

# 2

## A camada física

Neste capítulo, analisaremos a camada mais baixa da hierarquia em nosso modelo de referência, a camada física. Ela define as interfaces elétrica, de sincronização e outras, pelas quais os bits são enviados como sinais pelos canais. A camada física é o alicerce sobre o qual a rede é construída. Como as propriedades dos diferentes tipos de canais físicos determinam o desempenho (p. ex., throughput, latência e taxa de erros), este é um bom lugar para começar nossa jornada até a “terra das redes”.

Inicialmente, discutiremos três meios de transmissão: guiado ou cabeado (p. ex., fio de cobre, cabo coaxial, fibra óptica), sem fio (rádio terrestre) e satélite. Cada uma dessas tecnologias tem diferentes propriedades que afetam o projeto e o desempenho das redes que as utilizam. Esse material fornecerá informações fundamentais sobre as principais tecnologias de transmissão usadas em redes modernas.

Depois, faremos uma análise teórica da transmissão de dados, apenas para descobrir que a mãe natureza impõe limites sobre o que pode ser enviado por um canal (ou seja, um meio de transmissão físico usado para enviar bits). Em seguida, abordaremos a modulação digital, que trata de como os sinais analógicos são convertidos em bits digitais e em sinais novamente. A seguir, examinaremos os esquemas de multiplexação, explorando como várias conversas podem ser feitas no mesmo meio de transmissão ao mesmo tempo, sem interferir umas com as outras.

Por fim, veremos três exemplos de sistemas de comunicação usados na prática nas redes de computadores a longas distâncias: o sistema de telefonia (fixa), o sistema de telefonia móvel (ou celular) e o sistema de televisão a cabo. Como os três são muito importantes na prática, dedicaremos uma boa quantidade de espaço a cada um.

### 2.1 MEIOS DE TRANSMISSÃO GUIADOS

O objetivo da camada física é transmitir um fluxo bruto de bits de uma máquina para outra. Vários meios físicos podem ser usados para realizar a transmissão real. Quase sempre, o meio de transmissão que ocorre por um fio ou cabo físico é chamado de **meio de transmissão guiado**, pois as transmissões de sinal são guiadas ao longo de um caminho por um fio ou cabo. Os meios de transmissão guiados mais comuns são o cabo de cobre (na forma de cabo coaxial ou par trançado) e a fibra óptica. Cada tipo de meio de transmissão guiada tem seu próprio conjunto de desvantagens em termos de frequência, largura de banda, atraso, custo e facilidade de instalação e manutenção. A largura de banda é uma métrica da capacidade de transporte de um meio. Ela é medida em **Hz** (ou MHz ou GHz) e seu nome é uma homenagem ao físico alemão Heinrich Hertz. Isso será discutido em detalhes neste capítulo.

#### 2.1.1 Armazenamento persistente

Uma das formas mais comuns de transportar dados de um dispositivo para outro é gravá-los em armazenamento persistente, como o armazenamento magnético ou em estado sólido (p. ex., DVDs graváveis) e transportar fisicamente a fita ou os discos para a máquina de destino, onde eles finalmente serão lidos. Apesar de não ser tão sofisticado quanto a utilização de um satélite de comunicação geossíncrono, esse método costuma ser muito mais econômico, em especial nas aplicações em que a alta largura de banda ou o custo por bit transportado é o fator mais importante.

Um cálculo simples esclarecerá essa questão. Uma fita Ultrium de padrão industrial pode armazenar 30 terabytes. Uma caixa de  $60 \times 60 \times 60$  cm pode conter cerca de 1.000 fitas desse tipo, perfazendo uma capacidade total de 800 terabytes, ou 6.400 terabits (6,4 petabits). Uma caixa de fitas pode ser entregue em qualquer parte dos Estados Unidos em 24 horas pelo serviço de Sedex dos Correios, pela Federal Express e por outras transportadoras. A largura de banda efetiva dessa transmissão é de 6.400 terabits/86.400 s, ou um pouco mais de 70 Gbps. Se o destino estiver a apenas uma hora de distância, a largura de banda será ampliada para mais de 1.700 Gbps. Nenhuma rede de computadores consegue nem mesmo se aproximar desse desempenho. Logicamente, as redes estão ficando mais rápidas, mas as densidades das fitas também estão aumentando.

Se considerarmos o custo, obteremos um quadro semelhante. O custo de uma fita Ultrium é de aproximadamente US\$ 40 quando a compra é feita no atacado. Uma fita pode ser reutilizada pelo menos 10 vezes. Portanto, o custo das fitas passa a ser US\$ 4.000 por caixa, para cada utilização. Adicione a esse montante mais US\$ 1.000 pelo frete (provavelmente muito menos) e teremos um custo final aproximado de US\$ 5.000 para transportar 800 TB. Consequentemente, para transportar 1 gigabyte, gastaremos pouco mais de meio centavo de dólar. Nenhuma rede pode competir com esses valores. Moral da história:

*Nunca subestime a largura de banda de uma caminhonete cheia de fitas “voando” na estrada.*

Para mover quantidades de dados *muito* grandes, geralmente essa é a melhor solução. A Amazon tem o que chama de “Snowmobile”, que é um grande caminhão cheio de milhares de discos rígidos, todos conectados a uma rede de alta velocidade dentro do caminhão. A capacidade total do caminhão é de 100 PB (100.000 TB ou 100 milhões de GB). Quando uma empresa tem uma grande quantidade de dados para mover, ela pode levar o caminhão às suas instalações e se conectar à rede de fibra óptica da empresa e, em seguida, sugar todos os dados para o caminhão. Feito isso, o caminhão se dirige para outro local e despeja todos os dados. Por exemplo, uma empresa que deseja substituir seu próprio centro de dados completo pela nuvem da Amazon pode estar interessada nesse serviço. Para enormes volumes de dados, nenhum outro método de transporte de dados pode se aproximar disso.

## 2.1.2 Pares trançados

Embora as características de largura de banda do armazenamento persistente sejam excelentes, os aspectos de atraso são ruins. O tempo de transmissão é medido em horas ou dias, e não em milissegundos. Muitas aplicações, incluindo a Web, videoconferência e jogos on-line, contam com a transmissão de dados com pouco atraso. Um dos meios

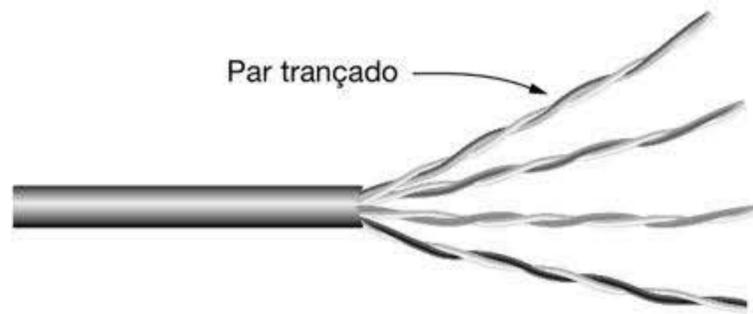
de transmissão mais antigos e ainda mais comuns é o **par trançado**, que consiste em dois fios de cobre encapados, em geral com cerca de 1 mm de espessura. Os fios são enrolados de forma helicoidal, assim como uma molécula de DNA. O trançado dos fios é feito porque dois fios paralelos formam uma antena simples. Quando os fios são trançados, as ondas de diferentes partes dos fios se cancelam, o que significa menor interferência. Um sinal normalmente é transportado a partir da diferença das tensões terminais (diferença de potencial – ddp) entre os dois fios no par. Essa transmissão do sinal como a diferença entre os dois níveis de tensão, em vez de uma tensão absoluta, oferece melhor imunidade ao ruído externo, pois o ruído tende a afetar a tensão trafegando pelos dois fios da mesma forma, mantendo a ddp relativamente inalterada.

A aplicação mais comum do par trançado é o sistema telefônico. Quase todos os telefones estão conectados à estação central da companhia telefônica por um par trançado. Tanto as chamadas telefônicas quanto o acesso à Internet por ADSL utilizam essas linhas. Os pares trançados podem se estender por diversos quilômetros sem amplificação, mas, quando se trata de distâncias mais longas, o sinal é atenuado e existe a necessidade de repetidores. Quando muitos pares trançados percorrem paralelamente uma distância muito grande, como acontece na ligação entre um prédio e a estação central da companhia telefônica, eles são reunidos e envolvidos por uma capa protetora. Se não estivessem trançados, esses pares provocariam muitas interferências. Em locais onde as linhas telefônicas são instaladas em postes, é comum vermos cabos de pares trançados com vários centímetros de diâmetro.

Os pares trançados podem ser usados na transmissão de sinais analógicos ou digitais. A largura de banda depende da espessura do fio e da distância percorrida, mas, em muitos casos, é possível alcançar diversos megabits/s por alguns quilômetros, e mais do que isso quando são usados alguns truques. Em virtude do desempenho adequado, da grande disponibilidade e do baixo custo, os pares trançados são usados em larga escala e é provável que permaneçam assim nos próximos anos.

O cabeamento de par trançado pode ser de vários tipos. A variedade mais comum empregada em muitos prédios é chamada cabeamento de **Categoria 5e**, ou “Cat 5e”. Um par trançado de Categoria 5e consiste em dois fios isolados e levemente trançados. Quatro pares desse tipo normalmente são agrupados em uma capa plástica para proteger os fios e mantê-los juntos. Esse arranjo pode ser visto na Figura 2.1.

Diferentes padrões de LAN podem usar os pares trançados de formas distintas. Por exemplo, a Ethernet de 100 Mbps usa dois (dos quatro) pares, um para cada direção. Para alcançar velocidades mais altas, a Ethernet de 1 Gbps usa todos os quatro pares nas duas direções



**Figura 2.1** Cabo UTP Categoria 5e com quatro pares trançados. Estes cabos podem ser usados para redes locais (LANs).

simultaneamente; isso requer que o receptor decomponha o sinal que é transmitido localmente.

Neste ponto, devemos explicar alguma terminologia geral. Os enlaces que podem ser usados nos dois sentidos ao mesmo tempo, como uma estrada de mão dupla, são chamados enlaces **full-duplex**. Ao contrário, os que são usados em qualquer sentido, mas apenas um de cada vez, como uma linha férrea de trilho único, são chamados enlaces **half-duplex**. Uma terceira categoria consiste em enlaces que permitem o tráfego em apenas uma direção, como uma rua de mão única, chamados enlaces **simplex**.

Retornando ao par trançado, os cabos Cat 5 substituíram os cabos **Categoria 3** mais antigos com um cabo semelhante que usa o mesmo conector, porém com mais voltas por metro. Mais voltas resultam em menos interferências e em um sinal de melhor qualidade por distâncias maiores, tornando os cabos mais adequados para a comunicação de computador de alta velocidade, especialmente LANs Ethernet de 100 Mbps e 1 Gbps.

Os fios mais novos provavelmente serão de **Categoria 6** ou mesmo de **Categoria 7**, as quais têm especificações mais rígidas para lidar com sinais de larguras de banda maiores. Alguns cabos na Categoria 6 e superiores podem aceitar os enlaces de 10 Gbps que agora são comuns em muitas redes, como em novos prédios comerciais. Os fios de **Categoria 8** trabalham com velocidades maiores que as categorias inferiores, mas operam somente a curtas distâncias, em torno de 30 m, e portanto são adequados apenas dentro dos centros de dados. O padrão de Categoria 8 tem duas opções: Classe I, que é compatível com a Categoria 6A, e Classe II, que é compatível com a Categoria 7A.

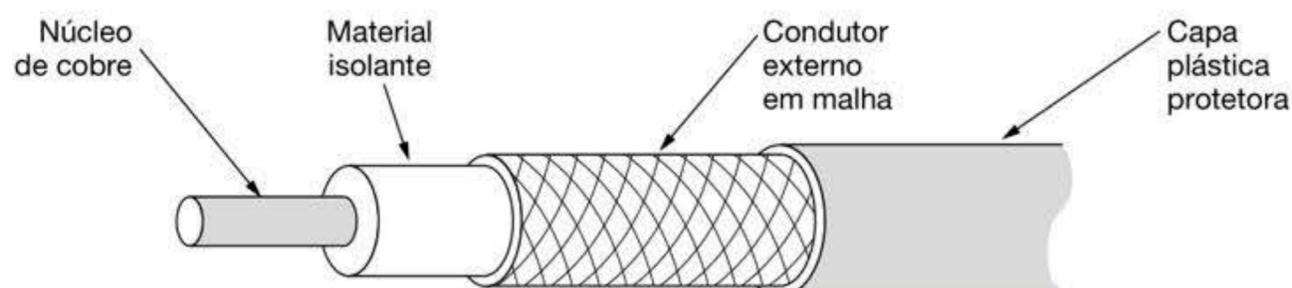
Até a Categoria 6, esses tipos de fios são conhecidos como **par trançado não blindado**, ou **UTP (Unshielded Twisted Pair)**, pois consistem simplesmente em fios e isolamento. Ao contrário, os cabos de Categoria 7 possuem uma blindagem nos pares de fios individuais e também ao redor do cabo inteiro (mas dentro da capa plástica protetora). A blindagem reduz a suscetibilidade à interferência externa e linha cruzada com outros cabos vizinhos, atendendo às especificações de desempenho exigidas. Os cabos são reminiscências dos cabos de pares trançados blindados de alta qualidade, porém grossos e caros, que a IBM introduziu no início da década de 1980, mas que não se tornaram populares fora das instalações da empresa. Evidentemente, é hora de tentar novamente.

### 2.1.3 Cabo coaxial

Outro meio de transmissão comum é o **cabo coaxial** (conhecido por muitos apenas como “coax”). Ele tem melhor blindagem e maior largura de banda que os pares trançados e, assim, pode se estender por distâncias mais longas em velocidades mais altas. Dois tipos de cabo coaxial são amplamente utilizados. Um deles, o cabo de 50 ohms, é comumente empregado nas transmissões digitais. O outro tipo, o cabo de 75 ohms, é usado com frequência nas transmissões analógicas e de televisão a cabo. Essa distinção se baseia mais em fatores históricos do que técnicos (p. ex., as primeiras antenas dipolo tinham uma impedância de 300 ohms e era fácil desenvolver transformadores de casamento de impedância de 4:1). Começando em meados da década de 1990, as operadoras de TV a cabo começaram a oferecer acesso à Internet por cabo, o que tornou o cabo de 75 ohms mais importante para a comunicação de dados.

Um cabo coaxial consiste em um fio de cobre esticado na parte central, protegido por um material isolante. O isolante é envolvido por um condutor cilíndrico, geralmente como uma malha sólida entrelaçada. O condutor externo é coberto por uma camada plástica protetora. A Figura 2.2 apresenta uma vista de corte de um cabo coaxial.

A construção e a blindagem do cabo coaxial proporcionam a ele uma boa combinação de alta largura de banda e excelente imunidade ao ruído (p. ex., contra controles de porta de garagem e fornos de micro-ondas). A largura de banda possível depende da qualidade e do tamanho do



**Figura 2.2** Um cabo coaxial.

cabo. O cabo coaxial possui largura de banda extremamente alta; os cabos modernos têm uma largura de banda de até 6 GHz, permitindo, assim, que muitas conversas sejam transmitidas simultaneamente por um único cabo coaxial (um único programa de televisão poderia ocupar aproximadamente 3,5 MHz). Os cabos coaxiais eram muito usados no sistema telefônico para linhas de longa distância, mas agora estão sendo substituídos por fibras ópticas nessas rotas. Porém, eles ainda são usados em larga escala pelas redes de televisão a cabo e em redes metropolitanas, além da conectividade com a Internet de alta velocidade para residências em muitas partes do mundo.

### 2.1.4 Linhas de energia elétrica

As redes de telefonia e de televisão a cabo não são as únicas fontes de fiação que podem ser reutilizadas para a comunicação de dados. Há um outro tipo de fiação ainda mais comum: as linhas de energia elétrica. Estas oferecem energia elétrica às casas, e a fiação elétrica dentro das casas distribui a energia às tomadas elétricas.

O uso das linhas de energia elétrica para comunicação de dados é uma ideia antiga. Elas têm sido usadas pelas companhias de eletricidade para a comunicação de baixo nível há muitos anos, como para fazer medição remota e controlar dispositivos em casa (p. ex., o padrão X10). Nos últimos anos, tem havido um interesse renovado na comunicação de alto nível por essas linhas, tanto dentro de casa, como uma LAN, quanto fora dela, para o acesso de banda larga à Internet. Vamos nos concentrar no cenário mais comum: usar fios elétricos dentro de casa.

A conveniência de usar linhas de energia para a rede de dados deve ser clara. Basta conectar uma TV e um receptor na parede, o que você precisa fazer de qualquer forma, pois ele precisa de energia, e ele poderá enviar e receber filmes pela fiação elétrica. Essa configuração pode ser vista na Figura 2.3. Não há outro conector ou rádio. O sinal de dados é sobreposto ao sinal de baixa frequência (ou no fio ativo, ou “quente”), enquanto os dois sinais usam a fiação ao mesmo tempo.

A dificuldade em usar a fiação elétrica domiciliar como uma rede é que ela foi projetada para distribuir energia elétrica. Essa tarefa é muito diferente de distribuir sinais de dados, algo para o qual a fiação doméstica é pouco

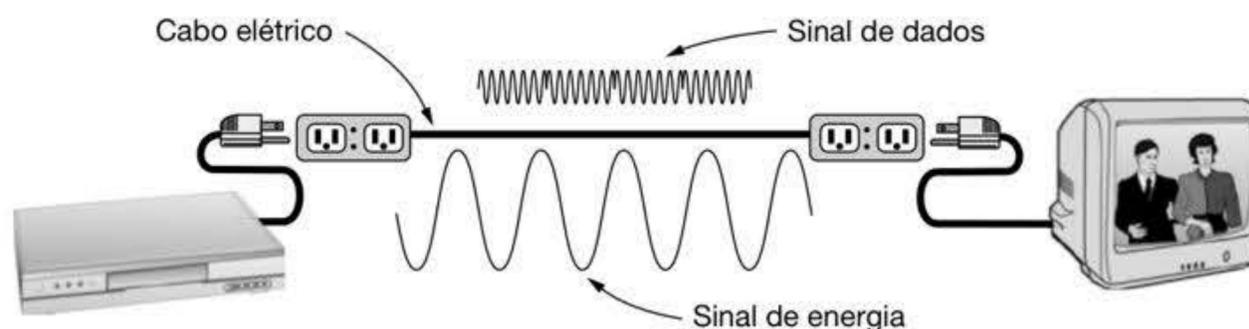
eficiente. Os sinais elétricos são enviados a 50-60 Hz e a fiação atenua os sinais de frequência muito mais alta (MHz) necessários para a comunicação de dados de alto nível. As propriedades elétricas da fiação variam de uma casa para outra e mudam à medida que os aparelhos são ligados e desligados, fazendo os sinais de dados oscilarem pela fiação. As correntes transitórias quando os aparelhos são ligados e desligados criam ruído por uma grande faixa de frequências. Sem o trançado cuidadoso dos pares trançados, a fiação elétrica atua como uma boa antena, apanhando sinais externos e emitindo sinais próprios. Esse comportamento significa que, para atender aos requisitos da regulamentação, o sinal de dados precisa excluir frequências licenciadas, como as faixas de radioamador.

Apesar dessas dificuldades, é possível enviar pelo menos 500 Mbps pela fiação elétrica doméstica usando esquemas de comunicação que resistem às frequências comprometidas e às sucessões de erros. Muitos produtos usam diversos padrões próprios para as redes de energia elétrica, de modo que há padrões em desenvolvimento ativo.

### 2.1.5 Fibra óptica

Muitas pessoas na indústria de informática se orgulham da rapidez com que a tecnologia usada nos computadores vem melhorando, conforme a lei de Moore, que prevê a duplicação do número de transistores por chip a cada dois anos aproximadamente (Kuszyk e Hammoudeh, 2018). O IBM PC original (de 1981) funcionava com uma velocidade de clock de 4,77 MHz; 40 anos depois, os PCs podiam usar uma CPU de quatro núcleos a 3 GHz. Esse aumento é um ganho de fator em torno de 2.500. Impressionante.

No mesmo período, enlaces de comunicação remotos passaram de 45 Mbps (uma linha T3 no sistema telefônico) para 100 Gbps (uma linha moderna de longa distância). Esse ganho também é impressionante, um fator de mais de 2.000, enquanto, no mesmo período, a taxa de erros passou de  $10^{-5}$  por bit para quase zero. Na última década, as CPUs isoladas se aproximaram dos limites físicos, motivo pelo qual agora é o número de CPUs que está sendo aumentado por chip. Por sua vez, a largura de banda alcançável com a tecnologia de fibra óptica pode ultrapassar a casa dos 50.000 Gbps (50 Tbps) e nem estamos perto de alcançar esses limites. O limite prático atual é de cerca de 100 Gbps,



**Figura 2.3** Rede de dados que usa a fiação elétrica domiciliar.

em razão de nossa incapacidade de realizar a conversão entre sinais elétricos e ópticos em uma velocidade maior. Para criar enlaces de maior capacidade, muitos canais simplesmente correm em paralelo por uma única fibra.

Nesta seção, estudaremos a fibra óptica para descobrir como funciona essa tecnologia de transmissão. Na corrida entre informática e comunicação, esta pode ganhar em virtude das redes de fibra óptica. As implicações reais disso seriam essencialmente largura de banda infinita e uma nova premissa de que os computadores são terrivelmente lentos e, por essa razão, as redes deveriam tentar evitar a computação a todo custo, independentemente do desperdício de largura de banda que isso signifique. Essa mudança levará algum tempo para entrar na cabeça de uma geração de cientistas da computação e engenheiros ensinados a pensar em termos dos baixos limites de transmissão impostos pelos fios de cobre.

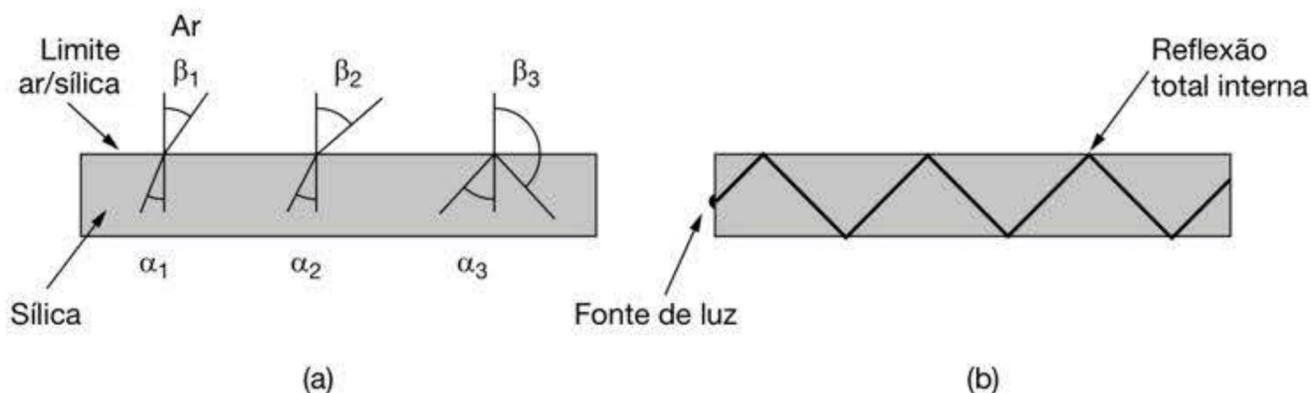
Naturalmente, esse cenário não diz tudo, pois não inclui custos. O custo para instalar fibra até chegar aos consumidores, evitando a baixa largura de banda dos fios e a disponibilidade limitada de espectro, é muito alto. A energia para a movimentação de bits também custa mais que a energia para sua computação. Sempre podemos ter ilhas de injustiça, onde a computação ou a comunicação é basicamente gratuita. Por exemplo, na borda da Internet, usamos computação e armazenamento para o problema de compactação e caching de conteúdo, tudo para fazer melhor uso dos enlaces de acesso à Internet. Dentro da Internet, podemos fazer o contrário, com empresas como o Google movendo grandes quantidades de dados pela rede até onde for mais barato armazená-los ou computá-los.

A fibra óptica é usada para transmissão por longa distância nos backbones da rede, LANs de alta velocidade (embora, até aqui, o cobre sempre tenha conseguido acompanhar) e acesso à Internet em alta velocidade, como **FTTH (Fiber to the Home)**. Um sistema de transmissão óptico tem três componentes-chave: a fonte de luz, o meio de transmissão e o detector. Convencionalmente, um pulso de luz indica um bit 1 e a ausência de luz indica um bit 0. O meio de transmissão é uma fibra de vidro ultrafina. O detector gera um pulso elétrico quando a luz incide sobre

ele. Conectando uma fonte de luz em uma ponta de uma fibra óptica e um detector na outra, temos um sistema de transmissão de dados unidirecional (i.e., simplex) que aceita um sinal elétrico, o converte e o transmite por pulsos de luz, para depois novamente converter a saída para um sinal elétrico na ponta receptora.

Esse sistema de transmissão perderia luz para o meio e seria inútil na prática se não fosse por um princípio interessante da física. Quando um raio de luz passa de um meio para outro – por exemplo, de sílica fundida para o ar –, o raio é refratado (inclinado) no limite sílica/ar, como mostra a Figura 2.4(a). Aqui, vemos um raio de luz incidindo no limite em um ângulo  $\alpha_1$  emergindo em um ângulo  $\beta_1$ . A quantidade de refração depende das propriedades dos dois meios (em particular, seus índices de refração). Para ângulos de incidência acima de um certo valor crítico, a luz é refratada de volta para a sílica; nada escapa para o ar. Assim, um raio de luz incidente em um ângulo crítico ou acima é interceptado dentro da fibra, como mostra a Figura 2.4(b), e pode se propagar por muitos quilômetros praticamente sem perdas.

O exemplo da Figura 2.4(b) mostra apenas um raio interceptado, mas, como qualquer raio de luz incidente no limite acima do ângulo crítico será refletido internamente, muitos raios distintos estarão ricocheteando em diferentes ângulos. Dizemos que cada raio tem um modo específico; assim, uma fibra que apresenta essa propriedade é chamada de **fibra multimodo**. Se o diâmetro da fibra for reduzido a alguns comprimentos de onda de luz (menos de 10 micra, ao contrário de mais de 50 micra para a fibra multimodo), a fibra agiria como um guia de onda e a luz só poderá se propagar em linha reta, sem ricochetear, produzindo, assim, uma **fibra de modo único** ou **fibra monomodo**. As fibras de modo único são mais caras, mas são amplamente utilizadas em distâncias mais longas; elas podem transmitir sinais para distâncias aproximadamente 50 vezes maiores que as fibras multimodo. As fibras de modo único disponíveis no momento podem transmitir dados a 100 Gbps por 100 km sem amplificação. Foram obtidas taxas de dados ainda mais altas em laboratório, para distâncias mais curtas. A decisão entre o uso de fibra monomodo ou multimodo depende da aplicação. A fibra multimodo pode ser usada



**Figura 2.4** (a) Três exemplos de um raio de luz dentro de uma fibra de sílica incidindo no limite fronteira ar/sílica em diferentes ângulos. (b) A luz interceptada pela reflexão total interna.

para transmissões de até cerca de 15 km e permitir o uso de equipamento de fibra óptica relativamente mais barato. Contudo, a largura de banda da fibra multimodo torna-se mais limitada com o aumento da distância.

### Transmissão de luz na fibra

As fibras ópticas são feitas de vidro, que, por sua vez, é produzido a partir da areia, uma matéria-prima abundante e de baixo custo. Os antigos egípcios já dominavam a manufatura do vidro, mas o material produzido por eles não podia ter mais de 1 mm de espessura para que a luz pudesse atravessá-lo. O vidro transparente usado nas janelas foi desenvolvido durante a Renascença. O vidro usado nas modernas fibras ópticas é tão transparente que se, em vez de água, os oceanos fossem cheios desse tipo de material, seria possível ver o fundo do mar da superfície, da mesma forma que é possível ver o solo quando voamos de avião em um dia claro.

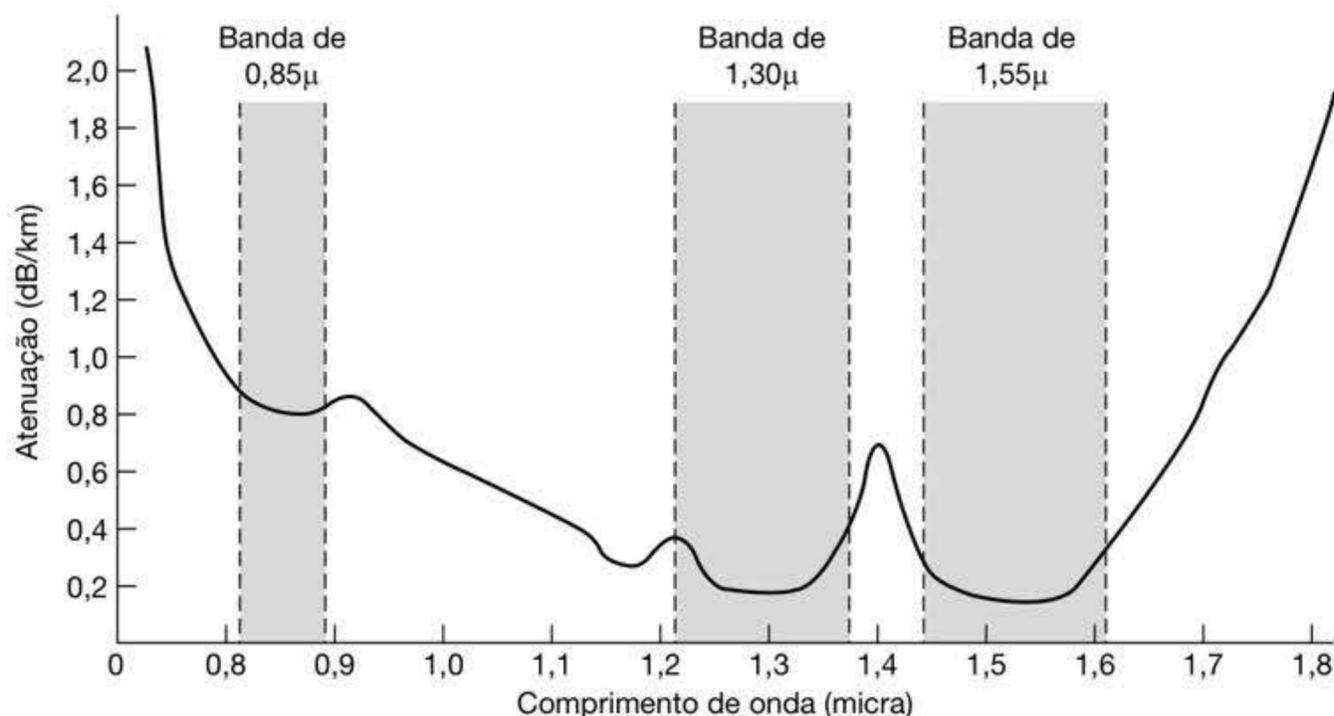
A atenuação de luz através do vidro depende do comprimento de onda da luz (bem como de algumas propriedades físicas do vidro). Ela é definida como a razão da potência do sinal de entrada e saída. Para o tipo de vidro usado nas fibras, a atenuação é mostrada na Figura 2.5, em decibéis (dB) por quilômetro linear de fibra. Por exemplo, quando o fator de potência é igual a dois, obtemos uma atenuação de  $10 \log_{10} 2 = 3$  dB. Discutiremos sobre decibéis mais adiante. Resumindo, essa é uma forma logarítmica de medir as taxas de potência, com 3 dB significando um fator de potência de dois. A figura mostra a parte do infravermelho do espectro que, na prática, é a utilizada. A luz visível tem comprimentos de onda ligeiramente mais curtos, que variam de 0,4 a 0,7 microm (1 microm é igual a  $10^{-6}$  m). Na verdade, esses comprimentos de onda seriam de 400 nm a 700 nm, mas manteremos a nomenclatura tradicional.

A comunicação óptica comumente utiliza três bandas de comprimentos de onda. Elas são centralizadas em 0,85, 1,30 e 1,55 microm, respectivamente. As três bandas têm entre 25.000 e 30.000 GHz de largura. A banda de 0,85 microm foi usada primeiro. Ela tem maior atenuação e, por isso, é usada para distâncias mais curtas, mas, nesse comprimento de onda, os lasers e os circuitos eletrônicos podem ser produzidos a partir do mesmo material (arseneto de gálio). As duas últimas bandas têm boas propriedades de atenuação (uma perda inferior a 5% por quilômetro). A banda de 1,55 microm agora é muito utilizada com amplificadores dopados com érbio, que atuam diretamente no domínio óptico.

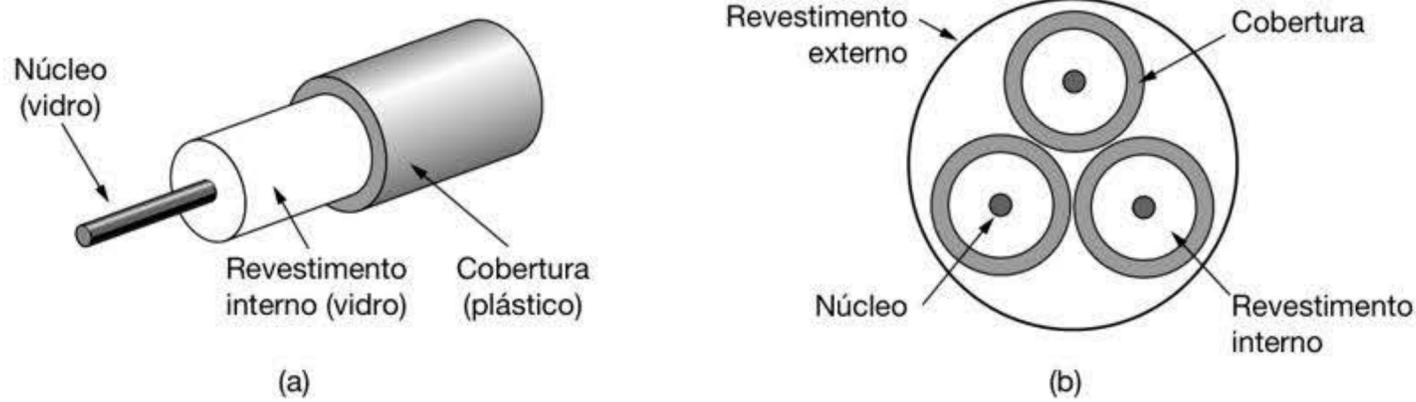
Os pulsos de luz enviados através de uma fibra se expandem à medida que se propagam. Essa expansão é chamada **dispersão cromática**. O volume da dispersão depende do comprimento de onda. Uma forma de impedir que esses pulsos dispersos se sobreponham é aumentar a distância entre eles, mas isso só pode ser feito reduzindo-se a taxa de sinalização. Felizmente, descobriu-se que, quando os pulsos são produzidos em uma forma especial relacionada ao recíproco do cosseno hiperbólico, praticamente todos os efeitos de dispersão são cancelados e é possível enviar pulsos por milhares de quilômetros sem que haja uma distorção significativa. Esses pulsos são chamados **sólitons**. Eles estão começando a ser muito utilizados na prática.

### Cabos de fibra

Os cabos de fibra óptica são semelhantes aos cabos coaxiais, exceto por não terem a malha metálica. A Figura 2.6(a) mostra a vista lateral de uma única fibra. No centro fica o núcleo de vidro através do qual a luz se propaga. Nas fibras multimodo, o núcleo normalmente tem 50 microm de diâmetro, o que corresponde à espessura de um fio de



**Figura 2.5** Atenuação da luz na fibra, na região do infravermelho.



**Figura 2.6** (a) Vista lateral de uma única fibra. (b) Vista da extremidade de um cabo com três fibras.

cabelo humano. Nas fibras de modo único, o núcleo tem entre 8 e 10 micra.

O núcleo é envolvido por um revestimento de vidro com um índice de refração inferior ao do núcleo, para manter toda a luz nele. Em seguida, há uma cobertura de plástico fino para proteger o revestimento interno. Geralmente, as fibras são agrupadas em feixes, protegidas por um revestimento externo. A Figura 2.6(b) mostra um cabo com três fibras.

Normalmente, os cabos de fibra terrestres são colocados no solo a um metro da superfície, onde ocasionalmente são atacados por retroescavadeiras ou roedores. Próximo ao litoral, cabos de fibra transoceânicos são enterrados em trincheiras por uma espécie de arado marítimo. Em águas profundas, eles são depositados no fundo, onde podem ser arrastados por redes de pesca ou comidos por lulas gigantes.

As fibras podem estar conectadas de três maneiras diferentes. Primeiro, elas podem ter conectores em suas extremidades e serem plugadas em soquetes de fibra. Os conectores perdem de 10 a 20% da luz, mas facilitam a reconfiguração dos sistemas. Segundo, elas podem ser unidas mecanicamente. Nesse caso, as duas extremidades são cuidadosamente colocadas uma perto da outra em uma luva especial e fixadas no lugar. O alinhamento pode ser melhorado fazendo-se a luz passar pela junção e, em seguida, realizando-se pequenos ajustes cuja finalidade é maximizar o sinal. As junções mecânicas são realizadas em cerca de 5 minutos por uma equipe treinada e resultam em uma perda de 10% da luz. Terceiro, duas peças de fibra podem ser fundidas de modo a formar uma conexão

sólida. A união por fusão é quase tão boa quanto uma fibra sem emendas; no entanto, mesmo nesse caso, há uma pequena atenuação. Nos três tipos de uniões podem ocorrer reflexões no ponto de junção, e a energia refletida pode interferir no sinal.

Dois tipos de fontes de luz geralmente são usados para fazer a sinalização: os diodos emissores de luz (Light Emitting Diodes, ou LEDs) e os lasers semicondutores. Eles têm diferentes propriedades, como mostra a Figura 2.7. O comprimento de onda desses elementos pode ser ajustado pela inserção de interferômetros de Fabry-Perot ou Mach-Zehnder entre a fonte e a fibra. Os interferômetros de Fabry-Perot são cavidades ressonantes simples que consistem em dois espelhos paralelos. A luz incide perpendicularmente aos espelhos. O comprimento da cavidade filtra os comprimentos de onda que cabem em um número inteiro de períodos. Os interferômetros de Mach-Zehnder separam a luz em dois feixes, os quais percorrem distâncias ligeiramente diferentes e são recombinados no destino, ficando em fase para certos comprimentos de onda.

A extremidade de recepção de uma fibra óptica consiste em um fotodiodo, que emite um pulso elétrico ao ser atingido pela luz. O tempo de resposta de um fotodiodo, que converte o sinal do domínio óptico para o elétrico, limita as taxas de dados a cerca de 100 Gbps. O ruído térmico também é importante, de modo que um pulso de luz deve conduzir energia suficiente para ser detectado. Com pulsos de potência suficiente, a taxa de erros pode se tornar muito pequena.

Item	LED	Laser semicondutor
Taxa de dados	Baixa	Alta
Tipo de fibra	Multimodo	Multimodo ou modo único
Distância	Curta	Longa
Vida útil	Longa	Curta
Sensibilidade à temperatura	Insignificante	Substancial
Custo	Baixo	Alto

**Figura 2.7** Uma comparação entre laser semicondutor e LEDs utilizados como fontes de luz.

## Comparação entre fibras ópticas e fios de cobre

É instrutivo comparar a fibra com o cobre. A fibra tem muitas vantagens. Para começar, ela pode gerenciar larguras de banda muito mais altas do que o cobre. Essa característica sozinha já justificaria seu uso nas redes de última geração. Em razão da baixa atenuação, os repetidores só são necessários a cada 50 km de distância em linhas longas, em comparação com a distância de 5 km no caso do cobre, resultando em uma economia de custo significativa. A fibra também tem a vantagem de não ser afetada por picos de tensão, interferência eletromagnética ou quedas no fornecimento de energia. Ela também está imune à ação corrosiva de alguns elementos químicos que pairam no ar, o que é importante em ambientes industriais desfavoráveis.

Por mais estranho que possa parecer, as empresas telefônicas gostam da fibra por outra razão: ela é fina e leve. Muitos dos dutos de cabos atuais estão completamente lotados, de modo que não há espaço para aumentar sua capacidade. Além da remoção e subsequente substituição de todo o cobre por fibras esvaziar os dutos, o cobre tem um excelente valor de revenda para as refinarias especializadas, pois trata-se de um minério de altíssima qualidade. Além disso, a fibra é muito mais leve que o cobre. Mil pares trançados com 1 km de comprimento pesam 8 toneladas. Duas fibras têm mais capacidade e pesam apenas 100 kg, reduzindo de maneira significativa a necessidade de sistemas mecânicos de suporte, que exigem mais manutenção. Nas novas rotas, as fibras são preferidas por terem um custo de instalação muito mais baixo. Por fim, elas não desperdiçam luz e dificilmente são interceptadas. Por essas razões, a fibra é uma alternativa com excelente nível de segurança contra possíveis escutas telefônicas.

No entanto, a fibra tem a desvantagem de ser uma tecnologia menos familiar, exigindo conhecimentos que nem todos os engenheiros têm e, além disso, as fibras podem ser danificadas com facilidade, se forem encurvadas demais. Como a transmissão óptica é basicamente unidirecional, a comunicação bidirecional exige duas fibras ou duas bandas de frequência em uma única fibra. Por fim, as interfaces de fibra são mais caras que as interfaces elétricas. Apesar disso, o futuro de toda a comunicação fixa de dados para distâncias superiores a alguns metros depende claramente da fibra. Para obter mais informações sobre todos os aspectos das fibras ópticas e de suas redes, consulte Pearson (2015).

## 2.2 TRANSMISSÃO SEM FIO

Muitas pessoas agora possuem conectividade sem fio para muitos dispositivos, desde notebooks e smartphones até relógios e refrigeradores inteligentes. Todos esses dispositivos contam com a comunicação sem fio para transmitir informações para outros dispositivos e extremidades da rede.

Nas próximas seções, examinaremos os conceitos básicos da comunicação sem fio em geral, pois ela tem muitas outras aplicações importantes além de oferecer conectividade aos usuários que desejam navegar na Web enquanto estão na praia. Existem algumas outras circunstâncias em que a comunicação sem fio apresenta vantagens até mesmo para dispositivos fixos. Por exemplo, quando há dificuldades para instalar cabos de fibra óptica em um prédio, por causa de acidentes geográficos (montanhas, florestas, pântanos, etc.), a tecnologia de transmissão sem fio pode ser mais apropriada. Vale a pena observar que a moderna comunicação digital sem fio teve início como um projeto de pesquisa do professor Norman Abramson, da University of Hawaii, na década de 1970, no qual os usuários estavam separados de seu centro de computação pelo Oceano Pacífico e o sistema de telefonia era inadequado. Discutiremos a respeito desse sistema, ALOHA, no Capítulo 4.

### 2.2.1 O espectro eletromagnético

Quando se movem, os elétrons criam ondas eletromagnéticas que podem se propagar pelo espaço livre (até mesmo no vácuo). Essas ondas foram previstas pelo físico inglês James Clerk Maxwell em 1865 e foram observadas pela primeira vez pelo físico alemão Heinrich Hertz em 1887. O número de oscilações de uma onda eletromagnética por segundo é chamado **frequência**,  $f$ , e é medido em Hz (em homenagem a Heinrich Hertz). A distância entre dois pontos máximos (ou mínimos) consecutivos é chamada **comprimento de onda**, designada universalmente pela letra grega  $\lambda$  (lambda).

Quando se instala uma antena de tamanho apropriado em um circuito elétrico, as ondas eletromagnéticas podem ser transmitidas e recebidas com eficiência por um receptor localizado a uma distância bastante razoável. Toda a comunicação sem fio é baseada nesse princípio.

No vácuo, todas as ondas eletromagnéticas viajam à mesma velocidade, independentemente de sua frequência. Essa velocidade, geralmente chamada **velocidade da luz**,  $c$ , é aproximadamente igual a  $3 \times 10^8$  m/s, ou cerca de 30 cm por nanossegundo. No cobre ou na fibra, a velocidade cai para cerca de 2/3 desse valor e se torna ligeiramente dependente da frequência. A velocidade da luz é o limite máximo que se pode alcançar no universo. Nenhum objeto ou sinal pode se mover com maior rapidez do que ela.

A relação fundamental entre  $f$ ,  $\lambda$  e  $c$  (no vácuo) é:

$$\lambda f = c \quad (2.1)$$

Como  $c$  é uma constante, se conhecermos  $f$ , chegaremos a  $\lambda$  e vice-versa. Via de regra, quando  $\lambda$  é medido em metros e  $f$  em MHz,  $\lambda f \approx 300$ . Por exemplo, ondas de 100 MHz têm cerca de 3 m de comprimento, ondas de 1.000 MHz têm 0,3 m e ondas com 0,1 metro têm uma frequência igual a 3.000 MHz.

O espectro eletromagnético é mostrado na Figura 2.8. As faixas de rádio, micro-ondas, infravermelho e luz visível do espectro podem ser usadas na transmissão de informações, por meio de modulação da amplitude, da frequência ou da fase das ondas. A luz ultravioleta, os raios X e os raios gama representariam opções ainda melhores, por terem frequências mais altas, mas são difíceis de produzir e modular, além de não se propagarem bem através dos prédios e de serem perigosos para os seres vivos.

As bandas (ou faixas) de frequências listadas na parte inferior da Figura 2.8 são os nomes oficiais definidos pela ITU (International Telecommunication Union) e se baseiam nos comprimentos de onda; portanto, a banda LF vai de 1 a 10 km (aproximadamente, de 30 kHz a 300 kHz). Os termos LF, MF e HF são as abreviaturas, em inglês, de baixa, média e alta frequências, respectivamente. É claro que, quando esses nomes foram criados, ninguém esperava ultrapassar 10 MHz, de forma que foram atribuídos os seguintes nomes às bandas mais altas surgidas posteriormente: Very, Ultra, Super, Extremely e Tremendously High Frequency. Além desses não há outros nomes, mas Incredibly, Astonishingly e Prodigiously High Frequency (IHF, AHF e PHF) também serviriam muito bem. Acima de  $10^{12}$  Hz, entramos no espectro do infravermelho, onde a comparação normalmente é com a luz, e não com o rádio.

A base teórica para a comunicação, que veremos mais adiante neste capítulo, afirma que o volume de informações que um sinal como uma onda eletromagnética é capaz de transportar depende da potência recebida e é proporcional à sua largura de banda. Observando a Figura 2.8, é possível entender com clareza por que as pessoas ligadas a redes gostam tanto das fibras ópticas. Muitos GHz de largura de banda estão disponíveis para a transmissão de dados

na banda de micro-ondas, e ainda mais na fibra, pois está mais à direita em nossa escala logarítmica. Como exemplo, considere a banda de 1,30 micron da Figura 2.5, que tem uma largura de 0,17 micra. Se usarmos a Equação 2.1 para encontrar a faixa de frequência dos comprimentos de onda inicial e final, descobrimos que ela é de aproximadamente 30.000 GHz. Com uma razoável relação sinal/ruído de 10 dB, teremos 300 Tbps.

A maioria das transmissões utiliza uma banda de frequência relativamente estreita (ou seja,  $\Delta f/f \ll 1$ ). Elas concentram seus sinais nessa banda estreita para usar o espectro com mais eficiência e obter taxas de dados razoáveis transmitindo com potência suficiente. O restante desta seção descreve três tipos diferentes de transmissão que utilizam bandas de frequência mais largas.

### 2.2.2 Espectro de dispersão por salto de frequência

No **espectro por salto de frequência**, o transmissor salta de uma frequência para outra centenas de vezes por segundo. Essa técnica é muito usada em comunicações militares, pois dificulta a detecção das transmissões e é praticamente impossível obstruí-las. Ela também oferece boa resistência ao enfraquecimento devido a sinais tomando múltiplos caminhos da origem ao destino e interferindo após a recombinação. Também oferece resistência à interferência de banda estreita, porque o receptor não fica restrito a uma frequência, quando impossibilitada, por tempo suficiente para encerrar a comunicação. Essa robustez a torna útil para as partes mais sobrecarregadas do espectro, como as bandas ISM que descreveremos em breve. Essa técnica também é

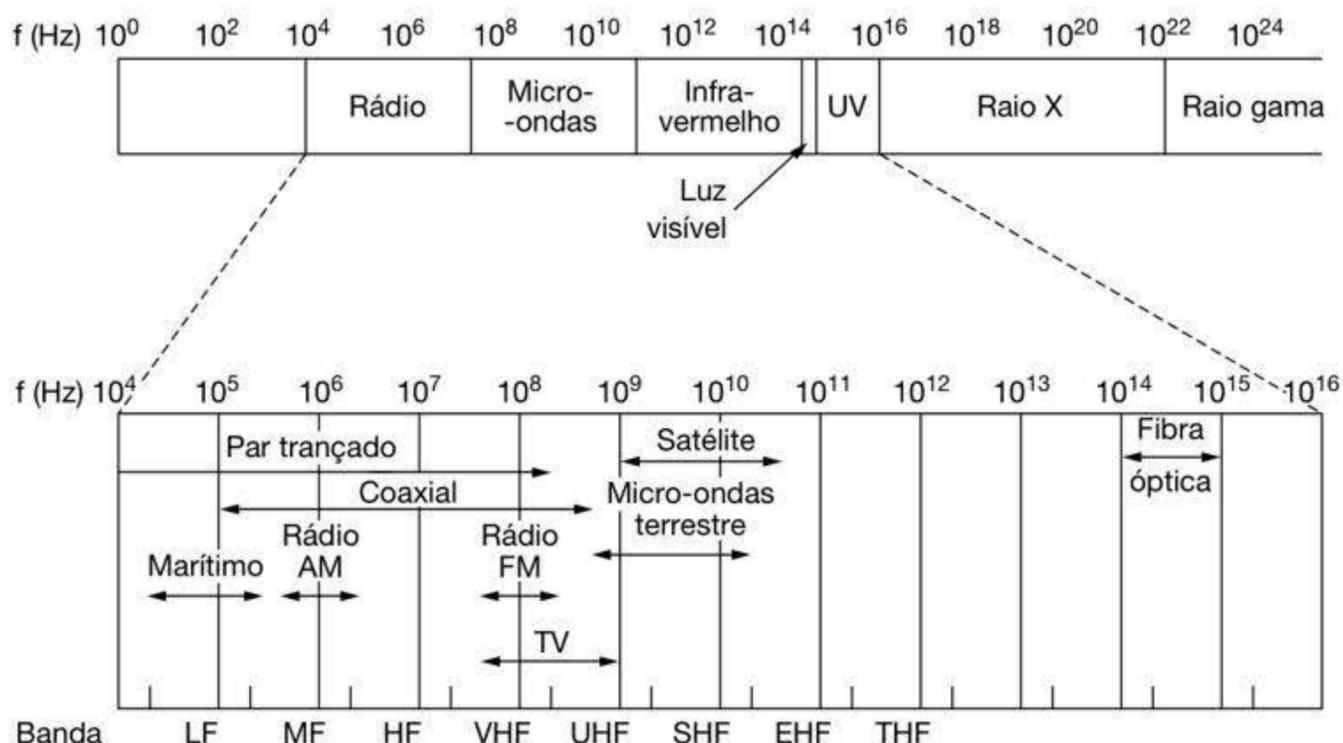


Figura 2.8 O espectro eletromagnético e a maneira como ele é usado na comunicação.

aplicada comercialmente, por exemplo, no Bluetooth e nas versões mais antigas das redes 802.11.

Como curiosidade, vale a pena mencionar que uma das pessoas que criaram essa técnica foi a atriz de cinema austríaca Hedy Lamarr, famosa por atuar em filmes europeus na década de 1930 com seu nome real Hedwig (Hedy) Kiesler. Seu primeiro marido era um rico fabricante de armamentos e mostrou a ela como era fácil bloquear os sinais de rádio então empregados para controlar torpedos. Quando descobriu que ele estava vendendo armas a Hitler, ela ficou horrorizada, se disfarçou de criada para escapar dele e fugiu para Hollywood, a fim de continuar sua carreira como atriz de cinema. Em seu tempo livre, Hedy inventou o salto de frequência para ajudar no esforço de guerra dos Aliados.

