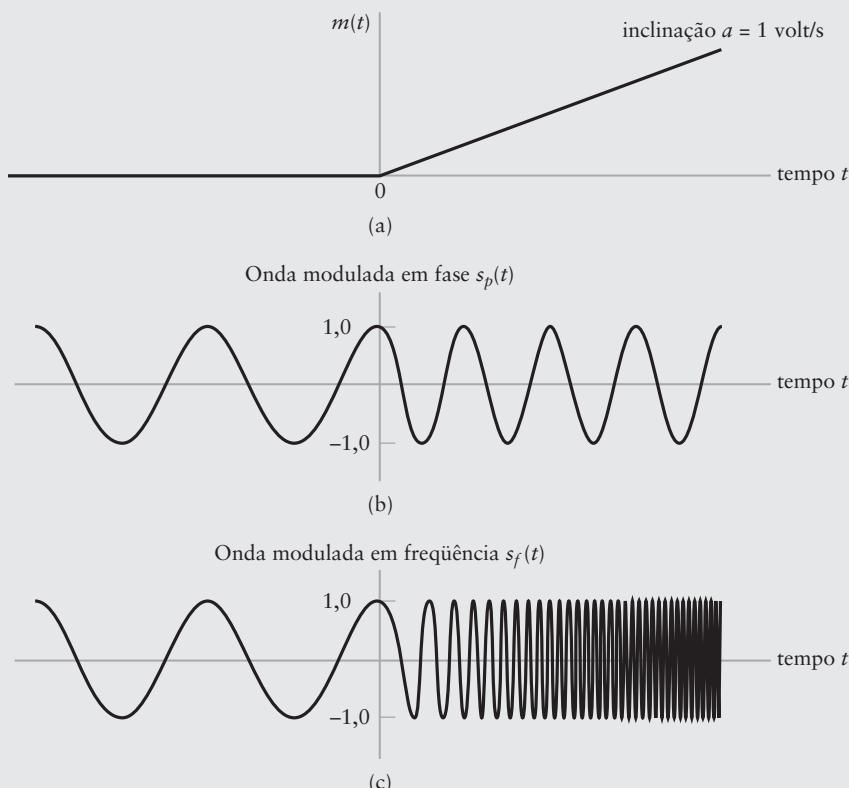


Figura 4.2 Começando no tempo $t = 0$, a figura mostra (a) aumento linear do sinal de mensagem $m(t)$, (b) onda modulada em fase e (c) onda modulada em freqüência.

- 1. Modulação em fase:** fator de sensibilidade de fase $k_p = \pi/2$ radianos/volt. Aplicando a Eq. (4.4) a $m(t)$, temos a onda PM

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t + k_p a t), & t \geq 0 \\ A_c \cos(2\pi f_c t), & t < 0 \end{cases}$$

a qual é mostrada na Fig. 4.2(b) para $A_c = 1$ volt.

Seja t_n o instante de tempo no qual a onda PM vivencia um cruzamento no zero. Isto ocorre sempre que o ângulo da onda PM for um múltiplo ímpar de $\pi/2$. Então, podemos fazer

$$2\pi f_c t_n + k_p a t_n = \frac{\pi}{2} + n\pi, \quad n = 0, 1, 2, \dots$$

como equação linear para t_n . Resolvendo esta expressão para t_n , obtemos a fórmula linear

$$t_n = \frac{\frac{1}{2} + n}{2f_c + \frac{k_p}{\pi} a}$$

Substituindo os valores dados de f_c , a e k_p nesta fórmula linear, obtemos

$$t_n = \frac{1}{2} + n, \quad n = 0, 1, 2, \dots$$

na qual t_n é medido em segundos.

- 2. Modulação em freqüência:** fator de sensibilidade de freqüência $k_f = 1$ Hz/volt. Aplicando a Eq. (4.7), obtemos a onda FM

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t + \pi k_f a t^2), & t \geq 0 \\ A_c \cos(2\pi f_c t), & t < 0 \end{cases}$$

a qual é mostrada na Fig. 4.2(c).

Invocando a definição de um cruzamento no zero, podemos fazer

$$2\pi f_c t_n + \pi k_f a t_n^2 = \frac{\pi}{2} + n\pi, \quad n = 0, 1, 2, \dots$$

como equação quadrática para t_n . A raiz positiva desta equação, ou seja,

$$t_n = \frac{1}{ak_f} \left(-f_c + \sqrt{f_c^2 + ak_f \left(\frac{1}{2} + n \right)} \right), \quad n = 0, 1, 2, \dots$$

define a fórmula para t_n . Substituindo os valores dados de f_c , a e k_f nesta equação quadrática, temos

$$t_n = \frac{1}{4} \left(-1 + \sqrt{9 + 16n} \right), \quad n = 0, 1, 2, \dots$$

na qual t_n novamente é medido em segundos.

Comparando os resultados de cruzamento por zero obtidos para ondas PM e FM, podemos fazer as seguintes observações uma vez que a onda modulante linear começa a agir na onda portadora senoidal:

- Para PM, a regularidade no cruzamento de zero é mantida. A freqüência instantânea varia do valor não modulado $f_c = 1/4$ Hz para o novo valor constante de

$$f_c + k_p(a/2\pi) = \frac{1}{2} \text{Hz.}$$

- Para FM, os cruzamentos no zero assumem uma forma irregular. Como esperado, a freqüência instantânea aumenta linearmente com o tempo t .

As formas de ondas moduladas em ângulo da Fig. 4.2 devem ser contrastadas com as correspondentes da Fig. 4.1. Enquanto no caso da modulação senoidal mostrada na Fig. 4.1 é difícil observar a diferença entre PM e FM, o mesmo não acontece no caso da Fig. 4.2. Em outras palavras, dependendo da onda modulante, é possível que PM e FM exibam formas de onda totalmente diferentes.

4.3 RELAÇÃO ENTRE ONDAS PM E FM

Examinando as definições das Eqs. (4.4) e (4.7), vemos que a onda FM pode ser vista como uma onda PM produzida pela onda modulante $\int_0^t m(\tau) d\tau$ no lugar de $m(t)$. Isto significa que a onda FM pode ser gerada integrando, primeiro, o sinal de mensagem $m(t)$ com relação ao tempo t e, então, utilizando o sinal resultante como a entrada de um modulador de fase, como mostrado na Fig. 4.3(a).

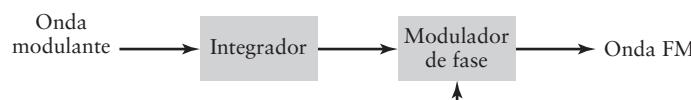
Alternativamente, uma onda PM pode ser vista com uma onda FM produzida pela onda modulante $dm(t)/dt$. Logo, a onda PM pode ser gerada diferenciando, primeiro, $m(t)$ com relação ao tempo t e, então, utilizando o sinal resultante como entrada de um modulador FM, como mostrado na Fig. 4.3(b).

Portanto, temos que a modulação em fase e a modulação em freqüência são unicamente relacionadas uma com a outra. Esta relação, por sua vez, significa que podemos deduzir as propriedades da modulação em fase das propriedades da modulação em freqüência e vice-versa. Por esta razão, neste capítulo iremos focalizar grande parte da discussão na modulação em freqüência.

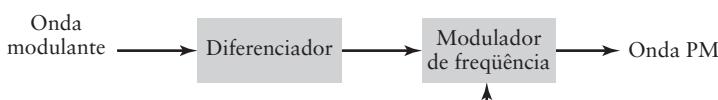
- **PROBLEMA 4.2** O esquema mostrado na Fig. 4.3(a) fornece a base para a geração indireta de uma onda FM. O modulador de fase é definido pela Eq. (4.4). Mostre que se a onda resultante deve estar exatamente na forma definida pela Eq. (4.7), então o fator de sensibilidade de fase k_p do modulador de fase é relacionado com o fator de sensibilidade de freqüência k_f da Eq. (4.7) pela fórmula

$$k_p = 2\pi k_f T$$

na qual T é o intervalo no qual a integração da Fig. 4.3(a) é executada. Justifique a dimensionalidade desta expressão. ◀



(a)



(b)

Figura 4.3 Ilustração da relação entre a modulação em freqüência e a modulação em fase. (a) Esquema para a geração da onda FM utilizando um modulador de fase. (b) Esquema para a geração de uma onda PM utilizando um modulador em freqüência.

4.4 MODULAÇÃO EM FREQÜÊNCIA DE FAIXA ESTREITA

Na Seção 4.2 ressaltamos o fato da onda FM ser uma função não linear da onda modulante. Esta propriedade faz com que a análise espectral da onda FM seja uma tarefa muito mais difícil do que para a onda AM correspondente.

Como, então, podemos fazer a análise espectral da onda FM? Propomos fornecer uma resposta empírica a esta importante questão procedendo da seguinte maneira:

- ▶ Consideraremos, primeiro, o caso simples de uma modulação de tom único que produz uma onda FM de faixa estreita.
- ▶ A seguir consideraremos o caso mais geral envolvendo também a modulação de tom único, mas desta vez a onda FM é faixa larga.

Podemos, obviamente, seguir em frente e considerar o caso mais elaborado de uma onda FM multiton. Entretanto, propomos não fazê-lo, pois nosso objetivo imediato é estabelecer uma relação empírica entre a largura de faixa de transmissão da onda FM e a largura de faixa da mensagem. Como iremos ver subsequentemente, a análise espectral em dois passos descrita anteriormente nos fornecerá material necessário para propor uma solução útil ao problema.

Considere, então, a onda modulante senoidal definida por

$$m(t) = A_m \cos(2\pi f_m t) \quad (4.9)$$

A freqüência instantânea da onda FM resultante é

$$\begin{aligned} f_i(t) &= f_c + k_f A_m \cos(2\pi f_m t) \\ &= f_c + \Delta f \cos(2\pi f_m t) \end{aligned} \quad (4.10)$$

na qual

$$\Delta f = k_f A_m \quad (4.11)$$

A grandeza Δf é chamada de *desvio de freqüência*, representando a diferença máxima da freqüência instantânea da onda FM da freqüência f_c da portadora. A característica fundamental da modulação em freqüência senoidal é que o desvio de freqüência Δf é proporcional à amplitude do sinal modulante e independente da freqüência da modulante.

Usando a Eq. (4.10) na primeira linha da Eq. (4.6), o ângulo $\theta_i(t)$ da onda FM é obtida como

$$\theta_i(t) = 2\pi f_c t + \frac{\Delta f}{f_m} \operatorname{sen}(2\pi f_m t) \quad (4.12)$$

A razão do desvio de freqüência Δf pela freqüência da modulante f_m é geralmente chamada de *índice de modulação* da onda FM. Representamos este novo parâmetro por β , escrevendo

$$\beta = \frac{\Delta f}{f_m} \quad (4.13)$$

e

$$\theta_i(t) = 2\pi f_c t + \beta \operatorname{sen}(2\pi f_m t) \quad (4.14)$$

A partir da Eq. (4.14), vemos que, em um sentido físico, o parâmetro β representa o desvio de fase da onda FM – ou seja, a distância máxima do ângulo $\theta_i(t)$ do ângulo $2\pi f_c t$ da portadora não modulada. Logo, β é medido em radianos.

A onda FM é dada por

$$s(t) = A_c \cos[2\pi f_c t + \beta \operatorname{sen}(2\pi f_m t)] \quad (4.15)$$

Para a onda FM $s(t)$ da Eq. (4.15) ser faixa estreita, o índice de modulação β deve ser pequeno quando comparado a um radiano. Para continuarmos, utilizamos a propriedade trigonométrica

$$\cos(A + B) = \cos A \cos B - \sin A \sin B$$

para expandir a Eq. (4.15) em

$$s(t) = A_c \cos(2\pi f_c t) \cos[\beta \sin(2\pi f_m t)] - A_c \sin(2\pi f_c t) \sin[\beta \sin(2\pi f_m t)] \quad (4.16)$$

Então, considerando a condição de que o índice de modulação β é pequeno quando comparado a um radiano, podemos utilizar as duas seguintes aproximações para todo tempo t :

$$\cos[\beta \sin(2\pi f_m t)] \approx 1$$

e

$$\sin[\beta \sin(2\pi f_m t)] \approx \beta \sin(2\pi f_m t)$$

Desta forma, a Eq. (4.16) é reduzida para

$$s(t) \approx A_c \cos(2\pi f_c t) - \beta A_c \sin(2\pi f_c t) \sin(2\pi f_m t) \quad (4.17)$$

A Eq. (4.17) define uma *forma aproximada* de onda FM de faixa estreita produzida pela onda modulante senoidal $A_m \cos(2\pi f_m t)$. Desta representação aproximada, podemos deduzir o modulador mostrado no diagrama em blocos da Fig. 4.4. Este modulador envolve a divisão da portadora $A_c \cos(2\pi f_c t)$ em dois caminhos. Um caminho é direto, o outro caminho contém um circuito deslocador de fase de -90° e um modulador de produto. A combinação deles gera uma onda modulada DSB-SC. A diferença entre estes dois sinais produz uma onda FM de faixa estreita, mas com alguma *distorção em amplitude*, como discutido a seguir.

Idealmente, a onda FM possui envelope constante e, para o caso de sinal modulante de freqüência f_m , o ângulo $\theta_i(t)$ também é senoidal com mesma freqüência. Mas a onda modulada produzida pelo modulador de faixa estreita da Fig. 4.4 difere desta condição ideal por dois pontos fundamentais:

1. O envelope contém uma modulação em amplitude residual que varia com o tempo.
2. O ângulo $\theta_i(t)$ contém distorção harmônica na forma de harmônicas de terceira ordem e superiores da freqüência de modulação f_m .

► **PROBLEMA 4.3** A representação Cartesiana de sinais passa-faixa discutida na Seção 3.8 é adequada para esquemas de modulação lineares exemplificados pela família de modulação em amplitude. Por outro lado, a representação polar

$$s(t) = a(t) \cos[2\pi f_c t + \phi(t)]$$

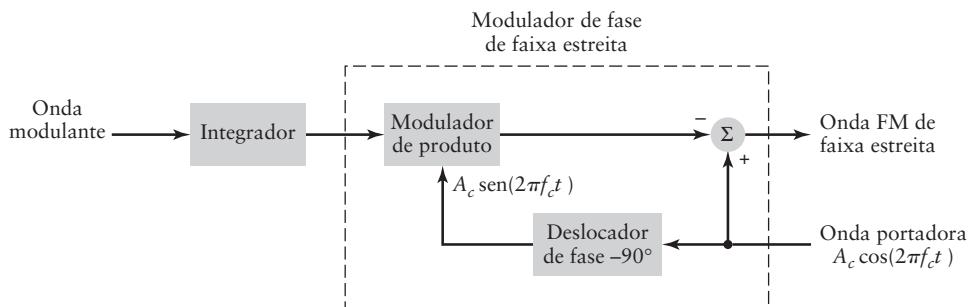


Figura 4.4 Diagrama em blocos para um método indireto de geração de uma onda FM de faixa estreita.

é adequada para esquemas de modulação não linear exemplificados pela família de modulação em ângulo. $a(t)$ nesta nova representação é o envelope de $s(t)$ e $\phi(t)$ é sua fase.

Começando com a representação [veja a Eq. (3.39)]

$$s(t) = s_I(t) \cos(2\pi f_c t) - s_Q(t) \sin(2\pi f_c t)$$

na qual $s_I(t)$ é a componente em fase e $s_Q(t)$ é a componente em quadratura, podemos escrever

$$a(t) = [s_I^2(t) + s_Q^2(t)]^{1/2}$$

e

$$\phi(t) = \operatorname{tg}^{-1} \left[\frac{s_Q(t)}{s_I(t)} \right]$$

Mostre que a representação polar de $s(t)$ em termos de $a(t)$ e $\phi(t)$ é exatamente equivalente a sua representação Cartesiana em termos de $s_I(t)$ e $s_Q(t)$. ◀

► **PROBLEMA 4.4** Considere a onda FM de faixa estreita aproximadamente definida pela Eq. (4.17). Utilizando o Problema 4.3:

- (a) Determine o envelope desta onda modulada. Qual é a razão do máximo ao mínimo valor deste envelope?
- (b) Determine a potência média da onda FM de faixa estreita, expressada como um percentual da potência média da onda portadora não modulada.
- (c) Expandindo o argumento angular $\theta(t) = 2\pi f_c t + \phi(t)$ da onda FM $s(t)$ de faixa estreita na forma de uma série de potências, e restringindo o índice de modulação β ao valor máximo de 0,3 radianos, mostre que

$$\theta(t) \approx 2\pi f_c t + \beta \sin(2\pi f_m t) - \frac{\beta^3}{3} \sin^3(2\pi f_m t)$$

Qual é o valor da distorção harmônica para $\beta = 0,3$ radianos?

Dica: Para x pequeno, a seguinte aproximação por série de potência

$$\operatorname{tg}^{-1}(x) \approx x - \frac{1}{3}x^3$$

é válida. Nesta aproximação, os termos a partir de x^5 são ignorados, o que é justificado quando x é pequeno se comparado à unidade. ◀

O ponto importante de nota do Problema 4.4 é que, restringindo-se o índice de modulação a $\beta \leq 0,3$ radianos, os efeitos da modulação em amplitude residual e a distorção harmônica são limitados a valores negligenciáveis. Estamos, portanto, capacitados a continuar a utilizar a Eq. (4.17) desde que $\beta \leq 0,3$ radianos. Em particular, podemos expandir a onda modulada em três componentes de freqüência:

$$s(t) \approx A_c \cos(2\pi f_c t) + \frac{1}{2} \beta A_c \{ \cos[2\pi(f_c + f_m)t] - \cos[2\pi(f_c - f_m)t] \} \quad (4.18)$$

Esta expressão é, de alguma forma, similar à equação correspondente na definição da onda AM, a qual é reproduzida do Exemplo 3.1 do Capítulo 3:

$$s_{\text{AM}}(t) = A_c \cos(2\pi f_c t) + \frac{1}{2} \mu A_c \{ \cos[2\pi(f_c + f_m)t] + \cos[2\pi(f_c - f_m)t] \} \quad (4.19)$$

na qual μ é o fator de modulação do sinal AM. Comparando as Eqs. (4.18) e (4.19) e ignorando as constantes respectivas β e μ , vemos que no caso da modulação senoidal, a diferença básica entre a onda AM e a onda FM de faixa estreita está no fato do sinal algébrico da freqüência lateral inferior na onda FM de faixa estreita ser inverso. De qualquer forma, a onda FM de faixa estreita requer, essencialmente, a mesma largura de faixa de transmissão (isto é, $2f_m$ para modulação senoidal) da onda AM.

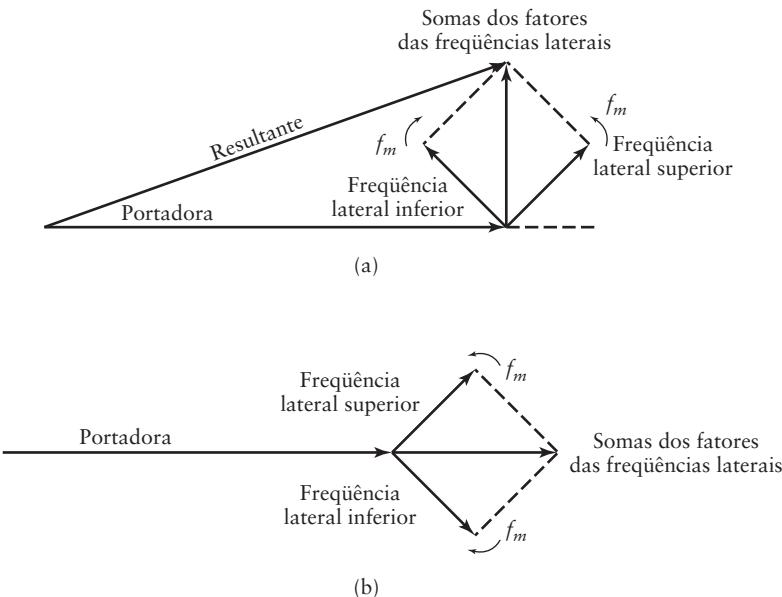


Figura 4.5 Comparação fasorial das ondas FM de faixa estreita e AM para modulação senoidal. (a) Onda FM de faixa estreita. (b) Onda AM.

■ Interpretação fasorial

Podemos representar a onda FM pelo diagrama fasorial mostrado na Fig. 4.5(a), no qual utilizamos o fasor da portadora como referência. Vemos que o resultante dos dois fatores das freqüências laterais está sempre em 90° com o fasor da portadora. O efeito desta geometria é a produção de um fasor resultante representando uma onda FM de faixa estreita aproximadamente de mesma amplitude do fasor da portadora, mas fora de fase com relação a ela.

O diagrama fasorial para a onda FM deve ser contrastado com o da Fig. 4.5(b) representando a onda AM correspondente. Neste último caso, vemos que o fasor resultante representando a onda AM possui amplitude diferente do fasor da portadora, mas sempre em fase com ele.

Apesar de tanto a onda FM de faixa estreita da Eq. (4.18) quanto a onda AM da Eq. (4.19) possuírem três componentes senoidais, as duas partes da Fig. 4.5 ilustram claramente as principais diferenças entre estas duas ondas moduladas. As diferenças são atribuídas à forma pela qual estas duas ondas moduladas são geradas.

4.5 MODULAÇÃO EM FREQÜÊNCIA DE FAIXA LARGA

A seguir queremos determinar o espectro de um onda FM de tom único definida pela fórmula exata da Eq. (4.15) para um valor arbitrário de índice de modulação β . Em geral, tal onda FM produzida por uma onda modulante senoidal é uma função periódica do tempo t somente quando a freqüência f_c da portadora é um múltiplo inteiro da freqüência modulante f_m .

- **PROBLEMA 4.5** Falando estritamente, a onda FM da Eq. (4.15) produzida por uma onda modulante senoidal é uma função não periódica do tempo t . Demonstre esta propriedade da modulação em freqüência. ◀

À luz deste problema, como podemos simplificar a análise espectral da onda FM de faixa larga definida na Eq. (4.15)? A resposta está na utilização da representação banda base complexa do sinal modulado (isto é, banda passante), a qual foi discutida na Seção 3.8. Especificamente, assuma que a freqüência f_c da portadora é grande o suficiente (comparada com a largura de faixa da onda FM) para justificar a escrita da Eq. (4.15) na forma

$$\begin{aligned}s(t) &= \operatorname{Re}[A_c \exp(j2\pi f_c t + j\beta \operatorname{sen}(2\pi f_m t))] \\ &= \operatorname{Re}[\tilde{s}(t) \exp(j2\pi f_c t)]\end{aligned}\quad (4.20)$$

na qual o operador $\operatorname{Re}[\]$ extrai a parte real da grandeza complexa contida dentro dos colchetes. O novo termo

$$\tilde{s}(t) = A_c \exp[j\beta \operatorname{sen}(2\pi f_m t)] \quad (4.21)$$

introduzido na Eq. (4.21) é o *envelope complexo* da onda FM $s(t)$. O ponto importante a notar da Eq. (4.21) é que, ao contrário da onda FM original $s(t)$, o envelope complexo $\tilde{s}(t)$ é uma função periódica do tempo com freqüência fundamental igual à freqüência f_m da modulante. Especificamente, substituindo o tempo t na Eq. (4.21) por $t + k/f_m$ para algum inteiro k , temos

$$\begin{aligned}\tilde{s}(t) &= A_c \exp[j\beta \operatorname{sen}(2\pi f_m(t + k/f_m))] \\ &= A_c \exp[j\beta \operatorname{sen}(2\pi f_m t + 2k\pi)] \\ &= A_c \exp[j\beta \operatorname{sen}(2\pi f_m t)]\end{aligned}$$

o que confirma f_m como a freqüência fundamental de $\tilde{s}(t)$. Podemos, portanto, expandir $\tilde{s}(t)$ na forma da série complexa de Fourier por:

$$\tilde{s}(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_m t) \quad (4.22)$$

na qual o coeficiente complexo de Fourier

$$\begin{aligned}c_n &= f_m \int_{-1/(2f_m)}^{1/(2f_m)} \tilde{s}(t) \exp(-j2\pi n f_m t) dt \\ &= f_m A_c \int_{-1/(2f_m)}^{1/(2f_m)} \exp[j\beta \operatorname{sen}(2\pi f_m t) - j2\pi n f_m t] dt\end{aligned}\quad (4.23)$$

define a nova variável:

$$x = 2\pi f_m t \quad (4.24)$$

Logo, podemos redefinir o coeficiente complexo de Fourier c_n da Eq. (4.23) na nova forma

$$c_n = \frac{A_c}{2\pi} \int_{-\pi}^{\pi} \exp[j(\beta \operatorname{sen} x - nx)] dx \quad (4.25)$$

A integral no lado direito da Eq. (4.25), exceto pela amplitude A_c da portadora, é referido como a *n-ésima ordem da função de Bessel* de primeiro tipo e argumento β . Esta função é geralmente representada pelo símbolo $J_n(\beta)$, portanto, podemos escrever

$$J_n(\beta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp[j(\beta \operatorname{sen} x - nx)] dx \quad (4.26)$$

Desta forma, podemos reescrever a Eq. (4.25) na forma compacta

$$c_n = A_c J_n(\beta) \quad (4.27)$$

Substituindo a Eq. (4.27) na Eq. (4.22), temos, em termos da função de Bessel $J_n(\beta)$, a seguinte expansão para o envelope complexo da onda FM:

$$\tilde{s}(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \exp(j2\pi n f_m t) \quad (4.28)$$

A seguir, substituindo a Eq. (4.28) na Eq. (4.20), temos

$$s(t) = \operatorname{Re} \left[A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \exp[j2\pi(f_c + nf_m)t] \right] \quad (4.29)$$

A amplitude A_c da portadora é constante e, portanto, pode ser retirada do operador de parte real $\operatorname{Re}[\cdot]$. Além disso, podemos alterar a ordem do somatório e da operação de parte real, pois os dois são operadores lineares. Desta forma, podemos reescrever a Eq. (4.29) na forma simplificada

$$s(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \cos[2\pi(f_c + nf_m)t] \quad (4.30)$$

A Eq. (4.30) é a forma desejada para a expansão em série de Fourier do sinal FM $s(t)$ de tom único para um valor arbitrário de índice de modulação β .

O espectro discreto de $s(t)$ é obtido calculando as transformadas de Fourier dos dois lados da Eq. (4.30), o que resulta em

$$S(f) = \frac{A_c}{2} \sum_{n=-\infty}^{\infty} J_n(\beta) [\delta(f - f_c - nf_m) + \delta(f + f_c + nf_m)] \quad (4.31)$$

na qual $s(t) \Leftrightarrow S(f)$ e $\cos(2\pi f_i t) \Leftrightarrow \frac{1}{2}[\delta(f - f_i) + \delta(f + f_i)]$ para um f_i arbitrário. A Eq. (4.31) mostra que o espectro de $s(t)$ é constituído por um número infinito de funções delta espaçadas em $f = f_c \pm nf_m$ para $n = 0, +1, +2, \dots$

■ Propriedades da FM de tom único para um índice de modulação β arbitrário

Na Fig. 4.6, traçamos a função de Bessel $J_n(\beta)$ em função do índice de modulação β para valores inteiros positivos diferentes de n . Podemos obter mais conhecimento sobre o comportamento da função de Bessel $J_n(\beta)$ utilizando as seguintes propriedades (veja o Apêndice 3 para mais detalhes):

1. Para valores inteiros (positivos e negativos) diferentes de n , temos

$$J_n(\beta) = J_{-n}(\beta), \quad \text{para } n \text{ par} \quad (4.32)$$

e

$$J_n(\beta) = -J_{-n}(\beta), \quad \text{para } n \text{ ímpar} \quad (4.33)$$

2. Para valores pequenos de índice de modulação β , temos

$$\left. \begin{array}{l} J_0(\beta) \approx 1, \\ J_1(\beta) \approx \frac{\beta}{2}, \\ J_n(\beta) \approx 0, \quad n > 2 \end{array} \right\} \quad (4.34)$$

3. A igualdade

$$\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1 \quad (4.35)$$

é válida exatamente para β arbitrário.

Portanto, utilizando as Eqs. (4.31) a (4.35) e as curvas da Fig. 4.6, podemos fazer as seguintes observações:

1. O espectro de uma onda FM contém a componente da portadora e um conjunto infinito de freqüência laterais localizadas simetricamente nos dois lados da freqüência

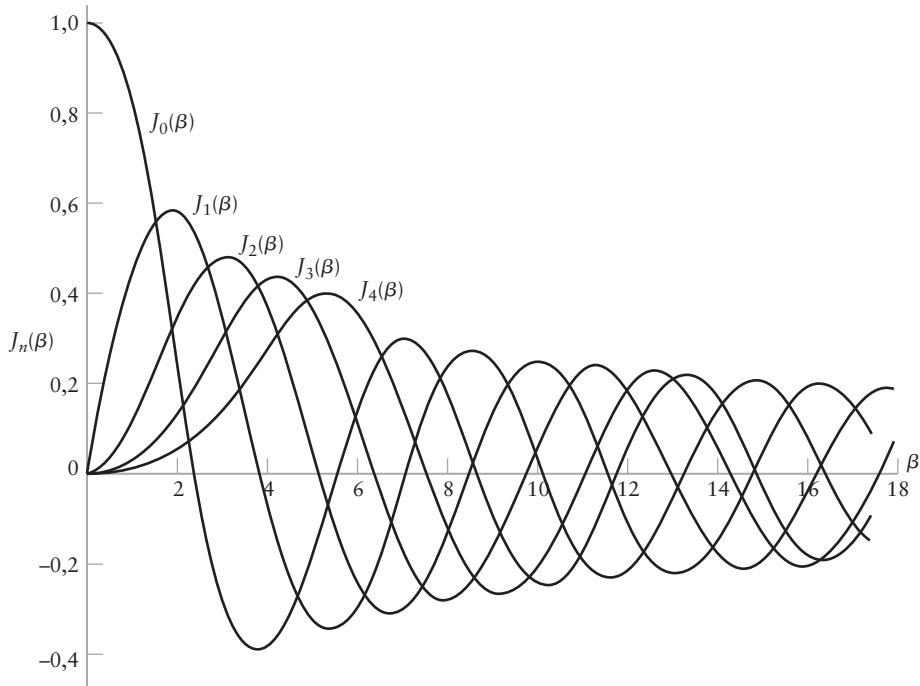


Figura 4.6 Gráfico da função de Bessel de primeiro tipo, $J_n(\beta)$, variando-se a ordem n .

da portadora, separadas por f_m , $2f_m$, $3f_m$, ... Com relação a este fato, o resultado é diferente da figura que prevalece em AM, pois no último caso uma onda modulante senoidal resulta em apenas um par de freqüências laterais.

2. Para o caso especial de β pequeno, comparado com a unidade, apenas os coeficiente de Bessel $J_0(\beta)$ e $J_1(\beta)$ possuem valores significativos, portanto a onda FM é efetivamente composta da portadora e um único par de freqüências laterais em $f_c \pm f_m$. Esta situação corresponde ao caso especial de FM de faixa estreita considerado na Seção 4.4
3. A amplitude da componente portadora varia com β de acordo com $J_0(\beta)$. Ou seja, ao contrário da onda AM, a amplitude da componente portadora da onda FM é dependente do índice de modulação β . A explicação física para esta propriedade é que o envelope da onda FM é constante, portanto a potência média de tal sinal desenvolvido em um resistor de 1 ohm também é constante, como na Eq. (4.8), a qual é reproduzida aqui por conveniência:

$$P_{\text{méd.}} = \frac{1}{2} A_c^2$$

Quando a portadora é modulada para gerar a onda FM, a potência das freqüências laterais podem aparecer apenas ao custo da potência original da portadora, fazendo, portanto, com que a componente da portadora seja dependente de β . Note que a potência média de uma onda FM também pode ser determinada da Eq. (4.30), mostrado por

$$P = \frac{1}{2} A_c^2 \sum_{n=-\infty}^{\infty} J_n^2(\beta) \quad (4.36)$$

Substituindo a Eq. (4.35) na Eq. (4.36), a expressão para a potência média $P_{\text{méd.}}$ se reduz para a Eq. (4.8), como deve ser.

EXEMPLO 4.2 Espectro FM para variação de amplitude e freqüência de uma onda modulante senoidal

Neste exemplo, queremos investigar as formas nas quais a variação de amplitude e freqüência de uma onda modulante senoidal afetam o espectro da onda FM. Considere primeiro o caso quando a freqüência da onda modulante é fixa, mas sua amplitude varia, produzindo uma variação correspondente no desvio de freqüência Δf . Portanto, mantendo a freqüência modulante f_m fixa, descobrimos que o espectro de amplitude da onda FM resultante é como mostrado na Fig. 4.7 para $\beta = 1, 2$ e 5 . Neste diagrama, normalizamos o espectro com relação à amplitude da portadora não modulada.

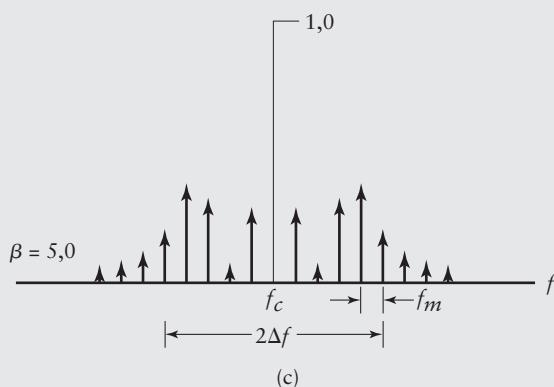
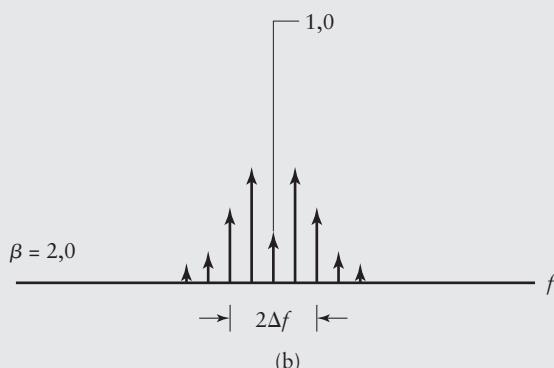
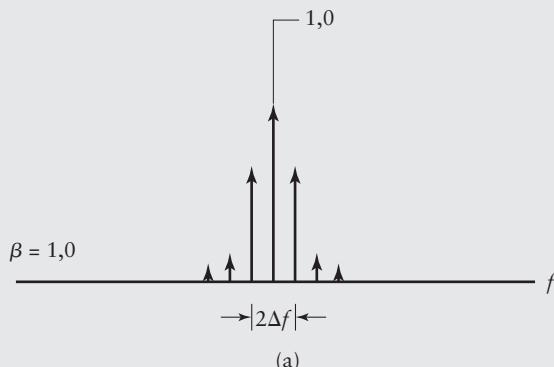


Figura 4.7 Espectro de amplitude discreto da onda FM, normalizada com respeito a amplitude da portadora não modulada, para o caso de modulação senoidal de freqüência fixa e amplitude variante. Apenas o espectro para freqüências positivas é mostrado.

Considere, a seguir, o caso quando a amplitude da onda modulante é fixada, ou seja, o desvio de freqüência Δf é mantido constante, e a freqüência de modulação f_m é variada. Neste segundo caso, obtemos o espectro de amplitude da onda FM resultante como mostrado na Fig. 4.8 para $\beta = 1, 2$ e 5 . Vemos, agora, que quando Δf é mantido fixo e β aumenta, temos um aumento no número de linhas espetrais aparecendo dentro do intervalo fixo de freqüência $f_c - \Delta f < |f| < f_c + \Delta f$. Ou seja, quando β tende ao infinito, a largura de faixa da onda FM se aproxima do valor limite de $2\Delta f$, o que é um importante ponto a ser lembrado.

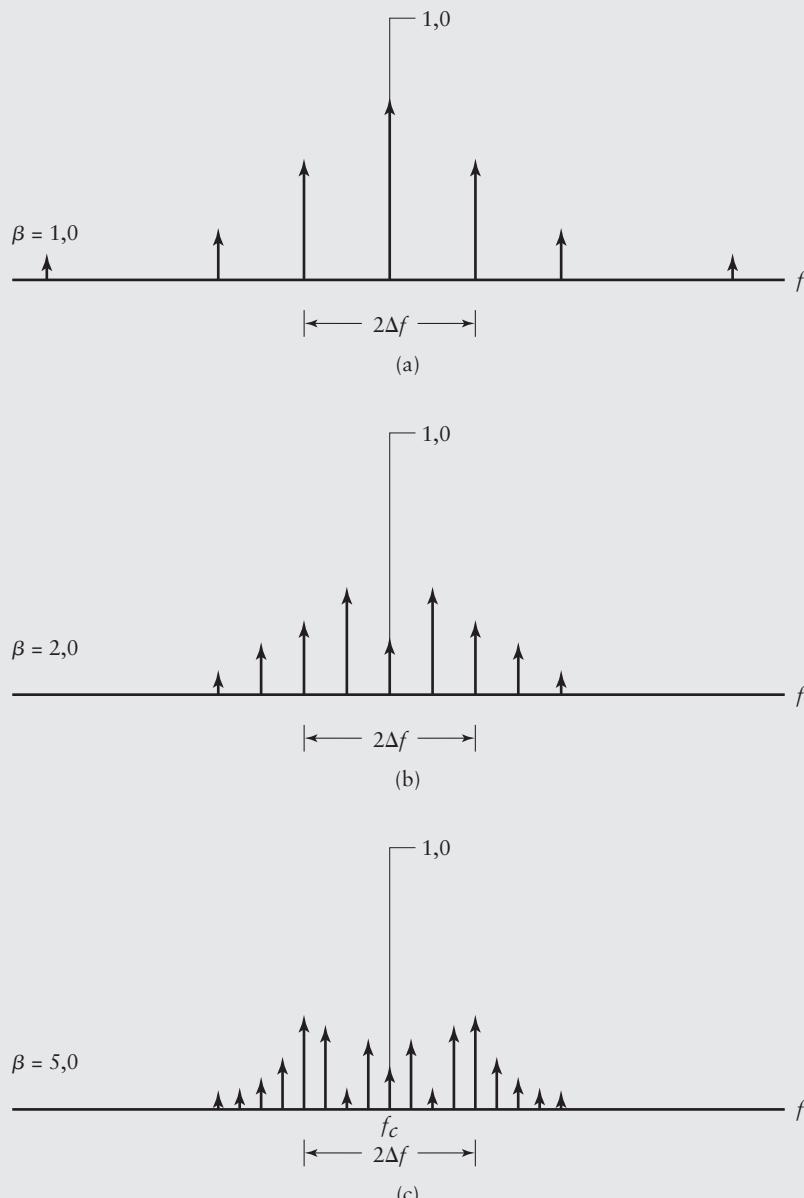


Figura 4.8 Espectro de amplitude discreto da onda FM, normalizado com respeito à amplitude da onda portadora não modulada, para o caso de modulação senoidal com freqüência variante e amplitude fixa. Apenas o espectro para freqüências positivas é mostrado.

4.6 LARGURA DE FAIXA DE TRANSMISSÃO DE ONDAS FM

■ Regra de Carson

Em teoria, a onda FM contém um número infinito de freqüências laterais, de tal forma que a largura de faixa necessária para transmitir esse tipo de onda modulada é aparentemente infinita em extensão. Na prática, entretanto, descobrimos que a onda FM é *efetivamente limitada a um número de freqüências laterais significativas compatíveis com o total especificado de distorção*. Podemos, portanto, utilizar esta idéia para especificar uma largura de faixa efetiva necessária para a transmissão de uma onda FM. Considere, primeiro, o caso de uma onda FM gerada por uma onda modulante de tom único e freqüência f_m . Nesta onda FM, as freqüências laterais que são separadas da freqüência portadora f_c por um total maior do que o desvio de freqüência Δf diminuem rapidamente em direção a zero, de tal forma que a largura de faixa sempre excede a excursão total de freqüência, mas, de qualquer forma, é limitada. Especificamente, podemos identificar dois casos limite:

1. Para grandes valores do índice de modulação β , a largura de faixa se aproxima, sendo um pouco maior apenas, do que a excursão total de freqüência $2\Delta f$, como ilustrado na Fig. 4.8(c).
2. Para pequenos valores do índice de modulação β , o espectro da onda FM é efetivamente limitado à freqüência da portadora f_c e um par de freqüências laterais em $f_c \pm f_m$, de tal forma que a largura de faixa se aproxima de $2f_m$, como ilustrado na Seção 4.4.

Considerando estes dois cenários, podemos definir uma regra aproximada para a largura de faixa de transmissão de uma onda FM gerada por uma onda modulante de tom único e freqüência f_m por

$$B_T \approx 2\Delta f + 2f_m = 2\Delta f \left(1 + \frac{1}{\beta} \right) \quad (4.37)$$

Esta relação empírica simples é chamada de *Regra de Carson*.

■ Curva universal para a largura de faixa de transmissão FM

A regra de Carson é simples de ser utilizada, mas, infelizmente, nem sempre fornece uma boa estimativa das necessidades de largura de faixa de sistemas de comunicação usando modulação em freqüência de faixa larga. Para uma medição mais adequada da largura de faixa FM, podemos utilizar a definição baseada na manutenção do número máximo de freqüências laterais significativas cujas amplitudes sejam todas maiores do que um valor selecionado. Uma escolha conveniente para este valor é 1% da amplitude da portadora não modulada. Podemos, portanto, definir a *largura de faixa de uma onda FM como a separação entre duas freqüências além das quais nenhuma freqüência lateral é maior do que 1% da amplitude da portadora obtida quando o sinal modulante é removido*. Ou seja, definimos a largura de faixa de transmissão como $2n_{\max} f_m$, na qual f_m é a freqüência modulante e n_{\max} é o maior valor do inteiro n que satisfaz a condição $|J_n(\beta)| > 0,01$. O valor de n_{\max} varia com o índice de modulação β e pode ser determinado facilmente dos valores tabulados da função de Bessel $J_n(\beta)$. A Tabela 4.2 mostra o número total de freqüências laterais significativas (incluindo tanto a freqüência lateral superior quanto a inferior) para diferentes valores de β , calculado com base no 1%. A largura de faixa de transmissão B_T calculada utilizando este procedimento pode ser apresentada na forma da curva universal normalizada com respeito ao desvio de freqüência Δf , traçando em função de β . Esta curva está mostrada na Fig. 4.9, a qual é traçada como a melhor aproximação para um conjunto de pontos obtidos usando a Tabela 4.2. Na Fig. 4.9, observamos que quando o índice de modulação é aumentado, a largura de faixa ocupada pelas freqüências laterais diminui em direção ao valor no qual a freqüência da portadora realmente desvia. Isto significa que os pequenos valores do índice de modulação β são relativamente mais extravagantes na largura de faixa de transmissão do que grandes valores de β .

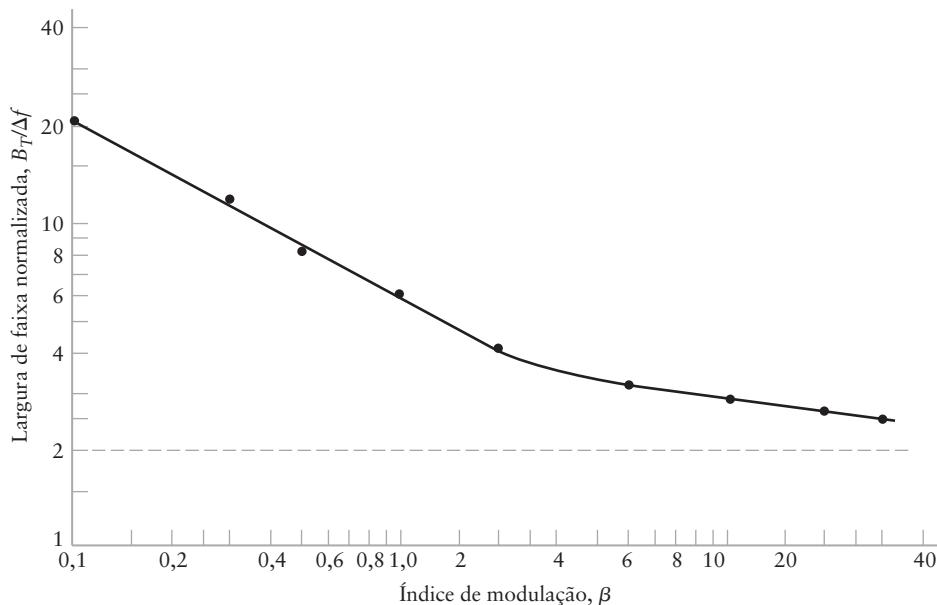


Figura 4.9 Curva universal para o cálculo da largura de faixa de 1% da onda FM.

Onda modulante arbitrária

Considere, a seguir, o caso mais geral de uma onda modulante arbitrária $m(t)$ com sua componente de freqüência mais alta representada por W , ou seja, W representa a largura de faixa da mensagem. Temos, agora, uma situação mais difícil de se trabalhar. Especificamente, a largura de faixa necessária para transmitir a onda gerada por uma onda modulante arbitrária é baseada na análise de pior caso de modulação por tom. Primeiro determinamos a chamada *razão de desvio* D , definida como a relação do desvio de freqüência Δf , o qual corresponde a máxima amplitude possível da onda modulante $m(t)$, para a mais alta freqüência modulante W . Esta condição representa o caso mais extremo possível. Podemos escrever, formalmente,

$$D = \frac{\Delta f}{W} \quad (4.38)$$

A razão de desvio D possui o mesmo papel para modulação não senoidal que o índice de modulação β possui para o caso de modulação senoidal. Logo, substituindo β por D e substituindo f_m por W , podemos generalizar a Eq. (4.7) como mostrado a seguir:

$$B_T = 2(\Delta f + W) \quad (4.39)$$

TABELA 4.2 Número de freqüências laterais significativas de um sinal FM de faixa larga para a variação do índice de modulação

Índice de modulação β	Número de freqüências laterais significativas $2n_{\max}$
0,1	2
0,3	4
0,5	4
1,0	6
2,0	8
5,0	16
10,0	28
20,0	50
30,0	70

Desta forma, nos referimos à Eq. (4.39) como a *regra de Carson generalizada* para a largura de faixa de transmissão de um sinal FM arbitrário. De forma similar, podemos generalizar a curva universal da Fig. 4.9 para obter um valor para a largura de faixa de transmissão do sinal FM. De um ponto de vista prático, a regra de Carson generalizada de alguma forma subestima a necessidade de largura de faixa de um sistema FM, enquanto que, de forma correspondente, a utilização da curva universal da Fig. 4.9 resulta em um resultado conservativo. Portanto, a escolha da largura de faixa de transmissão que estiver entre os limites fornecidos pelas duas regras é aceitável na maior parte dos casos práticos.

EXEMPLO 4.3 Transmissão FM comercial

Na América do Norte, o valor máximo de desvio de freqüência Δf é limitado em 75 kHz para a transmissão FM comercial por broadcasting. Se considerarmos a freqüência de modulação $W = 15$ kHz, a qual é normalmente o “máximo” de freqüência de áudio de interesse na transmissão FM, descobrimos que o valor correspondente da razão de desvio é [usando a Eq. (4.38)]

$$D = \frac{75}{15} = 5$$

Usando os valores $\Delta f = 75$ kHz e $D = 5$ na regra de Carson generalizada da Eq. (4.39), obtemos que o valor aproximado da largura de faixa de transmissão do sinal FM é igual a

$$B_T = 2(75 + 15) = 180 \text{ kHz}$$

Por outro lado, utilizando a curva universal da Fig. 4.9, a largura de faixa de transmissão do sinal FM é

$$B_T = 3,2 \Delta f = 3,2 \times 75 = 240 \text{ kHz}$$

Neste exemplo, a regra de Carson subestima a largura de faixa de transmissão por 25% quando comparado com o resultado utilizando a curva universal da Fig. 4.9.

4.7 GERAÇÃO DE ONDAS FM

De acordo com a Eq. (4.5), a freqüência instantânea $f_i(t)$ da onda FM varia linearmente com o sinal de mensagem $m(t)$. Para o projeto de um *modulador de freqüência*, precisamos, portanto, de um dispositivo que produz um sinal de saída cuja freqüência instantânea é sensível a variações de amplitude de um sinal, de forma linear.

Existem dois métodos básicos para a geração de ondas moduladas em freqüência, um direto e outro indireto.

■ Método direto

O *método direto* utiliza um oscilador senoidal, com um dos elementos reativos (por exemplo, elemento capacitivo) no circuito tanque de um oscilador sendo diretamente controlado pelo sinal de mensagem. Em termos conceituais, o método direto é, portanto, diretamente implementado. Além disso, ele é capaz de fornecer grandes desvios de freqüência. Entretanto, uma limitação séria ao método direto é a tendência da freqüência da portadora desviar, o que geralmente é inaceitável para aplicações comerciais de rádio. Para resolver esta limitação, é necessário implementar uma estabilização de freqüência do gerador FM, o que é realizado através do uso de realimentação do oscilador. Veja o Prob. 4.15 para a descrição de tal procedimento. Apesar do oscilador ser simples de se construir, a utilização da estabilização de freqüência adiciona complexidade ao projeto do modulador de freqüência.

Método indireto: modulador Armstrong

No *método indireto*, por outro lado, o sinal de mensagem é primeiro utilizado para produzir uma FM de faixa estreita, seguido por um multiplicador de freqüência para aumentar o desvio de freqüência para níveis desejados. Neste segundo método, o problema da estabilidade da freqüência da portadora é reduzido pela utilização de um oscilador altamente estável (por exemplo, um oscilador a cristal) na geração do FM de faixa estreita. Este esquema de modulação é chamado de *modulador Armstrong de freqüência de faixa larga*, em reconhecimento ao seu inventor.

Um diagrama em blocos simplificado deste sistema FM indireto é mostrado na Fig. 4.10. O sinal de mensagem $m(t)$ é primeiro integrado e, então, utilizado para modular em fase um oscilador controlado a cristal. A utilização do controle do cristal fornece estabilidade de freqüência. Para minimizar a distorção inherente do modulador de fase, o desvio de fase máximo, ou índice de modulação β , é mantido pequeno propositalmente, resultando, portanto, em uma onda FM de faixa estreita. Para a implementação do modulador de fase de faixa estreita, podemos utilizar o arranjo descrito na Fig. 4.4. A onda FM de faixa estreita é, a seguir, multiplicada em freqüência pelo uso de um multiplicador de freqüência, para produzir a onda FM de faixa larga desejada.

Um *multiplicador de freqüência* é constituído por um elemento não linear *sem memória* seguido por um filtro passa-faixa, como mostrado na Fig. 4.11. A implicação do dispositivo não linear sendo sem memória é que ele não possui elementos de armazenamento de energia. A relação de entrada/saída de tal dispositivo pode ser expressa na forma geral

$$v(t) = a_1 s(t) + a_1 s^2(t) + \cdots + a_n s^n(t) \quad (4.40)$$

na qual a_1, a_2, \dots, a_n são coeficientes determinados pelo ponto de operação do dispositivo e n é a *mais alta ordem de não linearidade*. Em outras palavras, um dispositivo não linear sem memória é um dispositivo de lei de n -ésima potência.

A entrada $s(t)$ é uma onda FM definida por

$$s(t) = A_c \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right] \quad (4.41)$$

na qual a freqüência instantânea é

$$f_i(t) = f_c + k_f m(t) \quad (4.42)$$

Suponha que (1) a freqüência de meia faixa do filtro passa-faixa da Fig. 4.11 é ajustada para ser igual a $n f_c$, na qual f_c é a freqüência da portadora da onda FM $s(t)$ de entrada e (2) a faixa passante do filtro é projetada para ter a largura de faixa igual a n vezes a largura de faixa de transmissão de $s(t)$. No Problema 4.24, trabalhando com os efeitos não lineares

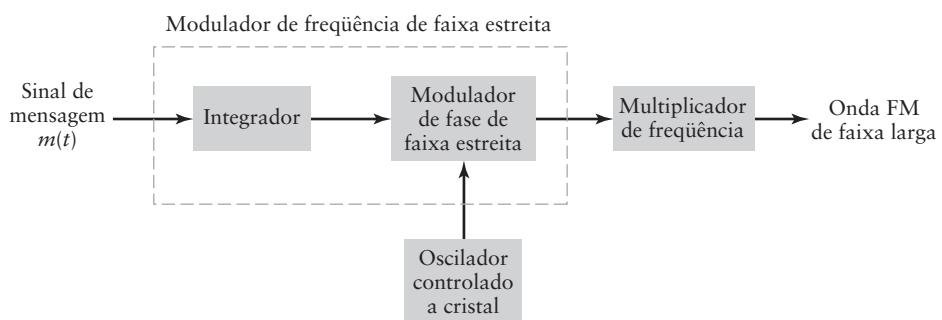


Figura 4.10 Diagrama em blocos do método indireto para a geração da onda FM de faixa larga.

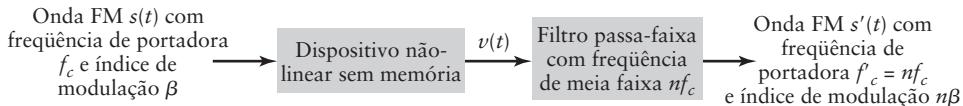


Figura 4.11 Diagrama em blocos de um multiplicador de freqüência.

de sistemas FM, lidamos com as contribuições espectrais de tais termos não lineares como termos de segunda e terceira ordem da relação entrada/saída da Eq. (4.40) Por enquanto, é suficiente dizer que após a filtragem passa-faixa da saída $v(t)$ do dispositivo não linear, teremos uma nova onda FM definida por

$$s'(t) = A_c \cos \left[2\pi f'_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right] \quad (4.43)$$

cuja freqüência instantânea é

$$f'_i(t) = n f_c + n k_f m(t) \quad (4.44)$$

Portanto, comparando a Eq. (4.44) com a (4.42), vemos que o subsistema não linear da Fig. 4.11 funciona com um multiplicador de freqüência com $f'_c = n f_c$ e $k'_f = n k_f$. A razão n de multiplicação é determinada pela mais alta potência n da relação de entrada/saída da Eq. (4.40), caracterizando o dispositivo não linear sem memória.

4.8 DEMODULAÇÃO DE SINAIS FM

A *demodulação de freqüência* é o processo pelo qual o sinal de mensagem original é recuperado de uma onda FM de entrada. Em outras palavras, a demodulação de freqüência é o inverso da modulação em freqüência. Com o modulador de freqüência sendo um dispositivo que produz um sinal cuja freqüência instantânea varia linearmente com a amplitude do sinal de mensagem de entrada, temos que, para a demodulação de freqüência, precisamos de um dispositivo cuja amplitude de saída seja sensível a variações na freqüência instantânea da onda FM de entrada, de forma linear também.

Iremos descrever, em seguida, dois dispositivos para a demodulação de freqüência. Um dispositivo é chamado de discriminador de freqüências, baseado na detecção de inclinação seguido por uma detecção de envelope. O outro dispositivo, chamado de PLL (*phase-locked loop*) executa a demodulação de freqüência de uma maneira indireta.

■ Discriminador de freqüência

Lembre-se que o sinal FM é dado por

$$s(t) = A_c \cos \left(2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right)$$

a qual é a Eq. (4.41), reproduzida por conveniência. A questão a ser resolvida é: Como podemos recuperar o sinal de mensagem $m(t)$ do sinal modulado $s(t)$? Podemos motivar a criação de um receptor para esta recuperação observando que se fizermos a derivada da Eq. (4.44) com relação ao tempo, obtemos

$$\frac{ds(t)}{dt} = -2\pi A_c [f_c + k_f m(t)] \operatorname{sen} \left(2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right) \quad (4.45)$$

Inspecionando a Eq. (4.45), observamos que a derivada é um sinal passa-faixa com modulação em amplitude definida pelo termo multiplicativo $[f_c + k_f m(t)]$. Conseqüente-

mente, se f_c for grande o suficiente tal que a portadora não seja sobremodulada, então podemos recuperar o sinal de mensagem $m(t)$ com um detector de envelope de forma similar ao descrito para sinais AM no Capítulo 3. Esta idéia fornece a motivação necessária para o *discriminador de freqüência*, o qual é basicamente um modulador constituído por um diferenciador seguido por um detector de envelope.

Entretanto, existem questões práticas relacionadas com a implementação do discriminador como descrito – particularmente, o diferenciador. No Capítulo 2, mostramos que a diferenciação corresponde a uma função de transferência linear no domínio da freqüência, ou seja,

$$\frac{d}{dt} \rightleftharpoons j2\pi f \quad (4.46)$$

na qual, como sempre, \rightleftharpoons representa a relação de transformação de Fourier. Em termos práticos, é difícil construir um circuito que possua a função de transferência equivalente ao lado direito da Eq. (4.46) para todas as freqüências. Em seu lugar, construímos um circuito que aproxima esta função de transferência na largura de faixa do sinal faixa-passante – em particular, para $f_c - (B_T/2) \leq |f| \leq f_c + (B_T/2)$, na qual B_T é a largura de faixa de transmissão do sinal FM $s(t)$ de entrada. Uma transformada de Fourier típica que satisfaz esta condição é descrita por

$$H_1(f) = \begin{cases} j2\pi[f - (f_c - B_T/2)], & f_c - (B_T/2) \leq |f| \leq f_c + (B_T/2) \\ 0, & \text{caso contrário} \end{cases} \quad (4.47)$$

A característica de transferência deste chamado *circuito de inclinação* é ilustrada na Fig. 4.12 para freqüências positivas. Um circuito de inclinação prático deve ter um ganho não unitário associado com a inclinação, mas, por simplicidade, assumimos que ele possui ganho unitário sem perda de generalização. O circuito não precisa ter, também, resposta nula fora da largura de faixa de transmissão, desde que o circuito seja precedido por um filtro passa-faixa centrado em f_c e com faixa passante B_T .

É mais simples continuarmos com a *representação banda base complexa* do processamento de sinal executado pelo discriminador. Especificamente, seguindo a teoria desta representação desenvolvida no Capítulo 3, determinamos que o envelope complexo do sinal FM $s(t)$ (reproduzido na página 187) é

$$\tilde{s}(t) = A_c \exp\left(j2\pi k_f \int_0^t m(\tau) d\tau\right) \quad (4.48)$$

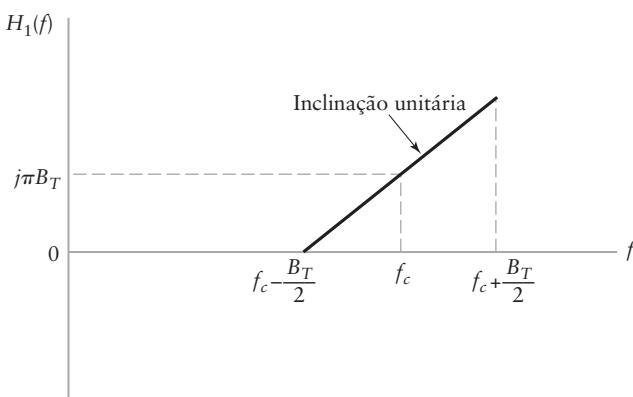


Figura 4.12 Resposta em freqüência de um circuito de inclinação ideal.

cuja aplicabilidade requer que a freqüência f_c da portadora seja grande quando comparada com B_T . De forma correspondente, podemos expressar o filtro banda base complexo (isto é, o circuito de inclinação) correspondente à Eq. (4.48) por

$$\tilde{H}_1(f) = \begin{cases} j2\pi[f + (B_T/2)], & -B_T/2 \leq f \leq B_T/2 \\ 0, & \text{caso contrário} \end{cases} \quad (4.49)$$

Seja $\tilde{s}_1(t)$ o envelope complexo da resposta do circuito de inclinação devido a $\tilde{s}(t)$. Então, de acordo com a transformação passa-faixa para passa-baixa descrita no Capítulo 3, podemos expressar a transformada de Fourier de $\tilde{s}_1(t)$ por

$$\begin{aligned} \tilde{S}_1(f) &= \frac{1}{2} \tilde{H}_1(f) \tilde{S}(f) \\ &= \begin{cases} j\pi \left(f + \frac{1}{2} B_T \right) \tilde{S}(f), & -\frac{1}{2} B_T \leq f \leq \frac{1}{2} B_T \\ 0, & \text{caso contrário} \end{cases} \end{aligned} \quad (4.50)$$

na qual $\tilde{S}(f)$ é a transformada de Fourier de $\tilde{s}(t)$. A razão da introdução do fator de multiplicação $\frac{1}{2}$ na primeira linha da Eq. (4.50) foi apresentada no Capítulo 3. Para determinar $\tilde{s}(t)$, a qual é a inversa de $\tilde{S}(f)$, invocamos duas propriedades pertinentes da transformada de Fourier, transcritas aqui (veja o Capítulo 2):

1. Multiplicação da transformada de Fourier $\tilde{S}(f)$ por $j2\pi f$ é equivalente à diferenciar a transformada inversa de Fourier $\tilde{s}(t)$ de acordo com a Propriedade 9, descrita na Eq. (2.33), mostrado por

$$\frac{d}{dt} \tilde{s}(t) \iff j2\pi f \tilde{S}(f)$$

2. Aplicação da propriedade de linearidade (isto é, Eq. (2.14)) à parte não nula de $\tilde{S}_1(f)$, resultando em

$$\tilde{s}_1(t) = \frac{1}{2} \frac{d}{dt} \tilde{s}(t) + \frac{1}{2} j\pi B_T \tilde{s}(t) \quad (4.51)$$

Substituindo a Eq. (4.48) na Eq. (4.51), temos

$$\tilde{s}_1(t) = \frac{1}{2} j\pi A_c B_T \left[1 + \left(\frac{2k_f}{B_T} \right) m(t) \right] \exp \left(j2\pi k_f \int_0^t m(\tau) d\tau \right) \quad (4.52)$$

Finalmente, a resposta atual do circuito de inclinação devido à onda FM $s(t)$ é dada por¹

$$\begin{aligned} s_1(t) &= \operatorname{Re}[\tilde{s}_1(t) \exp(j2\pi f_c t)] \\ &= \frac{1}{2} \pi A_c B_T \left[1 + \left(\frac{2k_f}{B_T} \right) m(t) \right] \cos \left(2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau + \frac{\pi}{2} \right) \end{aligned} \quad (4.53)$$

O próximo bloco funcional a ser considerado é o detector de envelope, o qual é alimentado por $s_1(t)$. A partir da Eq. (4.53), vemos que $s_1(t)$ é uma *onda modulada hibridamente*, exibindo tanto modulação em amplitude quanto modulação em freqüência do sinal de mensagem $m(t)$. Desde que a extensão da modulação em amplitude seja mantida, ou seja,

¹ Note que a primeira linha da Eq. (4.53) é uma repetição da Eq. (3.43) do Capítulo 3, a qual lida com a relação entre um sinal modulado $s(t)$ e sua representação complexa $\tilde{s}(t)$.

$$\left(\frac{2k_f}{B_T}\right)|m(t)|_{\max} < 1, \quad \text{para todo } t$$

então o detector de envelope recupera o sinal de mensagem $m(t)$, exceto por uma polarização (*bias*). Especificamente, sob condições ideais, a saída do detector de envelope é dada por

$$v_1(t) = \frac{1}{2}\pi A_c B_T \left[1 + \left(\frac{2k_f}{B_T}\right)m(t) \right] \quad (4.54)$$

A polarização em $v_1(t)$ é definida pelo termo constante da Eq. (4.54) – especificamente, $\pi A_c B_T / 2$.

Para remover a polarização, podemos utilizar um segundo circuito de inclinação seguido por um detector de envelope próprio. Desta vez, entretanto, projetamos o circuito de inclinação de tal forma que ele possua uma inclinação negativa. Desta forma, inferimos da Eq. (4.54) que a saída desta segunda configuração é dada por

$$v_2(t) = \frac{1}{2}\pi A_c B_T \left[1 - \left(\frac{2k_f}{B_T}\right)m(t) \right] \quad (4.55)$$

Assim, subtraindo a Eq. (4.55) da Eq. (4.54), obtemos uma *saída total sem polarização*, mostrada por

$$\begin{aligned} v(t) &= v_1(t) - v_2(t) \\ &= cm(t) \end{aligned} \quad (4.56)$$

na qual c é uma constante.

Utilizando as Eqs. (4.54) a (4.56), podemos construir o diagrama em blocos da Fig. 4.12 para o discriminador de freqüência ideal cuja composição é mostrada a seguir:

- O caminho superior da figura pertence à Eq. (4.54).
- O caminho inferior pertence à Eq. (4.55).
- O somador final representa a Eq. (4.56).

Este sistema particular de detecção é chamado de *discriminador balanceado de freqüência*, no qual o termo “balanceado” se refere ao fato de que os dois circuitos de inclinação são relacionados um com o outro da forma descrita pelas Eqs. (4.54) e (4.55).

De uma perspectiva prática, o desafio na implementação do discriminador balanceado de freqüência² da Fig. 4.13 é como construir os circuitos de inclinação de forma a satisfazer os requisitos de projeto das Eqs. (4.54) e (4.55).

■ Phase-locked loop (PLL)

A *phase-locked loop* (ou simplesmente PLL), é um sistema realimentado cuja operação é altamente associada com a modulação em freqüência. Ele é geralmente utilizado para a sincronização de portadora, e indireta demodulação de freqüência. Esta última aplicação é o objeto de interesse aqui.

Basicamente, o PLL é constituído de três principais componentes:

² Em Haykin, pp. 178-180, a realização prática do discriminador balanceado de freqüência é descrita utilizando um par de filtros RLC altamente ressonantes. Os dois filtros são projetados para ter um alto fator Q . O *fator de qualidade*, ou *fator Q*, de um filtro ressonante é uma medida de quão afiada a resposta em freqüência do filtro é. Ele é formalmente definido como 2π vezes a razão da máxima energia armazenada no filtro pela energia dissipada no filtro, os dois medidos por ciclo. Para esta aplicação, um filtro é sintonizado em uma freqüência acima da freqüência f_c da portadora não modulada e o outro filtro é sintonizado em uma freqüência abaixo de f_c . Fazendo o fator Q alto, a linearidade existente na porção necessária da resposta em freqüência total, centrada em f_c , é determinada pela separação das duas freqüências de ressonância.

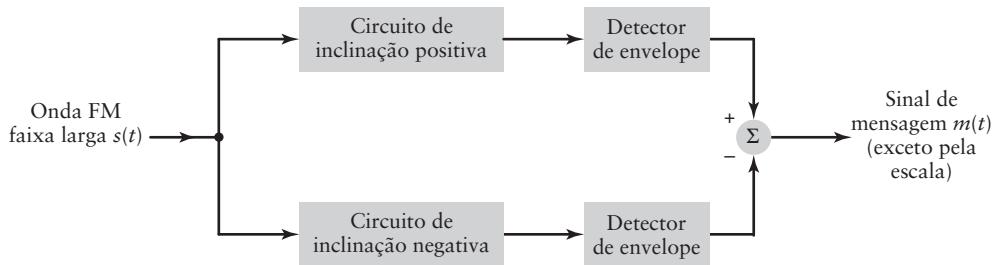


Figura 4.13 Diagrama em blocos do discriminador balanceado de freqüência.

- ▶ Oscilador controlado por tensão (VCO – do inglês *Voltage-Controlled Oscillator*), o qual executa a modulação em freqüência em seu próprio sinal de controle.
- ▶ Multiplicador, o qual multiplica a onda FM de entrada pela saída do oscilador controlado por tensão.
- ▶ Filtro de malha, do tipo passa-baixa, cuja função é remover as componentes de alta freqüência contidas no sinal de saída do multiplicador e, portanto, formatar a resposta de freqüência total do sistema.

Como mostrado no diagrama de blocos da Fig. 4.14, estes três componentes são conectados para formar um *sistema de malha fechada*.

Para demonstrar a operação do PLL como demodulador de freqüência, assumimos que o VCO foi ajustado de tal forma que quando o sinal de controle (isto é, a entrada) for zero, duas condições são satisfeitas:

1. A freqüência do VCO é ajustada precisamente na freqüência f_c da portadora não modulada da onda FM $s(t)$ de entrada.
2. A saída do VCO possui um deslocamento de fase de 90° com respeito a onda portadora não modulada.

Suponha, então, que a onda FM de entrada é definida por

$$s(t) = A_c \operatorname{sen}[2\pi f_c t + \phi_1(t)] \quad (4.57)$$

na qual A_c é a amplitude da portadora. Pela definição, o ângulo $\phi_1(t)$ é relacionado com o sinal de mensagem $m(t)$ pela integral

$$\phi_1(t) = 2\pi k_f \int_0^t m(\tau) d\tau \quad (4.58)$$

na qual k_f é o fator de sensibilidade de freqüência do modulador de freqüência responsável pela geração de $s(t)$. De forma correspondente, de acordo com os pontos (1) e (2), definimos a onda FM produzida pelo VCO como

$$r(t) = A_v \cos[2\pi f_c t + \phi_2(t)] \quad (4.59)$$

na qual A_v é a amplitude. O ângulo $\phi_2(t)$ é relacionado com o sinal de controle $v(t)$ do VCO pela integral

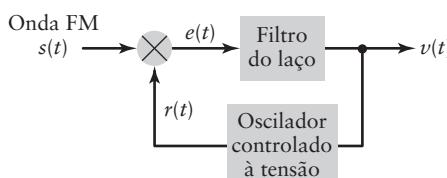


Figura 4.14 Diagrama em blocos do *phase-locked loop* (PLL).

$$\phi_2(t) = 2\pi k_v \int_0^t v(\tau) d\tau \quad (4.60)$$

na qual k_v é o fator de sensibilidade de freqüência do VCO.

A função da malha de realimentação no VCO é ajustar o ângulo $\phi_2(t)$ de tal forma que ele seja igual a $\phi_1(t)$, ajustando, portanto, o estágio para a demodulação de freqüência. Para aprofundarmos mais nesta função e como ela é realizada, precisamos desenvolver um modelo para o PLL, descrito a seguir.

Para isto, observamos, primeiro, que a multiplicação da onda FM $s(t)$ de entrada pela onda FM $r(t)$ gerada localmente produz duas componentes (exceto pelo fator de escala $\frac{1}{2}$):

1. A componente de alta freqüência, a qual é definida pelo *termo freqüência dobrada* – especificamente

$$k_m A_c A_v \sin[4\pi f_c t + \phi_1(t) + \phi_2(t)]$$

na qual k_m é o *ganho de multiplicação*.

2. Uma componente de baixa freqüência, a qual é definida pelo *termo de diferença de freqüência* – especificamente,

$$k_m A_c A_v \sin[\phi_1(t) - \phi_2(t)]$$

► **PROBLEMA 4.6** Utilizando uma identidade trigonométrica bem conhecida envolvendo o produto do seno de um ângulo e do cosseno de outro ângulo, demonstre os dois resultados descritos nos pontos 1 e 2. ◀

Com o filtro de malha projetado para suprimir as componentes de alta freqüência da saída do multiplicador, podemos descartar o termo de freqüência dobrada. Com isto, podemos reduzir o sinal aplicado ao filtro de malha a

$$e(t) = k_m A_c A_v \sin[\phi_e(t)] \quad (4.61)$$

na qual $\phi_e(t)$ é o *erro de fase*, definido por

$$\begin{aligned} \phi_e(t) &= \phi_1(t) - \phi_2(t) \\ &= \phi_1(t) - 2\pi k_v \int_0^t v(\tau) d\tau \end{aligned} \quad (4.62)$$

Quando o erro de fase $\phi_e(t)$ é zero, o PLL é dito estar *travado em fase*. É dito que ele está *quase travado em fase* quando o erro de fase $\phi_e(t)$ é pequeno quando comparado com um radiano, sob esta condição podemos utilizar a aproximação

$$\sin[\phi_e(t)] \approx \phi_e(t)$$

Esta aproximação está dentro de 4% desde que $\phi_e(t)$ seja menor do que 0,5 radianos. Desta forma, podemos aproximar o sinal de erro da Eq. (4.61) por

$$\begin{aligned} e(t) &\approx k_m A_c A_v \phi_e(t) \\ &= \frac{K_0}{k_v} \phi_e(t) \end{aligned} \quad (4.63)$$

na qual o novo parâmetro

$$K_0 = k_m k_v A_c A_v \quad (4.64)$$

é chamado de *parâmetro de ganho de malha* do PLL.

O sinal de erro $e(t)$ atua no filtro de malha para produzir a saída final $v(t)$. Seja $h(t)$ a resposta ao impulso do filtro de malha. Podemos, então, relacionar $v(t)$ com $e(t)$ pela integral de convolução

$$v(t) = \int_{-\infty}^{\infty} e(\tau)h(t - \tau) d\tau \quad (4.65)$$

As Eqs. (4.62), (4.63), (4.65) e (4.60), nesta ordem, constituem, o *modelo de realimentação linearizado* do PLL. O modelo está mostrado na Fig. 4.15(a) com o ângulo $\phi_1(t)$ da onda FM $s(t)$ de entrada atuando como entrada e a saída $v(t)$ do filtro de malha atuando como a saída final do PLL.

Da teoria de sistemas lineares, lembramos o importante teorema para sistemas realimentados:³

Quando a função de transferência de malha aberta de um sistema linear realimentado possui amplitude grande comparada com a unidade para todas as freqüências, a função de transferência de malha fechada do sistema é eficazmente determinada pelo inverso da função de transferência da malha de realimentação.

Dito de outra forma, a função de transferência de malha fechada do sistema realimentado se torna essencialmente independente da malha direta.

Do modelo realimentado linearizado da Fig. 4.15(a), observamos três pontos pertinentes ao problema:

1. A malha de realimentação é definida somente pelo integrador escalonado descrito pela Eq. (4.60), o qual é a contribuição do VCO para o modelo. Alternativamente, o inverso desta malha de realimentação é descrito no domínio do tempo pelo *diferenciador escalonado*

$$v(t) = \frac{1}{2\pi k_v} \left(\frac{d\phi_2(t)}{dt} \right) \quad (4.66)$$

2. O comportamento no domínio do tempo da malha fechada do PLL é descrito pela saída final $v(t)$ produzida em resposta ao ângulo $\phi_1(t)$ da onda FM $s(t)$ de entrada.
3. A amplitude da função de transferência de malha aberta do PLL é controlada pelo parâmetro de malha K_0 da Eq. (4.64).

Assumindo que o parâmetro de malha K_0 é grande quando comparado com a unidade, a aplicação do teorema de sistemas lineares realimentados ao modelo da Fig. 4.15(a) nos ensina que a função de transferência de malha fechada (isto é, o comportamento no domínio do tempo da malha fechada) do PLL é eficazmente determinada pelo inverso da função de transferência (isto é, comportamento no domínio do tempo) da malha de realimentação. Desta forma, utilizando o teorema de realimentação descrito anteriormente e a Eq. (4.66), podemos relacionar a saída final $v(t)$ com o ângulo de entrada $\phi_1(t)$ pela fórmula aproximada

$$v(t) \approx \frac{1}{2\pi k_v} \left(\frac{d\phi_1(t)}{dt} \right) \quad (4.67)$$

³ Considere o exemplo clássico de um *amplificador de realimentação negativa*, o qual é construído por dois componentes: um amplificador de ganho μ na malha direta e um ganho de circuito β na malha de realimentação. O *ganho de malha fechada do amplificador* é definido por

$$A = \frac{\mu}{1 + \mu\beta}$$

O termo em produto $\mu\beta$ no denominador é o *ganho de malha aberta do amplificador realimentado*. Quando $\mu\beta$ é grande quando comparado com a unidade, a fórmula para A é eficazmente determinada pelo inverso de β , como mostrado por

$$A \approx \frac{1}{\beta}$$

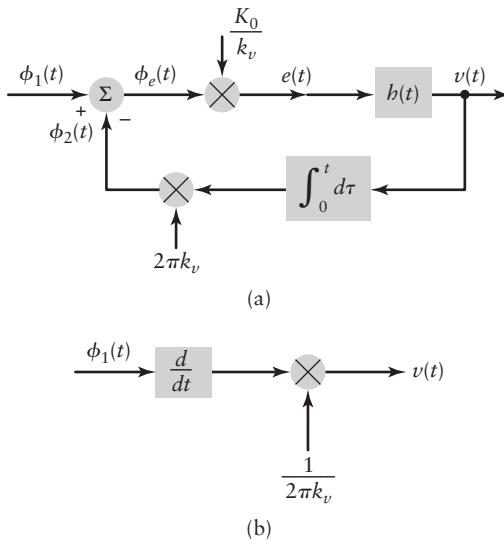


Figura 4.15 (a) Modelo linearizado do PLL. (b) Forma aproximada do modelo, assumindo que o ganho de malha K_0 é grande quando comparado com a unidade.

Permitir que K_0 assuma um valor grande possui o efeito de fazer com que o erro de fase $\phi_e(t)$ tenda a zero. Sob esta condição, temos $\phi_1(t) \approx \phi_2(t)$ de acordo com a primeira linha da Eq. (4.62). Esta condição de igualdade aproximada é o motivo da substituição de $\phi_2(t)$ por $\phi_1(t)$ na Eq. (4.67).

Com a relação aproximada descrita pela Eq. (4.67), podemos, agora, simplificar o modelo linearizado da Fig. 4.15(a) pela forma mostrada na parte (b) da figura. Logo, substituindo a Eq. (4.58) na Eq. (4.67), obtemos

$$\begin{aligned} v(t) &\approx \frac{1}{2\pi k_v} \cdot \frac{d}{dt} \left(2\pi k_f \int_0^t m(\tau) d\tau \right) \\ &= \frac{k_f}{k_v} m(t) \end{aligned} \quad (4.68)$$

A Eq. (4.68) afirma que quando o sistema opera no modo de fase travada ou quase fase travada e o parâmetro de ganho de malha K_0 é grande quando comparado com a unidade, a demodulação de freqüência da onda FM $s(t)$ de entrada é realizada, ou seja, o sinal de mensagem original $m(t)$ é recuperado de $s(t)$, exceto pelo fator de escala (k_f/k_v).

Uma importante característica do PLL, atuando como demodulador de freqüência, é que a largura de faixa da onda FM $s(t)$ de entrada pode ser muito maior do que a largura de faixa do filtro de malha caracterizado pela função de transferência $H(f)$; ou seja, a transformada da resposta ao impulso $h(t)$ do filtro de malha. A função de transferência $H(f)$ do filtro de malha pode e, portanto, deve ser restringida à banda base (isto é, à faixa de freqüência original ocupada pelo sinal de mensagem). Então o sinal de controle do VCO – ou seja, $v(t)$ – possui a largura de faixa da banda base (sinal de mensagem) $m(t)$, enquanto que a saída $r(t)$ do VCO é uma onda modulada em freqüência de faixa larga cuja freqüência instantânea rastreia as variações da freqüência instantânea da onda FM $s(t)$ de entrada devido a $m(t)$. Aqui estamos simplesmente reafirmando o fato de que a largura de faixa da onda FM de faixa larga é muito maior do que a largura de faixa do sinal de mensagem responsável por sua geração.

A complexidade do PLL é determinada pela função de transferência $H(f)$ do filtro de malha. A forma mais simples do PLL é obtida fazendo $H(f) = 1$, ou seja, sem filtro de malha.

Neste caso, o PLL é chamado de PLL de primeira ordem. Para circuitos de mais alta ordem, a função de transferência $H(f)$ assume uma forma mais complexa dependente da freqüência.

A principal limitação do PLL de primeira ordem é que o parâmetro de ganho de malha K_0 controla tanto a largura de faixa da malha quanto o limite da faixa de freqüências do circuito. O *limite da faixa de freqüências* se refere à faixa de freqüências para as quais a malha continua na condição de travada em fase com relação à onda FM de entrada. É por esta razão que, apesar de sua simplicidade, um PLL de primeira ordem é raramente utilizado na prática. Ao contrário, o procedimento recomendado é utilizar um PLL de segunda ordem, cuja implementação é satisfeita utilizando um filtro de malha de primeira ordem, veja o Problema 4.25.

- **PROBLEMA 4.7** Utilizando o modelo linearizado da Fig. 4.15(a), mostre que o modelo é aproximadamente governado pela equação integro-diferencial

$$\frac{d\phi_e(t)}{dt} + 2\pi K_0 \int_{-\infty}^{\infty} \phi_e(\tau) h(t - \tau) d\tau \approx \frac{d\phi_1(t)}{dt}$$

Logo, obtenha os dois resultados aproximados no domínio da freqüência:

$$(a) \Phi_e(f) \approx \frac{1}{1 + L(f)} \Phi_1(f)$$

$$(b) V(f) \approx \frac{jf}{k_v} \frac{L(f)}{1 + L(f)} \Phi_1(f)$$

nas quais

$$L(f) = K_0 \frac{H(f)}{jf}$$

é a função de transferência de malha aberta. Finalmente, mostre que quando $L(f)$ é grande quando comparado com a unidade para todas as freqüências dentro da faixa da mensagem, a versão no domínio do tempo da fórmula da parte (b) se reduz para a forma aproximada da Eq. (4.68). ◀

4.9 EXEMPLO DE TEMA: MULTIPLEXAÇÃO DE FM ESTÉREO

A *multiplexação estéreo* é uma forma de multiplexação por divisão de freqüência (FDM) projetada para transmitir dois sinais separados pela mesma portadora. Ela é amplamente utilizada na transmissão de rádio FM por broadcasting para enviar dois elementos diferentes de um programa (por exemplo, duas seções diferentes de uma orquestra, um vocalista e um acompanhante), fornecendo uma dimensão *especial* a um ouvinte no lado do receptor.

A especificação dos padrões para a transmissão de FM estéreo é influenciada por dois fatores:

1. A transmissão precisa operar dentro dos canais alocados para o broadcasting de FM.
2. Ela precisa ser compatível com receptores de rádio mono (não estéreos).

A primeira condição ajusta os parâmetros de freqüência permitidos, incluindo o desvio de freqüência. A segunda condição restringe a forma pela qual o sinal transmitido é configurado.

A Fig. 4.16(a) mostra o diagrama em blocos do sistema de multiplexação usado no transmissor de FM estéreo. Sejam $m_l(t)$ e $m_r(t)$ os sinais recebidos pelos microfones do lado esquerdo e direito do lado transmissor do sistema. Eles são aplicados a um *mixer* que gera o sinal soma $m_l(t) + m_r(t)$ e o sinal diferença $m_l(t) - m_r(t)$. O sinal soma é mantido não processado em sua forma banda base, disponibilizado para a recepção mono. O sinal diferença é uma subportadora de 38 kHz (obtida de um oscilador a cristal de 19 kHz pela multiplicação

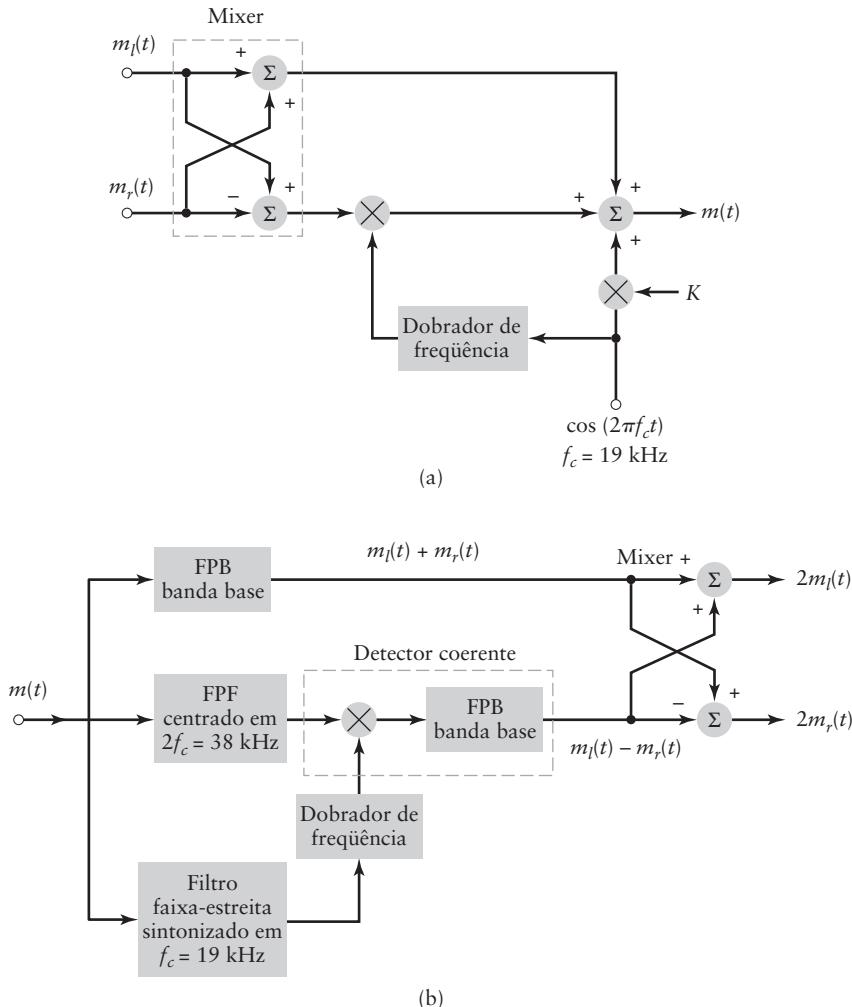


Figura 4.16 (a) Multiplexador para a transmissão de FM estéreo. (b) Demultiplexador no receptor de FM estéreo.

ção de freqüência) são aplicados a um modulador de produto, produzindo, portanto, uma onda modulada DSB-SC. Em adição ao sinal soma e esta onda modulada DSB-SC, o sinal multiplexado $m(t)$ também inclui um piloto de 19 kHz para fornecer uma referência para a detecção coerente do sinal diferença no receptor estéreo. Portanto, de acordo com a Fig. 4.16(a), o sinal multiplexado é descrito por

$$m(t) = [m_l(t) + m_r(t)] + [m_l(t) - m_r(t)] \cos(4\pi f_c t) + K \cos(2\pi f_c t) \quad (4.69)$$

na qual $f_c = 19$ kHz e K é a amplitude do tom piloto. O sinal multiplexado $m(t)$, então, modula em freqüência a portadora principal produzindo o sinal transmitido. Esta modulação em freqüência não é mostrada na Fig. 4.16(a). O piloto é posicionado entre 8 e 10 porcento do pico do desvio de freqüência. A amplitude K da Eq. (4.69) é escolhida para satisfazer esta condição.

No lado do receptor estéreo, antes de qualquer coisa, o sinal multiplexado $m(t)$ é recuperado pela demodulação de freqüência da onda FM de entrada. Então, $m(t)$ é aplicado ao sistema de demultiplexação mostrado na Fig. 4.16(b). Os componentes individuais do sinal multiplexado $m(t)$ são separados pela utilização de três filtros apropriados. O

piloto recuperado (usando um filtro de faixa estreita ajustado para 19 kHz) é dobrado em freqüência para produzir a subportadora desejada de 38 kHz. A disponibilidade desta subportadora possibilita a *detecção coerente* da onda modulada DSB-SC, veja a parte da Fig. 4.16(b) dentro do retângulo tracejado. O sinal diferença $m_l(t) - m_r(t)$ é, portanto, recuperado. O filtro passa-faixa banda base no caminho de cima da Fig. 4.16(b) é projetado para passar o sinal soma $m_l(t) + m_r(t)$. Finalmente, o *mixer* simples reconstrói o sinal original do lado esquerdo $m_l(t)$ e o sinal do lado direito $m_r(t)$, exceto pelo fator 2 de escala, e os aplica aos respectivos alto-falantes. A recepção FM estereofônica é, portanto, obtida.

4.10 RESUMO E DISCUSSÃO

No Capítulo 3, estudamos os princípios básicos da primeira família de modulação de ondas contínuas, baseado na modulação em amplitude e suas variantes. Neste capítulo, completamos o estudo dos princípios básicos da modulação de ondas contínuas, baseado na modulação em ângulo.

Fundamentalmente, existem dois tipos de modulação em ângulo:

- Modulação em fase (PM), na qual a fase instantânea da onda portadora senoidal é linearmente variada com o sinal de mensagem.
- Modulação em freqüência (FM), na qual a freqüência instantânea da onda portadora senoidal é linearmente variada com o sinal de mensagem.

Estes dois métodos de modulação são fortemente relacionados no sentido de que, se tivermos uma, podemos obter a outra. Por esta razão, focalizamos muito na discussão da modulação em freqüência.

A modulação em freqüência (FM) é tipificada pela equação

$$s(t) = A_c \cos\left(2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau\right) \quad (4.70)$$

Na qual $m(t)$ é o sinal de mensagem, $A_c \cos(2\pi f_c t)$ é a onda portadora senoidal e k_f é a sensibilidade de freqüência do modulador. A Eq. (4.70) é uma repetição da Eq. (4.7), reproduzida aqui por pura conveniência.

Ao contrário da modulação em amplitude, a partir da Eq. (4.70), vemos que a FM é um processo de modulação não linear. Desta forma, a análise espectral de FM é mais difícil do que a para AM. Mesmo assim, estudando FM de tom único, fomos capazes de desenvolver nosso conhecimento sobre as propriedades espectrais do FM. Em particular, obtemos uma regra empírica chamada de regra de Carson generalizada para o cálculo aproximado da largura de faixa B_T de transmissão de FM. De acordo com esta regra, B_T é controlado por um único parâmetro: o índice de modulação β para FM senoidal ou a razão de desvio D para FM não senoidal.

No FM, a amplitude da portadora é, portanto, a potência média transmitida, é mantida constante. Esta é uma importante vantagem do FM sobre AM no combate aos efeitos do ruído ou interferência na recepção, uma questão que estudaremos no Capítulo 9, após estarmos mais familiarizados com a teoria da probabilidade e processos aleatórios do Capítulo 8. Esta vantagem se torna ainda mais pronunciada quando o índice de modulação (razão de desvio) é aumentado, o que, de forma correspondente, possui o efeito de aumentar a largura de faixa de transmissão. Portanto, a modulação em freqüência fornece um método prático para o compromisso entre a largura de faixa do canal e a melhoria da performance frente ao ruído, o que não é possível na modulação em amplitude.

Um comentário final é importante. Tal como na modulação em amplitude, o desenvolvimento da família da modulação em ângulo foi motivado por sua relevância direta em comunicações analógicas, mas vários aspectos deste ramo da teoria de modulação são igualmente aplicáveis às comunicações digitais. Por exemplo, se o sinal de mensagem na Eq. (4.70) for restrito aos níveis -1 e $+1$, representando os símbolos binários 0 e 1 , respec-

tivamente, então temos a forma básica da modulação digital chamada de chaveamento por deslocamento em freqüência binário (BFSK), discutida no Capítulo 7.

Problemas adicionais

- 4.8 Trace as ondas PM e FM produzidas pela onda dente de serra mostrada na Fig. 4.17 como a fonte da modulação.

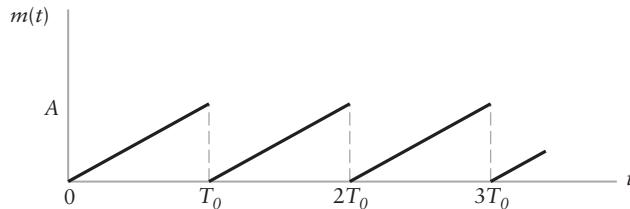


Figura 4.17 Problema 4.8

- 4.9 Em um radar *modulado em freqüência*, a freqüência instantânea da portadora transmitida é variada como na Fig. 4.18. Tal sinal é gerado pela modulação em freqüência com a onda modulante triangular periódica. A freqüência instantânea do sinal de eco recebido é mostrado em tracejado na Fig. 4.18, na qual τ é o atraso de tempo de retorno. Os sinais transmitidos e de eco são aplicados a um mixer e a componente de diferença de freqüência é mantida. Assumindo que $f_0\tau \ll 1$ para todo τ , determine o número médio em um segundo de ciclos de batimento na saída do mixer, em termos do pico de desvio Δf da freqüência de portadora, o atraso τ , e a freqüência de repetição f_0 do sinal transmitido. (O *batimento* se refere a um sinal cuja freqüência é a diferença entre as freqüências dos dois sinais de entrada.)

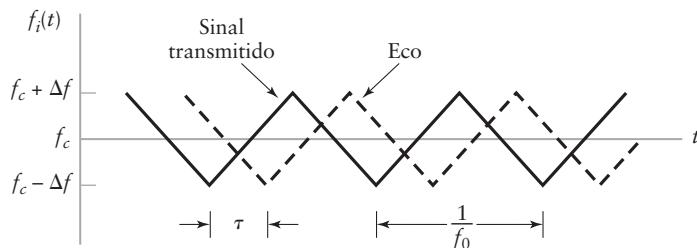


Figura 4.18 Problema 4.9

- 4.10 Considere um intervalo Δt de uma onda FM $s(t) = A_c \cos[\theta(t)]$ tal que $\theta(t)$ satisfaz a condição

$$\theta(t + \Delta t) - \theta(t) = \pi$$

Logo, mostre que se Δt é suficientemente pequeno, a freqüência instantânea da onda FM dentro deste intervalo é aproximadamente dada por

$$f_i \approx \frac{1}{2\Delta t}$$

- 4.11 A onda modulante senoidal

$$m(t) = A_m \cos(2\pi f_m t)$$

é aplicada a um modulador de fase com sensibilidade de fase k_p . A onda portadora não modulada possui freqüência f_c e amplitude A_c . Determine o espectro da onda modulada em fase resultante assumindo que o desvio de fase máximo $\beta = k_p A_m$ não excede 0,3 radianos.

- 4.12 Uma onda portadora é modulada em freqüência por um sinal senoidal com freqüência f_m e amplitude A_m .

- (a) Determine os valores do índice de modulação β para o qual a componente da portadora da onda FM é reduzida para zero. Para este cálculo você pode utilizar os valores de $J_0(\beta)$ dados no Apêndice 3.

- (b) Em um certo experimento realizado com $f_m = 1 \text{ kHz}$, e aumentando-se A_m (começando de zero volts), foi determinado que a componente da portadora da onda FM é reduzida para zero pela primeira vez quando $A_m = 2 \text{ volts}$. Qual é a sensibilidade de freqüência do modulador? Qual é o valor de A_m para o qual a componente da portadora é reduzida para zero pela segunda vez?

- 4.13 A onda portadora com freqüência 100 MHz é modulada em freqüência por uma onda senoidal de amplitude 20V e freqüência 100 kHz. A sensibilidade de freqüência do modulador é 25 kHz/V.

- (a) Determine a largura de faixa aproximada da onda FM utilizando a regra de Carson.

- (b) Determine a largura de faixa obtida pela transmissão apenas das freqüências laterais com amplitudes acima de 1% da amplitude da portadora não modulada. Utilize a curva universal da Fig. 4.9 para este cálculo.

- (c) Repita seus cálculos assumindo que a amplitude da onda modulante é dobrada.

- (d) Repita seus cálculos assumindo que a freqüência da modulante é dobrada.

- 4.14 Considere uma onda PM de faixa larga produzida pela onda modulante senoidal $A_m \cos(2\pi f_m t)$ utilizando um modulador com sensibilidade de fase igual a k_p radianos por volt.

- (a) Mostre que, se o desvio de fase máximo da onda PM for grande quando comparado com um radiano, a largura de faixa da onda PM varia linearmente com a freqüência de modulação f_m .

- (b) Compare esta característica de uma onda PM de faixa larga com a de uma onda FM de faixa larga.

- 4.15 A Fig. 4.19 mostra um diagrama de blocos de um sistema de malha fechada para a estabilização da freqüência da portadora de um modulador de freqüência de faixa larga. O oscilador controlado a tensão mostrado na figura constitui o modulador de freqüência. Utilizando as idéias de mixagem (isto é, translação de freqüência) (descritas no Capítulo 3) e discriminação de freqüência (descrita neste capítulo), discuta como o sistema realimentado da Fig. 4.19 é capaz de explorar a precisão de freqüência do oscilador a cristal para estabilizar o oscilador controlado a tensão.

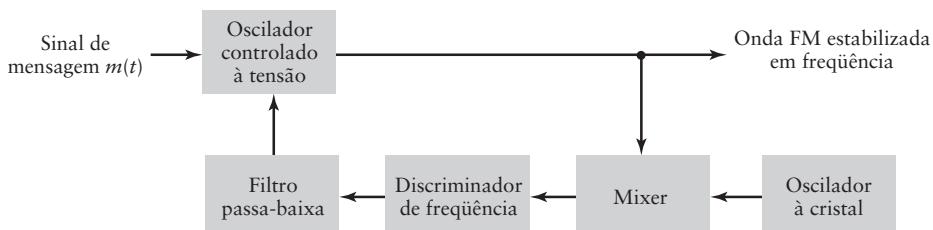


Figura 4.19 Problema 4.15

- 4.16 Considere o esquema de demodulação de freqüência mostrado na Fig. 4.20 no qual a onda FM $s(t)$ de entrada é passado em uma linha de atraso que produz um deslocamento de fase de $-\pi/2$ radianos na portadora com freqüência f_c . A saída da linha de atraso é subtraída de $s(t)$ e a onda composta resultante é, então, passada em um detector de envelope. Esta demodulação possui aplicação na demodulação de ondas FM com freqüências na faixa de microondas. Assumindo que

$$s(t) = A_c \cos[2\pi f_c t + \beta \sin(2\pi f_c t)]$$

analise a operação deste demodulador quando o índice de modulação β é menor do que a unidade e o atraso T produzido pela linha de atraso é suficientemente pequeno para justificar as aproximações:

$$\cos(2\pi f_m T) \approx 1$$

e

$$\sin(2\pi f_m T) \approx 2\pi f_m T$$

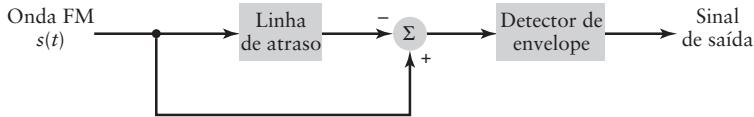


Figura 4.20 Problema 4.16.

- 4.17 Considere o seguinte par de sinais modulantes:

$$1. \quad m_1(t) = \begin{cases} a_1 t + a_0, & t \geq 0 \\ 0, & t = 0 \end{cases}$$

$$2. \quad m_2(t) = \begin{cases} b_2 t^2 + b_1 t + b_0, & t \geq 0 \\ 0, & t = 0 \end{cases}$$

no qual a s e b s são parâmetros constantes.

O sinal 1 é aplicado a um modulador de freqüência, enquanto o sinal 2 é aplicado a um modulador de fase. Determine as condições para as quais as saídas destes dois moduladores de ângulo são exatamente os mesmos.

- 4.18 Neste problema, iremos trabalhar nas especificações de um receptor FM super heterodino listado na Tabela 3.2. Em particular, dadas aquelas especificações, faça o seguinte:

- (a) Determine a faixa de freqüências fornecidas pelo oscilador local do receptor para acomodar a faixa de portadora RF de 88-108 MHz.
- (b) Determine a faixa correspondente de imagens de freqüência.

Problemas avançados

- 4.19 A freqüência instantânea de uma onda senoidal é igual a $f_c + \Delta f$ para $|t| < T/2$ e f_c para $|t| > T/2$. Determine o espectro desta onda modulada em freqüência. Dica: Divida o intervalo de tempo de interesse em três regiões não sobrepostas:

- (i) $-\infty < t < -T/2$
- (ii) $-T/2 \leq t \leq T/2$
- (iii) $T/2 < t < \infty$

- 4.20 A Fig. 4.21 mostra um diagrama em blocos de um *analizador espectral de tempo real* trabalhando com o princípio de modulação em freqüência. O sinal $g(t)$ e o sinal modulado em freqüência $s(t)$ são aplicados ao multiplicador e a saída $g(t)s(t)$ passa por um filtro com resposta $h(t)$ ao impulso. $s(t)$ e $h(t)$ são sinais FM lineares cujas freqüências instantâneas variam em taxas opostas, como mostrado por

$$s(t) = \cos(2\pi f_c t + \pi k t^2)$$

e

$$h(t) = \cos(2\pi f_c t - \pi k t^2)$$

na qual k é uma constante. Mostre que o envelope da saída do filtro é proporcional ao espectro de amplitude do sinal de entrada $g(t)$ com o termo em produto kt possuindo o papel de freqüência f . Dica: Utilize a notação complexa descrita na Seção 3.8 para a transmissão passa-faixa.

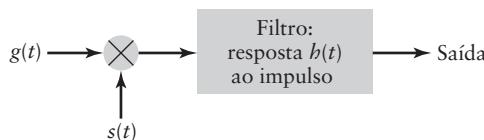


Figura 4.21 Problema 4.20.

4.21 Considere a onda modulada

$$s_1(t) = a(t) \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right]$$

na qual $a(t)$ é uma função envelope que varia lentamente, f_c é a freqüência da portadora, k_f é a sensibilidade de freqüência e $m(t)$ é o sinal de mensagem. A onda modulada $s(t)$ é processada por um *limitador passa-faixa*, o qual é constituído por um limitador seguido por um filtro passa-faixa. A função do limitador passa-faixa é remover flutuações de amplitude devido a $a(t)$. Especifique os parâmetros para o componente do filtro passa-faixa de forma a produzir a onda FM

$$s_2(t) = A \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right]$$

na qual A é uma amplitude constante.

- 4.22** A análise de distorção produzida em uma onda FM aplicada a um canal de comunicação linear é de importante interesse prático. Neste problema, iremos explorar a análise para o caso especial de uma onda FM de faixa larga produzida por uma onda modulante senoidal. Seja $H(f)$ a função de transferência do canal. Começando da Eq. (4.15):

- (a) Obtenha a expressão para o sinal modulado produzido na saída do canal.
- (b) Utilizando a expressão obtida na parte (a), discuta a distorção produzida pelo canal.

- 4.23** Na Seção 4.1, mostramos que o ângulo instantâneo $\theta_i(t)$ em ondas moduladas em ângulo pode ser variado de acordo com o sinal de mensagem $m(t)$ de infinitas formas. O tratamento da modulação em ângulo apresentado neste capítulo focalizou na modulação em fase e modulação em freqüência como dois importantes candidatos. O propósito deste problema é explorar outros métodos de produzir ondas moduladas em ângulo.

- (a) Para esta exploração, considere derivadas e integrais do sinal de mensagem $m(t)$ como possíveis funções responsáveis pelo processo de modulação.
- (b) Existiriam benefícios práticos nestes novos métodos de modulação em ângulo? Elabore sua resposta.

- 4.24** Neste problema, iremos explorar como o uso de FM pode superar distorção não linear. Considere um canal sem memória caracterizado pela seguinte relação não linear de entrada/saída:

$$v_o(t) = a_1 v_i(t) + a_2 v_i^2(t) + a_3 v_i^3(t)$$

na qual $v_i(t)$ é a entrada e $v_o(t)$ é a saída; a_1 , a_2 e a_3 são coeficientes constantes. A entrada é definida pelo sinal modulado em freqüência

$$v_i(t) = A_c \cos \left(2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right)$$

a largura de faixa da mensagem é definida por W , e o desvio de freqüência do sinal FM é Δf .

- (a) Calcule a saída $v_o(t)$.
- (b) Utilizando a regra de Carson generalizada, mostre que se a freqüência da portadora satisfizer a condição

$$f_c > 3 \Delta f + 2W$$

então o efeito da distorção não linear pode ser removido pela filtragem passa-faixa.

- (c) Especifique a freqüência de meia faixa e a largura de faixa do filtro da parte (b).

- 4.25** Considere um *PLL de segunda ordem* utilizando um filtro de malha com função de transferência

$$H(f) = 1 + \frac{a}{jf}$$

na qual a é um parâmetro do filtro.

- (a) Utilizando este filtro de malha na fórmula a seguir (veja a parte a do Problema 4.7)

$$\Phi_e(f) = \frac{1}{1 + L(f)} \Phi_1(f)$$

mostre que a transformada de Fourier resultante do erro de fase $\phi_e(t)$ é descrita por

$$\Phi_e(f) = \left(\frac{(jf/f_n)^2}{1 + 2\zeta(jf/f_n) + (jf/f_n)^2} \right) \Phi_1(f)$$

na qual f_n é a *freqüência natural* da malha e

$$\zeta = \sqrt{K_0/4a}$$

é o *fator de amortecimento*.

- (b) Logo, justifique as afirmativas de que pela escolha apropriada dos parâmetros f_n e ζ é possível para este PLL superar as limitações da versão de primeira ordem da malha.

MODULAÇÃO DE PULSO: TRANSIÇÃO DE COMUNICAÇÕES ANALÓGICAS PARA DIGITAIS

Na *modulação de onda contínua* (OC), que estudamos nos Capítulo 3 e 4, algum parâmetro da onda portadora senoidal é variado continuamente de acordo com o sinal de mensagem. Este é um contraste direto com a modulação de pulso, a qual é estudada neste capítulo. Na *modulação de pulso*, algum parâmetro de um *trem de pulsos* é variado de acordo com o sinal de mensagem. Neste contexto, podemos distinguir duas famílias de modulação de pulso, *modulação analógica de pulso* e *modulação digital de pulso*, dependendo de como a modulação é realizada. Na modulação analógica de pulso, um trem de pulso periódico é utilizado como onda portadora e alguma característica de cada pulso (isto é, amplitude, duração ou posição) é variada de forma contínua em função do valor correspondente da amostra do sinal de mensagem. Portanto, na modulação analógica de pulso, a informação é basicamente transmitida de forma analógica, mas a transmissão acontece em tempos discretos. Na modulação digital de pulso, por outro lado, o sinal de mensagem é representado em uma forma discreta tanto no tempo quanto na amplitude, permitindo, portanto, sua transmissão de forma digital como a seqüência de pulsos codificados. De forma simples, a modulação digital de pulso não possui a contra partida de onda contínua.

A utilização de pulsos codificados para a transmissão de sinais analógicos contendo informação representa o ingrediente básico na aplicação de comunicações digitais. Este capítulo pode, portanto, ser visto como a transição das comunicações analógicas para digitais em nosso estudo dos princípios de sistemas de comunicação.

Começamos o capítulo descrevendo o processo de amostragem, o qual é básico a todos os sistemas de modulação por pulso. Seguimos nossa discussão mostrando a modulação por amplitude de pulso, o qual é a forma mais simples de modulação analógica de pulso. Então continuamos com a descrição do processo de quantização, sendo seu uso a distinção da modulação digital de pulso da modulação analógica de pulso. Em particular, descrevemos três formas amplamente utilizadas de modulação digital de pulso – chamadas de modulação por codificação de pulso, modulação delta e modulação por codificação diferencial de pulso.

O material apresentado neste capítulo sobre modulação de pulso nos ensina as seguintes lições:

- *Lição 1: Dado um sinal de mensagem estritamente limitado em faixa, o teorema da amostragem apresenta a condição para uma versão uniformemente amostrada do sinal que preserva seu conteúdo de informação.*
- *Lição 2: Sistemas de modulação analógica de pulso são baseados no processo de amostragem para manter a representação contínua de amplitude do sinal de mensagem. Em contraste, sistemas de modulação digital de pulso utilizam não somente o processo de amostragem, mas também o processo de quantização, o qual é não reversível. A quantização possibilita a representação do sinal de mensagem na forma discreta tanto no tempo quanto em amplitude. Fazendo isto, a modulação digital de pulso possibilita explorar todo o poder de técnicas de processamento digital de sinais.*

5.1 PROCESSO DE AMOSTRAGEM

Muito do material sobre a representação de sinais e sistemas apresentado até aqui foi desenvolvido para sinais e sistemas contínuos tanto no tempo quanto na freqüência. Em vários pontos do Capítulo 2, entretanto, fizemos considerações sobre a representação de sinais periódicos. Em particular, lembre-se que a transformada de Fourier de um sinal periódico com período T_0 consiste de uma seqüência infinita de funções delta que ocorrem em múltiplos inteiros da freqüência fundamental $f_0 = 1/T_0$. Podemos, portanto, afirmar que *tornar um sinal periódico no domínio do tempo possui o efeito de amostrar o espectro do sinal no domínio da freqüência*. Podemos dar mais um passo invocando a propriedade da dualidade da transformada de Fourier e afirmar que *amostrar um sinal no domínio do tempo possui o efeito de tornar o espectro do sinal periódico no domínio da freqüência*. Esta última afirmativa é o assunto em estudo nesta seção.

O processo de amostragem é geralmente, mas não exclusivamente, descrito no domínio do tempo. Desta forma, ele é uma operação básica ao processamento digital de sinais e comunicações digitais. Através da utilização do processo de amostragem, um sinal analógico é convertido em uma seqüência correspondente de amostras que são geralmente uniformemente espaçadas no tempo. Claramente, para este procedimento ter utilidade prática, é necessário que escolhamos a taxa de amostragem adequadamente, de tal forma que a seqüência de amostrar definam unicamente o sinal analógico original. Esta é a essência do teorema da amostragem, a qual será obtida a seguir.

■ Amostragem instantânea e consequências no domínio da freqüência

Considere um sinal arbitrário $g(t)$ com energia finita, o qual é especificado para todo tempo t . Um segmento do sinal $g(t)$ é mostrado na Fig. 5.1(a). Suponha que amostremos o sinal $g(t)$ instantaneamente e a uma taxa uniforme, a cada T_s segundos. Conseqüentemente, obtemos uma seqüência infinita de amostras espaçadas T_s segundos uma da outra e representada por $\{g(nT_s)\}$, na qual n assume todos os possíveis valores inteiros, tanto positivos quanto negativos. Referimos a T_s como o *período de amostragem* ou *intervalo de amostragem* e seu recíproco $f_s = 1/T_s$ como *taxa de amostragem*. Esta forma ideal de amostragem é chamada de *amostragem instantânea*.

Seja $g_\delta(t)$ o sinal obtido pela ponderação individual dos elementos de uma seqüência periódica de funções delta de Dirac espaçados T_s segundos uns dos outros pela seqüência de números $\{g(nT_s)\}$, como mostrado por (veja a Fig. 5.1(b))

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} g(nT_s) \delta(t - nT_s) \quad (5.1)$$

Denominamos $g_\delta(t)$ como o *sinal instantaneamente (idealmente) amostrado*. O termo $\delta(t - nT_s)$ representa a função delta posicionada no tempo $t = nT_s$. Da definição da função delta apresentada na Seção 2.4, lembre-se que tal função idealizada possui área unitária. Podemos, portanto, ver o fator de multiplicação $g(nT_s)$ da Eq. (5.1) como uma “massa” designada à função delta $\delta(t - nT_s)$. A função delta ponderada desta forma é aproximada por um pulso retangular de duração Δt e amplitude $g(nT_s)/\Delta t$. Quanto menor o valor de Δt , melhor a aproximação.

O sinal instantaneamente amostrado $g_\delta(t)$ possui forma matemática similar à da transformada de Fourier de um sinal periódico. Este fato é facilmente estabelecido comparando a Eq. (5.1) para $g_\delta(t)$ com a transformada de Fourier de um sinal periódico dado pelo lado direito da Eq. (2.88). Esta correspondência sugere que podemos determinar a transformada de Fourier do sinal amostrado $g_\delta(t)$ invocando a propriedade da dualidade da transformada de Fourier, cuja essência está embutida na Eq. (2.24). De fato, aplicando esta propriedade para a transformada de Fourier da Eq. (2.88) e a Eq. (2.87) relacionada, podemos desenvolver a Tabela 5.1. As linhas listadas na tabela descrevem as relações duais entre a amostragem no domínio do tempo e sua contra-partida, amostragem no domínio da freqüência.

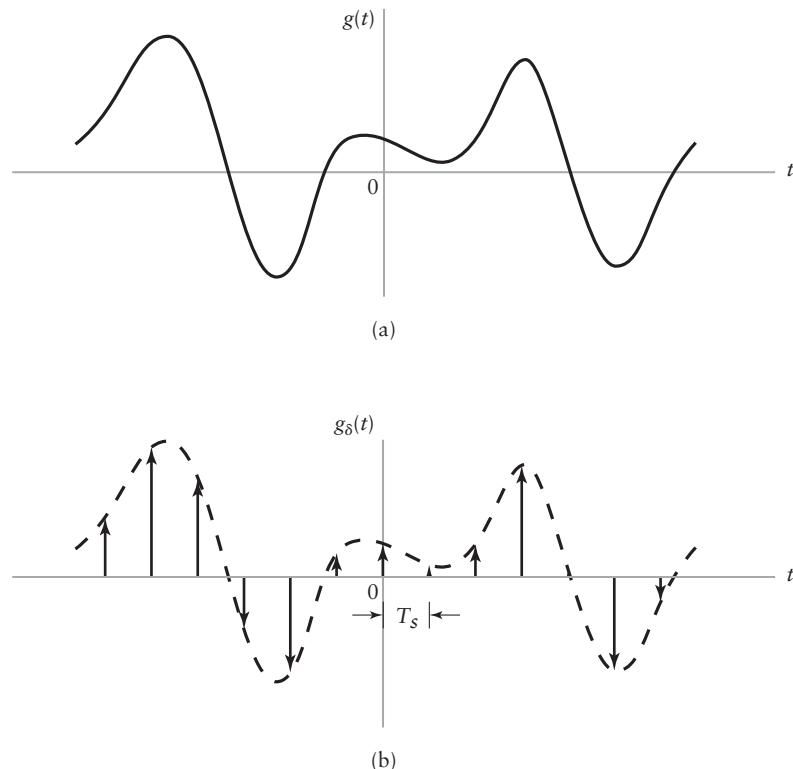


Figura 5.1 Ilustração do processo de amostragem. (a) Forma de onda analógica $g(t)$. (b) Representação instantaneamente amostrada de $g(t)$.

TABELA 5.1 Dualidade das relações de amostragem tempo-freqüência

Amostragem ideal no domínio da freqüência (espectro discreto); veja o Capítulo 2	Amostragem ideal no domínio do tempo (função no tempo discreto); veja este capítulo
Período fundamental $T_0 = 1/f_0$ Função delta $\delta(f - m f_0)$, na qual $m = 0, \pm 1, \pm 2, \dots$ Periodicidade no domínio do tempo Função limitada no tempo	Taxa de amostragem $f_s = 1/T_s$ Função delta $\delta(t - n T_s)$, na qual $n = 0, \pm 1, \pm 2, \dots$ Periodicidade no domínio da freqüência Espectro limitado em faixa

$$T_0 \sum_{m=-\infty}^{\infty} g(t - mT_0) = \sum_{n=-\infty}^{\infty} G(nf_0) e^{j2\pi n f_0 t}$$

||

$$\sum_{n=-\infty}^{\infty} G(nf_0) \delta(f - nf_0)$$

$$\sum_{n=-\infty}^{\infty} g(nT_s) \delta(t - nT_s)$$

||

$$\sum_{n=-\infty}^{\infty} g(nT_s) e^{-j2\pi n T_s f} = f_s \sum_{m=-\infty}^{\infty} G(f - mf_s)$$

► PROBLEMA 5.1

- Utilizando o material apresentado na Seção 2.5, justifique a relação matemática listada na parte inferior do lado esquerdo da Tabela 5.1 a qual pertence à amostragem ideal no domínio da freqüência.
- Aplicando a propriedade da dualidade da transformada de Fourier à parte (a), justifique as relações matemáticas listadas na parte inferior do lado direito da tabela, as quais pertencem à amostragem ideal no domínio do tempo. ◀

A motivação para a criação da Tabela 5.1 está no trabalho matemático básico para a formulação do teorema da amostragem no domínio do tempo. Para isto, reproduzimos as relações listadas na parte inferior do lado direito da tabela na forma

$$\sum_{n=-\infty}^{\infty} g(nT_s) \delta(t - nT_s) \iff f_s \sum_{m=-\infty}^{\infty} G(f - mf_s) = \sum_{n=-\infty}^{\infty} g(nT_s) \exp(-j2\pi nT_s f) = G_\delta(f) \quad (5.2)$$

na qual $G(f)$ é a transformada de Fourier do sinal original $g(t)$ e $f_s = 1/T_s$ é a taxa de amostragem. Em palavras, a Eq. (5.2) afirma que o *processo de amostragem uniforme de um sinal contínuo de energia finita resulta em um espectro periódico cuja freqüência de repetição é igual à taxa de amostragem*.

■ Teorema da amostragem

As relações da Eq. (5.2) se aplicam a qualquer sinal $g(t)$ contínuo no tempo de energia finita. Suponha, entretanto, que o sinal $g(t)$ é *estritamente limitado em faixa*, sem componentes de freqüência maiores do que W hertz. Ou seja, a transformada de Fourier $G(f)$ do sinal $g(t)$ possui a propriedade de $G(f)$ ser zero para $|f| \geq W$, como ilustrado na Fig. 5.2(a). A forma do espectro mostrado nesta figura é utilizada somente com o propósito ilustrativo. Suponha também que escolhamos o período de amostragem $T_s = 1/2W$, o qual, como iremos ver, é o máximo valor permitido. Então, o espectro $G_\delta(f)$ correspondente do sinal amostrado $g_\delta(t)$ é como mostrado na Fig. 5.2(b). Fazendo $T_s = 1/2W$ na Eq. (5.2) e usando $G_\delta(f)$ para representar a transformada de Fourier de $g_\delta(t)$, podemos escrever

$$G_\delta(f) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \exp\left(-j\frac{\pi n f}{W}\right) \quad (5.3)$$

A Eq. (5.3) define a transformada de Fourier $G_\delta(f)$ da seqüência $\{g(n/2W)\}_{n=-\infty}^{\infty}$, a qual é obtida pela amostragem uniforme do sinal $g(t)$ contínuo no tempo na taxa especial $(1/T_s) = f_s = 2W$. A fórmula obtida utilizando o período de amostragem $T_s = 1/2W$, mostrada na Eq. (5.3), é chamada de *transformada de Fourier em tempo discreto*¹ da seqüência $\{g(nT_s)\}_{n=-\infty}^{\infty}$.

- **PROBLEMA 5.2** Mostre que quando o período de amostragem T_s se aproxima de zero, a fórmula da transformada de Fourier em tempo discreto $G_\delta(f)$ dada na Eq. (5.3) se aproxima da fórmula da transformada de Fourier de $G(f)$. ◀

Retornando à Eq. (5.2), suponha que isolemos o termo correspondente a $m = 0$ do somatório e, portanto, escrevemos

$$G_\delta(f) = f_s G(f) + f_s \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} G(f - mf_s)$$

¹ Na fórmula de Fourier da Eq. (5.3), o tempo t é implicitamente discretizado. Se formos um passo a mais e discretizarmos a freqüência f também, fazendo $f = k \times 2W$, obtemos a transformada discreta de Fourier, a qual é periódica tanto no tempo quanto na freqüência. Especificamente,

$$G_k = \sum_{n=0}^{N-1} g_n \exp(-j2\pi nk), \quad k = 0, 1, \dots, N-1$$

na qual

$$g_n = g\left(\frac{n}{2W}\right) = g(nT_s)$$

e

$$G_k = G_\delta(2kW) = G_\delta\left(\frac{k}{T_s}\right)$$

O parâmetro N é o número de amostras em cada período, seja no domínio do tempo ou no domínio da freqüência. A transformada discreta de Fourier foi discutida no Capítulo 2.

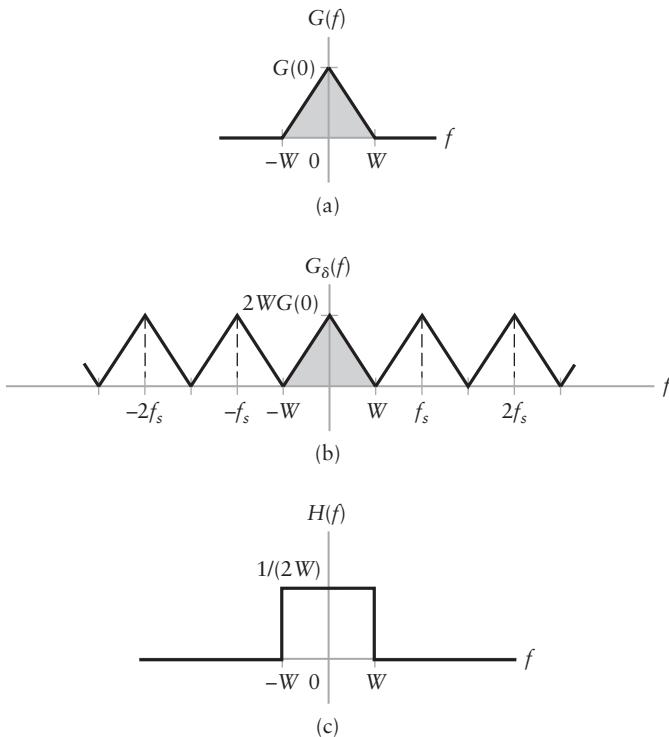


Figura 5.2 (a) Espectro de um sinal $g(t)$ estritamente limitado em faixa. (b) Espectro da versão instantaneamente amostrada de $g(t)$ para um período de amostragem $T_s = 1/2W$. (c) Resposta em freqüência de um filtro passa-baixa projetado para recuperar o sinal de mensagem original $g(t)$ de sua versão uniformemente amostrada.

Desta expressão vemos que, para um sinal estritamente limitado em faixa, sob as duas condições

1. $G(f) = 0$ para $|f| \geq W$
2. $f_s = 2W$

o termo do somatório é igual a zero. Então, resolvendo a expressão simplificada para $G(f)$, obtemos

$$G(f) = \frac{1}{2W} G_\delta(f), \quad -W < f < W \quad (5.4)$$

Eliminando $G_\delta(f)$ entre as Eqs. (5.3) e (5.4),

$$G(f) = \frac{1}{2W} \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \exp\left(-\frac{j\pi n f}{W}\right), \quad -W < f < W \quad (5.5)$$

Portanto, se os valores amostrados $g(n/2W)$ de um sinal $g(t)$ forem especificados para todo tempo, então a transformada de Fourier $G(f)$ do sinal $g(t)$ é unicamente determinada, exceto por um fator de escala de $1/2W$, pela *transformada de Fourier em tempo discreto* da Eq. (5.3) para o espectro $G_\delta(f)$ limitado ao intervalo $-W \leq f \leq W$. Como $g(t)$ é relacionado com $G(f)$ pela transformada inversa de Fourier, temos que o sinal $g(t)$ é, por ele mesmo, unicamente determinado pelos valores de amostra $g(n/2W)$ para $-\infty < n < \infty$. Em outras palavras, a sequência $\{g(n/2W)\}$ possui toda a informação contida em $g(t)$.

Considere, a seguir, o problema de reconstruir o sinal $g(t)$ da sequência de valores amostrados $\{g(n/2W)\}$. Substituindo a Eq. (5.5) na fórmula de transformada inversa de Fourier definindo $g(t)$ em termos de $G(f)$, temos

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft) df \\ &= \int_{-W}^{W} \frac{1}{2W} \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \exp\left(-j\frac{\pi n f}{W}\right) \exp(j2\pi ft) df \end{aligned}$$

Podemos alterar a ordem do somatório e integração, pois os dois são operadores lineares. Desta forma, podemos redefinir o sinal desejado $g(t)$ como

$$g(t) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \frac{1}{2W} \int_{-W}^{W} \exp\left[j2\pi f\left(t - \frac{n}{2W}\right)\right] df \quad (5.6)$$

► **PROBLEMA 5.3** Mostre que

$$\begin{aligned} \frac{1}{2W} \int_{-W}^{W} \exp\left[j2\pi f\left(t - \frac{n}{2W}\right)\right] df &= \frac{\sin(2\pi Wt - n\pi)}{(2\pi Wt - n\pi)} \\ &= \text{sinc}(2Wt - n) \end{aligned}$$

À luz do Problema 5.3, a fórmula da Eq. (5.6) se reduz para

$$g(t) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \text{sinc}(2Wt - n), \quad -\infty < t < \infty \quad (5.7)$$

A Eq. (5.7) é a *fórmula de interpolação* para a reconstrução do sinal original $g(t)$ da seqüência de valores amostrados $\{g(n/2W)\}$, com a função $\text{sinc}(2Wt)$ possuindo o papel de *função de interpolação*. Cada amostra é multiplicada por uma versão atrasada da função de interpolação e, então, todas as formas de onda resultantes são somadas para obtermos $g(t)$.

► **PROBLEMA 5.4** Este problema é dedicado à identificação do filtro linear que satisfaz a fórmula de interpolação da Eq. (5.7), apesar de ser em uma forma não realizável. A Eq. (5.7) é baseada na premissa de que o sinal $g(t)$ é estritamente limitado à faixa $-W \leq f \leq W$. Com esta especificação em mente, considere um filtro passa-baixa ideal cuja resposta em freqüência $H(f)$ é como mostrado na Fig. 5.2(c). A resposta ao impulso deste filtro é definida por (veja a Eq. (2.25))

$$h(t) = \text{sinc}(2Wt), \quad -\infty < t < \infty$$

Suponha que o sinal correspondente $g_{\delta}(t)$ instantaneamente amostrado definido na Eq. (5.1) é aplicado a este filtro passa-baixa ideal. Com este conhecimento prévio, utilize a integral de convolução para mostrar que a saída resultante do filtro é definida exatamente pela fórmula de interpolação da Eq. (5.7). ◀

À luz do Problema 5.4, podemos, agora, formalmente dizer que o *filtro de síntese* ou *filtro de reconstrução* projetado para reconstrução do sinal original $g(t)$ estritamente limitado em faixa de sua versão $g_{\delta}(t)$ instantaneamente amostrada de acordo com a Eq. (5.7), consiste de um filtro passa-baixa ideal cuja resposta em freqüência é limitada exatamente à mesma faixa do sinal $g(t)$, ou seja, $-W \leq f \leq W$. Este filtro de reconstrução é não causal e, portanto, não realizável fisicamente. Posteriormente nesta seção, iremos descrever como implementar fisicamente o filtro de reconstrução pelo relaxamento das especificações do sinal $g(t)$.

A transformada de Fourier em tempo discreto da Eq. (5.5) define o espectro $G(f)$ da mensagem em termos de valores amostrados $g(n/2W)$, uniformemente espaçados em $-\infty < n < \infty$. A fórmula de interpolação da Eq. (5.7) define o sinal de mensagem $g(t)$ em termos dos valores destas mesmas amostras. Com base nestas duas fórmulas, podemos apresentar o *teorema da amostragem* para sinais estritamente limitados em faixa de energia finita em duas partes equivalentes:

1. Análise. Um sinal limitado em faixa de energia finita que não possua componentes de freqüência maiores do que W hertz é completamente descrito pela especificação dos valores do sinal em instantes de tempo separados por $1/2W$ segundos.

2. Síntese. Um sinal limitado em faixa de energia finita que não possua componentes de freqüência maiores do que W hertz é completamente recuperado do conhecimento de suas amostras tiradas na taxa de $2W$ amostras por segundo.

A taxa de amostragem de $2W$ amostras por segundo para um sinal de largura de faixa W hertz é chamada de *taxa de Nyquist*, seu recíproco $1/2W$ (medido em segundos) é chamado de *intervalo de Nyquist*. A parte de análise do teorema da amostragem se aplica ao transmissor. A parte de síntese do teorema, por outro lado, se aplica ao receptor. Note que a taxa de Nyquist é a menor taxa de amostragem permitida.

► **PROBLEMA 5.5** Especifique a taxa de Nyquist e o intervalo de Nyquist para cada um dos sinais a seguir:

- (a) $g(t) = \text{sinc}(200t)$
- (b) $g(t) = \text{sinc}^2(200t)$
- (c) $g(t) = \text{sinc}(200t) + \text{sinc}^2(200t)$

► **PROBLEMA 5.6** Considere a amostragem uniforme da onda senoidal

$$g(t) = \cos(\pi t)$$

Determine a transformada de Fourier da forma de onda amostrada para cada um dos seguintes períodos de amostragem:

- (a) $T_s = 0,25\text{s}$
- (b) $T_s = 1\text{s}$
- (c) $T_s = 1,5\text{s}$

► **PROBLEMA 5.7** Considere um sinal contínuo no tempo definido por

$$g(t) = \frac{\sin(\pi t)}{\pi t}$$

O sinal $g(t)$ é uniformemente amostrado para produzir a seqüência infinita $\{g(nT_s)\}_{n=-\infty}^{\infty}$. Determine a condição que o período T_s deve satisfazer para que o sinal $g(t)$ seja unicamente recuperado da seqüência $\{g(nT_s)\}$.

■ Fenômeno de aliasing

A obtenção do teorema da amostragem, como descrito aqui, é baseada na consideração do sinal $g(t)$ ser estritamente limitado em faixa. Na prática, entretanto, nenhum sinal físico contendo informação é estritamente limitado em faixa, resultando em sempre haver alguma subamostragem. Conseqüentemente, o *aliasing* é produzido pelo processo de amostragem. Por aliasing nos referirmos ao fenômeno da componente de alta freqüência do espectro do sinal assumir a identidade de uma baixa freqüência no espectro de sua versão amostrada, como ilustrado na Fig. 5.3. O espectro com alias mostrado pela curva sólida na Fig. 5.3(b) pertence a uma versão “subamostrada” do sinal de mensagem representado pelo espectro da Fig. 5.3(a).

Para combater os efeitos de aliasing na prática, podemos utilizar duas *medidas corretivas*:

1. Antes da amostragem, um filtro passa-baixa anti-aliasing é utilizado para atenuar as componentes de alta freqüência do sinal de mensagem que não são essenciais à informação contida no sinal.
2. O sinal filtrado é amostrado a uma taxa um pouco maior do que a taxa de Nyquist.

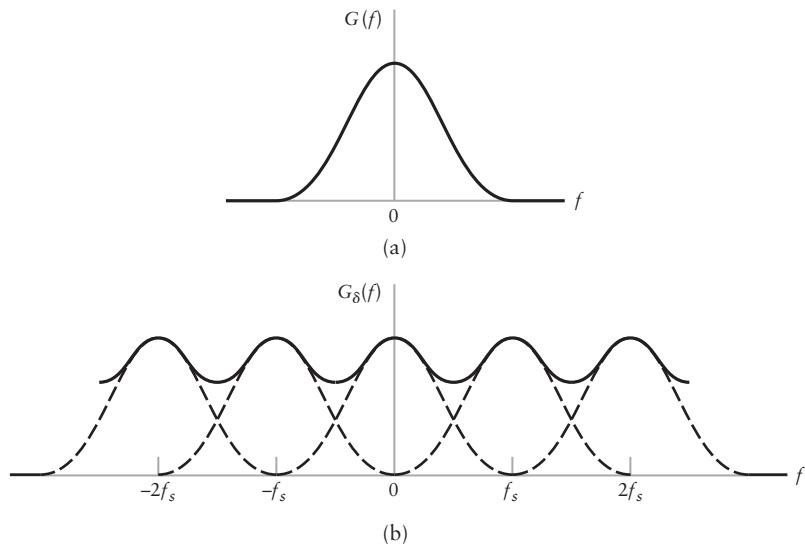


Figura 5.3 (a) Espectro do sinal. (b) Espectro de uma versão subamostrada do sinal, exibindo o fenômeno de aliasing.

A utilização da taxa de amostragem maior do que a taxa de Nyquist também possui efeitos benéficos de facilitar o projeto do *filtro de síntese* utilizado para recuperar o sinal original de sua versão amostrada. Considere o exemplo de um sinal de mensagem que foi filtrado por um filtro anti-aliasing (passa-baixa), resultando no espectro mostrado na Fig. 5.4(a). O espectro correspondente da versão instantaneamente amostrada do sinal é mostrado na Fig. 5.4(b), assumindo uma taxa de amostragem maior do que a taxa de Nyquist. De acordo com a Fig. 5.4(b), podemos facilmente ver que o projeto de um *filtro de reconstrução fisicamente realizable* objetivando a recuperação do sinal original de sua versão uniformemente amostrada pode ser realizado como mostrado a seguir (veja a Fig. 5.4(c)):

- ▶ O filtro de reconstrução é um filtro passa-baixa com faixa passante estendendo de $-W$ a W , o que, por si só, define o filtro anti-aliasing.
- ▶ O filtro possui faixa de transição não nula estendendo (para freqüências positivas) de W a $f_s - W$, na qual f_s é a taxa de amostragem.

A faixa de transição não nula do filtro garante que ele é fisicamente realizável, ela é mostrada tracejada para enfatizar a sua forma arbitrária.

5.2 MODULAÇÃO POR AMPLITUDE DE PULSO

Agora que já conhecemos a essência do processo de amostragem, estamos prontos para definir formalmente a modulação por amplitude de pulso, a qual é a forma mais simples e básica das técnicas de modulação analógica de pulso. Na *modulação por amplitude de pulso* (PAM – do inglês, *pulso-amplitude modulation*), a amplitude dos pulsos regularmente espaçados é variada proporcionalmente aos valores amostrados do sinal contínuo de mensagem. Os pulsos podem estar na forma retangular ou algum outro formato adequado. A modulação por amplitude de pulso é definida aqui de forma similar à *amostragem natural*, na qual o sinal de mensagem é multiplicado por um trem periódico de pulsos retangulares. Na amostragem natural, entretanto, o topo de cada pulso retangular modulado pode variar com o sinal de mensagem, enquanto que no PAM ele é mantido plano. (A amostragem natural é explorada posteriormente no Problema 5.26.)

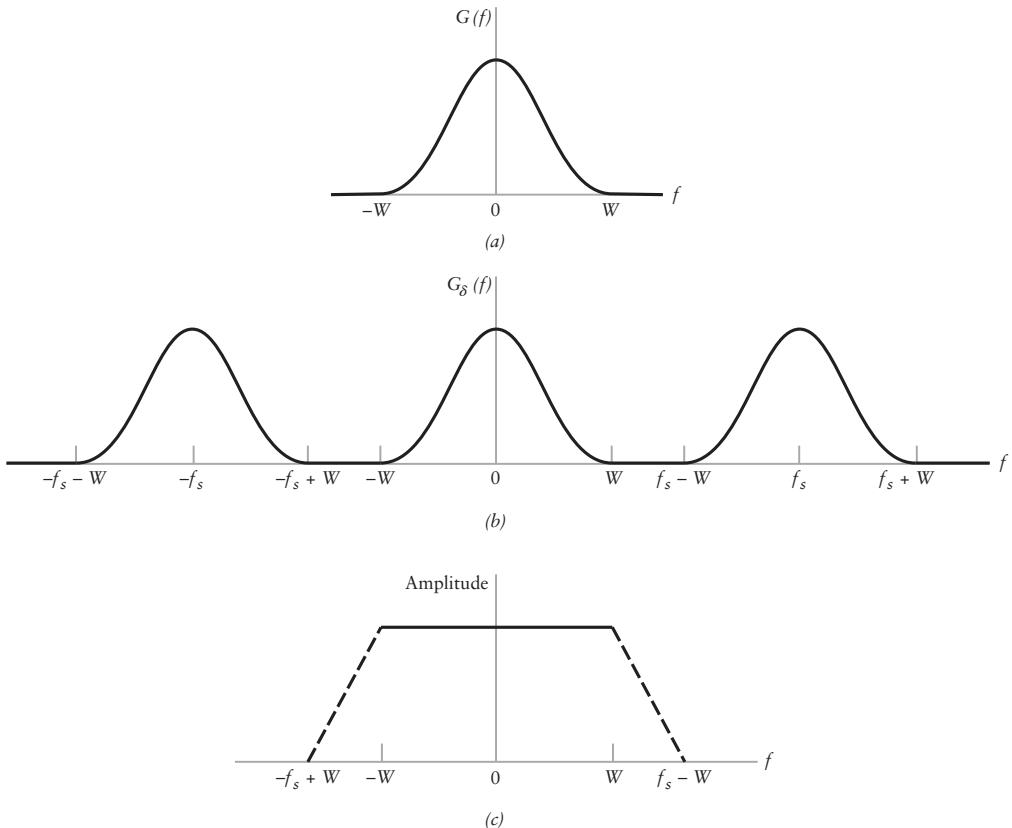


Figura 5.4 (a) Espectro apóis filtro anti-aliasing de um sinal contendo a informação. (b) Espectro da versão instantaneamente amostrada do sinal, assumindo a utilização de uma taxa de amostragem maior do que a taxa de Nyquist. (c) Resposta em amplitude idealizada de um filtro de reconstrução.

A forma de onda do sinal PAM é apresentada na Fig. 5.5. A curva tracejada nesta figura indica a forma de onda do sinal de mensagem $m(t)$, e a seqüência de pulsos retangulares modulados em amplitude mostrados como linha sólida representam o sinal PAM $s(t)$ correspondente. Existem duas operações necessárias para a geração do sinal PAM:

1. *Amostragem instantânea* do sinal de mensagem $m(t)$ a cada T_s segundos, com a taxa de amostragem $f_s = 1/T_s$ escolhida de acordo com o teorema da amostragem.
2. *Extensão* da duração de cada amostra, de tal forma que elas ocupem algum valor finito T .

Na tecnologia de circuitos digitais, estas duas operações conjuntas são chamadas de *sample-and-hold*. Uma importante razão para estender intencionalmente a duração de cada amostra é evitar a utilização extensiva da largura de faixa do canal, pois a largura de faixa é inversamente proporcional à duração do pulso. Entretanto, deve-se ter cuidado no total que estendemos a duração T da amostra, como a análise a seguir revela.

■ Filtragem sample-and-hold: Análise

Seja $s(t)$ a seqüência de pulsos de topo plano gerado pela maneira descrita na Fig. 5.5. Logo, podemos expressar o sinal PAM por

$$s(t) = \sum_{n=-\infty}^{\infty} m(nT_s)h(t - nT_s) \quad (5.8)$$

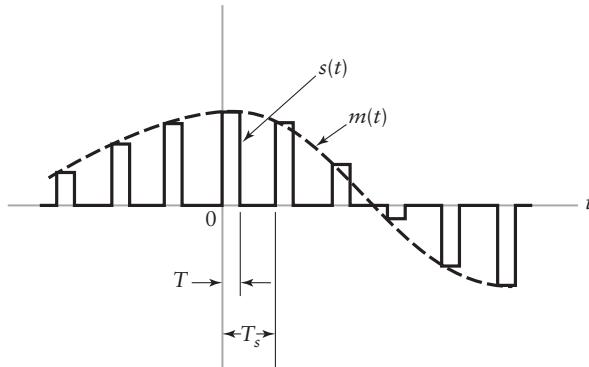


Figura 5.5 Amostragem de topo plano do sinal de mensagem.

na qual T_s é o período de amostragem e $m(nT_s)$ é o valor da amostra de $m(t)$ obtida no tempo $t = nT_s$. $h(t)$ é um pulso retangular padrão de amplitude unitária e duração T , definido por (veja a Fig. 5.6(a)):

$$h(t) = \text{ret}\left(\frac{t - \frac{T}{2}}{T}\right) = \begin{cases} 1, & 0 < t < T \\ \frac{1}{2}, & t = 0, t = T \\ 0, & \text{caso contrário} \end{cases} \quad (5.9)$$

Pela definição, a versão instantaneamente amostrada de $m(t)$ é dada por [veja a Eq. (5.1)]

$$m_\delta(t) = \sum_{n=-\infty}^{\infty} m(nT_s)\delta(t - nT_s) \quad (5.10)$$

na qual $\delta(t - nT_s)$ é a função delta deslocada no tempo. Para modificar $m_\delta(t)$ de tal forma que ela assuma a mesma forma do sinal PAM $s(t)$, convoluímos $m_\delta(t)$ com o pulso $h(t)$, obtendo

$$\begin{aligned} m_\delta(t) \star h(t) &= \int_{-\infty}^{\infty} m_\delta(\tau)h(t - \tau) d\tau \\ &= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} m(nT_s)\delta(\tau - nT_s)h(t - \tau) d\tau \\ &= \sum_{n=-\infty}^{\infty} m(nT_s) \int_{-\infty}^{\infty} \delta(\tau - nT_s)h(t - \tau) d\tau \end{aligned} \quad (5.11)$$

Na última linha, trocamos a ordem do somatório e da integral, pois os dois são operadores lineares. Utilizando a propriedade de deslocamento da função delta, ou seja,

$$\int_{-\infty}^{\infty} \delta(\tau - nT_s)h(t - \tau) d\tau = h(t - nT_s)$$

Desta forma, a Eq. (5.11) se reduz para

$$m_\delta(t) \star h(t) = \sum_{n=-\infty}^{\infty} m(nT_s)h(t - nT_s) \quad (5.12)$$

Os termos de somatório das Eqs. (5.8) e (5.12) são idênticos. Portanto, temos que o sinal PAM $s(t)$ é matematicamente equivalente à convolução de $m_\delta(t)$, a versão instantaneamente amostrada de $m(t)$, pelo pulso $h(t)$, mostrado por

$$s(t) = m_\delta(t) \star h(t) \quad (5.13)$$

Determinando a transformada de Fourier dos dois lados da Eq. (5.13) e lembrando que a convolução de duas funções no tempo é transformada em uma multiplicação de suas transformadas de Fourier, obtemos

$$S(f) = M_\delta(f)H(f) \quad (5.14)$$

na qual $S(f) = F[s(t)]$, $M_\delta(f) = F[m_\delta(t)]$ e $H(f) = F[h(t)]$. Da Eq. (5.2), vemos que a transformada de Fourier $M_\delta(f)$ é relacionada com o espectro do sinal original $m(t)$ por:

$$M_\delta(f) = f_s \sum_{k=-\infty}^{\infty} M(f - kf_s) \quad (5.15)$$

na qual $f_s = 1/T_s$ é a taxa de amostragem. Portanto, substituindo a Eq. (5.15) na Eq. (5.14), obtemos

$$S(f) = f_s \sum_{k=-\infty}^{\infty} M(f - kf_s)H(f) \quad (5.16)$$

- **PROBLEMA 5.8** Começando pela Eq. (5.9), mostre que a transformada de Fourier do pulso rectangular $h(t)$ é dada por

$$H(f) = T \operatorname{sinc}(fT) \exp(-j\pi fT) \quad (5.17)$$

o que acontece a $H(f)/T$ quando a duração T do pulso tende a zero? ◀

Dado o sinal PAM $s(t)$ cuja transformada de Fourier $S(f)$ é definida pela Eq. (5.16), como recuperamos o sinal de mensagem $m(t)$ original? Como primeiro passo para esta recuperação, podemos passar $s(t)$ por um filtro passa-baixa cuja resposta em freqüência é definida na Fig. 5.2(c); considera-se, aqui, que o sinal de mensagem $m(t)$ é limitado à largura de faixa W e que a taxa de amostragem f_s é maior do que a taxa de Nyquist $2W$. Então, da Eq. (5.16), obtemos o espectro da saída resultante do filtro igual a $M(f)H(f)$. Esta saída é equivalente a passar o sinal de mensagem $m(t)$ através de outro filtro passa-baixa com função de transferência $H(f)$. O passo a seguir na recuperação do sinal de mensagem $m(t)$ requer a utilização da *equalização*, discutida a seguir.

