

Gustavo Silva Oliveira

Formatos de Modulação de uma Portadora Óptica com Detecção Direta

Trabalho de Conclusão de Curso apresentado à
Escola de Engenharia de São Carlos, da
Universidade de São Paulo

Curso de Engenharia Elétrica com ênfase em
Eletrônica

Orientadora: Prof^a. Dra. Mônica de Lacerda Rocha

São Carlos

2011

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

Oliveira, Gustavo Silva.

O48f

Formatos de modulação de uma portadora óptica com detecção direta. / Gustavo Silva Oliveira ; orientador Mônica de Lacerda Rocha -- São Carlos, 2011.

Monografia (Graduação em Engenharia Elétrica com ênfase em Eletrônica) -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2011.

1. Comunicações ópticas. 2. Formatos de modulação de fase. 3. Formatos de modulação de intensidade. 4. DPSK. 5. DQPSK. 6. DB. 7. AMI. 8. CRZ. 9. RZ-OOK. 10. NRZ-OOK. I. Título.

FOLHA DE APROVAÇÃO

Nome: Gustavo Silva Oliveira

Título: "Formatos de Modulação de uma Portadora Óptica com Detecção Direta"

Trabalho de Conclusão de Curso defendido e aprovado
em 30/11/2011,

com NOTA 6,0 (Seis, Zero), pela comissão julgadora:



MSc. Anderson Leonardo Sanches - EESC/USP



MSc. Mariana Massimino Feres - EESC/USP



Prof. Associado Homero Schiabel
Coordenador da CoC-Engenharia Elétrica
EESC/USP

Agradecimentos

À minha mãe Shirley, pela dedicação e carinho que dedica em tempo integral a seus filhos; ao meu pai Romeu, pelo exemplo de pessoa batalhadora e íntegra que é; e ao meu irmão Júlio, pela imagem de força, persistência e coragem que sempre me passou.

À Prof^a Dra. Mônica de Lacerda Rocha, pela orientação e ajuda no desenvolvimento deste trabalho.

Aos meus amigos Victor Hugo (Furmiga), Diego Meirelles (Batata) e Bruno Rogani, por tornarem mais leve e divertida essa caminhada árdua que, às vezes, é a vida.

Ao meu amigo e futuro companheiro de trabalho na Petrobrás Igor (Castanha), pelas muitas e ótimas experiências que ainda certamente estão por vir.

Aos meus colegas de classe Stênio (Tung) e João Pedro (Toca), por termos trilhado juntos a ênfase em Telecomunicações e por terem compartilhado comigo seus conhecimentos e amizade preciosa.

Ao Saul (Biro), pelo companheirismo durante esses quase cinco anos que moramos juntos.

A todos os demais colegas de classe, pela convivência harmoniosa e enriquecedora que mantivemos nesses 5 anos de curso de Engenharia Elétrica.

Resumo

Este trabalho de conclusão de graduação analisa formatos de modulação óptica alternativos ao tradicional formato binário de intensidade OOK (*on-off keying*), tais como CSRZ (*carrier-suppressed return-to-zero*), CRZ (*chirped return-to-zero*), ACRZ (*alternate-chirp return-to-zero*), CNRZ (*chirped nonreturn-to-zero*), duobinário, AMI (*alternate-mark inversion*), DPSK (*differential phase shift keying*) e DQPSK (*differential quadrature phase shift keying*). Tais formatos alternativos, ditos avançados, são uma peça-chave para atender às demandas do mercado atual de telecomunicações por transporte de grande volume de dados e alta largura de banda, bem como por redução do custo de bit da informação transmitida.

Além desses formatos avançados, que são atualmente os mais investigados por oferecerem uma tolerância maior aos efeitos degradantes ao sinal que a oferecida pelo OOK, também se faz um estudo deste último formato por ele constituir a base da modulação utilizada nos sistemas que transmitem a taxas de até 10 Gb/s.

Também são revistos a teoria sobre sistemas de comunicação óptica e os componentes que fazem parte desse sistema, como o transmissor, a fibra óptica e o receptor, por exemplo.

Algumas simulações preliminares feitas no software *Optisystem* (versão 7.0), da empresa canadense *Optiwave Systems Inc.* Nessas simulações analisou-se o desempenho do formato DPSK na configuração *back-to-back*, através do diagrama de olho e de visualizadores da forma de onda e do espectro do sinal. Com essas simulações, pôde-se observar o sinal em vários “momentos” (a sequência de bits a ser enviada, o sinal produzido pelo MZM (*Mach-Zehnder modulator*) e o sinal demodulado no receptor) e a consequente degradação sofrida por ele.

A parte final deste trabalho resume o que foi apresentado e traz as conclusões deste trabalho, como a que realça a importância dos formatos avançados de modulação para a transmissão de taxas a partir de 40 Gb/s por canal.

Palavras-chave: comunicações ópticas, formatos de modulação de fase, formatos de modulação de intensidade, DPSK, DQPSK, DB, AMI, CRZ, RZ-OOK, NRZ-OOK.

Abstract

This undergraduate thesis analyzes the most important optical modulation formats as an alternative to the use of the traditional intensity binary format OOK (on-off keying), such as CSRZ (carrier-suppressed return-to-zero), CRZ (chirped return-to-zero), ACRZ (alternate-chirp return-to-zero), CNRZ (chirped nonreturn-to-zero), duobinário, AMI (alternate-mark inversion), DPSK (differential phase shift keying) and DQPSK (differential quadrature phase shift keying). Those formats, classified as advanced, play an important role in meeting telecommunications market current demands for high data volume transmission and broadband as well as for cost reduction of the transmitted information per bit.

In addition to those advanced formats, which are currently the most investigated ones for being able to provide higher tolerance to propagation impairments imposed to the information signal than the one provided by OOK, this last format is also investigated because it is widely used in systems transmitting at rates up to 10 Gb/s.

The theory about optical communication systems and its components (transmitter, optical fiber and receptor) is also reviewed in this work.

Some preliminary simulations were made using the software *Optisystem* (version 7.0), from the Canadian company *Optiwave Systems Inc.* In these simulations, the performance of the DPSK format in the configuration back-to-back was analyzed through its eye diagram, waveform and signal spectrum. And then the signal and its degradation were observed in some “instants” of time: the sequence of bits to be sent, the signal produced by MZM (Mach-Zehnder modulator) and the signal demodulated at the receiver.

The last part of this paper summarizes everything that was discussed and also features the conclusions of this work, the most important one being the importance of advanced modulation formats for systems transmitting beyond 40 Gb/s per channel.

Key-words: optical communications, phase modulation formats, intensity modulation formats, DPSK, DQPSK, DB, AMI, CRZ, RZ-OOK, NRZ-OOK.

Sumário

AGRADECIMENTOS.....	V
RESUMO	VII
ABSTRACT	IX
LISTA DE FIGURAS	XV
LISTA DE TABELAS	XIX
LISTA DE ACRÔNIMOS	XXI
1. INTRODUÇÃO	1
2. PRINCIPAIS FATORES DE DEGRADAÇÃO DO SINAL EM FIBRAS ÓPTICAS	3
2.1. PERDAS POR ABSORÇÃO	3
2.2. PERDAS POR ESPALHAMENTO RAYLEIGH	4
2.3. PERDAS POR CURVATURAS E MICROCURVATURAS NA FIBRA, EM EMENDAS E CONECTORES	4
2.4. ATENUAÇÃO E RUÍDO DE AMPLIFICAÇÃO	5
2.5. DISPERSÃO CROMÁTICA	6
2.6. DISPERSÃO POR MODO DE POLARIZAÇÃO	7
2.7. EFEITOS NÃO LINEARES	7
2.8. FILTRAGEM ÓPTICA DE BANDA ESTREITA E DIAFONIA ENTRE CANAIS WDM.....	9
3. COMPONENTES DE UM SISTEMA DE COMUNICAÇÃO ÓPTICA.....	11
3.1. TRANSMISSOR.....	12
3.1.1. FONTES ÓPTICAS	13

3.1.1.1. LED	13
3.1.1.2. LASER	15
3.1.2. MODULADOR ÓPTICO	16
3.1.2.1. MODULAÇÃO DIRETA	16
3.1.2.2. MODULAÇÃO EXTERNA	17
3.1.2.2.1. MODULADOR DE ELETROABSORÇÃO.....	17
3.1.2.2. 2. MODULADOR MACH-ZENDER.....	18
3.2. RECEPTOR	20
3.2.1. FOTODETECTOR.....	21
3.2.2. RECEPTOR POR DETECÇÃO DIRETA.....	24
3.2.3. RECEPTOR POR DETECÇÃO COERENTE.....	26
3.3.FIBRA ÓPTICA.....	27
4. CLASSIFICAÇÃO DOS FORMATOS DE MODULAÇÃO.....	29
5. FORMATOS AVANÇADOS DE MODULAÇÃO.....	31
5.1. MODULAÇÃO MULTINÍVEL	32
5.2. MODULAÇÃO COM MEMÓRIA	32
5.2.1. MODULAÇÃO PSEUDO-MULTINÍVEL	33
5.2.2. CODIFICAÇÕES CORRELATIVA E RESTRITIVA	33
6. FORMATOS DE MODULAÇÃO DE INTENSIDADE.....	35
6.1. CHAVEAMENTO ON-OFF SEM RETORNO AO ZERO (NRZ-OOK - <i>NONRETURN-TO-ZERO ON-OFF KEYING</i>)	35
6.2. CHAVEAMENTO ON-OFF COM RETORNO AO ZERO (RZ-OOK - <i>RETURN-TO-ZERO ON-OFF KEYING</i>).....	35

6.3. RETORNO AO ZERO COM PORTADORA SUPRIMIDA (<i>CSRZ – CARRIER SUPPRESSED RETURN-TO-ZERO</i>).....	39
6.4. RETORNO AO ZERO COM <i>CHIRP</i> (<i>CRZ – CHIRPED RETURN-TO-ZERO</i>), RETORNO AO ZERO COM <i>CHIRP</i> ALTERNADO (<i>ACRZ – ALTERNATE-CHIRP RETURN-TO-ZERO</i>).....	41
6.5. NÃO RETORNO AO ZERO COM <i>CHIRP</i> (<i>CNRZ – CHIRPED NONRETURN-TO-ZERO</i>).....	43
6.6. DUOBINÁRIO (DB, PSBT).....	44
6.7. INVERSÃO ALTERNADA DE SINAL (<i>AMI – ALTERNATE-MARK INVERSION</i>) ..	47
7. FORMATOS DE MODULAÇÃO DE FASE DIFERENCIAL.....	51
7.1. MODULAÇÃO POR DESLOCAMENTO DIFERENCIAL DE FASE (<i>DPSK – DIFFERENTIAL PHASE SHIFT KEYING</i>).....	51
7.2. MODULAÇÃO POR DESLOCAMENTO DIFERENCIAL DE FASE EM QUADRATURA (<i>DQPSK – DIFFERENTIAL QUADRATURE PHASE SHIFT KEYING</i>)	55
8. SIMULAÇÕES COM O <i>OPTISYSTEM</i>	59
9. CONCLUSÃO	65
REFERÊNCIAS BIBLIOGRÁFICAS.....	67

Lista de figuras

Figura 1 - Alcance e taxas de transmissão das diversas tecnologias de comunicação. [1]	1
Figura 2 - Fontes de perda nas fibras ópticas. [3]	3
Figura 3 - Dispersão na fibra monomodo e na fibra com dispersão deslocada. [4]	7
Figura 4 – Diagrama de blocos de um transmissor óptico. [3].....	12
Figura 5 – LED de emissão de superfície e LED de emissão lateral. [4].....	15
Figura 6 - Modulação direta.	16
Figura 7 - Função de transferência do EAM. [2]	17
Figura 8 - Modulador Mach-Zehnder.....	18
Figura 9 - Função de transferência do MZM. [2].....	19
Figura 10 - Diagrama de blocos de um receptor óptico. [3].....	20
Figura 11 - Figuras de mérito de diferentes fotodiodos. [6].....	22
Figura 12 - Coeficiente de absorção óptica em função do comprimento de onda. [6]	22
Figura 13 - Fotodetector de avalanche (APD). [6]	23
Figura 14 - Diagrama de um receptor por detecção direta. [3]	25
Figura 15 - Diagrama de olho ideal (esquerda) e degradado (direita) para o formato NRZ. [3].	26
Figura 16 - Esquema básico de detecção coerente.	27
Figura 17 - Diferentes perfis de índice de refração para as fibras ópticas. [3].....	28
Figura 18 - Classificação dos formatos de modulação de intensidade e de fase mais discutidos atualmente. [2].....	30
Figura 19 - Espectro óptico e diagrama de olho de um sinal NRZ. [2].....	35

Figura 20 - Sequência de bits ópticos (010110) transmitida nos formatos RZ e NRZ. [4]	36
Figura 21 - Geração de pulsos RZ, a partir de pulsos NRZ, utilizando um MZM como modelador de pulso. [2].....	36
Figura 22 - Formas de controle senoidal de um MZM atuando como modelador de pulso para sinais RZ com ciclo de trabalho de 33% (sólido) e 50% (tracejado). Os pontos de polarização são indicados por círculos abertos. [1]	38
Figura 23 - Espectro óptico e diagrama de olho de um sinal RZ com ciclo de trabalho de 50%. [2]	38
Figura 24 - Espectro óptico e diagrama de olho de um sinal RZ com ciclo de trabalho de 33%. [2]	39
Figura 25 - Controle senoidal de MZM atuando como modelador de pulso para sinais CSRZ com ciclo de trabalho de 67%. As curvas de transmissão sólidas e tracejadas referem-se à potência e ao campo óptico respectivamente. O ponto de polarização é indicado por um círculo aberto. [1]	40
Figura 26 - Espectro óptico e diagrama de olho de um sinal CSRZ. [2].....	40
Figura 27 - Espectro óptico e diagrama de olho de um sinal duobinário. [2]	44
Figura 28 - Visualização das sequências original e pré-codificada de dados e do sinal duobinário de 3 níveis que comanda o MZM entre seus máximos de transmissão. [1]	45
Figura 29 - Circuitos geradores de sinais DB: com filtro de atraso e adição (a) e com filtro passa-baixas (b). B: largura de banda do filtro passa-baixas. R: taxa de dados. [2]	46
Figura 30 - Espectro óptico e diagrama de olho de um sinal RZ-AMI com ciclo de trabalho de 33%. [2]	47
Figura 31- (a) Uma dos tipos de transmissor AMI. (b) Sinal no formato NRZ (linha cheia) interferindo com sua replica atrasada (linha tracejada) em um DI para produzir pulsos de fase alternada. [2]	49
Figura 32 - Constelações dos sinais OOK (superior) e DPSK (inferior). [7].....	52
Figura 33 - Duas configurações típicas de um transmissor RZ-DPSK: (a) utilizando um modulador de fase e (b) utilizando um MZM. [2].....	53
Figura 34 - Diagramas espectral e de olho: (a) NRZ-DPSK e (b) RZ-DPSK 33%. [2]	54

Figura 35 - Receptor DPSK balanceado. [2].....	55
Figura 36 - Diagrama de um transmissor DQPSK utilizando dois MZM em paralelo. [2].....	56
Figura 37 - Diagrama spectral do NRZ-DQPSK e do RZ-DQPSK 33%. [2]	56
Figura 38 - Configuração de um receptor DQPSK. [2].....	57
Figura 39 - Transmissor NRZ-DPSK.....	59
Figura 40 - Receptor NRZ-DPSK.	60
Figura 41 - Diagramas de olho de dois sinais DPSK: (a) um a 10 Gb/s e (b) outro a 40 Gb/s. ..	60
Figura 42 - Diagramas espectrais de dois sinais DPSK: (a) um a 10 Gb/s e (b) outro a 40 Gb/s.	61
Figura 43 – Sequência de bits enviados a 10 Gb/s (a) e sinal óptico produzido pelo MZM (b).	61
Figura 44 - Sequência de bits enviados a 40 Gb/s (a) e sinal óptico produzido pelo MZM (b)..	62
Figura 45 - Formas de onda de entrada (a) e saída (b) do sinal DPSK a 10 Gb/s.	62
Figura 46 - Formas de onda de entrada (a) e saída (b) do sinal DPSK a 10 Gb/s.	63

Lista de tabelas

Tabela 1 - Sequência de bits modulada em diferentes formatos multiníveis.	34
Tabela 2 - Sequência de bits, sequência de bits invertidos, dados pré-codificados e sinal duobinário transmitido.	46
Tabela 3 - Sequência de bits, dados pré-codificados e sinal AMI transmitido.	48

Lista de acrônimos

ACRZ: *alternate-chirp return-to-zero* (retorno ao zero com *chirp* alternado)

AMI: *alternate-mark inversion* (inversão alternada de sinal)

APol: *alternate polarization* (polarização alternada)

ASE: *amplified spontaneous emission* (emissão espontânea amplificada)

ASK: *amplitude shift keying* (chaveamento por mudança de amplitude)

BER: *bit error rate* (taxa de erro de bit)

CNRZ: *chirped nonreturn-to-zero* (sem retorno ao zero com *chirp*)

CRZ: *chirped return-to-zero* (retorno ao zero com *chirp*)

CSRZ: *carrier-suppressed return-to-zero* (retorno ao zero com portadora suprimida)

CW: *continuous wave* (onda contínua)

DCS: *duobinary-carrier-suppressed* (duobinário com portadora suprimida)

DI: *delay interferometer* (interferômetro de atraso)

DML: *directly modulated laser* (modulação direta de laser)

DPSK: *differential phase shift keying* (modulação por deslocamento diferencial de fase)

DQPSK: *differential quadrature phase shift keying* (modulação por deslocamento diferencial de fase em quadratura)

EAM: *electroabsorption modulator* (modulador de eletroabsorção)

EDFA: *erbium doped fiber amplifier* (amplificador a fibra dopada com érbio)

FEC: *forward error correction* (técnica de correção de erros)

FSK: *frequency shift keying* (chaveamento por mudança de frequência)

FWM: *four-wave mixing* (mistura de quatro ondas)

IFWM: *intrachannel four-wave mixing* (mistura de quatro ondas intracanal)

IXPM: *intrachannel cross-phase modulation* (modulação de fase cruzada intracanal)

LED: *light emitting diode* (diodo emissor de luz)

M-ASK: *multilevel amplitude shift keying* (chaveamento por mudança de amplitude multinível)

MZM: *Mach-Zehnder modulator* (modulador Mach-Zehnder)

NRZ: *nonreturn-to-zero* (sem retorno ao zero)

OADM: *optical add/drop multiplexer* (multiplexador óptico *add/drop*)

OOK: *on-off keying* (modulação liga-desliga)

PMD: *polarization mode dispersion* (dispersão por modo de polarização)

Pol-SK: *polarization shift keying* (chaveamento por mudança de polarização)

PSBT: *phase-shaped binary transmission* (transmissão binária com fase formatada)

PSK: *phase shift keying* (chaveamento por mudança de fase)

RZ: *return-to-zero* (com retorno ao zero)

SPM: *self-phase modulation* (automodulação de fase)

WDM: *wavelength division multiplexing* (multiplexação por divisão de comprimento de onda)

XPM: *cross-phase modulation* (modulação de fase cruzada)

1. Introdução

Nos últimos anos, devido principalmente ao crescimento de usuários e de empresas que utilizam a Internet e, conseqüentemente, ao maior tráfego de dados exigido, tem-se assistido ao aumento vertiginoso da demanda por largura de banda em sistemas de comunicação óptica, os quais, conforme é ilustrado na Figura 1, são os únicos capazes, comercial e tecnologicamente, de transmitir grandes volumes de dados a grandes distâncias. Para atender a essa crescente demanda, os sistemas de comunicação óptica têm experimentado um crescimento contínuo em sua capacidade de transmissão.