Seu esquema utilizava 88 frequências, o número de teclas (e frequências) do piano. Por sua invenção, ela e seu amigo, o compositor George Antheil, receberam a patente 2.292.387 dos Estados Unidos. Porém, eles não conseguiram convencer a Marinha americana de que sua invenção tinha alguma utilidade prática, e nunca receberam royalties por ela. Somente anos depois de expirar a patente, a técnica foi redescoberta e usada em dispositivos eletrônicos móveis, em vez de bloquear sinais para torpedos na época da guerra.

### 2.2.3 Espectro de dispersão de sequência direta

O **espectro de dispersão de sequência direta** usa uma sequência de código para dispersar o sinal de dados por uma banda de frequência mais ampla. É bastante usado comercialmente, como um meio eficiente na utilização do espectro, para permitir que vários sinais compartilhem a mesma banda de frequência. Esses sinais podem receber diferentes códigos, um método chamado acesso múltiplo por divisão de código, ao qual retornaremos mais adiante neste capítulo. Tal método aparece em contraste com o salto de frequência na Figura 2.9. Ele forma a base das redes de telefonia móvel 3G e também é usado no GPS (Global Positioning System). Mesmo sem códigos diferentes, o espectro de dispersão de sequência direta, assim como o espectro de dispersão por salto de frequência, pode tolerar interferência e enfraquecimento, pois apenas uma fração do sinal desejado é perdida. Ele é usado nesse papel nas versões mais

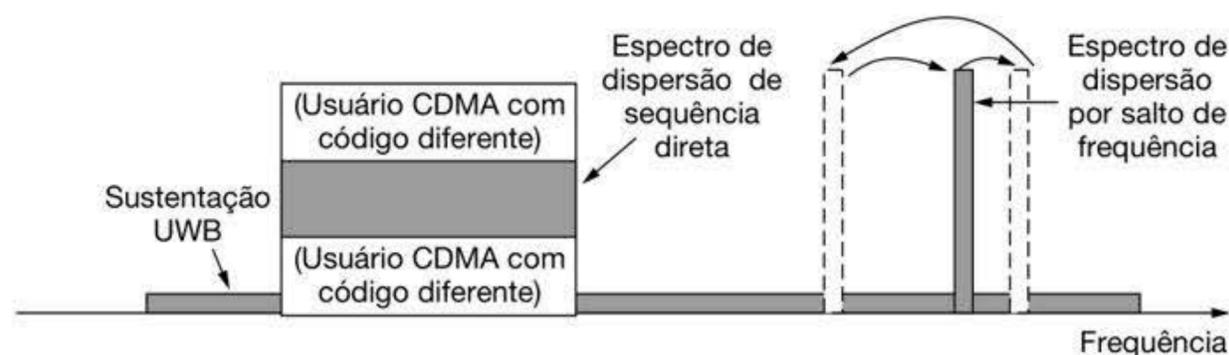
antigas do protocolo de LANs sem fio 802.11b. Para obter informações mais detalhadas e fascinantes sobre a história da comunicação por espectro de dispersão, consulte Walters (2013).

### 2.2.4 Comunicação de banda ultralarga

A comunicação de banda ultralarga, ou **UWB (Ultra-Wide-Band)** envia uma série de pulsos rápidos, de baixa energia, variando suas frequências de portadora para trocar informações. As rápidas transições levam a um sinal que se espalha estreitamente por uma banda de frequência muito larga. Esse método é definido como sinais que têm uma largura de banda de pelo menos 500 MHz ou pelo menos 20% da frequência central de sua banda de frequência. A UWB também aparece na Figura 2.9. Com tanta largura de banda, a UWB tem o potencial de se comunicar a várias centenas de megabits por segundo. Por se espalhar por uma ampla banda de frequência, ela pode tolerar uma grande quantidade de interferência relativamente forte de outros sinais de banda estreita. Tão importante quanto isso, como a UWB tem pouquíssima energia em qualquer frequência quando usada para transmissão em curta distância, ela não causa interferência prejudicial para outros sinais de rádio de banda estreita. Ao contrário da transmissão por espectro de dispersão, a comunicação UWB transmite de formas que não interferem com os sinais da portadora na mesma banda de frequência. Ela também pode ser usada para formar imagens através de objetos sólidos (solo, paredes e corpos), ou como parte de sistemas de localização precisos. A tecnologia é popular para aplicações internas a curta distância, além de imagens precisas de radar e tecnologias de rastreamento de local.

## 2.3 TRANSMISSÃO ATRAVÉS DO ESPECTRO

A seguir, mostraremos como as diversas partes do espectro eletromagnético da Figura 2.8 são usadas, começando pelo rádio. Partiremos da premissa de que todas as transmissões utilizam uma banda de frequência estreita, a menos que seja dito de outra forma.



**Figura 2.9** Comunicação por espectro de dispersão e banda ultralarga (UWB).

### 2.3.1 Transmissão de rádio

As ondas de rádio são fáceis de gerar, podem percorrer longas distâncias e penetrar facilmente em prédios, portanto, são muito utilizadas para comunicação, seja em ambientes fechados, seja em locais abertos. As ondas de rádio também são omnidirecionais, o que significa que elas viajam em todas as direções a partir da origem. Desse modo, o transmissor e o receptor não precisam estar cuidadosa e fisicamente alinhados.

Vale lembrar que o rádio omnidirecional nem sempre é bom. Na década de 1970, a General Motors decidiu equipar todos os seus novos Cadillacs com freios controlados por computador, que impediam o travamento das rodas. Quando o motorista pisava no pedal de freio, o computador prendia e soltava os freios, em vez de travá-los de verdade. Um belo dia, um policial rodoviário de Ohio começou a usar seu novo rádio móvel para falar com a central de polícia e, de repente, o Cadillac próximo a ele passou a se comportar como um cavalo selvagem. Depois de ser abordado pelo patrulheiro, o motorista disse que não tinha feito nada e que o carro tinha ficado louco de uma hora para outra.

Com o tempo, começou a surgir um padrão: às vezes, os Cadillacs enlouqueciam, mas somente quando trafegavam pelas principais estradas de Ohio, particularmente quando estavam sendo observados por um policial rodoviário. A General Motors demorou a entender o motivo pelo qual os Cadillacs funcionavam sem problemas nos outros estados e também em rodovias secundárias de Ohio. Só depois de muita pesquisa, eles descobriram que a fiação do Cadillac formava uma ótima antena que captava a frequência usada pelo novo sistema de rádio da Patrulha Rodoviária de Ohio.

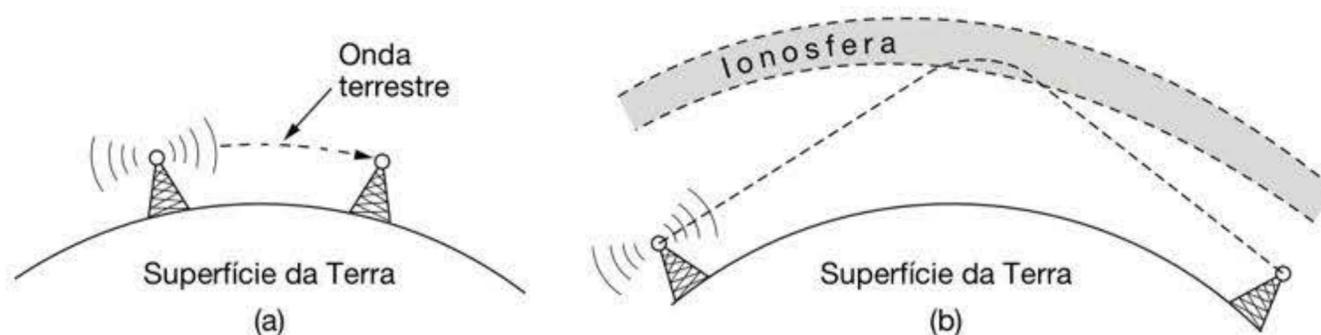
As propriedades das ondas de rádio dependem da frequência. Em baixas frequências, as ondas de rádio atravessam bem os obstáculos, mas a potência cai bruscamente à medida que a distância da origem aumenta – pelo menos cerca de  $1/r^2$  no ar –, pois a energia do sinal se espalha de forma mais estreita por uma superfície maior. Essa atenuação é chamada de **perda no caminho**. Em altas frequências, as ondas de rádio tendem a viajar em linha reta e a ricochetear nos obstáculos. A perda no caminho ainda reduz

a potência, embora o sinal recebido também possa depender muito das reflexões. Ondas de rádio de alta frequência também são absorvidas pela chuva e outros obstáculos, até certo ponto, mais do que as frequências baixas. Em todas as frequências, as ondas de rádio estão sujeitas à interferência de motores e outros equipamentos elétricos.

É interessante comparar a atenuação das ondas de rádio com a dos sinais nos meios guiados. Com fibra, cabo coaxial e par trançado, o sinal cai pela mesma fração por distância unitária, por exemplo, 20 dB por 100 m para o par trançado. Com o rádio, o sinal cai pela mesma fração enquanto a distância dobra, por exemplo, 6 dB pelo dobro do espaço livre. Isso significa que as ondas de rádio podem percorrer longas distâncias, e a interferência entre os usuários é um problema. Por essa razão, todos os governos exercem um rígido controle sobre o licenciamento do uso de transmissores de rádio, com apenas algumas exceções, descritas mais adiante neste capítulo.

Nas bandas VLF, LF e MF, as ondas de rádio se propagam perto do solo, como mostra a Figura 2.10(a). Elas podem ser detectadas dentro de um raio de, talvez, mil quilômetros nas frequências mais baixas; nas frequências mais altas, esse raio é menor. A radiodifusão em frequências AM utiliza a banda MF, razão pela qual as ondas de rádio produzidas pelas estações de rádio AM de Boston não podem ser captadas facilmente em Nova York. As ondas de rádio nessas bandas atravessam os prédios com facilidade; esse é o motivo por que os rádios portáteis funcionam em ambientes fechados. O principal problema relacionado à utilização dessas bandas para comunicação de dados diz respeito à baixa largura de banda que oferecem.

Nas bandas HF e VHF, as ondas terrestres tendem a ser absorvidas pela Terra. No entanto, as ondas que alcançam a ionosfera, uma camada de partículas carregadas situadas em torno da Terra a uma altura de 100 a 500 km, são refratadas por ela e enviadas de volta à Terra, como mostra a Figura 2.10(b). Em determinadas condições atmosféricas, os sinais podem ricochetear diversas vezes. Os operadores de radioamador utilizam essas bandas em comunicações de longa distância. Os militares também se comunicam nas bandas HF e VHF.



**Figura 2.10** (a) Nas bandas VLF, LF e MF, as ondas de rádio obedecem à curvatura da Terra. (b) Na banda HF, elas ricocheteariam na ionosfera.

### 2.3.2 Transmissão de micro-ondas

Acima de 100 MHz, as ondas trafegam praticamente em linha reta e, portanto, podem ser concentradas em uma faixa estreita. A concentração de toda a energia em um pequeno feixe através de uma antena parabólica (como a conhecida antena de TV por satélite) oferece uma relação sinal/ruído muito mais alta, mas as antenas de transmissão e recepção devem estar alinhadas com o máximo de precisão. Além disso, essa direcionalidade permite o alinhamento de vários transmissores em uma única fileira, fazendo eles se comunicarem com vários receptores também alinhados sem que haja interferência, desde que sejam observadas algumas regras mínimas de espaçamento. Antes da fibra óptica, durante décadas essas micro-ondas formaram o núcleo do sistema de transmissão telefônica de longa distância. Na verdade, a MCI, uma das primeiras concorrentes da AT&T após sua desregulamentação, construiu todo o seu sistema com comunicações de micro-ondas que percorriam dezenas de quilômetros entre uma torre e outra. Até mesmo o nome da empresa refletia isso (MCI significava Microwave Communications, Inc.). Há muito tempo, a MCI passou a utilizar as fibras ópticas e, por meio de uma longa série de fusões corporativas e falências no setor de telecomunicações, passou a fazer parte da Verizon.

Tendo em vista que as micro-ondas trafegam em linha reta, se as torres estiverem muito afastadas, a Terra acabará ficando entre elas (como acontece no caso de um enlace entre Seattle e Amsterdã). Assim, é preciso instalar repetidores em intervalos periódicos. Quanto mais altas as torres, mais distantes elas podem estar umas das outras. A distância entre os repetidores aumenta de acordo com a raiz quadrada da altura da torre. Torres com 100 m de altura devem ter repetidores a cada 80 km.

Ao contrário das ondas de rádio nas frequências mais baixas, as micro-ondas não atravessam muito bem as paredes dos prédios. Além disso, embora o feixe possa estar bem concentrado no transmissor, ainda há alguma divergência no espaço. Algumas ondas podem ser refratadas nas camadas atmosféricas mais baixas e, conseqüentemente, sua chegada pode ser mais demorada que a das ondas diretas. As ondas atrasadas podem chegar fora de fase em relação à onda direta, e assim cancelar o sinal. Esse efeito é chamado de **enfraquecimento por múltiplos caminhos (multipath fading)** e em geral é um problema sério. Ele depende das condições atmosféricas e da frequência. Algumas operadoras mantêm 10% de seus canais ociosos como sobressalentes, para serem usados quando o enfraquecimento por múltiplos caminhos eliminar temporariamente alguma banda de frequência.

A demanda por taxas de dados mais altas incentiva as operadoras a usarem transmissões com frequências cada vez mais altas. As bandas de até 10 GHz agora são de uso rotineiro, mas a partir de 4 GHz surge um novo problema: a absorção pela água. Essas ondas têm apenas alguns

centímetros de comprimento e são absorvidas pela chuva. Esse efeito não causaria qualquer problema se estivéssemos planejando construir um gigantesco forno de micro-ondas para ser usado a céu aberto mas, no caso das comunicações, trata-se de um grave problema. Assim como acontece com o enfraquecimento por múltiplos caminhos, a única solução é desligar os enlaces que estão sendo afetados pela chuva e criar uma nova rota que os contorne.

Em resumo, a comunicação por micro-ondas é tão utilizada na telefonia de longa distância, em telefones celulares, na distribuição de sinais de televisão e em outros usos que foi preciso desenvolver uma severa diminuição do espectro. Ela tem uma série de vantagens significativas sobre a fibra. A mais importante delas é que as micro-ondas dispensam a necessidade de se ter direitos sobre um percurso. Além do mais, quando se compra um pequeno lote de terra a cada 50 km e nele é instalada uma torre de micro-ondas, é possível ignorar o sistema telefônico e se comunicar diretamente. Foi por essa razão que a MCI passou a trabalhar tão rapidamente como uma companhia telefônica de longa distância. (A Sprint, outro antigo concorrente da AT&T, trilhou um caminho bem diferente: ela se formou a partir da Southern Pacific Railroad, que já detinha um grande número de concessões de direitos de percurso, e simplesmente enterrava os cabos de fibra ao lado das ferrovias.)

O uso de micro-ondas também é relativamente econômico. A instalação de duas torres simples (com alguns postes com quatro esteios) e a colocação de antenas em cada uma delas pode ser menos dispendiosa que enterrar 50 km de fibra em uma área urbana congestionada ou em uma região montanhosa, e talvez seja mais econômica que arrendar a rede de fibra da companhia telefônica, especialmente se esta ainda não tiver coberto totalmente os custos da retirada do cobre quando os cabos de fibra foram instalados.

### 2.3.3 Transmissão de infravermelho

As ondas de infravermelho não guiadas são bastante utilizadas na comunicação de curto alcance. Todos os dispositivos de controle remoto utilizados nos aparelhos de televisão, ar condicionado e aparelhos de som empregam a comunicação por infravermelho. Elas são relativamente direcionais, baratas e fáceis de montar, mas têm uma desvantagem importante: não atravessam objetos sólidos (para provar essa afirmação, posicione-se entre o controle remoto e o televisor e veja se ele funciona). Em geral, quando nos deslocamos do rádio de onda longa em direção à luz visível, as ondas assumem um comportamento cada vez mais parecido com o da luz, perdendo pouco a pouco as características de ondas de rádio.

No entanto, o fato de as ondas de infravermelho não atravessarem paredes sólidas pode ser visto como uma qualidade. É por essa razão que um sistema infravermelho

instalado em um ambiente fechado não interfere em um sistema semelhante instalado nas salas ou nos prédios adjacentes: não é possível controlar o aparelho de televisão do vizinho com o seu controle remoto. Além disso, a segurança dos sistemas de infravermelho contra espionagem é melhor que a dos sistemas de rádio, exatamente por essa razão. Portanto, não é necessária nenhuma licença do governo para operar um sistema de infravermelho, ao contrário dos sistemas de rádio, que devem ser licenciados fora das bandas ISM. A comunicação por infravermelho tem uso limitado em escritórios, por exemplo, para conectar notebooks e impressoras com o padrão **IrDA (Infrared Data Association)**, mas não tem um papel importante no jogo das comunicações.

### 2.3.4 Transmissão via luz

A transmissão óptica não guiada, ou **óptica do espaço livre**, vem sendo utilizada há séculos. Paul Revere usou a sinalização óptica binária a partir da Old North Church imediatamente antes de sua famosa corrida. Uma aplicação mais moderna consiste em conectar as LANs em dois prédios por meio de lasers instalados em seus telhados. Por sua própria natureza, a transmissão óptica usando raios laser é unidirecional; assim, cada prédio precisa de seu próprio raio laser e de seu próprio fotodetector. Esse esquema oferece uma largura de banda muito alta, tem baixo custo e é relativamente seguro, pois é difícil interceptar um raio laser estreito. Ele também é relativamente fácil de ser instalado e, ao contrário das micro-ondas, não precisa de uma licença da **FCC (Federal Communications Commission)** nos Estados Unidos e de outras agências governamentais em outros países.

A potência do laser, um feixe muito estreito, também é uma desvantagem aqui. Apontar um raio laser de 1 mm para um alvo do tamanho de uma cabeça de alfinete a 500 m de distância exige a mira de um pistoleiro dos filmes de faroeste. Normalmente, são colocadas lentes no sistema para tirar um pouco do foco do raio. Para aumentar a dificuldade, mudanças no vento e na temperatura podem distorcer o raio, e os feixes de raios laser de fato não podem atravessar chuva ou neblina espessa, mas normalmente funcionam bem em dias ensolarados. Contudo, muitos desses fatores não são problemas quando o uso é para conectar duas naves espaciais.

Um dos autores deste livro (AST) certa vez participou de uma conferência em um moderno hotel europeu cujos organizadores tiveram a felicidade de oferecer uma sala repleta de terminais para que os participantes pudessem ler seus e-mails durante as apresentações menos interessantes. Como a companhia telefônica local não havia se disposto a instalar um grande número de linhas telefônicas por apenas três dias, os organizadores colocaram um raio laser no telhado e o apontaram na direção do prédio de ciência da

computação da universidade onde trabalhavam, situada a alguns quilômetros dali. Eles testaram o sistema na noite anterior à conferência e ele havia funcionado perfeitamente. Às 9h da manhã seguinte, em um belo dia de sol, o sistema entrou em pane e ficou fora do ar durante todo o dia. Nos dois dias seguintes, o problema se repetiu. Após a conferência, os organizadores descobriram o problema. O calor do sol fez as correntes de convecção do telhado do prédio emanarem, como mostra a Figura 2.11. Esse ar turbulento desviou o feixe e fez ele dançar em torno do detector, algo parecido com uma estrada que parece cintilar em um dia muito quente. A lição aqui é que, para funcionar bem em condições difíceis e também em boas condições, os enlaces ópticos não guiados precisam ser elaborados com uma margem de erro suficiente.

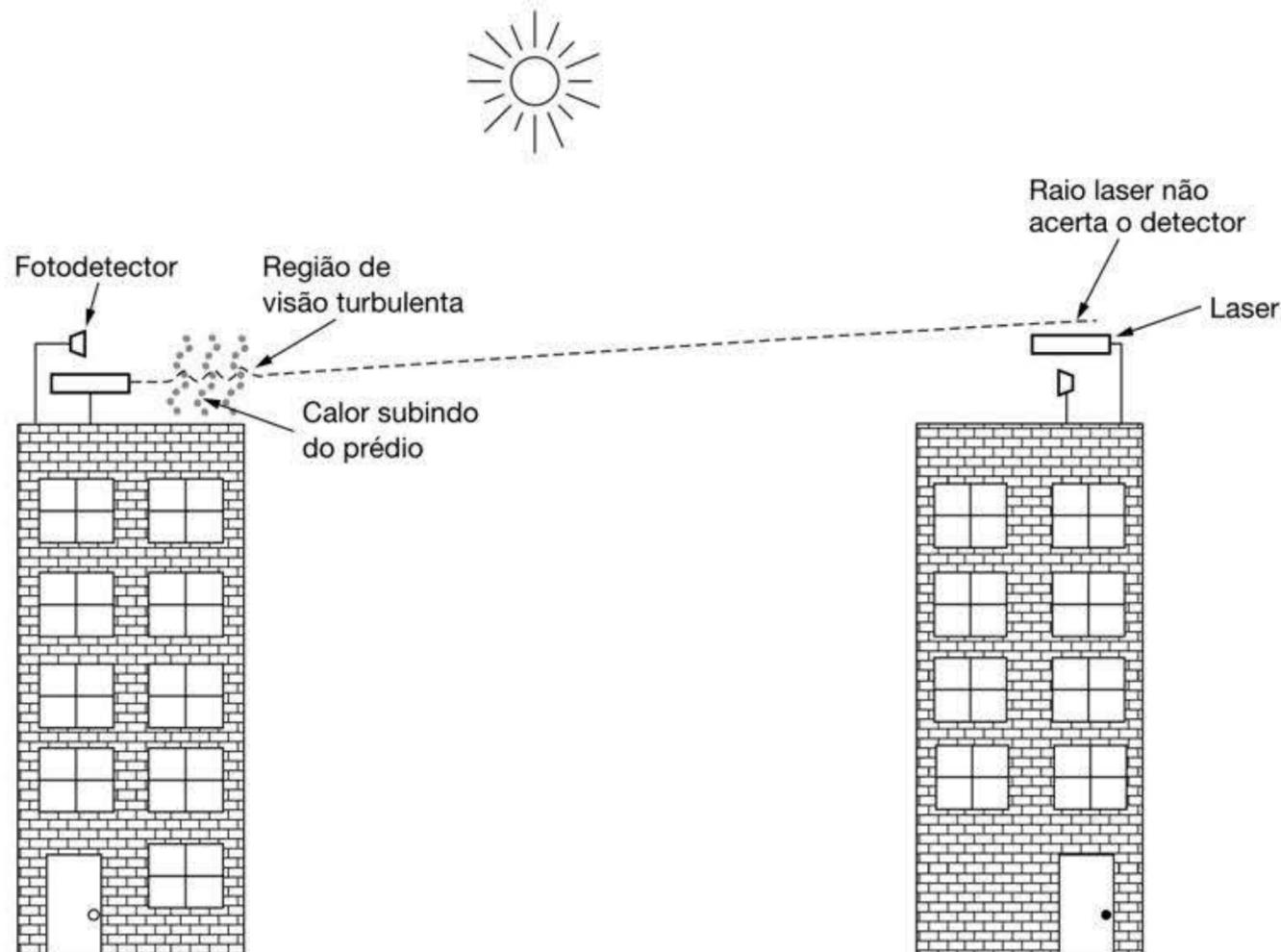
Atualmente, a comunicação óptica não guiada pode parecer uma tecnologia de rede exótica, mas logo poderá se tornar mais predominante. Em muitos lugares, estamos cercados de câmeras (que sentem a luz) e telas (que emitem luz usando LEDs ou outra tecnologia). A comunicação de dados pode ser disposta em cima dessas telas codificando informações no padrão em que os LEDs se acendem e apagam, abaixo do limiar da percepção humana. A comunicação via luz visível dessa maneira é inerentemente segura e cria uma rede de baixa velocidade na vizinhança imediata da tela. Isso poderia permitir todo o tipo de cenário de computação ubíqua. As luzes piscando nos veículos de emergência podem alertar os sinais de trânsito e veículos mais próximos para ajudar a liberar um caminho. Os sinais de informação poderiam transmitir mapas por difusão. As lâmpadas em ocasiões festivas poderiam até mesmo transmitir canções sincronizadas com sua exibição.

## 2.4 DE FORMAS DE ONDA A BITS

Nesta seção, descrevemos como os sinais são transmitidos através do meio físico que já discutimos. Vamos começar com uma discussão da base teórica para a comunicação de dados, seguindo com uma discussão sobre modulação (o processo de conversão de formas analógicas de onda em bits) e multiplexação (que permite que um único meio físico carregue várias transmissões simultâneas).

### 2.4.1 A base teórica da comunicação de dados

As informações podem ser transmitidas por fios, fazendo-se variar alguma propriedade física, como tensão ou corrente. Representando o valor dessa tensão ou corrente como uma função do tempo com um valor único,  $f(t)$ , podemos criar um modelo para o comportamento do sinal e



**Figura 2.11** Correntes de convecção podem interferir nos sistemas de comunicação a laser. A figura mostra um sistema bidirecional com dois lasers.

analisá-lo matematicamente. Essa análise será o assunto das próximas seções.

### Análise de Fourier

No início do século XIX, o matemático francês Jean-Baptiste Fourier provou que qualquer função periódica razoavelmente estável,  $g(t)$ , com período  $T$ , pode ser construída como a soma de um número (possivelmente infinito) de senos e cossenos:

$$g(t) = \frac{1}{2}c + \sum_{n=1}^{\infty} a_n \text{sen}(2\pi nft) + \sum_{n=1}^{\infty} b_n \text{cos}(2\pi nft) \quad (2.2)$$

onde  $f = 1/T$  é a frequência fundamental,  $a_n$  e  $b_n$  são as amplitudes do seno e do cosseno dos  $n$ -ésimos **harmônicos** (termos) e  $c$  é uma constante que determina o valor médio da função. Essa decomposição é chamada **série de Fourier**. A partir dela, a função pode ser reconstruída, ou seja, se o período  $T$  for conhecido e as amplitudes forem dadas, a função original no tempo poderá ser encontrada efetuando-se as somas da Equação 2.2.

Um sinal de dados com uma duração finita (como acontece com todos eles) pode ser tratado apenas com base na premissa de que ele repete o mesmo padrão indefinidamente (ou seja, o intervalo de  $T$  a  $2T$  é igual ao de  $0$  a  $T$ , etc.).

As  $a_n$  amplitudes podem ser calculadas para qualquer  $g(t)$  dada, multiplicando-se ambos os lados da Equação 2.2 por  $\text{sen}(2\pi kft)$  e, em seguida, integrando-se de  $0$  a  $T$ . Como

$$\int_0^T \text{sen}(2\pi kft) \text{sen}(2\pi nft) dt = \begin{cases} 0 & \text{para } k \neq n \\ T/2 & \text{para } k = n \end{cases}$$

apenas um termo do somatório permanece:  $a_n$ . O somatório  $b_n$  desaparece completamente. Da mesma forma, multiplicando a Equação 2.2 por  $\text{cos}(2\pi kft)$  e integrando entre  $0$  e  $T$ , podemos derivar  $b_n$ . Simplesmente integrando ambos os lados da equação tal como ela se encontra, podemos achar  $c$ . Os resultados dessas operações são:

$$a_n = \frac{2}{T} \int_0^T g(t) \text{sen}(2\pi nft) dt$$

$$b_n = \frac{2}{T} \int_0^T g(t) \text{cos}(2\pi nft) dt \quad c = \frac{2}{T} \int_0^T g(t) dt$$

### Sinais limitados pela largura de banda

A importância de tudo isso para a comunicação de dados é que os canais reais afetam diferentes sinais de frequência de formas distintas. Vamos analisar um exemplo específico: a transmissão do caractere ASCII “b” codificado como um

byte de 8 bits. O padrão de bits que deve ser transmitido é 01100010. A parte esquerda da Figura 2.12(a) mostra a saída de tensão do computador transmissor. A análise de Fourier desse sinal produz os seguintes coeficientes:

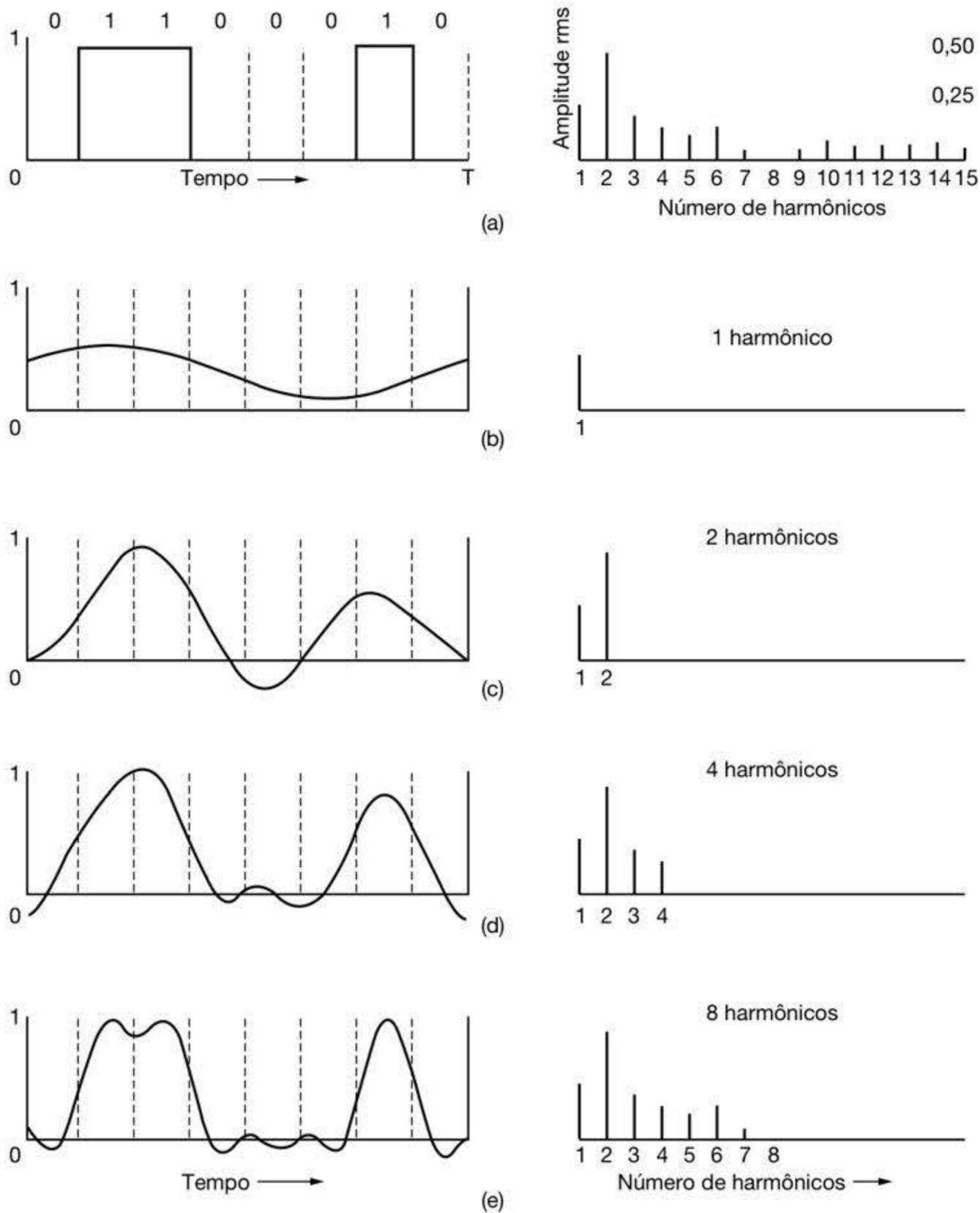
$$a_n = \frac{1}{\pi n} [\cos(\pi n/4) - \cos(3\pi n/4) + \cos(6\pi n/4) - \cos(7\pi n/4)]$$

$$b_n = \frac{1}{\pi n} [\sin(3\pi n/4) - \sin(\pi n/4) + \sin(7\pi n/4) - \sin(6\pi n/4)]$$

$$c = 3/4.$$

A raiz quadrada média das amplitudes,  $\sqrt{a_n^2 + b_n^2}$  para os primeiros termos são mostradas no lado direito da Figura 2.12(a). Esses valores são importantes porque seus quadrados são proporcionais à energia transmitida na frequência correspondente.

Nenhum recurso de transmissão é capaz de transmitir sinais sem perder parte da energia no processo. Se todos os coeficientes da série de Fourier fossem igualmente reduzidos, o sinal resultante seria reduzido em amplitude, mas não distorcido (i.e., ele teria a mesma forma quadrada mostrada na Figura 2.12(a)). Infelizmente, todos os meios de transmissão reduzem diferentes componentes de Fourier



**Figura 2.12** (a) Um sinal binário e a raiz quadrada média das amplitudes de Fourier. (b)-(e) Aproximações sucessivas do sinal original.

por diferentes valores e, em consequência disso, introduzem distorção. Em geral, para um fio, as amplitudes são transmitidas sem redução, de 0 até alguma frequência  $f_c$  (medida em ciclos/s ou Hz), com todas as frequências acima dessa frequência de corte (*cutoff*) sendo atenuadas. A faixa de frequências transmitidas sem serem fortemente atenuadas denomina-se **largura de banda**. Na prática, o corte não é nítido; assim, muitas vezes a largura de banda varia de 0 até a frequência em que a potência recebida caiu para a metade.

A largura de banda é uma propriedade física do meio de transmissão, que depende, por exemplo, da construção, da espessura e do comprimento do meio (fio ou fibra). Os filtros normalmente são usados para limitar ainda mais o volume de largura de banda de um sinal. Os canais sem fio 802.11 têm permissão para usar até aproximadamente 20 MHz, por exemplo, de modo que rádios 802.11 filtram a largura de banda do sinal para essa faixa (embora, em alguns casos, seja usada uma banda de 80 MHz).

Como outro exemplo, os canais de televisão tradicionais (analógicos) ocupam 6 MHz cada, por um fio ou pelo ar. Essa filtragem permite que mais sinais compartilhem determinada região do espectro, o que melhora a eficiência geral do sistema. Isso significa que a faixa de frequência para alguns sinais não começará em zero, mas em algum número maior. Porém, isso não importa. A largura de banda ainda é a largura da banda de frequências que são passadas, e a informação que pode ser transportada depende apenas dessa largura, e não das frequências inicial e final. Os sinais que vão de 0 para cima, até uma frequência máxima, são chamados sinais de **banda base**. Sinais que são deslocados para ocupar uma faixa de frequências mais alta, como acontece para todas as transmissões sem fio, são chamados sinais de **banda passante**.

Veamos agora como seria a forma do sinal da Figura 2.12(a) se a largura de banda fosse tão estreita que apenas as frequências mais baixas fossem transmitidas (ou seja, se a função estivesse sendo aproximada pelos primeiros termos da Equação 2.2). A Figura 2.12(b) mostra o sinal resultante de um canal pelo qual apenas o primeiro harmônico (o fundamental,  $f$ ) pode passar. Da mesma forma, a Figura 2.12(c)-(e) mostra os espectros e as funções reconstruídas para canais com uma largura de banda mais alta. Para a transmissão digital, o objetivo é receber um sinal com fidelidade suficiente para reconstruir a sequência de bits que foi enviada. Já podemos fazer isso com facilidade na Figura 2.12(e), de modo que é um desperdício usar mais harmônicos para receber uma réplica mais precisa.

Dada uma taxa de bits de  $b$  bits/s, o tempo necessário para o envio de 8 bits (em nosso exemplo, um bit de cada vez) é  $8/b$  s; assim, a frequência do primeiro harmônico desse sinal é  $b/8$  Hz. Uma linha telefônica comum, frequentemente chamada de **linha de qualidade de voz**, tem

uma frequência de corte artificialmente introduzida, pouco acima de 3.000 Hz. A presença dessa restrição significa que o número do harmônico mais alto transmitido é aproximadamente  $3.000/(b/8)$  ou  $24.000/b$  (o corte não é preciso).

Para algumas taxas de dados, os números funcionam de acordo com o padrão mostrado na Figura 2.13. Esses números deixam claro que, quando se tenta fazer uma transmissão a 9.600 bps usando uma linha telefônica de qualidade de voz, o modelo sugerido na Figura 2.12(a) assume a forma mostrada na Figura 2.12(c), tornando complicada a recepção precisa do fluxo original de bits. Podemos perceber também que, em taxas de dados bem mais altas que 38,4 kbps, não existe a menor possibilidade de que todos os sinais sejam *binários*, mesmo quando não há o menor ruído no meio de transmissão. Em outras palavras, limitando-se a largura de banda, limita-se a taxa de dados, mesmo nos canais perfeitos (sem ruídos). No entanto, sofisticados esquemas de codificação que usam diversos níveis de tensão possibilitam a existência e a utilização de taxas de dados mais altas. Vamos discutir essa questão ao longo deste capítulo.

Há muita confusão sobre largura de banda, pois ela significa coisas diferentes para engenheiros elétricos e para cientistas da computação. Para os engenheiros elétricos, a largura de banda (analógica) é, como descrevemos, uma quantidade medida em Hz. Para os cientistas da computação, a largura de banda (digital) é a taxa de dados máxima de um canal, uma quantidade medida em bits/s. Essa taxa de dados é o resultado final do uso da largura de banda analógica de um canal físico para a transmissão digital, e as duas estão relacionadas, conforme veremos a seguir. Neste livro, o contexto esclarecerá se queremos dizer largura de banda analógica (Hz) ou largura de banda digital (bits/s).

Bps	T (ms)	Primeiro harmônico (Hz)	Número de harmônicos enviados
300	26,67	37,5	80
600	13,33	75	40
1.200	6,67	150	20
2.400	3,33	300	10
4.800	1,67	600	5
9.600	0,83	1.200	2
19.200	0,42	2.400	1
38.400	0,21	4.800	0

**Figura 2.13** Relação entre as taxas de dados e os harmônicos em nosso exemplo muito simples.

## 2.4.2 A taxa de dados máxima de um canal

Em 1924, Henry Nyquist, um engenheiro da AT&T, percebeu que até mesmo um canal perfeito tem uma capacidade de transmissão finita. Ele derivou uma equação expressando a taxa máxima de dados de um canal sem ruído com largura de banda finita. Em 1948, Claude Shannon aprofundou o trabalho de Nyquist e o estendeu ao caso de um canal sujeito a ruído aleatório (i.e., ruído termodinâmico) (Shannon, 1948). Esse artigo é o mais importante de toda a teoria da informação. Veja a seguir um resumo dos resultados, agora clássicos.

Nyquist provou que, se um sinal arbitrário atravessar um filtro passa-baixa de largura de banda  $B$ , o sinal filtrado poderá ser completamente reconstruído a partir de apenas  $2B$  amostras (exatas) por segundo. Fazer uma amostragem da linha com uma rapidez maior que  $2B$  vezes por segundo seria inútil, pois os componentes de frequência mais alta que essa amostragem poderia recuperar já teriam sido filtrados. Se o sinal consistir em  $V$  níveis discretos, o teorema de Nyquist afirma que:

$$\text{taxa máxima de dados} = 2B \log_2 V \text{ bits/s} \quad (2.3)$$

Por exemplo, um canal de 3 kHz sem ruído não pode transmitir sinais binários (ou seja, de dois níveis) a uma taxa superior a 6.000 bps.

Até agora, só mencionamos os canais sem ruído. Se houver ruído aleatório, a situação se deteriorará com rapidez. Além disso, sempre há ruído aleatório (térmico), em virtude do movimento das moléculas no sistema. O volume de ruído térmico presente é medido pela relação entre a potência do sinal e a do ruído, chamada relação sinal/ruído, ou **SNR (Signal-to-Noise Ratio)**. Se representarmos a potência do sinal por  $S$  e a potência do ruído por  $N$ , a relação sinal/ruído será  $S/N$ . Em geral, a relação é expressa em uma escala logarítmica como a quantidade  $10 \log_{10} S/N$ , pois ela pode variar por uma faixa muito grande. As unidades da escala logarítmica são chamadas de **decibéis (dB)**, com “deci” significando 10 e “bel” escolhido em homenagem a Alexander Graham Bell, que patenteou o telefone. Uma relação  $S/N$  igual a 10 corresponde a 10 dB, uma relação igual a 100 equivale a 20 dB, uma relação igual a 1.000 equivale a 30 dB, e assim por diante. Com frequência, os fabricantes de amplificadores estereofônicos caracterizam a largura de banda (faixa de frequência) na qual seu produto é linear oferecendo a frequência de 3 dB em cada extremidade. Esses são os pontos em que o fator de amplificação foi dividido aproximadamente ao meio (porque  $10 \log_{10} 0,5 \approx -3$ ).

O principal resultado de Shannon é que a taxa máxima de dados ou **capacidade** de um canal com ruídos cuja largura de banda é  $B$  Hz, e cuja relação sinal/ruído é  $S/N$ , é dada por:

$$\text{Taxa máxima de dados} = B \log_2 (1 + S/N) \text{ bits/s} \quad (2.4)$$

Esta equação nos diz as melhores capacidades que os canais reais podem ter. Por exemplo, ADSL (Asymmetric Digital Subscriber Line), que oferece acesso à Internet por linhas telefônicas normais, usa uma largura de banda com cerca de 1 MHz. A SNR depende muito da distância da central telefônica até a casa, e uma SNR com cerca de 40 dB para linhas curtas de 1 a 2 km é muito boa. Com essas características, o canal nunca pode transmitir muito mais do que 13 Mbps, não importa quantos níveis de sinal sejam usados e não importa com que frequência as amostras são tomadas. Na prática, a ADSL é especificada para até 12 Mbps, embora os usuários normalmente vejam taxas mais baixas. Essa taxa de dados é realmente muito boa, com mais de 60 anos de técnicas de comunicação tendo reduzido bastante a lacuna entre a capacidade de Shannon e a capacidade dos sistemas reais.

O resultado de Shannon utilizou os argumentos da teoria da informação e se aplica a qualquer canal sujeito a ruído térmico. Os exemplos contrários devem ser tratados na mesma categoria das máquinas de movimento contínuo (ou moto-perpétuo). Para a ADSL ultrapassar os 12 Mbps, ela deve melhorar a SNR (p. ex., inserindo repetidores digitais nas linhas mais próximas do cliente) ou usar mais largura de banda, como é feito com a evolução para ADSL2+.

## 2.4.3 Modulação digital

Agora que já estudamos as propriedades dos canais com e sem fio, voltamos nossa atenção para o problema de enviar informações digitais. Os canais com fio e sem fio transportam sinais analógicos, como a tensão variando continuamente, a intensidade de luz ou a intensidade de som. Para enviar informações digitais, temos de criar sinais digitais para representar os bits. O processo de conversão entre bits e sinais que os representam é chamado de **modulação digital**.

Começaremos com esquemas que convertem diretamente os bits em um sinal. Eles resultam em **transmissão de banda base**, em que o sinal ocupa frequências de zero até um máximo que depende da taxa de sinalização. Ele é comum para fios. Depois, vamos considerar os esquemas que regulam a amplitude, a fase ou a frequência de um sinal da portadora para transportar bits. Esses esquemas resultam em **transmissão de banda passante**, em que o sinal ocupa uma banda de frequências em torno da frequência do sinal da portadora. Isso é comum para canais sem fio e ópticos, para os quais os sinais devem residir em uma determinada banda de frequência.

Os canais normalmente são compartilhados por vários sinais. Afinal, é muito mais conveniente usar um único fio para transportar vários sinais do que instalar um fio para cada sinal. Esse tipo de compartilhamento é chamado de **multiplexação** e pode ser realizado de diversas maneiras.

Apresentaremos métodos para multiplexação por divisão de tempo, frequência e código.

As técnicas de modulação e multiplexação que descrevemos nesta seção são todas bastante usadas para canais com fio, de fibra, terrestres sem fio e por satélite.

### Transmissão de banda base

A forma mais simples de modulação digital é usar uma tensão positiva para representar um bit 1 e uma tensão negativa para representar um bit 0, como pode ser visto na Figura 2.14(a). Para uma fibra óptica, a presença de luz poderia representar 1 e a ausência de luz, 0. Esse esquema é chamado **NRZ (Non-Return-to-Zero)**. O nome esquisito tem motivos históricos, e significa simplesmente que o sinal acompanha os dados. Um exemplo aparece na Figura 2.14(b).

Uma vez enviado, o sinal NRZ se propaga pelo fio. Na outra ponta, o receptor o converte para bits fazendo a amostragem do sinal em intervalos de tempo regulares. Esse sinal não ficará exatamente como o enviado, e será atenuado e distorcido pelo canal e pelo ruído no receptor. Para decodificar os bits, o receptor mapeia as amostras de sinal para os símbolos mais próximos. Para o NRZ, uma tensão positiva será considerada para indicar que foi enviado um bit 1, e uma tensão negativa será considerada para indicar que foi enviado um 0.

O NRZ é um bom ponto de partida para nossos estudos, pois é simples, mas na prática raramente é usado por si só. Esquemas mais complexos podem converter bits em sinais, os quais atendem melhor às considerações da engenharia. Esses esquemas são chamados de **códigos de linha**. A seguir, descreveremos os códigos de linha que ajudam na eficiência da largura de banda, recuperação de clock e equilíbrio na componente CC.

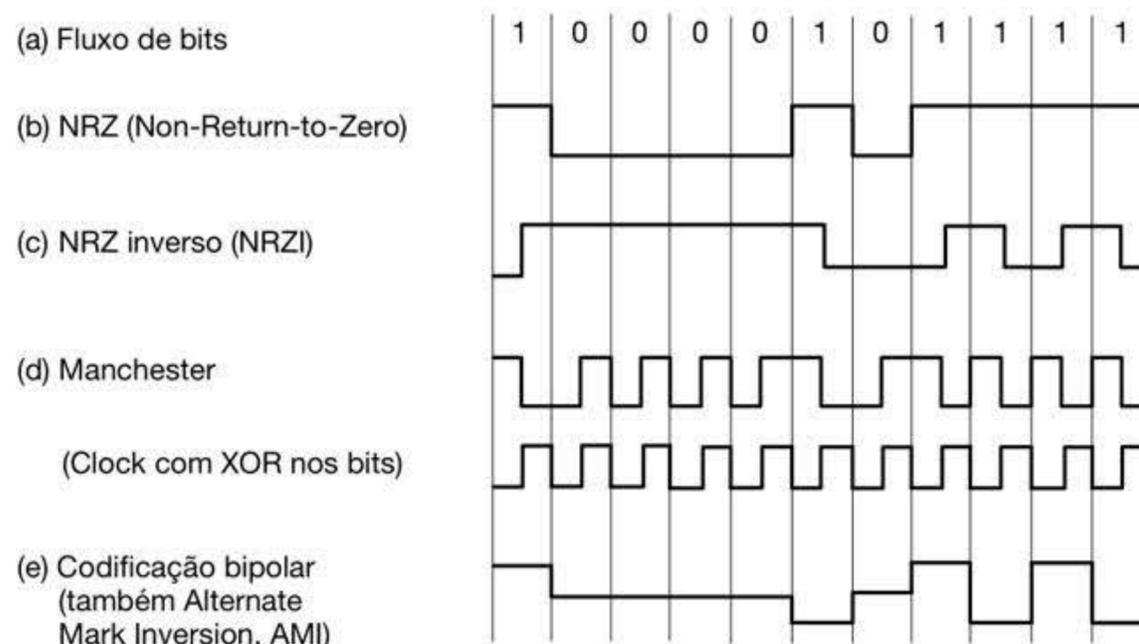
### Eficiência da largura de banda

Com o NRZ, o sinal pode alternar entre os níveis positivo e negativo a cada 2 bits (no caso de 1s e 0s alternados). Isso significa que precisamos de uma largura de banda de pelo menos  $B/2$  Hz quando a taxa de bits é  $B$  bits/s. Essa relação vem da taxa de Nyquist (Equação 2.3). Por ser um limite fundamental, não podemos usar o NRZ mais rápido sem usar mais largura de banda, sendo normalmente um recurso limitado, até mesmo para canais com fio. Sinais de maior frequência são cada vez mais atenuados, tornando-os menos úteis, e sinais de maior frequência também exigem circuitos eletrônicos mais rápidos.

Uma estratégia para usar largura de banda limitada com mais eficiência é usar mais de dois níveis de sinalização. Usando quatro níveis de tensão, por exemplo, podemos enviar 2 bits ao mesmo tempo como um único **símbolo**. Esse projeto funcionará desde que o sinal no receptor seja suficientemente forte para distinguir os quatro níveis. A taxa em que o sinal muda é, então, metade da taxa de bits, de modo que a largura de banda necessária foi reduzida.

Chamamos a taxa em que o sinal muda de **taxa de símbolos**, para distingui-la da **taxa de bits** (que é a taxa de símbolos multiplicada pelo número de bits por símbolo). Um nome mais antigo para a taxa de símbolos, principalmente no contexto de dispositivos chamados modems de telefone, que transmitem dados digitais pelas linhas telefônicas, é **taxa baud**. Na literatura, “taxa de bits” e “taxa baud” normalmente são usadas de modo incorreto.

Observe que o número de níveis de sinal não precisa ser uma potência de dois. Normalmente ele não é, com alguns dos níveis usados para proteção contra erros e simplificação do projeto do receptor.



**Figura 2.14** Códigos de linha: (a) Bits, (b) NRZ, (c) NRZI, (d) Manchester, (e) Bipolar ou AMI.

### Recuperação de clock

Para todos os esquemas que codificam bits em símbolos, o receptor precisa saber quando um símbolo termina e o próximo começa, para decodificar os bits corretamente. Com o NRZ, em que os símbolos são simplesmente níveis de tensão, uma longa sequência de 0s ou 1s deixa o sinal inalterado. Após um tempo, é difícil distinguir os bits, pois 15 zeros se parecem muito com 16 zeros, a menos que você tenha um clock muito preciso.

Os clocks precisos ajudariam com esse problema, mas eles são uma solução dispendiosa para equipamentos básicos. Lembre-se de que estamos temporizando bits em links que trabalham em muitos megabits/s, de modo que o clock teria de fluir por menos de uma fração de microssegundo pela maior sequência permitida. Isso pode ser razoável para enlaces lentos ou mensagens curtas, mas essa não é uma solução geral.

Uma estratégia é enviar um sinal de clock separado para o receptor. Outra linha de clock não é grande coisa para barramentos de computador ou cabos curtos, em que existem muitas linhas paralelas, mas é um desperdício para a maioria dos enlaces de rede, pois, se tivéssemos outra linha para enviar um sinal, poderíamos usá-la para enviar dados. Aqui, um truque inteligente é misturar o sinal de clock com o sinal de dados efetuando a operação XOR por ambos, de modo que nenhuma linha extra seja necessária. Os resultados aparecem na Figura 2.14(d). O clock faz uma transição de clock em cada tempo de bit, de modo que ele funciona no dobro da taxa de bits. Quando ele é XORado com o nível 0, faz a transição de baixo para alto, que é simplesmente o clock. Essa transição é o 0 lógico. Quando ele é XORado com o nível 1, é invertido e faz uma transição de alto para baixo. Essa transição é o 1 lógico. Esse esquema é conhecido como **codificação Manchester**, e foi usado para a Ethernet clássica.

A desvantagem da codificação Manchester é que ela exige duas vezes mais largura de banda que a NRZ, em virtude do clock, e aprendemos que a largura de banda normalmente importa. Uma estratégia diferente é baseada na ideia de que devemos codificar os dados para garantir que haja transições suficientes no sinal. Considere que o NRZ terá problemas de recuperação de clock somente para longas sequências de 0s e 1s. Se houver transições frequentes, será fácil para o receptor permanecer sincronizado com o fluxo de símbolos de chegada.

Como um passo na direção certa, podemos simplificar a situação codificando um 1 como uma transição e 0 como nenhuma transição, ou vice-versa. Essa codificação é chamada **NRZI (Non-Return-to-Zero Inverted)**, uma variação da NRZ. Um exemplo pode ser visto na Figura 2.14(c). O popular padrão **USB (Universal Serial Bus)** para a conexão de periféricos de computador usa NRZI. Com ele, longas sequências de 1s não causam problema.