Efeito da abertura e sua equalização

A Fig. 5.6(b) mostra o gráfico da amplitude e fase da transformada de Fourier $H(f)$ em função da freqüência f . A partir desta figura, vemos que, com a utilização de amostras de topo plano para gerar o sinal PAM, introduzimos uma *distorção de amplitude* e um *atraso* de $T/2$. Este efeito é, de alguma forma, similar à variação na transmissão com a freqüência que é causada pelo tamanho finito da abertura de busca em canais de televisão. Por esta razão, a distorção causada pelo uso de modulação por amplitude de pulso (baseada em amostragem de topo plano) para a transmissão de um sinal analógico contendo a informação é chamada de *efeito de abertura*.

Esta distorção pode ser corrigida conectando um *equalizador* em cascata com o filtro passa-baixa de reconstrução, como mostrado na Fig. 5.7. O equalizador possui o efeito de diminuir a atenuação dentro da faixa do filtro de reconstrução quando a freqüência aumenta, de tal forma a compensar o efeito da abertura. Idealmente, a resposta em amplitude do equalizador é dada por

$$\frac{1}{|H(f)|} = \frac{1}{T \operatorname{sinc}(fT)} = \frac{\pi f}{\operatorname{sen}(\pi fT)}$$

O total de equalização necessária na prática é geralmente pequeno. De fato, para um ciclo de trabalho $(T/T_s) \leq 0,1$, a distorção em amplitude é menor do que 5% e, neste caso, a necessidade de equalização pode ser totalmente omitida.

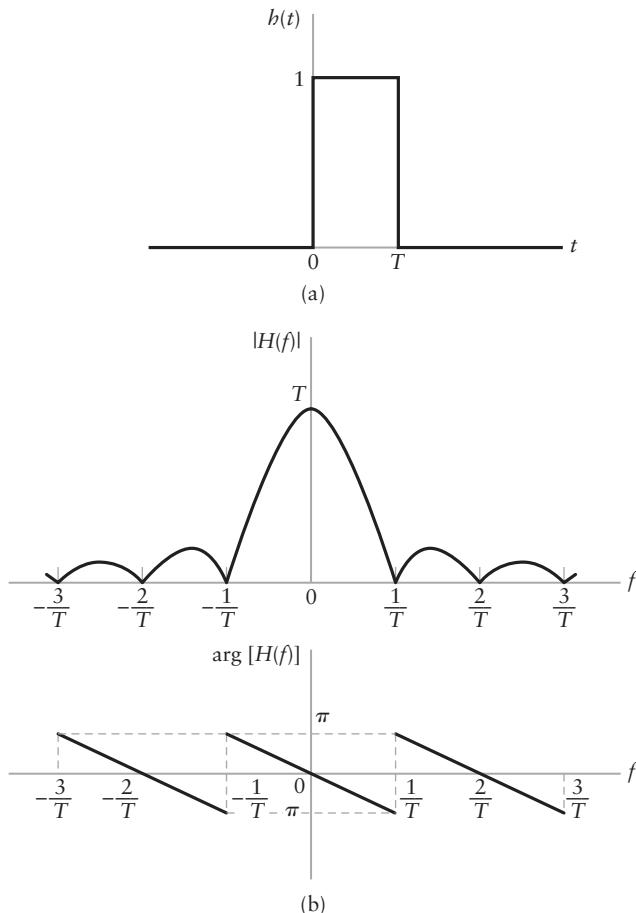


Figura 5.6 (a) Pulso retangular $h(t)$. (b) Espectro $H(f)$, definido em termos de sua amplitude e fase.



Figura 5.7 Recuperação do sinal de mensagem $m(t)$ do sinal PAM $s(t)$.

A transmissão do sinal PAM impõe restrições mais rígidas na resposta em amplitude e fase do canal, devido à relativa curta duração dos pulsos transmitidos. Além disso, pode ser mostrado que a performance frente ao ruído de um sistema PAM nunca pode ser melhor do que a transmissão direta do sinal de mensagem. Desta forma, para transmissões em longa distância, o sinal PAM deve ser utilizado somente como forma de processar a mensagem para uma multiplexação por divisão no tempo. O conceito de multiplexação por divisão no tempo é discutido posteriormente neste capítulo.

5.3 MODULAÇÃO POR POSIÇÃO DE PULSO

Na modulação por amplitude de pulso, a amplitude do pulso é o parâmetro variável. A duração do pulso é o próximo parâmetro lógico disponível para a modulação. Na modulação por duração de pulso (PDM – do inglês, *pulse-duration modulation*), as amostras do

sinal de mensagem são utilizadas para variar a duração dos pulsos individuais. Esta forma de modulação também é chamada de *modulação por largura de pulso* (PWM – do inglês, *pulse-width modulation*), ou *modulação por comprimento de pulso*. O sinal modulante pode variar no tempo da borda de subida, da borda de descida, ou em ambas as bordas. Na Fig. 5.8(c), a borda de descida é variada de acordo com o sinal de mensagem, o qual foi considerado como sendo senoidal na Fig. 5.8(a). A portadora de pulsos periódicos é mostrada na Fig. 5.8(b).

A PDM desperdiça potência, pois pulsos longos gastam uma potência considerável durante o tempo do pulso, sem acrescentar informação. Se esta potência inutilizada for subtraída do PDM, de tal forma que somente as transições no tempo sejam essencialmente preservadas, obtemos um tipo de modulação de pulso mais eficiente, chamada de modulação por posição de pulso (PPM – do inglês, *pulse-position modulation*). Na PPM, a posição de um pulso relativa ao seu tempo de ocorrência não modulado é variada de acordo com o sinal de mensagem, como ilustrado na Fig. 5.8(d) para o caso de uma modulação senoidal.

Seja T_s a duração da amostra. Utilizando a amostra $m(nT_s)$ do sinal de mensagem $m(t)$ para modular a posição do n -ésimo pulso, obtemos o sinal PPM

$$s(t) = \sum_{n=-\infty}^{\infty} g(t - nT_s - k_p m(nT_s)) \quad (5.18)$$

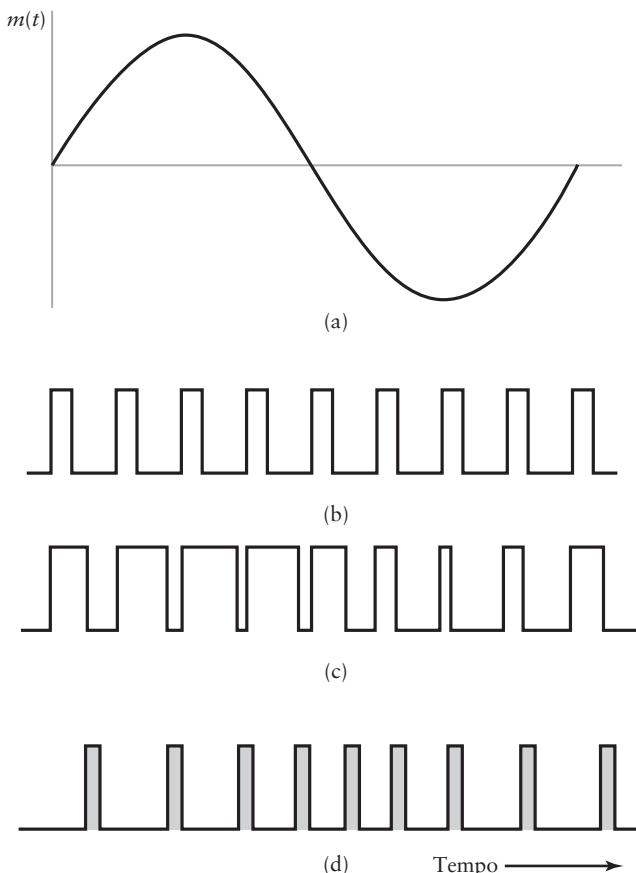


Figura 5.8 Ilustração de duas formas diferentes de modulação por tempo de pulso para o caso de uma onda modulante senoidal. (a) Onda modulante. (b) Portadora de pulso. (c) Onda PDM. (d) Onda PPM.

na qual k_p é o fator de sensibilidade do modulador por posição de pulso (em segundos por volt) e $g(t)$ representa o pulso padrão de interesse. Claramente, os pulsos diferentes que constituem o sinal PPM $s(t)$ devem ser estritamente não sobrepostos. Uma condição suficiente para esta restrição ser satisfeita é

$$g(t) = 0, \quad |t| > (T_s/2) - k_p|m(t)|_{\max} \quad (5.19)$$

a qual, por sua vez, requer que

$$k_p|m(t)|_{\max} < (T_s/2) \quad (5.20)$$

quanto mais perto $k_p|m(t)|_{\max}$ for da metade da duração de amostragem T_s , mais estreito deve ser o pulso padrão $g(t)$ para garantir que os pulsos individuais do sinal PPM $s(t)$ não interfiram uns com os outros, e maior será a largura de faixa ocupada pelo sinal PPM. Assumindo que a Eq. (5.19) é satisfeita e que não existe interferência entre pulsos adjacentes do sinal PPM $s(t)$, então as amostras do sinal $m(nT_s)$ podem ser recuperadas perfeitamente.² Além disso, se o sinal de mensagem $m(t)$ for estritamente limitado em faixa, a partir do teorema da amostragem, o sinal de mensagem $m(t)$ original pode ser recuperado do sinal PPM $s(t)$ sem distorção.

5.4 COMPLETANDO A TRANSIÇÃO DE ANALÓGICO PARA DIGITAL

Neste ponto do livro é instrutivo relembrarmos as técnicas de modulação que estudamos até este momento para a transmissão de sinais analógicos contendo informação (isto é, voz e sinais de vídeo) em um canal de comunicação, além de mostrarmos os que ainda serão considerados. As técnicas estudadas até agora são modulação de onda contínua e modulação analógica de pulso. Apesar destas duas famílias de técnicas de modulação serem de fato diferentes, elas compartilham atributos e limitações semelhantes.

Primeiro, é natural pensarmos na modulação por amplitude de pulso como a contrapartida da modulação em amplitude estudada no Capítulo 3. O que temos de contrapartida como modulação de pulso para a modulação em freqüência estudada no Capítulo 4? Na modulação em freqüência, os cruzamentos com zero da onda modulada variam no tempo de acordo com o sinal de mensagem. Na modulação por posição de pulso, as posições dos pulsos transmitidos variam no tempo de acordo com o sinal de mensagem. De forma fraca, podemos, portanto, pensar na modulação por posição de pulso como a contrapartida da modulação em freqüência.

Uma conclusão intuitiva desta analogia entre membros de modulação analógica de pulso e membros da modulação de onda contínua é que estas duas famílias de técnicas de modulação oferecem a mesma ordem de performance quando aplicadas à transmissão de sinais analógicos em canais de comunicação. Ao falarmos em performance, estamos pensando em termos da necessidade de largura de faixa de transmissão e comportamento do ruído no receptor. A característica operacional que distingue as duas famílias é que as técnicas de modulação de onda contínua operam no tempo contínuo, enquanto que a modulação analógica de pulso opera no tempo discreto.

Na mudança da modulação de onda contínua para a modulação analógica de pulso, nos movemos para o processamento de sinais em tempo discreto. Por que não ir um passo adiante e incorporar, também, a *discretização em amplitude*? De fato, isto é precisamente o que fazemos na *modulação digital de pulso*. Desta forma, temos uma nova família de técnicas de modulação para a transmissão de sinais analógicos em canais de comunicação. As vantagens oferecidas pelas técnicas de modulação digital de pulso incluem:

² A geração e detecção de ondas PPM é discutida em Haykin (1994), pp. 365-369.

1. *Performance.* Em um sistema analógico de comunicação, seja usando modulação de onda contínua ou modulação analógica de pulso, os efeitos de distorção do sinal e ruído do canal (os quais ocorrem ao longo do caminho da transmissão) são *acumulativos*. Estas fontes de degradação tendem, portanto, a se tornarem progressivamente fortes e, em última instância, sobrepujam a capacidade do sistema de comunicação de oferecer um nível aceitável de performance, da fonte ao destino. Infelizmente, a utilização de repetidores na forma de amplificadores, colocados em pontos diferentes ao longo do caminho de transmissão, é de pouca ajuda, pois o sinal de mensagem e o ruído são amplificados pelo mesmo fator. Em um forte contraste, a modulação digital de pulso permite a utilização de *repetidores regenerativos*, os quais, quando colocados ao longo do caminho de transmissão, em distâncias curtas o suficiente, podem praticamente eliminar os efeitos degradantes do ruído e da distorção do sinal.
2. *Robustez.* Ao contrário de sistemas analógicos de comunicação, um sistema digital de comunicação pode ser projetado para combater os efeitos do ruído do canal e distorção do sinal, desde que o ruído e a distorção sejam mantidos dentro de certos limites.
3. *Confiabilidade.* Sistemas digitais de comunicação podem ser projetados para serem altamente confiáveis, explorando-se técnicas poderosas de controle de erro, de tal forma que a estimativa do sinal de mensagem entregue ao usuário é praticamente indistinguível do sinal de mensagem entregue pela fonte de informação no outro lado do sistema. (Códigos de controle de erro são discutidos no Capítulo 10.)
4. *Segurança.* Pelo mesmo motivo, sistemas de comunicação digital podem ser projetados altamente seguros, explorando algoritmos de codificação baseados no processamento digital para sua implementação.
5. *Eficiência.* Sistemas de comunicação digital são inherentemente mais eficientes do que sistemas de comunicação analógicos no compromisso entre largura de faixa de transmissão e relação sinal/ruído.
6. *Integração de sistemas.* A utilização de comunicações digitais possibilita a integração digitalizada de sinais analógicos (isto é, voz e sinais de vídeo) com dados digitais de computador, o que não é possível utilizando-se comunicações analógicas.

Esta impressionante lista de vantagens fez com que as técnicas de modulação digital de pulso se tornassem a escolha para a transmissão de voz e sinais de vídeo em canais de comunicação.

Os benefícios da utilização da modulação digital de pulso, entretanto, são obtidos ao custo de aumento da complexidade do sistema. De qualquer forma, explorando-se o poder de computação de processadores digitais de sinais na forma de hardware e/ou software e a flexibilidade que esses processadores oferecem, os sistemas de comunicação digital podem ser projetados de maneira economicamente viáveis, graças aos contínuos melhoramentos em CIIs VLSI (*very-large-scale integrated*).

Agora que já identificamos a família de modulação digital de pulso como o método de escolha para comunicações, nossa próxima tarefa, neste capítulo, é descrever os três membros da família – modulação por codificação de pulso, modulação delta e modulação por codificação diferencial de pulso. O estudo da modulação por codificação de pulso ocupa as Seções 5.5 e 5.6, seguida pelos dois outros membros nas Seções 5.7 e 5.8, respectivamente. A modulação por codificação de pulso é o padrão contra o qual a modulação delta e a modulação por codificação diferencial de pulso são comparadas.

5.5 PROCESSO DE QUANTIZAÇÃO

Um sinal contínuo, tal como a voz, possui uma faixa contínua de amplitudes e, portanto, suas amostras possuem uma faixa contínua de valores. Em outras palavras, dentro de uma faixa finita de amplitudes do sinal, existe um número infinito de níveis de amplitude. Na realidade, entretanto, não é necessário transmitir as amplitudes exatas das amostras.

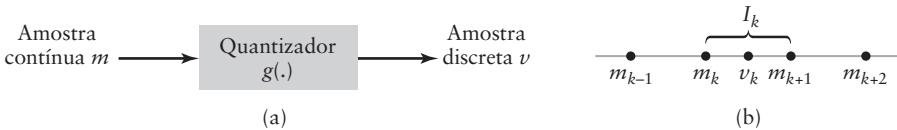


Figura 5.9 Descrição de um quantizador sem memória.

Dizemos isso porque para qualquer sentido humano (seja audição ou visão), o receptor final pode detectar apenas diferenças finitas de intensidade. Isto significa que o sinal original contínuo pode ser *aproximado* por um sinal construído pelas amplitudes discretas da modulação digital de pulso, selecionadas com base em uma função de erro mínimo. Claramente, se associarmos níveis discretos de amplitudes suficientemente próximos, podemos tornar o sinal aproximado indistinguível do sinal contínuo original, para todos os efeitos práticos. Observe, também que a quantização é um processo não reversível.

A *quantização em amplitude* é definida como o processo de transformar a amplitude da amostra $m(nT_s)$ de um sinal banda base $m(t)$ no tempo $t = nT_s$ em uma amplitude discreta $v(nT_s)$ tomada de um conjunto finito de possíveis níveis. Chamamos atenção para um processo de quantização que é sem memória e instantâneo, o que significa que a transformação no tempo $t = nT_s$ não é afetada por amostras anteriores ou posteriores do sinal de mensagem. Esta forma de quantização, apesar de não ótima, é geralmente utilizada na prática devido à sua simplicidade.

Quando lidamos com um quantizador sem memória, simplificamos a notação ignorando o índice tempo. Ou seja, utilizamos o símbolo m no lugar da amostra $m(nT_s)$, como indicado na Fig. 5.9(a). Então, como mostrado na Fig. 5.9(b), a amplitude m do sinal é especificada pelo índice k se ele estiver dentro do intervalo

$$I_k: \{m_k < m \leq m_{k+1}\}, \quad k = 1, 2, \dots, L \quad (5.21)$$

no qual L é o número total de níveis de amplitude utilizados no *quantizador*, o que se refere ao subsistema que executa o processo de quantização. As amplitudes m_k , $k = 1, 2, \dots, L$ são chamadas de *níveis de decisão* ou *limites de decisão*. Na saída do quantizador, o índice k é transformado na amplitude v_k que representa todas as amplitudes que estão dentro do intervalo I_k . As amplitudes v_k , $k = 1, 2, \dots, L$ são chamadas de *níveis de representação* ou *níveis de reconstrução*, e o espaçamento entre a representação de dois níveis adjacentes é chamada de *quantum* ou *tamanho do passo*. Portanto, a saída do quantizador v é igual a v_k se a amostra m do sinal de entrada pertencer ao intervalo I_k . O mapeamento

$$v = g(m) \quad (5.22)$$

é a *característica de quantização*. Esta característica é descrita por uma função em degrau. Quantizadores podem ser do tipo *uniforme* ou *não uniforme*. Em um quantizador uniforme, os níveis de representação são espaçados uniformemente. Caso contrário, o quantizador é *não uniforme*. Os quantizadores considerados nesta seção são do tipo uniforme. Quantizadores não uniformes são considerados na Seção 5.6. A característica de quantização também pode ser do tipo *midtread* ou *midrise*. A Fig. 5.10(a) mostra a característica entrada/saída de um quantizador uniforme do tipo *midtread*, o qual recebe este nome porque a origem está no meio do degrau da forma de onda do gráfico. A Fig. 5.10(b) mostra a característica entrada/saída correspondente de um quantizador uniforme do tipo *midrise*, no qual a origem está no meio da subida do degrau do gráfico. Note que os dois tipos de quantizador uniforme, ilustrados na Fig. 5.10, são *simétricos* com relação à origem.

5.6 MODULAÇÃO POR CODIFICAÇÃO DE PULSO

Com o processo de amostragem e quantização a nossa disposição, estamos prontos para descrever a modulação por codificação de pulso, que é a forma mais básica de modulação

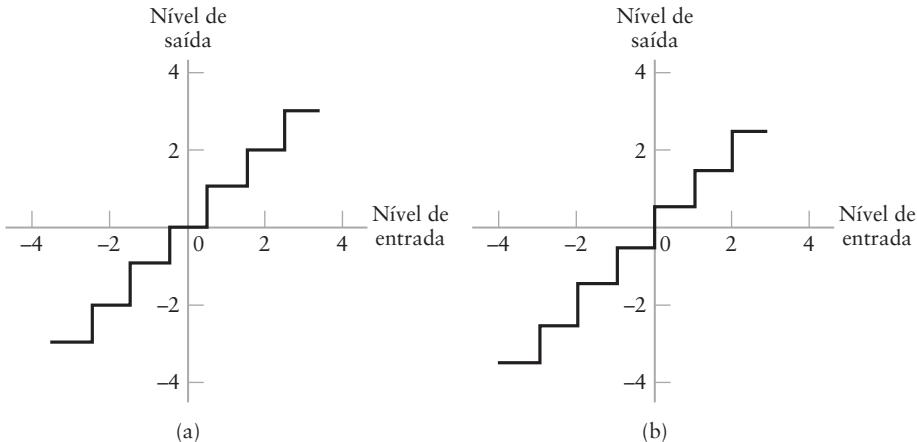


Figura 5.10 Dois tipos de quantizadores: (a) midtread e (b) midrise.

digital de pulso. Na *modulação por codificação de pulso* (PCM – do inglês, *pulse code modulation*), um sinal de mensagem é representado por uma seqüência de pulsos codificados, obtidos pela representação do sinal na forma discreta tanto no tempo quanto em amplitude.

As operações básicas realizadas no transmissor de um sistema PCM são *amostragem*, *quantização* e *codificação*, como mostrado na Fig. 5.11(a); as operações de quantização e codificação são, geralmente, executadas pelo mesmo circuito, o qual é chamado de *conversor analógico digital*.

As operações básicas no receptor são a *regeneração* de sinais distorcidos, *decodificação* e *reconstrução* do trem de amostras quantizadas, como mostrado na Fig. 5.11(c). A regeneração também ocorre em pontos intermediários ao longo do caminho de transmissão, quando necessário, como indicado na Fig. 5.11(b). Quando a multiplexação por divisão no tempo (discutida posteriormente neste capítulo) é utilizada, é necessário sincronizar o receptor com o transmissor, para que o sistema total opere satisfatoriamente. A seguir iremos descrever as operações de amostragem, quantização e codificação básicas ao sistema PCM.

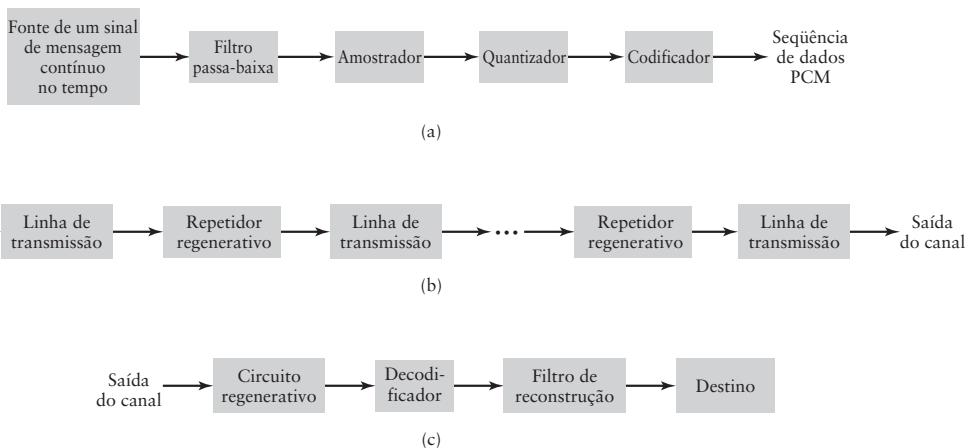


Figura 5.11 Elementos básicos de um sistema PCM. (a) Transmissor, (b) caminho de transmissão conectando o transmissor ao receptor e (c) receptor.

■ Operações no transmissor

(i) Amostragem

O sinal de entrada (banda base) é amostrado por um trem de pulsos retangulares, estreitos o suficiente para se aproximarem do processo de amostragem instantânea. Para garantir a reconstrução perfeita do sinal de mensagem no receptor, a taxa de amostragem deve de ser maior do que duas vezes a componente W do sinal de mensagem, de acordo com o teorema da amostragem. Na prática, um filtro anti-aliasing (passa-baixa) é utilizado na frente do amostrador, para excluir as freqüências maiores do que W antes da amostragem, como mostrado na Fig. 5.11(a). Portanto, a aplicação da amostragem permite reduzir o sinal de mensagem que varia continuamente (com alguma duração finita) a um número limitado de valores discretos por segundo.

(ii) Quantização não uniforme

A versão amostrada do sinal de mensagem é, então, quantizada, fornecendo uma nova representação do sinal que, agora, é discreto tanto no tempo quanto em amplitude. O processo de quantização pode seguir uma lei uniforme, como descrita na Seção 5.5. Em algumas aplicações, entretanto, é preferível utilizar uma separação variável entre os níveis de representação. Por exemplo, a faixa de tensões cobertas por sinais de voz, do pico de uma fala alta a uma passagem baixa de uma fala baixa, é da ordem de 1000 para 1. Utilizando-se um *quantizador não uniforme* com a característica do aumento do tamanho do degrau quando a separação da origem da característica entrada/saída é aumentada, os últimos degraus do quantizador podem assumir tais excursões que possibilitam o sinal de voz excursionar em grandes faixas de amplitude que ocorrem com baixa freqüência. Em outras palavras, passagens fracas que precisam de mais proteção são favorecidas ao custo de altas passagens. Desta forma, uma precisão percentual praticamente uniforme é atingida em grande parte da faixa de amplitude do sinal de entrada, utilizando menos degraus do que seriam necessários se utilizássemos um quantizador uniforme.

A utilização de um quantizador não uniforme é equivalente a passar o sinal de mensagem em um *compressor* e, então, aplicar o sinal comprimido a um quantizador uniforme. Uma forma particular de lei de compressão utilizada na prática é a chamada lei μ ,³ definida por

$$|v| = \frac{\log(1 + \mu|m|)}{\log(1 + \mu)} \quad (5.23)$$

na qual o logaritmo é o logaritmo natural, m e v são as tensões normalizadas de entrada e saída, respectivamente, e μ é uma constante positiva. Por conveniência de representação, a entrada do quantizador e sua saída são *normalizadas*, de forma a ocupar uma faixa de valores sem dimensão, de zero a um, como mostrado na Fig. 5.12(a). Neste caso, traçamos a lei μ variando o parâmetro μ . Valores práticos de μ tendem a estar na vizinhança de 255. O caso de quantização uniforme corresponde a $\mu = 0$. Para um dado valor de μ , a inclinação recíproca da curva de compressão, a qual define o quantum de degrau, é dada pela derivada de $|m|$ com respeito a $|v|$, ou seja,

$$\frac{d|m|}{d|v|} = \frac{\log(1 + \mu)}{\mu} (1 + \mu|m|) \quad (5.24)$$

³ A lei μ , utilizada para a compressão de sinais é descrita em Smith (1957). Esta lei de compressão é utilizada nos Estados Unidos, Canadá e Japão. Na Europa, a lei A é utilizada para a compressão do sinal. Esta segunda lei de compressão é descrita em Cattermole (1969, pp. 133-140). Para a discussão sobre a lei μ e lei A, veja também o artigo de Kaneko (1970).

Vemos, portanto, que a lei μ não é nem estritamente linear nem estritamente logarítmica, mas é aproximadamente linear em baixos níveis de entrada, correspondendo a $\mu|m| \ll 1$ e aproximadamente logarítmica em altos níveis de entrada, correspondendo a $\mu|m| \gg 1$.

Outra lei de compressão utilizada na prática é a chamada *lei A*, definida por

$$|v| = \begin{cases} \frac{A|m|}{1 + \log A}, & 0 \leq |m| \leq \frac{1}{A} \\ \frac{1 + \log(A|m|)}{1 + \log A}, & \frac{1}{A} \leq |m| \leq 1 \end{cases} \quad (5.25)$$

a qual é mostrada na Fig. 5.12(b). Valores típicos de A utilizados na prática tendem a estar na proximidade de 100. O caso de quantização uniforme corresponde a $A = 1$. A inclinação recíproca desta segunda curva de compressão é dada pela derivada de $|m|$ com respeito a $|v|$, mostrado por

$$\frac{d|m|}{d|v|} = \begin{cases} \frac{1 + \log A}{A}, & 0 \leq |m| \leq \frac{1}{A} \\ (1 + \log A)|m|, & \frac{1}{A} \leq |m| \leq 1 \end{cases} \quad (5.26)$$

Da primeira linha da Eq. (5.26), podemos inferir que o quantum de degrau sob o segmento linear central, o qual possui efeito dominante em pequenos sinais, é diminuído pelo fator $A/(1 + \log A)$. Isto ocorre, normalmente, em 25 dB na prática, quando comparado com a quantização uniforme.

- **PROBLEMA 5.9** Usando as Eqs. (5.23) e (5.25), respectivamente, obtenha as características de inclinação das Eqs. (5.24) e (5.26).

(iii) Codificação

Combinando-se o processo de amostragem e quantização, a especificação do sinal contínuo de mensagem (banda base) se torna um conjunto limitado de valores discretos, mas não na forma mais adequada para a transmissão em uma linha cabeada ou por rádio. Para explorar as vantagens da amostragem e quantização com o propósito

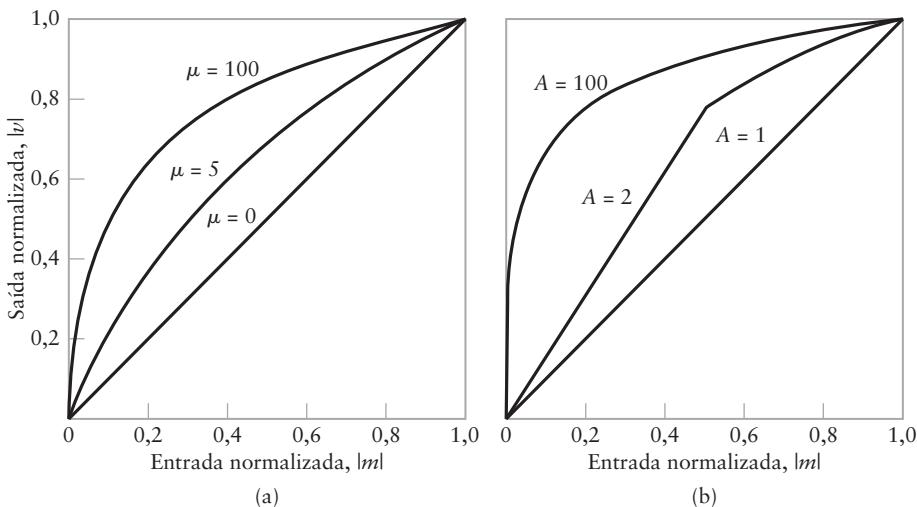


Figura 5.12 Leis de compressão. (a) Lei μ . (b) Lei A .

de tornar o sinal transmitido mais robusto ao ruído, interferência ou outras degradações do canal, podemos precisar utilizar um *processo de codificação* para transformar o conjunto discreto de valores amostrados em uma forma mais apropriada de sinal. Qualquer plano de representação de um conjunto discreto de valores em uma forma particular de eventos discretos é chamado de *código*. Um dos eventos discretos em um código é chamado de *elemento do código* ou *símbolo*. Por exemplo, a presença ou ausência de um pulso é um símbolo. Um arranjo particular de símbolos utilizados em um código para representar um valor único do conjunto discreto é chamado de *palavra de código* ou *caractere*.

Em um *código binário*, cada símbolo pode ser de um dentre dois valores distintos, tal como pulso negativo ou pulso positivo. Os dois símbolos de um código binário são, geralmente, representados por 0 e 1. Na prática, um código binário é preferido sobre outros códigos (por exemplo, o código trinário) por dois motivos:

1. A vantagem máxima sobre os efeitos de ruído em uma mídia de comunicação é obtida utilizando um código binário, pois o símbolo binário suporta um relativo nível alto de ruído.
2. O código binário é fácil de ser gerado e regenerado.

Suponha que, em um código binário, cada palavra de código consista de R bits. O bit é um acrônimo para *dígito binário*. Então R representa o número de *bits por amostra*. Logo, utilizando este código, podemos representar um total de 2^R números distintos. Por exemplo, uma amostra quantizada em um de 256 níveis pode ser representada por um código com 8 bits.

Existem várias formas de estabelecer uma correspondência de um-para-um entre os níveis de representação e palavras de código. Um método conveniente é expressar o número ordinal do nível de representação como um número binário. No sistema de numeração binário, cada dígito possui um peso que é uma potência de 2, como ilustrado na Tabela 5.2 para o caso de quatro bits por amostra (isto é, $R = 4$).

TABELA 5.2 Sistema de numeração binária para $R = 4$

Número ordinal do nível de representação	Número do nível expresso com a soma de potência de 2	Número binário
0		0000
1	2^0	0001
2	2^1	0010
3	$2^1 + 2^0$	0011
4	2^2	0100
5	$2^2 + 2^0$	0101
6	$2^2 + 2^1$	0110
7	$2^2 + 2^1 + 2^0$	0111
8	2^3	1000
9	$2^3 + 2^0$	1001
10	$2^3 + 2^1$	1010
11	$2^3 + 2^1 + 2^0$	1011
12	$2^3 + 2^2$	1100
13	$2^3 + 2^2 + 2^0$	1101
14	$2^3 + 2^2 + 2^1$	1110
15	$2^3 + 2^2 + 2^1 + 2^0$	1111

■ Regeneração ao longo do caminho de transmissão

A característica mais importante de um sistema PCM está na capacidade de *controlar* os efeitos de distorção e ruído, produzidos pela transmissão de um sinal PCM em um canal. Esta capacidade é implementada pela reconstrução do sinal PCM através de uma cadeia de *repetidores regenerativos* adequadamente espaçados ao longo do caminho de transmissão. Como ilustrado na Fig. 5.13, existem três funções básicas executadas por um repetidor regenerativo: *equalização*, *sincronização* e *tomada de decisão*. O equalizador formata os pulsos recebidos de forma a compensar os efeitos de distorção de amplitude e fase produzidas pelas características de transmissão do canal. O circuito de temporização fornece um trem de pulso periódico, obtido dos pulsos do receptor. Isto é realizado para a amostragem renovada dos pulsos equalizados em instantes de tempo nos quais a relação sinal/ruído é máxima. A amostra extraída é comparada com um *limiar* pré-determinado em um dispositivo de tomada de decisão. Em cada intervalo de bit, uma decisão é tomada indicando se o símbolo recebido é 1 ou 0, se o limiar é excedido ou não. Se o limiar for excedido, um novo e limpo pulso representando o símbolo 1 é transmitido ao próximo repetidor. Caso contrário, um novo e limpo pulso representando o símbolo 0 é transmitido. Desta forma, o acúmulo de distorção e ruído entre os repetidores espaçados é removido, desde que o distúrbio não seja grande o suficiente para causar um erro no processo de tomada de decisão. Idealmente, exceto pelo atraso, o sinal regenerado é exatamente o mesmo do sinal contendo a informação que foi originalmente transmitido. Na prática, entretanto, o sinal regenerado difere do sinal original por dois motivos principais:

1. A presença inevitável do ruído do canal e interferência fazem com que o repetidor tome uma decisão errada ocasionalmente, introduzindo, pois, *erros de bit* no sinal regenerado.
2. Se o espaçamento entre os pulsos recebidos desviar de seu valor atribuído, um *jitter* é introduzido na posição do pulso regenerado, causando, portanto, distorção.

■ Operações no receptor

(i) Decodificação e expansão

A primeira operação no receptor é a *regeneração* (isto é, a formatação e limpeza) dos pulsos recebidos uma última vez. Estes pulsos limpos são, então, reagrupados em palavras de código e decodificados (isto é, mapeados de volta) em um sinal PAM quantizado. O processo de decodificação envolve a geração de um pulso cuja amplitude é a soma linear de todos os pulsos da palavra de código. Cada pulso é ponderado por seu valor posicional ($2^0, 2^1, 2^2, 2^3, \dots, 2^{R-1}$) do código, no qual R é o número de bits por amostra.

A seqüência de amostras decodificadas é representada como uma *estimativa* da seqüência das amostras comprimidas produzidas pelo quantificador no transmissor. Utilizamos o termo “estimativa” para enfatizar o fato de que não existe como o receptor compensar a aproximação introduzida no sinal transmitido pelo quantificador. Além disso, outras fontes de ruído incluem erros de bits e jitter produzidos ao longo do caminho de transmissão. Para recuperar a seqüência de amostras decodificadas a seu nível relativo correto, devemos, obviamente, utilizar um subsistema no receptor

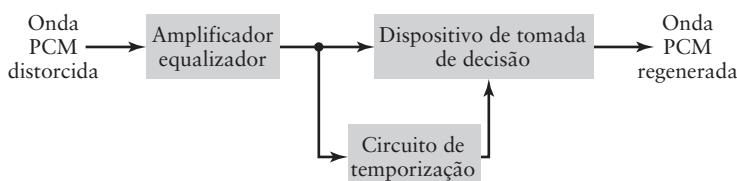


Figura 5.13 Diagrama em blocos de um repetidor regenerativo.

com característica complementar à do compressor utilizado no transmissor. Tal subsistema é chamado de *expansor*. Idealmente, as leis de compressão e expansão são exatamente *inversas*, ou seja, exceto pelo efeito de quantização, a saída do expansor é igual à entrada do compressor, se os dois dispositivos forem conectados diretamente. A combinação do compressor e expansor é chamada de *compander*.

(ii) Reconstrução

A operação final no receptor é recuperar o sinal de mensagem. Esta operação é realizada passando a saída do expansor em um filtro passa-baixa de reconstrução, cuja frequência de corte é igual à largura de faixa da mensagem. A recuperação do sinal de mensagem significa estimativa ao invés de reconstrução exata.

Um último comentário é necessário. O termo “modulação” na modulação por codificação de pulso é errôneo. Na realidade, a modulação por codificação de pulso é uma estratégia de codificação de fonte, pela qual um sinal analógico emitido por uma fonte é convertido em uma forma digital. A transmissão do dado digital produzido é outra questão, cujo tratamento será realizado no Capítulo 6.

5.7 MODULAÇÃO DELTA

Da discussão apresentada na Seção 5.6, vimos que o projeto de um sistema de modulação por codificação de pulso envolve várias operações, as quais tendem a resultar em uma implementação prática cara. Para simplificar o projeto do sistema, podemos utilizar outra técnica de modulação digital de pulso chamada de modulação delta, a qual é considerada nesta seção.

■ Considerações básicas

Na *modulação delta* (DM – do inglês, *delta modulation*), um sinal de mensagem de entrada é superamostrado (isto é, em uma taxa muito superior à taxa de Nyquist) para aumentar, propósitadamente, a *correlação* entre amostras adjacentes do sinal. O aumento da correlação é feito de forma a permitir a utilização de uma estratégia de quantização simples para a construção do sinal codificado.

Nesta forma básica, a DM fornece uma *aproximação em degrau* para a versão superamostrada do sinal de mensagem. Ao contrário do PCM, a diferença entre o sinal de entrada e sua aproximação é quantizada em apenas dois níveis, representados por $\pm\Delta$ e correspondendo a *diferenças* positiva e negativa. Portanto, se a aproximação está abaixo do sinal de entrada em qualquer instante amostrado, ela é aumentada por Δ . Por outro lado, se a aproximação está acima do sinal, ela é diminuída por Δ . Desde que o sinal de entrada não varie muito rapidamente de uma amostra para outra, vemos que a aproximação em degrau permanece dentro de $\pm\Delta$ do sinal de entrada.

Representamos o sinal de entrada por $m(t)$ e sua aproximação em degrau por $m_q(t)$. O princípio básico da modulação delta pode, então, ser formalizado pelo seguinte conjunto de três relações em tempo discreto:

$$e(nT_s) = m(nT_s) - m_q(nT_s - T_s) \quad (5.27)$$

$$e_q(nT_s) = \Delta \text{sgn}[e(nT_s)] \quad (5.28)$$

$$m_q(nT_s) = m_q(nT_s - T_s) + e_q(nT_s) \quad (5.29)$$

Na qual T_s é o período de amostragem; $e(nT_s)$ é o *sinal de erro* representando a diferença entre o valor amostrado atual $m(nT_s)$ do sinal de entrada e sua última aproximação, ou seja, $m(nT_s) - m_q(nT_s - T_s)$; $e_q(nT_s)$ é a versão quantizada de $e(nT_s)$; $\text{sgn}[\cdot]$ é a função sinal, assumindo que o valor +1 ou -1. A saída $e_q(nT_s)$ do quantizador é finalmente codificada para produzir o dado DM desejado.

A Fig. 5.14(a) ilustra a forma pela qual a onda em degrau segue aproximadamente as variações do sinal de entrada $m(t)$ de acordo com as Eqs. (5.27) a (5.29). A Fig. 5.19(b) mostra a seqüência binária correspondente à saída do modulador delta. É evidente que no sistema de modulação delta, a taxa da transmissão de informação é simplesmente igual à taxa de amostragem $f_s = 1/T_s$.

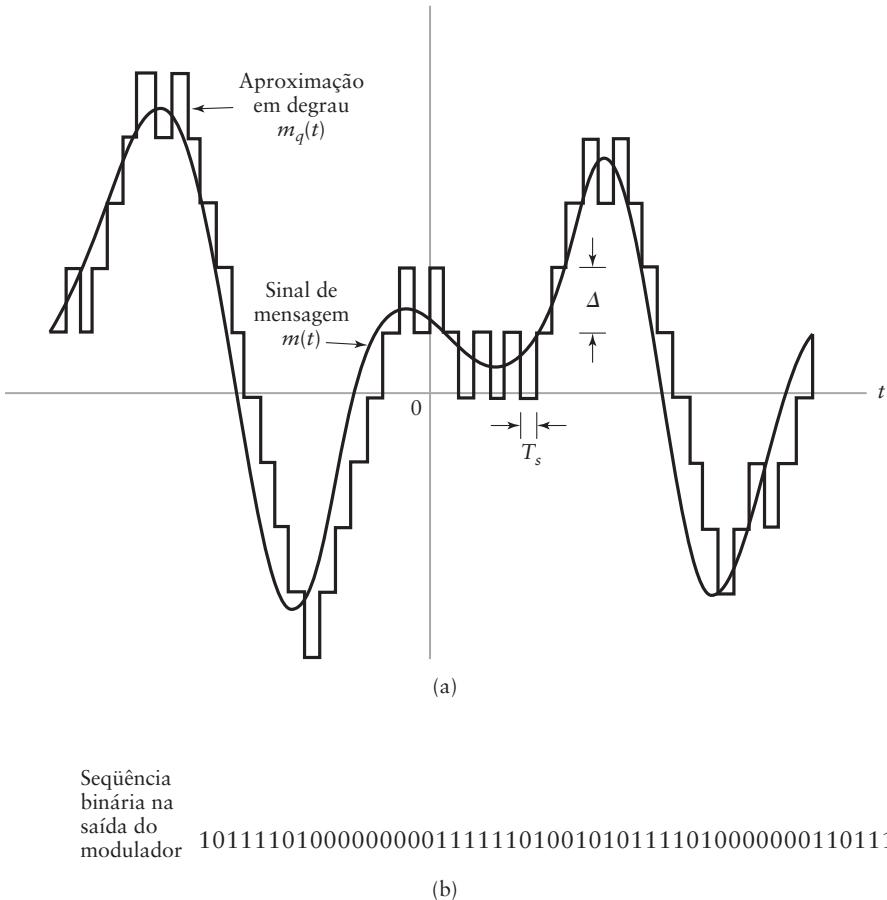


Figura 5.14 Ilustração da modulação delta. (a) Forma de onda $m(t)$ analógica e sua aproximação em degrau $m_q(t)$. (b) Seqüência binária na saída do modulador.

■ Detalhes do sistema

A principal vantagem da modulação delta é sua *simplicidade*. Ela pode ser implementada aplicando a versão amostrada do sinal de mensagem de entrada a um transmissor constituído por um *comparador*, um *quantizador* e um *acumulador*, conectados como mostrado na Fig. 5.15(a). Os detalhes do transmissor são obtidos diretamente das Eqs. (5.27) a (5.29). O comparador calcula a diferença entre duas entradas. O quantizador é constituído por um *limitador* cuja característica entrada/saída é uma versão escalonada da função sinal. O acumulador opera com a saída do quantizador para produzir uma aproximação do sinal de mensagem.

A Eq. (5.29) é chamada de *equação diferença de ordem um*, a ordem se refere ao fato da amostra atual $m_q(nT_s)$ diferir de uma amostra passada $m_q(nT_s - T_s)$ por um total igual ao erro de quantização $e_q(nT_s)$. Assumindo que o processo de acumulação começa no tempo zero, a solução desta equação leva ao resultado aproximado

$$\begin{aligned}
 m_q(nT_s) &= m_q(nT_s - T_s) + e_q(nT_s) \\
 &= m_q(nT_s - 2T_s) + e_q(nT_s - T_s) + e_q(nT_s) \\
 &\quad \vdots \\
 &= \sum_{i=1}^n e_q(iT_s)
 \end{aligned} \tag{5.30}$$

na qual $e_q(nT_s)$ é relacionada à amostra $m(nT_s)$ da mensagem pelas Eqs. (5.27) e (5.28).

Portanto, no instante de amostragem nT_s , o acumulador aumenta a aproximação pelo incremento Δ na direção positiva ou negativa. Se o sinal de entrada $m(nT_s)$ for maior do que a aproximação mais recente $m_q(nT_s)$, um incremento positivo $+\Delta$ é aplicado à aproximação. Se, por outro lado, o sinal de entrada for menor, um incremento negativo $-\Delta$ é aplicado à aproximação.

Desta forma, o acumulador faz o melhor possível para *seguir* as amostras de entrada um degrau (de amplitude $+\Delta$ ou $-\Delta$) por vez.

No receptor mostrado na Fig. 5.15(b), a aproximação em degrau $m_q(t)$ é reconstruída passando a seqüência de pulsos positivos ou negativos, produzidas na saída do decodificador, através de um acumulador, de forma similar à utilizada no transmissor. O ruído fora-de-faixa de quantização apresentado na forma de onda $m_q(t)$ em degrau em alta freqüência é rejeitado passando $m_q(t)$ em um filtro, como mostrado na Fig. 5.15(b). O filtro é do tipo passa-baixa, com largura de faixa igual à largura de faixa da mensagem original.

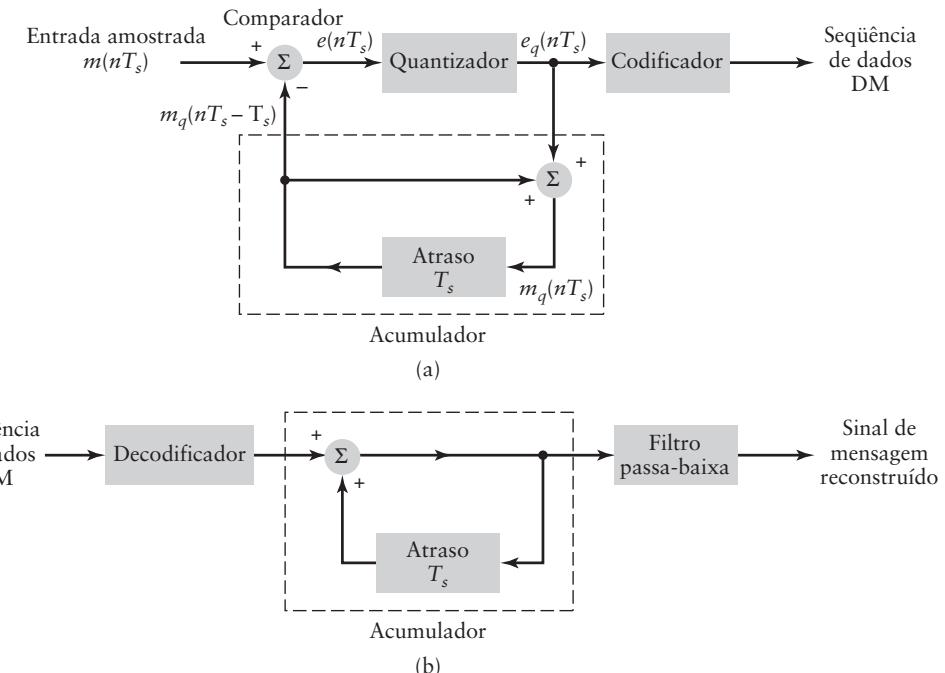


Figura 5.15 Sistema DM: (a) transmissor e (b) receptor.

■ Erros de quantização

A modulação delta é sujeita a dois tipos de erro de quantização: (1) distorção de sobre-inclinação e (2) ruído granular. Iremos discutir, primeiro, a causa da distorção de sobre-inclinação e, então, a do ruído granular.

Observe que a Eq. (5.29) é uma *integração digital equivalente*, no sentido de que ela representa o acúmulo de incrementos positivos e negativos de amplitude Δ . Além disso, representando o erro de quantização por $q(nT_s)$, mostrado por

$$m_q(nT_s) = m(nT_s) + q(nT_s) \quad (5.31)$$

observamos da Eq. (5.27) que a entrada do quantizador é

$$e(nT_s) = m(nT_s) - m(nT_s - T_s) - q(nT_s - T_s) \quad (5.32)$$

Portanto, exceto pelo erro atrasado de quantização $q(nT_s - T_s)$, a entrada do quantizador é uma *diferença atrasada de primeira ordem* do sinal de entrada, a qual pode ser vista como a aproximação digital da derivada do sinal de entrada ou, de forma equivalente, como o inverso do processo de integração digital. Se considerarmos, agora, a inclinação máxima do sinal original de mensagem $m(t)$, é claro que, para que a seqüência de amostras quantizadas $\{m_q(nT_s)\}$ aumente na mesma taxa da seqüência de amostras de entrada $\{m(nT_s)\}$ na região de inclinação máxima de $m(t)$, precisamos que a condição

$$\frac{\Delta}{T_s} \geq \max \left| \frac{dm(t)}{dt} \right| \quad (5.33)$$

seja satisfeita. Caso contrário, veremos que o degrau Δ é muito pequeno para que a aproximação em degrau $m_q(t)$ siga o segmento do sinal original de mensagem $m(t)$, fazendo com que $m_q(t)$ esteja atrás de $m(t)$, como ilustrado na Fig. 5.16. Esta condição é chamada de *sobre-inclinação*. Desta forma, o erro de quantização resultante é chamado de distorção (*ruído*) de sobre-inclinação. Note que, como a inclinação máxima da aproximação em degrau $m_q(t)$ é fixada pelo tamanho do degrau Δ , o aumento ou diminuição em $m_q(t)$ tende a ocorrer ao longo de linhas retas, como ilustrado no começo da curva da Fig. 5.16. Por esta razão, o modulador delta que utiliza um valor fixo Δ para o passo é geralmente chamado de *modulador delta linear*.

Em contraste com a distorção de sobremodulação, o *ruído granular* ocorre quando o tamanho do passo Δ é muito grande quando comparado com a característica de inclinação local do sinal original de mensagem $m(t)$. Esta segunda situação faz com que a aproximação em degrau $m_q(t)$ oscile ao redor de um segmento relativamente plano de $m(t)$, o qual é ilustrado no final da curva da Fig. 5.16. O ruído granular na modulação delta pode ser visto como o análogo do ruído de quantização na modulação por codificação de pulso.

- **PROBLEMA 5.10** O melhor que um sistema DM linear pode fazer é fornecer um compromisso entre a distorção de sobre-inclinação e o ruído granular. Justifique esta afirmativa. ◀

Desta discussão vemos que existe a necessidade de um grande passo para acomodar uma faixa dinâmica grande, enquanto que um pequeno passo é necessário para a representação adequada de sinais de baixo nível. É evidente, portanto, que se quisermos escolher

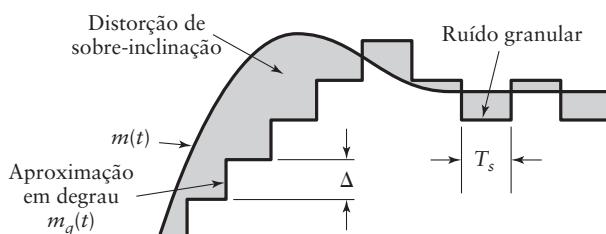


Figura 5.16 Ilustração dos erros de quantização, distorção de sobre-inclinação e ruído granular na modulação delta.

um tamanho de passo ótimo para minimizar a potência média⁴ do erro de quantização em um modulador delta, precisamos fazer com que o sistema DM seja adaptativo. Esta condição, por sua vez, significa que o tamanho do degrau precisa variar de acordo com o sinal de mensagem de entrada.

■ Modulação delta-sigma

Como mencionado previamente, a entrada do quantizador em uma forma convencional de modulação delta pode ser vista como uma aproximação de uma derivada do sinal de mensagem de entrada. Este comportamento resulta em uma desvantagem da modulação delta, pois distúrbios de transmissão, tais como ruído, resultam em um erro acumulativo no sinal demodulado. Esta desvantagem pode ser superada pela integração do sinal de mensagem antes da modulação delta. A utilização da integração também possui outros efeitos benéficos:

- ▶ O conteúdo de baixa freqüência do sinal de entrada é pré-enfatizado.
- ▶ A correlação entre amostras adjacentes da entrada do modulador delta é aumentada, o que tende a aumentar a performance total do sistema reduzindo a potência média do sinal de erro na entrada do quantizador.
- ▶ O projeto do receptor é simplificado.

O sistema de modulação delta que incorpora a integração em sua entrada é chamado de *modulação delta-sigma* (D-ΣM). Para ser mais preciso, entretanto, ela deve ser chamada de *modulação sigma-delta*, pois a integração é, de fato, realizada antes da modulação delta. E qualquer forma, a primeira terminologia é a mais utilizada na literatura.

A Fig. 5.17(a) mostra um diagrama em blocos de um sistema de modulação delta-sigma. Neste diagrama, o sinal de mensagem $m(t)$ é definido em sua forma contínua no tempo, o que significa que o modulador de pulso é constituído, agora, por um limitador seguido por um multiplicador. Este último componente é alimentado por um gerador de pulso externo (clock) para produzir um sinal codificado em um bit. A utilização da integração na entrada do transmissor requer, claramente, uma ênfase inversa do sinal – chamada de diferenciação – no receptor. A necessidade desta diferenciação é, entretanto, eliminada, devido ao seu cancelamento pela integração no receptor DM convencional. Portanto, o receptor de um sistema de modulação delta-sigma é constituído apenas de um filtro passa-baixa, como indicado na Fig. 5.17(a).

Além disso, observe que a integração é, basicamente, uma operação linear. Desta forma, podemos simplificar o projeto do transmissor combinando os integradores 1 e 2 da Fig. 5.17(a) em um único integrador posicionado após o comparador, como mostrado na Fig. 5.17(b).

Esta última forma de implementar a modulação delta-sigma é não somente mais simples do que a da Fig. 5.17(a), mas também fornece uma interpretação interessante da modulação delta-sigma como uma versão “suavizada” da modulação por codificação de pulso de 1 bit. Neste contexto, suavização se refere ao fato da saída do comparador ser integrada antes da quantização, e o termo modulação por codificação de pulso de 1 bit simplesmente reafirma o fato do quantizador ser constituído por um limitador com apenas dois níveis de representação.

5.8 MODULAÇÃO POR CODIFICAÇÃO DIFERENCIAL DE PULSO

Para outra forma de modulação digital de pulso, observamos que quando um sinal de voz ou vídeo é amostrado a uma taxa um pouco mais alta do que a taxa de Nyquist, o sinal

⁴ Em termos estatísticos, a potência média de um processo aleatório (exemplificado pelo erro de quantização) é igual ao valor de média quadrática do processo. Esta questão é discutida no Capítulo 8.

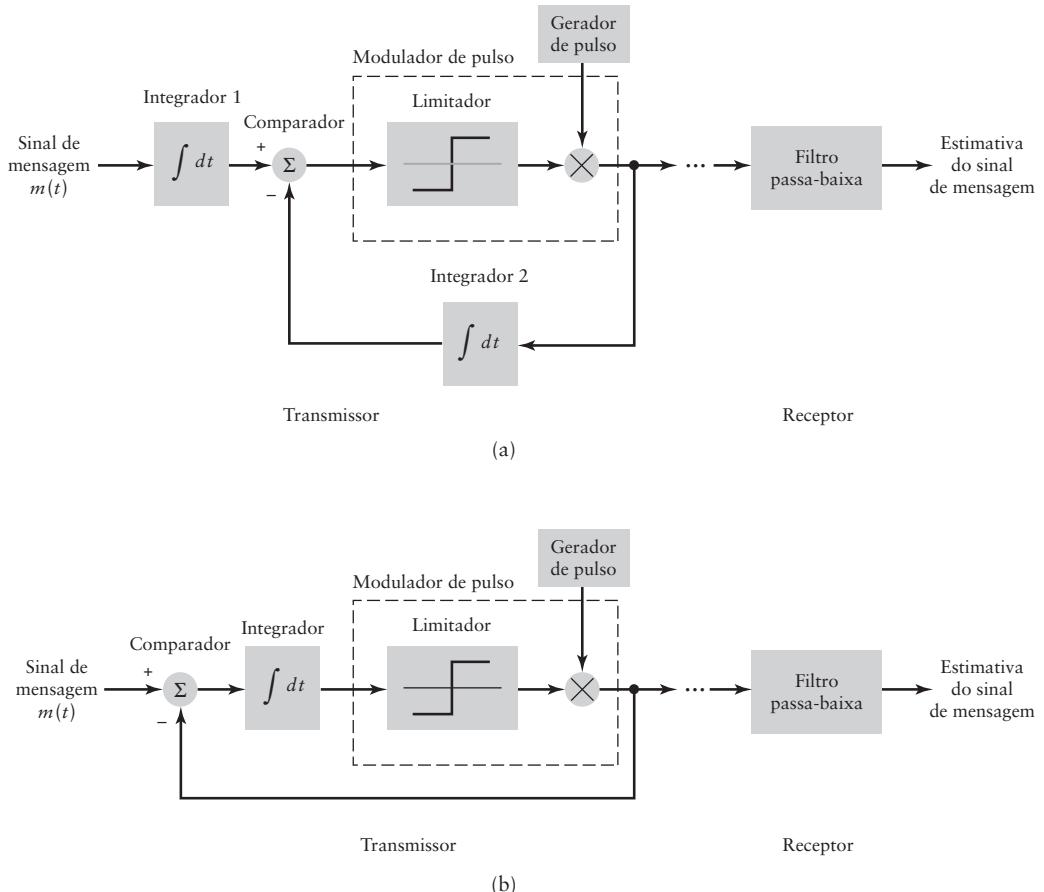


Figura 5.17 Duas versões equivalentes do sistema de modulação delta-sigma: O sistema mostrado na parte (b) da figura é uma versão simplificada do sistema na parte (a).

amostrado resultante exibe um alto grau de correlação entre amostras adjacentes. O significado desta alta correlação é que, na média, o sinal não varia rapidamente de uma amostra para a próxima, fazendo com que a diferença entre amostras adjacentes tenha uma potência média que é menor do que a potência média do próprio sinal. Quando estas amostras altamente correlacionadas são codificadas como um sistema PCM padrão, o sinal codificado resultante possui *informação redundante*. A redundância significa que símbolos que não são absolutamente essenciais à transmissão de informação são gerados como resultado do processo de codificação. Removendo esta redundância antes da codificação obtemos um sinal codificado mais *eficiente*, comparado ao PCM.

Agora, se conhecermos uma parte suficiente de um sinal redundante, podemos *inferir* o resto ou, ao menos, fazer uma estimativa altamente provável. Em particular, se soubermos o comportamento passado de um sinal até um certo ponto no tempo, é possível fazer alguma inferência sobre seus valores futuros, este processo é chamado de *predição*. Suponha, então, que um sinal de mensagem $m(t)$ é amostrado a uma taxa $f_s = 1/T_s$ para produzir uma seqüência de amostras correlacionadas separadas por T_s segundos. Esta seqüência é representada por $\{m(nT_s)\}$. O fato de ser possível prever os valores futuros do sinal $m(t)$ é a motivação para o esquema de *quantização diferencial* mostrado na Fig. 5.18(a). Neste esquema, o sinal de entrada do quantificador é definido por

$$e(nT_s) = m(nT_s) - \hat{m}(nT_s) \quad (5.34)$$

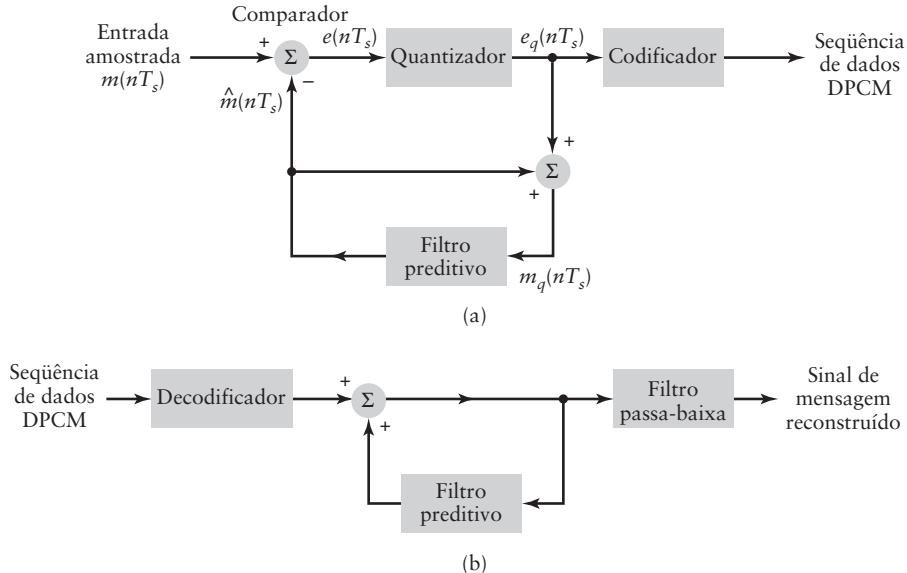


Figura 5.18 Sistema DPCM: (a) Transmissor e (b) receptor.

o qual é a diferença entre a amostra de entrada $m(nT_s)$ e a sua previsão, representada por $\hat{m}(nT_s)$. Este valor predito é produzido por um *filtro de predição*, cuja entrada, como iremos ver, consiste da versão quantizada de $m(nT_s)$. O sinal diferença $e(nT_s)$ é chamado de *erro de predição*, pois ele é o total pelo qual o filtro de predição falha em prever exatamente o sinal de mensagem de entrada. Uma abordagem simples, mas eficaz, para a implementação do filtro de predição é utilizar um *filtro de atraso de linha* ou um *filtro em tempo discreto*, com o atraso básico igual ao período de amostragem. O diagrama em blocos deste filtro é mostrado na Fig. 5.19, a qual mostra que a previsão $\hat{m}(nT_s)$ é modelada como a combinação linear de p valores passados da versão quantizada de $m(nT_s)$, na qual p é a *ordem de predição*.

Codificando a saída do quantizador da Fig. 5.18(a), obtemos a variação do PCM chamada de *modulação por codificação diferencial de pulso* (DPCM, do inglês *differential pulse-code modulation*). Este é o sinal codificado utilizado na transmissão.

A saída do quantizador pode ser descrita como

$$e_q(nT_s) = e(nT_s) + q(nT_s) \quad (5.35)$$

na qual $q(nT_s)$ é o erro de quantização. De acordo com a Fig. 5.18(a), a saída do quantizador $e_q(nT_s)$ é somada com o valor de previsão $\hat{m}(nT_s)$ para produzir a entrada do filtro de predição

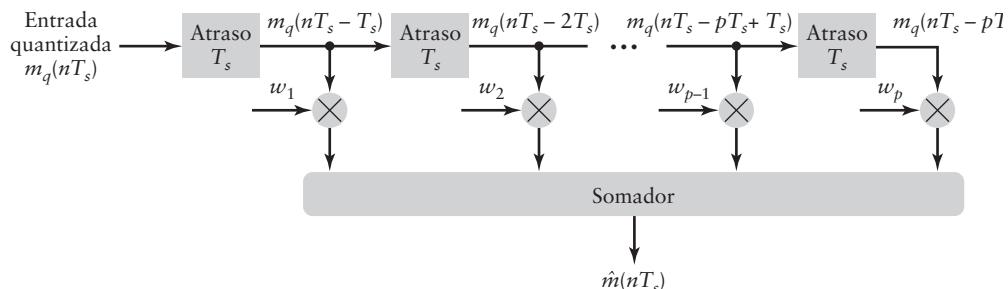


Figura 5.19 Filtro de linha de atraso utilizado como filtro de predição.

$$m_q(nT_s) = \hat{m}(nT_s) + e_q(nT_s) \quad (5.36)$$

Substituindo a Eq. (5.35) na Eq. (5.36), obtemos

$$m_q(nT_s) = \hat{m}(nT_s) + e(nT_s) + q(nT_s) \quad (5.37)$$

Entretanto, da Eq. (5.34) observamos que o termo de soma $\hat{m}(nT_s) + e(nT_s)$ é igual ao sinal de mensagem amostrado $m(nT_s)$. Portanto, podemos reescrever a Eq. (5.37) como

$$m_q(nT_s) = m(nT_s) + q(nT_s) \quad (5.38)$$

a qual representa a versão quantizada da amostra da mensagem $m(nT_s)$. Ou seja, independente das propriedades do filtro de predição, o sinal quantizado $m_q(nT_s)$ na entrada do filtro de predição difere do sinal de mensagem amostrado $m(nT_s)$ pelo erro de quantização $q(nT_s)$. Desta forma, se a predição for boa, a potência média do erro de predição $e(nT_s)$ será menor do que a potência média de $m(nT_s)$, de tal forma que o quantizador com um dado número de níveis pode ser ajustado para produzir um erro de quantização com a menor potência média que poderia ser possível se $m(nT_s)$ fosse quantizado diretamente utilizando PCM.

O receptor para a reconstrução da versão quantizada do sinal de mensagem é mostrado na Fig. 5.18(b). Ele é constituído por um decodificador para reconstruir o sinal de erro quantizado. A versão quantizada da entrada original é reconstruída da saída do decodificador utilizando o mesmo filtro de predição do transmissor da Fig. 5.18(a). Na ausência de ruído de canal, vemos que o sinal codificado na entrada do receptor é idêntico ao sinal codificado na saída do transmissor. Desta forma, a saída correspondente do receptor é igual a $m_q(nT_s)$, a qual difere do sinal de entrada original $m(nT_s)$ apenas pelo erro de quantização $q(nT_s)$ inserido como resultado da quantização do erro de predição $e(nT_s)$. Finalmente, uma estimativa do sinal de mensagem original $m(t)$ é obtida passando a sequência $m_q(nT_s)$ através de um filtro de reconstrução passa-baixa.

A partir desta análise, observamos que em um ambiente sem ruído os filtros de predição do transmissor e receptor operam na mesma sequência de amostras, $\{m_q(nT_s)\}$. Com este objetivo em mente, uma malha de realimentação é adicionada ao quantizado do transmissor, como mostrado na Fig. 5.18(a).

A modulação por codificação diferencial de pulso inclui a modulação delta como um caso especial. Em particular, quando comparamos o sistema DPCM da Fig. 5.18 com o sistema DM da Fig. 5.15, vemos que eles são, basicamente, similares exceto por duas importantes diferenças:

- A utilização do quantizador de um bit (dois níveis), no sistema DM.
- Substituição do filtro de predição no DPCM pelo elemento de atraso simples (isto é, ordem zero de predição).

Em outras palavras, DM é a versão de 1 bit da DPCM. Note, entretanto, que ao contrário do sistema PCM, os transmissores do DPCM e DM envolvem a utilização de uma *realimentação*.

Enquanto o ruído for considerado, podemos finalmente fazer duas afirmativas:

1. DPCM, tal como DM, é sujeito à distorção de sobre-inclinação sempre que o sinal de entrada mudar mais rapidamente do que o filtro de predição é capaz de acompanhar.
2. Tal como o PCM, o DPCM apresenta ruído de quantização.

- **PROBLEMA 5.11** Justifique estas duas afirmativas sobre a fonte de ruído em um sistema DPCM.

5.9 CÓDIGO DE LINHA

Na realidade, PCM, DM e DPCM representam estratégias diferentes para a codificação de fonte, pela qual um sinal analógico é convertido em uma forma digital. Entretanto, todos

os três compartilham uma característica comum: uma vez que a seqüência binária de 1s e 0s por produzida, um *código de linha* é necessário para a representação elétrica da seqüência binária. Existem vários códigos de linha que podem ser utilizados para esta representação, resumidos a seguir:

1. *Sinalização liga-desliga (on-off)*, na qual o símbolo 1 é representado pela transmissão de um pulso de amplitude constante para a duração do símbolo e o símbolo 0 é representado pela ausência do pulso, como mostrado na Fig. 5.20(a).
2. *Sinalização sem retorno para o zero (NRZ, nonreturn-to-zero)*, na qual os símbolos 1 e 0 são representados por pulsos com amplitude positiva e negativa igual, como ilustrado na Fig. 5.20(b).
3. *Sinalização com retorno para zero (RZ, return-to-zero)*, na qual o símbolo 1 é representado por um pulso retangular positivo com largura de meio símbolo, e o símbolo 0 é representado pela ausência de transmissão do pulso, como ilustrado na Fig. 5.20(c).
4. *Sinalização bipolar com retorno para zero (BRZ, bipolar return-to-zero)*, a qual utiliza três níveis de amplitude, como indicado na Fig. 5.20(d). Especificamente, pulsos positivo e negativo com amplitudes iguais são utilizados alternativamente para o símbolo 1 e nenhum pulso é utilizado para o símbolo 0. Uma propriedade útil da sinalização BRZ é que o espectro de potência do sinal transmitido não possui componente CC e possui componentes em baixa-freqüência relativamente insignificantes, quando o símbolo 1 e 0 ocorrem com mesma probabilidade.
5. *Fase dividida (Código Manchester)*, o qual é ilustrado na Fig. 5.20(e). Neste método de sinalização, o símbolo 1 é representado por um pulso positivo seguido por um pulso negativo, com os dois pulsos de amplitude igual e largura de meio símbolo. Para o símbolo 0, as polaridades dos dois pulsos são invertidas. O código Manchester suprime a componente CC e possui componentes em baixa-freqüência relativamente insignificantes, independente da estatística do sinal.
6. *Codificação diferencial*, na qual a informação é codificada em termos das transições do sinal, como ilustrado na Fig. 5.20(f). No exemplo do sinal PCM mostrado na figura, a transição é utilizada para representar o símbolo 0, enquanto a ausência de transição é utilizada para o símbolo 1. Evidentemente que o sinal codificado diferencialmente pode ser invertido sem afetar sua interpretação. A informação binária original é recuperada pela comparação da polaridade dos símbolos adjacentes para estabelecer se ocorreu ou não uma transição. Note que a codificação diferencial requer a utilização de um *bit de referência*, como indicado na Fig. 5.20(f).

As formas de onda mostradas nas partes (a) a (f) da Fig. 5.20 foram traçadas para a seqüência de dados binários 01101001. É importante observar que foi utilizada uma forma retangular do pulso para desenhar estas formas de onda, o que simplifica muito a representação elétrica. Os benefícios de se utilizar outras formas de pulso para a transmissão de um dado PCM são discutidas no Capítulo 6.

5.10 EXEMPLOS DE TEMAS

■ Multiplexação por divisão no tempo

O teorema da amostragem fornece a base para a transmissão da informação contida em um sinal de mensagem $m(t)$ limitado em faixa na forma de uma seqüência de amostra de $m(t)$ tomadas uniformemente a uma taxa que, geralmente, é um pouco maior do que a taxa de Nyquist. Uma característica importante do processo de amostragem é a *conservação do tempo*. Ou seja, a transmissão das amostras da mensagem utiliza o canal de comunicação periodicamente por apenas uma fração do intervalo de amostragem e, desta forma, algum intervalo de tempo entre amostras adjacentes é disponibilizado para outras fontes de mensagem independentes, em um sistema de compartilhamento de tempo. Obtemos, portanto,

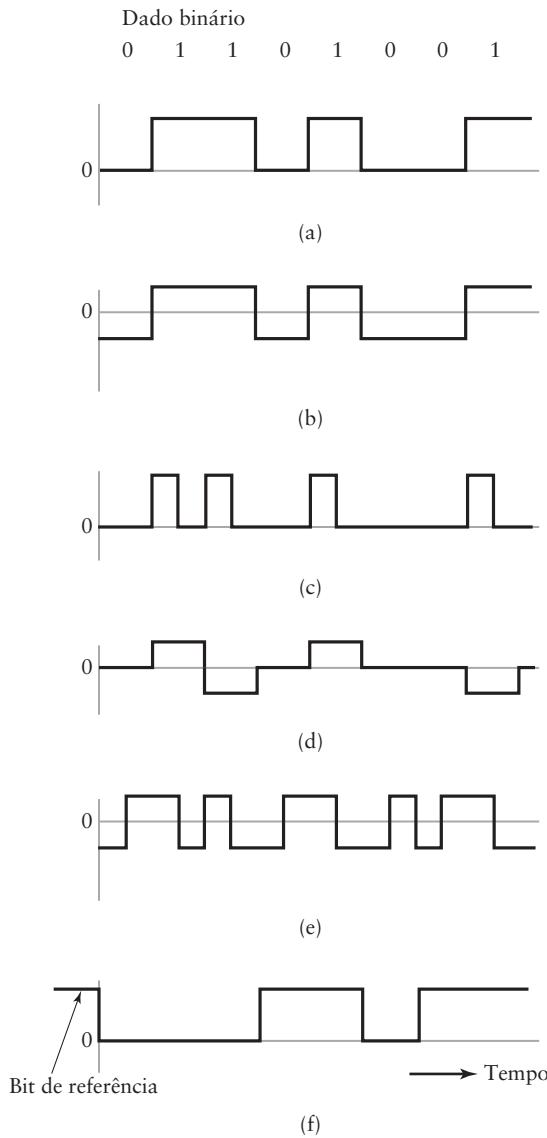


Figura 5.20 Códigos de linha. (a) sinalização liga-desliga. (b) sinalização sem retorno para zero. (c) sinalização com retorno para zero. (d) sinalização bipolar com retorno para zero. (e) Fase dividida ou Código Manchester. (f) Codificação diferencial.

um sistema de multiplexação por divisão no tempo (TDM – do inglês, *time-division multiplex*), o qual possibilita a utilização de um canal de comunicação comum por várias fontes de mensagem independentes sem interferência mútua entre elas.

O conceito da TDM é ilustrado pelo diagrama em blocos mostrado na Fig. 5.21. Cada sinal de mensagem de entrada é primeiro restrinido em largura de faixa por um filtro anti-aliasing passa-faixa para remover as freqüências que não são essenciais para uma representação adequada do sinal. As saídas dos filtros são, então, aplicadas a um *comutador*, o qual é geralmente implementado usando um circuito eletrônico de chaveamento. O comutador possui duas funções básicas: (1) tomar uma amostra estreita de cada uma das N mensagens de entrada a uma taxa f_s um pouco mais alta do que $2W$, onde W é a freqüência de corte do filtro anti-aliasing e (2) entrelaçar seqüencialmente estas N amostras dentro do

intervalo de amostragem T_s . De fato, esta última função é a essência da operação de multiplexação por divisão no tempo.

Seguindo o processo de comutação, o sinal multiplexado é aplicado a um *modulador de pulso*, com o propósito de transformar o sinal multiplexado em uma forma adequada para a transmissão em um canal comum. É claro que a utilização da multiplexação por divisão no tempo introduz um fator N de expansão de largura de faixa, pois o esquema deve apertar N amostras obtidas de N fontes independentes de mensagem em um slot de tempo igual a um intervalo de amostragem. No lado do receptor do sistema, o sinal recebido é aplicado a um *demodulador de pulso*, o qual executa a operação inversa do modulador de pulso. As amostras estreitas produzidas na saída do demodulador de pulso são distribuídas aos filtros de reconstrução passa-baixa apropriados através de um *outro comutador*, o qual opera em *sincronismo* com o comutador do transmissor. Esta sincronização é essencial para a operação satisfatória do sistema. A forma em que esta sincronização é implementada depende naturalmente do método de modulação de pulso utilizado para transmitir a sequência de amostras multiplexadas.

O sistema TDM é altamente sensível à dispersão do canal comum, ou seja, à resposta em amplitude não constante do canal e resposta de fase não linear, os dois medidos com respeito à freqüência. Desta forma, a *equalização* tanto da resposta de amplitude quanto de fase do canal é necessária para garantir a operação satisfatória do sistema. De fato, a equalização compensa a dispersão do canal. (Equalização de canal é discutida no Capítulo 6). Entretanto, ao contrário da multiplexação por divisão de freqüência (FDM) considerada no Capítulo 3, a aproximação de primeira ordem do TDM é imune às não linearidades do canal como fonte de cross-talk. A razão para este comportamento é que sinais de mensagem diferentes não são simultaneamente aplicados ao canal.

Sincronização

Em uma aplicação utilizando PCM, por exemplo, é natural multiplexar diferentes fontes de mensagem por uma divisão no tempo, na qual cada fonte mantém sua distinção de todas as outras fontes durante a viagem de transmissão ao receptor. Esta individualidade é responsável pela relativa facilidade na qual fontes de mensagem podem ser removidas ou inseridas em um sistema de multiplexação por divisão de tempo. Quando o número de fontes de mensagem independentes aumenta, o intervalo de tempo que pode ser alocado para cada fonte precisa ser reduzido, pois todas devem ser acomodadas em um intervalo de tempo igual ao recíproco da taxa de amostragem. Isto, por sua vez, significa que a duração permitida de uma representação de palavra de código de uma única amostra é reduzida. Entretanto, pulsos tendem a se tornar mais difíceis de serem gerados e transmitidos quando sua duração é reduzida. Além disso, se os pulsos se tornarem muito curtos, defeitos na mídia de transmissão começam a interferir na operação adequada do sistema. Assim sendo, na prática, é necessário restringir o número de fontes independentes de mensagem que podem ser incluídas em um grupo de divisão no tempo.

De qualquer forma, para um sistema PCM com multiplexação por divisão no tempo operar adequadamente, é necessário que as operações de temporização no receptor, exceto

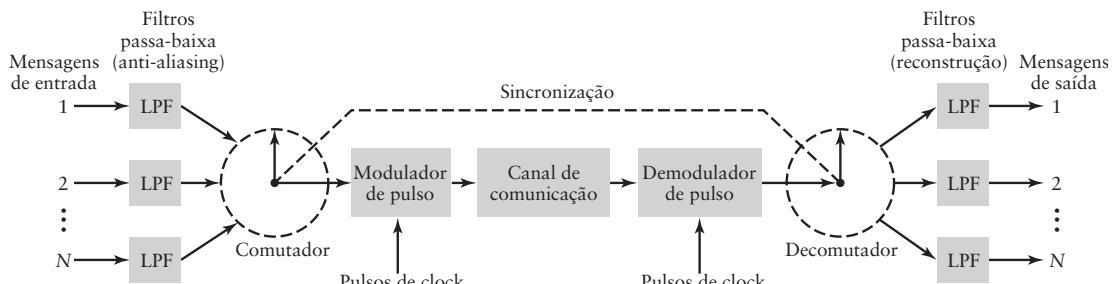


Figura 5.21 Diagrama de blocos de um sistema TDM.

pela perda de tempo na transmissão e repetição regenerativa, sigam as operações correspondentes do transmissor. De forma geral, isto representa a necessidade de um clock local ao receptor para manter o mesmo tempo do clock distante do transmissor, exceto pelo fato do clock local ser atrasado pelo total igual ao tempo necessário para transportar os sinais de mensagem do transmissor ao receptor. Este atraso, por sua vez, aumenta a diferença de fase entre os dois clocks. Um procedimento possível para a sincronização dos clocks do transmissor e receptor é enviar a parte um elemento de código ou pulso ao final de cada *quadro* (constituído por uma palavra de código obtida pela sucessão de cada fonte independente de mensagem) e pela transmissão desse pulso apenas a cada outro quadro. Neste caso, o receptor inclui um circuito que deve procurar pelo padrão de 1s e 0s alternando na metade da taxa de quadro e, portanto, estabelecendo a sincronização entre o transmissor e receptor.

Quando o caminho de transmissão é interrompido, é altamente improvável que os clocks do transmissor e receptor continuem indicando, por muito tempo, o mesmo tempo. Desta forma, implementando o processo de sincronização, devemos criar um procedimento regular para a detecção do pulso de sincronização. O procedimento é constituído por observar os elementos de código um por um até que o pulso de sincronização seja detectado. Ou seja, após observar um elemento de código particular por tempo suficiente para estabelecer a ausência do pulso de sincronização, o clock do receptor é mandado de volta por um elemento de código e o próximo elemento de código é observado. Este *processo de busca* é repetido até que o pulso de sincronização seja detectado. Claramente, o tempo necessário para a sincronização depende do período em que a transmissão adequada é restabelecida.

EXEMPLO 5.1 Sistema T1

Neste exemplo iremos descrever uma importante característica de um sistema PCM chamado de *sistema T1*, o qual transporta 24 canais de voz em pares de fios com repetidores regenerativos espaçados a, aproximadamente, intervalos de 2 km. O sistema de transporte T1 é básico ao sistema digital norte-americano de hierarquia⁵ de chaveamento em comunicação telefônica.

Um sinal de voz (masculino ou feminino) é essencialmente limitado a uma faixa de 200 a 3100 Hz, sendo que freqüências fora desta faixa não contribuem muito para o reconhecimento e compreensão da voz. De fato, circuitos de telefonia que respondem a esta faixa de freqüência fornecem um serviço satisfatório. Assim sendo, é comum passar o sinal de voz por um filtro passa-faixa com freqüência de corte em, aproximadamente, 3,1 kHz antes da amostragem. Logo, com $W = 3,1$ kHz, o valor nominal da taxa de Nyquist é 6,2 kHz. O sinal filtrado de voz é geralmente amostrado a uma taxa um pouco maior, em 8kHz, que é a taxa *padrão* de amostragem em sistemas de telefonia.

Para compressão/expansão, o sistema T1 utiliza uma característica *linear por partes* (constituída por 15 segmentos lineares) para aproximar a lei μ logarítmica da Eq. (5.23), com constante $\mu = 255$. Esta aproximação é construída de tal forma que os pontos finais dos segmentos estejam na curva de compressão calculada da Eq. (5.23) e suas projeções no eixo vertical são espaçadas uniformemente.

Existe um total de 255 níveis de representação associados com os 15 segmentos da lei de compressor. Para acomodar este número de níveis de representação, cada um dos 24 canais de voz utilizam um código binário com uma palavra de 8 bits. O primeiro bit indica se a amostra de voz de entrada é positiva ou negativa. Este bit é 1 se positivo e 0 se negativo. Os próximos 3 bits da palavra de código identificam o segmento particular dentro do qual a amplitude da amostra de voz de entrada está e os quatro últimos bits identificam a representação real do nível dentro do segmento.

Com uma taxa de amostragem de 8 kHz, cada quadro do sinal multiplexado T1 ocupa um período de 125 μ s. Em particular, ele é constituído por 24 palavras de 8 bits, mais um único bit adicionado ao final do quadro com o propósito de sincronização. Logo, cada quadro é constituído por um total de $(24 \times 8) + 1 = 193$ bits. Desta forma, a duração de cada bit é igual a 0,647 μ s, e a taxa resultante de transmissão é de 1,544 megabits por segundo (Mb/s).

⁵ Para a descrição da hierarquia de chaveamento utilizada na América do Norte, veja Haykin (2001), pp. 214-217.

■ Rádio de impulso

Sistemas de transmissão digital tradicionais tentam minimizar a largura de faixa do sinal transmitido. Logo, a filtragem é geralmente aplicada a pulsos retangulares para reduzir a largura de faixa ocupada. Entretanto, um método que não segue esta filosofia e que tem chamado a atenção recentemente é chamado de *rádio de impulso*. Nesta técnica, a informação é enviada através de pulsos muito estreitos que são muito separados no tempo. Como as larguras dos pulsos são muito estreitas, o espectro do sinal resultante é muito amplo, consequentemente, esta técnica é uma forma de *transmissão de rádio com faixa ultra larga* (UWB – do inglês, *ultra-wideband*), a qual forma o objeto de nosso terceiro e último exemplo de tema.

Especificamente, um tipo de pulso utilizado para o rádio de impulso é o *monociclo Gaussiano*. Esta forma de pulso é a derivada do pulso Gaussiano escalonado $g(t) = \exp(-\pi t^2)$, discutido no Capítulo 2. A forma de onda do monociclo Gaussiano é dada por

$$v(t) = A \left(\frac{t}{\tau} \right) \exp \left\{ -\pi \left(\frac{t}{\tau} \right)^2 \right\} \quad (5.39)$$

onde A é o fator de escala em amplitude e τ é a constante de tempo do pulso. Este sinal está mostrado na Fig. 5.22. Ele consiste de um lóbulo positivo seguido por um lóbulo negativo, com largura aproximada de pulso de τ . Para aplicações de rádio de impulso, a largura τ do pulso é normalmente entre 0,20 e 1,50 nanosegundos.

O espectro de uma seqüência destes pulsos pode ser obtido pela transformada de Fourier de um pulso individual, mostrado na Fig. 5.23. O eixo de freqüência na Fig. 5.23 foi normalizado em termos da constante de tempo τ ; para $\tau = 1,0$ nanosegundo, o eixo de freqüências varia de 0 a 4 GHz.

Existem diversos métodos para a modulação digital de uma onda impulsiva. Um método é a modulação por posição de pulso, como mostrado na Fig. 5.24. (A modulação por posição de pulso foi considerada na Seção 5.3). Neste método, existe uma separação nominal de tempo T_p entre pulsos sucessivos. Para transmitir o sinal binário 0, o pulso é transmitido um pouco antes, no tempo $t = -T_c$. Para transmitir o sinal binário 1, o pulso é transmitido um pouco depois, em $t = +T_c$. O receptor detecta este tempo antes/depois e demodula o dado adequadamente. Separações típicas entre pulsos (isto é, T_p) estão na faixa de 25 nanosegundos a 1000 nanosegundos, resultando em uma faixa de taxa de dados de 40 Mbit/s a 1 Mbit/s.

A natureza de faixa ultra larga do sinal modulado possui aspectos bons e ruins. Como a potência do sinal é espalhada em uma grande largura de faixa, o total de potência em

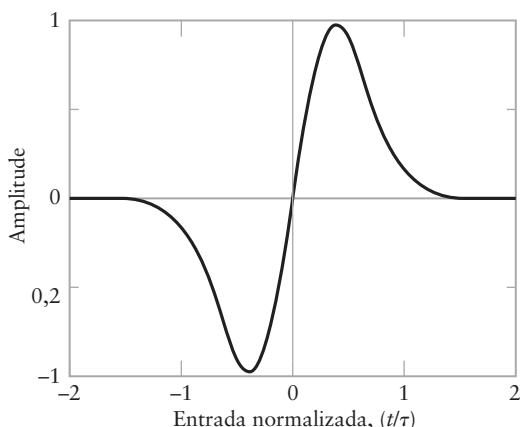


Figura 5.22 Gráfico no domínio do tempo de um monociclo Gaussiano utilizado em um rádio de impulso.

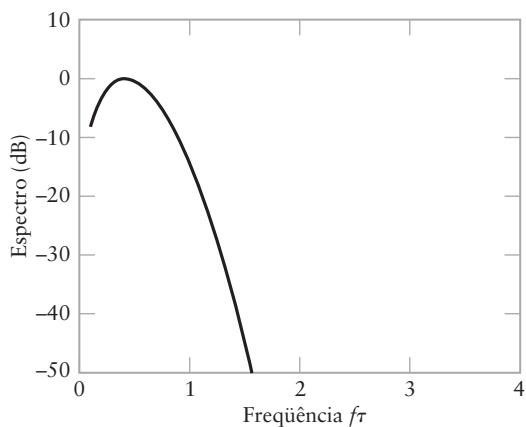


Figura 5.23 Espectro do monociclo Gaussiano.

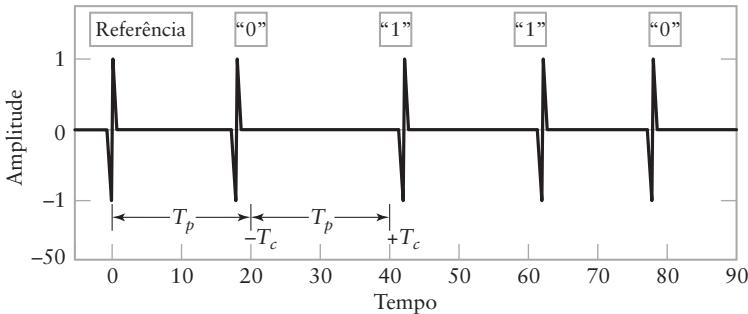


Figura 5.24 Modulação por posição de pulso em um rádio de impulso.

qualquer canal de faixa estreita particular é pequeno, o que é bom. Entretanto, a potência cai em todos os canais particulares de faixa estreita, o que é ruim. Em particular, existe a preocupação de rádios de faixa ultra larga causarem interferência destrutiva em serviços de rádio de faixa estreita que já existem operando no mesmo espectro de rádio. Como consequência, apesar de um rádio de faixa ultra larga ser permitido em várias jurisdições, existem limites restritos sobre o espectro de potência que pode ser transmitido. Devido a esta limitação na potência transmitida, rádios de faixa ultra larga são restringidos a aplicações de curta distância, normalmente menores do que poucas centenas de metros.⁶

5.11 RESUMO E DISCUSSÃO

Neste capítulo introduzimos dois processos fundamentais e complementares:

- *Amostragem*: a qual opera no domínio do tempo. O processo de amostragem é a ligação entre a forma de onda analógica e sua representação em tempo discreto.
- *Quantização*: a qual opera no domínio da amplitude. O processo de quantização é a ligação entre a forma de onda analógica e sua representação em amplitude discreta.

O processo de amostragem é construído sobre o *teorema da amostragem*, o qual afirma que um sinal estritamente limitado em faixa, sem componentes de freqüência superiores a W , é representado unicamente por uma seqüência de amostras tomadas a uma taxa uniforme igual ou maior do que a taxa de Nyquist de $2W$ amostras por segundo. Quanto ao processo de quantização, ele explora o fato de que qualquer sentido humano, como último receptor, pode detectar apenas diferenças de intensidade finitas.

O processo de amostragem é básico à operação de todos os sistemas de modulação de pulso, os quais podem ser classificados como modulação analógica de pulso ou modulação digital de pulso. A característica distinta entre eles é que a modulação analógica de pulso mantém uma representação de amplitude contínua do sinal de mensagem, enquanto que sistemas de modulação digital de pulso implementam a quantização para fornecer a representação do sinal de mensagem discreta tanto no tempo quanto na amplitude.

A modulação analógica de pulso resulta da variação de algum parâmetro dos pulsos transmitidos, tal como amplitude, duração ou posição, resultando na modulação por amplitude de pulso (PAM), modulação por duração de pulso (PDM) ou modulação por posição de pulso (PPM), respectivamente.

Sistemas de modulação digital de pulso, por outro lado, transmitem sinais de mensagem analógicos como a seqüência de pulsos codificados, possibilitado pela combinação da amostragem e quantização. A modulação por codificação de pulso é a forma de modulação digital de pulso que contém naturalmente algumas vantagens únicas do sistema, as quais,

⁶ Para uma discussão mais detalhada sobre rádio de faixa ultra larga, veja as duas referências a seguir: Win and Scholts (1998) e Cassioli, Win and Molisch (2002).

por sua vez, o tornaram o *método preferido de codificação* para a transmissão de sinais analógicos como voz e vídeo. As vantagens da modulação por codificação de pulso incluem robustez ao ruído e à interferência, regeneração eficiente dos pulsos codificados ao longo do caminho de transmissão e o formato uniforme para tipos diferentes de sinais de mensagem (isto é, voz, vídeo e dados).

A modulação delta e a modulação por codificação diferencial de pulso são duas outras formas úteis de modulação digital de pulso. A principal vantagem da modulação delta é o circuito simplificado. Entretanto, esta vantagem é paga com o aumento da taxa de transmissão de dados. Em contraste, a modulação por codificação diferencial de pulso possui uma complexidade maior de circuito para melhorar a performance do sistema. A melhoria é obtida ao custo de aumento da complexidade do sistema, o que facilita a idéia de predição para remover símbolos redundantes da seqüência de dados de entrada e, portanto, permitir a utilização de uma largura de faixa de canal reduzida quando comparada ao PCM.

Outras melhorias na operação da modulação delta e modulação por codificação diferencial de pulso podem ser realizadas através do uso de adaptação para considerar variações estatísticas nos dados de entrada. Especificamente, a adaptação é utilizada na modulação delta para melhorar a performance frente ao ruído. Por outro lado, a adaptação é utilizada na modulação por codificação diferencial de pulso para reduzir a necessidade de largura de faixa.

É importante perceber que as técnicas de modulação de pulso são *com perdas*, no sentido de que alguma informação é perdida como resultado da representação do sinal que elas executam. Por exemplo, na modulação por amplitude de pulso, a prática comum é utilizar um filtro anti-alias (passa-baixa) antes da amostragem. Fazendo isto, informação é perdida em virtude do fato das componentes de alta freqüência, consideradas como não essenciais, serem removidas pelo filtro. A natureza de perda da modulação de pulso é mais facilmente vista na modulação por codificação de pulso que prejudicada pelo ruído de quantização (isto é, distorção), o qual aparece em função do fato da seqüência transmitida de pulsos codificados não possuir uma precisão infinita necessária para representar exatamente amostras contínuas. De qualquer forma, a perda de informação que ocorre pela utilização do processo de modulação de pulso está *sobre o controle do projetista*, ou seja, ela pode ser feita pequena o suficiente para que seja imperceptível ao usuário final.

Um ponto que precisa ser ressaltado uma última vez: na realidade, PCM, DM e DPCM são *estratégias de codificação de fonte*, cujo propósito é converter um sinal analógico em uma forma digital. Para a transmissão real dos dados codificados em um canal de comunicação, a forma discreta da modulação por amplitude de pulso (PAM) é geralmente utilizada. (Detalhes desta aplicação de PAM são apresentados no próximo capítulo.)

Neste capítulo também incluímos três exemplos de temas, ressaltando importantes aplicações resumidas a seguir:

- ▶ Multiplexação por divisão no tempo, a qual possibilita a utilização comum de canal de comunicação por várias fontes independentes de mensagem, obtida em função de uma importante característica do processo de amostragem – a conservação de tempo.
- ▶ O sistema T1, o qual acomoda a transmissão PCM de 24 canais de voz em pares de fios com repetidores regenerativos espaçados a cada intervalo de 2 km, aproximadamente.
- ▶ Rádio de impulso, pelo qual a informação é enviada em um canal sem fio em banda base, usando pulsos muito estreitos.

Problemas adicionais

5.12 (a) Trace o espectro de uma onda PAM produzida pelo sinal modulante

$$m(t) = A_m \cos(2\pi f_m t)$$

assumindo freqüência de modulação $f_m = 0,2$ Hz, período de amostragem $T_s = 1$ s e duração de pulso

$$T = 0,45\text{s}$$

- (b) Utilizando um filtro ideal de reconstrução, trace o espectro da saída do filtro. Compare este resultado com a saída que deveria ser obtida se não houvesse o efeito da abertura.
- 5.13 Neste problema, calculamos a equalização necessária para o efeito de abertura de um sistema PAM. A freqüência operacional $f = f_s/2$, a qual corresponde a maior componente de freqüência do sinal de mensagem para uma taxa de amostragem igual à taxa de Nyquist. Trace $1/\text{sinc}(0,5T/T_s)$ versus T/T_s , e determine a equalização necessária quando $T/T_s = 0,25$.
- 5.14 Um sinal de *telemetria* PAM envolve a multiplexação de quatro sinais de entrada: $s_i(t)$, $i = 1, 2, 3$ e 4. Dois dos sinais $s_1(t)$ e $s_2(t)$ possuem larguras de faixa de 80 Hz, cada, enquanto que os dois outros sinais $s_3(t)$ e $s_4(t)$ possuem larguras de faixa de 1kHz, cada. Os sinais $s_3(t)$ e $s_4(t)$ são, cada um, amostrados a uma taxa de 2400 amostras por segundo. Esta taxa de amostragem é dividida por 2^R (isto é, uma potência inteira de 2) para obter a taxa de amostragem para $s_1(t)$ e $s_2(t)$.
- Determine o maior valor de R .
 - Utilizando o valor de R determinado na parte (a), projete um sistema de multiplexação que primeiro multiplexa $s_1(t)$ e $s_2(t)$ em uma nova seqüência $s_5(t)$ e, então, multiplexa $s_3(t)$, $s_4(t)$ e $s_5(t)$.
- 5.15 (a) Um sinal senoidal com amplitude de 3,25 volts é aplicado a um quantizador uniforme do tipo midtread cuja saída assume os valores $0, \pm 1, \pm 2, \pm 3$ volts. Trace a forma de onda da saída resultante do quantizador para um ciclo completo da entrada.
- (b) Repita o cálculo para quando o quantizador é do tipo midrise cuja saída assume os valores $\pm 0,5, \pm 1,5, \pm 2,5$ e $\pm 3,5$ volts.
- 5.16 Considere a seguinte seqüência de 1s e 0s:
- Uma seqüência alternada de 1s e 0s.
 - Uma seqüência longa de 1s seguida por uma seqüência longa de 0s.
 - Uma seqüência longa de 1s seguida por um único 0 e, então, uma seqüência longa de 1s.
- Trace a forma de onda para cada caso destas seqüências utilizando os seguintes métodos de representação dos símbolos 1 e 0:
- Sinalização liga-desliga.
 - Sinalização bipolar com retorno para zero.
- 5.17 A onda senoidal
- $$m(t) = 6 \text{ sen}(2\pi t) \text{ volts}$$
- é transmitida usando um sistema PCM binário de 4 bits. O quantizador é do tipo midrise com tamanho de degrau de 1 volt. Trace a onda PCM resultante para um ciclo completo da entrada. Assuma uma taxa de quatro amostras por segundo, com amostras tomadas em $t = \pm 1/8, \pm 3/8, \pm 5/8, \dots$, segundos.
- 5.18 Considere um compact disc (CD) que utiliza uma modulação por codificação de pulso para gravar sinais de áudio com largura de faixa $W = 15$ kHz. As especificações do modulador incluem:
- | | |
|--------------|-------------------------|
| Quantização: | uniforme com 512 níveis |
| Codificação: | binária |
- Determine (a) a taxa de Nyquist e (b) a menor taxa de bits possível.
- 5.19 Este problema lida com a digitalização de um sinal de televisão utilizando a modulação por codificação de pulso. A largura de faixa do sinal é 4,5 MHz. As especificações do modulador são:
- | | |
|--------------|------------------------------|
| Amostragem: | 15% acima da taxa de Nyquist |
| Quantização: | uniforme com 1024 níveis |
| Codificação: | binária |
- Determine (a) a taxa de Nyquist e (b) a menor taxa de bits possível.
- 5.20 A Fig. 5.25 mostra um sinal PCM no qual os níveis de amplitude de +1 volt e -1 volt são utilizados para representar os símbolos binários 1 e 0, respectivamente. A palavra de código utilizada é constituída por três bits. Obtenha a versão amostrada do sinal analógico a partir do qual este sinal PCM foi obtido.

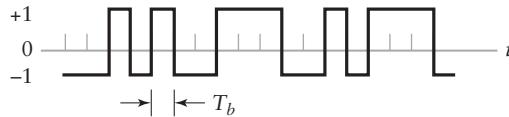


Figura 5.25 Problema 5.20.

- 5.21 Considere uma onda senoidal com freqüência f_m e amplitude A_m , aplicada a um modulador delta com tamanho de degrau Δ . Mostre que a distorção de sobre-inclinação ocorrerá se

$$A_m > \frac{\Delta}{2\pi f_m T_s}$$

na qual T_s é o período de amostragem. Qual é a potência máxima que pode ser transmitida sem a distorção de sobre-inclinação?

- 5.22 Considere um sistema de modulação delta (DM) utilizado para transmitir um sinal de voz, o qual é uniformemente amostrado a uma taxa de 64 kHz. Assuma as seguintes especificações:

Largura de faixa do sinal de voz = 3,1 kHz

Amplitude máxima do sinal = 10 volts

- (a) Para evitar a distorção de sobre-inclinação, qual é o menor valor possível para o tamanho do degrau Δ utilizado no sistema?
- (b) Determine a potência média do ruído granular.
- (c) Determine a menor largura de faixa do canal para transmitir o dado codificado DM.

- 5.23 Repita o Problema 5.22, mas desta vez utilize uma onda senoidal com freqüência igual a 3,1 kHz e amplitude de pico de 10 volts.

- 5.24 No sistema DPCM mostrado na Fig. 5.26, mostre que, na ausência de ruído de canal, os filtros de predição do transmissor e receptor operam em sinais de entrada um pouco diferentes.

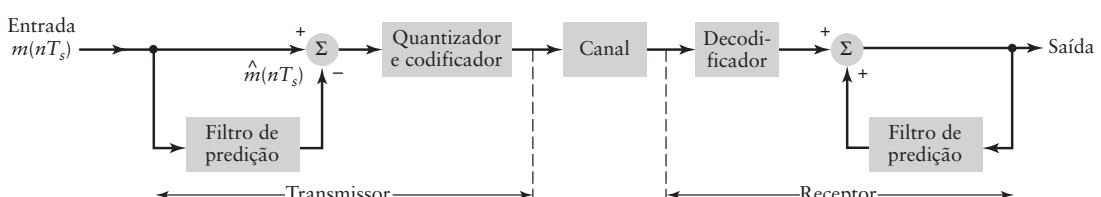


Figura 5.26 Problema 5.24.

Problemas avançados

- 5.25 (a) Dado um sinal físico qualquer, não é possível amostrar o sinal sem danificá-lo devido ao fenômeno de aliasing.
 (b) O melhor que o projetista pode fazer é gerenciar o problema de aliasing de tal forma que o dano não seja percebido pelo usuário humano.

Justifique a validade destas duas afirmativas.

- 5.26 Na *amostragem natural*, um sinal analógico $g(t)$ é multiplicado por um trem de pulsos retangulares periódico $c(t)$. Dado que a freqüência de repetição do pulso desse trem periódico é f_s e a duração de cada pulso retangular é T (com $f_s T \gg 1$):
- (a) Determine o espectro do sinal $s(t)$ resultante da utilização da amostragem natural. Você pode assumir que $t = 0$ corresponde ao ponto médio do pulso retangular em $c(t)$.
 - (b) Mostre que o sinal original $g(t)$ pode ser recuperado exatamente de sua versão amostrada naturalmente desde que as condições do teorema da amostragem sejam satisfeitas.
- 5.27 A Fig. 5.27 mostra o diagrama de blocos de um *chopper bipolar*. O chopper possui dois caminhos paralelos, um direto e outro invertido. O comutador chaveia entre estes dois caminhos com

uma freqüência representada por f_s . O chopper produz um sinal de saída $y(t)$ em resposta a um sinal de entrada $x(t)$.

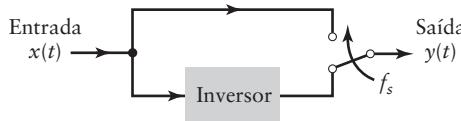


Figura 5.27 Problema 5.27.

- (a) Determine $y(t)$ em termos de $x(t)$.
- (b) Dado que $X(f)$ é a transformada de Fourier de $x(t)$, determine a transformada de Fourier de $y(t)$.

Utilize gráficos para ilustrar sua resposta.

- 5.28 A Fig. 5.28 mostra um subsistema constituído por um amostrador instantâneo seguido por um filtro de síntese, o qual é utilizado em um *osciloscópio amostrador*. O subsistema é projetado para explorar o fenômeno de aliasing para a amostra expandida de sinais periódicos arbitrários. Assuma que o sinal periódico de entrada $x(t)$ foi pré-filtrado (não mostrado na Fig. 5.28) para suprimir todas as componentes de freqüência maiores do que a m -ésima harmônica.

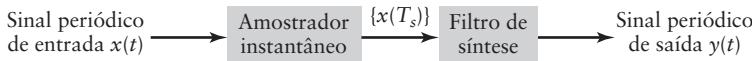


Figura 5.28 Problema 5.28.

Especificamente, o sinal periódico de entrada $x(t)$ é amostrado a uma taxa f_s um pouco maior do que sua freqüência fundamental f_0 , como mostrado por

$$f_s = (1 - \alpha)f_0$$

na qual o fator α está no intervalo $0 < \alpha < 1$. Desta forma, o alias é propositalmente introduzido na composição da seqüência amostrada $\{x(nT_s)\}$, sendo $n = 0, \pm 1, \pm 2, \dots$ e $T_s = 1/f_s$. A seqüência $\{x(nT_s)\}$ é, a seguir, processada pelo filtro passa-baixa de síntese com freqüência de corte $B = f_s/2$, produzindo, portanto, o sinal periódico de saída $y(t)$.

Utilize gráficos para ilustrar a relação entre $x(t)$ e $y(t)$ e seus respectivos espectros $X(f)$ e $Y(f)$, e incluir o seguinte:

- (a) Mostre que o sinal de saída é uma versão *expandida* do sinal de entrada, mostrado por

$$y(t) = x(at)$$

- (b) Para prevenir a sobreposição espectral, o fator de expansão a deve satisfazer a condição

$$a < \frac{1}{2m + 1}$$

- (c) O espectro $Y(f)$ contém a imagem *comprimida* do espectro $X(f)$.

- 5.29 Em um estúdio de televisão, o sinal de vídeo é produzido capturando 60 quadros estáticos da cena por segundo. Logo, o período de amostragem do sinal de vídeo é $T_s = 1/60$ segundos. Isto significa que um dado ponto da tela de televisão é, na realidade, preto a maior parte do tempo. O ponto é acendido periodicamente a cada $1/60$ segundos. A luz emitida pela televisão forma um interessante experimento do fenômeno de aliasing existente na amostragem uniforme da senóide complexa $\exp(j2\pi ft)$.

Suponha que a tela da televisão é toda mascarada, exceto por uma faixa estreita horizontal e que sentamos de costas para a televisão. Para ver o que está acontecendo na televisão utilizamos um espelho rodando no sentido anti-horário o eixo horizontal. Demonstre os dois possíveis resultados:

- (a) A faixa horizontal irá aparecer estática se a velocidade de rotação do espelho for igual à taxa de amostragem do sinal de vídeo.

- (b) Caso contrário, a faixa horizontal da televisão irá aparecer no espelho como se ela estivesse rodando ao contrário.

(Nota: O experimento descrito no Problema 5.29 mostra como a televisão pode ser utilizada para demonstrar a forma na qual um dispositivo chamado de *estroboscópio* trabalha).

- 5.30 Na Seção 5.2 discutimos a interpolação de uma sequência de amostras baseada no filtro de sample-and-hold. Na literatura de controle, este filtro de interpolação é chamado de amostrador de *ordem zero*. Um filtro de interpolação mais complexo, chamado de *amostrador de primeira ordem* pode ser preferível do que o amostrador de ordem zero. Como o nome indica, um amostrador de primeira ordem executa a interpolação entre pontos de dados por um polinômio de primeira ordem, ou seja, uma linha reta.

A Fig. 5.29 mostra a resposta $h(t)$ ao impulso do filtro amostrador de primeira ordem a um pulso de amplitude unitária e duração T .

- (a) Mostre que a resposta em freqüência do filtro amostrador de primeira ordem é dada por

$$H(f) = T \left(\frac{1 - \exp(-j2\pi fT)}{j2\pi fT} \right)^2 (1 + j2\pi fT)$$

- (b) Trace as respostas de amplitude e fase do filtro amostrador de primeira ordem. Compare-as com o filtro sample-and-hold.
(c) Determine a função de transferência do equalizador que deve ser colocado em série com o filtro amostrador de primeira ordem para a reconstrução perfeita do sinal de mensagem. Compare seu resultado com a equalização necessária para o filtro de sample-and-hold para um ciclo de trabalho (T/T_s) $\leq 0,1$. Comente seu resultado.
(d) Trace a resposta do filtro amostrador de primeira ordem à entrada senoidal $\cos(50t)$, assumindo $f_s = 100$ Hz e $T = 0,01$. Compare seu resultado com o produzido pelo filtro sample-and-hold. Novamente, comente o que a comparação mostrou.

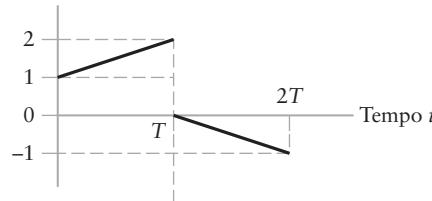


Figura 5.29 Problema 5.30.

- 5.31 Neste problema, iremos trabalhar na obtenção do monociclo Gaussiano $v(t)$ da Eq. (5.39) e seu espectro apresentado na Fig. 5.23.

Para sermos específicos, considere o pulso Gaussiano unitário

$$g(t) = \exp(-\pi t^2)$$

o qual é sua própria transformada de Fourier, mostrado por

$$G(f) = \exp(-\pi f^2)$$

(Para detalhes deste par transformada de Fourier, veja o Exemplo 2.6).

Diferenciando $g(t)$ com respeito ao tempo t , obtemos o monociclo Gaussiano correspondente

$$g'(t) = -2\pi t \exp(-\pi t^2)$$

no qual o sinal de primo representa diferenciação.

- (a) Aplicando as propriedades de linearidade e dilatação da transformada de Fourier a $g'(t)$, obtenha $v(t)$ da Eq. (5.39). Qual é o valor do parâmetro A do pulso na Fig. 5.22?
(b) Utilizando os resultados da parte (a), e utilizando a propriedade da diferenciação da transformada de Fourier, obtenha a fórmula utilizada para traçar o espectro de $v(t)$ mostrado na Fig. 5.23.

As propriedades da transformada de Fourier referidas nas partes (a) e (b) são discutidas no Capítulo 2.

TRANSMISSÃO DE DADOS BANDA BASE

A transmissão de dados digitais (independente de sua origem) em um canal de comunicação físico é limitada por dois fatores inevitáveis:

1. *Interferência intersimbólica*, a qual surge devido a imperfeições na resposta em frequência do canal.
2. *Ruído de canal*, o qual se refere a sinais elétricos indesejados que aparecem na saída do canal devido a fenômenos físicos aleatórios e imprevisíveis.

Neste capítulo iremos focalizar no problema de interferência intersimbólica. A discussão de ruído em receptores de comunicação digital será deixada para a parte final deste livro.

Como o nome indica, a interferência intersimbólica refere-se à interferência causada pela resposta no tempo do canal espalhando um símbolo em seus adjacentes. A interferência intersimbólica é problemática porque ela possui o efeito de introduzir desvio (isto é, erros) entre a seqüência de dados reconstruída na saída do receptor e a seqüência de dados original aplicada à entrada do transmissor. Logo, a não ser que medidas corretivas sejam tomadas, a interferência intersimbólica pode impor um limite na taxa máxima de transmissão de dados através do canal, a qual estará abaixo da capacidade física do canal.

Com a interferência intersimbólica atuando como questão de cunho prático, iremos estudar uma importante medida corretiva para lidar com esse problema, chamada de *formatação de pulso banda base*. Esta medida envolve a utilização de pulsos limitados em faixa que são formatados de forma especial para mitigar o problema de interferência intersimbólica. Esta medida corretiva baseia-se no conhecimento prévio do canal.

O capítulo irá nos ensinar três lições:

- *Lição 1: A compreensão do problema de interferência intersimbólica e como resolvê-lo é de fundamental importância para o projeto de sistemas de comunicação digital.*
- *Lição 2: O espectro de cosseno levantado fornece uma poderosa ferramenta matemática para a formatação de pulso banda base projetada para mitigar o problema de interferência intersimbólica.*
- *Lição 3: O padrão de olho é um indicador visual da performance, mostrando as limitações físicas de um sistema de transmissão de dados digitais de forma intuitiva.*

6.1 TRANSMISSÃO BANDA BASE DE DADOS DIGITAIS

Lembre-se, do Capítulo 1, que o termo “banda base” é utilizado para indicar a faixa de freqüências representando o sinal original entregue por uma fonte de informação. A fonte de informação, por exemplo, pode ser um computador que produz uma seqüência de dados binários construída com os símbolos 0 e 1. A tarefa do sistema de comunicação digital é transportar a seqüência de dados da fonte para seu destino, através de um canal, fazendo isto de forma confiável. Para executar esta tarefa, precisamos utilizar uma técnica de modulação que envolva a variação de amplitude, fase ou freqüência de pulsos transmitidos de acordo com o dado, de alguma forma discreta. Neste capítulo, enfatizamos a utilização de modulação por amplitude de pulso discreta, a qual é uma forma de modulação por amplitude sendo quantizada em um conjunto de níveis discretos. Existem três razões para esta ênfase:

1. A modulação por amplitude de pulso discreta é simples de ser analisada.
2. Ela é a forma mais eficiente de modulação de pulso discreta em termos tanto da potência quanto da utilização de largura de faixa.
3. As técnicas analíticas desenvolvidas para trabalhar com a modulação por amplitude de pulso discreta podem ser estendidas para outras técnicas de modulação por pulso discreta utilizando fase ou freqüência.

Na *modulação por amplitude de pulso discreta (PAM)*, a amplitude dos pulsos transmitidos é variada de forma discreta em função de uma seqüência de entrada de dados digitais. A Fig. 6.1(a) mostra os blocos funcionais básicos de um sistema PAM banda base. A seqüência de dados digitais de entrada é representada por $\{b_k\}$. No tempo $t = kT_b$, no qual T_b é a duração do bit e $k = 0, \pm 1, \pm 2, \dots$, o elemento b_k , representando o símbolo binário 1 ou 0, é emitido pela fonte de informação. A seqüência binária de dados $\{b_k\}$ é aplicada a um *codificador de linha*, cujo propósito é produzir um sinal codificado em nível representado por $\{a_k\}$. (Codificadores de linha são descritos no Capítulo 5.) Por exemplo, podemos definir o sinal $\{a_k\}$ codificado em nível em termos de pulsos positivos e negativos com amplitude fixa e curta duração (curta o suficiente para que os dois pulsos sejam vistos como impulsos unitários com polaridades invertidas). Especificamente, escrevemos

$$a_k = \begin{cases} +1 & \text{se a entrada } b_k \text{ for símbolo 1} \\ -1 & \text{se a entrada } b_k \text{ for símbolo 0} \end{cases} \quad (6.1)$$

O sinal $\{a_k\}$ codificado em nível é aplicado ao *filtro de transmissão* para produzir a seqüência de pulsos, cujo formato básico é representado nos domínios do tempo e freqüência por $g(t)$ e $G(f)$, respectivamente. Dessa forma, o sinal $\{a_k\}$ codificado em nível possui o papel de sinal modulante, resultando em um sinal PAM discreto definido por

$$s(t) = \sum_{k=-\infty}^{\infty} a_k g(t - kT_b) \quad (6.2)$$

O sinal PAM $s(t)$ é transmitido através de um *canal de comunicação linear*, o qual é descrito nos domínios do tempo e freqüência pela resposta $h(t)$ ao impulso e função de transferência $H(f)$, respectivamente. Ignorando o efeito de ruído aditivo de canal, podemos expressar a saída do canal por

$$x(t) = s(t) \star h(t) \quad (6.3)$$

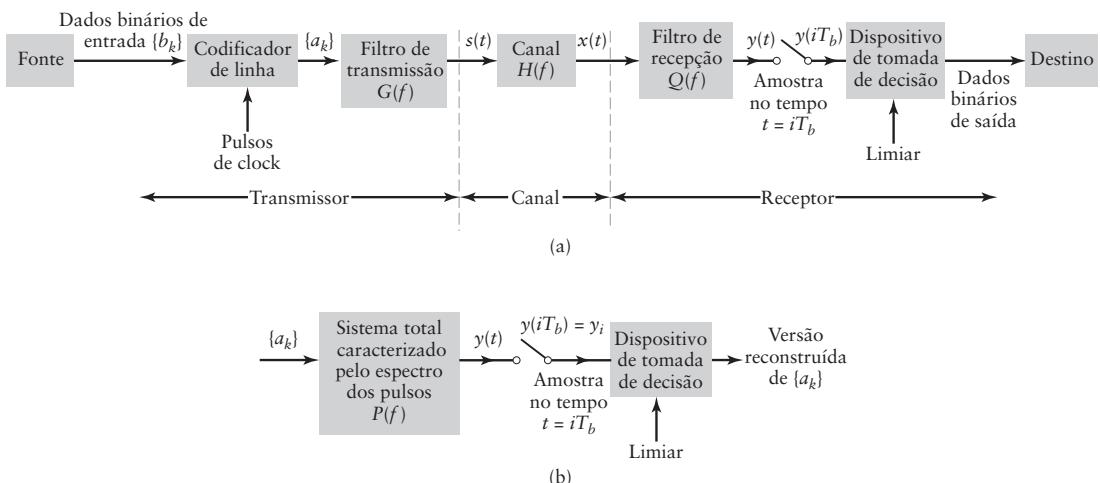


Figura 6.1 Transmissão banda base de dados binários. (a) Diagrama em blocos do sistema, mostrando seus elementos constituintes desde a fonte até o destino. (b) Representação simplificada do sistema.

na qual o símbolo \star representa a convolução no domínio do tempo.

A saída $x(t)$ do canal é processada por um *filtro receptor*, o qual é descrito nos domínios do tempo e freqüência pela resposta $q(t)$ ao impulso e função de transferência $Q(f)$, respectivamente. A saída resultante é, portanto, definida pela convolução de $x(t)$ e $q(t)$, ou seja,

$$y(t) = x(t) \star q(t) \quad (6.4)$$

A saída $y(t)$ do filtro é, a seguir, *amostrada em sincronia* com o *gerador de pulsos de clock* do transmissor. A sincronização é, geralmente, estabelecida pela extração de um *clock* ou *sinal de temporização* da saída do filtro de recepção. Finalmente, a seqüência de amostras obtida é utilizada para reconstruir a seqüência original de dados binários através de um *dispositivo de tomada de decisão*. Especificamente, a amplitude de cada amostra é comparada com um *limiar*. Se o limiar for excedido, a decisão é feita em favor do símbolo 1, digamos. Se o limiar não for excedido, a decisão é feita em favor do símbolo 0. Se a amplitude da amostra for exatamente igual ao limiar, o símbolo pode ser escolhido como 0 ou 1, aleatoriamente, sem afetar a performance geral do sistema. Quando os símbolos 0 e 1 possuem a mesma probabilidade de ocorrer, é razoável ajustar o limiar para o nível de amplitude zero.

O modelo mostrado na Fig. 6.1(a) representa não somente o sistema de transmissão de dados essencialmente banda base por natureza (por exemplo, a transmissão de dados em um cabo coaxial) mas também o *equivalente banda base* de um sistema de modulação linear utilizado para transmitir dados em um canal passa-faixa (por exemplo, um canal de telefonia). Neste último caso, o modelo equivalente banda base do sistema de transmissão de dados é desenvolvido utilizando as idéias apresentadas na Seção 3.8.

6.2 PROBLEMA DE INTERFERÊNCIA INTERSÍMBÓLICA

Como já mencionado, para a discussão atual iremos ignorar o efeito de ruído aditivo de canal. Fazemos isto para podermos focalizar nossa atenção nos efeitos de imperfeições na resposta em freqüência do canal (isto é, dispersão da forma de um pulso pelo canal) na transmissão de dados através do canal.

Usando as Eqs. (6.2) a (6.4), exceto pelo fator de escala, podemos expressar a *saída do filtro de recepção* como o sinal PAM modificado

$$y(t) = \sum_{k=-\infty}^{\infty} a_k p(t - kT_b) \quad (6.5)$$

A linearidade da transmissão de dados mostrada na Fig. 6.1 nos leva à expressão da *forma total do pulso* $p(t)$ da Eq. 6.5 como sendo o produto de convolução múltiplo

$$p(t) = g(t) \star h(t) \star q(t) \quad (6.6)$$

Naturalmente, o pulso $p(t)$ recebido possui forma diferente do sinal $s(t)$ transmitido da Eq. (6.2). Seja o espectro $P(f)$ a transformada de Fourier do pulso $p(t)$. Então, no domínio da freqüência, podemos escrever

$$P(f) = G(f)H(f)Q(f) \quad (6.7)$$

Como mencionado anteriormente, a saída do filtro de recepção $y(t)$ da Eq. (6.5) é amostrada em sincronismo com o transmissor. Seja¹

$$y(iT_b) = \sum_{k=-\infty}^{\infty} a_k p[(i - k)T_b], \quad i = 0, \pm 1, \pm 2, \dots$$

¹ Na prática, pode ser preferível amostrar a saída do filtro de recepção $y(t)$ em tempos $t = iT_b + t_0$ ao invés de $t = iT_b$. A razão para isto é a inevitável presença de atraso e distorção na resposta total do pulso $p(t)$. Para simplificar a apresentação, ignoramos este efeito.

a amostra de $y(t)$ produzida no tempo $t = iT_b$. Para simplificar a apresentação, introduzimos a notação em tempo discreto,

$$y_i = y(iT_b)$$

e, correspondentemente,

$$p_i = p(iT_b)$$

Portanto, reescrevemos a amostra de $y(t)$ nos instantes $t = iT_b$ de forma compacta como o somatório discreto de convolução

$$y_i = \sum_{k=-\infty}^{\infty} a_k p_{i-k}, \quad i = 0, \pm 1, \pm 2, \dots \quad (6.8)$$

Referindo à Fig. 6.1(a), vemos que $y_i = y(iT_b)$ é a entrada do dispositivo de tomada de decisão.

Definindo

$$p_0 = p(0) = \sqrt{E} \quad (6.9)$$

na qual E é a *energia por bit (símbolo) do sinal transmitido*. O índice i se refere ao instante no qual a saída do filtro de recepção é amostrada no receptor, enquanto que o índice k se refere a um símbolo na seqüência de dados produzida pela fonte de informação na entrada do transmissor. Portanto, isolando o termo representando $k = i$ na Eq. (6.8), podemos escrever, de forma equivalente,

$$y_i = \sqrt{E} a_i + \sum_{\substack{k=-\infty \\ k \neq i}}^{\infty} a_k p_{i-k}, \quad i = 0, \pm 1, \pm 2, \dots \quad (6.10)$$

na Eq. (6.10), o primeiro termo a_i representa o símbolo binário transmitido, exceto pelo fator de escala \sqrt{E} . O segundo termo, envolvendo o efeito combinatório de todos os outros símbolos binários transmitidos antes e depois de a_i , representa o fenômeno residual chamado de *interferência intersimbólica (IIS)*. Na ausência de IIS (e da já considerada ausência de ruído de canal), a Eq. (6.10) se reduz para a condição ideal

$$y_i = \sqrt{E} a_i, \quad \text{para todo } i$$

a qual, obviamente, representa a decodificação perfeita.

O problema de formatação de pulso envolvido no projeto do sistema PAM da Fig. 6.1(a) pode, agora, ser afirmado como (ignorando o efeito de ruído de canal):

Dada a função de transferência $H(f)$ do canal, determine o espectro do pulso transmitido $G(f)$ e a função de transferência do filtro de recepção $Q(f)$ de forma a satisfazer os dois requisitos básicos:

- (i) *A interferência intersimbólica é reduzida para zero.*
- (ii) *A largura de faixa de transmissão é conservada.*

Para satisfazer estes dois requisitos, precisamos ter o controle sobre a forma total do pulso $p(t)$ no domínio do tempo ou, de forma equivalente, o espectro total do pulso $P(f)$ no domínio da freqüência. A questão chave é como este controle é realmente exercido.

6.3 CANAL DE NYQUIST

À luz das Eqs. (6.5) e (6.7), podemos substituir o sistema PAM da Fig. 6.1(a) pela configuração PAM mais simples mostrada na parte (b) da figura. A característica distinta da Fig. 6.1(b) é seu foco no espectro total do pulso $P(f)$.

Para a *solução ótima* do problema de formatação de pulso, a condição para a interferência intersimbólica nula deve ser satisfeita para o *mínimo* possível de largura de faixa de transmissão. Com $y_i = a_p$, para todo i definindo a condição para interferência intersimbólica nula, inferimos da Eq. (6.10) que é necessário que, para o formato geral do pulso $p(t)$, a transformada inversa de Fourier do espectro do pulso $P(f)$ da Fig. 6.1(b) satisfaça a condição

$$p_i = p(iT_b) = \begin{cases} \sqrt{E}, & \text{para } i = 0 \\ 0, & \text{para todo } i \neq 0 \end{cases} \quad (6.11)$$

A Eq. (6.11) implica na amostragem de $p(t)$ em uma taxa uniforme igual à *taxa de bit* $1/T_b$. Suponha que $p(t)$ é limitado em faixa para freqüência no intervalo $-B_0 < f < B_0$, no qual B_0 deve ser definido. Então, invocando a *fórmula de interpolação* da Eq. (5.7), embutida no *teorema da amostragem*, podemos expressar a forma do pulso $p(t)$ em termos dos valores de suas amostras como

$$p(t) = \sum_{i=-\infty}^{\infty} p\left(\frac{i}{2B_0}\right) \operatorname{sinc}(2B_0t - i) \quad (6.12)$$

Suponha que a largura de faixa B_0 é relativa à taxa de bit $1/T_b$ por

$$B_0 = \frac{1}{2T_b} \quad (6.13)$$

Então, substituindo a Eq. (6.11) na Eq. (6.12), obtemos a função sinc

$$\begin{aligned} p_{\text{oti}}(t) &= \sqrt{E} \operatorname{sinc}(2B_0t) \\ &= \frac{\sqrt{E} \sin(2\pi B_0 t)}{2\pi B_0 t} \end{aligned} \quad (6.14)$$

como a *forma ótima de pulso*.

► **PROBLEMA 6.1** A forma do pulso $p(t)$ de um sistema PAM banda base binário é definida por

$$p(t) = \operatorname{sinc}\left(\frac{t}{T_b}\right)$$

na qual T_b é a duração do bit do dado binário de entrada. Os níveis de amplitude na saída do gerador de pulso são +1 e -1, dependendo se o símbolo binário na entrada for 1 ou 0, respectivamente. Descreva a forma na saída do filtro de recepção em resposta à seqüência de dados de entrada 001101001. ◀

O espectro geral do pulso é definido pela função ótima retangular, também chamada de *função parede de tijolos*

$$P_{\text{oti}}(f) = \begin{cases} \frac{\sqrt{E}}{2B_0}, & \text{para } -B_0 < f < B_0 \\ 0, & \text{caso contrário} \end{cases} \quad (6.15)$$

As partes (a) e (b) da Fig. 6.2, respectivamente, mostram o espectro ótimo $P_{\text{oti}}(f)$ e sua inversa $p_{\text{oti}}(t)$.

Os pontos importantes a serem observados dos gráficos idealizados da Fig. 6.2 são:

1. O espectro da função retangular $P_{\text{oti}}(f)$ define B_0 como a *menor largura de faixa de transmissão para interferência intersimbólica nula*. Ou seja, a solução ótima para o problema de formatação de pulso envolve nenhuma freqüência com valor absoluto excedendo metade da taxa de bit. O parâmetro B_0 definido de acordo com a Eq.

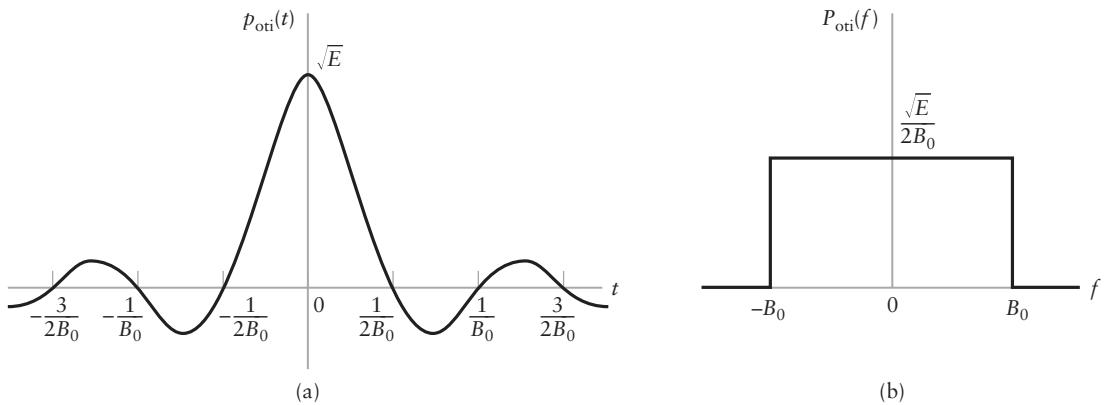


Figura 6.2 (a) Função sinc $p(t)$ como formato de pulso ótimo. (b) Espectro do pulso ótimo.

(6.13) é chamado de *largura de faixa de Nyquist*. Correspondentemente, o sistema PAM da Fig. 6.1(b) com o espectro de pulso ótimo $P_{\text{oti}}(f)$ definido pela Eq. (6.15) é chamado de *canal de Nyquist*.

2. A forma de pulso ótima $p_{\text{oti}}(t)$ é a resposta ao impulso de um canal passa-baixa ideal com resposta em amplitude de $P_{\text{oti}}(f)$ na faixa passante e largura de faixa B_0 . Sendo definida como a função sinc, $p_{\text{oti}}(t)$ possui seu valor de pico na origem e passa por zero em intervalos múltiplos à duração de bit T_b . Com esta base, os pulsos definidos por $p_{\text{oti}}(t - kT_b)$ na Eq. (6.5) com $k = 0, \pm 1, \pm 2, \dots$, não irão interferir um com o outro.

Resumindo, o canal de Nyquist definido pelo espectro geral do pulso $P(f)$ da Eq. (6.15) é a solução óptima para interferência intersimbólica nula na menor largura de faixa de transmissão possível em um ambiente sem ruído.

Apesar do canal de Nyquist ser de fato a solução óptima para o problema de formatação de pulso, existem duas dificuldades em seu uso que tornam o sistema PAM impraticável:

1. O sistema requer que o espectro $P(f)$ seja plano de $-B_0$ a B_0 e zero em todo o restante. Isto é fisicamente irrealizável e muito difícil de ser aproximado na prática, devido às transições abruptas em $\pm B_0$.
2. A função no tempo $p(t)$ diminui para $1/|t|$ para $|t|$ grande, resultando em uma lenta taxa de decaimento. Isto é causado pela descontinuidade de $P(f)$ em $\pm B_0$. Desta forma, praticamente não existe margem de erro nos tempos de amostragem no receptor.

Para entendermos o problema de erro de temporização mostrado no ponto 2, considere a Eq. (6.5) e a amostra de $y(t)$ no tempo $t = \Delta t$, na qual Δt é o erro de temporização. Para simplificar a análise, iremos colocar o termo de amostragem correto iT_b como sendo zero. Obtemos, portanto, (na ausência de ruído de canal)

$$y(\Delta t) = \sqrt{E} \sum_{k=-\infty}^{\infty} a_k p(\Delta t - kT_b), \quad \text{para } iT_b = 0$$

Fazendo $p(t)$ igual ao valor ótimo definido na primeira linha da Eq. (6.14), temos

$$y(\Delta t) = \sum_{k=-\infty}^{\infty} a_k \text{sinc}[2B_0(\Delta t - kT_b)]$$

com $2B_0 T_b = 1$, de acordo com a Eq. (6.13), podemos simplificar a expressão de $y(\Delta t)$ em

$$\begin{aligned} y(\Delta t) &= \sqrt{E} \sum_{k=-\infty}^{\infty} a_k \text{sinc}(2B_0 \Delta t - k) \\ &= \sqrt{E} a_0 \text{sinc}(2B_0 \Delta t) + \sqrt{E} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} a_k \frac{\text{sen}(2\pi B_0 \Delta t - \pi k)}{2\pi B_0 \Delta t - \pi k} \end{aligned}$$

na qual, na segunda linha, fizemos duas coisas distintas: isolamos o termo correspondente a $k = 0$ e utilizamos a fórmula da função sinc. A seguir, utilizando a identidade trigonométrica

$$\sin(2\pi B_0 \Delta t - \pi k) = \sin(2\pi B_0 \Delta t) \cos(\pi k) - \cos(2\pi B_0 \Delta t) \sin(\pi k)$$

com $\cos(\pi k) = (-1)^k$ e $\sin(\pi k) = 0$ para todo k , podemos ir um passo adiante e escrever

$$y(\Delta t) = \sqrt{E}a_0 \operatorname{sinc}(2B_0 \Delta t) + \sqrt{E} \left(\frac{\sin(2\pi B_0 \Delta t)}{\pi} \right) \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \frac{(-1)^k a_k}{2B_0 \Delta t - k} \quad (6.16)$$

O primeiro termo do lado direito da Eq. (6.16) define o símbolo binário desejado, enquanto que a série restante representa a *interferência intersimbólica causada pelo erro de temporização Δt* na amostragem do sinal $y(t)$. A interferência intersimbólica causada desta forma decai na taxa $1/\Delta t$. Além disso, dependendo do valor de Δt , é possível que esta série divirja, causando decisões erradas no receptor.

Para mitigar os problemas devido à realização física do canal de Nyquist e erro de temporização nulo, precisamos, portanto, olhar para outras formas de pulso. Na próxima seção, apresentaremos uma solução que relaxa a solução (ideal) de largura de faixa mínima embutida no canal de Nyquist.

6.4 ESPECTRO DO PULSO DE COSENTO LEVANTADO

Para garantir a realização física do espectro geral do pulso $P(f)$, precisamos de uma solução que seja diferente do canal de Nyquist por uma importante característica: a *$P(f)$ modificada diminui em direção a zero gradualmente e não abruptamente*. Em termos mais específicos, $P(f)$ proposto é constituído por duas partes:

1. *Porção plana*, a qual ocupa a faixa de freqüências $0 \leq |f| \leq f_1$ para algum parâmetro f_1 a ser definido.
2. *Porção de decaimento (roll-off)*, a qual ocupa a faixa de freqüências $f_1 < |f| < 2B_0 - f_1$.

O parâmetro f_1 é *ajustável* sob o controle do projetista. A porção plana pode, portanto, manter parte da solução retangular ideal. Para a porção de decaimento é necessário o decréscimo gradual de $P(f)$ para zero. A questão chave é como formular esta característica gradual de decaimento. Para a resposta, primeiro olhamos para *um ciclo completo da função cosseno definida no domínio da freqüência, a qual é levantada por um total igual a sua amplitude*. O que deve ser feito a seguir é separar esta função na origem em duas metades iguais, separadas por um total igual à largura $2f_1$ da porção plana do ponto 1 e, finalmente, utilizá-las para levar em consideração a porção de decaimento sob o ponto 2. Esta nova figura constitui a forma desejada para o espectro do pulso modificado $P(f)$. Por razões óbvias, $P(f)$ reconstruída desta maneira é chamada de *espectro do pulso de cosseno levantado*.

Para definirmos o espectro do pulso de cosseno levantado em termos matemáticos, escrevemos,

$$P(f) = \begin{cases} \frac{\sqrt{E}}{2B_0}, & 0 \leq |f| < f_1 \\ \frac{\sqrt{E}}{4B_0} \left\{ 1 + \cos \left[\frac{\pi(|f| - f_1)}{2(B_0 - f_1)} \right] \right\}, & f_1 \leq |f| < 2B_0 - f_1 \\ 0, & 2B_0 - f_1 \leq |f| \end{cases} \quad (6.17)$$

A freqüência f_1 e a largura de faixa de Nyquist B_0 são relacionadas pelo novo parâmetro

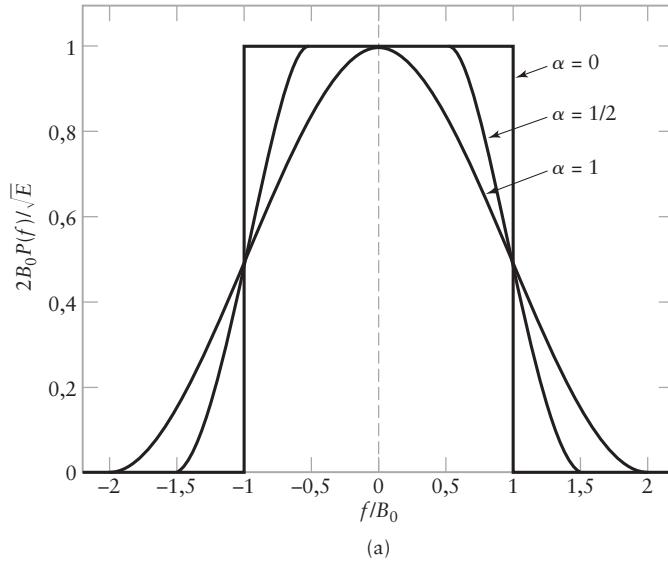
$$\alpha = 1 - \frac{f_1}{B_0} \quad (6.18)$$

o qual é chamado de *fator de roll-off* (ou fator de decaimento), para $\alpha = 0$, ou seja, $f_1 = B_0$, obtemos o canal de Nyquist descrito na Seção 6.3.

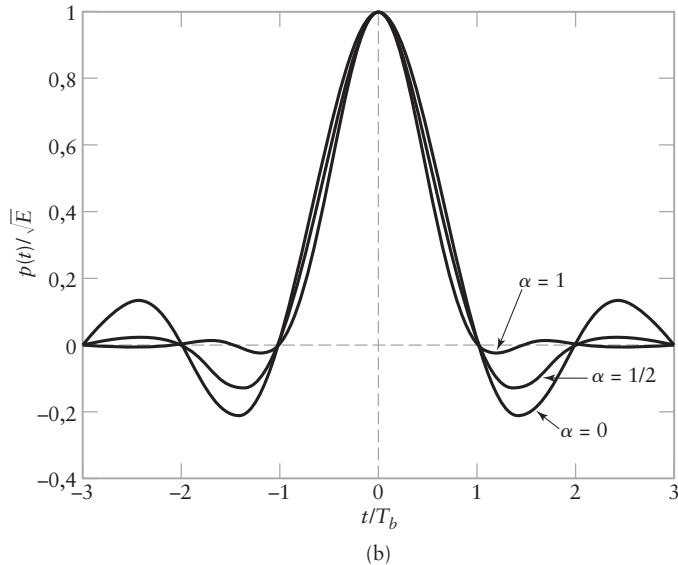
O espectro $P(f)$ normalizado pela multiplicação de $2B_0/\sqrt{E}$, é traçado na Fig. 6.3(a) em função da freqüência normalizada f/B_0 para três valores de α : 0, 0,5 e 1. Vemos que para $\alpha = 0,5$ ou 1, a função $P(f)$ corte gradualmente quando comparada com a solução retangular ideal (correspondendo a $\alpha = 0$).

A forma do pulso modificado $p(t)$, ou seja, a transformada inversa de Fourier do espe-
ctro do pulso de cosseno levantado $P(f)$, é definida por

$$p(t) = \sqrt{E} \operatorname{sinc}(2B_0t) \left(\frac{\cos(2\pi\alpha B_0 t)}{1 - 16\alpha^2 B_0^2 t^2} \right) \quad (6.19)$$



(a)



(b)

Figura 6.3 (a) Espectro do pulso de cosseno levantado para diferentes valores de roll-off. (b) Res-
posta do pulso $p(t)$ (isto é, transformada inversa de Fourier de $P(f)$ para taxas diferentes de roll-off).

A função no tempo da Eq. (6.19) é constituída pelo produto de dois fatores: o fator $\sqrt{E} \operatorname{sinc}(2B_0 t)$ que define o canal de Nyquist e um segundo fator que diminui para $1/|t|^2$ para $|t|$ grande. O primeiro fator garante os cruzamentos por zero de $p(t)$ nos instantes de amostragem $t = iT_b$ desejados, como sendo um inteiro (positivo ou negativo). O segundo fator reduz a cauda do pulso consideravelmente abaixo do canal de Nyquist, de tal forma que a transmissão de ondas binárias utilizando tais pulsos é relativamente insensível a erros de tempo de amostragem. De fato, o total de interferência intersimbólica resultante de um erro de temporização Δt diminui quando o fator α de roll-off aumenta de zero para a unidade.

A forma do pulso modificado $p(t)$ é traçada na Fig. 6.3(b) para $\alpha = 0, 0,5$ e 1 . Para o caso especial de $\alpha = 1$, a função $p(t)$ é simplificada para

$$p(t) = \sqrt{E} \left(\frac{\operatorname{sinc}(4B_0 t)}{1 - 16B_0^2 t^2} \right) \quad (6.20)$$

A função $p(t)$ da Eq. (6.20) exibe duas propriedades interessantes:

1. Para $t = \pm T_b/2 = \pm 1/(4B_0)$, temos $p(t) = 0,5\sqrt{E}$, ou seja, a largura do pulso medida na metade da amplitude é exatamente igual à duração T_b do bit.
2. Existem cruzamentos de zero em $t = \pm 3T_b/2, \pm 5T_b/2, \dots$, além dos cruzamentos com zero normais nos instantes de amostragem $t = \pm T_b, \pm 2T_b, \dots$

Estas duas propriedades são particularmente úteis para fornecer a base para a extração de sinais de temporização da saída $y(t)$ do filtro de recepção, os quais são utilizados para a sincronização do receptor com o transmissor.

- **PROBLEMA 6.2** Mostre que para freqüências positivas, a área sob a curva de cosseno levantado normalizado $P(f)/(\sqrt{E}/2B_0)$ em função de f/B_0 é unitária para todos os valores de fator de roll-off na faixa $0 \leq \alpha \leq 1$. Uma afirmativa similar é válida para freqüências negativas. ◀
- **PROBLEMA 6.3** Dado que $P(f)$ é a transformada de Fourier de uma função do tipo pulso $p(t)$, podemos apresentar o seguinte teorema:²

O pulso $p(t)$ diminui assintoticamente com o tempo para $1/t^{k+1}$, desde que as duas condições a seguir sejam válidas:

1. As primeiras $k - 1$ derivadas da transformada de Fourier $P(f)$ com respeito à freqüência f são contínuas.
2. A k -ésima derivada de $P(f)$ é descontínua.