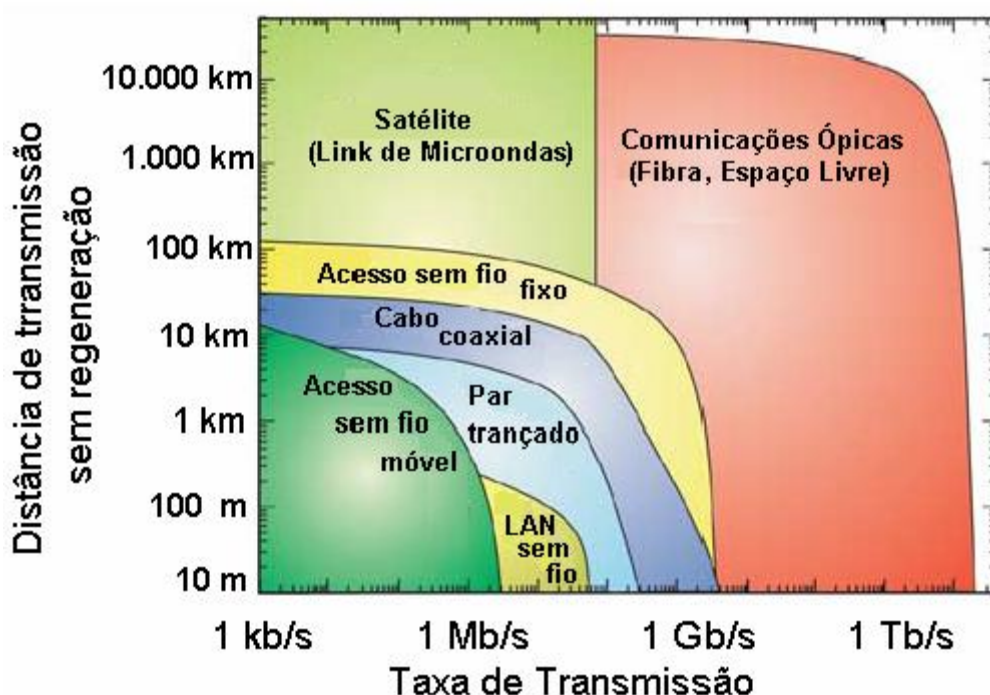


Figura 1 - Alcance e taxas de transmissão das diversas tecnologias de comunicação. [1]

O aumento da capacidade de transmissão de um sistema óptico é realizado basicamente sob as três formas a seguir: [2]

- instalação de novas fibras ópticas;
- aumento do número de canais do sistema e
- aumento da taxa de transmissão de cada canal.

A primeira forma é a mais simples e imediata para se aumentar a capacidade de um sistema, mas também é a mais dispendiosa, pois envolve obtenção de autorizações, pagamento por direito de passagem e gastos com obras civis. Já as duas últimas formas, por envolverem intensas pesquisas, são de mais longo prazo, mas acabam diminuindo o custo da informação transmitida.

O aumento do número de canais envolve tanto a implementação de sistemas WDM (*wavelength division multiplexing*) com grades de frequência cada vez menos espaçadas (o que resulta numa maior eficiência espectral, que é a razão entre a taxa de dados por canal e o espaçamento espectral) quanto a exploração de novas bandas de transmissão, como, por exemplo, as bandas S (1460 nm a 1530 nm) e L (1565 nm a 1625 nm), ou seja, outras bandas além da já comumente utilizada banda C (1530 nm a 1565 nm).

Além do aumento do número de canais, outra frente que tem sido exaustivamente explorada para aumentar a capacidade de um sistema de comunicação óptica é o aumento das taxas de transmissão de cada canal. Atualmente sistemas de 10 Gb/s são ainda bastante utilizados, mas já se está migrando rapidamente para sistemas de nova geração que operam a 40 Gb/s e além.

O aumento tanto do número de canais do sistema quanto da taxa de transmissão impõe alguns desafios tecnológicos complexos, e algumas das estratégias para superar esses desafios são as seguintes: [2]

- utilização de componentes ópticos de baixa perda, de dispositivos compensadores de dispersão e de elementos de chaveamento/roteamento óptico para minimizar o ruído gerado pela amplificação e para evitar conversões eletro-ópticas;
- uso de amplificadores ópticos de baixo ruído (amplificadores Raman distribuídos) para diminuir o ruído acumulado ao longo do sistema;
- instalação de fibras ópticas avançadas que introduzem menos distorções não lineares, permitindo assim potências maiores necessárias à transmissão de altas taxas;
- utilização de formatos avançados de modulação óptica para possibilitar mais resistência do sinal ao ruído e à filtragem óptica de banda estreita realizada através de diversos OADM (*optical add/drop multiplexer*) e também para possibilitar o envio de mais bits através de cada símbolo enviado.

Como se vê, uma das frentes a ser explorada para viabilizar os sistemas de nova geração é exatamente o uso dos formatos avançados de modulação digital, os quais, na verdade, já são empregados em sistemas de radiofrequência de comunicação. Por isso essa denominação “formatos avançados” deve ser usada no contexto do domínio óptico.

Até velocidades de 10 Gb/s, a modulação utilizada quase exclusivamente é a OOK (*on-off keying*), que consiste em representar os bits “1” e “0” respectivamente pela presença e ausência de intensidade de campo óptico. Esse é o esquema de modulação mais simples de ser implementado, mas ele não consegue fornecer as altas taxas de transmissão requeridas pelos sistemas da nova geração sem alta degradação do sinal durante a transmissão.

2. Principais fatores de degradação do sinal em fibras ópticas

Entre os vários desafios tecnológicos impostos pela migração para sistemas ópticos de 40 Gb/s está o de mitigar os efeitos degradantes ao sinal quando se transmite a taxas iguais ou superiores a essa. Entre esses efeitos que limitam a capacidade de transmissão do sistema e que serão abordados com mais detalhes logo a seguir, podemos destacar:

- perdas por absorção (extrínseca e intrínseca);
- perdas por espalhamento Rayleigh;
- perdas por curvaturas e microcurvaturas na fibra, em emendas e conectores;
- atenuação do sinal e o ruído introduzido para amplificá-lo;
- as dispersões do sinal conforme ele se propaga pela fibra;
- os efeitos não lineares da fibra no sinal;
- os efeitos da filtragem óptica de banda estreita e a diafonia entre canais WDM (*wavelength division multiplexing*).

2.1. Perdas por absorção

As perdas numa fibra óptica devem-se principalmente à absorção, ao espalhamento Rayleigh e às perdas por curvaturas. A Figura 2 mostra a contribuição de cada uma dessas parcelas para a perda total da fibra.

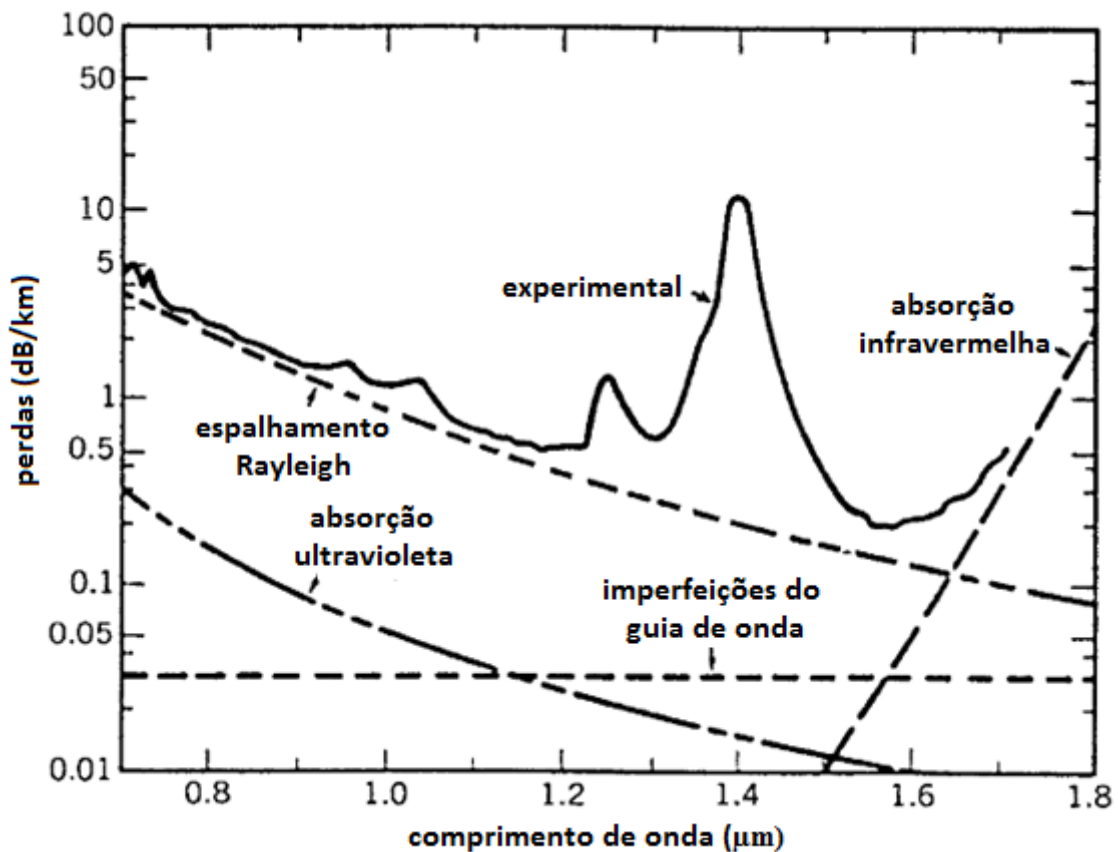


Figura 2 - Fontes de perda nas fibras ópticas. [3]

As absorções materiais representam a perda mínima fundamental atingível. Só podem ser superadas através da mudança do material da fibra. Essas absorções podem ser classificadas em extrínsecas e intrínsecas [3].

As extrínsecas são devidas principalmente a íons de impureza presentes na fibra óptica, tais como íons metálicos e íons hidroxilas de moléculas de vapor de água retidas na fibra.

As intrínsecas são provocadas pelo próprio material da fibra (SiO_2) e ocorrem quando um fóton do sinal óptico interage com um elétron na banda de valência e o excita para um nível de energia mais alto. São responsáveis por perdas menores que 0,03 dB/km na região que vai de 1300 nm a 1600 nm.

Nas fibras usadas atualmente, as perdas são influenciadas principalmente pelo íon hidroxila e pela absorção intrínseca do material da fibra na região do infravermelho próximo.

2.2. Perdas por espalhamento Rayleigh

Essas perdas decorrem de variações microscópicas na densidade do material, de flutuações de composição e de não homogeneidades estruturais (defeitos) durante a fabricação da fibra, as quais causam variações no índice de refração do material por distâncias pequenas comparadas ao comprimento de onda [3].

As perdas por espalhamento Rayleigh são inversamente proporcionais à quarta potência do comprimento de onda do sinal óptico usado. Por isso, para fibras de sílica, que utilizam comprimentos de onda na região de 1300 nm a 1600 nm, as perdas por espalhamento Rayleigh determinam um mínimo fundamental de perdas que é 0,16 dB/km. Entretanto essas perdas podem ser diminuídas abaixo de 0,01 dB/km em comprimentos de onda maiores que 3000 nm. Como as fibras de sílica não podem ser usadas nessa região já que as absorções intrínsecas passam a ser proibitivas além de 1600 nm, tem-se pesquisado outros materiais para compor as fibras [3].

2.3. Perdas por curvaturas e microcurvaturas na fibra, em emendas e conectores

As perdas por curvaturas ocorrem sempre que a fibra sofre uma curvatura de raio finito (efeitos perceptíveis para raio de curvatura maior que 5 mm) já que, quando isso ocorre, sempre se perde um pouco de energia para fora da fibra. Entretanto essas perdas são negligenciáveis na prática, pois dificilmente ocorrem curvas com raio excedendo a 5 mm.

Curvas microscópicas aleatórias (que geralmente são causadas quando as fibras são incorporadas nos cabos) provocam perdas da mesma forma que as curvaturas macroscópicas, mas normalmente essas perdas de potência acontecem para modos de mais alta ordem. Sendo

assim, fibras monomodo são menos suscetíveis a perdas por microcurvaturas. Todavia, mesmo em fibras monomodo, precauções são tomadas (como escolher o valor do parâmetro frequência normalizada V^1 o mais próximo possível do valor 2,405) no intuito de minimizar essas perdas, que podem chegar a 100 dB/km se nada for feito.

2.4. Atenuação e ruído de amplificação

As fibras ópticas monomodo usadas atualmente caracterizam-se por uma baixa atenuação (cerca de 0,2 dB/km) na região de 1550 nm, valor que é próximo ao limite fundamental de 0,16 dB/km (imposto pelo espalhamento Rayleigh) para fibras de sílica [3]. Entretanto mesmo essa pouca atenuação acaba reduzindo a intensidade do sinal, após ele percorrer longas distâncias, a um nível abaixo do limiar detectável pelos receptores ópticos, o que cria a necessidade de amplificação do sinal.

Essa amplificação é feita usualmente através de EDFA (*erbium doped fiber amplifier*), espaçados em torno de 80 km a 100 km, em sistemas terrestres, e em torno de 40 km a 60 km, em sistemas submarinos. Esse espaçamento é ditado pelas perdas da fibra: quanto menores elas forem, maior poderá ser a distância que um amplificador fica do outro, o que reduz custos. Também se tem pesquisado a amplificação óptica distribuída (usando amplificação Raman, por exemplo), a qual fornece ganho ao longo da própria fibra de transmissão.

O inconveniente quanto ao uso de amplificadores ópticos é que eles, em particular, introduzem ruído ASE (*amplified spontaneous emission*). Esse tipo de ruído, que apresenta densidade espectral de potência aproximadamente constante (assemelha-se a ruído branco [3]) e que em sistemas WDM (*wavelength division multiplexing*) com vários estágios de amplificadores cascadeados torna-se um sério problema devido aos efeitos cumulativos de cada estágio, degrada a relação sinal-ruído e adiciona flutuações ao sinal amplificado, as quais são convertidas em flutuações de corrente na fotodetecção. Por exemplo, na detecção quadrática (que é o tipo de detecção prevalente no domínio óptico), o receptor faz o batimento dos sinais que chegam a ele, gerando, além do sinal elétrico que representa o sinal óptico desejado, o batimento do ruído com ele mesmo e do ruído com o próprio sinal óptico [2].

Além das fontes de ruído citadas anteriormente, que estão sempre presentes, há uma fonte de ruído cujos efeitos são mais críticos em sistemas com maiores taxas de bit, que é o ruído térmico. O que ocorre é que, conforme a taxa de transmissão vai aumentando, a largura de banda dos dispositivos elétricos precisa também aumentar na mesma proporção, o que acaba

¹ A condição para que a fibra seja monomodo é que V seja menor que 2,405, e quanto mais próximo de 2,405, melhor, pois isso proporciona que o diâmetro do núcleo da fibra seja o maior possível ainda mantendo a condição monomodo, o que reduz a energia perdida para a casca.

introduzindo mais ruído térmico (que é proporcional à largura de banda desses dispositivos, conforme foi demonstrado por John B. Johnson e Harry Nyquist).

2.5. Dispersão cromática

A dispersão cromática expressa a dependência da velocidade de grupo de um pulso propagando-se por uma fibra, em relação à frequência óptica. Essa dependência faz que componentes espectrais diferentes se propaguem a velocidades diferentes, provocando o alargamento temporal do pulso após uma determinada distância. Esse alargamento, proporcional ao comprimento da fibra, pode provocar interferência inter-simbólica, o que, por sua vez, limita a taxa de dados máxima pela qual a informação pode ser transmitida em uma fibra óptica [2].

A dispersão cromática tem como principais causas:

- a abertura numérica;
- o diâmetro do núcleo, o perfil do índice de refração da fibra, o material e a geometria da fibra;
- o comprimento de onda utilizado e a largura espectral do pulso transmitido (que depende basicamente se o transmissor é a laser ou a LED - *light emitting diode*).

A dispersão cromática pode ser dividida em dois tipos: a dispersão material, devida à dependência do índice de refração com o comprimento de onda, e a dispersão de guia de onda, decorrente da estrutura física do guia de onda.

Numa fibra com perfil “degrau” de índice de refração, a dispersão não é um fator importante, mas em fibras com perfis de índice de refração mais complexos, a dispersão de guia de onda pode se tornar significativa. A dispersão material e a de guia de onda podem ter sinais opostos dependendo do comprimento de onda utilizado na transmissão e, em fibras monomodo de índice degrau, esses dois efeitos se cancelam em 1310 nm (ver Figura 3), produzindo uma fibra de dispersão zero. Isso viabiliza uma comunicação de alta largura de banda nesse comprimento de onda, porém a desvantagem é que a atenuação da fibra não é minimizada nessa frequência. As fibras de sílica exibem mínima atenuação em 1550 nm, o que, aliado ao fato de que os EDFA (*erbium doped fiber amplifier*) operam próximos a 1550 nm, torna óbvio que, se a propriedade de dispersão zero em 1310 nm pudesse ser mudada para coincidir com a janela de transmissão de 1550 nm, seria possível uma comunicação de alta largura de banda por longas distâncias. Foi com esse objetivo que a fibra de dispersão deslocada (*dispersion-shifted fiber*) foi desenvolvida [4]. Também se criou a fibra de dispersão plana (*dispersion-flattened fiber*), a qual possui uma dispersão cromática pequena por toda a janela de transmissão usada em sistemas ópticos (1300 nm a 1600 nm).