Naturalmente, longas sequências de 0s ainda causam um problema que precisamos consertar. Se fôssemos a companhia telefônica, poderíamos simplesmente solicitar que o transmissor não transmitisse muitos 0s. Para que funcionassem corretamente, as linhas telefônicas digitais mais antigas nos Estados Unidos, chamadas **linhas T1** (discutidas mais adiante), realmente exigiam que não mais do que 15 zeros consecutivos fossem enviados. Para resolver realmente o problema, podemos dividir sequências de 0s mapeando pequenos grupos de bits para serem transmitidos de modo que os grupos com 0s sucessivos sejam mapeados para padrões ligeiramente maiores, que não têm muitos 0s consecutivos.

Um código bem conhecido para fazer isso é chamado **4B/5B**. Cada 4 bits são mapeados para um padrão de 5 bits com uma tabela de tradução fixa. Os padrões de 5 bits são escolhidos de tal forma que nunca haverá uma sequência de mais de três 0s consecutivos. O mapeamento aparece na Figura 2.15. Esse esquema acrescenta 20% de overhead, que é melhor do que os 100% de overhead da codificação Manchester. Como existem 16 combinações de entrada e 32 combinações de saída, algumas das combinações de saída não são usadas. Deixando de lado as combinações com muitos 0s sucessivos, ainda existirão alguns códigos. Como um bônus, podemos usar esses códigos não de dados para representar sinais de controle da camada física. Por exemplo, em alguns casos, “11111” representa a linha ociosa, e “11000” representa o início de um quadro.

Uma alternativa é fazer os dados parecerem aleatórios, o que é conhecido como embaralhamento (ou scrambling). Nesse caso, provavelmente haverá transições frequentes. Um **embaralhador** funciona realizando o XOR dos dados com uma sequência pseudoaleatória antes de serem transmitidos. Esse embaralhamento tornará os dados tão aleatórios quanto a sequência pseudoaleatória (supondo que eles sejam independentes da sequência pseudoaleatória). O receptor, então, realiza o XOR dos bits de entrada com

Dados (4B)	Código (5B)	Dados (4B)	Código (5B)
0000	11110	1000	10010
0001	01001	1001	10011
0010	10100	1010	10110
0011	10101	1011	10111
0100	01010	1100	11010
0101	01011	1101	11011
0110	01110	1110	11100
0111	01111	1111	11101

**Figura 2.15** Mapeamento 4B/5B.

a mesma sequência pseudoaleatória para registrar os dados reais. Para que isso seja prático, a sequência pseudoaleatória precisa ser fácil de criar. Em geral, ela é dada como uma semente para um gerador de números aleatórios simples.

O embaralhamento é atraente porque não acrescenta largura de banda ou tempo. De fato, normalmente é importante condicionar o sinal de modo que ele não mantenha energia em componentes de frequência dominantes (causados por padrões de dados repetitivos) que poderiam radiar interferência eletromagnética. O embaralhamento ajuda porque os sinais aleatórios tendem a ser “brancos” ou ter a energia espalhada pelos demais componentes de frequência.

Contudo, ele não garante que não haverá sequências longas. Ocasionalmente, é possível não ter sorte. Se os dados forem iguais à sequência pseudoaleatória, com o XOR, eles serão todos 0. Esse resultado geralmente não ocorre com uma sequência pseudoaleatória longa, difícil de prever. Contudo, com uma sequência curta ou previsível, é possível que usuários maliciosos enviem padrões de bits que causem longas sequências de 0s depois do embaralhamento, causando falha nos enlaces. As antigas versões dos padrões para enviar pacotes IP por enlaces SONET no sistema telefônico tinham esse defeito (Malis e Simpson, 1999). Era possível que os usuários enviassem certos “pacotes fatais”, sabendo que isso causaria problemas.

## Sinais balanceados

Sinais que têm tanto tensão positiva quanto negativa, mesmo que seja por curtos períodos, são chamados de **sinais balanceados**. Sua média é zero, o que significa que eles não têm o componente elétrico CC (corrente contínua). A falta desse componente é uma vantagem, pois alguns canais, como o cabo coaxial ou as linhas com transformadores, atenuam bastante um componente CC, em virtude de suas propriedades físicas. Além disso, o método de conexão do receptor ao canal de comunicação, chamado de **acoplamento capacitivo**, deixa passar apenas a parte CA (corrente alternada) de um sinal. De qualquer modo, se enviarmos um sinal cuja média não seja zero, desperdiçamos energia, pois o componente CC será eliminado.

O balanceamento ajuda a oferecer transições para a recuperação de clock, pois existe uma mistura de tensões positiva e negativa. Ele também oferece um modo simples de calibrar receptores, pois a média do sinal pode ser medida e usada como um patamar de decisão para decodificar símbolos. Com sinais não balanceados, a média pode flutuar do nível de decisão verdadeiro, em razão de uma alta taxa de “1s”, por exemplo, que faria com que mais símbolos fossem decodificados com erros.

Um modo simples de construir um código balanceado é usar dois níveis de tensão para representar um “1” lógico e um 0 lógico. Por exemplo, +1 V para um bit 1 e -1 V para

um bit 0. Para enviar 1, o transmissor alterna entre os níveis +1 V e -1 V, de modo que eles sempre resultem na média. Esse esquema é chamado de **codificação bipolar**. Nas redes de telefone, ele é chamado **AMI (Alternate Mark Inversion)**, baseado em uma terminologia antiga, em que 1 é chamado “marca” e 0 é chamado “espaço”. Um exemplo pode ser visto na Figura 2.14(e).

A codificação bipolar acrescenta um nível de tensão para alcançar o equilíbrio. Como alternativa, podemos usar um mapeamento como 4B/5B para conseguir o equilíbrio (bem como as transições para a recuperação de clock). Um exemplo desse tipo de código balanceado é o código de linha **8B/10B**. Ele mapeia 8 bits de entrada para 10 bits de saída, de modo que é 80% eficiente, assim como o código de linha 4B/5B. Os 8 bits são divididos em um grupo de 5 bits, que é mapeado para 6 bits, e um grupo de 3 bits, que é mapeado para 4 bits. Os símbolos de 6 e 4 bits são, então, concatenados. Em cada grupo, alguns padrões de entrada podem ser mapeados para padrões de saída balanceados, com o mesmo número de 0s e 1s. Por exemplo, “001” é mapeado para “1001”, que é balanceado. Mas não existem combinações suficientes para todos os padrões de saída serem balanceados. Para esses casos, cada padrão de entrada é mapeado para dois padrões de saída. Um terá um 1 extra e a alternativa terá um 0 extra. Por exemplo, “000” é mapeado tanto para “1011” quanto para seu complemento “0100”. À medida que os bits de entrada são mapeados para bits de saída, o codificador recorda a **disparidade** do símbolo anterior. A disparidade é o número total de 0s ou 1s pelos quais o sinal está fora de equilíbrio. O codificador, então, seleciona um padrão de saída ou seu padrão alternativo, para reduzir a disparidade. Com 8B/10B, a disparidade será de no máximo 2 bits. Assim, o sinal nunca estará longe de ser balanceado. Nunca haverá mais do que cinco 1s ou 0s consecutivos, para ajudar com a recuperação do clock.

## Transmissão de banda passante

A comunicação por frequências de banda base é mais adequada para transmissões com fio, como par trançado, cabo coaxial ou fibra óptica. Em outras circunstâncias, particularmente aquelas envolvendo redes sem fio e transmissões de rádio, precisamos usar uma faixa de frequências que não começa com zero para enviar informações por um canal. Especificamente, para canais sem fio, não é prático enviar sinais de frequência muito baixa, pois o tamanho da antena precisa ser uma fração do comprimento de onda do sinal, que se torna grande em altas frequências de transmissão. De qualquer forma, restrições da regulamentação e a necessidade de evitar interferência normalmente ditam a escolha de frequências. Até mesmo para fios, colocar um sinal em determinada banda de frequência é útil para permitir que diferentes tipos de sinais coexistam no canal. Esse tipo de transmissão é chamado transmissão de banda

passante, pois uma banda de frequência arbitrária é usada para passar o sinal.

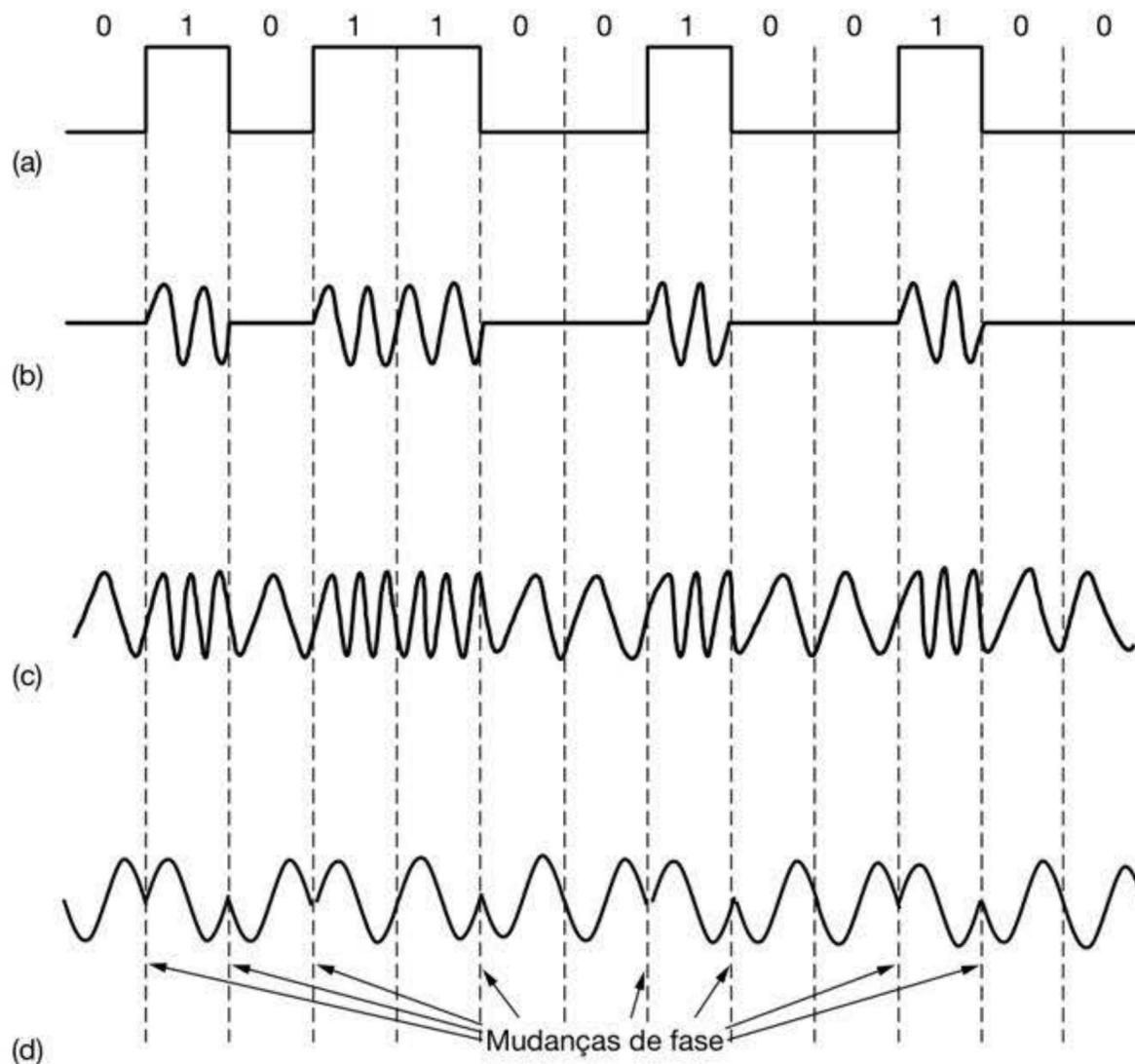
Felizmente, nossos resultados fundamentais, mostrados anteriormente, são todos em termos de largura de banda, ou da *largura* da banda de frequência. Os valores de frequência absoluta não importam para a capacidade. Isso significa que podemos apanhar um sinal de **banda base**, que ocupa de 0 a  $B$  Hz, e deslocá-lo para cima para ocupar uma banda passante de  $S$  a  $S + B$  Hz sem mudar a quantidade de informação que ele pode transportar, embora o sinal tenha aparência diferente. Para processar um sinal no receptor, podemos deslocá-lo de volta para a banda base, onde é mais conveniente detectar símbolos.

A modulação digital é realizada com a transmissão da banda passante regulando ou modulando um sinal de portadora sobreposto à banda passante. Podemos modular a amplitude, a frequência ou a fase do sinal da portadora. Cada um desses métodos tem um nome correspondente. No **ASK (Amplitude Shift Keying)**, duas amplitudes diferentes são usadas para representar 0 e 1. Um exemplo com um nível diferente de zero e zero aparece na Figura 2.16(b). Mais de dois níveis podem ser usados para codificar vários bits por símbolo.

De modo semelhante, com o chaveamento por deslocamento de frequência, ou **FSK (Frequency Shift Keying)**, dois ou mais tons diferentes são usados. O exemplo na

Figura 2.16(c) usa apenas duas frequências. Na forma mais simples do chaveamento por deslocamento de fase, ou **PSK (Phase Shift Keying)**, a onda da portadora é sistematicamente deslocada em 0 ou 180 graus em cada período de símbolo. Como existem duas fases, ela é chamada de **chaveamento binário por deslocamento de fase**, ou **BPSK (Binary Phase Shift Keying)**. O chaveamento “binário”, nesse caso, refere-se aos dois símbolos, não significa que eles representem 2 bits. Um exemplo aparece na Figura 2.16(d). Um esquema melhor, que usa a largura de banda do canal de modo mais eficiente, é usar quatro deslocamentos, por exemplo, 45, 135, 225 ou 315 graus, para transmitir 2 bits de informação por símbolo. Essa versão é chamada **chaveamento por deslocamento de fase em quadratura**, ou **QPSK (Quadrature Phase Shift Keying)**.

Podemos combinar esses esquemas e usar mais níveis para transmitir mais bits por símbolo. Somente a frequência ou a fase podem ser moduladas em determinado momento, pois elas estão relacionadas, normalmente sendo a taxa de mudança de fase com o tempo. Em geral, amplitude e fase são moduladas em combinação. Três exemplos aparecem na Figura 2.17 e, em cada um, os pontos dão as combinações válidas de amplitude e de fase de cada símbolo. Na Figura 2.17(a), vemos pontos equidistantes em 45, 135, 225 e 315 graus. A fase de um ponto é indicada pelo ângulo que uma linha a partir dele até a origem faz com o eixo  $x$



**Figura 2.16** (a) Um sinal binário. (b) Chaveamento por deslocamento de amplitude. (c) Chaveamento por deslocamento de frequência. (d) Chaveamento por deslocamento de fase.

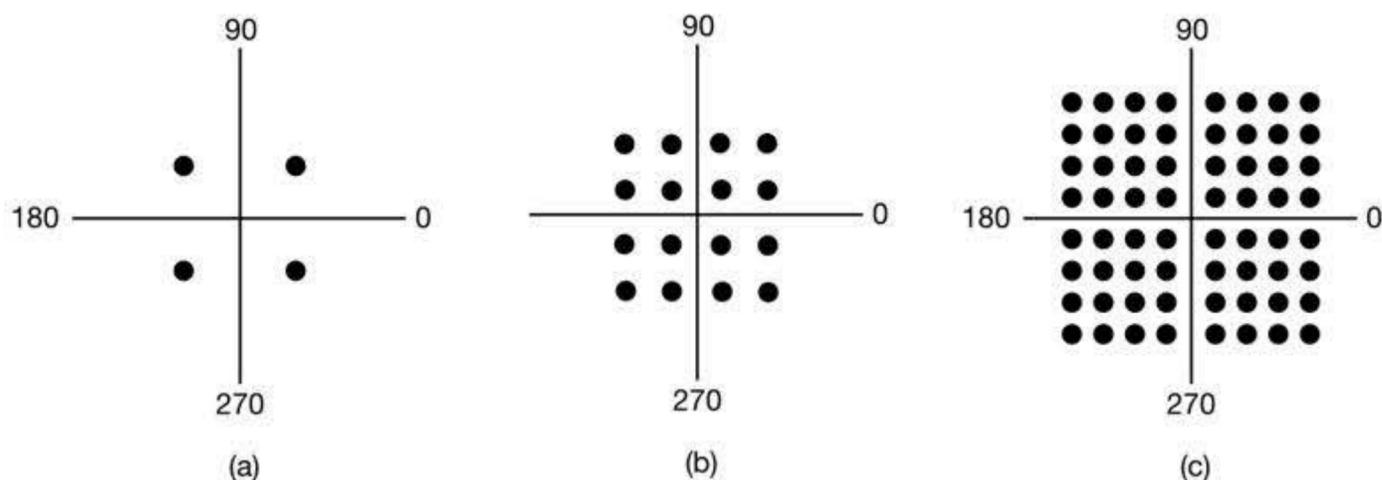


Figura 2.17 (a) QPSK. (b) QAM-16. (c) QAM-64.

positivo. A amplitude de um ponto é a distância a partir da origem. Essa figura é uma representação do QPSK.

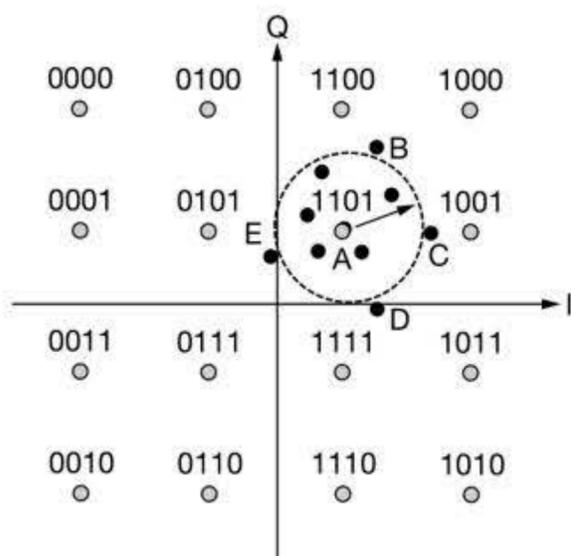
Esse tipo de diagrama é chamado de **diagrama de constelação**. Na Figura 2.17(b), vemos um esquema de modulação com um diagrama de constelação mais denso. Como são utilizadas 16 combinações de amplitudes e fase, o esquema de modulação pode ser usado para transmitir 4 bits por símbolo. Ele é chamado **QAM-16**, em que QAM significa **Quadrature Amplitude Modulation (modulação por amplitude de quadratura)**. A Figura 2.17(c) é outro esquema de modulação com 64 combinações diferentes, de forma que podem ser transmitidos 6 bits por símbolo. Ele é chamado **QAM-64**. Também são usadas QAMs de ordem mais alta. Como podemos imaginar a partir dessas constelações, é mais fácil criar circuitos eletrônicos para produzir símbolos como uma combinação de valores em cada eixo do que como uma combinação de valores de amplitude e fase. É por isso que os padrões se parecem com quadrados, em vez de círculos concêntricos.

As constelações que vimos até aqui não mostram de forma alguma como os bits são atribuídos aos símbolos. Ao fazer a atribuição, uma consideração importante é que uma pequena rajada de ruído no receptor não ocasiona muitos erros de bit. Isso poderia acontecer se atribuíssemos valores de bit consecutivos a símbolos adjacentes. Com a

QAM-16, por exemplo, se um símbolo designasse 0111 e o símbolo vizinho designasse 1000, se o receptor por engano apanhasse o símbolo adjacente, ele faria com que todos os bits ficassem errados. Uma solução melhor é mapear entre bits e símbolos, de modo que os símbolos adjacentes difiram em apenas 1 posição de bit. Esse mapeamento é chamado de **código Gray**. A Figura 2.18 mostra um diagrama de constelação QAM-16 que foi codificada pelo código Gray. Agora, se o receptor decodificar o símbolo com erro, ele cometerá apenas um erro de único bit no caso esperado em que o símbolo codificado está próximo do símbolo transmitido.

### 2.4.4 Multiplexação

Os esquemas de modulação que vimos nos permitem enviar um sinal para transmitir bits por um enlace com ou sem fio, mas eles só descrevem como transmitir um fluxo de bits por vez. Na prática, as economias de escala desempenham um papel importante no modo como usamos as redes. Custa basicamente o mesmo valor instalar e manter uma linha de transmissão com uma largura de banda alta e uma linha com largura de banda baixa entre dois escritórios diferentes (em outras palavras, os custos advêm de ter de cavar a



Quando 1101 é enviado:

Ponto	Decodifica como	Erros de bit
A	1101	0
B	1100	1
C	1001	1
D	1111	1
E	0101	1

Figura 2.18 QAM-16 em código Gray.

trincheira, não do tipo de cabo ou fibra que se utiliza). Consequentemente, esquemas de multiplexação têm sido desenvolvidos para compartilhar as linhas entre muitos sinais. As três principais formas de multiplexar uma única linha física são tempo, frequência e código; há também uma técnica chamada multiplexação por divisão de comprimento de onda, que é basicamente uma forma óptica de multiplexação por divisão de frequência. Discutiremos cada uma dessas técnicas a seguir.

### Multiplexação por divisão de frequência

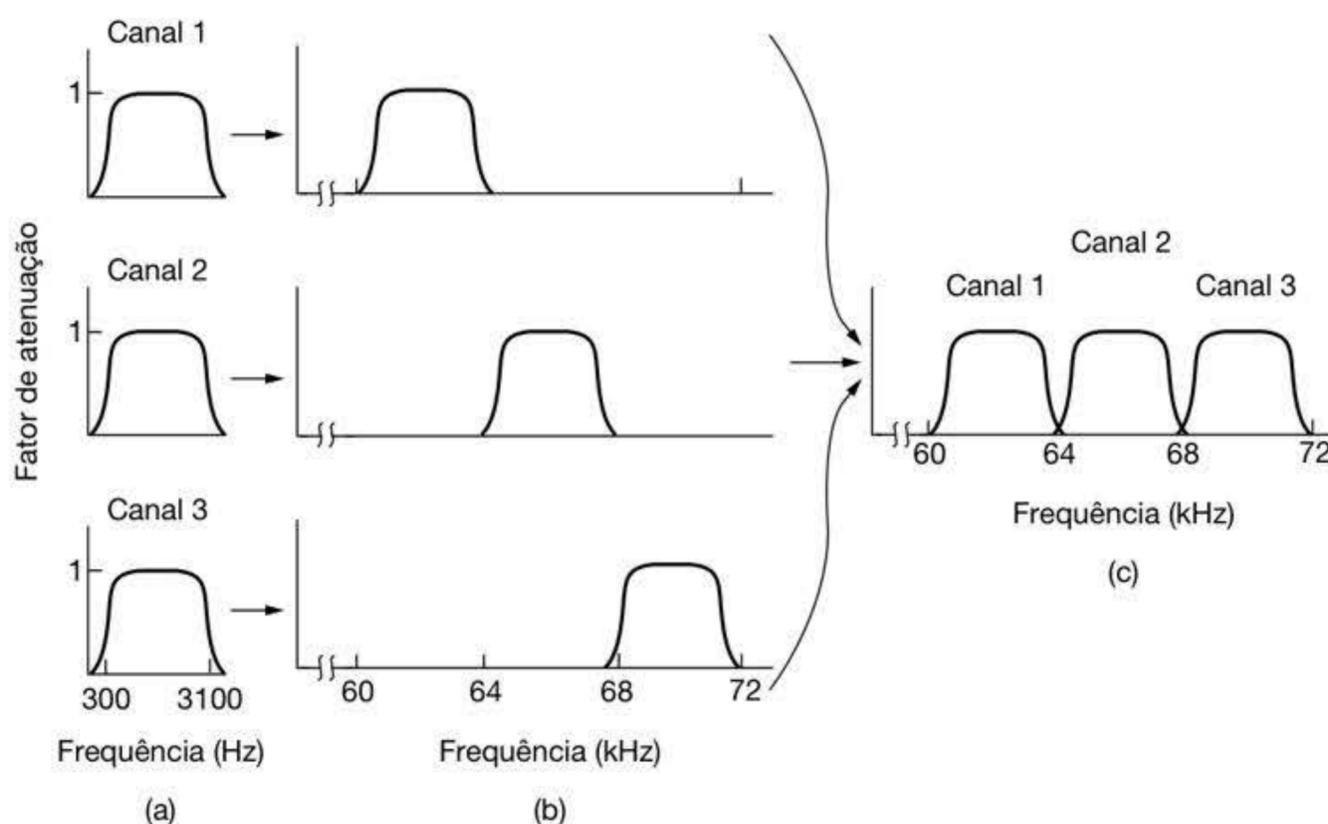
A **multiplexação por divisão de frequência**, ou **FDM (Frequency Division Multiplexing)**, tira proveito da transmissão de banda passante para compartilhar um canal. Ela divide o espectro em bandas de frequência, com cada usuário tendo posse exclusiva de alguma banda para enviar seu sinal. A transmissão de rádio AM ilustra a FDM. O espectro alocado é de cerca de 1 MHz, aproximadamente 500 a 1500 kHz. Diferentes frequências são alocadas a diferentes canais lógicos (estações), cada um operando em uma parte do espectro, com a separação entre canais grande o bastante para impedir interferência.

Para ver um exemplo mais detalhado, a Figura 2.19 mostra como três canais telefônicos de nível de voz são multiplexados com o uso da FDM. Os filtros limitam a largura de banda utilizável a cerca de 3.100 Hz por canal de qualidade de voz. Quando muitos canais são multiplexados ao mesmo tempo, são alocados 4.000 Hz para cada um. A largura de banda em excesso é chamada **banda de proteção**, e mantém os canais bem separados. Primeiro, os canais de voz têm sua frequência aumentada, cada qual com

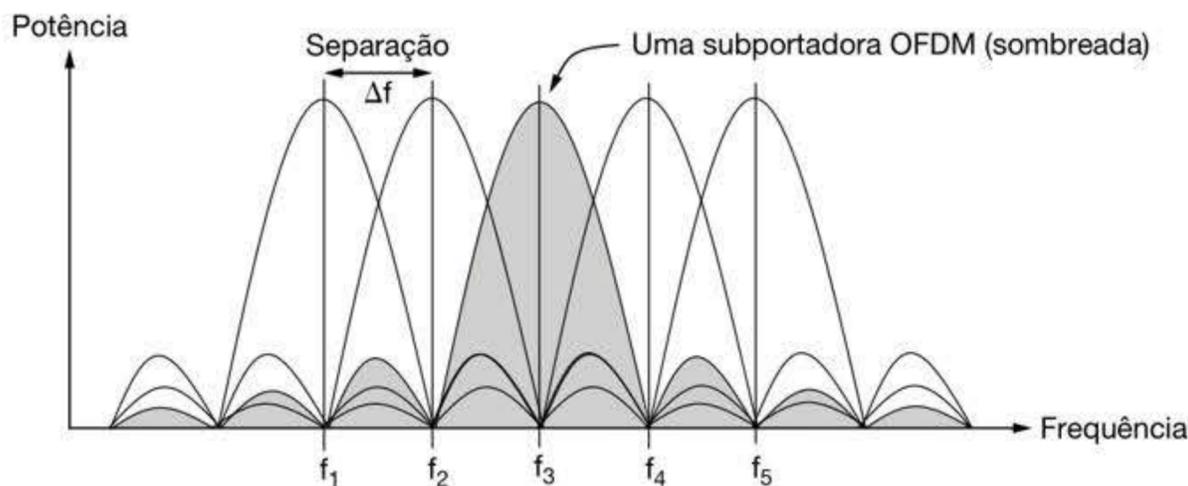
um valor diferente. Depois eles podem ser combinados, pois agora não há dois canais ocupando a mesma porção do espectro. Observe que, apesar de haver intervalos entre os canais graças às bandas de proteção, há certa sobreposição entre canais adjacentes, porque os filtros reais não têm limites nítidos ideais. Essa sobreposição significa que um forte pico na borda de um canal será sentido no canal adjacente como ruído não térmico.

Esse esquema tem sido usado para multiplexar chamadas no sistema telefônico há muitos anos, mas o esquema de multiplexação no tempo agora é preferido em lugar dele. Contudo, a FDM continua a ser usada nas redes telefônicas, assim como na telefonia celular, redes sem fio terrestres e redes por satélite em um nível de detalhamento mais alto.

Ao enviar dados digitais, é possível dividir o espectro de modo eficiente sem usar bandas de proteção. Na multiplexação ortogonal por divisão de frequência, ou **OFDM (Orthogonal Frequency Division Multiplexing)**, a largura de banda do canal é dividida em muitas subportadoras que enviam dados de maneira independente (p. ex., com QAM). As subportadoras compactam bastante o domínio de frequência, assim, os sinais de cada subportadora se estendem para as adjacentes. Contudo, como vemos na Figura 2.20, a resposta em frequência de cada subportadora é projetada de modo que seja zero no centro das subportadoras adjacentes. As subportadoras podem, portanto, ser amostradas em suas frequências centrais sem interferência de seus vizinhos. Para que isso funcione, um **tempo de proteção** é necessário para repetir uma parte dos sinais de símbolo no tempo, de modo que tenham a resposta de frequência desejada. Porém, esse overhead é muito menor que o necessário para muitas bandas de proteção.



**Figura 2.19** Multiplexação por divisão de frequência. (a) As larguras de banda originais. (b) As larguras de banda aumentaram em frequência. (c) O canal multiplexado.

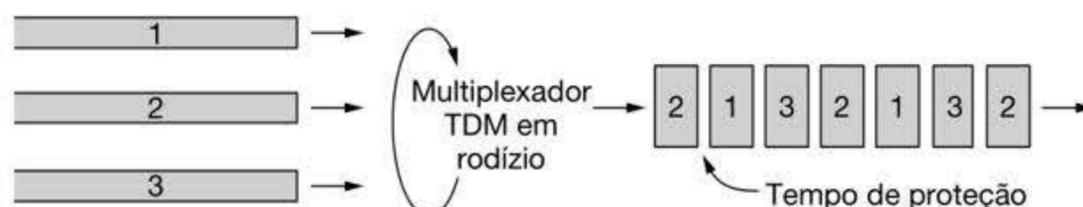


**Figura 2.20** Multiplexação ortogonal por divisão de frequência (OFDM).

A ideia da OFDM já existe há muito tempo, mas somente no início da década de 2000 ela foi amplamente adotada, seguindo a observação de que é possível implementá-la de modo eficiente em termos de uma transformada de Fourier dos dados digitais por todas as subportadoras (em vez de modular separadamente cada subportadora). A OFDM é usada em 802.11, redes a cabo, redes por linhas de energia elétrica e sistemas celulares de quarta geração (4G). Normalmente, o fluxo de informações digitais de alta velocidade é dividido em muitos fluxos de baixa velocidade, os quais são transmitidos nas subportadoras em paralelo. Essa divisão é valiosa porque é mais fácil lidar com as degradações do canal no nível da subportadora; algumas subportadoras podem ser muito degradadas e excluídas em favor de subportadoras que são bem recebidas.

### Multiplexação por divisão de tempo

Uma alternativa à FDM é a **multiplexação por divisão de tempo**, ou **TDM (Time Division Multiplexing)**. Aqui, os usuários se alternam (em um padrão de rodízio), cada um periodicamente usando a largura de banda inteira por um pequeno período. Um exemplo de três fluxos sendo multiplexados com TDM aparece na Figura 2.21. Os bits de cada fluxo de entrada são apanhados em um **slot de tempo** fixo e enviados para o fluxo agregado. Esse fluxo trabalha em uma velocidade que é a soma dos fluxos individuais. Para que isso funcione, os fluxos precisam estar sincronizados no tempo. Pequenos intervalos de tempo de proteção (semelhantes à banda de proteção de frequência) podem ser acrescentados para acomodar pequenas variações de sincronização.



**Figura 2.21** Multiplexação por divisão de tempo (TDM).

A TDM é bastante usada como parte das redes de telefone e celular. Para evitar um ponto de confusão, vamos esclarecer que isso é muito diferente da **multiplexação estatística por divisão de tempo**, ou **STDM (Statistical Time Division Multiplexing)**. O termo “estatística” é acrescentado para indicar que os fluxos individuais contribuem para o fluxo multiplexado, e *não* sobre uma programação fixa, mas de acordo com a estatística de sua demanda. A STDM é a comutação de pacotes com outro nome.

### Multiplexação por divisão de código

Existe um terceiro tipo de multiplexação, que funciona de modo completamente diferente da FDM e da TDM. A **multiplexação por divisão de código**, ou **CDM (Code Division Multiplexing)** é uma forma de comunicação por **dispersão espectral**, na qual um sinal de banda estreita é espalhado por uma banda de frequência mais larga. Isso pode torná-la ainda mais tolerante às interferências, além de permitir que vários sinais de diferentes usuários compartilhem a mesma banda de frequência. Como a CDM é usada principalmente para essa última finalidade, ela normalmente é chamada de **acesso múltiplo por divisão de código**, ou **CDMA (Code Division Multiple Access)**.

O CDMA permite que cada estação transmita por todo o espectro de frequência o tempo todo. Várias transmissões simultâneas são separadas usando a teoria da codificação. Antes de entrarmos no algoritmo, vamos considerar uma analogia: o saguão de um aeroporto com muitos pares de pessoas conversando. Com a TDM, todas as pessoas estariam no meio do saguão, mas conversariam por turnos, um par de pessoas de cada vez. Com a FDM, as pessoas formariam grupos bem separados, cada

um mantendo sua própria conversa ao mesmo tempo, alguns com uma altura maior e outros com uma altura menor, de modo que cada par pudesse manter sua própria conversa ao mesmo tempo, mas ainda independentemente dos outros grupos. Com o CDMA, todas as pessoas estariam no meio do saguão falando ao mesmo tempo, mas cada par de pessoas conversando em um idioma diferente. O par que estivesse falando em francês só reconheceria esse idioma, rejeitando tudo que não fosse francês como ruído. Desse modo, a chave do CDMA é a capacidade de extrair o sinal desejado e rejeitar todos os outros como ruído aleatório. A seguir, veremos uma descrição um pouco simplificada do sistema.

No CDMA, cada tempo de duração de um bit é subdividido em  $m$  intervalos curtos, denominados **chips**, que são multiplicados pela sequência de dados original (os chips são uma sequência de bits, mas são denominados chips para que não sejam confundidos com os bits da própria mensagem). Normalmente, existem 64 ou 128 chips por bit, mas, no exemplo apresentado a seguir, usaremos 8 chips/bit, para simplificar. A cada estação é atribuído um código de  $m$  bits exclusivo, chamado **sequência de chips**. Para fins pedagógicos, é mais conveniente escrever esses códigos como sequências de  $-1$  e  $+1$ . Mostraremos as sequências de chips entre parênteses.

Para transmitir um bit 1, uma estação envia sua sequência de chips. Para transmitir um bit 0, envia a negação de sua sequência de chips. Não são permitidos quaisquer outros padrões. Assim, para  $m = 8$ , se a estação  $A$  receber a atribuição da sequência de chips  $(-1 -1 -1 +1 +1 -1 +1 +1)$ , ela transmitirá um bit 1 ao enviar a sequência de chips, e 0 transmitindo seu complemento:  $(+1 +1 +1 -1 -1 +1 -1 -1)$ . Na realidade, são enviados sinais com esses níveis de tensão, mas isso é suficiente para pensarmos em termos das sequências.

O aumento do volume de informações a serem enviadas de  $b$  bits/s para  $mb$  chips/s só poderá ocorrer se a largura de banda disponível for  $m$  vezes maior que a largura de

banda necessária para uma estação que não usa CDMA (se não houver mudanças nas técnicas de modulação ou codificação). Se tivéssemos uma banda de 1 MHz disponível para 100 estações, com FDM, cada uma teria 10 kHz e poderia transmitir a uma velocidade de 10 kbps (supondo-se 1 bit por Hz). No CDMA, cada estação utiliza 1 MHz inteiro e, portanto, a taxa de chips é de 100 chips por bit, para espalhar a taxa de bits da estação de 10 kbps pelo canal.

Na Figura 2.22(a) e (b), mostramos as sequências de chips binárias atribuídas a quatro exemplos de estações e os sinais que elas representam. Cada estação tem sua própria sequência exclusiva de chips. Vamos usar o símbolo  $S$  para indicar o vetor de  $m$  chips correspondente à estação  $S$ , e  $\bar{S}$  para sua negação. Todas as sequências de chips são **ortogonais** par a par – isso significa que o produto interno normalizado de duas sequências de chips distintas,  $S$  e  $T$  (indicado como  $S \cdot T$ ), é 0. Sabemos como gerar tal sequência ortogonal de chips usando um método conhecido como **códigos de Walsh**. Em termos matemáticos, a ortogonalidade das sequências de chips pode ser expressa por:

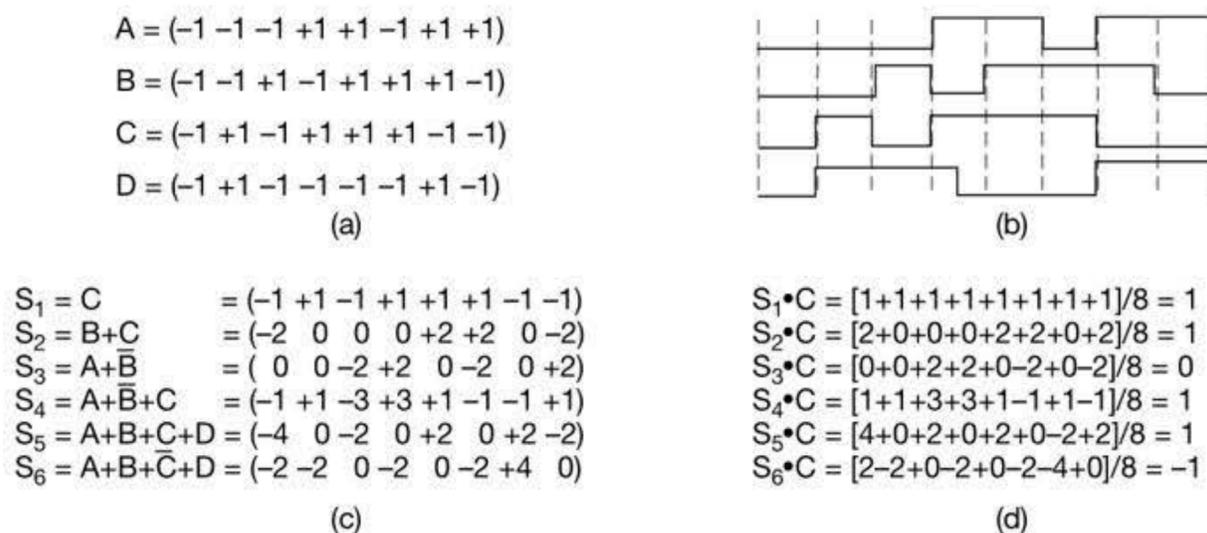
$$S \cdot T = \frac{1}{m} \sum_{i=1}^m S_i T_i = 0 \quad (2.5)$$

Em linguagem comum, o número de pares iguais é igual ao número de pares diferentes. Essa propriedade da ortogonalidade será essencial mais adiante. Observe que, se  $S \cdot T = 0$ , então  $S \cdot \bar{T}$  também será 0. O produto interno normalizado de qualquer sequência de chips por ela mesma é igual a 1:

$$S \cdot S = \frac{1}{m} \sum_{i=1}^m S_i S_i = \frac{1}{m} \sum_{i=1}^m S_i^2 = \frac{1}{m} \sum_{i=1}^m (\pm 1)^2 = 1$$

Isso ocorre porque cada um dos  $m$  termos do produto interno é 1 e, portanto, a soma é  $m$ . Observe também que  $S \cdot \bar{S} = -1$ .

Durante cada intervalo com duração de um bit, uma estação pode transmitir um bit 1 (enviando sua sequência de chips), pode transmitir um bit 0 (enviando a negativa



**Figura 2.22** (a) Sequências de chips binárias para quatro estações. (b) Sinais que as sequências representam. (c) Seis exemplos de transmissões. (d) Recuperação do sinal da estação C.

de sua sequência de chips), ou pode ficar inativa e não realizar nenhuma transmissão. Por enquanto, supomos que todas as estações estão sincronizadas e que a transmissão de todas as sequências de chips começa no mesmo instante. Quando duas ou mais estações transmitem simultaneamente, suas sequências bipolares somam-se linearmente. Por exemplo, se, durante um período de um chip, três estações transmitirem +1 e uma estação transmitir -1, o resultado +2 será recebido. Isso pode ser considerado como a soma de tensões sobrepostas no canal: se três estações transmitirem +1 volt e uma estação transmitir -1 volt como saída, o resultado será 2 volts. A Figura 2.22(c) apresenta seis exemplos em que uma ou mais estações transmitem um bit 1 ao mesmo tempo. No primeiro exemplo, *C* transmite um bit 1 e, assim, simplesmente obtemos a sequência de chips de *C*. No segundo, *B* e *C* transmitem bits 1 e obtemos a soma de suas sequências bipolares de chips da seguinte forma:

$$\begin{aligned} (-1 -1 +1 -1 +1 +1 +1 -1) + (-1 +1 -1 +1 +1 +1 -1 -1) = \\ (-2 \ 0 \ 0 \ 0 +2 +2 \ 0 -2) \end{aligned}$$

Para recuperar o fluxo de bits de uma estação individual, o receptor precisa conhecer com antecedência a sequência de chips da estação transmissora. Ele executa a recuperação calculando o produto interno normalizado da sequência de chips recebida e a sequência de chips da estação cujo fluxo de bits está tentando recuperar. Se a sequência de chips recebida for **S** e o receptor estiver tentando ouvir uma estação cuja sequência de chips é **C**, ele apenas calcula o produto interno normalizado, **S**•**C**.

Para entender por que esse procedimento funciona, imagine que as duas estações, *A* e *C*, transmitem um bit 1 ao mesmo tempo que *B* transmite um bit 0, como no caso do terceiro exemplo. O receptor percebe a soma, **S** = **A** + **B** + **C**, e calcula:

$$\mathbf{S} \cdot \mathbf{C} = (\mathbf{A} + \mathbf{B} + \mathbf{C}) \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C} + \mathbf{C} \cdot \mathbf{C} = 0 + 0 + 1 = 1$$

Os dois primeiros termos desaparecem porque todos os pares de sequências de chips foram cuidadosamente escolhidos para serem ortogonais, como mostra a Equação 2.5. Agora já deve estar claro por que essa propriedade precisa ser imposta às sequências de chips.

Para tornar mais concreto o processo de decodificação, vamos considerar novamente os seis exemplos da Figura 2.22(d). Suponha que o receptor esteja interessado em extrair o bit enviado pela estação *C* de cada uma das seis somas de *S*<sub>1</sub> a *S*<sub>6</sub>. Ele calcula o bit somando aos pares os produtos da **S** recebida com o vetor **C** da Figura 2.22(a), extraindo depois 1/8 do resultado (pois, neste caso *m* = 8). Os exemplos incluem casos em que *C* é silencioso, envia um bit 1 e um bit 0, individualmente e em combinação com outras transmissões. Como vemos, o bit correto é decodificado a cada vez. É como falar francês.

Em princípio, dada uma capacidade de computação suficiente, o receptor poderá escutar todos os transmissores ao mesmo tempo, executando o algoritmo de decodificação correspondente a cada um deles em paralelo. Na prática, é mais fácil falar do que fazer, e é útil saber quais transmissores poderiam estar transmitindo.

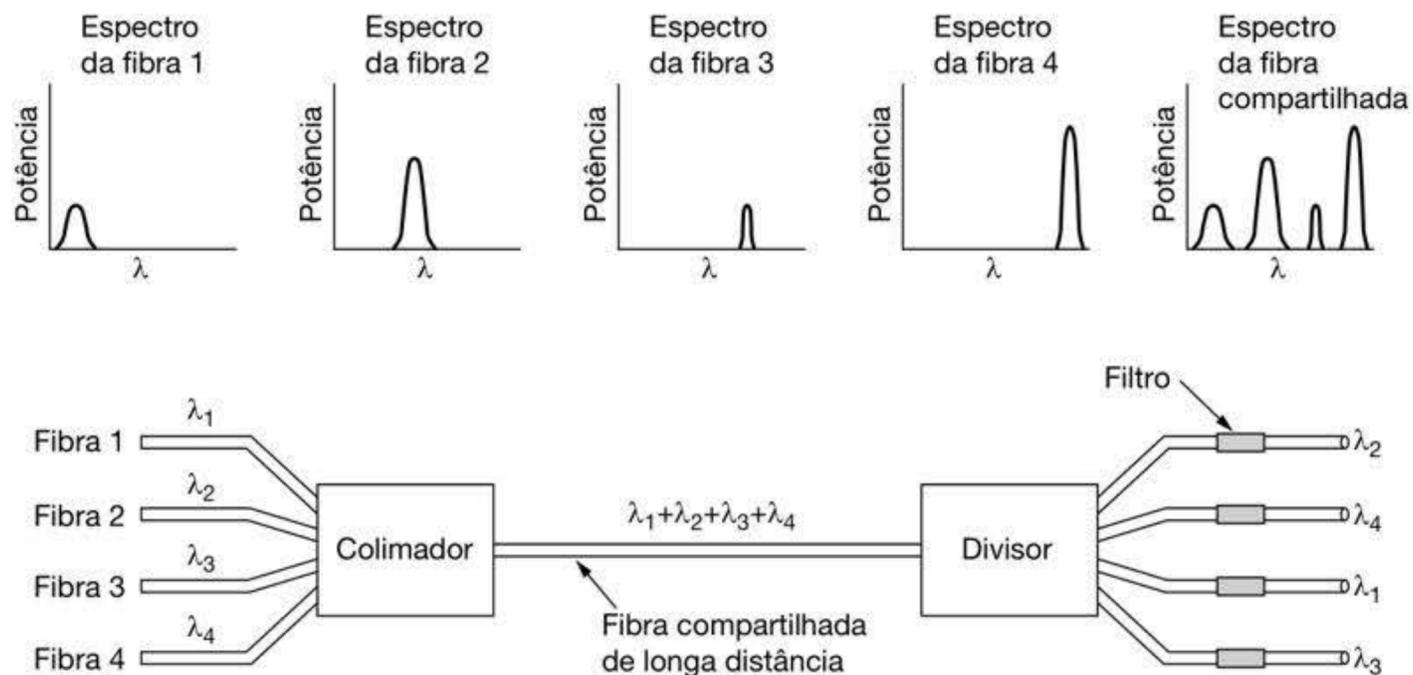
O ideal é que, no sistema CDMA sem ruído que estudamos aqui, o número de estações que enviam simultaneamente possa ser arbitrariamente grande, usando sequências de chips maiores. Para 2<sup>*n*</sup> estações, os códigos de Walsh podem oferecer 2<sup>*n*</sup> sequências de chips ortogonais de tamanho 2<sup>*n*</sup>. Contudo, uma limitação significativa é o fato de considerarmos que todos os chips são sincronizados no tempo no receptor. Essa sincronização nem mesmo é aproximadamente verdadeira em algumas aplicações, como em redes de celular (nas quais o CDMA foi bastante utilizado a partir da década de 1990). Isso leva a diferentes projetos.

Assim como as redes de celular, o CDMA é usado por satélites e redes a cabo. Nesta rápida introdução, passamos brevemente pelos muitos fatores complicadores. Os engenheiros que desejam obter um conhecimento profundo do CDMA deverão ler Viterbi (1995) e Harte et al. (2012). Entretanto, essas referências exigem muita base em engenharia de comunicação.

## Multiplexação por divisão de comprimento de onda

A **multiplexação por divisão de comprimento de onda**, ou **WDM (Wavelength Division Multiplexing)** é uma forma de FDM que multiplexa diversos sinais em uma fibra óptica usando diferentes comprimentos de onda da luz. Na Figura 2.23, quatro fibras chegam juntas a um colimador óptico, cada uma com sua energia presente em um comprimento de onda distinto. Os quatro feixes são combinados em uma única fibra compartilhada para transmissão a um destino remoto. Na extremidade remota, o feixe é dividido no mesmo número de fibras que havia no lado da entrada. Cada fibra de saída contém um núcleo curto, especialmente construído, que filtra todos os comprimentos de onda, com exceção de um. Os sinais resultantes podem ser roteados até seu destino ou recombinados de diferentes maneiras para transporte ou multiplexação adicional.

Realmente não há nada de novo aqui. Trata-se apenas da FDM em frequências muito altas, com o termo WDM devendo-se à descrição dos canais de fibra óptica em função de seu comprimento de onda ou “cor” em vez da frequência propriamente dita. Desde que cada canal tenha sua própria frequência dedicada (i.e., de comprimentos de onda) e todas as faixas sejam disjuntas, elas poderão ser multiplexadas na fibra de longa distância. A única diferença em relação à FDM elétrica é que um sistema óptico que utilize uma rede de difração será completamente passivo e, portanto, altamente confiável.



**Figura 2.23** Multiplexação por divisão de comprimento de onda.

A razão para a WDM ser popular é o fato de a energia em um único canal normalmente ter apenas alguns gigahertz de largura porque, no momento, é impossível realizar a conversão entre sinais elétricos e ópticos com rapidez maior que essa. Utilizando-se muitos canais em paralelo com diferentes comprimentos de onda, a largura de banda agregada aumenta de forma linear com o número de canais. Como a largura de banda de uma única banda de fibra é de aproximadamente 25.000 GHz (veja a Figura 2.5), teoricamente existe espaço para 2.500 canais de 10 Gbps, mesmo a 1 bit/Hz (e também são possíveis taxas mais altas).

A tecnologia WDM tem progredido a uma taxa que faria a tecnologia computacional se envergonhar. O sistema WDM foi criado por volta de 1990, com os primeiros sistemas comerciais tendo 8 canais de 2,5 Gbps por canal. Em 1998, havia sistemas com 40 canais de 2,5 Gbps no mercado. Em 2006, já eram utilizados produtos com 192 canais de 10 Gbps e 64 canais de 40 Gbps, capazes de mover até 2,56 Tbps. Em 2019, havia sistemas que podiam lidar com 160 canais, aceitando mais de 16 Tbps por um único par de fibra. Isso é 800 vezes mais capacidade que os sistemas de 1990. Os canais também estão bastante compactados na fibra, com 200, 100 ou mesmo 50 GHz de separação.

Estreitando o espaçamento para 12,5 GHz, é possível usar 320 canais em uma única fibra, aumentando ainda mais a capacidade de transmissão. Tais sistemas, com um grande número de canais pouco espaçados, são conhecidos como **DWDM (Dense WDM)**. Sistemas DWDM costumam ser mais caros, pois precisam manter comprimentos de onda e frequências estáveis, devido ao pouco espaçamento de cada canal. Como resultado, esses sistemas regulam de perto sua temperatura, a fim de garantir que as frequências sejam precisas.

Um dos impulsionadores da tecnologia WDM é o desenvolvimento de amplificadores totalmente ópticos. Antes, a cada 100 km era necessário dividir todos os canais

e converter cada um deles a um sinal elétrico para amplificação separada, antes de convertê-los novamente em sinais ópticos e combiná-los. Hoje, os amplificadores totalmente ópticos podem regenerar o sinal inteiro uma única vez a cada 1.000 km, sem a necessidade de várias conversões ópticas/elétricas.

No exemplo da Figura 2.23, temos um sistema de comprimento de onda fixo. Bits da fibra de entrada 1 vão para a fibra de saída 3, bits da fibra de entrada 2 vão para a fibra de saída 1, etc. Porém, também é possível criar sistemas WDM comutados no domínio óptico. Em dispositivos como esses, os filtros de saída são ajustáveis com o uso de interferômetros de Fabry-Perot ou de Mach-Zehnder. Esses dispositivos permitem que as frequências selecionadas sejam trocadas dinamicamente por um computador de controle. Essa capacidade oferece uma grande flexibilidade para a provisão de muitos caminhos de comprimento de onda diferentes através da rede telefônica a partir de um conjunto fixo de fibras. Para obter mais informações sobre redes ópticas e WDM, consulte Grobe e Eiselt (2013).

