

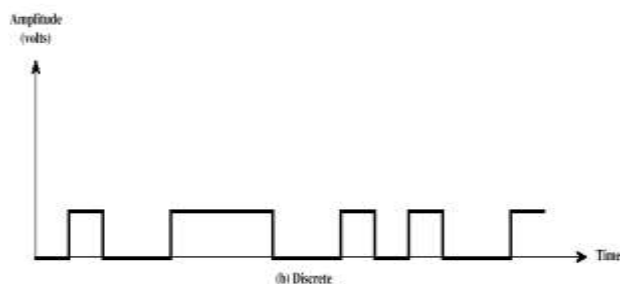
TRANSMISSÃO DE DADOS

O sucesso de uma transmissão de dados depende principalmente de dois fatores: (i) da qualidade do sinal sendo transmitido e, (ii) das características do meio de transmissão. O objetivo deste capítulo é prover o leitor com um sentimento intuitivo da natureza desses dois fatores.

1. Base Teórica sobre Sinais de Transmissão

Toda transmissão de dados ocorre pela inserção de sinais, em geral na forma de ondas eletromagnéticas, em algum sistema de transmissão. O sucesso dessa transmissão depende principalmente de dois fatores: (i) a qualidade do sinal transmitido e (ii) as características do meio de transmissão. Por isso, antes de formalizar os conceitos sobre sinais, é preciso primeiro repassar algumas terminologias utilizadas na comunicação de dados:

- ⇒ Sobre os meios de transmissão: Esses podem ser guiados se fizerem uso de cabos (tipo par trançado, coaxial ou fibra), ou podem ser não guiados se usarem meios de transmissão dispersos (como o espaço aéreo, água do mar, vácuo etc.) como caminho para a transmissão. Em ambos os casos a comunicação, em geral ocorre na forma de ondas eletromagnéticas.
- ⇒ Sobre a forma de utilização do meio de transmissão: Se os dispositivos comunicantes se interligam por um enlace direto (sem qualquer dispositivo ativo, a não ser amplificadores ou repetidores de sinais), diz-se que o caminho entre eles é ponto-a-ponto. Numa configuração multi-ponto ou *broadcast* mais de dois dispositivos compartilham o mesmo meio de transmissão.
- ⇒ Sobre a tipo de diálogo entre os dispositivos comunicantes: Uma transmissão pode ser *simplex*, *half-duplex* ou *full-duplex*. No primeiro caso, os sinais são transmitidos apenas em uma direção, configurando uma estação como transmissora e a outra como receptora. No segundo caso, ambas as estações podem transmitir, mas apenas uma por vez. Na transmissão full-duplex ambas as estações podem transmitir ao mesmo tempo.



sinal periódico digital (onda quadrada).

Matematicamente, um sinal $s(t)$ é definido como periódico se e somente se:

$$s(t + T) = s(t); \quad -\infty < t < +\infty,$$

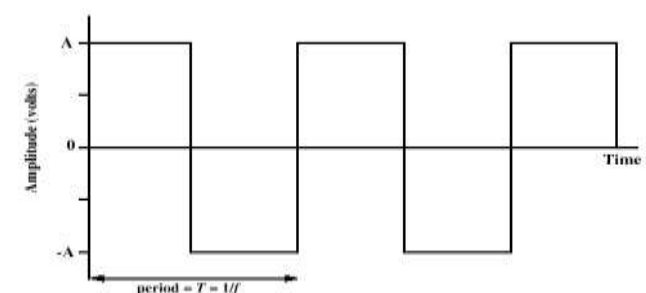
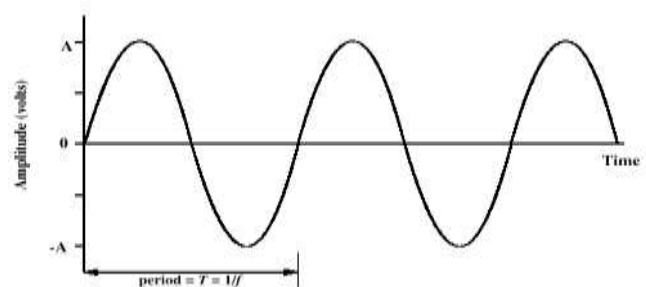
onde a constante T é o período do sinal, ou seja, é o menor valor que satisfaz a equação. O sinal será considerado aperiódico se não seguir esta formalização.