Demonstre a validade desse teorema para os três valores diferentes de α mostrados na Fig. 6.3(a). ◀

- **PROBLEMA 6.4** A Eq. (6.17) define o espectro $P(f)$ do pulso de cosseno levantado como sendo de valor real e, portanto, atraso nulo. Na prática, todo sistema de transmissão vivencia algum atraso finito. Para acomodar esta questão prática, podemos associar a $P(f)$ uma característica de fase linear na faixa de freqüência $0 \leq |f| \leq 2B_0 - f_1$.
 - (a) Mostre que esta modificação em $P(f)$ introduz um atraso finito em sua transformada inversa de Fourier, ou seja, na forma do pulso $p(t)$.
 - (b) De acordo com a Eq. (6.19), $p(t)$ representa uma resposta não causal no tempo. O atraso introduzido em $p(t)$ pela modificação de $P(f)$ também possui um efeito benéfico, tendendo a tornar $p(t)$ essencialmente causal. Para que isto aconteça, entretanto, o atraso não deve ser menor do que um certo valor que depende do fator α de roll-off. Sugira valores adequados para este atraso considerando $\alpha = 0, 1/2$ e 1 . ◀

² Para a discussão detalhada deste teorema, veja Gitlin, Hayes e Weinstein (1992), p. 258.

■ Requisito de largura de faixa de transmissão

Da Eq. (6.17) vemos que a porção não nula do espectro $P(f)$ do pulso de cosseno levantado é limitada ao intervalo $(0, 2B_0 - f_1)$ para freqüências positivas. Desta forma, a largura de faixa de transmissão necessária pela utilização do espectro do pulso de cosseno levantado é dada por

$$B_T = 2B_0 - f_1$$

Eliminando f_1 entre a Eq. (6.18) e esta fórmula para B_T , obtemos

$$B_T = B_0(1 + \alpha) \quad (6.21)$$

na qual B_0 é a largura de faixa de Nyquist e α é o fator de roll-off. Portanto, a largura de faixa necessária para o espectro de cosseno levantado excede o canal ótimo de Nyquist pelo total

$$f_v = \alpha B_0 \quad (6.22)$$

o qual é chamado de *excesso de largura de faixa*. Desta definição, facilmente vemos que a razão de excesso de largura de faixa f_v (resultante da utilização do espectro do pulso de cosseno levantado) para a largura de faixa B_0 de Nyquist (necessária pelo canal de Nyquist) é igual ao fator α de roll-off. Este é o motivo pelo qual o fator de roll-off é, algumas vezes, chamado de *fator de excesso de largura de faixa*. De qualquer forma, a necessidade de acomodar o excesso de largura de faixa f_v é o preço que precisamos pagar pela transmissão de dados binários em um canal a um taxa igual a $1/T_b$, com uma interferência intersimbólica nula e de forma fisicamente realizável.

Os dois casos a seguir, um ideal e outro prático, são de interesse particular:

1. Quando o fator α de roll-off é zero, o excesso de largura de faixa f_v é reduzido para zero, permitindo, portanto, que a largura de faixa de transmissão B_T assuma seu valor mínimo possível igual a $B_0 = 1/(2T_b)$.
2. Quando o fator α de roll-off é unitário, o excesso de largura de faixa é aumentado a B_0 . Desta forma, a largura de faixa de transmissão B_T é dobrada quando comparada ao caso (ideal) 1. Como mencionado anteriormente, a escolha de $\alpha = 1$ é a base para a sincronização entre o receptor e o transmissor.

EXEMPLO 6.1 Requisito de largura de faixa para o sistema T1

No Capítulo 5, descrevemos o formato do sinal para o *sistema de portadora T1* o qual é utilizado para multiplexar 24 entradas independentes de voz, baseado em uma palavra PCM de 8 bits. A duração do bit do sinal resultante multiplexado por divisão no tempo (incluindo o bit de quadro) é

$$T_b = 0,647 \mu\text{s}$$

A taxa de bit do sistema T1 é

$$R_b = \frac{1}{T_b} = 1,544 \text{ Mb/s}$$

Para $\alpha = 0$, a largura de faixa de Nyquist do sistema T1 é

$$B_0 = \frac{1}{2T_b} = 772 \text{ kHz}$$

a qual é a menor largura de faixa de transmissão do sistema T1 para interferência intersimbólica nula. Entretanto, um valor mais realístico para a largura de faixa de transmissão é obtido utilizando-se o espectro do pulso de cosseno levantado com fator roll-off $\alpha = 1$. Neste caso, a utilização da Eq. (6.21) resulta em

$$B_T = 0,772(1 + 1) = 1,544 \text{ MHz}$$

a qual é o dobro da largura de faixa B_0 de Nyquist.

■ Duas propriedades adicionais do espectro do pulso de cosseno levantado

Da definição do espectro $P(f)$ do pulso de cosseno levantado dada na Eq. (6.17), vemos que ela exibe duas outras propriedades interessantes, como descrito no material a seguir.

- **PROPRIEDADE 1** A porção amortecida (roll-off) do espectro $P(f)$ exibe simetria ímpar com relação ao ponto médio $f = \pm B_0$

Para provar esta propriedade, definimos a função em freqüência

$$P_v(f) = P_{\text{oti}}(f) - P(f) \quad (6.23)$$

a qual é uma caracterização única da porção amortecida do espectro de cosseno amortecido. Utilizando as Eqs. (6.15) e (6.17) nesta definição, temos

$$P_v(f) = \begin{cases} 0, & \text{para } 0 \leq |f| \leq f_1 \\ \frac{\sqrt{E}}{4B_0} \left\{ 1 - \cos \left[\frac{\pi(|f| - f_1)}{2(B_0 - f_1)} \right] \right\}, & \text{para } f_1 \leq |f| \leq B_0 \\ -\frac{\sqrt{E}}{4B_0} \left\{ 1 + \cos \left[\frac{\pi(|f| - f_1)}{2(B_0 - f_1)} \right] \right\}, & \text{para } B_0 \leq |f| \leq 2B_0 - f_1 \\ 0, & \text{para } 2B_0 - f_1 < |f| \leq 2B_0 \end{cases} \quad (6.24)$$

A Fig. 6.4(a) mostra $P_{\text{oti}}(f)$ e $P(f)$ para o fator de roll-off $\alpha = 1/2$. Apenas os gráficos para freqüências positivas são mostrados na figura. A função residual $P_v(f)$ é mostrada na Fig. 6.4(b). Desta figura, vemos imediatamente que

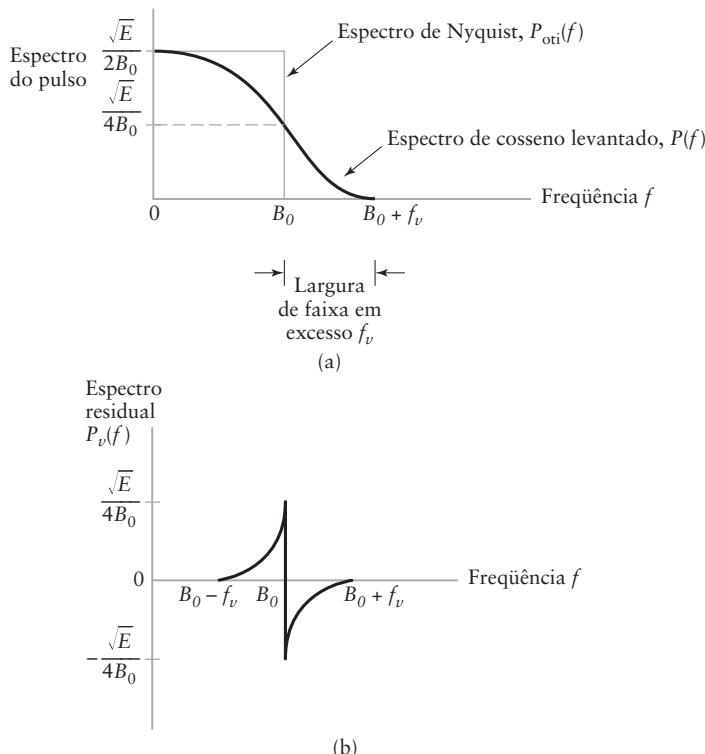


Figura 6.4 (a) Espectros de Nyquist e do pulso de cosseno levantado para freqüências positivas. (b) Espectro residual $P_v(f)$.

$$P_\nu(-f') = -P_\nu(f') \quad (6.25)$$

na qual

$$f' = f - B_0 \quad (6.26)$$

A Eq. (6.25) confirma a propriedade de simetria ímpar da porção amortecida com relação ao ponto médio $f = B_0$ para freqüências positivas. Similarmente, esta propriedade também é válida para o outro ponto médio $f = -B_0$ para freqüências negativas.

- **PROBLEMA 6.5** Começando da fórmula da Eq. (6.24) e utilizando a definição da Eq. (6.26), demonstre a propriedade da Eq. (6.25). ◀

A característica espectral mostrada na Fig. 6.4(a) nos lembra de uma situação similar pertencente à modulação de faixa lateral vestigial, a qual foi estudada no Capítulo 3. Especificamente, comparando as Figs. 3.24(a) e 6.4(a), facilmente vemos que apesar destas duas características se referirem a aplicações completamente diferentes, elas possuem, basicamente, a mesma formulação matemática, exceto por duas pequenas diferenças:

1. O espectro banda base $P(f)$ do pulso de cosseno levantado da Fig. 6.4(a) é centrado na origem em $f = 0$, enquanto que o espectro da faixa lateral vestigial da Fig. 3.24(a) é centrado na freqüência f_c da portadora senoidal.
2. O parâmetro f_v na Fig. 6.4(a) se refere ao excesso de largura de faixa medido com relação à solução retangular ideal para interferência intersimbólica nula, enquanto que o parâmetro f_v na Fig. 3.24(a) se refere ao excesso de largura de faixa medido com relação à largura de faixa ótima obtida pela modulação de faixa lateral única.

A diferença 1 é de natureza técnica e a diferença 2 simplesmente pertence à terminologia. O que realmente importa é o fato de que nas duas aplicações, na qual o espectro do pulso de cosseno levantado é utilizado para a transmissão de dados digitais banda base e o espectro da faixa lateral vestigial é geralmente (mas não sempre) utilizado para modulação analógica, a motivação é garantir a realização física.³ (O Exemplo de Tema sobre televisão digital a ser apresentado no Capítulo 7 utiliza a modulação por faixa lateral vestigial.)

- **PROPRIEDADE 2** O somatório infinito de réplicas do espectro do pulso de cosseno levantado, espaçados por $2B_0$ hertz, é igual a uma constante, como mostrado por

$$\sum_{m=-\infty}^{\infty} P(f - 2mB_0) = \frac{\sqrt{E}}{2B_0} \quad (6.27)$$

Para provar esta propriedade, lembre-se do Capítulo 2 que a amostragem de uma função que possui transformada de Fourier no domínio do tempo é transformada em periodicidade no domínio da freqüência, como descrito na Eq. (2.88). Adaptando esta equação para a situação em estudo, podemos escrever

$$\sum_{n=-\infty}^{\infty} p\left(\frac{n}{2B_0}\right) \delta\left(t - \frac{n}{2B_0}\right) \iff 2B_0 \sum_{m=-\infty}^{\infty} P(f - 2mB_0) \quad (6.28)$$

³ É ainda mais interessante observar que Harry Nyquist é o originador tanto da modulação por faixa lateral vestigial quanto do espectro do pulso de cosseno levantado:

O formato espectral para a modulação por faixa lateral vestigial aparece no artigo H. Nyquist e K. W. Pfleger, "Effect of the quadrature component in single-sideband transmission", *The Bell System Technical Journal*, vol 19, pp. 63-73, janeiro de 1940.

O espectro do pulso de cosseno levantado foi descrito no clássico artigo: H. Nyquist, "Certain topics in telegraph transmission theory", *Transactions of the American Institute of Electrical Engineers*, vol 47, pp. 617-644, abril de 1928.

O espectro $P(f)$ do pulso de cosseno levantado e sua inversa $p(t)$ são, respectivamente, definidos nas Eqs. (6.17) e (6.19). Em particular, amostrando a resposta $p(t)$ do pulso modificado na taxa $1/2B_0$, podemos escrever

$$p\left(\frac{n}{2B_0}\right) = \sqrt{E} \operatorname{sinc}(n) \left(\frac{\cos(\pi n \alpha)}{1 - 4n^2 \alpha^2} \right)$$

na qual $n = 0, \pm 1, \pm 2, \dots$. Observando dois pontos:

$$\begin{aligned} 1. \operatorname{sinc}(n) &= \frac{\sin(n\pi)}{n\pi} \\ &= \begin{cases} 1, & \text{para } n = 0 \\ 0, & \text{para } n = \pm 1, \pm 2, \dots \end{cases} \\ 2. \cos(\pi n \alpha) &= 1 \text{ para } n = 0, \end{aligned}$$

Temos que

$$p\left(\frac{n}{2B_0}\right) = \begin{cases} \sqrt{E}, & \text{para } n = 0 \\ 0, & \text{para } n = \pm 1, \pm 2, \dots \end{cases}$$

Desta forma, a Eq. (6.28) se reduz para

$$\sqrt{E}\delta(t) \iff 2B_0 \sum_{m=-\infty}^{\infty} P(f - 2mB_0)$$

ou, de forma equivalente,

$$\frac{\sqrt{E}}{2B_0}\delta(t) \iff \sum_{m=-\infty}^{\infty} P(f - 2mB_0) \quad (6.29)$$

Finalmente, observando que a transformada de Fourier da função delta $\delta(t)$ é unitária, a Eq. (6.29) é simplesmente outra forma de descrever a forma mostrada na Eq. (6.27).

Tendo provada a validade da Propriedade 2, podemos fazer o caminho inverso. Especificadamente, o critério para a formatação de pulso para a interferência intersimbólica está englobada pela seguinte afirmativa genérica:

Dado a forma do pulso modificado $p(t)$ para a transmissão de dado em um canal imperfeito utilizando a modulação por amplitude de pulso na taxa de dados de $1/T_b$, o formato do pulso $p(t)$ elimina a interferência intersimbólica se e somente se, seu espectro $P(f)$ satisfaz a condição

$$\sum_{m=-\infty}^{\infty} P\left(f - \frac{m}{T_b}\right) = \text{constante}, \quad |f| \leq \frac{1}{2T_b} \quad (6.30)$$

Esta afirmativa inclui o PAM binário como caso especial, para o qual a taxa de dados $1/T_b$ é igual a $2B_0$ de acordo com a Eq. (6.13). Além disso, o espectro do pulso de cosseno levantado é um exemplo, apesar de ser um importante exemplo, que satisfaz a Eq. (6.30).

■ Espectro do pulso da raiz do cosseno levantado

Uma forma de pulso mais sofisticada para a transmissão de dados digitais banda base é o espectro do pulso da raiz do cosseno levantado ao invés do espectro regular do pulso da Eq. (6.17). Especificamente, escrevemos,

$$G(f)H(f) = P^{1/2}(f) \quad (6.31)$$

na qual, como antes, $G(f)$ é a resposta em freqüência do filtro de transmissão e $H(f)$ é a resposta em freqüência do canal. Desta forma, a resposta em freqüência do filtro de recepção é definida por

$$Q(f) = P^{1/2}(f) \quad (6.32)$$

Multiplicando a Eq. (6.31) pela Eq. (6.32), temos

$$G(f)H(f)Q(f) = P(f)$$

a qual é a repetição da Eq. (6.7). Com esta base, o formato do pulso é partitionado igualmente entre duas entidades:

- A combinação do filtro de transmissão e do canal constitui uma entidade. Com $H(f)$ conhecido e $P(f)$ definido pela Eq. (6.17) para um fator α de roll-off preestabelecido, podemos utilizar a Eq. (6.31) para determinar a resposta em freqüência do filtro de transmissão.
- O filtro de recepção constitui a outra entidade. Logo, para o mesmo fator α de roll-off podemos utilizar as Eqs. (6.17) a (6.32) para determinar a resposta em freqüência do filtro de recepção.

Se o canal for afetado por ruído aditivo e o formato do pulso é partitionado igualmente entre o transmissor e receptor da forma descrita, então o receptor irá maximizar a razão sinal/ruído de saída nos instantes de amostragem. Outras discussões sobre este assunto serão postergadas para o Capítulo 10, o qual é dedicado ao ruído em receptores de comunicação digital.

6.5 TRANSMISSÃO BANDA BASE DE DADOS M-ÁRIO

No sistema PAM binário banda base da Fig. 6.1(a), a seqüência $\{b_k\}$ emitida pela fonte de informação consiste de símbolos binários que representam um dentre dois possíveis níveis de amplitude, -1 para o símbolo 0 e $+1$ para o símbolo 1 . Por outro lado, em uma versão *M-ário banda base* do sistema, a saída do codificador de linha assume um dentre M possíveis níveis de amplitude, com $M > 2$. Em um sistema *M-ário*, a fonte de informação emite a seqüência de símbolos de um alfabeto que consiste de M símbolos. Cada nível de amplitude da saída do codificador de linha corresponde a um símbolo distinto, de tal forma que existem M níveis distintos de amplitude para serem transmitidos.

Considere então, um sistema PAM *M-ário* com um alfabeto de sinal constituído por M símbolos, com a *duração do símbolo* representada por T segundos. Denominamos $1/T$ como a *taxa de sinalização* do sistema, a qual é expressa em *símbolos por segundo* ou, simplesmente, *bauds*. É informativo relacionar a taxa de sinalização desse sistema com a de um sistema PAM binário para o qual o valor de M é 2 e a duração de bit é T_b segundos. O sistema PAM binário transmite dados a uma taxa de $1/T_b$ bits por segundo. Também observamos que, no caso do sistema PAM *quaternário*, por exemplo, os quatro possíveis símbolos podem ser identificados pelos dibits 00 , 10 , 11 e 01 . Um *dibit* se refere a uma palavra constituída por dois bits. Vemos, portanto, que cada símbolo representa 2 bits de dados e 1 baud é igual a 2 bits por segundo. Podemos generalizar este resultado dizendo que em um sistema PAM *M-ário*, 1 baud equivale a $\log_2 M$ bits por segundo e que a duração do símbolo T do sistema PAM *M-ário* é relacionado com a duração T_b do bit de um sistema PAM binário com a taxa de bit equivalente como mostrado por:

$$T = T_b \log_2 M \quad (6.33)$$

Portanto, em uma dada largura de faixa do canal, percebemos que utilizando um sistema PAM *M-ário*, somos capazes de transmitir dados em uma taxa que é $\log_2 M$ mais rápida do que o correspondente em um sistema PAM binário.

Entretanto, esta melhoria na utilização da largura de faixa possui o seu preço. Especificamente, a potência transmitida deve ser aumentada por um fator igual a $M^2/\log_2 M$, comparada com o sistema PAM binário, se queremos ter a mesma performance na presença de ruído de canal. Esta questão é discutida no Capítulo 10. Além disso, a complexidade do sistema é aumentada.

6.6 PADRÃO DE OLHO

Até aqui discutimos o problema da interferência intersimbólica e como mitigá-lo. Nesta seção, iremos descrever uma ferramenta chamada de padrão de olho para sua determinação experimental.

O padrão de olho é produzido pela *superposição sincronizada de (a maior quantidade possível) intervalos de símbolos sucessivos da forma de onda distorcida que aparece na saída do filtro de recepção antes do limiar*. Como exemplo ilustrativo, considere a forma de onda distorcida, mas livre de ruído, mostrada na Fig. 6.5(a). A parte (b) da figura mostra a superposição sincronizada da forma de onda dos intervalos de 8 símbolos binários. A forma resultante é chamada de “padrão de olho” devido a sua semelhança com o olho humano. Pelo mesmo motivo, o interior do padrão de olho é chamado de *abertura do olho*.

Enquanto o ruído aditivo do canal não for muito grande, então o padrão de olho é bem definido e, portanto, pode ser estudado experimentalmente em um osciloscópio. A forma de onda em estudo é aplicada às placas de deflexão do osciloscópio com o circuito de base de tempo operando em condição sincronizada. Da perspectiva experimental, o padrão de olho possui duas virtudes:

- Simplicidade de geração.
- Possibilidade de fornecer muita informação sobre as características do sistema de transmissão de dados. Logo, ele é amplamente utilizado como indicador visual⁴ de

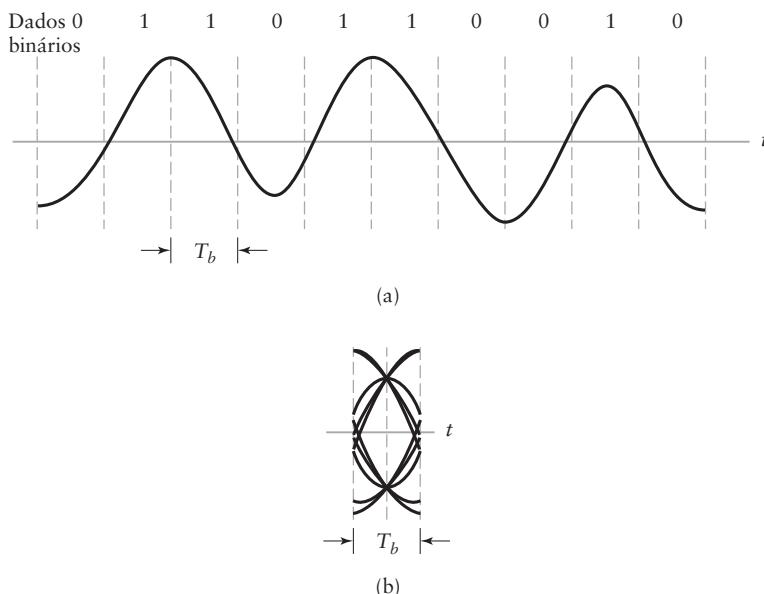


Figura 6.5 (a) Seqüência de dados binários e suas formas de onda. (b) Padrão de olho correspondente.

⁴ Outro indicador visual da performance do sistema é o chamado *diagrama de dispersão (scatter)*, o qual é obtido trançando a parte imaginária em função da parte real de um sinal complexo aparecendo na saída do filtro de recepção. Para mais detalhes, veja o livro de Jeruchim, Balaban e Shanmugan (2000), pp. 666-667.

quão bem ou mal o sistema de transmissão de dados executa a tarefa de transportar a seqüência de dados ao longo de um canal físico.

■ Características temporais

A Fig. 6.6 mostra um padrão *genérico de olho* para dados binários distorcidos, mas sem ruído. O eixo horizontal, representando o tempo, expande o intervalo de símbolo de $-T_b/2$ a $+T_b/2$, no qual T_b é a duração do bit.

A partir deste diagrama, podemos inferir três características temporais pertencentes ao sistema de transmissão de dados binários, exemplificado pelo sistema PAM da Fig. 6.1(a):

- (i) *Tempo de amostragem ótimo*: A largura da abertura de olho define o intervalo de tempo no qual a forma de onda binária distorcida que aparece na saída do filtro de recepção da Fig. 6.1(a) pode ser uniformemente amostrada sem erros de decisão. Claramente, o *tempo de amostragem ótimo* é o tempo no qual o olho está em sua maior abertura.
- (ii) *Jitter de cruzamento de zero*. Na prática, o sinal de tempo (para a sincronização do receptor com o transmissor) é extraído dos *cruzamentos de zero* da forma de onda que aparece na saída do filtro de recepção. Nesta forma de sincronização, sempre existem irregularidades no cruzamento com o zero, o que, por sua vez, resulta no *jitter* e, portanto, em tempos de amostragem não ótimos.
- (iii) *Sensibilidade de temporização*. Outra característica de temporização relacionada é a da sensibilidade do sistema a *erros de temporização*. Esta sensibilidade é determinada pela taxa na qual o padrão de olho é fechado quando o tempo de amostragem é variado.

A Fig. 6.6 indica como estas três características do sistema podem ser medidas do padrão de olho.

■ Distorção de pico para interferência intersimbólica

Até este momento temos considerado que a amplitude ideal do sinal ocupa a faixa de -1 a $+1$. Então, na ausência de ruído do canal, a abertura do olho assume dois valores extremos:

- (i) *Uma abertura de olho unitária*,⁵ a qual corresponde a interferência intersimbólica nula.

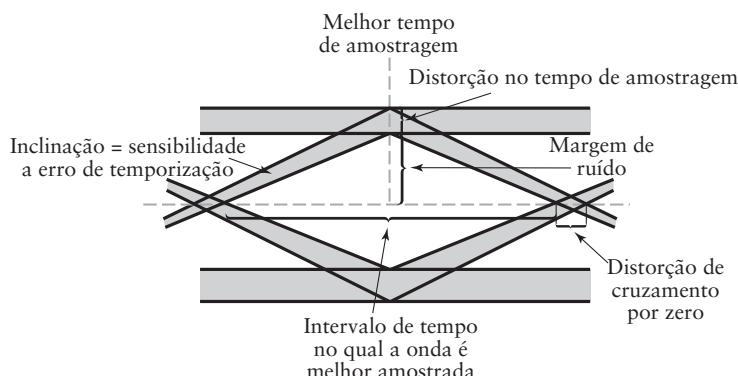


Figura 6.6 Interpretação do padrão de olho para um sistema de transmissão de dados binários banda base.

⁵ Em sentido restrito, um padrão de olho que é completamente aberto ocupa a faixa de -1 a $+1$. Com esta base, a interferência intersimbólica nula iria corresponder a uma abertura de olho ideal igual a dois. Entretanto, por duas razões, conveniência de representação e consistência com a literatura, escolhemos uma abertura de olho unitária para nos referirmos à condição ideal de interferência intersimbólica nula.

- (ii) *Uma abertura de olho nula*, a qual corresponde a um padrão de olho completamente fechado. Este segundo extremo ocorre quando o efeito da interferência intersimbólica é severa o suficiente para que alguns traços superiores do padrão de olho cruzem com os traços inferiores.

Na situação (ii), é realmente possível para o receptor cometer erros de decisão mesmo quando o canal é livre de ruído. Normalmente, *uma abertura de olho de 0,5 ou mais é considerada como resultante de uma transmissão de dados confiável*.

Em um ambiente ruidoso, a extensão da abertura de olho na tempo de amostragem ótimo fornece a medida da margem de operação sobre o ruído aditivo de canal. Esta medida, como ilustrada na Fig. 6.6, é chamada de *margem de ruído*.

A partir desta discussão, é evidente que o padrão de olho possui um importante papel na medição da performance do sistema. Logo, precisamos de uma definição formal para a abertura de olho. Para isto, oferecemos a seguinte definição:

$$(\text{abertura de olho}) = 1 - D_{\text{pico}} \quad (6.34)$$

na qual D_{pico} representa um novo critério chamado de *distorção de pico*. O ponto digno de nota aqui é que a distorção de pico é o critério de *pior caso* para a medição do efeito da interferência intersimbólica na performance (isto é, taxa de erro) de um sistema de transmissão de dados. A relação entre a abertura de olho e a distorção de pico é ilustrada na Fig. 6.7. Como a abertura de olho não possui dimensão, a distorção de pico também não possuirá dimensão. Para enfatizar esta afirmativa de ser sem dimensão, os dois valores extremos de abertura de olho definidos anteriormente são traduzidos para:

- (i) *Distorção de pico nula*, a qual ocorre quando a abertura de olho é unitária.
- (ii) *Distorção de pico unitária*, a qual ocorre quando o padrão de olho está completamente fechado.

Com isto, a distorção de pico é formalmente definida como *o valor máximo assumido pela interferência intersimbólica para todas as possíveis seqüências transmitidas, com seu cálculo dividido pelo fator de normalização igual ao valor absoluto do nível ideal de sinal para a interferência intersimbólica nula*. Referindo à Eq. (6.10), as duas componentes englobadas nesta definição são definidas por:

- (i) A componente do sinal idealizada da saída do filtro de recepção $y_i = y(iT_b)$ é definida pelo primeiro termo da Eq. (6.10), ou seja, a_i , na qual a_i é o i -ésimo símbolo codificado e energia de sinal transmitida por bit unitário.
- (ii) A interferência intersimbólica é definida pelo segundo termo, ou seja,

$$\sum_{\substack{k=-\infty \\ k \neq i}}^{\infty} a_k p_{i-k}$$

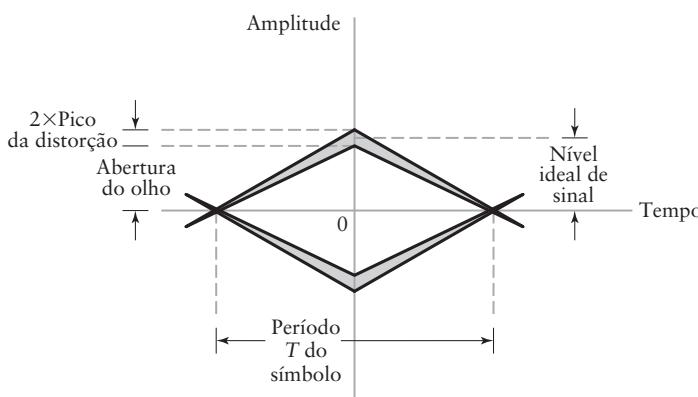


Figura 6.7 Ilustração da relação entre a distorção de pico e a abertura de olho. Nota: Os níveis ideais do sinal são escalonados para ficarem na faixa de -1 a $+1$.

O valor máximo para este somatório ocorre quando cada símbolo a_k codificado possui o mesmo sinal algébrico que p_{i-k} . Portanto,

$$(\text{IIS máximo}) = \sum_{\substack{k=-\infty \\ k \neq i}}^{\infty} |p_{i-k}|$$

Logo, invocando a definição de distorção de pico, chegamos à formula desejada:

$$\begin{aligned} D_{\text{pico}} &= \sum_{k=-\infty}^{\infty} |p_{i-k}| \\ &= \sum_{\substack{k=-\infty \\ k \neq i}}^{\infty} |p[(i-k)T_b]| \end{aligned} \quad (6.35)$$

na qual para todo $i = k$, $p_0 = p(0) = 1$. Note que, assumindo uma amplitude de sinal de -1 a $+1$, escalonamos a energia do sinal transmitido para o símbolo em $E = 1$.

Por sua própria natureza, a distorção de pico é o *criterio de pior caso para toda a transmissão de dados em um canal ruidoso*. A abertura de olho especifica a menor margem de ruído possível.⁶

■ Padrões de olho para transmissão M-ário

Como levantado anteriormente na Seção 6.5, um sistema de transmissão de dados M -ário utiliza M símbolos codificados no transmissor e $M - 1$ limiares no receptor. Desta forma, o padrão de olho para um sistema de transmissão de dados M -ário contém $(M - 1)$ aberturas de olho empilhadas verticalmente uma em cima da outra. Os limiares são definidos pelos níveis de transição de amplitude enquanto nos movemos para cima uma abertura de olho para sua adjacente. Quando os símbolos possuem mesma probabilidade, os limiares serão eqüidistantes uns dos outros.

Em um sistema restritamente linear de transmissão de dados com seqüência de dados transmitidos realmente aleatoriamente, todas as $M - 1$ aberturas de olho devem ser idênticas. Na prática, entretanto, é geralmente possível achar assimetrias no padrão de olho de um sistema de transmissão de dados M -ário, as quais são causadas por não linearidades no canal de comunicação ou por outras partes do sistema.

6.7 EXPERIMENTO DE COMPUTADOR: DIAGRAMAS DE OLHO PARA SISTEMAS BINÁRIOS E QUATERNÁRIOS

As Figs. 6.8(a) e 6.8(b) mostram diagramas de olho para sistema de transmissão PAM banda base usando $M = 2$ e $M = 4$, respectivamente. O canal não possui limite de largura de faixa e os símbolos fonte são aleatoriamente gerados no computador. Um pulso cosseno levantado é utilizado nos dois casos. Os parâmetros do sistema utilizados para a geração destes diagramas de olho são: taxa de bit = 1 Hz e fator de roll-off $\alpha = 0,5$. Para o caso binário $M = 2$ mostrado na Fig. 6.8(a), a duração T do símbolo e a duração do bit T_b são as mesmas, com $T_b = 1$ s. Para o caso de $M = 4$ mostrado na Fig. 6.8(b), temos $T = T_b \log_2 M = 2T_b$. Nos dois casos, vemos que os olhos estão abertos, indicando operação perfeita do sistema, perfeita no sentido da interferência intersímbolo ser nula.

As Figs. 6.9(a) e 6.9(b) mostram os diagramas de olho para estes dois sistemas de transmissão banda base de pulso utilizando os mesmos parâmetros anteriores para os sistemas, mas desta vez com condição de largura de faixa limitada. Especificamente, o canal

⁶ A menor margem possível de ruído coloca um *limite superior* na probabilidade do erro de símbolo devido a presença inevitável de ruído aditivo do canal. A noção de “probabilidade de erro de símbolo” é formalmente apresentada no Capítulo 10.

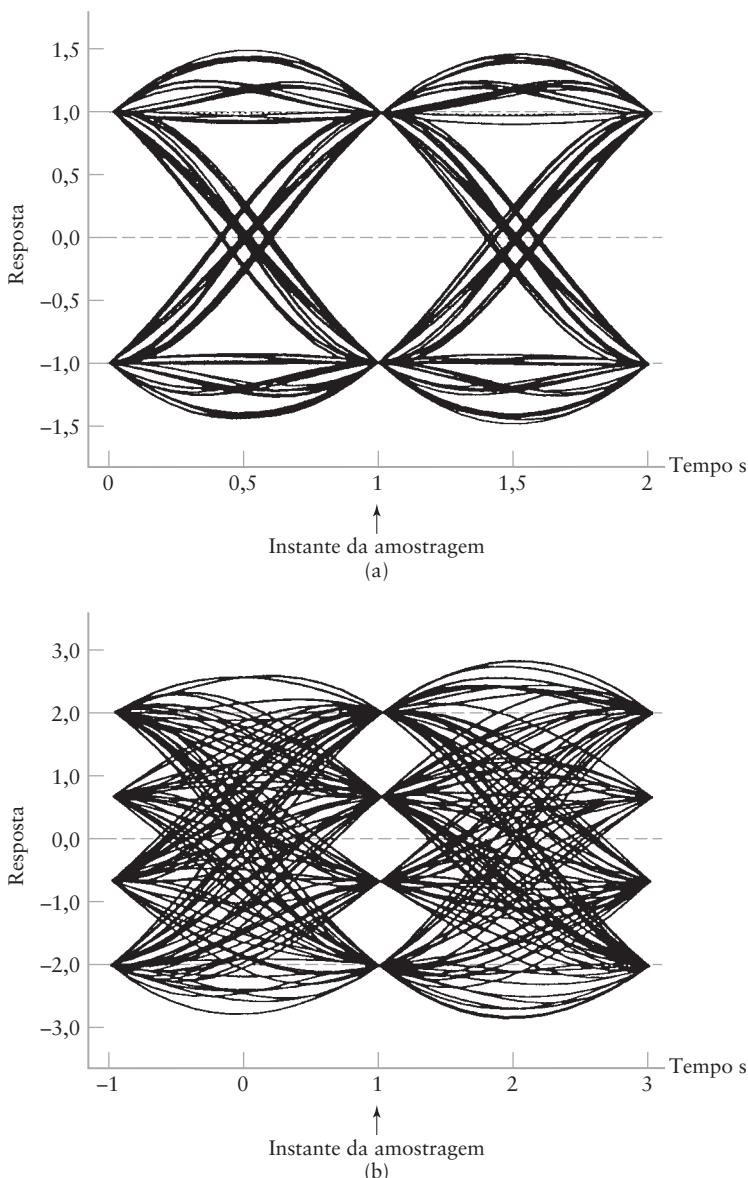


Figura 6.8 Diagrama de olho para o sinal recebido sem limitação de largura de faixa. (a) $M = 2$. (b) $M = 4$.

é, agora, modelado como um *filtro Butterworth* passa-baixa, cuja resposta em freqüência é definida por

$$|H(f)| = \frac{1}{1 + (f/f_0)^{2N}}$$

na qual N é a ordem do filtro e f_0 é a freqüência de corte de 3 dB. Para o resultado mostrado na Fig. 6.9, os seguintes parâmetros do filtro foram utilizados:

1. $N = 3$ e $f_0 = 0,6$ Hz para PAM binário
2. $N = 3$ e $f_0 = 0,3$ Hz para PAM quaternário

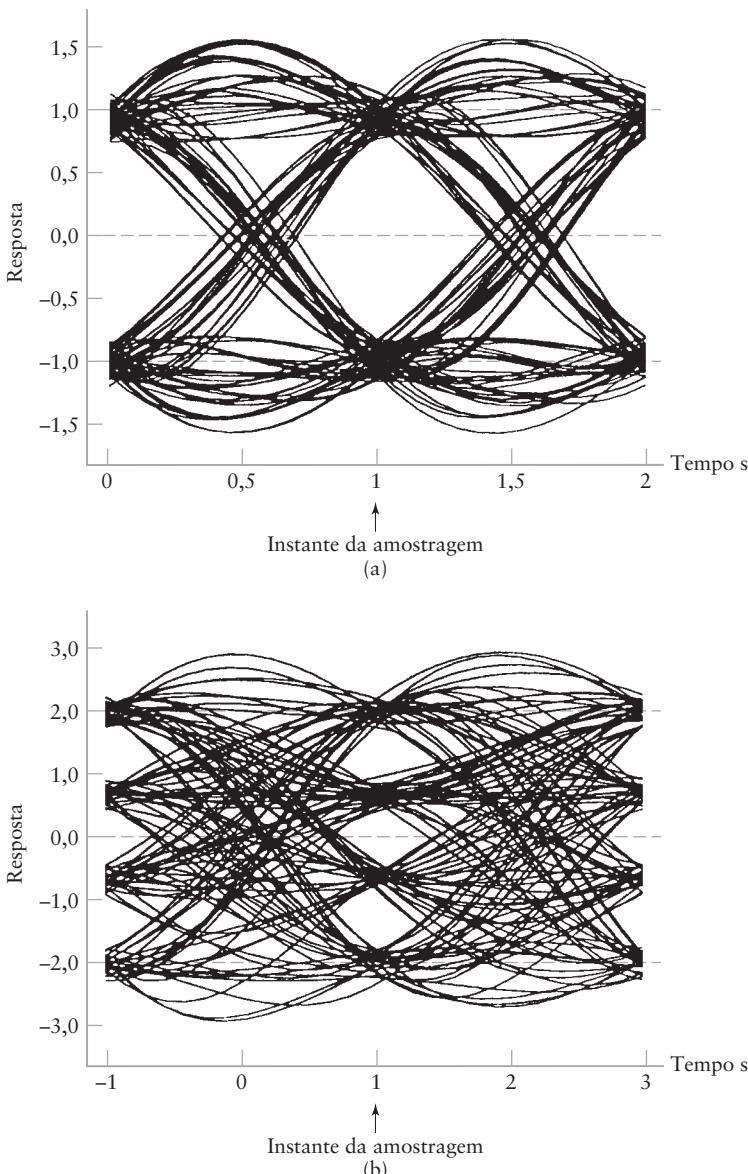


Figura 6.9 Diagrama de olho para o sinal recebido utilizando um canal com largura de faixa limitado. (a) $M = 2$. (b) $M = 4$.

Com o fator de roll-off $\alpha = 0,5$ e largura de faixa de Nyquist $B_0 = 0,5$ Hz, para PAM binário, a utilização da Eq. (6.21) define a largura de faixa de transmissão do sistema de transmissão PAM como sendo

$$B_T = 0,5(1 + 0,5) = 0,75 \text{ Hz}$$

Apesar da freqüência de corte do canal ser maior do que o absolutamente necessário, seu efeito na faixa passante é observado no decréscimo do tamanho da abertura de olho. Ao invés dos valores distintos no tempo $t = 1$ s (como mostrado nas Figs. 6.8(a) e 6.8(b)), agora existe uma região manchada. Se a largura de faixa do canal for reduzida, o olho fecharia ainda mais, até que não seria possível perceber a abertura do olho.

6.8 EXEMPLO DE TEMA: EQUALIZAÇÃO

Uma abordagem eficiente para a *transmissão em alta velocidade* de dados digitais em um canal de comunicação limitado em faixa (canais de telefonia, por exemplo), é utilizar a combinação de duas estratégias de processamento de sinal:

- ▶ *Modulação por amplitude de pulso discreta* (PAM), a qual codifica as amplitudes de pulsos sucessivos em um trem de pulso periódico por um conjunto discreto de possíveis níveis de amplitude.
- ▶ *Esquema de modulação linear*, o qual oferece a virtude de conservação de largura de faixa para transmitir o trem de pulso codificado em um canal.

Na parte do receptor do sistema de transmissão de dados, o sinal recebido é demodulado e amostrado em sincronia, então, finalmente, decisões são tomadas para os símbolos particulares que foram transmitidos. Quando a relação sinal/ruído de transmissão é alta, observamos que o número de níveis de amplitude que podem ser detectados é essencialmente limitado pela interferência intersimbólica ao invés do ruído aditivo de canal. Em princípio, se o canal for precisamente conhecido, então é praticamente possível fazer com que a interferência intersimbólica nos instantes de amostragem seja arbitrariamente pequena utilizando um par adequado de filtros de transmissão e recepção, de forma a controlar a forma do pulso da maneira descrita nas seções anteriores. Portanto, enquanto o problema de interferência intersimbólica for o ponto em questão, podemos considerar a tarefa de transmissão de dados em um canal como sendo banda base.

Na prática, entretanto, raramente temos o conhecimento prévio exato das características do canal. Além disso, existe o problema de imprecisão que aparece na implementação física de filtros de formatação de pulso. O resultado final destas questões práticas é que sempre teremos alguma *distorção residual* para a interferência intersimbólica que resultará em um fator limitante na taxa de transmissão que pode ser mantida pelo sistema. Para compensar esta distorção residual intrínseca, podemos utilizar um processo chamado de *equalização*. O filtro utilizado para executar este processo é chamado de *equalizador*.

Portanto, além da formatação de pulso executada pelo filtro de recepção, agora temos uma nova função a ser executada, a equalização da distorção residual. Como estas duas funções são lineares, propomos combiná-las em uma única estrutura. Além disso, reconhecendo a necessidade de uma estrutura com coeficientes *ajustáveis* para lidar com o processo de equalização, propomos a utilização de uma estrutura chamada de *filtro transversal*. Este filtro, mostrado na Fig. 6.10, é constituído pelos seguintes componentes:

- ▶ *Linha de atraso*, cujos pontos de tomada são uniformemente espaçados T segundos um do outro. T é a duração do símbolo.
- ▶ *Pesos ajustáveis*, os quais são conectados aos pontos de tomada da linha de atraso.
- ▶ *Somador*, o qual soma sucessivamente as versões atrasadas do sinal de entrada, após elas terem sido ponderadas individualmente.

Com a equalização de canal sendo a função de interesse e o filtro transversal com coeficientes ajustáveis à estrutura para esta função, é adequado nos referirmos a esta nova estrutura como o *equalizador transversal ajustável* ou simplesmente, *equalizador transversal*.

■ Equalização de zero forçado

Para levarmos a solução ao problema de equalização, considere o sistema composto mostrado na Fig. 6.11:

- ▶ O primeiro subsistema, caracterizado pela resposta $c(t)$ ao impulso, representa a ação combinada do filtro de transmissão e canal de comunicação.
- ▶ O segundo sistema, caracterizado pela resposta $h_{eq}(t)$ ao impulso, representa a formação de pulso combinada com a equalização de distorção residual no receptor.

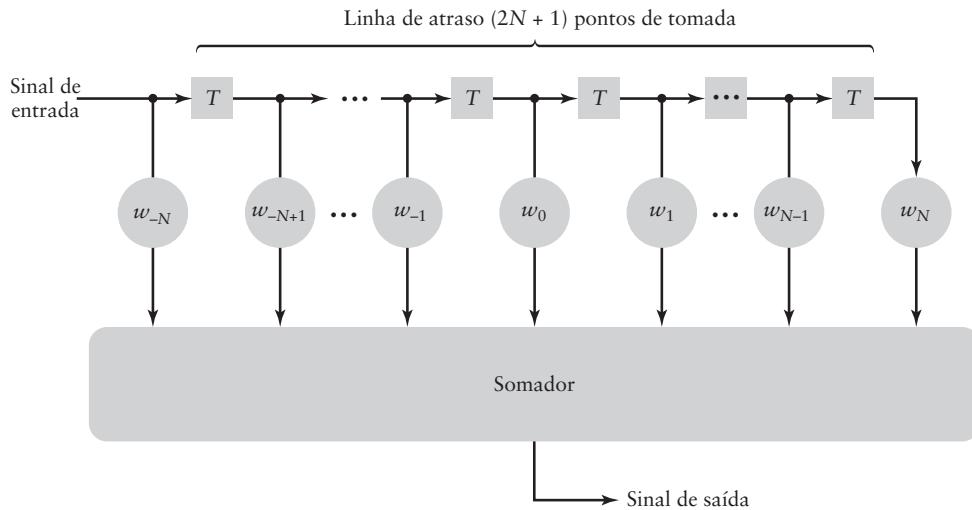


Figura 6.10 Filtro transversal.

Para simetria estrutural, com relação ao ponto central do equalizador transversal ajustável, o número total de pontos de tomada é escolhido como sendo $(2N + 1)$. Desta forma, sejam os pesos (coeficientes) do equalizador representado por $w_{-N}, \dots, w_{-1}, w_0, w_1, \dots, w_N$. Ou seja, podemos expressar a resposta ao impulso do equalizador como

$$h_{\text{eq}}(t) = \sum_{k=-N}^N w_k \delta(t - kT) \quad (6.36)$$

na qual $\delta(t)$ é a função delta de Dirac e w_k é o peso conectado ao k -ésimo ponto de tomada da linha de atraso do equalizador. De acordo com a Fig. 6.11, o equalizador transversal com resposta $h_{\text{eq}}(t)$ ao impulso é conectado em série com a combinação filtro de transmissão-canál de resposta $c(t)$ ao impulso. Seja $p(t)$ a resposta total ao impulso desta conexão em série. Podemos, então, expressar $p(t)$ como a convolução de $c(t)$ com $h_{\text{eq}}(t)$, mostrado por

$$\begin{aligned} p(t) &= c(t) \star h_{\text{eq}}(t) \\ &= c(t) \star \sum_{k=-N}^N w_k \delta(t - kT) \end{aligned} \quad (6.37)$$

na qual o símbolo \star representa a convolução. Alterando a ordem do somatório e da convolução, o que é permitido pois os dois subsistemas da Fig. 6.11 são lineares, podemos escrever

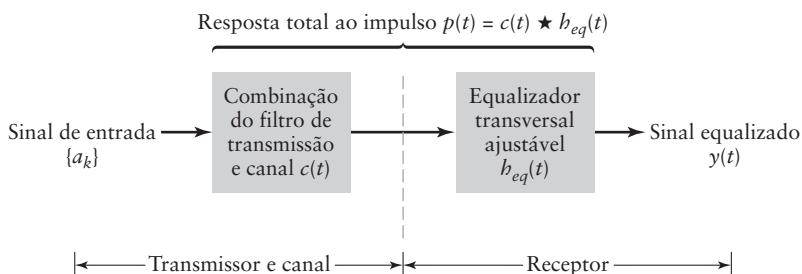


Figura 6.11 Representação simplificada do cenário para a resolução do problema de equalização de canal.

$$\begin{aligned}
p(t) &= \sum_{k=-N}^N w_k c(t) \star \delta(t - kT) \\
&= \sum_{k=-N}^N w_k c(t - kT)
\end{aligned} \tag{6.38}$$

na qual, na última linha, utilizamos a propriedade de deslocamento da função delta. Calculando a Eq. (6.38) nos instantes de amostragem $t = iT$, obtemos o *somatório discreto de convolução*.

$$p(iT) = \sum_{k=-N}^N w_k c((i - k)T) \tag{6.39}$$

Para simplificar a representação, seja $p_i = p(iT)$ e $c_i = c(iT)$. Podemos, então, reescrever a Eq. (6.39) na forma compacta

$$p_i = \sum_{k=-N}^N w_k c_{i-k} \tag{6.40}$$

A eliminação da interferência intersimbólica requer que o critério de Nyquist para a transmissão sem distorção, descrito na Eq. (6.11), seja satisfeito. Desta forma, podemos escrever

$$p_i = \begin{cases} \sqrt{E}, & \text{para } i = 0 \\ 0, & \text{para todo } i \neq 0 \end{cases}$$

na qual E é a energia por símbolo do sinal transmitido. Da Eq. (6.40) observamos que existem apenas $(2N + 1)$ pesos ajustáveis à nossa disposição. Logo, o critério de Nyquist para a transmissão sem distorção só pode ser aproximadamente satisfeito (dependendo de N), como mostrado por

$$p_i = \begin{cases} \sqrt{E}, & i = 0 \\ 0, & i = \pm 1, \pm 2, \dots, \pm N \end{cases} \tag{6.41}$$

Logo, impondo a condição da Eq.(6.39) no somatório discreto de convolução da Eq. (6.40), obtemos um sistema de $(2N + 1)$ equações simultâneas:

$$\sum_{k=-N}^N w_k c_{i-k} = \begin{cases} \sqrt{E}, & i = 0 \\ 0, & i = \pm 1, \pm 2, \dots, \pm N \end{cases} \tag{6.42}$$

De forma equivalente, na notação matricial, podemos escrever

$$\begin{bmatrix} c_0 & \cdots & c_{-N+1} & c_{-N} & c_{-N-1} & \cdots & c_{-2N} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ c_{N-1} & \cdots & c_0 & c_{-1} & c_{-2} & \cdots & c_{-N-1} \\ c_N & \cdots & c_1 & c_0 & c_{-1} & \cdots & c_{-N} \\ c_{N+1} & \cdots & c_2 & c_1 & c_0 & \cdots & c_{-N+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ c_{2N} & \cdots & c_{N+1} & c_N & c_{N-1} & \cdots & c_0 \end{bmatrix} \begin{bmatrix} w_{-N} \\ \vdots \\ w_{-1} \\ w_0 \\ w_1 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \sqrt{E} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{6.43}$$

Um equalizador transversal descrito pela Eq. (6.42), ou de forma equivalente pela Eq. (6.43), é chamado de *equalizador de zero forçado*, pois o equalizador forma a interferência intersimbólica a ser zero em $(2N + 1)$ instantes de amostragem do sinal recebido. O equalizador de zero forçado é ótimo no sentido de que ele minimiza a *distorção de pico* na ausê-

cia de ruído. A distorção de pico é definida na Eq. (6.34). Outra interessante característica do equalizador de zero forçado é que ele é relativamente simples de ser implementado. Além disso, em teoria, o quanto maior fizermos o equalizador (isto é, permitindo que N tenda ao infinito), mais perto o sistema equalizado estará da condição ideal especificada pelo critério de Nyquist para a transmissão sem distorção em um canal. Note, entretanto, que como o equalizador de zero forçado ignora o efeito do ruído aditivo de canal, o sistema equalizado não oferecerá sempre a melhor solução ao problema de interferência intersimbólica.⁷

► **PROBLEMA 6.6** Assuma as seguintes condições perfeitas:

- A distorção residual em um sistema de transmissão de dados é zero.
- A formatação de pulso é particionada igualmente entre a combinação filtro de transmissão-canal e o receptor.
- O equalizador transversal é infinitamente longo.
- (a) Obtenha o valor correspondente da função de transferência do equalizador em termos do espectro total do pulso $P(f)$.
- (b) Para um fator de roll-off $\alpha = 1$, demonstre que um equalizador de tamanho infinito pode, essencialmente, satisfazer a condição perfeita obtida na parte (a) deste problema. ◀

■ **Como o receptor pode determinar o $\{c_k\}$?**

Dados os coeficientes $c_{-N}, \dots, c_{-1}, c_0, c_1, \dots, c_N$, definindo a resposta ao impulso amostrada da combinação filtro de transmissão-canal, podemos, então, utilizar o sistema simultâneo de equações (6.43) para obter os $(2N + 1)$ pesos correspondentes do equalizador transversal.⁸ Este cálculo, entretanto, presume que o receptor já tem o conhecimento do conjunto de coeficientes $\{c_k\}_{k=-N}^N$. Mas como o receptor pode saber isso?

Um método geralmente utilizado para esta questão fundamental é a utilização de uma sessão *piloto assistida* que funciona como descrito a seguir:

1. Para a seqüência binária de dados $\{b_k\}$ aplicada à entrada do transmissor, utilize uma seqüência estatística de 1s e 0s que possua uma característica tipo ruído; logo, a referência para esta seqüência é uma *seqüência pseudo-ruído (PR)*. A razão para a utilização desta referência como piloto é possibilitar uma medida de unicidade.
2. A seqüência PR é conhecida *a priori* pelo receptor. Desta forma, com o receptor sincronizado com o transmissor, o receptor é capaz de saber quando iniciar a sessão de treinamento.
3. Finalmente, conhecendo a seqüência PR e medindo a saída correspondente do canal, o receptor pode estimar a seqüência $\{c_k\}$ representando a resposta amostrada ao impulso da combinação filtro de transmissão-canal.

⁷ Para um equalizador mais robusto do que o equalizador de zero forçado, olhamos para um critério de otimização diferente, chamado de *critério de erro médio quadrático*, o qual considera os efeitos combinados de distorção residual e ruído de canal. O erro médio quadrático (por exemplo, erro médio de potência) é definido como o valor esperado do erro quadrático (diferença) entre a resposta “desejada” e a resposta “real” do equalizador. A esperança é um operador estatístico, discutido no Capítulo 8.

Para uma discussão detalhada do equalizador de erro médio quadrático mínimo e sua implementação adaptativa, veja Haykin (2001). Ao contrário do equalizador fixo (isto é, equalizador transversal cujos pesos dos pontos de tomada sejam mantidos fixos uma vez que os ajustes individuais tenham sido calculados), um *equalizador adaptativo* é equipado com um mecanismo que ajusta continuamente os pesos dos pontos de tomada do equalizador, de forma a compensar a natureza variante no tempo dos canais de telefonia e sem fio, o que constitui uma outra realidade prática de comunicações digitais.

⁸ A solução da Eq. (6.43), ou seja, os $(2N + 1)$ pesos do equalizador transversal, é definida pela inversa da matriz $(2N + 1) \times (2N + 1)$ de coeficientes conhecidos do lado esquerdo da equação multiplicado pelo vetor coluna altamente esparsa do lado direito da equação. Para N grande, o cálculo da matriz inversa se torna problemático, e, neste caso, precisamos utilizar outra abordagem. Para detalhes, veja Haykin (2001).

6.9 RESUMO E DISCUSSÃO

O assunto de transmissão de dados digitais em um canal de comunicação envolve duas famílias diferentes de aplicações, dependendo do tipo de canal considerado:

- (i) *Transmissão banda base*, na qual o canal é do tipo passa-baixa.
- (ii) *Transmissão de dados passa-faixa*, na qual o canal é do tipo passa-faixa (a transmissão passa-faixa é chamada, também, de transmissão passa banda).

Neste capítulo, estudamos a transmissão passa-banda. O estudo da transmissão de dados passa-faixa é considerado no Capítulo 7.

Em particular, este capítulo focalizou no problema de interferência intersimbólica, o qual aparece devido a imperfeições na resposta em freqüência do canal, considerado como sendo linear. Com a atenção direcionada para um pulso de sinal de interesse na saída do canal, a interferência intersimbólica (IIS) se refere ao efeito no pulso devido ao *cross-talk* ou *spillover* (espalhamento) de todos os outros pulsos de sinal em uma seqüência de dados aplicada à entrada do canal.

Uma medida corretiva amplamente utilizada na prática é *formatar* o espectro geral do pulso do sistema banda base, começando da fonte do sinal de mensagem, por todo o caminho até o receptor. A solução ótima para o espectro do pulso é fornecida pelo canal de Nyquist, o qual garante interferência intersimbólica nula a uma taxa de bit igual a duas vezes a largura de faixa do canal. Entretanto, esta solução ótima para o problema de interferência intersimbólica é irrealizável, devido à sua característica retangular. Para contornar a questão de impossibilidade de realização, podemos utilizar o espectro do pulso de cosseno levantado, o qual fornece flexibilidade de projeto através do *fator de roll-off* que varia entre zero e um. Note, entretanto, que a utilização de um fator de roll-off não nulo é uma condição necessária mas não suficiente para a realização física da solução do problema de interferência intersimbólica nula. Além disso, o espectro do pulso de cosseno levantado deve ser associado com uma característica de fase linear cuja inclinação depende do fator de roll-off, como discutido no Problema 6.4. Também é válido mencionar que uma solução mais sofisticada para o problema de interferência intersimbólica é fornecida pelo particionamento igualitário da tarefa total de formação de pulso entre o transmissor e o receptor, utilizando o espectro do pulso de raiz de cosseno levantado.

Neste capítulo, também estudamos a questão prática relacionada com a interferência intersimbólica: como calculá-la experimentalmente. O padrão de olho é uma ferramenta intuitiva utilizada para isto. De fato, em um único gráfico, o padrão de olho mostra os efeitos de degradação do jitter, interferência intersimbólica (devido a imperfeições do sistema) e ruído do canal, todos de natureza aleatória. A interferência intersimbólica é um *fenômeno dependente do sinal*. Desaparecendo, portanto, quando o sinal contendo a informação é desligado. Por outro lado, o ruído está sempre lá, independente se existir a transmissão de dados ou não.

Um último comentário deve ser feito. Além da formatação de pulso, existe outra medida corretiva para lidar com o problema de interferência intersimbólica, chamada de equalização de canal. Esta segunda medida corretiva envolve a utilização de um filtro transversal com coeficientes ajustáveis. O filtro é colocado no receptor e é ajustado de maneira a compensar a distorção residual resultante das imperfeições da resposta ao impulso do canal.

Problemas adicionais

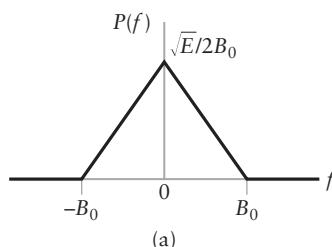
- 6.7 Começando da Eq. (6.17) para o espectro $P(f)$ do pulso de cosseno levantado, utilize a transformada inversa de Fourier para obter a resposta $p(t)$ correspondente, no tempo, definida na Eq. (6.19).
- 6.8 O espectro do pulso de cosseno levantado para um fator unitário de roll-off é dado por

$$P(f) = \begin{cases} \frac{\sqrt{E}}{2B_0} \cos^2\left(\frac{\pi f}{4B_0}\right), & 0 \leq |f| < 2B_0 \\ 0, & 2B_0 \leq |f| \end{cases}$$

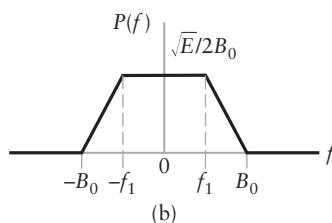
a qual é um caso especial da Eq. (6.17) para $\alpha = 1$. Mostre que a resposta $p(t)$ no tempo, a transformada inversa de Fourier de $P(f)$, é

$$p(t) = \sqrt{E} \frac{\operatorname{sinc}(4B_0 t)}{1 - 16B_0^2 t^2}$$

- 6.9 Um computador disponibiliza em sua saída um dado binário na taxa de 56 kilobits por segundo. A saída do computador é transmitida utilizando um sistema PAM binário banda base projetado para ter o espectro do pulso de cosseno levantado. Determine a largura de faixa necessária para cada um dos seguintes fatores de roll-off:
- (a) $\alpha = 0,25$
 - (b) $\alpha = 0,5$
 - (c) $\alpha = 0,75$
 - (d) $\alpha = 1,0$
- 6.10 Uma onda PAM binária deve ser transmitida em um canal passa-baixa com largura de faixa de 75 kHz. A duração do bit é de 10 μs . Obtenha o espectro de pulso de cosseno levantado que satisfaz estas condições.
- 6.11 Considere um canal, cuja largura de faixa é 3,0 kHz, disponível para a transmissão de dados usando PAM binário. Trace a taxa $1/T_b$ de bit possível (sinalização) em função do excesso de largura de faixa f_e , assumindo que o fator α de roll-off varia de zero à unidade e que o critério de interferência intersimbólica nula seja satisfeito.
- 6.12 Você possui um canal cuja largura de faixa é 3,0 kHz. A necessidade é transmitir dados no canal a uma taxa de 4,5 kilobits/s, utilizando PAM binário.
- (a) Qual é o máximo fator de roll-off no espectro de pulso de cosseno levantado que possibilita esta transmissão de dados?
 - (b) Qual é o excesso de largura de faixa correspondente?
- 6.13 Este problema surge do critério de formatação de pulso para interferência intersimbólica nula, descrito pela Eq. (6.30). Este critério pode ser satisfeito por um número infinito de espectro total de pulso representado por $P(f)$. O espectro retangular da Fig. 6.2(a) e o espectro do pulso de cosseno levantado da Fig. 6.4(a) são dois destes exemplos. O espectro de pulso mostrado nas



(a)



(b)

Figura 6.12 Problema 6.13.

partes (a) e (b) da Fig. 6.12 são dois outros exemplos que também podem satisfazer o critério de formatação de pulso da Eq. (6.30).

(a) Obtenha a condição que a largura de faixa B_0 da Fig. 6.12(a) deve satisfazer para que a condição de interferência intersimbólica nula seja satisfeita para PAM binário.

(b) Repita o problema para o espectro do pulso da Fig. 6.12(b).

(c) Dado os quatro espectro de pulso, os dois das Figs. 6.2(b) e 6.3(a) e os da Fig. 6.12, por que, então, o espectro do pulso de cosseno levantado da Fig. 6.3(a) é a escolha preferida na prática? Justifique sua resposta.

6.14 Repita o Problema 6.12, dado que cada conjunto de três dígitos binários sucessivos da saída de um computador é codificado em oito possíveis níveis de amplitude, e que o sinal resultante é transmitido utilizando um sistema PAM de 8 níveis projetado para ter o espectro do pulso de cosseno levantado.

6.15 Um sinal analógico é amostrado, quantizado e codificado em uma onda PCM binária. O número de níveis de representação utilizados é 128. Um pulso de sincronização é adicionado ao final de cada palavra de código. O sinal PCM resultante é transmitido em um canal com largura de faixa de 13 kHz utilizando um sistema PAM *quaternário* com espectro de pulso de cosseno levantado. O fator de roll-off é unitário.

(a) Determine a taxa (em bits por segundo) na qual a informação é transmitida através do canal.

(b) Determine a taxa na qual o sinal analógico é amostrado. Qual é o maior valor possível para a maior componente de frequência do sinal analógico?

6.16 Uma onda binária utilizando sinalização sem retorno para zero é gerada pela representação do símbolo 1 por um pulso de amplitude +1 e o símbolo 0 por um pulso de amplitude -1. Nos dois casos a duração do pulso é igual à duração do bit. Este sinal é aplicado a um filtro RC passa-baixa com função de transferência:

$$H(f) = \frac{1}{1 + jf/f_0}$$

Construa o padrão de olho da saída do filtro para as seguintes seqüências:

(a) 1s e 0s alternando.

(b) Uma longa seqüência de 1s seguida por uma longa seqüência de 0s.

(c) Uma longa seqüência de 1s seguida por um único 0 e, então, uma longa seqüência de 1s.

Assuma uma taxa de bit de $2B_0$ por segundo.

6.17 A seqüência binária 011010 é transmitida através de um canal contendo o espectro de pulso de cosseno levantado com fator de roll-off unitário. Assuma a utilização de sinalização sem retorno para zero, com os símbolos 1 e 0 representados por +1 e -1, respectivamente.

(a) Construa a onda recebida em escala e indique os melhores tempos de amostragem para a regeneração.

(b) Construa o padrão de olho para esta onda recebida e mostre que ele é completamente aberto.

(c) Determine os cruzamentos de zero da onda recebida.

6.18 A resposta ao impulso amostrada de um sistema de transmissão de dados (composto pelo filtro de transmissão e o canal) é definida por

$$c_n = \{0,0; 0,15; 0,68; -0,22; 0,08\}$$

Para a equalização de zero forçado do sistema, propõe-se a utilização de um filtro transversal com três pontos de tomada.

(a) Calcule os pesos ajustáveis do equalizador.

(b) Utilizando o equalizador determinado na parte (a), calcule a interferência intersimbólica residual na saída do equalizador.

(c) Identifique a amplitude da amostra que faz a maior contribuição à interferência intersimbólica residual.

- 6.19 Repita o Problema 6.18, mas desta vez com um filtro transversal com cinco pontos de tomada para a equalização de zero forçado do sistema. Compare a interferência intersimbólica residual na saída do equalizador com o Problema 6.18 e comente sobre os benefícios de se utilizar um filtro transversal maior para o processo de equalização.

Problemas avançados

O tratamento da interferência intersimbólica apresentado durante grande parte deste capítulo tratou a interferência intersimbólica como um fenômeno indesejado. Pela adição controlada de interferência intersimbólica ao sinal transmitido, é possível obter uma taxa de sinalização de $2B_0$ símbolos por segundo em um canal com largura de faixa de B_0 hertz. Tal abordagem é chamada de *codificação correlativa* ou *sinalização por resposta parcial*. Os quatro problemas a seguir abordam questões relacionadas com a codificação correlativa, resumidos por:

- ▶ Os Problemas 6.20 e 6.21 pertencem à sinalização duobinária. Em particular, o Problema 6.20 aborda o filtro de conversão duobinária, cuja utilização pode levar à propagação de erros. Para resolver este problema, utilizados o pré-codificador abordado no Problema 6.21.
- ▶ O Problema 6.22 analisa as modificações nos esquemas descritos nos Problemas 6.20 e 6.21.
- ▶ Finalmente, o Problema 6.23 aborda a utilização da sinalização duobinária modificada para a modulação por faixa lateral única aplicada à transmissão de dados.

- 6.20 A Fig. 6.13 mostra o esquema de sinalização duobinária em sua forma mais simples. A seqüência binária de entrada $\{b_k\}$ consiste de símbolos binários não correlacionados, cada um com duração T_b . Esta seqüência é primeiramente aplicada a um modulador por amplitude de pulso para produzir uma seqüência $\{a_k\}$ de dois níveis, constituído de pulsos (curtos o suficiente para serem vistos como impulsos unitários). Especificamente,

$$a_k = \begin{cases} +1, & \text{se símbolo } b_k \text{ é 1} \\ -1, & \text{se símbolo } b_k \text{ é 0} \end{cases}$$

A seqüência $\{a_k\}$ de dois níveis é, então, aplicada ao filtro de conversão duobinária dentro do retângulo tracejado da Fig. 6.13, no qual o canal de Nyquist $H_{\text{oti}}(f)$ é definido por (veja a Eq. (6.15))

$$H_{\text{oti}}(f) = \begin{cases} \frac{\sqrt{E}}{2B_0}, & \text{para } -B_0 < f < B_0 \\ 0, & \text{caso contrário} \end{cases}$$

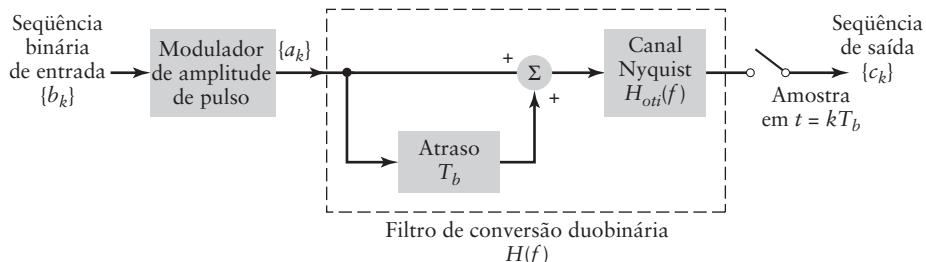


Figura 6.13 Problema 6.20.

No receptor, a saída do canal é amostrada a cada T_b segundos de forma sincronizada com o transmissor. A amostra produzida é representada por c_k . Uma estimativa \hat{a}_k do símbolo original a_k é produzida utilizando a regra

$$\hat{a}_k = c_k - \hat{a}_{k-1}$$

na qual \hat{a}_{k-1} é a estimativa anterior.

- (a) Determine a resposta total em freqüência $H(f)$ do filtro de conversão duobinária.

- (b) Determine a resposta $h(t)$ ao impulso deste filtro.
- (c) A regra descrita anteriormente utilizada no receptor é um exemplo de *realimentação de decisão*, a qual pode ser vista com o inverso do filtro de linha de atraso no transmissor. A principal desvantagem desta regra de decisão é que uma vez que erros tenham sido cometidos, eles tendem a *propagar* através da saída. Por quê? Justifique sua resposta.

- 6.21 Para diminuir o *fenômeno de propagação de erro* discutido na parte (c) do Problema 6.20, podemos utilizar um *pré-codificador*, como mostrado na Fig. 6.14. A pré-codificação é aplicada à seqüência binária de entrada $\{b_k\}$, produzindo uma nova seqüência binária $\{d_k\}$, definida por

$$d_k = b_k \oplus d_{k-1}$$

na qual o símbolo \oplus representa a *adição módulo dois*. Esta adição é equivalente à operação OU-EXCLUSIVO, a qual funciona como descrito a seguir. A saída d_k é o símbolo 1 se as entradas b_k ou d_{k-1} forem diferentes uma da outra, caso contrário, a saída d_k é o símbolo 0.

Como mostrado na Fig. 6.14, a seqüência binária pré-codificada $\{d_k\}$ é aplicada ao modulador por amplitude de pulso, seguido pelo filtro de conversão duobinária. Estes dois componentes seguem exatamente as mesmas descrições apresentadas no Problema 6.20. Desta vez, entretanto, a regra de decisão utilizada pelo receptor para a detecção da seqüência binária original $\{b_k\}$ a partir da seqüência de três níveis $\{c_k\}$ é definida por

Se $|c_k| < 1$, diga que o símbolo binário b_k é 1

Se $|c_k| > 1$, diga que o símbolo binário b_k é 0

De acordo com esta decisão, o detector consiste de um retificador seguido por um dispositivo de limiar. Mais importante, se qualquer c_k for recebido com erro, o erro é confinado ao instante.

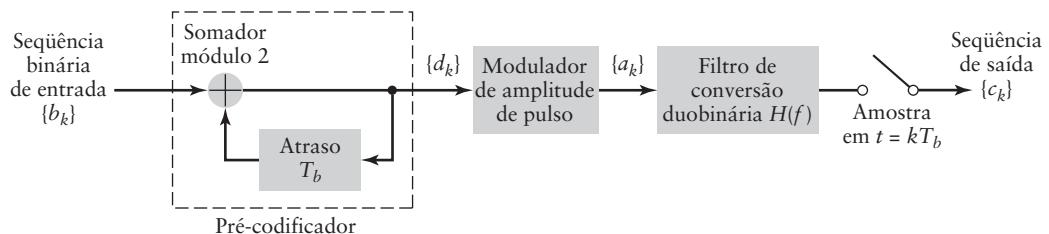


Figura 6.14 Problema 6.21.

Para ilustrar a operação do esquema de sinalização duobinária com pré-codificação, considere a seqüência de dados de entrada

$$\{b_k\} = 0010110$$

Para seguirmos com a pré-codificação desta seqüência, adicione um bit extra na saída do pré-codificador. Este bit extra é arbitrariamente escolhido como sendo 1. Logo, faça o seguinte:

- (a) Determine a seqüência pré-codificada $\{d_k\}$ e, então, a seqüência $\{a_k\}$ de dois níveis produzida pelo modulador de amplitude de pulso.
- (b) Determine a saída do código duobinário $\{c_k\}$.
- (c) Aplicando $\{c_k\}$ ao detector baseado no retificador, descrito anteriormente, determine a seqüência binária resultante. Portanto, demonstre a reconstrução correta da seqüência original.

- 6.22 Neste problema, iremos considerar outro esquema correlativo chamado de *sinalização duobinária modificada*, a qual é mostrada na Fig. 6.15. A estratégia de codificação correlativa é, agora, definida por

$$c_k = a_k - a_{k-2}$$

a qual envolve a correlação de dois símbolos binários atrasados.

- (a) Determine a resposta em freqüência $H(f)$ do filtro de conversão duobinária modificada dentro do segundo retângulo tracejado da Fig. 6.15.

- (b) Determine a resposta $h(t)$ ao impulso deste filtro e mostre que ele possui três níveis distintos nos instantes de amostragem.
 (c) Para a decodificação no receptor, demonstre a utilização com sucesso da regra de decisão:

$$\hat{b}_k = \begin{cases} \text{símbolo 1,} & \text{se } |c_k| > 1 \\ \text{símbolo 0,} & \text{se } |c_k| < 1 \end{cases}$$

Para a demonstração, utilize, novamente, a seqüência de entrada 0010110.

- (d) À luz de nossas descobertas nas partes (a) a (c), discuta as vantagens do codificador duobinário modificado com relação ao codificador duobinário discutido no Problema 6.21.

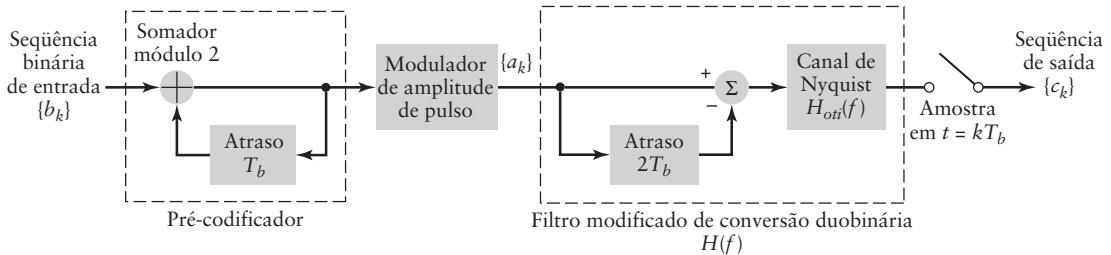


Figura 6.15 Problema 6.22.

- 6.23 O esquema de sinalização duobinária modificada apresentada no Problema 6.22 é adequada para a modulação por faixa lateral única aplicada à transmissão de dados binários em um canal linear. Justifique a aplicação prática desta afirmativa.

TÉCNICAS DE MODULAÇÃO PASSA-FAIXA

Na transmissão de dados banda base, estudada no Capítulo 6, toda seqüência serial de dados de entrada é representada na forma de uma onda discreta modulada por amplitude de pulso que pode ser transmitida em um canal passa-baixa (por exemplo, um cabo co-axial). Mas e se a condição for transmitir a seqüência de dados em um canal passa-faixa, exemplificado por um canal de comunicação sem fio ou por satélite? Em aplicações deste tipo, geralmente optamos por trabalhar com uma *estratégia de modulação configurada ao redor de uma portadora senoidal* cuja amplitude, fase ou freqüência é variada de acordo com a seqüência de dados contendo a informação. As técnicas de modulação digital que trabalham com transmissão de dados passa-faixa são estudadas neste capítulo.

O objetivo primário deste capítulo é descrever algumas técnicas de modulação digital passa-faixa utilizadas na prática. Em particular, descrevemos três esquemas básicos de modulação, chamados de chaveamento de amplitude, chaveamento de fase e chaveamento de freqüência, seguido por algumas de suas variantes. Outra questão que receberá atenção particular é a detecção coerente *versus* não coerente. Um sistema de comunicação digital é dito ser *coerente* se o receptor for sincronizado com o transmissor, com respeito à fase da portadora. Caso contrário, o sistema é chamado de *não coerente*. Naturalmente, um sistema não coerente oferece a vantagem prática de complexidade reduzida, mas ao custo de performance degradada. Considerações sobre a questão de performance frente ao ruído são deixadas para o Capítulo 10.

Este capítulo irá nos ensinar três lições:

- ▶ *Lição 1: Cada esquema de modulação passa-baixa é definido por um sinal transmitido com uma representação fasorial única.*
- ▶ *Lição 2: No lado do receptor, técnicas de demodulação são implementadas de forma diferente, dependendo se o receptor é coerente ou não coerente.*
- ▶ *Lição 3: Duas formas de se classificar um esquema de modulação digital são (a) pelo tipo de modulação utilizada e (b) se a seqüência de dados transmitida está na forma bimária ou M-ária.*

7.1 ALGUMAS PRELIMINARES

Dada uma fonte binária que emite os símbolos 0 e 1, o processo de modulação envolve o chaveamento da amplitude, fase ou freqüência de uma onda senoidal dentre um par de possíveis valores de acordo com os símbolos 0 e 1. Para sermos mais específicos, considere a portadora senoidal

$$c(t) = A_c \cos(2\pi f_c t + \phi_c) \quad (7.1)$$

na qual A_c é a amplitude da portadora, f_c é a freqüência da portadora e ϕ_c é a fase da portadora. Dados os três parâmetros da portadora $c(t)$, podemos, agora, identificar três formas distintas de modulação binária:

1. *Chaveamento binário de amplitude* (BASK, *binary amplitude shift-keying*), na qual a freqüência da portadora e a fase da portadora são mantidas constantes enquanto que a amplitude da portadora é chaveada em dois possíveis valores utilizados para representar os símbolos 0 e 1.

2. *Chaveamento binário de fase* (BPSK, *binary phase-shift keying*), na qual a amplitude da portadora e a freqüência da portadora são mantidas constantes enquanto que a fase da portadora é chaveada entre dois possíveis valores (por exemplo, 0° e 180°) para representar os símbolos 0 e 1.
3. *Chaveamento binário de freqüência* (BFSK, *binary frequency-shift keying*), na qual a amplitude da portadora e a fase da portadora são mantidas constantes enquanto que a freqüência da portadora é chaveada entre dois possíveis valores utilizados para representar os símbolos 0 e 1.

À luz destas definições, vemos que BASK, BPSK e BFSK são casos especiais de modulação em amplitude, modulação em fase e modulação em freqüência, respectivamente. De fato, foi devido a esta relação entre as técnicas de modulação analógicas e digitais que nas seções de “Resumo e Discussão” do Capítulo 3 sobre modulação em amplitude e no Capítulo 4 sobre modulação em ângulo que rapidamente mostramos as conexões entre os esquemas de modulação analógica e digital. Uma conclusão importante obtida desta relação entre técnicas de modulação analógica e digital é que, apesar de suas diferenças básicas, BASK, BPSK e BFSK compartilham uma característica comum: todos os três são exemplos de um *processo passa-faixa*.

Na literatura de comunicações analógicas, a portadora senoidal $c(t)$ é geralmente definida como na Eq. (7.1). Por outro lado, na literatura sobre comunicações digitais, a prática usual é assumir que a portadora $c(t)$ possui *energia unitária medida durante a duração de um símbolo (bit)*. Especificamente, do Problema 3.28, lembramos que a amplitude da portadora

$$A_c = \sqrt{\frac{2}{T_b}} \quad (7.2)$$

na qual T_b é a *duração do bit*. Utilizando a terminologia da Eq. (7.2), podemos, pois, expressar a portadora $c(t)$ na forma equivalente

$$c(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_c t + \phi_c) \quad (7.3)$$

A partir do material apresentado no Capítulo 2 sobre a representação de Fourier de sinais e sistemas, aprendemos que a diminuição da duração de um pulso retangular possui o efeito de alargar a faixa efetiva de freqüência contidas no pulso. De maneira correspondente, a diminuição da duração T_b do bit possui o efeito de aumentar a necessidade de largura de faixa de transmissão para uma onda modulada binária.

Uma outra lição aprendida do material apresentado nos capítulos anteriores – Capítulos 3 e 4 – é o fato da necessidade de largura de faixa de transmissão de uma onda modulada em ângulo ser maior do que a necessidade de largura de faixa de transmissão de uma onda modulada em amplitude. Com base nesta lição, podemos dizer que a necessidade de largura de faixa do BFSK é maior do que a do BASK para uma dada fonte binária. Entretanto, o mesmo *não* é válido para BPSK, como iremos ver no material apresentado neste capítulo. Esta é uma das várias diferenças que distinguem a modulação digital da modulação analógica.

■ Considerações passa-faixa

O espectro de uma onda modulada digitalmente, exemplificada pela BASK, BPSK e BFSK é centrada na freqüência f_c da portadora, implicitamente ou explicitamente. Além disso, tal como na modulação analógica, é prática normal considerar que a freqüência f_c da portadora é grande quando comparada com a “largura de faixa” da seqüência binária de dados de entrada que atua como sinal modulante. Esta consideração passa-faixa possui certas implicações que serão discutidas a seguir.

Para sermos específicos, considere um esquema de modulação linear para o qual a onda modulada é definida por

$$s(t) = b(t)c(t) \quad (7.4)$$

na qual $b(t)$ representa uma onda binária de entrada. Então, ajustando a fase da portadora $\phi_c = 0$ por conveniência de representação, podemos utilizar a Eq. (7.3) para expressar a onda modulada $s(t)$ por

$$s(t) = \sqrt{\frac{2}{T_b}} b(t) \cos(2\pi f_c t) \quad (7.5)$$

Com a consideração de $f_c \gg W$, na qual W é a largura de faixa da onda binária $b(t)$, não existirá sobreposição espectral na geração de $s(t)$ (isto é, o conteúdo espectral da onda modulada para freqüências positivas é essencialmente separada de seu conteúdo espectral para freqüências negativas).

Outra implicação da consideração passa-faixa é que podemos expressar a *energia por bit do sinal transmitido* como

$$\begin{aligned} E_b &= \int_0^{T_b} |s(t)|^2 dt \\ &= \frac{2}{T_b} \int_0^{T_b} |b(t)|^2 \cos^2(2\pi f_c t) dt \end{aligned} \quad (7.6)$$

Utilizando a identidade trigonométrica

$$\cos^2(2\pi f_c t) = \frac{1}{2}[1 + \cos(4\pi f_c t)]$$

podemos reescrever a Eq. (7.6) como

$$E_b = \frac{1}{T_b} \int_0^{T_b} |b(t)|^2 dt + \frac{1}{T_b} \int_0^{T_b} |b(t)|^2 \cos(4\pi f_c t) dt \quad (7.7)$$

A consideração passa-faixa implica que $|b(t)|^2$ é essencialmente constante em um ciclo completo da onda senoidal $\cos(4\pi f_c t)$, o que, por sua vez, significa que

$$\int_0^{T_b} |b(t)|^2 \cos(4\pi f_c t) dt \approx 0$$

Desta forma, podemos aproximar a Eq. (7.7) por

$$E_b \approx \frac{1}{T_b} \int_0^{T_b} |b(t)|^2 dt \quad (7.8)$$

Em palavras, para esquemas de modulação digital governados pela Eq. (7.5), a energia do sinal transmitido (por bit) é uma versão escalonada da energia da onda binária de entrada responsável por modular a portadora senoidal.

► **PROBLEMA 7.1** Com a consideração passa-faixa, mostre que

$$\int_0^{T_b} \sin(2\pi f_c t) \cos(2\pi f_c t) dt \approx 0$$

independentemente de como a duração T_b do bit é exatamente relacionada com f_c , desde que $f_c \gg 1/T_b$.

► **PROBLEMA 7.2** Mostre que a Eq. (7.8) é *invariante* com respeito à fase ϕ_c da portadora (isto é, ela é válida para todo ϕ_c).

7.2 CHAVEAMENTO BINÁRIO DE AMPLITUDE

O chaveamento binário de amplitude (BASK) é uma das primeiras formas de modulação digital utilizadas em rádio telegrafia no começo do século XX. Para descrevermos formalmente o BASK, considere a seqüência binária de dados $b(t)$ a qual é do tipo de *sinalização liga-desliga*. Ou seja, $b(t)$ é definida por

$$b(t) = \begin{cases} \sqrt{E_b}, & \text{para o símbolo binário 1} \\ 0, & \text{para o símbolo binário 0} \end{cases} \quad (7.9)$$

Então, multiplicando $b(t)$ pela onda portadora senoidal da Eq. (7.3) com fase ϕ_c ajustada para zero por conveniência de representação, temos a onda BASK

$$s(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t), & \text{para o símbolo 1} \\ 0, & \text{para o símbolo 0} \end{cases} \quad (7.10)$$

A freqüência f_c da portadora pode ter um valor arbitrário, consistente com a transmissão do sinal modulado em qualquer lugar no espectro eletromagnético de rádio, desde que ela satisfaça a consideração passa-faixa.

Quanto a duração do bit é ocupada pelo símbolo 1, a energia do sinal transmitido é E_b . Quando a duração do bit é ocupada pelo símbolo 0, a energia do sinal transmitido é zero. Com isto, podemos expressar a *energia média do sinal transmitido* como

$$E_{\text{méd.}} = \frac{E_b}{2} \quad (7.11)$$

Para esta fórmula ser válida, entretanto, os dois símbolos binários precisam ser *equi-prováveis*. Em outras palavras, em uma *longa* seqüência binária de dados, os símbolos 1 e 0 ocorrem essencialmente em números iguais na seqüência de dados. Para esta igualdade ser válida, entretanto, *não deve haver polarização* envolvida na geração da seqüência de bits, seja em favor do símbolo 1 ou do símbolo 0.

Geração e detecção de sinais ASK

Das Eqs. (7.9) e (7.10), rapidamente vemos que um sinal BASK é facilmente gerado utilizando um modulador de produto com duas entradas. Uma entrada, o sinal liga-desliga da Eq. (7.9), é o sinal modulante. A onda portadora senoidal

$$c(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_c t)$$

supre a outra entrada.

Uma propriedade da BASK imediatamente aparente da Fig. 7.1(b), a qual mostra a forma de onda BASK correspondente à seqüência binária de dados de entrada da Fig. 7.1(a), é a *não constância do envelope* da onda modulada. Desta forma, se considerarmos a detecção da onda BASK, a forma mais simples é utilizar um detector de envelope, explorando a propriedade de não constância do envelope do sinal BASK.

Experimento de computador I: análise espectral de BASK

Considere uma seqüência binária de dados constituída por uma *onda quadrada*, cuja amplitude alterna entre níveis constantes $\sqrt{E_b}$ e zero a cada T_b segundos. A onda quadrada é centrada na origem, por conveniência de apresentação. O objetivo deste experimento é:

- (i) Investigar o efeito da variação da freqüência f_c da portadora no espectro de potência do sinal $s(t)$ BASK, assumindo que a onda quadrada é mantida fixa. Lembre-se que

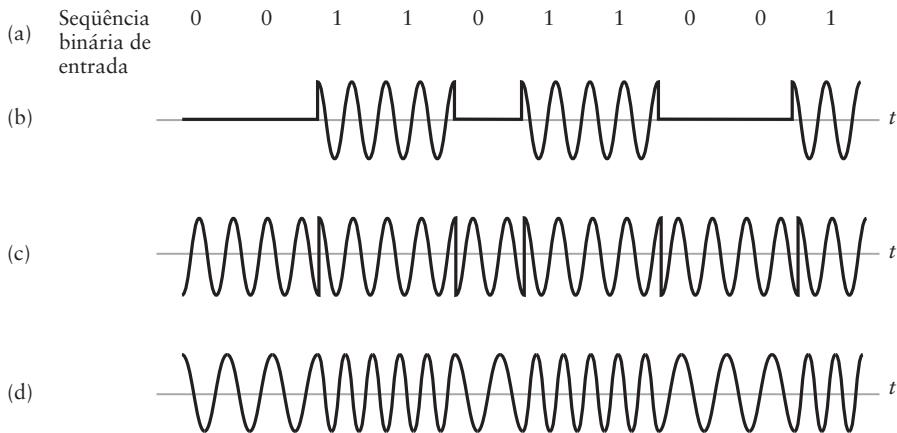


Figura 7.1 As três formas básicas de sinalização de informação binária. (a) Seqüência binária de dados. (b) Chaveamento de amplitude. (c) Chaveamento de fase. (d) Chaveamento de freqüência com fase contínua.

o espectro de potência de um sinal (expresso em decibels) é definido como 10 vezes o logaritmo (base 10) do quadrado do espectro da amplitude do sinal.

- (ii) Investigar o efeito de variação da freqüência da onda portadora no espectro do sinal BASK, assumindo que a onda senoidal é mantida constante.

Para o propósito de cálculo computacional, ajustamos a freqüência da portadora para $f_c = n/T_b$, na qual n é um inteiro. Esta escolha de freqüência f_c de portadora permite a simulação de um sistema passa-faixa em um computador digital sem a necessidade de $f_c \gg 1/T_b$. A única restrição na escolha é ter certeza de que a sobreposição espectral seja evitada (seguimos esta prática quando executamos experimentos de computador ao nos aprofundarmos no estudo de outros esquemas de modulação digital.)

Para traçar o espectro de potência (em decibels) das ondas moduladas digitalmente (neste experimento e nos outros a seguir) utilizamos o algoritmo da transformada rápida de Fourier (FFT), o qual foi discutido na Seção 2.10. Os parâmetros de simulação utilizados nos experimentos de computador (nestes e nos experimentos subsequentes) são:

$$\text{Número de bits de dados (1s e 0s)} = 100$$

$$\text{Freqüência de amostragem, } f_s = 100 \text{ Hz}$$

$$\text{Número de pontos de dados (amostras)} = 100 \times f_s = 10.000 \text{ pontos de dados}$$

$$\text{Tamanho do bloco para a FFT, } N = 4096$$

Os resultados dos cálculos são mostrados nas Figs. 7.2 e 7.3.

As duas partes da Fig. 7.2 correspondem ao objetivo (i) do experimento. Especificamente, os dois gráficos mostrados na Fig. 7.2 correspondem aos seguintes parâmetros:

$$\text{Duração do bit, } T_b = 1\text{s}$$

$$\text{Freqüência da portadora} \quad f_c = \begin{cases} 4 \text{ Hz para Fig. 7.2(a)} \\ 8 \text{ Hz para Fig. 7.2(b)} \end{cases}$$

O segundo conjunto de resultados mostrados na Fig. 7.3 corresponde ao objetivo (ii) do experimento, usando os seguintes parâmetros:

$$\text{Freqüência da portadora, } f_c = 8 \text{ Hz}$$

$$\text{Duração do bit, } T_b = \begin{cases} 1 \text{ s para Fig. 7.3(a)} \\ \frac{1}{2} \text{ s para Fig. 7.3(b)} \end{cases}$$

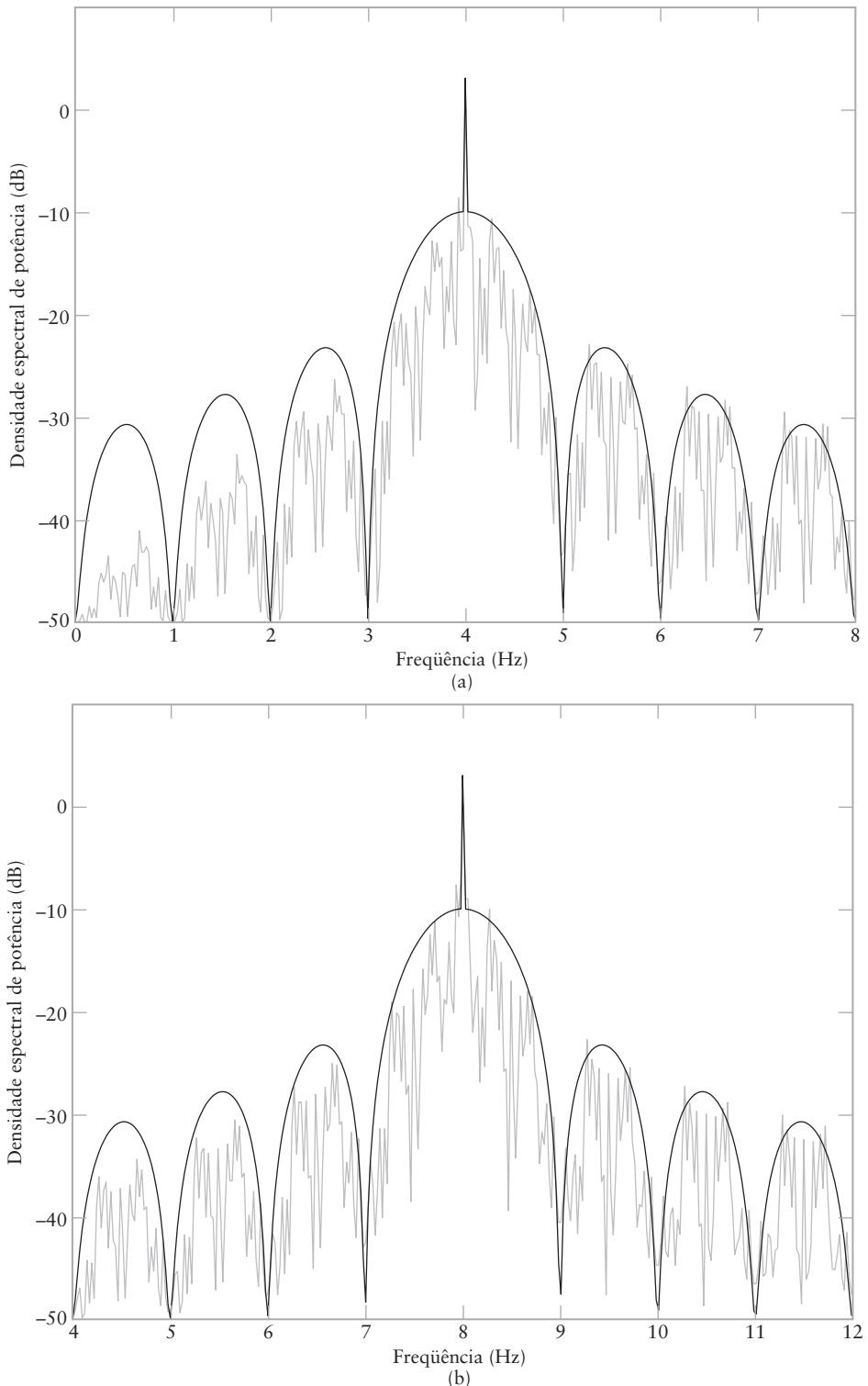


Figura 7.2 Espectro de potência do sinal BASK produzido por uma onda quadrada como sinal modulante para a variação da freqüência da portadora: (a) $f_c = 4 \text{ Hz}$ e $T_b = 1 \text{ s}$. (b) $f_c = 8 \text{ Hz}$ e $T_b = 1 \text{ s}$.

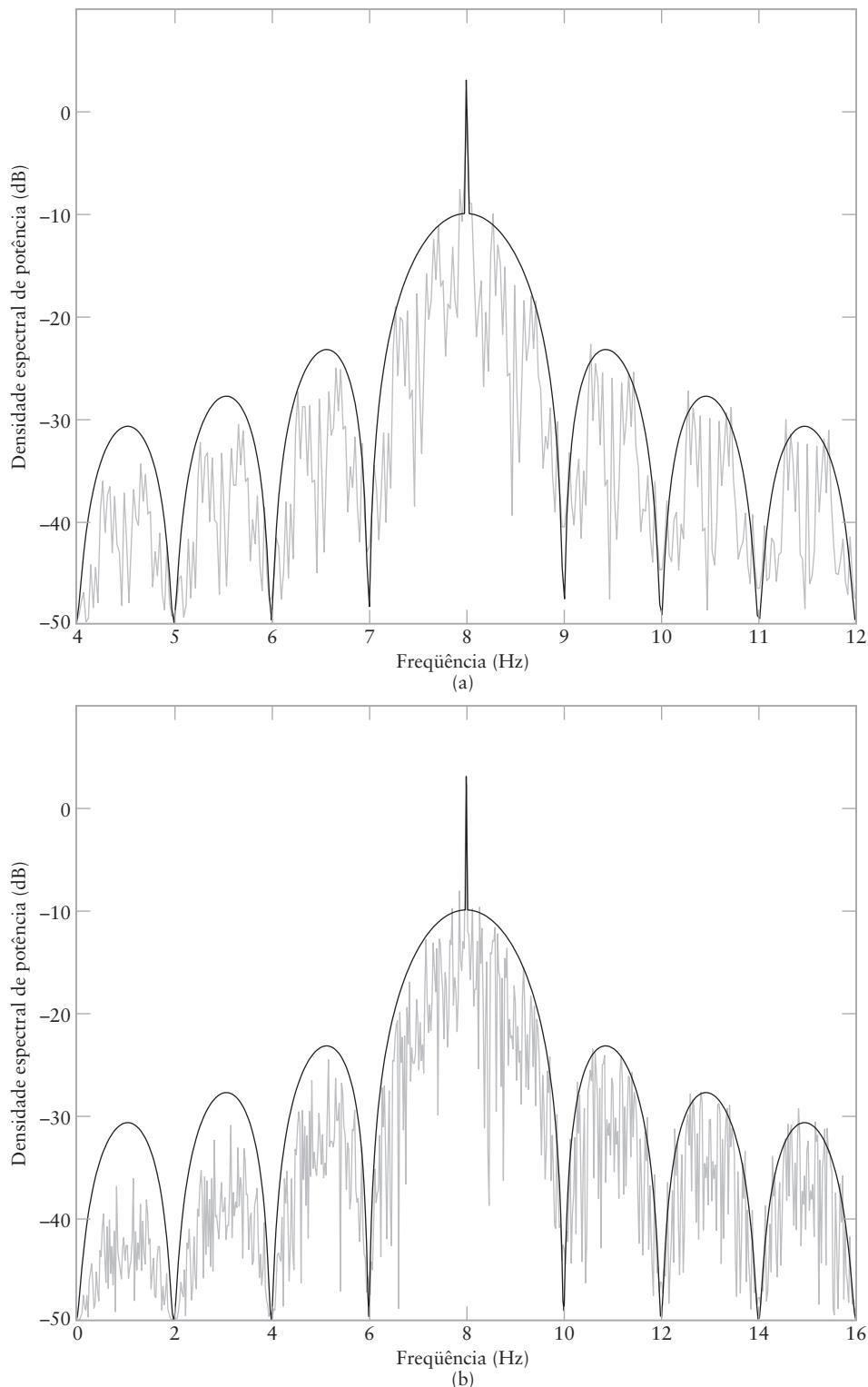


Figura 7.3 Espectro de potência do sinal BASK produzido por uma onda quadrada como sinal modulante para variação da duração do bit: (a) $f_c = 8\text{Hz}$ e $T_b = 1\text{s}$. (b) $f_c = 8\text{ Hz}$ e $T_b = 1/2\text{ s}$.

Em cada parte das Figs. 7.2 e 7.3, mostramos os dois gráficos de potência espectral: um sombreado, resultante do cálculo e o outro, uma curva bem definida, obtido da teoria. (Seguimos uma prática similar nos próximos experimentos de computador.) Devemos reconhecer que as sequências de dados binárias utilizadas no experimento possuem duração finita, por isso a forma irregular mostrada nas figuras. Em contraste, a teoria leva a uma curva sólida, a qual segue de perto o “envelope” dos resultados computacionais.

Com os resultados mostrados nas Figs. 7.2 e 7.3, podemos fazer as seguintes observações para freqüências positivas:

1. O espectro do sinal BASK contém uma componente em linha em $f = f_c$.
2. Quando a onda quadrada é mantida fixa e a freqüência da portadora é dobrada, a freqüência de meia faixa do sinal BASK é, também, dobrada.
3. Quando a portadora é mantida fixa e a duração do bit dividida pela metade, a largura do lóbulo principal da função sinc que define o envelope do espectro BASK é dobrado, o que, por sua vez, significa que a largura de faixa de transmissão do sinal BASK é dobrada.
4. A largura de faixa de transmissão do BASK, medida em termos do comprimento do lóbulo principal do espectro, é igual a $2/T_b$, na qual T_b é a duração do bit.

Estas observações suportam a afirmativa: O sinal BASK é um exemplo de modulação em amplitude que inclui a onda portadora e as componentes em sua composição.

7.3 CHAVEAMENTO DE FASE

■ Chaveamento binário de fase (BPSK)

Na forma mais simples de chaveamento de fase, chamado de *chaveamento binário de fase* (*BPSK, binary phase-shift keying*) o par de sinais $s_1(t)$ e $s_2(t)$ utilizados para representar os símbolos 1 e 0, respectivamente, são definidos por

$$s_i(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t), & \text{para o símbolo 1 correspondendo a } i = 1 \\ \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t + \pi) = -\sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t), & \text{para o símbolo 0 correspondendo a } i = 2 \end{cases} \quad (7.12)$$

na qual $0 \leq t \leq T_b$, com T_b representando a duração do bit e E_b representando a energia por bit do sinal transmitido. Veja a forma de onda da Fig. 7.1(c) para um exemplo de representação do BPSK. Um par de ondas senoidais, $s_1(t)$ e $s_2(t)$, os quais diferem apenas em deslocamento relativo de fase de π radianos, como definido pela Eq. (7.12), são chamados de *sinais antipodal*. Das duas linhas desta equação, vemos que o BPSK é, de fato, um caso especial de modulação por faixa lateral dupla-portadora suprimida (DSB-SC), um fato mostrado anteriormente na Seção 3.10.

O BPSK difere do BASK em um ponto importante: o envelope do sinal modulado $s(t)$ é mantido constante no valor $\sqrt{2E_b/T_b}$ por todo o tempo t . Esta propriedade, a qual segue diretamente da Eq. (7.12) possui duas consequências importantes:

1. A energia transmitida por bit, E_b , é constante. De forma equivalente, a potência média transmitida é constante.
2. A demodulação do BPSK não pode ser realizada usando uma detecção de envelope. Em vez dela, precisamos olhar para a detecção coerente, como descrito a seguir.

■ Geração e detecção coerente de sinais BPSK

(i) Geração

Para gerar um sinal BPSK, observamos que o sinal BPSK é um caso especial da modulação DSB-SC. Especificamente, utilizamos um modulador de produto constituído por dois componentes (veja a Fig. 7.4(a)):

- (i) *Codificador de nível sem retorno para zero*, no qual a seqüência binária de dados de entrada é codificada na forma polar com os símbolos 1 e 0 representados pelos níveis de amplitude constante: $\sqrt{E_b}$ e $-\sqrt{E_b}$, respectivamente.
- (ii) *Modulador de produto*, o qual multiplica a onda binária codificada em nível pela onda portadora senoidal $c(t)$ com amplitude $\sqrt{2/T_b}$ para produzir o sinal BPSK.

Os pulsos de temporização utilizados para gerar a onda binária codificada em nível e a onda portadora senoidal são, geralmente, mas não necessariamente, extraídos de um clock principal comum.

(ii) Detecção

Para detectar a seqüência binária original de 1s e 0s, o sinal BPSK $x(t)$ na saída do canal é aplicado a um receptor constituído de quatro seções, como mostrado na Fig. 7.4(b):

- (i) *Modulador de produto*, o qual também é suprido por um sinal de referência gerado localmente que é uma réplica da onda portadora $c(t)$.
- (ii) *Filtro passa-baixa*, projetado para remover as componentes de freqüência dobrada da saída do modulador de produto (isto é, as componentes centradas em $2f_c$) e passar as componentes de freqüência nula.
- (iii) *Amostrador*, o qual amostra uniformemente a saída do filtro passa-baixa em $t = iT_b$, na qual $t = 0, \pm 1, \pm 2, \dots$. O clock que governa a operação do amostrador é *sincronizado* com o clock responsável pela temporização do bit no transmissor.
- (iv) *Dispositivo de tomada de decisão*, o qual compara os valores amostrados da saída do filtro passa-baixa com um *limiar* fornecido externamente, a cada T_b segundos. Se o limiar for excedido, o dispositivo decide em favor do símbolo 1, caso contrário ele decide em favor do símbolo 0.

O receptor BPSK descrito na Fig. 7.4 é dito ser *coerente*, no sentido de que o sinal senoidal de referência aplicado ao modulador de produto no demodulador está em *sincronismo de fase* (e, obviamente, de freqüência) com a onda portadora utilizada no modulador.

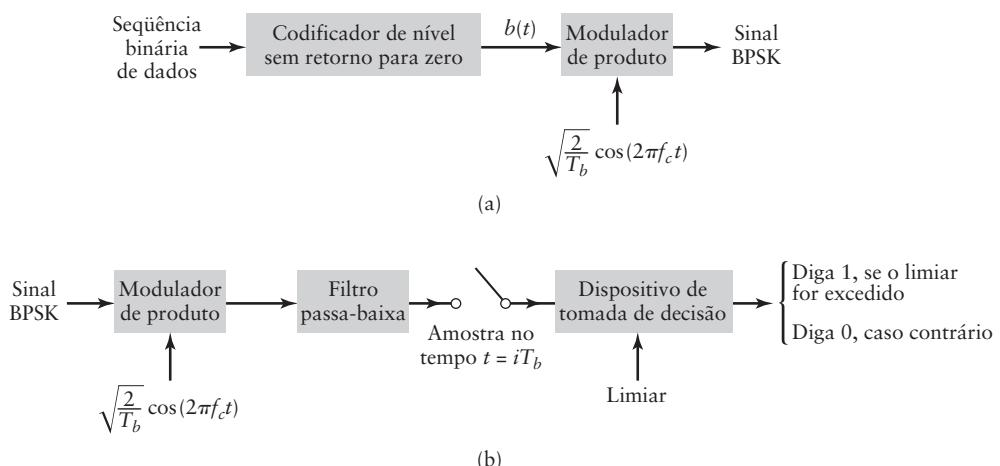


Figura 7.4 (a) Modulador BPSK. (b) Detector coerente para BPSK; para o amostrador, o inteiro $i = 0, \pm 1, \pm 2, \dots$

Esta condição pode ser atendida utilizando um PLL, o qual foi descrito na Seção 4.8. Além do sincronismo com relação a fase da portadora, o receptor também possui um conhecimento preciso do intervalo ocupado por cada símbolo binário.

A operação do receptor BPSK coerente da Fig. 7.(b) segue um procedimento similar ao descrito para a demodulação da onda modulada por faixa lateral dupla-portadora suprimida (DSB-SC), descrita na Seção 3.3, com duas importantes adições: O amostrador e o dispositivo de tomada de decisão. A razão para a similaridade está no que já foi afirmado: BPSK é simplesmente outra forma de modulação DSB-SC.

Entretanto, uma questão que precisa de atenção particular é como projetar o filtro passa-baixa da Fig. 7.4(b). Especificamente, qual deve ser a largura de faixa do filtro? Das conclusões obtidas dos resultados gráficos apresentados na Fig. 2.28 da resposta de um filtro passa-baixa ideal a um pulso retangular de entrada para um produto tempo-largura de faixa variante, lembramos que um produto tempo-largura de faixa maior ou igual a um é uma condição necessária para garantir que a forma de onda de entrada do filtro seja reconhecida na saída resultante. Para o problema em questão, podemos, portanto, afirmar que a largura de faixa do filtro passa-baixa no receptor BPSK coerente da Fig. 7.4(b) precisa ser igual ou maior do que o recíproco da duração T_b do bit para operação satisfatória do receptor.

■ Experimento de computador II: análise espectral do BPSK

Tal como no experimento sobre BASK, considere uma seqüência binária de dados que consiste de uma onda quadrada, cuja amplitude alterna entre $+\sqrt{E_b}$ e $-\sqrt{E_b}$ a cada T_b segundos. A onda quadrada é centrada na origem. Os objetivos deste segundo experimento são similares aos do Experimento de Computador I sobre BASK:

- (i) Avaliar o efeito da variação da freqüência f_c da portadora no espectro de potência do sinal BPSK, para uma onda quadrada modulante fixa.
- (ii) Avaliar o efeito da variação da freqüência de modulação no espectro de potência do sinal BPSK para uma freqüência de portadora fixa.

Para sermos consistentes com o experimento I sobre BASK, os objetivos (i) e (ii) são investigados para o mesmo conjunto de freqüências. Os resultados do experimento sobre BPSK são mostrados nas Figs. 7.5 e 7.6, nos quais, como antes, mostramos os resultados computacionais juntamente com os resultados teóricos correspondentes. Comparando estas duas figuras com as Figs. 7.2 e 7.3 para BASK, respectivamente, podemos fazer duas importantes observações:

1. Os sinais BASK e BPSK ocupam a mesma largura de faixa de transmissão, ou seja, $2/T_b$, a qual é definida pela largura do lóbulo principal do espectro de potência com forma sinc.
2. O espectro BASK inclui a componente da portadora, enquanto que esta componente é ausente do espectro BPSK. Com esta observação estamos simplesmente reafirmando o fato do BASK ser um exemplo de modulação em amplitude enquanto que o BPSK é um exemplo de modulação por faixa lateral dupla-portadora suprimida.

Esta segunda observação possui suas próprias implicações práticas:

- A presença da portadora no espectro BASK significa que a seqüência binária de dados pode ser recuperada pela detecção de envelope do sinal BASK.
- Por outro lado, a supressão da portadora no espectro BPSK resulta na utilização da detecção coerente para a recuperação da seqüência binária de dados do sinal BPSK, como discutido anteriormente.

■ Chaveamento em quadratura de fase

Um importante objetivo de comunicações digitais é a *utilização eficiente da largura de faixa do canal*. Este objetivo é atingido pelo esquema de modulação que conserva a largura de

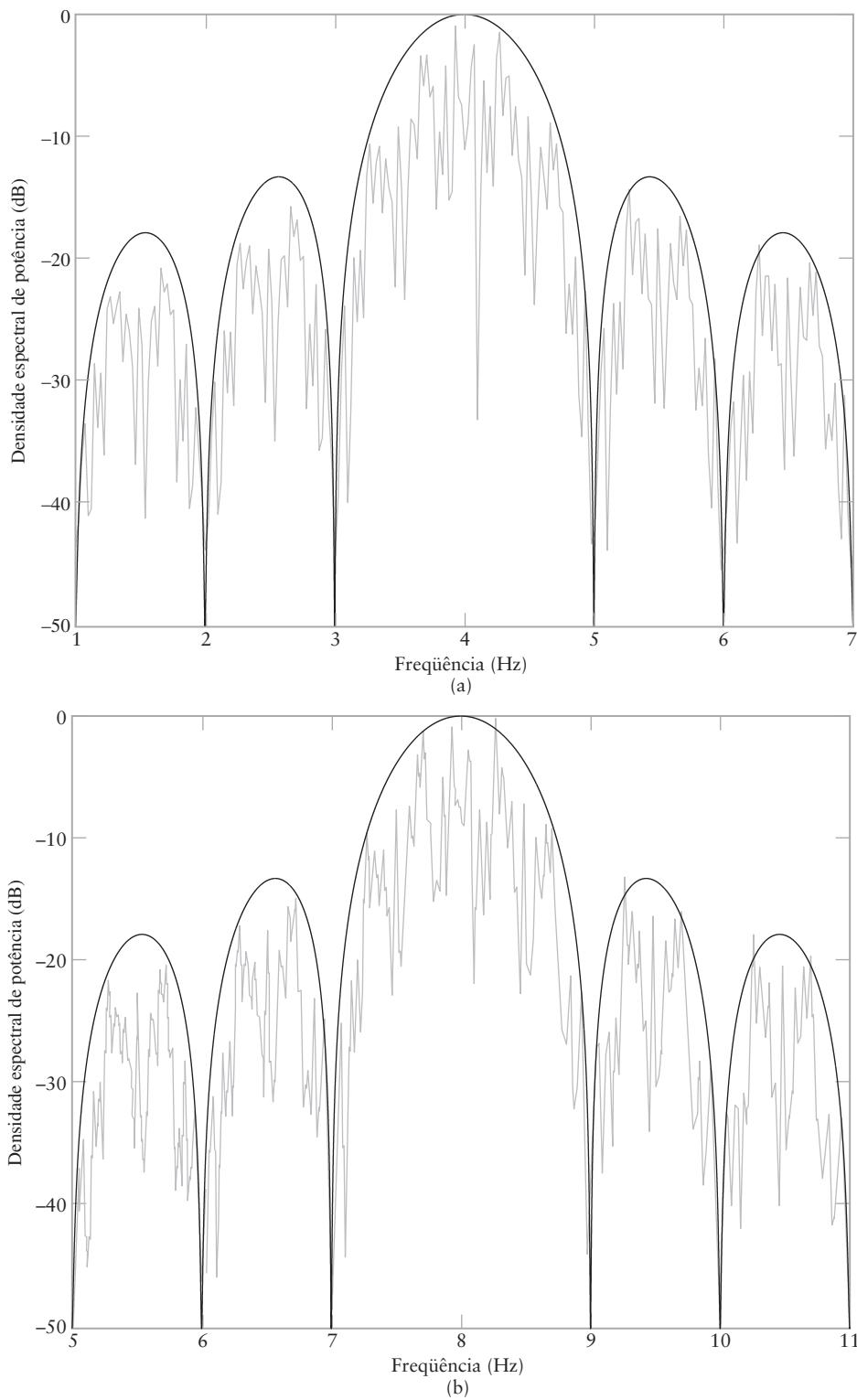


Figura 7.5 Espectro de potência do sinal BPSK produzido por uma onda quadrada como sinal modulante para a variação da freqüência da portadora: (a) $f_c = 4$ Hz e $T_b = 1$ s. (b) $f_c = 8$ Hz e $T_b = 1$ s.

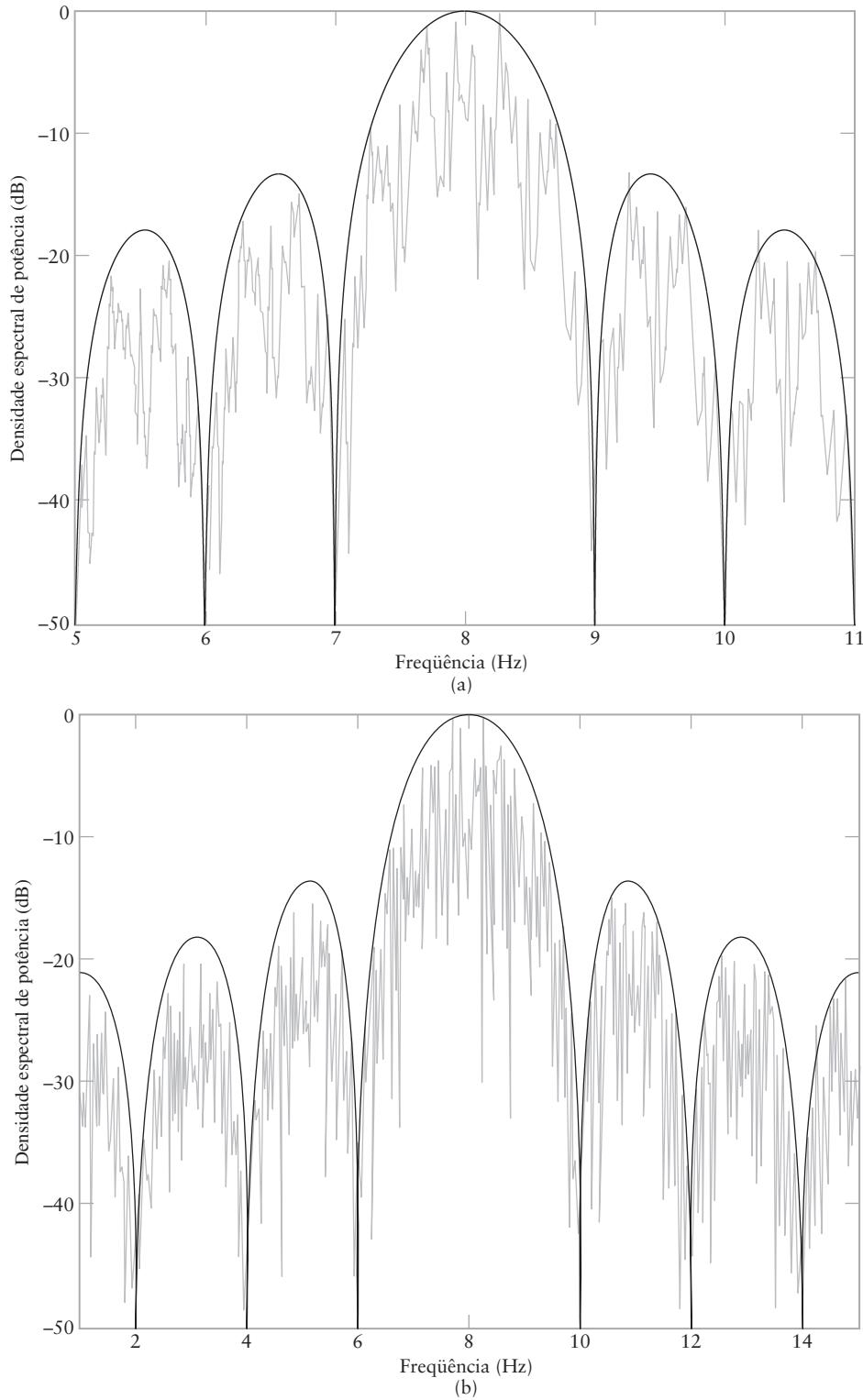


Figura 7.6 Espectro de potência do sinal BPSK produzido por uma onda quadrada como sinal modulante para variação da duração do bit: (a) $f_c = 8\text{Hz}$ e $T_b = 1\text{s}$. (b) $f_c = 8\text{ Hz}$ e $T_b = 1/2\text{ s}$.

faixa chamado de *chaveamento em quadratura de fase*, o qual é baseado nas mesmas idéias da multiplexação por portadora em quadratura que foi discutida na Seção 3.5.

Na modulação por chaveamento em quadratura de fase (QPSK, *quadriphase-shift keying*), tal como no BPSK, a informação transportada pelo sinal transmitido está contida na fase da portadora senoidal. Em particular, a fase da portadora senoidal assume um de quatro valores igualmente espaçados, tais como $\pi/4$, $3\pi/4$, $5\pi/4$ e $7\pi/4$. Para este conjunto de valores, definimos o sinal transmitido como

$$s_i(t) = \begin{cases} \sqrt{\frac{2E}{T}} \cos\left[2\pi f_c t + (2i - 1)\frac{\pi}{4}\right], & 0 \leq t \leq T \\ 0, & \text{caso contrário} \end{cases} \quad (7.13)$$

na qual $i = 1, 2, 3, 4$. E é a *energia por símbolo do sinal transmitido* e T é a *duração do símbolo*. Cada um dos quatro valores de fase igualmente espaçados corresponde a um único par de bits, chamado de *dibit*. Por exemplo, podemos escolher o conjunto apresentado de valores de fase para representar o conjunto do *código gray* de dibits: 10, 00, 01 e 11. Nesta forma de codificação, vemos que apenas um único bit é alterado de um dibit para o próximo. Note que a *duração do símbolo* (isto é, a duração de cada dibit) é duas vezes a duração do bit, mostrado por

$$T = 2T_b \quad (7.14)$$

Usando uma identidade trigonométrica bem conhecida, podemos reescrever o sinal transmitido no intervalo $0 \leq t \leq T$ na forma expandida

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos\left[(2i - 1)\frac{\pi}{4}\right] \cos(2\pi f_c t) - \sqrt{\frac{2E}{T}} \sin\left[(2i - 1)\frac{\pi}{4}\right] \sin(2\pi f_c t) \quad (7.15)$$

na qual $i = 1, 2, 3, 4$. Com base na forma expandida da Eq. (7.15), podemos fazer algumas importantes observações:

1. Na realidade, o sinal QPSK é constituído por dois sinais BPSK.
2. Um sinal BPSK, representado pelo primeiro termo

$$\sqrt{2E/T} \cos[(2i - 1)\pi/4] \cos[(2\pi f_c t)],$$

define o produto da modulação de uma onda binária pela portadora senoidal $\sqrt{2/T} \cos(2\pi f_c t)$, a qual possui energia unitária durante a duração T do símbolo. Também reconhecemos que

$$\sqrt{E} \cos\left[(2i - 1)\frac{\pi}{4}\right] = \begin{cases} \sqrt{E/2} & \text{para } i = 1, 4 \\ -\sqrt{E/2} & \text{para } i = 2, 3 \end{cases} \quad (7.16)$$

portanto, vemos que esta onda binária possui amplitude igual a $\pm\sqrt{E/2}$.

3. O outro sinal BPSK, representado pelo segundo termo

$$-\sqrt{2E/T} \sin\left[(2i - 1)\frac{\pi}{4}\right] \sin(2\pi f_c t),$$

define o produto da modulação de uma onda binária diferente pela portadora senoidal $\sqrt{2/T} \sin(2\pi f_c t)$, a qual também possui energia unitária por símbolo. Desta vez, observamos que

$$-\sqrt{E} \sin\left[(2i - 1)\frac{\pi}{4}\right] = \begin{cases} -\sqrt{E/2} & \text{para } i = 1, 2 \\ \sqrt{E/2} & \text{para } i = 3, 4 \end{cases} \quad (7.17)$$

e, portanto, vemos que esta segunda onda binária também possui amplitude igual a $\pm\sqrt{E}/2$, apesar de em uma forma diferente com respeito ao índice i .

4. As duas ondas binárias definidas nas Eqs. (7.16) e (7.17) compartilham um valor comum para a duração do símbolo, ou seja, T .
5. As duas ondas portadoras senoidais identificadas nos pontos 2 e 3 estão em quadratura de fase uma com respeito a outra. Além disso, as duas possuem energia unitária por duração de símbolo. Podemos, portanto, afirmar que estas duas ondas portadoras constituem um *par ortogonal de funções base*.
6. Para cada possível valor do índice i , as Eqs. (7.16) e (7.17) identificam o díbit correspondente, como indicado na Tabela 7.1. Esta tabela também inclui outras colunas relacionadas com a fase do sinal QPSK e amplitudes das duas ondas binárias identificadas nos pontos 2 e 3.

TABELA 7.1 Relações entre o índice i e identidade do díbit correspondente, e outras questões correlatas

<i>Índice i</i>	<i>Fase do sinal QPSK (radianos)</i>	<i>Amplitudes das ondas binárias constituintes</i>		<i>Díbit de entrada</i> $0 \leq t \leq T$
		<i>Onda binária 1</i> $a_1(t)$	<i>Onda binária 2</i> $a_2(t)$	
1	$\pi/4$	$+\sqrt{E}/2$	$-\sqrt{E}/2$	10
2	$3\pi/4$	$-\sqrt{E}/2$	$-\sqrt{E}/2$	00
3	$5\pi/4$	$-\sqrt{E}/2$	$+\sqrt{E}/2$	01
4	$7\pi/4$	$+\sqrt{E}/2$	$+\sqrt{E}/2$	11

■ Geração e detecção coerente de sinais QPSK

Considerando os seis pontos resumidos anteriormente e o material apresentado previamente nesta seção sobre BPSK, podemos construir os diagramas em bloco da Fig. 7.7 para a geração e detecção dos sinais QPSK, como descrito a seguir:

(i) Geração

Para geral o sinal QPSK, a seqüência binária de dados de entrada é primeiro convertida para a forma polar por um *codificador de nível sem retorno para zero*, a saída do codificador é representada por $b(t)$. Os símbolos 1 e 0 são, portanto, representados por $\sqrt{E_b}$ e $-\sqrt{E_b}$, na qual $E_b = E/2$. A onda binária resultante é, a seguir, dividida por um demultiplexador (construído por um conversor serial para paralelo) em duas ondas binárias separadas consistindo dos números pares e ímpares dos bits de entrada $b(t)$. Estas duas ondas binárias, chamadas de *componentes demultiplexadas* da onda binária de entrada, são representadas por $a_1(t)$ e $a_2(t)$. Em qualquer intervalo de sinalização, as amplitudes de $a_1(t)$ e $a_2(t)$ são determinadas de acordo com as colunas 3 e 4 da Tabela 7.1, dependendo do díbit em particular que será transmitido. As ondas binárias demultiplexadas $a_1(t)$ e $a_2(t)$ são utilizadas para modular o par de portadoras em quadratura, $\sqrt{2/T} \cos(2\pi f_c t)$ e $\sqrt{2/T} \sin(2\pi f_c t)$. Finalmente, os dois sinais BPSK são subtraídos para produzir o sinal QPSK desejado, como mostrado na Fig. 7.7(a).

(ii) Detecção

O receptor QPSK é constituído de um *canal-(I)* em fase e um *canal-(Q)* em quadratura com uma entrada em comum, como mostrado na Fig. 7.7(b). Cada canal, por si só é construído por modulador em produto, filtro passa-baixa, amostrador e dispositivo de tomada de decisão. Sob condições ideais, os canais *I* e *Q* do receptor, respectiva-

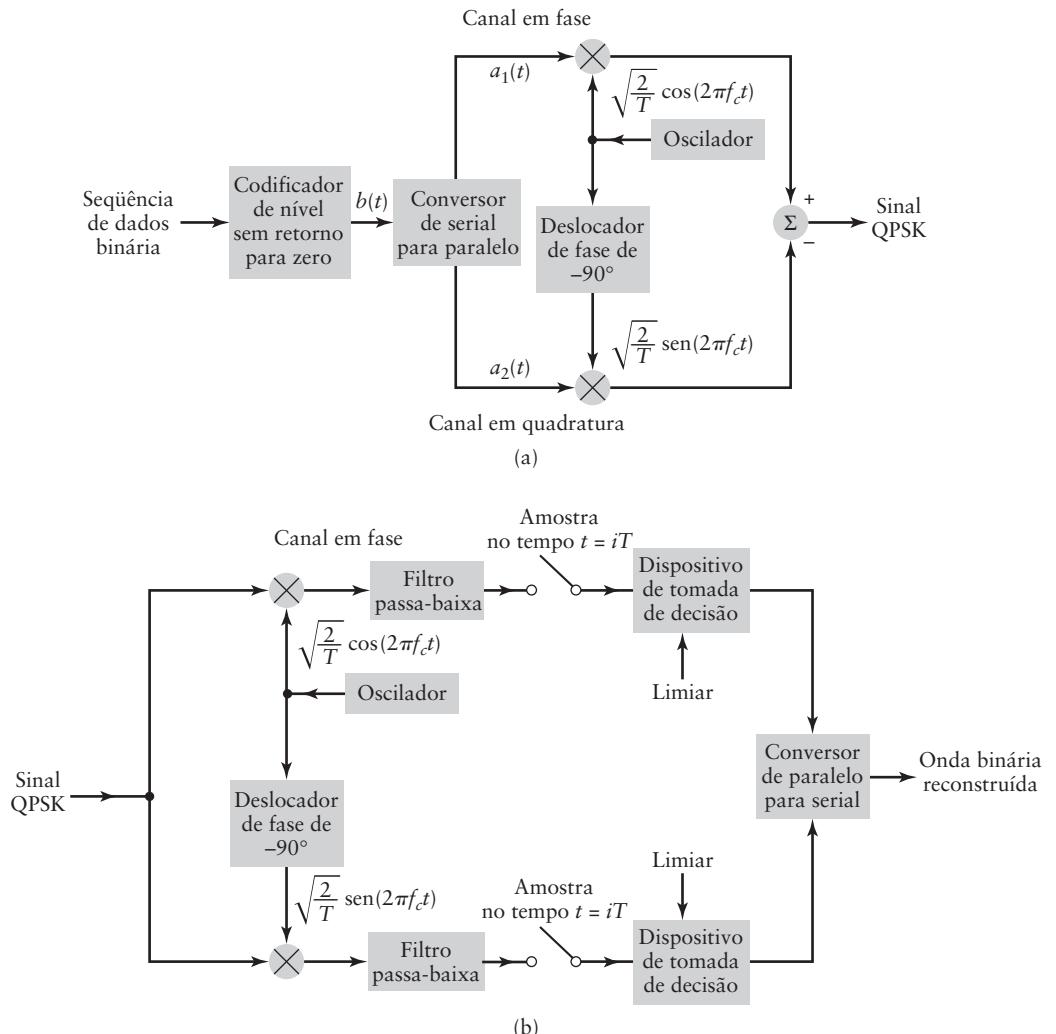


Figura 7.7 Diagrama em blocos do (a) transmissor QPSK, (b) receptor coerente QPSK, para os dois amostradores síncronos, $i = 0, \pm 1, \pm 2, \dots$

mente, recuperam as componentes demultiplexadas $a_1(t)$ e $a_2(t)$, responsáveis pela modulação do par ortogonal de portadoras no transmissor. Desta forma, aplicando as saídas destes dois canais a um *multiplexador* (construído com um conversor paralelo para série), o receptor recupera a seqüência binária original. (Iremos retornar à detecção coerente do QPSK no Capítulo 10.)

O projeto do receptor QPSK é construído com a estratégia descrita para o receptor coerente BPSK. Especificamente, cada um dos dois filtros passa-baixa no receptor coerente QPSK da Fig. 7.7(b) devem possuir uma largura de faixa igual ou maior do que o recíproco da duração do símbolo T para operação satisfatória do receptor.

É interessante compararmos o transmissor e o receptor QPSK apresentados na Fig. 7.8 com o transmissor e o receptor do sistema de multiplexação em quadratura da Fig. 3.17. Vemos que, apesar destas duas figuras serem utilizadas em aplicações diferentes, uma analógica e outra digital, os dois são baseados no mesmo princípio: *conservação de largura de faixa*. Desta forma, não é uma surpresa ver que seus respectivos transmissores e receptores compartilham estruturas similares.

■ Chaveamento em quadratura de fase deslocada

No QPSK, a amplitude da portadora é mantida constante. Entretanto, a fase da portadora pode pular por $\pm 90^\circ$ ou $\pm 180^\circ$ a cada duração de dois bits (dibit). Esta última propriedade pode ser de preocupação particular quando o sinal QPSK é filtrado durante o curso de transmissão em um canal de comunicação. Infelizmente, tal ação de filtragem pode fazer com que a amplitude da portadora, e portanto o envelope do sinal QPSK, flutue. Quando o sistema de transmissão de dados contém componentes não lineares, flutuações deste tipo são indesejáveis pois elas tendem a distorcer o sinal recebido. O resultado final é um diagrama de olho com abertura reduzida, a qual foi discutida na Seção 6.6.

A extensão das flutuações de amplitude exibidas por sinal QPSK podem ser reduzidas utilizando uma variante do chaveamento em quadratura de fase, chamado de *chaveamento em quadratura de fase deslocada* (OQPSK, offset quadriphase-shift keying).¹ Na OQPSK, a onda binária demultiplexada representada por $a_2(t)$ na Fig. 7.7(a) é atrasada (isto é, deslocada) pela duração de um bit com relação a outra onda binária demultiplexada representada por $a_1(t)$ na figura. Esta modificação possui o efeito de confinar a probabilidade de ocorrência de transições de fase a 0° e $\pm 90^\circ$. Entretanto, as transições de fase de $\pm 90^\circ$ no OQPSK ocorrem duas vezes mais freqüentemente, mas com uma faixa reduzida de flutuações de amplitude, quando comparado com o QPSK. Além das transições de fase de $\pm 90^\circ$, também existem transições de fase de $\pm 180^\circ$ na QPSK. Portanto, percebemos que flutuações de amplitude na OQPSK devido a filtragem possuem amplitude menor do que na QPSK. As considerações feitas aqui sobre transições de fase são ilustradas no próximo exemplo.

EXEMPLO 7.1 Transições de fase

As partes (a) e (b) da Fig. 7.8 mostram as formas de onda de QPSK e OQPSK, as duas produzidas pela seqüência binária de dados 0011011001 com a seguinte composição no intervalo $0 \leq t \leq T_b$:

- O díbit de entrada (isto é, o par de bits adjacentes na seqüência binária de dados) muda ao ir do intervalo $0 \leq t \leq 2T_b$ para o próximo intervalo $2T_b \leq t \leq 4T_b$.
- O díbit muda novamente do intervalo $2T_b \leq t \leq 4T_b$ para o próximo intervalo $4T_b \leq t \leq 6T_b$.

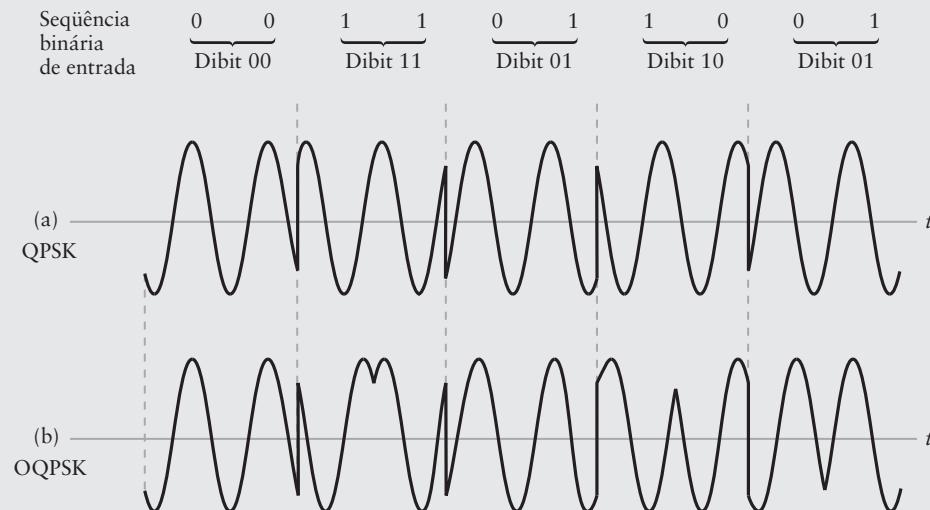


Figura 7.8 Comparação gráfica das transições de fase em QPSK e OQPSK.

¹ A modulação OQPSK também é chamada de *staggered quadriphase-shift keying* (SQPSK).

- O díbit muda uma vez mais ao ir do intervalo $4T_b \leq t \leq 6T_b$ para o próximo intervalo $6T_b \leq t \leq 8T_b$.
- Finalmente, o díbit muda pela última vez ao ir do intervalo $6T_b \leq t \leq 8T_b$ para o intervalo $8T_b \leq t \leq 10T_b$.

Examinando as duas formas de onda da Fig. 7.8, vemos que:

- (i) Na QPSK, a fase da portadora sofre saltos de $0^\circ, \pm 90^\circ, \pm 180^\circ$ a cada $2T_b$ segundos.
- (ii) Na OQPSK, por outro lado, a fase da portadora vivencia somente saltos de 0° ou $\pm 90^\circ$ a cada T_b segundos.

■ Experimento de computador III: espectro do QPSK e OQPSK

Para o nosso próximo experimento, calculamos o espectro de potência para sinais QPSK e OQPSK, assumindo a utilização de ondas quadradas como seqüência de dados de entrada. Novamente, o algoritmo de FFT foi utilizado nos cálculos. Especificamente, os cálculos são feitos para valor fixo de freqüência de portadora e duas durações de bit diferentes:

- (i) *Espectro QPSK*: As partes (a) e (b) da Fig. 7.9 mostram o espectro de potência (em decibels) do QPSK para os seguintes parâmetros:

Freqüência da portadora, $f_c = 8$ Hz

$$\text{Duração do bit, } T_b = \begin{cases} 1 \text{ s para a parte (a) da figura} \\ \frac{1}{2} \text{ s para a parte (b) da figura} \end{cases}$$

- (ii) *Espectro OQPSK*: As partes (a) e (b) da Fig. 7.10 mostram o espectro de potência correspondente para o OQPSK para os mesmos parâmetros utilizados no QPSK.

Como antes, cada parte das Figs. 7.9 e 7.10 incluem dois gráficos espetrais, um determinado computacionalmente e outro teoricamente, com uma boa concordância entre a teoria e o experimento.

A conclusão importante obtida pela comparação das duas partes da Fig. 7.9 para QPSK com as da Fig. 7.10 para o OQPSK é resumida a seguir: apesar dos dois métodos digitais de modulação naturalmente resultarem em formas de onda diferentes, o espectro de potência de QPSK e OQPSK são idênticos para o mesmo conjunto de parâmetros do sistema.

Além disso, comparando os gráficos de potência espectral da Fig. 7.9 para QPSK com os da Fig. 7.6 para BPSK, observamos que a QPSK ocupa a largura de faixa igual a metade da BPSK.

- **PROBLEMA 7.3** Apesar dos sinais QPSK e OQPSK possuírem formas de onda diferentes, seus espectros de amplitude são idênticos, mas seus espectros de fase são diferentes por uma componente de fase linear. Justifique a validade desta afirmativa. ◀

7.4 CHAVEAMENTO EM FREQÜÊNCIA

■ Chaveamento binário em freqüência

Na forma mais simples de chaveamento em freqüência, chamado de *chaveamento binário em freqüência* (BFSK, *binary frequency-shift keying*), os símbolos 0 e 1 são distinguidos um do outro pela transmissão de duas ondas senoidais que diferem em freqüência por um total fixo. Um par típico de ondas senoidais é descrito por

$$s_i(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_1 t), & \text{para o símbolo 1 correspondendo a } i = 1 \\ \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_2 t), & \text{para o símbolo 0 correspondendo a } i = 2 \end{cases} \quad (7.18)$$

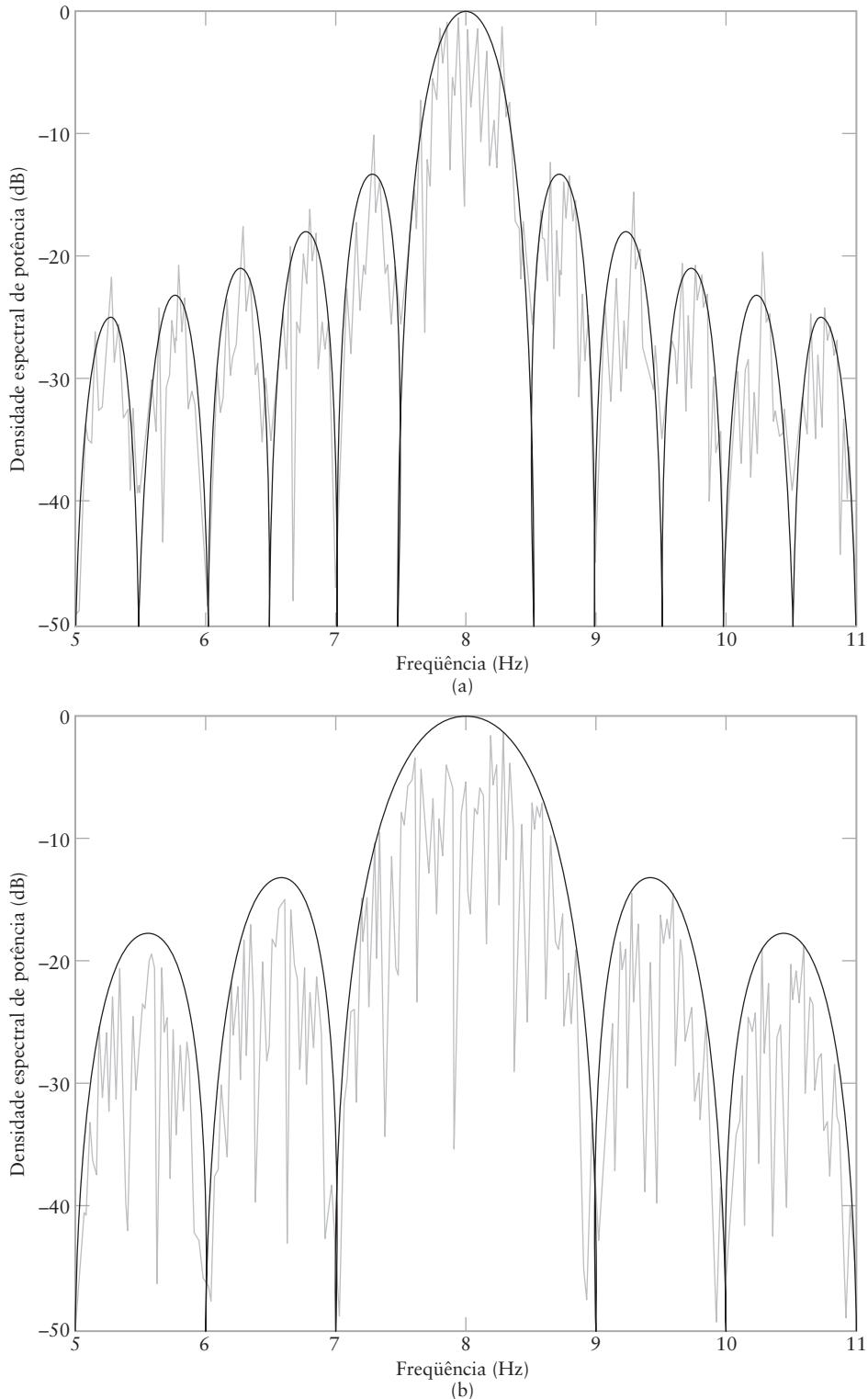


Figura 7.9 Espectro de potência do sinal QPSK produzido por uma onda quadrada como sinal modulante para freqüência de portadora fixa e duração de bit variável: (a) $f_c = 8 \text{ Hz}$ e $T_b = 1\text{s}$. (b) $f_c = 8\text{Hz}$ e $T_b = 1/2\text{s}$.