## 2.5 A REDE PÚBLICA DE TELEFONIA COMUTADA

Quando dois computadores, instalados perto um do outro, precisam se comunicar, geralmente é mais fácil conectá-los por meio de um cabo. As LANs funcionam dessa forma. No entanto, quando as distâncias começam a ficar grandes ou muitos computadores e cabos têm de atravessar uma estrada ou outra passagem pública, os custos de instalação de cabos privados costumam ser proibitivos. Além disso, em quase todos os países do mundo, também é ilegal estender linhas de transmissão privadas em (ou sob) propriedades públicas. Consequentemente, os projetistas de rede devem

confiar nos recursos de telecomunicações existentes, como a rede de telefonia, a rede celular ou a rede de televisão a cabo.

O fator limitante para a rede de dados tem sido a “última milha” sobre a qual os clientes se conectam, que pode depender de qualquer uma dessas tecnologias físicas, ao contrário da chamada infraestrutura de “backbone” para o restante da rede de acesso. Na última década, essa situação mudou drasticamente, com velocidades de 1 Gbps para casa se tornando cada vez mais comuns. Embora uma contribuição para velocidades de último quilômetro mais rápidas seja a implantação contínua de fibra na extremidade da rede, talvez um fator ainda mais significativo em alguns países seja a engenharia sofisticada das redes de telefone e cabo *existentes* para espremer cada vez mais largura de banda da infraestrutura. Acontece que trabalhar com a infraestrutura física existente para aumentar as velocidades de transmissão é muito mais barato do que colocar novos cabos (de fibra) no solo. Agora, vamos explorar as arquiteturas e características de cada uma dessas infraestruturas físicas de comunicação.

Esses recursos existentes, em particular a **PSTN (Public Switched Telephone Network)**, foram projetados há muitos anos, tendo em vista um objetivo completamente diferente: a transmissão da voz humana de uma forma mais ou menos reconhecível. Um cabo instalado entre dois computadores pode transferir dados a 10 Gbps ou mais; a rede telefônica, portanto, tem seu trabalho reduzido em termos de transmissão de bits em altas velocidades. As primeiras tecnologias DSL (Digital Subscriber Line) só podiam transmitir dados em velocidades de alguns Mbps; agora, as versões mais modernas da DSL podem alcançar velocidades que se aproximam de 1 Gbps. Nas próximas seções, descreveremos o sistema telefônico e mostraremos como ele funciona. Para obter informações adicionais sobre o funcionamento interno do sistema telefônico, consulte Laino (2017).

### 2.5.1 Estrutura do sistema telefônico

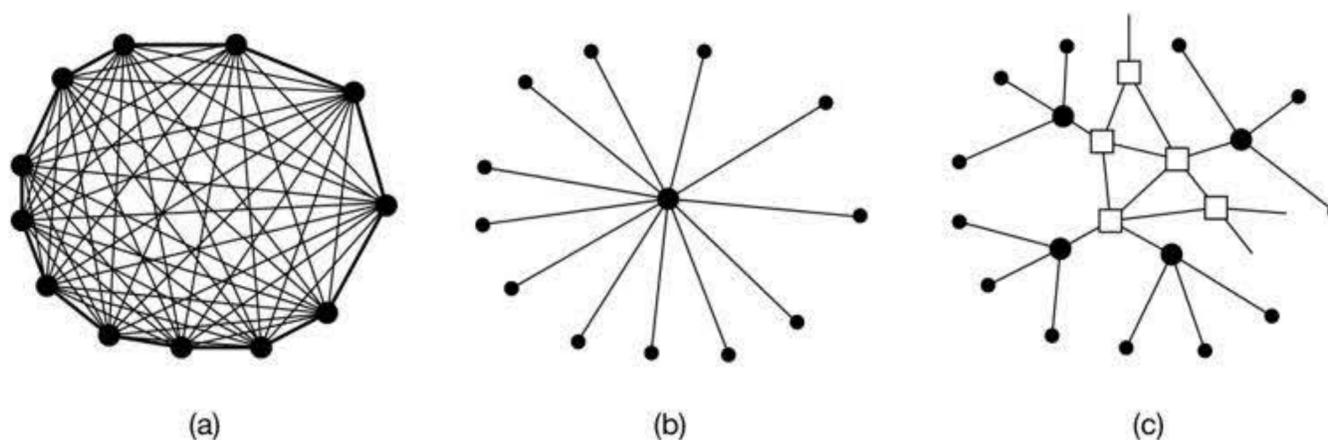
Logo depois que Alexander Graham Bell patenteou a invenção do telefone, em 1876 (apenas algumas horas antes

de seu concorrente, Elisha Gray), houve uma grande demanda por essa nova invenção. Inicialmente, o mercado estava voltado para a venda de telefones, que eram comercializados aos pares. Era o usuário quem tinha de conectar os dois aparelhos usando um fio. Se o proprietário de um telefone quisesse usar o aparelho para conversar com  $n$  outros proprietários de telefone, tinha de conectar fios em todas as  $n$  residências. Em um ano, as cidades ficaram tomadas por fios que passavam pelas casas e pelas árvores, em uma selva de fios. Logo ficou óbvio que o modelo de conexão de um telefone a outro, como é mostrado na Figura 2.24(a), não funcionaria.

Bell percebeu essa situação e criou a Bell Telephone Company, que abriu sua primeira estação de comutação (em New Haven, Connecticut) em 1878. A empresa ligava um fio até a casa ou o escritório de cada usuário. Para fazer uma chamada, o usuário girava a manivela, o que emitia um som na companhia telefônica e chamava a atenção de um operador. Este, por sua vez, conectava manualmente o emissor da chamada ao receptor usando um cabo curto. Observe, na Figura 2.24(b), o modelo de uma única estação de comutação.

Não demorou muito tempo para as estações de comutação da Bell System se espalharem por todos os locais, e logo as pessoas passaram a querer fazer chamadas interurbanas. Por isso, a Bell System passou a conectar uma estação de comutação a outra. Contudo, o problema original veio à tona mais uma vez: conectar cada estação de comutação a outra através de um fio logo se tornou inviável. Então, foram inventadas as estações de comutação de segundo nível. Depois de algum tempo, tornaram-se necessárias várias estações de segundo nível, como mostra a Figura 2.24(c). Mais tarde, a hierarquia cresceu até alcançar cinco níveis.

Em 1890, era possível notar a presença das três principais partes do sistema telefônico: as estações de comutação, os fios que ligavam os usuários a essas estações (agora já operando com cabos de pares trançados, isolados e balanceados em vez de cabos abertos com retorno por terra) e as conexões de longa distância existentes entre



**Figura 2.24** (a) Rede totalmente interconectada. (b) Switch centralizado. (c) Hierarquia de dois níveis.

as estações de comutação. Para obter breves informações técnicas sobre o sistema telefônico e sua história, consulte Hawley (1991).

Embora tenha havido melhorias em todas as três áreas desde então, o modelo básico da Bell System continuou praticamente intacto por mais de 100 anos. Embora seja bastante simplificada, a descrição a seguir apresenta a ideia básica do sistema telefônico. Cada telefone contém dois fios de cobre que saem do aparelho e se conectam diretamente à **estação final** mais próxima da companhia telefônica (também denominada **estação central local**). Em geral, a distância varia de 1 a 10 km, sendo menor nas cidades que nas regiões rurais. Só nos Estados Unidos, existem cerca de 22 mil estações finais. As conexões através de dois fios entre o telefone de cada assinante e a estação final são conhecidas no mercado como **circuito terminal**. Se todos os circuitos terminais existentes no mundo fossem esticados de uma extremidade a outra, seu comprimento equivaleria a mil vezes a distância da Terra à Lua e de volta à Terra.

Houve uma época em que 80% do capital da AT&T estava no cobre dos circuitos terminais, o que a tornava a maior mina de cobre do mundo. Felizmente, essa informação não foi muito difundida na comunidade financeira. Se tivessem conhecimento desse fato, alguns empresários poderiam comprar a AT&T, encerrar todos os serviços de telefonia dos Estados Unidos, descascar toda a fiação e vender os fios a uma refinaria de cobre para ter um retorno rápido do capital.

Se um assinante conectado a uma determinada estação final ligar para outro assinante da mesma estação, o mecanismo de comutação dentro da estação configurará uma conexão elétrica direta entre os dois circuitos terminais. Essa conexão permanecerá intacta durante a chamada.

Se o telefone chamado estiver conectado a outra estação final, outro procedimento terá de ser usado. Cada estação final contém uma série de linhas de saída para um ou mais centros de comutação vizinhos, denominados **estações interurbanas** (ou, se estiverem na mesma área, **estações Tandem**). Essas linhas são denominadas **troncos de conexão interurbana**. O número de diferentes tipos de

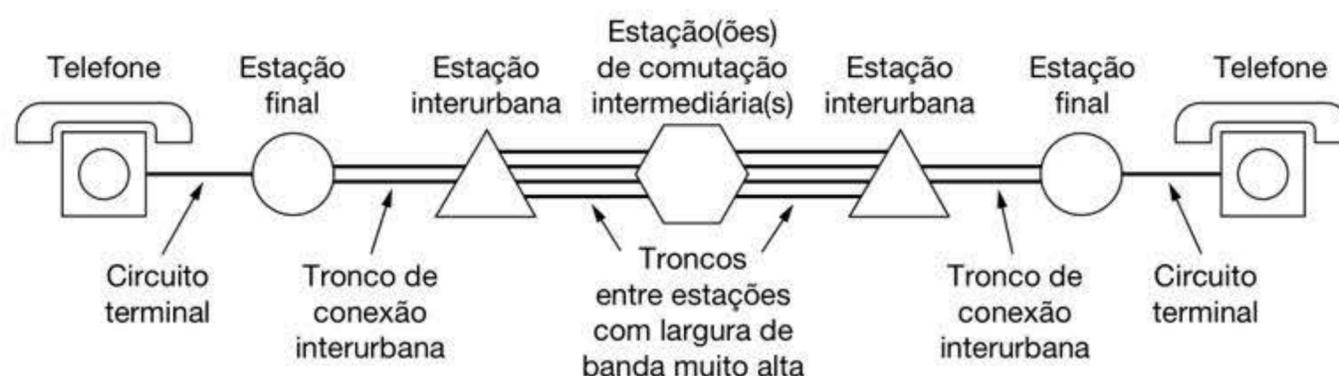
centros de comutação e sua topologia variam de um país para outro, dependendo da densidade da telefonia do país.

Se as estações finais do transmissor e do receptor tiverem um tronco de conexão interurbana ligado à mesma estação interurbana (uma situação bastante provável caso estejam geograficamente próximos), a conexão poderá ser estabelecida dentro da estação interurbana. Observe, na Figura 2.24(c), uma rede telefônica formada apenas por telefones (os pontos pequenos), estações finais (os pontos maiores) e estações interurbanas (os quadrados).

Se o transmissor e o receptor não compartilharem a mesma estação interurbana, terá de ser estabelecido um caminho entre duas estações interurbanas, as quais se comunicam entre si por meio de **troncos interurbanos** de alta largura de banda (também denominados **troncos entre estações**). Antes da dissolução da AT&T, em 1984, o sistema telefônico dos Estados Unidos usava o roteamento hierárquico para encontrar um caminho, indo para níveis mais altos da hierarquia até que houvesse uma estação de comutação em comum. Isso foi substituído pelo roteamento não hierárquico, mais flexível. A Figura 2.25 mostra como uma conexão de média distância pode ser roteada.

Nas telecomunicações, são usados vários meios de transmissão. Diferentemente dos prédios de escritórios modernos, nos quais a fiação normalmente é de Categoria 5 ou 6, os circuitos terminais são formados por cabos de pares trançados de Categoria 3, com a fibra apenas começando a aparecer. Entre as estações de comutação, o uso de cabos coaxiais, micro-ondas e, principalmente, de fibras ópticas, é bastante frequente.

No passado, a transmissão em todo o sistema telefônico era analógica, com o sinal de voz real sendo transmitido como uma tensão elétrica da origem até o destino. Com o advento da fibra óptica, da eletrônica digital e dos computadores, todos os troncos e switches são agora digitais, deixando o circuito terminal como o último fragmento de tecnologia analógica no sistema. A transmissão digital é preferida porque, em uma chamada longa, não é necessário reproduzir com precisão uma forma de onda analógica depois de ter passado por muitos amplificadores.



**Figura 2.25** Rota de um circuito típico para uma chamada de longa distância.

Ser capaz de distinguir corretamente 0 de 1 já é suficiente. Essa propriedade torna a transmissão digital mais confiável que a analógica, sendo também mais econômica e de manutenção mais fácil.

Em suma, o sistema telefônico é formado por três componentes principais:

1. Circuitos terminais (pares trançados analógicos indo para as residências e para as empresas).
2. Troncos (fibra óptica digital de altíssima largura de banda conectando as estações de comutação).
3. Estações de comutação (onde as chamadas são transferidas de um tronco para outro, elétrica ou opticamente).

Os circuitos terminais oferecem acesso ao sistema inteiro para todas as pessoas; assim, eles são críticos. Infelizmente, eles também constituem o elo mais fraco no sistema. Para os troncos de longa distância, a principal questão é reunir várias chamadas e transmiti-las ao mesmo tempo, pela mesma fibra, o que é feito usando a WDM. Por último, existem duas formas fundamentalmente distintas de executar a comutação: comutação de circuitos e comutação de pacotes. Portanto, analisaremos ambas.

## 2.5.2 O circuito terminal: modems, ADSL e fibra óptica

Nesta seção, estudaremos o circuito terminal, tanto o novo quanto o antigo. Veremos os modems de telefone, ADSL e FTTH. Em alguns lugares, o circuito terminal foi modernizado instalando-se fibra óptica até a residência (ou muito próximo dela). Essas instalações admitem redes de computadores desde o início, com o circuito terminal tendo ampla largura de banda para os serviços de dados. Infelizmente, o custo para levar a fibra até as residências é muito grande. Às vezes, isso é feito quando as ruas são escavadas para outras finalidades; alguns municípios, especialmente em áreas urbanas densamente povoadas, possuem circuitos terminais de fibra. Porém, em sua maioria, eles são a exceção, mas certamente farão parte do futuro.

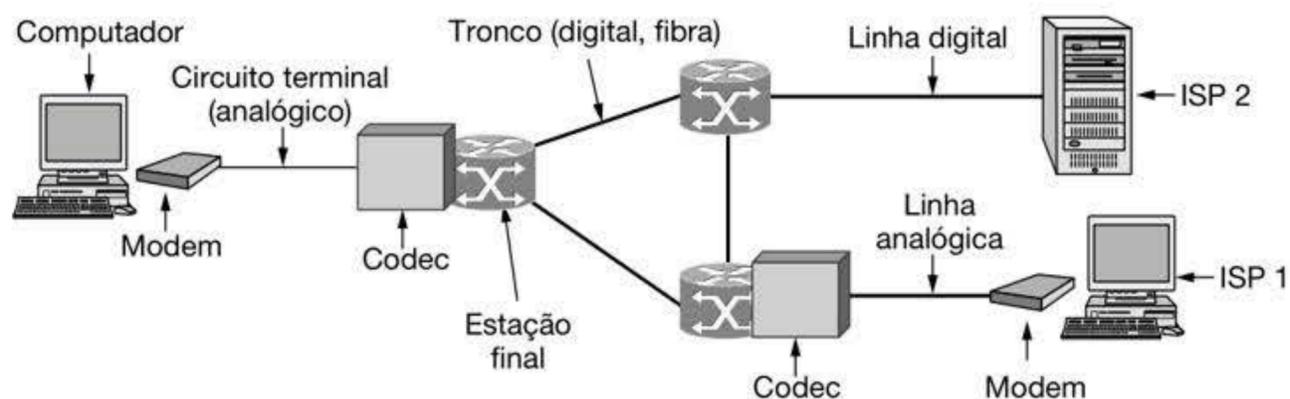
## Modems de telefone

A maioria das pessoas está familiarizada com o circuito terminal de dois fios, que vem da estação final de uma companhia telefônica até as residências. Com frequência, o circuito terminal também é chamado “o último quilômetro”, embora o comprimento real possa chegar a vários quilômetros. Muito esforço tem sido dedicado para espremer a rede de dados pelos circuitos terminais de cobre que já estão instalados. Os modems de telefone enviam dados digitais entre computadores pelo canal estreito que a rede telefônica oferece para uma chamada de voz. Eles já foram muito utilizados, mas foram substituídos em grande parte por tecnologias de banda larga, como ADSL, que reutilizam o circuito terminal para enviar dados digitais de um cliente para a estação final, onde são enviados para a Internet. Os modems e a ADSL precisam lidar com as limitações dos antigos circuitos terminais: largura de banda relativamente estreita, atenuação e distorção dos sinais, e suscetibilidade ao ruído elétrico, como linha cruzada.

Para enviar bits pelo circuito terminal, ou qualquer outro canal físico pelo mesmo motivo, eles precisam ser convertidos para sinais analógicos que podem ser transmitidos pelo canal. Essa conversão é realizada usando os métodos para modulação digital que estudamos na seção anterior. Na outra ponta do canal, o sinal analógico é convertido de volta para bits.

Um dispositivo que converte entre um fluxo de bits digitais e um sinal analógico que representa os bits é chamado **modem**, que é uma abreviação de “*modulador-demodulador*”. Os modems vêm em muitas variedades: modems de telefone, DSL, a cabo e sem fio. No caso de um modem a cabo ou DSL, o dispositivo geralmente é um hardware separado que fica entre a linha física entrando na casa e o restante da rede dentro da casa. Os dispositivos sem fio geralmente possuem seus próprios modems embutidos. Logicamente, o modem é inserido entre o computador (digital) e o sistema telefônico (analógico), como vemos na Figura 2.26.

Os modems de telefone são usados para enviar bits entre dois computadores por uma linha telefônica com



**Figura 2.26** O uso de transmissão analógica e digital para uma chamada de computador para computador. A conversão é feita por modems e codecs.

qualidade de voz, em lugar da conversação que normalmente preenche a linha. A principal dificuldade ao fazer isso é que uma linha de telefone para a voz humana é limitada a 3.100 Hz, o suficiente para transportar uma conversa. Essa largura de banda é mais de quatro ordens de grandeza menor que a largura de banda usada para Ethernet ou rede 802.11 (WiFi). Não é surpresa que as taxas de dados dos modems de telefone também sejam quatro ordens de grandeza menores que as da Ethernet e da rede 802.11.

Vamos analisar os números para ver por que isso acontece. O teorema de Nyquist nos diz que, mesmo com uma linha perfeita de 3.000 Hz (que uma linha telefônica decididamente não é), não há sentido em enviar símbolos em uma taxa mais rápida que 6.000 bauds. Por exemplo, vamos considerar um modem mais antigo enviando a uma taxa de 2.400 símbolos/s, ou 2.400 bauds, que captura vários bits por símbolo, enquanto permite o tráfego nas duas direções ao mesmo tempo (usando diferentes frequências para diferentes direções).

O humilde modem de 2.400 bps usa 0 volt para um valor lógico 0 e 1 volt para um valor 1, com 1 bit por símbolo. No próximo passo, ele pode usar quatro símbolos diferentes, como nas quatro fases do QPSK, de modo que, com 2 bits/símbolo, ele pode alcançar uma taxa de dados de 4.800 bps.

Uma longa progressão de taxas mais altas tem sido alcançada à medida que a tecnologia melhora. Taxas mais altas exigem um conjunto maior de símbolos (ver Figura 2.17). Com muitos símbolos, até mesmo uma pequena quantidade de ruído na amplitude ou fase detectada pode resultar em um erro. Para reduzir a chance de erros, os padrões para modems de velocidade mais alta utilizam alguns dos símbolos para correção de erros. Os esquemas são conhecidos como **modulação codificada por treliças**, ou **TCM (Trellis Coded Modulation)**. A Figura 2.27 mostra alguns padrões de modem comuns.

Por que parar em 33.600 bps? A razão para isso é que o limite de Shannon para o sistema telefônico é de aproximadamente 35 kbps, com base no comprimento médio dos circuitos terminais e na qualidade dessas linhas. Uma transmissão mais rápida que isso violaria as leis da física (departamento da termodinâmica) ou exigiria novos circuitos terminais (o que gradualmente está sendo feito).

Padrão de modem	Baud	Bits/símbolo	Bps
V.32	2.400	4	9.600
V.32 bis	2.400	6	14.400
V.34	2.400	12	28.800
V.34 bis	2.400	14	33.600

**Figura 2.27** Alguns padrões de modem e suas taxas de bits.

Contudo, existe um meio de mudar a situação. Na estação final da companhia telefônica, os dados são convertidos para a forma digital para transmissão dentro da rede telefônica (o núcleo da rede telefônica convertia de analógico para digital há muito tempo). O limite de 35 kbps é para a situação em que existem dois circuitos terminais, um em cada ponta. Cada um deles acrescenta ruído ao sinal. Se pudéssemos nos livrar de um desses circuitos terminais, aumentaríamos a SNR e a taxa máxima seria dobrada.

É com essa técnica que os modems de 56 kbps funcionam. Uma extremidade, normalmente um ISP, recebe uma entrada digital de alta qualidade da estação final mais próxima. Desse modo, quando uma extremidade da conexão é um sinal de alta qualidade, como ocorre com a maioria dos ISPs atuais, a taxa máxima de dados pode chegar a 70 kbps. Entre dois usuários domésticos com modems e linhas analógicas, o máximo ainda é 33,6 kbps.

A razão para o uso de modems de 56 kbps (em vez de modems de 70 kbps) está relacionada ao teorema de Nyquist. Um canal de telefonia é transportado dentro do sistema telefônico como amostras digitais. Cada canal telefônico tem cerca de 4.000 Hz de largura, incluindo as bandas de proteção. O número máximo de amostras independentes por segundo é, portanto, 8.000. O número de bits por amostra na América do Norte é 8, um dos quais pode ser usado para fins de controle, permitindo 56.000 bits/s de dados do usuário. Na Europa, todos os 8 bits estão disponíveis para os usuários, e assim poderiam ser usados modems de 64.000 bits/s; porém, para chegar a um acordo internacional sobre um padrão, o valor de 56.000 foi escolhido.

O resultado final são os modems padrão **V.90** e **V.92**. Eles são capazes de transmitir 56 kbps no canal downstream (ISP ao usuário) e 33,6 kbps e 48 kbps no canal upstream (usuário para ISP), respectivamente. A assimetria ocorre porque normalmente existem mais dados transportados do ISP para o usuário do que o contrário. Isso também significa que uma parte maior da largura de banda limitada pode ser alocada ao canal downstream, para aumentar as chances de que ele realmente funcione a 56 kbps.

## Digital Subscriber Lines (DSL)

Quando a indústria de telefonia finalmente conseguiu alcançar 56 kbps, ela se congratulou pelo serviço bem-feito. Enquanto isso, a indústria de TV a cabo estava oferecendo velocidades de até 10 Mbps sobre cabos compartilhados. À medida que o acesso à Internet se tornou uma parte cada vez mais importante de seus negócios, as companhias telefônicas locais começaram a perceber que precisavam de um produto mais competitivo. Sua resposta foi começar a oferecer novos serviços digitais ao circuito terminal do assinante.

Inicialmente, havia muitas ofertas sobrepostas, todas sob o nome genérico de **linha digital do assinante**, ou **xDSL (Digital Subscriber Line)**, para diversos *x*.

Os serviços com mais largura de banda do que o serviço telefônico padrão às vezes são chamados de **banda larga (broadband)**, embora o termo seja mais um conceito de marketing do que um conceito técnico específico. Mais adiante, vamos nos concentrar principalmente naquele que provavelmente se tornou o mais popular desses serviços, a **ADSL (Asymmetric DSL)**. Também usaremos o termo DSL ou xDSL como uma abreviação geral.

A razão para os modems serem tão lentos é que os telefones foram inventados para transportar a voz humana, e o sistema inteiro foi cuidadosamente otimizado para esse propósito. Os dados sempre estiveram em segundo plano. No ponto em que cada circuito terminal chega à estação final, o fio passa por um filtro que atenua todas as frequências abaixo de 300 Hz e acima de 3.400 Hz. O corte não é nítido – 300 Hz e 3.400 Hz são os pontos de 3 dB – assim, a largura de banda normalmente é mencionada como 4.000 Hz, embora a distância entre os pontos de 3 dB seja de 3.100 Hz. Portanto, os dados no fio também estão restritos a essa banda estreita.

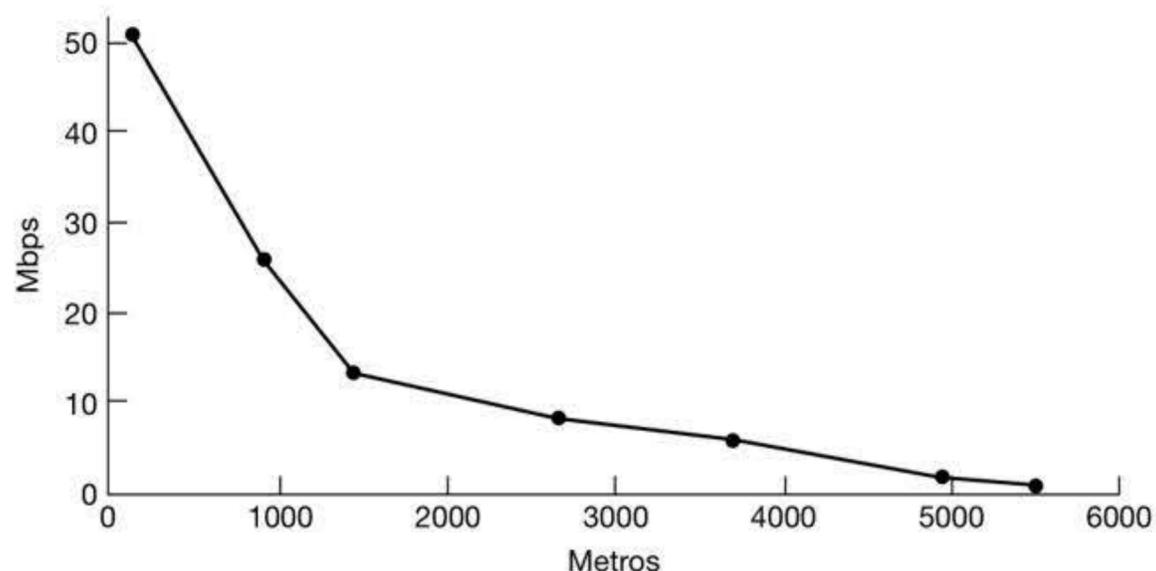
O artifício que faz o xDSL funcionar é o fato de, quando um cliente se inscreve nele, a linha de entrada ser conectada a um tipo diferente de switch, que não tem esse filtro, tornando disponível toda a capacidade do circuito terminal. Então, o fator limitador passa a ser a constituição física do circuito terminal, que aceita aproximadamente 1 MHz, e não a largura de banda artificial de 3.100 Hz criada pelo filtro.

Infelizmente, a capacidade do circuito terminal cai rapidamente com a distância da estação final, pois o sinal é degradado cada vez mais ao longo do fio. Ela também depende da espessura e da qualidade geral do par trançado. A Figura 2.28 mostra um esboço da largura de banda potencial como uma função da distância, pressupondo que todos os outros fatores estão otimizados (novos fios, pacotes de serviços, etc.).

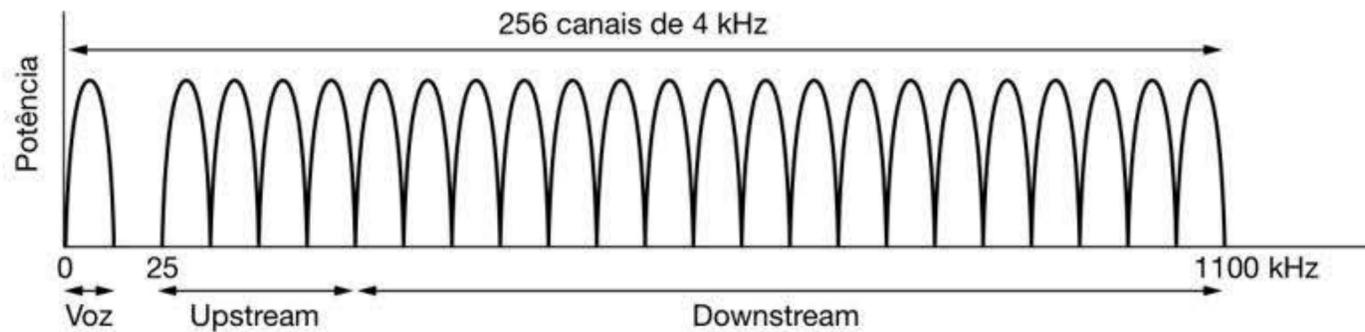
A implicação dessa figura cria um problema para a companhia telefônica. Quando escolhe uma velocidade para oferecer, ela está ao mesmo tempo escolhendo um raio a partir de suas estações finais, além do qual o serviço não poderá ser oferecido. Isso significa que, quando clientes distantes tentarem assinar o serviço, eles receberão a seguinte mensagem: “Muito obrigado por seu interesse, mas você está 100 m além da distância máxima da central mais próxima que poderia lhe oferecer o serviço. Você não gostaria de se mudar?”. Quanto mais baixa a velocidade escolhida, maior o raio e maior o número de clientes cobertos. Porém, quanto mais baixa a velocidade, menos atraente será o serviço e menor o número de pessoas que estarão dispostas a pagar por ele. É aqui que os negócios encontram a tecnologia.

Todos os serviços xDSL foram criados visando certos objetivos. Primeiro, os serviços devem funcionar nos circuitos terminais de pares trançados da Categoria 3 existente. Segundo, eles não devem afetar os telefones e os aparelhos de fax atuais dos clientes. Terceiro, eles devem ser muito mais rápidos que 56 kbps. Quarto, eles devem estar sempre ativos, apenas com uma tarifa mensal e nenhuma tarifa por minuto.

Para cumprir os objetivos técnicos, o espectro de 1,1 MHz disponível no circuito terminal é dividido em 256 canais independentes de 4.312,5 Hz cada um. Esse arranjo pode ser visto na Figura 2.29. O esquema OFDM, que vimos na seção anterior, é usado para enviar dados por esses canais, embora normalmente seja chamado **DMT (Discrete MultiTone)** no contexto da ADSL. O canal 0 é usado para o **POTS (Plain Old Telephone Service)**, e os canais de 1 a 5 não são usados, a fim de impedir que o sinal de voz e os sinais de dados interfiram uns com os outros. Dos 250 canais restantes, um é utilizado para o controle upstream e outro é empregado para o controle downstream. Os demais estão disponíveis para dados do usuário.



**Figura 2.28** Variação da largura de banda *versus* a distância sobre o UTP da categoria 3 para DSL.



**Figura 2.29** Operação ADSL usando modelagem discreta por multítono.

Em princípio, cada um dos canais restantes pode ser usado em um fluxo de dados full-duplex; porém, harmônicos, linhas cruzadas e outros efeitos mantêm a utilização de sistemas práticos bem abaixo do limite teórico. Cabe ao provedor definir quantos canais serão usados para upstream e quantos serão usados para downstream. Uma mistura de 50% para cada um é tecnicamente possível, mas a maioria dos provedores aloca algo como 80 a 90% da largura de banda ao canal downstream, pois a maioria dos usuários faz mais download do que upload de dados. Essa escolha deu origem à letra “A” (assimétrica) no acrônimo ADSL. Uma divisão comum reserva 32 canais para upstream e os restantes para downstream. Também é possível tornar bidirecionais alguns dos canais upstream mais altos para aumentar a largura de banda, embora essa otimização exija o uso de um circuito especial para cancelamento de ecos.

O padrão ADSL internacional, conhecido como **G.dmt**, foi aprovado em 1999 e permite velocidades de até 8 Mbps downstream e 1 Mbps upstream. Ele foi substituído por uma segunda geração em 2002, chamada ADSL2, com diversas melhorias para permitir velocidades de até 12 Mbps downstream e 1 Mbps upstream. ADSL2+ dobra a velocidade downstream para 24 Mbps ao dobrar a largura de banda para usar 2,2 MHz pelo par trançado.

Em 2006 ocorreu a melhoria seguinte, o **VDSL**, que empurrou a taxa de dados nos circuitos terminais mais curtos para 52 Mbps downstream e 3 Mbps upstream. Então, de 2007 a 2011, uma série de novos padrões, sob o nome de **VDSL2**, sobre circuitos terminais de alta qualidade, conseguiram usar largura de banda de 12 MHz e atingir taxas de dados de 200 Mbps downstream e 100 Mbps upstream. Em 2015, o **Vplus** foi proposto para circuitos terminais menores que 250 m. Em princípio, ele pode atingir 300 Mbps downstream e 100 Mbps upstream, mas não é fácil fazê-lo funcionar na prática. Podemos estar perto do fim da linha aqui para a fiação Categoria 3 existente, exceto talvez para distâncias ainda mais curtas.

Dentro de cada canal, a modulação QAM é usada a uma taxa de aproximadamente 4.000 símbolos/s. A qualidade da linha em cada canal é monitorada constantemente, e a taxa de dados é ajustada usando um diagrama de constelação maior ou menor, como os da Figura 2.17. Diferentes canais podem ter diferentes taxas de dados, com até 15 bits por símbolo enviado em um canal com uma SNR alta, e

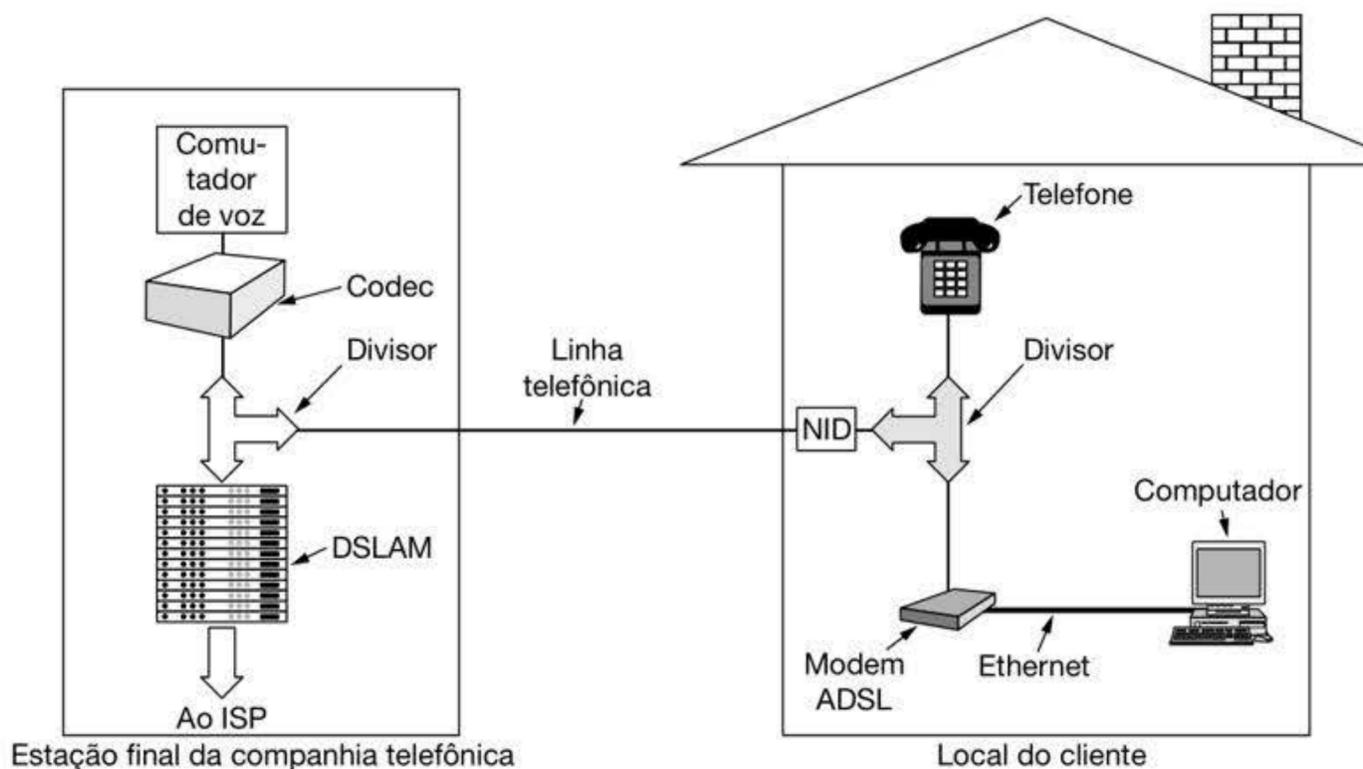
descendo até 2, 1 ou nenhum bit por símbolo enviado em um canal com uma SNR baixa, dependendo do padrão.

A Figura 2.30 mostra a organização de uma ADSL típica. Nesse esquema, um técnico da companhia telefônica deve instalar um **dispositivo de interface de rede**, ou **NID (Network Interface Device)** no local do cliente. Essa pequena caixa plástica marca o fim da propriedade da companhia telefônica e o início da propriedade do cliente. Próximo ao NID (ou às vezes combinado a ele) há um **divisor**, um filtro analógico que separa a banda de 0 a 4.000 Hz utilizada pelo POTS dos dados. O sinal do POTS é roteado até o telefone ou o equipamento de fax existente, e o sinal de dados é roteado até um modem ADSL, que usa o processamento do sinal digital para implementar OFDM. Como a maioria dos modems ADSL atuais é externa, o computador deve estar conectado ao modem em alta velocidade. Normalmente, isso é feito usando Ethernet, cabo USB ou rede 802.11.

Na outra extremidade do fio, ao lado da estação final, está instalado um divisor correspondente. Aqui, a porção de voz do sinal é filtrada e enviada ao switch de voz normal. O sinal acima de 26 kHz é roteado para um novo tipo de dispositivo, chamado **multiplexador de acesso à linha digital do assinante**, ou **DSLAM (Digital Subscriber Line Access Multiplexer)**, que contém a mesma espécie de processador de sinal digital que o modem ADSL. O DSLAM converte o sinal em bits e envia pacotes para a rede de dados do provedor de serviços da Internet.

Essa separação completa entre o sistema de voz e a ADSL torna relativamente fácil para uma companhia telefônica distribuir esse serviço. Basta adquirir um DSLAM e um divisor, e conectar os assinantes da ADSL ao divisor. Outros serviços de alta largura de banda entregues pela rede telefônica (p. ex., ISDN) exigem mudanças muito maiores no equipamento de comutação existente.

A próxima fronteira para a implantação da DSL é atingir velocidades de transmissão de 1 Gbps ou mais. Esses esforços estão se concentrando em uma série de técnicas complementares, incluindo uma técnica chamada **bonding** (ou **união**), que cria uma única conexão DSL virtual combinando duas ou mais conexões DSL físicas. Obviamente, se combinarmos dois pares trançados, será possível dobrar a largura de banda. Em alguns lugares, os fios telefônicos que entram nas casas usam um cabo que, na verdade, tem dois



**Figura 2.30** Configuração típica de equipamento ADSL.

pares trançados. A ideia original era permitir duas linhas e números de telefone separados na casa, mas usando a união dos pares, uma única conexão de Internet de alta velocidade pode ser alcançada. Um número cada vez maior de ISPs na Europa, Austrália, Canadá e Estados Unidos já está implantando uma tecnologia chamada **G.fast**, que usa união de pares. Tal como acontece com outras formas de DSL, o desempenho do G.fast depende da distância da transmissão; testes recentes viram velocidades simétricas se aproximando de 1 Gbps a distâncias de 100 m, quando acoplado a uma implantação de fibra conhecida como **FTTdp (Fiber to the Distribution Point)**, que leva a fibra a um ponto de distribuição de várias centenas de assinantes e usa cobre para transmitir dados pelo restante do caminho para a casa (em VDSL2, isso pode chegar a 1 km, embora a velocidades mais baixas). FTTdp é apenas um tipo de implantação de fibra que leva a fibra do núcleo da rede a algum ponto próximo à extremidade da rede. A próxima seção descreve diversos modos de implantação de fibra.

### Fiber To The X (FTTX)

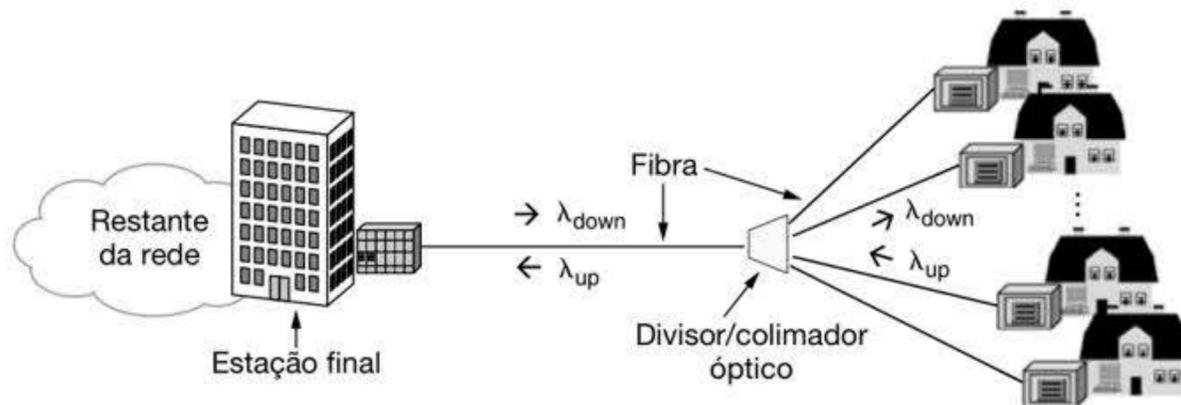
A velocidade dos circuitos terminais normalmente é restrita pelos cabos de cobre instalados na rede de telefonia convencional, que não podem transmitir dados em altas velocidades por uma distância tão grande quanto a fibra. Assim, um objetivo final, onde for economicamente viável, é levar a fibra até a casa do cliente, algo chamado de **FTTH (Fiber To The Home)**. As companhias telefônicas continuam a tentar melhorar o desempenho do circuito terminal, geralmente instalando a fibra o mais próximo possível da casa. Se não for até ela, eles podem oferecer **FTTN (Fiber to the Node)**, onde a fibra termina em um gabinete em uma rua, às vezes a alguns quilômetros da casa do cliente. Fiber

to the Distribution Point (FTTdp), como já dissemos, leva a fibra um passo a mais para a casa do cliente, geralmente levando-a para alguns metros do local desejado. Entre essas duas opções há o **FTTC (Fiber to the Curb)**. Todos esses projetos **FTTX (Fiber to the X)** às vezes são chamados de “fibra no terminal”, pois é usada uma certa quantidade de fibra no circuito terminal.

Existem diversas variações na forma “FTTX” (onde *X* significa porão, calçada ou vizinhança). Elas são usadas para indicar que a implantação da fibra pode chegar perto da casa. Nesse caso, o cobre (par trançado ou cabo coaxial) oferece velocidades rápidas o suficiente pela última distância curta. A escolha da extensão em que a fibra é disposta é uma questão econômica, balanceando custo com receita esperada. De qualquer forma, o importante é que a fibra óptica atravessou a barreira tradicional do “último quilômetro”. Focaremos o FTTH em nossa discussão.

Assim como os fios de cobre antes dele, o circuito terminal de fibra é passivo. Isso significa que nenhum equipamento energizado é necessário para amplificar ou processar os sinais de alguma outra forma. A fibra simplesmente transporta sinais entre a casa e a estação final. Isso, por sua vez, reduz o custo e melhora a confiabilidade. Normalmente, as fibras das casas são reunidas de modo que apenas uma fibra alcance a estação final por grupo de até 100 casas. Na direção downstream, os divisores ópticos dividem o sinal da estação final, de modo que alcance todas as casas. A criptografia é necessária por segurança se apenas uma casa puder ser capaz de decodificar o sinal. Na direção upstream, colimadores ópticos mesclam os sinais das casas para um único sinal, que é recebido na estação final.

Essa arquitetura é chamada **rede óptica passiva**, ou **PON (Passive Optical Network)**, mostrada na Figura 2.31. É comum usar um comprimento de onda compartilhado



**Figura 2.31** Rede óptica passiva para Fiber To The Home.

entre todas as casas para a transmissão downstream, e outro comprimento de onda para a transmissão upstream.

Mesmo com a divisão, uma enorme largura de banda e a baixa atenuação da fibra significam que as PONs podem oferecer altas velocidades aos usuários por distâncias de até 20 km. As taxas de dados reais e outros detalhes dependem do tipo de PON. Dois tipos são comuns: as **GPONs (Gigabit-capable PONs)** vêm do mundo das telecomunicações e são definidas por um padrão ITU; as **EPONs (Ethernet PONs)** estão mais ligadas ao mundo das redes e são definidas por um padrão IEEE. Ambas trabalham em torno de um gigabit e podem transportar tráfego para diferentes serviços, incluindo Internet, vídeo e voz. Por exemplo, as GPONs oferecem 2,4 Gbps downstream e 1,2 ou 2,4 Gbps upstream.

Para compartilhar a capacidade de uma única fibra na estação final entre diferentes casas, é preciso que haja protocolos adicionais. A direção downstream é fácil, pois a estação final pode enviar mensagens a cada casa diferente na ordem que desejar. Na direção upstream, porém, as mensagens de diferentes casas não podem ser enviadas ao mesmo tempo, ou diferentes sinais colidiriam. As casas também não podem escutar as transmissões umas das outras, de modo que não podem escutar antes de transmitir. A solução é que o equipamento nas casas solicite e receba fatias de tempo para utilizar o equipamento na estação final. Para que isso funcione, existe um processo de localização para ajustar os tempos de transmissão a partir das casas, de modo que todos os sinais recebidos na estação final sejam sincronizados. O projeto é semelhante aos modems a cabo, que explicaremos mais adiante neste capítulo. Para obter mais informações sobre PONs, veja Grobe e Elbers (2008) ou Andrade et al. (2014).

### 2.5.3 Troncos e multiplexação

Na rede telefônica, os troncos não são apenas muito mais rápidos do que os circuitos terminais; eles são diferentes em dois outros aspectos. O núcleo da rede telefônica transporta informações digitais, não informações analógicas, ou seja, bits e não voz. Isso necessita de uma conversão na estação final para a forma digital, para a transmissão por troncos de longa distância. Os troncos transportam milhares, ou mesmo milhões, de chamadas simultaneamente. Esse

compartilhamento é importante para economizar escalas, pois custa basicamente a mesma coisa instalar e manter um tronco de alta largura de banda e um de baixa largura de banda entre duas estações de comutação. Isso é realizado com as versões de multiplexação TDM e FDM.

A seguir, examinaremos rapidamente como os sinais de voz são digitalizados de modo que possam ser transportados pela rede telefônica. Depois, veremos como a TDM é usada para transportar bits nos troncos, incluindo o sistema TDM usado para fibra óptica (SONET). Em seguida, passaremos para FDM conforme se aplica à fibra óptica, chamada de WDM.

### Sinais de voz digitalizados

Desde cedo no desenvolvimento da rede telefônica, o núcleo lidava com chamadas de voz como informação analógica. As técnicas de FDM foram usadas por muitos anos para multiplexar canais de voz de 4.000 Hz (compostos de 3.100 Hz mais bandas de proteção) em unidades cada vez maiores. Por exemplo, 12 chamadas na banda de 60 a 108 kHz são conhecidas como um **grupo**, cinco grupos (um total de 60 chamadas) são conhecidos como um **supergrupo**, e assim por diante. Esses métodos de FDM ainda são usados em alguns canais de fio de cobre e micro-ondas. Contudo, a FDM requer circuitos analógicos e isso não é adequado para um computador digital. Ao contrário, a TDM pode ser tratada inteiramente pela eletrônica digital, de modo que se tornou muito mais usada nos últimos anos. Como a TDM só pode ser usada para dados digitais e os circuitos terminais produzem sinais analógicos, é preciso que haja uma conversão de analógico para digital na estação final, onde todos os circuitos terminais individuais se juntam para serem combinados nos troncos de saída.

Os sinais analógicos são digitalizados na estação final por um dispositivo chamado **codec** (abreviação de “codificador-decodificador”) usando uma técnica chamada **PCM (Pulse Code Modulation)**, que forma o coração do sistema telefônico moderno. O codec cria 8.000 amostras por segundo (125  $\mu$ s/amostra), pois o teorema de Nyquist diz que isso é suficiente para capturar todas as informações de largura de banda do canal telefônico de 4 kHz. Em uma taxa de amostragem mais baixa, as informações seriam perdidas; a uma

taxa mais alta, nenhuma informação extra seria obtida. Cada amostra da amplitude do sinal é quantizada para um número de 8 bits. Quase todos os intervalos de tempo no sistema telefônico são múltiplos de  $125 \mu\text{s}$ . A taxa de dados não compactada padrão para uma chamada telefônica com qualidade de voz é, portanto, de 8 bits a cada  $125 \mu\text{s}$ , ou 64 kbps.

Cada amostra da amplitude do sinal é quantizada para um número de 8 bits. Para reduzir o erro em virtude da quantização, seus níveis são espaçados de forma desigual. Uma escala logarítmica é usada, gerando relativamente mais bits para menores amplitudes de sinal e relativamente menos bits para grandes amplitudes de sinal. Desse modo, o erro é proporcional à amplitude do sinal. Duas versões de quantização são bastante utilizadas:  $\mu$ -law, usada na América do Norte e no Japão, e A-law, usada na Europa e no restante do mundo. As duas versões são especificadas no padrão ITU G.711. Um modo equivalente de pensar nesse processo é imaginar que a faixa dinâmica do sinal (ou a razão entre os maiores e menores valores possíveis) é compactada antes de ser quantizada (uniformemente), e então expandida quando o sinal analógico é recriado. Por esse motivo, ela é chamada **compactação/expansão**, ou **companding**. Também é possível compactar as amostras depois que elas são digitalizadas, de modo que exigem muito menos do que 64 kbps. Contudo, deixaremos esse assunto para quando explorarmos aplicações de áudio como VoIP.

Na outra extremidade da chamada, um sinal analógico é recriado a partir das amostras quantizadas, reproduzindo-as (e suavizando-as) ao longo do tempo. Não será exatamente igual ao sinal analógico original, embora o tenhamos amostrado na taxa de Nyquist, pois as amostras foram quantizadas.

### T-Carrier: multiplexação de sinais digitais pela rede de telefonia

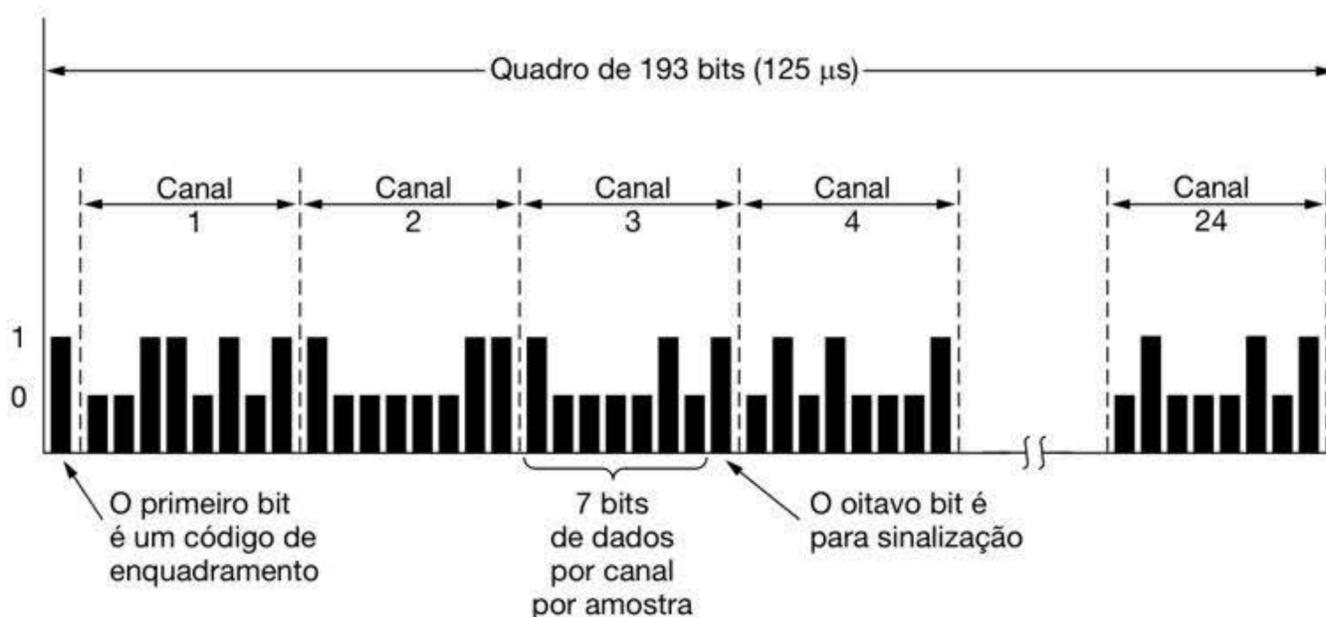
**T-Carrier** é uma especificação para a transmissão de vários canais TDM por um único circuito. A TDM baseada

na PCM é usada para transportar várias chamadas de voz por troncos, enviando uma amostra de cada chamada a cada  $125 \mu\text{s}$ . Quando a transmissão digital começou a surgir como uma tecnologia viável, a ITU (então chamada CCITT) foi incapaz de chegar a um acordo sobre um padrão internacional para a PCM. Consequentemente, diversos esquemas incompatíveis agora estão em uso em diferentes países.