Dentro do contexto da representação de sinais, a onda senoidal é o sinal contínuo fundamental e, de forma genérica, pode ser representada por três parâmetros:

1.1 Definição de Sinais no Domínio do Tempo

Visualizado como uma função do tempo, um sinal eletromagnético pode ser ou contínuo ou discreto. Um sinal contínuo é aquele no qual a sua intensidade varia suavemente no tempo, ou seja, não existem quebras ou descontinuidades no sinal¹. Um sinal discreto, por sua vez, é aquele no qual a intensidade desse sinal se mantém num nível constante por um período de tempo e então muda para um outro nível constante. A figura ao lado apresenta ambos os tipos de sinais. O sinal contínuo poderia representar a fala humana e o sinal discreto poderia representar os binários 1s e 0s.

O mais simples tipo de sinal é o sinal periódico, no qual um mesmo padrão de sinal se repete com o tempo. A figura abaixo apresenta um exemplo de um sinal periódico analógico (onda senoidal) e um



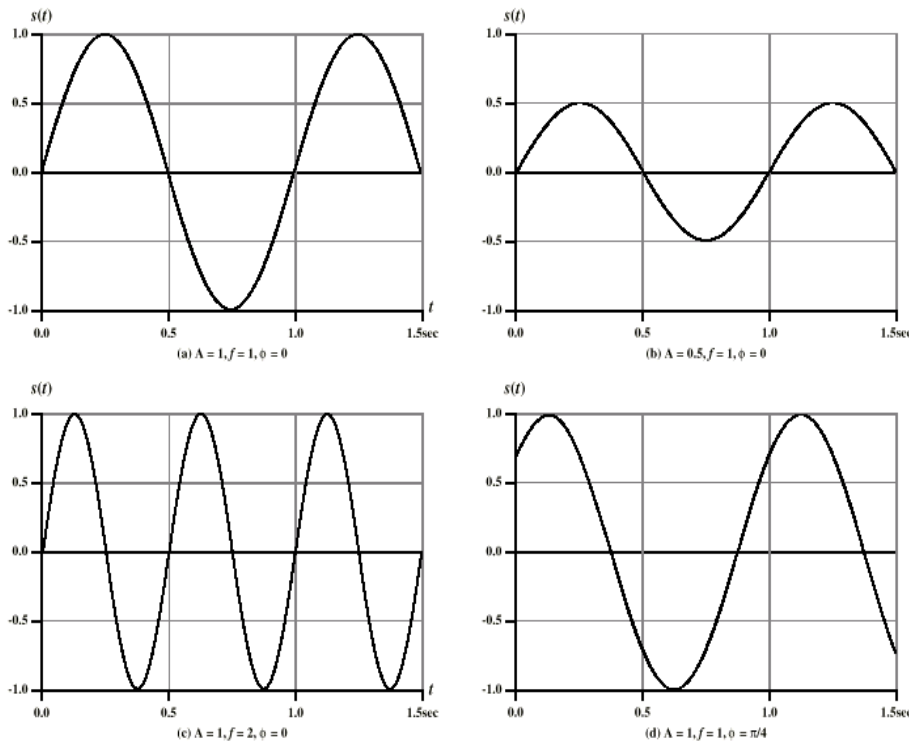
¹ Uma definição matemática: Um sinal $s(t)$ é contínuo se $\lim_{a \rightarrow t} s(a) = s(t)$ para todo $a, t \rightarrow a$.