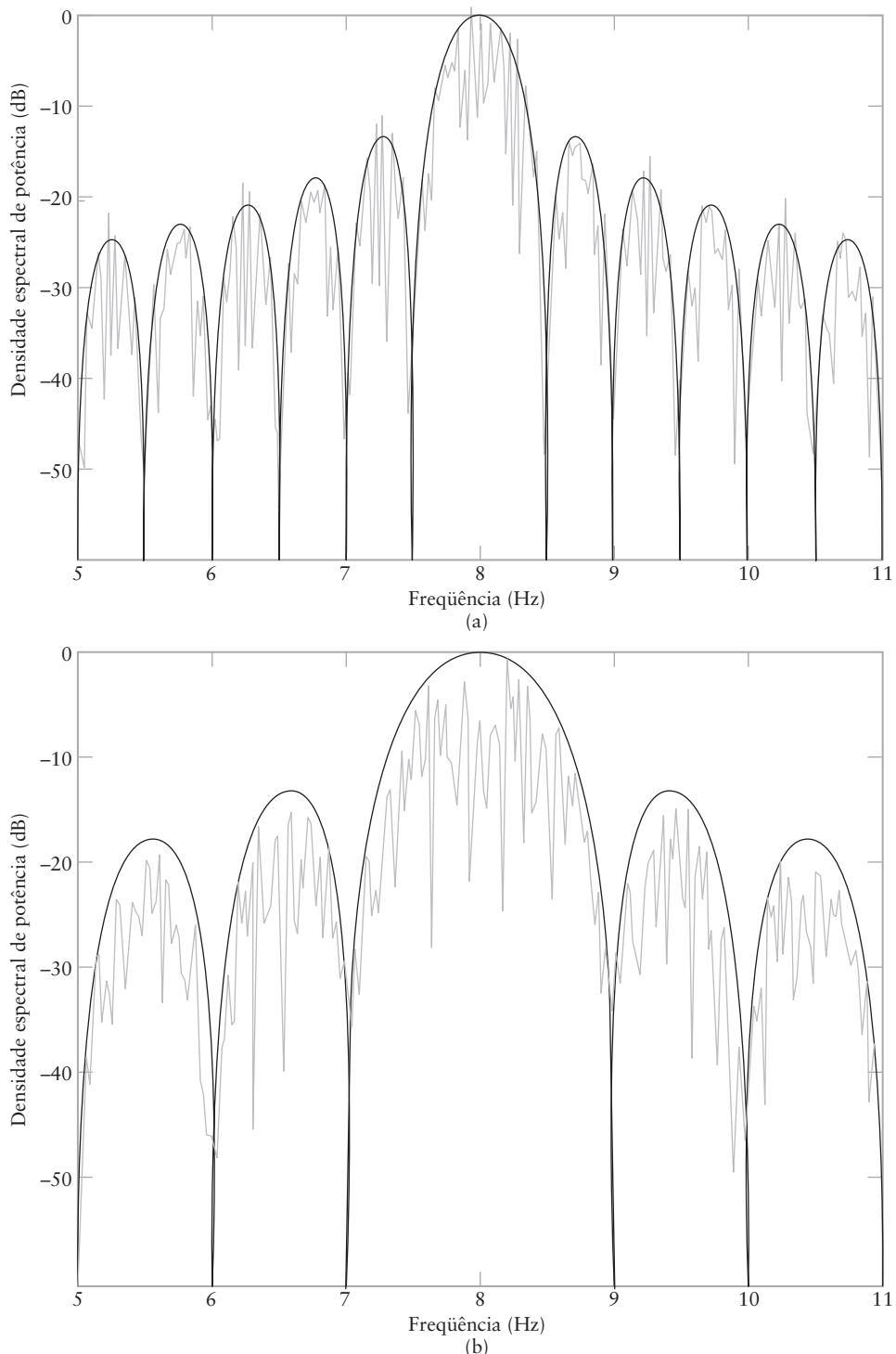


Figura 7.10 Espectro de potência do sinal OQPSK produzido por uma onda quadrada como sinal modulante para freqüência de portadora fixa e duração de bit variável: (a) $f_c = 8 \text{ Hz}$ e $T_b = 1\text{s}$. (b) $f_c = 8\text{Hz}$ e $T_b = 1/2\text{s}$.

na qual E_b é a energia por bit do sinal transmitido. Quando as freqüências f_1 e f_2 são escolhidas de tal forma que elas diferem uma da outra por um total igual ao recíproco da duração T_b do bit, o sinal BFSK é chamado de *BFSK de Sunde*, em homenagem a seu originador. Este sinal é um *sinal contínuo em fase* no sentido de que a continuidade de fase é sempre mantida, incluindo os tempos de chaveamento entre bits.

■ Experimento de computador IV: BFSK de Sunde

(i) Forma de onda

A Fig. 7.11 mostra a forma de onda do sinal BFSK de Sunde produzido pela seqüência binária de entrada 0011011001 para uma duração de bit $T_b = 1\text{s}$. A parte (a) da figura apresenta a forma de onda da seqüência de entrada e a parte (b) mostra a forma de onda correspondente do sinal BFSK. A última parte da figura mostra, claramente, a propriedade de continuidade de fase do BFSK de Sunde.

(ii) Espectro

A Fig. 7.12 mostra dois gráficos espectrais sobrepostos do BFSK de Sunde para uma onda quadrada de entrada para freqüências positivas. Como antes, um gráfico é computacional e o outro teórico. Os parâmetros para o cálculo computacional são:

$$\text{Duração de bit:} \quad T_b = 1\text{s}$$

$$\text{Freqüência da portadora:} \quad f_c = 8\text{Hz}$$

Examinando a Fig. 7.12 podemos fazer as seguintes observações para freqüências positivas:

- (i) O espectro contém duas componentes de linha na freqüência $f = f_c \pm 1/(2T_b)$, iguais a 7,5 Hz e 8,5 Hz para $f_c = 8\text{Hz}$ e $T_b = 1\text{s}$.
 - (ii) O lóbulo principal ocupa uma faixa com largura igual a $(3/T_b) = 3\text{Hz}$, centrado na freqüência da portadora $f_c = 8\text{Hz}$.
 - (iii) O maior lóbulo lateral está, aproximadamente, 21 dB abaixo do lóbulo principal.
- **PROBLEMA 7.4** Mostre que o processo de modulação envolvido na geração do BFSK de Sunde é não linear. ◀

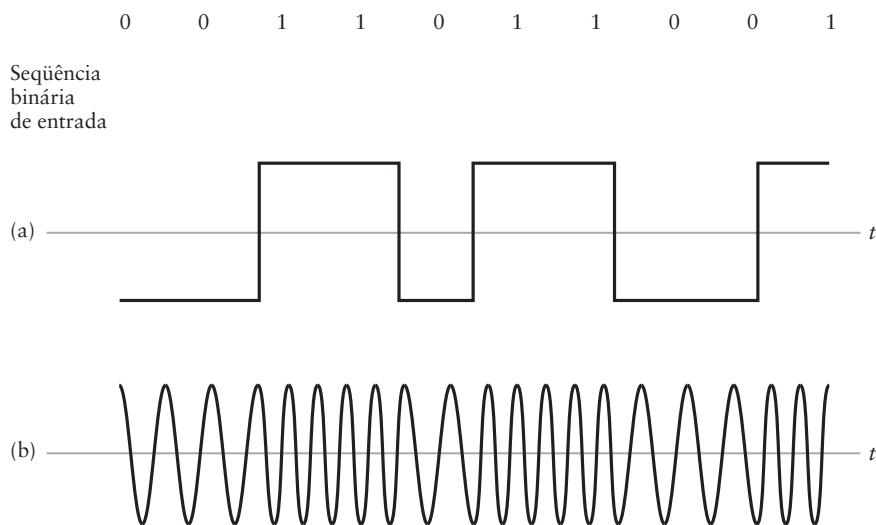


Figura 7.11 (a) Seqüência binária e sua forma de onda codificada sem retorno para zero. (b) Sinal BFSK de Sunde.

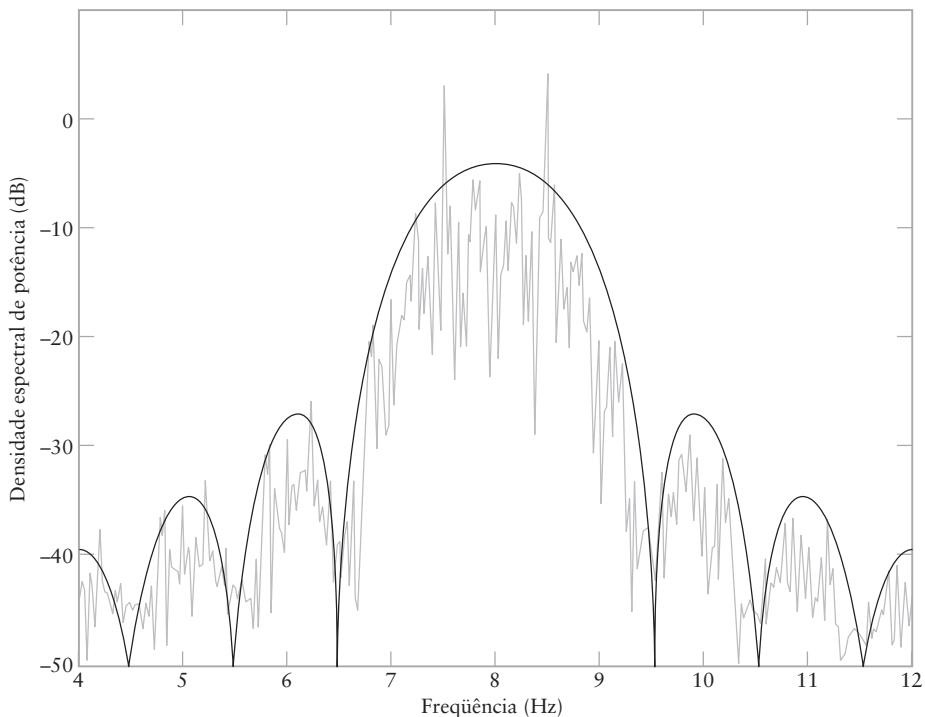


Figura 7.12 Espectro de potência para o BFSK de Sunde produzido por uma onda quadrada como sinal modulante e com os seguintes parâmetros: $f_c = 8\text{Hz}$ e $T_b = 1\text{s}$.

■ Chaveamento em freqüência com fase contínua

O BFSK de Sunde é a forma mais simples de uma família de sinais digitalmente modulados chamados de *chaveamento em freqüência com fase contínua* (CPFSK, continuous-phase frequency-shift keying), os quais exibem a seguinte propriedade distinta:

A onda modulada mantém a continuidade de fase em todos os pontos de transição, mesmo que nestes pontos a seqüência binária de dados de entrada mude indo e voltando entre os símbolos 0 e 1.

Em outras palavras, o sinal CPFSK é uma onda modulada por onda contínua, como qualquer outra onda modulada em ângulo do mundo analógico, apesar do fato da onda modulada ser descontínua.

No BFSK de Sunde, a excursão total δf na freqüência transmitida do símbolo 0 para o símbolo 1, e vice-versa, é igual a taxa de bit da seqüência de dados de entrada. Em outra forma especial de CPFSK chamada de *mínimo chaveamento* (MSK, minimum shift keying), o processo de modulação binário utiliza um valor diferente para a excursão de freqüência δf , com o resultado desta nova onda modulada possuir propriedades espectrais melhores do que o BFSK de Sunde.

■ Mínimo chaveamento

No MSK, a excursão total de freqüência δf do símbolo binário 1 para o símbolo 0, ou vice-versa, é metade da taxa de bit, mostrado por

$$\begin{aligned}\delta f &= f_1 - f_2 \\ &= \frac{1}{2T_b}\end{aligned}\tag{7.19}$$

A freqüência da portadora não modulada é a média aritmética das duas freqüências f_1 e f_2 transmitidas, ou seja,

$$f_c = \frac{1}{2}(f_1 + f_2) \quad (7.20)$$

Expressando f_1 e f_2 em termos da freqüência f_c da portadora e da excursão total δf de freqüência, temos

$$f_1 = f_c + \frac{\delta f}{2}, \quad \text{para o símbolo 1} \quad (7.21)$$

$$f_2 = f_c - \frac{\delta f}{2}, \quad \text{para o símbolo 0} \quad (7.22)$$

Desta forma, definimos formalmente o sinal MSK como a onda modulada em ângulo

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_c t + \theta(t)] \quad (7.23)$$

na qual $\theta(t)$ é a fase do sinal MSK. Em particular, quando a freqüência f_1 é transmitida, correspondendo ao símbolo 1, vemos das Eqs. (7.21) e (7.23) que a fase $\theta(t)$ assume o valor

$$\begin{aligned} \theta(t) &= 2\pi\left(\frac{\delta f}{2}\right)t \\ &= \frac{\pi t}{2T_b}, \quad \text{para o símbolo 1} \end{aligned} \quad (7.24)$$

Em palavras, isto significa que no tempo $t = T_b$, a transmissão do símbolo 1 aumenta a fase do sinal MSK $s(t)$ por $\pi/2$ radianos. Pelo mesmo motivo, quando a freqüência f_2 é transmitida, correspondendo ao símbolo 0, vemos das Eqs. (7.22) e (7.23) que a fase $\theta(t)$ assume o valor

$$\begin{aligned} \theta(t) &= 2\pi\left(-\frac{\delta f}{2}\right)t \\ &= -\frac{\pi t}{2T_b}, \quad \text{para o símbolo 0} \end{aligned} \quad (7.25)$$

Isto significa que no tempo $t = T_b$, a transmissão do símbolo 0 diminui a fase de $s(t)$ por $\pi/2$ radianos.

A mudança de fase descrita nas Eqs. (7.24) e (7.25) para o MSK é radicalmente diferente de sua contrapartida no BFSK de Sunde. Especificamente, a fase do sinal BFSK de Sunde muda por $-\pi$ radianos ao término do intervalo representando o símbolo 0 e $+\pi$ radianos para o símbolo 1. Entretanto, as mudanças $-\pi$ e $+\pi$ são *exatamente as mesmas*, módulo 2π . Esta observação, de fato, significa que o BFSK de Sunde não possui memória, em outras palavras, saber qual mudança em particular ocorreu no intervalo prévio de bit não ajuda em nada o intervalo atual de bit. Em contraste, temos uma situação completamente diferente no caso do MSK em virtude da forma diferente na qual as transmissões dos símbolo 1 e 0 afetam a fase $\theta(t)$, como mostrado nas Eqs. (7.24) e (7.25). Observe, também, que a excursão total de freqüência δf no MSK é o *menor espaçamento de freqüência* entre os símbolos 0 e 1 que permite que suas representações FSK sejam coerentemente ortogonais, por isso a terminologia “mínimo chaveamento”.

EXEMPLO 7.2 Relação entre formas de onda OQPSK e MSK

O propósito deste exemplo é ilustrar a relação que existe entre formas de onda OQPSK e MSK. As Figs. 7.13 e 7.14 mostram esta relação fundamental:

- As cinco formas de onda da Fig. 7.13 mostram as componentes do sinal OQPSK para a seqüência binária de dados de entrada 0011011001.
- As cinco formas de onda correspondentes da Fig. 7.14 apresentam as componentes do sinal MSK para a mesma seqüência binária de entrada 0011011001.

Comparando os resultados mostrados nas Figs. 7.13 e 7.14, podemos fazer a seguinte observação. Apesar do OQPSK e MSK serem obtidos por princípios de modulação diferentes, o MSK do chaveamento em freqüência e o OQPSK do chaveamento em fase, estas duas ondas moduladas digitalmente são, de fato, muito relacionadas. A diferença básica entre elas está simplesmente na forma pela qual os símbolos binários em suas componentes em fase e em quadratura são codificados em nível. No OQPSK, a codificação em nível é baseada em pulsos retangulares, com uma onda binária deslocada da outra onda binária pela duração de um bit. Por outro lado, no MSK, a codificação de nível é baseada em meio ciclo da cossenóide.

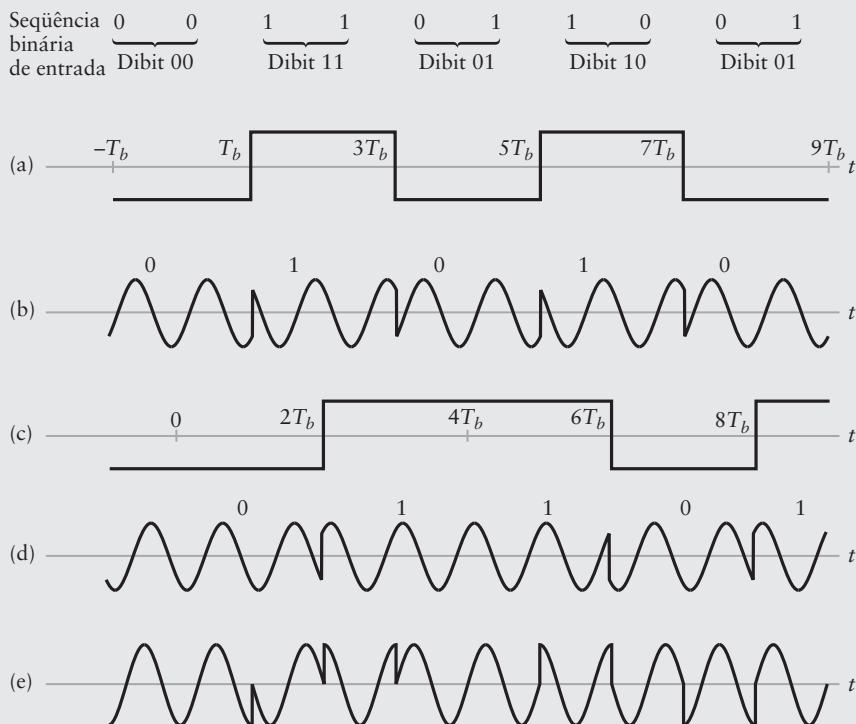


Figura 7.13 Componentes do sinal OQPSK: (a) Sinal modulante para a componente em fase. (b) Forma de onda modulada da componente em fase. (c) Sinal modulante para a componente em quadratura. (d) Forma de onda modulada da componente em quadratura. (e) Forma de onda do sinal OQPSK obtido subtraíndo-se (d) de (b).

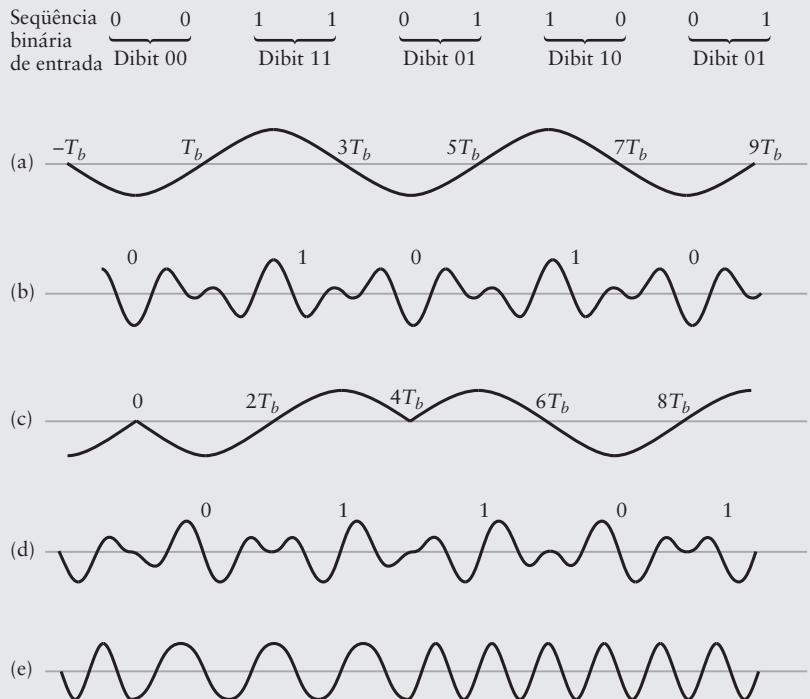


Figura 7.14 Componentes do sinal MSK. (a) Sinal modulante para a componente em fase. (b) Forma de onda modulada em fase. (c) Sinal modulante para a componente em quadratura. (d) Componente modulada em quadratura. (e) Forma de onda do sinal MSK obtida subtraindo-se (d) de (b).

A forte relação entre MSK e OQPSK mostrada no Exemplo 7.2 nos impele ao estágio de formulação analítica do sinal MSK, descrito a seguir.

■ Formulação do mínimo chaveamento²

Para procedermos com a formulação, nos referimos à Eq. (7.23) e utilizamos uma identidade trigonométrica bem conhecida para expandir a onda modulada em ângulo (isto é, o sinal MSK) $s(t)$ para

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos(\theta(t)) \cos(2\pi f_c t) - \sqrt{\frac{2E_b}{T_b}} \sin(\theta(t)) \sin(2\pi f_c t) \quad (7.26)$$

com esta equação, podemos identificar:

- (i) $s_I(t) = \sqrt{\frac{2E_b}{T_b}} \cos(\theta(t))$ é a componente em fase (I) associada com a portadora $\sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t)$. (7.27)

- (ii) $s_Q(t) = \sqrt{\frac{2E_b}{T_b}} \sin(\theta(t))$ é a componente em quadratura (Q) associada com a portadora deslocada por 90° . (7.28)

Para ressaltarmos a presença da seqüência binária de dados de entrada na composição de $s_I(t)$ e $s_Q(t)$, as reformulamos como

$$s_I(t) = a_1(t) \cos(2\pi f_0 t) \quad (7.29)$$

² A formulação analítica do MSK apresentada aqui segue Ziemer e Tranter (2002).

e

$$s_Q(t) = a_2(t) \operatorname{sen}(2\pi f_0 t) \quad (7.30)$$

Os sinais $a_1(t)$ e $a_2(t)$ são duas ondas binárias extraídas da seqüência binária de dados de entrada através de demultiplexação e deslocamento, de maneira similar ao OQPSK. Desta forma, eles assumem os valores +1 e -1 em intervalos de símbolo (isto é, díbit) de duração $T = 2T_b$, na qual T_b é a duração de bit da seqüência binária de dados de entrada. Os dois sinais $a_1(t)$ e $a_2(t)$ são respectivamente ponderados pelas funções senoidais $\cos(2\pi f_0 t)$ e $\operatorname{sen}(2\pi f_0 t)$, na qual a freqüência f_0 deve ser determinada. Para definir f_0 , utilizamos as Eqs. (7.29) e (7.30) para reconstruir a onda modulada em ângulo original $s(t)$ em termos dos sinais de dados $a_1(t)$ e $a_2(t)$. Desta forma, obtemos

$$\begin{aligned} \theta(t) &= -\operatorname{tg}^{-1} \left[\frac{s_Q(t)}{s_I(t)} \right] \\ &= -\operatorname{tg}^{-1} \left[\frac{a_2(t)}{a_1(t)} \operatorname{tg}(2\pi f_0 t) \right] \end{aligned} \quad (7.31)$$

Com base nesta equação, reconhecemos dois possíveis cenários:

1. $a_2(t) = a_1(t)$. Este cenário aparece quando dois símbolos binários sucessivos (constituindo um díbit) na seqüência de dados de entrada são os mesmos (isto é, os dois são 0s ou 1s). Logo, a Eq. (7.31) se reduz para

$$\begin{aligned} \theta(t) &= -\operatorname{tg}^{-1}[\operatorname{tg}(2\pi f_0 t)] \\ &= -2\pi f_0 t \end{aligned} \quad (7.32)$$

2. $a_2(t) = -a_1(t)$. Este segundo cenário aparece quando dois símbolos binários sucessivos (isto é, díbits) da seqüência de dados de entrada são diferentes. Logo, a Eq. (7.31) se reduz para

$$\begin{aligned} \theta(t) &= -\operatorname{tg}^{-1}[-\operatorname{tg}(2\pi f_0 t)] \\ &= 2\pi f_0 t \end{aligned} \quad (7.33)$$

As Eqs. (7.32) e (7.33) são, respectivamente, de forma matemática similar às Eqs. (7.25) e (7.24). Desta forma, podemos definir formalmente, agora,

$$f_0 = \frac{1}{4T_b} \quad (7.34)$$

Resumindo, dada uma onda binária $b(t)$ codificada em nível sem retorno para zero, de duração predeterminada T_b e uma onda portadora senoidal de freqüência f_c , podemos formular o sinal MSK pelo procedimento a seguir:

1. Utilize a onda binária $b(t)$ para construir as ondas binárias demultiplexadas e deslocadas $a_1(t)$ e $a_2(t)$.
2. Utilize a Eq. (7.34) para determinar a freqüência f_0 .
3. Utilize as Eqs. (7.29) e (7.30) para determinar as componentes em fase $s_I(t)$ e em quadratura $s_Q(t)$, respectivamente, das quais o sinal MSK $s(t)$ é obtido.

Este foi o procedimento utilizado no Exemplo 7.3 para construir o sinal MSK de suas componentes relacionadas.

- **PROBLEMA 7.5** Para resumir, podemos dizer que o MSK é um OQPSK no qual os símbolos das componentes em fase e em quadratura (com base em díbit-a-díbit), são ponderados pela função pulso básica

$$p(t) = \operatorname{sen}\left(\frac{\pi t}{2T_b}\right) \operatorname{ret}\left(\frac{t}{2T_b} - \frac{1}{2}\right)$$

na qual T_b é a duração do bit e $\operatorname{ret}(t)$ é a função retangular de duração unitária e amplitude unitária. Justifique este resumo.

- **PROBLEMA 7.6** A seqüência 11011100 é aplicada a um modulador MSK. Assumindo que o ângulo $\theta(t)$ do sinal MSK é zero no tempo $t = 0$, trace o diagrama que mostra a evolução de $\theta(t)$ para os oito símbolos binários da seqüência de entrada.
- **PROBLEMA 7.7** O processo de modulação em ângulo envolvida na geração de um sinal MSK é linear. Justifique esta afirmativa.
- **PROBLEMA 7.8** Uma forma simples de demodular um sinal MSK é utilizar um discriminador de freqüência, o qual foi discutido no Capítulo 4 sobre modulação em ângulo. Justifique esta utilização e especifique a característica linear de entrada/saída do discriminador.

■ Experimento de computador V: espectro MSK

Neste experimento, calcularemos o espectro de potência do sinal MSK produzido por uma onda quadrada de entrada. Os parâmetros para este experimento são:

Duração de bit	$T_b = 1$ s
Freqüência da portadora	$f_c = 8$ Hz

A Fig. 7.15 mostra dois gráficos espectrais de potência do sinal MSK, um computacional e outro teórico. Comparando esta figura com a do QPSK na Fig. 7.9(a) e do BFSK

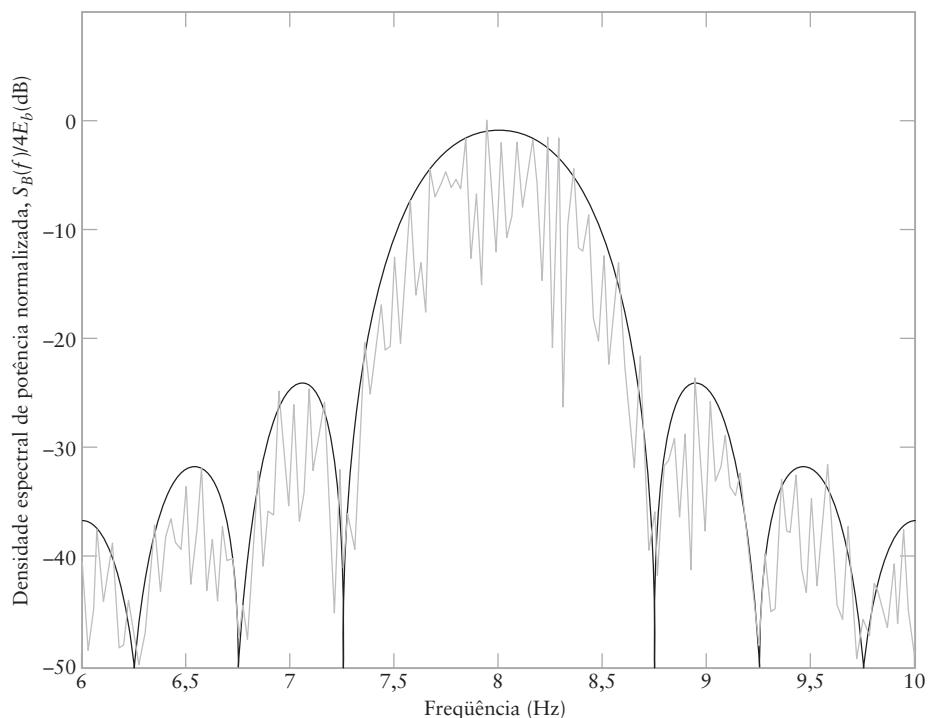


Figura 7.15 Espectro de potência para o MSK produzido por uma onda quadrada como sinal modulante e com os seguintes parâmetros: $f_c = 8$ Hz e $T_b = 1$ s.

de Sunde na Fig. 7.12 para o mesmo conjunto de parâmetros, podemos fazer as seguintes observações para freqüências positivas:

- (i) *MSK versus QPSK*: O lóbulo principal do MSK ocupa uma faixa de freqüência cuja largura é $1,5/T_b = 1,5\text{Hz}$, centrada na freqüência da portadora $f_c = 8\text{Hz}$. Com esta base, a largura de faixa de transmissão do MSK é 50% maior do que o QPSK. Entretanto, os lóbulos laterais do MSK são consideravelmente menores do que os do QPSK.
- (ii) *MSK versus BFSK de Sunde*: A largura de faixa de transmissão do MSK é metade do BFSK de Sunde. Além disso, o BFSK de Sunde exibe duas componentes de linha em $f = f_c \pm 1/(2T_b)$, enquanto que o espectro do MSK é contínuo ao longo de toda a faixa de freqüências.

A Fig. 7.16 mostra o espectro de potência do sinal MSK para o mesmo conjunto de parâmetros utilizados na Fig. 7.15. Na Fig. 7.16, a freqüência varre a faixa de 0 a 50 Hz, enquanto que na Fig. 7.15, a freqüência varre a faixa de 6 a 10 Hz. O motivo para a inclusão da Fig. 7.16 é simplesmente ilustrar que apesar da freqüência f_c da portadora não ser alta o suficiente para eliminar sobreposição espectral, a sobreposição é relativamente pequena, evidenciada por:

- O pequeno valor do espectro na freqüência zero.
- O pequeno grau de assimetria na freqüência da portadora $f_c = 8\text{Hz}$.

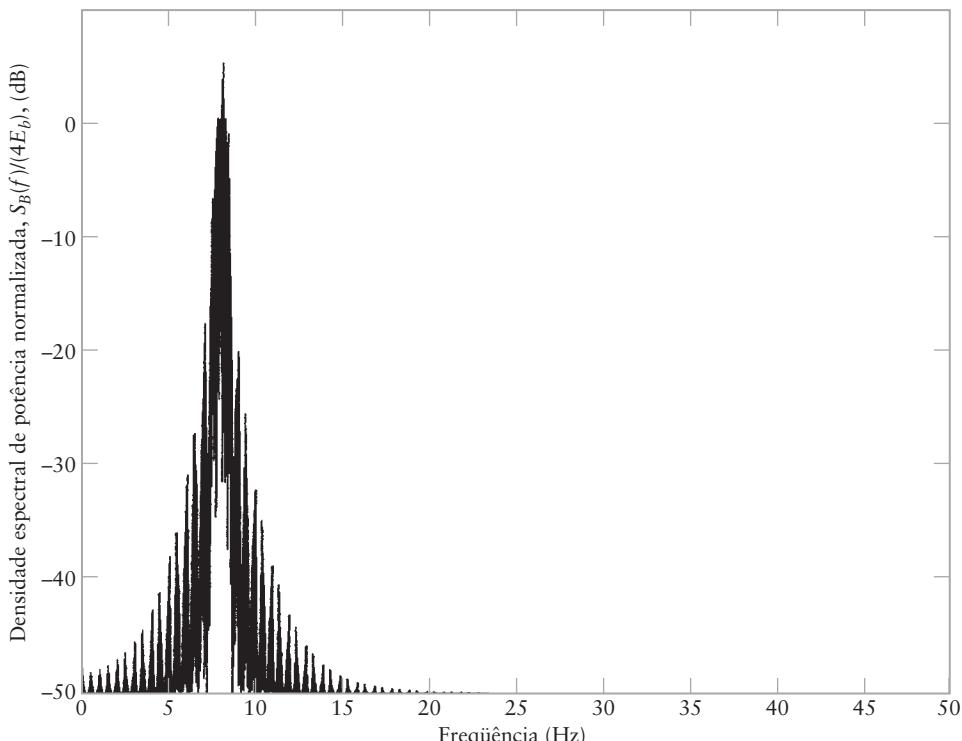


Figura 7.16 Confirmação ilustrativa de que, para o MSK produzido pela onda modulante quadrada, a freqüência de portadora $f_c = 8\text{Hz}$ é alta o suficiente para produzir uma sobreposição espectral insignificante para a duração de bit $T_b = 1\text{s}$.

7.5 RESUMO DOS TRÊS ESQUEMAS DE SINALIZAÇÃO BINÁRIA

As partes (b), (c) e (d) da Fig. 7.1 ilustram as formas de onda dos três esquemas básicos de sinalização, BASK, BPSK e BFSK, os quais são respectivamente produzidos pela seqüência binária de dados 0011011001 da parte (a) da figura. Esta figura claramente ilustra os seguintes pontos:

- (i) BASK, BPSK e BFSK são as contrapartidas digitais de modulação em amplitude, modulação em fase e modulação em freqüência, respectivamente. Este ponto reforça observações prévias.
- (ii) Tanto BASK quanto BPSK exibem *descontinuidade*. Em contraste, é possível configurar o BFSK de tal forma que a *continuidade de fase* seja mantida ao longo de toda a seqüência binária de dados de entrada. De fato, a forma de onda BFSK mostrada na parte (d) da figura é um exemplo de mínimo chaveamento, a qual exibe esta propriedade.

Além disso, a Tabela 7.2 apresenta um resumo dos três esquemas de modulação binária: BASK, BPSK e BFSK. As colunas 2 e 3 da tabela resumem os discriminantes que distinguem estes três esquemas básicos em termos matemáticos. A última coluna da tabela mostra as representações *fasoriais* dos três esquemas, mostrando, portanto, os discriminantes que os distinguem em termos gráficos.

7.6 ESQUEMAS DE MODULAÇÃO DIGITAL NÃO COERENTES

Receptores coerentes, exemplificados pelos esquemas mostrados nas Figs. 7.4(b) e 7.7(b), precisam conhecer uma referência da fase da onda portadora para estabelecer o sincronismo com seus respectivos transmissores. Entretanto, em alguns ambientes de comunicação, a sincronização de fase entre o receptor e seu transmissor ou é impraticável ou é muito cara. Em situações deste tipo, precisamos utilizar a *detecção não coerente* abandonando a necessidade de sincronização entre receptor e seu transmissor, sabendo que quando fazemos isto, a performance do receptor é degradada na presença de ruído de canal (como discutiremos no Capítulo 10).

Na discussão de técnicas de sinalização binária, descrevemos o BASK, BPSK e BFSK, nesta ordem. Este ordenamento foi motivado pelas seguintes considerações. Tanto o BASK quanto o BPSK são exemplos de modulação linear, com o aumento da complexidade ao se sair do BASK indo em direção ao BPSK. Por outro lado, BFSK é, em geral, um exemplo de modulação não linear, sendo deixado, portanto, por último. Nesta seção, entretanto, alteramos a ordem na qual os esquemas de detecção não coerentes são considerados. Especificamente, começaremos pela detecção não coerente do BASK, seguido pela do BFSK. A razão para esta nova ordem é guiada pelo fato dos dois utilizarem técnicas similares, baseadas em filtros passa-faixa, mas com o aumento de complexidade do BASK para o BFSK. A discussão é completada apresentando-se a detecção não coerente do BPSK, a qual requer considerações mais cuidadosas.

■ Detecção não coerente do sinal BASK

Referindo à Eq. (7.10), vemos que a geração de sinais BASK envolve a utilização de uma única portadora senoidal com freqüência f_c para o símbolo 1 e o desligamento da transmissão para o símbolo 0. Agora, o projetista do sistema precisa ter o conhecimento de dois parâmetros do sistema:

- A freqüência f_c da portadora.
- A largura de faixa de transmissão, a qual é determinada pela duração T_b do bit.

É, portanto, natural utilizarmos estes parâmetros conhecidos no projeto de um receptor não coerente para o BASK. Especificamente, o receptor é constituído por um filtro passa-faixa, seguido por um detector de envelope, então um amostrador e, finalmente, um dispositivo de tomada de decisão, como indicado na Fig. 7.17.

TABELA 7.2 Resumo dos três esquemas básicos de modulação binária

<i>Tipo do esquema de modulação</i>	<i>Parâmetro variável</i>	<i>Definição da onda modulada $s_1(t)$ ou $s_2(t)$ para $0 \leq t \leq T_b$</i>	<i>Representação fatorial da onda modulada</i>
1. Chaveamento binário em amplitude (BASK)	$\begin{pmatrix} \text{Amplitude da portadora} \\ A_c \end{pmatrix} = \begin{cases} \sqrt{\frac{2}{T_b}} & \text{para o símbolo 1} \\ 0 & \text{para o símbolo 0} \end{cases}$	$s_1(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t)$ para o símbolo 1 $s_2(t) = 0$ para o símbolo 0	Fator nulo para o símbolo 0 Fator para o símbolo 1
2. Chaveamento binário em fase (BASK)	$\begin{pmatrix} \text{Fase da portadora} \\ \phi_c \end{pmatrix} = \begin{cases} 0 & \text{para o símbolo 1} \\ \pi & \text{para o símbolo 0} \end{cases}$	$s_1(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t)$ para o símbolo 1 $s_2(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t + \pi)$ para o símbolo 0	Fator para o símbolo 0 Fator para o símbolo 1
3. Chaveamento binário em frequência (BASK)	$\begin{pmatrix} \text{Frequência da portadora} \\ f_c \end{pmatrix} = \begin{cases} f_1 & \text{para o símbolo 1} \\ f_2 & \text{para o símbolo 0} \end{cases}$	$s_1(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_1 t)$ para o símbolo 1 $s_2(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_2 t)$ para o símbolo 0	Fator para o símbolo 0 Fator para o símbolo 1

Notações:

T_b = duração do bit

E_b = energia por bit do sinal transmitido

Portadora: $c(t) = A_c \cos(2\pi f_c t + \phi_c)$

A fase ϕ_c da portadora é feita igual a zero para o BASK e BFSK.

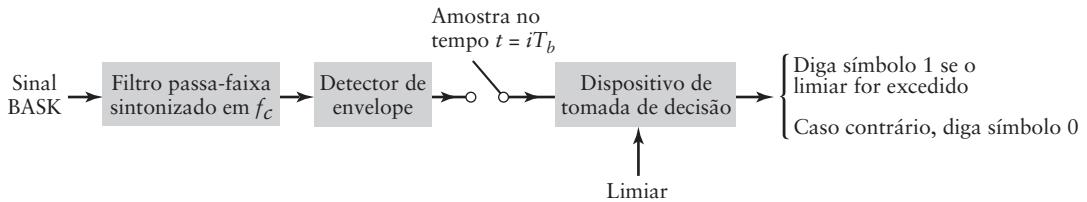


Figura 7.17 Receptor não coerente BASK, o inteiro i no amostrador é igual a $0, \pm 1, \pm 2, \dots$

O *filtro passa-faixa* é projetado para ter uma freqüência de meia faixa igual à freqüência f_c da portadora e largura de faixa igual à largura de faixa de transmissão do sinal BASK. Além disso, considera-se que a interferência intersimbólica (IIS) produzida pelo filtro é negligenciável, o que, por sua vez, requer que o tempo de subida e o tempo de descida da resposta do filtro a um pulso retangular sejam pequenos quando comparados com a duração T_b do bit. Sob estas condições, vemos que em resposta ao sinal BASK de entrada (considerando livre de ruído), o filtro passa-faixa produz uma senóide pulsada para o símbolo 1 e, idealmente, nenhuma saída para o símbolo 0. A seguir, o detector de envelope segue o envelope da versão filtrada do sinal BASK. Finalmente, o *dispositivo de tomada de decisão*, trabalhando em conjunto com o *amostrador*, regenera a seqüência binária de dados original comparando a amostra da saída do detector de envelope com um *limiar* predeterminado a cada T_b segundos. Esta operação assume a disponibilidade de uma temporização de bit no receptor. Se o limiar for excedido no tempo $t = iT_b$, $i = 0, \pm 1, \pm 2, \dots$, o receptor decide em favor do símbolo 1. Caso contrário, ele decide em favor do símbolo 0. Na ausência de ruído de canal e distorção de canal, a saída do receptor (bit-a-bit) deve ser uma réplica exata da seqüência binária de dados original, aplicada ao transmissor, sujeito às considerações anteriores do filtro passa-faixa.

■ Detecção não coerente de sinais BFSK

Para o caso do BFSK, lembramos da Seção 7.4 que a transmissão de símbolos 1 e 0 é representada por duas ondas portadoras com freqüências f_1 e f_2 , respectivamente, com adequado espaçamento entre elas. Com esta caracterização, podemos construir a detecção não coerente baseada no BASK formulando o receptor não coerente BFSK da Fig. 7.18. O receptor consiste de dois caminhos, um para lidar com a freqüência f_1 (isto é, símbolo 1) e o outro para lidar com a freqüência f_2 (isto é, o símbolo 0):

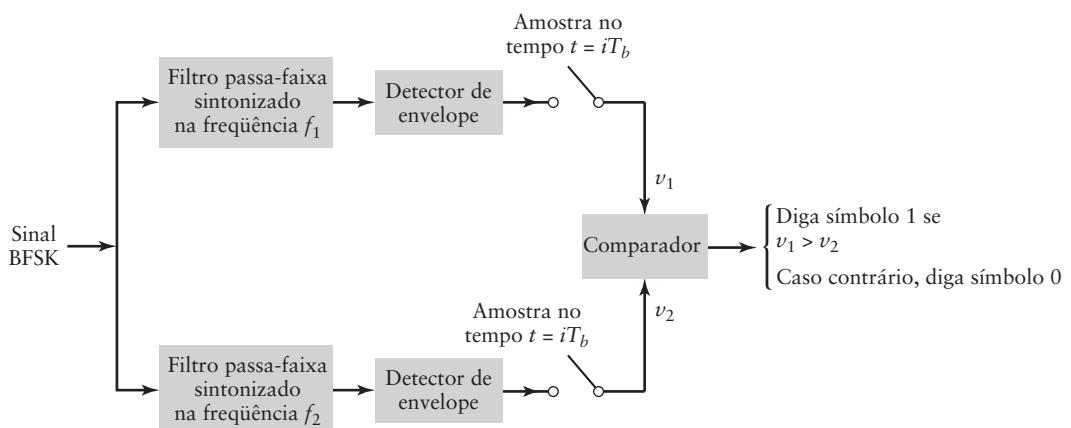


Figura 7.18 Receptor BFSK não coerente. Os dois amostradores operam em sincronismo com $i = 0, \pm 1, \pm 2, \dots$

- ▶ O caminho 1 utiliza um filtro passa-faixa com freqüência de meia faixa f_1 . A versão filtrada do sinal BFSK de entrada é detectada por envelope e , então, amostrada no tempo $t = iT_b$, $i = 0, \pm 1, \pm 2, \dots$, para produzir a saída v_1 .
- ▶ O Caminho 2 utiliza um filtro passa-faixa com freqüência de meia faixa f_2 . Tal como o caminho 1, a versão filtrada do sinal BFSK é detectada por envelope e , então, amostrada no tempo $t = iT_b$, $i = 0, \pm 1, \pm 2, \dots$, para produzir a saída v_2 .

Os dois filtros passa-faixa devem possuir a mesma largura de faixa, igual à largura de faixa de transmissão do sinal BFSK. Além disso, como mencionado ao se trabalhar com BASK, a interferência intersimbólica produzida pelos filtros é considerada como sendo negligenciável.

As saídas dos dois caminhos, v_1 e v_2 , são aplicadas a um *comparador*, no qual decisões sobre a composição do sinal BFSK são repetidas a cada T_b segundos. Novamente, considera-se a disponibilidade de uma temporização de bit no receptor. Reconhecendo que o caminho superior corresponde ao símbolo 1 e o caminho inferior ao símbolo 0, o comparador decide em favor do símbolo 1 se v_1 for maior do que v_2 no instante específico da temporização do bit. Caso contrário, a decisão é feita em favor do símbolo 0. Em um ambiente livre de ruído e sem distorção de canal, a saída do receptor (bit-a-bit) deve ser, novamente, uma réplica da seqüência binária de dados original aplicada à entrada do transmissor.

■ Chaveamento de fase diferencial

Da discussão anterior, vemos que tanto o chaveamento de amplitude quanto o chaveamento de freqüência levam naturalmente à detecção não coerente sempre que for impraticável manter a sincronização de fase da portadora do receptor com o transmissor. Mas este não é o caso para o chaveamento de fase. Não podemos ter uma detecção não coerente no sentido tradicional, pois o termo “não coerente” significa demodular sem a informação da fase da portadora. Para contornarmos esta dificuldade, implementamos uma técnica “pseudo PSK”, chamada de *chaveamento de fase diferencial* (*DPSK*, *differential phase-shift keying*), a qual, em sentido não restrito, permite a utilização de detecção não coerente.

A DPSK elimina a necessidade de um sinal de referência coerente no receptor pela combinação de duas operações básicas no transmissor:

- ▶ *Codificação diferencial* da onda binária de entrada (a qual foi discutida nos códigos de linha da Seção 5.9).
- ▶ *Chaveamento de fase* (a qual foi discutida na Seção 7.3).

É devido a esta combinação que falamos de “chaveamento de fase diferencial”. De fato, para enviar o símbolo 0, adiantamos a fase da forma de onda atual do sinal por 180 graus e para enviar o símbolo 1 deixamos a fase da forma de onda atual do sinal inalterada. Assim sendo, o receptor é equipado com um dispositivo de armazenamento (isto é, memória) projetado para medir a *diferença relativa de fase* entre as formas de onda recebidas durante dois intervalos sucessivos de bit. Desde que a fase θ varie lentamente (isto é, lento o suficiente para que ela seja considerada essencialmente constante dentre dois intervalos de bit), a diferença de fase entre as formas de onda recebidas em dois intervalos sucessivos de bit será, essencialmente, independente de θ .

Geração e detecção de sinais DPSK

(i) Geração

O processo de codificação diferencial na entrada do transmissor começa com um *primeiro bit arbitrário, servindo simplesmente como referência*. Seja $\{d_k\}$ a seqüência diferencialmente codificada com este bit de referência adicionado. Para gerar esta seqüência, o transmissor executa as duas operações a seguir:

- ▶ Se o símbolo b_k binário de entrada for 1, então o símbolo d_k é inalterado com relação ao símbolo anterior d_{k-1} .

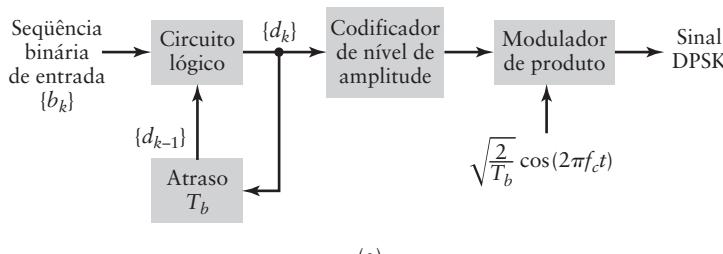
- Se o símbolo b_k de entrada for 0, então o símbolo d_k é alterado com relação ao símbolo anterior d_{k-1} .

A seqüência diferencialmente codificada $\{d_k\}$ gerada desta forma é utilizada para deslocar a fase de uma onda portadora senoidal por ângulos de fase 0 e π radianos, representando os símbolos 1 e 0, respectivamente.

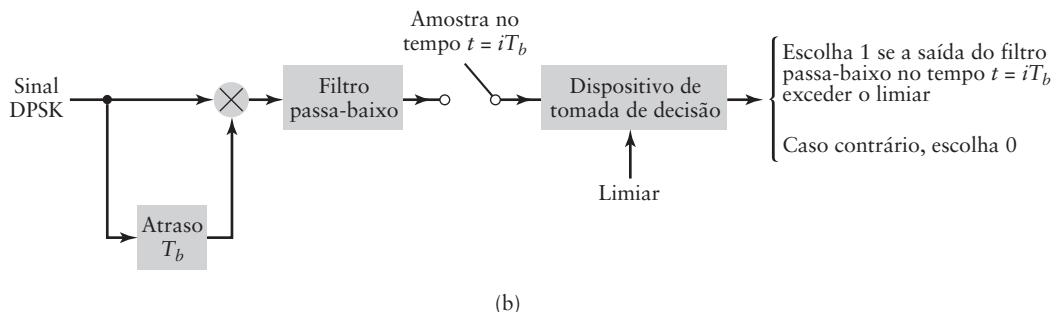
O diagrama em blocos do transmissor DPSK é mostrado na Fig. 7.19(a). Ele consiste, em parte, de um *circuito lógico* e um *elemento de atraso de um bit* (atuando como unidade de memória) interconectados de forma a converter a seqüência binária $\{b_k\}$ de entrada na seqüência diferencialmente codificada $\{d_k\}$. Esta seqüência é codificada em nível de amplitude e, então, utilizada para modular a onda portadora de frequência f_c , produzindo o sinal DPSK desejado.

(ii) Detecção

Para a detecção de sinais DPSK, utilizamos a vantagem do fato dos pulsos modulados em fase, pertencentes a dois bits sucessivos, serem idênticos, exceto por uma possível inversão de sinal. Logo, o pulso de entrada é multiplicado pelo pulso anterior, o que significa que o pulso anterior tem a função de ser um *sinal de referência gerado localmente*. Com esta base, podemos formular o receptor da Fig. 7.19(b) para a detecção de sinais DPSK. Comparando o detector DPSK da Fig. 7.19(b) com o detector de sinais BPSK da Fig. 7.4(b), vemos que as estruturas dos dois receptores são similares, exceto na fonte do sinal de referência gerado localmente. De acordo com a Fig. 7.19(b), o sinal DPSK é detectado, desde que o *bit de referência* seja conhecido, o qual, como mencionado anteriormente, é inserido no começo da seqüência binária de dados de entrada. Em particular, aplicando a saída amostrada do filtro passa-faixas a um dispositivo de tomada de decisão com um *limiar* fornecido, a detecção do sinal DPSK é realizada. Se o limiar for excedido, o receptor decide em favor do símbolo 1, caso contrário, a decisão é em favor do símbolo 0. Novamente, considera-se que o receptor possui uma informação de temporização de bit para que o amostrador possa funcionar corretamente.



(a)



(b)

Figura 7.19 Diagrama de blocos para (a) transmissor DPSK e (b) receptor DPSK; para o amostrador, o inteiro = 0, ±1, ±2,....

EXEMPLO 7.3 Da seqüência binária de dados para o sinal DPSK e o retorno

Começando com a seqüência binária de dados $\{b_k\}$ dada na primeira linha da Tabela 7.3 e utilizando o símbolo 1 como primeiro bit de referência, podemos construir a seqüência $\{d_k\}$ diferencialmente codificada na linha 3 da tabela. A segunda linha é a versão atrasada de $\{d_k\}$, por um bit. Note que para cada índice k , o símbolo d_k é o complemento da soma de módulo 2 de b_k e d_{k-1} . A quarta linha da Tabela 7.3 define a fase (em radianos) do sinal DPSK transmitido.

As duas últimas linhas da Tabela 7.3 pertencem ao receptor DPSK. A linha 5 define a polaridade (positiva ou negativa) da saída do filtro passa-baixa do receptor da Fig. 7.19(b). A linha final define a seqüência binária de dados produzida na saída do receptor, a qual é idêntica à seqüência binária de dados de entrada do topo da tabela, tal como deve ser em um ambiente livre de ruído.

TABELA 7.3 Ilustração da geração e detecção do sinal DPSK

$\{b_k\}$	1	0	0	1	0	0	1	1
$\{d_{k-1}\}$	1	1	0	1	1	0	1	1
Seqüência $\{d_k\}$ diferencialmente codificada	1	1	0	1	1	0	1	1
Fase transmitida (em radianos)	0	0	π	0	0	π	0	0
Saída do amostrador (polaridade)	+	-	-	+	-	-	+	+
Símbolo binário na saída do dispositivo de tomada de decisão	1	0	0	1	0	0	1	1

Nota: O símbolo 1 inserido no começo da seqüência d_k diferencialmente codificada é o bit de referência.

7.7 ESQUEMAS DE MODULAÇÃO DIGITAL M-ÁRIO

Por definição, em um *esquema de modulação digital M-ário*, enviamos qualquer um de M possíveis sinais $s_1(t), s_2(t), \dots, s_M(t)$ durante cada intervalo de sinalização (símbolo) de duração T . Em quase todas as aplicações, $M = 2^m$, onde m é um inteiro. Sob esta condição a duração do símbolo é $T = mT_b$ onde T_b é a duração do bit.

Esquemas de modulação M-ário são preferidos em detrimento de esquemas de modulação binário para a transmissão de dados digitais em canais passa-faixa quando a condição é *conservar largura de faixa ao custo de aumento de potência e complexidade do sistema*. Na prática, raramente encontramos um canal de comunicação que possui a largura de faixa exata necessária para a transmissão da saída de uma fonte de informação através de esquemas de modulação binária. Portanto, quando a largura de faixa do canal é menor do que o valor necessário, utilizamos um esquema de modulação M-ário para a conservação máxima de largura de faixa.

■ Chaveamento de fase M-ário

Para ilustrar a capacidade de esquemas de modulação M-ário para a conservação de largura de faixa, considere, primeiro, a transmissão de informação constituída por uma seqüência binária com duração T_b de bit. Se precisarmos transmitir esta informação por PSK binário, por exemplo, iremos precisar de uma largura de faixa do canal que é inversamente proporcional à duração T_b do bit. Entretanto, se tomarmos blocos de m bits para produzir um símbolo e utilizarmos um esquema PSK M-ário com $M = 2^m$ e duração de símbolo $T = mT_b$, então a largura de faixa necessária é proporcional a $1/(mT_b)$. Este simples argumento mostra que a utilização do PSK M-ário possibilita a redução na largura de faixa de transmissão por um fator $m = \log_2 M$ sobre PSK binário.

No PSK M-ário, a fase disponível de 2π radianos é dividida igualmente, de forma discreta, para todos os M sinais transmitidos, como mostrado pelo sinal modulado em fase

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t + \frac{2\pi}{M} i\right), \quad i = 0, 1, \dots, M-1 \quad 0 \leq t \leq T \quad (7.35)$$

no qual E é a energia por símbolo do sinal e f_c é a freqüência da portadora. Utilizando uma identidade trigonométrica bem conhecida, podemos expandir a Eq. (7.35) para

$$\begin{aligned} s_i(t) &= \left[\sqrt{E} \cos\left(\frac{2\pi}{M} i\right) \right] \left[\sqrt{\frac{2}{T}} \cos(2\pi f_c t) \right] \\ &\quad - \left[\sqrt{E} \sin\left(\frac{2\pi}{M} i\right) \right] \left[\sqrt{\frac{2}{T}} \sin(2\pi f_c t) \right], \quad i = 0, 1, \dots, M-1 \quad 0 \leq t \leq T \end{aligned} \quad (7.36)$$

Os coeficiente *discretos* $\sqrt{E} \cos\left(\frac{2\pi}{M} i\right)$ e $-\sqrt{E} \sin\left(\frac{2\pi}{M} i\right)$ são, respectivamente, relacionados com as *componentes em fase* e em *quadratura* dos sinais $s_i(t)$ PSK M -ário. Percebemos, agora, que

$$\left\{ \left[\sqrt{E} \cos\left(\frac{2\pi}{M} i\right) \right]^2 + \left[\sqrt{E} \sin\left(\frac{2\pi}{M} i\right) \right]^2 \right\}^{1/2} = \sqrt{E}, \quad \text{para todo } i \quad (7.37)$$

Desta forma, a modulação PSK M -ário possui a propriedade única das componentes em fase e em quadratura do sinal modulado $s_i(t)$ serem *inter-relacionados* de tal forma que o envelope discreto do sinal é restrito a manter-se no valor constante de \sqrt{E} para todo M . A estratégia de modulação do QPSK discutida na Seção 7.3 é um exemplo de PSK M -ário com o número de níveis de fase $M = 4$.

Diagrama de espaço de sinal

O resultado descrito na Eq. (7.37), combinado com o fato das componentes em quadratura e fase do PSK M -ário serem discretos, nos leva à representação geométrica do PSK M -ário. Para explicar, suponha que construamos um diagrama bidimensional com os eixos horizontal e vertical, respectivamente, definidos pelo seguinte par de *funções ortonormais*:

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), \quad 0 \leq t \leq T \quad (7.38)$$

e

$$\phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t), \quad 0 \leq t \leq T \quad (7.39)$$

na qual a consideração de passa-faixa implica em ortogonalidade; o fator $\sqrt{\frac{2}{T}}$ garante energia unitária no intervalo T para $\phi_1(t)$ e $\phi_2(t)$. Com esta base, podemos representar as componentes em fase $\sqrt{E} \cos\left(\frac{2\pi}{M} i\right)$ e em quadratura $-\sqrt{E} \sin\left(\frac{2\pi}{M} i\right)$ para $i = 0, 1, \dots, M-1$ como um conjunto de pontos neste diagrama bidimensional, como ilustrado na Fig. 7.20 para $M = 8$. Tal diagrama é chamado de *diagrama de espaço de sinal*.

A Fig. 7.20 nos leva a três importantes observações:

1. PSK M -ário é descrito em termos geométricos por uma *constelação de M pontos de sinais distribuídos uniformemente em um círculo de raio \sqrt{E}* .
2. Cada ponto de sinal na figura corresponde ao sinal $s_i(t)$ da Eq. (7.35) para um valor particular do índice i .
3. O quadrado da distância da origem a cada ponto do sinal é igual à energia E do sinal.

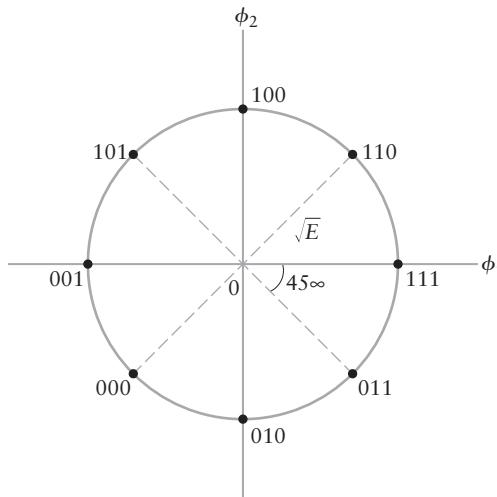


Figura 7.20 Diagrama de espaço de sinal para 8-PSK.

Com estas observações, podemos, agora, afirmar formalmente que o diagrama de espaço de sinal da Fig. 7.20 resume completamente a descrição geométrica do PSK M-ário de forma extremamente útil. Note que seqüências de 3 bits correspondendo aos 8 pontos de sinal estão no *código Gray*, com apenas um bit mudando quando nos movemos ao longo da constelação na figura, de um ponto de sinal para o seu adjacente.

■ Modulação em amplitude em quadratura M-ário

Suponha, a seguir, que a restrição da Eq. (7.37) que caracteriza a modulação PSK M-ário é removida. Então, as componentes em fase e em quadratura do sinal resultante modulado M-ário podem ser *independentes* uma da outra. Especificamente, a descrição matemática desse novo sinal modulado assume a forma

$$s_i(t) = \sqrt{\frac{2E_0}{T}} a_i \cos(2\pi f_c t) - \sqrt{\frac{2E_0}{T}} b_i \sin(2\pi f_c t), \quad i = 0, 1, \dots, M-1 \quad 0 \leq t \leq T \quad (7.40)$$

na qual o parâmetro de nível a_i na componente em fase e o parâmetro de nível b_i na componente em quadratura são *independentes um do outro para todo i*. Este novo esquema de modulação é chamado de *modulação em amplitude em quadratura M-ário* (QAM, *M-ary quadrature amplitude modulation*). Note também que a constante E_0 é a energia do sinal pertencente a um valor particular do índice i para o qual a amplitude do sinal modulado é a menor.

QAM M-ário é uma forma *híbrida* da modulação M-ário, no sentido de que ela combina o chaveamento de amplitude e o chaveamento de fase. Ele inclui dois casos especiais:

- (i) Se $b_i = 0$ para todo i , o sinal modulado $s_i(t)$ da Eq. (7.40) se reduz para

$$s_i(t) = \sqrt{\frac{2E_0}{T}} a_i \cos(2\pi f_c t) \quad i = 0, 1, \dots, M-1$$

a qual define o *chaveamento de amplitude M-ário* (ASK M-ário).

- (ii) Se $E_0 = E$ e a restrição

$$(Ea_i^2 + Eb_i^2)^{1/2} = \sqrt{E}, \quad \text{para todo } i$$

é satisfeita, então o sinal modulado $s_i(t)$ da Eq. (7.40) se reduz para PSK M-ário.

Diagrama de espaço de sinal

A Fig. 7.21 mostra a representação no espaço de sinal do QAM M -ário para $M = 16$, com cada ponto de sinal definido por um par de parâmetros de nível a_i e b_i , com $i = 1, 2, 3, 4$. Desta vez, vemos que os pontos de sinal estão distribuídos uniformemente em uma malha retangular. A propriedade retangular do diagrama de espaço de sinal é um testemunho do fato das componentes em fase e quadratura do QAM M -ário serem independentes uma da outra. Além disso, vemos da Fig. 7.21 que, ao contrário do PSK M -ário, pontos de sinal diferentes do QAM M -ário são caracterizados por diferentes níveis de energia, tal como eles devem ser. Note também que cada ponto de sinal na constelação corresponde a um *tetra-bit* específico, o qual é constituído por 4 bits. Assumindo a utilização do código de Gray, apenas um bit é alterado quando nos movemos de um ponto de sinal da constelação, horizontalmente ou verticalmente, para um ponto adjacente, como ilustrado na Fig. 7.21.

■ Chaveamento de freqüência M -ário

Entretanto, quando consideramos a versão M -ário do chaveamento de freqüência, a figura é um pouco diferente da que descreve PSK M -ário ou QAM M -ário. Especificamente, em uma forma de FSK M -ário, os sinais transmitidos são definidos para algum inteiro n fixo, como mostrado a seguir:

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos\left[\frac{\pi}{T}(n + i)t\right], \quad i = 0, 1, \dots, M - 1 \quad (7.41)$$

Todos os sinais M transmitidos possuem duração T igual e energia E igual. Com as freqüências individuais dos sinais separadas uma da outra por $1/(2T)$ hertz, os sinais da Eq. (7.41) são *ortogonais*, isto é, eles satisfazem a condição

$$\int_0^T s_i(t)s_j(t) dt = \begin{cases} E & \text{para } i = j \\ 0 & \text{para } i \neq j \end{cases} \quad (7.42)$$

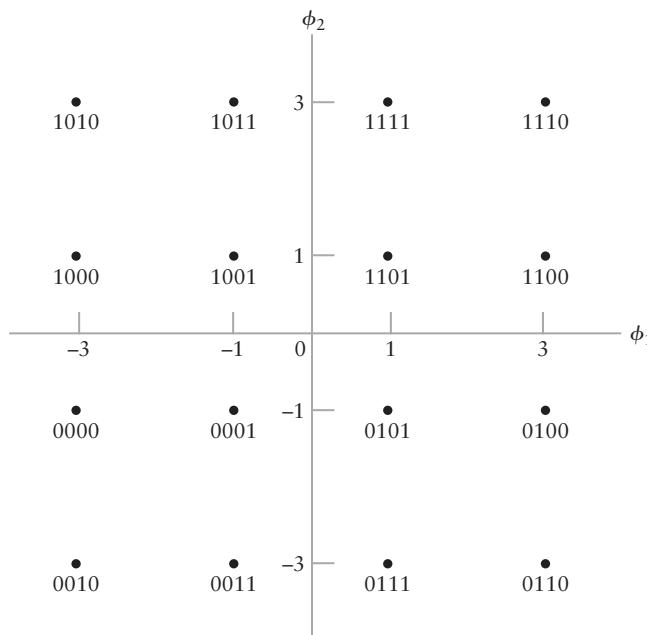


Figura 7.21 Diagrama de espaço de sinal para um QAM M -ário com código Gray para $M = 16$.

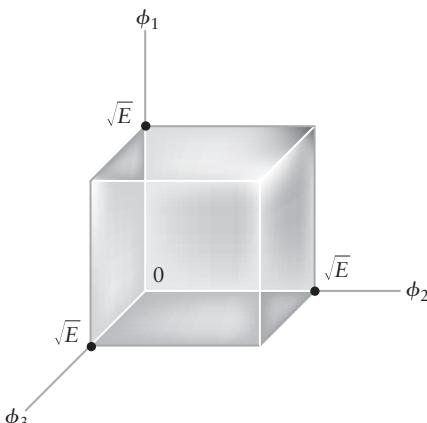


Figura 7.22 Constelação de sinal para FSK M -ário com $M = 3$.

Tal como o PSK M -ário, o envelope do FSK M -ário é constante para todo M , como mostrado na Eq. (7.41). Logo, as duas estratégias de modulação M -ário podem ser utilizadas em canais não lineares. Por outro lado, o QAM M -ário só pode ser utilizado em canais lineares, pois seu envelope varia com o índice i (isto é, o ponto de sinal particular escolhido para a transmissão).

Diagrama de espaço de sinal

Para desenvolver uma representação geométrica do FSK M -ário, começamos pela Eq. (7.41). Em termos dos sinais $s_i(t)$ definidos por ela, introduzimos um conjunto completo de funções ortonormais:

$$\phi_i(t) = \frac{1}{\sqrt{E}} s_i(t) \quad i = 0, 1, \dots, M - 1 \quad 0 \leq t \leq T \quad (7.43)$$

Ao contrário do PSK M -ário e QAM M -ário, observamos agora que o FSK M -ário é descrito por um *diagrama de espaço de sinal M -dimensional*, no qual o número de pontos de sinal é igual ao número de coordenadas. A visualização de tal diagrama é difícil além de $M = 3$. A Fig. 7.22 ilustra a representação geométrica de um FSK M -ário para $M = 3$.

- **PROBLEMA 7.9** Começando pela Eq. (7.41), prove a propriedade de ortogonalidade da Eq. (7.42) que caracteriza o FSK M -ário. ◀

7.8 MAPEAMENTO DE FORMAS DE ONDA DIGITALMENTE MODULADAS EM CONSTELAÇÕES DE PONTOS DE SINAIS

A idéia de diagramas de espaço de sinal mencionada na Seção 7.7 é de grande importância na teoria estatística de comunicações. Em particular, ela fornece a base matemática para a *representação geométrica de sinais de energia*, exemplificado pelas formas de onda digitalmente moduladas. Para um método específico de modulação, a representação geométrica é mostrada na forma de uma constelação de pontos no diagrama de espaço de sinal, a qual é única para aquele método.

O propósito desta seção é fazer duas coisas:

- Consolidar a idéia de diagrama de espaço de sinal.
- Discutir o que esta idéia nos ensina na análise de ruído em sistemas de comunicação digital, o qual será analisado posteriormente.

Com esta consolidação em mente, a Fig. 7.23 sobre BPSK mostra a forma na qual as duas formas de onda $s_1(t)$ e $s_2(t)$, respectivamente representando os símbolos binários 1 e 0,

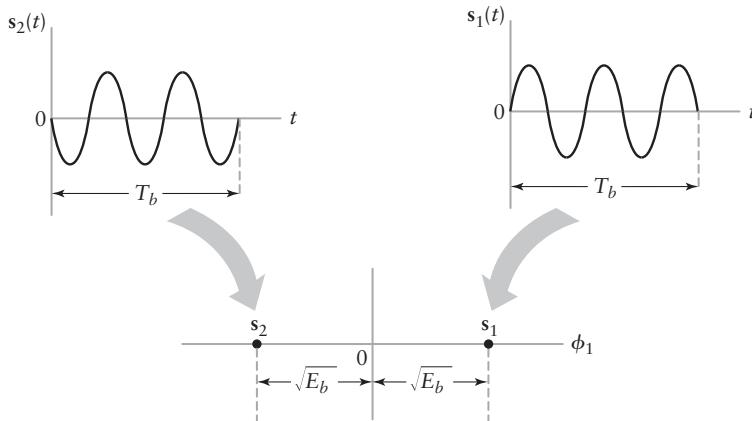


Figura 7.23 Mapeamento de um sinal BPSK em um diagrama de espaço de sinal unidimensional.

são mapeadas nos *pontos de sinal transmitidos* s_1 e s_2 . A questão é: Como o mapeamento é realizado? A resposta está na utilização de um *correlator*. Primeiro, observamos que a representação por espaço de sinal do BPSK é simples, envolvendo uma única função base:

$$\phi_1(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_c t) \quad (7.44)$$

Em termos mais específicos, cada uma das duas formas de onda de sinalização $s_1(t)$ e $s_2(t)$ é correlacionada com a função fase $\phi_1(t)$. A utilização da correlação fornece outra forma de projetar um receptor para a detecção coerente do BPSK. Para a questão em estudo, procedemos em dois passos, como mostrado a seguir:

1. Correlacionando o sinal

$$s_1(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t) \quad \text{para o símbolo 1} \quad (7.45)$$

com a função fase $\phi_1(t)$ no intervalo de tempo $0 \leq t \leq T_b$, obtemos o ponto de sinal

$$\begin{aligned} s_1 &= \int_0^{T_b} \phi_1(t) s_1(t) dt \\ &= \int_0^{T_b} \frac{2}{T_b} \sqrt{E_b} \cos^2(2\pi f_c t) dt \end{aligned} \quad (7.46)$$

Com a consideração de passa-faixa, a Eq. (7.46) é simplificada para

$$s_1 = \sqrt{E_b} \quad (7.47)$$

2. Similarmente, podemos mostrar que o sinal

$$s_2(t) = -\sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t) \quad \text{para o símbolo 0} \quad (7.48)$$

é representado pelo ponto de sinal

$$s_2 = -\sqrt{E_b} \quad (7.49)$$

Estes dois resultados são, de fato, os mostrados na Fig. 7.23.

- **PROBLEMA 7.10** Justifique as Eqs. (7.47) e (7.49). ◀

Considere a Fig. 7.24, representando o BFSK. Tal como o BPSK, o diagrama de espaço de sinal consiste de dois *pontos de sinal transmitidos*: o ponto s_1 , representando a transmissão do símbolo binário 1, e o ponto s_2 , representando a transmissão do símbolo binário 0. Seguindo um procedimento similar ao utilizado para o mapeamento do espaço de sinal do BPSK, podemos construir o mapeamento no espaço de sinal da Fig. 7.24 para BFSK no qual os dois pontos de sinal são definidos por

$$s_1 = \begin{bmatrix} \sqrt{E_b} \\ 0 \end{bmatrix} \quad (7.50)$$

e

$$s_2 = \begin{bmatrix} 0 \\ \sqrt{E_b} \end{bmatrix} \quad (7.51)$$

Os dois diagramas de espaço de sinal das Figs. 7.23 e 7.24 diferem em um importante ponto: *dimensionalidade*, a qual é definida pelo número de funções base. O diagrama de espaço de sinal do BPSK da Fig. 7.23 é unidimensional, pois ele utiliza uma única função base $\phi_1(t)$ da Eq. (7.44). Por outro lado, o diagrama de espaço de sinal da Fig. 7.24 para o BFSK é bidimensional, pois ele envolve o par de funções base

$$\phi_1(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_1 t) \quad (7.52)$$

e

$$\phi_2(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_2 t) \quad (7.53)$$

O que podemos aprender desses dois diagramas quando considerarmos o efeito de ruído aditivo de canal no BPSK e BFSK? Para responder a esta questão, precisamos notar primeiro que a separação entre os pontos de sinal para o BPSK é $2\sqrt{E_b}$, enquanto que para o BFSK a separação é $\sqrt{2E_b}$, assumindo que a energia por bit, E_b , e a duração do bit, T_b ,

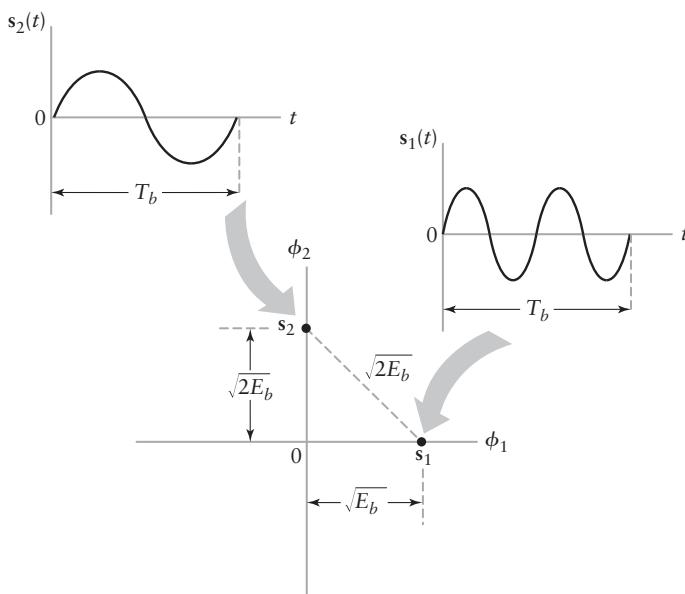


Figura 7.24 Mapeamento do BFSK em um diagrama de espaço de sinal bidimensional.

são as mesmas para os dois esquemas. Em outras palavras, podemos agora obter nossa primeira afirmativa:

1. A separação entre os pontos de sinal transmitidos para o BPSK é $\sqrt{2}$ vezes maior do que para o BFSK.

Quanto à presença de ruído aditivo de canal, ele possui o efeito de fazer com que o ponto de sinal recebido varie aleatoriamente ao redor do ponto do sinal transmitido. Desta forma, podemos obter nossa segunda afirmativa:

2. O ponto de sinal recebido está dentro de uma “nuvem” centrada no ponto de sinal transmitido.

À luz das afirmativas 1 e 2, estamos prontos para dizer que sob a consideração de caracterização estatística idêntica para o ruído aditivo de canal (e, portanto, uma nuvem idêntica de incerteza), tanto no esquema BPSK quanto no BFSK, o ruído degrada mais a performance do BFSK do que o BPSK. A validade desta observação é confirmada no Capítulo 10, no qual ela é matematicamente analisada.

7.9 EXEMPLOS DE TEMAS

Multiplexação ortogonal por divisão em freqüência

Na Seção 2.11, discutimos a resposta em freqüência de um par trançado utilizado para conectar casas à central telefônica de chaveamento. Verificamos que a resposta em freqüência para cabos curtos é bastante plana na faixa de voz, mas deteriora quando o comprimento do cabo aumenta. Para nosso primeiro exemplo de tema, iremos considerar como podemos transmitir dados em alta velocidade utilizando o par trançado de casa. Este serviço é geralmente chamado de *digital subscriber line (DSL)*.

Considere a mídia de transmissão que possui resposta em freqüência arbitrária mostrada na Fig. 7.25(a). É evidente que a mídia irá alterar o espectro transmitido da Fig. 7.25(b) e, portanto, irá causar uma interferência intersímbólica significativa em quase todas as técnicas de modulação discutidas neste capítulo. Estamos nos referindo a técnicas de modulação que possuem características similares à Fig. 7.25(b) como *sistemas de portadora única*.

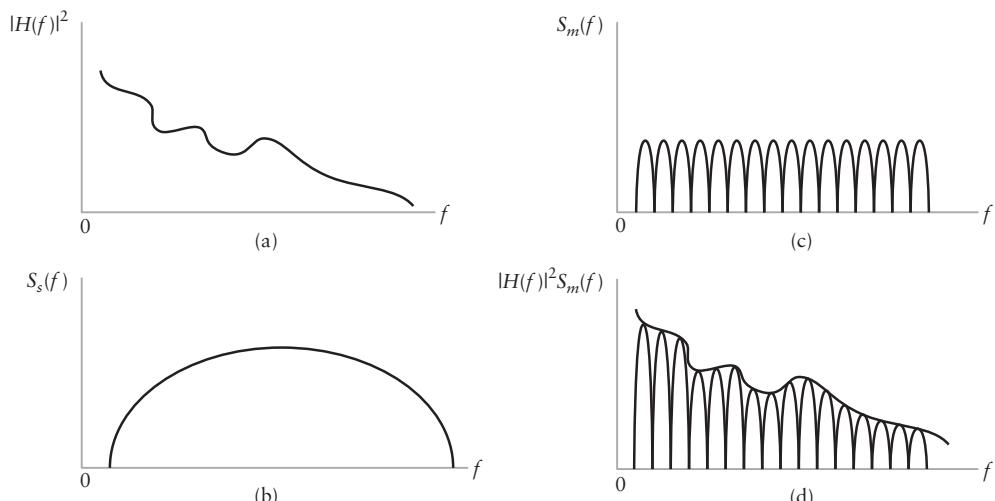


Figura 7.25 (a) Resposta em freqüência quadrática do canal. (b) Densidade espectral de potência da modulação de portadora única. (c) Densidade espectral de potência transmitida de uma modulação de múltiplas portadoras. (d) Densidade espectral de potência recebida da modulação de múltiplas portadoras.

Considere, a seguir, o *sistema de múltiplas portadoras*, mostrado na Fig. 7.25(c), na qual um único transmissor implementa múltiplas portadoras. Cada uma responsável por uma fração da taxa total de dados. Com a transmissão de múltiplas portadoras, a mídia terá o mesmo efeito no espectro total, mas o espectro das portadoras individuais é relativamente não distorcido exceto por um possível ganho e rotação de fase, como mostrado na Fig. 7.25(d).

Vamos, agora, considerar como podemos implementar tal sistema de múltiplas portadoras no qual as componentes individuais são chamadas de *subportadoras*. Como todas as subportadoras são geradas pela mesma fonte, assumimos que elas podem ser sincronizadas. Primeiro, considere um único período de símbolo e sejam N as subportadoras com freqüências $\{f_n\}$ e $\{b_n\}$ bits transportando informação. O n -ésimo sinal de subportadora pode, então, ser representado por

$$s_n(t) = b_n \exp(j2\pi f_n t), \quad 0 \leq t \leq T \quad (7.54)$$

Se b_n for ± 1 , então a Eq. (7.54) é simplesmente uma representação complexa banda base do BPSK na freqüência f_n da subportadora. Logo, o sinal de múltiplas portadoras em um período de símbolo pode ser representado como a soma da Eq. (7.54) para todas as subportadoras, como mostrado por

$$\begin{aligned} s(t) &= \sum_{n=0}^{N-1} s_n(t) \\ &= \sum_{n=0}^{N-1} b_n \exp(j2\pi f_n t) \quad 0 \leq t \leq T \end{aligned} \quad (7.55)$$

Agora engenhosamente tomamos as seguintes decisões:

- Amostramos este sinal de múltiplas portadoras em intervalos de T_s , no qual $T_s = T/N$.
- Escolhemos as freqüências das subportadoras para serem espaçadas em intervalos de $f_n = n/T$. Então, a Eq. (7.55), calculada para $t = kT_s$ se torna

$$s(kT_s) = \sum_{n=0}^{N-1} b_n \exp\left(j2\pi \frac{kn}{N}\right) \quad (7.56)$$

Comparando o que foi apresentado aqui com o material da Seção 2.10, vemos que o sinal amostrado de múltiplas portadoras é simplesmente a transformada discreta inversa de Fourier da seqüência de dados $\{b_n\}$. Conseqüentemente, podemos representar este sistema de transmissão de dados como mostrado na Fig. 7.26, na qual o transmissor inclui uma conversão série para paralelo da seqüência de dados seguido pelo sistema de transformada discreta inversa de Fourier. O receptor executa o processamento inverso do transmissor. Se considerarmos adicionalmente que N é uma potência de dois, então o sistema pode ser implementado eficientemente pelo algoritmo da transformada rápida de Fourier (FFT), o qual também foi discutido na Seção 2.10. Na Fig. 7.26, TDF é o acrônimo para transformada discreta de Fourier.

A decisão sobre os intervalos de amostragem e freqüência não somente facilita a implementação, mas também possui os seguintes benefícios práticos:

- Os canais adjacentes são espaçados em intervalos $1/T_s$. Esta escolha implica que existe uma significativa sobreposição espectral, e, portanto, o esquema de modulação minimiza o total de largura de faixa necessária.
- O fato das freqüências das subportadoras serem múltiplas do período do símbolo significa que elas são ortogonais³ em um período de símbolo. Conseqüentemente, não existe interferência entre as subportadoras mesmo que a sobreposição seja significativa.

³ A ortogonalidade significa que a condição

$$\int_0^T s_n(t)s_k^*(t) dt = 0$$

é válida para $n \neq k$.

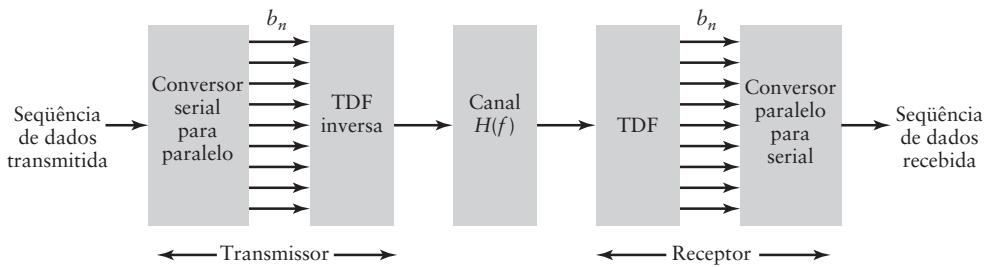


Figura 7.26 Diagrama de blocos do sistema OFDM (*orthogonal Frequency-division Multiplexing*).

- ▶ As freqüências das subportadoras são contínuas de um intervalo de símbolo para o próximo, significando que não existe distorção das propriedades espetrais quando transmitimos uma seqüência contínua de dados.
- ▶ Os símbolos podem ser continuamente estendidos pela simples repetição chamada de *prefix* ou *postfix cíclico*. Esta característica pode ser benificamente utilizada em canais que tendem a alargar o intervalo do pulso.

A combinação da multiplexação das subportadoras no domínio da freqüência e a ortogonalidade destas subportadoras dão o nome *multiplexação ortogonal por divisão de freqüência* (OFDM, *orthogonal Frequency-division Multiplexing*) para esta técnica de modulação de múltiplas portadoras.⁴

■ Comunicação GSM (*Global System for Mobile*)

No Capítulo 4, discutimos a idéia de modulação em freqüência baseada na definição

$$s(t) = A_c \cos\left(2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau\right) \quad (7.57)$$

na qual o sinal de mensagem $m(t)$ é um sinal analógico. Neste segundo exemplo de tema, iremos considerar a modulação em ângulo que é produzida pelo sinal de mensagem definido por

$$m(t) = \sum_{n=0}^N b_n p(t - nT) \quad (7.58)$$

na qual T é o período do símbolo, b_n é o símbolo do dado no n -ésimo período de símbolo e $p(t)$ é a forma do pulso aplicada a cada período de símbolo. No curso histórico dos eventos, a modulação analógica era a principal forma de modulação utilizada para comunicações via rádio por vários anos antes do advento das comunicações digitais. Conseqüentemente, nas comunicações digitais iniciais, era geralmente mais conveniente utilizar um transmissor FM, por isso o termo *FM digital*. Sistemas FM digitais possuem várias vantagens, não só a eficiência de potência devido ao seu envelope constante.

A eficiência de potência desejada tornou o FM digital um forte candidato para aplicações sem fio utilizando terminais pessoais (*handhelds*) com pequena capacidade de bateria. De fato, uma forma de FM digital foi escolhida como padrão europeu em 1991 e espalhou-se posteriormente em outras partes do mundo sob o nome GSM (*Global System for Mobile communications*).

⁴ Para uma discussão mais detalhada sobre OFDM, veja Bahai e Saltzberg (1999).

A modulação atual selecionada para o GSM é chamada de *mínimo chaveamento Gaussiano* (GMSK, *Gaussian minimum shift keying*), o qual é uma combinação especial das Eqs. (7.57) e (7.58). Em particular, introduzimos três condições:

1. A forma do pulso é escolhida como tendo a forma Gaussiana

$$p(t) = c \exp\{-\pi c^2 t^2\} \quad (7.59)$$

na qual o fator de escala é $c = B\sqrt{2\pi/\log(2)}$ e o logaritmo é o natural.

2. O fator B controla a largura de faixa do sinal, a qual, para um preestabelecido período T de símbolo, é escolhida tal que o produto tempo-largura de faixa $B_T = 0,3$.
3. O fator de sensibilidade de freqüência k_f do modulador é selecionado tal que a mudança de fase em um período de símbolo é $\pi/2$ radianos.

Com estas três condições, o GMSK possui um espectro compacto e pode ser coerentemente detectado com apenas 0,46 dB de degradação relativa à performance frente ao ruído do BPSK ideal.

A comunicação GSM (implementando o GMSK) pode, portanto, ser vista como um exemplo prático da evolução do sistema FM analógico para o sistema digital, com a tentativa de se manter as vantagens do primeiro.⁵

■ Televisão digital

Para o nosso terceiro e último exemplo de tema, considere a *televisão digital*. Existe um movimento mundial para mudança da televisão analógica para a digital, com o objetivo de fornecer vídeo e áudio de alta qualidade. Na América do Norte, o padrão de televisão analógico NTSC⁶ está sendo substituído pelo padrão *digital Advanced Television* (ATV, ou televisão digital avançada). Por questão de compatibilidade, o novo padrão ATV é projetado para caber na largura de faixa de 6 MHz utilizada no sistema NTSC atual. O sistema ATV fornece dois níveis de serviço:

- (i) 19 Mb/s de saída em um canal de broadcast de 6MHz para rádio terrestre.
- (ii) 28 Mb/s de saída em um canal de 6MHz utilizado para sistemas a cabo.

Similar à televisão analógica, a televisão digital é um sistema complexo. O diagrama em blocos da Fig. 7.27 é uma dica da complexidade envolvida. Vamos considerar os diferentes componentes mostrados na figura.

De importância primária estão as *fontes de áudio e vídeo*, sem as quais não existiria a necessidade do restante. Atualmente, estas fontes podem assumir uma dentre várias formas, seja analógica ou digital. Portanto, o transmissor deve suportar uma variedade de interfaces. O dado auxiliar pode incluir dados de controle e dados associados com o programa de serviço de áudio e vídeo, tal como *closed captioning*.

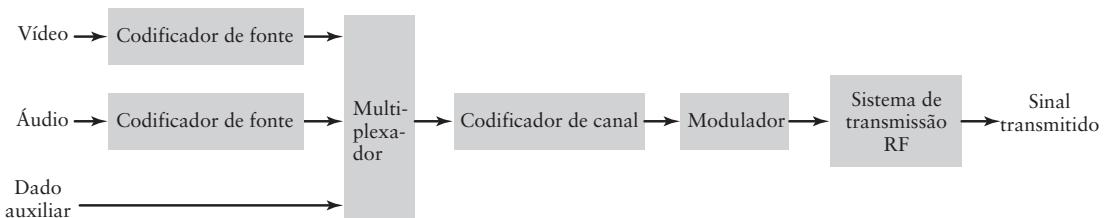


Figura 7.27 Modelo de broadcasting da televisão digital.

⁵ Para uma discussão detalhada do GMSK, veja Steele e Hanzo (1999).

⁶ NTSC é a sigla para *National Television System Committee*.

A primeira operação na fonte de dados é a *codificação de fonte*. Esta operação refere-se ao processamento aplicado às seqüências de dados de áudio e vídeo para eliminar *redundâncias* e minimizar a taxa de bit necessária para a transmissão. Ela também é chamada de *compressão de dados*. O total de compressão possível depende tanto da dinâmica do sinal de vídeo ou áudio quanto do total de distorção permitido. Em geral, o objetivo é tornar qualquer distorção imperceptível ao usuário humano. Como o sistema é projetado para transportar vídeo de alta definição, atingir este objetivo utilizando a largura de faixa disponível requer algoritmos complexos para a compressão de vídeo e áudio.

O sistema ATV implementa vários algoritmos de compressão e codificação de fonte. Para o vídeo, o algoritmo definido no padrão MPEG-2 é o utilizado. Este algoritmo de compressão é capaz de codificar a definição padrão de televisão a taxas de bits de quatro a nove megabits por segundo e televisão de alta definição de 15 a 25 megabits por segundo.

Para a codificação da fonte de áudio, um algoritmo do padrão Dolby AC-3 é utilizado. Este mesmo algoritmo é utilizado em DVDs. Para um sistema de canal 5.1, este segundo algoritmo pode obter taxas de compressão na faixa de dez para um para áudio de alta qualidade.

O *multiplexador* da Fig. 7.27 implementa três funções específicas:

1. Divisão da seqüência de dados digitais em “pacotes” de informação.
2. Fornecimento de um identificador único para cada pacote ou tipo de pacote.
3. Multiplexação dos pacotes de vídeo, áudio e dados auxiliares em uma única seqüência de dados.

Dada as limitações de largura de faixa, esta multiplexação deve ser feita de forma altamente eficiente.

Os últimos três blocos da Fig. 7.27 correspondem à camada física do sistema de transmissão digital. O codificador de canal pega a seqüência binária de dados e adiciona informação adicional de codificação de canal, a qual, por sua vez, é utilizada pelo receptor para corrigir erros que podem ocorrer devido a defeitos na transmissão (veja a Seção 10.8 sobre detecção e correção de erro). O modulador utiliza a informação da seqüência binária de dados para modular a portadora transmitida. O subsistema de modulação oferece dois modos de transmissão, correspondendo aos dois níveis de serviço:

1. *Modo de broadcast terrestre*, chamado de 8-VSB, utiliza um sinal PAM 8-ário com modulação por faixa lateral vestigial para transmitir 18 megabis/s.
2. *Modo de alta taxa de dados*, para transmissão via cabo, chamado de 16-VSB. Utiliza um sinal PAM 16-ário com modulação por faixa lateral vestigial para transmitir 38 megabits/s.

Estudamos a modulação por faixa lateral vestigial (VSB) aplicada a sinais analógicos no Capítulo 3. Isto levanta uma questão: como a VSB é utilizada para transportar um sinal digital?

Lembre-se que, na banda base complexa, a resposta em freqüência do filtro formador de baixa lateral em um modulador VSB se assemelha ao mostrado na Fig. 7.28(a); veja o Problema 3.32. O modulador passa a maior parte das freqüências positivas sem distorção e bloqueia a maior parte das componentes de freqüência negativa do sinal. Na região próxima à freqüência zero, a resposta em freqüência muda de 1 para 0, deixando um vestígio das componentes de freqüência negativa do sinal. Para analisar um modulador VSB digital, note que esta resposta em freqüência pode ser expressa como a soma de duas respostas em freqüências, representadas por $H_I(f)$ e $H_Q(f)$, como mostrado nas Figs. 7.28(b) e 7.28(c), respectivamente. A primeira resposta em freqüência, $H_I(f)$, corresponde ao filtro passa-tudo que passa o sinal sem distorção. Como $H_I(f)$ é uma função par da freqüência, ela possui uma resposta ao impulso puramente real, $h_I(t)$. A segunda resposta em freqüência, $H_Q(f)$, corresponde a um filtro que passa a maior parte das freqüências positivas sem distorção, mas inverte a fase das componentes de freqüência negativa. Como $H_Q(f)$ é uma função ímpar da freqüência, ela possui uma resposta ao impulso puramente imaginária $j h_Q(t)$.

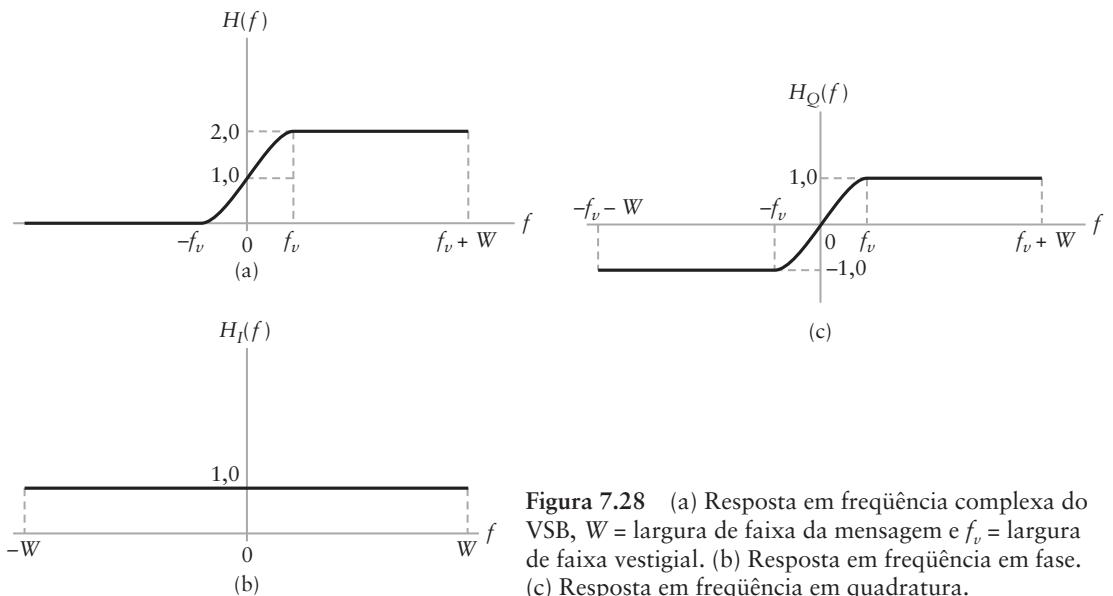


Figura 7.28 (a) Resposta em freqüência complexa do VSB, W = largura de faixa da mensagem e f_v = largura de faixa vestigial. (b) Resposta em freqüência em fase. (c) Resposta em freqüência em quadratura.

Seja a entrada do modulador VSB um sinal digital real definido por

$$m(t) = \sum_{n=0}^N a_n p(t - nT) \quad (7.60)$$

na qual a_n é um dentre $\pm 1, \pm 3, \pm 5$ ou ± 7 , para o sinal PAM 8-ário. A forma do pulso $p(t)$ corresponde a forma espectral da raiz do cosseno levantado. Esta forma de formato de pulso foi discutida na Seção 6.4. A representação banda base complexa da saída do modulador VSB é, portanto,

$$\tilde{s}(t) = (h_I(t) + j h_Q(t)) \star m(t) \quad (7.61)$$

na qual o símbolo \star representa a convolução. Como $h_I(t)$ é um filtro passa-tudo, temos que $h_I(t) \star m(t) = m(t)$, e seja $m_Q(t) = h_Q(t) \star m(t)$. Podemos, portanto, reescrever a Eq. (7.61) para

$$\tilde{s}(t) = m(t) + j m_Q(t) \quad (7.62)$$

O sinal passa-faixa transmitido é, portanto, dado por

$$s(t) = \operatorname{Re}[\tilde{s}(t) \exp(j2\pi f_c t)] \quad (7.63)$$

na qual f_c é a freqüência da portadora. No receptor, utilizamos um detector coerente para recuperar o sinal multiplicando $s(t)$ da Eq. (7.63) por $\cos(2\pi f_c t)$ (isto é, uma réplica da portadora gerada localmente) para obter (na ausência de ruído de canal)

$$\begin{aligned} x(t) &= s(t) \cos(2\pi f_c t) \\ &= \{m(t) \cos(2\pi f_c t) - m_Q(t) \sin(2\pi f_c t)\} \cos(2\pi f_c t) \end{aligned} \quad (7.64)$$

Utilizando identidades trigonométricas para ângulo dobrado, a Eq. (7.64) se reduz para

$$x(t) = \frac{1}{2}m(t)[1 + \cos(4\pi f_c t)] - \frac{1}{2}m_Q(t) \sin(4\pi f_c t) \quad (7.65)$$

a qual consiste de um termo banda base e dois termos de alta freqüência centrados em $2f_c$. Se filtrarmos passa-baixa o sinal $x(t)$, ficaremos com $\frac{1}{2}m(t)$, o que, exceto pelo fator de escala de 1/2, é o sinal de mensagem original.⁷

A abordagem descrita aqui pode ser aplicada a qualquer sinal e não depende do nível da modulação digital. Portanto, ela pode ser aplicada igualmente bem para o PAM 8-ário e 16-ário. Além disso, ela simplifica o receptor pois ele precisa processar apenas o canal real.

A taxa de símbolo ($1/T$) do ATV seja no VSB-8 ou VSB-16 é 10,76 megasímbolos por segundo. Como o VSB-8, existem três bits por símbolo e a taxa de bit do canal é três vezes a taxa do símbolo. Overhead e bits de redundância para a correção de erro posterior reduzem a informação do usuário para, aproximadamente, 19 megabits por segundo. A taxa de bit é duas vezes este valor para o VSB-16. Com a modulação VSB, a largura de faixa do canal é aproximadamente metade da taxa de símbolo, ou 5,38MHz. Portanto, o sinal modulado VSB cabe muito bem no canal de 6 MHz com alguma margem para a calda do filtro e erro de freqüência. Deve ser observado que o sistema ATV também inclui um tom piloto (um nível CC na banda base complexa). Este tom piloto é ajustado para estar 11,3 dB abaixo do nível do sinal, utilizado pelo receptor para a recuperação coerente do sinal original contendo a informação.

7.10 RESUMO E DISCUSSÃO

Neste capítulo, discutimos alguns pontos importantes de esquemas de modulação passa-faixa, resumidas aqui:

- (i) *Chaveamento de amplitude (ASK)*, com ênfase em sinalização binária.
- (ii) *Chaveamento em fase (PSK) com ênfase nas formas binária e quaternária de sinalização* (a sinalização quaternária trabalha com pares de bits, enquanto, obviamente, a sinalização binária trabalha com bits individuais). Na sinalização quaternária discutimos o chaveamento em quadratura de fase (QPSK) e sua variante chamada de chaveamento em quadratura de fase deslocada (OQPSK). A vantagem do OQPSK sobre o QPSK é que ele é menos sensível a não linearidades.
- (iii) *Chaveamento de freqüência (FSK)* com ênfase em sinalização binária. Também discutimos a modificação da FSK binária, chamada de mínimo chaveamento (MSK), a qual é um exemplo de chaveamento de freqüência com fase contínua (CPFSK). Uma importante propriedade da MSK é que ela utiliza o espaçamento de freqüência mínimo, o qual, por sua vez, permite que dois sinais FSK, representando os símbolos 0 e 1, sejam coerentemente ortogonais de tal forma a não interferirem um com o outro no processo de detecção coerente no receptor. Para sua geração, o MSK segue um procedimento similar ao OQPSK no que concerne a extração de seqüências deslocadas demultiplexadas da seqüência binária de dados de entrada. Eles diferem um do outro com relação a um simples e importante ponto: no OQPSK, estas duas ondas binárias auxiliares são ponderadas pela função retangular díbit-a-díbit. Enquanto que no MSK, a ponderação é executada pelo semicírculo positivo de uma cossenoide.
- (iv) *Detecção coerente*, o que significa que independente do esquema de modulação em estudo, o receptor deve estar sincronizado com o transmissor em dois pontos: fase da portadora e temporização do bit.
- (v) *Detecção não coerente*, a qual se refere ao esquema pelo qual o receptor ignora o conhecimento da fase da portadora entre a sua própria operação e a do transmissor. A implementação da detecção não coerente é direta para o chaveamento de amplitude e

⁷ É interessante notar que podemos utilizar a forma de pulso de raiz de cosseno levantado para a filtragem passa-baixa de $x(t)$ no receptor, o qual combina com a forma do pulso transmitido de forma a fornecer uma performance livre de interferência intersimbólica.

para o chaveamento de freqüência. Entretanto, a forma não coerente do chaveamento de fase requer uma atenção mais detalhada. Especificamente, uma pseudoforma de detecção não coerente é desenvolvida pela combinação de codificação diferencial e chaveamento de fase, a qual resulta no chaveamento de fase diferencial (DPSK). Note, entretanto, que todos os receptores não coerentes, independente de sua origem, requerem o conhecimento da informação de temporização de bit.

- (vi) *Análise espectral*, a qual foi demonstrada através de experimentos executados em exemplos de sinais *passa-faixa* representativos de diferentes estratégias de modulação. Os experimentos possuem valor ilustrativo em virtude do fato da taxa de bit ser uma pequena fração da freqüência da portadora. Na prática, entretanto, geralmente vemos que a taxa de bit é algumas ordens de magnitude menor do que a freqüência da portadora. Em tais situações, o procedimento recomendado é executar a análise espectral em *banda base*, na qual (sem perda de informação essencial) a presença da onda portadora é removida representando-se o sinal modulado em termos de suas componentes em fase e em quadratura. O leitor deve referir-se ao Problema 7.30 para uma repetição da análise espectral, desta vez utilizando a abordagem banda base.
- (vii) *Esquemas de sinalização M-ário*, cuja utilização é preferida frente a esquemas de modulação binária para a transmissão de dados digitais em canais passa-faixa quando a necessidade de conservação de largura de faixa do canal é de grande importância. Tanto na sinalização PSK M-ário quanto na FSK M-ário, o envelope do sinal transmitido é constante, permitindo, pois, a utilização de canais não lineares. Em contraste, no QAM M-ário, o envelope do sinal transmitido é variável, o que, portanto, limita sua utilização em canais passa-faixa.

Para os exemplos de temas, consideramos três sistemas digitais de grande uso em suas formas individuais:

- *Multiplexação ortogonal por divisão de freqüência* (OFDM), a qual fornece a base para a transmissão de dados de alta velocidade usando pares trançados que conectam casas a centrais de chaveamento de telefonia.
- *Sistema GSM (Global System for Mobile communication)*, o qual implementa uma modificação do MSK chamada de mínimo chaveamento Gaussiano (GMSK).
- *Televisão digital*, com o objetivo de prover vídeo e áudio de alta qualidade utilizando parâmetros de canal similares ao da televisão analógica.

Um último comentário deve ser feito. Ao longo da discussão sobre receptores neste capítulo, sejam eles do tipo coerente ou não coerente, ignoramos uma realidade prática: a inevitável presença de ruído aditivo de canal. Inevitavelmente, o sinal transmitido é perturbado por esta realidade, o que, por sua vez, possui o efeito de produzir erros na seqüência binária de dados estimada na saída do receptor. O efeito do ruído de canal é discutido no Capítulo 10, mas precisamos estudar, primeiro, a caracterização de sinais e ruído aleatórios, o que será feito no próximo capítulo.

Problemas adicionais

- 7.11** A seqüência binária 11100101 é aplicada a um modulador ASK. A duração do bit é $1 \mu\text{s}$ e a onda senoidal portadora utilizada para representar o símbolo 1 possui freqüência igual a 7 MHz.
- Obtenha a largura de faixa de transmissão do sinal transmitido.
 - Trace a forma de onda do sinal ASK transmitido.
- Assuma que o codificador de linha e o oscilador da onda portadora são controlados por um clock comum.
- 7.12** Repita o Problema 7.11 assumindo que o codificador de linha e o gerador da onda portadora operam independentemente um do outro. Comente seus resultados.

- 7.13 (a) Repita o Problema 7.11 para a seqüência binária 11100101 aplicada a um modulador PSK, assumindo que o codificador de linha e o oscilador da onda portadora são operados por um clock comum.
- (b) Repita os seus cálculos, assumindo que o codificador de linha e o oscilador da onda portadora operam independentemente.
- 7.14 A seqüência binária 11100101 é aplicada a um modulador QPSK. A duração do bit é $1\mu s$. A freqüência da portadora é 6 MHz.
- Calcule a largura de faixa de transmissão do sinal QPSK.
 - Trace a forma de onda do sinal QPSK.
- 7.15 Repita o Problema 7.14 para o caso de sinalização do OQPSK.
- 7.16 A seqüência binária 11100101 é aplicada ao modulador BFSK de Sunde. A duração do bit é $1\mu s$. As freqüências das portadoras utilizadas para representar os símbolos 0 e 1 são 2,5 MHz e 3,5 MHz, respectivamente.
- Calcule a largura de faixa de transmissão do sinal BFSK.
 - Trace a forma de onda do sinal BFSK.
- 7.17 Como mencionado anteriormente, a forma de onda apresentada na Fig. 7.1(d) é um exemplo de MSK. Determine
- A excursão de freqüência δf do MSK.
 - O parâmetro f_0 de freqüência.
- 7.18 A seqüência binária 11100101 é aplicada a um modulador MSK. A duração do bit é $1\mu s$. As freqüências da portadora utilizadas para representar os símbolos 0 e 1 são 2,5 MHz e 3MHz, respectivamente. Trace a forma de onda do sinal MSK.
- 7.19 Considere um modular MSK que utiliza uma portadora senoidal com freqüência $f_c = 50$ MHz. A taxa de bit da seqüência binária de entrada é 20×10^3 bits/s.
- Calcule a freqüência instantânea do modulador MSK para uma seqüência de dados na qual os símbolos 0 e 1 alternam.
 - Repita os cálculos da parte (a) para uma seqüência de dados constituída somente por 0s. E se a seqüência for constituída somente por 1s?
- 7.20 Suponha que você tenha um sinal $s(t)$ MSK. Como você pode extrair o sinal de temporização de bit de $s(t)$? Explique.
- 7.21 Considere o receptor não coerente da Fig. 7.17 para BASK e o da Fig. 7.18 para BFSK. Rapidamente falando, o receptor BFSK não coerente é duas vezes mais complexo do que o receptor BASK não coerente. Qual seria a vantagem do sistema BFSK sobre o sistema BASK, dado que eles operam na mesma seqüência binária de dados e no mesmo canal de comunicação? Justifique sua resposta.
- Dica:* Refira-se à representação fasorial da Tabela 7.2.
- 7.22 A Fig. 7.29 mostra um receptor para a detecção não coerente de sinais BFSK. O receptor é constituído por dois caminhos: O filtro passa-faixa do caminho superior é sintonizado para a freqüência de portadora f_1 representando o símbolo 1 e o filtro passa-faixa do caminho inferior é sintonizado para a portadora alternativa de freqüência f_2 representando o símbolo 0. Estes dois filtros são seguidos por detectores de nível de energia, cujas saídas são aplicadas ao comparador para recuperar a seqüência binária de dados original. Identifique as condições que o receptor da Fig. 7.29 deve satisfazer para que ele seja uma alternativa ao detector BFSK não coerente da Fig. 7.18.
- 7.23 A seqüência binária 11100101 é aplicada a um sistema DPSK. A duração do bit é $1\mu s$. A freqüência da portadora é 6 MHz.
- Calcule a largura de faixa de transmissão do sistema.
 - Trace a forma de onda do sinal transmitido.
 - Usando a seqüência transmitida da parte (b), trace a forma de onda produzida pelo receptor e compare-a com a seqüência binária original.

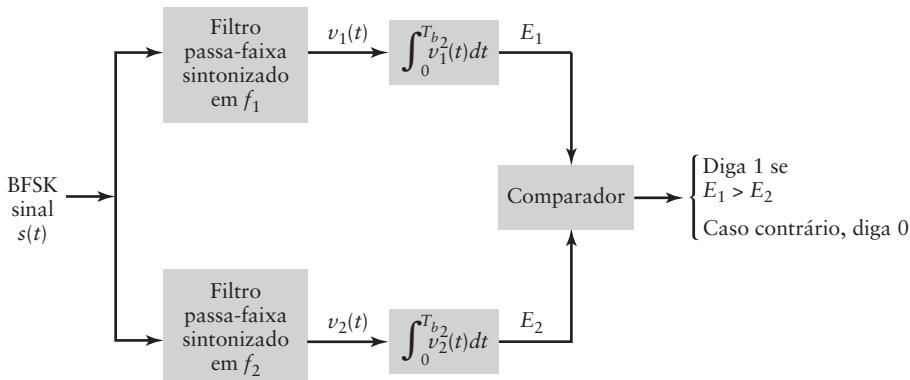


Figura 7.29 Problema 7.22

Problemas avançados

- 7.24** Considere o PSLL constituído por um multiplicador, um filtro de malha e um oscilador controlado a tensão (VCO). Refira-se à Seção 4.8 para uma descrição do PLL para a demodulação de modulação analógica em freqüência. Seja o sinal aplicado à entrada do multiplicador um sinal PSK definido por

$$s(t) = A_c \cos[2\pi f_c t + k_p b(t)]$$

na qual k_p é a sensibilidade de fase e o sinal de dados $b(t)$ assume o valor +1 para o símbolo binário 1 e -1 para o símbolo binário 0. A saída do VCO é

$$r(t) = A_c \sin[2\pi f_c t + \theta(t)]$$

- (a) Calcule a saída do filtro de malha, assumindo que o filtro remove apenas as componentes moduladas com o dobro da freqüência f_c da portadora.
- (b) Mostre que essa saída é proporcional ao sinal de dados $b(t)$ quando a malha está em fase travada, ou seja, $\theta(t) = 0$.

- 7.25** (a) Em um sistema de detecção coerente de FSK, os sinais $s_1(t)$ e $s_2(t)$ representando os símbolos 1 e 0, respectivamente, são definidos por

$$s_1(t) = A_c \cos\left[2\pi\left(f_c + \frac{\Delta f}{2}\right)t\right], \quad 0 \leq t \leq T_b \quad (1)$$

e

$$s_2(t) = A_c \cos\left[2\pi\left(f_c - \frac{\Delta f}{2}\right)t\right], \quad 0 \leq t \leq T_b \quad (2)$$

Assumindo que $f_c > \Delta f$, mostre que o coeficiente de correlação dos sinais $s_1(t)$ e $s_2(t)$, definido por

$$\rho = \frac{\int_0^{T_b} s_1(t)s_2(t) dt}{\sqrt{\int_0^{T_b} s_1^2(t) dt}}$$

é aproximadamente dado por

$$\rho \approx \text{sinc}(2 \Delta f T_b)$$

- (b) Qual é o valor mínimo de deslocamento de freqüência Δf para o qual os sinais $s_1(t)$ e $s_2(t)$ são ortogonais?

- 7.26** Um sinal BFSK com fase descontínua é gerado utilizando uma seqüência binária de dados de entrada para chavear entre dois osciladores senoidais independentes:

$$\text{Símbolo 1: } \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_1 t + \theta_1)$$

$$\text{Símbolo 0: } \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_2 t + \theta_2)$$

- (a) Mostre que este sinal BFSK pode ser descrito como a soma de dois sinais BASK independentes.
 (b) Para $\theta_1 = 30^\circ$ e $\theta_2 = 45^\circ$, trace a forma de onda BFSK para a seqüência binária de dados 01101001, e compare seu resultado com o mostrado na Fig. 7.1(d) para BFSK de fase contínua.

7.27 A Tabela 7.4 mostra as transições de fase vivenciadas por um sinal MSK. Verifique as linhas desta tabela.

TABELA 7.4 Transições de fase no MSK

Bit transmitido $0 \leq t \leq T_b$	Estado da fase (radianos)	
	$\theta(0)$	$\theta(T_b)$
0	0	$-\pi/2$
1	π	$-\pi/2$
0	π	$+\pi/2$
1	0	$+\pi/2$

7.28 A geração e detecção de QPSK pode ser vista como um exemplo da estratégia de “dividir-e-conquistar”. Em particular, a implementação do QPSK é obtida por

- ▶ Conhecer como gerar e detectar BPSK.
- ▶ Utilizar a idéia de multiplexação em quadratura.

Amplie o que foi aprendido neste exemplo e, portanto, desenvolva um transmissor para a geração de 8-PSK e o receptor para sua detecção.

■ Experimentos de computador

7.29 Neste problema, continuamos o Experimento de Computador III sobre a comparação espectral do QPSK e OQPSK. Especificamente, utilizando os parâmetros

Freqüência da portadora $f_c = 8\text{Hz}$

Duração do bit $T_b = 1\text{s}$

Calcule o espectro de fase destes dois métodos de modulação digital. Logo, justifique a afirmativa feita no Problema 7.3 de que estes dois métodos diferem pela componente linear de fase e determine a inclinação desta componente.

7.30 O objetivo deste experimento é demonstrar que a simulação de um sinal digitalmente modulado em um computador pode ser simplificado utilizando a idéia de envelope complexo introduzida no Capítulo 3. Em particular, o experimento compara as formas passa banda e banda base de transmissão de dados, suportada pelos scripts do Matlab do Apêndice 7. Os scripts fornecidos também possibilitam a geração de seqüências aleatórias de dados necessária para executar o experimento e, portanto, adicionam realismo prático aos experimentos.

- (a) Para proceder, gere duas seqüências binárias aleatórias $b_I(t)$ e $b_Q(t)$ e as utilize para formar o sinal passa banda multiplexado

$$s(t) = b_I(t)\cos(2\pi f_c t) - b_Q(t)\sin(2\pi f_c t)$$

Logo, calcule o espectro de amplitude de $s(t)$, utilizando os seguintes parâmetros:

Freqüência da portadora, $f_c = 10\text{ Hz}$

Duração do símbolo (bit), $T_b = 1\text{s}$

Assuma que as duas seqüências binárias são sincronizadas e utilizam níveis de amplitude de ± 1 para representar os símbolos 0 e 1.

- (b) Utilizando a notação complexa, defina o envelope complexo

$$\tilde{b}(t) = b_I(t) + jb_Q(t)$$

com o qual podemos reconstruir o sinal passa banda original

$$s(t) = \operatorname{Re}\{\tilde{b}(t)\exp(j2\pi f_c t)\}$$

Especificamente, calcule o espectro de magnitude do envelope complexo $b(t)$ e compare com o espectro de magnitude $s(t)$ calculado na parte (a). Comente sobre o significado computacional desta comparação.

- 7.31 Repita o Problema 7.30. Desta vez, utilizando uma forma de pulso de cosseno levantado com fator de roll-off $\alpha = 1$ para construir as seqüências binárias $b_I(t)$ e $b_Q(t)$. O Apêndice 7 fornece scripts do Matlab para a geração do pulso de cosseno levantado. Calcule o espectro de magnitude do envelope complexo $b(t)$ e compare-o com o espectro de magnitude do sinal passa banda $s(t)$. Comente seus resultados.

Capítulo 8

SINAIS E RUÍDOS ALEATÓRIOS

O termo “aleatório” é utilizado para descrever variações aparentemente imprevisíveis e erráticas de um sinal observado. Sinais aleatórios em uma ou outra forma são encontrados em praticamente todo sistema de comunicação. Considere a comunicação por voz, na qual a voz é geralmente convertida em um sinal elétrico através de um microfone antes de ser processada para a transmissão. Se este sinal elétrico for mostrado em um osciloscópio, podemos ser tentados a dizer, à primeira vista, que o sinal parece ser bem aleatório, ou seja, que é difícil predizê-lo ou reproduzi-lo. Similarmente, em comunicações digitais, se considerarmos uma seqüência de 0s e 1s que são transportados pela Internet, eles parecem ser bastante aleatórios. Eles são sempre 0 ou 1, mas sua ordem e localização são bem imprevisíveis. Esta aleatoriedade ou imprevisibilidade é uma propriedade fundamental da informação.¹ Se a informação fosse previsível, não seria necessário comunicá-la, pois o outro lado poderia predizer a informação antes de recebê-la.

Por outro lado, o ruído é o assassino da maioria dos sistemas de comunicação. O ruído limita a faixa e/ou a qualidade na qual os sinais contendo a informação podem ser transportados em um canal. O ruído pode ser definido como um sinal indesejado interferindo ou distorcendo o sinal sendo comunicado. O ruído é outro exemplo de sinal aleatório – se o ruído fosse previsível, poderíamos predizê-lo no receptor, removendo-o e negando seu efeito. Portanto, de alguma forma, o ruído é similar à informação em termos de sua natureza aleatória. O objetivo do engenheiro de comunicação é separar ao máximo um do outro.

Apesar da informação e ruído serem sinais aleatórios, eles ainda possuem propriedades que podem ser medidas e analisadas, em algum sentido médio. Este é o propósito deste capítulo, introduzir as ferramentas necessárias para analisar a informação e o ruído, necessárias para entender as técnicas de detecção descritas no restante do livro.

O material neste capítulo nos ensina as seguintes lições:

- ▶ *Lição 1: Eventos e sinais aleatórios podem ser modelados como a saída de experimentos aleatórios.*
- ▶ *Lição 2: Variáveis aleatórias fornecem uma representação geral para análise, comparação e processamento de saída de experimentos aleatórios. Propriedades estatísticas de eventos aleatórios podem ser obtidas da expectativa de várias funções destas variáveis aleatórias. A expectativa deve ser vista como um operador.*
- ▶ *Lição 3: Variáveis aleatórias Gaussianas possuem um papel chave na análise de sinais aleatórios devido ao teorema do limite central e sua capacidade de tratamento matemático.*
- ▶ *Lição 4: Um processo aleatório pode ser visto como uma família de variáveis aleatórias, indexadas por um parâmetro de tempo. Podemos, então, estender as técnicas de análise para variáveis aleatórias para estudar a variação no tempo de processos aleatórios.*
- ▶ *Lição 5: Ruído brando é um dos mais importantes processos aleatórios no estudo de sistemas de comunicação, tanto do ponto de vista prático quanto do ponto de vista de tratamento matemático de suas propriedades estatísticas.*
- ▶ *Lição 6: Ruído de faixa estreita pode ser analisado em termos de suas componentes em fase e em quadratura, de maneira similar aos sinais de comunicação de faixa estreita.*

¹ O conceito de aleatoriedade em comunicações é expandido pelo livro de Wozencraft e Jacobs (1965) sendo um dos pontos importantes da teoria da informação. Veja Cover e Thomas (1991).

8.1 PROBABILIDADE E VARIÁVEIS ALEATÓRIAS

A teoria da probabilidade é utilizada em situações que envolvem a execução de experimentos com uma saída sujeita à variação. Ou seja, se o experimento for repetido, a saída pode ser diferente devido à influência de um fenômeno aleatório em sua base. Tal experimento é chamado de *experimento aleatório*. Por exemplo, o experimento pode ser a observação do resultado de lançar uma moeda. Neste experimento, as possíveis saídas de um teste são “cara” ou “coroa”.

Dizemos que um experimento aleatório possui as seguintes propriedades:

1. Em qualquer tentativa do experimento, a saída é imprevisível.
2. Para uma grande quantidade de tentativas, a saída exibe uma *regularidade estatística*. Ou seja, um padrão médio definido de saída é observado se o experimento é repetido um grande número de vezes.

■ Abordagem em freqüência relativa

Para elaborar o conceito de regularidade estatística, seja o evento A uma possível saída de um experimento aleatório. Por exemplo, no experimento de lançar uma moeda, o evento A pode representar “cara”. Suponha que em n tentativas do experimento, o evento A ocorre n_A vezes. Podemos, então, associar a relação n_A/n ao evento A . Esta relação é chamada de *freqüência relativa* do evento A . Claramente, a freqüência relativa é um número real não negativo menor ou igual a um. Ou seja,

$$0 \leq \frac{n_A}{n} \leq 1 \quad (8.1)$$

Se o evento A não ocorre em nenhuma tentativa, $(n_A/n) = 0$. Se, por outro lado, o evento A ocorre em todas as tentativas, $(n_A/n) = 1$.

Dizemos que o experimento exibe regularidade estatística se, para qualquer seqüência de n tentativas, a freqüência relativa n_A/n converge para um limite quando n se torna grande. Definimos este limite

$$\mathbf{P}[A] = \lim_{n \rightarrow \infty} \left(\frac{n_A}{n} \right) \quad (8.2)$$

como a *probabilidade* do evento A . Portanto, no experimento de lançamento de moeda, esperamos que em um milhão de lançamentos de uma moeda, por exemplo, cerca de metade das saídas seja cara.

A probabilidade de um evento é utilizada para representar a possibilidade de uma tentativa do experimento resultar na ocorrência do evento. Em várias aplicações de engenharia e jogos de sorte, a utilização da Eq. (8.2) define a probabilidade de um evento ser aceitável. Por outro lado, considere a análise estatística do mercado de ações: Como podemos atingir a repetibilidade de tal experimento? Uma abordagem mais satisfatória é apresentar as propriedades que se espera que qualquer medida de probabilidade possua, postulando-as como *axiomas*, e, então, utilizar a interpretação de freqüência relativa para justificá-las.

■ Os axiomas da probabilidade

Quando executamos um experimento aleatório, é natural termos a consciência das várias saídas que podem aparecer. Neste contexto, é conveniente pensarmos no experimento e suas possíveis saídas como a definição de um espaço e de seus pontos. Para cada possível valor de saída do experimento, associamos um ponto chamado de *ponto de amostra*, o qual representamos por s_k . A totalidade de todos os pontos de amostra, correspondendo à união de todos as possíveis saídas de um experimento, é chamada de *espaço de amostra*, a qual representamos por S . Um *evento* corresponde a um único ponto de amostra ou a

um conjunto de pontos. Em particular, o espaço de amostra S total é chamado de *evento certeza*, e o conjunto nulo ϕ é chamado de *evento nulo ou impossível*. Um único ponto de amostra é chamado de *evento elementar*.

Considere, por exemplo, um experimento envolvendo o lançamento de um dado. Neste experimento, existem seis possíveis saídas: um, dois, três, quatro, cinco ou seis pontos na face superior do dado. Associando um ponto de amostra a cada uma destas seis possíveis saídas, temos um espaço de amostra que consiste de seus pontos de amostra, como mostrado na Fig. 8.1.

O evento elementar que descrever a afirmativa “seis mostrado” corresponde ao ponto de amostra $\{6\}$. Por outro lado, o evento descrevendo a afirmativa “um número par de pontos mostrados” corresponde ao subconjunto $\{2, 4, 6\}$ do espaço de amostra. Note que o termo “evento” é utilizado para descrever o subconjunto ou a afirmativa.

Agora estamos prontos para fazer uma definição formal de probabilidade. Um sistema de probabilidade consiste de:

1. O espaço de amostra S de eventos elementares (ou saídas).
2. Uma classe ϵ de eventos que são subconjuntos de S .
3. Uma média de probabilidade $P[A]$ associada a cada evento A da classe ϵ , a qual possui as seguintes propriedades:
 - (i) $P[S] = 1$
 - (ii) $0 \leq P[A] \leq 1$
 - (iii) Se $A \cup B$ é a união de dois eventos mutuamente exclusivos da classe ϵ , então

$$P[A \cup B] = P[A] + P[B] \quad (8.3)$$

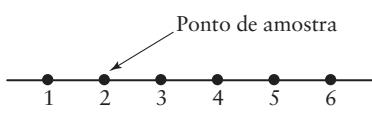
As Propriedades (i), (ii) e (iii) são chamadas de axiomas da probabilidade. O axioma (i) afirma que a probabilidade de um evento certo é um. O axioma (ii) afirma que a probabilidade de um evento é um número não negativo menor ou igual a um. O axioma (iii) afirma que a probabilidade da união de dois eventos mutuamente exclusivos é a soma das probabilidades dos eventos individuais.

Apesar da abordagem por axiomas para a teoria de probabilidade ser de natureza abstrata, todas os três axiomas possuem interpretações de freqüência relativa. O axioma (ii) corresponde à Eq. (8.1). O axioma (i) corresponde ao caso limite da Eq. (8.1) quando o evento A ocorre em todas as tentativas. Para interpretar o axioma (iii) notamos que se o evento A ocorre n_A vezes em n tentativas e o evento B ocorre n_B vezes, então a união do evento “ A ou B ” ocorre em $n_A + n_B$ tentativas (desde que A e B nunca ocorram na mesma tentativa). Logo, $n_{A+B} = n_A + n_B$, e, portanto, temos

$$\frac{n_{A \cup B}}{n} = \frac{n_A}{n} + \frac{n_B}{n} \quad (8.4)$$

a qual é a forma matemática similar a do axioma (iii).

Esta definição abstrata de um sistema de probabilidade é ilustrada na Fig. 8.2. O espaço de amostra S é mapeado em eventos via experimento aleatório. Os eventos podem ser saídas elementares do espaço de amostra ou subconjuntos grandes do espaço de amostra. A função de probabilidade associa um valor entre 0 e 1 a cada um destes eventos. O valor da probabilidade não é único ao evento, eventos mutuamente exclusivos podem possuir a mesma probabilidade. Entretanto, a probabilidade da união de todos os eventos, ou seja, o evento certo, é sempre unitário.



Espaço de amostras unidimensional

Figura 8.1 Espaço de amostra do experimento de lançamento de um dado.

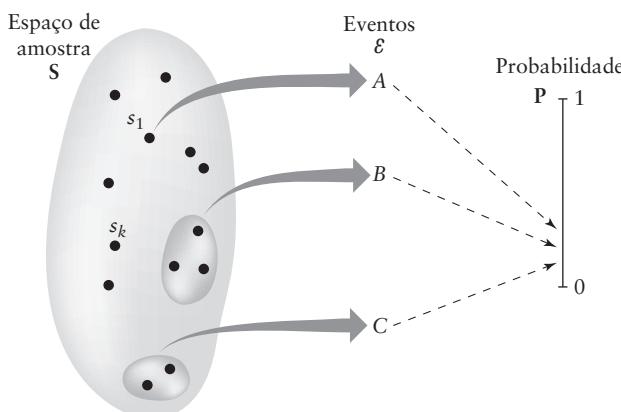


Figura 8.2 Ilustração da relação entre espaço de amostra, eventos e probabilidade.

■ Variáveis aleatórias

Apesar do significado da saída de um experimento aleatório ser claro, tais saídas geralmente não são as representações mais convenientes para análise matemática. Por exemplo, caras ou coroas não são representações matemáticas convenientes. Como outro exemplo, considere um experimento aleatório no qual retiramos bolas coloridas de uma urna. A cor não é capaz de ser diretamente analisada matematicamente.

Nestes casos, geralmente é mais conveniente associar um número ou faixa de valores às saídas de um experimento aleatório. Por exemplo, cara pode corresponder a 1 e coroa a 0. Utilizamos a expressão *variável aleatória* para descrever este processo de associação de um número a uma saída de um experimento aleatório.

Em geral, uma função cujo domínio é um espaço de amostra e cuja faixa é um conjunto de números reais é chamada de variável aleatória do experimento. Ou seja, para eventos em ϵ , uma variável aleatória associa um subconjunto da linha real. Portanto, se a saída de um experimento é s , representamos a variável aleatória como $X(s)$ ou somente X . Note que X é uma função, mesmo que ela seja, por razões históricas, chamada de variável aleatória. Representamos a saída particular de um experimento aleatório por x , ou seja, $X(s_k) = x$. *Pode existir mais de uma variável aleatória associada com o mesmo experimento aleatório.*

O conceito de variável aleatória é ilustrado na Fig. 8.3, na qual suprimimos os eventos mas mostramos os subconjuntos do espaço de amostra sendo mapeados diretamente a um subconjunto da linha real. A função de probabilidade se aplica a esta variável aleatória exatamente da mesma forma que ela se aplicaria aos eventos internos.

O benefício de se utilizar variáveis aleatórias é que a análise de probabilidade pode, agora, ser desenvolvida em termos de grandezas de valor real, independente da forma ou formato dos eventos internos do experimento aleatório. Variáveis aleatórias podem ser *discretas* e assumir apenas um número finito de valores, tais como no experimento de lançamento de moeda. Alternativamente, variáveis aleatórias podem ser *contínuas*, assumindo uma faixa de valores reais. Por exemplo, a variável aleatória que representa a amplitude de uma tensão de ruído em um instante particular no tempo é uma variável aleatória contínua porque, em teoria, ela pode assumir qualquer valor entre mais e menos infinito. Variáveis aleatórias podem também ser de valor complexo, mas uma variável aleatória de valor complexo pode sempre ser tratada como um vetor de duas variáveis aleatórias de valor real.

Para uma variável aleatória de valor discreto, a *função de probabilidade de massa* descreve a probabilidade de cada valor da variável aleatória. Para o experimento de lançamento de moeda, se for uma moeda justa, a função de probabilidade de massa da variável aleatória associada pode ser escrita por

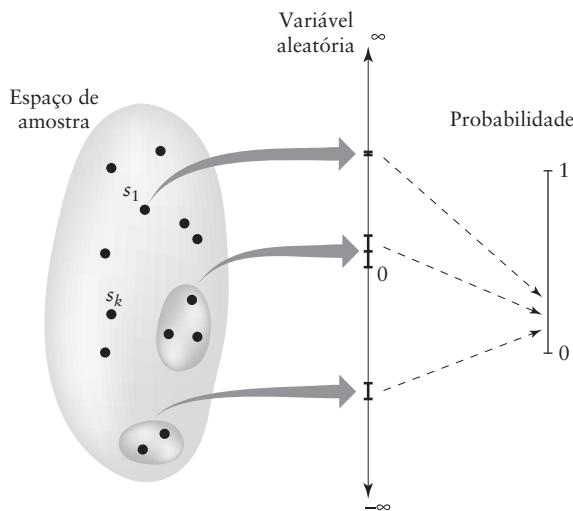


Figura 8.3 Ilustração da relação entre espaço de amostra, variáveis aleatórias e probabilidade.

$$P[X = x] = \begin{cases} \frac{1}{2}, & x = 0 \\ \frac{1}{2}, & x = 1 \\ 0, & \text{caso contrário} \end{cases} \quad (8.5)$$

Esta função de probabilidade de massa é ilustrada na Fig. 8.4, na qual funções delta de peso $\frac{1}{2}$ são utilizadas para representar a probabilidade de massa de cada um dos pontos, 0 e 1.

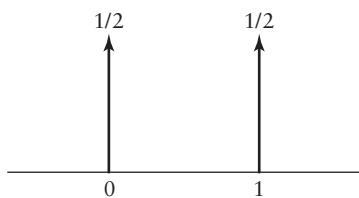


Figura 8.4 Ilustração da função de probabilidade de massa para o experimento de lançamento de moeda.

EXEMPLO 8.1 Variável aleatória de Bernoulli

Considere o experimento de lançamento de moeda na qual a probabilidade de sair cara é p . Seja X a variável aleatória que assume o valor 0 se for coroa e 1 se for cara. Dizemos que X é uma *variável aleatória de Bernoulli*. A função de probabilidade de massa de uma variável aleatória de Bernoulli é

$$P[X = x] = \begin{cases} 1 - p, & x = 0 \\ p, & x = 1 \\ 0, & \text{caso contrário} \end{cases} \quad (8.6)$$

■ Funções de distribuição

Muito relacionada com a função de probabilidade de massa está a *função de probabilidade de distribuição*. Ela é a probabilidade da variável aleatória X assumir qualquer valor menor ou igual a x . A função de distribuição é escrita como $F_X(x)$ tal que

$$F_X(x) = P[X \leq x] \quad (8.7)$$

A função $F_X(x)$ é uma função de x , não da variável aleatória X . Entretanto, ela depende da associação da variável aleatória X , representada pelo uso de X como subscrito. Para qualquer ponto x , a função de distribuição expressa a probabilidade do valor de X ser menor do que x .

A função de distribuição possui duas propriedades básicas, as quais são obtidas diretamente da Eq. (8.7):

1. A função de distribuição $F_X(x)$ é limitada entre zero e um.
2. A função de distribuição $F_X(x)$ é uma função monotônica não decrescente de x , ou seja,

$$F_X(x_1) \leq F_X(x_2) \quad \text{se } x_1 \leq x_2$$

Se X é uma variável aleatória de valor contínuo e $F_X(x)$ é diferenciável com respeito a x , então uma terceira função geralmente utilizada é a *função de densidade de probabilidade*, representada por $f_X(x)$,

$$f_X(x) = \frac{\partial}{\partial x} F_X(x) \quad (8.8)$$

A função densidade de probabilidade possui três propriedades básicas:

1. Como a função de distribuição é monotônica não decrescente, a função de densidade é não negativa para todos os valores de x .
2. A função de distribuição pode ser recuperada da função de densidade pela integração, mostrado por

$$F_X(x) = \int_{-\infty}^x f_X(s) \, ds \quad (8.9)$$

3. A Propriedade 2 implica em que a área total sob a curva da função de densidade seja unitária.

EXEMPLO 8.2 Função de distribuição de Bernoulli

Continuando com o Exemplo 8.1, vemos que a função de probabilidade de distribuição da variável aleatória de Bernoulli é dada por

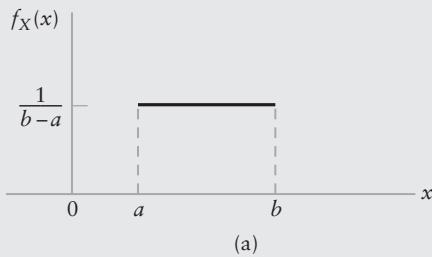
$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - p, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

EXEMPLO 8.3 Distribuição uniforme

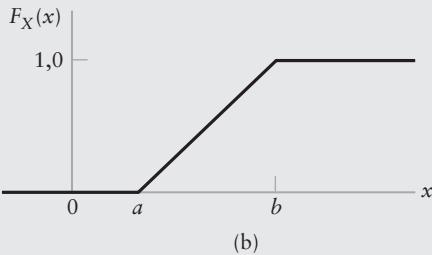
Considere uma variável aleatória X com função de densidade

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{caso contrário} \end{cases} \quad (8.10)$$

Esta função, mostrada na Fig. 8.5(a), satisfaz as condições de densidade de probabilidade porque $f_X(x) \geq 0$, e a área sob a curva é unitária. A variável aleatória possuindo a função densidade de probabilidade da Eq. (8.10) é dita ser *uniformemente distribuída*.



(a)



(b)

Figura 8.5 Distribuição uniforme. (a) Função densidade de probabilidade. (b) Função de distribuição.

A função de distribuição correspondente da variável aleatória uniformemente distribuída X é contínua em todo local, mostrado por

$$F_X(x) = \int_{-\infty}^x f_X(s) ds = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

Esta função de distribuição é traçada na Fig. 8.5(b).

■ Várias variáveis aleatórias

Até o momento, focalizamos em situações envolvendo uma única variável aleatória. Entretanto, freqüentemente nos deparamos com as saída de um experimento sendo descrita por várias variáveis aleatórias diferentes, e estamos interessados nas relações entre estas variáveis aleatórias. Nesta seção, iremos considerar o caso de duas variáveis aleatórias, mas a abordagem pode ser estendida para qualquer número de variáveis aleatórias.

Considere duas variáveis aleatórias X e Y . Definimos a *função de distribuição comum* $F_{X, Y}(x, y)$ como a probabilidade da variável aleatória X ser menor ou igual a um valor específico x e da variável aleatória Y ser menor ou igual a um valor específico y . A função de distribuição comum, $F_{X, Y}(x, y)$, é a probabilidade da saída de um experimento resultar em um ponto de amostra dentro do quadrante ($-\infty < X \leq x, -\infty < Y \leq y$) do espaço de amostra comum. Ou seja,

$$F_{X, Y}(x, y) = P[X \leq x, Y \leq y] \quad (8.11)$$

Suponha que a função de distribuição comum $F_{X, Y}(x, y)$ é sempre contínua, e que a derivada parcial

$$f_{X, Y}(x, y) = \frac{\partial^2 F_{X, Y}(x, y)}{\partial x \partial y} \quad (8.12)$$

existe e é sempre contínua. Chamamos a função $f_{X,Y}(x, y)$ de *função de densidade de probabilidade comum* das variáveis aleatórias X e Y . A função de distribuição comum $F_{X,Y}(x, y)$ é uma função monotônica não decrescente de x e y . Portanto, da Eq. (8.12), a função densidade de probabilidade comum $f_{X,Y}(x, y)$ é sempre não negativa. Além disso, o volume total sob o gráfico da função densidade de probabilidade comum é unitária.

A função densidade de probabilidade para uma única variável aleatória, digamos X , pode ser obtida de sua função densidade de probabilidade comum com uma segunda variável aleatória, digamos Y , pela seguinte forma. Primeiro observamos que

$$F_X(x) = \int_{-\infty}^{\infty} \int_{-\infty}^x f_{X,Y}(\xi, \eta) d\xi d\eta \quad (8.13)$$

Diferenciando os dois lados da Eq. (8.13) com respeito a x , obtemos a relação desejada

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, \eta) d\eta \quad (8.14)$$

Portanto, a função densidade de probabilidade $f_X(x)$ pode ser obtida da função densidade de probabilidade comum $f_{X,Y}(x, y)$ pela simples integração em todos os possíveis valores da variável aleatória Y . A utilização de argumentos similares no contexto da outra variável aleatória Y resulta em $f_Y(y)$. As funções densidade de probabilidade $f_X(x)$ e $f_Y(y)$ são chamadas de *densidades marginais*.

As variáveis aleatórias X e Y são *estatisticamente independentes* se a saída X não afetar a saída Y . Matematicamente, para X e Y independentes, a probabilidade comum $P[X \in A, Y \in B]$ é o produto das probabilidades individuais, ou seja,

$$P[X \in A, Y \in B] = P[X \in A]P[Y \in B] \quad (8.15)$$

para todos os conjuntos A e B nas faixas respectivas de X e Y . A propriedade da Eq. (8.15) se estende imediatamente para funções de distribuição fazendo $A = (-\infty, x]$ e $B = (-\infty, y]$, ou seja,

$$F_{X,Y}(x, y) = F_X(x)F_Y(y) \quad (8.16)$$

Uma propriedade similar é válida para a função de densidade comum. Em geral, para situações envolvendo duas ou mais variáveis aleatórias, as variáveis podem ser trabalhadas separadamente sem nos preocuparmos com a interação com as outras variáveis aleatórias. Para simplificar a notação quando consideramos as probabilidades de variáveis aleatórias, geralmente iremos suprimir o conjunto envolvido. Por exemplo, a Eq. (8.15) é geralmente escrita simplesmente como $P[X, Y] = P[X]P[Y]$.

EXEMPLO 8.4 Variável aleatória binomial

Considere uma seqüência de experimentos de lançamento de moeda na qual a probabilidade de sair cara é p e seja X_n a variável aleatória de Bernoulli representando a saída do n -ésimo lançamento. Como a saída de um lançamento de moeda não é influenciado pela saída dos lançamentos subsequentes, este conjunto é referenciado como um conjunto de *tentativas independentes de Benoulli*.

Seja Y o número de caras que ocorrem em N lançamentos de moedas. Esta nova variável aleatória pode ser descrita por

$$Y = \sum_{n=1}^N X_n \quad (8.17)$$

Qual é a função de probabilidade de massa de Y ?

Considere, primeiro, a probabilidade de se obter y caras em seqüência seguido por $N - y$ coroas. Usando a independência das tentativas, a aplicação repetida da Eq. (8.15) implica em que esta probabilidade é dada por

$$\begin{aligned} P[\text{caras } y \text{ seguida por } N - y \text{ coroas}] &= ppp \dots pp(1-p)(1-p) \dots (1-p) \\ &= p^y(1-p)^{N-y} \end{aligned}$$

Por simetria, esta é a probabilidade de qualquer seqüência de N tentativas que possui y caras. Para determinar a probabilidade de se obter y caras em qualquer lugar nas N tentativas, a definição de frequência relativa de probabilidade implica em que simplesmente precisamos determinar o número de possíveis arranjos de N lançamentos que possuam y caras e $N - y$ coroas. Ou seja, a probabilidade de $Y = y$ é dada por

$$P[Y = y] = \binom{N}{y} p^y (1-p)^{N-y} \quad (8.18)$$

na qual

$$\binom{N}{y} = \frac{N!}{y!(N-y)!}$$

é a função combinatória. A Eq. (8.18) define a função de probabilidade de massa de Y e a variável aleatória Y é dita possuir uma *distribuição binomial*. A distribuição binomial possui este nome porque os valores de $P[Y = y]$ são termos sucessivos na expansão da expressão binomial

$$[p + (1-p)]^N \quad (8.19)$$

na qual o $(y+1)$ -ésimo termo da expansão corresponde a $P[Y = y]$. A função de probabilidade de massa binomial para $N = 20$ e $p = \frac{1}{2}$ é ilustrada na Fig. 8.6.

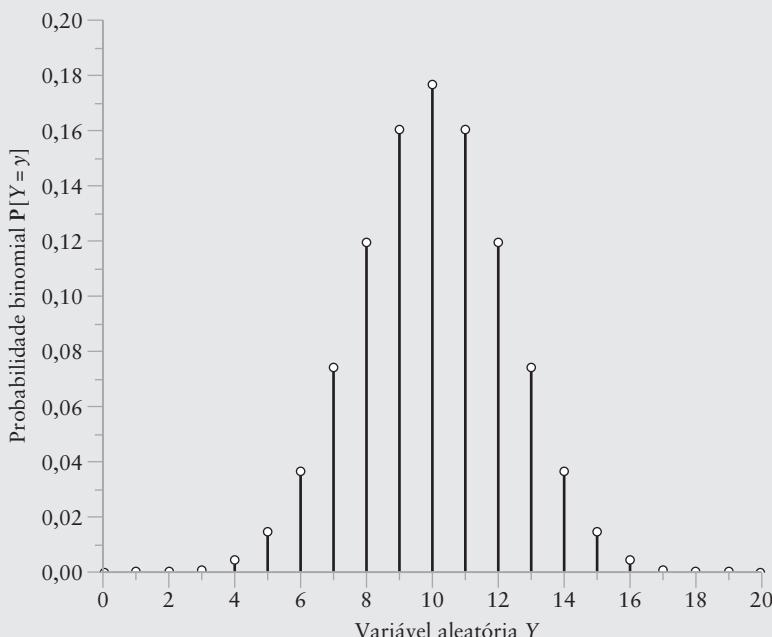


Figura 8.6 Função de probabilidade de massa binomial para $N = 20$ e $p = \frac{1}{2}$.

EXEMPLO 8.5 Função de distribuição binomial

Este exemplo trabalha com a função de distribuição de probabilidade de uma variável aleatória binomial.

Pela definição da Eq. (8.7), a função de distribuição é dada por $F_Y(y) = \mathbb{P}[Y \leq y]$, portanto, para uma variável aleatória binomial, temos

$$\begin{aligned} F_Y(y) &= \mathbb{P}[Y \leq y] \\ &= \sum_{k=0}^y \mathbb{P}[Y = k] \\ &= \sum_{k=0}^y \binom{N}{k} p^k (1-p)^{N-k} \end{aligned} \quad (8.20)$$

Esta função de distribuição de Y é ilustrada na Fig. 8.7 para $N = 20$ e $p = \frac{1}{2}$. Podemos utilizar a observação da Eq. (8.19) para mostrar que

$$\begin{aligned} F_Y(N) &= \sum_{k=0}^N \binom{N}{k} p^k (1-p)^{N-k} \\ &= [p + (1-p)]^N \\ &= 1 \end{aligned} \quad (8.21)$$

como deve ser para uma função de distribuição.

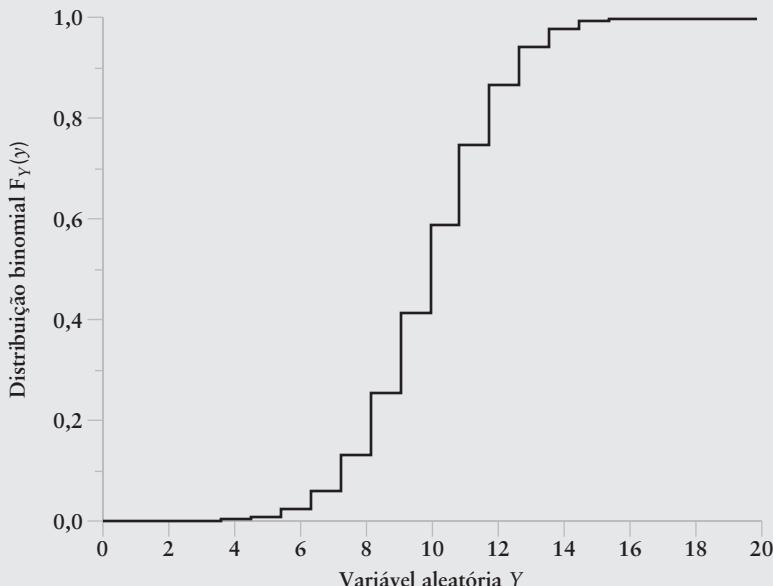


Figura 8.7 Função de distribuição binomial para $N = 20$ e $p = \frac{1}{2}$.