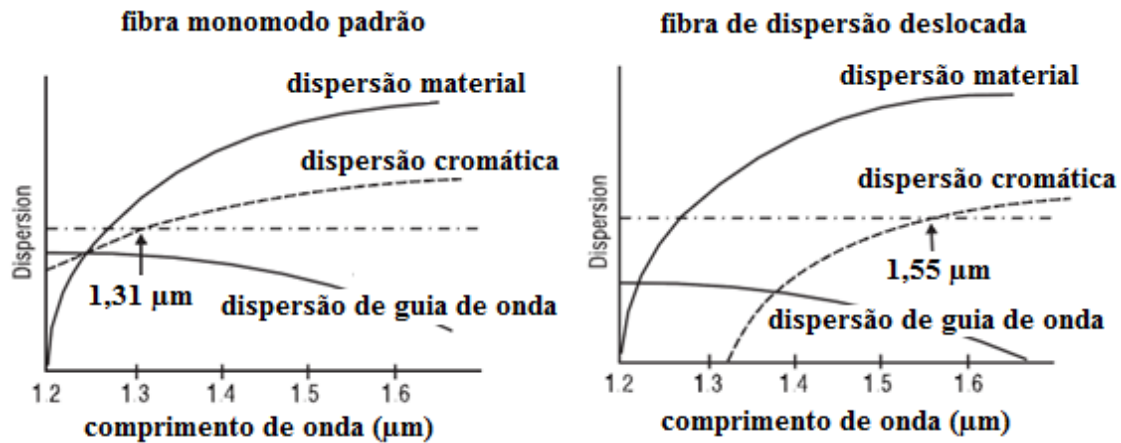


Figura 3 - Dispersão na fibra monomodo e na fibra com dispersão deslocada. [4]

2.6. Dispersão por modo de polarização

Outra fonte de alargamento do pulso óptico está relacionada com a birrefringência (diferença entre os índices de refração efetivos experimentados por cada modo) da fibra, a qual é devida a pequenos desvios da fibra em relação a uma geometria cilíndrica perfeita. Esses desvios são causados por imperfeições resultantes do processo de fabricação, por *stress* imposto pelas vibrações mecânicas durante o cabeamento ou manuseio, por emendas ou até por variações de temperatura [2].

Em uma fibra monomodo ideal, perfeita, devido à sua simetria circular, o modo de propagação transversal existe em dois modos de polarização degenerados, os quais têm idênticas propriedades de propagação. Entretanto numa fibra real, devido às deformações geométricas mencionados anteriormente, essas duas componentes ortogonalmente polarizadas experimentam índices de refração distintos e acabam se propagando com velocidades diferentes. Isso leva à dispersão dessas duas componentes e ao alargamento do pulso ao longo da fibra. Esse fenômeno é chamado PMD (*polarization mode dispersion*) e se torna um fator limitante em sistemas de longa distância que operam a altas taxas de bit [3].

2.7. Efeitos não lineares

A resposta à intensidade óptica de qualquer dielétrico se torna não linear para campos eletromagnéticos intensos. Apesar de a sílica da fibra não ser intrinsecamente um material altamente não linear, o forte confinamento do modo transversal do sinal óptico faz a intensidade óptica dentro da fibra exceder a MW/cm^2 . Em tais níveis de intensidade óptica, o índice de refração da fibra é afetado através do efeito Kerr (alteração do índice de refração com a intensidade óptica), o que provoca alterações na fase do sinal óptico, as quais, em conjunto com a dispersão provocada pela fibra, resultam em distorções na forma de onda dos sinais transmitidos. Por exemplo, dentre outros efeitos não lineares, podemos citar a automodulação de

fase (SPM - *self-phase modulation*), caracterizado pela alteração da fase óptica de um pulso devida à sua própria intensidade.

As não linearidades que ocorrem entre pulsos dentro de um mesmo canal WDM (*wavelength division multiplexing*) ou entre o sinal e o ruído ASE (*amplified spontaneous emission*) dentro de um mesmo canal WDM chamam-se não linearidades intracanaís, ao passo que as que ocorrem entre dois canais WDM são chamadas não linearidades intercanais. A importância de cada classe de não linearidades depende principalmente da taxa de transmissão por canal óptico. Como regra, pode-se dizer que os efeitos intercanais afetam sistemas WDM com taxas de transmissão por canal menores que 10 Gb/s e as não linearidades intracanaís afetam os sistemas com taxas mais altas que aquela. O impacto também depende da dispersão da fibra, pois fibras com baixa dispersão possuem efeitos intercanais mais fortes que as fibras de alta dispersão. Quanto às interações não lineares entre sinal e ruído, elas dependem significativamente do nível de ruído presente e se tornam mais fortes se a relação sinal-ruído é baixa durante a propagação.

Dentre os efeitos intercanais, podem-se citar a modulação cruzada de fase (XPM - *cross-phase modulation*), que é o deslocamento de fase de um sinal causado pelas flutuações de intensidade de outros canais WDM que compartilham a mesma fibra, e a mistura de quatro ondas (FWM - *four-wave mixing*), que é a interação não linear entre três frequências ópticas que dá origem a uma quarta frequência a qual pode se situar próxima à frequência de um canal WDM utilizado pelo sistema, o que prejudica a recepção.

Quanto às distorções intracanaís, causadas por pulsos muito curtos que, propagando-se numa fibra com dispersão local relativamente alta, alargam-se rapidamente, resultando em sobreposição de potência entre pulsos adjacentes, podem ser citadas a IXPM (*intrachannel cross-phase modulation*), que é a modulação de fase de um símbolo ocasionada por símbolos vizinhos em um mesmo canal, levando à geração de *jitter* temporal, e a IFWM (*intrachannel four-wave mixing*), que é a troca de potência entre símbolos diferentes, a qual ocorre quando três frequências regularmente espaçadas de três diferentes símbolos interagem para gerar energia numa quarta frequência sempre que todas as três frequências coexistem dentro do mesmo pequeno *slot* de tempo). Na extremidade do enlace, compensadores de dispersão reconstituem os pulsos à largura original, porém não podem desfazer os efeitos da IFWM, que se materializam como pulsos fantasmas em *slots* de tempo de bits “0”, ou como *ripple* na amplitude de bits “1” [2].

Como consequência dessas não linearidades, e em total contraste com sistemas de radiofrequência, o desempenho de um enlace de comunicação óptica exibe um máximo para um certo nível de potência, que representa um compromisso entre um mínimo de potência

abaixo do qual seriam necessários vários estágios de amplificação que introduziriam muito ruído ASE e um máximo acima do qual as não linearidades da fibra se tornariam limitantes [5].

2.8. Filtragem óptica de banda estreita e diafonia entre canais WDM

Nos últimos anos, os sistemas de comunicações ópticas têm sido impulsionados para eficiências espectrais de 0,8 b/s/Hz, os chamados sistemas WDM (*wavelength division multiplexing*) denso, ou DWDM. Os dois principais fatores de degradação oriundos de espaçamentos tão curtos de canais WDM são a diafonia (*crosstalk*) entre canais WDM e a filtragem óptica de banda estreita. A interferência entre canais depende da forma de onda da portadora óptica do canal interferente carregando trem de bits aleatoriamente diferentes, bem como da diferença de fase das portadoras dos canais WDM. O projeto cuidadoso de multiplexadores e demultiplexadores tem de assegurar a máxima supressão possível de diafonia entre canais WDM, ao mesmo tempo em que deve manter as distorções do sinal em decorrência da forte filtragem óptica ao mínimo [2].

Múltiplos OADM (*optical add/drop multiplexer*) em redes com alta eficiência espectral e roteadas opticamente representam uma concatenação de diversos filtros do tipo multiplexador-demultiplexador. Esta concatenação reduz a largura de banda da filtragem óptica como um todo e distorce o sinal. Desalinhamentos de comprimento de onda entre múltiplos filtros ou entre filtros e laser intensificam os efeitos de redução de largura de banda [2].

3. Componentes de um sistema de comunicação óptica

Os sistemas de comunicação óptica já passaram por pelo menos quatro grandes evoluções até chegarem a seu estágio atual. Na primeira geração, transmitiam em 850 nm usando fibras multimodo e laser de GaAlAs (gálio, alumínio e arsênio). A segunda geração (que passou a usar laser de InGaAsP – índio, gálio, arsênio e fósforo) mudou a janela de transmissão para 1310 nm, já que a fibra apresenta mínima dispersão nessa região, e passou a usar fibras monomodo, o que diminuiu sensivelmente a dispersão dos pulsos e possibilitou o aumento da taxa de bits. Na terceira geração, mudou-se a janela de transmissão novamente, desta vez para a região do infravermelho situada perto de 1550 nm, que é a região que apresenta a menor atenuação da fibra; porém para que isso não implicasse um aumento da dispersão, foram usadas fibras com dispersão deslocada (fibras com dispersão zero em 1550 nm). A partir da terceira geração, começou-se a pesquisar o uso de detecção coerente, que, entretanto, foi abandonado quando do descobrimento dos amplificadores ópticos (principalmente os EDFA - *erbium doped fiber amplifier*).

A atual geração (quarta geração) de sistemas de comunicação óptica faz uso de amplificação óptica para aumentar o espaçamento entre repetidores e de WDM para aumentar a taxa de bits. Nesses sistemas, as perdas na fibra são compensadas periodicamente usando EDFA espaçados entre 60 km a 80 km uns dos outros [3].

Um dos esforços atuais em relação a sistemas WDM é para aumentar a capacidade do sistema através da transmissão de cada vez mais canais. Um dos pontos a ser superados é que, com o crescente aumento da largura de banda do sistema, frequentemente não é possível amplificar todos os canais usando um único amplificador. Como consequência, novos tipos de amplificação têm sido pesquisados para abarcar toda a região espectral que se estende aproximadamente de 1300 nm a 1650 nm. A amplificação Raman, por exemplo, já consegue amplificar sinais das bandas S, C e L. Também há pesquisas sendo feitas para o desenvolvimento de uma nova fibra (*dry fiber*), denominada de “seca” por possuir baixíssima concentração de íons hidroxila e que possui a propriedade de ter uma atenuação baixa nos 350 nm de largura de banda citados anteriormente. Caso se obtenha êxito nessas empreitadas, os sistemas de quinta geração poderiam ter milhares de canais WDM, o que poderia resultar numa taxa total de bits de até 30 Tb/s [3].

Outra frente de pesquisa para implementar a quinta geração está relacionada ao aumento da taxa de bits de cada canal, e a esperança para se conseguir esse aumento está no uso de sólitons (pulsos ópticos que não têm sua forma alterada durante a propagação devido à interação balanceada de efeitos dispersivos lineares e não lineares) para transmitir informação nas redes ópticas.

Independente de a qual geração pertença um sistema óptico, ele é sempre constituído de três componentes essenciais, que serão detalhados a seguir: transmissor, canal de transmissão e receptor. O canal de transmissão é a fibra óptica, e os outros dois componentes são projetados para atender as especificidades desse tipo de canal.

3.1. Transmissor

O papel de um transmissor óptico é converter o sinal do domínio elétrico para o óptico para que ele possa ser inserido na fibra. A Figura 4 mostra o diagrama de blocos de um transmissor desse tipo, que consiste de uma fonte óptica, um modulador e um acoplador para o canal.

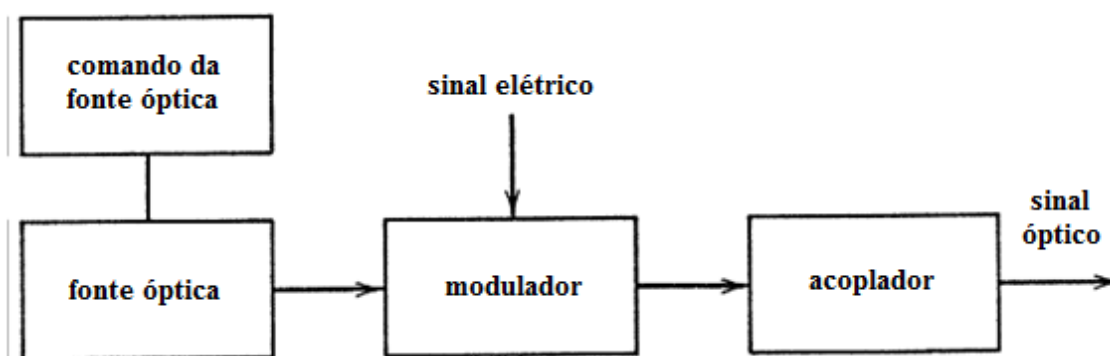


Figura 4 – Diagrama de blocos de um transmissor óptico. [3]

Laser semiconductor ou LED é usado como fonte óptica porque apresenta compatibilidade para com as fibras ópticas. Entre essas compatibilidades, podem-se citar: tamanho compacto, alta eficiência e confiabilidade, adequado comprimento de onda da luz emitida, pequena área de emissão, compatível com as dimensões do núcleo da fibra e possibilidade de modulação direta em frequências relativamente altas.

A emissão de luz pode ocorrer por dois processos fundamentais, conhecidos como emissão espontânea e emissão estimulada. No caso da emissão espontânea, os fótons de luz são emitidos em direções aleatórias, sem relação de fase entre si. A emissão estimulada, ao contrário, opera sob a condição denominada de inversão de população e é iniciada com um fóton já existente. O fóton emitido segue o fóton original não só em frequência, mas também em outras características, como a direção de propagação e a fase. Diferentemente dos LED, que emitem luz através de emissão espontânea, todos os lasers emitem luz através do processo de emissão estimulada.

O sinal óptico pode ser gerado de duas formas: modulando-se diretamente uma portadora óptica (que vem a ser a luz gerada pela fonte óptica) através da variação da corrente injetada na fonte óptica semicondutora ou usando um modulador externo. Embora o sinal

modulado externamente seja muito melhor, dispensa-se, para a diminuição de custos, o uso de um modulador externo quando a modulação direta é capaz de produzir um sinal de qualidade aceitável.

O acoplador tipicamente é uma microlente que foca o sinal óptico no plano de entrada da fibra com a máxima eficiência possível. A potência lançada é um importante parâmetro de projeto porque se pode aumentar o espaçamento entre amplificadores (ou regeneradores) aumentando a potência até certo ponto, a partir do qual os efeitos lineares começam a se sobressair [3].

Essa potência lançada é um tanto baixa para LED (menor que -10 dBm), mas lasers conseguem injetar em torno de 10 dBm de potência na fibra. Como LED também são limitados em suas capacidades de modulação, a maioria dos sistemas ópticos usa lasers como fonte óptica, nos quais a taxa de bits é limitada frequentemente mais pela eletrônica do que pelo laser em si. Quando projetados corretamente, transmissores ópticos conseguem operar a taxas de até 40 Gb/s [3].

3.1.1. Fontes ópticas

3.1.1.1. LED

Em sua forma mais simples, um LED é uma homojunção “pn” polarizada diretamente que emite luz por emissão espontânea (processo conhecido como eletroluminescência). Essa emissão de luz é provocada pela recombinação radiativa de pares elétron-lacuna na região de depleção do semiconductor. A luz emitida é incoerente, possui uma largura espectral relativamente grande (40 nm a 60 nm), assim como um espalhamento angular também alto [3].

Após gerada, parte dessa luz escapa do dispositivo e pode ser acoplada na fibra óptica. Verifica-se, entretanto, que apenas por volta de 1% da potência elétrica aplicada (10 mW) é convertida em potência luminosa de saída (100 μ W), ou seja, a eficiência global de LED é muito baixa.

Abaixo estão listadas algumas outras características dos LED:

- baixa potência de saída, se comparado à produzida pelos lasers;
- tempo de resposta de emissão lento, que é o atraso de tempo entre a aplicação de um pulso de corrente e o início da emissão de luz (este é o fator que limita a largura de banda de modulação direta de LED);
- eficiência quântica normalmente menor que a dos lasers (este parâmetro está relacionado com a fração de pares elétron-lacuna injetados que se recombina radiativamente).

A potência emitida depende de uma alta eficiência quântica, que é atingida quando a estrutura permite o confinamento de portadores de carga (utilizado para obter alto nível de recombinação radiativa na região ativa), e de que a emissão óptica estimulada ocorra dentro da região ativa da junção “pn”, que é onde a recombinação radiativa acontece.

Outros dois fatores que influenciam a potência (ou brilho) emitida por um LED são a eficiência de corrente e a eficiência externa ou de extração.

A eficiência de corrente é a fração da corrente total do diodo devida aos portadores minoritários desejados (elétrons injetados no lado “p” ou lacunas injetadas no lado “n”) que recombina antes de alcançar os contatos ôhmicos. Tal eficiência pode se aproximar de 100% desde que se utilizem dopagem assimétrica e heteroestruturas duplas (ou heterojunções), que recebem esse nome por utilizarem materiais de ligas diferentes em ambos os lados da região ativa, e desde se construam diodos largos (o que assegura que os portadores se recombinem antes de atingir os contatos).

A eficiência de extração (quanto de radiação de fato deixa o dispositivo) é o aspecto mais complicado para a maioria dos LED. Existem várias contribuições para este parâmetro, tais como reflexão interna total, absorção interna e bloqueios causados pelos contatos. Como os materiais semicondutores usados em LED possuem índice de refração alto (por volta de 3,5), o ângulo crítico de reflexão total é por volta de 16° , o que torna a eficiência de extração baixa (aproximadamente 2%). Para aumentá-la, busca-se realizar um confinamento óptico da luz emitida com o intuito de prevenir a absorção da radiação emitida pelo material que circunda a junção “pn” e também se procuram criar dispositivos finos com superfícies rugosas, o que aumenta as chances de a luz emitida atingir a superfície dentro do ângulo crítico.

Outro parâmetro usado para caracterizar um LED é sua responsividade, definida como a razão entre a potência óptica emitida e a corrente injetada no dispositivo. Um LED que opera na região de 1300 nm possui uma responsividade típica de 0,01 W/A, e esse valor cai para correntes acima de 80 mA porque, a partir de correntes dessa magnitude, começam a aumentar as recombinações não radiativas de pares elétron-lacuna.