- i. Amplitude - é o valor de pico ou a força do sinal em relação ao tempo. Tipicamente, este valor é medido em volts ou watts.
- ii. Frequência (f) - é a taxa (em ciclos por segundo ou Hertz) na qual o sinal se repete. Um parâmetro equivalente é o período (T) de um sinal, o qual é a quantidade de tempo que leva para que uma repetição ocorra. Portanto, $T = 1/f$.
- iii. Fase - é a medida da posição relativa dessa onda no tempo em relação a um único período do sinal, como ilustrado na figura abaixo. Para compreendê-la, devemos entender que uma onda senoidal genérica pode ser escrita da seguinte maneira:

$$s(t) = A \sin(2\pi ft + \phi),$$

onde: A = Amplitude do sinal; f = frequência do sinal; ϕ = Fase do sinal

A figura ao lado apresenta o efeito de variar cada um desses três parâmetros. No primeiro quadro, a frequência é 1 Hz e portanto o período é $T = 1$ segundo. O segundo quadro (da direita) tem a mesma frequência e fase, mas a amplitude é de $1/2$. No terceiro quadro, tem-se $f = 2$, o qual é equivalente a $T = 1/2$. Finalmente, o último quadro apresenta o efeito de uma fase alterada para $\pi/4$ radianos, a qual é equivalente a 45 graus (2π radianos é igual a 360 graus = 1 período).



1.1.1 Ondas Senoidais em Relação ao Espaço

Na figura anterior, houve a apresentação do sinal em função do tempo num dado ponto do espaço. Se, por outro lado, considerarmos o ponto zero do gráfico como sendo a origem do sinal e o eixo horizontal como sendo o eixo do espaço, será possível perceber o comportamento do sinal em relação à sua origem, como uma onda de rádio eletromagnética distante em relação a uma antena de rádio, ou o formato de um sinal sonoro alguma distância em relação à caixa de som que lhe deu origem. Portanto, pode-se concluir que em um instante particular de tempo a intensidade do sinal varia em uma forma senoidal como uma função da distância da fonte.

Sobre o ponto de vista do espaço, pode-se considerar o comprimento de onda, λ , de um sinal como sendo a distância ocupada por um único ciclo, ou, de outra maneira, como a distância entre dois pontos de fase correspondente de dois ciclos consecutivos. Assumindo que o sinal está viajando com uma velocidade v então o comprimento de onda é relacionado ao período da seguinte maneira: $\lambda = vT$, ou, equivalentemente, $\lambda f = v$. Essa discussão tem relevância particular para o caso onde $v = c$, ou seja, a velocidade da luz no espaço livre, a qual é igual a 3×10^8 m/s.

1.2 Definição de Sinais no Domínio da Frequência

A informação pode ser transmitida em cabos pela variação de alguma de suas propriedades físicas tais como a voltagem ou a corrente, permitindo formalizar sinais em função do tempo. No entanto, pode-se demonstrar também (usando uma disciplina conhecida como análise de Fourier) que qualquer sinal é formado por componentes de diferentes frequências, na qual cada componente é uma senoidal. Um exemplo disso está apresentado na próxima figura. As componentes deste sinal são exatamente as ondas senoidais de frequências f_1 e $3f_1$, nesse caso chamadas de componentes de $s(t)$. As partes (a) e (b) da figura a seguir apresentam essas componentes individualmente. A parte (c) é a onda resultante da soma das duas anteriores, representada pela seguinte fórmula:

$$s(t) = \sin(2\pi f_1 t) + 1/3 \sin(2\pi(3f_1)t)$$

Existem vários pontos de vista interessantes que podem ser percebidos nesta figura:

- A segunda frequência é um múltiplo inteiro da primeira frequência. Quando todas as componentes de frequência de um sinal são múltiplos inteiros de uma frequência, essa última é referenciada como sendo a frequência fundamental.
- O período do sinal resultante é igual ao período da frequência fundamental, ou seja, o período da componente $\sin(2\pi f_1 t)$ é $T = 1/f_1$ e o período de $s(t)$ é também T .