O método usado na América do Norte e no Japão é a portadora **T1**, representada na Figura 2.32. (Técnicamente falando, o formato é chamado DS1 e a portadora é chamada T1, mas, seguindo a tradição da indústria, aqui não faremos essa distinção sutil.) A portadora T1, introduzida em 1962, consiste em 24 canais de voz multiplexados. Por sua vez, cada um dos 24 canais consegue inserir 8 bits no fluxo de saída.

Um quadro consiste em  $24 \times 8 = 192$  bits, mais um bit extra para fins de controle, produzindo 193 bits a cada  $125 \mu\text{s}$ . Isso resulta em uma taxa de dados bruta de 1,544 Mbps, dos quais 8 kbps são para sinalização. O 193º bit é usado para sincronização e sinalização de quadros. Em uma variação, o 193º bit é usado por um grupo de 24 quadros, chamado **superquadro estendido**. Seis dos bits, na 4ª, 8ª, 12ª, 16ª, 20ª e 24ª posições, utilizam o padrão 001011... Normalmente, o receptor continua a conferir esse bit para garantir que não perdeu a sincronização. Seis outros bits são usados para enviar um código de verificação de erro, para ajudar o receptor a confirmar se está sincronizado. Se sair de sincronismo, o receptor poderá procurar por esse padrão e validar o código de verificação de erro para se resincronizar. Os 12 bits restantes são usados para informação de controle, para operar e manter a rede, como o relatório de desempenho da extremidade remota.

O formato T1 tem diversas variações. As versões mais antigas enviavam informações de sinalização **na própria banda**, significando o uso do mesmo canal que os dados, usando alguns bits de dados. Esse projeto é uma



**Figura 2.32** A portadora T1 (1,544 Mbps).

forma de **sinalização associada ao canal**, pois cada canal tem seu próprio subcanal de sinalização privado. Em um arranjo de bits, o bit menos significativo de uma amostra de 8 em cada canal é usado a cada seis quadros. Este recebe o nome sugestivo de **bit de sinalização roubado**. A ideia é que alguns bits roubados não influenciarão nas chamadas de voz. Ninguém notará qualquer diferença audível.

Contudo, para dados, a história é outra. Entregar os bits errados é, no mínimo, inútil. Se versões mais antigas da T1 forem usadas para transportar dados, somente 7 de 8 bits, ou 56 kbps, podem ser usados em cada um dos 24 canais. Em vez disso, versões mais novas da T1 oferecem canais limpos, em que todos os bits podem ser usados para enviar dados. Canais limpos são o que as empresas que alugam uma linha T1 desejam quando enviam dados pela rede telefônica no lugar das amostras de voz. A sinalização para quaisquer chamadas de voz é, então, tratada **fora da banda**, significando um canal próprio separado dos dados. Normalmente, a sinalização é feita com **sinalização de canal comum**, no qual existe um canal de sinalização compartilhado. Um dos 24 canais pode ser usado para essa finalidade.

Fora da América do Norte e do Japão, a portadora E1 de 2,048 Mbps é usada no lugar da T1. Essa portadora tem 32 amostras de dados de 8 bits compactadas no quadro básico de 125 μs. Trinta dos canais são usados para informações e até dois são usados para sinalização. Cada grupo de quatro quadros oferece 64 bits de sinalização, metade deles usada para sinalização (associada ao canal ou ao canal comum) e metade usada para sincronização de quadros ou reservada para cada país usar como preferir.

A TDM permite que várias portadoras T1 sejam multiplexadas em portadoras de ordem mais alta. A Figura 2.33 mostra como isso pode ser feito. À esquerda, vemos quatro canais T1 sendo multiplexados em um canal T2. A multiplexação em T2 e acima é feita bit a bit, e não byte a byte, com os 24 canais de voz que compõem um quadro T1. Quatro fluxos T1 a uma velocidade de 1,544 Mbps deveriam gerar 6,176 Mbps, mas T2, na verdade, tem 6,312 Mbps. Os bits extras são usados para enquadramento e recuperação, no caso de a portadora apresentar alguma falha.

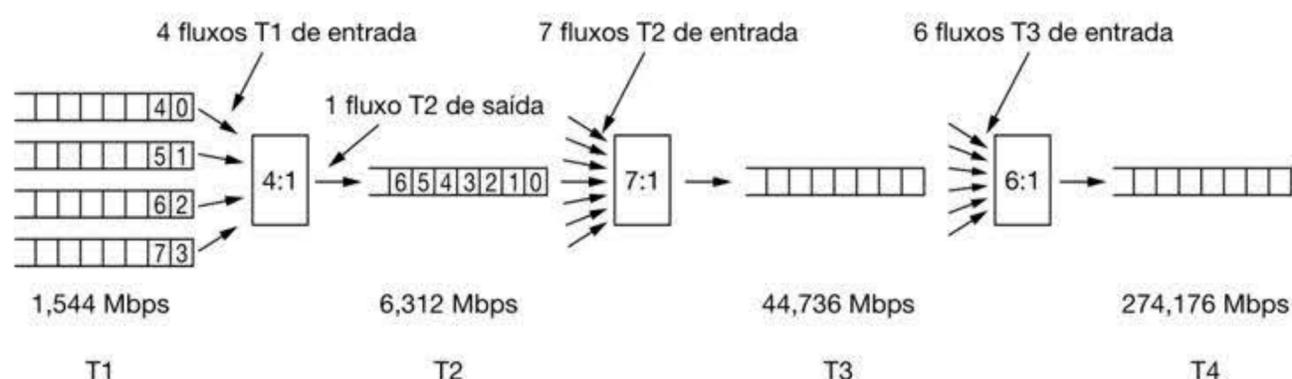
No nível seguinte, sete fluxos T2 são combinados bit a bit para formar um fluxo T3. Depois, seis fluxos T3 são unidos para formar um fluxo T4. Em cada etapa, um pequeno volume de overhead é adicionado para fins de enquadramento e recuperação, no caso de a sincronização entre transmissor e receptor ser perdida. T1 e T3 são extensamente utilizados pelos clientes, enquanto T2 e T4 são usados apenas dentro do sistema de telefonia propriamente dito e, portanto, não são bem conhecidos.

Da mesma forma que existe pouco consenso quanto à portadora básica entre os Estados Unidos e o restante do mundo, há igualmente pouco consenso sobre como ela será multiplexada em portadoras de largura de banda mais alta. O esquema norte-americano de avançar por 4, 7 e 6 não foi adotado por mais ninguém; assim, o padrão ITU requer multiplexação de quatro fluxos em um fluxo a cada nível. Além disso, o enquadramento e a recuperação de dados são diferentes entre os padrões dos Estados Unidos e da ITU. A hierarquia ITU para 32, 128, 512, 2.048 e 8.192 canais funciona em velocidades de 2,048, 8,848, 34,304, 139,264 e 565,148 Mbps.

### Multiplexação de redes ópticas: SONET/SDH

Nos primórdios da fibra óptica, cada companhia telefônica tinha seu próprio sistema óptico TDM patenteado. Depois que o governo dos Estados Unidos desmembrou a AT&T em 1984, as companhias telefônicas locais tiveram de se conectar a diversas concessionárias de comunicações de longa distância, todas com sistemas ópticos TDM de diversos vendedores e fornecedores, o que tornou óbvia a necessidade de padronização. Em 1985, a Bellcore, a unidade de pesquisa das RBOCs (Regional Bell Operating Companies), começou a trabalhar em um padrão, denominado **rede óptica síncrona**, ou **SONET (Synchronous Optical Network)**.

Mais tarde, a ITU também começou a participar desse trabalho, o que resultou em um padrão SONET e em um conjunto de recomendações paralelas (G.707, G.708 e G.709) em 1989. As recomendações da ITU são chamadas de **hierarquia digital síncrona**, ou **SDH (Synchronous Digital Hierarchy)**, mas diferem da SONET apenas em pequenos detalhes. Praticamente todo o tráfego telefônico



**Figura 2.33** Multiplexação de fluxos T1 em portadoras de velocidade mais alta.

de longa distância nos Estados Unidos e grande parte dele em outros lugares atualmente utiliza troncos que executam a SONET na camada física. Para obter informações adicionais sobre a SONET, consulte Perros (2005).

O projeto SONET tem quatro objetivos principais.

1. Interoperabilidade de portadora: a SONET tinha de tornar possível a rede interligada para diferentes concessionárias. A concretização desse objetivo exigia a definição de um padrão de sinalização comum, relacionado a comprimento de onda, sincronização, estrutura de enquadramento e outras questões.
2. Unificação pelas regiões: foram necessários alguns meios para unificar os sistemas digitais dos Estados Unidos, Europa e Japão, todos baseados em canais PCM de 64 kbps, mas combinados de formas diferentes (e incompatíveis).
3. Multiplexação de canais digitais: a SONET teve de proporcionar um modo de multiplexar vários canais digitais. No momento em que ela surgiu, a portadora digital de velocidade mais alta usada em todo o território dos Estados Unidos era a T3, a 44,736 Mbps. A T4 já havia sido definida, mas não era muito usada, e nada que ultrapassasse a velocidade da T4 havia sido definido. Parte da missão da SONET era dar continuidade à hierarquia até gigabits/s e proporcionar velocidades ainda maiores. Também era necessária uma forma padrão de multiplexar canais mais lentos em um canal SONET.
4. Suporte para gerenciamento: a SONET tinha de oferecer recursos de operação, administração e manutenção necessários para gerenciar a rede. Os sistemas anteriores não faziam isso muito bem.

Uma decisão inicial foi tornar a SONET um sistema TDM tradicional, com toda a largura de banda da fibra

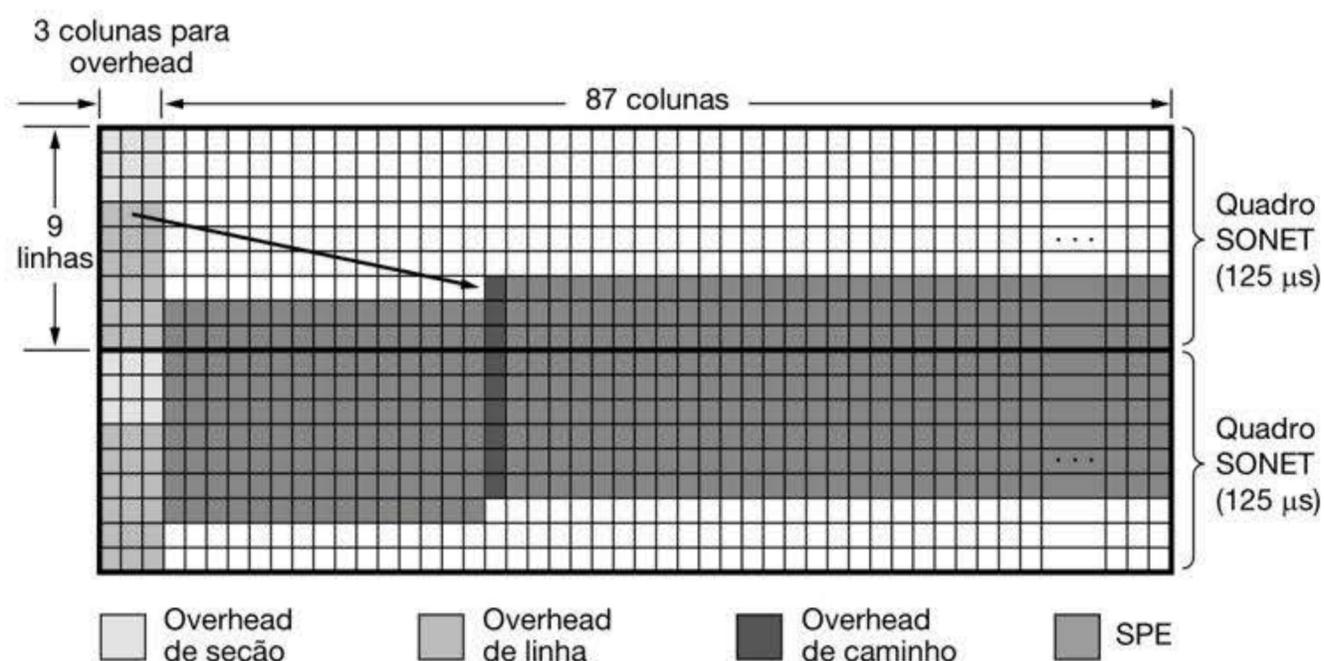
dedicada a um único canal contendo slots de tempo para os diversos subcanais. Portanto, a SONET é um sistema síncrono. Cada transmissor e receptor é ligado a um clock comum. O relógio mestre, que controla o sistema, tem uma precisão de aproximadamente uma parte em  $10^9$ . Os bits em uma linha SONET são transmitidos a intervalos extremamente precisos, controlados pelo clock mestre.

O quadro básico da SONET é um bloco de 810 bytes, transmitido a cada  $125 \mu\text{s}$ . Tendo em vista que a SONET é síncrona, os quadros são emitidos independentemente de haver ou não dados úteis a enviar. A taxa de 8.000 quadros/s corresponde exatamente à taxa de amostragem dos canais PCM utilizados em todos os sistemas de telefonia digital.

Os quadros de 810 bytes da SONET são mais bem descritos como um retângulo de bytes, com 90 colunas de largura por 9 linhas de altura. Desse modo,  $8 \times 810 = 6.480$  bits são transmitidos 8 mil vezes por segundo, o que resulta em uma taxa de dados bruta de 51,84 Mbps. Esse é o canal básico da SONET, chamado **STS-1 (Synchronous Transport Signal-1)**. Todos os troncos SONET são múltiplos do STS-1.

As três primeiras colunas de cada quadro são reservadas para as informações de gerenciamento do sistema, conforme ilustra a Figura 2.34. Nesse bloco, as três primeiras linhas contêm o overhead de seção, e as seis linhas seguintes contêm o overhead de linha. O overhead de seção é gerado e verificado no início e no fim de cada seção, enquanto o overhead de linha é gerado e verificado no início e no fim de cada linha.

Um transmissor SONET transmite quadros de 810 bytes em sequência, sem intervalos entre eles, mesmo quando não existem dados (e, nesse caso, ele transmite dados fictícios). Do ponto de vista do receptor, tudo o que ele vê é um fluxo de bits contínuo; assim, como saber onde começa cada quadro? A resposta é que os dois



**Figura 2.34** Dois quadros duplos na rede SONET.

primeiros bytes de cada quadro contêm um padrão fixo que o receptor procura. Se encontra esse padrão no mesmo lugar em um número grande de quadros consecutivos, o receptor pressupõe que está sincronizado com o transmissor. Em teoria, um usuário poderia inserir esse padrão como carga útil de uma maneira simples; porém, na prática, isso não pode ser feito em virtude da multiplexação usada por diversos usuários no mesmo quadro, além de outras razões.

As 87 colunas restantes de cada quadro contêm  $87 \times 9 \times 8 \times 8.000 = 50.112$  Mbps de dados do usuário, os quais poderiam ser amostras de voz, T1 e outras portadoras, ou pacotes. A SONET é simplesmente um contêiner conveniente para transportar bits. O **envelope síncrono de carga útil**, ou **SPE (Synchronous Payload Envelope)**, que transporta os dados do usuário, nem sempre começa na linha 1, coluna 4. O SPE pode começar em qualquer lugar do quadro. Um ponteiro para o primeiro byte está contido na primeira fileira do overhead de linha. A primeira coluna do SPE é o overhead de caminho (ou seja, o cabeçalho do protocolo da subcamada de caminho ponta a ponta).

A capacidade de permitir que o SPE comece em qualquer lugar dentro do quadro SONET e até mesmo se espalhe por dois quadros, como mostra a Figura 2.34, oferece flexibilidade adicional ao sistema. Por exemplo, se uma carga útil chega na origem enquanto um quadro SONET fictício está sendo construído, ele pode ser inserido no quadro atual em vez de ser mantido até o início do próximo.

A hierarquia de multiplexação da SONET/SDH é mostrada na Figura 2.35. Foram definidas taxas de STS-1 a STS-768, variando de aproximadamente uma linha T3 até 40 Gbps. Até mesmo taxas maiores certamente serão definidas com o tempo, com OC-3072 a 160 Gbps sendo o próximo na fila, se e quando isso se tornar tecnologicamente viável. A portadora óptica que corresponde a STS- $n$  é chamada OC- $n$ , porém sua configuração bit a bit é a mesma, exceto por uma certa reordenação de bits, necessária para sincronização. Os nomes da SDH são diferentes,

começando em OC-3 porque os sistemas baseados na ITU não possuem uma taxa próxima a 51,84 Mbps. Mostramos as taxas comuns, que desenvolvem-se a partir de OC-3 em múltiplos de quatro. A taxa de dados bruta inclui todo o overhead. A taxa de dados SPE exclui o overhead de linha e o de seção. A taxa de dados do usuário exclui todos os três tipos de overhead e só considera as 86 colunas de carga útil.

Em contrapartida, quando uma portadora como a OC-3 não é multiplexada, mas transporta os dados de uma única origem, a letra *c* (significando concatenado) é acrescentada à designação; assim, OC-3 indica uma portadora de 155,52 Mbps composta por três portadoras OC-1 distintas, mas OC-3c indica um fluxo de dados de uma única origem a uma velocidade de 155,52 Mbps. Os três fluxos OC-1 contidos em um fluxo OC-3c são entrelaçados por coluna: primeiro, a coluna 1 do fluxo 1, depois a coluna 1 do fluxo 2, a coluna 1 do fluxo 3, seguida pela coluna 2 do fluxo 1, e assim por diante, resultando em um quadro com 270 colunas de largura e 9 linhas de profundidade.

## 2.5.4 Comutação

Do ponto de vista do engenheiro de telefonia, o sistema telefônico é dividido em duas partes principais: a planta externa (os circuitos terminais e troncos, pois eles estão localizados fisicamente fora das estações de comutação) e a planta interna (os switches, que estão dentro das estações de comutação). Acabamos de estudar a planta externa. Agora vamos examinar a planta interna.

Hoje em dia, duas técnicas de comutação diferentes são usadas pela rede: comutação de circuitos e comutação de pacotes. O sistema telefônico tradicional é baseado na comutação de circuitos, embora a tecnologia VoIP conte com a comutação de pacotes. Veremos em detalhes a comutação de circuitos, comparando-a com a comutação de pacotes. Os dois tipos de comutação são tão importantes que voltaremos a eles quando entrarmos na camada de rede.

SONET		SDH	Taxa de dados (Mbps)		
Elétrico	Óptico	Óptico	Bruto	SPE	Usuário
STS-1	OC-1		51,84	50,112	49,536
STS-3	OC-3	STM-1	155,52	150,336	148,608
STS-12	OC-12	STM-4	622,08	601,344	594,432
STS-48	OC-48	STM-16	2.488,32	2.405,376	2.377,728
STS-192	OC-192	STM-64	9.953,28	9.621,504	9.510,912
STS-768	OC-768	STM-256	39.813,12	38.486,016	38.043,648

**Figura 2.35** Taxas de multiplexação da SONET e da SDH.

## Comutação de circuitos

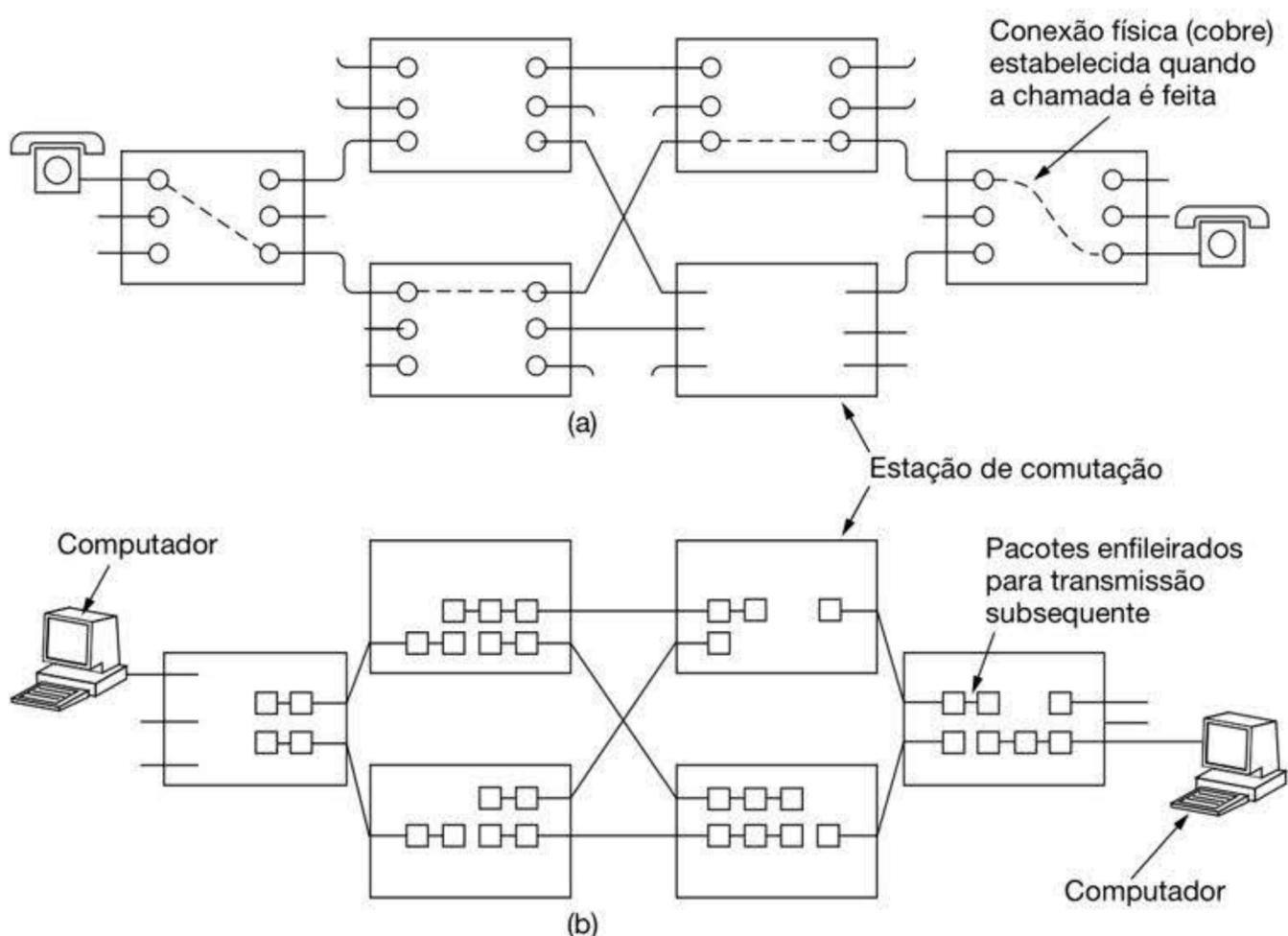
Tradicionalmente, quando você ou seu computador efetuavam uma chamada telefônica, o equipamento de comutação do sistema telefônico procurava um caminho físico desde o seu telefone até o telefone do receptor e o mantinha pela duração da chamada. Essa técnica, chamada **comutação de circuitos**, é apresentada esquematicamente na Figura 2.36(a). Cada um dos seis retângulos representa uma estação de comutação da concessionária de comunicações (estação final, estação interurbana, etc.). Nesse exemplo, cada estação tem três linhas de entrada e três linhas de saída. Quando uma chamada passa por uma estação de comutação, é estabelecida uma conexão física entre a linha que transportou a chamada e uma das linhas de saída, como mostram as linhas tracejadas.

No início da telefonia, a conexão era feita pela telefonista que conectava um cabo de ligação em ponte (jumper) aos soquetes de entrada e de saída. Na realidade, existe uma pequena história surpreendente associada à invenção do equipamento automático de comutação de circuitos. Esse dispositivo foi inventado por um agente funerário do século XIX, Almon B. Strowger. Logo depois que o telefone foi inventado, quando uma pessoa morria, alguém ligava para a telefonista do local e dizia: “Por favor, ligue-me com um agente funerário”. Infelizmente para o sr. Strowger, havia dois agentes funerários em sua cidade, e a esposa do outro agente era a telefonista do local. Ele percebeu rapidamente que teria de inventar um equipamento automático

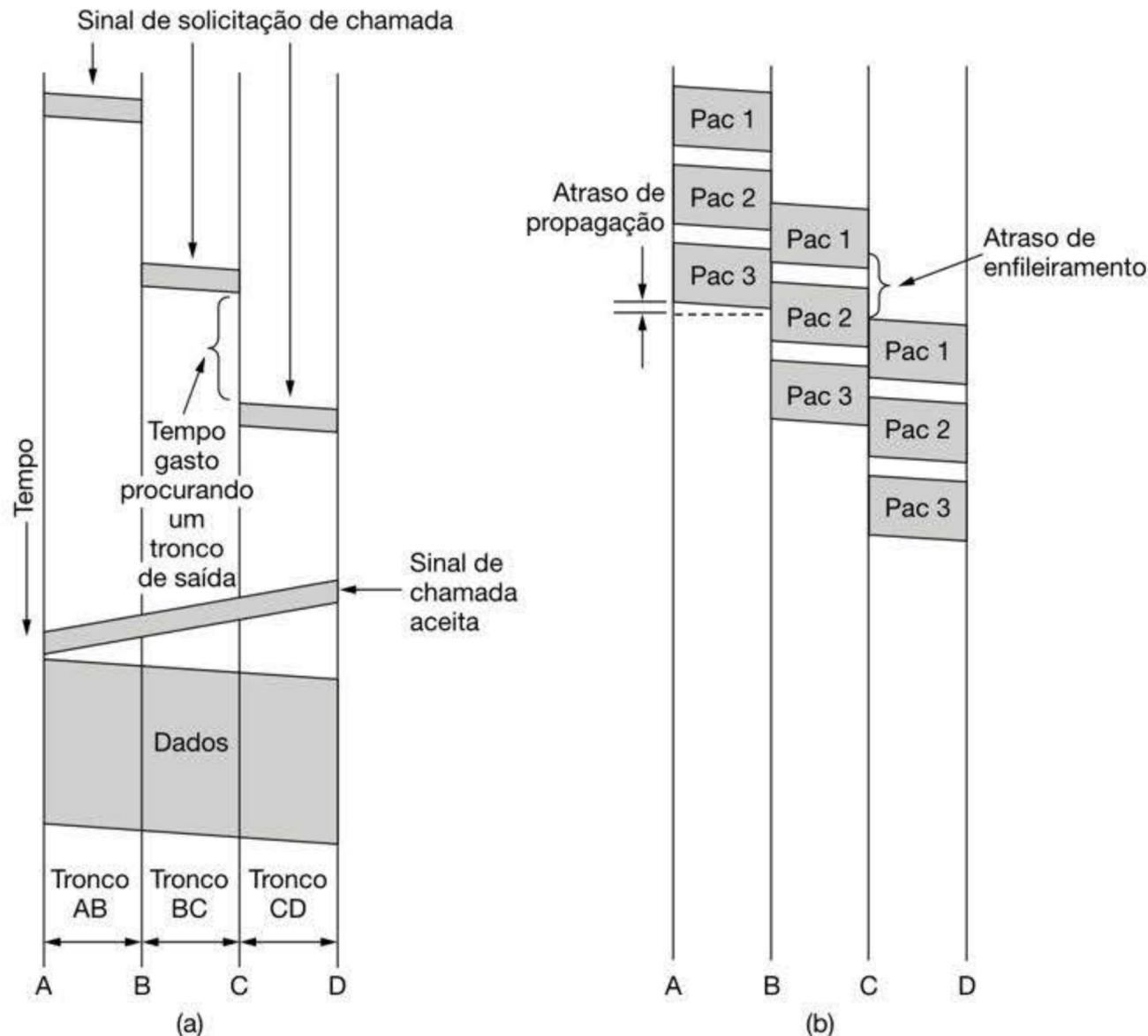
de comutação telefônica ou seu negócio iria à falência. Ele escolheu a primeira opção. Por cerca de 100 anos, o equipamento de comutação de circuitos usado em todo o mundo foi conhecido como **engrenagem de Strowger**. (A história não registra se a telefonista, desempregada, conseguiu emprego como operadora de informações, respondendo a perguntas como: “Qual é o número do telefone do agente funerário?”).

O modelo mostrado na Figura 2.36(a) é altamente simplificado, obviamente, porque partes do caminho físico entre os dois telefones podem de fato ser enlaces de micro-ondas ou de fibra, nos quais são multiplexadas milhares de chamadas. Entretanto, a ideia básica é válida: uma vez estabelecida uma chamada, haverá um caminho dedicado entre ambas as extremidades, e ele continuará a existir até que a chamada seja finalizada.

Uma propriedade importante da comutação de circuitos é a necessidade de se estabelecer um caminho ponta a ponta *antes* que qualquer dado possa ser enviado. O tempo decorrido entre o fim da discagem e o momento em que o telefone começa a tocar pode chegar a 10 segundos ou mais em chamadas interurbanas ou internacionais. Durante esse intervalo, o sistema telefônico procura uma conexão física, como mostra a Figura 2.37(a). Observe que, antes mesmo de se iniciar a transmissão de dados, o sinal de solicitação de chamada deve se propagar em todo o trajeto até o destino e lá ser reconhecido. Para muitas aplicações de informática (p. ex., a verificação de crédito em um ponto de venda), tempos de preparação longos são indesejáveis.



**Figura 2.36** (a) Comutação de circuitos. (b) Comutação de pacotes.



**Figura 2.37** Sincronização de eventos em (a) comutação de circuitos e (b) comutação de pacotes.

Como consequência do caminho reservado entre o transmissor e o receptor da chamada, uma vez estabelecida a configuração, o único atraso para a entrega dos dados é o tempo de propagação do sinal eletromagnético, cerca de 5 ms por 1.000 km. Outra consequência do caminho estabelecido é que não há perigo de congestionamento – ou seja, quando a chamada é feita, você nunca obtém sinal de ocupado. É claro que é possível receber um sinal de ocupado antes do estabelecimento da conexão, em decorrência da falta de capacidade de comutação ou de troncos.

### Comutação de pacotes

A alternativa à comutação de circuitos é a **comutação de pacotes**, mostrada na Figura 2.36(b) e descrita no Capítulo 1. Com essa tecnologia, os pacotes são enviados assim que estão disponíveis. Não é preciso estabelecer um caminho dedicado com antecedência, diferentemente da comutação de circuitos. A comutação de pacotes é semelhante a enviar uma série de cartas usando o sistema postal: cada uma segue independentemente das outras. Fica a critério dos roteadores usar a transmissão store-and-forward para enviar cada pacote em seu caminho ao destino por conta própria. Esse procedimento é diferente da comutação de circuitos, em que o resultado do estabelecimento da

conexão é a reserva de largura de banda desde o transmissor até o receptor, e todos os dados no circuito seguem esse caminho. Na comutação de circuitos, fazer todos os pacotes seguirem o mesmo caminho significa que eles não poderão chegar fora de ordem. Com a comutação de pacotes, não há caminho fixo e, assim, diferentes pacotes podem seguir caminhos distintos, dependendo das condições da rede no momento em que eles são enviados. Portanto, eles podem chegar fora de ordem.

As redes de comutação de pacotes impõem um limite superior apertado sobre o tamanho dos pacotes. Isso garante que nenhum usuário poderá monopolizar qualquer linha de transmissão por muito tempo (p. ex., muitos milissegundos), de modo que as redes de comutação de pacotes podem lidar com o tráfego interativo. Isso também reduz o atraso, pois o primeiro pacote de uma mensagem longa pode ser encaminhado antes que o segundo tenha chegado por completo. Contudo, o atraso store-and-forward de acumular um pacote na memória do roteador antes que ele seja enviado para o próximo roteador excede o da comutação de circuitos. Com a comutação de circuitos, os bits simplesmente fluem pelo fio de modo contínuo. Nada é armazenado e encaminhado mais tarde.

A comutação de pacotes e de circuitos também difere de outras maneiras. Como nenhuma largura de banda é

reservada com a comutação de pacotes, estes podem ter de esperar para serem encaminhados. Se muitos pacotes forem enviados ao mesmo tempo, isso introduz o **atraso de enfileiramento** e o congestionamento. Em contrapartida, não existe perigo de obter um sinal de ocupado e não conseguir usar a rede. Assim, o congestionamento ocorre em diferentes momentos com a comutação de circuitos (no momento da configuração) e com a comutação de pacotes (quando os pacotes são enviados).

Se um circuito tiver sido reservado para determinado usuário e não houver tráfego, sua largura de banda é desperdiçada. Ela não pode ser usada para outro tráfego. A comutação de pacotes não desperdiça largura de banda e, portanto, é mais eficiente do ponto de vista do sistema. Entender essa escolha é decisivo para compreender a diferença entre comutação de circuitos e comutação de pacotes. O dilema está entre garantir serviço e desperdiçar recursos *versus* não garantir serviço e não desperdiçar recursos.

A comutação de pacotes é mais tolerante a falhas que a comutação de circuitos. De fato, é por isso que ela foi criada. Se um switch ficar inativo, todos os circuitos que o utilizam serão encerrados, e nenhum tráfego poderá mais ser transmitido em qualquer um deles. Com a comutação de pacotes, os pacotes poderão ser roteados de modo a contornar switches inativos.

Outra diferença entre a comutação de circuitos e a de pacotes é o algoritmo de tarifação. Com a comutação de circuitos (i.e., as chamadas de telefonia de voz por PSTN), a tarifação se baseava historicamente na distância e no tempo. No caso dos telefones móveis, em geral a distância não é importante, exceto para chamadas internacionais, e o tempo desempenha apenas um papel secundário (p. ex., um plano de chamadas com 2 mil minutos gratuitos custa mais que um plano com mil minutos gratuitos e, algumas vezes, chamadas noturnas ou nos finais de semana são mais econômicas que o normal). Com a comutação de pacotes,

o tempo de conexão não é um problema, mas o volume de tráfego sim. Para usuários domésticos nos Estados Unidos e na Europa, os ISPs normalmente cobram uma tarifa fixa mensal, porque tal modelo é menos trabalhoso para eles e mais fácil de entender para os clientes. Em alguns países em desenvolvimento, a tarifação normalmente ainda é baseada no volume: os usuários podem comprar um “pacote de dados” de certo tamanho e usá-lo por um ciclo de cobrança. Certos horários do dia, ou mesmo certos destinos, podem não ser cobrados ou não contar para o limite de dados; esses serviços às vezes são chamados de **serviços com taxa zero**. Geralmente, os provedores de serviço de Internet da operadora no backbone da Internet cobram com base nos volumes de tráfego. Um modelo de tarifação típico é baseado no 95º percentil de amostras de 5 minutos: em um determinado enlace, um ISP medirá o volume de tráfego que passou pelo enlace nos últimos 5 minutos. Um ciclo de faturamento de 30 dias terá 8640 desses intervalos de 5 minutos e o ISP cobrará com base no 95º percentil dessas amostras. Essa técnica costuma ser chamada de **cobrança pelo percentil 95**.

As diferenças entre comutação de circuitos e comutação de pacotes estão resumidas na Figura 2.38. Tradicionalmente, as redes telefônicas têm usado a comutação de circuitos para oferecer chamadas telefônicas de alta qualidade, e as redes de computadores têm usado a comutação de pacotes por suas simplicidade e eficiência. Contudo, existem exceções dignas de nota. Algumas redes de computadores mais antigas têm sido comutadas por circuitos “por debaixo dos panos” (p. ex., X.25), e algumas redes telefônicas mais novas usam a comutação de pacotes com a tecnologia VoIP. Para os usuários, isso se parece externamente com uma chamada telefônica padrão, mas, internamente, os pacotes na rede com dados de voz são comutados. Com essa técnica, as chamadas internacionais com cartões de chamada

Item	Comutação de circuitos	Comutação de pacotes
Configuração de chamadas	Obrigatória	Não necessária
Caminho físico dedicado	Sim	Não
Cada pacote segue a mesma rota	Sim	Não
Os pacotes chegam em ordem	Sim	Não
A falha de um switch é fatal	Sim	Não
Largura de banda disponível	Fixa	Dinâmica
Momento de possível congestionamento	Durante a configuração	Em todos os pacotes
Largura de banda potencialmente desperdiçada	Sim	Não
Transmissão store-and-forward	Não	Sim
Tarifação	Por minuto	Por byte

**Figura 2.38** Comparação entre redes de comutação de circuitos e redes de comutação de pacotes.

podem se tornar mais baratas, embora talvez com uma qualidade de chamada inferior à do serviço tradicional.

## 2.6 REDES DE TELEFONIA MÓVEL

Mesmo que um dia o sistema de telefonia convencional receba fibra multigigabit de ponta a ponta, as pessoas agora esperam efetuar chamadas telefônicas e usar seus telefones para verificar e-mail e navegar pela Web em aviões, carros, piscinas e enquanto caminham no parque. Consequentemente, há um enorme interesse (e investimento) na telefonia sem fio.

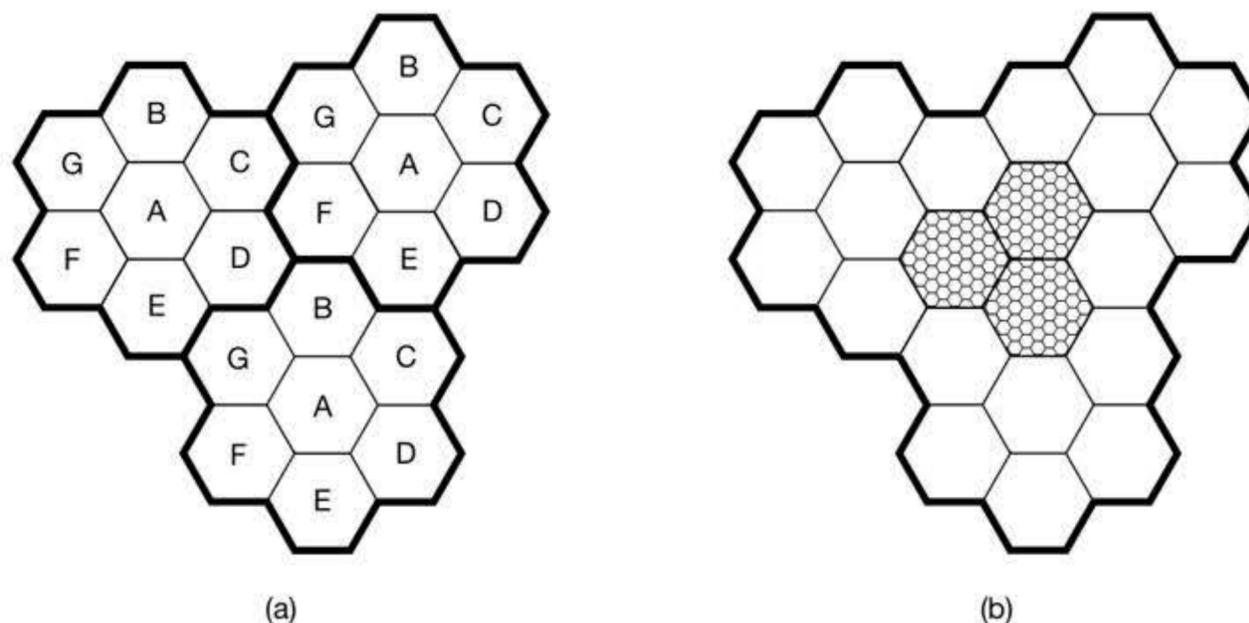
O sistema de telefonia móvel é usado para comunicação remota de voz e dados. Os **telefones móveis** (também chamados **telefones celulares**) passaram por cinco gerações distintas, normalmente chamadas 1G, 2G, 3G, 4G e 5G. As três gerações iniciais ofereciam voz analógica, voz digital e voz digital e dados (Internet, e-mail, etc.), respectivamente. A tecnologia 4G acrescenta outras capacidades, incluindo técnicas adicionais de transmissão pela camada física (p. ex., transmissões uplink por OFDM) e femtocélulas baseadas em IP (nós de celulares domésticos que são conectados à infraestrutura de Internet da linha fixa). A tecnologia 4G não admite a telefonia de comutação de circuitos, diferentemente de suas predecessoras, é baseada apenas na comutação de pacotes. A tecnologia 5G está sendo implantada agora, mas serão necessários muitos anos antes que ela substitua completamente as gerações anteriores em todos os lugares. A tecnologia 5G aceitará transmissões de até 20 Gbps, bem como implantações mais densas. Há também um foco na redução da latência da rede para dar suporte a uma maior gama de aplicações, como para jogos altamente interativos.

### 2.6.1 Conceitos comuns: células, handoff, paging

Em todos os sistemas de telefonia móvel, uma região geográfica é dividida em **células**, e é esse o motivo pelo qual esses dispositivos são chamados telefones celulares. Cada célula utiliza algum conjunto de frequências não utilizado por qualquer uma das células vizinhas. A ideia fundamental que dá aos sistemas celulares uma capacidade muito maior que a dos sistemas anteriores é o uso de células relativamente pequenas e a reutilização de frequências de transmissão em células próximas (mas não adjacentes). O projeto do celular aumenta a capacidade do sistema enquanto as células são menores. Além disso, células menores significam menor necessidade de energia, o que possibilita a existência de dispositivos transmissores e receptores menores e mais econômicos.

As células permitem a reutilização de frequências, que é ilustrada na Figura 2.39(a). Em geral, as células são razoavelmente circulares, porém, é mais simples representá-las como hexágonos. Na Figura 2.39(a), todas as células têm o mesmo tamanho. Elas são agrupadas em unidades de sete células. Cada letra indica um grupo de frequências. Observe que, para cada conjunto de frequências, existe um afastamento de aproximadamente duas células de extensão, no qual essa frequência não é reutilizada, o que proporciona boa separação e pouca interferência.

Em uma área em que o número de usuários cresce a ponto de o sistema ficar sobrecarregado, a potência pode ser reduzida, e as células sobrecarregadas são divididas em células menores, chamadas **microcélulas**, para permitir maior reutilização de frequências, como mostra a Figura 2.39(b). Algumas vezes, as empresas de telefonia criam microcélulas temporárias, utilizando torres portáteis com enlaces de satélite a fim de atender à demanda de eventos esportivos, shows de rock e outros eventos nos quais um



**Figura 2.39** (a) As frequências não são reutilizadas nas células adjacentes. (b) Para aumentar o número de usuários, podem ser utilizadas células menores.

grande número de usuários de telefones celulares se reúne por algumas horas.

No centro de cada célula há uma estação-base que recebe as transmissões de todos os telefones presentes na célula. Essa estação consiste em um computador e um transmissor/receptor, conectados a uma antena. Em um sistema de pequeno porte, todas as estações-base estão conectadas a um único dispositivo, chamado **centro de comutação móvel**, ou **MSC (Mobile Switching Center)**. Em um sistema maior, podem ser necessários vários MSCs, todos conectados a um MSC de segundo nível, e assim por diante. Basicamente, os MSCs são estações finais, como no sistema telefônico, e na verdade estão conectados a pelo menos uma estação final de um sistema telefônico. Os MSCs se comunicam com as estações-base, entre si e com a PSTN usando uma rede de comutação de pacotes.

Em qualquer instante, cada telefone móvel logicamente ocupa uma célula específica e está sob o controle da estação-base dela. Quando um telefone móvel deixa fisicamente uma célula, sua estação-base detecta que o sinal do telefone está enfraquecendo e questiona todas as estações-base vizinhas quanto à quantidade de energia que elas estão recebendo dele. Quando as respostas retornam, a estação-base faz a transferência para a célula que está obtendo o sinal mais forte; quase sempre, essa é a célula em que o telefone está localizado no momento. O telefone é, então, informado de quem é o seu novo chefe e, se houver uma chamada em andamento, ele será solicitado a passar para um novo canal (porque o antigo não é reutilizado em nenhuma das células adjacentes). Esse processo é chamado de **handoff** e leva cerca de 300 ms. A atribuição de canais é feita pelo MSC, o centro nervoso do sistema. Na verdade, as estações-base são apenas simples retransmissoras de rádio.

Encontrar locais altos para instalar antenas de estação-base é uma questão fundamental. Esse problema levou algumas concessionárias de telecomunicações a fazer alianças com a Igreja Católica Romana, que possui um número significativo de locais apropriados para a instalação de antenas em todo o mundo, todos convenientemente controlados por uma única administração.

As redes celulares normalmente têm quatro tipos de **canais**. **Canais de controle** (da base para a unidade móvel) são usados para gerenciar o sistema. **Canais de localização** (da base para a unidade móvel) alertam os usuários móveis a respeito de chamadas destinadas a eles. **Canais de acesso** (bidirecionais) são usados para estabelecimento de chamadas e atribuição de canais. Por fim, os **canais de dados** (bidirecionais) transportam voz, fax ou dados.

## 2.6.2 Tecnologia de primeira geração (1G): voz analógica

Vamos examinar a tecnologia da rede celular, começando pelo sistema mais antigo. Os radiotelefonos móveis eram

usados esporadicamente na comunicação militar e marítima, durante as primeiras décadas do século XX. Em 1946, foi criado em St. Louis, Estados Unidos, o primeiro sistema para telefones baseado em automóveis. O sistema utilizava um único transmissor grande no topo de um alto edifício e tinha um único canal, usado para transmissões e recepções. Para conversar, o usuário tinha de apertar um botão que ativava o transmissor e desativava o receptor. Tais sistemas, conhecidos como **sistemas “apertar para falar” (push-to-talk systems)**, foram instalados em diversas cidades a partir da década de 1950. Táxis e carros de polícia utilizam essa tecnologia com frequência.

Na década de 1960, foi instalado o **sistema de telefonia móvel aperfeiçoado**, ou **IMTS (Improved Mobile Telephone System)**. Ele também utilizava um transmissor de alta potência (200 watts) no topo de uma montanha, mas agora tinha duas frequências: uma para transmissão e outra para recepção. Por isso, o botão “apertar para falar” não era mais necessário. Como toda a comunicação dos telefones móveis utilizava um canal para transmissão e outro para recepção dos sinais, os usuários móveis não podiam ouvir uns aos outros (ao contrário do que acontecia com o sistema “apertar para falar” utilizado em táxis).

O IMTS admitia 23 canais espalhados pelas frequências de 150 a 450 MHz. Em virtude do pequeno número de canais, muitas vezes os usuários tinham de esperar muito tempo antes de obter um tom de discagem. Além disso, pela alta potência do transmissor, os sistemas adjacentes tinham de estar a vários quilômetros de distância uns dos outros para evitar interferência. Em suma, sua capacidade limitada tornou o sistema impraticável.

O **sistema avançado de telefonia móvel**, ou **AMPS (Advanced Mobile Phone System)**, bastante semelhante, inventado pelo Bell Labs e instalado inicialmente nos Estados Unidos em 1983, aumentou significativamente a capacidade da rede celular. Ele também foi usado na Inglaterra, onde recebeu o nome TACS, e no Japão, onde foi chamado MCS-L1. O uso do AMPS foi encerrado formalmente em 2008, mas vamos examiná-lo para entender o contexto para os sistemas 2G e 3G, que o aperfeiçoaram. No AMPS, as células normalmente ficam afastadas por 10 a 20 km; em sistemas digitais, as células são menores. Enquanto um sistema IMTS cobrindo 100 km só pode ter uma chamada em cada frequência, um sistema AMPS poderia ter 100 células de 10 km na mesma área, podendo ter de 10 a 15 chamadas em cada frequência, em células bastante separadas.

O AMPS utiliza FDM para separar os canais e usa 832 canais full-duplex, cada um consistindo em um par de canais simplex. Esse arranjo é conhecido como **duplex por divisão de frequência**, ou **FDD (Frequency Division Duplex)**. Os 832 canais simplex de 824 a 849 MHz são usados para transmissão do aparelho móvel à estação-base, e 832 canais simplex de 869 a 894 MHz são usados para

transmissão da estação-base ao aparelho móvel. Cada um desses canais simplex tem 30 kHz de largura.

Os 832 canais no AMPS estão divididos em quatro categorias. Como as mesmas frequências não podem ser reutilizadas em células vizinhas e 21 desses canais são reservados em cada célula para controle, o número real de canais de voz disponíveis por célula é bem menor que 832, normalmente em torno de 45.

### Gerenciamento de chamadas

Cada telefone móvel tem um número de série de 32 bits e um número de telefone de dez dígitos em sua PROM (memória programável somente de leitura). O número de telefone é representado como um código de área de 3 dígitos em 10 bits e um número de assinante de 7 dígitos em 24 bits. Quando um telefone é contatado, varre uma lista pré-programada de 21 canais de controle até encontrar o sinal mais forte. Em seguida, o telefone transmite seu número de série de 32 bits e o número de telefone de 34 bits. A exemplo de todas as outras informações de controle do AMPS, esse pacote é enviado várias vezes em formato digital e com um código de correção de erros, apesar de os próprios canais de voz serem analógicos.

Quando ouve a mensagem, a estação-base avisa ao MSC, que registra a existência de seu novo cliente e também informa a sua localização atual ao MSC local. Durante a operação normal, o telefone móvel repete o registro uma vez a cada 15 minutos, aproximadamente.

Para fazer uma chamada, o usuário móvel liga o telefone, digita no teclado o número a ser chamado (pelo menos, conceitualmente) e pressiona o botão LIGAR. Em seguida, o telefone transmite o número a ser chamado e sua própria identidade no canal de acesso. Se houver uma colisão, ele tenta novamente mais tarde. Ao receber a solicitação, a estação-base informa ao MSC. Se o chamador for um cliente da empresa do MSC (ou de uma de suas parceiras), o MSC procura um canal disponível para a chamada. Se encontrar algum, o número do canal será enviado de volta no canal de controle. Em seguida, o telefone móvel se conecta automaticamente ao canal de voz selecionado e aguarda até que a parte chamada atenda ao telefone.

As chamadas recebidas funcionam de forma diferente. Para começar, todos os telefones inativos ouvem continuamente o canal de localização para detectar as mensagens destinadas a eles. Quando é feita uma chamada para um telefone móvel (a partir de um telefone fixo ou de outro telefone móvel), um pacote é enviado ao MSC local do telefone chamado, para que ele seja localizado. Em seguida, é enviado um pacote à estação-base em sua célula atual, que, então, envia um pacote de difusão no canal de localização com o formato: "Unidade 14, você está aí?". O telefone chamado responde "Sim" no canal de acesso. Depois, a base transmite algo como: "Unidade 14, chamada para você no canal 3". Nesse momento, o telefone chamado se

conecta ao canal 3 e começa a emitir sinais sonoros (ou a tocar alguma melodia que o proprietário do telefone ganhou como presente de aniversário).