Como já dito, a largura espectral é expressa em nanômetros usando-se a fórmula $\Delta f = -\frac{c}{\lambda^2} * \Delta \lambda$. Por ela, pode-se ver que LED a InGaAsP (que emitem em 1300 nm) têm uma largura 1,7 vez maior que a de outros a GaAs (que emitem em torno de 850 nm). Devido a essa largura espectral grande (aproximadamente 5% do comprimento de onda da luz emitida), o produto distância-taxa de bits é limitado consideravelmente pela dispersão na fibra quando LED são usados em sistemas de comunicação óptica.

Os LED podem ser classificados como de emissão de superfície (*surface-emitting*) ou de emissão lateral (*edge-emitting*), dependendo respectivamente se o LED emite luz de uma superfície que é paralela ao plano de junção ou de uma borda da região de junção, conforme mostra a Figura 5. Os de superfície emitem raios de luz com uma maior abertura angular que os de borda e, por isso, acoplam menos potência luminosa na fibra (1% contra 10% aproximadamente), e também possuem uma largura espectral maior que a dos LED de borda [3].

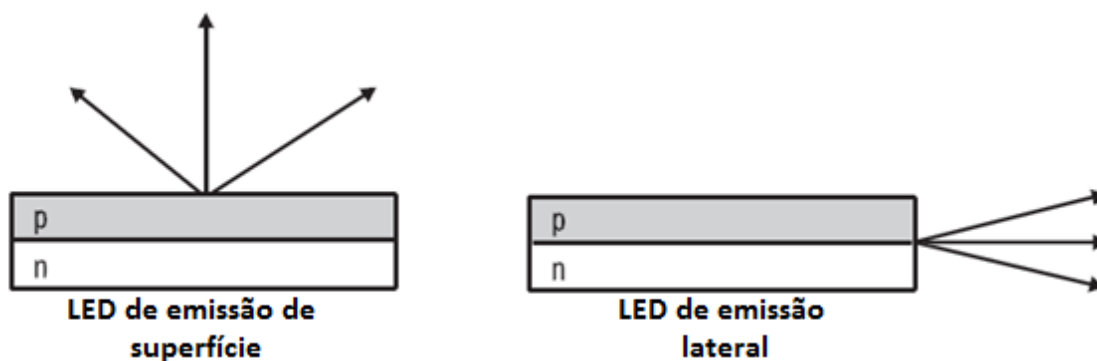


Figura 5 – LED de emissão de superfície e LED de emissão lateral. [4]

Devido a suas características, os LED são usados somente em redes de acesso local, ou seja, para comunicações a curta distância (alguns quilômetros), com pequena taxa de bits (10 Mb/s a 100 Mb/s) e que utilizam fibras multimodo.

3.1.1.2. Laser

Lasers semicondutores são as fontes de luz mais utilizadas em sistemas de comunicação óptica devido ao desempenho superior comparado ao dos LED. Eles emitem luz por emissão estimulada e são essencialmente um amplificador óptico inserido em uma cavidade refletiva que passa por um processo conhecido como bombeio para obter o ganho óptico. O tipo de laser mais comum em comunicações é o semiconductor, que usa semicondutores como meio de ganho. Devido às diferenças fundamentais entre emissão estimulada e espontânea, os lasers são capazes de emitir potências mais altas, emitem luz coerente e também possuem um espalhamento angular do feixe óptico de saída menor que o dos LED (já que somente as ondas que se propagam ao longo da direção da cavidade ou muito próximas a ela podem ser sustentadas em seu interior), o que permite uma alta eficiência (em torno de 50%) no acoplamento da luz em fibras monomodo. A largura espectral menor dos lasers possibilita operações em velocidades mais altas (aproximadamente 10 Gb/s), já que a dispersão se torna menor para esse tipo de fonte. Além disso, lasers podem ser modulados diretamente em frequências altas (até 25 GHz), pois seu tempo de chaveamento é menor que o dos LED.

Devido a suas características, os lasers são usados em sistemas de médias e longas distâncias que utilizam fibras monomodo e que operam a altas taxas de transmissão. O ponto negativo dos lasers é que eles são inerentemente não lineares, o que torna transmissões analógicas mais difíceis. Eles também são muito sensíveis a flutuações na temperatura e na corrente injetada, que causam alteração do comprimento de onda produzido. Em aplicações WDM (*wavelength division multiplexing*), a estabilidade das fontes ópticas é fundamental, o que exige circuitos mais complexos e mecanismos de realimentação para detectar e corrigir variações no comprimento de onda.

3.1.2. Modulador óptico

Um dos primeiros passos no projeto de um sistema de comunicação óptica é decidir como o sinal elétrico será convertido em uma sequência de bits ópticos. Normalmente a saída de uma fonte óptica é modulada aplicando-se o sinal elétrico ou diretamente na fonte óptica (modulação direta) ou a um modulador externo (modulação externa).

3.1.2.1. Modulação direta

A modulação direta de lasers (DML – *directly modulated laser*) é a maneira mais fácil de imprimir informação sobre uma portadora óptica. A informação é modulada sobre a corrente de controle do laser, resultando num formato de modulação em intensidade binária.

Conforme pode ser visto na Figura 6, os lasers são dispositivos polarizados por uma corrente I_b , e qualquer variação da corrente elétrica é correspondida por uma variação de sua potência óptica de saída.

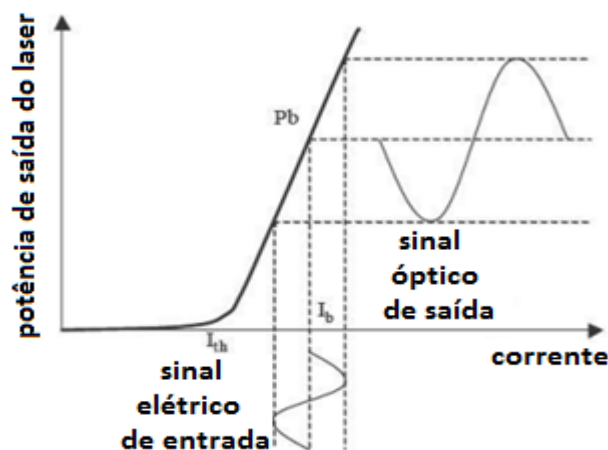


Figura 6 - Modulação direta.

A principal desvantagem de lasers modulados diretamente para transmissão de altas taxas de bits em aplicações que não sejam de curtas distâncias é seu *chirp*, isto é, a modulação de fase residual acompanhando a modulação em intensidade desejada. O *chirp* do laser alarga o

espectro óptico, o que dificulta a utilização de modulação direta em sistemas DWDM (*dense wavelength division multiplexing*), nos quais os canais têm seu espaçamento reduzido. Além disso, o *chirp* pode conduzir a distorções do sinal intensificadas pela interação com a dispersão cromática da fibra [5].

Atualmente, os lasers modulados diretamente são largamente empregados na modulação de taxas de 2,5 Gb/s e também estão disponíveis para algumas aplicações em 10 Gb/s [2].

3.1.2.2. Modulação externa

3.1.2.2.1. Modulador de eletroabsorção

Moduladores de eletroabsorção (EAM – *electroabsorption modulator*) são estruturas de semicondutor “pin” cujo limiar de absorção pode ser modulado pela aplicação de uma tensão externa, dessa forma alterando propriedades de absorção do dispositivo. Moduladores de absorção apresentam tensões de comando (ou *driving*) relativamente baixas (tipicamente 2 V). Atualmente existem EAM para modulação de taxas de 40 Gb/s e há demonstrações em pesquisas para taxas de até 80 Gb/s. Entretanto, assim como os lasers modulados diretamente, eles produzem um pouco de *chirp* residual. Eles possuem características de absorção que dependem do comprimento de onda, suas taxas de extinção (razão entre a máxima e a mínima potência do pulso modulado) não ultrapassam 10 dB e suas perdas de inserção atingem 10 dB [2].

Apesar dessas desvantagens, os EAM possuem uma vantagem, que é o fato de eles serem fabricados com o mesmo material semicondutor que é usado para a fabricação do laser, o que permite a compactação em circuito integrado dos dois dispositivos. Essa integração evita a alta perda na interface da fibra com o circuito integrado e resulta em transmissores compactos [1]. A Figura 7 mostra a curva de transmissão de potência característica de um EAM como função da tensão de controle.



Figura 7 - Função de transferência do EAM. [2]

3.1.2.2. 2. Modulador Mach-Zehnder

O modulador Mach-Zehnder (MZM – *Mach-Zehnder modulator*) é um modulador externo baseado no interferômetro Mach-Zehnder. Esta categoria de modulador faz uso de material eletro-óptico, cujo índice de refração pode ser alterado pela aplicação de uma tensão externa, e de um interferômetro Mach-Zehnder para, utilizando o princípio de interferências construtivas e destrutivas entre os sinais que se propagam por cada um de seus braços, produzir modulação em intensidade [2].

A estrutura desse modulador é mostrada na Figura 8. O sinal óptico que nele chega é dividido por um acoplador de entrada e percorre dois guias de onda (os braços do interferômetro). Na ausência de tensão externa, os campos nos dois braços experimentam deslocamentos de fase idênticos e interferem construtivamente na saída de outro acoplador. O deslocamento de fase adicional introduzido em um dos braços, através de mudanças no índice de refração induzidas por tensão, reduz a natureza construtiva da interferência, reduzindo assim a intensidade transmitida. Quando a diferença de fase entre os dois braços é igual a π , nenhuma luz é transmitida devido à interferência totalmente destrutiva entre os dois sinais. Dessa forma, a sequência de bits elétricos aplicada ao modulador produz uma réplica óptica dessa sequência de bits.

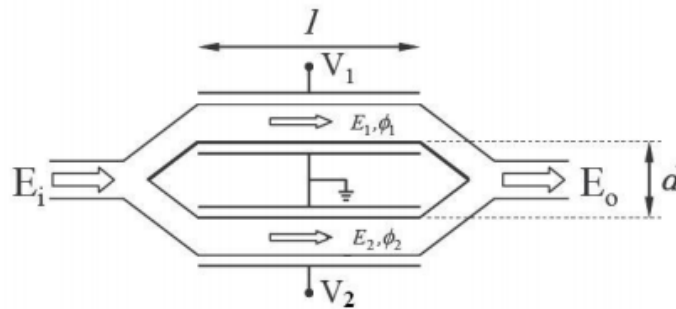


Figura 8 - Modulador Mach-Zehnder.

A função de transferência de um MZM para o campo de um sinal óptico é:

$$T_E(V_1, V_2) = \frac{1}{2} * (e^{i\phi(V_1)} + e^{i\phi(V_2)+i\psi}) = e^{\frac{i[\phi(V_1)+\phi(V_2)+\psi]}{2}} * \cos\left(\frac{\phi(V_1) - \phi(V_2) - \psi}{2}\right) \quad (1)$$

$\phi(V_1)$ e $\phi(V_2)$ são as fases adicionadas ao sinal óptico em cada um dos braços do MZM e ψ é uma mudança de fase adicional, constante no tempo, em um dos braços do modulador, conhecida como polarização do modulador [1]. Analisando-se a fórmula acima, conclui-se que a intensidade do campo óptico modulado é determinada pela diferença dos dois sinais de comando do MZM, ao passo que sua fase é controlada pela soma desses dois sinais.

Se a modulação de fase depende linearmente da tensão de controle ($\phi = kV$), o que é verdadeiro para a maioria dos materiais usados em MZM, a função de transferência de potência $T_P(V_1, V_2)$ depende somente da diferença de tensão ΔV aplicada entre os dois braços do modulador e é expressa por (V_{bias} é a tensão de polarização do MZM):

$$T_P(V_1, V_2) = |T_E(V_1, V_2)|^2 = \cos^2 \left(\frac{k\Delta V}{2} + \frac{kV_{bias}}{2} \right) \quad (2)$$

A função de transferência de potência do MZM é mostrada na Figura 9. A tensão de modulação requerida para mudar a fase de um braço do MZM de π , dessa forma chaveando o MZM do máximo ao mínimo da transmissão, é chamada de tensão de chaveamento V_π [1].

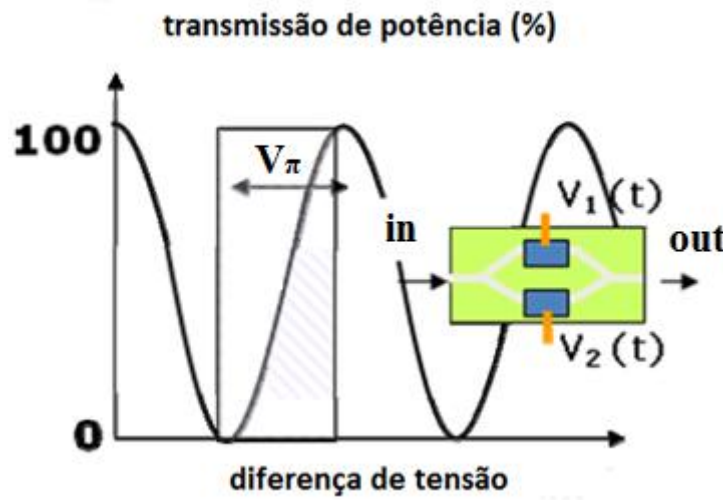


Figura 9 - Função de transferência do MZM. [2]

Para uma dada diferença de tensão ΔV , relacionada à intensidade da modulação desejada, há um grau de liberdade na escolha das tensões aplicadas em cada braço do MZM. Devido a isso, uma importante característica que pode ser explorada no MZM é a modulação de fase (*chirp*) que pode ser impressa ao sinal modulado. Conforme pode ser percebido através da análise da função de transferência do campo elétrico do sinal óptico no MZM, o termo de fase é função de $V_1(t) + V_2(t)$. Para modulação ou modelamento de pulso (*pulse carving*) livres de *chirp*, o que muitas vezes é desejado, os dois braços do modulador são controlados pela mesma amplitude de tensão, porém com sinais opostos ($V_1(t) = -V_2(t)$). Dessa forma, o termo de fase torna-se nulo. Esta condição é conhecida como polarização balanceada ou operação *push-pull*.

MZM são geralmente fabricados com niobato de lítio (LiNbO_3) e apresentam características de modulação pouco dependentes do comprimento de onda usado na transmissão, possuem também excelente taxa de extinção (por volta de 20 dB) e menor perda de inserção (aproximadamente 5 dB) que EAM (*electroabsorption modulator*) [1].

Devido ao seu bom desempenho em modulação e à possibilidade de modular de forma independente a intensidade e a fase do campo elétrico do sinal óptico, os MZM formam a base da maioria dos formatos avançados de modulação óptica [1].

3.2. Receptor

Um receptor óptico faz exatamente o contrário do transmissor, ou seja, converte o sinal óptico recebido de volta ao sinal elétrico original transmitido. Esse processo é realizado basicamente através do efeito fotoelétrico. A Figura 10 mostra o diagrama de blocos de um receptor óptico, que consiste de um acoplador, de um fotodetector e de um demodulador.

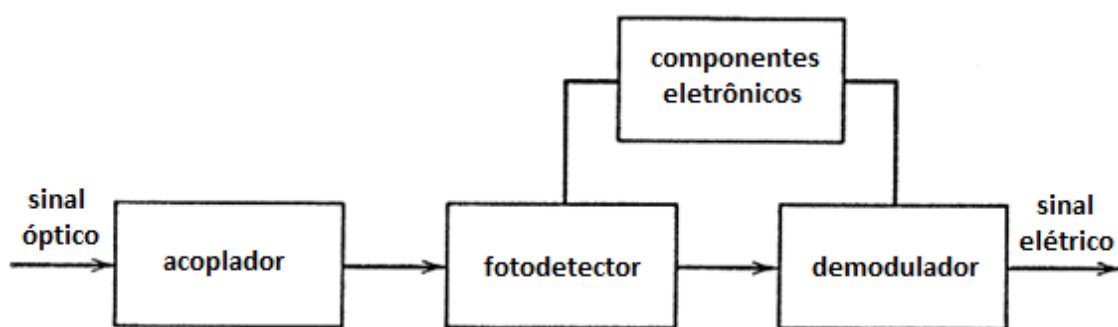


Figura 10 - Diagrama de blocos de um receptor óptico. [3]

O acoplador foca o sinal óptico recebido no fotodetector, que normalmente é um fotodiodo devido a sua compatibilidade com o resto do sistema. O fotodiodo converte a sequência de bits ópticos em um sinal elétrico variante no tempo. Já o projeto do demodulador depende do formato de modulação usado pelo sistema. O uso dos formatos FSK (*frequency shift keying*) e PSK (*phase shift keying*) geralmente exige técnicas de demodulação coerente (demodulação homódina ou heteródina). Entretanto a maioria dos sistemas atualmente emprega um tipo de modulação chamado de modulação de intensidade com detecção direta. Nesse caso, a demodulação é feita por um circuito de decisão que identifica bits como “1” ou “0” com base exclusiva na amplitude do sinal elétrico que sai do fotodetector. Por isso a exatidão do circuito de decisão depende da relação sinal-ruído do sinal elétrico que sai desse dispositivo.

Um parâmetro importante para qualquer receptor é sua sensibilidade, que é definida como a mínima potência óptica que deve incidir no receptor para que ele produza certa BER (*bit error rate*). Apesar de haver sistemas com BER tão baixas quanto 10^{-14} , os sistemas atuais operam, em geral, com BER de 10^{-9} , mas códigos corretores de erros acabam diminuindo ainda mais essa taxa. A sensibilidade do receptor depende da relação sinal-ruído, que, por sua vez, depende das várias fontes de ruído que corrompem o sinal ao longo da fibra. Mesmo para um receptor perfeito, um pouco de ruído sempre é introduzido durante o processo de fotodetecção. Esse ruído é denominado de ruído quântico ou ruído balístico (*shot*) por ter sua origem na própria natureza particular dos elétrons. O efeito cumulativo do ruído balístico junto com as

outras fontes de ruído (ruído térmico, ruído dos amplificadores, interferência inter-simbólica provocada pela dispersão cromática e por efeitos não lineares) que degradam a relação sinal-ruído no circuito de decisão determina a sensibilidade do receptor. Em geral a relação sinal-ruído depende também da taxa de bits usada já que a contribuição de algumas fontes de ruído (ruído térmico e balístico, por exemplo) aumenta com a largura de banda do sinal transmitido.

Há dois tipos de receptores ópticos, a saber: os de detecção direta e os de detecção coerente. No primeiro tipo, o processo consiste em detectar a presença ou não de luz em um intervalo de bit. São os chamados sistemas com modulação por intensidade e detecção direta (IM-DD). Já nos sistemas de detecção coerente, o sinal transmitido é modulado em amplitude, frequência ou fase e é detectado através de técnicas homódinas ou heteródinas.