- Seguindo o raciocínio proposto na figura, ao adicionarmos mais um ou dois termos ou componentes² à essa onda, as partes alta e baixa da onda resultante tenderá a um sinal discreto, se aproximando cada vez mais de uma onda quadrada. Para mais dois harmônicos a fórmula ficaria da seguinte maneira:

$$s(t) = \sin(2\pi f_1 t) + \frac{1}{3} \sin(2\pi(3f_1)t) + \frac{1}{5} \sin(2\pi(5f_1)t) + \frac{1}{7} \sin(2\pi(7f_1)t)$$

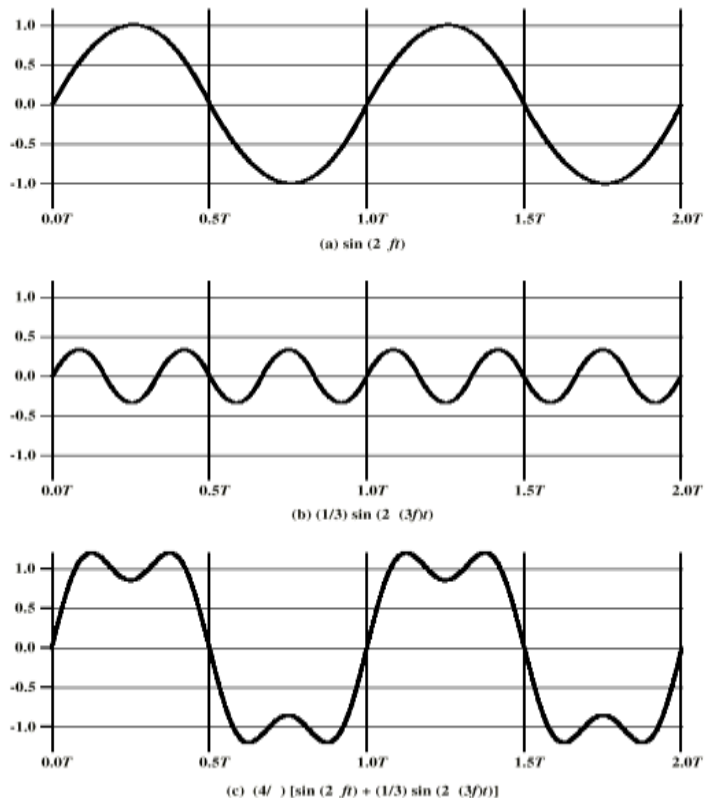
- Uma onda quadrada genérica possui a seguinte fórmula (originada da série de Fourier), onde a amplitude da k-ésima componente de frequência, Kf_1 , é apenas $1/k$ (nesse caso, A = Amplitude do sinal e f_1 é a frequência desse sinal):

$$s(t) = A \times \sum_{K=1,3,5,\dots}^{\infty} \frac{1}{K} \sin(2\pi K f_1 t),$$

Portanto, pode-se dizer que para cada sinal, existe uma função no domínio do tempo $s(t)$ que especifica a amplitude do sinal no instante do tempo e, similarmente, existe uma função $S(f)$ que especifica as frequências constituintes deste sinal.

Antes de continuarmos nossa discussão sobre as componentes de um sinal, é válido definir alguns conceitos no domínio da frequência cuja compreensão são importantes:

- ⇒ **Espectro** - O espectro de um sinal é o domínio de frequências que ele contém. Para o sinal da figura ao lado, o espectro se estende de f_1 até $3f_1$.
- ⇒ **Largura de banda absoluta** de um sinal é a largura do espectro. No caso da figura ao lado, a largura de banda é $2f_1$, ou seja, a diferença entre a maior e a menor frequência do sinal. Vale ressaltar que muitos sinais tem uma largura de banda infinita mas a maior parte da energia deles está contida nas primeiras frequências, formando uma banda relativamente estreita de frequências. Esta banda onde a maior parte da energia está concentrada é referenciada como a largura de banda efetiva, ou, simplesmente largura de banda da frequência fundamental.



1.2.1 Relação entre Taxa de Transmissão e Largura de Banda

É sabido que nenhum meio de transmissão consegue transmitir sinais sem perdas porque cada meio possui uma largura de banda inerente às suas propriedades físicas, o que limita a quantidade de frequências do sinal que podem ser transportadas. Por outro lado, baseado no que foi dito na seção anterior, a maior parte da energia do sinal se situa nas primeiras frequências ou componentes deste sinal. Portanto, é possível transportar um sinal digital, não como uma onda quadrada, mas como uma aproximação dessa onda quadrada (considerando somente as frequências que armazenam uma quantidade maior de energia e desconsiderando as demais) cujas bases superior e inferior serão mais retas ou abauladas em função das capacidades de percepção do receptor.