- **PROBLEMA 8.1** Um pacote de informação contém 200 bits. Este pacote é transmitido em um canal de comunicação no qual a probabilidade de erro de cada bit é 10^{-3} . Qual é a probabilidade do pacote ser recebido sem erro?
- **PROBLEMA 8.2** Suponha que o pacote do Problema 8.1 inclua um código de correção de erro que pode corrigir até três erros localizados em qualquer posição do pacote. Qual é a probabilidade de um pacote particular ser recebido com erro neste caso?

■ Probabilidade condicional

Suponha que estamos estudando um experimento ou sinal aleatório que seja caracterizado por duas variáveis aleatórias, X e Y , as quais não são independentes. Então, podemos esperar que conhecendo o valor de uma variável aleatória, digamos X , podemos influenciar os valores observados para a outra variável aleatória. Por exemplo, considere o lançamento de dois dados. Suponha que X representa o número mostrado no primeiro dado e Y representa a soma dos dois dados. Sabendo que $X = 3$ isso claramente irá influenciar o valor que pode ser esperado para Y .

Seja $P[Y|X]$ a função de probabilidade de massa de Y dado que X tenha ocorrido. A probabilidade $P[Y|X]$ é chamada de *probabilidade condicional de Y dado X* . Assumindo que X tenha probabilidade não nula, a probabilidade condicional é definida como

$$P[Y|X] = \frac{P[X, Y]}{P[X]} \quad (8.22)$$

na qual $P[X, Y]$ é a probabilidade comum das duas variáveis aleatórias. Esta definição pode ser justificada utilizando a interpretação de freqüência relativa.

Podemos reescrever a Eq. (8.22) como

$$P[X, Y] = P[Y|X] P[X] \quad (8.23)$$

Se considerarmos a probabilidade condicional de X dado Y , também podemos escrever

$$P[X, Y] = P[X|Y] P[Y] \quad (8.24)$$

As Eqs. (8.23) e (8.24) afirmam que a probabilidade comum de dois eventos pode ser expressa como o produto da probabilidade condicional de um evento dado o outro e a probabilidade elementar do outro. Note que as probabilidades condicionais $P[Y|X]$ e $P[X|Y]$ possuem as mesmas propriedades das várias probabilidades mencionadas anteriormente.

Situações podem existir nas quais a probabilidade condicional $P[X|Y]$ e as probabilidades $P[X]$ e $P[Y]$ sejam determinadas diretamente e facilmente, mas a probabilidade condicional $P[Y|X]$ ser desejada. Das Eqs. (8.23) e (8.24), desde que $P[X] \neq 0$, podemos determinar $P[Y|X]$ utilizando a relação

$$P[Y|X] = \frac{P[X|Y] P[Y]}{P[X]} \quad (8.25)$$

Esta relação é uma forma especial da *regra de Bayes*.

Suponha que a probabilidade condicional $P[Y|X]$ é simplesmente igual à probabilidade de ocorrência de Y , ou seja,

$$P[Y|X] = P[Y] \quad (8.26)$$

Com esta condição, a probabilidade comum de X e Y é igual ao produto das probabilidades individuais:

$$P[X, Y] = P[X] P[Y] \quad (8.27)$$

temos, então, da Eq. (8.24), que $P[X|Y] = P[X]$. Neste caso, o conhecimento da saída de uma variável aleatória não nos diz nada sobre a probabilidade da saída da outra variável aleatória além do que saberíamos sem esse conhecimento. Variáveis aleatórias X e Y que satisfazem esta condição são *estatisticamente independentes*.

EXEMPLO 8.6 Canal binário simétrico

Considere um *canal sem memória discreto* utilizado para transmitir dados binários. O canal é dito ser *discreto* no sentido de que é projetado para lidar com mensagens discretas. Ele é *sem memória* no sentido da saída do canal em qualquer tempo depender somente da entrada do canal naquele tempo. Devido à inevitável presença de *ruido* no canal, erros são feitos na seqüência binária de dados recebida. Especificamente, quando o símbolo 1 é enviado, *ocasionalmente*, um erro acontece e o símbolo 0 é recebido, e vice-versa. O canal é assumido como sendo simétrico, o que significa que a probabilidade de receber o símbolo 1 quando 0 é enviado é a mesma probabilidade de receber o símbolo 0 quando o símbolo 1 é enviado.

Para descrever completamente a natureza probabilística deste canal, precisamos de dois conjuntos de probabilidades:

1. A probabilidade *a priori* de enviar os símbolos binários 0 e 1 é dada por

$$P[X = x] = \begin{cases} p_0 & x = 0 \\ p_1 & x = 1 \end{cases}$$

na qual X é a variável aleatória representando o símbolo transmitido. Note que $p_0 + p_1 = 1$, tal que X é uma variável aleatória de Bernoulli.

2. A probabilidade condicional de erro é dada por

$$P[Y = 1|X = 0] = P[Y = 0|X = 1] = p$$

na qual Y é a variável aleatória representando o símbolo recebido. A probabilidade condicional $P[Y = 0|X = 1]$ é a probabilidade do símbolo 0 ser recebido quando o símbolo 1 foi enviado.

A condição é determinar as probabilidades *a posteriori* $P[X = 0|Y = 0]$ e $P[X = 1|Y = 1]$. A probabilidade condicional $P[X = 0|Y = 0]$ é a probabilidade do símbolo 0 ter sido enviado quando o símbolo 0 foi recebido e, similarmente, para $P[X = 1|Y = 1]$. As duas probabilidades condicionais se referem a eventos que são observados “após o fato”, pois isso é o nome de probabilidades “a posteriori”.

Como os eventos $Y = 0$ e $Y = 1$ são mutuamente exclusivos e a probabilidade de receber o símbolo 0 ou o símbolo 1 é unitária, temos, do axioma (iii):

$$P[Y = 0|X = 0] + P[Y = 1|X = 0] = 1$$

Dito isto,

$$P[Y = 0|X = 0] = 1 - p$$

Similarmente, podemos escrever

$$P[Y = 1|X = 1] = 1 - p$$

Desta forma, podemos utilizar o diagrama de transição de probabilidade mostrado na Fig. 8.8 para representar o canal binário de comunicação especificado neste exemplo. O termo “transição de probabilidade” se refere à probabilidade condicional de erro. A Fig. 8.8 claramente mostra a natureza (assumida) simétrica do canal. Por isso o nome “canal binário simétrico”.

Da Fig. 8.8 podemos deduzir os seguintes resultados:

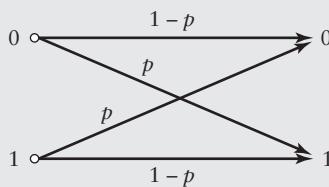


Figura 8.8 Diagrama de transição de probabilidade do canal binário simétrico.

1. A probabilidade de receber o símbolo 0 é dada por

$$\begin{aligned} P[Y = 0] &= P[Y = 0|X = 0]P[X = 0] + P[Y = 0|X = 1]P[X = 1] \\ &= (1 - p)p_0 + pp_1 \end{aligned} \quad (8.28)$$

2. A probabilidade de receber o símbolo 1 é dada por

$$\begin{aligned} P[Y = 1] &= P[Y = 1|X = 0]P[X = 0] + P[Y = 1|X = 1]P[X = 1] \\ &= pp_0 + (1 - p)p_1 \end{aligned} \quad (8.29)$$

Portanto, aplicando a regra de Bayes, obtemos,

$$\begin{aligned} P[X = 0|Y = 0] &= \frac{P[Y = 0|X = 0]P[X = 0]}{P[Y = 0]} \\ &= \frac{(1 - p)p_0}{(1 - p)p_0 + pp_1} \end{aligned} \quad (8.30)$$

e

$$\begin{aligned} P[X = 1|Y = 1] &= \frac{P[Y = 1|X = 1]P[X = 1]}{P[Y = 1]} \\ &= \frac{(1 - p)p_1}{pp_0 + (1 - p)p_1} \end{aligned} \quad (8.31)$$

Estes são os resultados desejados.

- **PROBLEMA 8.3** Continuando com o Exemplo 8.6, determine a seguintes probabilidades condicionais: $P[X = 0|Y = 1]$ e $P[X = 1|Y = 0]$. ◀
- **PROBLEMA 8.4** Considere um canal binário simétrico para o qual a probabilidade de erro é $p = 10^{-4}$ e os símbolos 0 e 1 ocorrem com probabilidade igual. Calcule as seguintes probabilidades:
 - (a) A probabilidade de receber o símbolo 0.
 - (b) A probabilidade de receber o símbolo 1.
 - (c) A probabilidade do símbolo 0 ter sido enviado dado que o símbolo 0 foi recebido.
 - (d) A probabilidade do símbolo 1 ter sido enviado dado que o símbolo 0 foi recebido. ◀

8.2 ESPERANÇA

A função distribuição de probabilidade pode incluir mais detalhes do que o necessário em alguns casos, enquanto ela fornecer uma descrição completa da variável aleatória. Podemos querer utilizar medidas estatísticas mas simples, tais como média e variância, para descrever a variável aleatória.

Média

Estas medidas estatísticas ou *esperanças* são representadas por, por exemplo, $E[g(x)]$ para o valor esperado de uma função $g(\cdot)$ da variável X . Para o caso da esperança do valor de X , representamos a média usando a notação μ_X . Para uma variável aleatória discreta X , a *média*, μ_X , é a soma ponderada de todas as possíveis saídas

$$\begin{aligned} \mu_X &= E[X] \\ &= \sum_X x P[X = x] \end{aligned} \quad (8.32)$$

na qual o somatório ocorre para todas as possíveis saídas da variável aleatória X . Para uma variável aleatória contínua com função de densidade $f_X(x)$, a definição análoga de valor esperado é

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf_X(x) dx \quad (8.33)$$

Geralmente o valor médio de uma variável aleatória é estimado de N observações da variável aleatória $\{x_1, x_2, \dots, x_N\}$, usando o *estimador*

$$\hat{\mu}_X = \frac{1}{N} \sum_{n=1}^N x_n \quad (8.34)$$

Ou seja, estimamos a média da distribuição calculando o valor médio em um determinado número de observações da variável aleatória. Este estimador é baseado na definição de frequência relativa de probabilidade. Por exemplo, se as possíveis saídas da variável aleatória Z são $1, 2, 3, \dots, M$, então o estimador se torna

$$\hat{\mu}_Z = \frac{1 \cdot n_1 + 2 \cdot n_2 + \dots + M \cdot n_M}{N}$$

no qual n_i é o número de vezes que a observação $Z = i$ ocorre. Podemos reescrever esta equação como

$$\begin{aligned}\hat{\mu}_Z &= \sum_{i=1}^M i \frac{n_i}{N} \\ &\approx \sum_{i=1}^M i \mathbb{P}[Z = i]\end{aligned}$$

Ou seja, esperamos que as saídas mais prováveis ocorram mais freqüentemente quando um número de observações é feito.

Retornando ao caso geral, se considerarmos X como variável aleatória representando as observações da tensão de um sinal aleatório, então o valor médio de X representa a tensão média ou *nível CC* do sinal.

■ Variância

A variância de uma variável aleatória é uma estimativa do espalhamento da distribuição de probabilidade ao redor da média. Para variáveis aleatórias, a variância σ_X^2 é dada pela esperança da distância quadrática de cada saída para o valor médio da distribuição.

$$\begin{aligned}\sigma_X^2 &= \text{Var}(X) \\ &= \mathbb{E}[(X - \mu_X)^2] \\ &= \sum_X (x - \mu_X)^2 \mathbb{P}[X = x]\end{aligned} \quad (8.35)$$

Para uma variável aleatória contínua com função de densidade $f_X(x)$, a definição análoga de variância é dada por

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx \quad (8.36)$$

Tal como no caso da média, podemos estimar a variância de uma variável aleatória de N observações independentes utilizando a fórmula

$$\hat{\sigma}_X^2 = \frac{1}{N-1} \sum_{n=1}^N (x_n - \hat{\mu}_X)^2 \quad (8.37)$$

Podemos justificar este estimador utilizando a interpretação de freqüência relativa da probabilidade. Por exemplo, se utilizarmos a variável Z definida anteriormente, teremos

$$\begin{aligned} \hat{\sigma}_Z^2 &= \frac{(1 - \hat{\mu}_Z)^2 \cdot n_1 + (2 - \hat{\mu}_Z)^2 \cdot n_2 + \cdots + (M - \hat{\mu}_Z)^2 \cdot n_M}{N-1} \\ &= \frac{N}{N-1} \sum_{n_i=1}^M (i - \hat{\mu}_Z)^2 \frac{n_i}{N} \\ &\approx \frac{N}{N-1} \sum_{i=1}^M (i - \hat{\mu}_Z)^2 \mathbf{P}[Z = i] \end{aligned} \quad (8.38)$$

O fator $N/(N-1)$ que aparece no lado direito da Eq. (8.38) é devido à consideração de que as mesmas N observações são utilizadas para estimar o valor médio $\hat{\mu}$. Com este fator, o valor esperado do lado direito é σ_Z^2 e, consequentemente, a Eq. (8.37) é um *estimador não polarizado* da variância. Dizemos que \hat{g} é um estimador não polarizado de g se $E[\hat{g}] = E[g]$.

Se considerarmos X como uma variável aleatória representando observações da tensão de um sinal aleatório, então a variância representa a *potência CA* do sinal. O segundo momento de X , $E[X^2]$, também é chamado *valor médio quadrático* do sinal aleatório e representa fisicamente a *potência total do sinal*.

EXEMPLO 8.7 Média e variância de uma variável aleatória de Bernoulli

Se X é uma variável aleatória de Bernoulli com parâmetro p , então o valor esperado de X é

$$\begin{aligned} E[X] &= \sum_{k=0}^1 k \mathbf{P}[X = k] \\ &= 0 \cdot (1 - p) + 1 \cdot p \\ &= p \end{aligned}$$

com $\mu_X = E[X]$, a variância de X é dada por

$$\begin{aligned} \sigma_X^2 &= \sum_{k=0}^1 (k - \mu_X)^2 \mathbf{P}[X = k] \\ &= (0 - p)^2(1 - p) + (1 - p)^2p \\ &= (p^2 - p^3) + (p - 2p^2 + p^3) \\ &= p(1 - p) \end{aligned}$$

Covariância

Também de importância na análise de sistemas de comunicação são as medidas estatísticas entre duas variáveis aleatórias. A *covariância* de duas variáveis aleatórias, X e Y , é dada pelo valor esperado do produto das duas variáveis aleatórias,

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \quad (8.39)$$

Podemos expandir esta equação para obter² (veja o Problema 8.23)

$$\text{Cov}(X, Y) = E[XY] - \mu_X\mu_Y \quad (8.40)$$

Se as duas variáveis aleatórias forem contínuas com densidade comum $f_{X, Y}(x, y)$, então o termo de esperança da Eq. (8.40) é dado por

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_{X, Y}(x, y) dx dy \quad (8.41)$$

Se as duas variáveis aleatórias forem independentes, então

$$\begin{aligned} E[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_X(x)f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} xf_X(x) dx \int_{-\infty}^{\infty} yf_Y(y) dy \\ &= E[X]E[Y] \end{aligned} \quad (8.42)$$

como poderia ser intuitivamente esperado. Substituindo este resultado na Eq. (8.40), vemos que a covariância de variáveis aleatórias independentes é zero. Deve ser observado, entretanto, que o oposto nem sempre é verdadeiro: covariância zero não implica, em geral, em independência.

- **PROBLEMA 8.5** Determinar a média e a variância de uma variável randômica uniformemente distribuída entre os pontos a e b. ◀

8.3 TRANSFORMAÇÃO DE VARIÁVEIS ALEATÓRIAS

Em sistema de comunicação, sinais aleatórios podem ser processados por diversos dispositivos, lineares ou não lineares, antes da saída final. Se representarmos um sinal aleatório por uma variável aleatória e soubermos sua função de distribuição antes do processamento, é lógico perguntar: Qual será a nova função de distribuição da variável aleatória após o processamento? Nesta seção iremos fornecer algumas respostas a esta questão quando a transformação é de *um-para-um*³. Por exemplo, assuma que a variável aleatória X com distribuição $f_X(x)$ é transformada para $Y = ax + b$. Qual é a função de distribuição de Y ?

Respondemos a esta questão retornando aos fundamentos da probabilidade. Considere a probabilidade de X pertencer ao conjunto A , sendo A um subconjunto da linha real. Se $X \in A$, então temos que $Y \in B$, sendo B definido por $B = aA + b$, logo, temos

$$P[X \in A] = P[Y \in B] \quad (8.43)$$

Suponha que o conjunto B é a linha real de $-\infty$ a y , que escrevemos por $(-\infty, y]$. Então, o conjunto A é dado por $A = (B - b)/a = (-\infty, (y - b)/a]$. Portanto, podemos escrever

$$\begin{aligned} F_Y(y) &= P[Y \in (-\infty, y)] \\ &= P[X \in (-\infty, (y - b)/a)] \\ &= F_X\left(\frac{y - b}{a}\right) \end{aligned} \quad (8.44)$$

o que define a relação entre as funções de distribuição da variável aleatória original X e da variável aleatória transformada Y . Em geral, se $Y = g(X)$, é uma transformação de um-

² Se as variáveis aleatórias forem de valor complexo, então esta expressão é modificada para $\text{Cov}(X, Y) = E[XY^*] - \mu_X\mu_Y^*$, na qual o asterisco representa conjugado complexo. Em tal situação, $\text{Cov}(X, Y)$ e $\text{Cov}(Y, X)$ são desiguais.

³ Para outros exemplos de funções de variáveis aleatórias, veja as páginas 119-126 de Leon-Garcia (1994) e as páginas 179-190 de Bertsekas e Tsitsiklis (2002).

para-um da variável aleatória X para a variável aleatória Y , então a função de distribuição de Y é dada por

$$F_Y(y) = F_X(g^{-1}(y)) \quad (8.45)$$

na qual o símbolo $g^{-1}(y)$ representa a *função inversa* de $g(y)$. Pode ser mostrado que se X

possui densidade $f(x)$, então a densidade de Y é $f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|$, desde que a transformação seja diferenciável.

EXEMPLO 8.8 A transformação cosseno

Seja X uma variável aleatória uniformemente distribuída em $[0, 2\pi]$. Seja Y a variável transformada $Y = \cos(X)$. Queremos determinar a função de distribuição de Y .

Primeiro observamos que a transformação $Y = \cos(X)$ não é de um-para-um, portanto não podemos utilizar a teoria descrita anteriormente. Entretanto, notamos que podemos dividir a faixa de X em dois intervalos $[0, \pi]$ e $[\pi, 2\pi]$, sendo que, em cada intervalo, a transformação é de um-para-um. Consequentemente, se Y pertence ao conjunto B , então X pertence ao conjunto $A_1 = \cos^{-1}(B) \cap [0, \pi]$ ou X pertence ao conjunto $A_2 = \cos^{-1}(B) \cap [\pi, 2\pi]$. Como os dois conjuntos A_1 e A_2 são disjuntos, podemos utilizar os axiomas da probabilidade para escrever

$$P[Y \in B] = P[X \in A_1] + P[X \in A_2] \quad (8.46)$$

Quando $B = (-\infty, y]$, então temos, da Fig. 8.9,

$$A_1 = \begin{cases} \phi, & y < -1 \\ [\cos^{-1}(y), \pi], & |y| \leq 1 \\ [0, \pi], & y > 1 \end{cases} \quad \text{e} \quad A_2 = \begin{cases} \phi, & y < -1 \\ [\pi, 2\pi - \cos^{-1}(y)], & |y| \leq 1 \\ [\pi, 2\pi], & y > 1 \end{cases}$$

na qual $\cos^{-1}(y)$ se refere ao valor principal, ou seja, $\cos^{-1}(y)$ está sempre em $[0, \pi]$. Calculando a Eq. 8.46, temos

$$\begin{aligned} F_Y(y) &= \begin{cases} P[\phi] + P[\phi], & y < -1 \\ P[A], & |y| \leq 1 \\ P[0, \pi] + P[\pi, 2\pi], & y > 1 \end{cases} \\ &= \begin{cases} 0, & y < -1 \\ P[A], & |y| \leq 1 \\ 1, & y > 1 \end{cases} \end{aligned} \quad (8.47)$$

na qual $A = [\cos^{-1}(y), 2\pi - \cos^{-1}(y)]$. Como X é uniformemente distribuído com densidade $1/2\pi$, esta probabilidade é proporcional ao comprimento do intervalo A , ou seja,

$$P[X \in A] = \frac{2\pi - 2\cos^{-1}(y)}{2\pi}$$

A substituição deste resultado na Eq. (8.47) define a distribuição de Y .

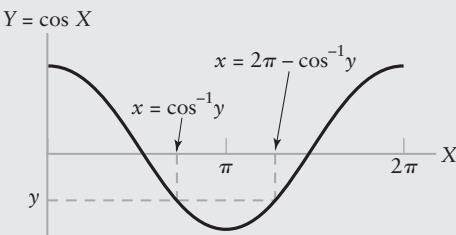


Figura 8.9 Ilustração da transformação de cosseno.

- **PROBLEMA 8.6** Seja X uma variável aleatória e seja $Y = (X - \mu_X)/\sigma_X$. Quais são as médias e variâncias da variável aleatória Y ? ◀
- **PROBLEMA 8.7** Qual é a função densidade de probabilidade da variável aleatória Y do Exemplo 8.8? Trace esta função de densidade. ◀

8.4 VARIÁVEIS ALEATÓRIAS GAUSSIANAS⁴

A variável aleatória *Gaussiana*⁵ possui um importante papel em diversas aplicações e é a variável aleatória mais geralmente encontrada na análise estatística de sistemas de comunicação. Uma variável aleatória Gaussiana é uma variável aleatória contínua com função de densidade dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left\{-\frac{(x - \mu_X)^2}{2\sigma_X^2}\right\} \quad (8.48)$$

na qual a variável aleatória Gaussiana X possui média μ_X e variância σ_X^2 . Esta função de densidade se estende de $-\infty$ a ∞ , sendo simétrica com relação a média μ_X . Uma variável aleatória Gaussiana possui diversas propriedades que serão ditas sem prova⁶:

1. Uma variável aleatória Gaussiana é completamente caracterizada por sua média e variância.
2. Uma variável aleatória Gaussiana mais uma constante é outra variável aleatória Gaussiana com a média ajustada pela constante.
3. Uma variável aleatória Gaussiana multiplicada por uma constante é outra variável aleatória Gaussiana na qual tanto a média quanto a variância são afetadas pela constante.
4. A soma de duas variáveis aleatórias Gaussianas também é uma variável aleatória Gaussiana.
5. A soma ponderada de N variáveis aleatórias Gaussianas é uma variável aleatória Gaussiana.
6. Se duas variáveis aleatórias Gaussianas possuem covariância nula (não correlacionadas), elas também são independentes.

Exceto pelas Propriedades 2 e 3, estas propriedades não são válidas, em geral, para outros tipos de variáveis aleatórias.

Para o caso especial de uma variável aleatória Gaussiana com média zero, $\mu_X = 0$, e variância unitária, $\sigma_X^2 = 1$, a função de densidade é dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2), \quad -\infty < x < \infty \quad (8.49)$$

a qual possui a forma familiar de sino mostrada na Fig. 8.10(a).

A função de distribuição desta *variável aleatória Gaussiana normalizada* é dada pela integral desta função

⁴ A distribuição Gaussiana recebeu este nome depois do grande matemático C. G. Gauss. Com 18 anos, Gauss inventou o método dos mínimos quadráticos para determinar a melhor estimativa de uma grandeza baseado em uma seqüência de medidas. Gauss, posteriormente, utilizou o método dos mínimos quadráticos para estimar órbitas de planetas com medidas ruidosas, um procedimento que foi publicado em 1809 em seu livro *Theory of Motion of Heavenly Bodies*. Em conexão com o erro da observação, ele desenvolveu a distribuição Gaussiana.

⁵ Variáveis aleatórias Gaussianas também são chamadas de *variáveis aleatórias normais*. Engenheiros e físicos tendem a utilizar o termo “Gaussiano” enquanto que matemáticos tendem a utilizar o termo “normal”.

⁶ A prova destas propriedades pode ser encontrada no Capítulo 4 de Leon-Garcia (1994) e Capítulo 3 de Bertsekas e Tsitsiklis (2002).

$$\begin{aligned}
 F_X(x) &= \int_{-\infty}^x f_X(s) \, ds \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-s^2/2) \, ds
 \end{aligned} \tag{8.50}$$

e ilustrada na Fig. 8.10(b). Não existe solução analítica para esta integral, mas, devido ao freqüente aparecimento de integrais deste tipo, várias funções relacionadas foram definidas e tabuladas. A função relacionada, geralmente utilizada no contexto de comunicações, é a *função Q*, definida por⁷

$$\begin{aligned}
 Q(x) &= \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-s^2/2) \, ds \\
 &= 1 - F_X(x)
 \end{aligned} \tag{8.51}$$

A última linha da Eq. (8.51) indica que a *função Q* é o *complemento da função de distribuição Gaussiana normalizada*. A função Q é mostrada na Fig. 8.11.

Reiterando: com variáveis aleatórias Gaussianas, a média e a variância possuem importância particular, pois elas caracterizam completamente a função de distribuição.

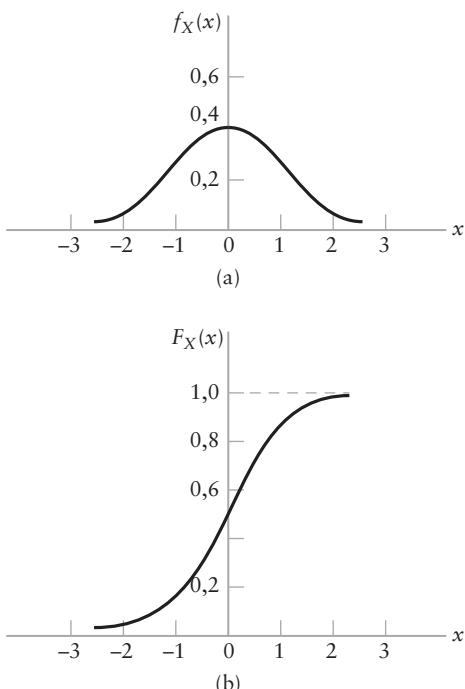


Figura 8.10 A distribuição Gaussiana normalizada. (a) A função densidade de probabilidade. (b) A função de distribuição.

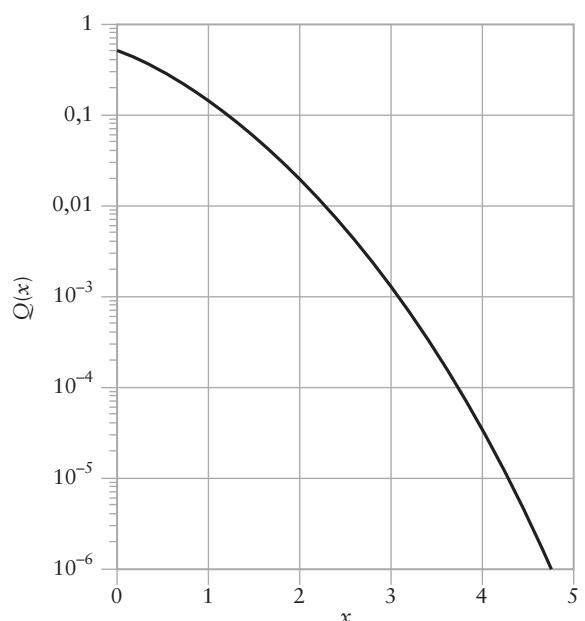


Figura 8.11 A função Q.

⁷ O Apêndice 4 apresenta valores tabulados Q(x) da função Q.

EXEMPLO 8.9 Probabilidade de erro de bit no PAM

Em um esquema de transmissão por modulação por amplitude de pulso (PAM), o dado binário é representado pelos níveis de tensão de $+A$ para um 1 e $-A$ para um 0. Suponha que um 1 é transmitido e recebido na presença de ruído Gaussiano com média zero e variância de σ^2 . Queremos determinar a probabilidade do bit ser incorretamente detectado.

O dado recebido pode ser representado pela variável aleatória Y , definida por

$$Y = A + N \quad (8.52)$$

na qual N é uma variável aleatória Gaussiana com média nula e variância σ^2 . Temos, então, das propriedades de variáveis aleatórias que Y também é uma variável aleatória Gaussiana mas com média de A e variância σ^2 , como ilustrado na Fig. 8.12.

A probabilidade de um erro ocorrer é a probabilidade de Y possuir valor menor do que zero. Esta probabilidade corresponde à área sombreada sob a curva da Fig. 8.12. Matematicamente, esta probabilidade é dada por

$$P[Y < 0] = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma} \exp\{-(y-A)^2/2\sigma^2\} dy \quad (8.53)$$

Se fizermos a mudança de variável fazendo

$$s = -\frac{y-A}{\sigma} \quad (8.54)$$

então a Eq. (8.53) se torna

$$\begin{aligned} P[Y < 0] &= \frac{1}{\sqrt{2\pi}} \int_{A/\sigma}^{\infty} \exp\{-s^2/2\} ds \\ &= Q\left(\frac{A}{\sigma}\right) \end{aligned} \quad (8.55)$$

na qual utilizamos a definição da função Q da Eq. (8.51). Iremos retornar a este exemplo no Capítulo 10.

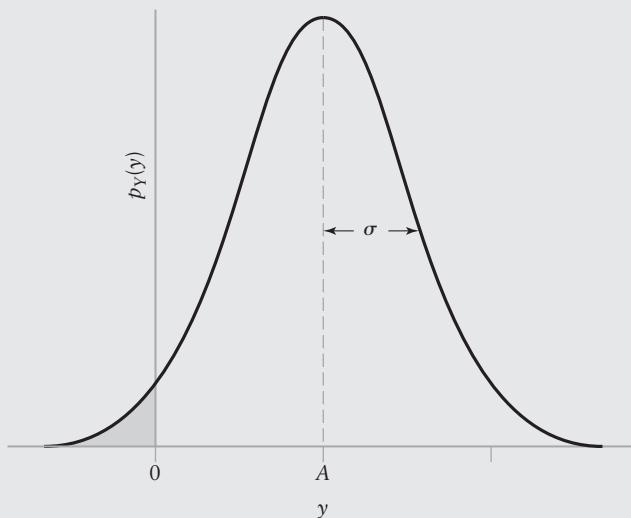


Figura 8.12 Função de densidade do sinal PAM, Y , ruidoso.

- **PROBLEMA 8.8** Mostre que a média e a variância de uma variável aleatória Gaussiana X com função de densidade dada pela Eq. (8.48) é μ_X e σ_X^2 , respectivamente. ◀
- **PROBLEMA 8.9** Mostre que para uma variável aleatória Gaussiana X com média μ_X e variância σ_X^2 , a transformação $Y = (X - \mu_X)/\sigma_X^2$ converte X em uma variável aleatória Gaussiana normalizada com média nula e variância unitária. ◀

8.5 TEOREMA DO LIMITE CENTRAL

Um resultado importante da teoria de probabilidade que está muito relacionado com a distribuição Gaussiana é o *teorema do limite central*. Seja X_1, X_2, \dots, X_n um conjunto de variáveis aleatórias com as seguintes propriedades:

1. Os X_k com $k = 1, 2, 3, \dots, n$ são estatisticamente independentes.
2. Todos os X_k possuem a mesma função densidade de probabilidade.
3. Tanto a média quanto a variância existem para cada X_k .

Não assumimos que a função de densidade de X_k é Gaussiana. Seja Y uma nova variável aleatória definida por

$$Y = \sum_{k=1}^n X_k \quad (8.56)$$

Então, de acordo com o teorema do limite central, a variável aleatória normalizada

$$Z = \frac{Y - E[Y]}{\sigma_Y} \quad (8.57)$$

aproxima-se de uma variável aleatória Gaussiana com média zero e variância unitária quando o número de variáveis aleatórias X_1, X_2, \dots, X_n aumenta sem limite. Ou seja, quando n se torna grande, a distribuição de Z aproxima-se da de uma variável aleatória Gaussiana com média nula e variância unitária, como mostrado por

$$F_Z(z) \rightarrow \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{s^2}{2}\right\} ds \quad (8.58)$$

Esta é a afirmação matemática do teorema do limite central. Em palavras, a *distribuição normalizada da soma de variáveis aleatórias independentes distribuídas identicamente aproxima-se de uma distribuição Gaussiana quando o número de variáveis aleatórias aumenta, independente das distribuições individuais*. Portanto, variáveis aleatórias Gaussianas são comuns porque elas caracterizam as propriedades assintóticas de vários outros tipos de variáveis aleatórias.

Quando n é finito, a aproximação Gaussiana da Eq. (8.58) é mais precisa na porção central da função de densidade (por isso limite central) e menos precisa nas “caudas” da função de densidade.

■ Experimento de computador: somas de variáveis aleatórias

No Problema 8.55, no final deste capítulo, descrevemos um experimento de computador para demonstrar o teorema do limite central. Neste experimento, consideraremos a variável aleatória

$$Z = \sum_{i=1}^N X_i$$

na qual X_i são variáveis aleatórias independentes, uniformemente distribuídas no intervalo de -1 a $+1$. No experimento de computador, calculamos 20.000 amostras de Z para $N = 5$

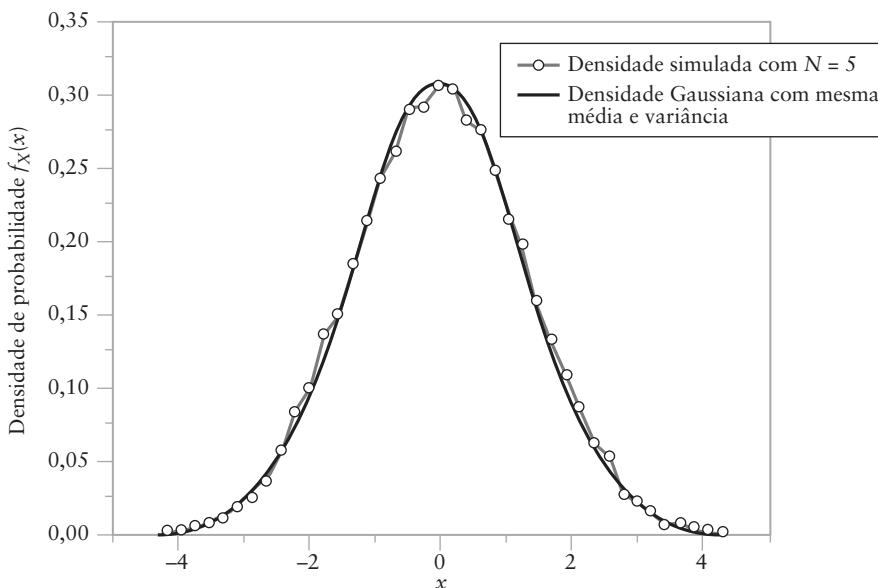


Figura 8.13 Comparação da densidade empírica de uma soma de cinco variáveis uniformes com uma densidade Gaussiana com mesma média e variância.

e estimamos a função de densidade correspondente formando um histograma do resultado. Na Fig. 8.13 comparamos este histograma (escalonado para área unitária) com a função de densidade Gaussiana com mesma média e variância. Como o gráfico indica, não é preciso muitas contribuições aleatórias para aproximar uma distribuição Gaussiana.

Os resultados deste experimento indicam quão poderoso é o teorema do limite central e explica porque modelos Gaussianos são amplamente utilizados na análise de sinais aleatórios em comunicações e várias outras áreas. Um bônus é a capacidade de tratamento matemático de variáveis aleatórias Gaussianas e o fato deste tipo de distribuição permanecer inalterado após o processamento linear, mas, obviamente, com a média e a variância da distribuição Gaussiana sujeita à mudança.

- **PROBLEMA 8.10** Determine a média e a variância da soma de cinco variáveis aleatórias independentes distribuídas uniformemente no intervalo de -1 a +1. ◀

8.6 PROCESSOS ALEATÓRIOS

Em um sistema de comunicação via rádio, o sinal recebido geralmente contém uma componente do sinal contendo a informação, uma componente de interferência aleatória e o ruído do canal. O sinal contendo a informação pode ser representado, por exemplo, por um sinal de voz que, tipicamente, consiste de rajadas espaçadas aleatoriamente de energia e com duração aleatória. A componente de interferência pode representar ondas eletromagnéticas espúrias produzidas por outros sistemas de comunicação operando na proximidade do rádio receptor. A principal fonte de ruído de canal é o ruído térmico, o qual é causado pelo movimento aleatório de elétrons em condutores e dispositivos na frente do receptor. Vemos, portanto, que o sinal variante no tempo recebido é aleatório por natureza. Nesta seção, iremos combinar os conceitos de variação temporal e variáveis aleatórias para introduzir o conceito de *processos aleatórios*. Apesar de não ser possível predizer exatamente o valor de sinais e processos aleatórios, é possível descrever o sinal em termos de seus parâmetros estatísticos, tal como potência média e densidade espectral de potência, como será mostrado neste capítulo.

Processos aleatórios representam o modelo matemático formal destes sinais aleatórios. Da discussão anterior, processos aleatórios possuem as seguintes propriedades:

1. Processos aleatórios são funções do tempo.
2. Processos aleatórios são aleatórios no sentido de não ser possível predizer exatamente qual forma de onda será observada no futuro.

Análogo a variáveis aleatórias, quando discutimos um experimento envolvendo processos aleatórios é conveniente pensar em termos de um espaço de amostra. Especificamente, cada saída do experimento é associado com um ponto de amostra. Entretanto, neste caso, cada ponto de amostra representa uma função variante no tempo. A agregação de todas as possíveis saídas do experimento recebe o nome de espaço de amostra, espaço amostral ou processo aleatório. Reiterando, cada ponto do espaço de amostra ou espaço *amostral* é uma função do tempo. Como parte integral da definição de processo aleatório, assumimos a existência de uma distribuição de probabilidade sobre os conjuntos apropriados do espaço de amostra, de tal forma que podemos falar da probabilidade de vários eventos.

Considere um experimento aleatório especificado pela saída s de um espaço amostral S e as probabilidades destes eventos. Suponha que associemos a cada ponto de amostra s uma função do tempo com o rótulo

$$X(t, s), \quad -T < t < T \quad (8.59)$$

na qual $2T$ é o período total de observação. Para um ponto de amostra s_j fixo, a função de $X(t, s_j)$ é chamada de *realização* ou *função amostra* do processo aleatório. Para simplificar a notação, representamos esta função amostra por

$$x_j(t) = X(t, s_j) \quad (8.60)$$

A Fig. 8.14 ilustra um conjunto de funções de amostra $\{s_j(t): j = 1, 2, \dots\}$. A partir desta figura, notamos que em um t_k fixo dentro da janela de observação, o conjunto de números

$$\{x_1(t_k), x_2(t_k), \dots, x_n(t_k)\} = \{X(t_k, s_1), X(t_k, s_2), \dots, X(t_k, s_n)\}$$

é uma variável aleatória. Portanto, temos um espaço amostral (família) de variáveis aleatórias $\{X(t, s)\}$, o qual é chamado de processo aleatório. Para simplificar a notação, é conveniente suprimir s e usar $X(t)$ para representar o processo aleatório.

Para comparar:

- Com uma variável aleatória, a saída do experimento aleatório é mapeada em um número real.
- Com um processo aleatório, a saída do experimento aleatório é mapeada em uma forma de onda que é uma função do tempo.

Em qualquer ponto da janela de observação, t_k , as possíveis saídas de um processo aleatório podem ser representadas por uma variável aleatória. A *família de todas as variáveis aleatórias, indexadas pela variável de tempo t, forma o processo aleatório..*

A faixa de possíveis processos aleatórios é muito grande. Para restringir esta faixa a processos aleatórios que são tanto: (i) típicos em situações reais e (ii) matematicamente tratáveis, precisamos de duas condições técnicas, *estacionariedade* e *ergodicidade*, as quais discutiremos a seguir.

■ Processos aleatórios estacionários

Com processos aleatórios reais, geralmente vemos que a caracterização estatística de um processo é independente do tempo no qual a observação ocorre. Ou seja, se um processo aleatório é dividido em vários intervalos de tempo, as várias sessões do processo exibem essencialmente as mesmas propriedades estatísticas. Tal processo é dito ser *estacionário*. Caso contrário, ele é *não estacionário*.

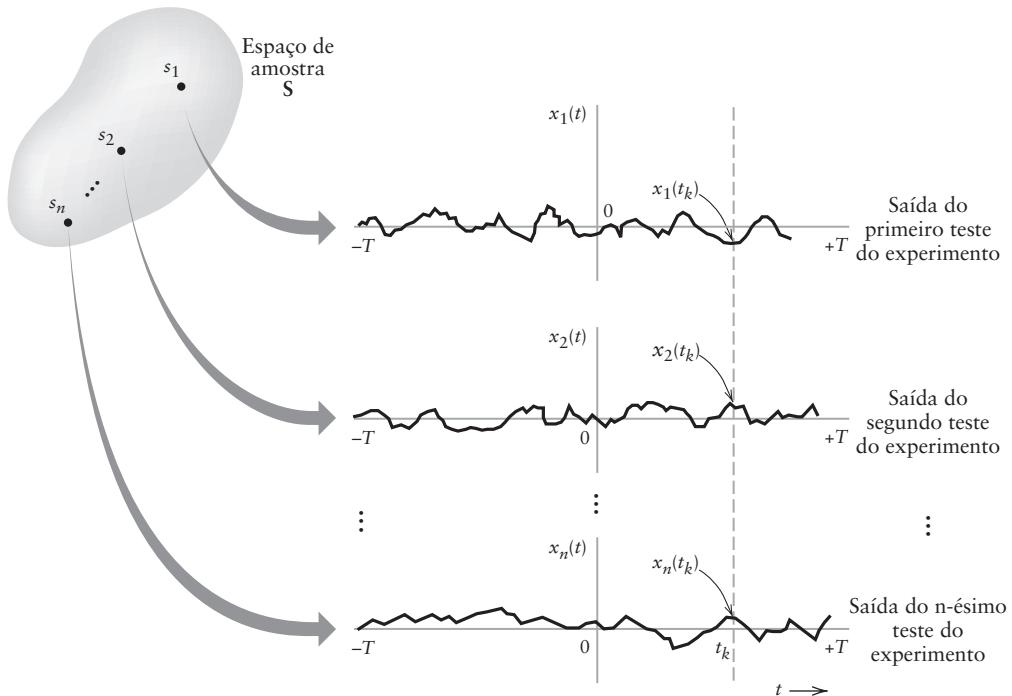


Figura 8.14 Ilustração da relação entre espaço de amostra e família de funções de amostra.

Esta caracterização é um pouco análoga à teoria de sistemas lineares, na qual definimos sistemas invariantes no tempo como aqueles cuja resposta ao impulso não varia com o tempo, em contraste a sistemas variantes no tempo, cuja resposta ao impulso varia com o tempo. Tal como em sistemas invariantes no tempo, processos aleatórios estacionários são, geralmente, mais fáceis de serem trabalhados matematicamente e mais simples de serem analisados.

Para sermos mais precisos, seja $X(t)$ um processo aleatório que é observado no tempo t_1 . Seja $F_{X(t_1)}(x)$ a função distribuição de probabilidade associada com observações de funções de amostra diferentes do processo aleatório no tempo t_1 . Suponha que o mesmo processo aleatório é observado no tempo $t_1 + \tau$, e que a função de distribuição correspondente é $F_{X(t_1 + \tau)}(x)$. Então, se

$$F_{X(t_1 + \tau)}(x) = F_{X(t_1)}(x) \quad (8.61)$$

para todo t_1 e todo τ , dizemos que o processo é *estacionário de primeira ordem*. Um processo aleatório estacionário de primeira ordem possui uma função de distribuição que é independente do tempo. Como consequência, parâmetros estatísticos tais como média e variância também são independentes do tempo para tal processo. Por exemplo, suponha que $F_{X(t_1)}(x)$ possua a densidade $f_{X(t_1)}(x)$. Então o valor médio

$$\mu_X = \int_{-\infty}^{\infty} sf_{X(t_1)}(s) ds = \int_{-\infty}^{\infty} sf_{X(t_1 + \tau)}(s) ds \quad (8.62)$$

não varia com o tempo, pois a função de distribuição (e logo a densidade) é invariante no tempo.

A seguir, considere a amostragem do processo aleatório $X(t)$ em dois pontos no tempo t_1 e t_2 , com a correspondente função de distribuição comum $F_{X(t_1), X(t_2)}(x_1, x_2)$. Suponha um segundo conjunto de observações feitas no tempo $t_1 + \tau$ e $t_2 + \tau$ e que a correspondente

função de distribuição comum é $F_{X(t_1+\tau), X(t_2+\tau)}(x_1, x_2)$. Então, se para todo t_1, t_2 e τ observamos que

$$F_{X(t_1+\tau), X(t_2+\tau)}(x_1, x_2) = F_{X(t_1), X(t_2)}(x_1, x_2) \quad (8.63)$$

dizemos que o processo é *estacionário de segunda ordem*. A estacionariedade de segunda ordem implica que quantidades estatísticas tais como covariância e correlação, a qual iremos discutir a seguir, não dependem do tempo absoluto.

A definição de processo aleatório de k -ésima ordem segue de maneira similar. Se a equivalência entre as funções de distribuição forem válidas para todos os deslocamentos de tempo τ , todo k e todos os possíveis tempos de observação t_1, t_2, \dots, t_k , então dizemos que o processo é *estritamente estacionário*. Em outras palavras, um processo aleatório $X(t)$ é estritamente estacionário se a distribuição comum de qualquer conjunto de variáveis aleatórias obtidas pela observação do processo aleatório $X(t)$ for invariante com respeito à localização da origem $t = 0$.

- **PROBLEMA 8.11** Um processo aleatório é definido pela função

$$X(t, \theta) = A \cos(2\pi ft + \theta)$$

na qual A e f são constantes e θ é uniformemente distribuído no intervalo de 0 a 2π . X é estacionário de primeira ordem? ◀

- **PROBLEMA 8.12** Mostre que se um processo aleatório é estacionário de segunda ordem ele também é estacionário de primeira ordem. ◀

8.7 CORRELAÇÃO DE PROCESSOS ALEATÓRIOS

Enquanto processos aleatórios são, por definição, imprevisíveis, geralmente observamos que as amostras do processo em tempos diferentes podem ser correlacionadas. Por exemplo, se $X(t_1)$ for grande, então também podemos esperar que $X(t_1 + \tau)$ seja grande, se τ for pequeno. Para quantificar esta relação, considere a covariância de duas variáveis aleatórias, definida na Seção 8.2, aplicada a amostras do processo aleatório $X(t)$ nos tempos t_1 e t_2 . Ou seja, a covariância de duas variáveis aleatórias $X(t_1)$ e $X(t_2)$ é dada por

$$\text{Cov}(X(t_1), X(t_2)) = \mathbb{E}[X(t_1)X(t_2)] - \mu_{X(t_1)}\mu_{X(t_2)} \quad (8.64)$$

Definimos o primeiro termo do lado direito da Eq. (8.64) como a *autocorrelação* do processo aleatório e utilizamos a notação genérica

$$R_X(t, s) = \mathbb{E}[X(t)X^*(s)] \quad (8.65)$$

na qual utilizamos o asterisco para representar o conjugado complexo para quando $X(t)$ for de valor complexo. Se $X(t)$ é estacionário de segunda ordem ou mais, então a Eq. (8.64) pode ser escrita por

$$\begin{aligned} R_X(t, s) &= \mathbb{E}[X(t)X^*(s)] \\ &= R_X(t - s) \end{aligned} \quad (8.66)$$

A estacionariedade de segunda ordem também implica que a média do processo aleatório é constante. Se a média é zero, então as funções de autocorrelação e covariância do processo aleatório são equivalentes. A seguir será mostrada a importância da função de autocorrelação como descritor de processos aleatórios.

Para várias aplicações, não precisamos que um processo aleatório tenha todas as propriedades necessárias para ser estacionário de segunda ordem. Em particular geralmente precisamos apenas que:

1. A média do processo aleatório é uma constante independente do tempo: $E[X(t)] = \mu_X$ para todo t .
2. A autocorrelação de um processo aleatório depende apenas da diferença de tempo: $E[X(t)X^*(t - \tau)] = R_X(\tau)$, para todo t e τ .

Se um processo aleatório possui estas duas propriedades, então podemos dizer que ele é *estacionário de sentido amplo ou fracamente estacionário*. Note que a estacionariedade de sentido amplo não implica em estacionariedade de segunda ordem. Nem estacionariedade de segunda ordem implica em estacionariedade de sentido amplo, pois os momentos de primeira e segunda ordem podem não existir. No restante do livro, iremos assumir que todos os processos de interesse são estacionários em sentido amplo.

■ Propriedades da função de autocorrelação

A função de autocorrelação de um processo aleatório estacionário em sentido amplo possui as seguintes propriedades para um processo de valor real:

- **PROPRIEDADE 1 – POTÊNCIA DE UM PROCESSO ESTACIONÁRIO EM SENTIDO AMPLO** O segundo momento ou valor médio quadrático de um processo aleatório de valor real é dado por

$$\begin{aligned} R_X(0) &= E[X(t)X(t)] \\ &= E[X^2(t)] \end{aligned} \quad (8.67)$$

O valor médio quadrático é, portanto, equivalente à potência média do processo.

- **PROPRIEDADE 2 – SIMETRIA** A correlação de um processo aleatório estacionário em sentido amplo de valor real possui simetria par.

Para mostrar isto, considere

$$\begin{aligned} R_X(-\tau) &= E[X(t)X(t + \tau)] \\ &= E[X(t + \tau)X(t)] \\ &= R_X(\tau) \end{aligned} \quad (8.68)$$

- **PROPRIEDADE 3 – VALOR MÁXIMO** A função de autocorrelação de um processo aleatório estacionário em sentido amplo é máxima na origem.

Para mostrar esta propriedade para um processo de valor real, forme quantidade não negativa

$$\begin{aligned} 0 &\leq E[(X(t) \pm X(t - \tau))^2] \\ &\leq E[X^2(t)] + E[X^2(t - \tau)] \pm 2E[X(t)X(t - \tau)] \\ &\leq 2R_X(0) \pm 2R_X(\tau) \end{aligned} \quad (8.69)$$

Reorganizando a última relação, temos $R_X(0) \geq |R_X(\tau)|$.

O significado físico da função de autocorrelação $R_X(\tau)$ é que ela fornece uma maneira de descrever a interdependência de duas variáveis aleatórias obtidas pela observação do processo aleatório $X(t)$ τ segundos separados uma da outra. É, portanto, aparente que quanto mais rapidamente o processo aleatório $X(t)$ varia como o tempo, mais rapidamente a função de autocorrelação $R_X(\tau)$ irá diminuir de seu máximo $R_X(0)$ quando τ aumenta, como ilustrado na Fig. 8.15. Este decaimento pode ser caracterizado pelo *tempo de decorrelação* τ_0 , especificamente, para $\tau > \tau_0$, a amplitude da correlação $R_X(\tau)$ permanece abaixo de um valor preestabelecido. Podemos, portanto, definir o tempo de decorrelação τ_0 de um processo estacionário $X(t)$ de média zero como o tempo para que a amplitude da função de autocorrelação $R_X(\tau)$ diminua, digamos, para 1% de seu valor máximo em $R_X(0)$.

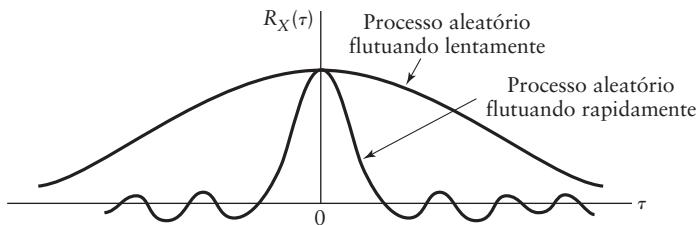


Figura 8.15 Ilustração de funções de autocorrelação de processos aleatórios com flutuações rápida e lenta.

EXEMPLO 8.10 Autocorrelação de um processo de cosseno aleatório

Seja $X(t)$ um processo aleatório definido por

$$X(t) = A \cos(2\pi ft + \theta)$$

na qual a amplitude A e a freqüência f são conhecidas, mas θ é uniformemente distribuído no intervalo entre 0 e 2π . Este é um tipo especial de processo aleatório no qual um único parâmetro θ define a função amostra para todo tempo. O requerido é determinar a autocorrelação de $X(t)$.

A autocorrelação é dada por

$$\begin{aligned} R_X(t, t - \tau) &= E[X(t)X(t - \tau)] \\ &= A^2 E[\cos(2\pi ft + \theta) \cos(2\pi f(t - \tau) + \theta)] \end{aligned}$$

Aplicando a propriedade trigonométrica $2 \cos A \cos B = \cos(A - B) + \cos(A + B)$ a esta relação, obtemos,

$$R_X(t, t - \tau) = \frac{A^2}{2} \cos(2\pi f\tau) + \frac{A^2}{2} E[\cos(4\pi ft - 2\pi f\tau + 2\theta)]$$

Como θ é uniformemente distribuído entre 0 e 2π , temos

$$\begin{aligned} E[\cos(4\pi ft - 2\pi f\tau + 2\theta)] &= \frac{1}{2\pi} \int_0^{2\pi} \cos(4\pi ft - 2\pi f\tau + 2\theta) d\theta \\ &= \frac{1}{4\pi} \sin(4\pi ft - 2\pi f\tau + 2\theta) \Big|_0^{2\pi} \\ &= 0 \end{aligned}$$

Conseqüentemente, a expressão para a autocorrelação se reduz para

$$R_X(t, t - \tau) = \frac{A^2}{2} \cos(2\pi f\tau)$$

A autocorrelação claramente depende apenas da diferença de tempo τ neste exemplo e o processo pode ser mostrado como sendo estacionário em sentido amplo.

► **PROBLEMA 8.13** Seja $X(t)$ um processo aleatório definido por

$$X(t) = A \cos(2\pi ft)$$

na qual A é uniformemente distribuído entre 0 e 1 e f é constante. Determine a função de autocorrelação de X . X é estacionário em sentido amplo? ◀

■ Ergodicidade

Para determinar as propriedades estatísticas de um processo aleatório, geralmente precisamos calcular esperanças. A esperança de um processo aleatório em um ponto particular no tempo requer realizações independentes separadas do processo aleatório. Por exemplo, para um processo aleatório $X(t)$ com N realizações equiprováveis $\{x_j(t): j = 1, 2, \dots, N\}$, o valor esperado e o segundo momento do processo aleatório no tempo $t = t_k$ são, respectivamente, dados pelas *médias da família*

$$\mathbb{E}[X(t_k)] = \frac{1}{N} \sum_{j=1}^N x_j(t_k) \quad (8.70)$$

e

$$\mathbb{E}[X^2(t_k)] = \frac{1}{N} \sum_{j=1}^N x_j^2(t_k) \quad (8.71)$$

Se o processo é estacionário em sentido amplo, então o valor médio e o segundo momento calculados por estas duas equações não dependem do tempo t_k .

Em problemas práticos envolvendo processos aleatórios, o que geralmente estará disponível para o usuário não é o processo aleatório, mas uma de suas funções de amostra $x(t)$. Em tais casos, os parâmetros mais facilmente medidos são as médias temporais. Por exemplo, a *média temporal* de uma função amostra contínua obtida de um processo de valor real é dada por

$$\mathcal{E}[x] = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt \quad (8.72)$$

e a *autocorrelação no tempo da função amostra* é dada por

$$\mathcal{R}_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t - \tau) dt \quad (8.73)$$

Então a questão é: *Quando as médias temporais de uma função amostra são iguais às médias do espaço amostral do processo aleatório correspondente?* Intuitivamente, se as estatísticas do processo aleatório $X(t)$ não mudarem com o tempo, então podemos esperar que as médias temporais e as médias do espaço amostral sejam equivalentes.

Dependendo das propriedades estacionárias do processo aleatório, várias médias temporais das funções de amostra podem ser utilizadas para aproximar as médias ou esperanças do espaço amostral correspondente. Processos aleatórios para os quais esta equivalência é válida são ditos serem *ergódicos*. Na maioria das aplicações físicas, os *processos estacionários em sentido amplo são ergódicos*, e, neste caso, as *médias temporais e esperanças podem ser utilizadas uma no lugar da outra*.

O leitor alerta irá notar que, tal como a estacionariedade, existem vários graus de ergodicidade. As equivalências das Eqs. (8.70) e (8.72) por um lado e as Eqs. (8.66) e (8.73) correspondentes por outro lado são análogas à forma de ergodicidade de primeira e segunda ordem.

Além disso, se assumirmos que o processo aleatório de valor real é ergódico, então podemos expressar a função de autocorrelação por

$$\begin{aligned} R_X(\tau) &= \mathbb{E}[X(t)X(t - \tau)] \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t - \tau) dt \end{aligned} \quad (8.74)$$

na qual $x(t)$ é uma função amostra do processo aleatório $X(t)$. Esta definição de autocorrelação é idêntica à definição de correlação para sinais de potência determinísticos como

descrito no Capítulo 2. Como consequência, a autocorrelação de um processo aleatório ergódico possui as mesmas propriedades da autocorrelação de sinais determinísticos.

O conceito de ergodicidade também leva naturalmente à idéia de estimadores para a função de autocorrelação. Em particular, se $x(t)$ é uma função amostra de um processo ergódico estacionário em sentido amplo $X(t)$, então uma *estimativa da autocorrelação de um processo de valor real para o atraso $\tau = \tau_0$* é (veja a Fig. 2.29)

$$\hat{R}_X(\tau_0) = \frac{1}{N} \sum_{n=1}^N x(t_n)x(t_n - \tau_0) \quad (8.75)$$

na qual $\{t_n\}$ é um conjunto conveniente de tempos de amostragem, uniformemente espaçados ou não. Similar aos estimadores de média e variância de uma variável aleatória, esta estimativa de autocorrelação é motivada pela definição de freqüência relativa da probabilidade.

EXEMPLO 8.11 Autocorrelação em tempo discreto

Em sistemas de comunicação digital, geralmente amostramos um sinal contínuo em tempos discretos $t_n = nT_s$, na qual T_s é o intervalo de amostragem. Este sinal amostrado $\{X(t_1), X(t_2), \dots\}$ é um processo aleatório em tempo discreto para o qual podemos definir a função de autocorrelação em tempo discreto

$$R_X(kT_s) = E[X(nT_s)X((n - k)T_s)] \quad (8.76)$$

na qual a esperança é a média usual do espaço amostral. Como sugerido nesta seção, podemos querer estimar esta autocorrelação calculando a média temporal em alguma função amostra particular. Esta média temporal, em N amostras da função amostra $x(t)$, é definida por

$$\mathcal{R}_x(kT_s) = \frac{1}{N} \sum_{n=1}^N x(nT_s)x((n - k)T_s) \quad (8.77)$$

Na Fig. 8.16(a), mostramos o sinal aleatório como uma função do tempo. Para este sinal, a autocorrelação da média temporal do sinal aleatório é mostrada na Fig. 8.16(b). O eixo horizontal representa

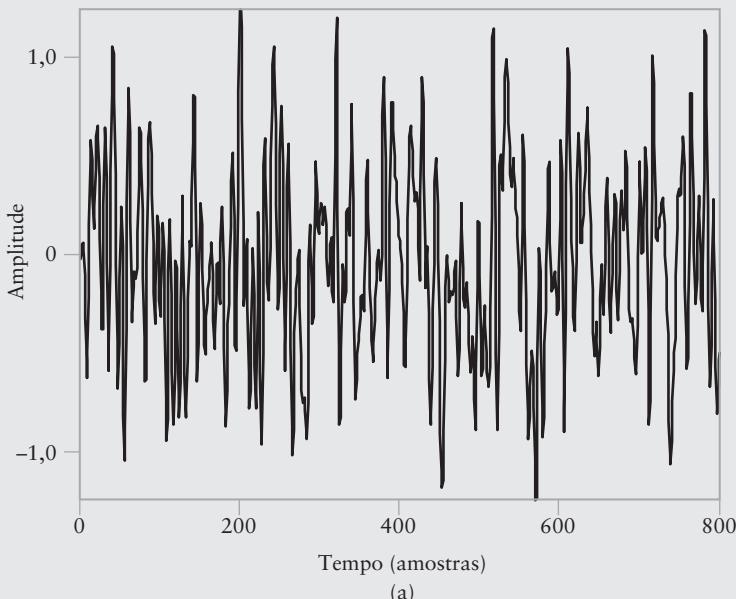


Figura 8.16 Ilustração de (a) sinal aleatório e (b) sua autocorrelação (*continua*).

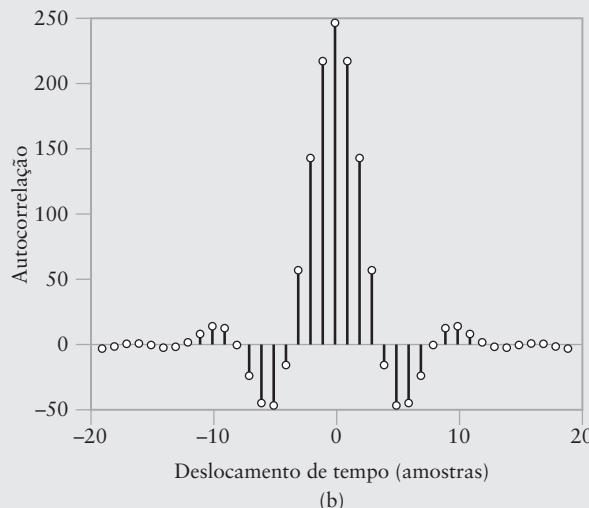


Figura 8.16 Continuação.

o atraso de tempo τ em períodos de amostragem. Apesar do sinal ser bem aleatório no domínio do tempo, ele possui uma função de autocorrelação suave, como mostrado na Fig. 8.16(b).

- **PROBLEMA 8.14** Um processo aleatório em tempo discreto $\{Y_n; n = \dots, -1, 0, 1, 2, \dots\}$ é definido por

$$Y_n = \alpha_0 Z_n + \alpha_1 Z_{n-1}$$

na qual $\{Z_n\}$ é um processo aleatório com função de autocorrelação $R_Z(n) = \sigma^2 \delta(n)$, na qual $\delta(n)$ é a função delta. Qual é a função de autocorrelação $R_Y(n, m) = E[Y_n Y_m]$? O processo $\{Y_n\}$ é estacionário em sentido amplo? ◀

8.8 ESPECTRO DE SINAIS ALEATÓRIOS

Como descrito anteriormente, um sinal aleatório pode ser visto como pertencendo a uma família de sinais, gerado por algum mecanismo probabilístico. Logo, nenhum par de sinais exibe as mesmas variações no tempo. Seja $x(t)$ uma função amostra do processo aleatório $X(t)$. A Fig. 8.17 mostra um gráfico da forma de onda de $x_T(t)$ no intervalo $-T < t < T$. Podemos definir a transformada de Fourier da função amostra $x_T(t)$ como

$$\xi_T(f) = \int_{-\infty}^{\infty} x_T(t) \exp(-j2\pi ft) dt \quad (8.78)$$

Esta transformada converte a função amostra $x_T(t)$ em uma nova função amostra $\xi_T(f)$.⁸ Definimos a coleção (família) de todas estas novas funções amostra como o novo processo aleatório $\Xi_T(f)$. Efetivamente, a transformada de Fourier converteu a família de variáveis aleatórias $X(t)$ indexadas pelo parâmetro t para uma nova família de variáveis aleatórias $\Xi_T(f)$ indexadas pelo parâmetro f .

⁸ Utilizamos anteriormente a notação $x(t) = X(f)$ para representar um par transformada de Fourier. Entretanto, a utilização de $X(t)$ para representar o processo aleatório introduz a possibilidade de ambigüidade na notação. Logo, nesta seção, utilizaremos $x(t) = \xi(f)$ como par transformada de Fourier de funções amostra e $X(t) = \Xi_T(f)$ como par transformada de Fourier de processos aleatórios.

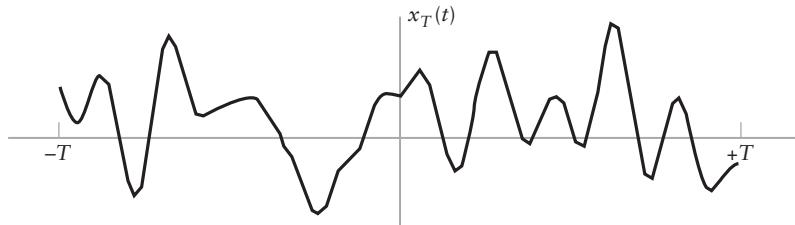


Figura 8.17 Função amostra de um processo aleatório.

Da discussão do Capítulo 2, lembramos que a densidade espectral de potência da função amostra $x(t)$ no intervalo $-T < t < T$ é $|\xi_T(f)|^2/2T$, na qual $\xi_T(f)$ é a transformada de Fourier de $x_T(t)$. Esta densidade espectral de potência irá depender da função amostra $x(t)$ particular retirada do espaço amostral. Desta forma, para obter a densidade espectral de potência de um processo aleatório, devemos executar uma operação de média de espaço amostral e, então, calcular o limite quando T tende ao infinito.

A operação de média do espaço amostral requer a utilização da distribuição de probabilidade do espaço. O valor de f é mantido fixo, enquanto calculamos a média para a família. Para a discussão atual, é suficiente saber que a operação de média no espaço amostral é realizada utilizando o operador E , de esperança. Portanto, escrevemos o valor da média da família do novo processo aleatório $|\Xi_T(f)|^2$ como sendo $E[|\Xi_T(f)|^2]$ e a densidade espectral de potência correspondente do processo aleatório $X(t)$ como sendo

$$S_X(f) = \lim_{T \rightarrow \infty} \frac{1}{2T} E[|\Xi_T(f)|^2] \quad (8.79)$$

É importante notar que na Eq. (8.79), a média no espaço amostral deve ser calculada antes do limite ser determinado. O fato do processo ser estacionário em sentido amplo também é crítico para esta definição.

Esta abordagem de determinação da densidade espectral de potência através da média do espaço amostral pode ser utilizada para motivar um estimador da densidade espectral de potência dada uma função amostra de um processo estacionário ergódico. Na Seção 2.10, vimos como a transformada discreta de Fourier (TDF) pode ser utilizada para aproximar numericamente a transformada de Fourier. Em particular, se $\{x_n: n = 0, 1, \dots, N-1\}$ são amostras uniformemente espaçadas de uma função $x(t)$ para $t = nT_s$, então a transformada discreta de Fourier é definida por

$$\xi_k = \sum_{n=0}^{N-1} x_n W^{kn} \quad (8.80)$$

na qual $W = \exp(-j2\pi/N)$ e $\{\xi_k\}$ são amostras do domínio da freqüência em $f = k/NT_s$. Conseqüentemente, podemos estimar a densidade espectral de potência de um processo aleatório seguindo três passos:

1. Particionar a função amostra $x(t)$ em M seções de tamanho NT_s e amostrar-las em intervalos T_s .
2. Executar a TDF em cada seção de tamanho NT_s . Seja $\{\xi_{k-mN}\}$, $m = 0, \dots, M-1$, a representação de M saídas da TDF, um conjunto para cada seção.
3. Calcular a média do quadrado da amplitude de cada TDF e, então, a estimativa da densidade espectral de potência será dada por

$$\begin{aligned} \hat{S}_X\left(\frac{k}{NT_s}\right) &= \frac{1}{M} \sum_{m=0}^{M-1} |\xi_{k+mN}|^2 \\ &= \frac{1}{M} \sum_{m=0}^{M-1} \left| \sum_{n=0}^{N-1} x_{n+mN} W^{kn} \right|^2, \quad k = 0, \dots, M-1 \end{aligned} \quad (8.81)$$

Este resultado é baseado, claramente, na consideração do processo ser ergódico.

■ Propriedades da densidade espectral de potência

A densidade espectral de potência $S_X(f)$ e a função de autocorrelação $R_X(\tau)$ de um processo aleatório estacionário em sentido amplo formam um par transformada de Fourier nas variáveis f e τ . Em particular, as mesmas relações de Weiner-Khintchine aplicadas a processos determinísticos também relacionam a densidade espectral de potência e a função de autocorrelação de um processo aleatório, como mostrado por

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f\tau) d\tau \quad (8.82)$$

e

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f) \exp(j2\pi f\tau) df \quad (8.83)$$

As Eqs. (8.82) e (8.83) são as relações básicas da teoria de análise espectral de processos aleatórios. As relações de Weiner-Khintchine mostram que se a autocorrelação ou a densidade espectral de potência de um processo aleatório for conhecida, a outra pode ser exatamente determinada.

Podemos utilizar este par de relações para obter algumas propriedades genéricas da densidade espectral de potência de um processo estacionário em sentido amplo.

- **PROPRIEDADE 1 – VALOR MÉDIO QUADRÁTICO** *O valor médio quadrático de um processo estacionário é igual à área total sob o gráfico da densidade espectral de potência, ou seja,*

$$\mathbb{E}[|X(t)|^2] = \int_{-\infty}^{\infty} S_X(f) df \quad (8.84)$$

Esta propriedade é obtida diretamente da Eq. (8.83) fazendo $\tau = 0$ e observando que

$$R_X(0) = \mathbb{E}[|X(t)|^2]$$

- **PROPRIEDADE 2 – NÃO NEGATIVIDADE** *A densidade espectral de potência de um processo aleatório estacionário é sempre não negativa, ou seja,*

$$S_X(f) \geq 0, \quad \text{para todo } f \quad (8.85)$$

Esta propriedade é obtida da definição de densidade espectral de potência dada pela Eq. (8.79).

- **PROPRIEDADE 3 – SIMETRIA** *A densidade espectral de potência de um processo aleatório real é uma função par da freqüência, ou seja,*

$$S_X(-f) = S_X(f) \quad (8.86)$$

Esta propriedade é facilmente obtida substituindo, primeiro, $-f$ por f na Eq. (8.82).

$$S_X(-f) = \int_{-\infty}^{\infty} R_X(\tau) \exp(j2\pi f\tau) d\tau$$

A seguir, substituindo $-\tau$ por τ e percebendo que $R_X(-\tau) = R_X(\tau)$ para todo processo real, obtemos

$$S_X(-f) = \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f\tau) d\tau = S_X(f)$$

o qual é o resultado desejado.

- **PROPRIEDADE 4 – PROCESSOS ALEATÓRIOS FILTRADOS** Se um processo aleatório estacionário $X(t)$ com espectro $S_X(f)$ for passado através de um filtro linear com resposta em freqüência $H(f)$, o espectro do processo aleatório estacionário de saída $Y(t)$ é dado por

$$S_Y(f) = |H(f)|^2 S_X(f) \quad (8.87)$$

Este resultado é análogo ao obtido para sinais determinísticos.⁹

EXEMPLO 8.12 Filtragem de uma senóide aleatória

Um sinal aleatório $X(t)$ com função de autocorrelação

$$R_X(\tau) = \frac{A^2}{2} \cos(2\pi f_c \tau)$$

é processado por um filtro causal com resposta $h(t)$ ao impulso que é zero para tempos negativo, para produzir um novo processo aleatório

$$Y(t) = \int_0^t h(t-s)X(s) ds$$

Se o filtro possuir resposta ao impulso correspondendo à resposta em freqüência

$$H(f) = \frac{1}{1 + j2\pi RCf} \quad (8.88)$$

qual é a função de autocorrelação de $Y(t)$?

Da Propriedade 4 de densidades espetrais de potência de processos aleatórios, temos que a densidade espectral de $Y(t)$ é dada por

$$S_Y(f) = |H(f)|^2 S_X(f)$$

Podemos utilizar o par transformada de Fourier

$$\cos(2\pi f_c \tau) \iff \frac{1}{2} [\delta(f - f_c) + \delta(f + f_c)] \quad (8.89)$$

para obter a densidade espectral $S_X(f)$. Logo,

$$\begin{aligned} S_Y(f) &= \frac{1}{2} |H(f)|^2 [\delta(f - f_c) + \delta(f + f_c)] \\ &= \frac{1}{2} |H(f_c)|^2 [\delta(f - f_c) + \delta(f + f_c)] \end{aligned}$$

na qual a última linha antecipa a propriedade de deslocamento da função delta de Dirac. Utilizando a transformada inversa de Fourier de $S_Y(f)$ e calculando $H(f)$ da Eq. (8.88) para $f = f_c$, obtemos

$$R_Y(\tau) = \frac{1}{1 + (2\pi RCf_c)^2} \cos(2\pi f_c \tau)$$

Vemos, portanto, que, de forma similar a senóides determinísticas, a filtragem linear não afeta a freqüência da senóide aleatória.

⁹ A prova deste resultado para sinais determinísticos é obtida do teorema da convolução da Seção 2.2. A prova para processos aleatórios pode ser encontrada no Capítulo 1 de Haykin (2001).

- **PROBLEMA 8.15** Para o processo em tempo discreto do Problema 8.14, utilize a transformada discreta de Fourier para aproximar o espectro correspondente. Ou seja, calcule e trace

$$S_Y(k) = \sum_{n=0}^{N-1} R_Y(n) W^{kn}$$

Se a amostragem no domínio do tempo acontece em n/T_s , onde $n = 0, 1, 2, \dots, N-1$, a qual frequência o inteiro k corresponde? ◀

8.9 PROCESSOS GAUSSIANOS

Até aqui, apresentamos a teoria de processos aleatórios em termos gerais. No restante deste capítulo, iremos considerar esta teoria no contexto de alguns processos aleatórios importantes que são geralmente encontrados no estudo de sistemas de comunicação.

Vimos como variáveis aleatórias Gaussianas possuem um papel importante em sistemas de comunicação pois (i) vários processos físicos que geram ruído em sistemas de comunicação podem ser considerados como tendo, aproximadamente, uma distribuição Gaussiana e (ii) variáveis aleatórias Gaussianas podem ser matematicamente tratadas e, portanto, são convenientes de serem utilizadas.

Similarmente, um processo Gaussiano possui um importante papel no estudo de processos aleatórios por duas razões. Primeiro, processos Gaussianos possuem várias propriedades que tornam resultados analíticos possíveis. Segundo, os processos aleatórios produzidos por fenômenos físicos são, geralmente, tais que o modelo Gaussiano é apropriado.

Antes de definirmos um processo Gaussiano, devemos fornecer alguma base sobre a distribuição comum de múltiplas variáveis aleatórias Gaussianas. A distribuição comum de N variáveis aleatórias Gaussianas pode ser escrita como¹⁰

$$f_X(\mathbf{x}) = \frac{1}{(2\pi)^{N/2} |\Lambda|^{1/2}} \exp\{-(\mathbf{x} - \boldsymbol{\mu}) \Lambda^{-1} (\mathbf{x} - \boldsymbol{\mu})^T / 2\} \quad (8.90)$$

a qual é chamada de *distribuição Gaussiana multivariável*. Na Eq. 8.90, as grandezas respectivas são:

$\mathbf{X} = (X_1, X_2, \dots, X_N)$ representa o vetor N -dimensional de variáveis aleatórias Gaussianas.

$\mathbf{x} = (x_1, x_2, \dots, x_N)$ é o vetor correspondente de indeterminadas.

$\boldsymbol{\mu} = (\mathbb{E}[X_1], \mathbb{E}[X_2], \dots, \mathbb{E}[X_N])$ é o vetor N -dimensional de médias.

Λ é a matriz N por N de covariância com elementos individuais dados por $\Lambda_{ij} = \text{Cov}(X_i, X_j)$.

A notação $|\Lambda|$ representa o determinante da matriz Λ . Com esta definição de distribuição Gaussiana multivariável temos a base para a definição de um processo aleatório Gaussiano.

Um processo aleatório $X(t)$, com t assumindo os valores do conjunto T , é dito ser um *processo Gaussiano* se, para qualquer inteiro k e qualquer subconjunto $\{t_1, t_2, \dots, t_k\}$ de T , as k variáveis aleatórias $\{X(t_1), X(t_2), \dots, X(t_k)\}$ possuem distribuição Gaussiana comum. Ou seja, se para qualquer k , $f_{X(t_1), X(t_2), \dots, X(t_k)}(x_1, \dots, x_k)$ possui densidade equivalente à Eq. (8.90).

Um processo Gaussiano possui as seguintes propriedades:

1. Se um processo Gaussiano é estacionário em sentido amplo, então ele também é estacionário em sentido estrito.
2. Se um processo Gaussiano é aplicado a um filtro linear estável, então o processo aleatório $Y(t)$ produzido na saída do filtro também é Gaussiano.

¹⁰ Veja o Capítulo 4 de Leon-Garcia (1994).

3. Se a integração é definida no sentido de média quadrática, então podemos alterar a ordem das operações de integração e esperanças com um processo aleatório Gaussiano.¹¹

A primeira propriedade vem da observação que se um processo Gaussiano é estacionário em sentido amplo, então (i) sua média não varia como o tempo e (ii) os elementos de Λ_{ij} da matriz de covariância dependem apenas da diferença de tempo $t_i - t_j$ e não do t absoluto. Como a distribuição comum N -dimensional de amostras de processos Gaussianos depende apenas da média e da covariância, pela Eq. (8.90), um processo Gaussiano estacionário em sentido amplo também é estritamente estacionário.

A segunda propriedade vem da observação de que a operação de filtragem pode ser escrita por

$$Y(t) = \int_0^t h(t-s)X(s) ds \quad (8.91)$$

se utilizarmos os três fatos:

1. A integral da Eq. (8.91) é definida como o limite médio quadrático do somatório

$$Y(t) = \lim_{\Delta s \rightarrow 0} \sum_i h(t-i\Delta s)x(i\Delta s) \Delta s$$

Logo, observamos que o lado direito é a soma ponderada de variáveis aleatórias Gaussianas $X(i\Delta s)$.

2. Lembre-se das propriedades de variáveis aleatórias Gaussianas que a soma ponderada de variáveis aleatórias Gaussianas é outra variável aleatória Gaussiana.
3. Se uma seqüência de variáveis aleatórias Gaussianas converge no sentido de média quadrática, então o resultado é uma variável aleatória Gaussiana.

Estes três fatos¹² juntos podem ser utilizados para provar que $Y(t)$ também é um processo aleatório Gaussiano.

A terceira propriedade de processos Gaussianos implica em que se $Y(t)$ é dado pela Eq. (8.91), então a média da saída é dada por

$$\begin{aligned} E[Y(t)] &= E\left[\int_0^t h(t-s)X(s) ds\right] \\ &= \int_0^t h(t-s)E[X(s)] ds \\ &= \mu_Y(t) \end{aligned}$$

Estes resultados são muito úteis em sistemas de comunicação, nos quais a filtragem linear de processos aleatórios ocorre bem freqüentemente.

- **PROBLEMA 8.16** O processo em tempo discreto $\{Y_n; n = 0, 1, 2, \dots\}$ definido por $Y_0 = 0$ e

$$Y_{n+1} = \alpha Y_n + W_n$$

é um processo Gaussiano, se W_n é Gaussiano? Assuma que $|\alpha| < 1$. Justifique sua resposta. ◀

¹¹ Veja o Capítulo 6 de Leon-Garcia (1994). A convergência de média quadrática de $\{Y_n\}$ para Y implica que

$$\lim_{n \rightarrow \infty} E[(Y - Y_n)^2] = 0$$

¹² A invariância do tipo de distribuição Gaussiana após a filtragem é mostrada no Capítulo 11 de Thomas (1971).

8.10 RUÍDO BRANCO

A análise de ruído em sistemas de comunicação é geralmente baseada em um processo de ruído idealizado, chamado de ruído branco. A *densidade espectral de potência do ruído branco é independente da freqüência*. O ruído branco é análogo ao termo “luz branca” no sentido de que todas as componentes de freqüência estão presentes em quantidade iguais. Representamos a densidade espectral de potência $W(t)$ de um processo de ruído branco por

$$S_W(f) = \frac{N_0}{2} \quad (8.92)$$

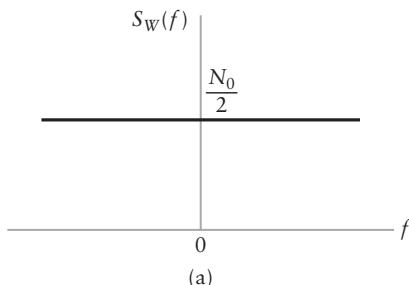
na qual o fator $\frac{1}{2}$ foi incluído para indicar que metade da potência é associada com as freqüências positivas e metade com as freqüências negativas, como ilustrado na Fig. 8.18(a). A dimensão de N_0 é watts por hertz. O parâmetro N_0 é geralmente medido no estágio de entrada de um receptor de comunicação. Como não existe função delta na origem na densidade espectral de potência da Fig. 8.18(a), o ruído branco não possui potência CC. Ou seja, sua esperança ou valor médio é nulo.

Como a função de autocorrelação é a transformada inversa de Fourier da densidade espectral de potência, temos que, da Eq. (8.92), a autocorrelação do ruído branco é dada por

$$R_W(\tau) = \frac{N_0}{2} \delta(\tau) \quad (8.93)$$

A função de autocorrelação do ruído branco consiste de uma função delta ponderada pelo fator $N_0/2$ localizado em $\tau = 0$, tal como na Fig. 8.18(b). Observamos que $R_W(\tau)$ é zero para $\tau \neq 0$. Conseqüentemente, quaisquer duas amostras diferentes do ruído branco, não importa quanto perto elas estejam no tempo, são não correlacionadas.

Estritamente falando, o ruído branco possui potência média infinita e, desta forma, não é fisicamente realizável. Entretanto, o ruído branco possui propriedades matemáticas convenientes sendo, portanto, útil na análise de sistemas. A utilização do processo de ruído branco é análoga à da função impulso ou função delta na análise de sistemas lineares. O efeito do impulso é observado apenas após ele ter passado através de um sistema com largura de faixa finita. De forma semelhante, o efeito do ruído branco é observado apenas após a sua passagem através de um sistema com largura de faixa finita. Podemos, pois, afirmar que enquanto a largura de faixa do processo de ruído na entrada de um sistema é



(a)

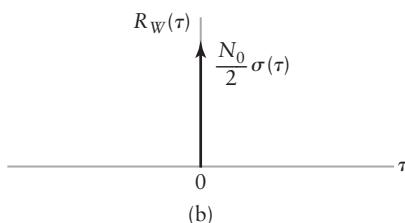


Figura 8.18 Características do ruído branco. (a) Densidade espectral de potência. (b) Função de autocorrelação.

apreciavelmente maior do que a do próprio sistema, podemos modelar o processo de ruído por um ruído branco. Este geralmente é o caso em sistemas de comunicação práticos.

EXEMPLO 8.13 Ruído branco idealmente filtrado por passa-baixa

Suponha que um processo de ruído branco $W(t)$ de média nula e densidade espectral de potência $N_0/2$ é aplicado a um filtro passa-baixa ideal com largura de faixa B com ganho unitário de amplitude na faixa passante. A densidade espectral de potência do processo de ruído $N(t)$ que aparece na saída do filtro é, portanto (veja a Fig. 8.19(a)),

$$S_N(f) = \begin{cases} \frac{N_0}{2}, & |f| < B \\ 0, & |f| > B \end{cases} \quad (8.94)$$

A função de autocorrelação de $N(t)$ é a transformada inversa de Fourier da densidade espectral de potência mostrada na Fig. 8.19(a).

$$\begin{aligned} R_N(\tau) &= \int_{-B}^B \frac{N_0}{2} \exp(j2\pi f_c \tau) df \\ &= N_0 B \operatorname{sinc}(2B\tau) \end{aligned} \quad (8.95)$$

Esta função de autocorrelação é mostrada na Fig. 8.19(b). Vemos que $R_N(\tau)$ possui seu valor máximo de $N_0 B$ na origem, é simétrico em τ , e passa por zero em $\tau = \pm n/2B$ onde $n = 1, 2, 3, \dots$

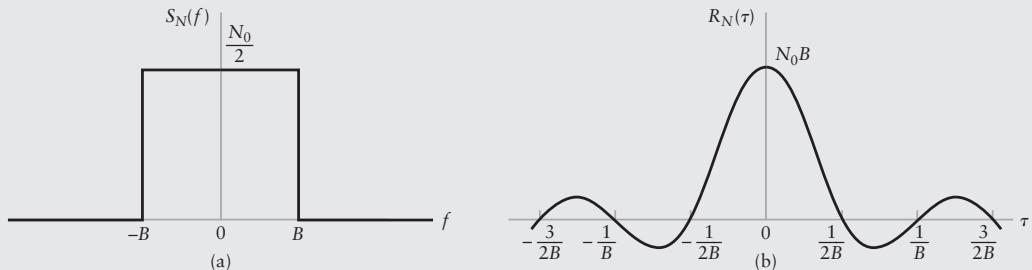


Figura 8.19 Características de um ruído branco filtrado por um filtro passa-baixa. (a) Densidade espectral de potência. (b) Função de autocorrelação.

EXEMPLO 8.14 Ruído branco filtrado por RC passa-baixa

Considere um processo de ruído branco $W(t)$ de média nula e densidade espectral de potência $N_0/2$ aplicado a um filtro RC passa-baixa, como mostrado na Fig. 8.20(a). A função de transferência do filtro é

$$H(f) = \frac{1}{1 + j2\pi f RC} \quad (8.96)$$

Da Eq. (8.87), obtemos que a densidade espectral de potência da saída do filtro passa-baixa RC é, portanto (veja a Fig. 8.20(b)),

$$S_N(f) = \frac{N_0/2}{1 + (2\pi f RC)^2}$$

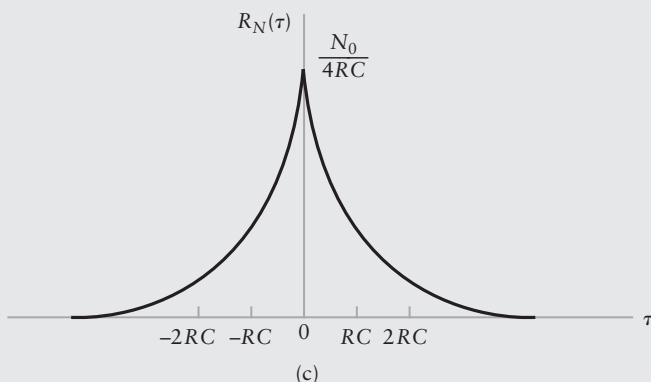
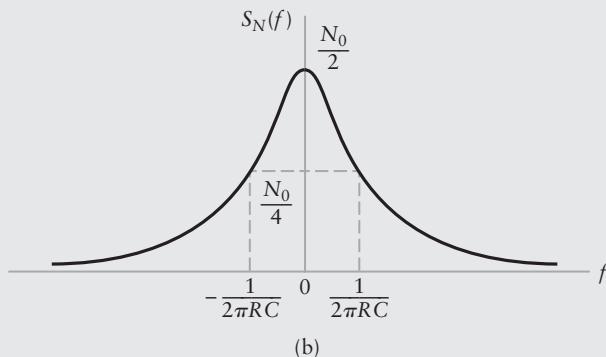
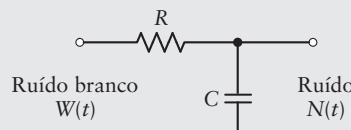


Figura 8.20 Características do ruído branco filtrado por RC. (a) Filtro RC passa-baixa. (b) Densidade espectral de potência da saída $N(t)$. (c) Função de autocorrelação de $N(t)$.

Da Eq. (2.16) do Capítulo 2, temos o par transformada de Fourier

$$\exp(-a|\tau|) \iff \frac{2a}{a^2 + (2\pi f)^2}$$

Portanto, usando a propriedade de dilatação da transformada de Fourier (veja a Seção 2.2), vemos que a autocorrelação do processo de ruído filtrado $N(t)$ é

$$R_N(\tau) = \frac{N_0}{4RC} \exp\left(-\frac{|\tau|}{RC}\right)$$

a qual é mostrada na Fig. 8.20(c).

► **PROBLEMA 8.17** Um processo de ruído branco em tempo discreto $\{W_n\}$ possui função de auto-correlação dada por $R_W(n) = N_0\delta(n)$.

- Usando a transformada discreta de Fourier, determine a densidade espectral de potência da $\{W_n\}$.
- O processo de ruído branco é passado através de um filtro em tempo discreto com resposta em freqüência discreta

$$H(k) = \frac{1 - (\alpha W^k)^N}{1 - \alpha W^k}$$

na qual, para uma transformada de Fourier discreta de N pontos, $W = \exp\{-j2\pi/N\}$. Qual é a densidade espectral de potência da saída do filtro? ◀

8.11 RUÍDO DE FAIXA ESTREITA

Um receptor de comunicação inclui vários estágios de processamento de sinal. Um estágio comum de processamento de sinal para sistemas banda passante é um filtro de faixa estreita cuja largura de faixa é apenas larga o suficiente para passar a componente modulada do sinal recebido, essencialmente não distorcida, mas não grande o suficiente para admitir ruído excessivo para dentro do receptor. O processo de ruído que aparece na saída de tal filtro é chamado de *ruído de faixa estreita*. Se um ruído de faixa estreita possui espectro centrado nas freqüências de meia faixa $\pm f_c$, como ilustrado na Fig. 8.21(a), vemos que uma função amostra do processo de ruído de faixa estreita é, de alguma forma, similar a forma de onda senoidal de freqüência f_c que varia lentamente em amplitude e fase, como ilustrado na Fig. 8.21(b).

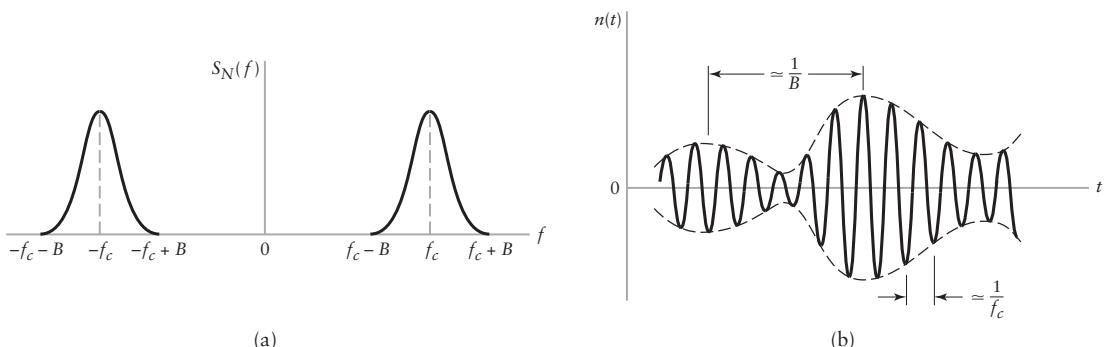


Figura 8.21 (a) Densidade espectral de potência de um ruído de faixa estreita. (b) Função amostra do ruído de faixa estreita.

O ruído de faixa estreita pode ser representado matematicamente utilizando *componentes em fase e em quadratura*, tal como as utilizamos para representar sinais de faixa estreita no Capítulo 2 e capítulos subsequentes. Para o processo de ruído de faixa estreita $N(t)$ de largura de faixa $2B$ e centrada na freqüência f_c da Fig. 8.21, podemos representar $N(t)$ na forma

$$N(t) = N_I(t) \cos(2\pi f_c t) - N_Q(t) \sin(2\pi f_c t) \quad (8.97)$$

na qual $N_I(t)$ é chamada de *componente em fase* de $N(t)$ e $N_Q(t)$ é a *componente em quadratura*. Tanto $N_I(t)$ quanto $N_Q(t)$ são processos aleatórios passa-baixa. Ou seja, seus espectros são confinados a $0 < |f| < B$. O conhecimento das componentes em fase e quadratura, além da freqüência central f_c , caracteriza completamente o ruído de faixa estreita.

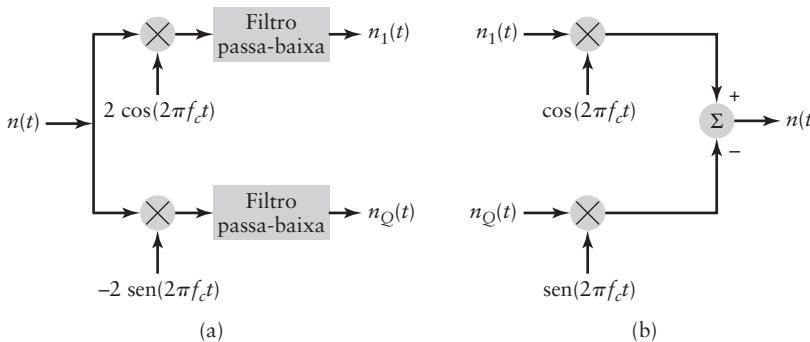


Figura 8.22 (a) Extração das componentes em fase e quadratura do processo de ruído de faixa estreita. (b) Geração do processo de ruído de faixa estreita de suas componentes em fase e quadratura.

Dada a função amostra $n(t)$ do ruído de faixa estreita, as componentes em fase e quadratura podem ser extraídas utilizando o esquema mostrado na Fig. 8.22(a). Os dois filtros passa-baixa são considerados como ideais com largura de faixa igual a B . Este esquema é obtido diretamente da representação de ruído de faixa estreita dado pela Eq. (8.97). Alternativamente, se tivermos as componentes em fase e quadratura, podemos gerar o ruído de faixa estreita $n(t)$ utilizando a Fig. 8.22(b).

As componentes em fase e quadratura do ruído de faixa estreita possuem as seguintes importantes propriedades:¹³

1. A componente em fase $N_I(t)$ e a componente em quadratura $N_Q(t)$ do ruído de faixa estreita $N(t)$ possuem média nula.
2. Se o ruído de faixa estreita $N(t)$ for Gaussiano, então suas componentes em fase e quadratura serão Gaussianas.
3. Se o ruído de faixa estreita $N(t)$ for estacionário, então suas componentes em fase e quadratura serão estacionárias.
4. Tanto a componente em fase $N_I(t)$ quanto em quadratura $N_Q(t)$ possuem a mesma densidade espectral de potência. Esta densidade espectral é relacionada com a densidade espectral de potência da densidade de faixa estreita $S_N(f)$ por

$$S_{N_I}(f) = S_{N_Q}(f) = \begin{cases} S_N(f - f_c) + S_N(f + f_c), & -B \leq f \leq B \\ 0, & \text{caso contrário} \end{cases} \quad (9.98)$$

5. A componente em fase $N_I(t)$ e a componente em quadratura $N_Q(t)$ possuem a mesma variância do ruído de faixa estreita $N(t)$.

Como ilustração destas propriedades, considere o ruído de faixa estreita com densidade espectral de potência mostrada na Fig. 8.23(a). De acordo com a Propriedade 4, o espectro da componente em fase do ruído de faixa estreita é dado por

$$S_{N_I}(f) = \begin{cases} N_0, & -B \leq f \leq B \\ 0, & \text{caso contrário} \end{cases}$$

e a componente em quadratura possui espectro similar. Portanto, a densidade espectral da componente em fase é duas vezes a do ruído de faixa estreita. Entretanto, como o ruído de faixa estreita possui densidade espectral não nula em duas faixas de largura $2B$ centradas em $\pm f_c$, temos que

$$\int_{-\infty}^{\infty} S_N(f) df = \int_{-\infty}^{\infty} S_{N_I}(f) df = 2N_0B$$

¹³ A justificativa destas propriedades é discutida no Capítulo 1 de Haykin (2001).

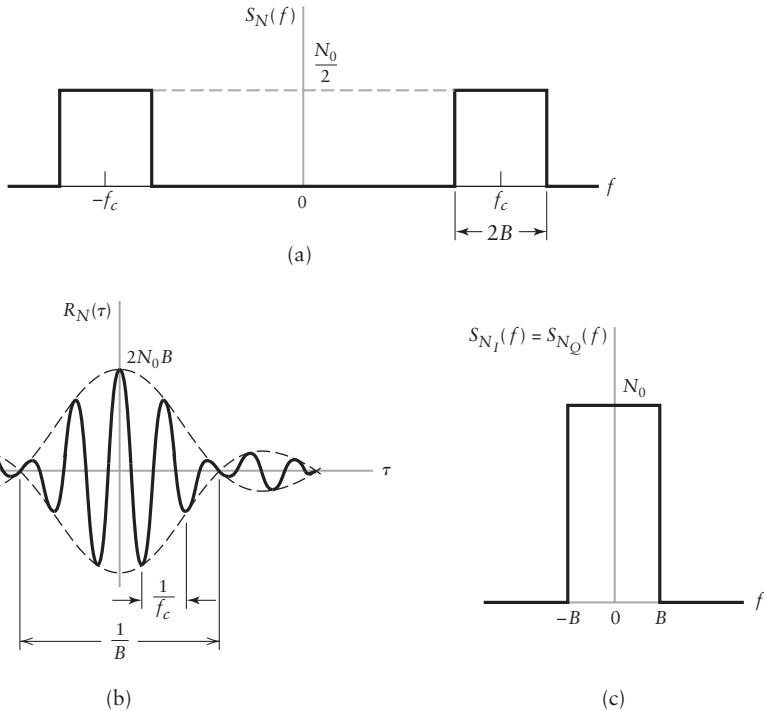


Figura 8.23 Características de um ruído branco idealmente filtrado por passa-faixa.
 (a) Densidade espectral de potência. (b) Função de autocorrelação. (c) Densidade espectral de potência das componentes em fase e quadratura.

a qual confirma a Propriedade 5 de que a componente em fase possui a mesma variância (potência) do ruído de faixa estreita.

EXEMPLO 8.15 Ruído branco idealmente filtrado por passa-faixa

Considere o ruído branco Gaussiano de média nula e densidade espectral de potência $N_0/2$, o qual é passado através de um filtro passa-faixa ideal com freqüência central f_c , largura de faixa $2B$ e resposta em amplitude da faixa passante unitária. A densidade espectral de potência do ruído branco filtrado é, então,

$$S_N(f) = \begin{cases} N_0/2, & |f - f_c| < B \\ N_0/2, & |f + f_c| < B \\ 0, & \text{caso contrário} \end{cases} \quad (8.99)$$

como ilustrado na Fig. 8.23(a). O problema é determinar a função de autocorrelação de $N(t)$ e suas componentes em fase e quadratura.

A função de autocorrelação de $n(t)$ é a transformada inversa de Fourier da densidade espectral de potência características da Eq. (8.99), mostrado por

$$\begin{aligned} R_N(\tau) &= \int_{-f_c-B}^{-f_c+B} \frac{N_0}{2} \exp(j2\pi f_c \tau) df + \int_{f_c-B}^{f_c+B} \frac{N_0}{2} \exp(j2\pi f_c \tau) df \\ &= N_0 B \operatorname{sinc}(2B\tau) [\exp(-j2\pi f_c \tau) + \exp(2\pi f_c \tau)] \\ &= 2N_0 B \operatorname{sinc}(2B\tau) \cos(2\pi f_c \tau) \end{aligned} \quad (8.100)$$

Esta função de autocorrelação do ruído de faixa estreita ideal é mostrada na Fig. 8.23(b).

Da Propriedade 4 do ruído de faixa estreita, as componentes em fase e em quadratura possuem densidade espectral idênticas, como mostrado na Fig. 8.23(c). A função de autocorrelação de $N_I(t)$ ou $N_Q(t)$ é, portanto (veja o Exemplo 8.13),

$$R_{N_I}(\tau) = R_{N_Q}(\tau) = 2N_0B \operatorname{sinc}(2B\tau) \quad (8.101)$$

■ Largura de faixa equivalente ao ruído

No Exemplo 8.13, observamos que quando a fonte de ruído branco de média nula e densidade espectral de potência $N_0/2$ é conectada na entrada de um filtro passa-baixa ideal com largura de faixa B e ganho de faixa passante unitário, a potência média de saída (ou, de forma equivalente, $R_N(0)$) é N_0B . No Exemplo 8.14, observamos que quando uma fonte similar é conectar a um filtro RC passa-baixa, a potência média de saída é $N_0/(4RC)$. Para este filtro, a largura de faixa de 3dB é igual a $1/(2\pi RC)$. Podemos, portanto, fazer duas importantes observações. Primeiro, o ruído branco filtrado possui potência média finita. Segundo, a potência média é proporcional à largura de faixa. Podemos generalizar estas observações para incluir todos os tipos de filtros passa-baixa, definindo a *largura de faixa equivalente ao ruído*, como mostrado a seguir.

Suponha que uma fonte de ruído branco com espectro $S_W(f) = N_0/2$ é conectada à entrada de um filtro arbitrário com função de transferência $H(f)$. Das Propriedades 1 e 4 da densidade espectral de potência, a potência média do ruído de saída é

$$\begin{aligned} P_N &= \int_{-\infty}^{\infty} |H(f)|^2 S_W(f) df \\ &= \int_{-\infty}^{\infty} |H(f)|^2 \frac{N_0}{2} df \\ &= N_0 \int_0^{\infty} |H(f)|^2 df \end{aligned} \quad (8.102)$$

na qual, na última linha, utilizamos o fato da resposta em amplitude $|H(f)|$ ser uma função par da freqüência. Considere, agora, a mesma fonte de ruído conectada à entrada de um filtro passa-baixa ideal com resposta e freqüência zero $H(0)$ e largura de faixa B_N como ilustrado pela Fig. 8.24. Neste caso, a potência média do ruído de saída é

$$P_N = N_0 B_N |H(0)|^2 \quad (8.103)$$

Igualando as Eqs. (8.102) e (8.103), determinamos a largura de faixa B_N do filtro ideal que produz a mesma potência do filtro arbitrário. Fazendo isto, obtemos

$$B_N = \frac{\int_0^{\infty} |H(f)|^2 df}{|H(0)|^2} \quad (8.104)$$

A largura de faixa B_N é chamada de *largura de faixa equivalente ao ruído* para um filtro passa-baixa. Portanto, o procedimento para o cálculo da largura de faixa equivalente ao ruído consiste em substituir o filtro passa-baixa arbitrário com função de transferência $H(f)$ por um filtro ideal passa-baixa equivalente com resposta em freqüência zero $H(0)$ e largura de faixa B_N , como ilustrado na Fig. 8.24.

De maneira similar, podemos definir a *largura de faixa equivalente ao ruído para um filtro passa-faixa*, como ilustrado na Fig. 8.25. Esta figura mostra somente a resposta em

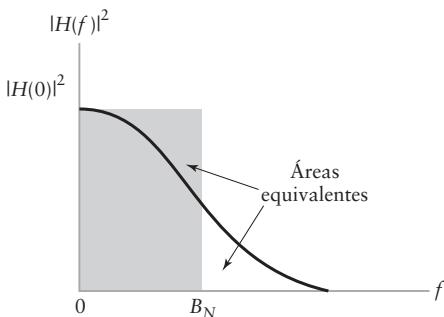


Figura 8.24 Ilustração de um filtro passa-baixa arbitrário $H(f)$ e um filtro passa-baixa ideal com largura de faixa B_N .

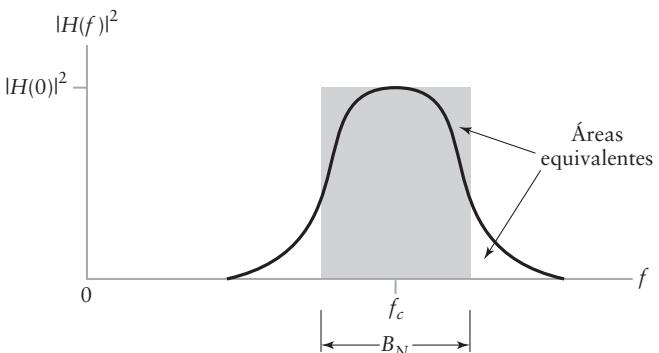


Figura 8.25 Ilustração de um filtro passa-faixa arbitrário $H(f)$ e o filtro passa-faixa ideal com largura de faixa B_N .

amplitude quadrada do filtro para freqüências positivas. Portanto, a largura de faixa equivalente ao ruído para um filtro passa-faixa pode ser definida por

$$B_N = \frac{\int_0^\infty |H(f)|^2 df}{|H(f_c)|^2} \quad (8.105)$$

na qual $|H(f_c)|^2$ é a resposta em amplitude de freqüência central do filtro. De fato, observamos que a Eq. (8.105) pode ser utilizada para representar os dois casos fazendo $f_c = 0$ para filtros passa-baixa. Então, temos o resultado genérico

$$P_N = N_0 |H(f_c)|^2 B_N \quad (8.106)$$

e o efeito da passagem do ruído branco através do filtro pode ser separado em duas partes:

- O ganho de potência de freqüência central $|H(f_c)|^2$.
- A largura de faixa B_N equivalente ao ruído, representando a *seletividade em freqüência* do filtro.

Esta separação se aplica se o filtro for passa-baixa ou passa-faixa. Logo, como regra geral, podemos dizer que o efeito do ruído no sistema é reduzido estreitando-se a largura de faixa do sistema.

EXEMPLO 8.16 Largura de faixa equivalente ao ruído do filtro RC

Considere o filtro passa-baixa de pólo único do Exemplo 8.14. A função de transferência deste filtro é

$$H(f) = \frac{1}{1 + j2\pi f RC}$$

Queremos determinar a largura de faixa equivalente ao ruído para este filtro.

A largura de faixa equivalente ao ruído para este filtro é

$$\begin{aligned} B_N &= \frac{1}{|H(0)|^2} \int_0^\infty |H(f)|^2 df \\ &= 1 \times \int_0^\infty \frac{df}{1 + (2\pi f RC)^2} \end{aligned}$$

Percebendo que o integrando é a derivada escalonada de $\operatorname{tg}^{-1}(f)$, obtemos

$$\begin{aligned} B_N &= \frac{1}{2\pi RC} \operatorname{tg}^{-1}(2\pi fRC) \Big|_0^\infty \\ &= \frac{1}{2\pi RC} \left(\frac{\pi}{2} - 0 \right) \\ &= \frac{1}{4RC} \end{aligned}$$

Devido ao baixo amortecimento do filtro de pólo único, a largura de faixa do ruído é um pouco maior do que sua largura de faixa de 3dB, $B_{3\text{dB}} = 1/(2\pi RC)$.

8.12 RESUMO E DISCUSSÃO

Começamos este capítulo indicando que a informação e o ruído são sinais aleatórios. De fato, esta aleatoriedade ou imprevisibilidade é a propriedade chave de sistemas de comunicação. Entretanto, mesmo sinais aleatórios possuem características estatísticas que podem ser medidas. Subseqüentemente, apresentamos várias ferramentas importantes para a caracterização de sinais aleatórios.

Introduzimos um experimento aleatório como modelo para fenômenos imprevisíveis e a definição de freqüência relativa de probabilidade como meio de associar a probabilidade à saída de um experimento aleatório. Isto levou aos três axiomas básicos da teoria de probabilidade.

Variáveis aleatórias foram apresentadas como uma função cujo domínio e o espaço de amostra do experimento aleatório e cuja faixa são os números reais. Variáveis aleatórias fornecem um método de unificar o tratamento de uma grande variedade de experimentos aleatórios. Funções de distribuição de probabilidade e de densidade foram mostradas como métodos fundamentais para a caracterização de uma variável aleatória.

O estudo de funções de variáveis aleatórias levou naturalmente ao conceito de esperança e momentos estatísticos e covariância de variáveis aleatórias.

Variáveis aleatórias Gaussianas foram apresentadas como um tipo particular importante de variável aleatória no estudo de sistemas de comunicação.

A consideração do parâmetro tempo em sinais aleatórios levou ao estudo de processos aleatórios. Um processo aleatório foi definido como uma família de variáveis aleatórias indexadas pelo tempo como parâmetro. Processos aleatórios estacionários, ergódicos e estacionários em sentido amplo foram apresentados como modelos da maioria dos processos físicos que exibem um comportamento aleatório. Foi mostrado que processos aleatórios estacionários em sentido amplo possuem várias das propriedades de sinais de potência determinísticos, incluindo o fato de que as fórmulas de Weiner-Khintchine relacionam o espectro de processos aleatórios com sua autocorrelação.

Processos Gaussianos e ruído branco foram apresentados como importantes processos aleatórios na análise de sistemas de comunicação.

Finalmente, foi mostrado que, semelhante a sinais determinísticos, podemos considerar versões passa-faixa ou faixa estreita do ruído. Este ruído de faixa estreita possui componentes em fase e quadratura, semelhante a sinais determinísticos.

Este capítulo foi uma breve e certamente não completa introdução a sinais aleatórios e ruído que são geralmente encontrados em sistemas de comunicação, mas seu tratamento apresentado aqui é adequado para um tratamento introdutório à teoria estatística de comunicação. Os próximos dois capítulos irão ilustrar a importância do material apresentado neste capítulo no projeto de receptores e no cálculo da performance de sistemas de comunicação.

Problemas adicionais

- 8.18 Considere um baralho com 52 cartas, dividido em quatro naipes, cada um com 13 cartas variando de dois até o Ás. Assuma que todas as cartas possuem a mesma probabilidade de serem escolhidas.
- Suponha que uma única carta é retirada de um baralho completo. Qual é a probabilidade desta carta ser um Ás de ouros? Qual é a probabilidade da carta ser um Ás de qualquer um dos quatro grupos?
 - Suponha que duas cartas são retiradas do baralho completo. Qual é a probabilidade das cartas retiradas serem um Ás e um Rei, não necessariamente do mesmo naipe? E se eles forem do mesmo naipe?
- 8.19 Suponha que um jogador possui um dado vermelho e um branco. Quantas saídas são possíveis no experimento aleatório de lançar os dois dados? Supondo que os dados sejam indistinguíveis, quantas saídas são possíveis?
- 8.20 Refira-se ao Problema 8.19.
- Qual é a probabilidade de lançar um 5 vermelho e um 2 branco?
 - Se os dados são indistinguíveis, qual é a probabilidade de lançar uma soma de 7? Se eles forem distinguíveis, qual é a probabilidade?
- 8.21 Considere uma variável aleatória X uniformemente distribuída entre os valores de 0 e 1 com probabilidade de $\frac{1}{4}$, assumindo o valor 1 com probabilidade $\frac{1}{4}$, e uniformemente distribuída entre os valores 1 e 2 com probabilidade $\frac{1}{2}$. Determine a função de distribuição da variável aleatória X .
- 8.22 Considere uma variável aleatória X definida pela densidade de exponencial dobrada

$$f_X(x) = a \exp(-b|x|), \quad -\infty < x < \infty$$

na qual a e b são constantes.

- Determine a relação entre a e b tal que $f_X(x)$ é uma função densidade de probabilidade.
 - Determine a função de distribuição $F_X(x)$ correspondente.
 - Calcule a probabilidade da variável aleatória X estar entre 1 e 2.
- 8.23 Mostre que a expressão para a variância de uma variável aleatória pode ser expressa em termos do primeiro e segundo momentos por

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

- 8.24 Uma variável aleatória R possui distribuição de Rayleigh com sua função densidade de probabilidade dada por

$$f_R(r) = \begin{cases} \frac{r}{b} \exp\left(-\frac{r^2}{2b}\right), & 0 < r < \infty \\ 0, & \text{caso contrário} \end{cases}$$

- Determine a função de distribuição $F_R(r)$ correspondente.
- Mostre que a média de R é igual a $\sqrt{b\pi/2}$.
- Qual é o valor médio quadrático de R ?
- Qual é a variância de R ?

- 8.25 Considere uma variável aleatória Z uniformemente distribuída, definida por

$$f_Z(z) = \begin{cases} \frac{1}{2\pi}, & 0 \leq z \leq 2\pi \\ 0, & \text{caso contrário} \end{cases}$$

As duas variáveis aleatórias X e Y são relacionadas com Z por $X = \sin(Z)$ e $Y = \cos(Z)$.

- Determine as funções densidade de probabilidade de X e Y .
- Mostre que X e Y são variáveis aleatórias não correlacionadas.
- X e Y são estatisticamente independentes? Por quê?

- 8.26 Uma variável aleatória Gaussiana possui média nula e desvio padrão de 10V. Uma tensão constante de 5V é adicionada a esta variável.
- Determine a probabilidade da medida deste sinal composto resultar em um valor positivo.
 - Determine a probabilidade da média aritmética de duas medidas independentes deste sinal ser positiva.

- 8.27 Considere um processo aleatório $X(t)$ definido por

$$X(t) = \sin(2\pi Wt)$$

no qual a freqüência W é uma variável aleatória com função densidade de probabilidade

$$f_W(w) = \begin{cases} \frac{1}{B}, & 0 \leq w \leq B \\ 0, & \text{caso contrário} \end{cases}$$

Mostre que $X(t)$ é não estacionário.

- 8.28 Considere o processo senoidal

$$X(t) = A \cos(2\pi f_c t)$$

no qual a freqüência f_c é constante e a amplitude A é uniformemente distribuída:

$$f_A(a) = \begin{cases} 1, & 0 \leq a \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

Determine se este processo é ou não estacionário em sentido estrito.

- 8.29 Um processo aleatório $X(t)$ é definido por

$$X(t) = A \cos(2\pi f_c t)$$

no qual A é uma variável aleatória Gaussiana de média nula e variância σ_A^2 . Este processo aleatório é aplicado a um integrador ideal, produzindo uma saída $Y(t)$ definida por

$$Y(t) = \int_0^t X(\tau) d\tau$$

- Determine a função densidade de probabilidade da saída $Y(t)$ no tempo particular t_k .
- Determine se $Y(t)$ é estacionário ou não.

- 8.30 Prove as duas propriedades seguintes da função de autocorrelação $R_X(\tau)$ de um processo aleatório $X(t)$:

- Se $X(t)$ contém uma componente CC igual a A , então $R_X(\tau)$ contém uma componente constante igual a A^2 .
- Se $X(t)$ contém uma componente senoidal, então $R_X(\tau)$ também contém uma componente senoidal de mesma freqüência.

- 8.31 Um processo aleatório em tempo discreto $\{Y_n\}$ é definido por

$$Y_n = \alpha Y_{n-1} + W_n, \quad n = \dots, -1, 0, +1, \dots$$

na qual o processo aleatório de média nula $\{W_n\}$ é estacionário com função de autocorrelação $R_W(k) = \sigma^2 \delta(k)$. Qual é a função de autocorrelação $R_Y(k)$ de Y_n ? Y_n é um processo estacionário em sentido amplo? Justifique sua resposta.

- 8.32 Determine a densidade espectral de potência do processo que possui função de autocorrelação

$$R_X(\tau) = \begin{cases} \sigma^2(1 - |\tau|), & \text{para } |\tau| \leq 1 \\ 0, & \text{para } |\tau| > 1 \end{cases}$$

- 8.33 Um pulso aleatório possui amplitude A e duração T mas começa em um tempo arbitrário t_0 . Ou seja, o processo aleatório é definido por

$$X(t) = A \text{ ret}(t + t_0)$$

na qual $\text{ret}(t)$ é definida na Seção 2.9. A variável aleatória t_0 é considerada uniformemente distribuída em $[0, T]$, com densidade

$$f_{t_0}(s) = \begin{cases} \frac{1}{T}, & 0 \leq s \leq T \\ 0, & \text{caso contrário} \end{cases}$$

- (a) Qual é a função de autocorrelação do processo aleatório $X(t)$?
- (b) Qual é o espectro do processo aleatório $X(t)$?

8.34 Dado que um processo aleatório estacionário $X(t)$ possui função de autocorrelação $R_X(\tau)$ e densidade espectral de potência $S_X(f)$, mostre que:

- (a) A função de autocorrelação de $dx(t)/dt$, a primeira derivada de $X(t)$, é igual ao negativo da segunda derivada de $R_X(\tau)$.
- (b) A densidade espectral de potência de $dX(t)/dt$ é igual a $4\pi^2 f^2 S_X(f)$.
Dica: Veja a solução do Problema 2.24.

8.35 Considere um processo estacionário em sentido amplo $X(t)$ com densidade espectral de potência $X(t)$ mostrado na Fig. 8.26. Determine a função de autocorrelação $R_X(\tau)$ do processo $X(t)$.

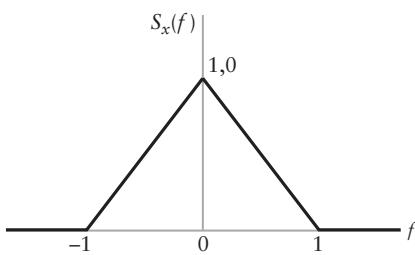


Figura 8.26 Problema 8.35

8.36 A densidade espectral de potência de um processo aleatório $X(t)$ é mostrada na Fig. 8.27.

- (a) Determine e trace a função de autocorrelação $R_X(\tau)$ de $X(t)$.
- (b) Qual é a potência CC contida em $X(t)$?
- (c) Qual é a potência CA contida em $X(t)$?
- (d) Quais taxas de amostragem irão gerar amostras não correlacionadas de $X(t)$? As amostras são estatisticamente independentes?

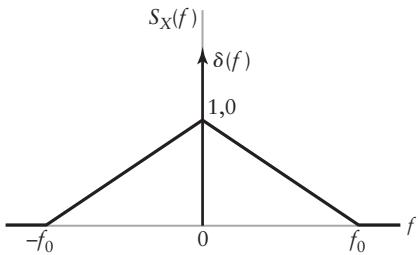


Figura 8.27 Problema 8.36.

8.37 Considere os dois filtros lineares mostrados em série na Fig. 8.28. Seja $X(t)$ um processo estacionário com função de autocorrelação $R_X(\tau)$. O processo aleatório que aparece na saída do primeiro filtro é $V(t)$ e o da saída do segundo filtro é $Y(t)$.

- (a) Determine a função de autocorrelação de $V(t)$.
- (b) Determine a função de autocorrelação de $Y(t)$.

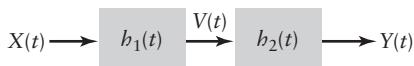


Figura 8.28 Problema 8.37.

- 8.38 A densidade espectral de potência de um processo aleatório de faixa estreita $X(t)$ é como mostrado na Fig. 8.29. Determine as densidades espectrais de potência das componentes em fase e quadratura de $X(t)$, assumindo $f_c = 5\text{Hz}$.

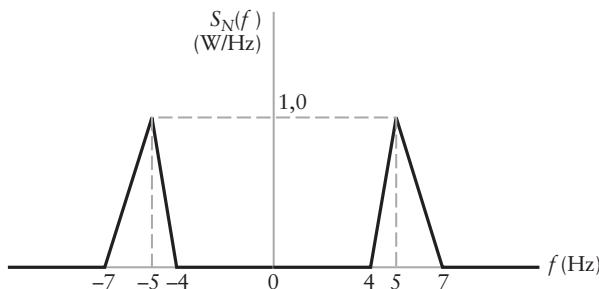


Figura 8.29 Problema 8.38.

- 8.39 Assuma que o processo $X(t)$ de faixa estreita descrito no Problema 8.38 é Gaussiano com média nula e variância σ_x^2 .
- Calcule σ_x^2 .
 - Determine a função densidade de probabilidade comum das variáveis aleatórias Y e Z obtidas pela observação das componentes em fase e quadratura de $X(t)$ em algum tempo fixo.

Problemas avançados

- 8.40 Determine a probabilidade dos dois últimos dígitos do cubo de um número natural (1, 2, 3,...) ser 01.
- 8.41 Considere um experimento aleatório de seleção de um número uniformemente distribuído na faixa $\{1, 2, 3, \dots, 120\}$. Sejam A , B e C os eventos nos quais o número selecionado é múltiplo de 3, 4 e 6, respectivamente.
- Qual é a probabilidade do evento A , ou seja, $P[A]$?
 - Qual é $P[B]$?
 - Qual é $P[A \cap B]$?
 - Qual é $P[A \cup B]$?
 - Qual é $P[A \cap C]$?
- 8.42 Uma mensagem é constituída por dez 0s e 1s.
- Quantas mensagens existem?
 - Quantas destas mensagens contêm exatamente quatro 1s?
 - Suponha que o décimo bit não é independente dos outros, mas é escolhido de tal forma que a soma de módulo 2 de todos os bits seja zero. Isto é referenciado como uma seqüência de paridade par. Quantas seqüências de paridade par existem?
 - Se esta seqüência de paridade par com dez bits for transmitida em um canal com probabilidade p de erro para cada bit, qual é a probabilidade da seqüência recebida conter um erro não detectado?
- 8.43 A probabilidade de um evento ocorrer ao menos uma vez em quatro testes independentes é igual a 0,59. Qual é a probabilidade de ocorrência do evento em uma tentativa, se as probabilidades são iguais em todas as tentativas?
- 8.44 Os tempos de chegada de dois sinais em um receptor são uniformemente distribuídos no intervalo $[0, T]$. O receptor ficará emperrado se a diferença de tempo entre chegadas for menor do que τ . Determine a probabilidade do receptor ficar emperrado.
- 8.45 Um sistema de telégrafo (uma versão antiga de comunicação digital) transmite um sinal de ponto ou traço. Assuma que as propriedades de transmissão são tais que 2/5 dos pontos e 1/3 dos

traços são recebidos incorretamente. Suponha que a razão de pontos transmitidos para traços transmitidos é de 5 para 3. Qual é a probabilidade do sinal recebido for o transmitido se:

- (a) O sinal recebido é um ponto?
- (b) O sinal recebido é um traço?

8.46 Quatro sinais de rádio são emitidos sucessivamente. A probabilidade de recepção de cada um deles é independente da recepção dos outros e igual, respectivamente, a 0,1; 0,2; 0,3 e 0,4. Determine a probabilidade de k sinais serem recebidos, sendo $k = 1, 2, 3, 4$.

8.47 Em uma rede de comunicação de computadores, o tempo τ de chegada de mensagens é modelado por uma função de distribuição exponencial tendo densidade

$$f_T(\tau) = \begin{cases} \frac{1}{\lambda}e^{-\lambda\tau}, & \tau \geq 0 \\ 0, & \text{caso contrário} \end{cases}$$

- (a) Qual é o tempo médio entre mensagens com esta distribuição?
- (b) Qual é a variância do tempo entre mensagens?

8.48 Se X possui densidade $f_X(x)$, determine a densidade da variável aleatória Y definida por:

- (a) $Y = aX + b$ para a e b constantes.
- (b) $Y = X^2$.
- (c) $Y = \sqrt{X}$, assumindo X como uma variável aleatória não negativa.

8.49 Sejam X e Y duas variáveis aleatórias independentes com funções de densidade $f_X(x)$ e $f_Y(y)$, respectivamente. Mostre que a variável aleatória $Z = X + Y$ possui função de densidade dada por

$$f_Z(z) = \int_{-\infty}^0 f_Y(z-s) f_X(s) ds$$

Dica: $P[Z \leq z] = P[X \leq z, Y \leq z - X]$

8.50 Determine a densidade espectral $S_Z(f)$ se

$$Z(t) = X(t) + Y(t)$$

na qual $X(t)$ e $Y(t)$ são processos aleatórios independentes de média nula com

$$R_X(\tau) = \alpha_1 e^{-\alpha_1 |\tau|}, \quad \text{e} \quad R_Y(\tau) = \alpha_2 e^{-\alpha_2 |\tau|}.$$

8.51 Considere um processo aleatório $X(t)$ definido por

$$X(t) = \operatorname{sen}(2\pi f_c t)$$

na qual a freqüência f_c é uma variável aleatória uniformemente distribuída no intervalo $[0, W]$. Mostre que $X(t)$ é não estacionário. Dica: Examine funções amostra específicas do processo aleatório $X(t)$ para, digamos, as freqüências $W/4$, $W/2$ e W .

8.52 Os osciladores utilizados em sistemas de comunicação não são ideais e geralmente sofrem de uma distorção chamada de *ruído de fase*. Tal oscilador pode ser modelado pelo processo aleatório

$$Y(t) = A \cos(2\pi f_c t + \phi(t))$$

no qual $\phi(t)$ é um processo aleatório que varia lentamente. Descreva e justifique as condições ao processo $\phi(t)$ tal que $Y(t)$ seja estacionário em sentido amplo.

8.53 Um sinal banda base é perturbado por um processo de ruído aditivo $N(t)$ como mostrado por

$$X(t) = A \operatorname{sen}(0,3\pi t) + N(t)$$

na qual $N(t)$ é um processo Gaussiano estacionário de média nula e variância σ^2 .

- (a) Determine as funções de densidade das variáveis aleatórias X_1 e X_2 , tais que

$$X_1 = X(t)|_{t=1}$$

$$X_2 = X(t)|_{t=2}$$

- (b) A processo de ruído $N(t)$ possui função de autocorrelação dada por

$$R_N(\tau) = \sigma^2 \exp(-|\tau|)$$

Determine a função de distribuição comum de X_1 e X_2 , ou seja, $f_{X_1, X_2}(x_1, x_2)$.

■ Experimentos de computador

- 8.54 Para este experimento de lançamento de dados, um script MATLAB é incluído no Apêndice 7. O script MATLAB simula o lançamento de um dado polarizado 1000 vezes. Os passos do script são:

- ▶ Role o dado N vezes e salve os resultados em X .
- ▶ Calcule o histograma de X para obter as probabilidades das diferentes faces.

Repita o experimento para $N = 10, 100, 1000$ e 10.000 . Comente sobre a definição de freqüência relativa de probabilidade como função de N , o número de lançamentos.

- 8.55 Para demonstrar o teorema do limite central, calculamos 20.000 amostras de Z para $N = 5$ e estimamos a função densidade de probabilidade correspondente formando um histograma dos resultados. Um script MATLAB para a execução deste experimento é incluído no Apêndice 7. Compare este histograma (escalonado para área unitária) com a função densidade Gaussiana tendo mesma média e variância.

- 8.56 Revise o script para o Problema 8.55 para determinar a distribuição da soma de variáveis aleatórias de Bernoulli. Compare-o com a distribuição Gaussiana quando N se torna grande.

- 8.57 Revise o script para o Problema 8.55 de forma que os valores médios não sejam idênticos, mas que também possuam uma distribuição aleatória, mas com a mesma média final. Calcule a distribuição da soma.

- 8.58 Neste experimento de computador, iremos simular digitalmente um processo aleatório Gaussiano. Um script MATLAB no Apêndice 7 gera um processo Gaussiano branco em tempo discreto e o filtra com um filtro da raiz de cosseno levantado em tempo discreto (como discutido no Capítulo 6). No script, executamos os seguintes passos:

- ▶ Geramos um processo Gaussiano branco em tempo discreto.
- ▶ Filtramos este processo Gaussiano com um filtro da raiz de cosseno levantado com 25% de excesso de largura de faixa.
- ▶ Calculamos o espectro do processo resultante em tempo discreto.
- ▶ Calculamos a autocorrelação do processo resultante em tempo discreto.

- (a) Determine a autocorrelação da seqüência filtrada.
 (b) Determine o espectro da seqüência filtrada.

Comente seus resultados.

Capítulo 9

RUÍDO EM COMUNICAÇÕES ANALÓGICAS

Neste capítulo, iremos revisar os métodos de modulação analógica do Capítulo 3 e 4 considerando os conceitos relacionados com o ruído introduzidos no Capítulo 8. Na prática, vemos que sinais modulados, independente do tipo de modulação, são perturbados pelo ruído e pelas características imperfeitas do canal durante a transmissão. O ruído pode ser amplamente definido como qualquer sinal desconhecido que afeta a recuperação do sinal desejado. Podem haver várias fontes de ruído em um sistema de comunicação, mas, geralmente, as principais fontes são os próprios dispositivos de comunicação encontrados durante o curso da transmissão. Existem várias formas pelas quais o ruído pode afetar o sinal desejado, mas uma das formas mais comuns é uma distorção aditiva. Ou seja, o sinal recebido é modelado por¹

$$r(t) = s(t) + w(t) \quad (9.1)$$

na qual $s(t)$ é o sinal transmitido e $w(t)$ é o ruído aditivo. Se conhecermos exatamente o ruído, então podemos subtraí-lo de $r(t)$ e recuperar exatamente o sinal transmitido. Infelizmente, este é raramente o caso. Grande parte do projeto de um sistema de comunicação é relacionado com o processamento do sinal $r(t)$ recebido, de forma a minimizar a presença do ruído aditivo.

Este capítulo irá focalizar na detecção de sinais analógicos na presença de ruído aditivo. O material deste capítulo nos ensina as seguintes lições.

- *Lição 1: A minimização dos efeitos do ruído é a principal preocupação em comunicações analógicas e, consequentemente, a relação da potência do sinal pela potência do ruído é uma medida importante para avaliar a qualidade da comunicação analógica.*
- *Lição 2: A modulação em amplitude pode ser detectada coerentemente, com a utilização de um oscilador sincronizado, ou não coerentemente, através de um simples detector de envelope. Entretanto, existe uma penalidade na performance a ser paga para a detecção não coerente.*
- *Lição 3: A modulação em freqüência é não linear e o espectro do ruído de saída é parabólico quando o espectro do ruído de entrada é plano. A modulação em freqüência possui a vantagem de permitir a negociação da largura de faixa para melhoria de performance.*
- *Lição 4: A filtragem pré e de-ênfase é um método de redução do ruído de saída em um demodulador FM sem a distorção do sinal. Esta técnica pode ser utilizada para melhorar significativamente o desempenho dos sistemas de modulação em freqüência.*

9.1 RUÍDO EM SISTEMAS DE COMUNICAÇÃO

Antes de olharmos para as técnicas de minimização dos efeitos do ruído, precisamos rever algumas das propriedades do ruído. No Capítulo 8, identificamos o ruído como um processo aleatório, mas, mesmo que o processo de ruído seja imprevisível, existem métodos de caracterização de seu comportamento. Em particular, podemos utilizar parâmetros estatísticos tais como:

¹ Neste capítulo e no seguinte, utilizamos letras minúsculas para sinais aleatórios, sabendo que eles representam funções amostra do processo aleatório.

- *A média do processo aleatório.* Para o ruído, o valor médio corresponde ao *deslocamento CC*. Na maioria dos sistemas de comunicação, deslocamentos CC são removidos pelo projeto pois eles necessitam de potência e transportam pouca informação. Conseqüentemente, tanto o ruído quanto o sinal são geralmente considerados como tendo média nula.
- *A autocorrelação do processo aleatório.* No Capítulo 11, iremos descrever modelos físicos para o ruído do receptor e meios de caracterizá-lo. Como mencionado anteriormente, o *ruído branco*, como descrito no Capítulo 8, é geralmente um bom modelo matemático para o ruído recebido e iremos utilizar este modelo extensivamente nos Capítulos 9 e 10. Com o ruído branco, amostras em um instante de tempo são não correlacionadas com aquelas em outro instante de tempo, independente de sua separação. Ou seja, a autocorrelação do ruído brando é descrita por

$$R_w(\tau) = \frac{N_0}{2} \delta(\tau) \quad (9.2)$$

na qual $\delta(\tau)$ é a função delta de Dirac e $N_0/2$ é densidade espectral de potência dos dois lados.

- *O espectro de um processo aleatório.* Para ruído branco Gaussiano aditivo, o espectro é plano e definido por

$$S_w(f) = \frac{N_0}{2} \quad (9.3)$$

As relações de Weiner-Khintchine afirmam que o espectro da Eq. (9.3) é a transformada de Fourier da função de autocorrelação da Eq. (9.2).

Dadas as características do ruído, devemos determinar como ele afeta o sinal recebido. Para calcular a potência do ruído, devemos medir o ruído em uma largura de faixa específica. Vimos que a potência do ruído na saída de um filtro com largura de faixa B_T equivalente ao ruído é

$$N = N_0 B_T \quad (9.4)$$

na qual, por conveniência de representação, assumimos que $|H(f_c)| = 1$. Claramente, quanto menor a largura de faixa B_T , menor será a potência N do ruído. Relacionando isto de volta à detecção do sinal recebido $r(t) = s(t) + w(t)$ da Eq. (9.1), parece ser intuitivo que devemos fazer B_T o menor possível para minimizar o ruído, mas ela não deve ser menor do que a largura de faixa de $s(t)$, caso contrário iremos distorcer o sinal desejado.

Esta situação é ilustrada na Fig. 9.1, na qual o sinal transmitido é distorcido pelo ruído branco aditivo e a combinação passa por um filtro com largura de faixa B_T . Se a largura de faixa do filtro for maior do que a largura de faixa do sinal, então iremos manter toda a energia do sinal desejado. Entretanto, se o filtro não for tão largo quanto o necessário para passar o sinal sem distorção, então iremos minimizar o total de ruído que passa pelo filtro. Conseqüentemente, a largura de faixa B_T é chamada de largura de faixa de transmissão do

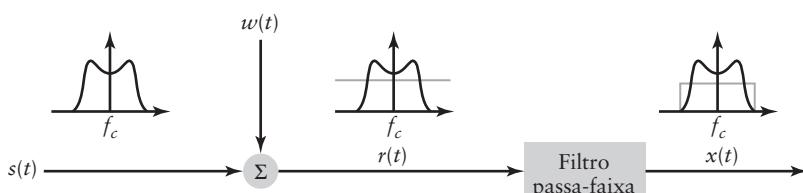


Figura 9.1 Diagrama em blocos do sinal mais o ruído antes e após a filtragem, mostrando o espectro em cada ponto.

sinal. O casamento do filtro do receptor com a largura de faixa do sinal transmitido é a base de vários esquemas ótimos de detecção.

A seguir, iremos representar o sinal antes da filtragem inicial por $x(t) = s(t) + n(t)$, na qual $n(t)$ é o ruído de faixa estreita, em contraste com $w(t)$, considerado como ruído branco.

9.2 RAZÕES SINAL/RUÍDO

Sabendo que a comunicação lida com sinais aleatórios – com a aleatoriedade sendo uma qualidade tanto da informação sendo transmitida quanto do ruído interferindo na transmissão – como quantificamos a performance de um sistema particular de comunicação? Neste capítulo, iremos focar na *razão sinal/ruído* (*RSR*), como uma medida da qualidade de sistemas analógicos. Este parâmetro estatístico também possui importância em sistemas digitais, como iremos ver no Capítulo 10.

Como descrito acima e repetido abaixo por conveniência, o sinal recebido em vários sistemas de comunicação pode ser modelado como a soma do sinal desejado, $s(t)$, e um sinal de ruído de faixa estreita, $n(t)$, como mostrado por

$$x(t) = s(t) + n(t) \quad (9.5)$$

Os dois termos do lado direito da Eq. (9.5) são aleatórios. O sinal é aleatório devido à imprevisibilidade de seu conteúdo de informação e o ruído é aleatório pelas razões descritas na Seção 9.1. No Capítulo 8, vimos que os dois parâmetros mais simples para descrever parcialmente uma variável aleatória são a média e a variância. Pelas razões descritas anteriormente, deslocamentos CC são considerados como sendo nulos. Conseqüentemente, para processos de média nula, uma simples medida da qualidade do sinal é a razão das variâncias dos sinais desejado e não desejado. Com esta base, a razão sinal/ruído é formalmente definida por

$$\text{RSR} = \frac{\mathbb{E}[s^2(t)]}{\mathbb{E}[n^2(t)]} \quad (9.6)$$

na qual \mathbb{E} é operador esperança. Para um sinal de comunicação, o nível do sinal quadrático é geralmente proporcional à potência. Conseqüentemente, a razão sinal/ruído é geralmente considerada como a relação da energia média do sinal por unidade de tempo pela energia média do ruído por unidade de tempo. Esta última interpretação é mais comum em sistemas de comunicação digital. Se por acaso utilizarmos a representação complexa do sinal e ruído, então, em vez de valores quadráticos, a Eq. (9.6) deve utilizar valores quadráticos da amplitude.

EXEMPLO 9.1 Razão sinal/ruído senoidal

Considere o caso quando o sinal transmitido da Eq. (9.5) é

$$s(t) = A_c \cos(2\pi f_c t + \theta)$$

na qual a fase θ é desconhecida no receptor. O sinal é recebido na presença de ruído aditivo, como mostrado na Fig. 9.2. O ruído é branco e Gaussiano, com densidade espectral de potência $N_0/2$.

Neste caso, apesar do sinal ser aleatório, ele também é periódico. Conseqüentemente, podemos estimar sua potência média integrando-o em um período (isto é, igualando uma média do espaço amostral com uma média temporal).

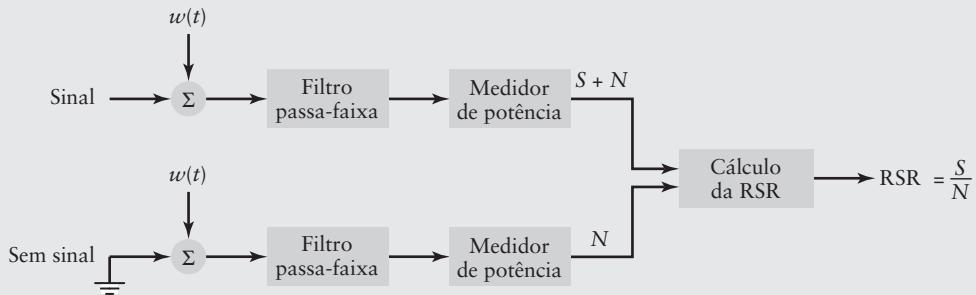


Figura 9.2 Esquema para a medição da RSR par ao Exemplo 9.1.

$$\begin{aligned}
 E[s^2(t)] &= \frac{1}{T} \int_0^T (A_c \cos(2\pi f_c t + \theta))^2 dt \\
 &= \frac{A_c^2}{2T} \int_0^T (1 + \cos(4\pi f_c t + 2\theta))^2 dt \\
 &= \frac{A_c^2}{2T} \left[t + \frac{\sin(4\pi f_c t + 2\theta)}{4\pi f_c} \right]_0^T \\
 &= \frac{A_c^2}{2}
 \end{aligned} \tag{9.7}$$

Teoricamente, o ruído branco se estende para freqüência infinita. O processo ruído de faixa estreita, $n(t)$, é o resultado da passagem do processo ruído branco através de um filtro passa-faixa com largura de faixa B_T equivalente ao ruído. Com esta consideração, calculamos a potência do ruído

$$\begin{aligned}
 E[n^2(t)] &= N \\
 &= N_0 B_T
 \end{aligned} \tag{9.8}$$

Substituindo os resultados das Eqs. (9.7) e (9.8) na Eq. (9.6), a razão sinal/ruído se torna

$$\text{RSR} = \frac{A_c^2}{2N_0 B_T} \tag{9.9}$$

Como a largura de faixa do sinal é arbitrariamente estreita neste exemplo, a escolha da largura de faixa B_T é, de certa forma, arbitrária. Conseqüentemente, na prática, a razão *densidade portadora/ruído* é definida por

$$\frac{C}{N_0} = \frac{A_c^2}{2N_0} \tag{9.10}$$

a qual não é adimensional como a RSR, mas é independente da escolha da largura de faixa. A razão C/N_0 possui unidade de hertz, na qual N_0 é medida em watts por hertz e a potência da portadora $C = A_c^2/2$ é medida em watts.

A razão sinal/ruído é claramente medida no receptor, mas existem vários pontos no receptor no qual a medida pode acontecer. De fato, medidas em pontos particulares do receptor possuem sua própria importância e valor. Por exemplo:

- Se a razão sinal/ruído for medida na frente do receptor, então ela é geralmente uma medida da qualidade do canal de transmissão e da recepção do receptor.

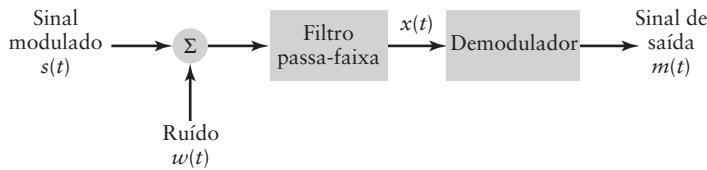


Figura 9.3 Diagrama de blocos de alto nível de um receptor genérico de um sistema de comunicação.

- Se a razão sinal/ruído for medida na saída do receptor, então ela é uma medida da qualidade do sinal de informação recuperado seja ele áudio, vídeo ou qualquer outro tipo.

Para ilustrar estes dois pontos, considere o diagrama em blocos de um receptor de comunicação analógico típico apresentado na Fig. 9.3. O sinal mais o ruído Gaussiano passam por um filtro passa-faixa para produzir o sinal banda passante $x(t)$. O sinal $x(t)$ é processado pelo demodulador para recuperar o sinal original de mensagem $m(t)$. A RSR medida na entrada do demodulador é chamada de *razão sinal/ruído pré-detecção*.

De importância igual ou superior, está a razão sinal/ruído da mensagem recuperada na saída do demodulador. Esta métrica define a qualidade do sinal que é entregue ao usuário final. Chamamos esta RSR de saída de *razão sinal/ruído pós-detecção*. Deve ser observado que as características do sinal e ruído podem diferir significativamente entre os cálculos pré e pós-detecção.

Os cálculos da razão sinal/ruído pós-detecção envolvem a utilização de um modelo idealizado do receptor, cujos detalhes dependem naturalmente do ruído do canal e do tipo de demodulação utilizado pelo receptor. Teremos mais para dizer sobre estas questões em seções subsequentes deste capítulo. Para compararmos diferentes esquemas de modulação-demodulação, introduzimos a idéia de um *modelo de transmissão de referência*, como mostrado na Fig. 9.4. Este modelo de referência é equivalente à transmissão da mensagem banda base. Neste modelo, duas considerações são feitas:

1. A potência da mensagem é a mesma da potência do sinal modulado do esquema de modulação em estudo.
2. O filtro passa-baixa banda base passa o sinal de mensagem e rejeita o ruído fora de faixa. Desta forma, podemos definir a *razão sinal/ruído de referência*, RSR_{ref} , como

$$RSR_{ref} = \frac{\text{potência média do sinal de mensagem modulado}}{\text{potência média do ruído medida na largura de faixa da mensagem}} \quad (9.11)$$

A razão sinal/ruído de referência da Eq. (9.11) pode ser utilizada para comparar esquemas de modulação-demodulação diferentes utilizando-a para normalizar as razões sinal/ruído pós-detecção. Ou seja, podemos definir uma *figura de mérito* para um esquema de modulação/demodulação em particular, como mostrado por

$$\text{Figura de mérito} = \frac{RSR \text{ pós-detecção}}{RSR \text{ de referência}}$$

Claramente, quanto maior o valor da figura de mérito, melhor será a performance frente ao ruído do receptor.

Para resumir nossa consideração sobre razões sinal/ruído:

- A RSR pré-detecção é medida antes do sinal ser demodulado.
- A RSR pós-detecção é medida após o sinal ser demodulado.
- A RSR de referência é definida com base no modelo de transmissão banda base.
- A figura de mérito é uma métrica adimensional para a comparação de esquemas diferentes de modulação/demodulação analógica, sendo definida como a razão entre a RSR pós-detecção e a RSR de referência.