3.2.1. Fotodetector

Alguns requisitos a que os fotodetectores devem obedecer são alta sensibilidade, resposta rápida, baixos ruído e custo e alta confiabilidade. Esses requisitos são mais adequadamente alcançados com fotodetectores construídos com semicondutores. Todos os fotodetectores podem ser caracterizados pelas figuras de mérito da eficiência quântica, responsividade e sensibilidade.

A eficiência quântica (η) determina quão eficientemente o dispositivo converte fótons em elétrons de condução na banda de condução. Para um dispositivo ideal, sem ganho óptico, a eficiência quântica se aproxima do valor unitário.

Por outro lado, do ponto de vista experimental, o parâmetro mais importante é a responsividade (\mathfrak{R}), que representa o quanto de fotocorrente (I) é gerado pela potência óptica total (P) que incide na área de detecção, e é dado por:

$$\mathfrak{R} = \frac{I}{P} = \frac{\eta q}{hf} = \frac{\eta \lambda}{1,24} \quad (3)$$

Na equação acima, q é a carga elétrica do elétron, h é a constante de Planck e f e λ (expresso em micrometros) são a frequência e o comprimento de onda da luz incidente respectivamente. Pode-se ver que a responsividade de um fotodetector aumenta com o comprimento de onda simplesmente porque mais fótons estão presentes para a mesma potência óptica. Tal dependência linear não continua para sempre porque há um momento em que a energia do fóton fica muito pequena para gerar elétrons e, então, a responsividade cai a zero. Em semicondutores, isso ocorre quando a energia do fóton fica menor que o *bandgap* do semicondutor. A responsividade e a eficiência quântica de diversos fotodiodos PIN podem ser vistas na Figura 11. Nessa figura, percebe-se que a eficiência quântica dos lasers a InGaAs, utilizados na terceira janela de transmissão (1550 nm), aproxima-se bastante de 100%.

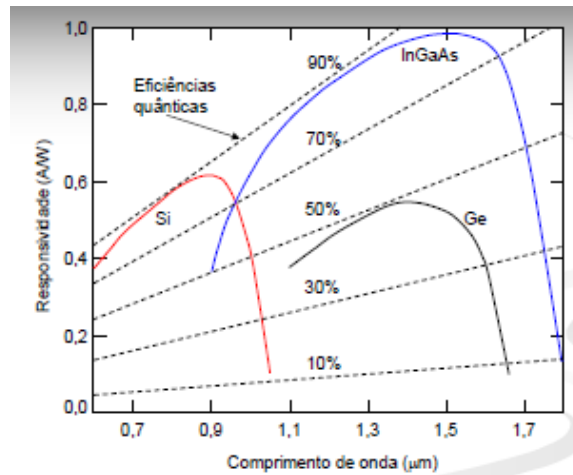


Figura 11 - Figuras de mérito de diferentes fotodiodos. [6]

Já a sensibilidade do fotodetector é descrita como a mínima potência incidente necessária para que aconteça uma relação sinal-ruído aceitável. Isto é, a partir de um determinado valor de potência óptica, pode-se distinguir adequadamente o sinal do ruído no detector. Vale mencionar que esse ruído provém basicamente de quatro processos aleatórios e independentes que são o ruído térmico, o balístico, o ruído de amplificação e a corrente de escuro, que é a corrente produzida pelo fotodetector mesmo na ausência de iluminação. Ainda para os fotodiodos APD, existe o ruído proveniente da flutuação estatística do ganho interno desses dispositivos decorrente do efeito avalanche.

Uma característica que ilustra a eficiência de semicondutores para o propósito de fotodetecção pode ser vista na Figura 12, que mostra a dependência do coeficiente de absorção óptica para diversos materiais normalmente usados para fabricar fotodetectores. Nessa figura se pode notar que, desde que o comprimento de onda da luz usada seja inferior ao comprimento de onda de corte do material, a partir do qual a absorção caia drasticamente, altos valores de coeficiente de absorção ($\sim 10^4 \text{ cm}^{-1}$) podem ser obtidos para a maioria dos semicondutores, o que permite a eficiência quântica aproximar-se de 100%.

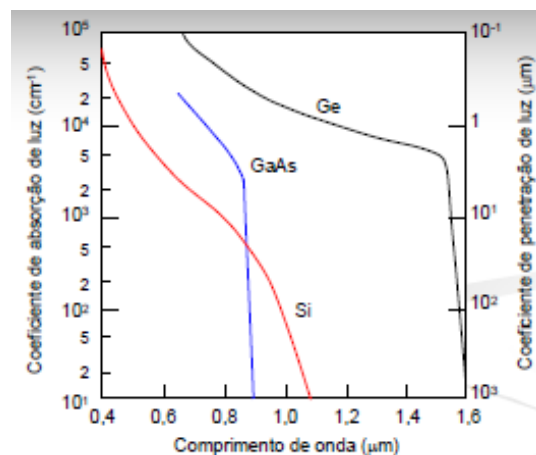


Figura 12 - Coeficiente de absorção óptica em função do comprimento de onda. [6]

Existem três tipos de fotodetectores usuais para recepção óptica: o MSM (metal-semicondutor-metal), o PIN e fotodetector de avalanche (APD).

O MSM é um dispositivo planar que consiste de dois contatos metálicos do tipo Schottky na forma de dedos interdigitais. Esse fotodetector permite o aumento da área efetiva de absorção sem que haja considerável aumento da capacitância intrínseca e da distância a ser percorrida pelos portadores gerados, o que possibilita um aumento na velocidade de resposta do dispositivo. As maiores vantagens do MSM em relação aos demais fotodetectores são:

- simplicidade de fabricação, pois apenas um tipo de semicondutor e de dopante são necessários;
- baixo ruído e baixa capacitância;
- compatibilidade com circuitos integrados.

O fotodiodo PIN apresenta uma camada espessa de material intrínseco entre as camadas “p” e “n”. O material intrínseco garante o alargamento da região de depleção por boa parte de sua extensão, o que constitui uma vantagem do PIN sobre fotodetectores sem a camada intrínseca porque esse alargamento diminui a capacitância da junção, já que a componente difusiva da fotocorrente (que é a parte “lenta” da fotocorrente) diminui consideravelmente, e isso melhora a largura de banda da resposta do diodo. Outra vantagem da colocação da camada intrínseca é que a responsividade do fotodiodo aumenta, entretanto sua largura não pode ser muito grande porque isso aumentaria o tempo de trânsito dos portadores gerados na região de depleção, o que acabaria diminuindo a largura de banda do dispositivo.

Já o fotodetector de avalanche é essencialmente um fotodiodo PIN no qual a tensão reversa aplicada é tão alta que o dispositivo opera no limiar de ruptura. Ele é constituído por uma camada “p” sobre uma camada intrínseca, a qual, por sua vez, está montada sobre uma camada “p⁺”. Ainda sobre a camada tipo “p” fica uma camada tipo “n⁺”, conforme pode ser visto na Figura 13.



Figura 13 - Fotodetector de avalanche (APD). [6]

Com essa operação no limiar de ruptura, elétrons e lacunas podem ionizar, por impacto, átomos do semiconductor e criar pares elétron-lacuna adicionais. Esse processo introduz ganho interno (o qual ocorre na junção “pn⁺”) e a fotocorrente se torna:

$$I = M * \mathfrak{R} * P \quad (4)$$

M é o ganho proporcionado pelo efeito avalanche e é definido como a razão entre a corrente de saída I_M e a fotocorrente primária não multiplicada I_P :

$$M = \frac{I_M}{I_P} \quad (5)$$

Com essa multiplicação interna da fotocorrente, o APD aumenta a sensibilidade e a responsividade do receptor, por isso é usado normalmente quando há pouca potência chegando ao receptor. Comparando-se as responsividades de um APD e um PIN, obtém-se:

$$\mathfrak{R}_{APD} = M * \mathfrak{R}_{PIN} \quad (6)$$

Devido a esse processo de ganho interno, os receptores que empregam um APD geralmente fornecem uma relação sinal-ruído maior para a mesma potência óptica incidente. Essa relação deveria melhorar por um fator de M^2 se o ruído no receptor não fosse afetado pelo mecanismo de ganho interno. Entretanto isso não acontece e a melhora na relação sinal-ruído é consideravelmente reduzida.

O APD é intrinsecamente ruidoso devido à natureza estatística do processo de geração de portadores. Isso implica que a corrente total fluindo no circuito externo sofre flutuações (ruído) em torno do valor médio, ou seja, o ganho M é um valor médio. Essa é a principal fonte de ruído para fotodetectores APD (o ruído térmico não tem a mesma importância que tem para os fotodetectores PIN).

Assim como ocorre nos fotodetectores PIN, os projetistas de APD também enfrentam um *trade-off* entre velocidade e sensibilidade, pois quanto maior for o ganho do dispositivo, menor será sua largura de banda.

3.2.2. Receptor por detecção direta

O projeto de um receptor depende do formato de modulação usado pelo transmissor e pode-se dizer que, atualmente, a maioria dos sistemas de transmissão óptica emprega modulação de intensidade, o que implica que o receptor fará uso de detecção direta. Esse tipo de receptor é capaz de identificar apenas variações na intensidade do sinal luminoso, não podendo, portanto, ser empregado em comunicações que utilizam modulação por fase ou frequência. A Figura 14 mostra o diagrama de blocos de um receptor desse tipo dividido em três partes.

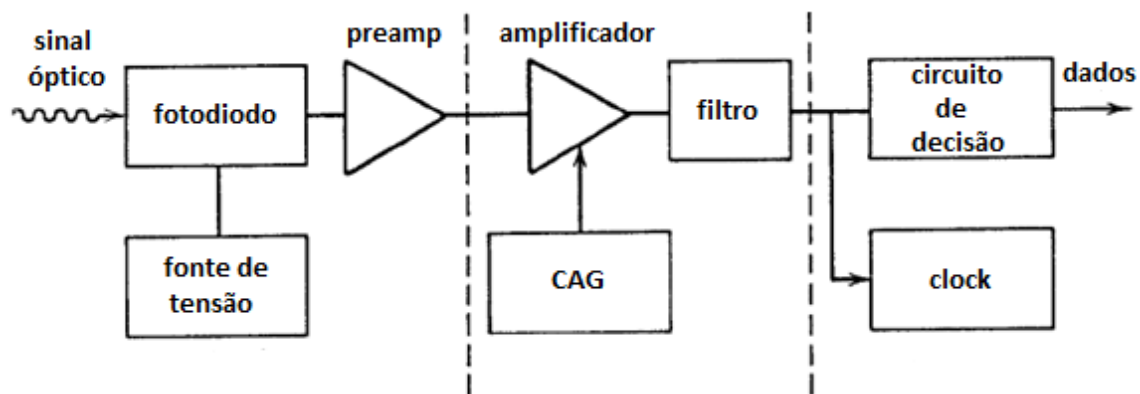


Figura 14 - Diagrama de um receptor por detecção direta. [3]

A primeira parte do receptor consiste de um fotodiodo e de um pré-amplificador (que pode ser um pré-amplificador óptico antes do fotodetector ou um pré-amplificador elétrico depois dele) para amplificar o sinal antes de ele ser processado. Um resistor de carga de alta resistência normalmente é usado para aumentar a tensão do sinal que alimentará o pré-amplificador, o que melhora a sensibilidade do receptor. Além disso, essa alta resistência reduz o ruído térmico, mas tem a desvantagem de diminuir a largura de banda do dispositivo. Por isso uma resistência alta não pode ser usada se a largura de banda for consideravelmente menor que a taxa de bits. Se necessário, um equalizador é usado para atenuar mais as componentes de baixa frequência que as de alta, o que aumenta a largura de banda, a qual também pode ser aumentada com a diminuição da resistência de carga, se a sensibilidade do receptor não for uma preocupação.

A segunda parte é constituída de um amplificador de alto ganho (amplificador principal) e de um filtro passa-baixas. O ganho é controlado automaticamente para limitar a tensão de saída média a um nível fixo independentemente da potência óptica incidente no receptor. Já o filtro formata o pulso recebido no intuito de reduzir o ruído sem introduzir muita interferência inter-simbólica (essa filtragem faz cada bit se espalhar para além de seu *slot* de tempo). Essa interferência é minimizada quando a transformada de Fourier da tensão de saída do filtro tem a forma do cosseno levantado. Levando isso em consideração e mais o formato do pulso dos bits transmitidos, os projetistas de receptores por detecção direta projetam o dispositivo para ter uma função de transferência que, sob condições ideais, produza a tensão de saída desejada.

A última parte do receptor, que é a seção de recuperação dos dados, consiste de um circuito de decisão e de um circuito de recuperação de *clock*. O propósito deste último é isolar a componente espectral de frequência igual à taxa de bits do sinal recebido. Essa componente fornece ao circuito de decisão informação a respeito da duração do *slot* de tempo de cada bit, ajudando, portanto, no processo de decisão. Porém, devido à natureza ruidosa dessa componente que alimenta o circuito de recuperação de *clock*, o instante de amostragem flutua de bit a bit

(*timing jitter*), o que degrada a relação sinal-ruído porque essas flutuações no instante de amostragem levam a flutuações adicionais no sinal. Isso pode ser entendido se se considerar que, quando o bit não é amostrado no centro de seu *slot*, o valor amostrado é reduzido de um valor que depende do *timing jitter*.

Já o circuito de decisão compara, em instantes de amostragem determinados pelo circuito de recuperação de *clock*, a saída do filtro passa-baixas com um nível-limite, decidindo então se o sinal corresponde ao bit “1” ou ao bit “0” conforme o nível detectado seja maior ou menor, respectivamente, que o nível-limite (o qual é calculado para otimizar a BER). Erros ocorrem quando, devido à presença de ruído, o nível do bit “1” fica menor que o nível-limite ou o nível do bit “0” fica maior que o nível-limite. O melhor instante de amostragem corresponde à situação em que a diferença entre os níveis do sinal do bit “1” e do bit “0” é máxima. Esse instante pode ser determinado através do diagrama de olho, formado pela superposição de duas sequências de bits, cada uma com três bits de duração. A Figura 15 mostra um diagrama de olho ideal e outro degradado por ruído e *jitter*, degradações essas que levam a um fechamento parcial do olho. O melhor instante de amostragem é o de máxima abertura do olho.

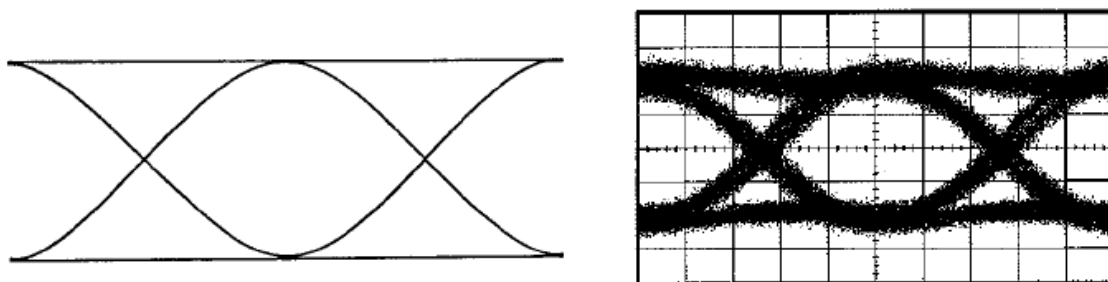


Figura 15 - Diagrama de olho ideal (esquerda) e degradado (direita) para o formato NRZ. [3]

Devido ao ruído inerente a todos os receptores (e, na prática, para os receptores por detecção direta com fotodiodos PIN, a principal fonte de ruído é o ruído térmico), sempre há uma probabilidade finita de um bit ser identificado incorretamente pelo circuito de decisão. Receptores digitais costumam ter uma probabilidade de erro muito baixa (tipicamente de 10^{-9}), e seu desempenho pode ser monitorado visualmente pelo diagrama de olho que eles fornecem, pois um fechamento do olho nesse diagrama é uma indicação de que o receptor não está operando bem.

3.2.3. Receptor por detecção coerente

Na recepção com detecção coerente, o sinal óptico recebido é misturado ao sinal óptico de um laser de pequena largura espectral que atua como um oscilador local, conforme é mostrado na Figura 16.

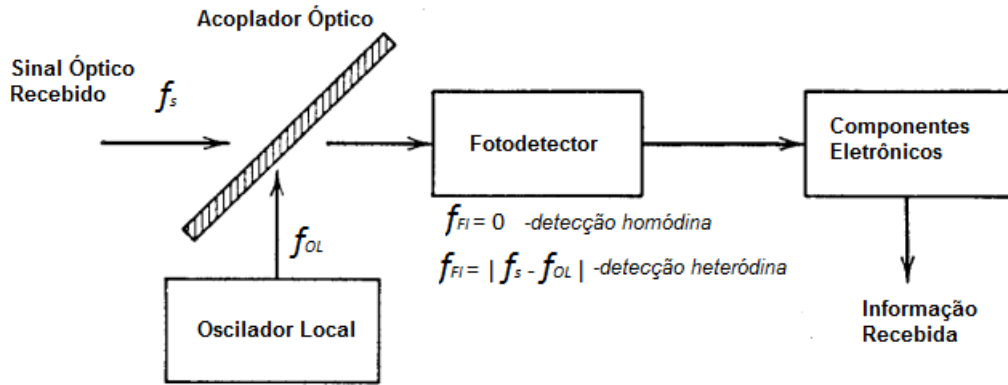


Figura 16 - Esquema básico de detecção coerente.

O batimento entre os dois campos produz um novo sinal:

$$P(t) = P_s + P_{OL} + 2\sqrt{P_s P_{OL}} \cos(\omega_{FI} t + \varphi_s - \varphi_{OL}) \quad (7)$$

$P_s = kA_s^2$; $P_{OL} = kA_{OL}^2$; $\omega_{FI} = \omega_0 - \omega_{OL}$; ω_0 , A_s e φ_s são respectivamente a frequência, a amplitude e a fase do sinal no instante da detecção; e ω_{OL} , A_{OL} e φ_{OL} são respectivamente a frequência, a amplitude e a fase do oscilador local no instante de detecção.