Em função disso, pode-se dizer que a taxa de transmissão de um sinal depende basicamente de dois fatores:

- i. Do método de codificação que é usado para representar os bits. Existem vários métodos de representação de bits, sendo que os pioneiros são aqueles em que os bits são representados por voltagens fixas pré-estabelecidas. Esse assunto será visto com detalhes mais adiante e vamos, por agora, considerar apenas a situação na qual o bit zero é representado pela parte baixa de uma onda quadrada (ou aproximada) e o bit um é representado pela parte alta dessa mesma onda quadrada.
- ii. Da quantidade de alterações que o sinal sofre no tempo³, ou seja, quantas vezes ele oscila entre as voltagens alta e baixa no tempo. No caso da Figura anterior, considerando que o pulso positivo representa o binário 1 e o pulso negativo representa o binário zero tem-se a representação da sequência 1010.... A duração de cada pulso é igual a $1/2f_1$; portanto, a taxa de dados é $2f_1$ bits por segundo (bps). Em outras palavras, a taxa de transmissão está diretamente relacionada à frequência fundamental do sinal, que é a região de maior energia do sinal.

Portanto, um meio de transmissão com determinada largura de banda pode produzir diferentes taxas de transmissão em função dos seguintes fatores:

- i. Da qualidade do sinal implícita no número de componentes que o sinal necessita para que o receptor possa identificar os bits. Significa dizer que, para dada uma largura de banda, quanto maior o número de componentes, menor deve ser a frequência fundamental e, consequentemente, a taxa de transmissão;

² Os termos são relacionados às funções seno e co-seno da Série de Fourier. Portanto cada termo possui dois harmônicos.

³ Essa definição é conhecida como baud.

- ii. Da frequência fundamental necessária para gerar os bits embutidos nas voltagens do sinal. Ou seja, quanto maior for a frequência fundamental maior a taxa de transmissão. Em contrapartida, é possível provar que a frequência fundamental não pode ser mais que a metade da largura de banda do meio.

Em outras palavras, se o sinal contiver muitas componentes, a qualidade do sinal é boa, mas a taxa de transmissão pode ser baixa. Em contrapartida, se o número de componentes for menor, a qualidade do sinal diminui mas a taxa de transmissão pode aumentar⁴. Vejamos alguns exemplos:

1) Vamos supor que estamos utilizando um sistema de transmissão digital que é capaz de transmitir sinais com uma largura de banda de 4MHz e vamos transmitir uma sequência alternada de 1s e 0s como uma onda quadrada considerando apenas duas componentes. Qual a taxa de dados que pode ser alcançada? Embora esta forma de onda seja distorcida de uma onda quadrada, ela é suficiente para que o receptor possa identificar um bit zero de um bit 1. Agora, se tomarmos a frequência fundamental $f_1 = 10^6$ ciclos/segundo = 1 MHz, então a largura de banda do sinal:

$$s(t) = \sin(2\pi \times 10^6 t) + \frac{1}{3} \sin(2\pi(3 \times 10^6) t) + \frac{1}{5} \sin(2\pi(5 \times 10^6) t)$$

é $(5 \times 10^6) - 10^6 = 4\text{MHz}$. Note que para $f_1 = 1\text{MHz}$ o período da frequência fundamental é $T = 1/10^6 = 1\mu\text{segundo}$ (microsegundo). Se tratarmos este comprimento de onda como uma sequência de 1s e 0s, um bit ocorre a cada $0,5 \mu\text{s}$, formando uma taxa de dados de $2 \times 10^6 = 2\text{Mbps}$. Portanto, para uma largura de banda de 4MHz, uma taxa de dados de 2 Mbps pode ser alcançada.