### 2.6.3 Tecnologia de segunda geração (2G): voz digital

A primeira geração de telefones celulares era analógica; a segunda geração é digital. A troca para digital tem diversas vantagens. Ela oferece ganhos de capacidade, permitindo que os sinais de voz sejam digitalizados e compactados. Ela melhora a segurança, permitindo que sinais de voz e de controle sejam criptografados. Isso, por sua vez, impede fraude e espionagem, seja por varredura intencional, seja por ecos de outras chamadas, em virtude da propagação de ondas de rádio. Por fim, ela capacita novos serviços, como mensagens de texto.

Assim como não havia padronização internacional durante a primeira geração, também não havia padronização internacional durante a segunda. Vários sistemas diferentes foram desenvolvidos, e três foram amplamente implementados. O **sistema avançado de telefonia móvel digital**, ou **D-AMPS (Digital Advanced Mobile Phone System)** é uma versão digital do AMPS que coexiste com ele e usa TDM para fazer várias chamadas no mesmo canal de frequência. Ele é descrito no padrão internacional IS-54 e seu sucessor, o IS-136. O **sistema global para comunicações móveis**, ou **GSM (Global System for Mobile Communications)** apareceu como o sistema dominante, e, embora tenha demorado para ser aceito nos Estados Unidos, agora é usado praticamente em todo o mundo. Assim como o D-AMPS, o GSM é baseado em uma mistura de FDM e TDM. O **acesso múltiplo por divisão de código**, ou **CDMA (Code Division Multiple Access)**, descrito no **padrão internacional IS-95**, é um tipo de sistema completamente diferente, que não é baseado nem em FDM nem em TDM. Embora o CDMA não tenha se tornado o sistema 2G dominante, sua tecnologia se tornou a base para os sistemas 3G.

Além disso, o nome **serviços de comunicações pessoais**, ou **PCS (Personal Communications Services)** às vezes é usado na literatura de marketing para indicar um sistema de segunda geração (ou seja, digital). Inicialmente, isso indicava um telefone móvel usando a banda de 1.900 MHz, mas essa distinção raramente é feita nos dias atuais. O sistema 2G dominante em quase todo o mundo é o GSM, que descrevemos em seguida com detalhes.

### 2.6.4 GSM: Global System for Mobile Communications

O GSM surgiu na década de 1980 como um esforço para produzir um único padrão 2G europeu. A tarefa foi atribuída a um grupo de telecomunicações chamado (em

francês) Groupe Spécialé Mobile. Os primeiros sistemas GSM foram implantados a partir de 1991 e experimentaram um sucesso repentino. Logo, ficou claro que o GSM seria mais do que um sucesso europeu, sendo absorvido até mesmo em países como a Austrália, de modo que o sistema foi renomeado para que tivesse um apelo mais global.

O GSM e outros sistemas de telefonia móvel que estudaremos retêm, dos sistemas 1G, um projeto baseado em células, reutilização de frequência pelas células e mobilidade com handoffs à medida que os assinantes se movem. São os detalhes que diferem. A seguir, descreveremos algumas das principais propriedades do GSM. Entretanto, o padrão GSM impresso tem mais de 5.000 [sic] páginas. Uma grande fração desse material se relaciona aos aspectos de engenharia do sistema, em especial ao projeto dos receptores para tratar da propagação de sinais por vários caminhos, e a sincronização de transmissores e receptores. Nada disso será mencionado aqui.

A Figura 2.40 mostra que a arquitetura do GSM é semelhante à arquitetura do AMPS, embora os componentes tenham nomes diferentes. O próprio aparelho móvel agora é dividido em um aparelho e em um chip removível, com informações do assinante e da conta contidas em um **cartão SIM**, uma abreviação de **Subscriber Identity Module (módulo de identidade do assinante)**. É o cartão SIM que ativa o aparelho e contém segredos que permitem que o aparelho e a rede se identifiquem e codifiquem as conversas. Um cartão SIM pode ser removido e colocado em um aparelho diferente, para que este se torne seu aparelho móvel em relação à rede.

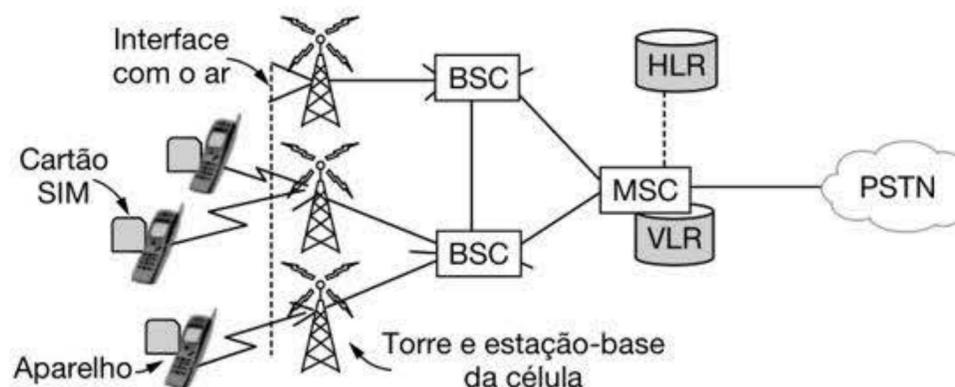
O telefone móvel fala com as estações-base da célula por uma **interface com o ar**, que descreveremos em breve. As estações-base da célula estão conectadas a um **controlador de estação-base**, ou **BSC (Base Station Controller)**, que controla os recursos de rádio das células e cuida do handoff. O BSC, por sua vez, está conectado a um MSC (como no AMPS), que direciona as chamadas e as conecta à rede de telefonia pública comutada, ou PSTN.

Para direcionar as chamadas, o MSC precisa saber onde os aparelhos podem ser encontrados atualmente. Ele mantém um banco de dados dos aparelhos nas vizinhanças,

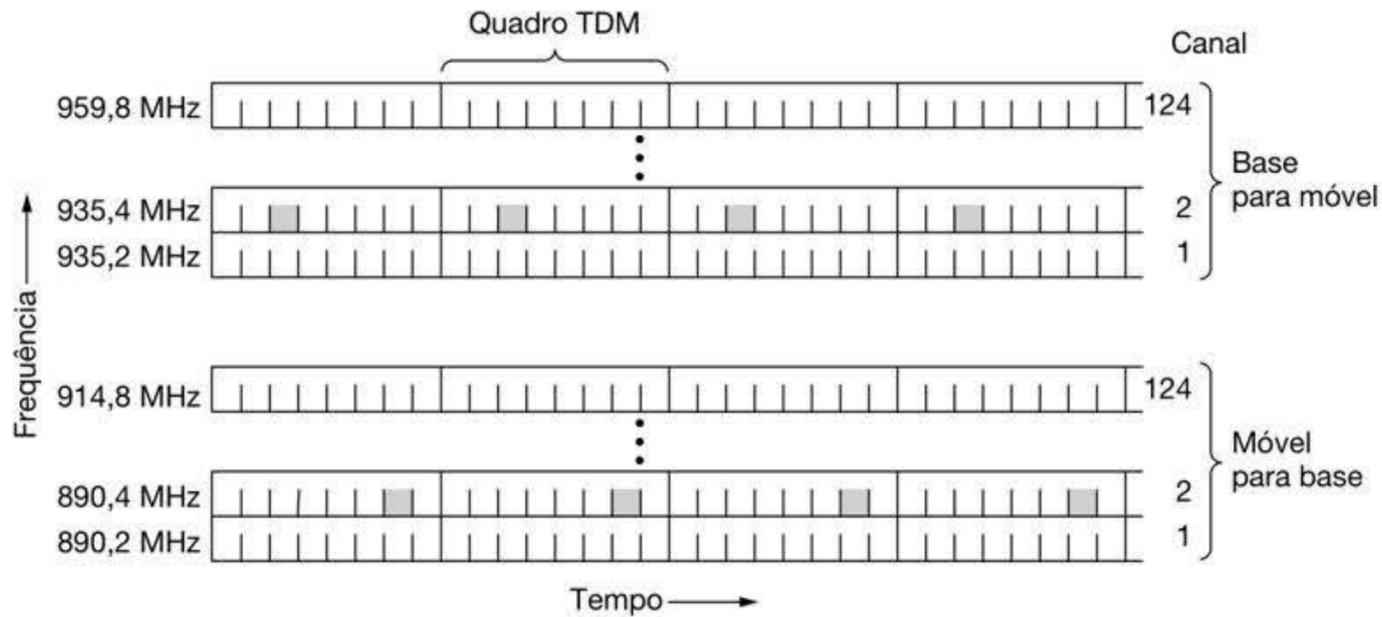
que estão associados às células que ele controla. Esse banco de dados é chamado **registrador de local do visitante**, ou **VLR (Visitor Location Register)**. Também há um banco de dados na rede móvel que indica o último local conhecido de cada aparelho, chamado **registrador de local inicial**, ou **HLR (Home Location Register)**, usado para direcionar as chamadas que chegam para os locais corretos. Os dois bancos de dados devem ser mantidos atualizados enquanto os aparelhos passam de uma célula para outra.

Agora, vamos descrever a interface com o ar em alguns detalhes. O GSM trabalha em uma faixa de frequências internacional, incluindo 900, 1.800 e 1.900 MHz. Foi alocado um espectro maior que o AMPS, para dar suporte a um número muito maior de usuários. O GSM é um sistema celular duplex por divisão de frequência, como o AMPS. Ou seja, cada aparelho móvel transmite em uma frequência e recebe em outra mais alta (55 MHz mais alta para GSM, contra 80 MHz mais alta para AMPS). Contudo, diferentemente do AMPS, com o GSM, um único par de frequências é dividido pela TDM em slots de tempo. Desse modo, ele é compartilhado por vários aparelhos.

Para lidar com vários aparelhos, os canais GSM são muito mais largos que os canais AMPS (200 kHz contra 30 kHz). Um canal de 200 kHz é representado na Figura 2.41. Um sistema GSM operando na região de 900 MHz tem 124 pares de canais simplex. Cada canal simplex tem 200 kHz de largura e aceita oito conexões separadas nele, usando a TDM. Cada estação atualmente ativa recebe um slot de tempo em um par de canais. Teoricamente, 992 canais podem ser aceitos em cada célula, mas muitos deles não estão disponíveis, a fim de evitar conflitos de frequência com células vizinhas. Na Figura 2.41, todos os oito slots de tempo sombreados pertencem à mesma conexão, quatro deles em cada direção. A transmissão e a recepção não acontecem no mesmo slot de tempo, pois os rádios GSM não podem transmitir e receber ao mesmo tempo, e leva algum tempo para passar de um para o outro. Se a unidade móvel atribuída à faixa de 890,4/935,4 MHz e ao slot de tempo 2 quisesse transmitir algo para a estação-base, ela usaria os quatro slots sombreados inferiores (e os slots depois deles no tempo), inserindo alguns dados em cada um até que todos os dados fossem enviados.



**Figura 2.40** Arquitetura móvel da rede GSM.



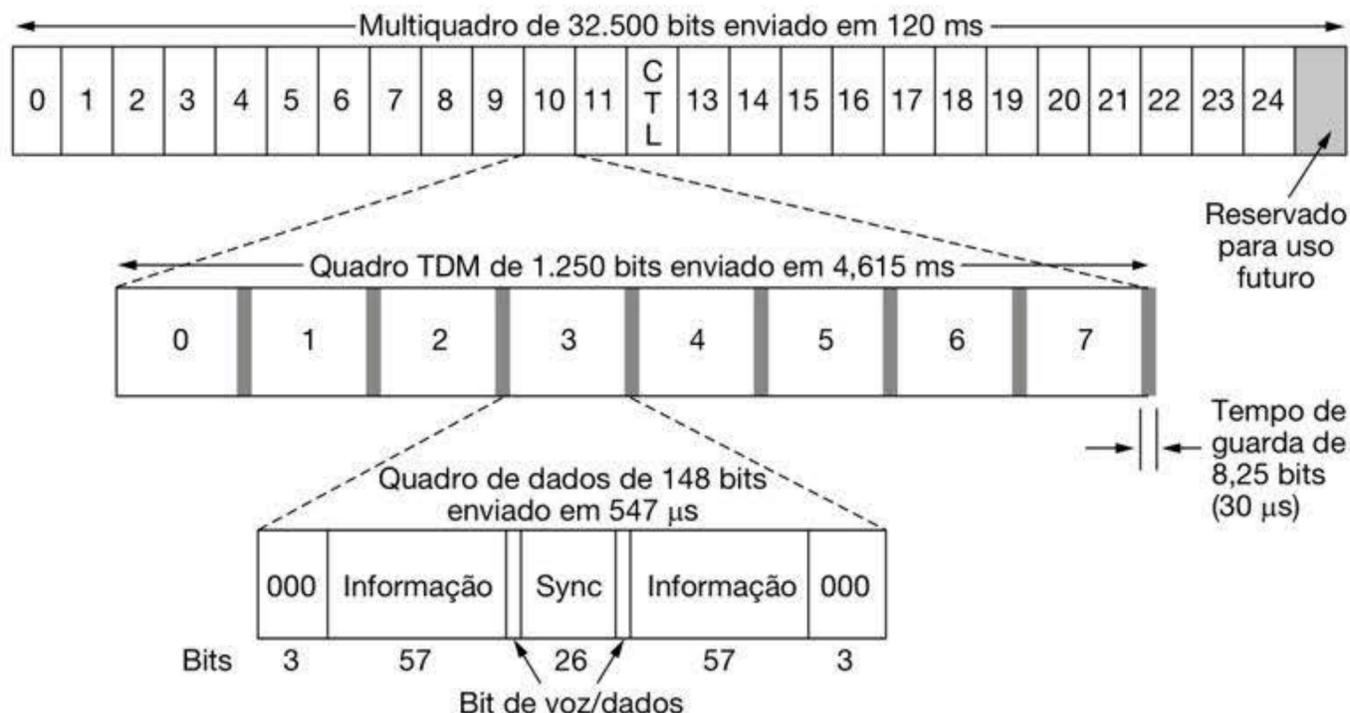
**Figura 2.41** O sistema GSM utiliza 124 canais de frequência, cada um usando um sistema TDM de oito slots.

Os slots TDM mostrados na Figura 2.41 fazem parte de uma complexa hierarquia de enquadramento. Cada slot TDM tem uma estrutura específica, e grupos de slots TDM formam multiquadros, também com uma estrutura específica. Uma versão simplificada dessa hierarquia é mostrada na Figura 2.42. Aqui, podemos ver que cada slot TDM consiste em um quadro de dados de 148 bits que ocupa o canal por 577  $\mu$ s (incluindo um tempo de proteção de 30  $\mu$ s depois de cada slot). Cada quadro de dados começa e termina com três bits 0, para fins de delimitação de quadros. Ele também contém dois campos *Informação* de 57 bits, cada um com um bit de controle que indica se o campo *Informação* seguinte se refere a voz ou a dados. Entre os campos *Informação* há um campo (de treinamento) *Sync* de 26 bits, usado pelo receptor para realizar a sincronização até os limites de quadro do transmissor.

Um quadro de dados é transmitido em 547  $\mu$ s, mas um transmissor só pode enviar um quadro de dados a cada

4,615 ms, pois ele está compartilhando o canal com sete outras estações. A taxa bruta de cada canal é de 270.833 bps, dividida entre oito usuários. Porém, como ocorre com o AMPS, o overhead consome uma grande fração da largura de banda, deixando, em última análise, 24,7 kbps de carga útil por usuário antes da correção de erros. Após essa correção, restam 13 kbps para voz. Embora isso seja substancialmente menor que os 64 kbps do PCM para sinais de voz não compactados na rede de telefonia fixa, a compactação no dispositivo móvel pode alcançar esses níveis com pouca perda de qualidade.

Como podemos ver na Figura 2.42, oito quadros de dados formam um quadro TDM, e 26 quadros TDM formam um multiquadro de 120 ms. Dos 26 quadros TDM em um multiquadro, o slot 12 é usado para controle e o slot 25 é reservado para uso futuro; assim, somente 24 estão disponíveis para tráfego do usuário.



**Figura 2.42** Uma parte da estrutura de enquadramento GSM.

Porém, além do multiquadro de 26 slots mostrado na Figura 2.42, também é usado um multiquadro de 51 slots (não mostrado). Alguns desses slots são empregados para guardar diversos canais de controle usados para gerenciar o sistema. O **canal de controle de broadcast** é um fluxo contínuo de saída da estação-base, contendo a identidade da estação-base e o status do canal. Todas as estações móveis monitoram a intensidade de seu sinal para verificar quando elas são transferidas para uma nova célula.

O **canal de controle dedicado** é usado para atualização de local, registro e estabelecimento de chamadas. Em particular, cada estação-base mantém o VLR, um banco de dados das estações móveis que atualmente estão sob sua jurisdição. As informações necessárias para manter o VLR são enviadas no canal de controle dedicado.

Por fim, existe o **canal de controle comum**, dividido em três subcanais lógicos. O primeiro deles é o **canal de localização**, que a estação-base utiliza para anunciar as chamadas recebidas. Cada estação móvel monitora continuamente esse canal para verificar se há chamadas a que ela deva responder. O segundo é o **canal de acesso aleatório**, que permite aos usuários solicitarem um slot no canal de controle dedicado. Se duas solicitações colidirem, elas serão adulteradas e terão de ser repetidas mais tarde. Usando o slot do canal de controle dedicado, a estação pode estabelecer uma chamada. O slot atribuído é anunciado no terceiro subcanal, o **canal de concessão de acesso**.

Por último, o GSM difere do AMPS no modo como o handoff é tratado. No AMPS, o MSC o controla totalmente, sem ajuda dos dispositivos móveis. Com os slots de tempo no GSM, na maior parte das vezes o dispositivo móvel não está nem enviando nem recebendo. Os slots ociosos são uma oportunidade para o dispositivo móvel medir a qualidade do sinal até outras estações-base nas proximidades. Ele faz isso e envia essa informação ao BSC, que pode utilizá-la para determinar quando um dispositivo móvel está saindo de uma célula e entrando em outra, de modo que possa realizar o handoff. Esse projeto é conhecido como **handoff auxiliado pela unidade móvel**, ou **MAHO (Mobile Assisted HandOff)**.

### 2.6.5 Tecnologia de terceira geração (3G): voz e dados digitais

A primeira geração de telefones móveis foi a voz analógica, e a segunda foi a voz digital. A terceira geração de telefones móveis, ou **3G**, como é chamada, trata de voz e dados digitais. Diversos fatores direcionaram o setor para a tecnologia 3G. Primeiro, na época do 3G, o tráfego de dados começou a ultrapassar o tráfego de voz na rede fixa; tendências semelhantes começaram a surgir para dispositivos móveis. Segundo, os serviços de telefone, Internet e vídeo começaram a convergir. O aumento do uso de smartphones, a partir do iPhone da Apple, que foi lançado em 2007, acelerou a

mudança para dados móveis. Os volumes de dados estão subindo bastante com a popularidade dos iPhones. Quando o iPhone foi lançado, ele usava uma rede **2,5G** (basicamente uma rede 2G melhorada), que não tinha capacidade de dados suficiente. Os usuários do iPhone, ávidos por dados, impulsionaram ainda mais a transição para tecnologias 3G, para dar suporte a velocidades de transmissão de dados mais altas. Um ano depois, em 2008, a Apple lançou uma versão atualizada de seu iPhone, que podia usar a rede de dados 3G.

Inicialmente, as operadoras deram passos pequenos em direção ao 3G, chegando ao que se costuma chamar de **2,5G**. Um sistema desse tipo utiliza **taxas de dados aprimoradas para evolução do GSM**, ou **EDGE (Enhanced Data Rates for GSM Evolution)**, que é simplesmente o GSM com mais bits por símbolo. O problema é que mais bits por símbolo também significa mais erros por símbolo, e assim o EDGE tem nove esquemas diferentes para modulação e correção de erros, que se distinguem pela proporção da largura de banda dedicada à correção dos erros introduzidos pela velocidade mais alta. Trata-se de um passo em direção a um caminho evolutivo que é definido de GSM a outras tecnologias 3G que discutiremos nesta seção.

A ITU tentou ser um pouco mais específica em relação à visão do 3G em 1992. Ela apresentou um projeto para alcançá-lo, denominado **IMT-2000**, em que IMT significava **International Mobile Telecommunications (telecomunicações móveis internacionais)**. Os serviços básicos que a rede IMT-2000 deveria oferecer a seus usuários eram:

1. Transmissão de voz de alta qualidade.
2. Serviço de mensagens (substituindo e-mail, fax, SMS, bate-papo, etc.).
3. Multimídia (reprodução de música, exibição de vídeos, filmes, televisão, etc.).
4. Acesso à Internet (navegação na Web, incluindo páginas com áudio e vídeo).

Outros serviços poderiam ser: videoconferência, telepresença, jogos em grupo e m-commerce (comércio móvel, bastando utilizar seu telefone no caixa para pagar as compras feitas em uma loja). Além disso, todos esses serviços deveriam estar disponíveis em âmbito mundial (com conexão automática via satélite, quando não for possível instalar nenhuma rede terrestre), de forma instantânea (sempre ativos) e com garantias de qualidade de serviço. Em outras palavras, um sonho.

A ITU previu uma única tecnologia mundial para o IMT-2000, de forma que os fabricantes fossem capazes de construir um único dispositivo que pudesse ser vendido e utilizado em qualquer lugar do mundo. Ter uma única tecnologia também facilitaria bastante a vida dos operadores de redes e encorajaria mais pessoas a usarem os serviços.

Acontece que isso foi um pouco otimista. O número 2000 significou três coisas: (1) o ano em que o serviço

deveria ser iniciado; (2) a frequência em que ele deveria operar (em MHz); e (3) a largura de banda que o serviço deveria ter (em kbps). Nenhum desses três se concretizou. Nada foi implementado em 2000. A ITU recomendou que todos os governos reservassem o espectro em 2 GHz para que os dispositivos pudessem passar de um país para outro de forma transparente. A China reservou a largura de banda exigida, mas ninguém mais fez isso. Finalmente, reconheceu-se que 2 Mbps não são atualmente viáveis para usuários que se movimentam *muito* (em virtude da dificuldade de realizar handoffs com rapidez suficiente). O mais realista é 2 Mbps para usuários que não estão em movimento, 384 kbps para pessoas andando e 144 kbps para conexões em carros.

Apesar desses problemas iniciais, muito foi realizado desde então. Várias propostas do IMT-2000 foram feitas e, após uma seleção, elas se reduziram a duas principais: (1) o **CDMA de banda larga**, ou **WCDMA (Wideband CDMA)**, foi proposto pela Ericsson e adotado pela União Europeia, que o chamou de **sistema universal de telecomunicações móveis**, ou **UMTS (Universal Mobile Telecommunications System)**; e (2) o **CDMA2000**, proposto pela Qualcomm nos Estados Unidos.

Esses dois sistemas são mais semelhantes do que diferentes, pois são baseados no CDMA de banda larga; o WCDMA usa canais de 5 MHz e o CDMA2000 usa canais de 1,25 MHz. Se os engenheiros da Ericsson e da Qualcomm fossem confinados em uma sala e solicitados a apresentar um projeto comum, eles provavelmente conseguiriam fazê-lo rapidamente. A dificuldade é que o problema real não é de engenharia, mas político (como sempre). A Europa queria um sistema que trabalhasse junto com o GSM, enquanto os Estados Unidos queriam um sistema que fosse compatível com um sistema já amplamente desenvolvido no país (o IS-95). Cada lado (naturalmente) também apoiava sua empresa local (a Ericsson está sediada na Suécia; a Qualcomm, na Califórnia). Por fim, a Ericsson e a Qualcomm estavam envolvidas em numerosos processos relacionados a suas respectivas patentes de CDMA. Para aumentar a confusão, o UMTS se tornou um padrão 3G único com múltiplas opções incompatíveis, incluindo o CDMA2000. Essa mudança foi um esforço para unificar os vários campos, mas ele apenas encobre as diferenças técnicas e oculta o foco dos esforços em andamento. Usaremos o UMTS para indicar o WCDMA, uma forma diferente com origem no CDMA2000.

Outra melhoria em relação ao esquema CDMA básico que descrevemos anteriormente é permitir que diferentes usuários enviem dados em diferentes taxas, independentes uma da outra. Esse truque é realizado naturalmente no CDMA fixando a taxa em que os chips são transmitidos e atribuindo aos usuários sequências de chips de diferentes tamanhos. Por exemplo, no WCDMA, a taxa de chip é de 3,84 Mchips/s, e o espalhamento de códigos varia de 4 a 256 chips. Com um código de 256 chips, restam cerca

de 12 kbps após a correção de erro, e essa capacidade é suficiente para uma chamada de voz. Com um código de 4 chips, a taxa de dados do usuário é próxima de 1 Mbps. Códigos de tamanho intermediário geram taxas intermediárias; para conseguir múltiplos Mbps, a unidade móvel precisa usar mais de um canal de 5 MHz ao mesmo tempo.

Vamos voltar nossa discussão para o uso do CDMA em redes celulares, por ser o fator de distinção dos dois sistemas. O CDMA não é FDM nem TDM, mas um tipo de mistura em que cada usuário envia na mesma banda de frequência ao mesmo tempo. Quando ele foi proposto inicialmente para sistemas de celular, a indústria teve aproximadamente a mesma reação que Colombo provocou na Rainha Isabel quando ele propôs alcançar as Índias navegando na direção errada. Porém, pela persistência de uma única empresa, a Qualcomm, o CDMA teve sucesso como um sistema 2G (IS-95) e amadureceu ao ponto de se tornar a base técnica para o 3G.

Para fazer o CDMA funcionar no ambiente de telefonia móvel, é preciso mais do que a técnica de CDMA básica que descrevemos na Seção 2.4. Especificamente, descrevemos um sistema chamado **CDMA síncrono**, no qual as sequências de chips são exatamente ortogonais. Esse projeto funciona quando todos os usuários estão sincronizados no momento inicial de suas sequências de chips, como no caso da estação-base transmitindo para unidades móveis. A estação-base pode transmitir as sequências de chips começando ao mesmo tempo, de modo que os sinais sejam ortogonais e capazes de ser separados. Contudo, é difícil sincronizar as transmissões de telefones móveis independentes. Sem alguns esforços especiais, suas transmissões chegariam na estação-base em momentos diferentes, sem nenhuma garantia de ortogonalidade. Para que as unidades móveis enviem para a estação-base sem sincronização, queremos codificar sequências que são ortogonais umas às outras em todos os deslocamentos possíveis, não apenas quando elas estão alinhadas no início.

Embora não seja possível encontrar sequências exatamente ortogonais para esse caso geral, longas sequências pseudoaleatórias chegam perto o suficiente. Elas têm a propriedade de que, com alta probabilidade, tenham uma baixa **correlação cruzada** entre si em todos os deslocamentos. Isso significa que, quando uma sequência é multiplicada por outra e somada para calcular o produto interno, o resultado será pequeno; ele seria zero se todas fossem ortogonais. (Intuitivamente, as sequências aleatórias sempre deverão parecer diferentes uma da outra. Multiplicá-las deverá, então, produzir um sinal aleatório, que se somará a um resultado pequeno.) Isso permite que um receptor filtre transmissões indesejadas do sinal recebido. Além disso, a **autocorrelação** de sequências pseudoaleatórias também é pequena, com alta probabilidade, exceto em um deslocamento zero. Isso significa que, quando uma sequência é multiplicada por uma cópia atrasada de si mesma e somada, o resultado será pequeno, exceto quando o atraso é

zero. (Intuitivamente, uma sequência aleatória atrasada se parece com uma sequência aleatória diferente, e voltamos ao caso da correlação cruzada.) Isso permite que um receptor intercepte o início da transmissão desejada no sinal recebido.

O uso de sequências pseudoaleatórias permite que a estação-base receba mensagens CDMA de unidades móveis não sincronizadas. Contudo, uma suposição implícita em nossa discussão do CDMA é que os níveis de potência de todas as unidades móveis são iguais no receptor. Se não forem, uma pequena correlação cruzada com um sinal poderoso poderia superar uma grande autocorrelação com um sinal fraco. Assim, a potência de transmissão nas unidades móveis deve ser controlada para reduzir a interferência entre sinais concorrentes. É essa interferência que limita a capacidade de sistemas CDMA.

Os níveis de potência recebidos em uma estação-base dependem da distância em que os transmissores se encontram e também de quanta potência eles transmitem. Pode haver muitas estações móveis em distâncias variadas da estação-base. Uma boa heurística para equalizar a potência recebida é que cada estação móvel transmita para a estação-base no inverso do nível de potência que ela recebe dessa estação-base. Em outras palavras, uma estação móvel recebendo um sinal fraco da estação-base usará mais potência do que outra obtendo um sinal forte. Para aumentar a precisão, a estação-base também oferece feedback a cada unidade móvel para aumentar, diminuir ou manter constante sua potência de transmissão. O feedback é frequente (1.500 vezes por segundo), pois o bom controle de potência é importante para reduzir a interferência.

Agora, vamos descrever as vantagens do CDMA. Primeiro, esse sistema pode melhorar a capacidade ao tirar proveito de pequenos períodos quando alguns transmissores estão silenciosos. Nas chamadas de voz, em conversas educadas, uma parte fica em silêncio enquanto a outra fala. Em média, a linha está ocupada apenas 40% do tempo. Contudo, as pausas podem ser pequenas e são difíceis de prever. Com sistemas TDM ou FDM, não é possível retribuir slots de tempo ou canais de frequência com rapidez suficiente para se beneficiar desses pequenos silêncios. Contudo, no CDMA, um usuário reduz a interferência para outros usuários simplesmente não transmitindo, e é provável que alguma fração dos usuários não esteja transmitindo

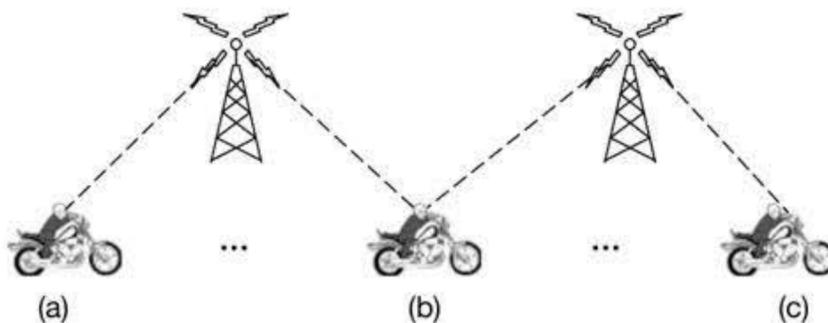
em uma célula ocupada em determinado momento. Assim, o sistema tira proveito dos silêncios esperados para permitir um número maior de chamadas simultâneas.

Em segundo lugar, com o CDMA, cada célula usa as mesmas frequências. Diferentemente de GSM e AMPS, a FDM não é necessária para separar as transmissões de diferentes usuários. Isso elimina complicadas tarefas de planejamento de frequência e melhora a capacidade, tornando mais fácil para a estação-base usar várias antenas direcionais, ou **antenas setorizadas**, em vez de uma antena omnidirecional. As antenas direcionais concentram um sinal na direção desejada e o reduzem, diminuindo, portanto, o sinal (e a interferência) em outras direções. Isso, por sua vez, aumenta a capacidade. Esquemas com três setores são comuns. A estação-base precisa rastrear a unidade móvel enquanto ela se move de um setor para outro. Esse rastreamento é fácil com o CDMA, pois todas as frequências são usadas em todos os setores.

Em terceiro lugar, o CDMA facilita o **soft handoff**, em que a unidade móvel é aceita pela nova estação-base antes de a anterior se desconectar. Desse modo, não existe perda de continuidade. O soft handoff aparece na Figura 2.43. Ele é fácil com o CDMA, pois todas as frequências são usadas em cada célula. Como opção há um **hard handoff**, em que a estação-base antiga libera a chamada antes de ela ser aceita pela nova. Se a nova não for capaz de aceitá-la (p. ex., porque não existe nenhuma frequência disponível), a chamada será desconectada de forma brusca. Os usuários tendem a notar essa interrupção, mas ela ocasionalmente é inevitável com a estrutura atual. O hard handoff é a norma nos projetos FDM para evitar o custo de fazer a unidade móvel transmitir ou receber em duas frequências simultaneamente.

## 2.6.6 Tecnologia de quarta geração (4G): comutação de pacotes

Em 2008, a ITU especificou um conjunto de padrões para sistemas 4G. O **4G**, que também costuma ser chamado de **IMT Advanced**, é totalmente baseado na tecnologia de rede de comutação de pacotes, incluindo seus predecessores. Seu predecessor imediato foi uma tecnologia normalmente chamada de **LTE (Long Term Evolution)**. Outro precursor e tecnologia relacionados ao 4G foi o 3GPP



**Figura 2.43** Soft handoff (a) antes, (b) durante e (c) depois.

LTE, também chamado de “4G LTE”. A terminologia é um pouco confusa, pois “4G” efetivamente se refere a uma geração de comunicações móveis, em que qualquer geração pode, de fato, incluir vários padrões. Por exemplo, a ITU considera o IMT Advanced como um padrão 4G, embora também aceite LTE como um padrão 4G. Outras tecnologias, como o malfadado WiMAX (IEEE 802.16), também são consideradas 4G. Tecnicamente, LTE e “4G verdadeiro” são versões diferentes do padrão 3GPP (versões 8 e 10, respectivamente).

A principal inovação em relação aos sistemas 3G anteriores é que as redes 4G utilizam comutação de pacotes, em vez de comutação de circuitos. A inovação que permite a comutação de pacotes é chamada de **EPC (Evolved Packet Core)**, que é basicamente uma rede IP simplificada, que separa o tráfego de voz da rede de dados. A rede EPC transporta voz e dados em pacotes IP. Assim, ela é uma rede de **voz sobre IP (VoIP)**, com recursos alocados usando as técnicas de multiplexação estatística descritas anteriormente. Como tal, a EPC deve gerenciar os recursos de forma que a qualidade da voz permaneça alta através dos recursos de rede que são compartilhados entre muitos usuários. Os requisitos de desempenho para LTE incluem, entre outras coisas, throughput máximo para upload de 100 Mbps e download de 50 Mbps. Para alcançar essas taxas mais altas, as redes 4G usam um conjunto de frequências adicionais, incluindo 700 MHz, 850 MHz, 800 MHz e outras. Outro aspecto do padrão 4G é a “eficiência espectral”, ou quantos bits podem ser transmitidos por segundo para uma determinada frequência; para tecnologias 4G, a eficiência espectral máxima deve ser 15 bps/Hz para downlink e 6,75 bps/GHz para uplink.

A arquitetura LTE inclui os seguintes elementos como parte do EPC, conforme mostrado no Capítulo 1, na Figura 1.19.

1. **Serving Gateway (S-GW)**. O S-GW encaminha pacotes de dados para garantir que os pacotes continuem a ser encaminhados para o dispositivo do usuário ao passar de um eNodeB para outro.
2. **Mobility Management Entity (MME)**. O MME rastreia, pagina o dispositivo do usuário e escolhe o S-GW para um dispositivo quando ele se conecta à rede pela primeira vez, assim como durante as transferências. Ele também autentica o dispositivo do usuário.
3. **Packet Data Network Gateway (P-GW)**. Faz a interface entre o dispositivo do usuário e uma rede de dados de pacote (ou seja, uma rede comutada por pacote) e pode executar funções como alocação de endereço para essa rede (p. ex., via DHCP), limitação de taxa, filtragem, inspeção profunda de pacotes e interceptação legal do tráfego. Os dispositivos do usuário estabelecem um serviço orientado à conexão com o gateway de pacotes usando o chamado **portador de EPS**, estabelecido quando o dispositivo do usuário se conecta à rede.
4. **Home Subscriber Server (HSS)**. O MME consulta o HSS para determinar se o dispositivo do usuário corresponde a um assinante válido.

A rede 4G também possui uma **RAN (Radio Access Network)** expandida. A rede de acesso por rádio para LTE introduz um nó de acesso denominado **eNodeB**, que realiza operações na camada física (como enfocaremos neste capítulo), bem como as camadas **MAC (Medium Access Control)**, **RLC (Radio Link Control)** e **PDCP (Packet Data Control Protocol)**, muitas delas específicas para a arquitetura de rede celular. O eNodeB executa o gerenciamento de recursos, o controle de admissão, a programação e outras funções do plano de controle.

Nas redes 4G, o tráfego de voz pode ser transportado pelo EPC usando uma tecnologia chamada **VoLTE (Voice over LTE)**, possibilitando às operadoras transmitir o tráfego de voz pela rede comutada por pacotes e tirando qualquer dependência da rede de voz legada, comutada por circuitos.

### 2.6.7 Tecnologia de quinta geração (5G)

Por volta de 2014, o sistema LTE atingiu a maturidade e as pessoas começaram a pensar no que viria em seguida. Obviamente, depois de 4G vem o 5G. A verdadeira questão, claro, é “o que será o 5G?”, que Andrews et al. (2014) discutem bastante. Anos depois, 5G passou a significar muitas coisas diferentes, dependendo do público e de quem está usando o termo. Basicamente, a próxima geração de tecnologia de rede celular móvel se resume a dois fatores principais: taxas de dados mais altas e latência menor do que as tecnologias 4G. É claro que existem tecnologias específicas que permitem alcançar maior velocidade e menor latência, que discutiremos a seguir.

O desempenho da rede celular é geralmente medido em termos de **taxa de dados agregada** ou **capacidade de área**, que é a quantidade total de dados que a rede pode atender em bits por unidade de área. Um dos objetivos do 5G é melhorar a capacidade de área da rede em três ordens de grandeza (mais de 1000 vezes a do 4G), usando uma combinação de tecnologias:

1. **Ultra-densificação e descarregamento**. Uma das maneiras mais simples de melhorar a capacidade da rede é adicionar mais células por área. Enquanto os tamanhos das células 1G eram da ordem de centenas de quilômetros quadrados, o 5G visa tamanhos de células menores, incluindo **picocélulas** (com menos de 100 m de diâmetro) e até mesmo **femtocélulas** (com alcance semelhante ao WiFi, de dezenas de metros). Um dos benefícios mais importantes da redução do tamanho da célula é a capacidade de reutilizar o espectro em

uma determinada área geográfica, reduzindo assim o número de usuários que estão competindo por recursos em qualquer estação-base. É claro que a redução do tamanho da célula traz suas próprias complicações, incluindo maior dificuldade para gerenciamento de mobilidade e handoff.

2. Largura de banda aumentada com ondas milimétricas. A maior parte do espectro de tecnologias anteriores estava na faixa de várias centenas de MHz a alguns GHz, correspondendo a comprimentos de onda que variam de centímetros a cerca de um metro. Esse espectro tem se tornado cada vez mais congestionado, especialmente nos principais mercados durante os horários de pico. Existem quantidades consideráveis de espectro não utilizado na faixa de ondas milimétricas de 20 a 300 GHz, com comprimentos de onda menores que 10 mm. Até pouco tempo, esse espectro não era considerado adequado para comunicação sem fio porque comprimentos de onda mais curtos não se propagam tão bem. Uma das maneiras como os desafios de propagação estão sendo enfrentados é usando grandes matrizes de antenas direcionais, o que é uma grande mudança arquitetônica em relação às gerações anteriores de redes celulares: tudo é diferente, desde propriedades de interferência até o processo de associar um usuário a uma estação-base.
3. Maior eficiência espectral por meio de avanços na tecnologia **MIMO (Multiple-Input Multiple-Output)**. MIMO melhora a capacidade de um enlace de rádio, usando várias antenas de transmissão e recepção para tirar vantagem da propagação de multicaminhos, através da qual o sinal de rádio transmitido chega ao receptor por meio de dois ou mais caminhos. MIMO foi introduzido na comunicação WiFi e nas tecnologias de celular 3G por volta de 2006, e possui algumas variações; os padrões celulares anteriores tiram proveito do **MU-MIMO (Multi-User MIMO)**. Geralmente, essas tecnologias aproveitam a diversidade espacial dos usuários para cancelar a interferência que pode ocorrer em qualquer uma das extremidades da transmissão sem fio. **Massive MIMO** é um tipo de MU-MIMO que aumenta o número de antenas da estação-base para que haja muito mais antenas do que terminais. Existe até a possibilidade de usar um arranjo de antenas tridimensionais, no chamado **FD-MIMO (Full-Dimension MIMO)**.

Outro recurso que acompanhará o 5G é o **fatiamento de rede**, que permitirá às operadoras de celular criarem várias redes virtuais na mesma infraestrutura física compartilhada, dedicando partes de sua rede a casos de uso específicos do cliente. Frações distintas da rede (e seus recursos) podem ser dedicadas a diferentes provedores de aplicações, as quais, por sua vez, podem ter diferentes requisitos. Por exemplo, aplicações que exigem alto throughput podem ser

alocadas a uma fatia de rede diferente daquelas que admitem um throughput mais baixo. **SDN (Software-Defined Networking)** and **NFV (Network Functions Virtualization)** são tecnologias emergentes que ajudarão a dar suporte ao fatiamento. Discutiremos essas tecnologias em capítulos mais adiante.

## 2.7 REDES POR CABO

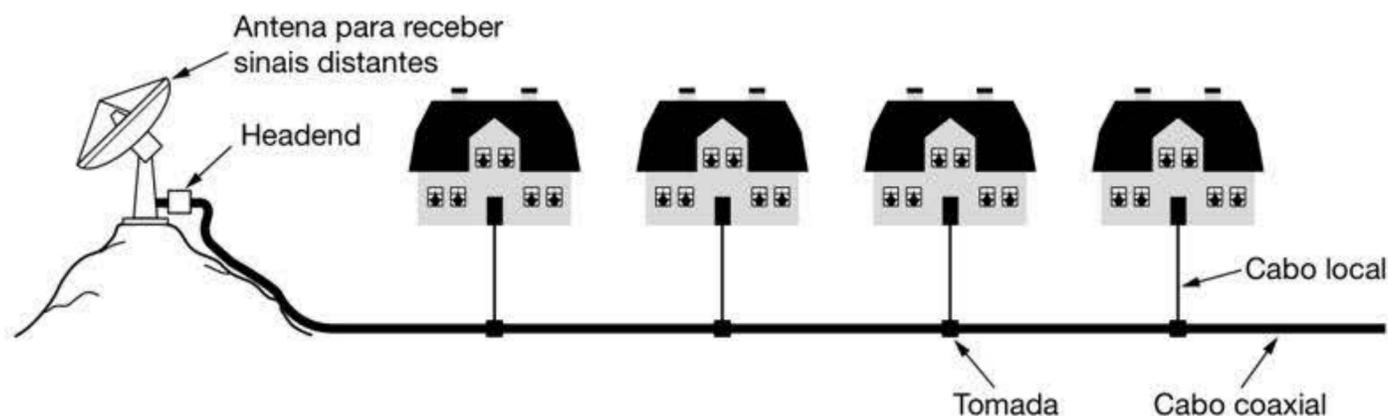
Certamente, os sistemas de telefonia fixa e móvel desempenharão um papel importante nas futuras redes, mas as redes por cabo também participarão das redes de acesso de banda larga do futuro. Hoje mesmo, muitas pessoas recebem serviços de televisão, telefone e Internet via cabo. Nas próximas seções, examinaremos com mais detalhes a televisão a cabo como uma rede e vamos compará-la com os sistemas de telefonia que acabamos de estudar. Para obter mais informações, consulte Harte (2017). O padrão DOCSIS de 2018 também tem informações úteis, principalmente em relação às modernas arquiteturas de rede por cabo.

### 2.7.1 História das redes por cabo: TV por antena comunitária

A televisão a cabo foi concebida no final da década de 1940 como uma forma de proporcionar melhor recepção de TV às pessoas que viviam em áreas rurais ou montanhosas. No início, o sistema consistia em uma grande antena situada no alto de uma colina para captar o sinal de televisão que se propaga pelo ar, um amplificador chamado **headend**, para reforçar o sinal, e um cabo coaxial para distribuí-lo pelas casas das pessoas, como ilustra a Figura 2.44.

Nos primeiros anos, a TV a cabo era chamada de **televisão de antena comunitária**, ou **CATV (Community Antenna Television)**. Sua operação era muito simples – qualquer pessoa que tivesse alguma prática em eletrônica era capaz de instalar um serviço para sua cidade, e os usuários se reuniam para pagar os custos. À medida que o número de assinantes crescia, outros cabos eram conectados ao cabo original e eram acrescentados outros amplificadores conforme a necessidade. A transmissão era unidirecional, do headend para os usuários. Em 1970, havia milhares de sistemas independentes.

Em 1974, a Time Inc. lançou um novo canal, denominado Home Box Office, com novo conteúdo (filmes) e distribuído somente por cabo. Seguiram-se outros canais dedicados apenas a notícias, esportes, culinária, história, filmes, ciência, crianças e muitos outros temas. Esse desenvolvimento ocasionou duas mudanças na indústria. Primeiro, as grandes corporações começaram a adquirir os sistemas a cabo existentes e estender novos cabos para conquistar novos assinantes. Segundo, agora havia necessidade



**Figura 2.44** Um antigo sistema de TV a cabo.

de conectar vários sistemas, normalmente em cidades distantes, a fim de distribuir o conteúdo dos novos canais de TV a cabo. As empresas especializadas começaram a estender cabos entre as cidades para conectar todas elas em um único sistema. Esse padrão era semelhante ao que ocorreu na indústria de telefonia 80 anos antes, com a conexão de estações finais anteriormente isoladas para tornar possível a comunicação interurbana.

### 2.7.2 Acesso de banda larga à Internet por cabo: redes HFC

Com o passar dos anos, o sistema de TV a cabo cresceu, e os cabos entre as várias cidades foram substituídos por fibra óptica de alta largura de banda, de forma semelhante ao que ocorreu no sistema telefônico. Um sistema com fibra nas linhas principais e cabo coaxial nas ligações para as residências é chamado de **sistema híbrido de cabo coaxial e fibra**, ou **HFC (Hybrid Fiber Coax)**, e é a arquitetura predominante nas redes a cabo atuais. A tendência de levar a fibra para mais perto do assinante doméstico continua, conforme descrevemos na seção sobre FTTH. Os conversores eletro-ópticos que constituem a interface entre as partes óptica e elétrica do sistema são chamados **nós de fibra**. Pelo fato de a largura de banda da fibra ser muito maior que a dos cabos coaxiais, um único nó de fibra pode alimentar vários cabos coaxiais. A Figura 2.45(a) mostra uma parte de um sistema HFC moderno.

No final da década de 1990, muitas operadoras de TV a cabo decidiram entrar no ramo de acesso à Internet e, muitas vezes, também no ramo de telefonia. No entanto, diferenças técnicas entre as instalações de cabo e de telefonia têm efeito sobre o que deve ser realizado para alcançar esses objetivos. Por um lado, todos os amplificadores unidirecionais no sistema tinham de ser substituídos por amplificadores bidirecionais, para dar suporte a transmissões tanto upstream quanto downstream. Enquanto isso estava acontecendo, os primeiros sistemas de Internet por cabo usavam a rede de televisão para as transmissões downstream e uma conexão discada via rede telefônica para as transmissões upstream. Essa foi uma alternativa arriscada, se é que houve uma, mas até que funcionou.

Eliminar todos os canais de TV e usar a infraestrutura de cabo estritamente para acesso à Internet provavelmente geraria um número razoável de clientes insatisfeitos (principalmente os mais antigos, pois os mais novos já haviam rompido os laços); assim, as empresas de TV a cabo hesitam em fazê-lo. Além disso, a maioria das cidades tem uma regulamentação bastante pesada sobre o que é transmitido por cabo e, portanto, as operadoras de serviços não teriam permissão para fazer isso, ainda que desejassem. Como consequência, elas precisaram encontrar um modo de fazer a televisão e a Internet coexistirem no mesmo cabo.

A solução é contar com a FDM. Os canais de TV a cabo da América do Norte ocupam a região de 54 a 550 MHz (com exceção do rádio FM, que ocupa a faixa de 88 a 108 MHz). Esses canais têm 6 MHz de largura, incluindo as bandas de proteção, e podem transportar um canal de TV analógica tradicional ou vários canais de TV digital. Na Europa, a extremidade inferior em geral é de 65 MHz, e os canais têm de 6 a 8 MHz de largura, em virtude da maior resolução exigida pelos sistemas PAL e SECAM, mas o esquema de alocação é semelhante nos outros aspectos. A parte baixa da banda não é usada. Os cabos modernos também operam bem acima de 550 MHz, chegando frequentemente a 750 MHz ou mais. A solução escolhida foi introduzir canais de upstream na banda de 5 a 42 MHz (um pouco mais alta na Europa) e usar as frequências na extremidade alta para os sinais downstream. O espectro dos serviços de cabo é ilustrado na Figura 2.46.

Como todos os sinais de televisão são downstream, é possível usar amplificadores upstream que só funcionam na região de 5 a 42 MHz e amplificadores downstream que só funcionam na frequência de 54 MHz e acima desta, como mostra a figura. Desse modo, obtemos uma assimetria nas larguras de banda upstream e downstream, porque está disponível uma parte maior do espectro acima da faixa de TV do que abaixo dela. Em contrapartida, a maior parte do tráfego provavelmente será downstream, e, assim, as operadoras de serviços a cabo não estão insatisfeitas com esse fato. Como vimos, em geral as companhias telefônicas oferecem um serviço DSL assimétrico, embora não tenham nenhuma razão técnica para fazê-lo. Além de atualizar os

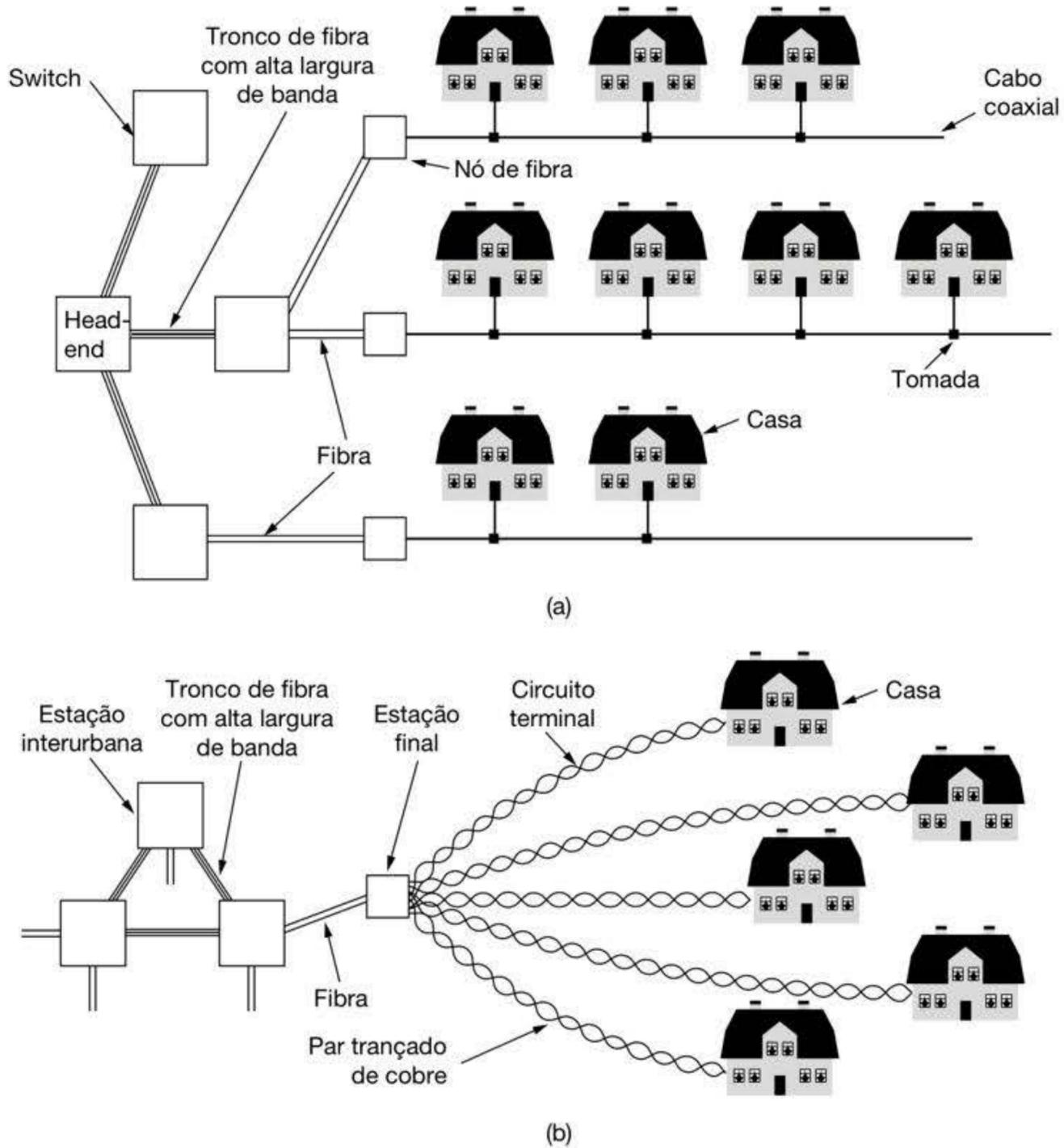


Figura 2.45 (a) Rede híbrida de cabo coaxial e fibra. (b) Sistema de telefonia fixa.