Esse novo sinal óptico passará pelo fotodetector, o qual gerará uma fotocorrente que possui informações sobre a amplitude, a frequência e a fase do campo do sinal óptico recebido. Convém lembrar que existem dois tipos de receptor coerente: o homódino, para o qual a frequência do oscilador local coincide com a frequência da portadora óptica, portanto não há translação para uma frequência intermediária ($\omega_{FI} = 0$), e o heteródino, para o qual a frequência do oscilador local é diferente da frequência da portadora óptica, e a frequência intermediária é da ordem da região de microondas (~ 1 GHz).

Uma das vantagens de se utilizar detecção coerente é que a sensibilidade do receptor pode ser melhorada teoricamente em torno de 20 dB em comparação com os sistemas de detecção direta. Outra grande vantagem é que essa técnica abre portas para o uso de diferentes tipos de modulação do sinal, além da modulação em amplitude.

Em sistemas de detecção coerente, a potência do oscilador local é alta, e o ruído térmico se torna muito menor do que o produto do batimento entre o sinal do oscilador local e o sinal recebido. Portanto seu efeito é consideravelmente reduzido com o uso de detecção coerente.

3.3.Fibra óptica

O papel da fibra óptica é transportar o sinal óptico do transmissor ao receptor com a mínima distorção possível. As fibras usadas atualmente têm perdas (ver seções 2.1, 2.2 e 2.3) de 0,2 dB/km, o que significa que, após percorrer 100 km, o sinal conserva somente 1% de sua

energia original. Por essa razão, as perdas na fibra continuam a ser um parâmetro importante no projeto de redes ópticas e determina o espaçamento entre amplificadores e regeneradores nesses sistemas. Outro fator que afeta o desempenho da fibra é a dispersão que ela apresenta (ver seções 2.5 e 2.6). Em fibras multimodo, o sinal se dispersa rapidamente (por volta de 10 ns/km), por isso esse tipo de fibra só é usado em lances curtos e de baixa taxa de bits. Já em fibras monomodo, o sinal se dispersa de 1 ps/(km.nm) aproximadamente. Essa dispersão é pequena o suficiente para ser aceitável para a maioria das aplicações e ainda pode ser diminuída controlando-se a largura espectral da fonte óptica [3].

As fibras podem ser fabricadas com diferentes perfis de índice de refração, como pode ser visto na Figura 17.

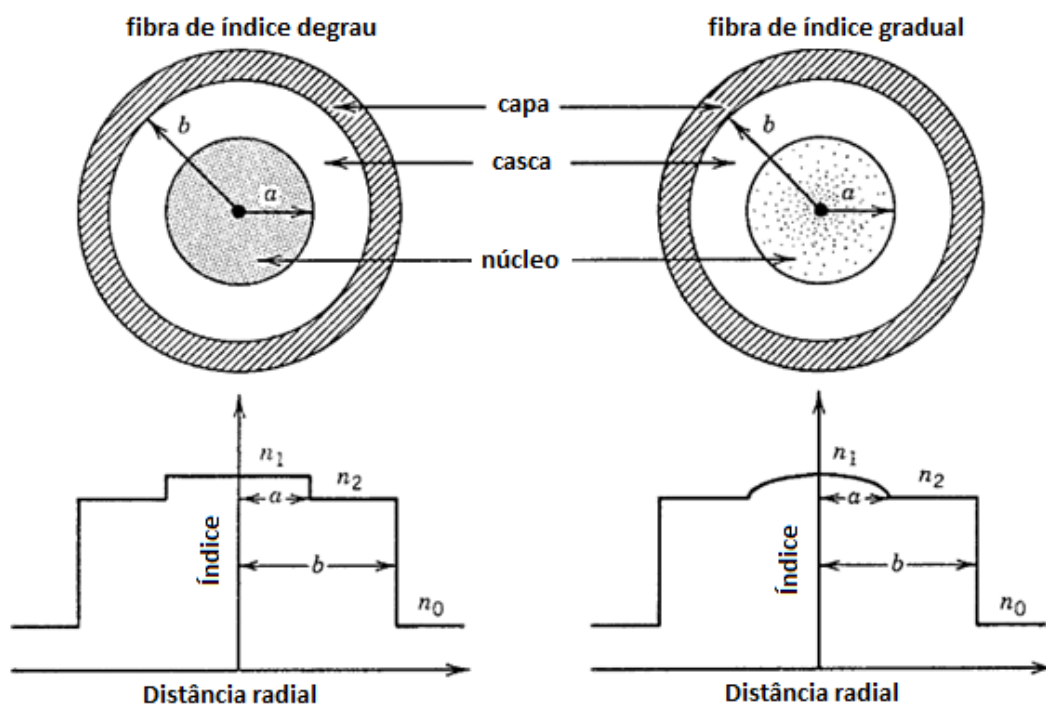


Figura 17 - Diferentes perfis de índice de refração para as fibras ópticas. [3]

Nas fibras do primeiro tipo, quanto menor for a mudança de índice de refração do núcleo para a casca, menor é a dispersão que um pulso óptico sofre ao se propagar por ela, mas torna-se mais difícil acoplar toda a potência gerada pela fonte óptica nesse caso. Já com as fibras de índice de refração variável ao longo do raio do núcleo, consegue-se reduzir a dispersão modal.

4. Classificação dos formatos de modulação

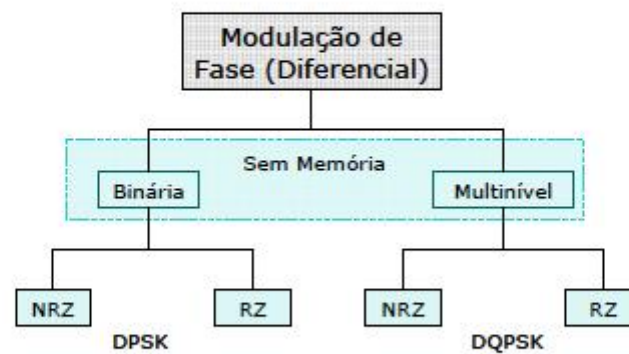
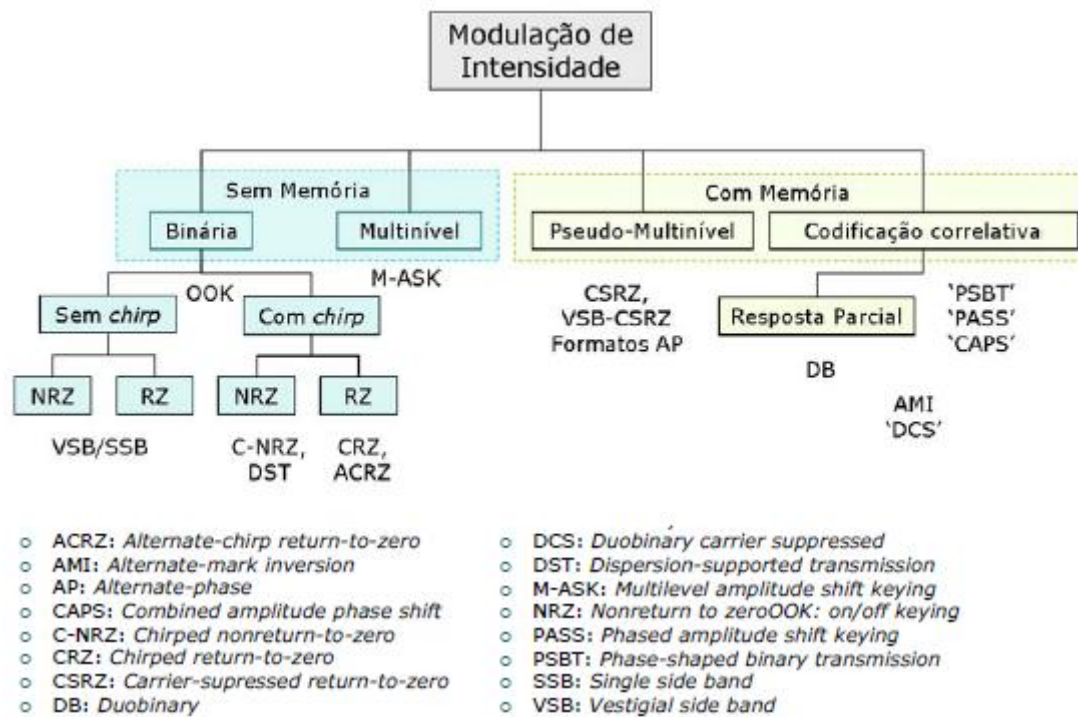
O campo óptico (que pode ser tanto o elétrico quanto o magnético) de um sinal de luz pode ser escrito como:

$$E(t) = A(t) * \cos[\omega(t) * t + \varphi(t)] \quad (8)$$

Há basicamente quatro características que podem ser exploradas para modular esse sinal: amplitude, fase e frequência, além da polarização da luz, resultando respectivamente nos formatos ASK (*amplitude shift keying*), PSK (*phase shift keying*), FSK (*frequency shift keying*) e Pol-SK (*polarization shift keying*). Dependendo de qual das quatro características físicas é usada para transmitir informações, o formato de modulação de dados será classificado como sendo de intensidade, fase, frequência ou polarização. Esta classificação não requer que um campo óptico modulado em fase tenha necessariamente uma envoltória constante nem que uma modulação em intensidade tenha fase constante. O fator determinante nessa classificação é a entidade física que é usada para transportar a informação de dados [2].

Enquanto os formatos de modulação que codificam informação na intensidade e na fase de uma portadora óptica já foram extensivamente estudados, pouca atenção ainda foi dada para os formatos que usam a polarização da luz para carregar dados, pois o uso de tais formatos requereria a correção de rotações aleatórias que ocorrem naturalmente na polarização da luz durante seu trajeto pela fibra e implicaria, portanto, uma maior complexidade do receptor, o que poderia ser justificável se a modulação de polarização oferecesse uma melhora significativa da sensibilidade do receptor em relação à modulação de intensidade, o que não ocorre na prática. Entretanto a modulação de polarização pode ser usada para melhorar as propriedades de propagação de um formato e também é usada em pesquisas, às vezes, para aumentar a eficiência espectral, seja transmitindo dois sinais diferentes no mesmo comprimento de onda mas em duas polarizações ortogonais (multiplexação em polarização), seja transmitindo canais WDM (*wavelength division multiplexing*) adjacentes em polarizações alternadas para reduzir a diafonia (*crosstalk*) ou interações não lineares entre eles (entrelaçamento por polarização – *polarization interleaving*).

Na Figura 18 estão esquematizados os mais importantes formatos de modulação discutidos atualmente em comunicações ópticas, que são os formatos de modulação de intensidade e os de fase. Dentro de cada grande grupo há subgrupos baseados nas características de modulação auxiliares, tais como: modulação com memória ou sem memória, modulação pulsada, com *chirp* ou sem *chirp*.



DPSK: Differential phase shift keying

DQPSK: Differential quadrature phase shift keying

Figura 18 - Classificação dos formatos de modulação de intensidade e de fase mais discutidos atualmente. [2]

5. Formatos avançados de modulação

Até alguns anos atrás, a maioria dos sistemas de comunicação óptica realizavam modulação binária da luz emitida por um laser e faziam a detecção dos símbolos transmitidos com base em um limiar fixo de intensidade. Ou seja, tais sistemas usavam modulação binária de intensidade no transmissor e realizavam fotodetecção quadrática no receptor, sem o emprego de técnicas coerentes de demodulação. Uma das razões de se ter preterido o uso de detecção coerente e o de formatos avançados de modulação foi o advento de amplificadores ópticos eficientes, o que possibilitou superar um dos principais limites que enfrentavam os pesquisadores da época, que eram os limites de alcance das redes ópticas imposto pela atenuação na fibra. Então, devido à invenção dos EDFA (*erbium doped fiber amplifier*), o uso do formato OOK (*on-off keying*) no transmissor e da detecção direta no receptor ganhou novo fôlego e deixaram-se de lado técnicas mais avançadas.

Hoje, entretanto, as pesquisas relacionadas à detecção coerente e aos formatos avançados de modulação (que, em comunicações ópticas, são todos os formatos que vão além da modulação OOK) têm ganhado renovado interesse devido, principalmente, a duas importantes vantagens que alguns desses formatos avançados apresentam em relação ao OOK e à sua demodulação binária insensível à fase, que são a possibilidade de aumentar as taxas de transmissão das redes ópticas e também a de prover mais resistência do sinal ao ruído.

Os formatos de modulação de intensidade avançados podem ser classificados em alguns subgrupos, tais como:

- modulação multinível;
- modulação pseudo-multinível;
- codificação correlativa.

Os formatos de modulação cujo alfabeto de símbolos consiste de mais de dois símbolos são geralmente denominados formatos multiníveis. Se a alocação dos símbolos é independente dos símbolos anteriormente enviados, a modulação é denominada sem memória (exemplo: modulação multinível). Se, ao invés disso, os símbolos enviados obedecem a determinadas regras para a sucessão de diferentes bits, o formato é chamado de modulação com memória (exemplos: modulação pseudo-multinível e codificação correlativa). Entre esses formatos de modulação que possuem mais de dois símbolos no alfabeto de símbolos, a codificação correlativa e a modulação pseudo-multinível têm recebido ultimamente mais atenção que a modulação multinível.

5.1. Modulação multinível

Usando modulação multinível, codificam-se $\log_2 M$ bits de dados sobre M símbolos, os quais são então transmitidos a uma taxa reduzida de símbolos de $R/\log_2 M$, sendo R a taxa de bits. A modulação multinível tem diversas implicações, sendo as mais importantes [5]:

- A largura espectral de sinais com modulação multinível é menor que a de sinais binários à mesma taxa de bits, permitindo filtragem mais intensa em redes roteadas opticamente e também permitindo a implementação de redes com eficiência espectral mais alta.
- A taxa de símbolos reduzida, comparada à do formato binário, possibilita maior robustez às distorções do sinal oriundas de efeitos dispersivos, tais como a dispersão cromática ou a PMD (*polarization mode dispersion*).
- A taxa de símbolos reduzida de sinais multiníveis contribui para transpor os limites da tecnologia optoeletrônica de alta velocidade.
- Uma desvantagem de sinais multiníveis está associada à tolerância reduzida ao ruído. Devido à adição de pontos na constelação de símbolos, reduz-se a distância entre eles para uma mesma potência média do sinal.
- Outra desvantagem da sinalização multinível é que ela geralmente conduz a uma reduzida tolerância às não linearidades da fibra, fato que se deve também à distância reduzida entre os pontos da constelação.

Modulação de intensidade multinível, modulação de fase multinível e modulação híbrida (intensidade / fase) multinível têm sido discutidos no contexto de transmissão de taxas multigigabit. A modulação de intensidade multinível (M-ASK – *multilevel amplitude shift keying*) não tem provado ser vantajosa para aplicações de transporte a fibra óptica até agora, principalmente devido à penalidade imposta à sensibilidade do receptor. Por exemplo, o formato 4-ASK incorre em uma penalidade de cerca de 8 dB em relação ao formato OOK binário devido à exigência, imposta pelo ruído que surge naturalmente da detecção direta, de espaçamentos desiguais entre os níveis de amplitude. O formato de modulação multinível que tem recebido grande atenção é o DQPSK (*differential quadrature phase shift keying*) [1].

5.2. Modulação com memória

Sinais digitais modulados com memória são aqueles cuja sucessão de símbolos obedece a determinadas regras ditadas por um código. Na modulação com memória utilizam-se mais de dois símbolos na codificação dos bits. Essa ampliação no alfabeto de símbolos não tem o mesmo objetivo da modulação multinível, que é o de aumentar a taxa de dados transmitidos ou reduzir, em comparação aos formatos binários, a taxa de símbolos para uma mesma taxa de bits. No caso da sinalização com memória, os símbolos adicionais têm a função de introduzir redundância para formatar o espectro ou evitar determinadas configurações adversas de

símbolos, melhorando a tolerância do formato a degradações específicas que ocorrem durante a transmissão. A finalidade principal da introdução de memória na modulação é prevenir erros de detecção.

Dentro da classe de formatos de modulação com memória, existe o formato codificação correlativa e o pseudo-multinível. Ambos caracterizam-se pela utilização de mais de dois símbolos para representar um único bit, fazendo uso de alguma modulação auxiliar. Por exemplo, utilizando a modulação auxiliar de fase, o bit “0” é representado pelo símbolo 0, enquanto que o bit “1” pode ser representado pelos símbolos $+1$ ou -1 . Devido à característica de lei quadrática dos receptores por detecção direta, que são incapazes de distinguir entre os símbolos $+1$ e -1 (já que ambos são entendidos como o bit “1” nesse tipo de receptor), o conjunto ternário de símbolos ópticos $\{-1, 0, +1\}$ é mapeado no conjunto binário de símbolos elétricos $\{0, 1\}$ [1].

Na modulação pseudo-multinível, a alocação dos símbolos redundantes independe da sequência de bits transmitidos, ao passo que, na codificação correlativa, o símbolo atribuído a determinada sequência de bits depende dos bits enviados anteriormente.

5.2.1. Modulação pseudo-multinível

O formato pseudo-multinível mais difundido (por ser o mais facilmente gerado) é o retorno ao zero com portadora suprimida (CSRZ – *carrier-suppressed return-to-zero*), no qual a informação é codificada sobre os níveis de intensidade $\{0,1\}$, mas a fase é deslocada de π a cada bit, independentemente da sequência de bits, conforme pode ser visualizado na Tabela 1. O CSRZ tem sido proposto principalmente para combater não linearidades da fibra. Já foi experimentalmente mostrado que uma escolha ótima de fases relativas entre pulsos adjacentes em um conjunto de quatro bits consecutivos pode ser $\{0, \pi/2, 0, \pi/2\}$ ou $\{0, 0, \pi, \pi\}$ para suprimir IFWM (*intrachannel four-wave mixing*) através da interferência destrutiva de seus efeitos [2].

Outro exemplo de codificação pseudo-multinível é a modulação auxiliar de polarização, na qual o estado de polarização é rotacionado de 90° a cada símbolo. Este tipo de modulação pseudo-multinível, conhecido como APol (*alternate polarization*), é particularmente eficiente para também reduzir o impacto de IFWM.

5.2.2. Codificações correlativa e restritiva

O representante mais importante dos formatos com codificação correlativa é o duobinário (DB) óptico ou PSBT (*phase-shaped binary transmission*), o qual será discutido mais detalhadamente na seção 6.6. Assim como no caso do CSRZ, a informação é transmitida pelos níveis de intensidade $\{0,1\}$, mas os deslocamentos de fase π ocorrem somente para bits

“1” separados por um número ímpar de bits “0”. Existe assim uma correlação entre as alterações auxiliares de fase e a codificação da informação. Deve-se ressaltar, entretanto, que a fase não é geralmente usada na detecção devido à insensibilidade dos receptores por detecção direta a ela.