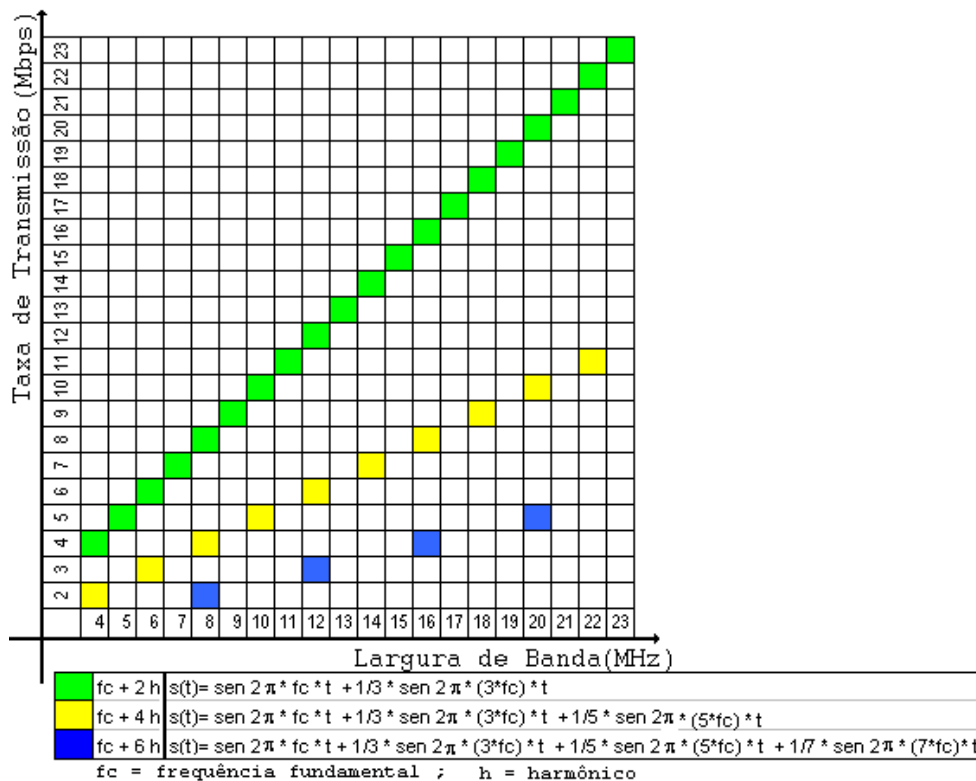
2) Considere o exemplo anterior, mas com largura de banda igual a 8MHz e frequência fundamental $f_1 = 2\text{MHz}$. Usando o mesmo raciocínio (número de componentes), a largura de banda do sinal é $(5 \times 2 \times 10^6) - 2 \times 10^6 = 8\text{MHz}$. Mas, neste caso $T = 1/f_1 = 0,5\mu\text{s}$. Como resultado disto, um bit ocorre a cada $0,25\mu\text{s}$ para uma taxa de dados de 4Mbps. A partir deste resultado, considerando constante os demais fatores, pode-se concluir que, ao dobrar a largura de banda do meio, dobra-se também a potencial taxa de dados.

3) Considerando um sistema de transmissão com capacidade de transmitir sinais com largura de banda de 4MHz, se fizermos a frequência fundamental $f_1 = 2\text{MHz}$ e com apenas uma componente extra (mínimo aproximado de onda quadrada para que os bits sejam perceptíveis), usando a mesma linha de raciocínio, a largura de banda do sinal é $(3 \times 2 \times 10^6) - 2 \times 10^6 = 4\text{MHz}$. Mas, neste caso, $T = 1/f_1 = 0,5\mu\text{s}$. Como resultado, um bit ocorre a cada $0,25\mu\text{s}$, totalizando uma taxa de dados na ordem de 4Mbps.