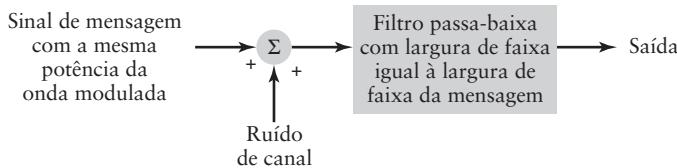


Figura 9.4 Modelo de transmissão de referência para comunicações analógicas.

- **PROBLEMA 9.1** Na prática, não podemos medir somente o sinal, mas o sinal acrescido pelo ruído. Explique como a RSR pode ser calculada neste caso. ◀

9.3 ESTRUTURAS DE RECEPTOR PASSA-FAIXA

Em um sistema de comunicação passa-faixa (ou banda passante), a informação é transmitida em alguma freqüência de portadora, tipicamente utilizando um arranjo similar ao do lado esquerdo da Fig. 9.5. O transmissor inclui um modulador que produz uma saída em uma freqüência intermediária (FI) padrão e um mixer local translada (converte para cima) esta saída para um “canal” ou rádio freqüência (RF).

O lado direito da Fig. 9.5 mostra um exemplo de um *receptor super heterodino* que foi discutido na Seção 3.9. No receptor, o oscilador de freqüência sintonizável translada (converte para baixo) esta freqüência de canal para uma freqüência intermediária (FI) padrão para demodulação. Exemplos comuns são transmissões de rádio AM, no qual a freqüência RF do canal está na faixa entre 510 e 1600 kHz, sendo a FI comum 455 kHz. Outro exemplo é o rádio FM, no qual os canais RF estão na faixa de 88 a 108 MHz e a FI típica é 10,7 MHz.

No Capítulo 3, vimos como representar sinais passa-faixa utilizando a representação em fase e quadratura por

$$s(t) = s_I(t) \cos(2\pi f_c t) - s_Q(t) \sin(2\pi f_c t) \quad (9.12)$$

na qual $s_I(t)$ é a componente em fase de $s(t)$ e $s_Q(t)$ é a componente em quadratura. Vários receptores limitam imediatamente a potência do ruído branco processando o sinal recebido com um filtro passa-faixa, como mostrado na Fig. 9.5. Tipicamente, existe um filtro passa-faixa antes e depois do oscilador local. O filtro antes do oscilador local é centrado em uma freqüência RF mais alta e, geralmente, possui largura de faixa muito maior para caber todos os canais RF que o receptor pode receber. Por exemplo, em um receptor FM, o filtro passa-faixa antes do oscilador local passa todas as freqüências de 88 a 108 MHz. O filtro

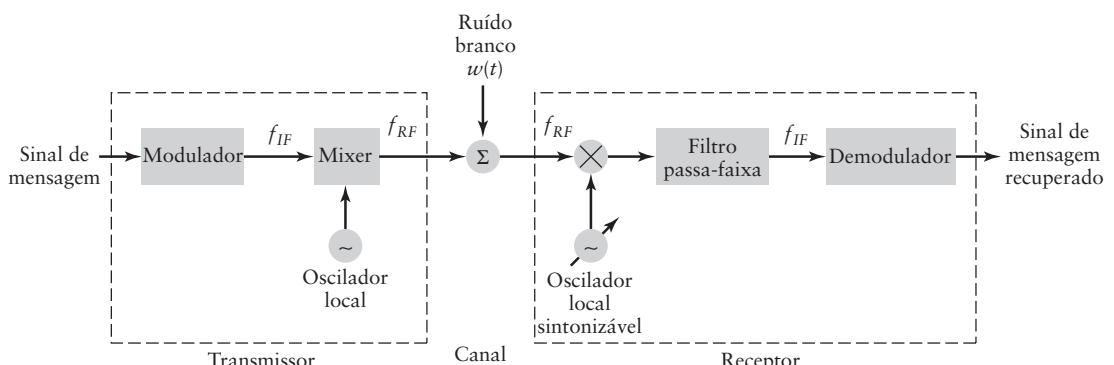


Figura 9.5 Diagrama em blocos de uma transmissão passa-faixa mostrando um receptor super heterodino.

passa-faixa após o oscilador passa o sinal de um único canal RF relativamente não distorcido, mas limita o ruído àquelas componentes dentro da faixa passante do filtro. No mesmo receptor FM, o filtro passa-faixa após o oscilador local tem largura de faixa aproximada de 200 kHz. Os efeitos mais interessantes para nós são os deste filtro de faixa estreita.

9.4 RUÍDO EM RECEPTORES LINEARES UTILIZANDO DETECÇÃO COERENTE

No Capítulo 3, olhamos para a geração e detecção de sinais modulados em amplitude. Uma variedade de sinais modulados em amplitude foi considerada, com sua demodulação dependendo se a portadora estava presente no sinal transmitido ou não. No caso da modulação por faixa lateral dupla-portadora suprimida (DSB-SC), o sinal modulado é representado por

$$s(t) = A_c m(t) \cos(2\pi f_c t + \theta) \quad (9.13)$$

na qual f_c é a freqüência da portadora e $m(t)$ é o sinal de mensagem. A fase θ da portadora é uma variável aleatória, mas invariante durante o curso da transmissão. Para sinais com portadora suprimida, a detecção linear coerente foi identificada como a estratégia de demodulação apropriada. Em particular, um receptor linear para este sinal pode ser implementado como mostrado na Fig. 9.6.

Na Fig. 9.6, o sinal RF recebido é a soma do sinal modulado e ruído branco Gaussiano $w(t)$. O sinal recebido é convertido para básico para uma FI por uma multiplicação por uma senóide de freqüência $f_c - f_{RF}$. Esta conversão para baixo é executada pelo modulador de produto mostrado como mixer 1 na Fig. 9.6. Após a filtragem passa-faixa, o sinal resultante é

$$x(t) = s(t) + n(t) \quad (9.14)$$

na qual $s(t)$ é o sinal modulado não distorcido e $n(t)$ é o ruído passa-faixa na saída do filtro. A densidade espectral de potência considerada para o ruído passa-faixa é ilustrada na Fig. 9.7. A Eq. (9.14) assume que o filtro passa-faixa possui largura suficiente e faixa plana para não causar qualquer distorção significativa ao sinal modulado. Esta equação também assume que o ganho do filtro é unitário. Na prática, o ganho do filtro geralmente é diferente do unitário, mas irá afetar o sinal e o ruído igualmente, portanto não há perda de generalidade na consideração de ganho unitário.

■ RSR pré-detecção

Para sinais analógicos, nossa medida de qualidade é a razão sinal/ruído. Para o sinal $s(t)$ da Eq. (9.13), a potência média das componentes do sinal é dada pela esperança do quadrado da amplitude. Como a portadora e o sinal modulante são independentes, ele pode ser dividido em duas componentes, como mostrado a seguir

$$\text{E}[s^2(t)] = \text{E}[(A_c \cos(2\pi f_c t + \theta))^2] \text{E}[m^2(t)] \quad (9.15)$$

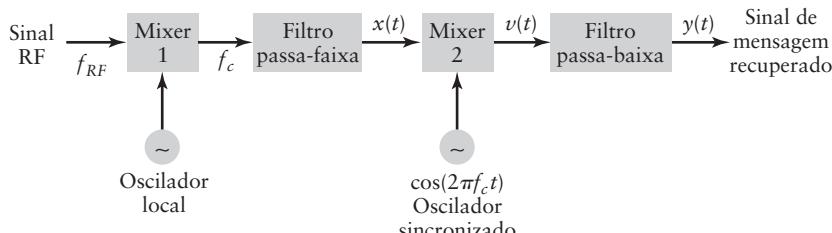


Figura 9.6 Receptor linear DSB-SC utilizando a detecção coerente.

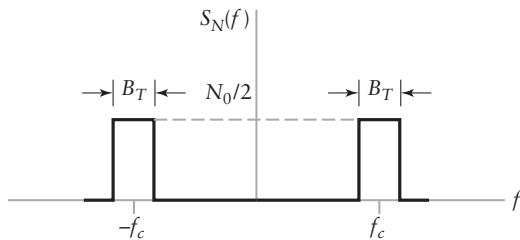


Figura 9.7 Densidade espectral de potência de um ruído passa-faixa.

Se

$$P = \mathbb{E}[m^2(t)] \quad (9.16)$$

for a potência média do sinal (mensagem), utilizando o Exemplo 9.1 para o primeiro fator da Eq. (9.15), temos

$$\mathbb{E}[s^2(t)] = \frac{A_c^2 P}{2} \quad (9.17)$$

Ou seja, a potência média do sinal recebido devido a componente modulada é $A_c^2 P/2$. O sinal $s(t)$ é a saída do modulador de produto que possui entradas $A_c \cos(2\pi f_c t + \theta)$ e $m(t)$, ambos com unidades de volts. Devido ao escalonamento interno, a saída deste modulador de produto possui unidades de volts, não volts ao quadrado. Conseqüentemente, a expressão para $\mathbb{E}[s^2(t)]$ possui unidades de potência e não potência ao quadrado.

Se o filtro passa-faixa possui largura de faixa de ruído B_T , então a potência do ruído que passa por este filtro é, por definição, $N_0 B_T$. Conseqüentemente, a razão sinal/ruído do sinal é

$$\text{RSR}_{\text{pré}}^{\text{DSB}} = \frac{A_c^2 P}{2 N_0 B_T} \quad (9.18)$$

Está é a *razão sinal/ruído pré-detecção* do sistema DSB-SC, pois ela é medida no ponto do sistema antes do sinal $m(t)$ ser demodulado.

■ RSR pós-detecção

A seguir, queremos determinar a RSR pós-detecção ou de saída do sistema DSB-SC. A razão sinal/ruído pós-detecção é a razão da potência do sinal de mensagem pela potência do ruído após a demodulação/detecção. A RSR pós-detecção depende tanto da técnica de modulação quanto da de demodulação.

Utilizando a representação faixa estreita do ruído passa-faixa, o sinal na entrada do detector coerente da Fig. 9.6 pode ser representado por

$$x(t) = s(t) + n_I(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \quad (9.19)$$

na qual $n_I(t)$ e $n_Q(t)$ são as componentes em fase e quadratura de $n(t)$ com respeito à portadora. A saída do mixer 2 da Fig. 9.6 é dada por

$$\begin{aligned} v(t) &= x(t) \cos(2\pi f_c t) \\ &= \frac{1}{2}(A_c m(t) + n_I(t)) \\ &\quad + \frac{1}{2}(A_c m(t) + n_I(t)) \cos(4\pi f_c t) - \frac{1}{2}n_Q(t) \sin(4\pi f_c t) \end{aligned} \quad (9.20)$$

na qual utilizamos as fórmulas de ângulo dobrado

$$\cos A \cos A = \frac{1 + \cos 2A}{2} \quad \text{e} \quad \sin A \cos A = \frac{\sin 2A}{2}$$

Note que a segunda linha da Eq. (9.20) possui duas partes: A primeira parte representa o sinal banda base e a componente em fase do ruído, enquanto que a segunda parte representa a componente em quadratura do ruído centrada em uma freqüência mais alta de $2f_c$. Estas componentes de alta freqüência são removidas pelo filtro passa-baixa da Fig. 9.6, e o resultado é

$$y(t) = \frac{1}{2}(A_c m(t) + n_I(t)) \quad (9.21)$$

Nas Eqs. (9.20) e (9.21), assumimos que os ganhos do segundo mixer e do filtro passa-baixa são unitários, sem perda de generalidade. Duas observações podem ser feitas:

- O sinal de mensagem $m(t)$ e a componente em fase $n_I(t)$ do ruído filtrado aparecem somadas na saída.

- A componente em quadratura do ruído é completamente rejeitada pelo demodulador.

Da Eq. (9.21), podemos calcular a relação sinal/ruído de saída ou pós-detecção observando que:

- A componente de mensagem é $\frac{1}{2}A_c m(t)$, portanto, de forma análoga ao cálculo da potênciia do sinal na pré-detecção, a potênciia do sinal na pós-detecção é $\frac{1}{4}A_c^2 P$, na qual P é a potênciia média da mensagem definida na Eq. (9.16).
- A componente de ruído é $\frac{1}{2}n_I(t)$ após a filtragem passa-baixa. Como descrito na Seção 8.11, a componente em fase possui densidade espectral de N_0 em toda a largura de faixa de $-B_T/2$ à $B_T/2$. Se o filtro passa-baixa possui largura de faixa W , correspondendo à largura de faixa da mensagem, a qual é menor ou igual a $B_T/2$, então a potênciia de saída do ruído é

$$\begin{aligned} \mathbf{E}[n_I^2(t)] &= \int_{-W}^W N_0 df \\ &= 2N_0 W \end{aligned} \quad (9.22)$$

Portanto a potênciia em $n_I(t)$ é $2N_0 W$.

Combinando estas observações, obtemos a RSR pós-detecção como sendo

$$\begin{aligned} \text{RSR}_{\text{pós}}^{\text{DSB}} &= \frac{\frac{1}{4}(A_c^2)P}{\frac{1}{4}(2N_0 W)} \\ &= \frac{A_c^2 P}{2N_0 W} \end{aligned} \quad (9.23)$$

Conseqüentemente, se $W \approx B_T/2$, a RSR pós-detecção é duas vezes a RSR pré-detecção. Isto ocorre em função da componente em quadratura do ruído ter sido removida pelo demodulador síncrono.

■ Figura de mérito

Deve estar claro do modelo de referência de transmissão definido na Seção 9.2, que a potênciia média do ruído para uma mensagem de largura de faixa W é $N_0 W$. Para a modulação DSB-SC, a potênciia média da mensagem modulada é dada pela Eq. (9.17) e, consequentemente, a RSR de referência para este esquema de transmissão é $\text{RSR}_{\text{ref}} = A_c^2 P / (2N_0 W)$. A figura de mérito correspondente para este receptor é

$$\text{Figura de mérito} = \frac{\text{RSR}_{\text{pós}}^{\text{DSB}}}{\text{RSR}_{\text{ref}}} = 1$$

Isto ilustra que não perdemos em performance quanto utilizamos o esquema de modulação passa-faixa quando comparado ao esquema de modulação banda base, mesmo que a largura de faixa do primeiro seja o dobro do sistema banda base. Conseqüentemente, a

modulação DSB-SC fornece a base contra a qual iremos comparar os outros esquemas de detecção de modulação em amplitude.

- **PROBLEMA 9.2** Um sinal DSB-SC é transmitido em um canal ruidoso, com densidade espectral de ruído $N_0/2$ de 2×10^{-17} watts por hertz. A largura de faixa da mensagem é 4 kHz e a frequência da portadora é 200 kHz. Assuma que a potência média recebida do sinal é -80 dBm. Determine a razão sinal/ruído pós-detectação do receptor. ◀

9.5 RUÍDO EM RECEPTORES AM UTILIZANDO A DETECÇÃO DE ENVELOPE

Na Seção 3.1, discutimos a detecção de envelope da modulação em amplitude com portadora suprimida. A detecção de envelope resulta em um receptor mais simples do que a abordagem coerente, pois ele não precisa do circuito para produzir a portadora sincronizada para a demodulação. Lembre-se da Seção 3.1 que o envelope do sinal modulado é representado por

$$s(t) = A_c(1 + k_a m(t)) \cos(2\pi f_c t) \quad (9.24)$$

na qual $A_c \cos(2\pi f_c t)$ é a onda portadora, $m(t)$ é o sinal de mensagem e k_a é a sensibilidade de amplitude do modulador. Para a detecção de envelope, o modelo do receptor é mostrado na Fig. 9.8. A frente deste receptor é idêntica à do receptor coerente. Ou seja, o sinal recebido, incluindo o ruído aditivo, passa pelo filtro passa-faixa.

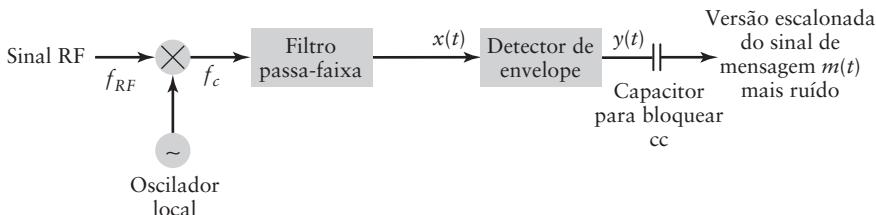


Figura 9.8 Modelo de um receptor AM utilizando a detecção de envelope.

■ RSR pré-detectação

Antes de discutirmos o restante do circuito, vamos considerar a RSR (pré-detectação) deste sinal passa-faixa. Na Eq. (9.24), a potência média da componente de portadora é $A_c^2/2$ devido a natureza senoidal da portadora. A potência na parte modulada do sinal é

$$\begin{aligned} E[(1 + k_a m(t))^2] &= E[1 + 2k_a m(t) + k_a^2 m^2(t)] \\ &= 1 + 2k_a E[m(t)] + k_a^2 E[m^2(t)] \\ &= 1 + k_a^2 P \end{aligned} \quad (9.25)$$

na qual assumimos que o sinal de mensagem possui média nula, $E[m(t)] = 0$, e a potência P da mensagem é definida como na Eq. (9.16). Conseqüentemente, a potência do sinal recebido é $A_c^2(1 + K_a^2 P)/2$. Tal como no receptor linear, assumimos, sem perda de generalidade, que o ganho do filtro passa-faixa é unitário, de tal forma que a razão sinal/ruído é dada por

$$RSR_{\text{pre}}^{\text{AM}} = \frac{A_c^2(1 + k_a^2 P)}{2N_0 B_T} \quad (9.26)$$

na qual B_T é a largura de faixa do ruído do filtro passa-faixa.

■ RSR pós-detectação

Para determinar a razão sinal/ruído pós-detectação, devemos analisar os efeitos do restante do circuito da Fig. 9.8 – em particular, o detector de envelope, o qual pode ser modelado

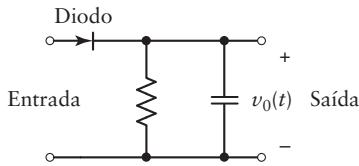


Figura 9.9 Circuito do detector de envelope.

como mostrado na Fig. 9.9 e cuja operação foi descrita na Seção 3.1. Tal como no receptor linear da Seção 9.4, podemos representar o ruído em termos de suas componentes em fase e quadratura e, consequentemente, modelar a entrada do detector de envelope por

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= [A_c + A_c k_a m(t) + n_I(t)] \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \end{aligned} \quad (9.27)$$

O objetivo do detector de envelope é recuperar as variações de amplitude de baixa freqüência do sinal de alta freqüência mostrado na Fig. 9.10. Conceitualmente, isto pode ser representado em um diagrama fasorial como mostrado na Fig. 9.11, no qual a componente de sinal do fasor é $A_c(1 + k_a m(t))$, e o ruído possui duas componentes fasoriais ortogonais, $n_I(t)$ e $n_Q(t)$.

Da Fig. 9.11, a saída do detector de envelope é a amplitude do fasor que representa $x(t)$, sendo dada por

$$\begin{aligned} y(t) &= \text{envelope de } x(t) \\ &= \{[A_c(1 + k_a m(t)) + n_I(t)]^2 + n_Q^2(t)\}^{1/2} \end{aligned} \quad (9.28)$$

A expressão acima é a saída de um detector de envelope ideal. Note que a fase de $x(t)$ não aparece nesta expressão, ela foi eliminada porque não é de nosso interesse. Esta expressão para $y(t)$ é, de alguma forma, complexa, mas pode ser simplificada para fornecer algum resultado intuitivo. Se assumirmos que o sinal é muito maior do que o ruído, então, utilizando a aproximação $\sqrt{A^2 + B^2} \approx$ quando $A \gg B$, podemos escrever

$$y(t) \approx A_c + A_c k_a m(t) + n_I(t) \quad (9.29)$$

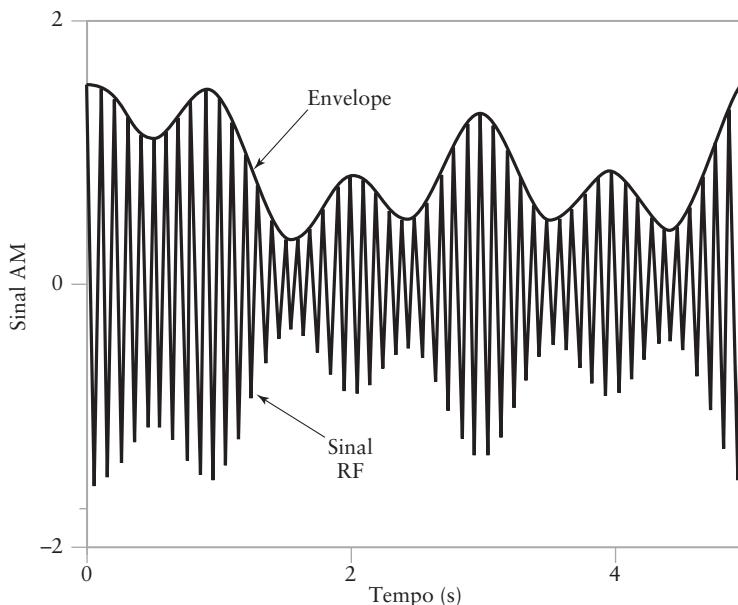


Figura 9.10 Ilustração do envelope de uma portadora de alta freqüência.

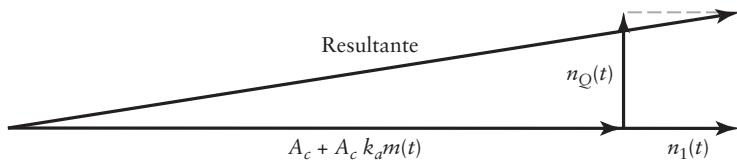


Figura 9.11 Diagrama fasorial para uma onda AM acrescida de ruído de faixa estreita.

sob altas condições de RSR. Esta nova expressão para o sinal demodulado possui três componentes: a componente CC, a componente de sinal e a componente de ruído. O termo CC pode ser removido utilizando um capacitor, como mostrado na Fig. 9.8, deixando apenas os termos de sinal e ruído. O resultado obtido até agora é similar ao observado para o receptor linear.

Desta forma, a RSR pós-detecção para a detecção de envelope de AM, utilizando uma largura de faixa W de mensagem, é dada por

$$\text{RSR}_{\text{pós}}^{\text{AM}} = \frac{A_c^2 k_a^2 P}{2N_0 W} \quad (9.30)$$

na qual o numerador representa a potência média da mensagem $A_c K_a m(t)$ e o denominador representa a potência média de $n_I(t)$ da Eq. (9.22).

Este cálculo da RSR de saída só é válido com duas condições:

- A RSR é alta.
- k_a é ajustado para 100% de modulação ou menos, ou seja, não há distorção no envelope do sinal.

Tal como a modulação em amplitude com portadora suprimida, a largura de faixa W da mensagem é aproximadamente metade da largura de faixa B_T de transmissão.

■ Figura de mérito

Para a modulação AM, a potência média transmitida é dada pelo produto da Eq. (9.25) e da potência da portadora $A_c^2/2$. Conseqüentemente, a RSR de referência é $A_c^2(1+k_a^2P)/(2N_0W)$. Combinando este resultado com a Eq. (9.30), a figura de mérito para este esquema de demodulação de modulação AM é

$$\text{Figura de mérito} = \frac{\text{RSR}_{\text{pós}}^{\text{AM}}}{\text{RSR}_{\text{ref}}} = \frac{k_a^2 P}{1 + k_a^2 P} \quad (9.31)$$

Como o produto $k_a^2 P$ é sempre menor do que a unidade (caso contrário o sinal seria sobremodulado), a figura de mérito para este sistema é sempre menor do que 0,5. Logo, a performance de ruído de um receptor por detecção de envelope é sempre inferior ao receptor DSB-SC, a razão é que ao menos metade da potência é perdida na transmissão da portadora como componente do sinal modulado (transmitido).

Isto completa a análise de ruído para a modulação por envelope, mas notamos que a análise matemática também pode ser executada quando a razão sinal/ruído é baixa.² Da Eq. (9.30), vemos que a RSR pós-detecção possui dependência linear inversa à densidade espectral do ruído. Com condições de baixa razão sinal/ruído, os efeitos não lineares aparecem. A detecção síncrona não possui este comportamento não linear, mas a qualidade da voz ainda é baixa para uma baixa RSR pós-detecção.

■ Experimento de computador: detecção de envelope

No Problema 9.26, descrevemos um experimento de computador que simula a detecção de envelope utilizando a notação fasorial apresentada anteriormente. No experimento, a

² Veja o Capítulo 2 de Haykin (2001).

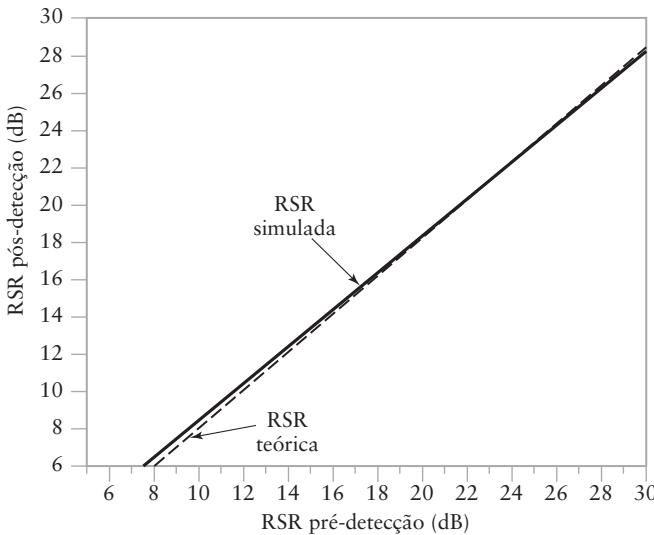


Figura 9.12 Comparação das RSRs pré-detecção e pós-detecção para uma detecção simulada de envelope AM.

mensagem é uma onda senoidal $m(t) = A \operatorname{sen}(2\pi f_m t)$, para a qual criamos um fasor variante no tempo e somamos um fasor correspondente ao ruído. Neste experimento, calculamos as RSRs pré e pós-detecção para amostras deste sinal. Estas duas medidas são traçadas uma contra a outra na Fig. 9.12 para $k_a = 0,3$. Além disso, traçamos também as performances teóricas destas duas medidas, dadas pelas Eqs. (9.26) e (9.30).

A RSR pós-detecção é computada da seguinte maneira:

- ▶ A potência do sinal de saída é determinada passando o sinal sem ruído em um detector de envelope e medindo a potência de saída.
- ▶ A potência de saída do ruído é calculada passando o sinal mais o ruído em um detector de envelope e subtraindo a saída obtida pelo sinal sem ruído. Com esta abordagem, qualquer distorção devido ao produto das componentes de ruído e sinal é incluída na contribuição do ruído.

A razão destas duas quantidades é utilizada como estimativa da RSR de saída.

Como podemos ver pela Fig. 9.12, existe uma concordância entre a teoria e o experimento para altos valores de RSR, como era esperado. Existem algumas pequenas discrepâncias, mas elas podem ser atribuídas a limitações da simulação em tempo discreto. Para RSR baixo, existe alguma variação da teoria, como também era esperado.

- ▶ **PROBLEMA 9.3** Para a mesma potência do sinal recebido, compare as RSRs pós-detecção do DSB-SC com detecção coerente e detecção de envelope para $k_a = 0,2$ e $0,4$. Assuma que a potência média da mensagem é $P = 1$.
- ▶ **PROBLEMA 9.4** Na prática, existe uma fase arbitrária θ na Eq. (9.24). Como ela irá afetar os resultados da Seção 9.5?

9.6 RUÍDO EM RECEPTORES SSB

Iremos considerar, agora, o caso de um receptor coerente de uma onda SSB de entrada. Usando as definições da Seção 3.6, assumimos que apenas a faixa lateral inferior é transmitida, tal que podemos expressar a onda modulada por

$$s(t) = \frac{A_c}{2} m(t) \cos(2\pi f_c t) + \frac{A_c}{2} \hat{m}(t) \operatorname{sen}(2\pi f_c t) \quad (9.32)$$

na qual $\hat{m}(t)$ é a transformada de Hilbert do sinal de mensagem $m(t)$. Podemos fazer as seguintes observações sobre as componentes em fase e quadratura de $s(t)$ na Eq. (9.32):

1. As duas componentes $m(t)$ e $\hat{m}(t)$ são não correlacionadas uma com a outra. Portanto, suas densidades espetrais de potência são aditivas.
2. A transformada de Hilbert $\hat{m}(t)$ é obtida passando $m(t)$ em um filtro linear com função de transferência $-j \operatorname{sgn}(f)$. A amplitude quadrática desta função de transferência é igual a um para todo f . Desta forma, $m(t)$ e $\hat{m}(t)$ possuem a mesma potência média P .

Portanto, procedendo de maneira similar à do receptor DSB-SC, vemos que as componentes em fase e quadratura da onda $s(t)$ modulada SSB contribuem com uma potência média de $A_c^2 P/8$, cada. A potência média de $s(t)$ é, portanto, $A_c^2 P/4$. Este resultado é metade do encontrado para o caso de DSB-SC, sendo intuitivamente satisfatório.

■ RSR pré-detecção

Para o sinal SSB, a largura de faixa B_T de transmissão é aproximadamente igual à largura de faixa W da mensagem. Conseqüentemente, utilizando os cálculos da potência do sinal da seção anterior, a razão sinal/ruído pré-detecção de um receptor coerente com modulação SSB é

$$\text{RSR}_{\text{pre}}^{\text{SSB}} = \frac{A_c^2 P}{4N_0 W} \quad (9.33)$$

■ RSR pós-detecção

Usando o mesmo receptor super heterodino da Fig. 9.6, o sinal passa-faixa após a multiplicação pela saída do oscilador síncrono $\cos(2\pi f_c t)$ é

$$\begin{aligned} v(t) &= x(t) \cos(2\pi f_c t) \\ &= \frac{1}{2} \left(\frac{A_c}{2} m(t) + n_I(t) \right) \\ &\quad + \frac{1}{2} \left(\frac{A_c}{2} m(t) + n_I(t) \right) \cos(4\pi f_c t) - \frac{1}{2} \left(\frac{A_c}{2} \hat{m}(t) + n_Q(t) \right) \sin(4\pi f_c t) \end{aligned} \quad (9.34)$$

Após a filtragem passa-baixa de $v(t)$, ficamos com

$$y(t) = \frac{1}{2} \left(\frac{A_c}{2} m(t) + n_I(t) \right) \quad (9.35)$$

como esperado, vemos que a componente em quadratura $\hat{m}(t)$ do sinal de mensagem foi eliminada da saída do detector. Com um filtro passa-faixa formador para o sinal SSB, o ruído $n(t)$ passa-faixa também será de natureza faixa única. Entretanto, como notado na Eq. (8.98), o espectro da componente em fase $n_I(t)$ do ruído é dada por

$$S_{N_I}(f) = \begin{cases} S_N(f - f_c) + S_N(f + f_c), & -B \leq f \leq B \\ 0, & \text{caso contrário} \end{cases} \quad (9.36)$$

para o caso de faixa lateral única, $S_N(f)$ é $N_0/2$ para $f_c < f < f_c + W$ e para $-f_c - W < f < -f_c$. Conseqüentemente,

$$S_{N_I}(f) = \begin{cases} \frac{N_0}{2}, & -W \leq f \leq W \\ 0, & \text{caso contrário} \end{cases} \quad (9.37)$$

Em outras palavras, a densidade espectral de $n_t(t)$ é de faixa dupla, tal como no caso DSB-SC, mas com metade da potência. Em aplicações práticas, o filtro SSB irá atenuar as componentes de baixa freqüência do sinal mais ruído. Logo, na prática, o espectro de $n_t(t)$ geralmente terá uma falha em CC.

Da Eq. (9.35), a componente de mensagem da saída do receptor é $A_c m(t)/4$, tal que a potência média da mensagem recuperada é $A_c^2 P/16$. A potência do ruído correspondente das Eqs. (9.35) e (9.37) é $\frac{1}{4}N_0 W$. Desta forma, a razão sinal/ruído pós-detecção de um sistema utilizando a modulação DSB no transmissor e detecção coerente no receptor é a razão destas duas potências, ou seja,

$$\text{RSR}_{\text{pós}}^{\text{SSB}} = \frac{A_c^2 P}{4N_0 W} \quad (9.38)$$

■ Figura de mérito

A potência média do sinal de um sistema SSB, como discutida acima, é $A_c^2 P/4$. Conseqüentemente, a RSR de referência é $A_c^2 P/(4N_0 W)$. A figura de mérito para o sistema SSB é a razão da Eq. (9.38) por esta última expressão, ou seja,

$$\text{Figura de mérito} = \frac{\text{RSR}_{\text{pós}}^{\text{SSB}}}{\text{RSR}_{\text{ref}}} = 1 \quad (9.39)$$

Conseqüentemente, a transmissão SSB tem a mesma figura de mérito do DSB-SC. A performance da faixa lateral vestigial com detecção coerente é similar a do SSB.

Comparando os resultados para esquemas diferentes de modulação em amplitude, vemos que existe um número de compromissos de projeto. A modulação por faixa lateral dupla-portadora suprimida fornece a mesma RSR do modelo de referência banda base, mas requer um circuito de sincronização para executar a detecção coerente. O AM sem portadora suprimida simplifica o projeto do receptor significativamente quando ela implementa um detector de envelope. Entretanto, o AM sem portadora suprimida requer mais potência transmitida para obter a mesma performance do modelo de referência banda base. A modulação de faixa lateral única obtém a mesma performance de RSR do modelo de referência banda base, mas requer apenas metade da largura de faixa de transmissão do sistema DSB-SC. Por outro lado, a SSB requer mais processamento no transmissor. *Estas observações são a primeira indicação de que o projeto de um sistema de comunicação envolve o compromisso entre potência, largura de faixa e complexidade de processamento.*

- **PROBLEMA 9.5** O sinal de mensagem do Problema 9.2 tendo largura de faixa W de 4 kHz é transmitido no mesmo canal ruidoso com densidade espectral de ruído $N_0/2$ de 2×10^{-17} watts por hertz utilizando a modulação por faixa lateral única. Se a potência média recebida do sinal é -80 dBm, qual é a razão sinal/ruído pós-detecção do receptor? Compare a largura de faixa do receptor SSB com a do receptor DSB-SC. ◀

9.7 DETECÇÃO DE MODULAÇÃO EM FREQÜÊNCIA (FM)

Agora, nos voltamos, para a detecção de portadoras moduladas em freqüência, com ruído. Lembre-se da Seção 4.1 que o sinal modulado em freqüência é dado por

$$s(t) = A_c \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right] \quad (9.40)$$

na qual A_c é a amplitude da portadora, f_c é a freqüência da portadora, k_f é o fator de sensibilidade de freqüência do modulador e $m(t)$ é o sinal de mensagem. O sinal FM $s(t)$ recebido possui freqüência de portadora f_c e largura de faixa de transmissão B_T , tal que um

total negligenciável de potência está fora da faixa de freqüência $f_c \pm B_T/2$ para freqüências positivas e, de forma similar, para freqüências negativas.

■ RSR pré-detecção

Para a detecção FM, assumimos um modelo de receptor como mostrado na Fig. 9.13. Tal como antes, assumimos que o ruído $w(t)$ é um processo Gaussiano branco de média nula com densidade espectral de potência $N_0/2$. O detector FM consiste de um filtro passa-faixa, um limitador, um discriminador e um filtro passa-baixa. O filtro passa-faixa possui freqüência central f_c e largura de faixa B_T , tal que ele passa o sinal FM sem distorção. Ordinariamente, B_T é pequeno quando comparado com a freqüência central f_c , de tal forma que podemos utilizar a representação faixa estreita para $n(t)$, a versão filtrada do ruído de canal $w(t)$. A RSR pré-detecção neste caso é simplesmente a potência da portadora $A_c^2/2$ dividida pelo ruído passado pelo filtro passa-faixa, $N_0 B_T$, ou seja,

$$\text{RSR}_{\text{pré}}^{\text{FM}} = \frac{A_c^2}{2N_0 B_T}$$

Em um sistema FM, o sinal de mensagem está embutido na freqüência instantânea da portadora. A amplitude da portadora é constante. Portanto, qualquer variação da amplitude da portadora na entrada do receptor deve resultar do ruído ou interferência. O *limitador* de amplitude, seguindo o filtro passa-faixa no modelo do receptor da Fig. 9.13, é utilizado para remover variações de amplitude ceifando a onda modulada. A onda resultante é quase retangular. Esta onda é suavizada por outro filtro passa-faixa que é parte integral do limitador, suprimindo, portanto, as harmônicas da freqüência da portadora que são causadas pelo ceifamento.

O discriminador do modelo da Fig. 9.13 consiste de duas componentes (veja o Capítulo 4):

1. *Circuito de inclinação ou diferenciador* com resposta em freqüência puramente imaginária que varia linearmente com a freqüência. Ele produz uma onda hibridamente modulada, na qual tanto a amplitude quanto a freqüência variam de acordo com o sinal de mensagem.
2. *Um detector de envelope* que recupera a variação de amplitude e reproduz o sinal de mensagem.

O circuito de inclinação e o detector de envelope são geralmente implementados como partes integrais de uma única unidade física. Um exemplo de resposta em amplitude do circuito de inclinação composto é mostrado na Fig. 9.14. Um exemplo de detector de envelope foi discutido na seção anterior. Veja também o Capítulo 3. A utilização combinada do circuito de inclinação e do detector de envelope como demodulador FM foi discutida na Seção 4.8.

O filtro de pós-detecção, chamado de “filtro passa-baixa” na Fig. 9.13, possui largura de faixa que é grande o suficiente para passar a mais alta componente de freqüência do sinal de mensagem. Este filtro remove as componentes fora de faixa do ruído na saída do discriminador e, portanto, mantém o efeito do ruído de saída em um mínimo.

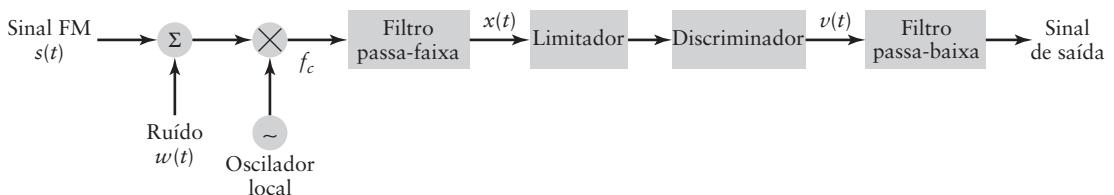


Figura 9.13 Modelo de um receptor FM.

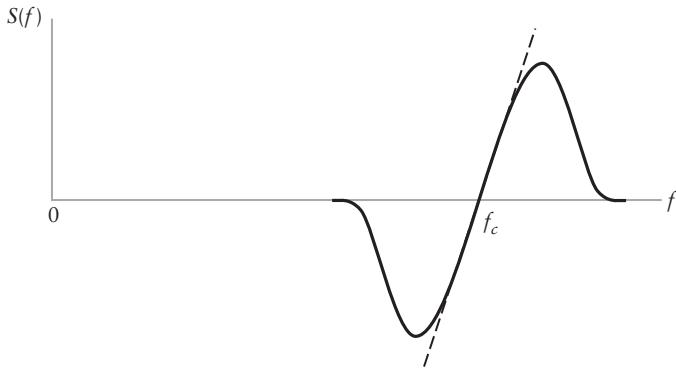


Figura 9.14 Resposta em amplitude do circuito de inclinação utilizado no discriminador FM.

■ RSR pós-detecção

O sinal FM ruidoso após a filtragem passa-faixa pode ser representado por

$$x(t) = s(t) + n(t) \quad (9.41)$$

na qual $s(t)$ é dada pela Eq. (9.40). Nos desenvolvimentos anteriores, expressamos o ruído filtrado $n(t)$ na saída do filtro passa-faixa da Fig. 9.13 em termos de suas componentes em fase e quadratura

$$n(t) = n_I(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \quad (9.42)$$

Podemos expressar, de forma equivalente, $n(t)$ em termos de seu envelope e fase como (veja o Problema 4.3.)

$$n(t) = r(t) \cos[2\pi f_c t + \phi_n(t)] \quad (9.43)$$

na qual o envelope é

$$r(t) = [n_I^2(t) + n_Q^2(t)]^{1/2} \quad (9.44)$$

e a fase é

$$\phi_n(t) = \operatorname{tg}^{-1} \left(\frac{n_Q(t)}{n_I(t)} \right) \quad (9.45)$$

Uma das propriedades desta representação polar é que a fase $\phi_n(t)$ é uniformemente distribuída entre 0 e 2π radianos.

Para procedermos, notamos que a fase de $s(t)$ é

$$\phi(t) = 2\pi k_f \int_0^t m(\tau) d\tau \quad (9.46)$$

Combinando as Eqs. (9.40), (9.43) e (9.46), o sinal ruidoso na saída do filtro passa-faixa pode ser expresso por

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= A_c \cos[2\pi f_c t + \phi(t)] + r(t) \cos[2\pi f_c t + \phi_n(t)] \end{aligned} \quad (9.47)$$

É interessante representar $x(t)$ por um diagrama fasorial, tal como na Fig. 9.15, na qual utilizamos o termo de sinal $s(t)$ como referência. Na Fig. 9.15, a amplitude de ruído é

$r(t)$ e a diferença de fase $\psi_n(t) = \phi_n(t) - \phi(t)$ é o ângulo entre o fasor ruído e o fasor sinal. A fase $\theta(t)$ do resultante é dada por

$$\theta(t) = \phi(t) + \operatorname{tg}^{-1} \left\{ \frac{r(t) \operatorname{sen}(\psi_n(t))}{A_c + r(t) \cos(\psi_n(t))} \right\} \quad (9.48)$$

O envelope de $x(t)$ não é de interesse para nós, pois as variações na saída do filtro passa-faixa são removidas pelo limitador.

Para obter resultados úteis, fazemos algumas aproximações considerando $\theta(t)$. Primeiro, assumimos que a razão portadora/ruído medida na entrada do discriminador é grande. Se R representa observações da função amostra $r(t)$ do envelope do ruído, então em grande parte do tempo a variável aleatória R é pequena quando comparada com a amplitude A_c da portadora. Com esta condição, e observando que $\operatorname{tg}^{-1} \xi \approx \xi$ desde que $\xi \ll 1$, a expressão para a fase é simplificada para

$$\theta(t) = \phi(t) + \frac{r(t)}{A_c} \operatorname{sen}[\psi_n(t)] \quad (9.49)$$

Simplificamos ainda mais esta expressão ignorando a componente de modulação no segundo termo da Eq. (9.49) e substituindo $\psi_n(t) = \phi_n(t) - \phi(t)$ por $\phi_n(t)$. Isto é justificado porque a fase $\phi_n(t)$ é uniformemente distribuída entre 0 e 2π radianos e, como $\phi(t)$ é independente de $\phi_n(t)$, é razoável assumir que a diferença de fase $\phi_n(t) - \phi(t)$ também é uniformemente distribuída em 2π radianos. Considerações teóricas mostram que esta consideração é justificada desde que a razão portadora/ruído seja alta.³ Então, notando que a componente de quadratura do ruído é $n_Q(t) = r(t) \operatorname{sen}[\phi_n(t)]$, podemos simplificar a Eq. (9.49) para

$$\theta(t) = \phi(t) + \frac{n_Q(t)}{A_c} \quad (9.50)$$

Usando a expressão para $\phi(t)$ dada pela Eq. (9.46), a Eq. (9.50) pode ser expressa por

$$\theta(t) \approx 2\pi k_f \int_0^t m(\tau) d\tau + \frac{n_Q(t)}{A_c} \quad (9.51)$$

Nosso objetivo é determinar o erro na freqüência instantânea da onda portadora causado pela presença do ruído filtrado $n(t)$. Com um discriminador ideal, sua saída é proporcional à derivada $d\theta(t)/dt$. Usando a expressão para $\theta(t)$ na Eq. (9.51), a saída do discriminador ideal, escalonada por 2π , é, portanto,

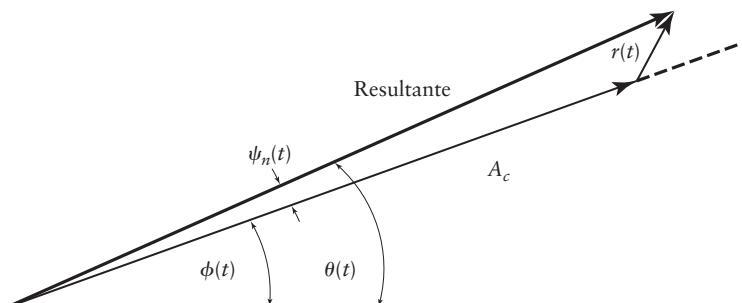


Figura 9.15 Diagrama fasorial para o sinal FM mais o ruído de faixa estreita, assumindo alta razão portadora/ruído.

³ Veja o Capítulo 5 de Downing (1964).

$$\begin{aligned} v(t) &= \frac{1}{2\pi} \frac{d\theta(t)}{dt} \\ &= k_f m(t) + n_d(t) \end{aligned} \quad (9.52)$$

na qual o ruído $n_d(t)$ é definido por

$$n_d(t) = \frac{1}{2\pi A_c} \frac{dn_Q(t)}{dt} \quad (9.53)$$

Vemos, agora, que, desde que a razão portadora/ruído seja alta, a saída $v(t)$ do discriminador consiste do sinal de mensagem original $m(t)$ multiplicado pelo fator constante k_f mais uma componente de ruído aditivo $n_d(t)$. O ruído aditivo na saída do discriminador é determinado essencialmente pela componente em quadratura $n_Q(t)$ do ruído de faixa estreita $n(t)$.

Desta forma, podemos utilizar a razão sinal/ruído pós-detecção, como definida anteriormente, para avaliar a qualidade da saída do receptor FM. A razão sinal/ruído pós-detecção é definida como a razão a potência média do sinal de saída pela potência média do ruído de saída. Da Eq. (9.52), vemos que a componente de mensagem da saída do discriminador e, portanto, a saída do filtro passa-baixa, é $k_f m(t)$. Logo, a potência média do sinal de saída é igual a $k_f^2 P$, na qual P é a potência média do sinal de mensagem $m(t)$.

Para determinar a potência média do ruído de saída, observamos que o ruído $n_d(t)$ na saída do discriminador é proporcional à derivada temporal da componente em quadratura $n_Q(t)$ do ruído. Como a diferenciação de uma função com respeito ao tempo corresponde a uma multiplicação de sua transformada de Fourier por $j2\pi f$, podemos obter o processo ruído $n_d(t)$ passando $n_Q(t)$ em um filtro linear com resposta em freqüência igual a

$$G(f) = \frac{j2\pi f}{2\pi A_c} = \frac{jf}{A_c} \quad (9.54)$$

Isto significa que a densidade espectral de potência $S_{N_d}(f)$ do ruído $n_d(t)$ é relacionada com a densidade espectral de potência $S_{N_Q}(f)$ da componente em quadratura $n_Q(t)$ por

$$\begin{aligned} S_{N_d}(f) &= |G(f)|^2 S_{N_Q}(f) \\ &= \frac{f^2}{A_c^2} S_{N_Q}(f) \end{aligned} \quad (9.55)$$

Com o filtro passa-faixa do modelo de receptor da Fig. 9.13 tendo uma resposta em freqüência ideal, caracterizada por uma largura de faixa B_T e freqüência de meia faixa f_c , temos que o ruído faixa estreita $n(t)$ terá uma característica de densidade espectral com forma similar. Se o ruído de entrada for branco, então, das propriedades das componentes em fase e quadratura do ruído de faixa estreita descritas na Seção 8.11, a densidade espectral de potência de $n_Q(t)$ será o equivalente passa-baixa da soma das respostas de freqüência positiva e negativa do filtro passa-faixa. Isto significa que a componente de quadratura $n_Q(t)$ do ruído de faixa estreita terá uma característica passa-baixa ideal mostrada na Fig. 9.16(a). A densidade espectral de potência correspondente do ruído $n_d(t)$ é mostrada na Fig. 9.16, ou seja,

$$S_{N_d}(f) = \begin{cases} \frac{N_0 f^2}{A_c^2}, & |f| < \frac{B_T}{2} \\ 0, & \text{caso contrário} \end{cases} \quad (9.56)$$

No modelo do receptor da Fig. 9.13, a saída do discriminador é seguida por um filtro passa-baixa com largura de faixa igual à largura de faixa W da mensagem. Para um FM de faixa larga, geralmente vemos que W é menor do que $B_T/2$, na qual B_T é a largura de

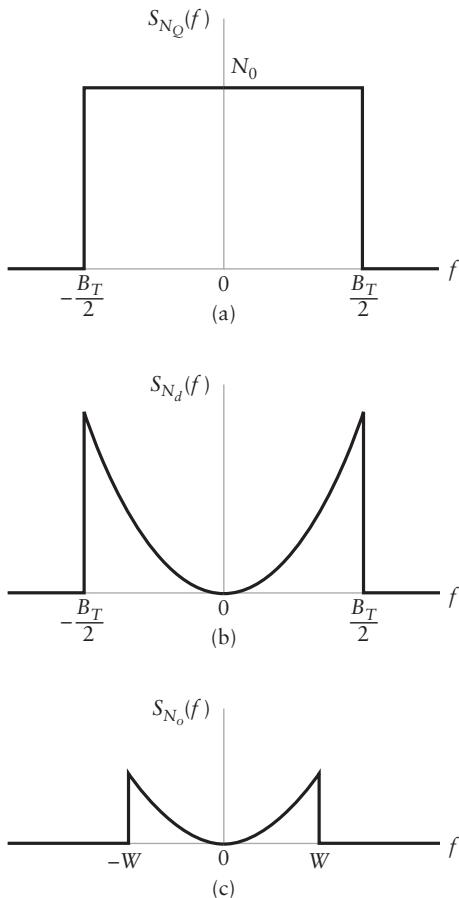


Figura 9.16 Análise de ruído de um receptor FM. (a) Densidade espectral de potência da componente em quadratura $n_Q(t)$ do ruído de faixa estreita $n(t)$. (b) Densidade espectral de potência $n_d(t)$ na saída do discriminador. (c) Densidade espectral de potência do ruído $n_o(t)$ na saída do receptor.

faixa de transmissão do sinal FM. Isto significa que as componentes fora de faixa do ruído $n_d(t)$ serão rejeitadas. Portanto, a densidade espectral de potência $S_{N_o}(f)$ do ruído $n_o(t)$ que aparece na saída do receptor é definida por

$$S_{N_o}(f) = \begin{cases} \frac{N_0 f^2}{A_c^2}, & |f| < W \\ 0, & \text{caso contrário} \end{cases} \quad (9.57)$$

como mostrado na Fig. 9.16(c). A potência média do ruído de saída é determinada pela integração da densidade espectral de potência $S_{N_o}(f)$ de $-W$ a W . Fazendo isto, obtemos o seguinte resultado:

$$\begin{aligned} \text{Potência média do ruído pós-detecção} &= \frac{N_0}{A_c^2} \int_{-W}^W f^2 df \\ &= \frac{2N_0 W^3}{3A_c^2} \end{aligned} \quad (9.58)$$

Das Eqs. (9.52) e (9.16), a potência média do sinal de saída é $k_f^2 P$. Como mencionado anteriormente, desde que a razão portadora/ruído seja alta, a razão sinal/ruído pós-detecção é igual à razão de $k_f^2 P$ pelo lado direito da Eq. (9.58), logo,

$$\text{RSR}_{\text{pós}}^{\text{FM}} = \frac{3A_c^2 k_f^2 P}{2N_0 W^3} \quad (9.59)$$

Logo, a RSR pós-detecção de um demodulador FM possui uma dependência não linear tanto com a sensibilidade de freqüência quanto com a largura de faixa da mensagem.

■ Figura de mérito

Com a modulação FM, a potência do sinal modulado é simplesmente $A_c^2/2$, logo a RSR de referência é $A_c^2/(2N_0 W)$. Conseqüentemente, a figura de mérito para este sistema é dada por

$$\begin{aligned} \text{Figura de mérito} &= \frac{\text{RSR}_{\text{pós}}^{\text{FM}}}{\text{RSR}_{\text{ref}}} = \frac{\frac{3A_c^2 k_f^2 P}{2N_0 W^3}}{\frac{A_c^2}{2N_0 W}} \\ &= 3 \left(\frac{k_f^2 P}{W^2} \right) \\ &= 3D^2 \end{aligned} \quad (9.60)$$

na qual, na última linha, introduzimos a definição $D = k_f P^{1/2}/W$ como desvio de razão do sistema FM à luz do material apresentado na Seção 4.6. Lembre-se daquela seção que a regra generalizada de Carson resulta na largura de faixa de transmissão $B_T = 2(k_f P^{1/2} + W) \approx 2k_f P^{1/2}/W$ para um sinal FM. Portanto, substituindo $B_T/2$ por $k_f P^{1/2}$ na definição de D , a figura de mérito de um sistema FM é aproximadamente dada por

$$\text{Figura de mérito} \approx \frac{3}{4} \left(\frac{B_T}{W} \right)^2 \quad (9.61)$$

Conseqüentemente, um aumento na largura de faixa B_T de transmissão fornece um aumento quadrático correspondente na razão sinal/ruído de saída de um sistema FM comparado com o sistema de referência. Portanto, quando o nível portadora/ruído está alto, ao contrário do sistema de modulação em amplitude, o sistema FM permite *trocar largura de faixa por melhoria na performance* de acordo com esta lei quadrática.

EXEMPLO 9.2 Ruído em canais FM multiplexados

Na estratégia de FM estéreo multiplexado descrito na Seção 4.9, queremos determinar a RSR pós-detecção do sinal estéreo diferença $m_l(t) - m_r(t)$, assumindo que a largura de faixa de transmissão é 200 kHz, a largura de faixa banda base é 19 kHz e a RSR pré-detecção é 12 dB.

Do material apresentado anteriormente nesta seção, a RSR pré-detecção é

$$\text{RSR}_{\text{pre}} = \frac{A_c^2}{2N_0 B_T}$$

Ignorando a componente piloto, a potência média do ruído de pós-detecção no canal superior do FM multiplexado é

$$\begin{aligned} N &= \frac{N_0}{A_c^2} \times 2 \int_W^{3W} f^2 df \\ &= \frac{52 W^3 N_0}{3 A_c^2} \end{aligned}$$

A RSR pós-detecção do canal superior é

$$\begin{aligned} \text{RSR}_{\text{pós}} &= \frac{3A_c^2 k_f^2 (P/2)}{52N_0 W^3} \\ &= \frac{A_c^2}{2N_0 B_T} \frac{3}{52} \frac{k_f^2 P B_T}{W^3} \\ &\approx \frac{A_c^2}{2N_0 B_T} \frac{3/4}{52} \left(\frac{B_T}{W} \right)^3 \\ &= 266,6 \\ &\approx 24,3 \text{ dB} \end{aligned}$$

na qual utilizamos a aproximação para a regra de Carson de $B_T \approx k_f P^{1/2}$ e assumimos que metade da potência está no canal superior.

Efeito limiar

A fórmula da Eq. (9.59), definindo a RSR pós-detecção de um receptor FM, é válida somente se a RSR pré-detecção, medida na entrada do discriminador, for alta comparada com a unidade. Se a RSR pré-detecção for diminuída, o receptor FM deixa de funcionar. No início, clicks individuais são ouvidos na saída do receptor e, quando a RSR pré-detecção diminui ainda mais, os clicks se fundem em um sinal ininteligível. E abaixo deste valor de falha, a Eq. (9.59) falha em prever precisamente a RSR pós-detecção. Este fenômeno é chamado de *efeito limiar*, sua determinação, entretanto, está além do escopo deste livro.

Experimento de computador: efeito limiar em FM

No Problema 9.27 descrevemos um experimento de computador para simular a detecção de um sinal FM com ruído. O sinal e o ruído são gerados utilizando a notação fasorial complexa, mas eles são subsequentemente convertidos para cima para uma FI, criando um sinal passa-faixa, aplicado à detecção por discriminador.

O fasor complexo do sinal FM é dado por

$$s(t) = A_c \exp \left\{ -j2\pi k_f \int_0^t m(\tau) d\tau \right\}$$

no qual (para os resultados a seguir), temos $m(t) = \sin(2\pi t)$ e $k_f = 20$ hertz por unidade de amplitude. Este sinal foi transmitido com uma largura de faixa passa-faixa de transmissão de $B_T = 62,5$ Hz e uma largura de faixa de detecção banda base de $W = 5,5$ Hz.

Similar ao experimento de computador AM, medimos as RSRs pré e pós-detecção do sinal e comparamos os resultados da teoria desenvolvida nesta seção. Estes resultados estão mostrados na Fig. 9.17 para o conjunto acima de parâmetros. Os resultados mostram que a teoria claramente acompanha a performance medida em alto RSR. Para baixo RSR, temos o efeito limiar mencionado, no qual a RSR pós-detecção degrada rapidamente quando a RSR pré-detecção diminui abaixo de 10dB.

- **PROBLEMA 9.6** O sinal $m(t) = \cos(2000\pi t)$ é transmitido através de modulação em freqüência. Se a sensibilidade de freqüência k_f for 2 kHz por volt, qual é a largura de faixa da regra de Carson do sinal FM? Se a RSR pré-detecção é 17 dB, calcule a RSR pós-detecção. Assuma que o demodulador inclui um filtro passa-baixa ideal com largura de faixa de 3,1 kHz. ◀
- **PROBLEMA 9.7** Calcule a RSR pós-detecção no canal inferior do Exemplo 9.2 e compare com a do canal superior. ◀

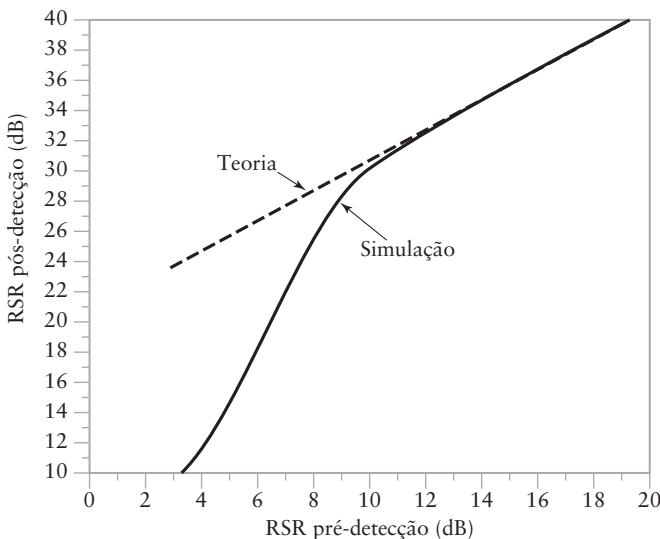


Figura 9.17 Comparação das RSRs pré-detecção e pós-detecção para um receptor FM simulado. ($f_m = 1$ Hz, $k_f = 20$ Hz, $B_T = 62,5$ Hz, $W = 5,5$ Hz, $A_c = 1$)

9.8 PRÉ-ÊNFASE E DE-ÊNFASE FM

Devido à natureza da lei quadrática do espectro do ruído de saída de um receptor FM, o ruído é mais severo em grandes valores de $|f|$. Isto se torna uma questão significante na transmissão de FM estéreo na qual o canal superior ($m_l(t) - m_r(t)$) sofre significativamente mais ruído do que o canal inferior ($m_l(t) + m_r(t)$).

Suponha que o demodulador inclua um filtro passa-baixa que gradualmente aumenta a atenuação quando $|f|$ aproxima-se de W ao invés de ficar aproximadamente plano para $|f| < W$ e corta rapidamente para W . Tal filtro, com função de transferência $H_{de}(f)$, é apresentado na Fig. 9.18(b). Este filtro irá *de-enfatizar* os efeitos do ruído em altas freqüências, como ilustrado na figura.

Além de reduzir o ruído, o filtro de de-ênfase também irá distorcer o sinal recebido. Para compensar esta distorção, pré-distorcemos, ou *pré-enfatizamos*, adequadamente o sinal banda base no transmissor, antes da modulação FM, utilizando um filtro com resposta em freqüência

$$H_{pre}(f) = \frac{1}{H_{de}(f)} \quad |f| < W \quad (9.62)$$

Com uma combinação de pré-ênfase de de-ênfase, como descrito pelas Eqs. (9.61) e (9.62), o sinal é recuperado sem distorção e, mais importante, com níveis reduzidos de ruído.

O *filtro de de-ênfase* é geralmente um simples circuito resistor-capacitor (RC) com

$$H_{de}(f) = \frac{1}{1 + j \frac{f}{f_{3dB}}} \quad (9.63)$$

Este filtro é aproximadamente plano para $|f| < f_{3dB}$, a largura de faixa de 3 dB do filtro. Com esta escolha, o espectro do ruído para $|f| > f_{3dB}$ se torna plano para a maior parte da largura de faixa da mensagem, como mostrado na Fig. 9.18(c).

No fim da transmissão, o filtro de pré-ênfase é

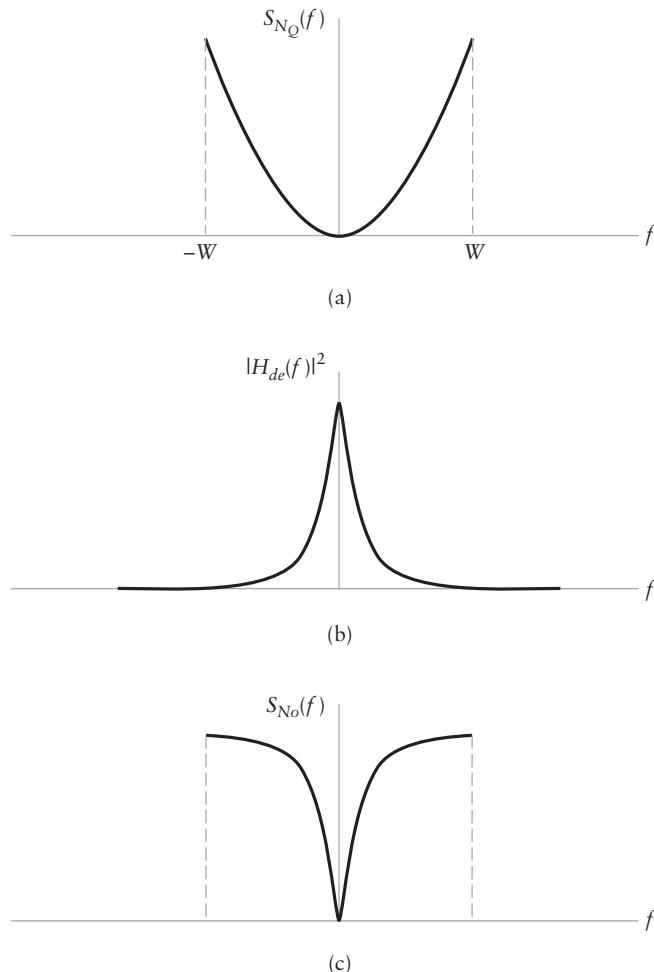


Figura 9.18 Utilização de pré-ênfase e de-ênfase em um sistema FM. (a) Espectro do ruído de saída antes da de-ênfase. (b) Resposta em freqüência do filtro de de-ênfase. (c) Espectro do ruído após a de-ênfase.

$$H_{\text{pre}}(f) = 1 + j \frac{f}{f_{3\text{dB}}} \quad (9.64)$$

Este segundo filtro possui pouco efeito em baixas freqüências na qual \$|f| < f_{3\text{dB}}\$. Em freqüências \$|f| > f_{3\text{dB}}\$, o filtro da Eq. (9.64) é aproximadamente equivalente a um diferenciador (lembre-se que a função de transferência de um diferenciador é \$j2\pi f\$). Portanto, o sinal pré-enfatizado é a soma do sinal original com a sua derivada. Conseqüentemente, o sinal modulado é aproximadamente,

$$\begin{aligned} s(t) &= A_c \cos \left(2\pi f_c t + 2\pi k_f \int_0^t \left(m(s) + \alpha \frac{dm(s)}{ds} \right) ds \right) \\ &= A_c \cos \left(2\pi f_c t + 2\pi k_f \int_0^t m(s) ds + 2\pi k_f \alpha m(t) \right) \end{aligned} \quad (9.64)$$

na qual \$\alpha = 1/f_{3\text{dB}}\$. Portanto, o FM pré-enfatizado é realmente uma combinação de modulação em freqüência e modulação em fase.

A pré-ênfase é utilizada em várias aplicações além da transmissão broadcasting de FM. A pré-ênfase pode ser utilizada para melhorar qualquer porção da faixa de mensagem que estiver degradada em relação às outras. Ou seja, porções da faixa de mensagem que são mais sensíveis ao ruído são amplificadas (enfatizadas) antes da transmissão. No receptor, o sinal é desenfatizado para reverter a distorção introduzida pelo transmissor. Ao mesmo tempo, a de-ênfase reduz o ruído que está na parte mais sensível da faixa de mensagem. Por exemplo, o sistema Dolby para gravação em fita pré-enfatiza altas freqüências para a gravação do som, de tal forma que o ruído de superfície pode ser desenfatizado quando a fita for tocada.

EXEMPLO 9.3 Melhoria pré-ênfase

Neste exemplo, trabalhamos com a melhoria na pós-detecção da razão sinal/ruído de um receptor FM com circuitos de pré e de-ênfase das Eqs. (9.64) e (9.63).

Da Eq. (9.58), a potência do ruído sem a de-ênfase é dada por

$$N = \frac{2}{3} \frac{N_0 W^3}{A_c^2}$$

A potência de ruído correspondente com a de-ênfase é dada por

$$\begin{aligned} N_{de} &= \int_{-W}^{W} \frac{N_0 f^2 / A_c^2}{1 + (f/f_{3dB})^2} df \\ &= 2 \frac{N_0}{A_c^2} [(W/f_{3dB}) - \operatorname{tg}^{-1}(W/f_{3dB})] (f_{3dB})^3 \end{aligned}$$

A melhoria devido à de-ênfase é a razão destas duas últimas expressões,

$$I = \frac{N}{N_{de}} = \frac{(W/f_{3dB})^3}{3[(W/f_{3dB}) - \operatorname{tg}^{-1}(W/f_{3dB})]}$$

Na transmissão broadcasting comercial de FM, normalmente temos $f_{3dB} = 2,1$ kHz e podemos razoavelmente assumir $W = 15$ kHz. Este conjunto de valores leva a $I = 22$, o que corresponde a uma melhoria de 13dB na razão sinal/ruído pós-detecção do receptor. Este exemplo ilustra que uma melhoria significativa na performance de ruído de um sistema FM pode ser obtida usando filtros pré-ênfase e de-ênfase construídos por circuitos RC simples.

- **PROBLEMA 9.8** Um sistema FM possui uma RSR pré-detecção de 15dB. Se a largura de faixa de transmissão é 30 MHz e a largura de faixa da mensagem é 6 MHz, qual é a RSR pós-detecção? Suponha que o sistema inclui filtros de pré-ênfase e de-ênfase como descrito pelas Eqs. (9.63) e (9.64). Qual é a RSR pós-detecção se a f_{3dB} do filtro de de-ênfase for 800 kHz? ◀

9.9 RESUMO E DISCUSSÃO

Neste capítulo, estudamos a detecção de vários sinais modulados analógicamente na presença de ruído aditivo. Começamos revisando as características de ruído Gaussiano, o que mais prejudica em sistemas de comunicação. Em seguida, introduzimos os conceitos de razão sinal/ruído (RSR) pré-detecção e pós-detecção e estabelecemos a razão da RSR pós-detecção para a RSR de um modelo de referência banda base como figura de mérito para comparar sistemas de comunicação analógica.

Subseqüentemente, analisamos a performance de ruído para vários esquemas diferentes de modulação em amplitude e percebemos:

- (i) A detecção de DSB-SC com receptor linear coerente possui a mesma performance de RSR do modelo de referência banda base, mas requer um circuito de sincronização para recuperar a portadora coerente para a demodulação.
- (ii) Sistemas AM com portadora não suprimida permitem um projeto simples de receptor incluindo a utilização de um detector de envelope, mas resultam em perda significativa de potência transmitida comparada com sistemas coerentes.
- (iii) A modulação SSB analógica possui a mesma performance de RSR do DSB-SC mas requer apenas metade da largura de faixa de transmissão.

A análise da performance de ruído em sinais FM indica que o espectro do ruído de saída possui uma forma parabólica diferente da observada na detecção AM. Este comportamento resulta na capacidade do FM de trocar largura de faixa de sinal por melhoria na performance de ruído de acordo com um lei quadrática. A utilização de pré e de-ênfase pode melhorar significativamente as razões sinal/ruído pós-detecção de sistemas FM. Também observamos que a detecção por discriminador de FM exibe um fenômeno de limiar no qual a performance rapidamente diminui para uma baixa RSR pré-detecção.

Neste capítulo, mostramos a importância da análise do ruído baseado na razão sinal/ruído para o cálculo da performance de sistemas de comunicação analógicas. Este tipo de análise de ruído é fundamental para o entendimento e projeto de qualquer sistema de comunicação, seja ela analógica ou digital.

Problemas adicionais

9.9 Uma função amostra

$$x(t) = A_c \cos(2\pi f_c t) + w(t)$$

é aplicada a um filtro RC passa-baixa. A amplitude A_c e a freqüência f_c da componente senoidal são constantes e $w(t)$ é ruído branco de média nula e densidade espectral de potência $N_0/2$. Encontre uma expressão para a razão sinal/ruído de saída com a componente senoidal de $x(t)$ considerada como sinal de interesse.

- 9.10 Um sinal modulado DSB-SC é transmitido em um canal ruidoso com densidade espectral de potência do ruído mostrada na Fig. 9.19. A largura de faixa da mensagem é 4 kHz e a freqüência da portadora é 200 kHz. Assuma que a potência média recebida do sinal é -80 dBm e determine a razão sinal/ruído de saída deste receptor.

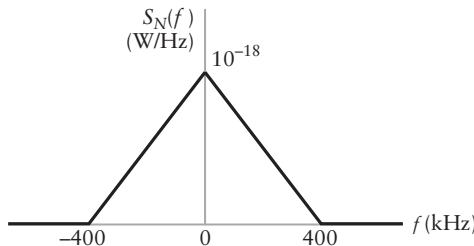


Figura 9.19 Problema 9.10.

- 9.11 Obtenha uma expressão para a razão sinal/ruído pós-detecção do receptor coerente da Fig. 9.6, assumindo que o sinal modulado $s(t)$ é produzido pela onda senoidal modulante

$$m(t) = A_m \cos(2\pi f_m t)$$

Execute seus cálculos para os dois tipos de receptor a seguir:

- (a) Receptor DSB-SC coerente
- (b) Receptor SSB coerente

- 9.12 Calcule a função de autocorrelação das componentes em fase e quadratura de um ruído de faixa estreita na entrada de um detector coerente de um sistema DSB-SC. Assuma que a densidade espectral do ruído é $S_N(f) = N_0/2$ para $|f - f_c| < B_T$.
- 9.13 Assuma que um sinal de mensagem $m(t)$ possui densidade espectral de potência

$$S_M(f) = \begin{cases} a \frac{|f|}{W}, & |f| \leq W \\ 0, & \text{caso contrário} \end{cases}$$

na qual a e W são constantes. Obtenha a expressão para a RSR pós-detecção do receptor quando:

- (a) O sinal é transmitido por DSB-SC.
- (b) O sinal é transmitido por modulação em amplitude com sensibilidade em amplitude $k_a = 0,3$.
- (c) O sinal é transmitido usando modulação em freqüência com sensibilidade em freqüência $k_f = 500$ hertz por volt.

Assuma que ruído branco Gaussiano de média nula e densidade espectral de potência $N_0/2$ é adicionado ao sinal na entrada do receptor.

- 9.14 Um transmissor de 10 kW modula em amplitude uma portadora com um tom $m(t) = \sin(2000\pi t)$, utilizando 50% de modulação. Perdas de propagação entre o transmissor e o receptor atenuam o sinal por 90 dB. O receptor possui um ruído na sua frente com densidade espectral $N_0 = -113$ watts/Hz e inclui um filtro passa-faixa com largura de faixa $B_T = 2 W = 10$ kHz. Qual é a razão sinal/ruído pós-detecção, assumindo que o receptor utiliza um detector de envelope?
- 9.15 A potência média de ruído por unidade de largura de faixa medida na frente de um receptor AM é 10^{-6} watts por Hz. O sinal modulante é senoidal, com uma potência de portadora de 80 watts e uma potência de faixa lateral de 10 watts por faixa lateral. A largura de faixa de mensagem é 4 kHz. Assumindo a utilização de um detector de envelope no receptor, determine a relação sinal/ruído de saída do sistema. Quantos decibels este sistema é inferior ao sistema de modulação DSB-SC?
- 9.16 Um receptor AM, operando com onda modulante senoidal e 80% de modulação, possui uma razão sinal/ruído pós-detecção de 20 dB. Qual é a razão sinal/ruído pré-detecção correspondente?
- 9.17 O sinal $m(t) = \cos(400\pi t)$ é transmitido via FM. Existe um filtro passa-faixa ideal passando $100 \leq |f| \leq 200$ na saída do discriminador. Calcule a RSR pós-detecção dado que $k_f = 1$ kHz por volt e a RSR pré-detecção é 500. Utilize a regra de Carson para estimar a largura de faixa pré-detecção.
- 9.18 Suponha que o espectro de um sinal modulante ocupa a faixa de freqüências $f_1 \leq |f| \leq f_2$. Para acomodar este sinal, o receptor de um sistema FM (sem pré-ênfase) utiliza um filtro passa-faixa ideal conectado à saída do discriminador de freqüência. O filtro passa as freqüências no intervalo $f_1 \leq |f| \leq f_2$. Determine a razão sinal/ruído de saída e a figura de mérito do sistema na presença de ruído branco aditivo na entrada do receptor.
- 9.19 Um sistema FM, operando com uma RSR pré-detecção de 14 dB, requer uma RSR pós-detecção de 30 dB e possui uma potência de mensagem de 1 watt e largura de faixa de 50 kHz. Usando a regra de Carson, estime qual deve ser a largura de faixa de transmissão do sistema. Suponha que este sistema inclui circuitos de pré-ênfase e de-ênfase com f_{3dB} de 10 kHz. Qual é a largura de faixa de transmissão necessária neste caso?

Problemas avançados

- 9.20 Assuma que um ruído de faixa estreita $n(t)$ é Gaussiano e que sua densidade espectral de potência $S_N(f)$ é simétrica com relação à freqüência de meia faixa f_c . Mostre que as componentes em fase e quadratura de $n(t)$ são estatisticamente independentes.
- 9.21 Suponha que a resposta em amplitude do filtro passa-faixa do receptor $|H_{BP}(f)|$ possui simetria em $\pm f_c$ e largura de faixa de ruído B_T . Das propriedades de ruído de faixa estreita descritas na Seção 8.11, qual é a densidade espectral $S_N(f)$ das componentes em fase e quadratura do ruído de faixa estreita $n(t)$ na saída do filtro? Mostre que a autocorrelação de $n(t)$ é

$$R_N(\tau) = \rho(\tau) \cos(2\pi f_c \tau)$$

na qual $\rho(\tau) = F^{-1}[S_N(f)]$. Justifique a aproximação $\rho(\tau) \approx 1$ para $|\tau| \ll 1/B_T$.

- 9.22 Assuma que, no demodulador DSB-SC da Fig. 9.6, existe um erro de fase ϕ no oscilador sincronizado, tal que a sua saída é $\cos(2\pi f_c t + \phi)$. Obtenha uma expressão para a saída do detector coerente e mostre que a RSR pós-detecção é reduzida pelo fator $\cos^2 \phi$.
- 9.23 Em um receptor utilizando detecção coerente, a onda senoidal gerada pelo oscilador local sofre de erro de fase $\theta(t)$ com respeito a onda portadora $\cos(2\pi f_c t)$. Assumindo que $\theta(t)$ é um processo Gaussiano de média nula e variância σ_θ^2 e que na maior parte do tempo o valor máximo de $\theta(t)$ é pequeno comparado com a unidade, obtenha o erro médio quadrático da saída do receptor para modulação DSB-SC. O erro médio quadrático é definido como o valor esperado da diferença quadrática entre a saída do receptor e a componente do sinal de mensagem na saída do receptor síncrono.
- 9.24 A Eq. (9.59) é o ruído FM pós-detecção de um filtro passa-baixa ideal. Obtenha o ruído pós-detecção para um sinal FM quando o filtro de pós-detecção é um filtro passa-baixa de segunda ordem com resposta em amplitude

$$|H(f)| = \frac{1}{(1 + (f/W)^4)^{1/2}}$$

Assuma: $|H_{BP}(f + f_c)|^2 \approx 1$ para $|f| < B_T/2$ e $B_T \gg 2W$.

- 9.25 Considere um sistema de comunicação com perda de transmissão de 100 dB e densidade espectral de ruído de 10^{-14} W/Hz na saída do receptor. Se a potência média de mensagem é $P = 2$ watt e a largura de faixa é 10 kHz, obtenha a potência média transmitida (em kW) necessária para uma RSR pós-detecção de 40 dB ou melhor quando a modulação é:
- (a) AM com $k_a = 1$. Repita os cálculos para $k_a = 0,1$.
 - (b) FM com $k_f = 10, 50$ e 100 kHz por volt.
- No caso do FM, verifique a limitação de limiar confirmando que a RSR de pré-detecção é maior do que 12 dB.

■ Experimentos de computador

- 9.26 Neste experimento, investigamos a performance da modulação em amplitude com ruído. O script MATLAB para experimento AM é fornecido no Apêndice 7. Ele simula a modulação por envelope por uma onda senoidal com índice de modulação de 0,3, adiciona ruído e, então, faz a detecção de envelope da mensagem. Usando este script:
- (a) Trace o envelope do sinal modulado
 - (b) Usando a função “spectra”, trace seu espetro.
 - (c) Trace o sinal detectado por envelope antes da filtragem passa-baixa.
 - (d) Compare a RSR pós-detecção com a teoria.
- 9.27 Neste segundo experimento de computador, investigamos a performance de FM com ruído. Usando o script MATLAB para o experimento FM fornecido no Apêndice 7:
- (a) Trace o espectro do sinal FM banda base.
 - (b) Trace o espectro do FM passa-faixa mais ruído.
 - (c) Trace o espectro do sinal detectado antes da filtragem passa-baixa.
 - (d) Trace o espectro do sinal detectado após a filtragem passa-baixa.
 - (e) Compare as RSRs pré-detecção e pós-detecção para o receptor FM.

Capítulo 10

RUÍDO EM COMUNICAÇÕES DIGITAIS

Nos últimos vinte anos, as comunicações digitais têm substituído as comunicações analógicas existentes em quase todas as instâncias. Novos serviços de comunicação normalmente consideram apenas a comunicação digital. Duas razões técnicas são a grande tolerância a erro fornecida pelas comunicações digitais e a capacidade de reprodução quase exata de seqüências digitais no receptor. Duas fortes razões externas para o aumento do domínio das comunicações digitais são o rápido crescimento de comunicações máquina a máquina, tal como a Internet e a espetacular evolução da eletrônica digital. Além disso, a habilidade de melhorar ainda mais a robustez de comunicações digitais através da utilização de códigos de correção de erro é uma importante vantagem sobre as comunicações analógicas. Como as comunicações digitais possuem uma tolerância ao erro maior do que as analógicas, as comunicações digitais são aplicadas em situações nas quais as comunicações analógicas nunca poderiam estar. Nessas situações difíceis, o ruído pode ainda ter um efeito significativo nas comunicações digitais; isso será investigado neste capítulo.

Falando abertamente, o objetivo da detecção é estabelecer a presença ou ausência de sinal contendo informação com ruído. Por esta razão, começamos nossa análise dos efeitos de ruído em comunicações digitais considerando a transmissão de um único pulso. As características de ruído em sistemas digitais são similares àquelas em sistemas analógicos. Especificamente, o sinal recebido pode ser modelado por

$$r(t) = s(t) + w(t) \quad (10.1)$$

na qual $s(t)$ é o sinal transmitido e $w(t)$ é o ruído aditivo. Como observado nos capítulos anteriores, existem várias similaridades entre técnicas de modulação analógicas e digitais. Iremos verificar que também existem similaridades entre estruturas de receptores analógicos e digitais.

O material apresentado neste capítulo nos ensina as seguintes lições.

- *Lição 1: A taxa de erro de bit é a medida primária da qualidade da performance de um sistema de comunicação digital, sendo, normalmente, uma função não linear da razão sinal/ruído.*
- *Lição 2: A análise da detecção de um único pulso permite a obtenção simples do princípio de filtragem casada. A filtragem casada pode ser aplicada à detecção ótima em vários esquemas de modulação digital linear.*
- *Lição 3: A performance da taxa de erro de bit da modulação por amplitude de pulso (PAM) melhora exponencialmente com a razão sinal/ruído com ruído branco Gaussiano.*
- *Lição 4: Receptores para esquemas de modulação linear passa-faixa binário e quaternário são diretamente desenvolvidos dos princípios de filtro casado e suas performances são similares à do PAM binário.*
- *Lição 5: A detecção não coerente de sinais digitais resulta em uma estrutura de receptor mais simples, mas ao custo de degradar a performance da taxa de erro de bit.*
- *Lição 6: O fornecimento de redundância no sinal transmitido através da adição de bits de verificação de paridade pode ser utilizada para a correção de erro à frente. A correção de erro à frente é um poderoso método para melhorar a performance de esquemas de modulação digital.*

10.1 TAXA DE ERRO DE BIT

Em sistemas digitais, a qualidade da saída da informação é a preocupação primária. Como a informação é digital e geralmente possui uma representação binária, esta qualidade é medida em termos da *taxa média de erro de bit* (TEB). Um erro de bit ocorre sempre que o bit transmitido e o bit correspondente recebido forem diferentes. Isto é um evento aleatório. Seja n o número de erros de bit observados em uma seqüência de bits de tamanho N . Então, a definição de freqüência relativa da TEB é

$$\text{TEB} = \lim_{N \rightarrow \infty} \left(\frac{n}{N} \right) \quad (10.2)$$

Em alguns sistemas digitais, outras medidas de qualidade, relacionadas com a taxa de erro de bit, são geralmente utilizadas. Por exemplo, vários sistemas digitais transferem a informação em *pacotes* e, independentemente se existe um ou centenas de erros no pacote, todo o pacote deve ser descartado. Nestes sistemas, a medida de qualidade é geralmente a *taxa de erro de pacote* (TEP), a qual pode ser diretamente relacionada com a TEB se os erros de bit forem estatisticamente independentes.

Naturalmente, a taxa de erro de bit necessária em um sistema digital depende da aplicação. Por exemplo:

- ▶ Para vocoders¹ de fala, uma TEB de 10^{-2} a 10^{-3} é geralmente considerada como suficiente.
- ▶ Para a transmissão de dados em canais sem fio, uma taxa de erro de bit de 10^{-5} a 10^{-6} é geralmente o objetivo.
- ▶ Para a transmissão de vídeo, uma TEB de 10^{-7} a 10^{-12} é geralmente o objetivo, dependendo da qualidade desejada e do método de codificação.
- ▶ Para dados financeiros, uma TEB de 10^{-11} , ou melhor, é geralmente necessária.

Dada a disponibilidade de diferentes estratégias de modulação/demodulação, precisamos comparar suas performances. Conceitualmente, gostaríamos de uma figura de mérito tal como a utilizada em sistemas de modulação/demodulação analógicos que pudéssemos associar a cada esquema de transmissão digital. Infelizmente, isto não é fácil em sistemas digitais, pois a qualidade geralmente não é uma função linear da razão sinal/ruído.

Entretanto, podemos definir uma equivalência à RSR de referência para sistemas digitais. Em particular, para sistemas digitais, a RSR de referência é a razão da *energia modular por bit de informação* pela *densidade espectral de ruído de uma faixa*, ou seja,

$$\begin{aligned} \text{RSR}_{\text{ref}}^{\text{digital}} &= \frac{\text{Energia modular por bit}}{\text{Densidade espectral de ruído}} \\ &= \frac{E_b}{N_0} \end{aligned}$$

Esta definição difere da definição de RSR de referência analógica em três pontos.

1. A definição analógica era a razão de potências. A definição digital é a razão de energias, pois as unidades de densidade espectral de ruído são watts/Hz, o que é equivalente à energia. Conseqüentemente, a definição digital é adimensional, tal como na definição analógica.

¹ Vocoders de fala analisam as características de curtos segmentos da fala, e parâmetros descrevendo estas características são transmitidos aos invés das amostras reais da fala. O tamanho destes segmentos de fala é normalmente 20 milissegundos. A vantagem dos vocoders frente às técnicas PCMA descritas no Capítulo 5 é a significativa redução da taxa de bit necessária para uma representação precisa. Existe uma grande variedade de algoritmos de vocoders, e mais detalhes podem ser obtidos em Gold e Morgan (1999).

2. A definição utiliza a densidade espectral de ruído de uma faixa, ou seja, ela assume que todo o ruído ocorre em freqüências positivas. Esta consideração é uma simples questão de conveniência.
3. A RSR de referência é independente da taxa de transmissão. Como ela é uma razão de energias, ela foi essencialmente normalizada pela taxa de bits.

Para comparar as estratégias de modulação/demodulação digital, o objetivo é determinar a performance da taxa de erro de bit como uma função da RSR de referência, representada por E_b/N_0 . Este *modelo de referência digital* fornece um quadro de referência para a comparação justa de diferentes esquemas.

EXEMPLO 10.1 Calculando a taxa de erro de pacote

Um sistema de transmissão é projetado para transferir dados com uma TEB de 10^{-5} , ou melhor. Se este sistema for utilizado para transmitir pacotes de 1000 bits, qual é a taxa de erro de pacote esperada? Assuma que os erros de bit são estatisticamente independentes.

Com a consideração de erros de bit independentes, os bits em um pacote formam o equivalente a uma seqüência de Bernoulli. Conseqüentemente, o número de erros de bit possui uma distribuição binomial como descrita no Exemplo 8.5. A probabilidade de um erro de pacote é simplesmente um menos a probabilidade do pacote não ter erros. Ou seja, eles são eventos complementares. Conseqüentemente, se n é o número de erros de bit, a taxa de erro de pacote é dada por

$$\begin{aligned} \text{TEP} &= 1 - P_{\text{binomial}}(n = 0) \\ &= 1 - \binom{1000}{0} (1 - 10^{-5})^{1000} (10^{-5})^0 \\ &= 9,95 \times 10^{-3} \end{aligned}$$

► **PROBLEMA 10.1** Seja H_0 o evento de um 0 ser transmitido e R_0 o evento de um 0 ser recebido. Defina H_1 e R_1 de forma similar para um 1. Expresse a TEB em termos da probabilidade destes dois eventos quando:

- (a) A probabilidade de um erro 1 é a mesma probabilidade de um erro 0.
- (b) A probabilidade de um 1 ser transmitido não é a mesma probabilidade de um 0 ser transmitido.

10.2 DETECÇÃO DE UM ÚNICO PULSO COM RUÍDO

Com esquemas analógicos, vimos que é desejável que as larguras de faixa dos filtros sejam as menores possíveis para minimizar o ruído, mas não pequenas demais para distorcerem o sinal desejado. Em um sentido amplo, podemos dizer que os filtros são casados com o sinal no contexto do domínio da freqüência. Com esquemas digitais existe uma definição ainda mais precisa de filtragem casada, como iremos ver a seguir.

Começamos com a situação mostrada na Fig. 10.1 para a detecção de um único pulso transmitido em banda base. O sinal recebido é primeiro processado por um detector linear. A saída do detector é amostrada e comparada com um limiar. Este comparador de limiar é utilizado para determinar qual das seguintes situações ocorreu:

1. O sinal $r(t)$ recebido consiste somente de ruído branco Gaussiano $w(t)$.
2. O sinal $r(t)$ recebido consiste de $w(t)$ mais um sinal $s(t)$ de forma conhecida.

O ruído é considerado como tendo média nula e densidade espectral $N_0/2$. No receptor, queremos determinar qual destas duas situações é verdadeira processando o sinal $r(t)$ de tal forma que, se o sinal $s(t)$ estiver presente, a saída do receptor em algum tempo arbitrário $t = T$ será consideravelmente maior do que se $s(t)$ estiver ausente.

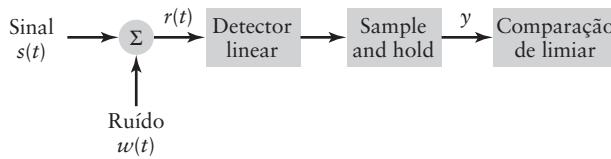


Figura 10.1 Processamento de um único pulso.

Um exemplo prático da situação descrita anteriormente é a modulação por amplitude de pulso (PAM) binária utilizando sinalização liga-desliga. Um pulso $s(t)$ pode representar o símbolo 1, enquanto que a ausência do pulso pode representar o símbolo 0. O objetivo é maximizar a saída do receptor quando o pulso estiver presente e minimizar a saída quando apenas o ruído estiver presente.

Para o esquema de transmissão de pulso único descrito, nosso objetivo atual é *maximizar a razão sinal/ruído* na saída do receptor. Matematicamente, existem duas formas possíveis para o sinal recebido:

$$r(t) = \begin{cases} s(t) + w(t), & \text{pulso presente} \\ w(t), & \text{pulso ausente} \end{cases} \quad (10.3)$$

Para este desenvolvimento, assumimos que $s(t)$ é um pulso banda base, sendo não nulo somente no intervalo $0 \leq t \leq T$. Nossa estratégia de detecção é filtrar o sinal recebido e amostrar a saída do filtro no tempo T . O filtro, descrito pela resposta ao impulso $g(t)$, é considerado como sendo linear e invariante no tempo. Logo, a variável aleatória para a determinação se o pulso está presente é definida pela forma da integral de convolução:

$$Y = \int_0^T g(T-t)r(t) dt \quad (10.4)$$

O objetivo é determinar o filtro $g(t)$ que maximiza a razão sinal/ruído da saída Y . Para analisar a Eq. (10.4), utilizamos a primeira linha da Eq. (10.3) para expandi-la:

$$Y = \int_0^T g(T-t)s(t) dt + \int_0^T g(T-t)w(t) dt \quad (10.5)$$

A primeira integral do lado direito da Eq. (10.5) é o termo de sinal, o qual será zero se o pulso estiver ausente, e a segunda integral é o termo de ruído o qual sempre estará lá. Começamos analisando o tempo de ruído. Seja a contribuição do ruído representada por

$$N = \int_0^T g(T-t)w(t) dt \quad (10.6)$$

Esta amostragem da saída do filtro converte o processo aleatório $w(t)$ em uma variável aleatória N . Se o processo de ruído $w(t)$ possui unidades de volts, então a Eq. (10.6) pode ser pensada como sendo o somatório ponderado do ruído no tempo e, neste caso, N possui unidades de volts-segundos. O valor esperado de N , definido por

$$\begin{aligned} E[N] &= \int_0^T g(T-t)E[w(t)] dt \\ &= 0 \end{aligned} \quad (10.7)$$

é, portanto, zero. Este resultado é obtido da troca da esperança e integração discutida na Seção 8.9 e do fato do ruído branco Gaussiano $w(t)$ ser considerado como tendo média nula. A variância do ruído de saída é dada por

$$\begin{aligned}
\text{Var}(N) &= \mathbf{E}[N^2] \\
&= \mathbf{E}\left[\int_0^T g(T-t)w(t) dt \int_0^T g(T-\tau)w(\tau) d\tau\right] \\
&= \int_0^T \int_0^T g(T-t)g(T-\tau)\mathbf{E}[w(t)w(\tau)] dt d\tau
\end{aligned} \tag{10.8}$$

na qual, novamente, alteramos a ordem da integração e da esperança. Podemos fazer isto porque as duas operações são lineares. Assumimos que o processo de ruído é branco, tal que $\mathbf{E}[w(t)w(\tau)] = N_0/2 \delta(t-\tau)$, na qual $\delta(t)$ é a função delta de Dirac. Podemos, então, reescrever a Eq. (10.8) como

$$\begin{aligned}
\mathbf{E}[N^2] &= \int_0^T \int_0^T g(T-t)g(T-\tau) \frac{N_0}{2} \delta(t-\tau) dt d\tau \\
&= \frac{N_0}{2} \int_0^T |g(T-\tau)|^2 d\tau \\
&= \frac{N_0 T}{2}
\end{aligned} \tag{10.9}$$

A segunda linha da Eq. (10.9) é obtida da primeira pela propriedade de deslocamento da função delta. A terceira linha é obtida com a consideração de que o filtro $g(t)$ foi normalizado, ou seja,

$$\int_0^T |g(t)|^2 dt = T \tag{10.10}$$

A unidade da densidade espectral de potência é watts por hertz. Portanto, $N_0 T$ da Eq. (10.9) possui unidade de watt-(segundo)². Se assumirmos que a medida é executada em um resistor de 1 ohm, então isto é equivalente a (volt-segundos)². Esta observação mostra que a unidade de $\mathbf{E}[N^2]$ é consistente com a de N . Para resumir, a amostra de ruído na saída do receptor linear possui

- ▶ Média nula
- ▶ Variância de $N_0 T/2$
- ▶ Distribuição Gaussiana, pois um processo Gaussiano filtrado também é Gaussiano (veja a Seção 8.9)

Deve ser observado que, apesar da restrição de normalização da Eq. (10.10), esta análise da componente de ruído não coloca nenhuma restrição adicional à escolha da resposta $g(t)$ ao impulso do filtro.

A seguir consideraremos a componente de sinal da Eq. (10.5), dada por

$$S = \int_0^T g(T-\tau)s(\tau) d\tau \tag{10.11}$$

Para maximizar a razão sinal/ruído, escolhemos $g(t)$ para maximizar a Eq. (10.11) sujeita a restrição da Eq. (10.10). Para resolver este problema de maximização, utilizamos a desigualdade de Schwartz para integrais (veja o Apêndice 5). A *desigualdade de Schwartz* para integrais é definida por

$$\left| \int_{-\infty}^{\infty} g(T-t)s(t) dt \right|^2 \leq \int_{-\infty}^{\infty} |g(T-t)|^2 dt \int_{-\infty}^{\infty} |s(t)|^2 dt \tag{10.12}$$

a qual é validada com a igualdade se

$$g(T - t) = cs(t) \quad (10.13)$$

para algum escalar c . Conseqüentemente, o sinal da Eq. (10.11) e a razão sinal/ruído são maximizadas se a Eq. (10.13) for satisfeita e o escalar c for escolhido tal que a restrição da Eq. (10.10) seja satisfeita.

A Eq. (10.13) implica que o filtro de recepção seja casado com a forma do pulso transmitido. Este é o *princípio da detecção de filtro casado* que foi referenciada nas seções anteriores. Ou seja, na transmissão de um único pulso, o *processamento do sinal recebido por um filtro casado para o sinal transmitido maximiza a razão sinal/ruído*.

Se substituirmos a Eq. (10.13) no detector da Eq. (10.4), obtemos

$$Y = c \int_0^T s(t)r(t) dt$$

No caso geral, no qual a temporização do pulso recebido não é exatamente conhecida, podemos calcular várias saídas em diferentes tempos τ deslocados, como mostrado por:

$$Y(\tau) = c \int_{-\infty}^{\infty} s(t)r(t - \tau) dt$$

Lembre-se do Capítulo 8 que esta expressão é equivalente à correlação cruzada de dois sinais ergódicos $s(t)$ e $r(t)$. Conseqüentemente, a estrutura do receptor da Eq. (10.4) com $g(T - t) = cs(t)$ também é chamada de *receptor de correlação*. Com $r(t)$ definido pela Eq. (10.1), a componente de sinal desta correlação é maximizada em $\tau = 0$. Isto enfatiza a importância da sincronização quando executamos a detecção ótima.

- **PROBLEMA 10.2** Suponha que, na Eq. (10.4), $r(t)$ representa um sinal banda base complexo ao invés de um sinal real. Qual seria a escolha ideal para $g(t)$ neste caso? Justifique sua resposta. ◀

10.3 DETECÇÃO ÓTIMA DE PAM BINÁRIO COM RUÍDO

A transmissão de pulso único, apesar de conveniente para análise, possui potencial limitado de comunicação. Estendemos, agora, os resultados da seção anterior para o caso no qual pulsos são enviados em intervalos consecutivos. Em particular, considere a transmissão PAM binária com sinalização liga-desliga como mostrado na Fig. 10.2. Esta forma de sinalização pode ser representada por

$$s(t) = A \sum_{k=0}^{\infty} b_k h(t - kT) \quad (10.14)$$

na qual b_k é zero se o k -ésimo bit for 0, b_k é um se o k -ésimo bit for 1 e $h(t)$ é um pulso retangular de duração T mas centrado em $t = T/2$. Com a forma do pulso transmitido $h(t) = \text{ret}[t - T/2]/T]$, o filtro casado para esta forma de pulso é $g(T - t) = h(t)$. (Veja o Problema 10.3).

A vantagem dos pulsos retangulares é que o k -ésimo pulso é contido dentro do intervalo $(k - 1)T < t < kT$, e não interfere com os pulsos nos intervalos adjacentes. Conseqüentemente, o filtro casado para o pulso no intervalo $(k - 1)T < t < kT$ é o filtro casado para um único pulso deslocado no tempo para este intervalo. Ou seja, no k -ésimo intervalo de

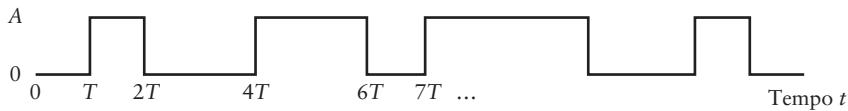


Figura 10.2 Exemplo de transmissão de PAM binário com sinalização liga-desliga. T representa o intervalo de símbolo.

símbolo, o filtro casado é $g_k(T-t) = h(t-kT)$. A saída do filtro casado de recepção no final do k -ésimo intervalo de símbolo é

$$\begin{aligned} Y_k &= \int_{-\infty}^{kT} g_k(T-t)r(t) dt \\ &= \int_{(k-1)T}^{kT} h(t-kT)r(t) dt \end{aligned} \quad (10.15)$$

na qual $r(t) = s(t) + w(t)$ com $s(t)$ dado pela Eq. (10.14). A segunda linha da Eq. (10.15) é obtida do fato de $h(t-kT)$ ser apenas não nulo no intervalo $(k-1)T < t < kT$.

Como o filtro casado $h(t) = 1$ sempre que ele for não nulo, temos

$$\begin{aligned} Y_k &= \int_{(k-1)T}^{kT} \text{ret}\left(\frac{t-(k+1/2)T}{T}\right)r(t) dt \\ &= \int_{(k-1)T}^{kT} r(t) dt \end{aligned} \quad (10.16)$$

A Eq. (10.16) é a descrição matemática do dispositivo físico chamado de *detector integra-e-descarta*. Este dispositivo simplesmente integra o sinal recebido durante o intervalo de símbolo, amostra a saída, e então começa novamente para o próximo intervalo, como mostrado na Fig. 10.3. Este simples detector é ótimo para a sinalização retangular, como segue da Eq. (10.16).

- **PROBLEMA 10.3** Se $g(t) = c \text{ ret}\left[\frac{\alpha(t-T/2)}{T}\right]$, determine c tal que $g(t)$ satisfaça a Eq. (10.10) com $\alpha > 1$. ◀

■ Performance TEB

Apesar de o filtro casado ser ótimo em termos de maximizar a razão sinal/ruído, queremos estimar a performance fornecida por tal esquema. Como descrito na Seção 10.2, a figura de mérito de sistemas digitais é freqüentemente a taxa de erro de bit (TEB), ou seja, a fração média dos bits recebidos que foram erroneamente detectados.

Com o esquema de transmissão liga-desliga descrito anteriormente, o receptor deve fazer a decisão entre duas hipóteses:

$$\begin{aligned} H_0: \quad b_k &= 0 \text{ foi transmitido} \\ H_1: \quad b_k &= 1 \text{ foi transmitido} \end{aligned} \quad (10.17)$$

baseado no sinal recebido Y_k durante o k -ésimo intervalo de símbolo. Como mostramos que, com a forma do pulso retangular, a detecção é independente de um intervalo de símbolo para o próximo, podemos descartar o subscrito k .

Um critério razoável para a escolha entre as duas hipóteses é escolher a *hipótese mais provável com base na observação*. Ou seja, comparamos as duas probabilidades condicionais (na qual y é o valor da variável aleatória Y):

- $P(H_0|y)$ é a probabilidade de um 0 ter sido transmitido se y for recebido.
- $P(H_1|y)$ é a probabilidade de um 1 ter sido transmitido se y for recebido.

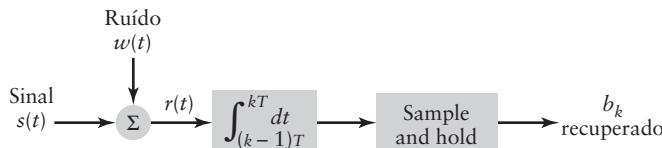


Figura 10.3 Detector integra-e-descarta.

A maior destas duas probabilidades fornece a decisão para o bit em análise. Esta situação é bem diferente de comunicações analógicas, na qual tentamos minimizar a distorção do sinal recebido devido ao ruído. Em comunicações digitais, se decisões confiáveis são possíveis, então a mensagem transmitida é recuperada sem distorção.

Praticamente, queremos uma simples regra para decidir entre duas hipóteses, H_0 e H_1 . Um exemplo de uma regra de decisão é escolher H_0 , ou seja, escolher 0 se y for menor do que algum limiar γ e escolher 1 caso contrário. Intuitivamente, se b_k é 0 ou 1, então podemos ajustar o limiar γ em $\frac{1}{2}$ e comparar y com este limiar. Se y for maior do que γ , então a decisão é H_1 , caso contrário é H_0 . Esta regra de decisão intuitiva é ótima em várias situações, mas vamos mostrar o por quê.

Considere a probabilidade de fazer um erro com esta regra de decisão baseada em probabilidades condicionais. Se um 1 é transmitido, a probabilidade de erro é

$$P[0 \text{ decidido} | H_1] = P[Y < \gamma | H_1] \quad (10.18)$$

na qual Y é a variável aleatória associada com a observação y . Referimos a este erro como um *erro Tipo I*. Para calcular esta probabilidade, lembre-se da discussão da Seção 10.2 que a variável aleatória na saída de um filtro casado com a qual baseamos nossa decisão, ou seja,

$$Y = S + N \quad (10.19)$$

possui duas componentes. Das Eqs. (10.7) e (10.9), N é uma variável aleatória Gaussiana de média nula com variância $\sigma^2 = N_0 T/2$. Da Eq. (10.14), a componente de sinal de Y possui uma parte determinística, a qual é a forma do pulso, e uma componente aleatória, a qual é o bit de modulação, b_k . Entretanto, para um intervalo de símbolo em particular, o bit de modulação é fixo e S é obtido pela substituição da Eq. (10.14) na Eq. (10.11), especificamente, para o nominal $k = 0$, temos

$$\begin{aligned} S &= Ab \int_0^T h(t)h(t) dt \\ &= ATb \end{aligned} \quad (10.20)$$

na qual a segunda linha é obtida da propriedade normalizada de filtro casado (veja a Eq. (10.10)) e $b = 0$ ou 1 dependendo se um 0 ou 1 estiver sendo transmitido. Note que S possui unidades de volts-segundos, consistente com os desenvolvimentos anteriores.

Da Eq. (10.20), vemos que S possui média $\mu = AT$ quando $b = 1$ estiver sendo transmitido. Correspondentemente, Y possui distribuição Gaussiana com média $\mu = AT$ e sua função de densidade é dada por

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{-(y - \mu)/2\sigma^2\} \quad (10.21)$$

a qual é mostrada na Fig. 10.4.

A probabilidade de um erro Tipo I, como especificado pela Eq. (10.18) é a probabilidade da saída Y estar na área sombreada, abaixo de γ , na Fig. 10.4. Matematicamente, esta probabilidade é a integral da área sombreada da função de densidade Gaussiana

$$\begin{aligned} P[Y < \gamma | H_1] &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\gamma} \exp(-(y - \mu)^2/2\sigma^2) dy \\ &= Q\left(\frac{\mu - \gamma}{\sigma}\right) \end{aligned} \quad (10.22)$$

na qual a função Q foi definida na Seção 8.4.

A Eq. (10.22) não é a única contribuição aos erros de bit. Um erro também pode ocorrer se um 0 for transmitido e um 1 for detectado. Nos referimos a este erro como um *erro Tipo II*. Pode ser mostrado que a probabilidade de um erro Tipo II é (veja o Problema 10.4)

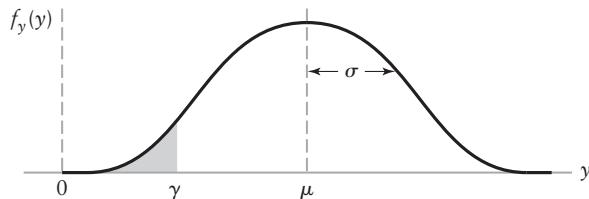


Figura 10.4 Função de densidade Gaussiana para a detecção de um 1 com sinalização liga-desliga.

$$P[Y > \gamma | H_0] = Q\left(\frac{\gamma - \mu}{\sigma}\right) \quad (10.23)$$

As regiões de probabilidade associadas com os erros Tipo I e Tipo II são ilustradas na Fig. 10.5. A probabilidade combinada de erro é dada pela regra de Baye (veja a Seção 8.1)

$$P_e = P[Y < \gamma | H_1] \cdot P[H_1] + P[Y > \gamma | H_0] \cdot P[H_0] \quad (10.24)$$

na qual $P[H_i]$ é a probabilidade *a priori* de um 0 ou 1 ser transmitido. Normalmente, os bits transmitidos são equiprováveis, o que significa que

$$P[H_1] = P[H_0] = \frac{1}{2} \quad (10.25)$$

Conseqüentemente, a probabilidade média de erro é dada por

$$P_e = \frac{1}{2}Q\left(\frac{\mu - \gamma}{\sigma}\right) + \frac{1}{2}Q\left(\frac{\gamma}{\sigma}\right) \quad (10.26)$$

Geralmente, queremos que as probabilidades dos dois tipos de erro, $P[\text{erro Tipo I}]$ e $P[\text{erro Tipo II}]$ sejam iguais, de forma a escolher $\gamma = \mu/2$. Pode ser mostrado que esta escolha de γ minimiza a probabilidade de erro descrita pela Eq. (10.26) (veja o Problema 10.22). Conseqüentemente, temos que a probabilidade média de erro é dada por

$$P_e = Q\left(\frac{\mu}{2\sigma}\right) \quad (10.27)$$

Ou seja, a probabilidade de um erro de bit depende da razão da média do sinal μ quando 1 é transmitido pelo desvio padrão σ do ruído através da função não linear Q .

Nosso próximo passo é expressar esta probabilidade de erro de bit em termos do modelo de referência digital. Este cálculo pode ser dividido em duas componentes:

- Para expressar a variância em termos da densidade espectral de ruído, temos, da Eq. (10.9), que $\sigma^2 = N_0 T/2$.

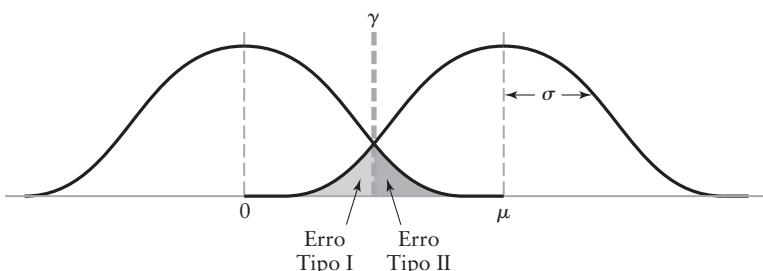


Figura 10.5 Ilustração dos erros de decisão para 0 e 1 com sinalização liga-desliga.

- Para expressar a amplitude A do sinal em termos da energia por bit E_b , assumimos que 0 e 1 possuem a mesma probabilidade de serem transmitidos. Então, a energia média por bit na entrada do receptor é

$$\begin{aligned} E_b &= \mathbb{E}\left[\int_0^T |s(t)|^2 dt\right] \\ &= A^2 \mathbb{E}[b^2] \int_0^T |h(t)|^2 dt \\ &= A^2 \left\{ \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 \right\} T \\ &= \frac{A^2 T}{2} \end{aligned} \quad (10.28)$$

na qual, na segunda linha, separamos as partes determinística e aleatória do sinal e a terceira linha foi obtida do fato de $h(t) = \text{rect}[(t - T/2)/T]$. Conseqüentemente, temos que $A = \sqrt{2E_b/T}$ para a sinalização PAM liga-desliga.

Assumindo estes valores para σ e $\mu = AT$ na Eq. (10.27), o resultado obtido é

$$P_e^{\text{liga-desliga}} = Q\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (10.29)$$

É interessante notar que se ao invés dos níveis 0 e A da sinalização liga-desliga utilizarmos os níveis bipolares de $-A/2$ e $+A/2$, então a análise acima não seria alterada exceto pelo fato do limiar, agora, ser $\gamma = 0$. Entretanto, com o esquema bipolar, a energia por bit é $E_b = A^2 T/4$ e a taxa de erro de bit em termos do modelo de referência digital é

$$P_e^{\text{bipolar}} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \quad (10.30)$$

Como a função Q é uma função monotonicamente decrescente de seu argumento, a Eq. (10.30) implica em que uma razão E_b/N_0 menor irá produzir a mesma taxa de erro do esquema liga-desliga caracterizado pela Eq. (10.29). Portanto, vemos que sinalização bipolar é um método mais eficiente em potência para a obtenção da mesma performance de taxa de erro de bit do que o esquema de sinalização liga-desliga.

- **PROBLEMA 10.4** Mostre que, com a sinalização liga-desliga, a probabilidade do erro Tipo II da Eq. (10.23) é dada por

$$P[Y > \gamma | H_0] = Q\left(\frac{\gamma}{\sigma}\right)$$



■ Formas de pulso não retangulares

Vimos no Capítulo 6 que devido a limitações espectrais e efeitos do canal, o sinal recebido pode não ter a forma de um pulso retangular. Se combinarmos esta observação com o ruído, o sinal recebido pode ser representado por

$$r(t) = A \sum_{k=0}^{\infty} b_k p(t - kT) + w(t) \quad (10.31)$$

na qual $p(t)$ é a forma do pulso. Como esta forma de pulso não é necessariamente confinada ao intervalo de duração T , pode-se esperar que a interferência intersimbólica seja uma questão importante no esquema de transmissão contínua.

Em particular, consideramos o caso quando $p(t)$ é um *pulso de raiz de cosseno levantado normalizado*, descrito na Seção 6.4. Para um único pulso tendo a forma espectral da raiz do cosseno levantado, a teoria de filtro casado da Seção 10.2 ainda se aplica. Além disso, esta forma de pulso possui a seguinte propriedade de ortogonalidade quando $E = 1$ (veja os Problemas 10.5 e 10.12):

$$\int_{-\infty}^{\infty} p(kT - t)p(t - lT) dt = \delta(k - l) \quad (10.32)$$

na qual $\delta(t)$ é a função delta de Dirac. Aplicando o filtro casado para o k -ésimo símbolo de $r(t)$ da Eq. (10.31), temos

$$\begin{aligned} Y_k &= \int_{-\infty}^{\infty} p(kT - t)r(t) dt \\ &= Ab_k \int_{-\infty}^{\infty} p(kT - t)p(t - kT) dt + N_k + \sum_{l \neq k} Ab_l \int_{-\infty}^{\infty} p(kT - t)p(t - lT) dt \end{aligned} \quad (10.33)$$

na qual o k -ésimo termo de ruído é

$$N_k = \int_{-\infty}^{\infty} p(kT - t)w(t) dt \quad (10.34)$$

Os dois primeiros termos da Eq. (10.33) representam o termo sinal mais ruído usual associado com a detecção de um único pulso. O terceiro termo representa a interferência devido aos símbolos adjacentes. Entretanto, se substituirmos a Eq. (10.32) na Eq. (10.33), vemos que a saída do detector se reduz para a da detecção de um único pulso. Ou seja, não existe interferência intersimbólica com a escolha adequada do formato do pulso. Com estas condições, $Y_k = A b_k + N_k$, e a performance TEB é a mesma com o formato de pulso retangular, o que é um resultado maravilhoso. A implementação de tal esquema é mostrada na Fig. 10.6, a qual envolve a utilização de um formador de pulso de raiz de cosseno levando.

► **PROBLEMA 10.5** Prove a propriedade da forma do pulso de raiz de cosseno levantado $p(t)$ dada pela Eq. (10.32), utilizando os seguintes passos:

- Se $R(f)$ é a representação da transformada de Fourier de $p(t)$, qual é a representação da transformada de Fourier de $p(t - lT)$?
- Qual é a transformada de Fourier de $q(\tau) = \int_{-\infty}^{\infty} p(\tau - t)p(t - lT) dt$? Qual é a sua forma espectral?
- O que é $q(\tau)$? O que é $q(kT)$?

Utilize estes resultados para mostrar que a Eq. (10.32) é válida. ◀

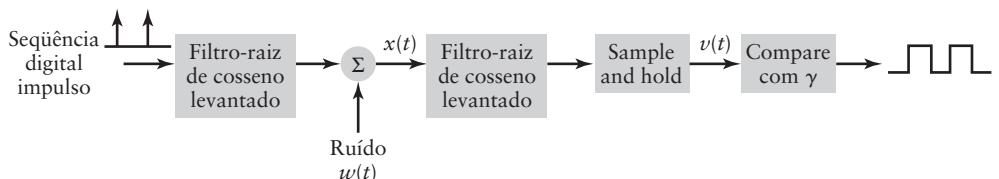


Figura 10.6 Implementação do PAM com filtragem de raiz de cosseno levantado. (A forma de um pulso de raiz de cosseno levantado foi discutida na Seção 6.4.)

10.4 DETECÇÃO ÓTIMA DE BPSK

Como descrito no Capítulo 7, uma das formas mais simples de comunicação passa-faixa digital é o chaveamento binário de fase. Com o BPSK, o sinal transmitido é

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t), & 0 \leq t \leq T \text{ se um } 1 \text{ é enviado} \\ A_c \cos(2\pi f_c t + \pi), & 0 \leq t \leq T \text{ se um } 0 \text{ é enviado} \end{cases}$$

Ou seja, fases opostas da mesma portadora são transmitidas para representar os símbolos binários. A detecção do BPSK é simplificada observando que como $\cos(2\pi f_c t + \pi) = -\cos(2\pi f_c t)$, o sinal transmitido pode ser igualmente representado por

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t), & 0 \leq t \leq T \text{ se um } 1 \text{ é enviado} \\ -A_c \cos(2\pi f_c t), & 0 \leq t \leq T \text{ se um } 0 \text{ é enviado} \end{cases} \quad (10.35)$$

Logo, em geral, podemos escrever

$$s(t) = A_c m(t) \cos(2\pi f_c t) \quad (10.36)$$

na qual, para a transmissão de um único pulso, $m(t) = +1$ para um 1 e $m(t) = -1$ para um 0, com $0 \leq t \leq T$. No caso mais geral, quando transmitimos múltiplos bits, temos

$$m(t) = \sum_{k=0}^N b_k h(t - kT) \quad (10.37)$$

na qual $h(t)$ é o pulso retangular $\text{ret}\left(\frac{t - T/2}{T}\right)$. Portanto, o BPSK possui uma forma muito similar à modulação por faixa lateral dupla-portadora suprimida (DSB-SC) utilizada para comunicações analógicas, um ponto que já foi apresentado no Capítulo 3. Conseqüentemente, para recuperar a mensagem, podemos utilizar uma estrutura de receptor análoga à estrutura do receptor coerente utilizado para DSB-SC. Uma estrutura típica de receptor para o BPSK é mostrada na Fig. 10.7, na qual o sinal RF mais o ruído branco Gaussiano, $s(t) + w(t)$, é transladado em freqüência para uma FI na qual ele sofre uma filtragem passa-faixa. O sinal passa-faixa, $x(t)$, é a entrada do detector coerente BPSK.

■ Detecção de BPSK com ruído

Seguindo o desenvolvimento para a demodulação DSB-SC da Seção 9.4, o sinal mais o ruído passa-faixa na entrada do detector coerente BPSK da Fig. 10.7 pode ser representado por

$$x(t) = s(t) + n_I(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \quad (10.38)$$

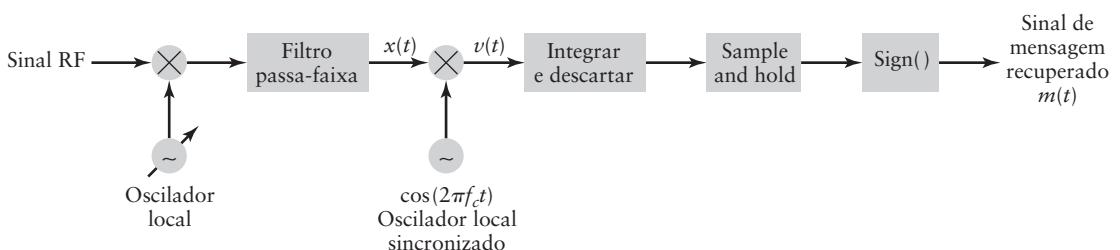


Figura 10.7 Estrutura típica do receptor BPSK.

na qual $n_I(t)$ e $n_Q(t)$ são as componentes em fase e quadratura do ruído passa-faixa $n(t)$, respectivamente. A saída do modulador de produto da Fig. 10.7 é dada por

$$\begin{aligned} v(t) &= x(t) \cos(2\pi f_c t) \\ &= \frac{1}{2} [A_c m(t) + n_I(t)] + \left[\frac{1}{2} (A_c m(t) + n_I(t)) \cos(4\pi f_c t) - \frac{1}{2} n_Q(t) \sin(4\pi f_c t) \right] \end{aligned} \quad (10.39)$$

na qual utilizamos as fórmulas: $\cos A \cos B = \frac{1}{2}[\cos(A-B) + \cos(A+B)]$ e $\sin A \sin B = \frac{1}{2}[\sin(A-B) + \sin(A+B)]$. Vamos considerar as duas partes da segunda linha da Eq. (10.39). A primeira parte representa o sinal banda base mais uma componente banda base do ruído, enquanto que a segunda parte representa o sinal e o ruído centrados em uma freqüência mais alta de $2f_c$.

Com o sinal DSB-SC analógico, utilizamos um filtro passa-baixa para remover as componentes de ruído de alta freqüência e recuperarmos o sinal desejado. Com sinais digitais, sabemos que o filtro casado é o método ótimo para recuperação dos dados. Podemos combinar o filtro casado com um filtro passa-baixa para obter os resultados desejados, desde que o filtro não distorça o sinal desejado. Entretanto, se percebermos que o filtro casado

$$\begin{aligned} g(T-t) &= h(t) \\ &= \text{ret}\left(\frac{t-T/2}{T}\right) \end{aligned}$$

já é um filtro passa-baixa, então não há necessidade para o segundo filtro.

O leitor atento irá notar que, na obtenção do filtro casado, assumimos que o ruído é branco. A componente em fase $n_I(t)$ é ruído de faixa estreita e, portanto, não branco. Entretanto, pode ser mostrado que enquanto o filtro passa-faixa possuir uma largura de faixa maior do que o sinal e a densidade espectral de ruído for aproximadamente plana na largura de faixa do sinal, o princípio de filtro casado continua válido.

No caso em que $m(t)$ é constituído por pulsos retangulares, o detector ótimo após a conversão para baixo é o filtro integra-e-descarta, como mostrado no caso de modulação por amplitude de pulso (PAM). Neste caso, a saída do detector integra-e-descarta é, portanto,

$$\begin{aligned} Y_k &\approx \frac{1}{2} \int_{(k-1)T}^{kT} [A_c m(t) + n_I(t)] dt \\ &= \frac{A_c T}{2} b_k + N_k \end{aligned} \quad (10.40)$$

A Eq. (10.40) assume que os termos de alta freqüência possuem impacto negligenciável na saída do detector. O termo de ruído na Eq. (10.40) é dado por

$$N_k \approx \frac{1}{2} \int_{(k-1)T}^{kT} n_I(t) dt \quad (10.41)$$

Lembre-se da Seção 8.11 que se a densidade espectral de potência do ruído de faixa estreita $n(t)$ for $N_0/2$, então a densidade espectral de potência da componente em fase $n_I(t)$ do ruído é N_0 , em sua respectiva faixa de freqüência. Desta forma, podemos escrever

$$\begin{aligned} E[N_k^2] &= \frac{1}{4} E \left[\int_{(k-1)T}^{kT} \int_{(k-1)T}^{kT} n_I(t) n_I(s) dt ds \right] \\ &= \frac{1}{4} \left[\int_{(k-1)T}^{kT} \int_{(k-1)T}^{kT} N_0 \delta(t-s) dt ds \right] \\ &= \frac{1}{4} \int_{(k-1)T}^{kT} N_0 ds \\ &= \frac{N_0 T}{4} \end{aligned}$$

Portanto, a variância σ^2 de N_k da Eq. (10.41) é $\frac{1}{4}N_0T$.

Na Fig. 7.4(b), mostramos um demodulador BPSK similar ao mostrado na Fig. 10.7. A principal diferença entre os dois demoduladores é que o filtro passa-baixa do primeiro é substituído pelo integrador no segundo. No caso sem ruído do Capítulo 7, o propósito do filtro passa-baixa é remover as componentes de alta freqüência, enquanto mantém o sinal banda base desejado. O integrador da Fig 10.7 também atua como um filtro passa-baixa removendo as componentes de alta freqüência, mas, devido a sua propriedade casada, ele adicionalmente remove a maior quantidade possível de ruído do sinal e, portanto, maximiza a razão sinal/ruído na saída do demodulador. Note que em comunicações digitais, o objetivo é recuperar a informação, 0s e 1s, da forma mais confiável possível. Ao contrário de comunicações analógicas, não existe a necessidade da forma de onda transmitida ser recuperada com um mínimo de distorção.

■ Análise de performance

A análise de taxa de erro de bit com BPSK é similar a análise de sinalização bipolar da Seção 10.3. O limiar para a decisão entre 0 e 1 na saída do filtro casado é ajustado para zero. Um limiar de zero possui a vantagem prática de não precisar ser calibrado se o caminho de transmissão possuir um ganho desconhecido.

Na Eq. (10.40), o valor médio do sinal é $A_cT/2$ ou $-A_cT/2$, dependendo se b_k é +1 ou -1, respectivamente. Se assumirmos que um 1 foi transmitido e fizermos $\mu = A_cT/2$, então a probabilidade de erro é

$$\begin{aligned} P_e &= P[Y < 0] \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^0 \exp\left\{-\frac{(y - \mu)^2}{2\sigma^2}\right\} dy \end{aligned} \quad (10.42)$$

Conseqüentemente, análogo à Eq. (10.22) com $\gamma = 0$, a taxa de erro de bit é dada por

$$P_e = Q\left(\frac{\mu}{\sigma}\right) \quad (10.43)$$

Por simetria, este resultado é válido se um +1 ou -1 for transmitido. Em termos de modelo de referência digital, observamos que a energia por bit com BPSK, assumindo que a energia de pulsos em $m(t)$ é normalizada, é simplesmente $E_b = A_c^2T/2$. Usando este resultado na definição de μ , com $b_k = +1$, temos

$$\mu = \frac{A_cT}{2} = \sqrt{\frac{2E_b}{T}} \frac{T}{2} = \sqrt{\frac{E_bT}{2}}$$

Substituindo o valor de μ obtido e $\sigma = \frac{1}{2}\sqrt{N_0T}$ na Eq. (10.43), obtemos o resultado

$$P_e^{\text{BPSK}} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \quad (10.44)$$

Esta performance de TEB é exatamente a mesma do esquema bipolar banda base, dada pela Eq. (10.30). Ou seja, não existe diferença na performance de TEB entre a sinalização PAM bipolar passa-baixa e a modulação BPSK passa-faixa. Portanto, tal como no caso analógico, não existe alteração na figura de mérito quando mudamos de um sistema banda base para um sistema equivalente passa-faixa.

A análise do BPSK pode ser estendida para o formato de pulso não retangular, de forma similar a apresentada para banda base. Para o formato de pulso não retangular, representamos o sinal transmitido por

$$s(t) = A_c \sum_{k=-\infty}^{\infty} b_k p(t - kT) \cos(2\pi f_ct) \quad (10.45)$$

na qual $p(t)$ é o formato do pulso e $b_k = \pm 1$ representa o dado. Esta combinação de formato de pulso de modulação BPSK forma um importante método para a comunicação de informação binária e controle da largura de faixa necessária para a transmissão. Em particular, a modulação BPSK permite que múltiplos usuários utilizem a mesma mídia de transmissão pela escolha das freqüências de portadora e o formato do pulso limita a largura de faixa do sinal transmitido e, portanto, reduz ou elimina a interferência entre links de comunicação utilizando diferentes f_c .

- **PROBLEMA 10.6** Compare a largura de faixa de transmissão necessária para o PAM binário e a modulação BPSK se os dois sinais tiverem uma taxa de dados de 9600 bps e utilizarem espectro de raiz de cosseno levantado com fator de roll-off de 0,5.
- **PROBLEMA 10.7** Trace um diagrama de blocos para um sistema de transmissão incluindo tanto o receptor quanto o transmissor de uma modulação BPSK com formato de pulso de raiz de cosseno levantado.

10.5 DETECÇÃO DE QPSK E QAM COM RUÍDO

Com o conhecimento da detecção de dados binários com ruído para banda base (PAM) e banda passante (BPSK) a nossa disposição, agora estendemos estes resultados para esquemas de modulação mais complexos, tais como chaveamento em quadratura de fase (QPSK) e modulação em amplitude em quadratura (QAM).

■ Detecção de QPSK com ruído

Vimos na Seção 7.3 que o sinal $s(t)$ modulado em QPSK pode ser representado na forma

$$s(t) = A_c \cos(2\pi f_c t + \phi(t)) \quad (10.46)$$

para $0 \leq t \leq T$. A portadora é transmitida em uma de quatro fases, cada fase representando um par de bits (dibit), como dado por

$$\phi(t) = \begin{cases} \frac{3\pi}{4}, & \text{dibit 00} \\ \frac{\pi}{4}, & \text{dibit 10} \\ -\frac{\pi}{4}, & \text{dibit 11} \\ -\frac{3\pi}{4}, & \text{dibit 01} \end{cases} \quad (10.47)$$

Também foi mostrado que a Eq. (10.46) pode ser representada em uma forma de quadratura pela expansão do tempo em cosseno, obtendo

$$\begin{aligned} s(t) &= A_c \cos[\phi(t)] \cos(2\pi f_c t) - A_c \sin[\phi(t)] \sin(2\pi f_c t) \\ &= A_c m_I(t) \cos(2\pi f_c t) - A_c m_Q(t) \sin(2\pi f_c t) \end{aligned} \quad (10.48)$$

Na Eq. (10.48), identificamos $\cos[\phi(t)]$ como a componente em fase $m_I(t)$ do sinal de mensagem e $\sin[\phi(t)]$ como a componente em quadratura $m_Q(t)$. A similaridade da Eq. (10.48) com a multiplexação por portadora em quadratura discutida na Seção 2.5 para sinais analógicos leva ao receptor em quadratura mostrado na Fig. 10.8. Utilizando a representação em fase e quadratura para o ruído passa-faixa, vemos que a entrada QPSK do detector coerente da Fig. 10.8 é descrita por

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= (A_c m_I(t) + n_I(t)) \cos(2\pi f_c t) - (A_c m_Q(t) + n_Q(t)) \sin(2\pi f_c t) \end{aligned} \quad (10.49)$$

Análogo aos resultados obtidos com a detecção BPSK, a saída intermediária do ramo superior da Fig. 10.8 é

$$\begin{aligned} v_I(t) &= \frac{1}{2}(A_c m_I(t) + n_I(t)) \\ &\quad + \frac{1}{2}(A_c m_I(t) + n_I(t)) \cos(4\pi f_c t) - \frac{1}{2}(A_c m_Q(t) + n_Q(t)) \sin(4\pi f_c t) \end{aligned} \quad (10.50)$$

na qual o termo do lado direito da primeira linha é um sinal passa-baixa e os termos na segunda linha são sinais de alta freqüência. Tal como no BPSK utilizando o formato de pulso retangular, como $m_I(t)$ é constante durante a duração do pulso, um filtro casado com formato retangular (integra-e-descarta) pode ser utilizado para recuperar $m_I(t)$ maximizando a razão sinal/ruído da saída. Como mencionado anteriormente, tal filtro também rejeita os termos de alta freqüência.

Os comentários feitos na Seção 10.4 também se aplicam à Fig. 10.8 quando comparada com a Fig. 7.7(b). Na Fig. 7.7(b), o filtro passa-baixa remove as componentes de alta-freqüência. Na Fig. 10.8, o integrador possui este propósito além de maximizar a razão sinal/ruído.

Similarmente, a saída do ramo inferior do detector em quadratura é

$$\begin{aligned} v_Q(t) &= -\frac{1}{2}(A_c m_Q(t) + n_Q(t)) \\ &\quad + \frac{1}{2}(A_c m_Q(t) + n_Q(t)) \cos(4\pi f_c t) + \frac{1}{2}(A_c m_I(t) + n_I(t)) \sin(4\pi f_c t) \end{aligned} \quad (10.51)$$

O sinal banda base $m_Q(t)$ pode ser recuperado de maneira similar a $m_I(t)$.

Para a componente em fase, se o primeiro bit do díbit for um 0, então a saída média é

$$\mu = \frac{A_c T}{2} \cos\left(\frac{\pm 3\pi}{4}\right) = -\frac{A_c T}{2\sqrt{2}}$$

Por outro lado, se o primeiro bit do díbit for um 1, então

$$\mu = \frac{A_c T}{2} \cos\left(\frac{\pm \pi}{4}\right) = \frac{A_c T}{2\sqrt{2}}$$

Após a filtragem passa-baixa, a forma da Eq. (10.50) é a mesma que encontramos para o BPSK na Eq. (10.40). Conseqüentemente, a probabilidade de erro no ramo em fase do sinal QPSK é

$$P_e = Q\left(\frac{|\mu|}{\sigma}\right) \quad (10.52)$$

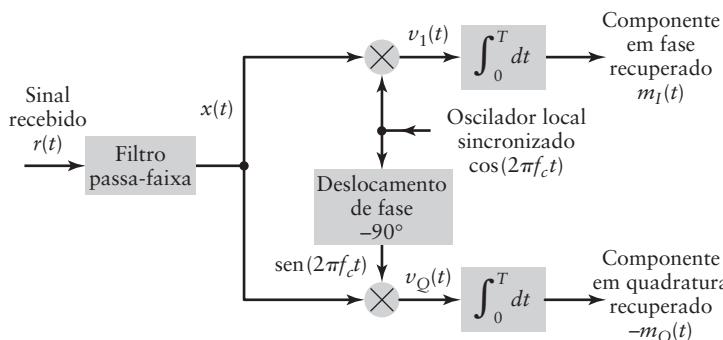


Figura 10.8 Receptor em quadratura para QPSK.

na qual σ é a raiz quadrada da variância do erro (isto é, o desvio padrão). O $|\mu|$ indica simetria entre o bit 0 e o bit 1. O mesmo resultado é válido para a componente em quadratura. Para expressar este resultado em termos do modelo de referência digital, observamos que, com a modulação QPSK, dois bits são transmitidos em um intervalo de símbolo de largura T . Conseqüentemente, a energia média por bit pode ser determinada por

$$\begin{aligned} 2E_b &= \mathbf{E}\left[\int_0^T s^2(t) dt\right] \\ &= A_c^2 \int_0^T \cos^2(2\pi f_c t + \phi(t)) dt \\ &= A_c^2 \int_0^T \frac{1 + \cos(4\pi f_c t + 2\phi(t))}{2} dt \\ &\approx \frac{A_c^2 T}{2} \end{aligned} \quad (10.53)$$

na qual, com a consideração banda passante, a integral do termo de alta freqüência é aproximadamente zero. Observamos também que a variância do ruído na saída de cada ramo é inalterada do BPSK com $\sigma^2 = \frac{1}{4}N_0T$. Com esta diferença, a taxa de erro de bit com $v_i(t)$ após a filtragem casada é dada por

$$\begin{aligned} P_e^{\text{QPSK}} &= Q\left(\frac{|\mu|}{\sigma}\right) \\ &= Q\left(\frac{A_c T / 2 \sqrt{2}}{\sqrt{1/4 N_0 T}}\right) \\ &= Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \end{aligned} \quad (10.54)$$

na qual fizemos a substituição $A_c = \sqrt{4E_b/T}$ da Eq. (10.53). Um resultado similar é obtido para $v_Q(t)$ por simetria.

Conseqüentemente, em termos de energia por bit, a performance do QPSK é exatamente a mesma do BPSK, mesmo *transmitindo duas vezes mais bits pelo mesmo canal*. Este importante resultado é devido a inerente natureza ortogonal nas componentes em fase e quadratura. Mais uma vez, vemos uma analogia com a comunicação analógica quando compararmos a transmissão de faixa lateral dupla e faixa lateral única e percebemos que podemos obter a mesma qualidade de performance mas com metade da largura de faixa de transmissão. Com a modulação QPSK, utilizamos a mesma largura de faixa de transmissão da BPSK, mas transmitimos duas vezes mais bits com a mesma confiabilidade.

Como descrito na Seção 7.3, o QPSK deslocado ou OQPSK é uma variante da modulação QPSK na qual a componente em quadratura é atrasada por metade de um período de símbolo relativa a componente em fase. Com a consideração de banda passante, atrasar a componente em quadratura não afeta a ortogonalidade com a componente em fase. Conseqüentemente, podemos utilizar o mesmo detector em quadratura para a recuperação do OQPSK que foi utilizado para o QPSK. A única diferença é que a amostragem do detector integra-e-descarta da componente em quadratura ocorre meio símbolo depois que a amostragem da componente em fase.

Como resultado desta similaridade entre o OQPSK e QPSK, a performance da taxa de erro de bit dos dois esquemas é idêntica se o caminho de transmissão não distorcer o sinal. Como mencionado no Capítulo 7, uma vantagem do OQPSK é sua reduzida variação de fase e ter potencialmente menos distorção se o caminho de transmissão inclui compo-

nentes não lineares tais como um amplificador operando próximo da saturação. Com estas condições não lineares, OQPSK pode ter uma performance melhor do que QPSK.

- **PROBLEMA 10.8** Mostre que a integral do termo de alta freqüência da Eq. (10.53) é aproximadamente zero. ◀

■ Detecção de QAM com ruído

A modulação em amplitude em quadratura (QAM), discutida na Seção 7.7, também é uma estratégia de modulação passa-faixa devido ao uso das componentes em fase e quadratura da portadora. A QAM pode ser vista como um híbrido de PAM multinível e QPSK. Em particular, QAM utiliza a componente em fase e quadratura para a transmissão tal como a QPSK faz. Entretanto, em cada componente em fase e quadratura na QAM, o modulador utiliza múltiplos níveis, ou seja, mais do que os dois níveis de $\pm\cos \phi$ ou $\pm\sin \phi$ utilizados no QPSK.

Primeiro, vamos considerar o PAM multinível. Vamos representar o sinal modulado banda base por

$$s(t) = A \sum_{k=0}^{\infty} c_k h(t - kT) \quad (10.55)$$

na qual c_k representa os diferentes níveis de modulação. Por exemplo, com um PAM de quatro níveis, o escalar c_k pode ser selecionado do conjunto $\{-3, -1, +1, +3\}$. O diagrama de olho para este sinal multinível é semelhante ao mostrado na Fig. 10.9.

Se revisarmos a análise do procedimento ótimo que levou ao detector de filtro casado, veremos que ele não depende da amplitude do sinal. Conseqüentemente, o projeto do detector para o receptor PAM multinível é idêntico ao do PAM liga-desliga. A única diferença entre o PAM binário e o PAM multinível são as comparações de limiar utilizadas para determinar qual dos níveis foi transmitida, como mostrado na Fig. 10.10.

Para cada valor de símbolo y na saída do detector, a probabilidade de erro é intuitivamente minimizada pela escolha do nível de modulação mais próxima – por exemplo, de $\{-3, -1, +1, +3\}$. Fazemos isto comparando a saída y com uma série de limiares antes de tomar a decisão. Para o diagrama de olho de quatro níveis mostrado na Fig. 10.9, os limiares

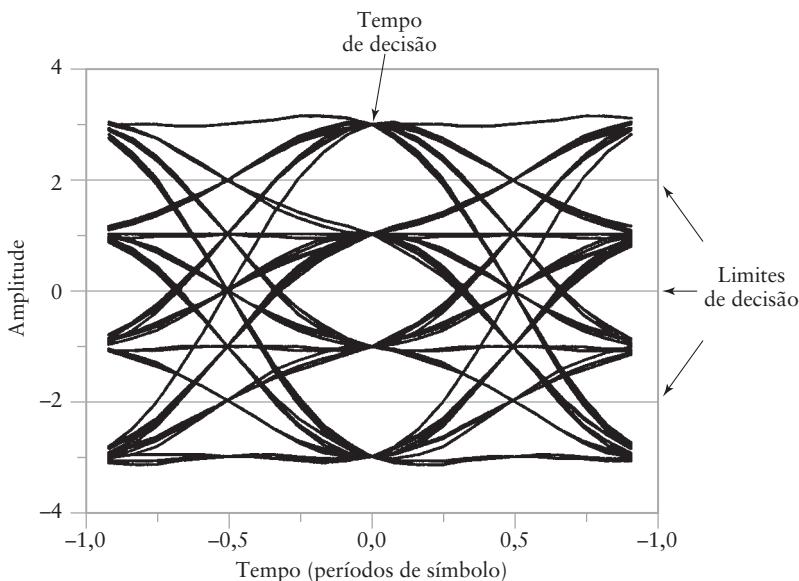


Figura 10.9 Diagrama de olho para o PAM de quatro níveis.

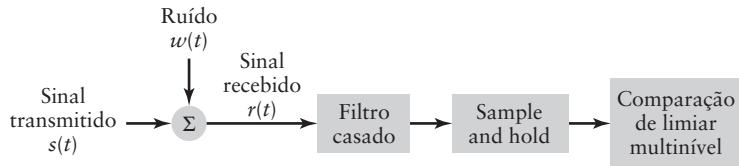


Figura 10.10 Detector PAM multinível.

são os valores 0 e ± 2 . Estes três limiares são mostrados na Fig. 10.11. Então, por exemplo, se $0 < y < 2$, estimamos $c_k = +1$. Se $-2 < y < 0$, estimamos $c_k = -1$.

Iremos, agora, discutir a performance deste detector, mas por simplicidade iremos assumir que o período T de símbolo é um. Para a maioria dos cenários, erros acontecerão com o vizinho mais próximo. Por exemplo, para o símbolo +1, a decisão errada mais provável são os símbolos vizinhos +3 e -1. Se, em geral, assumirmos que o nível do sinal transmitido é μ e a separação entre dois vizinhos próximos é $2A$, então, para aqueles símbolos que possuem dois vizinhos próximos, a probabilidade de erro destes símbolos é

$$\begin{aligned}
 P_e &= \mathbf{P}[y < \mu - A \text{ ou } y > \mu + A] \\
 &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\mu-A} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\} dy + \frac{1}{\sqrt{2\pi}\sigma} \int_{\mu+A}^{\infty} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\} dy \\
 &= 2Q\left(\frac{A}{\sigma}\right)
 \end{aligned} \tag{10.56}$$

na qual σ é o desvio padrão do erro. Este resultado se aplica a todos os símbolos que possuem dois vizinhos. Para os símbolos externos, os quais possuem apenas um vizinho, a probabilidade de erro é metade do valor definido na Eq. (10.56). Combinando estes dois casos, pode ser mostrado que (veja o Problema 10.23) a probabilidade de erro par ao PAM M-ário é, com os níveis de modulação separados por uma distância $2A$,

$$P_e^{\text{PAM}} = 2\left(\frac{M-1}{M}\right)Q\left(\frac{A}{\sigma}\right) \tag{10.57}$$

Apesar desta expressão parecer similar à obtida para a transmissão binária, existem duas importantes diferenças que devem ser notadas:

1. A Eq. (10.57) representa a *taxa de erro de símbolo*. Cada símbolo representa mais do que um bit. Conseqüentemente, cada erro de símbolo pode corresponder a mais do que um erro de bit, apesar dos símbolos serem geralmente arranjados, ou seja, codificados como o código de Gray, de forma que isto não ocorra.
2. Com a transmissão binária com níveis $+A$ e $-A$, a potência média transmitida é A^2 . Com a estratégia PAM M -ário, assumindo que todos os níveis são equiprováveis e separados por $2A$, podemos mostrar que a potência média transmitida é $(M^2 - 1)A^2/3$ (veja

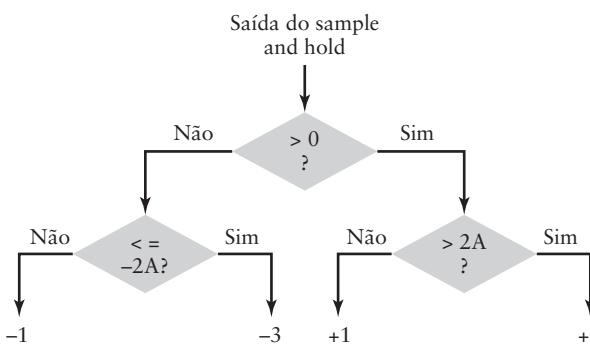


Figura 10.11 Ilustração das regras de decisão para o PAM de quatro níveis.

o Problema 10.23). Conseqüentemente, a produção de saída extra fornecida pelo PAM multinível requer significativamente mais potência para obter a mesma performance.

Assumindo a transmissão banda base, podemos relacionar a performance do PAM multinível com o modelo de referência digital da forma mostrada a seguir:

- Se assumirmos que o número de níveis M é uma potência de 2, ou seja, $M = 2^\beta$, então cada símbolo representa β bits.
- A energia média recebida por bit é $E_b = \frac{(M^2 - 1)A^2T}{3\beta}$.
- A variância do ruído na saída do filtro casado é $\sigma^2 = N_0T/2 = N_0/2$ com a consideração de $T = 1$. Isto é inalterado do caso PAM binário da Eq. (10.9).