Outro esquema de codificação correlativa é o chamado “inversão alternada de sinal” (AMI – *alternate-mark inversion*), o qual contém o mesmo conjunto de símbolos que o DB, mas com uma regra de correlação diferente: as mudanças de fase ocorrem para cada bit “1” (mesmo para bits “1” adjacentes) independente do número de bits “0” entre eles. O resultado da codificação de uma sequência de bits, tanto em AMI quanto em DB, pode ser visualizado na Tabela 1.

Tabela 1 - Sequência de bits modulada em diferentes formatos multiníveis.

Sequência de dados	0	1	0	1	1	1	0	0	1	1	0
CSRZ	0	-1	0	-1	+1	-1	0	0	+1	-1	0
DB	0	-1	0	-1	-1	-1	0	0	+1	+1	0
AMI	0	+1	0	-1	+1	-1	0	0	+1	-1	0

A codificação correlativa tem sido estudada no contexto das não linearidades intracanalais que dependem da sequência dos bits. Nesse caso, pode-se introduzir modulação de fase correlacionada ao bit ou podem-se remover determinadas configurações de bits que intensificam grande parte das distorções não lineares intracanalais do sinal. Esse tipo de esquema é chamado geralmente de codificação restritiva. Por exemplo, o IFWM depende da configuração dos dados transmitidos e das fases relativas entre símbolos. No formato OOK binário, a configuração de bits que deteriora a BER mais severamente é aquela que inclui muitos bits “0” isolados entre diversos bits “1” vizinhos. Então, removendo tais bits “0” isolados, pode-se melhorar a BER de forma significativa. Uma variedade de códigos de linha tem sido estudada para minimizar o efeito do IFWM e melhoras significativas têm sido previstas. Entretanto deve-se ressaltar que a codificação restritiva geralmente requer um aumento da taxa de símbolos, o que pode reduzir as vantagens oferecidas pelo código devido à exigência de uma maior relação sinal-ruído ou devido a uma maior largura espectral do sinal, situação que pode não ser possível em sistemas com alta eficiência espectral [2].

6. Formatos de modulação de intensidade

6.1. Chaveamento *on-off* sem retorno ao zero (NRZ-OOK - *nonreturn-to-zero on-off keying*)

O modo mais simples de gerar modulação óptica é usando o formato NRZ-OOK, muitas vezes denominado simplesmente de NRZ. Nesse formato, o pulso óptico permanece constante por toda a duração do *slot* do bit e sua amplitude não cai a zero entre dois ou mais bits “1” sucessivos. Como consequência a largura do pulso varia dependendo da sequência de bits enviada.

A partir de taxas de 10 Gb/s, o NRZ é mais convenientemente gerado usando DML (*directly modulated laser*) ou EAM (*electroabsorption modulator*) para distâncias curtas e médias ou usando MZM (*Mach-Zehnder modulator*) livre de *chirp* para longas distâncias. Quando é usado um modulador Mach-Zehnder, ele é polarizado a 50% da curva de transmissão (ponto de quadratura) e é controlado do mínimo ao máximo dessa curva através de uma variação de tensão de V_{π} . A Figura 19 apresenta o espectro óptico e o diagrama de olho típicos de um sinal NRZ.

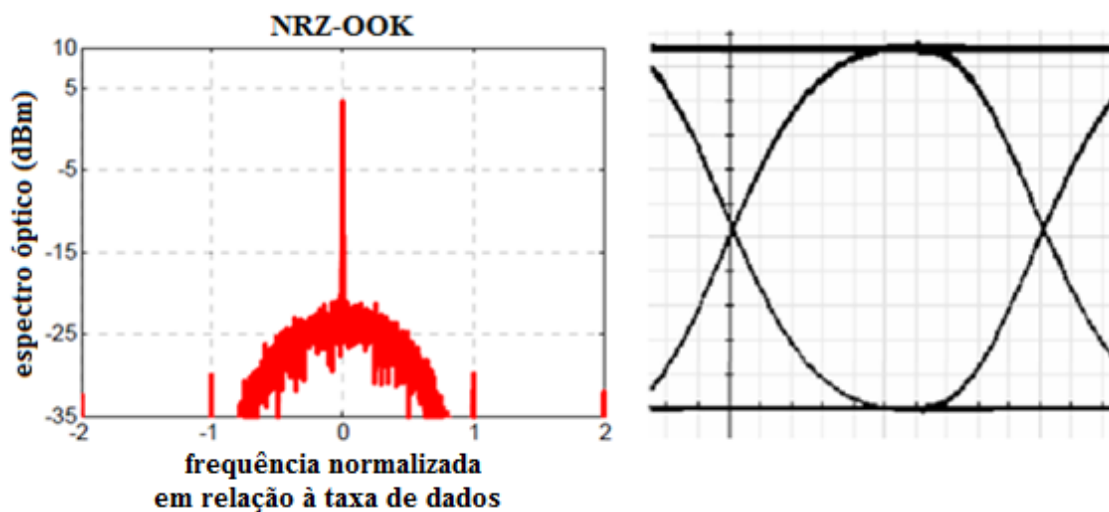


Figura 19 - Espectro óptico e diagrama de olho de um sinal NRZ. [2]

6.2. Chaveamento *on-off* com retorno ao zero (RZ-OOK - *return-to-zero on-off keying*)

No formato RZ, também conhecido como pulsado, cada pulso óptico que representa o bit “1” é mais curto que o *slot* de tempo desse bit e sua amplitude volta a zero antes do término da duração do bit. Portanto, ao contrário do formato NRZ, a largura dos pulsos permanece constante, independentemente da sequência de bits enviada. Uma desvantagem do formato RZ é que a largura de banda da sequência de bits enviada é, aproximadamente, o dobro da do formato

NRZ simplesmente porque transições *on-off* ocorrem mais vezes. Já a desvantagem do NRZ é que ele exige um controle mais apurado da largura do pulso, o qual, se não for bem feito, pode causar efeitos deletérios entre pulsos adjacentes se o pulso se espalhar durante a transmissão. As diferenças entre esses dois formatos podem ser visualizadas na Figura 20.

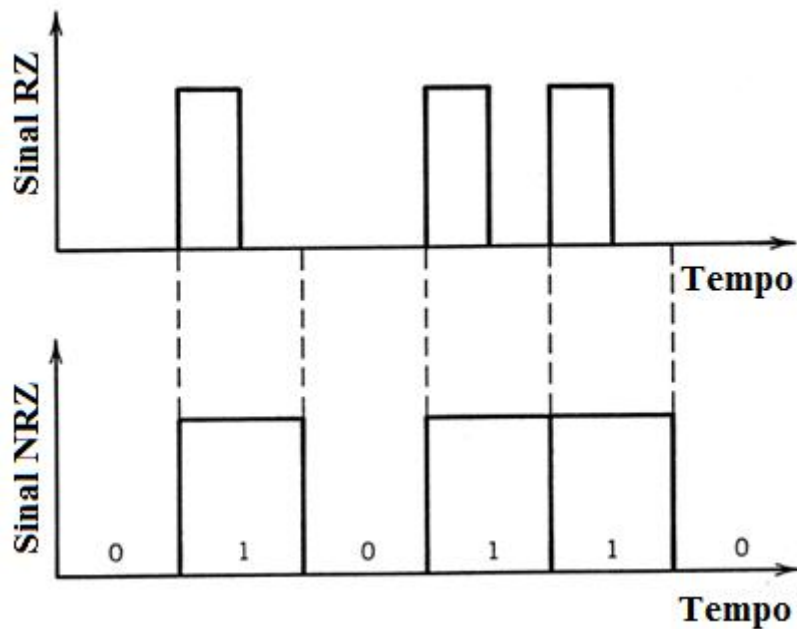


Figura 20 - Sequência de bits ópticos (010110) transmitida nos formatos RZ e NRZ. [4]

Transmissores RZ-OOK podem ser implementados pela geração eletrônica de formas de onda RZ, as quais, por sua vez, modulam uma portadora óptica. Outra alternativa, cujo diagrama de blocos é mostrado na Figura 21, consiste na utilização de pulsos externos para modelar um sinal NRZ óptico com o uso de um modulador adicional, o qual atua, nesse caso, como um modelador de pulso (*pulse carver*). Enquanto a primeira opção é confiável, com a tecnologia atual, para taxas de dados de até 10 Gb/s, um modelador de pulso tem de ser empregado para taxas de 40 Gb/s ou mais.

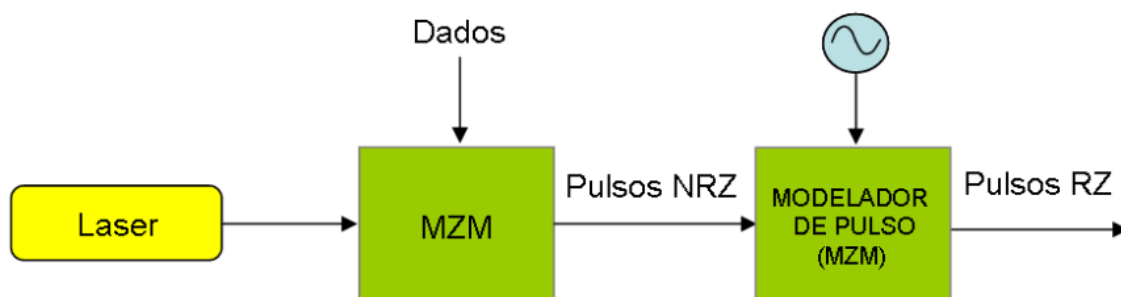


Figura 21 - Geração de pulsos RZ, a partir de pulsos NRZ, utilizando um MZM como modelador de pulso. [2]

Tipicamente os modeladores de pulso são implementados na forma de moduladores de eletroabsorção ou de moduladores Mach-Zehnder, controlados por um sinal senoidal, já que

sinais senoidais de vários gigahertz e de amplitude apreciável são facilmente gerados. Com EAM (*electroabsorption modulator*), pulsos ópticos com ciclo de trabalho pequeno podem ser gerados, o que torna os EAM adequados para sistemas TDM. Entretanto, devido às características de absorção variável e *chirp* residual dos EAM, os formatos avançados de modulação RZ são geralmente implementados por meio de modeladores de pulso baseados em MZM (*Mach-Zehnder modulator*).

Sinais RZ podem ser gerados com diferentes larguras de pulso, dependendo de como se polariza o MZM que atua como modelador de pulso e da frequência do sinal senoidal que o controla. Os sinais RZ mais difundidos são os com largura de pulso de 50%, 33% e 67%, os quais podem ser gerados através das maneiras listadas abaixo.

- Para gerar o sinal RZ com ciclo de trabalho de 50%, o MZM que atua como modelador de pulso tem de ser polarizado no ponto médio da curva de transmissão (ponto de quadratura) e ser controlado por um sinal senoidal com frequência igual à taxa de dados e cuja variação de tensão (*voltage swing*) se dá entre os pontos mínimos e máximos da curva de transmissão. Se essa variação de tensão for diminuída e, ao mesmo tempo, o ponto de polarização do modulador for ajustado, o ciclo de trabalho pode ser reduzido a até 36% (o que provoca perda de inserção, já que o modulador não é levado mais a seu máximo de transmissão) ou aumentado a 55% à custa de uma redução na taxa de extinção.
- Sinais RZ com ciclo de trabalho de 33% são gerados usando um MZM como modelador de pulso, controlado entre seus mínimos de transmissão por um sinal senoidal com frequência igual à metade da taxa de dados. Esses pulsos podem ser alargados à custa da redução da taxa de extinção se se diminuir a variação de tensão (*voltage swing*).
- Um MZM operando como modelador de pulso e controlado entre os máximos de sua função de transmissão por um sinal senoidal com frequência igual à metade da taxa de dados resulta em pulsos RZ com ciclo de trabalho de 67% e com fase alternada. Este formato, mais detalhadamente explicado na seção 6.3, é mais comumente chamado de RZ com portadora suprimida (CSRZ – *carrier-suppressed return-to-zero*).

Os métodos utilizados no controle senoidal de um MZM como modelador de pulso para sinais RZ 33% e 50% são mostrados na Figura 22.

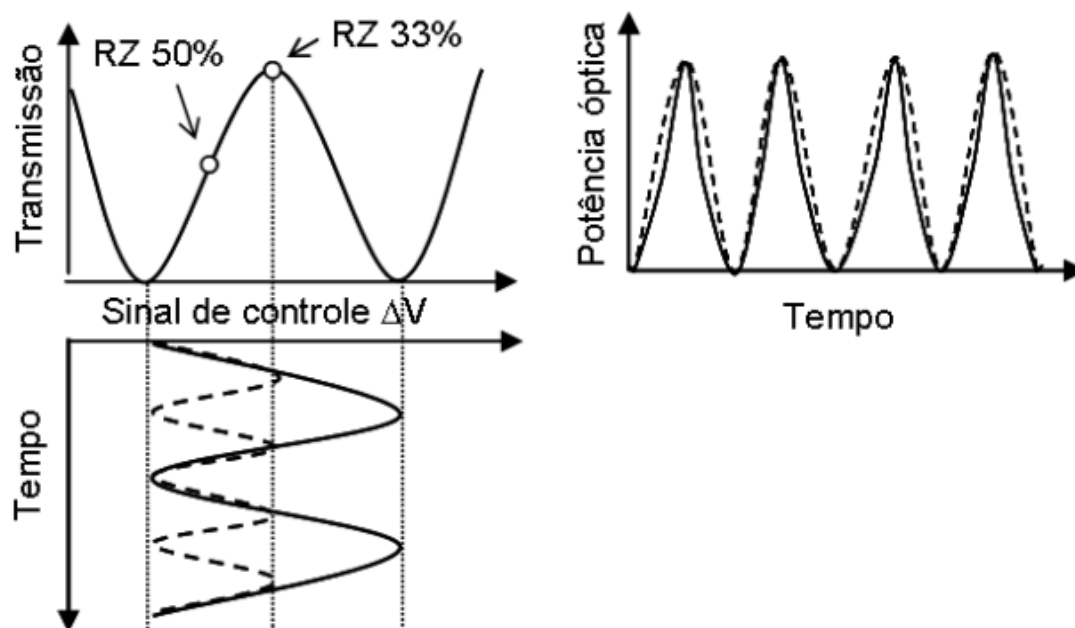


Figura 22 - Formas de controle senoidal de um MZM atuando como modelador de pulso para sinais RZ com ciclo de trabalho de 33% (sólido) e 50% (tracejado). Os pontos de polarização são indicados por círculos abertos. [1]

Os espectros ópticos e os diagramas de olho dos formatos RZ com ciclo de trabalho de 50% e 33%, conforme produzidos por um MZM em operação *push-pull*, são mostrados respectivamente na Figura 23 e na Figura 24.

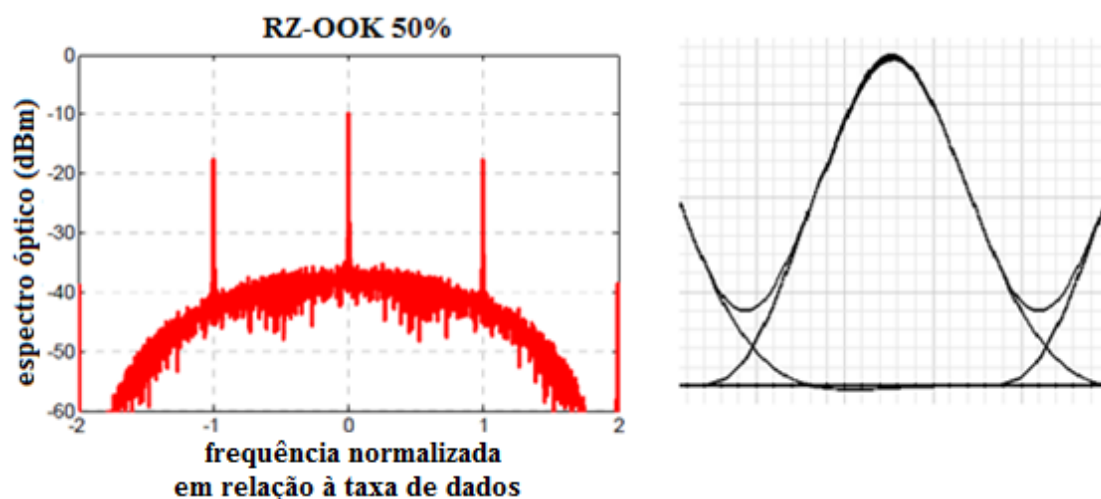


Figura 23 - Espectro óptico e diagrama de olho de um sinal RZ com ciclo de trabalho de 50%. [2]

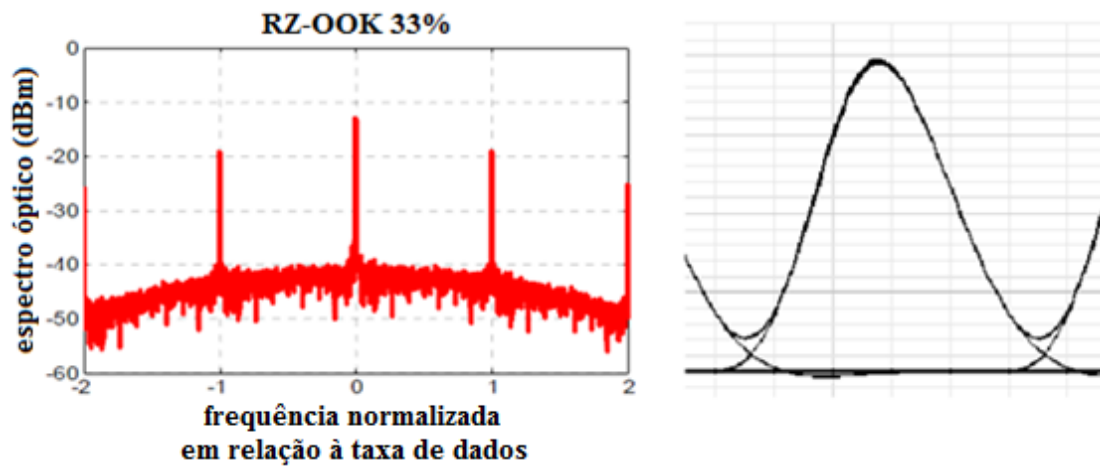


Figura 24 - Espectro óptico e diagrama de olho de um sinal RZ com ciclo de trabalho de 33%. [2]

6.3. Retorno ao zero com portadora suprimida (CSRZ – *carrier suppressed return-to-zero*)

O CSRZ é um formato de modulação pseudo-multinível caracterizado pela reversão do campo óptico a cada transição de bit (e essa reversão ocorre independentemente da sequência de bits anteriores). O CSRZ é mais convenientemente gerado pelo comando senoidal de um MZM (*Mach-Zehnder modulator*) entre seus máximos de transmissão a uma frequência igual à metade da taxa de dados, conforme mostra a Figura 25. Visto que existe um deslocamento de fase em π nos pontos de nulos da função de transmissão do MZM, inversões de fase ocorrem entre bits adjacentes. Assim, na média, o campo óptico de metade dos bits “1” tem sinal positivo enquanto a outra metade tem sinal negativo, resultando em uma envoltória do campo óptico com média zero. Como consequência, há uma supressão da portadora na frequência central, fato que dá nome ao formato.