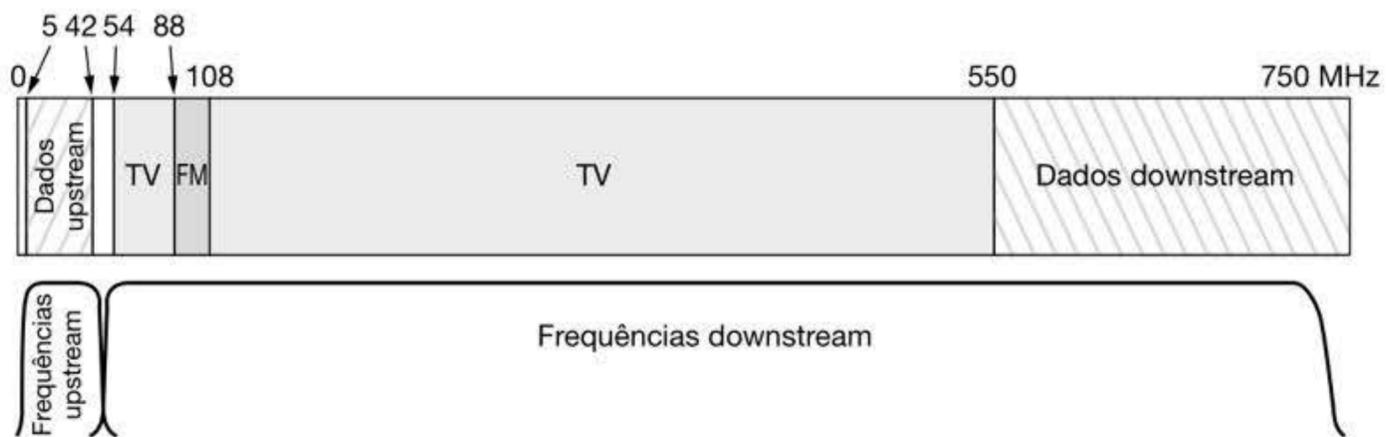


Figura 2.46 Alocação de frequências em um sistema típico de TV a cabo usado para acesso à Internet.

amplificadores, a operadora também tem de atualizar o headend, que deve passar de um amplificador não inteligente para um sistema computadorizado digital inteligente, com uma interface de fibra de alta largura de banda para um ISP.

Esse headend atualizado passou a se chamar **sistema de terminação de modem a cabo**, ou **CMTS (Cable Modem Termination System)**. CMTS e headend referem-se ao mesmo componente.

### 2.7.3 DOCSIS

As companhias de cabo operam redes que incluem a tecnologia da camada física HFC para a conectividade do último quilômetro, além de conexões por fibra e sem fio para o último quilômetro. A parte HFC dessas redes é bastante empregada nos Estados Unidos, Canadá, Europa e outros mercados, e usa os padrões **DOCSIS (Data Over Cable Service Interface Specification)** da CableLabs.

O padrão DOCSIS versão 1.0 apareceu em 1997. DOCSIS 1.0 e 1.1 tinham um limite funcional de 38 Mbps downstream e 9 Mbps upstream. Em 2001, DOCSIS 2.0 triplicou a largura de banda upstream. Mais tarde, DOCSIS 3.0 (2006) introduziu o suporte para IPv6 e permitiu a ligação de canais para as comunicações downstream e upstream, aumentando bastante a capacidade potencial para cada residência atendida para centenas de megabits por segundo. DOCSIS 3.1 (2013), que introduziu a Orthogonal Frequency Division Multiplexing (OFDM), largura de banda mais larga para os canais e maior eficiência, permitiu mais de 1 Gbps de capacidade downstream por residência. Extensões ao DOCSIS 3.1 foram acrescentadas por meio de atualizações ao padrão DOCSIS 3.1, incluindo a operação Full Duplex (2017), que permitirá capacidades downstream e upstream simétricas de vários gigabits, além do DOCSIS Low Latency (2018) e outros recursos para reduzir a latência.

Na camada de fibra coaxial híbrida, ou HFC, a rede é altamente dinâmica, com operadoras de rede a cabo realizando divisões de nó de fibra regularmente, o que empurra a fibra para mais perto da casa e reduz o número de residências atendidas por cada nó, disponibilizando mais capacidade para cada casa atendida. Em alguns casos, a HFC do último quilômetro é substituído por FTTH, e muitas novas construções também são para uso doméstico.

Os assinantes de Internet a cabo exigem um modem a cabo DOCSIS para servir como interface entre a rede doméstica e a rede do ISP. Cada modem a cabo envia dados em um canal upstream e um canal downstream, e cada um é alocado usando FDM. DOCSIS 3.0 utiliza vários canais. O esquema normal é apanhar cada canal downstream de 6 ou 8 MHz e modulá-lo com QAM-64 ou, se a qualidade do cabo for excepcionalmente boa, com QAM-256. Com um canal de 6 MHz e QAM-64, obtemos cerca de 36 Mbps. Quando o overhead da sinalização é subtraído, a largura de banda resultante é de cerca de 27 Mbps. Com QAM-256, a carga útil resultante é de cerca de 39 Mbps. Os valores europeus são 1/3 maiores devido à maior disponibilidade de largura de banda.

A interface entre o modem e a residência é simples; em geral, ela é uma conexão Ethernet. Hoje em dia, muitos usuários domésticos da Internet conectam o modem a cabo a um ponto de acesso WiFi para montar uma rede sem fio doméstica. Em alguns casos, o provedor de serviços de Internet (ISP) do usuário fornece um único

dispositivo de hardware que combina o modem a cabo e o ponto de acesso sem fio. A interface entre o modem a cabo e o restante da rede ISP é mais complicada, pois envolve a coordenação do compartilhamento de recursos entre muitos assinantes de cabo que podem estar conectados ao mesmo headend. Esse compartilhamento de recursos ocorre tecnicamente na camada de enlace, não na camada física, mas iremos abordá-lo neste capítulo para garantir a continuidade.

### 2.7.4 Compartilhamento de recursos em redes DOCSIS: nós e minislots

Há uma diferença fundamental importante entre o sistema HFC da Figura 2.45(a) e o sistema telefônico da Figura 2.45(b). Em determinados bairros, um único cabo é compartilhado por muitas casas, ao passo que, no sistema telefônico, cada casa tem seu próprio circuito terminal privado. Quando é utilizado para difusão de televisão, esse compartilhamento é natural. Todos os programas são transmitidos no cabo e não importa se existem 10 ou 10.000 espectadores. No entanto, quando o mesmo cabo é usado para acesso à Internet, faz uma grande diferença a existência de 10 ou de 10.000 usuários. Se um usuário decidir baixar um arquivo muito grande ou fazer o stream de um filme de 8K, essa largura de banda não estará disponível a outros usuários. Quanto mais usuários compartilhando um único cabo, maior a competição pela largura de banda. O sistema de telefonia não tem essa propriedade específica: a transferência de um grande arquivo por uma linha ADSL não reduz a largura de banda do seu vizinho. Contudo, a largura de banda do cabo coaxial é muito mais alta do que a dos pares trançados. Basicamente, a largura de banda que qualquer assinante recebe em determinado momento depende muito do uso dos outros assinantes que estejam compartilhando o mesmo cabo, conforme descrevemos com mais detalhes a seguir.

A estratégia usada pelos ISPs para resolver esse problema é desmembrar cabos longos e conectar cada um deles diretamente a um nó de fibra. A largura de banda do headend até cada nó de fibra é significativa e, se não existirem muitos assinantes em cada segmento de cabo, o volume de tráfego será gerenciável. Os cabos típicos há cerca de 10 a 15 anos conectavam de 500 a 2.000 casas, embora o número de casas por nó continue a diminuir conforme a implantação até o limite continua, em um esforço para aumentar a velocidade para os assinantes. O aumento de assinantes de Internet a cabo na última década, juntamente com a crescente demanda de tráfego, criou a necessidade cada vez maior de dividir esses cabos e adicionar mais nós de fibra. Em 2019, um tamanho de nó típico era de cerca de 300-500 casas, embora em algumas áreas os ISPs estejam construindo arquiteturas HFC N+0 (também conhecidas como “Fiber Deep”), o que pode reduzir esse

número para até 70, eliminando a necessidade de amplificadores de sinal em cascata e de passar cabos de fibra direto dos headends da rede para os nós no último segmento do cabo coaxial.

Quando um modem a cabo é conectado e ligado, ele percorre os canais downstream procurando por um pacote especial emitido periodicamente pelo headend para fornecer parâmetros de sistema aos modems que acabaram de se conectar. Ao encontrar esse pacote, o novo modem anuncia sua presença em um dos canais upstream. O headend responde atribuindo o modem a seus canais upstream e downstream. Essas atribuições podem ser alteradas mais tarde, se o headend julgar necessário equilibrar a carga.

Para upstream, há mais ruído de ondas de rádio, pois o sistema não foi projetado originalmente para dados, e o ruído de vários assinantes é afunilado no headend, de modo que é utilizado um esquema mais conservador. Este varia de QPSK a QAM-128, em que alguns dos símbolos são usados para proteção de erro com a modulação codificada por treliças. Com menos bits por símbolo no upstream, a assimetria entre as taxas upstream e downstream é muito mais do que é sugerido pela Figura 2.46.

Os modems DOCSIS atuais exigem um tempo para transmitir, e então o CMTS concede um ou mais períodos de tempo para que o modem possa transmitir, com base na disponibilidade; todos os usuários simultâneos disputam pelo acesso upstream e downstream. A rede usa TDM para compartilhar largura de banda no upstream para diversos assinantes. O tempo é dividido em **minislots**, e diferentes assinantes podem enviar em diferentes minislots. O headend anuncia periodicamente o início de uma nova rodada de minislots, mas o tiro de partida não é ouvido em todos os modems ao mesmo tempo, em virtude do tempo de propagação no cabo. Conhecendo a que distância está do headend, cada modem pode calcular há quanto tempo o primeiro minislot realmente começou.

É importante que o modem saiba sua distância até o headend para obter a sincronização correta. Primeiro, o modem determina sua distância até o headend, enviando-lhe um pacote especial e verificando quanto tempo demora para receber a resposta. Esse processo é chamado de **verificação do alcance** (ou **ranging**). Cada pacote upstream deve caber em um ou mais minislots consecutivos no headend quando é recebido. A extensão do minislot depende da rede. Uma carga útil típica é de 8 bytes.

Durante a inicialização, o headend também atribui cada modem a um minislot, que será usado para solicitar largura de banda upstream. Quando um computador quer enviar um pacote, ele o transfere ao modem, que então solicita o número necessário de minislots. Se a solicitação for aceita, o headend colocará uma confirmação no canal downstream, informando ao modem quais minislots foram reservados para seu pacote. Este é então enviado, a partir do minislot alocado a ele. Pacotes adicionais podem ser solicitados com a utilização de um campo no cabeçalho.

Em geral, vários modems receberão o mesmo minislot, o que leva à disputa (vários modems tentando enviar dados upstream ao mesmo tempo). O CDMA pode permitir que vários assinantes compartilhem o mesmo minislot, embora com uma taxa reduzida por usuário. A segunda opção é que o CDMA não seja utilizado, quando não haverá confirmação, em decorrência de uma colisão. Nesse caso, quando ocorrem colisões, o modem simplesmente esperará um tempo aleatório e tentará de novo. Após cada falha sucessiva, esse tempo aleatório é dobrado. (Para os leitores que já estão um pouco familiarizados com as redes, esse algoritmo é simplesmente o modelo ALOHA adotado com a recuperação de erro por backoff exponencial binário. A Ethernet não pode ser usada em redes a cabo, pois as estações não conseguem detectar o meio compartilhado. Voltaremos a analisar essas questões no Capítulo 4.)

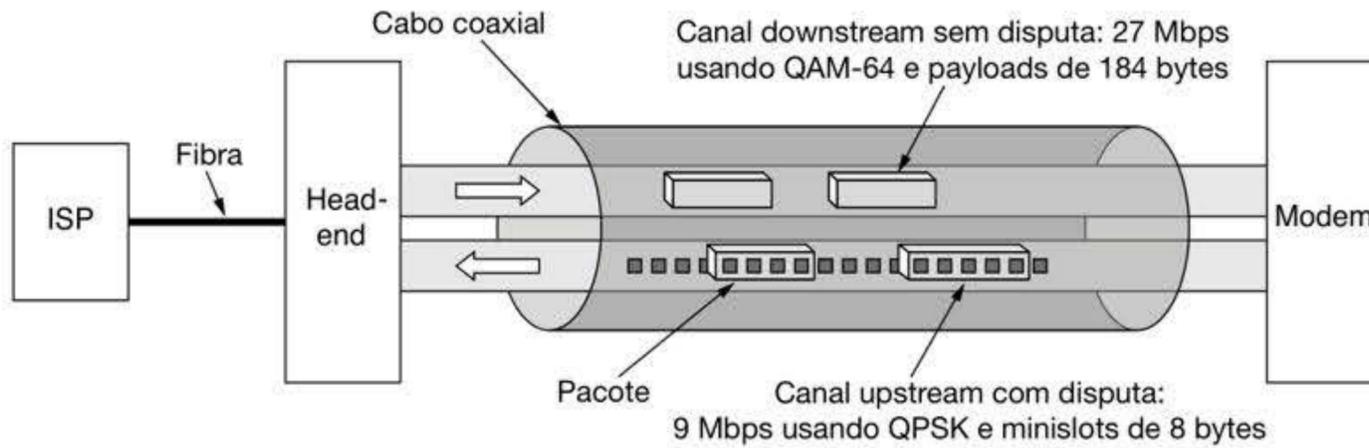
Os canais downstream são gerenciados de modo diferente dos canais upstream. Por um lado, só existe um transmissor (o headend) e, assim, não há disputa nem a necessidade de minislots. Por outro lado, o tráfego downstream em geral é muito maior que o upstream, e então é usado um tamanho de pacote fixo, de 204 bytes. Uma parte desse código é um código de correção de erros Reed-Solomon e algumas outras fontes de overhead, restando uma carga útil do usuário igual a 184 bytes. Esses números foram escolhidos para manter a compatibilidade com a televisão digital usando MPEG-2, de forma que os canais de TV e os canais de dados downstream sejam formatados de maneira idêntica. As conexões lógicas estão representadas na Figura 2.47.

## 2.8 SATÉLITES DE COMUNICAÇÕES

Na década de 1950 e no início dos anos 1960, as pessoas tentavam montar sistemas de comunicações emitindo sinais que se refletiam em balões meteorológicos metalizados. Infelizmente, os sinais recebidos eram muito fracos para que tivessem algum uso prático. Em seguida, a Marinha dos Estados Unidos detectou uma espécie de balão meteorológico que ficava permanentemente no céu – a Lua – e criou um sistema operacional para comunicações entre o navio e a base, utilizando a Lua em suas transmissões.

O progresso no campo da comunicação celeste precisou esperar até que o primeiro satélite de comunicações fosse lançado. A principal diferença entre um satélite artificial e um real é que o artificial pode amplificar os sinais antes de enviá-los de volta, transformando uma estranha curiosidade em um poderoso sistema de comunicações.

Os satélites de comunicações têm algumas propriedades interessantes, que os tornam atraentes para muitas aplicações. Em sua forma mais simples, um satélite de comunicações pode ser considerado um grande repetidor de micro-ondas no céu. Ele contém diversos **transponders**, cada um deles ouve uma parte do espectro, amplifica os



**Figura 2.47** Detalhes típicos dos canais upstream e downstream na América do Norte.

sinais de entrada e os transmite novamente em outra frequência, para evitar interferência com o sinal de entrada. Esse modo de operação é conhecido como um **canal em curva (bent pipe)**. O processamento digital pode ser acrescentado para manipular ou redirecionar separadamente os feixes de dados na banda geral, ou informações digitais ainda podem ser recebidas pelo satélite e retransmitidas. A regeneração de sinais dessa maneira melhora o desempenho em comparação com um canal em curva, pois o satélite não amplifica o ruído no sinal ascendente. Os feixes descendentes podem ser largos, cobrindo uma fração substancial da superfície terrestre, ou estreitos, cobrindo uma área com apenas centenas de quilômetros de diâmetro.

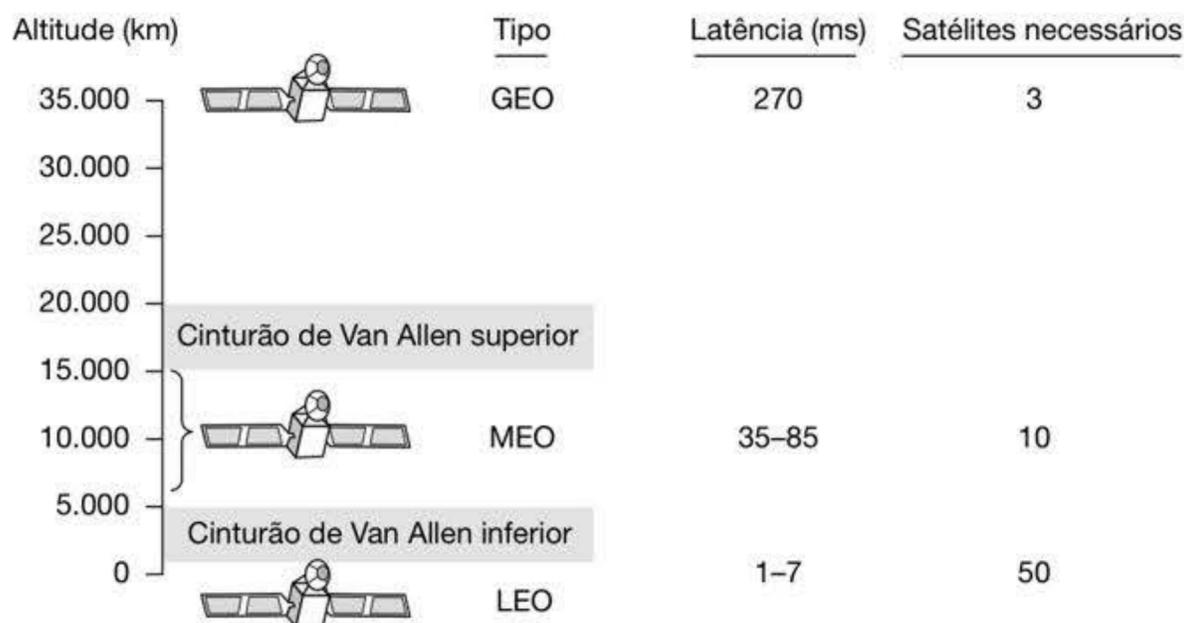
De acordo com a lei de Kepler, o período orbital de um satélite varia de acordo com o raio da órbita elevado à potência 3/2. Quanto mais alto o satélite, mais longo o período. Perto da superfície da Terra, o período é de cerca de 90 minutos. Consequentemente, os satélites de órbita baixa saem de visão com bastante rapidez (devido ao movimento dos satélites); assim, são necessários muitos deles para proporcionar cobertura contínua, e as antenas terrestres precisam acompanhá-los. A uma altitude de aproximadamente

35.800 km, o período é de 24 horas. Em uma altitude de 384.000 km, o período é de cerca de um mês, como pode atestar qualquer pessoa que observe a Lua regularmente.

O período do satélite é importante, mas não é o único fator para determinar onde posicioná-lo. Outra questão é a presença dos cinturões de Van Allen, camadas de partículas altamente carregadas que são capturadas pelo campo magnético terrestre. Qualquer satélite em órbita dentro deles seria destruído com bastante rapidez pelas partículas. Esses fatores nos levam a identificar três regiões nas quais os satélites podem ser posicionados com segurança. Essas regiões e algumas de suas propriedades estão ilustradas na Figura 2.48. A seguir, descreveremos rapidamente os satélites que habitam cada uma dessas regiões.

### 2.8.1 Satélites geoestacionários

Em 1945, o escritor de ficção científica Arthur C. Clarke calculou que um satélite na altitude de 35.800 km em uma órbita circular equatorial pareceria permanecer imóvel no céu, e assim não precisaria ser rastreado (Clarke, 1945). Ele continuou descrevendo um sistema de comunicação



**Figura 2.48** Satélites de comunicações e algumas de suas propriedades, inclusive altitude acima da Terra, tempo de atraso de ida e volta, e o número de satélites necessários para cobertura global.

completo que usava esses **satélites geoestacionários** (tripulados), incluindo as órbitas, os painéis solares, as frequências de rádio e os procedimentos de lançamento. Infelizmente, ele concluiu que os satélites eram impraticáveis em virtude da impossibilidade de colocar em órbita amplificadores a válvulas, frágeis e gastadores de energia; assim, nunca levou sua ideia adiante, embora tenha escrito algumas histórias de ficção científica sobre ela.

A invenção do transistor mudou tudo isso, e o primeiro satélite artificial de comunicações, chamado Telstar, foi lançado em julho de 1962. Desde então, os satélites de comunicações se transformaram em um negócio de vários bilhões de dólares, e o único aspecto do espaço sideral que se tornou altamente lucrativo. Esses satélites de alta órbita normalmente são chamados **satélites geoestacionários**, ou **GEO (Geostationary Earth Orbit)**.

Com a tecnologia atual, não é muito inteligente ter satélites geoestacionários com espaçamento muito menor que 2 graus entre eles no plano equatorial de 360 graus, a fim de evitar interferência. Com um espaçamento de 2 graus, só pode haver  $360/2 = 180$  desses satélites no céu ao mesmo tempo. No entanto, cada transponder pode usar várias frequências e polarizações, com a finalidade de aumentar a largura de banda disponível.

Para evitar o caos total no céu, a alocação de slots de órbitas é feita pela ITU. Esse processo é altamente político, com países que mal saíram da idade da pedra exigindo “seus” slots de órbitas (com a finalidade de arrendá-los pela melhor oferta). Contudo, outros países sustentam que os direitos nacionais de propriedade não se estendem para cima até a Lua e que nenhum país tem direito legal sobre os slots de órbita acima de seu território. Para aumentar a disputa, as telecomunicações comerciais não são a única aplicação. Emissoras de televisão, governos e instituições militares também querem ter uma fatia dessa torta orbital.

Os satélites modernos podem ser muito grandes, pesando até 5.000 kg e consumindo vários quilowatts de energia elétrica produzida pelos painéis solares. Os efeitos das gravidades solar, lunar e planetária tendem a movê-los para fora de seus slots de órbita e de suas orientações, um efeito compensado por motores de foguetes a bordo. Essa atividade de ajuste fino é chamada de **manutenção da estação**. Porém, quando o combustível para os motores tiver

se esgotado (em geral no período de 10 anos), o satélite fica sem controle, e, portanto, tem de ser desativado. Por fim, a órbita decai, o satélite entra de novo na atmosfera e é totalmente queimado ou (muito raramente) colide com a Terra.

Os slots de órbita não são o único ponto de discórdia. As frequências também o são, porque as transmissões do satélite para a Terra (downlink) interferem com usuários de micro-ondas. Conseqüentemente, a ITU alocou certas bandas de frequência para usuários de satélites. As principais estão listadas na Figura 2.49. A banda C foi a primeira a ser designada para tráfego comercial de satélite. Duas faixas de frequências são atribuídas nessa banda, a inferior para tráfego downlink (descendo do satélite) e a superior para tráfego uplink (subindo para o satélite). Para permitir que o tráfego ocorra em ambos os sentidos ao mesmo tempo, são necessários dois canais, os quais já estão sobrecarregados, porque também são usados pelas concessionárias de telecomunicações nos enlaces terrestres de micro-ondas. As bandas L e S foram acrescentadas por um acordo internacional em 2000. No entanto, elas são estreitas e também estão lotadas.

A próxima banda mais alta disponível para concessionárias de telecomunicações comerciais é a Ku (K under). Essa banda (ainda) não está congestionada e, em suas frequências mais altas, os satélites podem ficar à distância de apenas 1 grau, com velocidades de transmissão podendo atingir mais de 500 Mbps. Entretanto, existe outro problema: a chuva. A água absorve bastante essas micro-ondas curtas. Felizmente, em geral as tempestades fortes costumam ser localizadas; assim, o uso de várias estações terrestres separadas por uma grande distância, em lugar de apenas uma, contorna o problema, mas ao custo de antenas, cabos e equipamentos eletrônicos extras para permitir a comutação rápida entre estações. Na banda Ka (K above), também foi alocada uma largura de banda para o tráfego de satélite comercial, mas o equipamento necessário para usá-la é caro. Além dessas bandas comerciais, também existem muitas bandas governamentais e militares.

Um satélite moderno tem cerca de 40 transponders, a maioria normalmente com uma largura de banda de 36 MHz. Em geral, cada transponder opera como um canal em curva, mas satélites recentes têm alguma capacidade de processamento a bordo, permitindo uma operação mais

Banda	Downlink	Uplink	Largura de banda	Problemas
L	1,5 GHz	1,6 GHz	15 MHz	Baixa largura de banda; lotada
S	1,9 GHz	2,2 GHz	70 MHz	Baixa largura de banda; lotada
C	4,0 GHz	6,0 GHz	500 MHz	Interferência terrestre
Ku	11 GHz	14 GHz	500 MHz	Chuva
Ka	20 GHz	30 GHz	3.500 MHz	Chuva; custo do equipamento

**Figura 2.49** Principais bandas de satélite.

sofisticada. Nos primeiros satélites, a divisão dos transponders em canais era estática: a largura de banda simplesmente era dividida em bandas de frequências fixas. Hoje em dia, o feixe de cada transponder é dividido em slots de tempo, com diversos usuários alternando atividades. Mais uma vez, vemos como TDM e FDM são usados em muitos contextos.

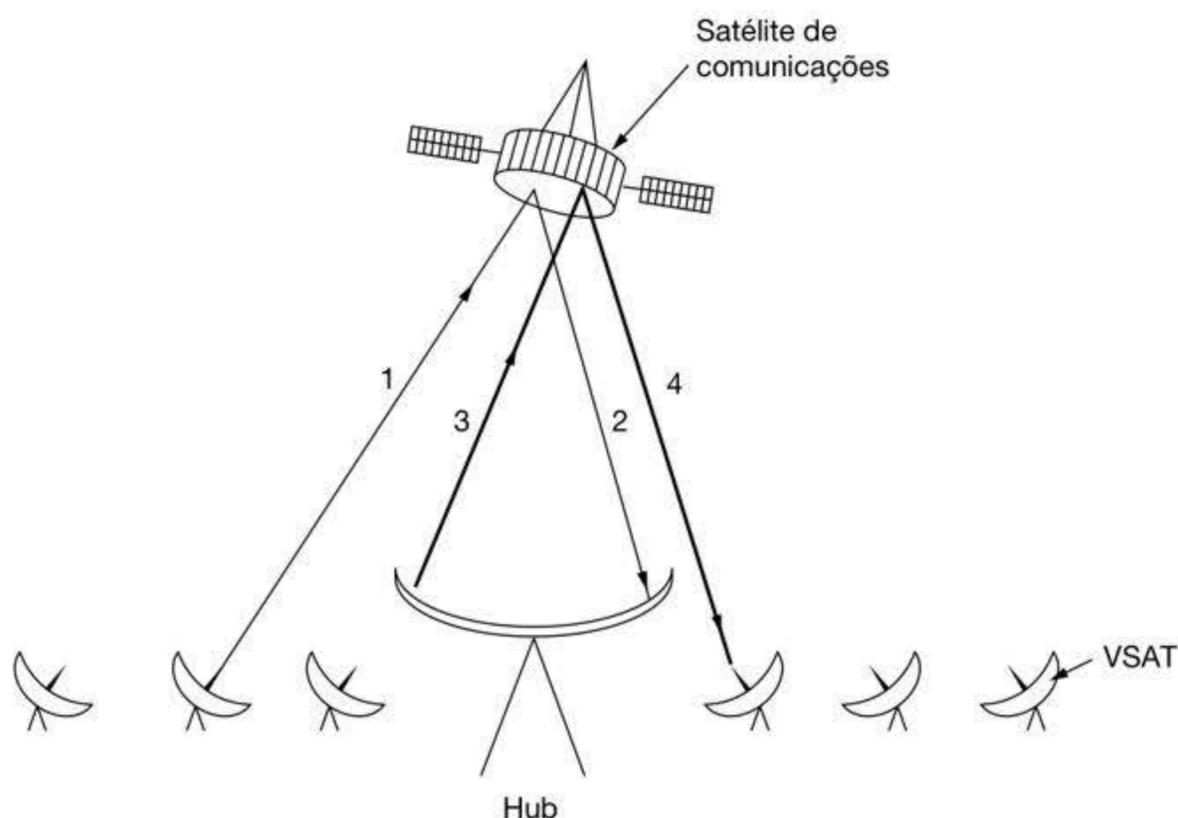
Os primeiros satélites geoestacionários tinham um único feixe espacial que iluminava cerca de 1/3 da superfície da Terra, denominado sua **área de cobertura (footprint)**. Com o enorme declínio de preço, tamanho e requisitos de potência dos equipamentos microeletrônicos, uma estratégia de transmissão muito mais sofisticada tornou-se viável. Cada satélite é equipado com diversas antenas e vários transponders. Cada feixe descendente pode ser focalizado em uma pequena área geográfica; portanto, podem acontecer diversas transmissões ascendentes e descendentes ao mesmo tempo. Em geral, esses chamados **feixes pontuais** têm forma elíptica e podem ter apenas algumas centenas de quilômetros de diâmetro. Em geral, um satélite de comunicações para os Estados Unidos tem um único feixe para os 48 estados contíguos, além de feixes pontuais para o Alasca e o Havai.

Um desenvolvimento importante no mundo dos satélites de comunicações foi a criação de microestações de baixo custo, às vezes chamadas **VSATs (Very Small Aperture Terminals)** (Abramson, 2000). Esses pequenos terminais têm antenas de 1 metro ou menos (em comparação com 10 m para uma antena de GEO padrão) e podem emitir cerca de 1 watt de energia. Geralmente, o uplink é adequado para 1 Mbps, mas o downlink normalmente exige vários megabits/s. A televisão transmitida por satélite utiliza essa tecnologia na transmissão de mão única.

Em muitos sistemas VSAT, as microestações não têm energia suficiente para se comunicarem diretamente umas com as outras (via satélite, é óbvio). Em vez disso, é necessária uma estação terrestre especial, o **hub**, com uma grande antena de alto ganho para retransmitir o tráfego entre VSATs, como mostra a Figura 2.50. O compromisso é um atraso mais longo em troca de estações mais econômicas para o usuário final.

Os VSATs apresentam um grande potencial em áreas rurais, especialmente nos países em desenvolvimento. Em grande parte do mundo, não existem linhas terrestres ou torres de celular. Estender fios telefônicos até milhares de pequenas aldeias é algo que vai muito além do orçamento da maioria dos governos de países em desenvolvimento. Erguer torres de celular é mais fácil, mas as torres precisam de conexões cabeadas para a rede telefônica nacional. Contudo, a instalação de antenas VSAT de 1 metro de diâmetro, alimentadas por células solares, geralmente é algo viável. Os VSATs fornecem a tecnologia que irá acabar de conectar o mundo e também podem oferecer acesso à Internet para usuários de smartphone em áreas onde não existe infraestrutura terrestre, o que acontece em grande parte do mundo em desenvolvimento.

Os satélites de comunicações têm diversas propriedades radicalmente diferentes dos enlaces terrestres ponto a ponto. Para começar, embora os sinais enviados e recebidos por um satélite trafeguem à velocidade da luz (aproximadamente 300.000 km/s), a longa distância de ida e volta introduz um atraso substancial para os satélites GEO. Dependendo da distância entre o usuário e a estação terrestre, e também da elevação do satélite acima do horizonte, a latência de ponta a ponta está entre 250 e 300 ms. Um valor



**Figura 2.50** VSATs utilizando um hub.

típico de ida e volta é 270 ms (540 ms, no caso de um sistema VSAT com um hub).

Para fins de comparação, os enlaces de micro-ondas terrestres têm um atraso de propagação de aproximadamente  $3 \mu\text{s}/\text{km}$ , e os enlaces de cabo coaxial ou fibra óptica geram um atraso de cerca de  $5 \mu\text{s}/\text{km}$ . Neste último caso, o atraso é maior porque os sinais eletromagnéticos trafegam com maior rapidez no ar do que em materiais sólidos.

Outra propriedade importante dos satélites é que eles basicamente são meios de difusão. Enviar uma mensagem para milhares de estações localizadas na área de cobertura de um transponder não custa mais do que enviar a mensagem para apenas uma estação. Para algumas aplicações, essa propriedade é muito útil. Por exemplo, poderíamos imaginar um satélite transmitindo páginas da Web comuns para os caches de um grande número de computadores espalhados por uma extensa área. Mesmo quando o broadcasting pode ser simulado com o uso de linhas ponto a ponto, o broadcasting por satélite pode ser muito mais econômico. Entretanto, do ponto de vista da privacidade, os satélites são um completo desastre: todo mundo pode ouvir tudo. A criptografia é essencial para a confidencialidade.

Nos satélites, o custo de transmissão de uma mensagem é independente da distância percorrida. O serviço de uma chamada transcontinental não custa mais do que uma chamada entre um lado e outro da rua. Os satélites também proporcionam excelentes taxas de erros e podem ser implementados quase instantaneamente, um detalhe fundamental para a comunicação militar.

### 2.8.2 Satélites terrestres de órbita média

Em altitudes muito mais baixas, entre os dois cinturões de Van Allen, encontramos os **satélites de órbita média**, ou **MEO (Medium-Earth Orbit)**. Vistos da Terra, esses satélites se deslocam lentamente em longitude, levando cerca de 6 horas para circular a Terra. Consequentemente, eles devem ser acompanhados à medida que se movem pelo céu. Pelo fato de estarem em órbitas mais baixas que os GEOs, têm uma área de cobertura menor no solo e exigem transmissores menos potentes para alcançá-los. Atualmente, esses satélites não são usados para telecomunicações, portanto, não os examinaremos mais aqui. A constelação de cerca de 30 satélites **GPS (Global Positioning System)** que estão em órbita a cerca de 20.200 km de altitude são exemplos de satélites MEO.

### 2.8.3 Satélites terrestres de órbita baixa

A uma altitude menor, encontramos os **satélites de órbita baixa**, ou **LEO (Low-Earth Orbit)**. Em razão de seu rápido movimento, são necessárias grandes quantidades desses satélites para formar um sistema completo. No entanto, pelo fato de os satélites estarem muito próximos da

Terra, as estações terrestres não precisam de muita potência, e o atraso de ida e volta é muito menor: as implementações veem latências de ida e volta em qualquer lugar entre cerca de 40 e 150 milissegundos. O custo de lançamento também é muito mais baixo. Nesta seção, examinaremos dois exemplos de constelações de satélites para o serviço de voz: Iridium e Globalstar.

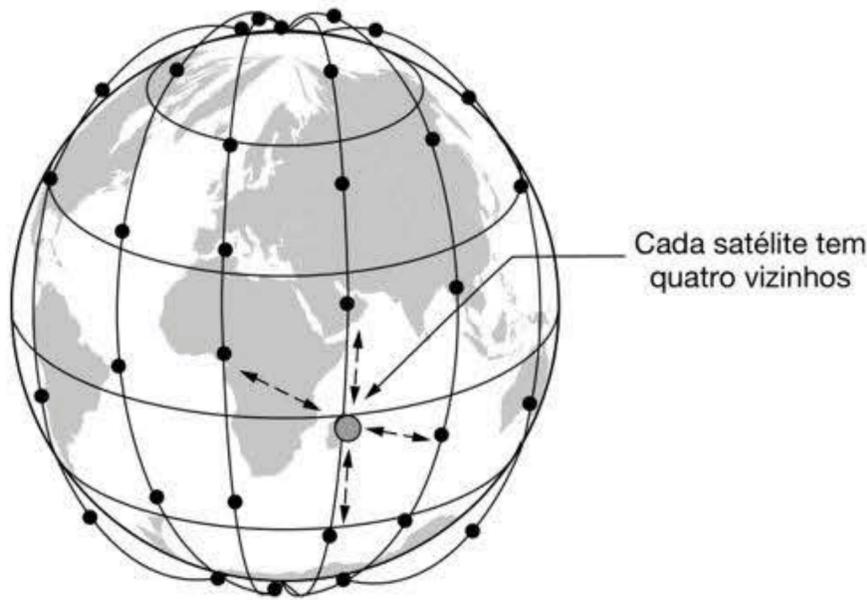
Durante os primeiros 30 anos da era do satélite, os de baixa órbita raramente eram usados, porque apareciam e desapareciam de vista com muita rapidez. Em 1990, a Motorola deu início a um novo empreendimento e enviou um requerimento à FCC solicitando permissão para lançar 77 satélites de baixa órbita para o projeto **Iridium** (o elemento 77 é o irídio). Mais tarde, o plano foi revisto para que fossem usados apenas 66 satélites; assim, o projeto deveria ter seu nome alterado para Dysprosium (o elemento 66), mas esse nome provavelmente lembrava muito mais uma doença do que um satélite. A ideia era que, assim que um satélite estivesse fora de vista, outro o substituiria. Essa proposta criou uma agitação entre outras empresas de comunicações. De repente, todas elas quiseram lançar uma cadeia de satélites de baixa órbita.

Após 7 anos reunindo parceiros e financiamentos, o serviço de comunicação teve início em novembro de 1998. Infelizmente, a demanda comercial por grandes e pesados telefones via satélite era desprezível, porque a rede de telefonia móvel (celular) havia crescido de modo espetacular desde 1990. Como consequência, o Iridium não gerou lucro e foi à falência em agosto de 1999, em um dos mais espetaculares fiascos corporativos da história. Os satélites e outros bens (no valor de 5 bilhões de dólares) foram adquiridos mais tarde por um investidor por 25 milhões de dólares, em uma espécie de venda de garagem extraterrestre. Outros empreendimentos comerciais logo se seguiram.

O serviço Iridium foi reiniciado em março de 2001, e tem crescido desde então. Há serviços de voz, dados, busca, fax e navegação em qualquer lugar do mundo, seja em terra, seja em mar e ar, com dispositivos portáteis que se comunicam diretamente com os satélites Iridium. Os clientes incluem as indústrias marítima, de aviação e de exploração de petróleo, bem como pessoas que viajam para regiões do mundo que não têm uma infraestrutura de telecomunicações (p. ex., desertos, montanhas, o Polo Sul e alguns países em desenvolvimento).

Os satélites Iridium estão posicionados a uma altitude de 670 km, em órbitas polares circulares. Eles estão organizados em eixos norte-sul, com um satélite a cada 32 graus de latitude, conforme mostra a Figura 2.51. Cada satélite tem no máximo 48 células (feixes pontuais), com uma capacidade de 3.840 canais, alguns deles usados para busca e navegação, enquanto outros são empregados para dados e voz.

Com seis eixos de satélite, a Terra inteira é coberta, como sugere a Figura 2.51. Uma propriedade interessante do Iridium é que a comunicação entre clientes distantes ocorre no espaço, como ilustra a Figura 2.52(a). Na figura,



**Figura 2.51** Os satélites Iridium formam seis eixos em torno da Terra.

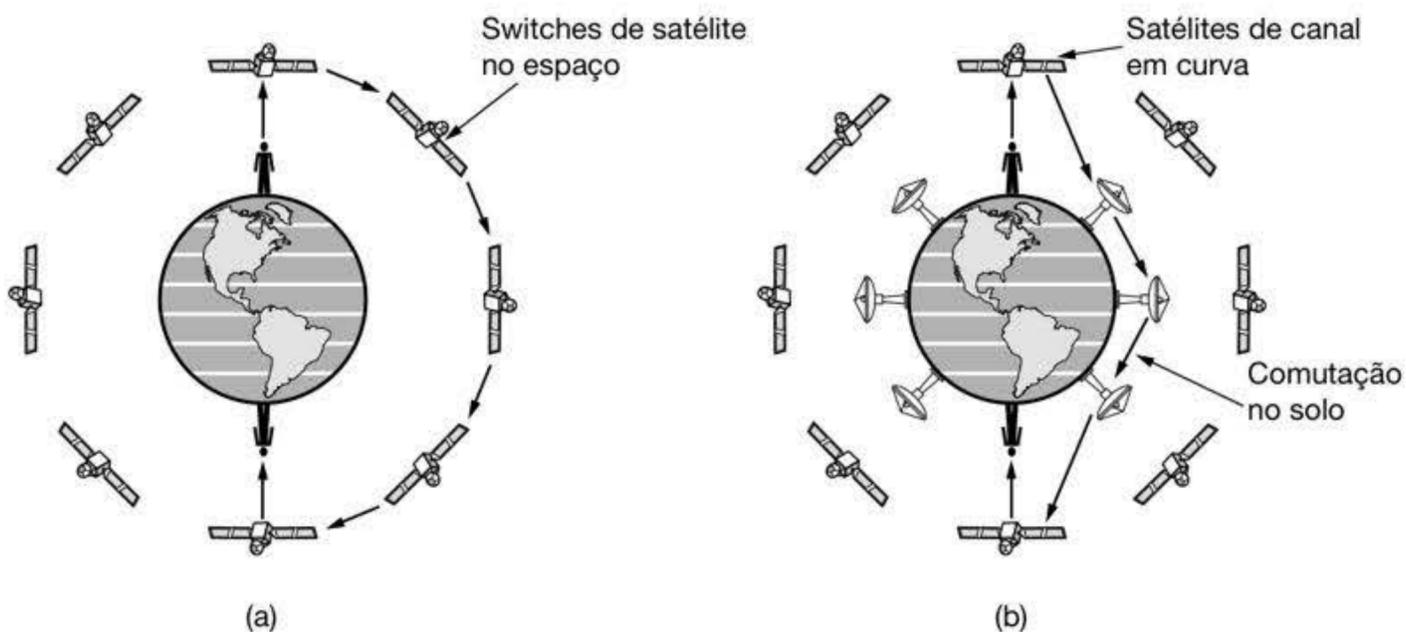
vemos um chamador no Polo Norte entrando em contato com um satélite situado diretamente acima dele. Cada satélite tem quatro vizinhos com os quais pode se comunicar, dois no mesmo eixo (mostrado) e dois em eixos adjacentes (não mostrados). Os satélites repassam a chamada por essa grade até ela finalmente ser enviada para o destinatário no Polo Sul.

Um projeto alternativo para o Iridium é o **Globalstar**. Ele se baseia em 48 satélites LEO, mas utiliza um esquema de comutação diferente do que é usado no Iridium. Enquanto este retransmite as chamadas de satélite para satélite, o que exige sofisticado equipamento de comutação nos satélites, o Globalstar utiliza um projeto tradicional de canal em curva. A chamada originada no Polo Norte na Figura 2.52(b) é enviada de volta à Terra e recebida pela grande estação terrestre na fábrica de brinquedos do Papai Noel. A chamada é, então, roteada por uma rede terrestre até a estação terrestre mais próxima ao destino, e é entregue por uma conexão de canal em curva da maneira ilustrada.

A vantagem desse esquema é que ele coloca a maior parte da complexidade no solo, onde é mais fácil de administrar. Além disso, o uso de grandes antenas nas estações terrestres, capazes de emitir um sinal potente e receber um sinal fraco, significa que podem ser utilizados telefones de potência mais baixa. Afinal, o telefone emite apenas alguns miliwatts de potência e, assim, o sinal que volta para a estação terrestre é bastante fraco, mesmo depois de ter sido amplificado pelo satélite.

Os satélites continuam a ser lançados a uma taxa de algo em torno de 20 por ano, incluindo satélites cada vez maiores, que agora pesam mais de 5.000 kg. Mas há também satélites muito pequenos para organizações mais preocupadas com o orçamento. Para tornar a pesquisa do espaço mais acessível, em 1999, os acadêmicos da Califórnia Polytechnic University e Stanford se reuniram para definir um padrão para satélites em miniatura e um disparador associado que reduziria bastante os custos de lançamento (Nugent et al., 2008). **Cubesats** são satélites em unidades de cubos de 10 cm × 10 cm × 10 cm, cada um pesando menos de 1 kg, que podem ser lançados a partir de US\$ 40.000 cada. O disparador voa como um segundo payload nas missões espaciais comerciais. Ele é basicamente um tubo que ocupa três unidades de cubesats e usa molas para lançá-los em órbita. Cerca de 20 cubesats foram lançados até agora, com muito mais em andamento. A maioria deles se comunica com estações terrestres nas faixas de UHF e VHF.

Outra implantação de satélites LEO é uma tentativa de rede backbone de Internet baseada em satélite. A implantação do OneWeb envolverá inicialmente uma constelação de várias centenas de satélites. Se for bem-sucedido, o projeto promete levar acesso de alta velocidade à Internet para lugares que ainda não o têm. Os satélites irão operar na banda Ku e usarão uma técnica chamada “pitch progressivo”, em que os satélites são ligeiramente girados para evitar interferência com satélites geoestacionários que estão transmitindo na mesma banda.



**Figura 2.52** (a) Retransmissão no espaço. (b) Retransmissão no solo.

## 2.9 COMPARAÇÃO DE DIFERENTES REDES DE ACESSO

Agora, vamos comparar as propriedades dos diferentes tipos de redes de acesso que pesquisamos.

### 2.9.1 Redes de acesso terrestres: cabo, fibra e ADSL

Cabo, FTTH e ADSL têm muito mais semelhanças do que diferenças. Eles oferecem um serviço comparável e, à medida que a concorrência entre eles aumenta, provavelmente terão preços comparáveis. Todas as tecnologias de acesso à rede, incluindo cabo, ADSL e FTTH, utilizam agora a fibra no backbone; elas diferem na tecnologia de acesso do último quilômetro, nas camadas física e de enlace. Provedores de fibra e ADSL costumam oferecer largura de banda mais consistente a cada assinante, pois cada usuário possui capacidade dedicada. Relatórios contínuos e recentes nos Estados Unidos, como a iniciativa Measuring Broadband America (MBA), da FCC, realizado anualmente, informam que os provedores de acesso à Internet geralmente cumprem suas velocidades anunciadas.

À medida que um sistema ADSL ou FTTH conquista mais usuários, seus números crescentes têm pouco efeito sobre os usuários existentes, pois cada usuário tem uma conexão dedicada até sua casa. Por sua vez, os assinantes de sistemas a cabo compartilham a capacidade de um único nó e, como resultado, quando um ou mais usuários em um nó aumentam seu uso, outros usuários podem sofrer com o congestionamento. Consequentemente, os provedores de serviço a cabo atualmente oferecem mais capacidade do que eles vendem a cada assinante. Os padrões DOCSIS mais modernos, como DOCSIS 3.0, exigem que os modems a cabo sejam capazes de ligar pelo menos quatro canais, para alcançar aproximadamente 170 Mbps downstream e 120 Mbps upstream (com cerca de 10% desse throughput dedicado ao overhead de sinalização).

Em última análise, as velocidades máximas que um assinante de serviço a cabo pode atingir se limitam à capacidade do cabo coaxial, mas, em comparação, a quantidade do espectro utilizável na fibra é muito maior. Com o cabo, à medida que mais assinantes se registram no serviço de Internet, o desempenho de outros usuários no mesmo nó diminui. Em resposta, os provedores de cabo dividem os cabos mais ocupados, conectando cada um diretamente a um nó de fibra (essa prática às vezes é chamada de **divisão de nó**). Como já dissemos, o número de residências por nó continua a diminuir, conforme os ISPs de cabo continuam a levar a fibra para mais perto da borda da rede.

Cabo, fibra e ADSL estão disponíveis em diferentes regiões e o desempenho dessas redes difere de acordo com a própria tecnologia e na forma como cada uma é implantada. A maioria dos usuários domésticos nos países

desenvolvidos pode ter uma linha telefônica se quiserem, mas nem todos estão próximos o suficiente de suas estações finais para obter ADSL. Alguns estão presos a linhas discadas de 56 kbps, especialmente nas áreas rurais. Na verdade, até mesmo nos Estados Unidos, existem grandes áreas nas quais uma linha T1 de 1,544 Mbps é um luxo raramente alcançado. Nas grandes cidades da Europa, com sua maior densidade populacional, a Internet por fibra óptica de 500 Mbps é muito comum. Alguns locais até possuem serviço de 1 Gbps disponível.

Além disso, nem todos têm o serviço a cabo. Se você o tiver e a empresa fornecer acesso à Internet, você poderá obtê-lo; a distância até o nó de fibra ou headend não é um problema. A disponibilidade de cabos e fibras em certas regiões, particularmente nas pouco povoadas, continua sendo uma preocupação. Em última análise, o acesso de alta velocidade à Internet ainda depende da implantação de fibra ou cabo nas residências. No caso de redes de cabo, aumentar as divisões de nós exige a implantação de mais fibra na vizinhança, em vez de depender da infraestrutura de cabo coaxial existente. Mesmo no caso de ADSL, a velocidade cai significativamente além de alguns quilômetros de uma estação central, de modo que mesmo ADSL requer algum tipo de aumento de fibra na borda (p. ex., FTTN) para oferecer alta velocidade em áreas pouco povoadas. Todas essas propostas são caras.

Historicamente, a infraestrutura de telefone (e redes DSL) tem sido geralmente mais confiável do que o cabo, embora os dados do projeto MBA da FCC mostrem que a lacuna foi reduzida, com a maioria dos serviços de cabo e DSL alcançando pelo menos “dois noves” de confiabilidade (ou seja, 99% de tempo de atividade, ou dezenas de horas de tempo de inatividade por ano). As redes sem fio de satélite e de área metropolitana têm um desempenho menos confiável. Em comparação, a rede telefônica convencional atinge “cinco noves” de confiabilidade, o que corresponde a apenas alguns minutos de indisponibilidade a cada ano (Bischof et al., 2018).

Por ser um meio ponto a ponto, ADSL é inerentemente mais segura que o cabo. Qualquer usuário de serviço a cabo pode facilmente ler todos os pacotes que chegam pelo cabo, não importa para quem eles se destinam. Por esse motivo, qualquer provedor de serviço a cabo decente codificará todo o tráfego nos dois sentidos. Apesar disso, ter um vizinho recebendo todas as suas mensagens criptografadas ainda é menos seguro do que se ele não recebesse nada.

### 2.9.2 Redes terrestres e por satélite

Uma comparação entre as comunicações por satélite e terrestre é instrutiva. Há algum tempo, pensava-se que o futuro da comunicação residia nos satélites de comunicações. Afinal, o sistema telefônico mudou muito pouco nos últimos 100 anos e não mostrou sinais de mudança para os próximos 100. Esse movimento glacial foi causado, em grande

parte, pelo ambiente regulador no qual se esperava que as companhias telefônicas fornecessem bons serviços de voz a preços razoáveis (o que elas fizeram) e, em troca, tivessem lucro garantido sobre seu investimento. Havia modems de 1.200 bps disponíveis para as pessoas que precisavam transmitir dados. Isso era praticamente tudo o que existia na época.