Substituindo estes resultados na Eq. (10.57), a probabilidade de erro de símbolo em termos da RSR de referência digital é

$$P_e^{\text{PAM}} = 2\left(\frac{M-1}{M}\right)\mathcal{Q}\left(\sqrt{\frac{6\beta}{M^2-1}}\sqrt{\frac{E_b}{N_0}}\right) \quad (10.58)$$

A vantagem do esquema de modulação M -ário é o aumento do número de bits, $\beta = \log_2 M$, transmitidos ao longo do canal em cada símbolo modulado. Entretanto, a desvantagem do PAM é a penalidade de performance a ser paga quando aumentamos o número de níveis de modulação. Se fizermos $\rho = \sqrt{6\beta}/(M^2 - 1)$ na Eq. (10.58), então para $M = 2$, temos $\rho = \sqrt{2}$, o qual é o resultado padrão para PAM bipolar, BPSK e QPSK. Para $M = 4$, $\rho = \sqrt{4/5}$, o que implica que significativamente mais energia por bit deve ser transmitida para obter a mesma taxa de erro. Para $M = 8$, $\rho = \sqrt{2/7}$, o que implica que ainda mais energia por bit seja necessária para obter a mesma taxa de erro.

Após esta breve introdução a performance de PAM multinível com ruído, reforçamos que várias formas de modulação em amplitude em quadratura (QAM) consistem de sinais PAM multiníveis tanto para a componente em fase quanto em quadratura, $m_I(t)$ e $m_Q(t)$, de um sinal passa-faixa e, portanto, podemos fazer as seguintes observações:

- Podemos utilizar esquemas PAM independentes nas componentes em fase e quadratura. Dito isto, um sinal PAM $m_I(t)$ modula a portadora em fase $\cos(2\pi f_c t)$ e o segundo sinal PAM $m_Q(t)$ modula a portadora em quadratura $\sin(2\pi f_c t)$.
- Devido à ortogonalidade das componentes em fase e quadratura, a taxa de erro é a mesma para as duas, e a mesma para um sistema PAM banda base.

Com o esquema QAM, duas vezes mais dados podem ser transmitidos na mesma largura de faixa que o esquema PAM banda base com a mesma eficiência de potência. Esta propriedade é uma simples extensão da comparação que fizemos entre a modulação BPSK e QPSK.

10.6 DETECÇÃO ÓTIMA DE FSK BINÁRIO

Do Capítulo 7, outra forma simples de modulação passa-faixa é o chaveamento de freqüência (FSK). Os esquemas anteriores de modulação e estruturas de receptores considerados neste capítulo são, de várias formas, análogos à modulação em amplitude mesmo que alguns sejam chamados de chaveamento de fase. O chaveamento de freqüência, entretanto, é mais relacionado com a modulação em freqüência. No FSK, diferentes freqüências são utilizadas para representar os bits de dados e o sinal transmitido em $0 \leq t \leq T$ é

$$s(t) = \begin{cases} A_c \cos(2\pi f_1 t), & 0 \leq t \leq T \text{ se um } 1 \text{ é enviado} \\ A_c \cos(2\pi f_2 t), & 0 \leq t \leq T \text{ se um } 0 \text{ é enviado} \end{cases} \quad (10.59)$$

Nas seções anteriores nas quais a forma de onda tinha a amplitude modulada, observamos que o sinal podia ser detectado otimamente com um simples filtro casado. No sistema de transmissão FSK da Eq. (10.59), temos, na realidade, duas formas de onda senoidais com f_1 e f_2 diferentes. Parece ser lógico, então, projetar filtros casados para cada uma destas duas formas de onda e selecionar a que produzir a maior saída, como mostrado na Fig. 10.12. Em particular, os dois filtros casados são

$$g_1(t) = \sqrt{2} \cos(2\pi f_1 t)$$

e

$$g_2(t) = \sqrt{2} \cos(2\pi f_2 t) \quad (10.60)$$

correspondendo a 1 e 0, respectivamente.

Para analisar a performance do detector FSK da Fig. 10.12, fazemos a consideração de que as freqüências f_1 e f_2 foram selecionadas de tal forma que as formas de onda correspondentes são ortogonais, ou seja,

$$\int_0^T \cos(2\pi f_1 t) \cos(2\pi f_2 t) dt = 0 \quad (10.61)$$

Tecnicamente, isto implica em $(f_1 - f_2)/T$ ser um número inteiro na consideração banda passante (veja o Problema 10.16). Esta consideração não é necessária na prática, mas a sua violação resulta em performance degradada (veja o Problema 10.24). Seja o sinal recebido igual a

$$r(t) = s(t) + w(t) \quad (10.62)$$

Suponha que um 1 tenha sido transmitido. Então, a saída do filtro casado correspondente a um 0 é

$$\begin{aligned} Y_2 &= \int_0^T r(t) \sqrt{2} \cos(2\pi f_2 t) dt \\ &= \int_0^T (A_c \cos(2\pi f_1 t) + w(t)) \sqrt{2} \cos(2\pi f_2 t) dt \\ &= 0 + N_2 \end{aligned} \quad (10.63)$$

logo, da Eq. (10.61), a primeira parte da integral é zero devido a ortogonalidade das duas formas do pulso. A componente de ruído na saída do filtro casado para um 0 é

$$N_2 = \int_0^T w(t) \sqrt{2} \cos(2\pi f_2 t) dt \quad (10.64)$$

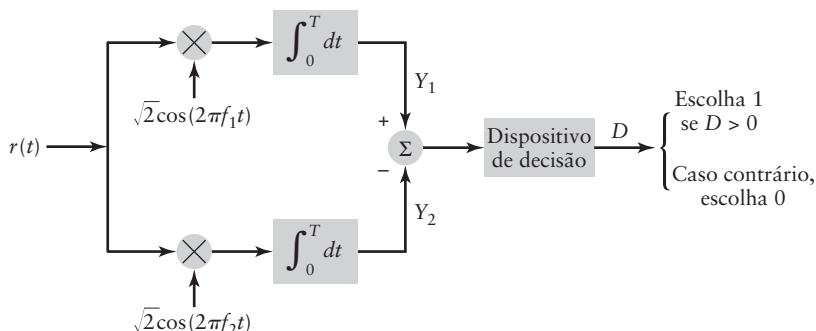


Figura 10.12 Detecção coerente de FSK binário.

Conseqüentemente, quando um 1 é transmitido, a saída do filtro casado para um 0 possui média zero e variância correspondente a N_2 . Por outro lado, a saída do filtro casado para um 1 é

$$\begin{aligned} Y_1 &= \int_0^T r(t) \sqrt{2} \cos(2\pi f_1 t) dt \\ &= \int_0^T (A_c \cos(2\pi f_1 t) + w(t)) \sqrt{2} \cos(2\pi f_1 t) dt \\ &\approx \frac{A_c T}{\sqrt{2}} + N_1 \end{aligned} \quad (10.65)$$

O passo da segunda linha para a terceira linha da Eq. (10.65) utiliza a fórmula de ângulo dobrado $2\cos^2 A = 1 + \cos 2A$. A aproximação da Eq. (10.65) é a consideração banda passante na qual a integral da componente de alta freqüência é aproximadamente zero no intervalo de 0 a T . A componente de ruído da Eq. (10.65) é definida por

$$N_1 = \int_0^T w(t) \sqrt{2} \cos(2\pi f_1 t) dt \quad (10.66)$$

Quando um 1 é transmitido, a saída do filtro casado para um 1 possui média $\mu = A_c T / \sqrt{2}$ e variância correspondente a N_1 . Um resultado simétrico é obtido quando um 0 é transmitido.

Para determinar qual bit foi transmitido, comparamos a saída dos dois filtros casados. A comparação mais simples é formar a diferença entre as saídas $D = Y_1 - Y_2$. Se um 1 é transmitido, então o valor médio de D é $\mu = A_c T / \sqrt{2}$. Se um 0 é transmitido, então o valor médio de D é $\mu = -A_c T / \sqrt{2}$. A regra de decisão óbvia, neste caso, é escolher 1 se D for maior do que zero e 0 caso contrário.

A variável aleatória D contém a diferença de N_1 e N_2 . afirmamos que estes dois termos de ruído são variáveis aleatórias Gaussianas independentes (veja o Problema 10.9). Como D é a diferença entre duas variáveis aleatórias Gaussianas independentes, ele possui variância $\sigma^2 = \text{Var}(N_1) + \text{Var}(N_2) = 2\text{Var}(N_1)$. Por analogia com o PAM bipolar da Seção 10.3, a probabilidade de erro para FSK binário é

$$P_e = Q\left(\frac{|\mu|}{\sigma}\right) \quad (10.67)$$

Colocando esta equação em termos da RSR de referência digital, vemos por inspeção da Eq. (10.59) que a energia por bit é $E_b = A_c^2 T / 2$ para o FSK. A variância do ruído da combinação dos dois termos N_1 e N_2 é $\sigma^2 = 2(N_0 T / 2) = N_0 T$ (veja o Problema 10.9). Desta forma, a TEB em termos do modelo de referência digital é

$$P_e^{\text{FSK}} = Q\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (10.68)$$

Esta performance de taxa de erro de bit é similar à da sinalização liga-desliga que vimos na Seção 10.3. Tanto o PAM liga-desliga quanto o FSK são formas de *sinalização ortogonal*, enquanto que as técnicas de modulação tais como PAM bipolar e BPSK são chamadas de *sinalização antipodal*. Em geral, vemos que a sinalização antipodal fornece uma vantagem de $10 \log_{10} 2 = 3\text{dB}$ sobre a sinalização ortogonal na performance da taxa de erro de bit.

- **PROBLEMA 10.9** Utilize as Eqs. (10.61), (10.64) e (10.66) para mostrar que N_1 e N_2 são não correlacionadas e, portanto, variáveis aleatórias Gaussianas independentes. Calcule a variância de $N_1 - N_2$.

10.7 DETECÇÃO DIFERENCIAL COM RUÍDO

No Capítulo 7, introduzimos a detecção diferencial de BPSK como um método simples de recuperação de dados sem a complexidade da detecção coerente. Uma estrutura típica de um receptor para a detecção diferencial de BPSK na presença de ruído é mostrada na Fig. 10.13. O sinal RF recebido é primeiro convertido para baixo para a freqüência FI f_c e, então, sofre uma filtragem passa-faixa. O sinal passa-faixa na saída do filtro pode ser representado por

$$x(t) = A_c d_k \cos(2\pi f_c t + \theta) + n(t), \quad (k-1)T < t \leq kT \quad (10.69)$$

na qual T é o período de bit e θ é a fase desconhecida da portadora. O bit diferencialmente codificado d_k é definido pela equação $d_k = b_k d_{k-1}$ na qual b_k é o k -ésimo bit de informação.² Na detecção coerente, circuitos de sincronização devem estimar a fase θ para produzir a referência coerente para a conversão para baixo para banda base. Com a detecção diferencial, utilizamos o circuito de *atrasa-e-multiplica* mostrado na Fig. 19.14. A saída do circuito de atrasa-e-multiplica é

$$\begin{aligned} v(t) &= x(t)x(t-T) \\ &= A_c^2 d_k d_{k-1} \cos(2\pi f_c t + \theta) \cos(2\pi f_c(t-T) + \theta) + \omega(t) \end{aligned} \quad (10.70)$$

na qual

$$\omega(t) = A_c d_k \cos(2\pi f_c t + \theta) n(t-T) + A_c d_{k-1} \cos(2\pi f_c(t-T) + \theta) n(t) + n(t)n(t-T) \quad (10.71)$$

é a soma dos termos de sinal vezes ruído e ruído vezes ruído. Implicitamente assumimos que a fase θ desconhecida permanece essencialmente constante durante o intervalo de dois bits. Utilizando a identidade trigonométrica $\cos A \cos B = \frac{1}{2}[\cos(A-B) + \cos(A+B)]$, podemos expandir a Eq. (10.70) para obter

$$v(t) = \frac{A_c^2}{2} d_k d_{k-1} \{\cos(2\pi f_c T) + \cos(4\pi f_c t - 2\pi f_c T + 2\theta)\} + \omega(t) \quad (10.72)$$

O detector integra-e-descarta mostrado na Fig. 10.14 remove as componentes de alta freqüência da Eq. (10.72) e o termo restante é

$$\begin{aligned} Y_k &= \int_{(k-1)T}^{kT} v(t) dt \\ &= \frac{A_c^2 T}{2} d_k d_{k-1} \cos(2\pi f_c T) + N_k \end{aligned} \quad (10.73)$$

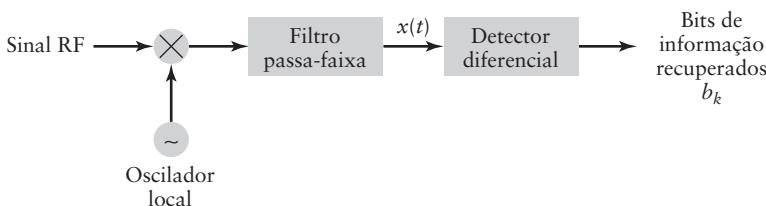


Figura 10.13 Processamento passa-faixa do detector diferencial

² Na discussão sobre codificação diferencial na Seção 7.6, b_k foi considerado como assumindo os valores lógicos 0 e 1. Os símbolos diferencialmente codificados eram dados por $d_k = b_k \oplus d_{k-1}$ os quais eram, então, modulados. Nesta seção, assumimos que b_k assume os valores reais de +1 e -1, e $d_k = d_{k-1} b_k$ fornece diretamente os níveis PAM. Estas duas representações são equivalentes.

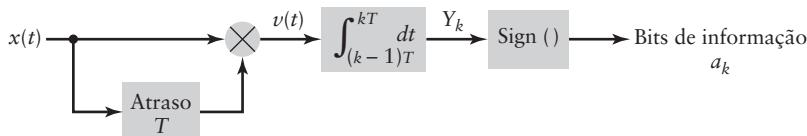


Figura 10.14 Detector diferencial.

na qual a variável aleatória aditiva N_k é devido a $\omega(t)$ da Eq. (10.71). Se a freqüência f_c for escolhida tal que $f_c T$ é aproximadamente um inteiro, então $\cos(2\pi f_c T) \approx 1$ e

$$\begin{aligned} Y_k &= \frac{A_c^2 T}{2} d_k d_{k-1} + N_k \\ &= \frac{A_c^2 T}{2} b_k + N_k \end{aligned} \quad (10.74)$$

na qual utilizamos o fato do dado de informação b_k ter sido diferencialmente codificado, tal que $b_k = -d_k d_{k-1}$.

Conseqüentemente, em termos de performance de ruído, a maior diferença entre um sistema DPSK e o PSK binário coerente não é a codificação diferencial, a qual pode ser utilizada em qualquer um dos casos, mas está na forma pela qual o sinal de referência é obtido para a detecção de fase do sinal recebido. Especificamente, em um receptor DPSK, a referência é contaminada pelo ruído aditivo na mesma extensão do pulso de informação. Ou seja, eles possuem a mesma razão sinal/ruído. Isto é evidente na presença dos termos tanto do ruído vezes sinal quanto ruído vezes ruído na expressão de $\omega(t)$ da Eq. (10.71). Isto faz com que a caracterização estatística da variável aleatória N_k e, portanto, a determinação da probabilidade total de erro em receptores DPSK seja, de alguma forma, complicada e além do escopo deste livro. Entretanto, a fórmula é³

$$P_e^{\text{DPSK}} = 0,5 \exp\left(-\frac{E_b}{N_0}\right) \quad (10.75)$$

É interessante notar que, como em um receptor DPSK as decisões são feitas com base no sinal recebido em dois intervalos sucessivos de bit, existe uma tendência de ocorrerem bits de erro em pares. A principal condição de $f_c T$ ser aproximadamente um inteiro em um receptor diferencial pode ser aliviada em vários casos de a FI for escolhida para ser aproximadamente zero, ou seja, $f_c \approx 0$ Hz.

10.8 RESUMO DE PERFORMANCE DIGITAL

■ Codificação Gray

Nesta seção, iremos resumir a performance de estratégias digitais de modulação/demodulação. Entretanto, antes de fazermos isto, devemos elaborar um comentário feito sobre PAM. Em particular, com PAM M -ário, calculamos a probabilidade de um erro de símbolo. Como, para $M > 2$, um símbolo é composto por dois ou mais bits, um erro de símbolo pode implicar em mais de um erro de bit. Entretanto, pela associação adequada de bits aos símbolos, a probabilidade disto ocorrer pode ser pequena. Especificamente, nos referimos ao *código de Gray* para codificar os símbolos, no qual existe apenas um bit de diferença entre símbolos adjacentes. Na Fig. 10.15 mostramos as estratégias de código de Gray para PAM 8-ário e 8-PSK. Observe que o erro mais comum nas duas constelações é selecionar o

³ Para a obtenção da fórmula dada na Eq. (10.75), veja Haykin (2001), pp. 407-417.

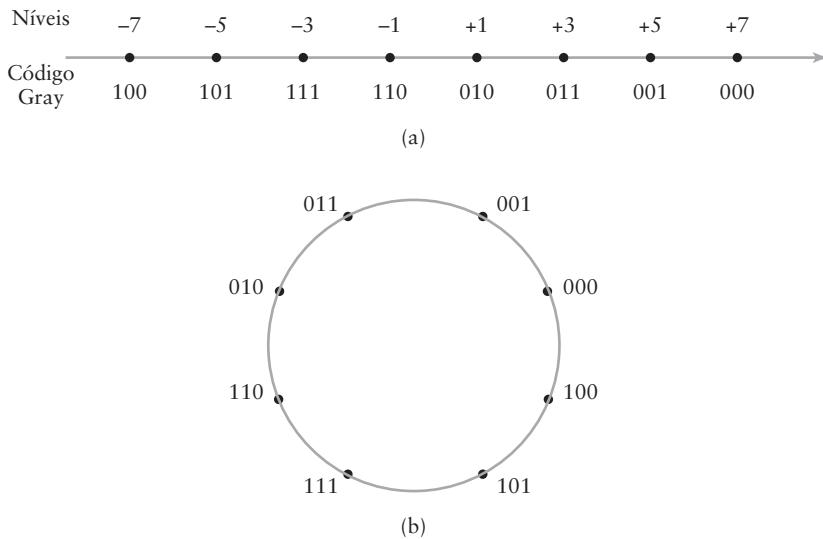


Figura 10.15 Código de Gray para (a) PAM 8-ário e (b) 8-PSK.

vizinho mais próximo. Logo, para esta codificação, a taxa de erro de símbolo e a taxa de erro de bit são quase idênticas. A exceção a esta igualdade ocorre em canais muito ruidosos nos quais a probabilidade de seleção de um símbolo que não é o vizinho mais próximo se torna grande.

■ Comparação de performance

Com a consideração de que a taxa de erro de símbolo é igual a taxa de erro de bit, podemos formular a comparação de performance de diferentes estratégias de modulação/demodulação mostrada na Tabela 10.1. Dividimos a tabela entre modulação banda base e banda passante. Apesar do PAM M -ário possuir um equivalente banda base, ele é raramente utilizado, pois geralmente se utiliza o QAM em seu lugar. Na Tabela 10.1, extrapolamos os resultados do PAM M -ário para o QAM M -ário como discutido no texto, baseado na analogia entre sinalização banda base bipolar e modulação QPSK.

Para os resultados da sinalização M -ário da Tabela 10.1, assumimos que:

- Para $M > 2$, a razão sinal/ruído é suficiente para que a taxa de erro de símbolo seja aproximadamente igual à taxa de erro de bit quando se utiliza o código de Gray.
- A ordem M de modulação é uma potência de 2, ou seja, $M = 2^\beta$ para β inteiro. Para QAM M -ário, existe uma consideração implícita de β ser par.
- A estratégia QAM M -ário utiliza um PAM M -ário em cada componente, em fase e quadratura.

Na Fig. 10.16, comparamos graficamente a performance de TEB de várias estratégias de transmissão diferentes. Note que a razão sinal/ruído de referência, E_b/N_0 , é medida em decibéis.

Se compararmos a performance de sistemas de transmissão digital com a de sistemas de transmissão analógica, podemos, então, fazer as seguintes observações:

- Em sistemas analógicos, geralmente precisamos de razões sinal/ruído pós-detecção de 20 a 20 dB para uma transmissão aceitável de voz. Para a modulação AM, a largura de faixa do canal é normalmente de 3 a 5 kHz. Para a transmissão FM, a largura de faixa do canal é normalmente de 25 a 30 kHz.
- Com o PCM codificando a voz, o sinal é geralmente amostrado a 8 kHz, com 8 bits de resolução (256 níveis). A taxa de dados resultante é 64 quilobits por segundo. Se esta seqüência de dados é enviada em um canal com um esquema de modulação em

TABELA 10.1 Comparação da performance da TEB para várias estratégias de modulação-demodulação. A detecção é coerente a não ser que indicado. (O parâmetro β é o número de bits por símbolos na dimensão em fase ou quadratura)

Modulação-Demodulação		Taxa de erro de bit
Banda base	Banda passante	
PAM liga-desliga	ASK	$Q\left(\sqrt{\frac{E_b}{N_0}}\right)$
PAM bipolar	BPSK	$Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$
PAM M -ário	QPSK	$2\frac{M-1}{M}Q\left(\sqrt{\frac{6\beta}{M^2-1}}\sqrt{\frac{E_b}{N_0}}\right)$
	QAM M -ário	$2\frac{\sqrt{M}-1}{\sqrt{M}}Q\left(\sqrt{\frac{6\beta}{M-1}}\sqrt{\frac{E_b}{N_0}}\right)$
	FSK binário	$Q\left(\sqrt{\frac{E_b}{N_0}}\right)$
	BPSK diferencial	$\frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$

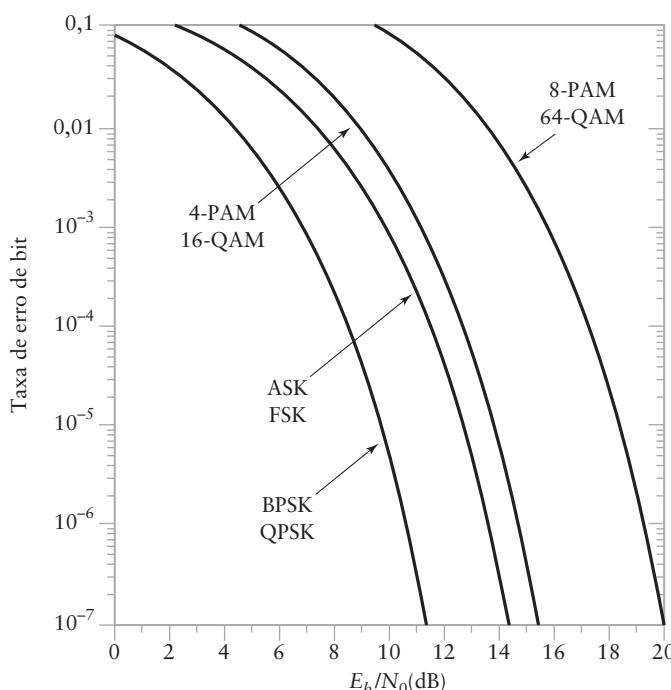


Figura 10.16 Comparação da TEB versus E_b/N_0 para várias estratégias de transmissão.

quadratura tal como QPSK com 32 kpbs para cada componente em fase e quadratura, a largura de faixa nominal do sinal digital é 32kHz. Se a taxa de erro tolerável para este dado for 10^{-5} , então, da Fig. 10.16, a E_b/N_0 pós-detecção deve ser 9,6 dB ou maior.

- Com o avanço das técnicas de codificação de voz (vocoders), a fala humana pode ser representada com razoável qualidade com apenas 4 kbps. Isto requer uma largura de faixa de transmissão de, aproximadamente, 2 kHz. Estes vocoders podem operar com taxas de erro de bit tão altas quanto 10^{-2} com pouca degradação na qualidade, reduzindo, portanto, a E_b/N_0 necessária para níveis tão baixos quanto 4dB.

A conclusão é que na mudança de técnicas analógicas para digitais, melhoramos a robustez da transmissão. Por robustez queremos dizer relativa insensibilidade ao ruído. Com o progresso no processamento de sinais, também reduzimos a necessidade de largura de faixa.

- **PROBLEMA 10.10** Trace a performance de TEB para BPSK diferencial e compare com os resultados da Fig. 10.16. ◀

Ruídos em modelos de espaço de sinal

No Capítulo 7, vimos como estratégias de modulação diferentes podem ser representadas em termos de seus diagramas de espaço de sinal. Com o diagrama de espaço de sinal, ou constelação, cada símbolo é representado como um ponto de sinal com a escolha apropriada de funções base. Por exemplo, o QPSK pode ser representado por

$$s(t) = \pm \sqrt{E_b} \phi_1(t) \pm \sqrt{E_b} \phi_2(t) \quad (10.76)$$

na qual a escolha dos sinais indica a polaridade de bit. As funções base são

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), \quad 0 \leq t \leq T \quad (10.77)$$

e

$$\phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t), \quad 0 \leq t \leq T \quad (10.78)$$

as quais são ortogonais na consideração banda passante. Um diagrama de espaço de sinal com ruído para o QPSK é mostrado na Fig. 10.17, na qual os quatro pontos da constelação são $(+\sqrt{E_b}, +\sqrt{E_b})$, $(+\sqrt{E_b}, -\sqrt{E_b})$, $(-\sqrt{E_b}, -\sqrt{E_b})$ e $(-\sqrt{E_b}, +\sqrt{E_b})$. Quando existe ruído de canal presente, o ponto de sinal recebido é aleatoriamente localizado ao redor do ponto de sinal transmitido. Na Fig. 10.17, o ponto de sinal recebido com ruído está dentro da “nuvem de ruído”, centrada no ponto de sinal transmitido em $(+\sqrt{E_b}, +\sqrt{E_b})$.

Se o ruído do canal é Gaussiano, então a nuvem da Fig. 10.17 representa uma distribuição Gaussiana bidimensional com média $(+\sqrt{E_b}, +\sqrt{E_b})$. A variância desta distribuição, σ^2 , depende da razão sinal/ruído do canal. Considere, agora, a probabilidade de um erro quando detectamos este ponto de sinal. Existem dois tipos de erros que podem ocorrer: o ruído de canal pode fazer com que o ponto do sinal recebido esteja no lado oposto do eixo $\phi_1(t)$. Isto causaria um erro no segundo bit do par. Alternativamente, se o ruído do canal fizer com que o ponto do sinal recebido esteja na esquerda do eixo $\phi_2(t)$, então o erro irá ocorrer no primeiro bit. No evento menos provável, no qual o ponto do sinal recebido está no terceiro quadrante, quando $(+\sqrt{E_b}, +\sqrt{E_b})$ é transmitido, então um erro irá ocorrer nos dois bits. Claramente, a probabilidade do tipo de erro está relacionada com a razão entre $\sqrt{E_b}$ e o desvio padrão do ruído, σ . Apesar de não ser óbvio do diagrama, a probabilidade de um tipo de erro é independente de outro tipo de erro.

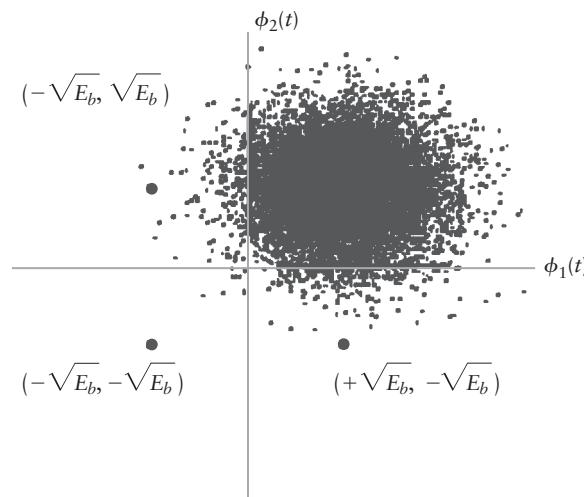


Figura 10.17 Representação por espaço de sinal de uma constelação ruidosa.

Para esta constelação, existe um círculo de raio $\sqrt{E_b}$ ao redor de cada ponto de sinal no qual nenhum erro pode ocorrer. Ou seja, $\sqrt{E_b}$ é a distância máxima de erro na qual podemos sempre garantir detecção perfeita. Este espaço de sinal pode ser facilmente estendido para várias dimensões. Por exemplo, se os pontos de sinal forem $(\pm\sqrt{E_b}, \pm\sqrt{E_b}, \pm\sqrt{E_b})$, então, ao redor de cada ponto de sinal existe uma esfera de raio $\sqrt{E_b}$ na qual garante-se que o ruído não causa um erro.

Como podemos gerar tais sinais de alta dimensão na prática? Considere o sinal

$$s(t) = s_1\phi_1(t) + s_2\phi_2(t) + s_3\phi_3(t) + s_4\phi_4(t)$$

no qual $\phi_1(t)$ e $\phi_2(t)$ são definidos nas Eqs. (10.77) e (10.78), mas estendidos de tal forma que eles são não nulos no intervalo $T < t \leq 2T$. As duas funções ortonormais restantes são dadas por

$$\phi_3(t) = \begin{cases} 0, & 0 < t \leq T \\ \sqrt{\frac{2}{T}} \cos(2\pi f_c t), & T < t \leq 2T \end{cases} \quad (10.79)$$

e

$$\phi_4(t) = \begin{cases} 0, & 0 < t \leq T \\ \sqrt{\frac{2}{T}} \sin(2\pi f_c t), & T < t \leq 2T \end{cases} \quad (10.80)$$

Com a consideração banda passante, as funções $\phi_3(t)$ e $\phi_4(t)$ são claramente ortogonais entre si, como elas são zero sempre que $\phi_1(t)$ e $\phi_2(t)$ são não nulas, elas também são ortogonais a estas funções. O ponto de sinal resultante é definido pelo vetor de dimensão quatro (s_1, s_2, s_3, s_4) , no qual s_i vale $\pm\sqrt{E_b}$. Apesar deste espaço de dimensão 4 poder ser visto como algo artificial, sua utilidade ficará mais evidente uma vez que tenhamos discutido códigos de correção de erro na próxima seção.

10.9 DETECÇÃO E CORREÇÃO DE ERRO

Uma característica fundamental de sistemas de comunicação é que o receptor não possui o conhecimento prévio da informação que está sendo transmitida através do canal. Consequentemente, no esquema de transmissão digital, quando o canal produz erros nos bits detectados, ele degrada a qualidade da comunicação. Nesta seção iremos ver algumas formas elementares de detecção e correção de erros que ocorrem na transmissão.⁴

Nas seções anteriores, mostramos que a performance da taxa de erro de bit em um sistema digital depende da RSR pós-detectação. Se a única fonte de degradação do canal é ruído branco Gaussiano aditivo, então esta TEB geralmente possui uma dependência exponencial com a RSR, como exemplificado pela função Q . Como resultado, a TEB rapidamente se torna pequena quanto a RSR melhora, geralmente fornecendo uma qualidade de transmissão melhor do que a que pode ser obtida por métodos analógicos. Em algumas instâncias, a performance digital ainda é insuficiente e queremos melhorá-la. Alguns exemplos são:

- ▶ Em canais de rádio (sem fio), a potência do sinal recebido pode variar com o tempo devido ao desvanecimento. Portanto, mesmo que a potência média do sinal resulte em uma boa performance em um canal sem desvanecimento, o sistema vivencia uma performance muito pior quando o sinal é desvanecido.
- ▶ Em aplicações via satélite, o satélite possui potência de transmissão limitada. Em tais casos, queremos obter uma performance TEB similar com menos energia transmitida.
- ▶ Em alguns sistemas de transmissão via cabo, os cabos podem ser agrupados tão próximos que pode existir crosstalk (interferência) entre os fios. Em tais circunstâncias, pode ser útil haver uma redundância no sistema para corrigir erros introduzidos por essa interferência ocasional.

Podemos atingir o objetivo de melhorar a performance de taxa de erro de bit pela adição de alguma *redundância* na seqüência transmitida. O propósito desta redundância é permitir que o receptor detecte e/ou corrija erros que foram introduzidos durante a transmissão. Este tipo de operação é chamado de *correção de erro à frente* (CEF) e sua posição em um sistema de transmissão digital é mostrada na Fig. 10.18:

- ▶ A mensagem digital de entrada (*bits de informação*) é *codificada* para produzir os *bits de canal*. Os bits de canal incluem os bits de informação, possivelmente em uma forma modificada, mais bits adicionais que são utilizados para a correção de erro.
- ▶ Os bits de canal são *modulados* e transmitidos pelo canal.
- ▶ O sinal mais ruído recebido é *demodulado* para produzir uma estimativa dos bits de canal.
- ▶ Os bits de canal estimados são, então, *decodificados* para fornecer uma estimativa da mensagem digital original.

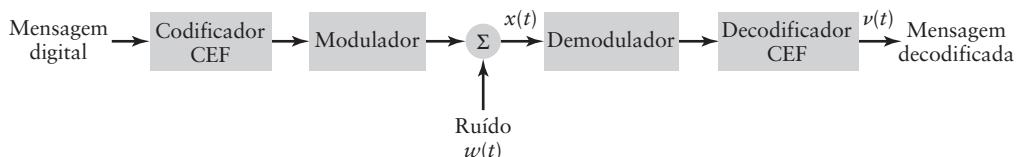


Figura 10.18 Diagrama em blocos do sistema incluindo codificação e decodificação.

⁴ A área de correção e detecção de erros nasceu, essencialmente, nos idos de 1940 com os artigos clássicos de R. Hamming e C. Shannon. Desde então uma enorme quantidade de pesquisa tem sido feita nesta área. Alguns livros introdutórios na área incluem Clark e Cain (1981), Lin e Costello (2004) e Blahut (1983). O material desta seção é adaptado de Arazi (1988).

Por detecção de erro, entendemos a capacidade do receptor detectar quando um ou mais erros tiverem ocorrido durante a transmissão dos dados. Isto não implica que o receptor irá corrigir os erros detectados, nem que o receptor será capaz de detectar todas as situações nas quais erros acontecerem.

Considere a transmissão de um bloco de dados contendo k bits de informação. Como afirmado acima, o receptor não sabe coisa alguma sobre os k bits de informação e, portanto, o bloco precisa de propriedades que permitirão ajudar o receptor a detectar quando um erro ocorre. Vamos assumir que cada bloco contém $n > k$ bits. Os $(n - k)$ bits extras são *redundantes*, ou seja, eles não transportam nova informação. Entretanto, o receptor os utiliza para determinar quando um erro ocorreu.

Para explicar como isto pode ser feito, iremos apresentar os seguintes conceitos:

- Sejam os bits representados por valores 0 e 1.
- Seja \oplus a representação da *soma de módulo 2*. Quando esta operação é aplicada a pares de bits, obtemos os resultados: $0 \oplus 0 = 1 \oplus 1 = 0$ e $0 \oplus 1 = 1 \oplus 0 = 1$.
- O operador \oplus também pode ser aplicado a blocos de bits, nos quais ele significa a soma de módulo 2 dos respectivos elementos do bloco, bit a bit. Por exemplo, $[1001] \oplus [0101] = [1100]$.

Agora, para o bloco de k bits de informação $[x_1 x_2 \dots x_k]$, suponha que adicionemos um bit p de *verificação de paridade*, tal que

$$x_1 \oplus x_2 \oplus \dots \oplus x_k \oplus p = 0 \quad (10.81)$$

e façamos a transmissão de $n = k + 1$ bits $[x_1 x_2 \dots x_k, p]$. Como este bloco estendido satisfaz a Eq. (10.81), dizemos que o bloco de bits possui *paridade par*.

No receptor, executamos a soma de módulo 2 do bloco de n bits recebidos. Se o resultado for 1 (ou seja, ele possui paridade ímpar), então podemos afirmar que existe um erro na transmissão. Desta afirmativa, deve ficar claro que um único bit de paridade irá detectar se 1, 3, 5, ..., ou seja, um número ímpar de erros ocorre durante a transmissão. Se um número par (2, 4, 6,...) de erros ocorrerem, a soma de módulo 2 (a verificação de paridade) será zero e os erros não serão detectados. Isto reforça a afirmativa anterior de que o procedimento de detecção de erros pode não detectar todos os erros. Portanto, o esquema de detecção de erro consiste de duas componentes. O *codificador* que processa os bits de informação, calcula os bits de paridade utilizando a informação e formata a combinação da informação com a paridade para a transmissão. O *decodificador* faz a verificação dos bits recebidos e declara se um erro foi detectado. Com esta base, definimos que a *capacidade de detecção de erro de um código* é o *número máximo de erros que o receptor pode sempre detectar na palavra de código transmitida*.

■ Detecção de erros com códigos de bloco

Um *código de bloco* é uma coleção de blocos binários chamados de *palavras de código*, todos de mesmo tamanho. Na Tabela 10.2, mostramos todos os blocos binários de tamanho quatro com paridade par. Uma palavra de código de tamanho n inclui $(n - k)$ bits de paridade que são calculados dos bits de informação. Conseqüentemente, o número de palavras de código em um bloco (n, k) é 2^k .

TABELA 10.2 Todas as palavras de código de tamanho 4 com paridade par

<i>Palavra de código</i>	
0000	1001
0011	1010
0101	1100
0110	1111

Sejam \mathbf{A} e \mathbf{B} duas palavras de código de tamanho n pertencentes a um código de bloco. Dizemos que o código de bloco é *linear* se a soma das duas palavras de código $\mathbf{A} \oplus \mathbf{B} = \mathbf{C}$ também for uma palavra de código. Note que um código linear sempre contém todas as palavras de código nulas, pois $\mathbf{A} \oplus \mathbf{A} = 0$.

Para analisar a detecção de erro e a capacidade de correção de códigos de bloco, precisamos de várias definições. Definimos o *peso de Hamming* de um bloco binário como o número de 1s do bloco. Para um bloco binário \mathbf{A} , iremos representar o peso de Hamming de \mathbf{A} por $w_H(\mathbf{A})$. Por exemplo, o peso de Hamming de $[1\ 0\ 0\ 1]$ é dois.

A *distância de Hamming* entre dois blocos binários é o número de lugares nos quais eles diferem. Como a expressão $\mathbf{A} \oplus \mathbf{B}$ possui 1 em cada posição na qual \mathbf{A} e \mathbf{B} diferem e zeros em todos os outros locais, a distância de Hamming $d_H(\mathbf{A}, \mathbf{B})$ pode ser expressa como

$$d_H(\mathbf{A}, \mathbf{B}) = w_H(\mathbf{A} \oplus \mathbf{B}) \quad (10.82)$$

A *distância mínima de Hamming* de um código é o menor valor obtido pela medição da distância entre todos os possíveis pares das palavras de código. Matematicamente, a distância mínima de Hamming, d_{\min} , é dada por

$$d_{\min} = \min_{\mathbf{A}, \mathbf{B}} d_H(\mathbf{A}, \mathbf{B}) \quad (10.83)$$

Esta definição assume que \mathbf{A} e \mathbf{B} não são iguais. Como um código linear sempre contém todas as palavras de código nulas, temos que, fazendo $\mathbf{B} = 0$ na Eq. (10.83), d_{\min} é o número mínimo de 1s em qualquer palavra de código não nula.

Seja \mathbf{A} a palavra de código transmitida e \mathbf{C} o bloco binário recebido. Podemos dizer que o bloco binário \mathbf{C} sempre será detectado como errado a não ser que ele seja outra palavra de código. Para mostrar isto, seja \mathbf{E} o vetor de erro – ou seja, o bloco binário com 1s indicando a localização de erros. Então, podemos escrever

$$\mathbf{C} = \mathbf{A} \oplus \mathbf{E} \quad (10.84)$$

pois somar 1 módulo 2 a um bit sempre irá alterar o bit. Somando \mathbf{A} aos dois lados, a Eq. (10.84) pode ser reescrita como

$$\mathbf{A} \oplus \mathbf{C} = \mathbf{E} \quad (10.85)$$

Com um código linear, se \mathbf{C} é uma palavra de código, então \mathbf{E} também será. Esta é uma importante propriedade de códigos de bloco lineares. Conseqüentemente, para garantirmos que um erro pode ser detectado, o número de bits errados deve ser menor do que d_{\min} :

$$\max w_H(\mathbf{E}) < d_{\min} \quad (10.86)$$

Ou seja, o peso máximo do erro que pode ser sempre detectado é $d_{\min} - 1$. Um código de bloco (k, k) que não possui bits de paridade possui uma distância de Hamming mínima de 1. Logo o receptor não pode detectar nenhum erro no bloco binário recebido. Um código de verificação com único bit de paridade possui uma distância mínima de Hamming de 2. Portanto, um código de verificação com um único bit de paridade pode sempre detectar um único bit de erro no bloco binário recebido.

EXEMPLO 10.2 Propriedades de códigos de bloco lineares

Para o código de tamanho 4 da Tabela 10.2, mostramos por exemplo que:

- (a) A soma de duas palavras de código é uma palavra de código:

$$[0011] \oplus [1100] = [1111]$$

- (b) A soma de uma palavra de código com ela mesmo é a palavra de código nula:

$$[1010] \oplus [1010] = [0000]$$

- (c) A soma de qualquer palavra de código e um vetor erro correspondendo a um erro não é uma palavra de código:

$$[1001] \oplus [0001] = [1000]$$

- (d) Ao menos alguns vetores de erro com peso 2 não são detectáveis:

$$[1010] \oplus [0011] = [1001]$$

■ Correção de erro

A correção de erro requer tanto a detecção dos erros quanto a determinação da posição dos erros. Definimos a *capacidade de correção de erro* de um código como o número máximo de erros no bloco binário recebido que o receptor *sempre* irá ser capaz de corrigir.

Sejam \mathbf{A} e \mathbf{B} duas palavras de código separadas por d_{\min} . Assuma que a palavra de código \mathbf{A} seja transmitida e que o bloco binário \mathbf{C} seja recebido. Suponha que o bloco binário \mathbf{C} recebido esteja a uma distância igual de cada uma das palavras de código \mathbf{A} e \mathbf{B} , ou seja, que ele possua a propriedade de

$$d_H(\mathbf{A}, \mathbf{C}) = d_H(\mathbf{B}, \mathbf{C}) = d_{\min}/2 \quad (10.87)$$

Logo, \mathbf{C} estará no meio do caminho entre \mathbf{A} e \mathbf{B} . Neste caso, apesar de \mathbf{A} ter sido transmitido, o receptor não tem justificativa para escolher \mathbf{A} no lugar de \mathbf{B} como palavra de código correta, pois a distância das duas para o bloco recebido é a mesma. Seja $\mathbf{C} = \mathbf{A} \oplus \mathbf{E}$, na qual \mathbf{E} é o vetor erro. Então para \mathbf{C} ser corrigível, ou seja, para recuperar \mathbf{A} de \mathbf{C} , a discussão da Eq. (10.87) implica em que

$$\max w_H(\mathbf{E}) < d_{\min}/2 \quad (10.88)$$

Podemos, portanto, afirmar que a correção de erro é possível se o número de erros for menor do que a metade da distância mínima de Hamming do código.

Claramente, então, o objetivo quando projetamos um código de correção de erro à frente é adicionar bits de paridade para aumentar a distância mínima de Hamming, para melhorar tanto a capacidade de detecção quanto de correção do código.

EXEMPLO 10.3 Três bits de paridade

Neste exemplo, construímos um código $(7, 4)$. Ou seja, as palavras de código possuem 7 bits e cada palavra de código possui quatro bits de informação. Seja $\mathbf{X} = [x_1, x_2, x_3, x_4]$ os bits de informação, construa os bits de paridade como mostrado a seguir

$$p_1 = x_1 \oplus x_2 \oplus x_3 \oplus x_4$$

$$p_2 = x_1 \oplus x_2$$

$$p_3 = x_3 \oplus x_4$$

As palavras de código para este código são mostradas na Tabela 10.3.

TABELA 10.3 As dezesseis palavras de código para o Exemplo 10.3

$[x_1, x_2, x_3, x_4, p_1, p_2, p_3]$		
0000000	1000101	
0001110	1001011	
0010110	1010011	
0011010	1011101	
0100101	1100000	
0101011	1101110	
0110011	1110110	
0111101	1111000	

Este código contém a palavra de código [1100000], a qual possui peso 2. Portanto, a distância mínima deste código é 2. Concluímos, portanto, que adicionar três bits de paridade nem sempre é suficiente para obter $d_{\min} = 3$. Nesta seção, iremos considerar um código (7, 4) melhor.

■ Códigos de Hamming

Nesta seção, iremos considerar alguns códigos simples mas eficientes chamados de códigos de Hamming, os quais estão entre os primeiros códigos de correção concebidos. Códigos de Hamming são uma família de códigos com tamanhos de bloco $n = 2^m - 1$ para $m = 3, 4, 5, \dots$. Para um tamanho n específico de bloco, existem m bits de paridade e $2^m - 1 - m$ bits de informação. Exemplos de código de Hamming incluem códigos (7, 4), (15, 11) e (31, 26). Todos estes códigos possuem uma distância mínima de Hamming $d_{\min} = 2$ e, portanto, podem corrigir um único bit de erro.

Como descrito anteriormente, existem dois passos no processo de codificação. O primeiro passo é a *codificação* que ocorre no transmissor. Isto requer o cálculo dos bits de paridade baseados nos bits de informação. O segundo passo é a *decodificação* que ocorre no receptor, isto requer a avaliação da verificação de paridade para determinar se, e onde, as equações de paridade foram violadas.

Iremos considerar primeiro o código de Hamming (7, 4) e seja $\mathbf{X} = [x_1, x_2, x_3, x_4]$ a representação dos quatro bits de informação. Para codificar estes bits, definimos a *matriz geradora k por n* do código de Hamming (7, 4) dada por

$$\mathbf{G} = \left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right] \quad (10.89)$$

Usando esta matriz geradora, a palavra de código de sete elementos $\mathbf{C} = [c_1, c_2, c_3, c_4, c_5, c_6, c_7]$ a ser transmitida é dada por

$$\mathbf{C} = \mathbf{X} \cdot \mathbf{G} \quad (10.90)$$

na qual o ponto \cdot representa a multiplicação vetor-matriz utilizando a aritmética de módulo-2; ou seja, os elementos c_k de \mathbf{C} são dados por

$$c_k = \bigoplus_{j=1}^4 x_j G_{jk}$$

Se expandirmos a Eq. (10.90), veremos que os elementos da palavra de código são

$$\begin{aligned} c_1 &= x_1 \\ c_2 &= x_2 \\ c_3 &= x_3 \\ c_4 &= x_4 \\ c_5 &= x_1 \oplus x_2 \oplus x_3 \\ c_6 &= x_2 \oplus x_3 \oplus x_4 \\ c_7 &= x_1 \oplus x_2 \oplus x_4 \end{aligned} \quad (10.91)$$

Comparando as Eqs. (10.89) e (10.91), observamos que as quatro primeiras colunas de \mathbf{G} são a matriz identidade e mapeiam os bits de informação neles mesmos. As três últimas colunas de \mathbf{G} formam as equações de verificação de paridade do código. Cada código de bloco possui sua própria e única matriz geradora, apesar de matrizes geradoras diferentes produzirem códigos com propriedades equivalentes. Por exemplo, se alterarmos quais-

quer duas colunas de \mathbf{G} , iremos obter um código de bloco diferente, mas com as mesmas propriedades de correção de erro do código anterior.

Seja $\mathbf{R} = [r_1, r_2, r_3, r_4, r_5, r_6, r_7]$ os bits recebidos, e possivelmente errados, do código de Hamming (7,4). Como podemos decodificar estes bits? Uma estratégia de decodificação possível é um *decodificador de lista*. Com esta estratégia, fazemos uma lista de todos os 2^7 possíveis blocos recebidos e calculamos a palavra de código mais próxima, em termos da distância de Hamming, antes de qualquer outra coisa. Então, para qualquer bloco recebido, olhamos nesta lista e determinamos a palavra de código mais próxima. Este decodificador de lista irá trabalhar conceitualmente para qualquer código. Entretanto, quando n se torna grande, a lista se torna grande demais para ter valor prático.

Uma estratégia de decodificação computacional é, portanto, uma alternativa preferível ao invés do decodificador de lista. Para isto, para o código de Hamming (7,4), podemos definir a *matriz de verificação* ($n - k$) por n , \mathbf{H} ,

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ \hline 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (10.92)$$

Note que a porção 3×3 inferior de \mathbf{H} forma uma matriz unitária e a parte acima da matriz unitária é a porção direita da matriz geradora da Eq. (10.89). Logo, as colunas da matriz de verificação \mathbf{H} formam os coeficientes das equações de paridade que as palavras do código de Hamming devem satisfazer. Para um código linear (n, k) , as colunas da matriz \mathbf{H} descrevem $(n - k)$ equações independentes pelas quais os $(n - k)$ bits de paridade são calculados.

Suponha que o bloco binário recebido seja $\mathbf{R} = \mathbf{C} \oplus \mathbf{E}$, na qual \mathbf{C} é o código transmitido e \mathbf{E} é o vetor erro. Calculamos o produto do vetor recebido e matriz de verificação de paridade

$$\begin{aligned} \mathbf{S} &= \mathbf{R} \cdot \mathbf{H} \\ &= (\mathbf{C} + \mathbf{E}) \cdot \mathbf{H} \\ &= \mathbf{C} \cdot \mathbf{H} + \mathbf{E} \cdot \mathbf{H} \end{aligned} \quad (10.93)$$

Como \mathbf{C} é a palavra de código, ela satisfaz a equação de verificação de paridade e $\mathbf{C} \cdot \mathbf{H} = 0$. Portanto,

$$\mathbf{S} = \mathbf{E} \cdot \mathbf{H} \quad (10.94)$$

O vetor \mathbf{S} de dimensão $(n - k)$ é chamado de *síndrome do erro*. Suponha que \mathbf{E} corresponde a um único erro de bit no m -ésimo bit da palavra de código. A síndrome \mathbf{S} irá, então, corresponder a m -ésima linha da matriz de verificação de paridade \mathbf{H} . Note que as linhas de \mathbf{H} são únicas. Portanto, podemos utilizar a síndrome para determinar a posição de qualquer erro único de bit. Este mapeamento é mostrado na Tabela 10.4. A síndrome de $[0, 0, 0]$ significa que nenhum erro ocorreu.

Portanto, mostramos que o código de Hamming (7,4) pode corrigir qualquer erro de um único bit. Ele pode corrigir mais do que um erro? Se \mathbf{E} contém dois erros, então a síndrome será constituída da soma de módulo 2 de duas linhas de \mathbf{H} . Esta soma irá produzir uma terceira linha em \mathbf{H} e pode, portanto, ser indistinguível do erro de um único bit que também produz esta linha. Conseqüentemente, o código de Hamming (7,4) pode corrigir todos os erros de um único bit, mas não erros de dois bits. Em outras palavras, a d_{\min} do código deve ser 3.

TABELA 10.4 Decodificação de síndrome para o código de Hamming (7, 4)

Erro (E)	Síndrome (S)	Rep. decimal
0000000	000	0
1000000	101	5
0100000	111	7
0010000	110	6
0001000	011	3
0000100	100	4
0000010	010	2
0000001	001	1

No lado da detecção de erro, novamente notamos que quaisquer dois bits errados irão produzir uma terceira linha de H. Como esta terceira linha é não nula, isto significa que podemos detectar (mas não corrigir) todos os erros de dois bits com o código de Hamming (7, 4). Note, entretanto, que se três bits errados E = [1011000] ocorrerem, então a síndrome S = 0. Portanto, o código não pode detectar todos os erros de três bits. Isto confirma ainda mais que d_{\min} é 3 para este código.

EXEMPLO 10.4 Correção de erro com o Código de Hamming

Quatro bits de informação são codificados em um código de Hamming (7, 4) cuja matriz geradora é dada pela Eq. (10.89). A seqüência de bits recebida é R = [1101110]. Qual seqüência foi transmitida se ocorreu um erro durante a transmissão?

Multiplicando R pela matriz de verificação de paridade, obtemos a síndrome do erro

$$\begin{aligned} S &= R \cdot H \\ &= [111] \end{aligned}$$

Como a síndrome do erro não é nula, sabemos que um ou mais erros ocorreram. Da Tabela 10.4, um único erro no segundo bit da esquerda teria causado esta seqüência. Logo, concluímos que a seqüência transmitida é [1001110] e os bits de informação são [1001].

EXEMPLO 10.5 Probabilidade de erro de bloco

Suponha que uma estratégia de transmissão BPSK é codificada com um código de bloco de Hamming (7,4). Compare a taxa de erro de bloco com e sem decodificação.

Da Seção 10.4, a probabilidade de erro de bit com BPSK é $P_e = Q(\sqrt{2E_b/N_0})$. Um bloco de sete bits não codificados é recebido corretamente se ele não contiver erros. A probabilidade de um erro de bloco – ou seja, um ou mais bits errados – é o complemento disto, ou seja,

$$P_b^{\text{não codificado}} = 1 - (1 - P_e)^7 \quad (10.95)$$

Com o código de Hamming (7, 4), a probabilidade de um erro de bloco é a probabilidade de dois ou mais erros em um bloco de sete bits. Esta probabilidade é dada por

$$P_b^{\text{codificado}} = 1 - (1 - P'_e)^7 - \binom{7}{1}(1 - P'_e)^6 P'_e \quad (10.96)$$

na qual P'_e é a probabilidade de um erro de bit com o código de Hamming. Se $P'_e = P_e$, então a probabilidade de um erro de bloco é claramente reduzida com o código de Hamming. Entretanto, existe uma nuance que deve ser considerada. Este é o motivo pelo qual utilizamos P'_e na Eq. (10.96) ao invés de P_e . Em particular, no modelo de referência digital, E_b se refere a *energia por bit de informação*. Com o código de Hamming (7, 4), estamos transmitindo quatro bits de informação e três bits de paridade, para um total de sete *bits de canal* por bloco. Fazendo E_c representar a energia por bit de canal, vemos que a energia de sete bits de canal corresponde a energia de quatro bits de informação:

$$E_c = \frac{4}{7} E_b \quad (10.97)$$

Além disso, a probabilidade de erro P'_e na Eq. (10.96) representa a probabilidade de um erro de bit de canal, portanto, utilizando esta probabilidade de erro para o BPSK e a Eq. (10.97), temos

$$\begin{aligned} P'_e &= Q\left(\sqrt{\frac{2E_c}{N_0}}\right) \\ &= Q\left(\sqrt{\frac{8E_b}{7N_0}}\right) \end{aligned} \quad (10.98)$$

Neste caso, podemos fazer uma comparação justa entre as estratégias de transmissão codificadas e não codificadas. As taxas de erro de bloco para estas duas estratégias são mostradas na Figura 10.19. A codificação provê uma melhora de aproximadamente 1 dB na performance de taxa de erro de bloco.

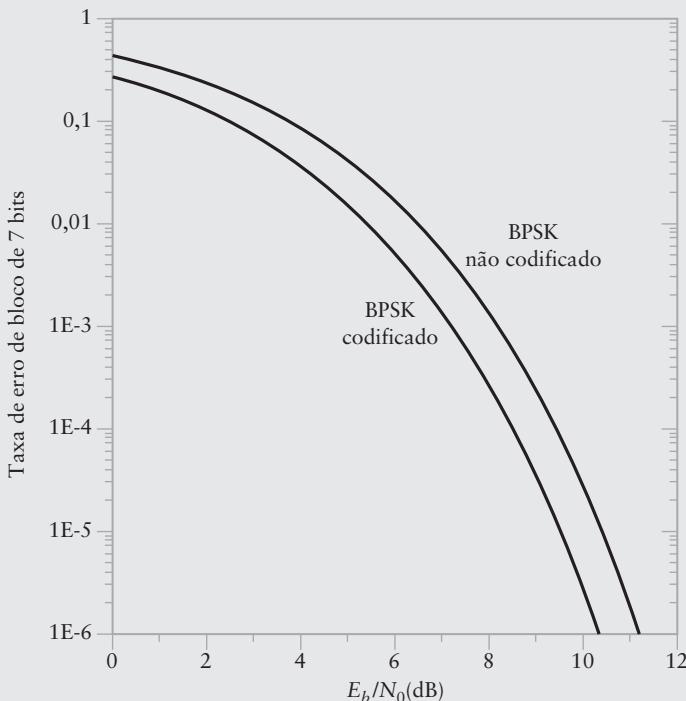


Figura 10.19 Comparação de uma taxa de erro de bloco de 7 bits de BPSK não codificado e BPSK codificado com Hamming (7, 4).

■ Códigos mais poderosos

Apesar dos códigos de Hamming serem códigos fáceis de serem explicados e muito eficientes, eles não são, de forma alguma, os únicos códigos de correção à frente. Outros códigos CEF incluem:

- ▶ Códigos de bloco de *Reed-Solomon* e *Bose-Chaudhuri-Hocquenghem* (BCH). Tal como os códigos de Hamming, eles são códigos (n, k) nos quais existem k bits de informação e um total de n bits, incluindo $n - k$ bits de paridade.⁵ Os valores de n e k não são restringidos como nos códigos de Hamming, o que possibilita maior flexibilidade.
- ▶ Códigos convolucionais. Como o nome sugere, estes códigos são o resultado da convolução de uma ou mais equações de verificação de paridade com os bits de informação. Por exemplo, se os bits de informação forem x_1, x_2, x_3, \dots , então geramos a seqüência de bits usando a relação

$$p_k = x_k \oplus x_{k-2} \oplus x_{k-3} \oplus x_{k-5} \quad (10.99)$$

Portanto, cada bit de entrada x_k resulta em dois bits x_k e p_k a serem transmitidos pelo canal. A transmissão de x_k e p_k para cada bit de informação é um exemplo de uma taxa $\frac{1}{2}$, ou seja, existem dois bits de canal para cada bit de informação. Este é um exemplo de um código *convolucional sistemático*, no qual o bit de informação x_k é explicitamente transmitido. Em códigos convolucionais não sistemáticos podem haver duas equações de paridade tais como a Eq. (10.99), e dois bits de paridade produzidos para cada bit de informação x_k são transmitidos, mas o bit de informação x_k não é transmitido. A capacidade de correção de erro de um código convolucional é determinada, geralmente, pelo atraso (span) da equação de geração de paridade, a qual é referenciada como a *memória* do código. Infelizmente, a complexidade da decodificação aumenta exponencialmente com a memória do código e, portanto, códigos práticos são geralmente limitados a memórias menores do que dez bits.

- ▶ Códigos Turbo. Códigos Turbo são códigos de bloco, mas utilizam uma estratégia de codificação similar aos códigos convolucionais. Por exemplo, com um código Turbo de taxa $\frac{1}{3}$, os três bits transmitidos para cada bit x_k de informação podem ser

$$\begin{aligned} x_k \\ p_k &= p_{k-1} \oplus x_k \oplus x_{k-2} \oplus x_{k-4} \\ q_k &= q_{k-1} \oplus x_k^i \oplus x_{k-2}^i \oplus x_{k-4}^i \end{aligned} \quad (10.100)$$

com códigos Turbo, a paridade é determinada por uma equação recursiva, ou seja, o k -ésimo bit de paridade depende do $(k - 1)$ -ésimo bit além dos bits de entrada. Note que na Eq. (10.110), a fórmula para os bits de paridade p e q são as mesmas. A diferença é que a fórmula para q utiliza x_k^i , o que se refere aos mesmos bits de informação, mas em uma ordem entrelaçada. O entrelaçamento significa a alteração da ordem dos bits de forma pseudo-aleatória.

Tanto para os códigos de bloco quanto para os códigos convolucionais, o princípio geral é adicionar bits de paridade de forma a maximizar a distância mínima d_{\min} . Entretanto, a maximização da distância mínima é sujeita a duas importantes restrições:

1. A *taxa de codificação*, ou seja, a razão entre o número de bits de informação e o número total de bits em uma palavra de código deve ser razoável. Isto garante que o sistema de comunicação possui uma significativa produção de saída de informação.
2. O código deve permitir uma estratégia de decodificação prática. É relativamente fácil construir códigos com distâncias mínimas grandes, mas as estratégias de decodificação correspondentes são geralmente muito complicadas para terem aplicação prática.

⁵ Nos códigos de Reed-Solomon, n e k em (n, k) geralmente se referem ao total de símbolos e símbolos de informação em uma palavra de código, na qual um símbolo corresponde a m bits.

Mesmo com estratégias de decodificação práticas e grandes distâncias mínimas, existem limites de como dados confiáveis podem ser transmitidos. Os limites teóricos caem na área da teoria de capacidade de canal de Shannon. A teoria de Shannon⁶ nos diz que *se a taxa de transmissão está abaixo da capacidade do canal, então existe um código de correção a frente que permite a transmissão sem erros*. Enormes esforços na aproximação destes limites teóricos têm sido realizados nos anos recentes com a descoberta dos códigos Turbo. Códigos Turbo possuem características de códigos convolucionais e códigos de bloco. Apesar de d_{\min} ser importante no projeto de códigos Turbo, outras características destes códigos melhoram a performance em razões sinal/ruído baixas.

■ Interpretação por espaço de sinal de códigos de correção de erro à frente

Na Seção 10.8, vimos como podemos construir um vetor de quatro dimensões com a representação de espaço de sinal de

$$\begin{aligned} \mathbf{s} &= [\pm \sqrt{E_b}, \pm \sqrt{E_b}, \pm \sqrt{E_b}, \pm \sqrt{E_b}] \\ &= \sqrt{E_b}[\pm 1, \pm 1, \pm 1, \pm 1] \end{aligned} \quad (10.101)$$

no qual existem dezesseis pontos na constelação definida por \mathbf{s} . Para sermos específicos, consideramos a construção de um vetor de sinal de dimensão N no contexto de códigos de correção de erro.

Para um código linear, definimos a palavra de código \mathbf{c} de tamanho N composta por bits de informação e bits de paridade. Em geral, para um vetor binário de tamanho N , sempre existem 2^N possíveis valores. Entretanto, em uma palavra de código, a relação especial entre bits de informação e bits de paridade implica que existirão somente 2^k possíveis valores para \mathbf{c} , na qual k é o número de bits de informação e $N - k$ é o número de bits de paridade.

Para a transmissão, devemos associar valores reais para cada palavra de código. Seja $\mathbf{s} = [s_1, s_2, s_3, \dots, s_N]$ uma palavra de código de valor real na qual assumimos que s_i assume um valor de $+\sqrt{E_b}$ ou $-\sqrt{E_b}$, dependendo se o símbolo de dado é 0 ou 1. Podemos considerar \mathbf{s} como sendo um vetor no espaço de sinal. Análogo à palavra de código \mathbf{c} , existem apenas 2^k possíveis valores para o vetor no espaço de sinal em um espaço de dimensão N .

Com esta nova definição de espaço de sinal, podemos explicar geometricamente alguns dos conceitos da teoria de codificação, como mostrado na Fig. 10.20. Em particular, para um código linear, definimos a distância de Hamming como o número de posições nas quais as palavras do código binário diferem. Em termos geométricos, se \mathbf{c}_1 e \mathbf{c}_2 são duas palavras de código com distância de Hamming mínima, então elas podem ser ilustradas, em um sentido bidimensional, como mostrado na Fig. 10.20(a). Os círculos de raio $t = [(d_{\min} - 1)/2]$, representando o menor inteiro menor do que $d_{\min}/2$, engloba os vetores binários que podem ser corretamente decodificados para as palavras de código em seus respectivos centros. Na realidade, estes círculos representam esferas N -dimensionais.

No espaço de sinal, a distância geométrica quadrada (Euclidiana) entre dois vetores vizinhos no espaço de sinal é definida pela distância Euclidiana quadrática

$$\begin{aligned} d^2 &= \|\mathbf{s}_1 - \mathbf{s}_2\|^2 \\ &= \sum_{i=1}^N (s_{1i} - s_{2i})^2 \\ &= d_{\min}(2\sqrt{E_b})^2 \\ &= 4d_{\min}E_b \end{aligned} \quad (10.102)$$

⁶ Para uma discussão da teoria de Shannon e suas implicações, veja o livro de Cover e Thomas (1991).

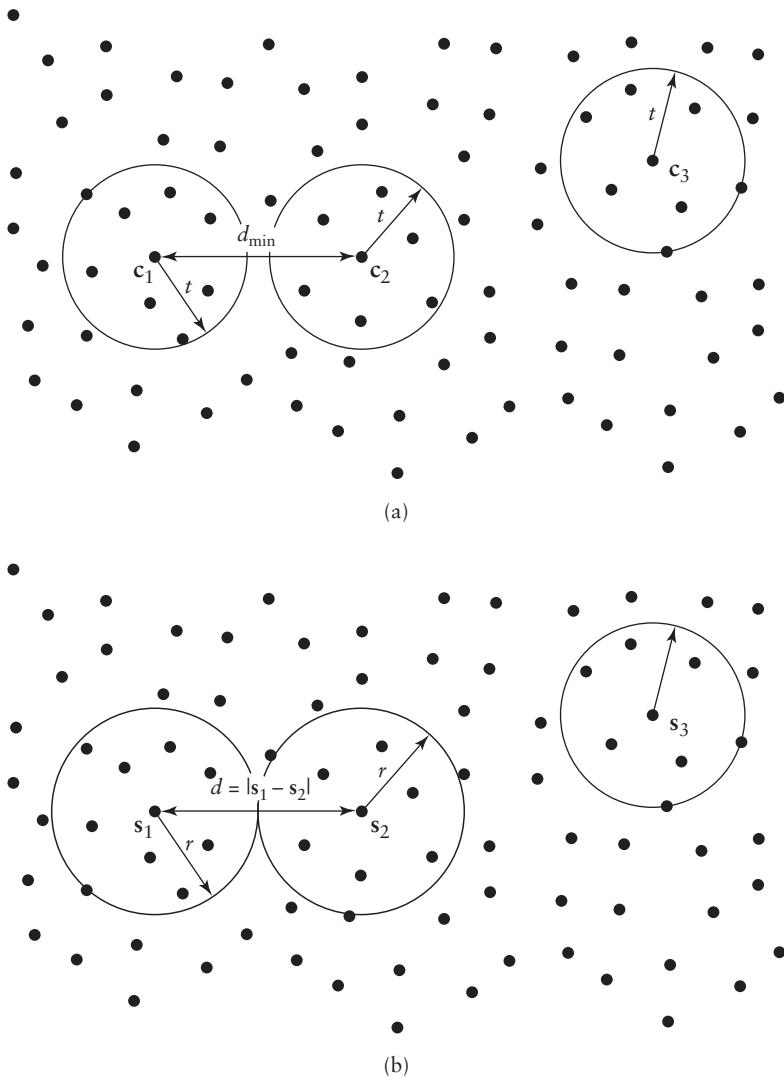


Figura 10.20 Ilustração da distância mínima em (a) espaço de palavra de código com $t = [(d_{\min} - 1)/2]$; e (b) espaço de sinal com $d = 2\sqrt{d_{\min}E_b}$ e $r = \sqrt{d_{\min}E_b}$.

na qual utilizamos o fato de dois vetores vizinhos no espaço de sinal, s_1 e s_2 , diferirem d_{\min} elementos se eles representam palavras de código. Portanto, a distância de Hamming entre palavras de código e a distância geométrica entre vetores correspondentes no espaço de sinal é diretamente relacionada. Em espaço de código, podemos selecionar duas palavras com separação mínima entre elas e desenhar esferas de raio t ao redor delas. Como o código possui distância mínima d_{\min} , isto implica em não existir nenhuma palavra de código dentro desta esfera. Qualquer ponto dentro desta esfera estará mais próximo de seu centro do que de qualquer outra palavra de código. Similarmente, no espaço de sinal, podemos desenhar esferas de raio $r = \sqrt{d_{\min}E_b}$ ao redor de cada vetor no espaço de sinal representando uma palavra de código, como mostrado na Fig. 10.20(b) e não existirá nenhum outro vetor no espaço de sinal representando palavras de código dentro desta esfera.

Esta análise está em conformação com nossa definição de capacidade de correção de um código: Qualquer sinal recebido que estiver dentro de uma distância $d_{\min}/2$ do código transmitido pode ser unicamente corrigido para aquela palavra de código. Sinais recebidos que estão além desta distância podem ser erroneamente corrigidos se caírem em outras esferas ou podem ser corretamente corrigidos se caírem nos espaços entre esferas e estiverem

mais próximos da palavra de código transmitida do que de qualquer outra. A capacidade de detecção de erro significa que o sinal recebido não cai diretamente em qualquer outra palavra de código.

O código de Hamming ($7, 4$) é um exemplo que possui 2^4 palavras de códigos (pontos) em um espaço de sinal de dimensão 7 e distância mínima entre duas palavras de código igual a três. Similarmente, o código de Hamming ($15, 11$) é um exemplo que possui 2^{11} palavras de código (pontos) em um espaço de dimensão 15 e, novamente, distância mínima igual a três.

10.10 RESUMO E DISCUSSÃO

Neste capítulo, investigamos a detecção de sinais digitalmente modulados com ruído. Começamos estabelecendo a taxa de erro de bit (TEB) como figura de mérito para a comparação de sistemas de comunicação digital e observamos que TEB é tipicamente uma função não linear da razão sinal/ruído. Depois, trabalhamos com a detecção de vários esquemas de modulação digital com a presença de ruído e fizemos as seguintes observações:

- (i) A análise da detecção de um único pulso com ruído mostra a filtragem casada ótima. O princípio de filtro casado é o degrau principal para as técnicas digitais de detecção subsequentes. Mostramos que a performance da taxa de erro de bit utilizando a filtragem casada é muito relacionada com a função Q definida no Capítulo 8.
- (ii) Mostramos como o princípio de filtragem casada pode ser estendido para a detecção da modulação por amplitude de pulso e que a performance da taxa de erro de bit pode ser determinada de maneira similar à da detecção de um único pulso.
- (iii) Discutimos como a estrutura do receptor para a detecção coerente de estratégias de modulação passa-faixa tais como BPSK, QPSK e QAM é similar à detecção coerente de sinais analógicos correspondentes. Após a conversão para baixo para banda base, o princípio da filtragem casada também se aplica a estas estratégias de modulação e a performance da TEB equivalente aos sistemas banda base correspondentes é, portanto, obtida.
- (iv) Mostramos como esquemas de modulação em quadratura, tais como QPSK e QAM, fornecem a mesma performance de suas contrapartidas unidimensionais BPSK e PAM, devido a ortogonalidade das componentes em fase e quadratura dos sinais modulados. Em particular, mostramos como a modulação QPSK fornece a mesma TEB da BPSK para o mesmo E_b/N_0 , as possibilita o dobro de throughput (produção de saída).
- (v) Também mostramos que estratégias antipodais tais como BPSK são mais eficientes em termos de potência do que estratégias de transmissão ortogonais, tais como a sinalização liga-desliga em FSK.
- (vi) Introduzimos o conceito de detecção não coerente quando mostramos que o BPSK pode ser detectado utilizando uma nova abordagem na qual os bits transmitidos são diferencialmente codificados. A simplicidade desta técnica de detecção resultou em uma pequena penalidade na performance da TEB.

Os conceitos de espaço de sinal introduzidos no Capítulo 7 possibilitam uma interpretação geométrica intuitiva da performance relativa de diferentes estratégias de modulação digital coerentes com ruído. Finalmente, fechamos o capítulo após uma breve introdução a códigos de correção de erro à frente, os quais podem ser combinados com qualquer uma das estratégias de modulação digital mencionadas anteriormente para reduzir a potência necessária para atingir a mesma performance de taxa de erro de bit – entretanto, existe geralmente um compromisso entre potência e largura de faixa que deve ser satisfeita quando utilizamos códigos de correção de erro à frente no projeto de sistema de comunicação digital.

Problemas adicionais

- 10.11 Um sistema de comunicação que transmite pulsos isolados e únicos é sujeito ao multicaminho, tal que se o pulso transmitido é $p(t)$ de duração T , o sinal recebido é

$$s(t) = p(t) + \alpha p(t - \tau)$$

Assumindo que α e τ são conhecidos, determine o filtro ótimo no receptor para o sinal com a presença de ruído branco Gaussiano de densidade espectral de potência $N_0/2$. Qual é a RSR pós-deteção na saída deste filtro?

- 10.12 A resposta ao impulso correspondente ao espectro de raiz de cosseno levantado, normalizado para satisfazer a Eq. (10.10) é dada por

$$g(t) = \left(\frac{4\alpha}{\pi}\right) \frac{\cos\left[\frac{(1+\alpha)\pi t}{T}\right] + \frac{T}{4\alpha t} \sin\left[\frac{(1-\alpha)\pi t}{T}\right]}{1 - \left(\frac{4\alpha t}{T}\right)^2}$$

na qual $T = 1/(2B_0)$ é o período do símbolo e α é o fator de roll-off. Obtenha uma representação em tempo discreto desta resposta ao impulso amostrando-a em $t = 0,1nT$ para o inteiro n tal que $-3T < t < 3T$. Aproxime numericamente a filtragem casada (por exemplo, com MATLAB) executando a convolução em tempo discreto

$$q_k = 0,1 \sum_{n=-60}^{60} g_n g_{k-n}$$

na qual $g_n = g(0,1nT)$. Qual é o valor de $q_k = q(0,1kT)$ para $k = \pm 20, \pm 10$ e 0?

- 10.13 Determine a função de autocorrelação em tempo discreto da seqüência de ruído $\{N_k\}$ definida pela Eq. (10.34)

$$N_k = \int_{-\infty}^{\infty} p(kT - t)w(t)dt$$

na qual $w(t)$ é o processo de ruído Gaussiano e o pulso $p(t)$ corresponde ao espectro da raiz de cosseno levantado. Como as amostras de ruído correspondente a intervalos de bit adjacentes se relacionam?

- 10.14 Desenhe uma constelação codificada por Gray (diagrama de espaço de sinal) para 16-QAM e 64-QAM. Você pode sugerir uma constelação para 32-QAM?

- 10.15 Escreva uma equação de definição para um sinal modulado QAM. Baseado na discussão de QPSK e PAM multinível, desenhe o diagrama em blocos para o receptor QAM coerente.

- 10.16 Mostre que se T é um múltiplo do período de f_c , então os termos $\sin(2\pi f_c t)$ e $\cos(2\pi f_c t)$ são ortogonais no intervalo $[t_0, T + t_0]$.

- 10.17 Para uma forma de pulso retangular, por quanto a largura de faixa de transmissão de nulo-a-nulo aumenta se a taxa de transmissão aumenta por um fator de três?

- 10.18 Com a consideração bandapassante, determine as condições para as quais dois sinais $\cos(2\pi f_0 t)$ e $\cos(2\pi f_1 t)$ serão ortogonais no intervalo de 0 a T .

- 10.19 Codifique a seqüência 1101 com um código de bloco de Hamming (7, 4).

- 10.20 A seqüência codificada 1001000, usando Hamming (7, 4) foi recebida. Se o número de erros de transmissão é menor do que dois, qual é a seqüência transmitida?

- 10.21 Um código de bloco de Hamming (15, 11) é aplicado ao esquema de transmissão BPSK. Compare a performance da taxa de erro de blocos dos sistemas não codificado e codificado. Explique o que seria diferente se a estratégia de modulação fosse QPSK.

Problemas avançados

- 10.22 Mostre que a escolha $\gamma = \mu/2$ minimiza a probabilidade de erro dada pela Eq. (10.26).
Dica: A função Q é continuamente diferenciável.

10.23 Para um PAM M -ário:

- (a) Mostre que a fórmula para probabilidade de erro, ou seja,

$$P_e = 2 \left(\frac{M-1}{M} \right) Q\left(\frac{A}{\sigma}\right)$$

é válida para $M = 2, 3$ e 4 . Por indução matemática, mostre que ela é válida para todo M .

- (b) Mostre que a fórmula para potência média,

$$P = \frac{(M^2 - 1)A^2}{3}$$

é válida para $M = 2$ e 3 . Mostre que ela é válida para todo M .

10.24 Considere a transmissão FSK binária na qual $(f_1 - f_2)T$ não é um inteiro.

- (a) Qual é a saída média do correlador superior da Fig. 10.12 se um 1 é transmitido? Qual é a saída média para o correlador inferior?
- (b) As variáveis aleatórias N_1 e N_2 são independentes nestas condições? Qual é a variância de $N_1 - N_2$?
- (c) Descreva as propriedades da variável aleatória D da Fig. 10.12, neste caso.

10.25 Mostre que a variância do ruído da componente em fase $n_I(t)$ do ruído passa-faixa é a mesma variância do ruído passa-baixa $n(t)$. Ou seja, para o ruído passa-faixa de largura de faixa B_N , temos

$$\mathbb{E}[n_I^2(t)] = N_0 B_N$$

10.26 Neste problema, investigamos os efeitos quanto os filtros de transmissão e recepção não combinam para formar um pulso livre de IIS. Para sermos específicos, dados são transmitidos em banda base usando PAM binário com um formato de pulso exponencial $g(t) = \exp(t/T)/u(t)$ na qual T é o período do símbolo (veja o Exemplo 2.2). O receptor detecta o dado usando um detector integra-e-descarta.

- (a) Com os símbolos de dado representados por ± 1 , qual é a amplitude da componente do sinal na saída do detector?
- (b) Qual é o pior caso para a amplitude da interferência intersimbólica na saída do detector. (Assuma que a seqüência de dados possui tamanho infinito). Usando o valor obtido na parte (a) como referência, qual é porcentagem pela qual a abertura do olho é reduzida por essa interferência?
- (c) Qual é a amplitude rms da interferência intersimbólica na saída do detector? Se esta interferência for tratada como equivalente a ruído, qual é a razão sinal/ruído equivalente na saída do detector? Comente sobre como isto afetará a performance da taxa de erro de bit deste sistema quando também existir ruído recebido.

10.27 Um sinal BPSK é aplicado a um receptor de filtro casado no qual falta a sincronização perfeita de fase com o transmissor. Especificamente, ele é alimentado por uma portadora local cuja fase difere da fase da portadora utilizada na transmissão por ϕ radianos.

- (a) Determine o efeito do erro de fase ϕ na probabilidade média de erro neste receptor.
- (b) Como verificação da fórmula obtida na parte (a), mostre que quando o erro de fase é zero, a fórmula se reduz para a mesma forma da Eq. (10.44).

10.28 Um sistema FSK binário transmite dados a uma taxa de 2,5 vezes megabits por segundos. Durante o percurso de transmissão, ruído branco Gaussiano de média nula e densidade espectral de potência 10^{-20} watts por hertz é adicionado ao sinal. Na ausência do ruído, a amplitude do sinal recebido é $1 \mu\text{V}$ em uma impedância de 50 ohms. Determine a probabilidade média de erro assumindo a detecção coerente do sinal FSK binário.

10.29 Uma das formas mais simples de código de correção de erro à frente é a repetição do código. Com um código de repetição N , o mesmo bit é enviado N vezes e o decodificador decide em favor do bit que for detectado na maioria dos testes (assumindo N é ímpar). Para um esquema de transmissão BPSK, determine a performance da TEB para um código de repetição 3.

■ Experimentos de computador

- 10.30 Neste experimento, simulamos a performance da sinalização bipolar com ruído branco Gaussiano aditivo. O script MATLAB incluído no Apêndice 7 para este experimento executa as seguintes tarefas:
- ▶ Gera uma seqüência aleatória com formato de pulso retangular.
 - ▶ Adiciona ruído Gaussiano.
 - ▶ Detecta o dado com um detector integra-e-descarta simulado.
- Com este script MATLAB:
- (a) Calcule o espectro do sinal transmitido e compare com o teórico.
 - (b) Explique o cálculo da variância dada uma razão E_b/N_0 .
 - (c) Confirme a taxa de erro de bit teoricamente predita para E_b/N_0 de 0 a 10 dB.
- 10.31 Neste experimento, simulamos a performance da sinalização bipolar com ruído branco Gaussiano aditivo, mas com o formato de pulso de raiz de cosseno levantado. Um script MATLAB é incluído no Apêndice 7 para esta simulação. Logo,
- (a) Calcule o espectro do sinal transmitido e compare com o teórico. Compare, também, com o espectro transmitido com formato de pulso retangular.
 - (b) Trace o diagrama de olho para o sinal recebido com a condição sem ruído. Explique a relação da abertura de olho com a performance da taxa de erro de bit.
 - (c) Confirme a taxa de erro de bit teoricamente predita para E_b/N_0 de 0 a 10 dB.
- 10.32 Neste experimento, simulamos o efeito de vários não-casamentos em sistemas de comunicação e seus efeitos na performance. Em particular, modifique o script MATLAB dos dois problemas anteriores para fazer o seguinte:
- (a) Simule a performance do sistema usando um formato de pulso retangular no transmissor e um formato de pulso de raiz de cosseno levantado no receptor. Comente sobre a degradação da performance.
 - (b) No caso de filtragem casada por raiz de cosseno levantado, inclua uma rotação de fase complexa $\exp(j\theta)$ no canal. Trace o diagrama de olho resultante para θ equivalente a 5° , 10° , 20° e 45° . Compare com o caso de $\theta = 0^\circ$. Faça o mesmo para a performance da TEB. Qual modificação à fórmula teórica da TEB poderia modelar adequadamente este comportamento?

CÁLCULOS DE SISTEMAS E RUÍDOS

A presença inevitável de ruído degrada a transmissão confiável de sinais através de um sistema de comunicação. Nos capítulos anteriores, modelamos o ruído como branco, Gaussiano e aditivo. Este modelo foi matematicamente tratado e nos permitiu obter vários resultados analíticos sobre a performance de comunicações digitais e analógicas com ruído. Nesses capítulos, afirmamos que o modelo branco Gaussiano era um modelo razoável da realidade. Agora, iremos considerar, mais detalhadamente as fontes potenciais de ruído e como eles aparecem. As fontes do ruído podem ser externas ao sistema (por exemplo, ruído atmosférico, ruído galáctico, ruído humano) ou internas ao sistema. A segunda categoria inclui um tipo importante de ruído que aparece devido a flutuações espontâneas da corrente ou tensão em circuitos elétricos. Este tipo de ruído, de uma forma ou outra, está presente em todo sistema de comunicação e representa uma limitação fundamental para a transmissão e detecção de sinais.

Neste capítulo, discutimos brevemente as fontes físicas de ruído em circuitos elétricos e desenvolvemos modelos quantitativos para a medição e predição da presença de ruído em um sistema de comunicação. Então, consideramos o efeito da propagação na potência recebida de ondas de rádio em um ambiente terrestre móvel. Quando os dois conceitos de nível de sinal recebido e nível de ruído de sinal são combinados, temos a razão sinal/ruído, a qual representa a figura de mérito fundamental para a qualidade da transmissão através de um canal de comunicação.

Este capítulo nos ensina as seguintes lições:

- ▶ *Lição 1: Ruído em sistemas de comunicação podem ser gerados por várias fontes, mas geralmente as fontes são os próprios dispositivos de comunicação. O ruído térmico e o shot noise são exemplos de processos de ruído branco gerados por circuitos elétricos.*
- ▶ *Lição 2: Em um ambiente de espaço livre, a potência do sinal recebido é atenuada proporcionalmente ao quadrado da distância da transmissão. Entretanto, a potência do sinal pode ser aumentada por antenas direcionais tanto no transmissor quanto no receptor.*
- ▶ *Lição 3: Em um ambiente terrestre, a comunicação via rádio pode ocorrer em vários caminhos. A interferência construtiva e destrutiva entre os diferentes caminhos leva, em geral, ao chamado fenômeno multcaminho, o qual causa perdas de propagação muito maiores do que o predito pelo modelo de espaço livre. Além disso, o movimento do terminal transmissor ou do terminal receptor resulta em mais variações da potência do sinal recebido.*

11.1 RUÍDO ELÉTRICO

Em um circuito elétrico, o ruído é gerado por vários fenômenos físicos.¹ Temos o ruído térmico produzido pelo movimento aleatório de elétrons em uma mídia de condição e o chamado *shot noise* produzido por flutuações aleatórias do fluxo de corrente em dispositivos eletrônicos. Estes dois tipos de ruído são os exemplos mais comuns de ruído de flutuação espontânea encontrados em circuitos elétricos.

Além do ruído térmico e do *shot noise* (ou ruído *shot*), transistores exibem um fenômeno de ruído de baixa freqüência chamado de *ruído flicker*. O valor médio quadrático

¹ Para um tratamento detalhado sobre ruído elétrico, veja Bell (1985), Van Der Ziel (1970) e Robinson (1974).

do ruído flicker é inversamente proporcional à freqüência, por isso ele também é chamado como ruído “um-por- f ”. Outro tipo de ruído encontrado em dispositivos semicondutores é o ruído *burst*, cujo valor médio quadrático tende a $1/f^2$.

O ruído flicker e burst são não brancos, com efeitos degradantes observados em baixas freqüências. Ordinariamente, eles podem ser ignorados acima de poucos quilohertz. Por outro lado, o ruído térmico e o ruído shot são brancos em todos os propósitos práticos. Logo, as influências degradantes destes ruídos na operação de um sistema de comunicação se estendem em toda a faixa de freqüência de interesse. Uma breve discussão sobre ruído térmico e ruído shot é, portanto, necessária.

Ruído térmico

O ruído térmico é uma fonte constantemente encontrada de ruído oriunda do movimento termicamente induzido de elétrons em um meio de condução. Em um condutor, existe um grande número de elétrons livres e um número igualmente grande de íons limitados por fortes forças moleculares. Os íons viram aleatoriamente ao redor de suas posições normais. Os elétrons livres se movem ao longo de caminhos orientados aleatoriamente, devido a colisões contínuas com os íons vibrantes. O efeito líquido do movimento aleatório é uma corrente elétrica que é do tipo aleatório. O valor médio da corrente é zero, pois, na média, existem tantos elétrons se movendo em uma direção quanto elétrons se movendo na outra.

Uma análise completa do ruído térmico requer a utilização de considerações termodinâmicas e mecânicas que estão além do escopo deste livro. Para o propósito da discussão apresentada aqui, é suficiente dizer que a densidade espectral de potência do ruído térmico produzido por um resistor é dada por²

$$S_{TN}(f) = \frac{2h|f|}{\exp(h|f|/kT) - 1} \quad (11.1)$$

na qual T é a *temperatura absoluta* em Kelvin, k é a *constante de Boltzmann* e h é a *constante de Planck*. Note que a densidade espectral de potência $S_{TN}(f)$ é medida em watts por hertz. Para “baixas” freqüências definidas por

$$f \ll \frac{kT}{h} \quad (11.2)$$

podemos utilizar a aproximação

$$\exp\left(\frac{h|f|}{kT}\right) \approx 1 + \frac{h|f|}{kT} \quad (11.3)$$

Substituindo a fórmula aproximada na Eq. (11.1), obtemos a fórmula aproximada correspondente para a densidade espectral de potência do ruído térmico:

$$S_{TN}(f) \approx 2kT \quad (11.4)$$

Substituindo os valores da constante de Boltzmann, $k = 1,38 \times 10^{-23}$ watts/Hz°K, constante de Planck, $h = 6,63 \times 10^{-34}$ watt-(segundo)², vemos que na temperatura nominal de $T = T_0 = 290$ K (17°C),

$$\frac{kT}{h} = 6 \times 10^{12} \text{ Hz} \quad (11.5)$$

Este limite superior de freqüência está na região infravermelho do espectro eletromagnético que está bem abaixo das freqüências encontradas em sistemas elétricos de comunicação

² Para uma discussão sobre as questões físicas envolvidas na formulação da Eq. (11.1), e para uma análise histórica da literatura pertinente, veja Bell (1985).

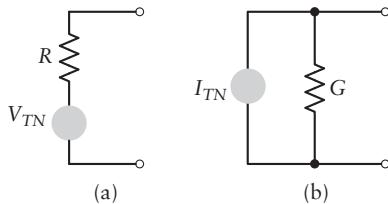


Figura 11.1 Modelos de um resistor ruidoso: (a) Circuito equivalente de Thévenin e (b) Circuito equivalente de Norton.

convencionais. Portanto, para todos os propósitos práticos, a utilização da fórmula da Eq. (11.4) é justificada.

Portanto, dado um resistor de R ohms, vemos da Eq. (11.4) que o valor médio quadrático da tensão do ruído térmico medida nos terminais do resistor é igual a

$$\begin{aligned}\mathbb{E}[V_{TN}^2] &= 2RB_N S_{TN}(f) \\ &= 4kTRB_N \text{ volts}^2\end{aligned}\quad (11.6)$$

na qual B_N é a largura de faixa (em hertz) na qual a tensão do ruído é medida. O fator 2 na primeira linha da Eq. (11.6) vem do fato da densidade espectral possuir tanto freqüências positivas quanto negativas. Podemos, pois, modelar o resistor com ruído pelo *circuito equivalente de Thévenin* constituído por uma fonte de tensão de ruído com valor médio quadrático de $\mathbb{E}[V_{TN}^2]$ em série com um resistor sem ruído, como mostrado na Fig. 11.1(a). Alternativamente, podemos utilizar o *circuito equivalente de Norton* constituído por um gerador de corrente de ruído em paralelo com uma condutância sem ruído $G = 1/R$, como mostrado na Fig. 11.1(b). O valor médio quadrático do gerador de corrente de ruído é

$$\begin{aligned}\mathbb{E}[I_{TN}^2] &= \frac{1}{R^2} \mathbb{E}[V_{TN}^2] \\ &= 4kTGB_N \text{ amps}^2\end{aligned}\quad (11.7)$$

Como o número de elétrons em um resistor é muito grande e seus movimentos aleatórios dentro do resistor são estatisticamente independentes um do outro, o teorema do limite central indica que o ruído térmico possui distribuição Gaussiana. Desta forma, para a faixa de freqüências encontradas em sistemas elétricos de comunicação, podemos modelar o ruído como sendo um *ruído branco Gaussiano* de média nula.

EXEMPLO 11.1 Tensão de ruído em rádio AM

O primeiro filtro de um rádio passa a faixa AM de broadcast de 535 kHz a 1605 kHz. A entrada do rádio possui uma resistência efetiva de 200 ohms. Qual é a tensão rms do ruído que devemos observar devido a esta resistência?

Da Eq. [11.6], a tensão média quadrática do ruído térmico em temperatura ambiente é

$$\begin{aligned}\mathbb{E}[V_{TN}^2] &= 4kTRB_N \\ &= 4 \times (1,38 \times 10^{-23}) \times 290 \times 300 \times (1070 \times 10^3) \\ &= 5,14 \times 10^{-12} \text{ volts}^2\end{aligned}$$

Logo, o tensão de raiz média quadrática do ruído térmico é 2,3 microvolts.

■ Potência disponível de ruído

Os cálculos de ruído envolvem a transferência de potência de uma fonte para uma carga; portanto, podemos ver que o *teorema da máxima transferência de potência* pode ser aplica-

do. Este teorema afirma que a máxima potência possível é transferida de uma fonte com resistência interna R para uma carga com resistência R_l quando $R_l = R$. Com esta condição de casamento, a potência produzida pela fonte é dividida igualmente entre a resistência interna da fonte e a resistência da carga. A potência entregue à carga é chamada de *potência disponível*. Aplicando o teorema da máxima transferência de potência ao circuito equivalente de Thévenin da Fig. 11.1(a) ou ao circuito equivalente de Norton da Fig. 11.1(b), facilmente vemos que um resistor ruidoso produz uma *potência disponível* igual a kTB_N watts.

EXEMPLO 11.2 Potência de ruído

Qual é a potência disponível de ruído produzida por um resistor de 10 quilo-ohms na largura de faixa de 0 a 10 MHz?

A potência disponível de ruído na temperatura ambiente é

$$\begin{aligned} P_N &= kTB_N \\ &= (1,38 \times 10^{-23}) \times 290 \times 10^7 \\ &= 4,0 \times 10^{-14} \text{ watts} \end{aligned}$$

A tensão de raiz média quadrática (rms) que esta potência produz em uma carga casada de 10 quilo-ohms é

$$\begin{aligned} V_{\text{rms}} &= \sqrt{P_N R} \\ &= 20,0 \text{ microvolts} \end{aligned}$$

- **PROBLEMA 11.1** Qual é a tensão rms em um resistor de 10 mega-ohms na temperatura ambiente se ela for medida em uma largura de faixa de 1 GHz? Qual é a potência disponível de ruído? ◀

Ruído shot

O ruído shot aparece em dispositivos eletrônicos devido a natureza discreta do fluxo de corrente em um dispositivo. O processo assume a existência de um fluxo médio de corrente que se manifesta na forma de elétrons fluindo de um catodo de uma placa em um tubo de vácuo, lacunas e elétrons fluindo em um dispositivo semicondutor e fôtons emitidos em fotodiodos. Apesar do número médio de partículas se movimentando através do dispositivo por unidade de tempo ser considerado constante, o processo de fluxo de corrente através do dispositivo exibe flutuações ao redor do valor médio. A maneira pela qual estas flutuações aparecem varia de um dispositivo para outro. Em um dispositivo de tubo de vácuo, as flutuações são produzidas pela emissão aleatória de elétrons do catodo. Em um dispositivo semicondutor, a causa é a difusão aleatória de elétrons ou a recombinação aleatória de elétrons e lacunas. Em um fotodiodo, é a emissão aleatória de fôtons. Em todos estes dispositivos, o mecanismo físico que controla o fluxo de corrente através do dispositivo apresenta flutuações intrínsecas ao redor de algum valor médio. O ruído shot produzido por estas flutuações é, portanto, dependente do valor médio da corrente.

Considere, por exemplo, um diodo a vácuo limitado em temperatura, mostrado na Fig. 11.2. O diodo consiste de dois eletrodos dentro de um tubo de vácuo: um catodo que é aquecido de forma a emitir elétrons e um anodo, ou placa, o qual é mantido em um potencial positivo com respeito ao catodo, de forma a atrair os elétrons. Assumimos que a diferença de potencial catodo-placa é grande o suficiente para fazer com que os elétrons termicamente emitidos pelo catodo aquecido sejam atraídos para a placa em tão alta velocidade que os efeitos de carga espacial sejam negligenciáveis. A corrente de placa é, então, determinada efetivamente pela taxa na qual os elétrons são emitidos do catodo. Considerando a corrente de placa como a soma de pulsos sucessivos de correntes, cada pulso cau-

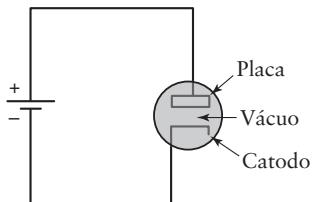


Figura 11.2 Diodo a vácuo limitado em temperatura.

sado pelo trânsito de um elétron através do espaço catodo-placa, vemos que o valor médio quadrático da componente de flutuação aleatória da corrente é dado por

$$\mathbb{E}[I_{\text{shot}}^2] = 2qIB_N \text{ amps}^2 \quad (11.8)$$

na qual q é a carga do elétron, igual a $1,60 \times 10^{-19}$ Coulombs, I é o valor médio da corrente em amperes e B_N é a largura de faixa do instrumento de medição, em hertz. A Eq. (11.8) é chamada de *fórmula de Schottky*. O tempo de trânsito típico de um elétron de um catodo para a placa é da ordem de 10^{-9} segundos. A fórmula de Schottky é válida desde que a frequência de operação seja pequena comparada com o recíproco do tempo de transmissão, de tal forma que podemos negligenciar os efeitos do tempo de trânsito.

Outra característica importante do ruído shot é que ele possui uma distribuição Gaussiana com média zero. Esta característica vem do fato do número de elétrons que contribui para a corrente de ruído shot ser muito grande e suas emissões aleatórias do catodo serem, para propósitos práticos, estatisticamente independentes uma da outra. Logo, o teorema do limite central prevê uma distribuição Gaussiana para o ruído shot.

A fórmula de Schottky também é válida para a junção de um diodo semicondutor. Neste caso, o valor médio da corrente é dado pela *equação de diodo*

$$I = I_s \exp\left(\frac{qV}{kT}\right) - I_s \quad (11.9)$$

na qual V é a tensão aplicada no diodo e I_s é a corrente de saturação. As outras constantes já foram previamente definidas. As duas componentes da corrente I produzem contribuições de ruído shot estatisticamente independentes uma da outra, como mostrado por

$$\begin{aligned} \mathbb{E}[I_{\text{shot}}^2] &= 2qI_s \exp\left(\frac{qV}{kT}\right) B_N + 2qI_s B_N \\ &= 2q(I + 2I_s)B_N \end{aligned} \quad (11.10)$$

A Fig. 11.3 mostra o modelo de ruído shot para um diodo de junção.³ O modelo inclui a resistência dinâmica do diodo, definida por

$$r = \frac{\delta V}{\delta I} = \frac{kT}{q(I + I_s)}$$

Note que, ao contrário de um resistor ordinário, a resistência dinâmica r é sem ruído, pois ela não envolve a dissipação de potência.

Em um transistor de junção bipolar, o ruído shot é gerado tanto nas junções de emissor quanto de coletores. Por outro lado, em transistores de efeito de campo, a utilização da estrutura de porta isolada evita o ruído shot de junção. Mesmo assim, o ruído shot é produzido pelo fluxo de corrente de porta. Obviamente, nos dois dispositivos, o ruído térmico aparece em função da resistência ôhmica interna: *resistência de base* em um transistor bipolar e *resistência de canal* em um transistor de efeito de campo.

³ Para detalhes de modelos de ruído para transistores de dispositivos semicondutores, veja Robinson (1974), pp. 93-116.

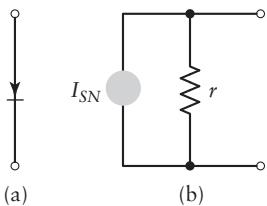


Figura 11.3 (a) Diodo de junção. (b) Modelo de ruído shot.

► **PROBLEMA 11.2** Qual é a potência disponível de ruído em 1 MHz devido ao ruído shot de um diodo de junção que possui uma tensão diferencial de 0,7 volts e conduz uma corrente média de 0,1 miliamperes? Assuma que a fonte de corrente do circuito equivalente de Norton da Fig. 11.3(b) possua uma resistência de 250 ohms. ◀

11.2 FIGURA DE RUÍDO

Uma medida conveniente da performance de ruído de um dispositivo linear de duas portas é obtida pela chamada *figura de ruído*. Considere um dispositivo linear de duas portas conectada a uma fonte de sinal com impedância interna $Z(f) = R(f) + jX(f)$, como mostrado na Fig. 11.4. A tensão de ruído $w(t)$ representa o ruído térmico associado com a resistência interna $R(f)$ da fonte. O ruído de saída do dispositivo de duas portas é composto por duas contribuições, uma devida a fonte de sinal e outra devido ao próprio dispositivo. Definimos a potência disponível de ruído em uma faixa de largura Δf centrada na freqüência f como a potência média máxima do ruído nesta faixa que pode ser obtida na saída do dispositivo. A potência máxima de ruído que o dispositivo de duas portas pode entregar a uma carga externa é obtida quando a impedância da carta é o conjugado complexo da impedância de saída do dispositivo. Ou seja, quando a resistência é casada e a reatância é removida. Definimos a *figura de ruído do dispositivo de duas portas como a razão da potência total de ruído de saída disponível (devido ao dispositivo e à fonte) por unidade de largura de faixa da porção devida à fonte*.

Seja a densidade espectral da potência total disponível de ruído da saída do dispositivo representada por $S_{NO}(f)$ e a densidade espectral da potência total de ruído devido à fonte na entrada do dispositivo representada por $S_{NS}(f)$. Seja $G(f)$ o ganho de potência disponível do dispositivo de duas portas, definido como a razão da potência do sinal disponível na saída do dispositivo pela potência do sinal disponível da fonte quando o sinal é uma onda senoidal com freqüência f . Então, podemos expressar a figura de ruído F do dispositivo por

$$F(f) = \frac{S_{NO}(f)}{G(f)S_{NS}(f)} \quad (11.11)$$

Se o dispositivo fosse livre de ruído, $S_{NO}(f) = G(f)S_{NS}(f)$ e a figura de ruído seria, então, unitária. Em um dispositivo físico, entretanto, $S_{NO}(f)$ é maior do que $G(f)S_{NS}(f)$, de tal forma que a figura de ruído é sempre maior do que a unidade. A figura de ruído geralmente é expressa em decibels, ou seja, $10 \log_{10}(F)$.

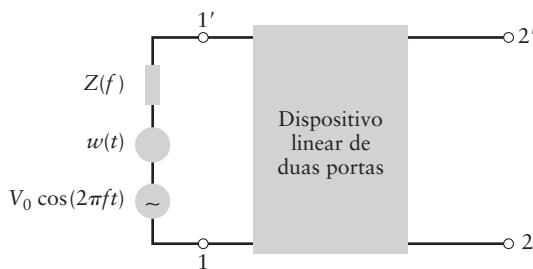


Figura 11.4 Dispositivo linear de duas portas.

A figura de ruído F da Eq. (11.11) é uma função da freqüência de operação f , sendo, portanto, geralmente chamada de *figura de ruído do ponto*. Em contraste, podemos definir a *figura de ruído médio* F_0 de um dispositivo de duas portas como a razão da potência de ruído total na saída do dispositivo pela potência de ruído de saída devido somente a fonte; ou seja,

$$F_0 = \frac{\int_{-\infty}^{\infty} S_{NO}(f) df}{\int_{-\infty}^{\infty} G(f)S_{NS}(f) df} \quad (11.12)$$

É evidente que, no caso do ruído térmico no circuito de entrada corresponder a uma resistência constante $R(f)$ e do dispositivo possuir um ganho constante em toda a faixa fixa com ganho zero em todas as outras freqüências, a figura de ruído de ponto F e a figura de ruído médio F_0 serão idênticas.

EXEMPLO 11.3 Cálculo da figura de ruído

A entrada do circuito de duas portas com ganho de 10 dB e a figura de ruído constante de 8 dB é conectada a um resistor que gera uma densidade espectral de potência $S_{NS}(f) = kT_0$ na qual T_0 é a temperatura nominal. Qual é a densidade espectral de ruído na saída do circuito de duas portas?

Da Eq. (11.11) aplicada a uma freqüência arbitrária f , vemos que a densidade espectral de ruído de saída é

$$S_{NO}(f) = F(f)G(f)S_{NS}(f)$$

Expressando em decibels, temos

$$\begin{aligned} S_{NO} &= F_{dB} + G_{dB} + (kT_0)_{dBm/Hz} \\ &= 10 + 8 + (-174) \\ &= -156 \text{ dBm/Hz} \end{aligned}$$

na qual utilizamos o fato de $kT_0 = 1,38 \times 10^{-23} \times 290 \text{ watts/Hz} \approx -174 \text{ dBm/Hz}$.

11.3 TEMPERATURA EQUIVALENTE DE RUÍDO

Uma desvantagem da figura de ruído F é que, quando utilizada para comparar dispositivos de baixo ruído, os valores são todos próximos da unidade, o que torna a comparação difícil. Em tais casos, é preferível utilizar a *temperatura equivalente de ruído*. Considere um dispositivo linear de duas portas cuja resistência de entrada é casada com a resistência interna de uma fonte, como mostrado na Fig. 11.5. Neste diagrama, também incluímos o gerador de tensão de ruído associado com a resistência interna R_s da fonte. Da Eq. (11.3), o valor médio quadrático desta fonte de tensão ruidosa é $4kT_sB_N$ na qual k é a constante de Boltzmann. Logo, a potência disponível de ruído na entrada do dispositivo é

$$N_s = kTB_N \quad (11.13)$$

Seja N_d a potência de ruído contribuída pelo dispositivo de duas portas para a potência total de ruído de saída N_t . Definimos N_d por

$$N_d = GkT_eB_N \quad (11.14)$$

na qual G é o ganho de potência disponível do dispositivo e T_e é a temperatura equivalente de ruído do dispositivo. Então, temos que a potência total de ruído de saída é

$$\begin{aligned} N_t &= GN_s + N_d \\ &= Gk(T + T_e)B_N \end{aligned} \quad (11.15)$$

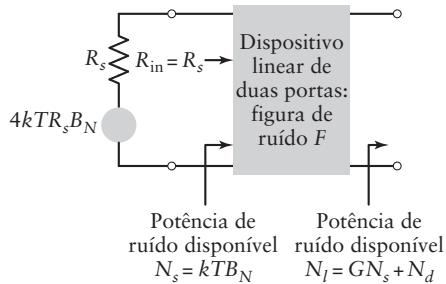


Figura 11.5 Dispositivo linear de duas portas casado com a resistência interna de uma fonte conectada na entrada.

A figura de ruído do dispositivo é, então,

$$F = \frac{N_l}{GN_s} = \frac{T + T_e}{T} \quad (11.16)$$

Resolvendo a Eq. (11.16) para a temperatura equivalente de ruído, obtemos

$$T_e = T(F - 1) \quad (11.17)$$

na qual a figura de ruído é medida com a condição de entrada casada, e a fonte de ruído é ajustada para a temperatura T . A Eq. (11.17) fornece uma relação entre a figura de ruído e a temperatura equivalente de ruído de um circuito de duas portas.

- **PROBLEMA 11.3** Um dispositivo eletrônico possui uma figura de ruído de 10 dB. Qual é a temperatura equivalente de ruído? ◀
- **PROBLEMA 11.4** O dispositivo do Problema 11.3 possui um ganho de 17dB e é conectado a um analisador de espectro. Se a entrada do dispositivo possui uma temperatura equivalente de 290 K e o analisador de espectro é sem ruído, expresse a densidade espectral de potência medida em dBm/Hz. Se o analisador de espectro possui uma figura de ruído de 25 dB, qual será a densidade espectral de potência medida neste caso? ◀

Densidade espectral de ruído

Um circuito de duas portas com temperatura equivalente de ruído T_e (referida à entrada) produz a potência disponível de ruído

$$N_{\text{méd}} = kT_eB_N \quad (11.18)$$

Percebendo que $N_{\text{méd}} = N_0B_N$, vemos que o ruído pode ser modelado como um ruído branco Gaussiano aditivo com média nula e densidade espectral de potência $N_0/2$, na qual

$$N_0 = kT_e \quad (11.19)$$

A densidade espectral de potência do ruído modelado desta forma depende apenas da constante de Boltzmann e da temperatura equivalente de ruído T_e . É a simplicidade deste modelo que faz com que a temperatura equivalente de ruído de um circuito composto seja um conceito útil.

11.4 CONEXÃO EM SÉRIE DE CIRCUITOS DE DUAS PORTAS

Geralmente é necessário calcular a figura de ruído de uma conexão em série de circuitos de duas portas cujas figuras de ruído individuais são conhecidas. Considere a Fig. 11.6, constituída por um par de circuitos de duas portas com figuras de ruído F_1 e F_2 e ganhos de potência G_1 e G_2 , conectados em série. Assume-se que os dispositivos são casados e que a figura de ruído F_2 do segundo circuito é definida assumindo uma potência de ruído de entrada N_1 .

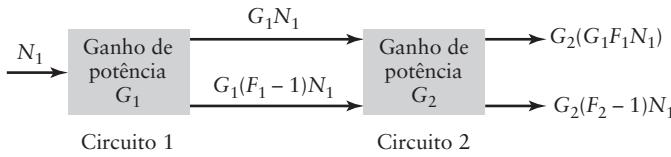


Figura 11.6 Conexão em série de dois circuitos ruidosos.

Na entrada do primeiro circuito, temos uma potência de ruído N_1 contribuída pela fonte mais uma potência de ruído equivalente $(F_1 - 1)N_1$ contribuída pelo próprio circuito. A potência de ruído de saída do primeiro circuito é, portanto, $F_1N_1G_1$. Adicionada a esta potência de ruído na entrada do segundo circuito, temos a potência extra equivalente $(F_2 - 1)N_1$ contribuída pelo próprio segundo circuito. A potência de ruído de saída desse segundo circuito é, portanto, igual a $F_1G_1N_1G_2 + (F_2 - 1)N_1G_2$. Podemos considerar a figura de ruído F como a razão da potência de ruído de saída real pela potência de ruído de saída se os circuitos forem sem ruído. Podemos, portanto, expressar a figura de ruído total da conexão em série da Fig. 11.6 como

$$\begin{aligned} F &= \frac{F_1G_1N_1G_2 + (F_2 - 1)N_1G_2}{N_1G_1G_2} \\ &= F_1 + \frac{F_2 - 1}{G_1} \end{aligned} \quad (11.20)$$

O resultado pode ser rapidamente estendido para a conexão em série de qualquer número de circuitos de duas portas, como mostrado por

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1G_2} + \frac{F_4 - 1}{G_1G_2G_3} + \dots \quad (11.21)$$

na qual F_1, F_2, F_3, \dots , são as figuras de ruído individuais e G_1, G_2, G_3, \dots , são os ganhos de potência disponível, respectivamente. A Eq. (11.21) mostra que se o primeiro estágio da conexão em série da Fig. 11.6 possui um alto ganho, a figura de ruído total F é dominada pela figura de ruído do primeiro estágio.

Correspondentemente, podemos expressar a temperatura equivalente de ruído total da conexão em série de qualquer número de circuitos ruidosos de duas portas como mostrado a seguir:

$$T_e = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1G_2} + \frac{T_4}{G_1G_2G_3} + \dots \quad (11.22)$$

na qual T_1, T_2, T_3, \dots são as temperaturas equivalentes de ruído dos circuitos individuais. Da Eq. (11.22), vemos que se o primeiro estágio da conexão em série dos circuitos ruidosos de duas portas possuir um ganho de potência grande, então o valor total da temperatura equivalente de ruído T_e é praticamente o mesmo da temperatura equivalente de ruído do primeiro estágio. Por esta razão, em um receptor de baixo ruído, tomamos cuidado extra quando projetamos o pré-amplificador ou o amplificador de baixo ruído no início do receptor.

EXEMPLO 11.4 Temperatura de ruído do sistema em um receptor de satélite

Um receptor de satélite consiste de uma antena que possui temperatura de ruído de 50K. Ela é seguida por um amplificador de baixo ruído com ganho de 30 dB e temperatura de ruído de 70K. O amplificador de baixo ruído, por sua vez, é seguido por um segundo estágio de amplificação com ganho de 40dB e temperatura de ruído de 1500K. Qual é a temperatura de ruído do sistema?

Para este sistema de três estágios, a Eq. (11.22) mostra que a temperatura de ruído do sistema é dada por

$$T_e = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1G_2}$$

na qual T_1 , T_2 e T_3 são as temperaturas de ruído da antena, amplificador de baixo ruído e segundo estágio de amplificação, respectivamente. Como a antena não fornece ganho elétrico algum ao sinal recebido, $G_1 = 1$. Conseqüentemente, a temperatura de ruído efetiva é

$$\begin{aligned} T_e &= 50 + \frac{70}{1} + \frac{1500}{1 \times 1000} \\ &= 121,5 \text{ K} \end{aligned}$$

- **PROBLEMA 11.5** Um receptor de televisão é constituído por uma antena com temperatura de ruído de 290K e um pré-amplificador com ganho de potência de 20 dB e uma figura de ruído de 9 dB. Um segundo estágio de amplificação no receptor fornece outros 20 dB de ganho e possui uma figura de ruído de 20 dB. Qual é a figura de ruído total do sistema? ◀

11.5 CÁLCULOS DE CONEXÃO EM ESPAÇO LIVRE

Nesta seção, iremos estudar a questão de cálculos de potências de ruído e sinal para canais de rádio que se baseiam na propagação em visada direta (ou linha direta) através do espaço. Tais cálculos são encontrados em comunicações via satélite,⁴ por exemplo. Em um sistema via satélite, o sinal de mensagem é transmitido de uma estação terrestre pelo uplink para um satélite geoestacionário, amplificado em um transponder e, então, retransmitido do satélite via downlink para outra estação terrestre. O satélite é posicionado em uma órbita geoestacionária (ao redor da terra) de tal forma que ele está sempre visível para diferentes estações terrestres localizadas dentro da zona de cobertura da antena do satélite na superfície da terra. O satélite, na realidade, funciona como um repetidor no céu. Outra aplicação com características de visada direta é a comunicação em espaço profundo⁵ de informação entre uma espaçonave e uma estação terrestre. Nesta segunda aplicação, o sistema inclui uma capacidade de rastreamento de tal forma que a espaçonave está sempre visível para uma dentre várias estações terrestres. Para a análise apresentada aqui, iremos considerar o link de satélite como ilustração de aplicações espaciais.

■ Cálculo da potência do sinal recebido

A Fig. 11.7 ilustra um downlink (canal de descida) entre uma espaçonave e uma estação terrestre. O canal inclui uma fonte de transmissão (na espaçonave) com sua saída irradiada pela antena da espaçonave. Na estação terrestre, a antena de recepção é utilizada para coletar a potência do sinal da onda eletromagnética de chegada, alimentando-a em um receptor de baixo ruído através de um guia de ondas.

Considerando que a fonte de transmissão irradia uma potência total P_T , se a potência é irradia *isotropicamente* (isto é, uniformemente em todas as direções), a densidade do fluxo de potência na distância r da fonte é $P_T/(4\pi r^2)$, na qual $4\pi r^2$ é a área superficial de uma esfera de raio r . Na prática, utilizamos uma antena altamente direcional para que a potência transmitida seja irradiada, primariamente, ao longo de uma direção particular de interesse. A antena possui um ganho que é definido como a razão da potência irradiada por unidade sólida de ângulo em uma dada direção pela potência média irradiada por ângulo sólido. Seja G_T o ganho de transmissão da antena na direção da qual a potência máxima é irradiada. Esta direção é chamada de *bore sight* (ou lóbulo principal de irradiação) da

⁴ Para um tratamento detalhado de cálculos de links para comunicação via satélite, veja o Capítulo 4 de Pratt e Bostian (1986).

⁵ Para um tratamento detalhado sobre cálculos de links para comunicações em espaço profundo, veja o Capítulo 3 de Yuen (1983).

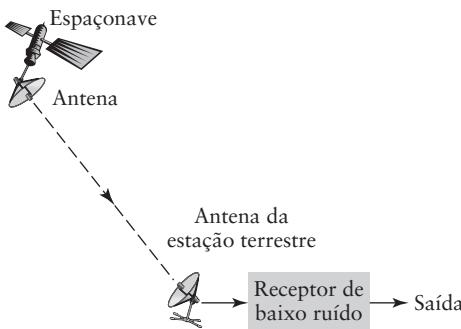


Figura 11.7 Canal de descida (downlink) de uma comunicação espacial.

antena. O ganho G_T é uma medida do aumento da potência irradiada pela antena sobre a potência irradiada por uma fonte isotrópica. Portanto, para uma potência total transmitida P_T alimentando uma antena sem perdas com ganho G_T , a densidade de fluxo de potência em uma distância r na direção do lóbulo principal da antena é dada por

$$\Phi = \frac{P_T G_T}{4\pi r^2} \quad (11.23)$$

Seja A_{ef} a *área de abertura efetiva* da antena de recepção. Esta área é relacionada com a área A de abertura física da antena pela fórmula

$$A_{\text{ef}} = \eta A \quad (11.24)$$

na qual η é a *eficiência de abertura*. Tipicamente, η está na faixa percentual de 40 a 90, dependendo do tipo de antena utilizada. O ganho da antena de recepção G_R é definido em termos da abertura efetiva A_{ef} por

$$G_R = \frac{4\pi A_{\text{ef}}}{\lambda^2} \quad (11.25)$$

na qual λ é o comprimento de onda da onda eletromagnética transmitida. λ é igual a c/f onde c é a velocidade de propagação (a qual é a mesma da velocidade da luz) e f é a frequência de transmissão. De forma equivalente, temos

$$A_{\text{ef}} = \frac{G_R \lambda^2}{4\pi} \quad (11.26)$$

Logo, dada a densidade Φ de fluxo de potência na antena de recepção com área efetiva de abertura A_{ef} , a potência recebida é

$$P_R = A_{\text{ef}} \Phi \quad (11.27)$$

Portanto, substituindo as Eqs. (11.23) e (11.26) na Eq. (11.27), obtemos o resultado

$$P_R = P_T G_T G_R \left(\frac{\lambda}{4\pi r} \right)^2 \quad (11.28)$$

A Eq. (11.28) é chamada de *fórmula de Friis* para transmissão, a qual relaciona a potência recebida com a potência transmitida. O produto $P_T G_T$ nesta equação é chamado de *potência efetiva isotrópica irradiada* (PEII). Ela descreve a combinação da fonte de transmissão e antena em termos de uma fonte isotrópica efetiva cuja potência $P_T G_T$ é irradiada uniformemente em todas as direções. O termo $(4\pi r/\lambda)^2$ é chamado de *caminho de perda* ou *perda em espaço livre*. Ele pode ser visto como a razão da potência transmitida pela potência recebida entre duas antenas que estão separadas por uma distância r .

Da Eq. (11.28), vemos que para valores dados de comprimento de onda λ e distância r , a potência recebida pode ser aumentada por três métodos:

1. A potência P_T transmitida pela espaçonave é aumentada. A potência P_T transmitida pode ser 20 watts ou menos. Mesmo que esta potência transmitida pareça ser baixa, a potência de entrada necessária para sua geração representa uma fração substancial da potência total disponível na espaçonave. Logo, existe um limite físico de quão grande podemos fazer a potência transmitida P_T .
2. O ganho G_T da antena de transmissão é aumentado. Isto irá ajudar a concentrar a potência transmitida na direção da antena de recepção. Entretanto, grandes valores de G_T requerem a utilização de uma antena grande. A escolha de G_T é, portanto, limitada pelas restrições de tamanho e peso da espaçonave.
3. O ganho G_R da antena de recepção é aumentado. Isto irá permitir que o receptor colete a maior quantidade possível da potência irradiada. Novamente, as restrições de tamanho e peso colocam um limite físico no tamanho da antena da estação terrestre, apesar destas limitações serem muito menos restritivas do que as da antena da espaçonave. Tipicamente⁶ temos $G_R \gg G_T$.

O ganho da antena de recepção e a perda em espaço livre são expressos em decibels (dB). Da mesma forma, a potência efetiva irradiada é expressa em *decibels relativos a um watt* (dBW). Então, podemos reescrever a fórmula de Friis para transmissão na forma

$$P_R = \text{EIRP} + G_R - L_p \quad (11.29)$$

na qual

$$\text{EIRP} = 10 \log_{10}(P_T G_T) \quad (11.30)$$

$$G_R = 10 \log_{10}\left(\frac{4\pi A_{\text{ef}}}{\lambda^2}\right) \quad (11.31)$$

$$L_p = 20 \log_{10}\left(\frac{4\pi r}{\lambda}\right) \quad (11.32)$$

Na Eq. (11.25), G_R é representado como uma razão de potências, enquanto que na Eq. (11.31), ele é expresso em decibels. A Eq. (11.29) é idealizada no sentido de não levar em conta todas as perdas no link. Para corrigi-la, devemos alterá-la para incluir um termo que representa o efeito combinado das perdas na atmosfera devido a atenuação da chuva, perdas nas antenas de transmissão e recepção e possíveis perdas de ganho devido ao não posicionamento correto das antenas. Seja L_0 a representação de todas estas perdas expresso em decibels. Este termo é, algumas vezes, chamado de *margem do sistema*. Então podemos modificar a expressão para a potência do sinal recebido como

$$P_R = \text{EIRP} + G_R - L_p - L_0 \quad (11.33)$$

A Eq. (11.33) representa a *quantidade de potência do link* no sentido de permitir ao projetista do sistema de um link de telecomunicações ajustar parâmetros controláveis tais como a PEII ou o ganho G_R da antena de recepção, fazendo um cálculo rápido da potência recebida.

A potência recebida P_R é geralmente chamada de *potência da portadora*. Esta nomenclatura é obtida dos primeiros sistemas via satélite nos quais a modulação em fase ou freqüência eram geralmente utilizadas. Como estas estratégias mantêm o envelope da portadora senoidal constante, a potência da portadora é normalmente igual à potência recebida.

⁶ Existem exceções para esta regra, principalmente em comunicações entre satélite e terminal móvel de comunicação e transmissão de televisão diretamente para as residências.

EXEMPLO 11.5 Quantidade de potência do link

Um satélite geoestacionário está localizado a uma distância de 40.000 km da estação terrestre. No satélite, uma fonte de freqüência 4GHz irradia uma potência de 10 watts através de uma antena com ganho de 20 dB. Assuma que a área de abertura efetiva da antena de recepção é 10 m². Calcule a potência do sinal recebido, ignorando perdas não ideais no link.

A potência efetiva irradiada com $P_T = 10$ W e $G_T = 20$ dB é

$$\begin{aligned} \text{PEII} &= 10 \log_{10}(P_T G_T) \\ &= 10 \log_{10} P_T + 10 \log_{10} G_T \\ &= 10 \log_{10} 10 + 20 \\ &= 30 \text{ dBW} \end{aligned}$$

A velocidade de propagação é igual a velocidade da luz: $c = 3 \times 10^8$ m/s. O comprimento de onda da onda eletromagnética transmitida em 4GHz é, portanto,

$$\begin{aligned} \lambda &= \frac{c}{f} \\ &= \frac{3 \times 10^8}{4 \times 10^9} \\ &= 0,075 \text{ m} \end{aligned}$$

Logo, a perda de caminho é

$$\begin{aligned} L_p &= 20 \log_{10} \left(\frac{4\pi r}{\lambda} \right) \\ &= 20 \log_{10} \left(\frac{4\pi \times 4 \times 10^7}{0,075} \right) \\ &= 196,5 \text{ dB} \end{aligned}$$

O ganho da antena de recepção com $A_{\text{ef}} = 10$ m² é

$$\begin{aligned} G_R &= 10 \log_{10} \left(\frac{4\pi A_{\text{ef}}}{\lambda^2} \right) \\ &= 10 \log_{10} \left(\frac{4\pi \times 10}{0,075^2} \right) \\ &= 43,5 \text{ dB} \end{aligned}$$

Portanto, utilizando a Eq. (11.29) e ignorando a margem do sistema L_0 , vemos que a potência do sinal recebido é igual a

$$\begin{aligned} P_R &= 30 + 43,5 - 196,5 \\ &= -123 \text{ dBW} \end{aligned}$$

De forma equivalente, a potência recebida pode ser escrita como $P_R = 5 \times 10^{-13}$ watts. O baixo valor de P_R indica que o sinal recebido na estação terrestre é extremamente fraco. Outras perdas no link (representadas pela margem do sistema) na estação terrestre farão com que o sinal recebido seja ainda mais fraco.

- **PROBLEMA 11.6** Uma antena de satélite possui diâmetro de 4,6 metros e opera em 12 GHz. Qual é o ganho da antena se a eficiência de abertura for 60%? Se a mesma antena for utilizada em 4GHz, qual será o ganho correspondente? ◀

- **PROBLEMA 11.7** Um satélite em uma distância de 40.000 quilômetros transmite um sinal em 12 GHz com PEII de 10 watts em direção a uma antena de 4,6 metros que possui eficiência de abertura de 60%. Qual é o nível de sinal recebido na saída da antena? ◀

■ Razão portadora/ruído

A razão portadora/ruído (RPR) é definida como a razão da potência da portadora pela potência de ruído disponível, medidas na entrada do receptor. Como mencionado anteriormente, a potência da portadora é a mesma potência P_R do sinal recebido. A fórmula para P_R é descrita pela equação de Friis para transmissão, Eq. (11.28). Para calcular a potência de ruído disponível na entrada do receptor, utilizamos a expressão kT_sB_N , na qual k é a constante de Boltzmann, T_s é a temperatura de ruído do sistema e B_N é a largura de faixa de ruído. Portanto, temos

$$\text{RPR} = \frac{P_R}{kT_sB_N} \quad (11.34)$$

Novamente, a simplicidade desta fórmula mostra que a utilização de temperatura de ruído é uma medida de quão ruidoso o sistema é. A razão portadora/ruído é geralmente a mesma da RSR de pré-detecção discutida no Capítulo 9.

EXEMPLO 11.6 Cálculos da RPR

Considere um receptor em uma estação terrestre com um amplificador criogenicamente resfriado. A temperatura equivalente de ruído do receptor é 20K. A estação terrestre utiliza uma grande antena operando em uma freqüência de 4GHz e com uma elevação de 15 graus. A temperatura de ruído da antena é estimada em 50K. Calcule a temperatura de ruído do sistema, assumindo que não existem perdas na antena alimentada e no guia de ondas. Calcule, então, a razão portadora/ruído, assumindo uma potência de portadora de -123 dBW (tal como no Exemplo 11.5) e uma largura de faixa de ruído de 36MHz.

Da Eq. (11.22), vemos que com $G_1 = 1$, a temperatura de ruído do sistema é

$$\begin{aligned} T_s &= T_a + T_e \\ &= 50 + 20 \\ &= 70 \text{ K} \end{aligned}$$

A potência de ruído disponível em 36 MHz de largura de faixa é, portanto,

$$\begin{aligned} kT_sB_N &= 1,38 \times 10^{-23} \times 70 \times (36 \times 10^6) \\ &= 3,48 \times 10^{-14} \text{ W} \\ &= -134,6 \text{ dBW} \end{aligned}$$

Logo, utilizando a Eq. (11.34), temos o seguinte valor para a razão portadora/ruído, expressa em decibels:

$$\begin{aligned} 10 \log_{10}(\text{CNR}) &= -123 + 134,6 \\ &= 11,6 \text{ dB} \end{aligned}$$

Lembre-se de nossa discussão sobre receptores analógicos no Capítulo 9 que esta RPR pode ser suficiente apenas para operar no limiar de um sistema FM. Em contraste, esta RPR pode fornecer uma taxa de erro de bit satisfatória para vários sistemas digitais.

- **PROBLEMA 11.8** A antena do Problema 11.7 possui uma temperatura de ruído de 70K e é diretamente conectada a um receptor com temperatura equivalente de ruído de 50K e um ganho de 60 dB. Qual é a temperatura de ruído do sistema? Se o sinal transmitido possuir uma largura de faixa de 100 kHz, qual é a razão portadora/ruído? Se o sinal digital possuir uma taxa de bit de 150 kbps, qual é a E_b/N_0 ? ◀

11.6 RÁDIO MÓVEL TERRESTRE

A propagação em espaço livre é geralmente um bom método para comunicações via satélite, nas quais existe uma visada direta limpa entre o transmissor e o receptor. Ela também possibilita a análise intuitiva pois existe uma explicação teórica simples para as perdas de propagação.

Em comunicações terrestres, por outro lado, as duas antenas estão relativamente próximas do chão. Conseqüentemente, prédios, terrenos e vegetação podem obstruir o caminho da visada direta. A comunicação via rádio terrestre geralmente baseia-se na *reflexão* de ondas eletromagnéticas em várias superfícies ou na *difração* destas ondas ao redor de vários objetos. Com estes modos adicionais de propagação, existem vários caminhos de propagação entre o transmissor e o receptor, e o receptor geralmente recebe um sinal significativo de mais do que um único caminho. Este fenômeno é chamado de propagação *multipath*. Com várias ondas chegando na mesma posição, propriedades de propagação podem diferir significativamente das propriedades da propagação em espaço livre.

Existem três modos básicos de propagação que se aplicam à propagação terrestre: espaço livre, reflexão e difração.⁷

1. A *propagação em espaço livre* depende do caminho de visada direta entre o transmissor e o receptor e uma certa área limpa ao redor do caminho. A área limpa necessária, ilustrada na Fig. 11.8, é relacionada com a separação entre os dois e o comprimento de onda da transmissão. Uma regra empírica é que o volume chamado de *primeira zona de Fresnel* deve ser mantido limpo de objetos para uma propagação aproximada à de espaço livre. A zona de Fresnel define uma elipsóide de revolução. Objetos dentro da primeira zona de Fresnel irão afetar a transmissão e causarão desvios do modelo de propagação em espaço livre. O raio da primeira zona de Fresnel depende da posição entre a antena de transmissão e recepção, sendo dado por

$$b = \sqrt{\frac{\lambda d_1 d_2}{d_1 + d_2}} \quad (11.35)$$

na qual λ é o comprimento de onda de transmissão, d_1 é a distância para o transmissor e d_2 é a distância para o receptor, para um ponto em particular ao longo do caminho.

2. A *reflexão* se refere ao batimento de ondas eletromagnéticas em objetos tais como edifícios, montanhas e veículos. Em comunicações terrestres sem fio, geralmente não existe uma visada direta entre o transmissor e o receptor e a comunicação se baseia na reflexão e no modo de difração a seguir. A reflexão geralmente atenua ainda mais o sinal e pode, também, afetar sua fase.
3. A *difração* se refere à capacidade das ondas eletromagnéticas de contornar objetos, como um edifício, ou terrenos, como morros, e atravessar objetos, como árvores e outras formas de vegetação.

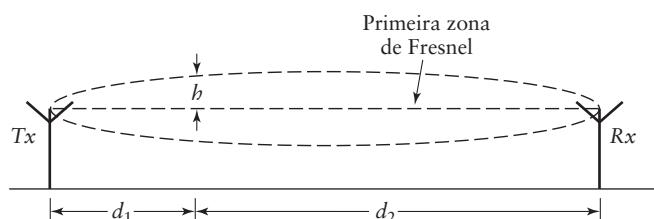


Figura 11.8 A primeira zona de Fresnel para a propagação em espaço livre.

⁷ Veja o Capítulo 2 de Haykin e Moher (2004) para mais detalhes.

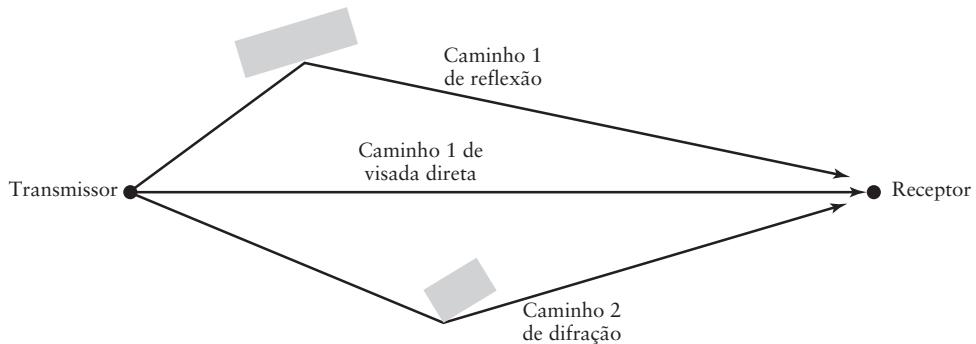


Figura 11.9 Ilustração da propagação multicaminho entre transmissor e receptor.

Na propagação terrestre, o sinal recebido é geralmente a combinação de vários destes modos de propagação. Ou seja, o sinal transmitido pode chegar no receptor por vários caminhos, como ilustrado na Fig. 11.9. Os sinais destes caminhos diferentes podem interferir construtivamente ou destrutivamente, uns com os outros. Portanto, a transmissão multicaminho pode ter propriedades bem diferentes da propagação em espaço livre. Medidas indicam que a propagação terrestre pode ser separada em várias componentes. A seguir, iremos descrever duas destas componentes: a perda média do caminho e o desvanecimento rápido.

EXEMPLO 11.7 Fantasmas

Antes do advento da televisão via cabo e via satélite, os sinais de televisão eram principalmente recebidos *no ar* via links terrestres. Um fenômeno, algumas vezes chamado de *fantasmas*, podia ser observado algumas vezes em estações de televisão fracas (isto é, mais distantes). Com os fantasmas, eram observados a imagem e a réplica da imagem em uma posição um pouco deslocada. Este fantasma é a visualização do multicaminho. Ele é uma cópia do sinal original recebido com um pequeno atraso de tempo.

- **PROBLEMA 11.9** As antenas de transmissão e recepção para um sinal de 4GHz são localizadas no topo de edifícios de 20 metros, separadas por dois quilômetros. Para a propagação em espaço livre, qual é a altura máxima permitida para um objeto localizado no meio do caminho entre as duas torres? ◀

■ Perda média do caminho⁸

Com a propagação em espaço livre, existe um modelo relativamente simples para a predição da perda de propagação em função da distância. Com outros modos de propagação – reflexão e difração – os cálculos não são tão simples, especialmente quando existem múltiplas reflexões e/ou difrações. Para cenários simples, podemos desenvolver modelos matemáticos para as perdas de propagação devido a reflexão e difração. Para cenários mais práticos, ainda podemos construir modelos físicos do ambiente incluindo terrenos, edifícios e vegetação e, utilizando uma análise por computador, estimar as perdas de propagação para caminhos particulares.

Alternativamente, podemos utilizar uma *abordagem estatística* na qual as características de propagação são empiricamente aproximadas com base em medidas de certos

⁸ Nesta seção, apresentamos um modelo relativamente simples para a perda média do caminho. Este modelo simples é adequado para o entendimento de fatores que afetam a perda do caminho. Para aplicações práticas, modelos mais sofisticados foram desenvolvidos, os quais dependem da aplicação. Um exemplo é o modelo de Okumura-Hata para a propagação de rádio terra-móvel em freqüências entre 150MHz e 1GHz. Veja o Capítulo 2 de Haykin e Moher (2004) para mais detalhes.

tipos gerais de ambientes, tais como urbano, suburbano e rural. A abordagem estatística é dividida em duas componentes: uma estimativa da *média da perda do caminho* e uma componente aleatória que depende das características físicas do terreno.

As medidas da potência do campo em vários ambientes em função da distância r , do transmissor para o receptor, motiva um modelo simples de propagação para a perda média do caminho, tendo a forma

$$\frac{P_R}{P_T} = \frac{\beta}{r^n} \quad (11.36)$$

no qual, tipicamente, o *expoente de perda do caminho* n varia de 2 a 5, dependendo do ambiente de propagação. O parâmetro β representa o ganho que está relacionado com a freqüência e ganhos da antena e também pode ser relacionado com a altura das antenas e outros fatores. O lado direito da Eq. (11.36) é, algumas vezes, escrito na forma alternativa

$$\frac{P_R}{P_T} = \frac{\beta_0}{(r/r_0)^n} \quad (11.37)$$

Nesta representação, β_0 representa a medida do ganho do caminho da distância de referência r_0 . Na ausência de outras informações, β_0 pode ser assumido como o ganho do caminho de espaço livre na distância de um metro. (A perda do caminho de referência é o inverso β_0 .)

Vários estudos de propagação foram realizados tentando identificar os diferentes efeitos ambientais. A conclusão obtida por estes estudos é que a maioria destes efeitos são localmente dependentes e difíceis, se não impossíveis, de serem genericamente caracterizados. Alguns valores de exemplo para o expoente n no modelo da Eq. (11.36) são dados na Tabela 11.1.

Apesar desta tabela indicar a tendência geral, existem exceções. Por exemplo, se o caminho de propagação é ao longo de uma rua reta com arranha-céus nos dois lados, pode haver um efeito de condução (guia de onda) e as perdas de propagação podem ser similares à do espaço livre ou até menores.

Este modelo de perda média do caminho é bem flexível e utilizado para o estudo analítico de problemas, pois ele nos permite parametrizar a performance de vários fatores relacionados com o sistema. Para aplicações comerciais em ambiente terrestre, um trabalho de medição em campo ou a modelagem detalhada do ambiente é preferível.

TABELA 11.1 Exemplo do expoente de perda do caminho

Ambiente	n
Espaço livre	2
Área rural plana	3
Área rural montanhosa	3,5
Suburbano, baixa altura	4
Urbano denso, arranha-céus	4,5

EXEMPLO 11.8 Potência do link terrestre

Uma campanha de medição de propagação indica que, em uma área urbana, a perda média do caminho pode ser modelada por

$$L_p = 41 + 31 \log_{10}(r)$$

na qual a perda do caminho está em decibel e r em metros. Além da perda média do caminho, o sistema deve incluir uma margem de 12 dB para compensar as variações da potência do sinal. Se o transmissor PEII é cinco watts, determine a potência mínima do sinal esperado em uma distância de cinco quilômetros.

Resolvemos este problema apelando para a fórmula de Friis para transmissão em decibel, ou seja,

$$P_R = \text{PEII} - L_p - L_0 + G_R$$

O cabeçalho do problema fornece todas as grandezas do lado direito, exceto G_R . Fazemos a consideração de que o receptor é móvel e o ganho da antena aproximadamente unitário. Ou seja, $G_R = 0$ dB. Então, em uma distância de 5 km,

$$\begin{aligned} P_R &= 10 \log_{10} 5 - 41 - 31 \log_{10} 5000 - 12 + 0 \\ &= -160,7 \text{ dBW} \\ &= -130,7 \text{ dBm} \end{aligned}$$

- **PROBLEMA 11.10** Uma campanha de medição indica que a perda média do caminho em 900 MHz em uma área suburbana pode ser modelada com um expoente de perda de caminho de 2,9. Qual é a perda média do caminho a uma distância de três quilômetros usando este modelo? Como esta perda se compara com a perda em espaço livre para a mesma distância? ◀

■ Perdas aleatórias do caminho

A perda média do caminho é simplesmente a atenuação média em função da distância. Cinquenta por cento das localidades irão ter uma perda maior e 50% uma perda menor. A perda real irá depender do terreno, edifícios e vegetação exatos ao longo do caminho. Um dos vários contribuintes para estas variações sobre o valor médio é o movimento do terminal transmissor ou receptor. Por exemplo, o nível de potência recebida em um terminal móvel que está a uma distância relativamente constante do transmissor pode ser mostrada na Fig. 11.10. Para esta figura, a média da potência do sinal recebido é aproximadamente -95 dBm, mas variações acima de 7 dB e abaixo de 35 dB podem ser observadas.

Estas variações rápidas da potência do sinal são devidas a reflexões de objetos locais que mudam rapidamente a fase da portadora em pequenas distâncias. A alteração da fase das reflexões faz com que elas se somem algumas vezes construtivamente e algumas vezes

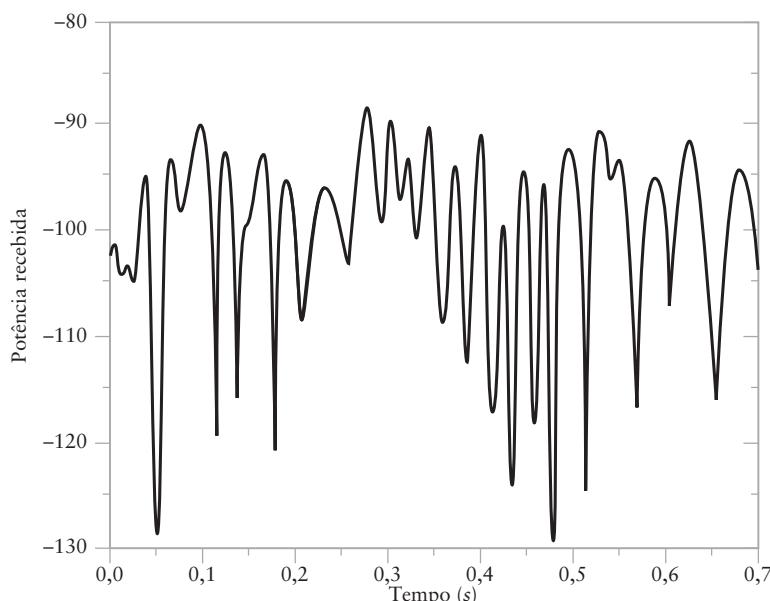


Figura 11.10 Ilustração das variações da potência do sinal recebido em um desvanecimento Rayleigh.

destrutivamente. Este rápido desvanecimento é geralmente chamado de *desvanecimento Rayleigh*, devido a função de distribuição de probabilidade que é utilizada para modelar as variações de amplitude. Se R é a amplitude do sinal, com raiz média quadrática do nível de R_{rms} , então a probabilidade da amplitude estar abaixo de um dado nível r é dada pela *função de distribuição de Rayleigh*

$$P[R < r] = 1 - \exp\left(-\frac{r^2}{R_{rms}^2}\right) \quad (11.38)$$

A distribuição de Rayleigh é mostrada na Fig. 11.11. A mediana da distribuição é aproximadamente R_{rms} . Isto significa que para 50% das localizações existe uma interferência construtiva ($R > R_{rms}$) e para 50% das localizações existe uma interferência destrutiva ($R < R_{rms}$).

Desvanecimentos profundos de 20 dB ($R < 0,1 R_{rms}$) ou mais ocorrem apenas raramente (com uma probabilidade de 1%). Entretanto, a Fig. 11.11 indica que pode haver uma ampla variação na potência do sinal recebido devido a reflexões locais.

EXEMPLO 11.9 Margem para o desvanecimento Rayleigh

Suponha que um rádio móvel possui uma sensibilidade do receptor de -90 dBm. Ou seja, -90 dBm é o nível mínimo do sinal que o receptor pode demodular com qualidade. O sistema opera em 900 MHz e a distância operacional é de 2 quilômetros. Obtenha a potência necessária do transmissor deste sistema se o ambiente de propagação possuir um expoente de perda de caminho $n = 2,9$ e o receptor deve tolerar 90% de desvanecimento.

Se o receptor deve tolerar 90% de desvanecimento, então a função de distribuição de Rayleigh da Fig. 11.11 indica que o sistema deve incluir uma margem de 10 dB. Conseqüentemente, o nível médio de sinal deve estar 10 dB acima da sensibilidade do receptor no limite de cobertura, ou seja, -80 dBm. Então, utilizando a Eq. (11.37), temos

$$P_T = \frac{P_R}{\beta_0} \frac{1}{(r/r_o)^{2,9}}$$

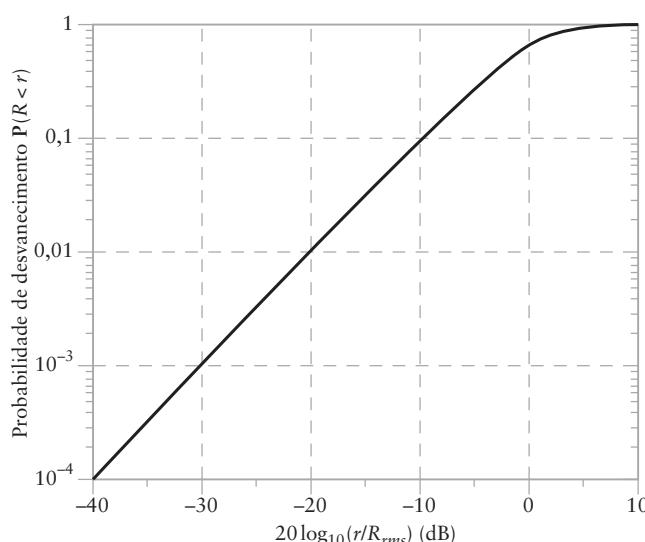


Figura 11.11 Distribuição normalizada de amplitude de Rayleigh (r/R_{rms}).

assumindo que β_0 é o ganho de caminho em espaço livre em um metro, então

$$\begin{aligned}\beta_o (dB) &= -L_p (1m) = -20 \log_{10} \left(\frac{4\pi}{\lambda} \right) \\ &= -31,5 \text{ dB}\end{aligned}$$

em 900 MHz. Expressando a equação acima de P_T no formato de decibel, obtemos

$$\begin{aligned}P_T &= P_R - \beta_0 + 29 \log_{10} \left(\frac{2000}{1} \right) \\ &= -80 + 31,5 + 95,7 \\ &= +47 \text{ dBm}\end{aligned}$$

Isto corresponde a, aproximadamente, um transmissor de 50 watts.