Pode-se desenhar as conclusões gerais a partir do que foi exposto baseado nos resultados dos exemplos:

- Uma dada largura de banda pode suportar várias taxas de dados dependendo dos requerimentos do receptor, de acordo com o número de componentes do sinal e da largura de banda (frequência fundamental) definida para o mesmo. O gráfico adiante apresenta esta constatação.
- Qualquer forma de onda digital terá largura de banda infinita. Portanto, se de um lado as razões práticas e econômicas apontam para que a informação digital seja aproximada por um sinal com largura de banda limitada, por outro lado, limitar a largura de banda cria distorções, as quais tornam a tarefa de interpretar o sinal recebido mais difícil. Quanto mais limitada a largura de banda, maior a distorção e maior o potencial de erros pelo receptor.
- Se a taxa de dados do sinal digital é W bps, então uma boa representação pode ser alcançada com uma largura de banda em torno de $2W\text{Hz}$. Contudo, se o padrão de ruídos não for muito severo, é possível recuperar o padrão de bits com uma largura de banda menor do que essa.
- Existe uma relação direta entre taxa de dados e largura de banda: quanto maior a taxa de dados de um sinal, maior a largura de banda efetiva. Olhando de uma outra maneira, quanto maior a largura de banda de um sistema de transmissão, maior é a taxa de dados que pode ser transmitida sobre este sistema.
- Uma outra observação que pode ser feita é: se pensarmos que a largura de banda de um sinal está centrada sobre alguma frequência, referenciada como frequência central, então quanto maior a frequência central, maior a largura de banda potencial e, portanto, maior a potencial taxa de dados. Considerando que se um sinal está centrado em 2MHz, sua largura de banda máxima é 4MHz.

⁴ Essa diminuição, no entanto, deve ser controlada de tal modo que haja um número mínimo de componentes para que o sinal possa ser perceptível pelo receptor.



2. Imperfeições na Transmissão

Com qualquer sistema de comunicação deve estar claro para o leitor que o sinal recebido irá diferir do sinal transmitido devido às várias imperfeições na transmissão. Para sinais analógicos, essas imperfeições introduzem várias modificações randômicas que degradam a qualidade do sinal. Para sinais digitais, os erros são introduzidos pela inversão de bits, ou seja, bit zero se transforma em bit 1 e vice-versa. Nesta seção examinar-se-á as várias imperfeições e far-se-ão comentários sobre seus efeitos na capacidade de transportar informações de um link de comunicação. As imperfeições mais significantes são a atenuação, a distorção por atraso e os ruídos.

a) Atenuação - A força de um sinal cai com a distância sobre qualquer meio de transmissão. Para meios guiados, esta redução na força, ou atenuação, é geralmente logarítmica e é tipicamente expressa como um número constante de decibéis por unidade de distância. Para meios não guiados (meio aéreo), a atenuação é uma função mais complexa de distância e de makeup da atmosfera. A atenuação introduz três considerações para os engenheiros de transmissão. A primeira delas é que o sinal recebido deve ser suficientemente forte para de modo que o circuito eletrônico no receptor possa detectar e interpretar o sinal. Segundo, o sinal deve manter um nível suficientemente maior que o ruído inerente ao meio para que possa ser recebido sem erros. Terceiro, a atenuação é uma função crescente em relação à frequência. O primeiro e o segundo problemas são tratados pelo controle da força do sinal e pela utilização de amplificadores ou repetidores. Para um enlace ponto-a-ponto, a força do sinal do transmissor deve ser forte o suficiente para que a inteligibilidade do sinal recebido, mas não tão forte para gerar uma sobrecarga no circuito do transmissor, o qual poderia causar uma distorção do sinal a ser gerado. Em distâncias muito grandes, a atenuação é muito grande e, nesse caso, os repetidores e amplificadores são usados para reforçar o sinal de tempos em tempos. Esses problemas são mais complexos para enlaces multiponto onde a distância do transmissor para o receptor é variável, como é o caso das redes Cable Modem, por exemplo. O terceiro problema é particularmente notável para sinais analógicos. Por que a atenuação varia em função da frequência, o sinal recebido é distorcido, reduzindo a sua inteligibilidade. Para contornar este problema, técnicas estão disponíveis para equalizar a atenuação através de uma banda de frequências. Isto é comumente feito para linhas telefônicas de grade-voz por usar bobinas de carga enroladas que mudam as propriedades elétricas da linha; o resultado é uma diminuição dos efeitos da atenuação. Uma outra abordagem é usar amplificadores capazes de amplificar mais as frequências maiores do que as menores.