Com o surgimento da concorrência, em 1984 nos Estados Unidos e um pouco mais tarde na Europa, esse quadro se alterou radicalmente. As companhias telefônicas começaram a substituir suas redes de longa distância por fibra óptica e introduziram serviços de alta largura de banda, como ADSL. Essas empresas também interromperam sua antiga prática de cobrar preços artificialmente elevados a usuários de serviços de longa distância, a fim de subsidiar o serviço local. Subitamente, as conexões terrestres de fibra pareciam ser a opção vencedora.

Apesar disso, os satélites de comunicações têm alguns segmentos de mercado muito importantes, que a fibra óptica não é capaz de alcançar. Primeiro, quando a implantação rápida é crítica, os satélites ganham facilmente. Uma resposta rápida é útil para sistemas de comunicação militares em tempos de guerra e, em tempos de paz, para resposta a desastres. Em Sumatra, em dezembro de 2004, após o grande terremoto e o tsunami subsequente, por exemplo, os satélites de comunicações foram capazes de restaurar as comunicações com os primeiros respondedores dentro de 24 horas. Essa resposta rápida foi possível porque existe um mercado de serviço de satélite desenvolvido em que grandes participantes, como Intelsat, com mais de 50 satélites, podem arrendar capacidade onde quer que ela seja necessária. Para clientes atendidos por redes de satélite existentes, um VSAT alimentado por energia solar pode ser preparado fácil e rapidamente para fornecer um enlace de megabits/s.

Um segundo nicho de mercado ocorre em lugares onde a infraestrutura terrestre é pouco desenvolvida. Muitas pessoas querem se comunicar enquanto se movimentam. As redes de telefone móvel abrangem locais com boa densidade populacional, mas não realizam um trabalho adequado em outros lugares (p. ex., no mar ou no deserto). Ao contrário, o Iridium oferece serviço de voz em qualquer lugar do planeta, até mesmo no Polo Sul. A infraestrutura em terra pode ser cara para instalar, dependendo do terreno e dos direitos necessários para viabilizar o meio. A Indonésia, por exemplo, tem seu próprio satélite para o tráfego de telefone doméstico. Lançar um satélite foi mais barato do que esticar milhares de cabos submarinos entre as 13.677 ilhas do arquipélago.

Um terceiro nicho se relaciona a situações em que a difusão é essencial. Uma mensagem enviada por satélite pode ser recebida por milhares de estações terrestres ao mesmo tempo. Os satélites são usados para distribuir grande parte da programação da TV para estações locais por esse motivo. Agora, existe um grande mercado para

transmissões de rádio e TV por satélite, diretamente para usuários finais com receptores de satélite em suas casas e carros. Vários outros tipos de conteúdo também podem ser transmitidos. Por exemplo, uma empresa que transmite um fluxo de preços de ações, apólices ou mercadorias a milhares de corretores deve considerar que um sistema de satélite é mais econômico que simular a difusão no solo.

Os Estados Unidos têm alguns provedores concorrentes de Internet baseada em satélite, incluindo Hughes (muitas vezes comercializado como DISH, anteriormente EchoStar) e Viasat, que operam principalmente com satélites geoestacionários ou MEO, com alguns provedores passando para LEO. Em 2016, o projeto MBA da FCC relatou que esses provedores baseados em satélite estavam entre os poucos provedores de serviços de Internet que estavam vendo uma queda no desempenho ao longo do tempo, provavelmente por causa do aumento de assinantes e largura de banda limitada. O relatório descobriu que esses provedores não eram capazes de oferecer velocidades superiores em torno de 10 Mbps.

Apesar disso, nos últimos anos, o acesso à Internet por satélite tem tido um interesse crescente, especialmente em nichos de mercado, como durante um voo. Alguns acessos à Internet durante o voo envolvem comunicação direta com torres de banda larga móvel, mas para voos sobre os oceanos, isso não funciona. Outro método que ajuda a lidar com a largura de banda limitada em aviões envolve a transmissão de dados para uma coleção de satélites em órbita geoestacionária. Algumas empresas, incluindo a OneWeb, conforme discutido anteriormente, e a Boeing estão trabalhando na construção de um backbone de Internet baseado em satélite usando satélites LEO. Os mercados ainda serão um pouco restritos, já que o throughput será de aproximadamente 50 Mbps, muito inferior ao da Internet no solo.

De maneira resumida, parece que a comunicação do futuro será feita por fibras ópticas terrestres combinadas com rádio celular, mas, para algumas aplicações específicas, os satélites são melhores. Entretanto, existe um motivo que se aplica a tudo isso: a economia. Embora a fibra ofereça mais largura de banda, é muito provável que a comunicação terrestre e por satélite entre em uma concorrência agressiva por melhores preços. Se os avanços tecnológicos reduzirem radicalmente o custo de lançamento de um satélite (p. ex., se no futuro algum veículo espacial puder lançar dezenas de satélites de uma só vez), ou se os satélites de baixa órbita se desenvolverem, não é certo que a fibra vencerá em todos os mercados.

## 2.10 POLÍTICA NA CAMADA FÍSICA

Diversos aspectos da camada física envolvem decisões políticas e regulamentadoras que, por fim, afetam o modo como essas tecnologias são usadas e desenvolvidas. Vamos

discutir rapidamente sobre a atividade política em andamento nas redes terrestres (i.e., as redes de telefone e cabo) e nas redes sem fio.

### 2.10.1 Alocação do espectro

O maior desafio com relação ao espectro eletromagnético refere-se à realização da **alocação do espectro** de modo eficiente e justo. Se várias partes puderem transmitir dados na mesma faixa do espectro na mesma região geográfica, haverá um potencial significativo para que as partes que se comunicam interfiram umas com as outras. Para evitar o caos total, têm sido feitos acordos nacionais e internacionais a respeito de quem terá o direito de usar cada uma das frequências. Como todos querem uma taxa de dados mais alta, todos desejam um espectro maior. Os governos nacionais alocam bandas do espectro para rádios AM e FM, televisão e telefones celulares, assim como para as empresas de telefonia, a polícia, os usuários marítimos, de navegação, militares, do governo e para muitos outros usuários concorrentes. Em termos mundiais, uma agência da ITU-R (WARC) tenta coordenar essa alocação de forma que possam ser fabricados dispositivos que funcionem em vários países. Porém, os países não são limitados pelas recomendações da ITU-R, e a FCC, que realiza a alocação para os Estados Unidos, ocasionalmente as têm rejeitado (em geral porque elas exigiam que algum grupo politicamente poderoso desistisse de alguma fração do espectro).

Até mesmo quando uma parte do espectro é alocada para algum uso, como telefones celulares, existe a questão adicional de decidir qual concessionária terá permissão para usar quais frequências. Três algoritmos foram muito utilizados no passado. O algoritmo mais antigo, frequentemente chamado de **concurso de beleza**, exige que cada concessionária explique por que sua proposta serve melhor ao interesse público. Então, os funcionários do governo decidem qual entre as belas histórias mais lhes agrada. Fazer algum funcionário do governo oferecer como prêmio a propriedade de bilhões de dólares à sua empresa favorita em geral leva a suborno, corrupção, nepotismo e crimes piores. Além disso, até mesmo um funcionário do governo honesto e escrupuloso que imagine que uma companhia estrangeira poderia realizar um trabalho melhor que qualquer uma das empresas nacionais teria muito a explicar.

Essa observação levou ao segundo algoritmo, que realiza um **sorteio** entre as empresas interessadas. O problema com essa ideia é que empresas que não têm qualquer interesse em usar o espectro podem participar desse sorteio. Se, digamos, um restaurante ou uma cadeia de sapatarias ganhasse, a empresa poderia revender o espectro a uma concessionária com um enorme lucro e sem nenhum risco.

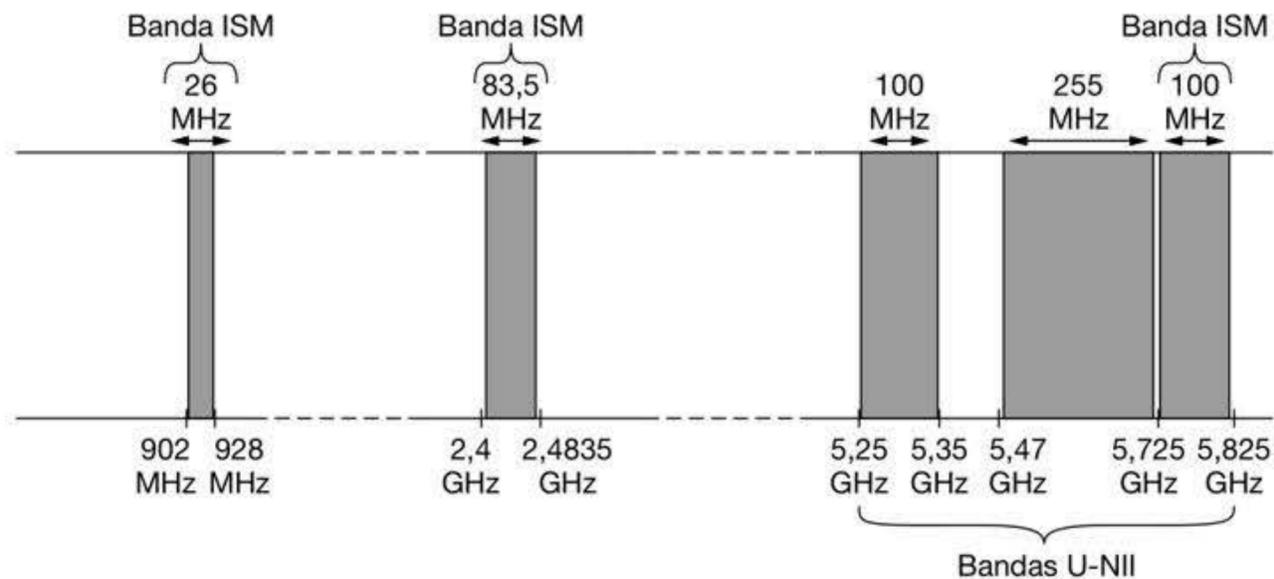
A ideia de conceder fatias do espectro a empresas com uma enorme dose de sorte mas sem nenhum método tem sido severamente criticada por muitos, o que levou ao

terceiro algoritmo: realizar **leilões** e conceder a largura de banda à empresa que fizer a melhor proposta. Quando o governo britânico leiloou as frequências necessárias para os sistemas de telefonia móvel em 2000, ele esperava obter aproximadamente 4 bilhões de dólares. Na realidade, recebeu cerca de 40 bilhões de dólares, pois as concessionárias entraram em uma disputa frenética, mortas de medo de perder o barco da telefonia móvel. Esse evento despertou a ganância dos governos vizinhos e os inspirou a realizar seus próprios leilões. Isso funcionou, mas também deixou algumas concessionárias tão endividadas que elas chegaram perto da falência. Até mesmo nos melhores casos, muitos anos serão necessários para essas empresas recuperarem o custo do licenciamento.

Uma abordagem muito diferente para alocar frequências é simplesmente não alocá-las. Em vez disso, basta deixar todo mundo transmitir à vontade, mas regular a potência utilizada, de forma que as estações tenham um alcance tão pequeno que não possam interferir umas com as outras. De acordo com essa proposta, a maioria dos governos reserva algumas bandas de frequência, chamadas **bandas ISM (Industrial, Scientific, Medical)** para uso sem licença. Sistemas para abertura de portas de garagens, telefones sem fio, brinquedos controlados por rádio, dispositivos tipo mouse sem fio e vários outros aparelhos domésticos sem fio utilizam as bandas ISM. Para diminuir a interferência entre esses dispositivos não coordenados, a FCC estabelece que todos os dispositivos nas bandas ISM devem limitar sua potência de transmissão (p. ex., para 1 watt) e usar técnicas para dispersar seus sinais por uma faixa de frequências. Os dispositivos também podem ter de evitar interferência com instalações de radar.

A localização das bandas ISM varia um pouco de país para país. Por exemplo, nos Estados Unidos, as bandas que os dispositivos em rede utilizam na prática, sem exigir uma licença da FCC, são mostradas na Figura 2.53. A banda de 900 MHz era usada para as primeiras versões do 802.11, mas está sobrecarregada. A banda de 2,4 GHz está disponível na maioria dos países e é bastante usada para 802.11b/g e Bluetooth, mas é sujeita a interferências de fornos de micro-ondas e instalações de radar. A banda de 5 GHz do espectro inclui bandas **U-NII (Unlicensed National Information Infrastructure)**. As bandas de 5 GHz são relativamente pouco desenvolvidas, mas, por terem a maior largura de banda e serem usadas por especificações WiFi, como a 802.11ac, rapidamente estão se tornando mais populares e sobrecarregadas também.

As bandas não licenciadas foram um grande sucesso na década passada. A capacidade de usar o espectro livremente ocasionou uma grande inovação nas LANs e PANs sem fio, evidenciada pela implantação generalizada de tecnologias como 802.11 e Bluetooth. Até mesmo alguns provedores estão entrando no jogo com tecnologias como LTE-U, que envolve a implantação de uma rede de celular LTE no espectro não licenciado. Essa tecnologia poderia



**Figura 2.53** As bandas ISM e U-NII usadas nos Estados Unidos por dispositivos sem fio.

permitir que dispositivos móveis operassem nesse espectro não licenciado, além das partes do espectro que são explicitamente alocadas para operar com redes de celular. LTE-U pode permitir que os ISPs de linha fixa que estão implantando pontos de acesso WiFi em centenas de milhões de lares transformem sua rede de pontos de acesso em uma rede de estações-base de celular. É claro que permitir que telefones celulares usem o espectro não licenciado traz o seu próprio conjunto de complicações. Por exemplo, dispositivos que operam no espectro não licenciado devem respeitar outros dispositivos que estão usando o mesmo espectro e tentar não interferir com os dispositivos considerados “titulares”. LTE-U também pode enfrentar seus próprios desafios de confiabilidade e desempenho, pois deve recuar para interagir bem com outros dispositivos que usam o espectro não licenciado, de outros dispositivos WiFi a monitores de bebê.

Vários desenvolvimentos na política durante os últimos 10 anos continuam a possibilitar mais inovação nas tecnologias sem fio. Um desenvolvimento interessante nos Estados Unidos é a potencial alocação futura de mais espectro não licenciado. Em 2009, a FCC decidiu permitir o uso não licenciado de **espaços vazios** em torno de 700 MHz. Os espaços vazios são bandas de frequência que foram alocadas mas não estão sendo usadas localmente. A mudança das transmissões de televisão de analógicas para digitais nos Estados Unidos, em 2010, liberou os espaços vazios em torno de 700 MHz. Uma dificuldade é que, para utilizar esses espaços, dispositivos não licenciados precisam poder detectar quaisquer transmissores licenciados vizinhos, incluindo microfones sem fio, que têm os direitos iniciais de usar a banda de frequência. A FCC também abriu de 57 GHz a 64 GHz para operação não licenciada em 2001. Essa faixa é uma parte enorme do espectro, mais do que todas as outras bandas ISM combinadas, de modo que pode aceitar o tipo de rede de alta velocidade que seria necessária para enviar streaming de TV de alta definição através do ar a sua sala de estar. A 60 GHz, as ondas de rádio são absorvidas pelo oxigênio. Isso significa que os sinais não

se propagam longe, tornando-os bem adequados a redes de curta distância. As altas frequências (60 GHz é uma banda de frequência extremamente alta, ou de “milímetro”, logo abaixo da radiação de infravermelho) representaram um desafio inicial para os fabricantes de equipamentos, mas os produtos atualmente estão no mercado.

Nos Estados Unidos, outras bandas do espectro também estão sendo reaproveitadas e leiloadas para operadoras, incluindo 2,5 e 2,9 GHz, a Banda C (anteriormente usada para comunicações por satélite) na faixa de 3,7–4,2 GHz, bem como outras, incluindo 3,5, 6, 24, 28, 37 e 49 GHz. Para a comunicação de curto alcance, a FCC também está considerando o uso de certas bandas muito altas, como a faixa de 95 GHz. No final de 2018, a FCC lançou seu primeiro leilão 5G, com outros planejados para os próximos anos. Esses leilões abrirão uma quantidade significativa de espectro para banda larga móvel, possibilitando larguras de banda maiores que seriam necessárias para streaming de vídeo e aplicativos da Internet das Coisas. O espectro de 24 e 28 GHz tem, cada um, aproximadamente 3.000 licenças à venda. A FCC também está oferecendo descontos para pequenas empresas e fornecedores rurais. Leilões de partes das bandas do espectro de 37, 39 e 49 GHz também estão programados. Em outros países, algumas dessas bandas do espectro podem operar como espectro não licenciado. Por exemplo, a indústria automotiva na Alemanha fez lobby com sucesso para permitir a banda de 3,5 GHz para uso de empresas privadas; outros países europeus provavelmente seguirão o exemplo.

## 2.10.2 A rede celular

É interessante observar como pequenas decisões políticas e de marketing podem ter um enorme impacto sobre a implantação de redes celulares nos Estados Unidos e na Europa. O primeiro sistema móvel foi criado nos Estados Unidos pela AT&T e regulamentado para todo o país pela FCC. Como resultado, todo o território do país tinha um único

sistema (analógico), e um telefone móvel adquirido na Califórnia também funcionava em Nova York. Ao contrário, quando a tecnologia de telefonia móvel chegou à Europa, cada país criou seu próprio sistema, o que resultou em um fiasco.

A Europa aprendeu com seus erros e, ao surgir a tecnologia digital, as PTTs estatais se juntaram e padronizaram um único sistema (GSM); portanto, qualquer telefone móvel europeu funcionará em qualquer lugar da Europa. Na época, os Estados Unidos haviam decidido que o governo não deveria participar do esforço de padronização, e assim a padronização da tecnologia digital ficou a cargo do mercado. Essa decisão resultou em diferentes fabricantes de equipamentos produzindo tipos distintos de telefones móveis. Em consequência, os Estados Unidos agora têm dois importantes sistemas de telefonia móvel digital totalmente incompatíveis em operação, além de outros sistemas secundários.

Apesar da liderança inicial dos Estados Unidos, a propriedade e a utilização da telefonia móvel na Europa é agora muito maior que naquele país. O fato de haver um único sistema para toda a Europa explica em parte esse fato, mas há outras razões. Um segundo ponto em que os Estados Unidos e a Europa divergiram foi a questão dos números de telefone. Nos Estados Unidos, os telefones móveis têm números misturados com os telefones comuns (fixos). Desse modo, quem faz a ligação não tem como saber se, digamos, (212) 234-5678 é um telefone fixo (com uma ligação de baixo custo ou gratuita) ou um telefone móvel (com uma tarifa cara). Para impedir que as pessoas ficassem receosas de usar o telefone, as empresas de telefonia decidiram fazer o proprietário do telefone móvel pagar pelas chamadas recebidas. Em consequência disso, muitas pessoas hesitaram em comprar um telefone móvel por medo de terem de pagar uma conta enorme apenas por receberem ligações. Na Europa, os telefones móveis têm um código de área especial (semelhante aos números 800 e 900) e, assim, podem ser reconhecidos instantaneamente. Como resultado, a regra habitual de “fazer o chamador pagar” também se aplica aos telefones móveis da Europa (com exceção das ligações internacionais, cujos custos são divididos).

Uma terceira questão que teve grande impacto na adoção da telefonia móvel foi o uso difundido de telefones pré-pagos na Europa (até 75% em algumas regiões). Esses telefones podem ser adquiridos em muitas lojas, até mesmo on-line. Eles são pré-carregados com, por exemplo, com um saldo de 20 ou 50 euros em ligações e podem ser recarregados (com a utilização de um código PIN secreto) quando o saldo termina. Por essa razão, praticamente todos os adolescentes e muitas crianças pequenas na Europa têm telefones móveis (em geral, pré-pagos) para que seus pais possam localizá-los, sem o perigo de terem de pagar uma conta enorme. Se o telefone móvel for usado apenas ocasionalmente, seu uso será quase gratuito, pois não haverá tarifa mensal nem por chamadas recebidas.

O leilão de cobiçadas bandas do espectro para 5G, juntamente com muitos dos avanços tecnológicos discutidos anteriormente neste capítulo, está prestes a sacudir os limites da rede celular pelos próximos anos. Já estamos vendo o aumento de operadoras de rede virtual móvel, ou **MVNOs (Mobile Virtual Network Operators)**, que são operadoras sem fio que não possuem a infraestrutura de rede sobre a qual prestam serviço a seus clientes. À medida que o tamanho das células continua a diminuir com frequências mais altas e o hardware para células pequenas continua a ser comoditizado, as MVNOs pagam para compartilhar a capacidade em uma infraestrutura que é operada por outra operadora. Elas decidem se irão operar seus próprios componentes de uma arquitetura LTE ou usar a infraestrutura de propriedade da operadora subjacente. MVNOs que operam sua própria rede principal às vezes são chamadas de MVNOs “completas”. Empresas como Qualcomm e Intel estão montando um projeto de referência para hardware de pequenas células, que pode resultar na desagregação completa da borda da rede, especialmente quando combinada com o uso do espectro não licenciado. A indústria também está começando a se voltar para a infraestrutura com eNodeBs de “caixa branca”, que se conectam a um escritório central que possui serviços EPC virtuais; o projeto M-CORD da Open Networking Foundation implementou uma arquitetura assim.

### 2.10.3 A rede telefônica

Por muitas décadas até 1984, a Bell System foi a responsável pelo serviço de chamadas locais e interurbanas em quase todos os Estados Unidos. Na década de 1970, o governo norte-americano concluiu que esse era um monopólio ilegal e promoveu uma ação para desmembrá-lo. O governo foi vitorioso e, em 1º de janeiro de 1984, a AT&T foi dividida na AT&T Long Lines, em 23 **BOCs (Bell Operating Companies)** e em algumas outras partes. As 23 BOCs foram agrupadas em sete BOCs regionais (RBOCs), o que as tornou economicamente viáveis. Toda a natureza do sistema de telecomunicações norte-americano foi alterada da noite para o dia por uma ordem judicial (e não por um ato do Congresso norte-americano).

As especificações exatas dessa ruptura foram descritas no conhecido **juízo final modificado**, ou **MFJ (Modified Final Judgement)**, um oxímoro, se é que houve um. Esse fato provocou o aumento da concorrência, a melhoria dos serviços e a redução das tarifas de longa distância para consumidores e empresas. Entretanto, os preços para o serviço local cresceram à medida que os subsídios cruzados das chamadas de longa distância foram eliminados, e o serviço local teve de se tornar autossuficiente. Hoje, muitos outros países estão considerando a abertura à concorrência em termos semelhantes.

De relevância direta para nossos estudos é que a nova estrutura competitiva introduziu um recurso técnico

fundamental para a arquitetura da rede telefônica. Para determinar precisamente a quem cabiam as responsabilidades, o território dos Estados Unidos foi dividido em 164 **áreas de transporte e acesso local**, ou **LATAs (Local Access and Transport Areas)**. De forma bem genérica, uma LATA corresponde à região coberta por um único código de área. Em geral, dentro de uma LATA existia um **operador local de troca**, ou **LEC (Local Exchange Carrier)** que detinha o monopólio do sistema telefônico convencional dentro de sua área. Os LECs mais importantes eram as BOCs, embora algumas LATAs contivessem uma ou mais das 1.500 companhias telefônicas independentes que operavam como LECs.

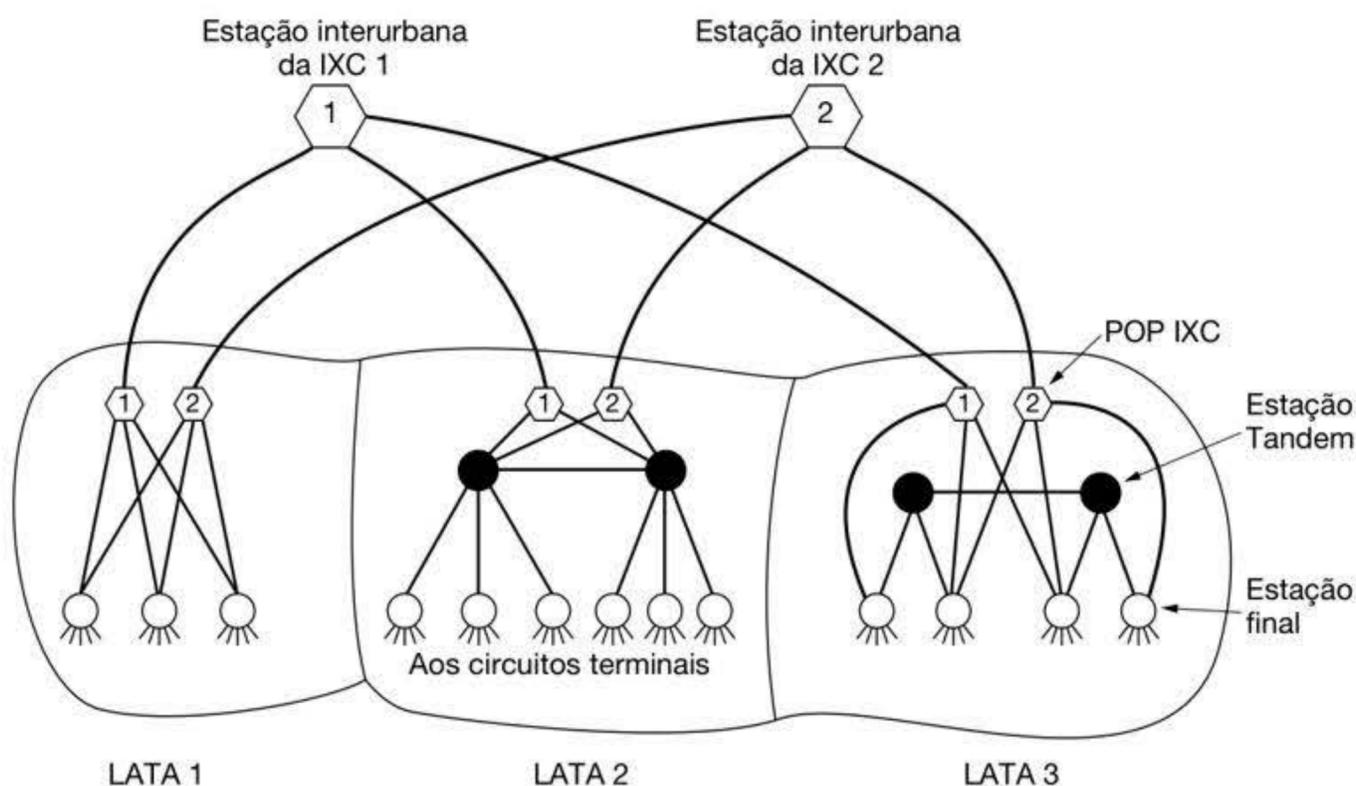
A nova característica era que todo o tráfego entre LATAs passou a ser manipulado por um tipo diferente de empresa, uma **operadora de intercâmbio**, ou **IXC (InterExchange Carrier)**. Originalmente, a AT&T Long Lines era a única IXC segura, mas hoje existem concorrentes fortes, como Verizon e Sprint, no ramo das IXCs. Uma das preocupações ao ocorrer o desmembramento foi assegurar-se de que todas as IXCs seriam tratadas igualmente em termos de qualidade das linhas, das tarifas e do número de dígitos que seus clientes teriam de teclar para usá-las. Observe, na Figura 2.54, como essa situação é tratada. Nela, vemos três exemplos de LATAs, cada uma com várias estações finais. As LATAs 2 e 3 também têm uma pequena hierarquia com estações Tandem (estações interurbanas intraLATA).

Qualquer IXC que deseje se encarregar de chamadas provenientes de uma LATA pode criar uma estação de comutação denominada **ponto de presença**, ou **POP (Point of Presence)**. O LEC é necessário para conectar cada IXC a cada estação final, seja diretamente, como nas LATAs 1 e 3,

seja indiretamente, como na LATA 2. Além disso, as condições da conexão, tanto técnicas quanto financeiras, têm de ser idênticas para todas as IXCs. Dessa forma, um assinante da LATA 1, por exemplo, pode escolher qual IXC usará para entrar em contato com assinantes que façam parte da LATA 3.

Como parte do MFJ, as IXCs foram proibidas de prestar serviços telefônicos locais, e os LECs foram proibidos de prestar serviços telefônicos entre LATAs, apesar de todas serem livres para atuar em quaisquer outros ramos, como a operação de restaurantes. Em 1984, essa era uma condição razoavelmente não ambígua. Infelizmente, a tecnologia tem uma forma interessante de tornar a lei obsoleta. Nem a TV a cabo nem os telefones celulares foram cobertos pelo acordo. À medida que a TV a cabo passou de unidirecional para bidirecional e a popularidade dos telefones celulares explodiu, os LECs e as IXCs começaram a comprar ou a se associar às operadoras de TV a cabo ou de telefones celulares.

Em 1995, o Congresso dos Estados Unidos percebeu que tentar manter uma distinção entre os vários tipos de empresas não era mais sustentável e elaborou um projeto de lei para preservar a acessibilidade para competição, mas que permitiria às empresas de TV a cabo, companhias telefônicas locais, concessionárias de comunicação de longa distância e operadoras de sistemas celulares entrarem nos ramos de negócio umas das outras. A ideia era que qualquer empresa poderia oferecer a seus clientes um único pacote integrado contendo serviços de TV a cabo, telefone e dados, e que diferentes empresas seriam concorrentes em serviços e preços. O projeto de lei foi sancionado em fevereiro de 1996, como uma reestruturação importante da regulamentação das telecomunicações. Como resultado, algumas



**Figura 2.54** Relacionamento entre LATAs, LECs e IXCs. Todos os círculos são estações de comutação de LECs. Cada hexágono pertence à IXC indicada pelo número.

BOCs se tornaram IXC's e algumas outras empresas, como operadoras de TV a cabo, começaram a oferecer serviços de telefonia local, competindo com os LECs.

Uma propriedade interessante da lei de 1996 foi a exigência de que os LECs implementassem **portabilidade de número local**. Isso significava que um cliente podia mudar de companhia telefônica local sem ter de receber um novo número de telefone. A portabilidade para números de telefone móvel (e entre linhas fixas e móveis) seguiu o exemplo em 2003. Essas providências removeram um enorme obstáculo para muitas pessoas, tornando-as muito mais inclinadas a mudar de LECs. Como resultado, o cenário das telecomunicações dos Estados Unidos se tornou muito mais competitivo, e outros países estão seguindo o mesmo caminho. Com frequência, outros países esperam para ver como esse tipo de experiência funciona nos Estados Unidos. Se isso funciona bem, eles fazem o mesmo; se funciona mal, eles tentam algo diferente.

Recentemente, a política de telecomunicações tem estado relativamente quieta, no que diz respeito às companhias telefônicas, com a maioria das ações e atividades passando para provedores de serviço de Internet. Porém, dois desenvolvimentos recentes envolvem atividades políticas em torno das inseguranças de um protocolo de sinalização chamado **SS7 (Signaling System 7)**, que permite que as redes celulares conversem entre si. O protocolo é inseguro, e o Congresso solicitou à FCC que tome medidas para resolver algumas dessas inseguranças. Outro desenvolvimento interessante relacionado à Lei de Telecomunicações de 1996 é a forma como as mensagens de texto são classificadas; ao contrário do tráfego de voz na rede telefônica, que é classificado como um serviço de comunicações (como chamadas telefônicas), as mensagens de SMS ("mensagens de texto") são classificadas como um serviço de informação (semelhante a mensagens instantâneas ou outros serviços de comunicação da Internet), que as sujeita a conjuntos de regulamentos muito diferentes sobre tudo, desde como elas podem ser cobradas às regras de privacidade que as regem.

## 2.11 RESUMO

A camada física é a base de todas as redes. A natureza impõe dois limites fundamentais sobre todos os canais, e estes determinam sua largura de banda. Esses limites são o limite de Nyquist, que trata de canais sem ruídos, e o limite de Shannon, que trata de canais com ruídos.

Os meios de transmissão podem ser guiados ou não guiados. Os principais meios guiados são o par trançado, o cabo coaxial e a fibra óptica. Dentre os meios não guiados estão o rádio terrestre, as micro-ondas, os raios infravermelhos, os raios laser através do ar e os satélites.

Métodos de modulação digitais enviam bits por meios guiados e não guiados como sinais analógicos. Os códigos de linha operam na banda base, e os sinais podem ser colocados em uma banda passante por meio da modulação da amplitude, da frequência e da fase de uma portadora. Os canais podem ser compartilhados entre usuários com a multiplexação por divisão de tempo, frequência e código.

Um elemento-chave na maioria das redes a longa distância é o sistema de telefonia implantado. Seus principais componentes são os circuitos terminais, troncos e switches. A ADSL oferece velocidades de até 40 Mbps ao circuito terminal do assinante, dividindo-o em muitas subportadoras que trabalham em paralelo. Isso ultrapassa muito as taxas dos modems de telefone. PONs levam a fibra até a residência, tornando as taxas de acesso ainda maiores do que a ADSL. Os troncos transportam informações digitais e são multiplexados com WDM para oferecer muitos enlaces de alta capacidade pelas fibras individuais, bem como TDM para compartilhar cada enlace com taxa alta entre os usuários. A comutação de circuitos e a comutação de pacotes desempenham papéis importantes.

Outro sistema para acesso à rede é a infraestrutura de cabo, que gradualmente foi evoluída de cabo coaxial para cabo híbrido de coaxial e fibra, em que muitos provedores de serviço de Internet a cabo agora oferecem até 1 Gbps (e, dentro de alguns anos, provavelmente 10 Gbps) aos assinantes. A arquitetura dessas redes, no entanto, é muito diferente, pois a capacidade da rede é compartilhada entre os assinantes no mesmo nó de serviço.

Para aplicações móveis, o sistema telefônico fixo não é adequado. Hoje, os telefones celulares estão sendo amplamente utilizados para voz e dados; desde o 4G, toda a voz, na verdade, é transportada pela rede de comutação de pacotes. A primeira geração, 1G, era analógica e dominada pelo AMPS. A 2G era digital, com GSM sendo atualmente o sistema de telefonia móvel mais utilizado no mundo. A 3G é digital e se baseia no CDMA de banda larga. A principal inovação do 4G foi a mudança para um núcleo comutado por pacotes. 5G é definido por células de menor tamanho, MIMO maciço e o uso de um espectro significativamente maior.

Em última análise, muitos aspectos da camada física são determinados não apenas pelas próprias tecnologias, mas também por organizações políticas, como órgãos de padronização e agências reguladoras. Uma área da camada física que é bastante dinâmica no cenário político é o espectro sem fio, em sua grande parte altamente regulamentado. À medida que cresce a necessidade de mais largura de banda para comunicações de dados, as agências reguladoras estão ativamente procurando maneiras de usar o espectro existente de forma mais eficiente, como reapropriação e leilão de partes do espectro previamente alocado.

## PROBLEMAS

- Um oleoduto é um sistema simplex, um sistema half-duplex, um sistema full-duplex ou nenhum dos anteriores? E um rio ou uma comunicação no estilo walkie-talkie?
- Quais são as vantagens da fibra óptica em relação ao cobre como meio de transmissão? Existe alguma desvantagem no uso da fibra óptica em relação ao cobre?
- Qual é a largura de banda existente em 0,1 micron de espectro em um comprimento de onda de 1 micron?
- Queremos enviar uma sequência de imagens de tela de computador por uma fibra óptica. A tela tem  $3.840 \times 2.160$  pixels, e cada pixel tem 24 bits. Há 60 imagens de tela por segundo. Qual é a taxa de dados necessária?
- Na Figura 2.5, a banda da esquerda é mais estreita que as outras. Por quê?
- Imagine que as operações realizadas pelos computadores digitais, atualmente implementadas por meio de sinais elétricos, pudessem ser implementadas com eficiência por meio de feixes de luz. Como isso afetaria a comunicação digital? Por que os computadores modernos não trabalham dessa forma?
- Em geral, as antenas de rádio funcionam melhor quando seu diâmetro é igual ao comprimento das ondas de rádio. Uma variação razoável para o diâmetro das antenas é de 1 cm a 1 m. Que faixa de frequências é coberta por esse intervalo?
- O enfraquecimento por múltiplos caminhos é maximizado quando os dois raios chegam com uma defasagem de 180 graus. Quanta diferença de caminho é necessária para maximizar o enfraquecimento para um enlace de micro-ondas de 1 GHz por 100 km?
- Um feixe de raios laser de 1 mm está orientado para um detector de 1 mm localizado a 100 m de distância no telhado de um edifício. Quanto desvio angular (em graus) o laser precisa ter antes de perder o detector?
- Calcule os coeficientes de Fourier para a função  $f(t) = t$  ( $0 \leq t \leq 1$ ).
- Um sinal binário de 5 GHz é enviado por um canal com uma relação sinal-ruído de 40 dB. Qual é o limite superior mais baixo na taxa de dados máxima? Explique sua resposta.
- Um canal sem ruído de 3 kHz tem uma amostra a cada 1 ms. Qual é a taxa máxima de dados desse canal? Como a taxa máxima de dados muda se o canal tiver ruído, com uma relação sinal-ruído de 30 dB?
- O teorema de Nyquist também se aplica à fibra óptica de alta qualidade em modo único, ou somente ao fio de cobre?
- Os canais de televisão têm uma largura de 6 MHz. Quantos bits/s poderão ser enviados, se forem usados sinais digitais de quatro níveis? Considere um canal sem ruído.
- Se um sinal binário for enviado sobre um canal de 3 kHz cuja relação sinal-ruído é de 20 dB, qual será a taxa máxima de dados que poderá ser alcançada?
- Um canal usando a codificação 4B/5B envia dados a uma taxa de 64 Mbps. Qual é a largura de banda mínima usada por esse canal?
- Em um diagrama de constelação, todos os pontos se encontram no eixo horizontal. Que tipo de modulação está sendo usado?
- Uma estação usando QAM-16 pode enviar 3 bits por símbolo? Explique por que (ou não).
- Qual é a largura de banda mínima necessária para alcançar uma taxa de dados de  $B$  bits/s se o sinal for transmitido usando as codificações NRZ, MLT-3 e Manchester? Justifique sua resposta.
- Prove que, nos dados mapeados em 4B/5B com a codificação NRZI, uma transição de sinal ocorrerá em tempos de pelo menos 4 bits.
- Um diagrama de constelação para modems semelhante ao da Figura 2.17 tem pontos de dados nas seguintes coordenadas: (1,1), (1, -1), (-1, 1) e (-1, -1). Quantos bps um modem com esses parâmetros pode alcançar a uma taxa de transmissão de 1.200 símbolos/s?
- Quantas frequências um modem full-duplex QAM-64 utiliza?
- Dez sinais, cada um exigindo 4.000 Hz, são multiplexados em um único canal utilizando FDM. Qual é a largura de banda mínima exigida para o canal multiplexado? Suponha que as bandas de proteção tenham 400 Hz de largura.
- Suponha que  $A$ ,  $B$  e  $C$  estejam simultaneamente transmitindo bits 0, usando um sistema CDMA com as sequências de chips da Figura 2.22(a). Qual é a sequência de chips resultante?
- Na discussão sobre ortogonalidade das sequências de chips CDMA, foi dito que, se  $\mathbf{S} \cdot \mathbf{T} = 0$ , então  $\mathbf{S} \cdot \mathbf{T}$  também é 0. Prove isto.
- Considere um modo diferente de observar a propriedade de ortogonalidade das sequências de chips do CDMA. Cada bit em um par de sequências pode coincidir ou não. Expresse a propriedade de ortogonalidade em termos de correspondências e não correspondências.
- Um receptor CDMA recebe os seguintes chips: (-1 +1 -3 +1 -1 -3 +1 +1). Supondo as sequências de chips definidas na Figura 2.22(a), quais estações transmitiram, e quais bits cada uma enviou?
- Na Figura 2.22, existem quatro estações que podem transmitir. Suponha que mais quatro estações sejam acrescentadas. Forneça as sequências de chips dessas estações.
- Qual é a probabilidade de que duas sequências de chips aleatórias com comprimento 128 tenham um produto interno normalizado de 1/4 ou mais?
- Nas redes de telefonia (fixa) e de televisão, vários usuários finais ainda estão conectados a uma única estação final, headend ou nó de fibra. Esses sistemas podem ser mais tolerantes a falhas do que o telefone tradicional discutido no Capítulo 1?
- Quantos códigos de estações finais existiam antes de 1984, quando cada estação final era identificada por seu código de área de três dígitos e pelos três primeiros dígitos do número local? Os códigos de área se iniciavam com um dígito no intervalo de 2 a 9, tinham 0 ou 1 como o segundo e terminavam com qualquer dígito. Os dois primeiros dígitos de um

- número local sempre estavam no intervalo de 2 a 9. O terceiro dígito podia ser qualquer um.
32. Um sistema telefônico simples consiste em duas estações finais e uma única estação interurbana, à qual cada estação final está conectada por um tronco full-duplex de 1 MHz. Um telefone comum é usado para fazer quatro ligações em um dia útil de 8 horas. A duração média de cada chamada é de 6 minutos e 10% das chamadas são interurbanas (ou seja, passam pela estação interurbana). Qual é o número máximo de telefones que uma estação final pode aceitar? (Suponha 4 kHz por circuito.) Explique por que a companhia telefônica pode decidir dar suporte a um número menor de telefones do que esse número máximo na estação final.
  33. Uma companhia telefônica regional tem 15 milhões de assinantes. Cada um de seus telefones está conectado a uma estação central por fios de cobre em par trançado. O comprimento médio desses pares trançados é de 10 km. Quanto vale o cobre contido nos circuitos terminais? Suponha que a seção transversal de cada fio seja um círculo com 1 mm de diâmetro, a densidade específica do cobre seja  $9,0 \text{ gramas/cm}^3$  e que o cobre seja vendido ao preço de 6 dólares por kg.
  34. Qual é a taxa de bits máxima alcançável em um modem padrão V.32 se a taxa baud for 4.800 e nenhuma correção de erro for usada?
  35. O custo de um microprocessador rápido diminuiu tanto que agora é possível incluir um em cada modem. De que maneira isso afeta o tratamento de erros na linha telefônica? Isso evita a necessidade de verificação/correção de erros na camada 2?
  36. Um sistema ADSL usando DMT aloca 3/4 dos canais de dados disponíveis para o enlace downstream. Ele utiliza modulação QAM-64 em cada canal. Qual é a capacidade do enlace downstream?
  37. Por que o tempo de amostragem do PCM foi definido como  $125 \mu\text{s}$ ?
  38. Qual é a relação sinal-ruído necessária para colocar uma portadora T1 em uma linha de 1 MHz?
  39. Compare a taxa máxima de dados de um canal sem ruído de 4 kHz usando:
    - (a) Codificação analógica (p. ex., QPSK) com 2 bits por amostra.
    - (b) O sistema T1 PCM.
  40. Se um sistema de portadora T1 apresentar uma falha e perder o controle de onde está, ele tentará se sincronizar novamente usando o primeiro bit de cada quadro. Em média, quantos quadros terão de ser examinados para que seja feita a resincronização com uma probabilidade de erro de 0,001?
  41. Qual é o percentual de overhead em uma portadora T1, ou seja, que percentagem dos 1,544 Mbps não é entregue ao usuário final? Como isso se relaciona ao percentual de overhead nas linhas OC-1 e OC-768?
  42. Os clocks da SONET têm uma taxa de variação de aproximadamente uma parte em  $10^9$ . Quanto tempo a variação leva para igualar a largura de 1 bit? Quais são as implicações desse cálculo, se houver?
  43. Na Figura 2.35, a taxa de dados do usuário para OC-3 é de 148,608 Mbps. Mostre como esse número pode ser derivado dos parâmetros OC-3 da SONET. Quais serão as taxas de dados bruta, SPE e do usuário de uma linha OC-3072?
  44. Para acomodar taxas de dados mais baixas que STS-1, a SONET tem um sistema de tributários virtuais (VTs). Um VT é uma carga útil parcial que pode ser inserida em um quadro STS-1 e combinada com outras cargas úteis parciais para preencher o quadro de dados. O VT1.5 utiliza 3 colunas, o VT2 usa 4 colunas, o VT3 usa 6 colunas e o VT6 usa 12 colunas de um quadro STS-1. Qual VT pode acomodar:
    - (a) Um serviço DS-1 (1,544 Mbps)?
    - (b) Um serviço europeu CEPT-1 (2,048 Mbps)?
    - (c) Um serviço DS-2 (6,312 Mbps)?
  45. Qual é a largura de banda disponível para o usuário em uma conexão OC-12c?
  46. Qual é a diferença, se houver, entre a parte demoduladora de um modem e a parte codificadora de um codec? (Afinal, ambos convertem sinais analógicos em sinais digitais.)
  47. Três redes de comutação de pacotes possuem  $n$  nós cada uma. A primeira rede tem uma topologia em estrela com um switch central, a segunda é um anel (bidirecional) e a terceira é totalmente interconectada, com um fio interligando cada nó. Quais são as opções de caminhos de transmissão em hops no melhor caso, no caso médio e no pior caso?
  48. Compare o atraso no envio de uma mensagem de  $x$  bits sobre um caminho de  $k$  hops em uma rede comutada por circuitos e em uma rede (levemente carregada) comutada por pacotes. O tempo de configuração de circuitos é de  $s$  segundos, o atraso de propagação é de  $d$  segundos por hop, o tamanho do pacote é de  $p$  bits e a taxa de dados é de  $b$  bps. Sob quais condições a rede de pacotes tem um atraso mais baixo? Além disso, explique as condições sob as quais uma rede de comutação de pacotes é preferível a uma rede de comutação de circuitos.
  49. Suponha que  $x$  bits de dados do usuário tenham de ser transmitidos por um caminho de  $k$  hops em uma rede comutada por pacotes como uma série de pacotes, cada um contendo  $p$  bits de dados e  $h$  bits de cabeçalho, sendo  $x \gg p + h$ . A taxa de bits das linhas é  $b$  bps e o atraso de propagação é desprezível. Que valor de  $p$  minimiza o atraso total?
  50. Em um sistema telefônico típico com células hexagonais, é proibido reutilizar uma banda de frequências em uma célula adjacente. Se 840 frequências estão disponíveis, quantas podem ser utilizadas em uma determinada célula?
  51. O layout real de células raramente é tão regular quanto o da Figura 2.39. Mesmo as formas de células individuais em geral são irregulares. Apresente uma possível razão para isso. Como essas formas irregulares afetam a atribuição de frequência de cada célula?
  52. Faça uma estimativa do número de microcélulas PCS com 100 m de diâmetro que seriam necessárias para cobrir a cidade de São Francisco ( $120 \text{ km}^2$ ).

53. Às vezes, quando um usuário móvel cruza o limite de uma célula para outra, a chamada atual é encerrada de forma brusca, embora todos os transmissores e receptores estejam funcionando perfeitamente. Por quê?
54. Na extremidade baixa, o sistema telefônico tem forma de estrela, com todos os circuitos terminais em uma vizinhança convergindo em uma estação final. Ao contrário, a televisão a cabo consiste em um único cabo longo que passa por todas as casas no mesmo bairro. Suponha que um cabo de TV do futuro fosse uma fibra de 10 Gbps, em vez de um fio de cobre. Ele poderia ser usado para simular o modelo de telefonia em que todos têm sua própria linha privada até a estação final? Nesse caso, quantas casas com um telefone poderiam ser conectadas a uma única fibra?
55. Um sistema de TV a cabo tem 100 canais comerciais, todos eles alternando programas com os anúncios. Isso tem mais semelhança com TDM ou FDM?
56. Uma empresa de serviços a cabo decide oferecer acesso à Internet por cabo em um bairro que tem 5 mil casas. A empresa utiliza um cabo coaxial e uma alocação de espectro que permite alcançar a largura de banda de 100 Mbps downstream por cabo. Para atrair clientes, a empresa decide garantir pelo menos 2 Mbps de largura de banda downstream para cada casa em qualquer instante. Descreva o que a empresa de serviços a cabo precisa fazer para fornecer essa garantia.
57. Usando a alocação espectral mostrada na Figura 2.46 e as informações dadas no texto, quantos Mbps um sistema de cabo aloca para o tráfego upstream e quantos para o tráfego downstream?
58. Com que velocidade um usuário de cabo recebe dados se a rede estiver ociosa? Suponha que a interface do usuário seja:
- Ethernet a 10 Mbps.
  - Ethernet a 100 Mbps.
  - Rede sem fio a 54 Mbps.
59. Os 66 satélites de baixa órbita do projeto Iridium estão divididos em seis eixos em torno da Terra. Na altitude em que eles se encontram, o período é de 90 minutos. Qual é o intervalo médio entre handoffs no caso de um transmissor estacionário?
60. Considere um satélite na altitude de satélites geoestacionários, mas cujo plano orbital está inclinado em relação ao plano equatorial por um ângulo  $\phi$ . Para um usuário estacionário na superfície da Terra na latitude norte  $\phi$ , esse satélite parece imóvel no céu? Se não, descreva seu movimento.
61. Calcule o tempo de trânsito de ponta a ponta para um pacote trafegar pelos satélites GEO (altitude: 35.800 km), MEO (altitude: 18.000 km) e LEO (altitude: 750 km).
62. Qual é a latência de uma chamada originada no Polo Norte para alcançar o Polo Sul se a chamada for roteada por satélites Iridium? Suponha que o tempo de comutação nos satélites seja 10 microssegundos e o raio da Terra seja 6.371 km.
63. Quanto tempo levará para transmitir um arquivo de 1 GB de um VSAT para outro usando um hub como o mostrado na Figura 2.50? Suponha que o uplink seja de 1 Mbps, o downlink seja de 7 Mbps e a comutação de circuitos seja usada com um tempo de configuração de circuito de 1,2 s.
64. Calcule o tempo de trânsito no problema anterior se for usada a comutação de pacotes. Suponha que o tamanho do pacote seja de 64 KB, o atraso de comutação no satélite e no hub seja de 10 microssegundos e o tamanho do cabeçalho do pacote seja 32 bytes.
65. A multiplexação de vários fluxos de dados STS-1, chamados tributários, desempenha um papel importante na SONET. Um multiplexador 3:1 efetua a multiplexação de três tributários STS-1 de entrada em um único fluxo STS-3 de saída. Essa multiplexação é feita byte a byte, isto é, os três primeiros bytes de saída são os primeiros bytes dos tributários 1, 2 e 3, respectivamente. Os três bytes de saída seguintes são os próximos bytes dos tributários 1, 2 e 3, respectivamente, e assim por diante. Escreva um programa que simule esse multiplexador 3:1. Seu programa deve consistir em cinco processos. O principal cria quatro processos, um para cada um dos três tributários STS-1 e um processo para o multiplexador. Cada processo tributário lê um quadro STS-1 de um arquivo de entrada como uma sequência de 810 bytes. Eles enviam seus quadros (byte por byte) ao processo multiplexador. Este recebe esses bytes e transmite como saída um quadro STS-3 (byte por byte), gravando-o na saída padrão. Utilize pipes para efetuar a comunicação entre os processos.
66. Escreva um programa para implementar CDMA. Suponha que a extensão de uma sequência de chips seja oito e o número de estações transmitindo seja quatro. Seu programa deve consistir em três conjuntos de processos: quatro processos transmissores ( $t_0, t_1, t_2$  e  $t_3$ ), um processo de junção e quatro processos receptores ( $r_0, r_1, r_2$  e  $r_3$ ). O programa principal, que também atua como o processo de junção, primeiro lê quatro sequências de chips (notação bipolar) da entrada padrão e uma sequência de 4 bits (1 bit por processo transmissor a ser transmitido), e se bifurca em quatro pares de processos transmissores e receptores. Cada par de processos transmissor/receptor ( $t_0, r_0; t_1, r_1; t_2, r_2; t_3, r_3$ ) é atribuído a uma sequência de chips e cada processo transmissor é atribuído a um 1 bit (primeiro bit para  $t_0$ , segundo bit para  $t_1$ , e assim por diante). Em seguida, cada processo transmissor calcula o sinal a ser transmitido (uma sequência de 8 bits) e o envia para o processo de junção. Depois de receber sinais de todos os processos transmissores, o processo de junção combina os sinais e envia o sinal combinado aos quatro processos receptores. Então, cada processo receptor calcula o bit que recebeu e o imprime na saída padrão. Use pipes para a comunicação entre processos.