- **PROBLEMA 11.11** Expresse a mediana verdadeira da distribuição de Rayleigh como uma fração do valor R_{rms} . Qual é o erro em decibel na aproximação $R_{mediana} \approx R_{rms}$? ◀

11.7 RESUMO E DISCUSSÃO

Neste capítulo, caracterizamos fontes específicas de ruído e atenuação de sinal em sistemas de comunicação práticos. Elas são necessárias para calcular as razões sinal/ruído utilizadas para caracterizar a performance do receptor nos capítulos anteriores. Em particular:

- (i) Várias fontes de ruído térmico foram identificadas e caracterizamos suas densidades espectrais de potência através da simples relação $N_0 = kT$, na qual k é a constante de Boltzmann e T é a temperatura absoluta em Kelvin.
- (ii) Os conceitos relacionados com a figura de ruído de temperatura equivalente de ruído foram definidos e utilizados para caracterizar as contribuições que vários dispositivos eletrônicos ou fontes de ruído fazem para o ruído total. Mostramos, então, como o ruído total é calculado em uma série de sistemas de duas portas.
- (iii) A equação de Friis foi desenvolvida para modelar matematicamente a relação entre as potências do sinal transmitido e recebido como uma função dos ganhos das antenas de transmissão e recepção e da perda do caminho.
- (iv) Perdas de propagação variando de condições de espaço livre, típicos de canais de satélite, até aplicações típicas de comunicação móvel terrestre foram descritas. Variações da perda média do caminho foram identificadas como uma importante consideração na propagação terrestre.

Este capítulo forneceu uma breve introdução a dois canais de comunicação – satélite e terrestre – e sobre os efeitos que a propagação e o ruído tem na performance do sistema. As perdas de propagação do canal são uma componente chave na determinação da potência do sinal no receptor. As fontes físicas de ruído contribuem para a potência do ruído no receptor. A combinação destas duas quantidades – ou seja, a razão da potência do sinal pela potência do ruído – forma o método primário de medição da qualidade de um link de comunicação.

Problemas adicionais

- 11.12** Calcule a densidade espectral de ruído em watts por hertz de:

- (a) Um resistor ideal em temperatura nominal de 290 K.
- (b) Um amplificador com temperatura equivalente de ruído de 22.000 K.

- 11.13** Para os dois casos do Problema 11.12, calcule a RSR pré-detecção quando a potência do sinal recebido é:
- 60 dBm e a largura de faixa do receptor é 1 MHz.
 - 90 dBm e a largura de faixa do receptor é 30 kHz.
- 11.14** Uma rede local sem fio transmite um sinal que possui largura de faixa de ruído de aproximadamente 6MHz. Se a potência do sinal nos terminais de entrada do receptor é -90 dBm e a figura de ruído do receptor é 8 dB, qual é a razão sinal/ruído de pré-detecção?
- 11.15** Um receptor de comunicação inclui uma antena reta cuja temperatura de ruído é aproximadamente a da terra, ou seja, 290 K. O pré-amplificador do receptor possui uma figura de ruído de 4 dB e um ganho de 25 dB. Calcule (a) a temperatura equivalente de ruído da antena e do pré-amplificador e (b) a figura de ruído combinada.
- 11.16** Uma antena parabólica com diâmetro de 0,75 metros é utilizada para receber um sinal de satélite de 12 GHz. Calcule o ganho em decibels desta antena. Assuma que a eficiência da antena é 60%.
- 11.17** Se a temperatura de ruído do sistema de um receptor de satélite é 300 K, qual é a potência necessária do sinal recebido para produzir uma C/N_0 de 80 dB?
- 11.18** Se um satélite está a 40.000 Km da antena do Problema 11.16, qual PEII do satélite irá produzir uma potência de sinal de -110 dBm nos terminais da antena? Assuma que a freqüência de transmissão é 12 GHz.
- 11.19** Antenas são colocadas em dois edifícios comerciais de 35 metros e separadas por 10 quilômetros. Determine a altura mínima de um edifício entre as duas torres que irá perturbar a consideração de propagação em espaço livre.
- 11.20** Se um receptor possui sensibilidade de -90 dBm e uma figura de ruído de 12 dB, qual é a razão sinal/ruído pré-detecção mínima de um sinal em 8-kHz?

Problemas avançados

- 11.21** Uma antena de satélite é instalada no topo de uma espaçonave e possui temperatura de ruído de 100 K. A antena é conectada por um cabo coaxial até um amplificador de baixo ruído no equipamento instalado na frente da espaçonave. O cabo causa uma atenuação de 2-dB no sinal. O amplificador de baixo ruído possui um ganho de 60 dB e temperatura de ruído de 120 K. Qual é a temperatura de ruído do sistema? Existe uma posição melhor para a instalação do amplificador de baixo ruído?
- 11.22** Um transmissor de uma rede local sem fio irradia 200 miliwatts. Experimentos indicam que a perda do caminho é precisamente descrita por

$$L_p = 31 + 33 \log_{10}(r)$$

na qual a perda do caminho está em decibels e r é a distância em metros. Se a sensibilidade mínima do receptor é -85 dBm, qual é a faixa de ação do transmissor?

- 11.23** Um rádio móvel transmite 20 watts e a perda média do caminho pode ser aproximada por

$$L_p = 69 + 31 \log_{10}(r)$$

na qual a perda do caminho está em decibels e r é a distância em quilômetros. A sensibilidade do receptor é -110 dBm e uma margem de 12 dB deve ser incluída para compensar as variações da perda média do caminho. Determine a faixa de ação do transmissor.

- 11.24** Um telefone celular transmite 600 miliwatts de potência.
- A sensibilidade do receptor é -90 dBm, qual será a faixa do telefone considerando propagação em espaço livre? Assuma que as antenas de transmissão e recepção possuem ganho unitário e que a transmissão ocorre em 900 MHz.
 - Se as condições de propagação mostram um expoente de perda de caminho de 3,1 com um ganho fixo $\beta = -36$ dB, qual será a faixa neste caso?

- 11.25 Um link de rádio em visada direta de 10 quilômetros é necessário para transmitir dados a uma taxa de 1 megabit por segundo com uma freqüência central de 4 GHz. O transmissor utiliza uma antena com ganho de 10 dB e modulação QPSK com espectro de pulso de raiz de cosseno levantado com um fator de roll-off de 0,5. O receptor também possui uma antena com ganho de 10 dB e possui uma temperatura de ruído de sistema de 900 K. Qual é a potência mínima transmitida necessária para atingir uma taxa de erro de bit de 10^{-5} ?
- 11.26 Um rádio terra-móvel transmite 128 kbps em uma freqüência de 700 MHz. O transmissor utiliza uma antena omni-direcional e modulação 16-QAM com espectro de pulso de raiz de cosseno levantado com fator de roll-off de 0,5. O receptor também utiliza uma antena omni-direcional e possui figura de ruído de 6 dB. Se a perda de caminho entre o transmissor e o receptor for dada por

$$L_p(r) = 30 + 28 \log_{10}(r) \text{ dB}$$

na qual r está em metros, calcule a faixa máxima na qual a taxa de erro de bit de 10^{-4} pode ser alcançada.

RAZÕES DE POTÊNCIA E DECIBEL

Em cálculos de sistemas e medidas envolvendo a utilização de razões de potência, é prática comum utilizar uma unidade chamada *decibel*. O decibel, normalmente abreviado por dB, é um décimo de uma unidade maior, o *bel*.¹ Na prática, entretanto, vemos que na maioria das aplicações o bel é uma unidade muito grande, por isto utilizamos o dB como unidade para expressar razões de potência. O dB é particularmente apropriado para medidas de som, pois o ouvido responde ao som de forma aproximadamente logarítmica. Portanto, incrementos iguais de dB são percebidos pelo ouvido como incrementos iguais no som.

Seja P a potência em algum ponto de interesse em um sistema. Seja P_0 o nível de potência de *referência* com relação ao qual a potência P é comparada. O número de decibels na razão de potência P/P_0 é definido como $10 \log_{10}(P/P_0)$. Por exemplo, uma razão de potência de 2 corresponde, aproximadamente a 3 dB e uma razão de potência de 10 corresponde exatamente a 10 dB.

Podemos expressar, também, a própria potência do sinal P em dB se dividirmos P por 1 watt ou um miliwatt. No primeiro caso, expressamos a potência P do sinal em dBW como $10 \log_{10}(P/1W)$, na qual W é a abreviação para watt. No segundo caso, expressamos a potência P do sinal em dBm como $10 \log_{10}(P/1mW)$, na qual mW é a abreviação para miliwatt.

¹ A unidade bel possui este nome em homenagem a Alexander Graham Bell. Além de inventar o telefone, Bell foi o primeiro a utilizar medidas de potência em logaritmo em pesquisa de som e audição.

Apêndice 2

SÉRIE DE FOURIER

Neste apêndice, iremos revisar a formulação da série de Fourier e iremos desenvolver a transformada de Fourier com uma generalização da série de Fourier.

Seja $g_{T_0}(t)$ um sinal periódico com período T_0 . Utilizando a *expansão em série de Fourier* neste sinal, somos capazes de calculá-lo como uma soma infinita de termos em seno e cosseno. A expansão pode ser expressa na forma trigonométrica:

$$g_{T_0}(t) = a_0 + 2 \sum_{n=1}^{\infty} [a_n \cos(2\pi n f_0 t) + b_n \sin(2\pi n f_0 t)] \quad (\text{A2.1})$$

na qual f_0 é a *freqüência fundamental*:

$$f_0 = \frac{1}{T_0} \quad (\text{A2.2})$$

Os coeficientes a_n e b_n representam as amplitudes dos termos em cosseno e seno, respectivamente. A quantidade $n f_0$ representa a n -ésima harmônica da freqüência fundamental f_0 . Cada um dos termos $\cos(2\pi n f_0 t)$ e $\sin(2\pi n f_0 t)$ são chamados de *funções base*. Estas funções base formam um conjunto *ortogonal* no intervalo T_0 no qual eles satisfazem o seguinte conjunto de relações:

$$\int_{-T_0/2}^{T_0/2} \cos(2\pi m f_0 t) \cos(2\pi n f_0 t) dt = \begin{cases} T_0/2, & m = n \\ 0, & m \neq n \end{cases} \quad (\text{A2.3})$$

$$\int_{-T_0/2}^{T_0/2} \cos(2\pi m f_0 t) \sin(2\pi n f_0 t) dt = 0, \quad \text{para todo } m \neq n \quad (\text{A2.4})$$

$$\int_{-T_0/2}^{T_0/2} \sin(2\pi m f_0 t) \sin(2\pi n f_0 t) dt = \begin{cases} T_0/2, & m = n \\ 0, & m \neq n \end{cases} \quad (\text{A2.5})$$

Para determinarmos o coeficiente a_0 , integramos os dois lados da Eq. (A2.1) em um período completo. Vemos, portanto, que a_0 é o *valor médio* do sinal periódico $g_{T_0}(t)$ em um período, como mostrado pela *média temporal*,

$$a_0 = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_{T_0}(t) dt \quad (\text{A2.6})$$

Para terminarmos o coeficiente a_n , multiplicamos os dois lados da Eq. (A2.1) por $\cos(2\pi n f_0 t)$ e integramos no intervalo $-T_0/2$ a $T_0/2$. Então, utilizando as Eqs. (A2.3) e (A2.4), temos

$$a_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_{T_0}(t) \cos(2\pi n f_0 t) dt, \quad n = 1, 2, \dots \quad (\text{A2.7})$$

Similarmente, temos

$$b_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_{T_0}(t) \operatorname{sen}(2\pi n f_0 t) dt, \quad n = 1, 2, \dots \quad (\text{A2.8})$$

Uma questão básica neste ponto é: Dado o sinal periódico $g_{T_0}(t)$ com período T_0 , como sabemos que a expansão da série de Fourier da Eq. (A2.1) é *convergente* no sentido de que o somatório infinito de termos desta expansão leva exatamente ao valor igual a $g_{T_0}(t)$? Para resolver esta questão, precisamos mostrar que, para os coeficientes a_0 , a_n e b_n calculados de acordo com as Eqs. (A2.6) a (A2.8), esta série irá, de fato, convergir para $g_{T_0}(t)$. Em geral, para um sinal periódico $g_{T_0}(t)$ com forma de onda arbitrária, não existe garantia que a série da Eq. (A2.1) irá convergir para $g_{T_0}(t)$ ou mesmo que os coeficientes a_0 , a_n e b_n irão existir. Uma prova rigorosa de convergência das série de Fourier está além do escopo deste livro. Aqui iremos simplesmente afirmar que o sinal periódico $g_{T_0}(t)$ pode ser expandido em uma série de Fourier se o sinal $g_{T_0}(t)$ satisfizer as *condições de Dirichlet*:

1. A função $g_{T_0}(t)$ é de valor único dentro do intervalo T_0 .
2. A função $g_{T_0}(t)$ possui no máximo um número finito de descontinuidades no intervalo T_0 .
3. A função $g_{T_0}(t)$ possui um número finito de máximos e mínimos no intervalo T_0 .
4. A função $g_{T_0}(t)$ é absolutamente integrável, ou seja,

$$\int_{-T_0/2}^{T_0/2} |g_{T_0}(t)| dt < \infty$$

As condições de Dirichlet são satisfeitas pelos sinais periódicos geralmente encontrados em sistemas de comunicação. Em um ponto de descontinuidade, a série de Fourier converge para o valor médio entre o ponto a esquerda e a direita da descontinuidade.

Série exponencial complexa de Fourier

A série de Fourier da Eq. (A2.1) pode ser colocada em uma forma muito mais simples e elegante com a utilização de exponenciais complexas. Fazemos isto substituindo na Eq. (A2.1) a forma exponencial para cosseno e seno, ou seja,

$$\cos(2\pi n f_0 t) = \frac{1}{2} [\exp(j2\pi n f_0 t) + \exp(-j2\pi n f_0 t)]$$

$$\operatorname{sen}(2\pi n f_0 t) = \frac{1}{2j} [\exp(j2\pi n f_0 t) - \exp(-j2\pi n f_0 t)]$$

Obtemos, portanto,

$$\begin{aligned} g_{T_0}(t) = a_0 + \sum_{n=1}^{\infty} [(a_n - jb_n) \exp(j2\pi n f_0 t) \\ + (a_n + jb_n) \exp(-j2\pi n f_0 t)] \end{aligned} \quad (\text{A2.9})$$

Seja c_n um coeficiente complexo relacionado com a_n e b_n através de

$$c_n = \begin{cases} a_n - jb_n, & n > 0 \\ a_0, & n = 0 \\ a_n + jb_n, & n < 0 \end{cases} \quad (\text{A2.10})$$

Então, podemos simplificar a Eq. (A2.9) como mostrado a seguir:

$$g_{T_0}(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_0 t) \quad (\text{A2.11})$$

na qual

$$c_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_{T_0}(t) \exp(-j2\pi n f_0 t) dt, \quad n = 0, \pm 1, \pm 2, \dots \quad (\text{A2.12})$$

A expansão em série da Eq. (A2.11) é chamada de *Série Exponencial Complexa de Fourier*. c_n é chamado de *coeficientes complexos de Fourier*. A Eq. (A2.12) afirma que, dado um sinal periódico $g_{T_0}(t)$, podemos determinar o conjunto completo de coeficientes complexos de Fourier. Por outro lado, a Eq. (A2.11) afirma que, dado este conjunto de coeficientes, podemos reconstruir o sinal periódico original $g_{T_0}(t)$ exatamente. A partir da matemática da análise real e complexa, a Eq. (A2.12) é um *produto interno* do sinal pela função base $\exp(j2\pi n f_0 t)$, e por cuja combinação linear todas as funções quadráticas integráveis podem ser expressas usando a Eq. (A2.11).

De acordo com esta representação, um sinal periódico contém todas as freqüências (tanto positivas quanto negativas) que são harmonicamente relacionadas com a fundamental. A presença de freqüências negativas é simplesmente um resultado do fato do modelo matemático do sinal descrito pela Eq. (A2.11) necessitar de freqüências negativas. De fato, esta representação não possui significado físico. A razão para a utilização de funções base de valor complexo e componentes de freqüências negativa é simplesmente para fornecer uma descrição matemática compacta de um sinal periódico, a qual é adequada tanto para o trabalho teórico quanto prático.

■ Espectro discreto

A representação de um sinal periódico por uma série de Fourier é equivalente à resolver o sinal em suas várias componentes harmônicas. Portanto, utilizando a série exponencial complexa de Fourier, vemos que um sinal periódico $g_{T_0}(t)$ com período T_0 possui componentes em freqüências $0, \pm f_0, \pm 2f_0, \pm 3f_0, \dots$, e assim por diante, na qual $f_0 = 1/T_0$ é a freqüência fundamental. Ou seja, enquanto o sinal $g_{T_0}(t)$ existe no domínio do tempo, podemos dizer que sua descrição no domínio da freqüência consiste de componentes nas freqüências $0, \pm f_0, \pm 2f_0, \dots$, chamado de *espectro*. Se especificarmos o sinal periódico $g_{T_0}(t)$, podemos determinar seu espectro, por outro lado, se especificarmos o espectro, podemos determinar o sinal correspondente. Isto significa que um sinal periódico $g_{T_0}(t)$ pode ser especificado em duas formas equivalentes:

1. Uma *representação no domínio do tempo*, na qual $g_{T_0}(t)$ é definido como uma função do tempo.
2. Uma *representação no domínio da freqüência*, na qual o sinal é definido em termos de seu espectro.

Apesar destas duas descrições serem aspectos separados de um dado fenômeno, elas não são independentes uma da outra, mas relacionadas, como mostra a teoria de Fourier.

Em geral, o coeficiente de Fourier c_n é um número complexo e podemos expressá-lo na forma:

$$c_n = |c_n| \exp[j \arg(c_n)] \quad (\text{A2.13})$$

$|c_n|$ define a amplitude da n -ésima componente de harmônica do sinal periódico $g_{T_0}(t)$, de tal forma que um gráfico de $|c_n|$ em função da freqüência resulta em um *espectro discreto em amplitude* do sinal. Um gráfico de $\arg(c_n)$ pela freqüência resulta em um *espectro discreto de fase* do sinal. Nos referimos ao espectro como um *espectro discreto* porque tanto a amplitude quanto a fase de c_n possuem valores não nulos somente para freqüências discretas que são múltiplos inteiros (tanto positivos quanto negativos) da freqüência fundamental.

Para uma função periódica de valor real $g_{T_0}(t)$, vemos pela definição do coeficiente de Fourier c_n dada pela Eq. (A2.12), que

$$c_{-n} = c_n^* \quad (\text{A2.14})$$

na qual c_n^* é o conjugado complexo de c_n . Portanto, temos

$$|c_{-n}| = |c_n| \quad (\text{A2.15})$$

e

$$\arg(c_{-n}) = -\arg(c_n) \quad (\text{A2.16})$$

ou seja, o espectro de amplitude de um sinal periódico de valor real possui simetria par (uma função par de n) e o espectro de fase possui simetria ímpar (uma função ímpar de n) com relação ao eixo vertical passando pela origem.

EXEMPLO Trem de pulso periódico

Considere um trem periódico de pulsos retangulares de duração T e período T_0 , como mostrado na Fig. A2.1. Por conveniência de análise, a origem foi escolhida como sendo o centro do pulso. Este sinal pode ser analiticamente descrito em um período por:

$$g_{T_0}(t) = \begin{cases} A, & -\frac{T}{2} \leq t \leq \frac{T}{2} \\ 0, & \text{para o restante do período} \end{cases} \quad (\text{A2.17})$$

Utilizando a Eq. (A2.12) para calcular o coeficiente complexo de Fourier c_n , obtemos

$$\begin{aligned} c_n &= \frac{1}{T_0} \int_{-T/2}^{T/2} A \exp(-j2\pi n f_0 t) dt \\ &= \frac{A}{n\pi} \operatorname{sinc}\left(\frac{n\pi T}{T_0}\right), \quad n = 0, \pm 1, \pm 2, \dots \end{aligned} \quad (\text{A2.18})$$

na qual T/T_0 é o chamado *duty cycle* (ciclo de trabalho).

Podemos simplificar a notação utilizando a *função sinc*:

$$\operatorname{sinc}(\lambda) = \frac{\sin(\pi\lambda)}{\pi\lambda} \quad (\text{A2.19})$$

Portanto, podemos reescrever a Eq. (A2.18) como:

$$c_n = \frac{TA}{T_0} \operatorname{sinc}\left(\frac{nT}{T_0}\right) = \frac{TA}{T_0} \operatorname{sinc}(f_n T) \quad (\text{A2.20})$$

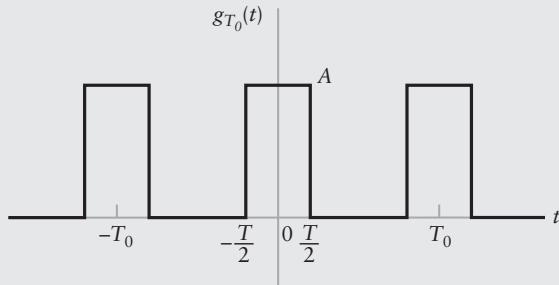


Figura A2.1 Trem periódico de pulsos retangulares de amplitude A , duração T e período T_0 .

Na Fig. A2.2, traçamos o espectro de amplitude $|c_n|$ e o espectro de fase $\arg(c_n)$ em função da freqüência discreta $fn = n/T_0$ para um ciclo de trabalho T/T_0 igual a 0,2. Baseado nesta figura, podemos observar:

1. O espaçamento de linha no espectro de amplitude da Fig. A2.2(a) é determinado pelo período T_0 .
2. O envelope do espectro de amplitude é determinado pela amplitude A do pulso, duração T do pulso e ciclo de trabalho T/T_0 .
3. Os cruzamentos com o zero ocorrem no envelope do espectro de amplitude em freqüências que são múltiplos inteiros de $1/T$.
4. O espectro de fase assume o valor de 0 graus e ± 180 graus, dependendo da polaridade de $\text{sinc}(nT/T_0)$. Na Fig. A2.2(b), utilizamos tanto 180 graus e -180 graus para preservar a simetria ímpar.

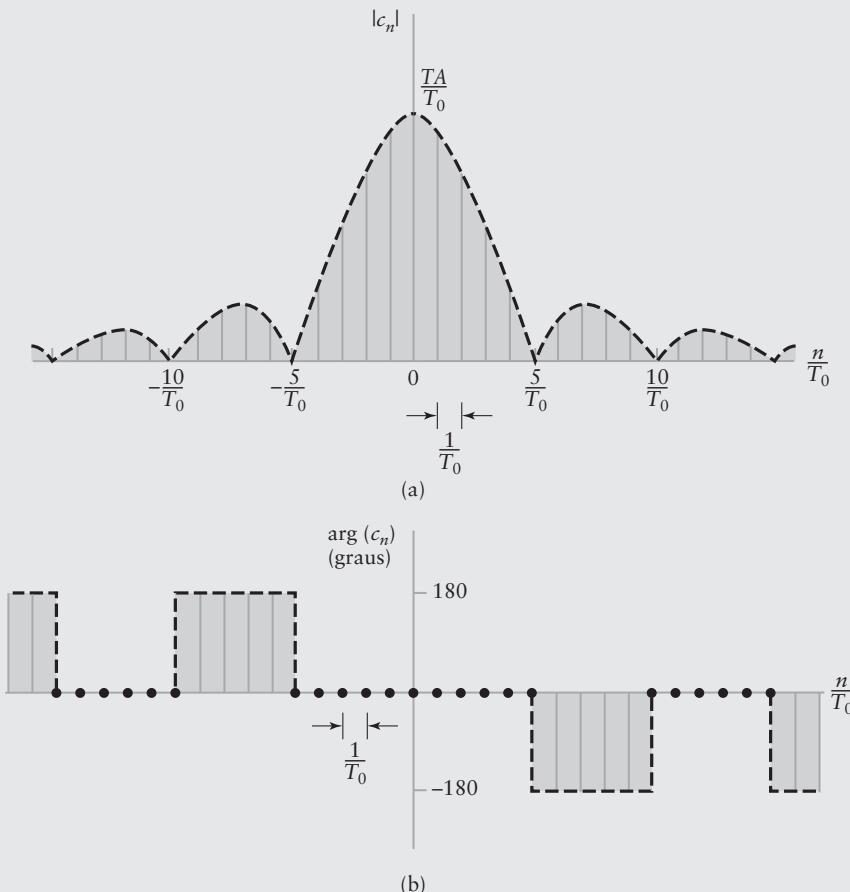


Figura A2.2 Espectro discreto de um trem periódico de pulsos retangulares para um ciclo de trabalho $T/T_0 = 0,2$. (a) Espectro de amplitude. (b) Espectro de fase.

A2.1 TRANSFORMADA DE FOURIER

Na seção anterior, utilizamos a série de Fourier para representar um sinal periódico. Queremos, agora, desenvolver uma representação similar para um sinal $g(t)$ que não é periódico,

em termos de sinais exponenciais complexos. Para isto, primeiro construímos uma função periódica $g_{T_0}(t)$ de período T_0 de tal forma que $g(t)$ define um ciclo desta função periódica, como ilustrado na Fig. A2.3. No limite, fazemos o período T_0 tender ao infinito, de tal forma que podemos escrever

$$g(t) = \lim_{T_0 \rightarrow \infty} g_{T_0}(t) \quad (\text{A2.21})$$

Representando a função periódica $g_{T_0}(t)$ em termos da forma exponencial complexa da série de Fourier, temos

$$g_{T_0}(t) = \sum_{n=-\infty}^{\infty} c_n \exp\left(\frac{j2\pi nt}{T_0}\right) \quad (\text{A2.22})$$

na qual

$$c_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_{T_0}(t) \exp\left(-\frac{j2\pi nt}{T_0}\right) dt \quad (\text{A2.23})$$

Escrevemos os expoentes propositadamente como mostrado nas Eqs. (A2.22) e (A2.23) porque queremos fazer T_0 tender ao infinito de acordo com a Eq. (A2.21). Definindo

$$\Delta f = \frac{1}{T_0} \quad (\text{A2.24})$$

$$f_n = \frac{n}{T_0} \quad (\text{A2.25})$$

e

$$G(f_n) = c_n T_0 \quad (\text{A2.26})$$

Portanto, alterando a notação na representação da série de Fourier de $g_{T_0}(t)$ dada nas Eqs. (A2.22) e (A2.23), temos as seguintes relações para o intervalo $-T_0/2 \leq t \leq T_0/2$:

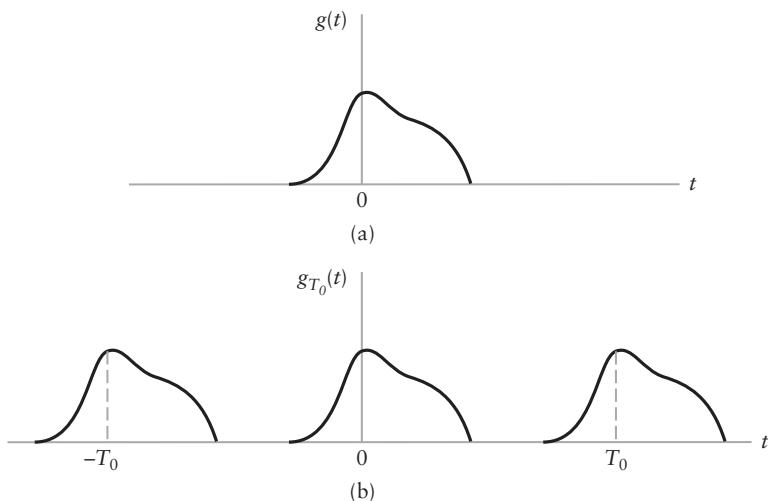


Figura A2.3 Ilustração da utilização de uma função arbitrariamente definida para construir uma forma de onda periódica. (a) Função no tempo $g(t)$ definida arbitrariamente. (b) Forma de onda periódica $g_{T_0}(t)$ baseada em $g(t)$.

$$g_{T_0}(t) = \sum_{n=-\infty}^{\infty} G(f_n) \exp(j2\pi f_n t) \Delta f \quad (\text{A2.27})$$

na qual

$$G(f_n) = \int_{-T_0/2}^{T_0/2} g_{T_0}(t) \exp(-j2\pi f_n t) dt \quad (\text{A2.28})$$

Agora, fazemos o período T_0 tender ao infinito ou, de forma equivalente, seu recíproco Δf tender a zero. Então, vemos que, no limite, a freqüência discreta f_n se aproxima da variável de freqüência contínua f e que o somatório da Eq. (A2.27) se torna uma integral definindo a área sob a função contínua de freqüência f – ou seja, $G(f)\exp(j2\pi ft)$. Além disso, quando T_0 tender ao infinito, a função $g_{T_0}(t)$ se aproxima de $g(t)$. Portanto, no limite, as Eqs. (A2.27) e (A2.28) se tornam, respectivamente,

$$g(t) = \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft) df \quad (\text{A2.29})$$

na qual

$$G(f) = \int_{-\infty}^{\infty} g(t) \exp(-j2\pi ft) dt \quad (\text{A2.30})$$

Portanto, atingimos nosso objetivo de representar um sinal $g(t)$ arbitrariamente definido em termos de funções exponenciais no intervalo completo ($-\infty < t < \infty$). Dada a função $g(t)$, a Eq. (A2.30) define a transformada de Fourier $G(f)$. Por outro lado, a Eq. (A2.29) define a transformada inversa de Fourier de $G(f)$.

FUNÇÕES DE BESSSEL

A3.1 SOLUÇÃO EM SÉRIE DA EQUAÇÃO DE BESSSEL

Em sua forma mais básica, uma *equação de Bessel de ordem n* é escrita como

$$x^2 \frac{d^2y}{dx^2} + x \frac{dy}{dx} + (x^2 - n^2)y = 0 \quad (\text{A3.1})$$

a qual é uma das mais importantes equações diferenciais com todos os coeficientes variáveis. Para cada ordem n , uma solução desta equação é definida pela série de potência

$$J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{1}{2}x\right)^{n+2m}}{m!(n+m)!} \quad (\text{A3.2})$$

A função $J_n(x)$ é chamada de *função de Bessel de ordem n do primeiro tipo*. A Eq. (A3.1) possui duas funções de coeficiente, $1/x$ e $(1 - n^2/x^2)$. Logo, ela não possui pontos singulares finitos exceto na origem. Temos, portanto, que a expansão em série da Eq. (A3.2) converge para todo $x > 0$. A Eq. (A3.2) pode, pois, ser utilizada para calcular numericamente $J_n(x)$ para $n = 0, 1, 2, \dots$. A Tabela A3.1 apresenta os valores de $J_n(x)$ para diferentes ordens n e variando x . É interessante notar que os gráficos de $J_0(x)$ e $J_1(x)$ são semelhantes aos gráficos de $\cos x$ e $\sin x$, respectivamente. Veja os gráficos da Fig. 4.6 no Capítulo 4.

TABELA A3.1 Tabela das funções de Bessel^a

$n \setminus x$	$J_n(x)$									
	0,5	1	2	3	4	6	8	10	12	
0	0,9385	0,7652	0,2239	-0,2601	-0,3971	0,1506	0,1717	-0,2459	0,0477	
1	0,2423	0,4401	0,5767	0,3391	-0,0660	-0,2767	0,2346	0,0435	-0,2234	
2	0,0306	0,1149	0,3528	0,4861	0,3641	-0,2429	-0,1130	0,2546	-0,0849	
3	0,0026	0,0196	0,1289	0,3091	0,4302	0,1148	-0,2911	0,0584	0,1951	
4	0,0002	0,0025	0,0340	0,1320	0,2811	0,3576	-0,1054	-0,2196	0,1825	
5	-	0,0002	0,0070	0,0430	0,1321	0,3621	0,1858	-0,2341	-0,0735	
6		-	0,0012	0,0114	0,0491	0,2458	0,3376	-0,0145	-0,2437	
7			0,0002	0,0025	0,0152	0,1296	0,3206	0,2167	-0,1703	
8			-	0,0005	0,0040	0,0565	0,2235	0,3179	0,0451	
9				0,0001	0,0009	0,0212	0,1263	0,2919	0,2304	
10				-	0,0002	0,0070	0,0608	0,2075	0,3005	
11					-	0,0020	0,0256	0,1231	0,2704	
12						0,0005	0,0096	0,0634	0,1953	
13						0,0001	0,0033	0,0290	0,1201	
14						-	0,0010	0,0120	0,0650	

^a Para uma tabela mais completa das funções de Bessel, veja Abramowitz e Stegun (1965, pp. 358-406).

A função $J_n(x)$ também por ser descrita na forma de uma integral por

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sen \theta - n\theta) d\theta \quad (\text{A3.3})$$

ou, de forma equivalente, utilizando a notação complexa, por

$$J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(jx \sen \theta - jn\theta) d\theta \quad (\text{A3.4})$$

A3.2 PROPRIEDADES DA FUNÇÃO DE BESSSEL

A função de Bessel $J_n(x)$ possui certas propriedades:

► **PROPRIEDADE 1**

$$J_n(x) = (-1)^n J_{-n}(x) \quad (\text{A3.5})$$

Para provar esta relação, substituímos θ por $(\pi - \theta)$ na Eq. (A3.3). Então, observando que $\sen(\pi - \theta) = \sen \theta$, temos

$$\begin{aligned} J_n(x) &= \frac{1}{\pi} \int_0^\pi \cos(x \sen \theta + n\theta - n\pi) d\theta \\ &= \frac{1}{\pi} \int_0^\pi [\cos(n\pi) \cos(x \sen \theta + n\theta) + \sen(n\pi) \sen(x \sen \theta + n\theta)] d\theta \end{aligned}$$

Para valores inteiros de n , temos

$$\begin{aligned} \cos(n\pi) &= (-1)^n \\ \sen(n\pi) &= 0 \end{aligned}$$

Portanto,

$$J_n(x) = \frac{(-1)^n}{\pi} \int_0^\pi \cos(x \sen \theta + n\theta) d\theta \quad (\text{A3.6})$$

Da Eq. (A3.3), também vemos que, substituindo n por $-n$:

$$J_{-n}(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sen \theta + n\theta) d\theta \quad (\text{A3.7})$$

A resultado desejado segue imediatamente da Eq. (A3.6) e (A3.7).

► **PROPRIEDADE 2**

$$J_n(x) = (-1)^n J_n(-x) \quad (\text{A3.8})$$

Esta relação é obtida substituindo x por $-x$ na Eq. (A3.3) e, então, usando a Eq. (A3.6).

► **PROPRIEDADE 3** *Para pequenos valores de x , temos*

$$J_n(x) \simeq \frac{x^n}{2^n n!} \quad (\text{A3.9})$$

Esta relação é obtida mantendo, simplesmente, apenas o primeiro termo na série de potência da Eq. (A3.2), e ignorando os termos de mais alta ordem. Portanto, quando x é pequeno, temos

$$J_0(x) \simeq 1$$

$$J_1(x) \simeq \frac{x}{2}$$

$$J_n(x) \simeq 0 \quad \text{para } n > 1 \quad (\text{A3.10})$$

► **PROPRIEDADE 4** $\sum_{n=-\infty}^{\infty} J_n^2(x) = 1 \quad \text{para todo } x \quad (\text{A3.11})$

Para provar esta propriedade, procedemos como mostrado a seguir. Observamos que $J_n(x)$ é real. Logo, multiplicando a Eq. (A3.4) por seu próprio complexo conjugado e somando em todos os possíveis valores de n , temos

$$\sum_{n=-\infty}^{\infty} J_n^2(x) = \frac{1}{(2\pi)^2} \sum_{n=-\infty}^{\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \exp(jx \sin \theta - jn\theta - jx \sin \phi + jn\phi) d\theta d\phi$$

Alterando a ordem da integração dupla e do somatório:

$$\sum_{n=-\infty}^{\infty} J_n^2(x) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} d\theta \int_{-\pi}^{\pi} d\phi \exp[jx(\sin \theta - \sin \phi)] \sum_{n=-\infty}^{\infty} \exp[jn(\phi - \theta)] \quad (\text{A3.12})$$

Usando a seguinte relação da teoria de transformada de Fourier para a função delta (veja o Capítulo 2)

$$\delta(\phi) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \exp(jn\phi) \quad (\text{A3.13})$$

na Eq. (A3.12) e, então, aplicando a propriedade de peneiramento da função delta, finalmente obtemos

$$\sum_{n=-\infty}^{\infty} J_n^2(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta = 1$$

a qual é o resultado desejado.

A função de Bessel $J_n(x)$ possui várias outras propriedades. Entretanto, considerando o escopo deste livro, as propriedades 1 a 4 discutem tudo o que precisamos.

Apêndice 4

A FUNÇÃO Q E SUA RELAÇÃO COM A FUNÇÃO ERRO

A4.1 A FUNÇÃO Q

Considere a variável aleatória Gaussiana *normalizada* u com média nula e variância unitária. A probabilidade de um valor observado de u ser maior do que x define a função Q .

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{u^2}{2}\right) du, \quad x \geq 0 \quad (\text{A4.1})$$

Em palavras, a função Q é igual a área sob a calda positiva da função densidade Gaussiana de média nula e variância unitária.

Da equação de definição (A4.1), as seguintes propriedades da função Q são obtidas:

- **PROPRIEDADE 1** Para $x = 0$, temos o valor exato

$$Q(0) = \frac{1}{2} \quad (\text{A4.2})$$

- **PROPRIEDADE 2** Para $x < 0$, o valor correspondente da função Q é dado pela relação

$$Q(x) = 1 - Q(|x|) \quad (\text{A4.3})$$

- **PROPRIEDADE 3** Um limite útil da função Q é dado por

$$Q(x) \leq \frac{1}{2} \exp\left(-\frac{1}{2}x^2\right), \quad x \geq 0 \quad (\text{A4.4})$$

A Tabela A4.1 mostra alguns poucos valores tabulados de $Q(x)$ para $0 \leq x \leq 5$. As propriedades descritas nas Eqs. (A4.2) a (A4.4) são confirmadas examinando-se as linhas desta tabela.

Os valores da função Q listados nesta tabela para $0 \leq x \leq 3,7$ foram arredondados para 5 casas decimais significativas. Para $3,8 \leq x \leq 5$, fizemos o arredondamento para 7 casas decimais.

A4.2 RELAÇÃO DA FUNÇÃO Q COM A FUNÇÃO ERRO COMPLEMENTAR

A função Q foi apresentada no Capítulo 8, sendo utilizada para calcular o efeito do ruído em comunicações digitais como discutido no Capítulo 10. Na literatura, geralmente vemos este efeito formulado em termos de outra função, chamada de função erro complementar.

Neste contexto, primeiramente definimos a função *erro* como

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-u^2) du \quad (\text{A4.5})$$

A função erro possui duas propriedades úteis

TABELA A4.1 Tabela de valores da função Q^a

x	$Q(x)$	x	$Q(x)$
0,0	0,50000	2,2	0,01390
0,1	0,46017	2,3	0,01072
0,2	0,42074	2,4	0,00820
0,3	0,38209	2,5	0,00621
0,4	0,34458	2,6	0,00466
0,5	0,30854	2,7	0,00347
0,6	0,27425	2,8	0,00256
0,7	0,24196	2,9	0,00187
0,8	0,21186	3,0	0,00135
0,9	0,18406	3,1	0,00097
1,0	0,15866	3,2	0,00069
1,1	0,13567	3,3	0,00048
1,2	0,11507	3,4	0,00034
1,3	0,09680	3,5	0,00023
1,4	0,08076	3,6	0,00016
1,5	0,06681	3,7	0,00011
1,6	0,05480	3,8	$7,24 \times 10^{-5}$
1,7	0,04457	3,9	$4,81 \times 10^{-5}$
1,8	0,03593	4,0	$3,17 \times 10^{-5}$
1,9	0,02872	4,30	$0,85 \times 10^{-5}$
2,0	0,02275	4,65	$0,17 \times 10^{-5}$
2,1	0,01786	5,00	$0,03 \times 10^{-5}$

^a A Tabela A4.1 é adaptada de Abramowitz e Stegun (1964), pp. 966-972. Este livro de consulta tabula a função densidade de probabilidade Gaussiana (normal).

► **PROPRIEDADE 1** $\text{erf}(-x) = -\text{erf}(x)$ (A4.6)

Esta propriedade é chamada de *relação de simetria*.

► **PROPRIEDADE 2** $\frac{2}{\sqrt{\pi}} \int_0^\infty \exp(-u^2) du = 1$ (A4.7)

Uma função relacionada, chamada de *função erro complementar*, é definida por

$$\begin{aligned} \text{erfc}(x) &= \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-u^2) du \\ &= 1 - \text{erf}(x) \end{aligned} \quad (\text{A4.8})$$

Examinando as Eqs. (A4.1) e a primeira linha da Eq. (A4.8), rapidamente vemos que a função Q e a função erro complementar são relacionadas por

$$Q(x) = \frac{1}{2} \text{erfc}\left(\frac{x}{\sqrt{2}}\right) \quad (\text{A4.9})$$

O inverso desta relação é dado por

$$\text{erfc}(x) = 2Q\left(\sqrt{2}x\right) \quad (\text{A4.10})$$

Logo, dada a função Q , podemos utilizar a Eq. (A4.10) para calcular o valor correspondente da função erro complementar para o valor x . Alternativamente, dada a função erro complementar, podemos calcular a função Q correspondente utilizando a Eq. (A4.9).

Apêndice 5

DESIGUALDADE DE SCHWARZ

Sejam $g_1(t)$ e $g_2(t)$ funções da variável real t no intervalo $a \leq t \leq b$. Assumindo que $g_1(t)$ e $g_2(t)$ satisfazem as condições

$$\int_a^b |g_1(t)|^2 dt < \infty \quad (\text{A5.1})$$

$$\int_a^b |g_2(t)|^2 dt < \infty \quad (\text{A5.2})$$

Então, de acordo com a *desigualdade de Schwarz*, temos

$$\left| \int_a^b g_1(t)g_2(t) dt \right|^2 \leq \int_a^b |g_1(t)|^2 dt \int_a^b |g_2(t)|^2 dt \quad (\text{A5.3})$$

Para provar esta desigualdade, primeiro formamos a integral

$$\int_a^b [\lambda g_1^*(t) + g_2^*(t)][\lambda g_1(t) + g_2(t)] dt = \lambda^2 A + \lambda(B + B^*) + C \quad (\text{A5.4})$$

na qual λ é uma variável real, o asterisco representa complexo conjugado e

$$A = \int_a^b |g_1(t)|^2 dt \geq 0 \quad (\text{A5.5})$$

$$B = \int_a^b g_1^*(t)g_2(t) dt \quad (\text{A5.6})$$

$$C = \int_a^b |g_2(t)|^2 dt \geq 0 \quad (\text{A5.7})$$

A integral da Eq. (A5.4) existe, é real e é uma função não negativa de λ , dita $f(\lambda)$. Como $f(\lambda)$ é não negativa, ela não deve ter raízes reais exceto, possivelmente, uma raiz dupla. Da fórmula quadrática, devemos ter, então,

$$(B + B^*)^2 \leq 4AC \quad (\text{A5.8})$$

Note que $(B + B^*)/2$ é igual à parte real de B . Substituindo as Eqs. (A5.5) a (A5.7) em (A5.8), temos

$$\left\{ \int_a^b [g_1^*(t)g_2(t) + g_1(t)g_2^*(t)] dt \right\}^2 \leq 4 \int_a^b |g_1(t)|^2 dt \int_a^b |g_2(t)|^2 dt \quad (\text{A5.9})$$

Esta é a forma mais geral da desigualdade de Schwarz apropriada para funções complexas $g_1(t)$ e $g_2(t)$. Quando tanto $g_1(t)$ quanto $g_2(t)$ são reais, temos

$$g_1^*(t)g_2(t) + g_1(t)g_2^*(t) = 2g_1(t)g_2(t) \quad (\text{A5.10})$$

e a Eq. (A5.3) é obtida imediatamente.

Note que a igualdade é obtida [além do caso trivial quando tanto $g_1(t)$ quanto $g_2(t)$ são nulas] quando a raiz dupla existe na Eq. (A5.4), ou seja, quando

$$\lambda g_1(t) + g_2(t) = \lambda g_1^*(t) + g_2^*(t) = 0 \quad (\text{A5.11})$$

Como λ é real, $g_1(t)$ e $g_2(t)$ são linearmente relacionadas. Olhando para o problema de um ponto de vista um pouco diferente, vemos que existe um valor real de λ para o qual a Eq. (A5.4) é zero e para o qual a primeira derivada com respeito a λ desaparece.

$$2\lambda A + (B + B^*) = 0 \quad (\text{A5.12})$$

ou

$$\lambda = -\frac{B + B^*}{2A} = -\frac{\int_a^b [g_1^*(t)g_2(t) + g_1(t)g_2^*(t)] dt}{2 \int_a^b |g_1(t)|^2 dt} \quad (\text{A5.13})$$

Esta relação é válida se e somente se

$$g_2(t) = -\lambda g_1(t) \quad (\text{A5.14})$$

Esta última relação é equivalente à Eq. (A5.11).

Apêndice 6

TABELAS MATEMÁTICAS

Este apêndice apresenta tabelas para (1) teoremas da transformada de Fourier, (2) pares transformada de Fourier, (3) pares transformada de Hilbert, (4) identidades trigonométricas, (5) expansões em série, (6) integrais indefinidas e definidas, (7) somatórios, (8) constantes úteis e (9) prefixos de unidades recomendados.

TABELA A6.1 Teoremas da transformada de Fourier

Propriedades	Descrição matemática
1. Linearidade	$ag_1(t) + bg_2(t) \Leftrightarrow aG_1(f) + bG_2(f)$ na qual a e b são constantes
2. Dilatação (escalonamento no tempo)	$g(at) \Leftrightarrow \frac{1}{ a } G\left(\frac{f}{a}\right)$ na qual a é uma constante
3. Dualidade	Se $g(t) \Leftrightarrow G(f)$, então $G(t) \Leftrightarrow g(-f)$
4. Deslocamento no tempo	$g(t - t_0) \Leftrightarrow G(f) \exp(-j2\pi f t_0)$
5. Deslocamento na freqüência	$\exp(j2\pi f_c t)g(t) \Leftrightarrow G(f - f_c)$
6. Área sob $g(t)$	$\int_{-\infty}^{\infty} g(t) dt = G(0)$
7. Área sob $G(f)$	$g(0) = \int_{-\infty}^{\infty} G(f) df$
8. Diferenciação no domínio do tempo	$\frac{d}{dt}g(t) \Leftrightarrow j2\pi f G(f)$
9. Integração no domínio do tempo	$\int_{-\infty}^t g(\tau) d\tau \Leftrightarrow \frac{1}{j2\pi f} G(f) + \frac{G(0)}{2} \delta(f)$
10. Funções conjugadas	Se $g(t) \Leftrightarrow G(f)$, então $g^*(t) \Leftrightarrow G^*(-f)$,
11. Multiplicação no domínio do tempo	$g_1(t)g_2(t) \Leftrightarrow \int_{-\infty}^{\infty} G_1(\lambda)G_2(f - \lambda) d\lambda$
12. Convolução no domínio do tempo	$\int_{-\infty}^{\infty} g_1(\tau)g_2(t - \tau) d\tau \Leftrightarrow G_1(f)G_2(f)$
13. Teorema da correlação	$\int_{-\infty}^{\infty} g_1(t)g_2^*(t - \tau) dt \Leftrightarrow G_1(f)G_2^*(f)$
14. Teorema de Rayleigh da energia	$\int_{-\infty}^{\infty} g(t) ^2 dt = \int_{-\infty}^{\infty} G(f) ^2 df$

TABELA A6.2 Pares transformada de Fourier

<i>Função no tempo</i>	<i>Transformada de Fourier</i>
$\text{ret}\left(\frac{t}{T}\right)$	$T \text{sinc}(fT)$
$\text{sinc}(2Wt)$	$\frac{1}{2W} \text{ret}\left(\frac{f}{2W}\right)$
$\exp(-at)u(t), \quad a > 0$	$\frac{1}{a + j2\pi f}$
$\exp(-a t), \quad a > 0$	$\frac{2a}{a^2 + (2\pi f)^2}$
$\exp(-\pi t^2)$	$\exp(-\pi f^2)$
$\begin{cases} 1 - \frac{ t }{T}, & t < T \\ 0, & t \geq T \end{cases}$	$T \text{sinc}^2(fT)$
$\delta(t)$	1
1	$\delta(f)$
$\delta(t - t_0)$	$\exp(-j2\pi f t_0)$
$\exp(j2\pi f_c t)$	$\delta(f - f_c)$
$\cos(2\pi f_c t)$	$\frac{1}{2} [\delta(f - f_c) + \delta(f + f_c)]$
$\text{sen}(2\pi f_c t)$	$\frac{1}{2j} [\delta(f - f_c) - \delta(f + f_c)]$
$\text{sgn}(t)$	$\frac{1}{j\pi f}$
$\frac{1}{\pi t}$	$-j \text{ sgn}(f)$
$u(t)$	$\frac{1}{2} \delta(f) + \frac{1}{j2\pi f}$
$\sum_{i=-\infty}^{\infty} \delta(t - iT_0)$	$\frac{1}{T_0} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_0}\right)$

Notas: $u(t)$ = função degrau unitário $\delta(t)$ = função delta de Dirac $\text{ret}(t)$ = função retangular $\text{sgn}(t)$ = função sinal $\text{sinc}(t)$ = função sinc**TABELA A6.3** Pares transformada de Hilbert^a

<i>Função no tempo</i>	<i>Transformada de Hilbert</i>
$m(t) \cos(2\pi f_c t)$	$m(t) \text{ sen}(2\pi f_c t)$
$m(t) \text{ sen}(2\pi f_c t)$	$-m(t) \cos(2\pi f_c t)$
$\cos(2\pi f_c t)$	$\text{sen}(2\pi f_c t)$
$\text{sen}(2\pi f_c t)$	$-\cos(2\pi f_c t)$
$\delta(t)$	$\frac{1}{\pi t}$
$\frac{1}{t}$	$-\pi \delta(t)$

^a Nos dois primeiros pares, foi assumido que $m(t)$ é limitado em faixa ao intervalo $-W \leq f \leq W$, na qual $W < f_c$.

TABELA A6.4 Identidades trigonométricas

$$\begin{aligned}
 \exp(\pm j\theta) &= \cos \theta \pm j \sin \theta \\
 \cos \theta &= \frac{1}{2}[\exp(j\theta) + \exp(-j\theta)] \\
 \sin \theta &= \frac{1}{2j}[\exp(j\theta) - \exp(-j\theta)] \\
 \sin^2 \theta + \cos^2 \theta &= 1 \\
 \cos^2 \theta - \sin^2 \theta &= \cos(2\theta) \\
 \cos^2 \theta &= \frac{1}{2}[1 + \cos(2\theta)] \\
 \sin^2 \theta &= \frac{1}{2}[1 - \cos(2\theta)] \\
 2 \sin \theta \cos \theta &= \sin(2\theta) \\
 \sin(\alpha \pm \beta) &= \sin \alpha \cos \beta \pm \cos \alpha \sin \beta \\
 \cos(\alpha \pm \beta) &= \cos \alpha \cos \beta \mp \sin \alpha \sin \beta \\
 \operatorname{tg}(\alpha \pm \beta) &= \frac{\operatorname{tg} \alpha \pm \operatorname{tg} \beta}{1 \mp \operatorname{tg} \alpha \operatorname{tg} \beta} \\
 \sin \alpha \sin \beta &= \frac{1}{2}[\cos(\alpha - \beta) - \cos(\alpha + \beta)] \\
 \cos \alpha \cos \beta &= \frac{1}{2}[\cos(\alpha - \beta) + \cos(\alpha + \beta)] \\
 \sin \alpha \cos \beta &= \frac{1}{2}[\sin(\alpha - \beta) + \sin(\alpha + \beta)]
 \end{aligned}$$

TABELA A6.5 Expansões em série

Série de Taylor	$f(x) = f(a) + \frac{f'(a)}{1!}(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + \cdots$
na qual	$f^{(n)}(a) = \frac{d^n f(x)}{dx^n} \Big _{x=a}$
Série de MacLaurin	$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots$
na qual	$f^{(n)}(0) = \frac{d^n f(x)}{dx^n} \Big _{x=0}$
Série binomial	$(1 + x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \cdots, \quad nx < 1$
Série exponencial	$\exp x = 1 + x + \frac{1}{2!}x^2 + \cdots$
Série logarítmica	$\ln(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots$
Série trigonométrica	$\begin{aligned} \operatorname{sen} x &= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots \\ \cos x &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots \\ \operatorname{tg} x &= x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots \\ \operatorname{sen}^{-1} x &= x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \cdots \\ \operatorname{tg}^{-1} x &= x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots, \quad x < 1 \\ \operatorname{sinc} x &= 1 - \frac{1}{3!}(\pi x)^2 + \frac{1}{5!}(\pi x)^4 - \cdots \end{aligned}$

TABELA A6.6 Integrais*Integrais indefinidas*

$$\int x \operatorname{sen}(ax) dx = \frac{1}{a^2} [\operatorname{sen}(ax) - ax \cos(ax)]$$

$$\int x \cos(ax) dx = \frac{1}{a^2} [\cos(ax) + ax \operatorname{sen}(ax)]$$

$$\int x \exp(ax) dx = \frac{1}{a^2} \exp(ax) (ax - 1)$$

$$\int x \exp(ax^2) dx = \frac{1}{2a} \exp(ax^2)$$

$$\int \exp(ax) \operatorname{sen}(bx) dx = \frac{1}{a^2 + b^2} \exp(ax) [a \operatorname{sen}(bx) - b \cos(bx)]$$

$$\int \exp(ax) \cos(bx) dx = \frac{1}{a^2 + b^2} \exp(ax) [a \cos(bx) + b \operatorname{sen}(bx)]$$

$$\int \frac{dx}{a^2 + b^2 x^2} = \frac{1}{ab} \operatorname{tg}^{-1}\left(\frac{bx}{a}\right)$$

$$\int \frac{x^2 dx}{a^2 + b^2 x^2} = \frac{x}{b^2} - \frac{a}{b^3} \operatorname{tg}^{-1}\left(\frac{bx}{a}\right)$$

Integrais definidas

$$\int_0^\infty \frac{x \operatorname{sen}(ax)}{b^2 + x^2} dx = \frac{\pi}{2} \exp(-ab), \quad a > 0, b > 0$$

$$\int_0^\infty \frac{\cos(ax)}{b^2 + x^2} dx = \frac{\pi}{2b} \exp(-ab), \quad a > 0, b > 0$$

$$\int_0^\infty \frac{\cos(ax)}{(b^2 - x^2)^2} dx = \frac{\pi}{4b^3} [\operatorname{sen}(ab) - ab \cos(ab)], \quad a > 0, b > 0$$

$$\int_0^\infty \operatorname{sinc} x dx = \int_0^\infty \operatorname{sinc}^2 x dx = \frac{1}{2}$$

$$\int_0^\infty \exp(-ax^2) dx = \frac{1}{2} \sqrt{\frac{\pi}{a}}, \quad a > 0$$

$$\int_0^\infty x^2 \exp(-ax^2) dx = \frac{1}{4a} \sqrt{\frac{\pi}{a}}, \quad a > 0$$

TABELA A6.7 Somatórios

$$\sum_{k=1}^K k = \frac{K(K+1)}{2}$$

$$\sum_{k=1}^K k^2 = \frac{K(K+1)(2K+1)}{6}$$

$$\sum_{k=1}^K k^3 = \frac{K^2(K+1)^2}{4}$$

$$\sum_{k=0}^{K-1} x^k = \frac{(x^K - 1)}{x - 1}$$

TABELA A6.8 Constantes úteis*Constantes físicas*

Constante de Boltzmann	$k = 1,38 \times 10^{-23}$ joule.Kelvin
Constante de Planck	$h = 6,626 \times 10^{-34}$ joule.segundo
Carga do elétron (fundamental)	$q = 1,602 \times 10^{-19}$ coulombs
Velocidade da luz no vácuo	$c = 2,998 \times 10^8$ metros/segundo
Temperatura padrão (absoluta)	$T_0 = 273$ Kelvin
Tensão térmica	$V_T = 0,026$ volts em temperatura ambiente
Energia térmica kT em temperatura padrão	$k T_0 = 3,77 \times 10^{-21}$ joule
Um hertz (Hz) = 1 ciclo/segundo; 1 ciclo = 2π radianos	
Um watt (W) = 1 joule/segundo	

Constantes matemáticas

Base do logaritmo natural	$e = 2,7182818$
logaritmo de e na base 2	$\log_2 e = 1,442695$
Logaritmo de 2 na base e	$\ln 2 = \log_e 2 = 0,693147$
Logaritmo de 2 na base 10	$\log_{10} 2 = 0,30103$
Pi	$\pi = 3,1415927$

TABELA A6.9 Prefixos de unidades recomendados

Múltiplos e submúltiplos	Prefixos	Símbolos
10^{12}	tera	T
10^9	giga	G
10^6	mega	M
10^3	quilo	K (k)
10^{-3}	mili	m
10^{-6}	micro	μ
10^{-9}	nano	n
10^{-12}	pico	p

RESPOSTAS DOS PROBLEMAS

NOTAS

1. O Capítulo 1 não possui problemas e, portanto, não possui respostas.
2. Vários dos problemas dos Capítulos 2 a 11 possuem respostas nos cabeçalhos dos próprios problemas e, portanto, não são referenciados aqui.

CAPÍTULO 2

2.1 $G(f) = \frac{2\pi f_c}{1 + 4\pi^2(f - f_c)^2}$

2.2 $g(t) = \frac{1}{\pi t}(e^{-j2\pi Wt} - 1)$

2.3 A parte imaginária de $G(0)$ deve ser zero.

2.4 A parte imaginária de $G(f)$ deve ser uma função ímpar de f .

2.11 $1 + T \sum_{m=1}^{\infty} \cos(2\pi m f_0 t) = T_0 \sum_{n=-\infty}^{\infty} \delta(f - n f_0), \quad f_0 = 1/T_0$

2.12 (a) É possível para um sistema ser causal mas instável.

(b) Pelo mesmo motivo, é possível para o sistema ser estável mas não causal.

2.13 $H(f) = \sqrt{2\pi\tau} \exp(-2\pi^2\tau^2f^2)$

2.14 (a) $|H(f)| = w_{\frac{N-1}{2}} + 2 \sum_{n=0}^{\frac{1}{2}(N-1)-1} w_n \cos(2\pi n f \Delta\tau)$

(b) $\arg(H(f)) = \exp\left(-j2\pi\left(\frac{N-1}{2}\right)f\Delta\tau\right)$

A implicação deste resultado é um atraso constante na transmissão do sinal ao longo do filtro, e, portanto, sem distorção de fase.

(c) O atraso de tempo é igual a $(N - 1)\Delta\tau/2$.

2.15 $R_{yx}^*(-\tau) = R_{xy}(\tau)$

2.16 $E_g(f) = \frac{1}{a^2 + 4\pi^2 f^2}, \quad -\infty < f < \infty$

2.17 $E_g(f) = \frac{4a^2}{(a^2 + 4\pi^2 f^2)^2}, \quad -\infty < f < \infty$

CAPÍTULO 3

- 3.1 O envelope do sinal AM pode assumir valor zéro se $k_a m(t) = -1$ para algum t , onde k_a é o fator de sensibilidade de amplitude e $m(t)$ é o sinal de mensagem.

3.2 (a) 98%

(b) 1%

3.3 $f_c > W$, onde f_c é a freqüência da portadora e W é a largura de faixa da mensagem.

3.4 (a) O conteúdo espectral da saída $v_2(t)$ consiste de duas partes:
Parte AM:

$$F[(a_1 + 2a_2 m(t)) \cos(2\pi f_c t)] = \frac{1}{2} a_1 (\delta(f - f_c) + \delta(f + f_c)) + a_2 (M(f - f_c) + M(f + f_c))$$

Parte indesejável:

$$F[a_1 m(t) + a_2 A_c^2 \cos^2(2\pi f_c t) + a_2 m^2(t)]$$

- (b) O filtro passa-faixa necessário deve ter largura de faixa $3W$ centrada em f_c , ou seja, as freqüências de corte do filtro são $f_c - W$ e $f_c + W$.
(c) Para extrair a onda AM, precisamos $f_c > 3W$.

3.5 A potência média na freqüência lateral superior ou inferior, expressa como percentual da potência média da onda modulada DSB-SC é 50%.

3.6 (a) $v(t) = \frac{1}{4} A_c A_m [\cos(2\pi(2f_c + f_m)t) + \cos(2\pi f_m t)] + \frac{1}{4} A_c A_m [\cos(2\pi(2f_c - f_m)t) + \cos(2\pi f_m t)]$

- (b) Os dois termos senoidais dentro do primeiro conjunto de colchetes são produzidos pela freqüência lateral superior em $f_c + f_m$. Os outros dois termos senoidais dentro do segundo conjunto de colchetes são produzidos pela freqüência lateral inferior em $f_c - f_m$.

3.7 A sobreposição espectral ocorre quando $f_c < W$. Quando isto acontece, parte da faixa lateral inferior move para a região de freqüência negativa e a parte correspondente da imagem da faixa lateral se move para a região de freqüência positiva, causando a sobreposição espectral. A detecção coerente falha pois o detector não é capaz de resolver parte do espectro no qual a sobreposição aconteceu.

3.12 Para o detector coerente recuperar o sinal de mensagem original (exceto por uma escala), precisamos que $f_c > W$.

CAPÍTULO 4

4.2 O fator de sensibilidade de fase k_p é medido em radianos/volt. O fator de sensibilidade de freqüência k_f é medido em hertz/volt. Portanto, a dimensão do termo do produto $2\pi k_f T$ é

$$(hertz/volt) \times (segundos) = \left(\frac{\text{ciclos}}{\text{segundos volt}} \right) \times (segundos) \\ = \text{radianos/volt}$$

a qual é a mesma dimensão de k_p .

4.4 (a) O envelope é

$$\begin{aligned} a(t) &= A_c \sqrt{1 + \beta^2 \sin^2(2\pi f_m t)} \\ &\approx A_c \left(1 + \frac{1}{2} \beta^2 \sin^2(2\pi f_m t) \right) \quad \text{para } \beta \text{ pequeno.} \\ &= A_c \left(1 + \frac{1}{4} \beta^2 - \frac{\beta^2}{4} \sin(4\pi f_m t) \right) \end{aligned}$$

O envelope máximo ocorre quando $\sin(2\pi f_m t) = \pm 1$, obtendo, portanto, o valor $A_c \left(1 + \frac{1}{2}\beta^2\right)$.

O envelope mínimo ocorre quando $\sin(2\pi f_m t) = 0$, obtendo, portanto, o valor A_c . Logo,

$$\frac{(\text{Envelope})_{\text{máx}}}{(\text{Envelope})_{\text{mín}}} = 1 + \frac{1}{2}\beta^2$$

(b) $\frac{\text{Potência média do FM de faixa estreita}}{\text{Potência média da portadora não modulada}} = \left(1 + \frac{1}{2}\beta^2\right)$

(c) A distorção harmônica é produzida pelo termo $\frac{\beta^3}{3} \sin^3(2\pi f_m t)$. Para $\beta = 0,3$ radianos, a distorção é $9 \times 10^{-3} \approx 1\%$, a qual é pequena suficiente para ser ignorada.

CAPÍTULO 5

5.5 (a) Taxa de Nyquist = 200 Hz, intervalo de Nyquist = 5ms.

(b) Taxa de Nyquist = 400 Hz, intervalo de Nyquist = 2,5 ms.

(c) A taxa de Nyquist e intervalo de Nyquist são as mesmas da parte (b).

5.6 (a) Para $T_s = 0,25\text{s}$

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} \cos\left(\frac{n\pi}{4}\right) \delta(t - 0,25n).$$

(b) Para $T_s = 1\text{s}$,

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} (-1)^n \delta(t - n).$$

(c) Para $T_s = 1,5\text{s}$

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} \cos(1,5n\pi) \delta(t - 1,5n).$$

5.7 A taxa de Nyquist deve exceder 1 Hz, correspondentemente, o intervalo de Nyquist deve ser menor do que 1s.

CAPÍTULO 6

6.1 Para a entrada 001101001, a forma de onda na saída do receptor consiste de $\text{sinc}(t/T_b)$ sempre que o símbolo 1 for transmitido e $-\text{sinc}(t/T_b)$ sempre que o símbolo 0 for transmitido. Não existirá interferência intersimbólica pois o pulso sinc vai a zero sempre que outro símbolo é transmitido.

6.4 (b) O valor associado ao atraso τ deve diminuir com o aumento de α :

- (i) $\tau \approx 5\text{s}$ para $\alpha = 0$
- (ii) $\tau \approx 3\text{s}$ para $\alpha = 1/2$
- (iii) $\tau \approx 2,5\text{s}$ para $\alpha = 1$

CAPÍTULO 7

- 7.6 Começando em zero, o diagrama aumenta linearmente atingindo π radianos na transmissão de 11, então diminui linearmente para $\pi/2$ na transmissão de 0, aumenta para π na transmissão de 1. O nível π é igual a $-\pi$ módulo 2π . Ele, então, aumenta linearmente de $-\pi$ para $\pi/2$ na transmissão de 111. Finalmente, ele diminui linearmente para $-\pi/2$ na transmissão de 00.
- 7.8 A característica linear entrada/saída do discriminador de freqüência possui as seguintes especificações:

- ▶ Ele ocupa a faixa de freqüência $f_c - \frac{1}{4}T_b < f < f_c + \frac{1}{4}T_b$.
- ▶ Ele vai a zero em $f = f_c$.

CAPÍTULO 8

8.1 0,82

8.2 $5,5 \times 10^{-5}$

8.3 $\frac{pp_0}{pp_0 + (1-p)p_1} \text{ e } \frac{pp_1}{pp_1 + (1-p)p_0}$

8.4 (a) 0,5

(b) 0,5

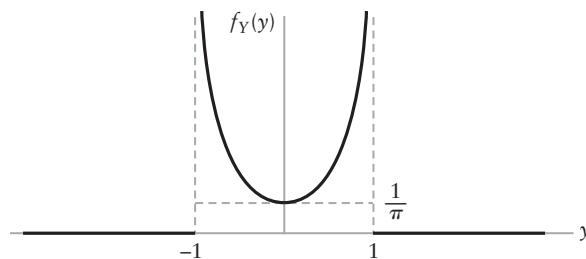
(c) 0,9999

(d) 0,0001

8.5 $(b+a)/2 \text{ e } (b-a^2)/12$

8.6 0 e 1

8.7
$$f_Y(y) = \begin{cases} 0, & y < -1 \\ \frac{1}{\pi\sqrt{1-y^2}}, & |y| < 1 \\ 0, & y > 1 \end{cases}$$



8.10 0 e 5/3.

8.11 Sim

8.13 $R_X(t_1, t_2) = \frac{1}{6} [\cos(2\pi f(t_1 - t_2)) + \cos(2\pi f(t_1 + t_2))]$; não.

8.14 $R_Y(n) = (\alpha_0^2 + \alpha_1^2)\sigma^2\delta(n) + \alpha_0\alpha_1\sigma^2(\delta(n-1) + \delta(n+1))$; sim.

8.15 $S_Y(k) = \beta_0 + 2\beta_1 \cos\left(\frac{2\pi k}{N}\right)$ e $S_Y(k)$ corresponde a freqüência $\frac{kf_s}{N}$, na qual f_s é a taxa de amostragem, $\beta_0 = (\alpha_0^2 + \alpha_1^2)\sigma^2$ e $\beta_1 = \alpha_0\alpha_1\sigma^2$.

8.16 Sim

8.17 (a) N_0

$$(b) S_Y(k) = N_0 \left| \frac{1 - (\alpha W^k)^N}{1 - \alpha W^k} \right|^2$$

CAPÍTULO 9

9.1 $\text{SNR} = \frac{(S + N) - (N)}{(N)}$ onde () indica medida de potência independente.

9.2 18 dB

9.3 $\text{SNR}_{\text{POST}}^{\text{AM}}/\text{SNR}_{\text{POST}}^{\text{DSB}}$ é 0,04 e 0,14.

9.4 A detecção de envelope é insensível ao deslocamento de fase.

9.5 10 dB e a largura de transmissão SSB é metade da DSB-SC.

9.6 8 kHz e 20,1 dB

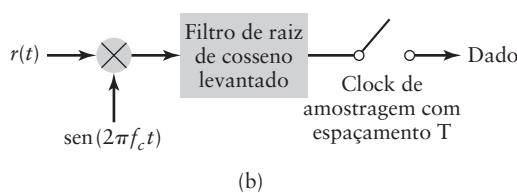
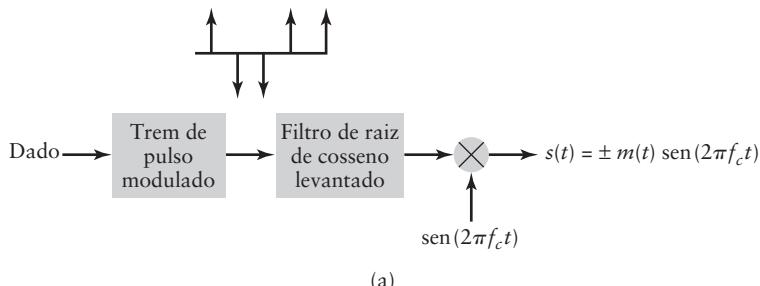
9.7 Canal superior é pior por 17,1 dB

9.8 34,7 dB e 48,3 dB

CAPÍTULO 1010.1 (a) $P(R_1|H_0)$ (b) $P[R_0|H_1]P[H_1] + P[R_1|H_0]P[H_0]$ 10.2 $cs^*(t)$ 10.3 $\sqrt{\alpha/2}$ 10.5 (a) $R(f) \exp(-j2\pi f/T)$ (b) $R^2(f) \exp(-j2\pi f/T)$ (c) $q(\tau) = m(\tau - lT)$ onde $m(\tau)$ é o formato de pulso do cosseno levantado
 $q(kT) = m((k - l)T) = \delta(k - l)$.

10.6 Largura de faixa PAM é 7,2 kHz. Largura de faixa BPSK é 14,4 kHz.

10.7



- 10.9 (a) $E[N_1 N_2] = 0$
(b) $\text{Var}(N_1 - N_2) = N_0 T$

CAPÍTULO 11

- 11.1 4×10^{-12} watts e 6,3 milivots
11.2 8×10^{-15} watts
11.3 2610°K
11.4 -147 dBm/Hz e -122 dBm/Hz
11.5 9,97dB
11.6 53,0 dBm/Hz e -122 dBm/Hz
11.7 -143 dBW
11.8 120°K, 147 dB e 13 db
11.9 13,9 m
11.10 132 dB e 101 dB
11.11 $R_{\text{médio}} = 0,83R_{\text{rms}}$

SCRIPTS MATLAB PARA EXPERIMENTOS DE COMPUTADOR DOS PROBLEMAS DOS CAPÍTULOS 7 A 10

PROBLEMA 7.30

```
%-----
% Script para o Problema 7.30(a)
% Modifique este script para o Problema 7.30(b)
%-----
Fs      = 32;          % (Hz) Taxa de amostragem (amostras por símbolo)
fe      = 10;          % (Hz) Freqüência da portadora
Nbts   = 500;         % Número de bits na seqüência
PulseShape = ones(1,Fs); % Forma de pulso retangular

%--- Gera duas seqüências aleatórias -----
bI      = sign(rand(1,Nbts)-0,5);
bQ      = sign(rand(1,Nbts)-0,5);

%--- Cria o trem de impulso -----
bI_t    = [1 zeros(1,Fs-1)] * bI;
bI_t    = bI_t(:);
bQ_t    = [1 zeros(1,Fs-1)] * bQ;
bQ_t    = bQ_t(:);

%--- Formata pulso e cria o sinal banda passante -----
bIp     = filter(PulseShape, 1, bI_t);
bQp     = filter(PulseShape, 1, bQ_t);

t = [1:length(bIp)]' / Fs; % Escala de tempo para o sinal banda passante
s = bIp.* cos(2*pi*fc*t) - bQp .* sen(2*pi*fc*t);

%--- Mostra os resultados -----
subplot(2,1,1), plot(t,s); % Mostra no tempo
ylabel('Amplitude'), xlabel('Tempo (s)'), axis([0, 50, -2, 2]) .

nFFT = 2^256; %Mostra na freqüência
[spec,freq] = spectrum (s,nFFT,nFFT / 4,nFFT/2,F s);
subplot(2,1,2), plot(freq, 10*log10(spec(:,1)/sqrt(nFFT)) )
ylabel('Espectro (dB)'), xlabel('Freqüência (Hz)'), axis([0, 16, -40,10])
```

PROBLEMA 7.31

```
%-----
% Modifique o script para o Problema 7.30 utilizando o seguinte
% formato de pulso de cosseno levantado
% Utilize a escala de tempo deslocada para evitar divisão por zero
%-----
B0      = 0,5;          % (Hz)
t       = [-2,001: 1/Fs : + 2,001]    % escala de tempo deslocada para o
                                         formato do pulso
rcos    = sinc(4*B0*t) ./ (1-16*B0^2*t.^2); % Da Eq. (6.20)
PulseShape = rcos;
```

PROBLEMA 8.54

```
%-----
% Distribuição de probabilidade com dado batizado
%-----
function a=Problem8_54;
N = 1000

for i=1:N
    X(i) = LoadedDie;
end
[F,X] = hist(X,[1:6]); % Traça o histograma dos resultados
stem(X,F);

function b = LoadedDie
% gera uma distribuição não uniforme em 1..6
b = floor(6*rand(1,1)^3) + 1;
```

PROBLEMA 8.55

```
%-----  
% distribuição de probabilidade de cinco variáveis uniformes  
%-----  
  
N      = 5;          % número de variáveis aleatórias a serem somadas  
nSmpls = 20000;      % número de amostras a serem geradas  
  
%---- Gera amostras de variáveis uniformes aleatórias ----  
Y      = 2*rand(N,nSmpls)-1;    % variável aleatória em [-1, +1]  
cSmpls = sum(Y);  
  
%---- Calcula e traça o histograma ----  
[N1,X] = hist(cSmpls,40); % histograma com 40 bins  
Delta   = X(2)-X(1);  
plot(X, N1/nSmpls/Delta); % normaliza o gráfico  
  
%---- compara com a teoria----  
hold on  
x      = [-5:0.01:5]  
sigma2 = N*(1/3);        % variância de uma uniforme é 1/3  
Gauss  = exp(-x.^2/2/sigma2) / sqrt(2*pi*sigma2);  
plot(x,Gauss,'r');  
hold off
```

PROBLEMA 8.58

```
N = 100000;      %---- número de amostras  
x = randn(1,N); %---- gera a seqüência Gaussiana  
  
%---- aproximação para o formato de pulso de um filtro de pulso de raiz de cosseno  
%    levantado com 50% de amortecimento (freqüência de corte em 1/8 da  
%    freqüência de amostragem  
rrc = [ 0,0015 -0,0082 -0,0075  0,0077  0,0212  0,0077...  
       -0,0375 -0,0784 -0,0531  0,0784  0,2894  0,4873...  
       0,5684  0,4873  0,2894  0,0784 -0,0531 -0,0784...  
       -0,0375  0,0077  0,0212  0,0077 -0,0075 -0,0082...  
       0,0015];  
z = filter(rrc,1,x);           % filtra o processo aleatório  
[P,F] = spectrum(z,256,0,Hanning(256),Fs); % calcula e traca o espectro  
plot(F,P(:,1));  
figure(2);  
Az = xcorr(z,25);            %---- calcula a autocorrelação para 25 atrasos  
plot(Az/max(abs(Az)));      %---- traça a autocorrelação normalizada  
grid
```

PROBLEMA 9.26

```

Fs = 1000; % taxa de amostragem
fe = 100; % freqüência da portadora
%-----
% filtro passa-faixa – Chebychev 1 de 6a. ordem
% passa-faixa de 50 a 150 Hz
%-----
Bt = 100; % largura de faixa da banda passante
b = [ 1 0 -3 0 3 0 -1];
a = [1,0000 -4,3705 8,6828 -10,0000 7,0486 -Z,8825 0,5398];
G = 0,0115;
%-----
SNRdBr = 30;
ka = 0,3;
t = [0:1/Fs:2];
A = sqrt(2); % para potência unitária da portadora
sigma2 = (1+ka^2*0,5) * 10^(-SNRdbr/10) * (Fs/2)/Bt;
sigma = sqrt(sigma2);
m = sen(2*pi*2*t);

%---- Modulador -----
e = cos(2*pi*fc*t);
AM = A*(1+ka*m).*c;

%---- adiciona o ruído -----
AMn = AM + sigma * randn(size(AM));

%---- Filtro passa-faixa -----
RxAMn = G*filter(b,a,AMn);
RxAM = G*filter(b,a,AM); % sem ruído

%---- RSR pré-detecção -----
C_N = 10*log10(sum(RxAM.^2/sum((RxAMn-RxAM).^2))

%---- Detector de envelope -----
BB_sig = EnvelopeDetector(RxAM); % sinal banda base limpo
BB_sign = EnvelopeDetector(RxAMn); % sinal banda base com ruído

%---- Calcula a RSR pós-detecção -----
error = sum((BB_sig - BB_sign).^2);
C = sum(BB_sig.^2);
SNRdBpost = 10*log10(C/error)

%---- traça a saída -----
figure(l);
plot(t, BB_sig); xlabel('tempo (s)'), ylabel('Amplitude')
hold on, plot(t, BB_sign,'g'), hold off

% veja Miscelânea para a descrição do detector de envelope

```

PROBLEMA 9.27

```

CNdB = [3:1:22]; % RSR pré-detecção
Fs = 500; % taxa de amostragem (Hz)
t = [0: (1/Fs) : 500-1/Fs); % instantes de amostragem do sinal

%--- Modulador FM -----
A = 1;
kf = 20,0; % sensibilidade FM
fm = 1; % freqüência do tom modulante (Hz)
m = sen(2*pi*fm*t); % sinal de mensagem
Int_m = cos(2*pi*fm*t) /2/pi; % sinal integrado
FM = A * exp(j*2*pi*kf*Int_m); % representação fasorial

for i=1:length(CNdB)
    %--- Gera o ruído banda base limitado em faixa ---
    Noise = LowPassNoise(CNdB(i), length(FM));

    %--- Gera a portadora -----
    fc = Fs/10; % freqüência da portadora para a demodulação
    Carrier = exp(j*2*pi*fc * [1:length(FM)]/Fs );

    %--- Converte para cima e demodula -----
    Noisy_message = FMdiscriminator( (FM + Noise) . * Carrier, Fs);
    Clean_message = FMdiscriminator( (FM) . * Carrier, Fs);

    PreSNR(i) = 20*log10(std(FM)/std(Noise));
    %--- Calcula a RSR pós-detecção -----
    Noise_power = sum((Noisy_message - Clean_message).^2);
    Signal_power = sum(Clean_message.^2);
    SNRdB(i) = 10*log10(Signal_power/Noise_power);

    W = 5,5; % largura de faixa do ruído do filtro passa-baixa utilizado no discriminador
    Bt = Fs/8; % largura de faixa de transmissão
    Theory(i) = 10*log10 ( 3* A^2*kf^2*0,5 / (2* (std(Noise)^2/Bt)*W^3));

    [CNdB(i) SNRdB(i) Theory(i)]
end

%--- traça os resultados -----
plot(PreSNR, SNRdB)
hold on, plot(PreSNR, Theory,'g'), hold off
grid on, xlabel ('C/N (dB)'), ylabel ('RSR pós-detecção (dB)')

```

% veja Miscelânea para a descrição do FMdiscriminator

PROBLEMA 10.30

```

Eb_N0 = 2;
Nbts = 10000; % número de bits, aumente para Eb/N0 maior
Fs = 4; % número de amostras por bit
b = sign(randn(1,Nbts)); % dado aleatório
pulse = ones(1,Fs);
Eb = sum(pulse.^2);
N0 = Eb / Eb_N0;
%---- executa a formatação de pulso retangular-----
S = pulse' * b;
S = S(:);
plot(S(1:100));
[P,F] = spectrum(S,256,0,Hanning(256),Fs);
plot(F,P(:,l));
%---- adiciona o ruído Gaussiano
Noise = sqrt(N0/2)*randn(size(S));
R = S + Noise;

%---- integra e descarta ---
D = sum(reshape(R,Fs,Nbts));
D = sign(D); % faz a decisão do bit
%---- conta os erros-----
Nerrs = (Nbts - sum(D.*b))/2;
BER = Nerrs/Nbts

```

PROBLEMA 10.31

```

Eb_N0 = 2
Nbts = 10000; % número de bits para simular
Fs = 4; % número de amostras por bit
b = sign(randn(1,Nbts)); % dado aleatório

%---- forma do pulso de raiz de cosseno levantado: 100% rolloff ----
pulse = [ 0,0064 0,0000 -0,0101 0,0000 0,0182 -0,0000 -0,0424...
          0,0000 0,2122 0,5000 0,6367 0,5000 0,2122 -0,0000...
          0,0424 0,0000 0,0182 -0,0000 -0,0101 0,0000 0,0064];
Delay = length(pulse);
Eb = sum(pulse.^2);
N0 = Eb / Eb_N0;

%---- executa a formatação de pulso de raiz de cosseno levantado ----
b_delta = [1 zeros(1,Fs-1)]' * b; % transforma b em uma seqüência de funções delta
b_delta = b_delta(:)';
S = filter(pulse,1,[b_delta zeros(1,Delay)]); % adiciona zeros para atraso do filtro

%---- adiciona ruído Gaussiano ----
Noise = sqrt(N0/2)*randn(size(S));
R = S + Noise;

```

```
%---- detecção por filtro casado -----
De = filter(pulse,1,R);
D = sign(De(Delay:Fs:end)); % amostra a saída no instante adequado
D = D(1:Nbits)';
%---- conta os erros-----
Nerrs = (Nbites - sum(D.*b))/2;
BER = Nerrs/Nbites

%---- traça o diagrama de olho-----
ploteye(De(Delay:end), Fs)
```

MISCELÂNEA

```
function BB_sig = EnvelopeDetector(AM_sig);
%-----
% Filtro passa-baixa – Butterworth de 4a. ordem com freqüência de corte = 50Hz e
% taxa de amostragem de 1 kHz
%-----
bLP = [1 4 6 4 1];
aLP = [1,0000 -3,1806 3,8612 -2,1122 0,4383];
GLP = 4,1660e-004;

%---- modelo do detector de envelope -----
AM_env(1) = AM_sig(1);
decay = 0,01; % decaimento do sinal em um período de amostragem

for i= 2:length(AM_sig);
    AM_env(i) = max(AM_sig(i), AM_env(i-1) - decay);
end
AM_rec = AM_env - mean(AM_env);

%---- filtro passa-baixa-----
BB_sig = GLP * filter(bLP, aLP, AM_rec);

function Message = FMdiscriminator(FM,Fsample);

%---- filtro passa-baixa: Resposta finita ao impulso
%---- (36 taps Fsample = 125 Hz, Fpass=5Hz Fstop= 10Hz, 40dB )
FIR_LP = [ -0,0009 -0,0109 -0,0102 -0,0140 -0,0167 -0,0180 -0,0171...
           -0,0133 -0,0064 0,0038 0,0172 0,0329 0,0502 0,0676...
           0,0839 0,0975 0,1074 0,1126 0,1126 0,1074 0,0975...
           0,0839 0,0676 0,0502 0,0329 0,0172 0,0038 -0,0064...
           -0,0133 -0,0171 -0,0180 -0,0167 -0,0140 -0,0102 -0,0109...
           -0,0009];

%figure(2),spectrum(FM,512,0,Hanning( 512),Fsample);
```

```
%---- Detecção por discriminador de FM -----
FMc = FM ./ abs(FM); % limitador
FMdis = (FMc(2:end)- FMc(1:end-1)) * Fsample; % diferenciador
FMenv = abs(FMdis); % detector de envelope
FMIp = FMenv - mean(FMenv); % remove o offset CC
BBdee = decimate(FMIp,2); % reduz a taxa de amostragem
BBdec = decimate(BBdee,2);
Message = filter(FIR_LP, 1, BBdec); % filtro passa-baixa

%plot(Message);

function Noise = LowPassNoise (SNRdB, Nsmpls)
%
% Gera ruído passa-baixa pela interpolação de ruído branco com três estágios de
% interpolação por 2
%
sigma2 = 10^(-SNRdB/10);
sigma = sqrt(sigma2/2);

%---- gera o ruído banda base limitado em faixa ----
Noise = sigma * (randn(1,Nsmpls/8)+ j*randn(1,Nsmpls/8)); % ruído branco
Noise = interp(Noise,2); % interpola por dois
Noise = interp(Noise,2);
Noise = interp(Noise,2);

Function [T,EyeSig] = ploteye(s,ups,offset);
%
% Função para traçar o diagrama de olho de s
%
% Entradas
%   s – sinal real
%   ups – taxa de sobreamostragem
%
if (nargin [less] 3), offset = 0; end
s = interp(s,4);
ups = ups^4;
f = mod(length(s),ups);
s = real(s(f+1 +offset:end-(ups-offset)));

%
EyeSigRef = reshape(real(s),ups,length(s)/ups);
EyeSigm1 = EyeSigRef(ups/2+2:ups,1:end-2);
EyeSig0 = EyeSigRef(:,2:end-1);
EyeSigp1 = EyeSigRef(1:ups/2,3:end);
EyeSig = [EyeSigm1; EyeSig0; EyeSigp1];

T = [-1 + 1/(ups) : 1/(ups) : + 1-1I(ups)];
pIot(T,EyeSig)
xLabel('Períodos de símbolo')
yLabel('Amplitude')
```

Bibliografia

- Abbate, J., *Inventing the Internet*, Cambridge, MA: The MIT Press, 1999.
- Anderson, J. B., *Digital Transmission Engineering*, IEEE Press, 1999.
- Abramowitz, M., L A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, New York: Dover, 1965.
- Arazi, B., *A Commonsense Approach to the Theory of Error Correcting Codes*, Cambridge, MA; The MIT Press, 1988.
- Bahai, A. R. S., B. R. Saltzberg, *Multi-carrier Digital Communications: Theory and Applications of OFDM*, Kluwer, 1999.
- Baran, P., "Packet switching," in J. C. McDonald, editor, *Fundamentals of Digital Switching*, 2a. ed., pp. 193-235, New York: Plenum, 1990.
- Bell, D. A., *Noise and Solid State*, Plymouth, Eng.: Pentech Press, 1985.
- Bennett, W. R., *Introduction to Signal Transmission*, New York: McGraw-Hill, 1970.
- Bertsekas, D. M., J. N. Tsitsiklis, *Introduction to Probability*, Athena Scientific, 2002.
- Blackman, N. M., *Noise and Its Effect on Communication*, New York: McGraw-Hill, 1966.
- Black, H. S., *Modulation Theory*, Princeton, NJ: Van Nostrand, 1953.
- Blahut, R. E., *Theory and Practice of Error Control Codes*, Reading, MA: Addison-Wesley, 1983.
- Bracewell, R. N., *The Fourier Transform and Its Applications*, 2a. ed., New York: McGraw-Hill, 1986.
- Carlson, A. B., *Communications Systems*, 2a. ed., New York: McGraw-Hill, 1975.
- Cassoli, D., M. Z. Win, A. F. Molisch, "The ultra-wide bandwidth indoor channel: from statistical model to simulations," *IEEE Journal Selected Areas in Communications*, vol. 20, pp. 1247-1257, 2002.
- Chung, K. L., *Elementary Probability Theory with Stochastic Processes*, New York: Springer-Verlag, 1979.
- Clark, G. C. Jr., J. B. Cain, *Error Correction Coding for Digital Communications*, New York: Plenum, 1981.
- Cooley, J. W., "How the FFT gained acceptance," *IEEE Signal Processing Magazine*, vol. 9, no. 1, pp. 10-13, January 1992.
- Cattornole, K. W., *Principles of Pulse-code Modulation* (New York: American Elsevier, 1969).
- Cover, T., J. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.
- Downing, J. J., *Modulation Systems and Noise*, Englewood Cliffs, NJ: Prentice Hall, 1964.
- Erickson, G., "A fundamental introduction to the compact disc player," <http://www.te.umn.edu/~erick20S/Papers/paper.html#strengths>.
- Gitlin, R. A., J. F. Hayes, S. B. Weinstein, *Data Communication Principles*, New York: Plenum, 1992.
- Gold, B., N. Morgan, *Speech and Audio Signal Processing*, New York: Wiley, 1999.
- Guttman, I., S. S. Wilks, J. S. Hunter, *Introductory Engineering Statistics*, 2a. ed., New York: Wiley, 1971.
- Haykin, S., *Adaptive Filter Theory*, 4a. ed., Upper Saddle River, NJ: Prentice-Hall, 2002.
- Haykin, S., *Communications Systems*, 4a. ed., New York: Wiley, 2001.
- Haykin, S., *Communication Systems*, 3a. ed., New York: Wiley, 1994.
- Haykin, S., and M. Moher, *Modern Wireless Communications*, New Jersey: Prentice Hall, 2005.
- James, J. H., B. Chen, L. Garrison, "Implementing VoIP: a voice transmission performance progress report," *IEEE Communications Magazine*, vol. 42, no. 7, pp. 36-41, July 2004.
- Jayant, N. S., P. Noll, *Digital Coding of Waveforms: Principles and Applications in Speech and Video*, Englewood Cliffs, NJ: Prentice-Hall, 1984.
- Jeruchim, M. C., P. Balaban, K. S. Shanmugan, *Simulation of Communication Systems*, 2a. ed., New York: Plenum, 2000.

- Kammler, D. W., *A First Course in Fourier Analysis*, Upper Saddle River, NJ: Prentice-Hall, 2000.
- Kaneko, H. "A unified formulation of segment companding laws and synthesis of codes and digital companders", *Bell System Tech. J.*, vol. 49, pp. 1555-1588, 1970.
- Karlin, S. , H. M. Taylor, *A First Course in Stochastic Processes*, 2a. ed., New York: Academic Press, 1975.
- Lathi, B. P., *Modern Digital and Analog Communication Systems*, 3a. ed., Oxford: Oxford University Press, 1998.
- Leon-Garcia, A., *Probability and Random Processes for Electrical Engineering*, 2a. ed., Reading, MA: Addison-Wesley, 1994.
- Lin, S., D. J. Costello, *Error Control Coding*, 2a. ed., Upper Saddle River, NJ: PrenticeHall, 2004.
- Lucky, R. W., J. Salz, E. J. Weldon, *Principles of Data Communication*, New York: McGraw-Hill, 1968.
- Madisetti, V. K., D. B. Williams, editors, *The Digital Signal Processing Handbook*, CRC Press, 1998.
- McDonald, J. C. ed. *Fundamentals of Digital Switching*, 2a. ed New York: Plenum, 1990.
- McEliece, R. J., *The Theory of Information and Coding*, Reading, MA: Addison-Wesley, 1977.
- Middleton, D., *An Introduction to Statistical Communication Theory*, New York: McGrawHill, 1960.
- Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, 2a. ed., New York: Mc-
Graw-Hill, 1984.
- Pratt, T., C. W. Bostian, *Satellite Communications*, New York: Wiley, 1986.
- Proakis, J. G., *Digital Communications*, 3a. ed., New York: McGraw-Hill, 1999. Robinson, E
N., *Noise and Fluctuations in Electronic Devices and Circuits*, Oxford: Clarendon Press, 1974.
- Shannon, C. E., "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379-423, 623-256, July, October, 1948.
- Smith, B. "Instantaneous companding of quantized signals", *Bell System Tech. J.*, vol. 36, pp.
653-709, 1957.
- Special Issue, "The Internet: past, present, and future," *IEEE Communications Magazine*, vol.
40, no. 7, pp. 42-76, July 2002.
- Steele, R. E., L. Hanzo, *Mobile Radio Communications*, 2a. ed., New York: Wiley, 1999.
- Sveshnikov, A. A. (ed.), *Problems in Probability Theory, Mathematical Statistics and Theory of
Random Functions*, New York: Dover Publications, 1968.
- Takahashi, A., H. Yoshino, "Perceptual QoS assessment of technologies for VoIP," *IEEE Com-
munications Magazine*, vol. 42, no. 7, pp. 28-34, July 2004.
- Tannenbaum, A. S., *Computer Networks*, 3a. ed., Englewood Cliffs, NJ: Prentice-Hall,1996.
- Thomas, J. B., *An Introduction to Applied Probability and Random Processes*, New York: Wi-
ley, 1971.
- Van Der Ziel, A., *Noise: Sources, Characterization, Measurement*, Englewood Cliffs, NJ: Pren-
tice Hall, 1970.
- Willner, A. E., "Mining the optical bandwidth for a terabit per second," *IEEE Spectrum*, vol.
34, pp. 32-41, April 1997.
- Win, M. Z., R. A. Scholtz, "Impulse radio: how it works," *IEEE Comm. Letters*, vol. 2,
pp.36-38,1998.
- Wozencraft, J. M., L M. Jacobs, *Principles of Communication Engineering*, New York: Wiley,
1965. Reprinted by Waveland Press, Prospect Heights: Illinois, 1990.
- Wylie, C. R., L. C. Barrett, Advanced Engineering Mathematics, 5a. ed (New York: McGraw-
Hill, 1982).
- Yuen, J. H. (ed.), *Deep Space Telecommunications Systems Engineering*, New York: Plenum,
1983.
- Zeimer, R. E., and W. H. Tranter, *Principles of Communications: Systems, Modulation and Noi-
se*, 5a. ed., New York: Wiley, 2002.

Glossário

FUNÇÕES

1. Função retangular	$\text{ret}(t) = \begin{cases} 1, & -\frac{1}{2} < t < \frac{1}{2} \\ 0, & t > \frac{1}{2} \end{cases}$
2. Função degrau unitário	$u(t) = \begin{cases} 1, & t > 0 \\ 0, & t < 0 \end{cases}$
3. Função sinal	$\text{sgn}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases}$
4. Função delta (de Dirac)	$\delta(t) = 0, \quad t \neq 0$ $\int_{-\infty}^{\infty} \delta(t) dt = 1$ ou, de forma equivalente, $\int_{-\infty}^{\infty} g(t)\delta(t - t_0) dt = g(t_0)$
5. Função sinc	$\text{sinc}(x) = \frac{\text{sen}(\pi x)}{\pi x}$
6. Integral seno	$\text{Si}(x) = \int_0^x \frac{\text{sen } u}{u} du$
7. Função Q	$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{u^2}{2}\right) du$
8. Coeficiente binomial	$\binom{n}{k} = \frac{n!}{(n-k)!k!}$
9. Função de Bessel de ordem n de primeiro tipo	$J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(jx \text{sen } \theta - jn\theta) d\theta$

Abreviações

A	ampere
ADC	conversor analógico/digital
AM	modulação em amplitude
ASK	chaveamento de amplitude
b/s	bits/segundo
BFSK	chaveamento binário em freqüência
BPSK	chaveamento binário de fase
CA	corrente alternada
CC	corrente contínua
codec	codificador/decodificador
CSB	canal simétrico binário
DAC	conversor digital/analógico
dB	decibel
dbmW	decibel referenciado a 1 miliwatt
dbW	decibel referenciado a 1 watt
DM	modulação delta
DPCM	modulação por codificação de pulso diferencial

DPSK	chaveamento de fase diferencial
DSB-SC	faixa lateral dupla – portadora suprimida
exp	exponencial
FDM	multiplexação por divisão na freqüência
FFT	transformada rápida de Fourier
FI	freqüência intermediária
FM	modulação em freqüência
FPB	filtro passa-baixa
FPF	filtro passa-faixa
FSK	chaveamento de freqüência
GMSK	mínimo chaveamento Gaussiano
GSM	sistema global para comunicação móvel
HDTV	televisão de alta definição
Hz	hertz
IIS	interferência intersimbólica
IP	Internet protocol
ISO	International Organization for Standardization
LDM	modulação delta linear
log	logaritmo
modem	modulador/demodulador
ms	milissegundo
MSK	mínimo chaveamento
NRZ	sem retorno para zero
NTSC	National Television Systems Committee
OC	onda contínua
OFDM	modulação por divisão em freqüência ortogonal
OOK	chaveamento liga-desliga
OQPSK	chaveamento em quadratura de fase deslocada
OSI	open systems interconnection
PAM	modulação por amplitude de pulso
PCM	modulação por codificação de pulso
PDM	modulação por duração de pulso
PLL	phase-locked loop
PM	modulação por fase
PPM	modulação por posição de pulso
PSK	chaveamento de fase
PWM	modulação por largura de pulso
QAM	modulação em quadratura de amplitude
QPSK	chaveamento em quadratura de fase
RF	rádio freqüência
RSR	razão sinal/ruído
RZ	retorno a zero
s	segundo
TDF	transformada discreta de Fourier
TDIF	transformada discreta inversa de Fourier
TDM	multiplexação por divisão no tempo
TEB	taxa de erro de bit
TV	televisão
UWB	faixa ultra larga
V	volt
VLSI	very large scale integration
W	watt
μs	microsssegundo

A

Adição em módulo-dois, 271
 Advanced Research Projects Agency Network (ARPA-NET), 21-22
 Algoritmo de decimação no tempo, 104
 Algoritmos de decimação em freqüência, 104
 Aliasing, 208-211
 AM, *veja* Modulação em amplitude
 Amostragem instantânea, 203-206
 Amostragem natural, 240-241
 Amplificador com realimentação negativa, 192-193n.3
 Analisador, 154-155
 Analisador de espectro em tempo real, 200
 Análise de Fourier, 32-33
 Área:
 sob $G(f)$, 49
 sob $g(t)$, 49
 Área efetiva de abertura, 454-455
 Armazenamento de dados, 29-30
 Armstrong, Edwin H., 19
 ARPANET, 21-22
 ARQ (requisição automática de repetição), 21-22
 Arquitetura em camadas, 26
 Atraso de conexão (Internet), 28-29
 ATV (Advanced television), padrão, 315-316

B

Banda base, 151-152
 Banda passante, 76-77
 Baran, P., 24-25n.2
 Bardeen, John, 20
 BASK, *veja* Chaveamento binário em amplitude
 Batimento (sinal), 163-164
 BBC, *veja* British Broadcasting Corporation
 Bel, 65-66
 Bell, Alexander Graham, 20, 467n.1
 Berners-Lee, Tim, 21-22
 BFSK, *veja* Chaveamento binário em freqüência
 BFSK de Sunde, 289, 292-293
 BIBO, *veja* Entrada limitada-saída limitada (BIBO), critérios de estabilidade
 Bits de canal, 430-431, 436-437
 Bits de informação, 430-431
 Boresight, 454-455
 BPSK, *veja* Chaveamento binário de fase
 Brattain, Walter H., 20
 British Broadcasting Corporation (BBC), 20
 Broadcasting, 22
 BRZ (retorno para zero bipolar), sinalização, 231-232

C

Cálculos de propagação em espaço livre, 453-459
 potência do sinal recebido, 454-458
 razão portadora/ruído, 457-459

Camada física, 27
 Caminho em fase (moduladores SSB), 142-143
 Caminho em quadratura (moduladores SSB), 142-143
 Canais:
 definidos, 69-70
 limitados em faixa, 30-31
 limitados em potência, 30-31
 Canais de satélite, 30-31
 Canais de telefone, 30-31
 Canais de televisão, 21-22
 Canais discretos sem memória, 334-335
 Canais sem fio, 30-31
 Canal de Nyquist, 246-249
 Canal de subida (comunicações), 23-24
 Canal-I, 134-137
 Capacidade de correção de erro, 432-433
 Características de quantizadores, 217-218
 Chaveamento de amplitude M-ário (ASK M-ário), 306-308
 CDs (compact discs), 29-30
 Chave de Strowger, 20
 Chaveamento, 325-326
 círculo, 24-25
 pacote, 24-26
 Chaveamento binário de fase (BPSK), 273-274, 280-282, 285
 análise espectral de, 281-282, 285
 detecção ótima de, 412-416
 geração/detecção coerente de sinais, 280-282
 Chaveamento binário em amplitude (BASK), 273-279
 análise espectral de, 276-279, 406-412
 geração/detecção de sinais, 275-277
 Chaveamento binário em freqüência (BFSK), 274-275, 289, 292-293
 detecção de, 421-423
 Chaveamento de fase, 280-289, 292
 binário, 280-282, 285
 em quadratura, *ver* Chaveamento em quadratura de fase (QPSK)
 Chaveamento de freqüência, 289, 292-300
 binário, 289, 292-293
 com fase contínua, 292-294
 detecção de, 421-424
 mínimo chaveamento, 293-300
 Chaveamento digital, 24-25
 Chaveamento em quadratura de fase (QPSK), 282, 285-289, 292
 deslocada, 287-289, 292
 detecção de, com ruído, 415-419
 espectro de potência de sinais, 288-289, 292
 geração/detecção de sinais, 285-289
 performance de, 426-428
 Chaveamento mínimo (MSK), 293-300
 cálculo do espectro de potência, 297-300
 formulação do, 296-299

- Chaveamento passo-a-passo, 20
Chaveamento por circuito, 24-25
Chaveamento por pacote, 21-22, 24-26
Chopper bipolar, 240-241
CI (círculo integrado), 20
Ciclo de trabalho, 471-472
Círcuito chaveado por pacote, 26
Círculo de inclinação, 187-188
Círculo equivalente de Norton, 446-447
Círculo equivalente de Thévenin, 446-447
Círculo integrado (CI), 20
Círculo lógico, 303-304
Circuitos de duas portas, conexão em série de, 452-454
Circuitos integrados em escala muito alta (VLSI), 20
Circuitos virtuais, 27-28
Circuitos VLSI, 20
Clark, Arthur C., 22
Codificação correlacional, 270
Codificação diferencial, 231-232
Codificador, 431-432
Codificador de nível sem retorno para zero, 280-281, 285-287
Código de Gray, 425-426
Código de linha (modulação por pulso), 231-233
Código Manchester, 231-232
Código Morse, 19
Códigos convolucionais, 437-439
Códigos de bloco, 431-434, 437-439
Códigos de bloco Bose-Chaudhuri-Hocquenghem, 437-438
Códigos de bloco Reed-Solomon, 437-438
Códigos de correção de erro a frente, 438-441
Códigos de Hamming, 433-438
Códigos turbo, 438-439
Coeficiente complexo de Fourier, 66-67
Compact discs (CDs), 29-30
Componentes demultiplexados, 286-287
Computação no local, 104
Computador pessoal (PCs), 21-22
Comte, Auguste, 19
Comunicação digital:
 história da, 20-22
 problema de detecção de sinal em, 32-33
 ruído em, 402-442
Comunicação sem fio, 23-24
Comunicação virtual, 27
Comunicações analógicas:
 problema de detecção de sinal em, 32-33
 ruído em, 374-399
Comunicações ópticas, história da, 22
Comunicações ponto-a-ponto, 22-23, 160-162
Comunicações via satélite:
 cálculos em espaço livre para, 453-459
 como sistema ponto-a-ponto, 23-24
 componentes de, 23-24
 história da, 22
 temperatura de ruído do sistema, de receptores, 453-454
Condições de Dirichlet, 36-37, 58-59
Consideração passa-faixa, 274-276
Constantes, 486
- Conteúdo de freqüência (de um sinal), 35
Controle de eco inicial (em VoIP), 29-30
Correlação, 54-55
Covariância, 338-340
CPFSK, *veja* Chaveamento de freqüência com fase contínua
Critério de erro médio quadrático, 266n.7
Critério de Paley-Wiener, 74-75
Cruzamentos com zero, 168-171
- D**
- Datagramas, 27-28
dB, *veja* Decibel
De Forest, Lee, 20
Decibel (dB), 30-31, 467
Decodificador, 431-432
Decodificador de lista, 434-435
Demodulação em freqüência, 186-195
 discriminador de freqüência, 186-191
 e phase-locked loop (PLL), 190-195
Demodulação síncrona, 131-132
Densidade espectral:
 cruzada, 94-95
 de energia, *ver* Densidade espectral de energia
 de potência, 95-98, 354-357
 de ruído, 451-453
Densidade espectral de energia, 87-94
 de sinais de energia, 87-94
 definição, 87-88
 e utilização de filtros, 88-92
 efeito da filtragem em, 88-92
 interpretação de, 91-72
Densidade marginal, 330-331
Desigualdade de Schwarz, 480-481
Deslocador de fase de faixa larga, 142-143
Desvanecimento de Rayleigh, 461-464
Desvio de freqüência, 174-175
Detecção coerente, 131-132, 148-149
 chaveamento binário de fase, 280-282
 com modulação por faixa lateral única, 142-144
 com modulação por faixa lateral vestigial, 148-150
 da onda modulada DSB-SC, 196-197
 modulação por faixa lateral dupla-portadora suprimida, 131-134
 ruído em receptores lineares utilizando, 379-383
Detecção de sinal:
 chaveamento binário de amplitude, 275-277
 chaveamento em quadratura de fase, 285-289
 coerente:
 chaveamento binário de fase, 280-282
 modulação por faixa lateral dupla-portadora suprimida, 131-134, 379-383
 receptores lineares, 379-383
 com modulação em freqüência, 388-396
 de BPSK, 412-416
 de FSK, 421-424
 de PAM binário com ruído, 406-413
 de QAM, 418-422
 de QPSK, 415-419
 diferencial, 423-426

- não coerente, 300-306
 chaveamento de fase diferencial, 303-305
 sinais BASK, 300-303
 sinais BFSK, 302-304
 problema de detecção de sinal em, 32-33
 teoria da detecção, 32-33
- Detecção e correção de erro (ruído em sistemas digitais), 429-441**
- códigos convolucionais, 437-439
 - códigos de bloco, 431-434, 437-439
 - códigos de bloco de Bose-Chaudhuri-horcquenghem, 437-438
 - códigos de bloco de Reed-Solomon, 437-438
 - códigos de Hamming, 433-438
 - códigos turbo, 438-439
 - interpretação de códigos por espaço de sinal, 438-441
- Detecção não coerente, 300-306**
- chaveamento de fase diferencial, 303-305, 423-426
 - sinais BASK, 300-303
 - sinais BFSK, 302-304
- Detector coerente de fase em quadratura (canal Q), 135-137**
- Detector coerente em fase (canal-I), 134-137**
- Detector integra-e-descarta, 407-408**
- Detectores/detecção de envelope, 117-120**
- de VSB mais portadora, 150-151
 - modulação em amplitude, 126-129
 - para AM senoidal, 128-129
 - para demodulação de onda AM, 126-128
 - ruído em receptores AM utilizando, 382-386
- Diagrama de dispersão, 257-258n.4**
- Diagrama de espaço de sinal, 306-307**
- Dibit, 282, 285**
- Diferenciação no domínio do tempo (transformada de Fourier), 49-51**
- Difração, em rádio móvel terrestre, 458-460**
- Digital subscriber line (DSL), 28-29, 312-313**
- Digital versatile discs (DVDs), 29-30**
- Diodo por tubo de vácuo, 20**
- Dirac Comb, 67-69**
- Discriminador de fase (canais I e Q), 135-137**
- Discriminador de frequência, 187-188**
- Discriminador de frequência balanceado, 190-191**
- Dispositivos de duas portas:**
- figura de ruído médio de, 450-451
 - temperatura equivalente de ruído de, 450-453
- Dispositivos hand-held (para acesso à Internet), 28-29**
- Distância de Hamming, 431-433**
- Distorção da forma de onda, 158-160**
- Distorção de envelope, 117-120**
- Distorção de pico nula, 259-260**
- Distorção de pico unitária, 259-260**
- Distorção em amplitude, 175-176**
- Distorção espectral, evitando, 124-126**
- Distorção harmônica, 175-176**
- Distribuição binomial, 331-332**
- DM, veja Modulação delta**
- DPCM, veja Modulação por codificação diferencial de pulso**
- DSL, veja Digital subscriber line**
- Duração de bit, 273-274**
- Duração de raiz média quadrática (rms), 114**
- DVDs (digital versatile discs), 29-30**
- E**
- Eckert, J. Presper, Jr., 21-22**
- Efeito de abertura, 213-214**
- Efeito limiar, 394-396**
- Efeito nulo de quadratura, 132-133**
- Eficiência de abertura, 454-455**
- Elemento de atraso de um bit, 303-304**
- ENIAC, 21-22**
- Entradas limitada-saída limitada (BIBO), critérios de estabilidade, 71-73**
- Envelope complexo:**
- da onda modulada, 152-154
 - de ondas FM, 178
- Equação de análise, 35-36**
- Equação de síntese, 35-36**
- Equação do diodo, 448-449**
- Equalização, 262-266**
- adaptativa, 21-22, 266n.7
 - de zero forçado, 263-266
 - na modulação por amplitude de pulso, 212-215
- Equalizador transversal ajustável, 263-264**
- Erro de fase, 191-192**
- Erro tipo 1, 408-409**
- Espectro contínuo, 67-68**
- Espectro contínuo de amplitude, 37-38**
- Espectro contínuo de fase, 37-38**
- Espectro da mensagem, 117-120**
- Espectro de pulso de cosseno levantado, 249-257**
- espectro de pulso de raiz de cosseno levantado, 255-257, 411-413
 - necessidade de largura de faixa de transmissão, 251-253
 - porção amortecida do espectro, 252-255
 - propriedades de, 252-256
- Espectro discreto, 67-68, 58-59, 472-473**
- Espectro senoidal DSB-SC, 131-132**
- Esperança, média estatística de, 336-340**
- Esquemas de modulação digital M-ário:**
- chaveamento de fase, 305-308
 - chaveamento de frequência, 308-310
 - modulação em quadratura de amplitude, 306-309, 418-422
- Estação base (comunicação via satélite), 23-24**
- Estacionário em sentido amplo, 348-349**
- Evento certo, 325-326**
- Evento elementar, 325-326**
- Evento impossível, 325-326**
- Evento nulo, 325-326**
- Excesso de largura de faixa, 252-253**
- Excitação, 68-69**
- Expansão em séries, 484**
- Experimento aleatório, 324-325**
- F**
- Faixa estreita, 126-128**
- Faixa lateral inferior, para freqüências positivas, 117-120**

- Faixa lateral superior, 117-120
 Família de modulação em amplitude, 116-162
 estratégias de modulação linear em, 116
 faixa lateral dupla – portadora suprimida, 129-137
 modulação em amplitude, 116-130
 modulação por faixa lateral única, 137-146
 modulação por faixa lateral vestigial, 145-152
 multiplexação de portadora em quadratura, 135-139
 Fantasmas, 459-460
 Farnsworth, Philo T., 20
 Fase dividida (código Manchester), 231-232
 Fase travada, 191-192
 Fator de modulação, 119-120
 Fator de qualidade (fator Q), 190-191n.2
 Fator de roll-off, 249-250, 267
 Fator de rotação, 101, 103
 Fator de sensibilidade de fase, 166-167
 Fator de sensibilidade de freqüência, 167-168
 Fator Q , 190-191n.2
 FDM, *veja* Multiplexação por divisão em freqüência
 Fenômeno de propagação de erro, 271
 Fenômeno dependente do sinal, 267
 Fenômeno multi-caminho (comunicação sem fio), 23-24, 458-464
 Ferramentas de busca, 28-29
 Fessenden, Reginald, 19
 FFT, *veja* Transformada rápida de Fourier
 Figura de mérito, 382-383, 385-388, 393-394
 Figura de ruído, 450-451
 Figura de ruído médio, 450-451
 Filtro anti-aliasing, 209-210
 Filtro casado, 21-22, 403-405
 Filtro de atraso de linha, 70-72, 229-230
 Filtro de de-ênfase, 396-397
 Filtro de malha, 190-191
 Filtro de síntese, 207-210
 Filtro em tempo discreto, 229-230
 Filtro passa-baixa complexo, 155-156
 Filtro passa-baixa ideal, 76-87
 aproximação de, 85-87
 resposta ao pulso do, 76-86
 Filtro transversal, 263-264
 Filtro VSB, 158-159
 Filtros, 76-77
 analógicos, 86-87
 Butterworth, 86-87, 260-261
 Chebyshev, 86-87
 definição, 69-70, 76-77
 digitais, 86-87
 e densidade espectral de energia, 88-92
 ideais passa-baixa, 76-87
 largura de faixa equivalente de ruído, 363-367
 para receptores super heterodinos, 156-157
 Filtros pára-faixa, 76-77
 Filtros passa-alta, 76-77
 Filtros passa-baixa, 76-77
 Filtros passa-faixa, 76-77, 154-156
 Fleming, John Ambrose, 20
 Fluxo de calor, 35-36n.1
 FM, *veja* Modulação em freqüência
 Forma de pulso de raiz de cosseno levantado normalizado, 411-412
 Formato ótimo de pulso, 246-247
 Fórmula de Euler, 48, 62-63
 Fórmula de Friis para transmissão, 455-456, 461-462
 Fórmula de interpolação, 207-208, 246-247
 Fórmula de Schottky, 448-449
 Fórmula de soma de Poisson, 67-68
 Fourier, Joseph, 35-36n.1
 Freqüência:
 fundamental, 66-67
 relação inversa entre tempo e, 55-58
 Freqüências laterais, 137-139
 Função de amostragem ideal (Dirac comb), 67-69
 Função de autocorrelação, 86-88
 Função de Bessel de ordem n , 178-180
 Função de correlação cruzada, 93-95
 Função de distribuição de Bernoulli, 329-330
 Função de distribuição de Rayleigh, 462-463
 Função de transferência passa-baixa, 146-148
 Função degrau unitário, 64-66
 Função delta, *veja* Função delta de Dirac
 Função delta de Dirac, 58-67
 aplicações de, 62-67
 como limitante do pulso Gaussiano, 59-63
 e extensão da transformada de Fourier, 58-59
 no domínio da freqüência, 117-120
 propriedade da replicação da, 59-60
 propriedade de peneiramento, 58-59
 Função densidade de probabilidade comum, 330-331
 Função erro complementar, função Q e, 478-479
 Função exponencial complexa, aplicação da função delta a, 62-63
 Função no tempo:
 com simetria ímpar, 44
 diferenciação de uma, 49-50
 integração de uma, 51-53
 partes real e imaginária de um, 52-53
 Função Q , 341-342, 409-410, 478-479
 Função sinal, 44, 63-65
 Função sinc, 46-47, 54-56
 Funções de Bessel, 178, 475-477
 Funções de distribuição, 328-331
 Funções de geração, 66-67
 Funções generalizadas (distribuições generalizadas), 59-60
 Funções senoidais, aplicação da função delta em, 62-64

G

- Ganho de malha aberta, 192-193n.3
 Ganho de malha fechada, 192-193n.3
 Gap de energia (em modulação SSB), 129-130
 Gauss, C. G., 340-341n.4
 Gráfico de fluxo de sinal, 101, 103-103
 Grupo, sinal, 33-34
 Grupo básico (multiplexação), 159-162
 Grupos master (multiplexação), 160-162
 Grupos muito largos (multiplexação), 160-162
 GSM (sistema global para comunicações móveis), 314-315

H

- Hamming, R., 429-430*n.3*
 HDTV, 23-24
 Hertz, Herinrich, 19
 História de dispositivos eletrônicos, 20
 Hockham G. A., 22
 Hosts, 24-25, 27-28
 HTML, 21-22
 http, 21-22

I

- Identidades trigonométricas, 484
 Imagem de sinais:
 como função bidimensional do tempo, 23
 supressão, 156-158
 Impulso unitário, 58-59
 Índice de modulação, 174-175
 Integração no domínio do tempo (transformada de Fourier), 51-53, 65-67
 Integrais, 485
 Integral de convolução, 53-54, 70-71
 Interfaces (modelo OSI), 26
 Interferência de imagem (em receptores super heterodí-nos), 156-158
 Internet, 27-30
 Internet protocol (IP), 27-28
 Interpretação de espaço de sinal de códigos, 438-441

K

- Kao, K. C., 22

L

- Laboratórios Bell, 20, 22
 Largura de faixa, 55-57
 de canal, 30-32
 definição de, 56-57
 e produto tempo-largura de faixa, 57-58
 e relação inversa entre tempo e freqüência, 56-58
 e resposta em amplitude, 74-75
 Largura de faixa da mensagem, 117-120, 145-146
 Largura de faixa de 3 dB, 56-57
 Largura de faixa de raiz média quadrática, 56-58, 114
 Largura de faixa de transmissão (ondas AM), 119-120
 Largura de faixa equivalente do ruído, 449-451
 Largura de faixa nulo-a-nulo, 56-57
 Largura de faixa rms, *veja* Largura de faixa de raiz média quadrática
 Largura de faixa vestigial, 145-146
 Laser, 22
 Lei quadrática, 124-126
 Limiar/nível de decisão, 217-218
 Limitador passa-faixa, 200-201
 Limite de faixa de freqüências, 194-195
 Linearidade da transformada de Fourier, 42-45
 Linguagem de formatação de hipertexto (HTML), 21-22

- Linhas de transmissão, 105-106
 Link de descida (comunicações), 23-24
 Links de espaço profundo, 30-31
 Lóbulo principal, 56-57
 Lodge, Oliver, 19
 Lucky, Robert, 21-22

M

- Marconi, Guglielmo, 19
 Margem do sistema, 456-457
 Matriz geradora *k-por-n*, 434-435
 Mauchly, John W., 21-22
 Maxwell, James Clerk, 19
 Mensagens de e-mail, 28-29
 Método de discriminação de fase (geração de SSB), 142-143
 Método de discriminação em freqüência, 140-143
 Método direto (geração de onda FM), 184-187
 Método indireto (geração de ondas FM), 185-187
 Microprocessadores, 20
 Middleton, D., 21-22
 Modelo de referência OSI (Open systems interconnection), 26, 27
 Modelo de transmissão de referência, 377-378
 Modelo Okumira-Hata, 460-461*n.8*
 Modelo OSI, *veja* Modelo de referência OSI (Open systems interconnection)
 Modelos de espaço de sinal, ruído em, 428-430
 Modems de banda larga, 28-29
 Modulação. *Veja também o tipo específico*, por exemplo:
 modulação de pulso
 definição, 31-32, 116
 motivação primária para, 116
 Modulação de amplitude em quadratura (QAM), 135-139, 421-422
 detecção de, com ruído, 418-422
 performance de, 426-428
 Modulação de onda contínua (OC), 116
 famílias de, 116
 modulação de pulso *versus*, 203
 modulação em amplitude, 116-162
 modulação em ângulo, 166-198
 Modulação delta (DM), 223-229
 erros de quantização, 226-228
 modulação delta-sigma, 227-229
 Modulação de pulso, 203-238
 análogica, 203
 digital, 203
 famílias de, 203
 linhas de código em, 231-233
 modulação delta, 223-229
 modulação por amplitude de pulso, *veja* Modulação por amplitude de pulso
 modulação por codificação de pulso, *veja* Modulação por codificação de pulso
 modulação por codificação diferencial de pulso, 228-232
 modulação por duração de pulso, 214-216
 modulação por onda contínua *versus*, *veja* Modulação de onda contínua

- modulação por posição de pulso, 215-216
 processo de quantização em, 217-220
 processos de amostragem para:
 amostragem instantânea, 203-206
 e fenômeno de aliasing, 208-211
 teorema da amostragem, 205-209
- Modulação digital de pulso, 203, 216-217
- Modulação DSB-SC, *veja* Modulação por faixa lateral dupla-portadora suprimida
- Modulação em amplitude (AM), 116-130
 base teórica, 116-122
 detecção de envelope, 126-129
 e ondas portadoras, 160-162
 experimentos de computador, 121-126, 128-129
 superação de limitações de, 128-130
- Modulação em ângulo, 166-198
 compromisso com, 166
 demodulação em freqüência, 186-195
 discriminador de freqüência, 186-191
 e PLL (*phase-locked loop*), 190-195
 e relação de ondas PM-FM, 172-174
 modulação em freqüência:
 faixa estreita, 173-177
 faixa larga, 177-182
 ondas moduladas em freqüência:
 geração de, 184-187
 largura de faixa de transmissão de, 183-185
 Relação de PM com, 172-174
 propriedades de ondas moduladas em ângulo, 167-173
- Modulação em fase (PM), 166-167
- Modulação em freqüência (FM), 167-168
 detecção de, 388-396
 faixa estreita, 173-177
 faixa larga, 177-182
 origem do conceito, 19
- Modulação faixa lateral vestigial (VSB), 129-130, 145-152
 de sinais de televisão, 156-160
 detecção coerente de, 148-150, 316-318
 filtro de formatação de faixa lateral para, 146-149
 motivação para utilização de, 145-146
 SSB *versus*, 145-146
- Modulação OC, *veja* Modulação de onda contínua
- Modulação por amplitude de pulso (PAM), 210-215
 binária, detecção ótima da, 406-413
 efeito da abertura e equalização, 212-215
 filtros sample-and-hold, 211-213
 performance da, 426-428
- Modulação por codificação de pulso (PCM), 218-224
 invenção da, 20
 operações do receptor, 223-224
 operações do transmissor, 218-223
 regeneração ao longo do caminho de transmissão, 221-223
- Modulação por faixa lateral dupla – portadora suprimida (DSM-SC), 128-137
 base teórica, 129-132
 detecção coerente, 131-134, 379-383
 experimento de computador, 133-136
 receptor Costas para detecção de, 134-137
- Modulação por faixa lateral única (SSB), 128-130, 137-146
 base teórica, 137-142
 detecção coerente de, 142-144, 346-388
 e modulação em amplitude, 160-162
 moduladores para, 140-143
 translação em freqüência, 143-146
- Modulação por faixa lateral vestigial, 163-164
- Modulação SSB, *veja* Modulação por faixa lateral única
- Modulação VSB por onda senoidal, 148
- Modulação VSB. *Veja* Modulação por faixa lateral vestigial
- Modulador de lei quadrática, 124-128
- Modulador de produto, 129-130
- Modulador delta linear, 226-227
- Modulador em freqüência de faixa larga de Armstrong, 185-187
- Monociclo Gaussiano, 235-236
- Morse, Samuel, 19
- MSK, *veja* Chaveamento mínimo
- Multiplexação, 158-160
- Multiplexação por divisão em freqüência (FDM), 158-160
 multiplexação estéreo, 194-197
 ortogonal, 312-315
 passos de modulação para, 159-162
- Multiplexação por divisão no tempo (TDM), 158-160, 232-236
- Multiplexação por portadora em quadratura, 135-139
- Multiplicador (PLL), 190-191
- Multiplicador de freqüência, 185-187
- N**
- Níveis de reconstrução, 217-218
- Níveis de representação, 217-218
- Nível de potência de referência, 467
- North D. O., 21-22
- Nós (gráfico de fluxo de sinal), 101, 103-103
- Noyce, Robert, 20
- Nulos, 56-57
- Nyquist, Harry, 20, 254-255*n.3*
- O**
- Onda modulada híbrida, 189-191
- Onda modulante senoidal, modulação SSB e, 137-139
- Onda portadora, 31-32
- Onda portadora complexa, 152-154
- Onda portadora senoidal, 31-32, 116. *Veja também* Modulação de onda contínua
- Onda quadrada, 276-277
- Ondas FM, *veja* Ondas moduladas em freqüência
- Ondas moduladas, representação banda base de, 151-155
- Ondas moduladas em fase (PM), 167-173
- Ondas moduladas em freqüência (FM), 167-173
 geração de, 184-187
 largura de faixa de transmissão de, 183-185
 relação de PM com, 172-174
- Ondas PM, *veja* Ondas moduladas em fase
- Órbita geoestacionária, 23-24, 453-454
- Oscilador controlado por tensão (VCO):
 e o receptor Costas, 135-137
 e PLL, 190-191

P

Pacotes, 26-28
 Padrão de olho, 257-263
 características de tempo para, 258-259
 distorção de pico para interferência intersimbólica, 258-261
 para transmissão M-ário, 260-261
 PAM, *veja* Modulação por amplitude de pulso
 Pára-faixa, 76-77
 Parâmetro de ganho de malha, 192-193
 Pares trancados, 28-29, 105-106
 Pares transformada de Fourier, 35-38, 483
 Pares transformada de Hilbert, 483
 PBXs (private branch exchanges), 29-30
 PCM, *veja* Modulação por codificação de pulso
 PCs (personal computer), 21-22
 PDM, *veja* Modulação por duração de pulso
 Percentagem de modulação, 117-120
 Perda de caminho, 455-456, 460-462
 Perda em espaço livre, 455-456
 Perda média do caminho, 460-462
 Perdas de transmissão, 22
 Performance do sistema, 32-33
 Periodicidade, 67-68, 100-101
 Periodograma, 96-97
 Peso de Hamming, 431-432
 Phase-locked loop, 190-195
 Pierce, John R., 22
 Piloto, para detecção de onda modulada VSB, 150-151n.3
 Pixels, 23
 PLL de primeira ordem, 194-195
 PM (modulação em fase), 166-167
 Ponto de amostra, 325-326
 Potência, de transmissão 456-457
 Potência da portadora (sistemas via satélite), 456-457
 Potência de sinal recebido, 454-458
 Potência disponível (ruído), 447-448
 Potência efetiva isotrópica irradiada (PEII), 455-456
 Potência média, 95-96
 Potência transmitida, 30-31
 PPM (modulação por posição de pulso), 215-216
 Predição, 229-230
 Prefixos de unidades, 486
 Primeiro detector, 156-158
 Private branch exchanges (PBXs), 29-30
 Probabilidade condicional, 333-337
 Problema de detecção de sinal, 32-33
 Problema de interferência intersimbólica, 244-246
 Processamento de sinal em tempo discreto, 216-217
 Processo aleatório, 33-34, 345-354, 374-375
 correlação de, 348-354
 estacionário, 346-349
 Processo de mapeamento, 23-24
 Processo de quantização:
 na modulação delta, 226-228
 na modulação por pulso, 217-220
 Processo Gaussiano, 357-446
 Processos de ponta, 26, 27
 Processos estocásticos, 33-34. *Veja também* Processo aleatório
 Produto duração-largura de faixa, 57-58

Produto tempo-largura de faixa, 57-58, 85-86
 Propagação em espaço livre, 458-459
 Propagação por visada direta, 23-24
 Propriedade da dilatação (transformada de Fourier), 45
 Propriedade de deslocamento na frequência (transformada de Fourier), 47-49
 Propriedade de deslocamento no tempo (transformada de Fourier), 47
 Propriedade de peneiramento (da função delta de Dirac), 58-60
 Propriedade de reflexão (transformada de Fourier), 45
 Propriedade de similaridade (transformada de Fourier), 45
 Protocolo (redes de dados), 26
 Protocolo de transferência de hipertexto (http), 21-22
 Pulso de rádio freqüência, (RF), 48-49
 Pulso dobrado, 51
 Pulso Gaussiano:
 como pulso unitário, 51
 definição, 49-50
 função delta como forma limitante do, 59-61
 Pulso retangular, 38-40
 Pulso RF, *veja* Pulso de rádio freqüência
 Pulso triangular, 51-52
 Pulso unitário, 51
 Pulso(s) exponencial, 40-42
 crescente truncado, 40-41
 doblado, 43-44
 decrescente truncado, 41-42

Q

QAM, *veja* Modulação de amplitude em quadratura
 QoS, *veja* Quality of service
 QPSK, *veja* Chaveamento em quadratura de fase
 Quality of service (QoS – qualidade de serviço), 28-29
 Quantidade de potência do link, 456-458
 Quantizador não uniforme, 219-220
 Quantum, 217-218
 Quase travado em fase, 191-192

R

Radar modulado em freqüência, 197-198
 Rádio, 23-24
 digital, 23
 história do, 19
 impulso, 235-237
 móvel terrestre, 62-63, 458-460
 Rádio AM, 23
 Rádio FM, 19, 23
 Razão de desvio, 183-184
 Razão de perda de pacote, 28-29
 Razão portadora/ruído (RPR), 457-459
 Razão sinal/ruído (RSR), 30-31
 como parâmetro de projeto de sistema, 31-32
 em comunicação analógica, 375-379
 pós-detecção, *veja* Razão sinal/ruído pós-detecção
 Razão sinal/ruído de referência, 377-378
 Razão sinal/ruído pós-detecção, 377-378, 381-390, 393-394
 em receptores AM utilizando detecção de envelope, 383-386

- em receptores lineares, 381-383
 em receptores SSB, 387-388
 na modulação FM, 389-394
Razão sinal/ruído pré-detecção, 377-378
 em receptores AM utilizando detecção de envelope, 383-384
 em receptores lineares, 380-381
 em receptores SSB, 387
 na modulação FM, 388-390
Razões de potências, 467
Realimentação, 185-187
Receptor Costas, 134-137
Receptor de correlação, 406-407
Receptores de faixa lateral única, ruído em, 386-388
Receptores de rádio, super heterodinos, 19
Receptores lineares, ruído em, 379-383
Receptores passa-faixa, ruído com, 378-380
Receptores super heterodinos, 19, 156-158, 378-379
Rede de comunicação, 23-26
Rede de inclinação, 388-389
Redes de computadores, 21-26
Redes de dados, 26-27
Reeves, Alex, 20
Reflexão (rádio móvel terrestre), 458-460
Regra da conjugação (transformada de Fourier), 45-46
Regra de Carson, 183
Relações de Wiener-Khitchine, 87-89, 255-256
Representação banda base:
 de filtros passa-faixa, 154-156
 de ondas moduladas, 151-155
Representação cartesiana, 175-176
Representação polar, 175-176
Requisição automática de repetição (ARQ), 21-22
Resistência de base, ruído térmico de, 449-450
Resolução de freqüência, 97-98
Resposta, definição, 68-69
Resposta ao impulso, 69-70
Resposta ao pulso (filtros passa-baixa ideais), 76-86
Resposta de fase, 73-74, 76-77
Resposta em amplitude, 73-77
Resposta em freqüência (em sistemas lineares), 72-75
Resposta no tempo, transmissão de sinais através de um sistema linear e, 69-72
Retorno a zero bipolar (BRZ), sinalização, 231-232
Reversão de bit, 104
Roteadores, 24-25, 27-28
RPR, *veja Razão portadora/ruído*
RSR, *veja Razão sinal/ruído*
Ruído, 445-465
 branco, 358-362, 445-446
 cálculos em comunicação em espaço livre, 453-459
 conexão em série de circuitos de duas portas, 452-454
 definição, 30-31, 324, 374
 elétrico, *veja Ruído elétrico*
 em comunicação analógica, 374-399
 com estruturas de receptores passa-faixa, 378-380
 detecção de modulação em freqüência, 388-396
 em receptores AM utilizando detecção de envelope, 382-386
 em receptores lineares utilizando detecção coerente, 379-383
 em receptores SSB, 386-388
 pré-ênfase/de-ênfase FM, 395-399
 razão sinal/ruído, 375-379
em comunicação digital, 402-442
 código Gray, 425-426
 comparação de performance das estratégias, 426-428
 detecção de pulso único, 403-407
 detecção de QAM, 418-422
 detecção de QPSK, 415-419
 detecção diferencial, 423-426
 detecção e correção de erro, 429-441
 detecção ótima de BPSK, 412-416
 detecção ótima de FSK, 421-424
 detecção ótima de PAM binário, 406-413
 em modelos de espaço de sinal, 428-430
 taxa de erro de bit, 402-404
em rádio móvel terrestre, 458-464
em sistemas de comunicação, 374-376
externo versus interno, 445
 faixa estreita, 362-367
 figura de ruído, 449-451
 fontes de, 445
 temperatura equivalente de ruído, 450-453
Ruído “um-sobre-”, 445-446
Ruído branco Gaussiano, 446-447
Ruído de canal, 30-31
Ruído de faixa estreita, 362-367
Ruído de fase, 372-373
Ruído de rajada, 445-446
Ruído elétrico, 445-450
 potência disponível de ruído, 447-448
 shot, 445-450
 térmico, 445-448
Ruído externo, 445
Ruído Flicker, 445-446
Ruído interno, 445
- ## S
- Saltos (gráfico de fluxo de sinal)**, 101, 103-103
Satélite orbitando a terra, 22
Scripts MATLAB (experimentos de computador), 487-494
Segundo detector, 156-158
Série de Fourier, 35-36n.1, 468-473
 e espectro discreto, 470-473
 exponencial complexa de, 66-67, 469-471
Serviço de melhor esforço (Internet), 27-29
Shannon, Claude, 21-22, 429-430n. 3
Shockley, William, 20
Sinais, *veja também tipos específicos*, por exemplo: sinais periódicos
 conteúdo de freqüência de, 35
 e largura de faixa, 56-57
 valor real, simetria de conjugado do espectro, 38-39
Sinais aleatórios, 324-359
 definição, 33-34, 324
 e processos aleatórios, 345-354
 correlação de, 348-354
 estacionário, 346-349
 e processos Gaussianos, 357-359
 e teoria da probabilidade, 324-346
 espectro de, 353-357

- Sinais banda larga, transmissão de, 145-146
 Sinais de energia, 86-97
 correlação cruzada de, 93-95
 definição, 36-37
 densidade espectral de energia de, 87-94
 existência da transformada de Fourier de, 36-37
 função de autocorrelação de, 86-88
 relações de Wiener-Khitchine para, 87-89
 Sinais de televisão, modulação VSB para, 156-160
 Sinais de voz, 23
 Sinais de voz, codificação digital de, 20
 Sinais passa-baixa, 56-58, 114
 Sinais passa-faixa, 56-57
 Sinais periódicos, 139-140
 em sistemas lineares, 68-76
 transformada de Fourier de, 66-69
 Sinais/sinalização antipodal, 280-281, 423-424
 Sinal CC, aplicação da função delta a, 62-63
 Sinal contendo a informação (mensagem), 116-117
 Sinal de fase contínua, 289, 292
 Sinal de mensagem, 116-117
 Sinal de vídeo, geração de, 23-24
 Sinalização com retorno para zero (RZ), 231-232
 Sinalização liga-desliga, 231-232
 Sinalização ortogonal, 423-424
 Sinalização por resposta parcial, 270
 Sinalização sem retorno para zero (NRZ), 231-232
 Síndrome de erro, 435-436
 Sintetizador, 154-155
 Sintonização da frequência da portadora, 156-157
 Sistema, definição, 68-69
 Sistema de demultiplexação, 196-197
 Sistema de portadora T-1, 20, 234-236, 252-253
 Sistema de programa armazenado (telefone), 20
 Sistema global para comunicações móveis (GSM), 314-315
 Sistema linear, 68-69
 Sistema linear invariante no tempo (LIT), 35, 70-72
 Sistema multi-portadoras, 312-313
 Sistemas causais, 71-72
 Sistemas de comunicação:
 armazenamento de dados, 29-30
 base teórica, 31-34
 composição do, 23
 condições operacionais para, 30-32
 conhecimento histórico prévio de, 19-22
 integração de telefone e Internet, 28-30
 Internet, 27-29
 melhoramento da performance de, 31-32
 parâmetros de projeto para, 31-32
 rádio, 23-24
 recursos primários em, 30-31
 redes de comunicação, 23-26
 redes de dados, 26-27
 ruído e, 374-376
 tipos básicos de, 22-23
 Sistemas de conservação de largura de faixa, 135-137
 Sistemas de realimentação em malha fechada, 190-191
 Sistemas de telefonia:
 história do, 20
 integração da Internet e, 28-30
 pares trançados, 105-106
 primeiro serviço comercial, 20
 Sistemas estáveis, 71-73
 Sistemas LIT, *veja* Sistema linear invariante no tempo
 Sobreposição espectral, 124-126
 Somatórios, 485
 Strowger, A. B., 20
 Sub-redes, Internet, 27-28
 Supergrupo (multiplexação), 160-162
 Superposição, princípio da, 68-69
 Superposição da transformada de Fourier, 42-45
- T**
- Tabelas matemáticas, 482-486
 Tamanho do passo, 217-218
 Taxa de código, 438-439
 Taxa de erro de bit (TEB), 402-404, 426-428
 Taxa de erro de símbolo, 420-421
 Taxa média de erro de bit, 402-403
 TDF, *veja* Transformada discreta de Fourier
 TDIF, *veja* Transformada discreta inversa de Fourier
 TDM, *veja* Multiplexação por divisão no tempo
 TEB, *veja* Taxa de erro de bit
 Técnicas de modulação passa-faixa digital, 273-320
 chaveamento binário de amplitude, 275-279
 chaveamento de fase, 280-289, 292
 chaveamento de frequência, 289, 292-300
 consideração passa-faixa em, 274-276
 esquemas de detecção não coerentes, 300-306
 chaveamento diferencial de fase, 303-305
 sinais BASK, 300-303
 sinais BFSK, 302-304
 esquemas de modulação digital M-ário, 305-310
 mapeamento de formas de onda em constelações de pontos de sinal, 209-312
 Telégrafo, 19
 Televisão avançada (ATV), padrão, 315-316
 Televisão avançada digital (ATV), padrão, 315-316
 Televisão de alta definição (HDTV), 23-24
 Televisão digital, 23-24, 315-318
 Televisão:
 alta definição, 23-24
 digital, 315-318
 história da, 20
 transmissão para, 23
 Temperatura equivalente de ruído, 450-453
 Tempo, relação inversa entre frequência e, 55-58
 Tentativas independentes de Bernoulli, 331-332
 Teorema da amostragem, 205-209, 246-247
 Teorema da convolução (transformada de Fourier), 53-54
 Teorema da correlação (transformada de Fourier), 53-55
 Teorema da energia de Rayleigh, 54-56
 Teorema da máxima transferência de potência, 447-448
 Teorema da modulação (transformada de Fourier), 47, 53-54
 Teorema do limite central, 343-346
 Teoria da modulação, 31-32
 Teoria da probabilidade, 33-34, 324-346
 abordagem de frequência relativa, 324-326
 axiomas da probabilidade, 325-327
 e médias estatísticas da esperança, 336-340
 e transformação de variáveis aleatórias, 339-341
 funções de distribuição, 328-331

- probabilidade condicional, 333-337
 teorema do limite central, 343-346
 várias variáveis aleatórias, 330-334
 variáveis aleatórias, 327-329
 variável aleatória Gaussiana, 340-344
- Teoria de detecção, 32-33
 Teoria eletromagnética da luz, 19
 Termo de frequência dobrada, 191-192
 Termos de diferença em frequência, 191-192
 Teste de hipótese, 32-33
 Transformação passa-faixa para passa-baixa (banda base), 155-156
 Transformada contínua de Fourier, *veja* Transformada de Fourier (TF)
 Transformada de Fourier, 35-69, 472-474
veja também tipos específicos, por exemplo: transformada rápida de Fourier
 cálculo numérico da, 97-106
 algoritmos FFT, 99-104
 cálculo da TDIF, 104-106
 TDF e TDIF, interpretação da, 98-100
 de sinais periódicos, 66-69
 definição, 32-33, 35-36
 dualidade da, 46-47
 e área sob $G(f)$, 49
 e área sob $g(t)$, 49
 e diferenciação no domínio do tempo, 49-51
 e função delta de Dirac, 58-67
 e integração no domínio do tempo, 51-53
 e regra da conjugação, 45-46
 e relação inversa entre tempo e frequência, 55-58
 e sinal de pulso como soma contínua de funções exponenciais, 37-42
 e teorema da convolução, 53-54
 e teorema da correlação, 53-55
 e teorema da energia de Rayleigh, 54-56
 e teorema da modulação, 53-54
 inversa, 35-36
 linearidade da, 42-45
 notações em, 36-38
 propriedade da dilatação da, 45
 propriedade de deslocamento no tempo da, 47
 propriedade de deslocamento no tempo da, 47-49
 propriedades da, 482
 teoremas, 482
- Transformada de Fourier em tempo discreto 206-209
 Transformada de Hilbert, 114-115, 153n.4(i)
 Transformada discreta de Fourier (TDF), 35, 97-98
 e FFT, 99-101
 interpretação da, 98-100
 Transformada discreta inversa de Fourier (TDIF), 98-99
 cálculo da, 104-106
 interpretação da, 98-100
 Transformada inversa de Hilbert, 114
 Transformada rápida de Fourier (FFT), 32-33, 99-104
- Transistor, invenção do, 20
 Transmissão de dados banda base, 243-267
 de dados M-ário, 256-257
 e canal de Nyquist, 246-249
 e espectro de pulso de cosseno levantado, 249-257
 espectro de pulso de raiz de cosseno levantado, 255-257
 experimento de computador, 260-263
 necessidade de largura de faixa de transmissão, 251-253
 padrão de olho em, 257-263
 porção amortecida do espectro, 252-255
 problema de interferência intersimbólica em, 244-246
- Transmissão de dados M-ário, 260-261
 banda base, 256-257
 padrão de olho para, 260-261, 418-420
- Transmissão de dados passa-faixa, 267
- Transmissão de rádio com faixa ultra larga (UWB), 236
- Transmissores, 23
- Treinamento assistido por piloto, 266, 317-318
- Triodo tubo de vácuo, 20
- U**
- Ungerboeck, G., 21-22
- V**
- Valor médio quadrático, 355-356
 Van Duren, H. C. A., 21-22
 Van Vleck, J. H., 21-22
 Variáveis aleatórias, 327-329, 339-341
 experiments com várias, 330-334
 Gaussiano, 340-344
 transformação de, 339-341
- Variável aleatória de Bernoulli, 328-329, 338-339
 Variável aleatória Gaussiana, 340-344
 Variável aleatória normal, 340-341n. 5
- Variável de busca, 86-87
 Variável de varredura, 86-87
 Varredura, 23
 VCO, *veja* Oscilador controlado por tensão
 Vocoders de fala, 305-306n.1
 VoIP, *veja* Voz sobre IP
 Von Neumann, John, 21-22
 Voz sobre IP (VoIP), 28-30
 VSB mais sinal da portadora, 150-151
- W**
- Weaver, Warren, 21-22
 World Wide Web, 21-22
- Z**
- Zworykin, Vladimir K., 20