b) Distorção por atraso - É um fenômeno peculiar a meios de transmissão guiados. A distorção é causada pelo fato de que a velocidade de propagação de um sinal através do meio guiado varia com a frequência. Para um sinal com largura de banda limitada, a velocidade tende a ser a maior próxima da frequência central e menor fora das duas arestas da banda. Portanto, várias componentes de frequência de um sinal chegarão no receptor em tempos diferentes. Este efeito é referenciado como uma distorção por atraso, por que o sinal recebido é distorcido devido ao atraso variável de suas componentes. Distorção por atraso é particularmente crítico para dados digitais. Considerando o caso que uma sequência de bits está sendo transmitida, usando sinais analógicos ou digitais. Por causa da distorção por atraso, alguns dos componentes do sinal de uma posição de bit irão pular sobre outras posições de bits causando interferência intersimbólica, a qual é a maior limitação para a taxa máxima de bits sobre um controle de transmissão.

c) Ruídos - Em qualquer transmissão de dados, o sinal recebido consistirá do sinal transmitido, modificado pelas várias distorções impostas pelo sistema de transmissão, mais os sinais indesejados que são inseridos durante a transmissão. Alguns tipos de ruídos podem ser mencionados:

- ⇒ Ruído Termal - Ocorre devido à agitação de elétrons no condutor. Está presente em todos os dispositivos e meios de transmissão e é uma função da temperatura. É uniformemente distribuído através do espectro de frequências e é normalmente referenciado como ruído branco. Não pode ser eliminado e impõe limites superiores à performance dos sistemas de comunicação.
- ⇒ Ruído de Intermodulação - Ocorre quando sinais de diferentes frequências compartilham o mesmo meio de transmissão. O efeito é produzir sinais numa frequência que é a soma, a diferença ou é um múltiplo dos dois sinais originais.
- ⇒ Crosstalk - É experimentado por qualquer um que, enquanto usando o telefone, consegue ouvir uma segunda conversação (linha cruzada). Isto é, na verdade, um acoplamento não desejado entre caminhos de sinais. Pode ocorrer por acoplamento elétrico, geralmente em par trançado ou cabo coaxial.
- ⇒ Ruído Impulsivo - São alterações de energia no sinal que ocorrem de forma aleatória e podem provocar uma alteração considerável no valor da informação contida no sinal, seja ele analógico ou digital. Em particular, os sinais digitais sofrem mais com esse tipo de ruído. Por exemplo, a transmissão de voz pode ser corrompida por pequenas alterações mas as entidades comunicantes ainda conseguem compreender o diálogo. Já numa transmissão de bits, seja em sinal analógico ou digital não há como ignorar estes surtos de voltagem/ruído na linha.

2.1 Capacidade do Canal

Em função das imperfeições naturais que podem distorcer ou corromper um sinal. Para dados digitais, a questão que então surge é que a extensão dessas imperfeições limita a taxa de dados que podem ser alcançados.

A taxa na qual o dado pode ser transmitido sobre um meio de comunicação qualquer, ou canal, sobre determinadas condições, é chamada capacidade do canal. Na prática um sinal não pode ter uma taxa de transmissão infinita, por que existem as imperfeições do meio de transmissão que a impedem. Dentro deste contexto, existem quatro conceitos relacionados entre si:

- ⇒ Taxa de Dados - Isto é a taxa de dados em bits por segundo na qual um dado pode ser comunicado
- ⇒ Largura de banda - É a largura do sinal transmitido como restringido pelo transmissor e pela natureza do meio de transmissão, expresso em ciclos por segundo ou Hertz.
- ⇒ Ruído - O nível médio de interferências não desejadas no caminho ou meio de transmissão.
- ⇒ Taxa de erros - A taxa na qual os erros ocorrem, onde um erro é definido como sendo a transformação indesejada de um bit 1 transmitido em zero e vice-versa.