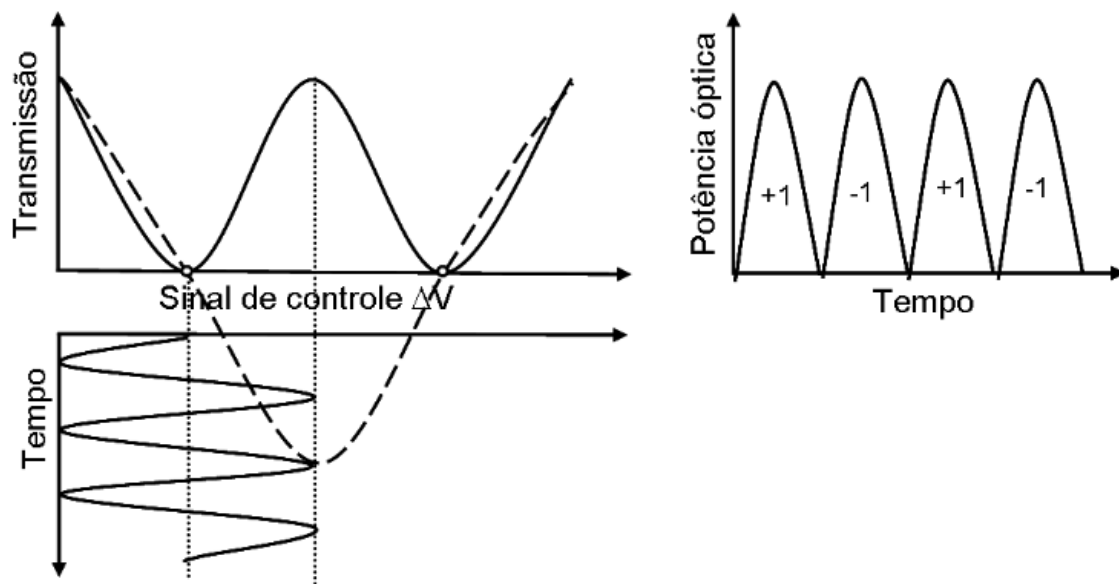


Figura 25 - Controle senoidal de MZM atuando como modelador de pulso para sinais CSRZ com ciclo de trabalho de 67%. As curvas de transmissão sólidas e tracejadas referem-se à potência e ao campo óptico respectivamente. O ponto de polarização é indicado por um círculo aberto. [1]

Os diagramas espectral e de olho do sinal CSRZ com ciclo de trabalho de 67%, gerado por um MZM na configuração *push-pull*, são mostrados na Figura 26. Pode-se ver que, devido ao fato de a fase óptica nesse sinal ser periódica com frequência igual à metade da taxa de dados, seu espectro exibe tons característicos em $\pm R/2$, sendo R a taxa de dados. Esse ciclo de 67% pode ser baixado para 50%, à custa de um excesso de perda de inserção, através da redução da amplitude dos sinais que controlam o MZM.

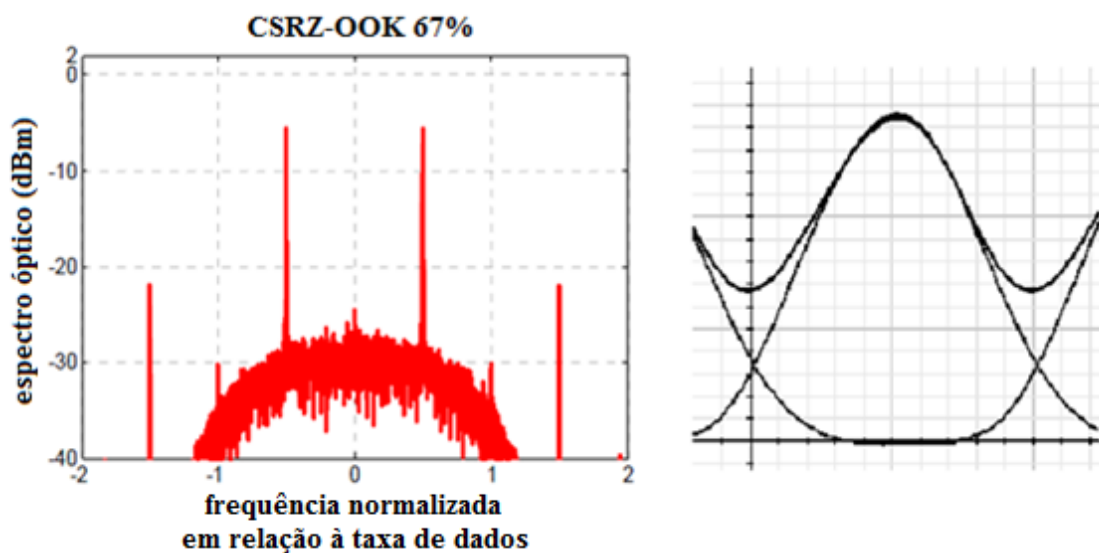


Figura 26 - Espectro óptico e diagrama de olho de um sinal CSRZ. [2]

O CSRZ apresenta importantes características de desempenho de transmissão. Sua propriedade de supressão da portadora pode ser benéfica pelo fato de que a frequência da portadora é a componente que contém a mais alta energia no espectro de sinais RZ. Portanto sua supressão pode significar maior tolerância a efeitos não lineares que a provida por outros formatos, o que permite um sistema com maior potência óptica transmitida na fibra, sendo uma opção interessante para sistemas WDM (*wavelength division multiplexing*) denso [2]. Essa é uma vantagem advinda da supressão da portadora no sinal CSRZ, mas é importante salientar que, quando se compara o CSRZ com outros formatos RZ, outras diferenças de desempenho podem surgir devido simplesmente às diferenças entre os ciclos de trabalho de cada sinal.

6.4. Retorno ao zero com *chirp* (CRZ – *chirped return-to-zero*), retorno ao zero com *chirp* alternado (ACRZ – *alternate-chirp return-to-zero*)

A modulação pulsada (formato RZ) provê um desempenho melhor ao receptor e frequentemente é mais tolerante a efeitos não lineares e dispersivos que formatos NRZ (*nonreturn-to-zero*). Além disso, uma quantidade controlada de *chirp* (isto é, um pouco de modulação analógica de fase que não contém informação) pode ser usada para melhorar a robustez a distorções não lineares. O desempenho superior no combate a essas distorções, obtido com a introdução de *chirp*, pode ser explanado de forma simplificada tanto no domínio do tempo quanto no domínio da frequência. No domínio do tempo, a modulação de fase adicional, em associação com a dispersão da fibra, resulta no alargamento dos pulsos, o que faz diminuir seus picos de intensidade. No domínio da frequência, a modulação de fase alarga o espectro do sinal, reduzindo assim os picos de intensidade das componentes espectrais do sinal. Essa redução nos picos de amplitude, tanto no espectro quanto no tempo, diminui as interações não lineares entre diferentes pulsos [2].

Dependendo se o *chirp* é o mesmo para todos os bits ou se toma valores opostos para bits adjacentes, o formato resultante recebe a denominação de RZ com *chirp* (CRZ) ou RZ com *chirp* alternado (ACRZ – *alternate-chirp return-to-zero*). Ambos os formatos são geralmente gerados usando modulação em três estágios: um modulador de dados NRZ, um modelador de pulso RZ controlado senoidalmente, livre de *chirp*, e um modulador de fase adicional controlado por um sinal senoidal. Este último modulador gera o termo de modulação de fase $e^{i*m*\cos(w_o t)}$, sendo m o índice de modulação de fase e w_o a frequência de modulação, que é igual à taxa de dados (no caso do CRZ) ou à sua metade (no caso do ACRZ).

A arquitetura mencionada acima de três moduladores é a mais complexa por requerer cuidadosa sincronização dos três sinais de controle, mas é também a que permite a operação mais flexível. Métodos alternativos de implementação que exploram o fato de que tanto a amplitude quanto a fase de um campo óptico podem ser modulados por um MZM de controle

dual (*dual-drive*) têm sido propostos. Neles, as funções do modelador de pulso e do modulador de fase podem ser combinadas em um único modulador, resultando num transmissor CRZ ou ACRZ composto de apenas dois moduladores.

Para entender como *chirp* é introduzido no sinal, basta analisar a função de transferência de um MZM, que, equivalentemente à função dada na seção 3.1.2.2. 2. Modulador Mach-Zender, também pode ser escrita assim:

$$T_E(V_1, V_2) = e^{\frac{i\pi * [V_1(t) + V_2(t)]}{2 * V_\pi}} * \cos\left(\pi * \frac{V_1(t) - V_2(t)}{2 * V_\pi} + \phi\right) \quad (9)$$

ϕ é a fase óptica relativa entre os dois braços do MZM na ausência das tensões senoidais de comando, que são dadas por:

$$V_1(t) = U_1 * \cos(2\pi f_0 t) \quad (10)$$

$$V_2(t) = U_2 * \cos(2\pi f_0 t + \delta) \quad (11)$$

δ é o atraso relativo de fase entre os dois sinais.

É sabido que para transformar o sinal NRZ em RZ e deixá-lo livre de *chirp*, basta fazer $V_1(t) = -V_2(t)$, adotando, para isso, $\delta = \pi$. Então, para introduzir *chirp* simétrico em relação ao bit, mantém-se fixa a diferença entre os sinais de comando do MZM ($V_1(t) - V_2(t)$), mas se introduz um desequilíbrio em suas amplitudes, o que pode acontecer tanto na configuração *push-pull* ($U_1 \neq U_2$, $\delta = \pi$) como na *push-push* ($U_1 \neq U_2$, $\delta = 0$). O índice de modulação de fase resultante é dado por:

$$m = \begin{cases} \frac{\pi * (U_1 + U_2)}{2 * V_\pi}, \delta = 0 \\ \frac{\pi * |U_1 - U_2|}{2 * V_\pi}, \delta = \pi \end{cases} \quad (12)$$

Por exemplo, pode-se obter sinal RZ 50% com índice de modulação de fase de 1,5 se as seguintes condições forem satisfeitas: $U_1 = 0,73 * V_\pi$ e $U_2 = 0,23 * V_\pi$.

No caso do CSRZ (RZ com ciclo de trabalho de 67%), o *chirp* produzido muda de fase a cada bit (*bit-alternating*), já que o sinal de modulação do MZM é periódico com frequência igual à metade da taxa de dados. Nesse caso, o desvio máximo de fase provocado pelo *chirp* no centro o pulso sempre reduz a diferença de fase existente inerentemente à modulação livre de *chirp* entre bits adjacentes. Por exemplo, o comando do MZM através de um único braço ($U_1 = V_\pi$ e $U_2 = 0$) resulta em $m = \frac{\pi}{2}$ rad, o que compensa totalmente a mudança de fase do

CSRZ no pico dos pulsos, deixando a fase óptica de todos os bits igual nesse instante de máxima amplitude.

Já para o formato RZ 33%, ao contrário das técnicas anteriores de introdução de *chirp*, o ajuste não só das amplitudes dos sinais $V_1(t)$ e $V_2(t)$ mas também da fase relativa entre eles precisa ser feito para se obter *chirp* simétrico em relação ao bit. A simetria é obtida impondo-se as seguintes condições:

$$U_1 = U_2 = V_\pi * \sqrt{\frac{m^2}{\pi^2} + 0,25} \quad (13)$$

$$\delta = 2 * \arccos\left(\frac{m * V_\pi}{\pi * U_1}\right) \quad (14)$$

Essa escolha de condições de operação leva ao seguinte:

$$V_1(t) - V_2(t) = 2U_1 * \sin\left(2\pi f_0 t + \frac{\delta}{2}\right) * \sin\left(\frac{\delta}{2}\right) \quad (15)$$

$$V_1(t) + V_2(t) = 2U_1 * \cos\left(2\pi f_0 t + \frac{\delta}{2}\right) * \cos\left(\frac{\delta}{2}\right) \quad (16)$$

Nota-se que o termo que modula a intensidade ($V_1(t) - V_2(t)$) está em perfeita quadratura com o que modula a fase ($V_1(t) + V_2(t)$). Portanto, como os picos dos pulsos sempre ocorrem quando o termo modulador de intensidade vale zero, há uma perfeita sincronização do *chirp* com o pico de intensidade do pulso, o que torna o *chirp* simétrico em relação ao bit.

Neste formato (RZ 33%), assim como acontece no caso do CSRZ, o *chirp* também muda de fase a cada bit. Por exemplo, para se ter um índice de modulação igual a 1,5 rad, $U_1(t) - U_2(t) \cong 0,7 * V_\pi$ e $\delta = 93^\circ$.

Tanto no caso do CRZ quanto no do ACRZ, a introdução de *chirp* alarga espectralmente o sinal. Embora isto reduza a adequação do formato para sistemas WDM com alta eficiência espectral, também aumenta sua robustez a não linearidades da fibra. O CRZ é usado predominantemente em comunicações a fibra de ponto a ponto que cobrem ultra-longas distâncias, como os sistemas submarinos transoceânicos, e costuma ter índice de modulação de fase de cerca de 1 radiano.

6.5. Não retorno ao zero com *chirp* (CNRZ – *chirped nonreturn-to-zero*)

Às vezes introduz-se *chirp* num formato NRZ para combater a dispersão provocada pela fibra óptica, o que dá origem ao formato CNRZ. Essa dispersão normalmente faz com que as componentes espectrais de maior frequência do pulso propaguem-se mais rapidamente que as de

menor frequência, resultando num alargamento temporal do pulso. Este efeito pode ser combatido pela introdução de uma quantidade apropriada de *chirp* em cada pulso óptico, de tal forma que se diminua a frequência da borda dianteira do pulso e se aumente a frequência da sua borda traseira, neutralizando assim o efeito da dispersão cromática.

6.6. Duobinário (DB, PSBT)

Em um sinal duobinário, o bit “1” é codificado com um pulso de luz com fase 0 ou π e o bit “0” é codificado com a ausência de um pulso ou com um pulso de baixa intensidade, dependendo do método utilizado para gerar o sinal DB. A fase dos bits “1” é escolhida de tal modo que exista uma alteração de fase igual a π entre bits “1” separados por um número ímpar de bits “0”.

O principal benefício de sinais DB é sua maior tolerância, quando comparada à provida por formatos de sinalização binária, à dispersão cromática e à filtragem óptica de banda estreita. Tais vantagens podem ser explicadas tanto no domínio do tempo quanto no domínio da frequência. A explicação no domínio do tempo considera a sequência de bits {1, 0, 1}, que em duobinário é codificada como {+1, 0, -1}. Se, em virtude de dispersão ou de filtragem óptica, os dois pulsos que representam os bits “1” se alargarem para dentro do bit “0”, a codificação DB faz com que ocorra interferência destrutiva entre esses bits “1”, mantendo baixo o nível do bit “0”. Os formatos OOK, por outro lado, fazem com que os pulsos interfiram construtivamente, levantando assim o nível do bit “0” e fechando o diagrama de olho. A explicação no domínio da frequência baseia-se na largura espectral menor de sinais DB adequadamente filtrados, o que reduz distorções provocadas por dispersão. Essa compressão espectral resulta das transições mais suaves que ocorrem na codificação duobinária {+1, 0, -1}, se comparadas às transições mais acentuadas do formato OOK {+1, 0, +1} [1]. Tanto o diagrama espectral quanto o de olho do formato DB podem ser vistos na Figura 27.

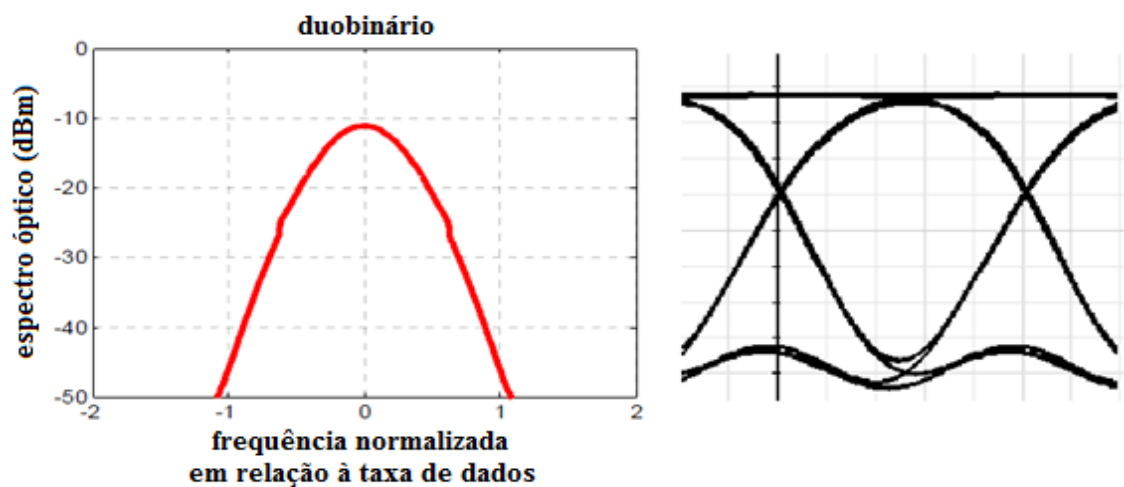


Figura 27 - Espectro óptico e diagrama de olho de um sinal duobinário. [2]

No formato duobinário, os transmissores usam uma versão pré-codificada da sequência de bits, conforme mostrado na Figura 28. Esta sequência de dados pré-codificada exibe uma mudança de nível para cada bit “0” contido na sequência de dados original. Então essa sequência pré-codificada é convertida, por meio de intensa filtragem passa-baixas no domínio elétrico, em um sinal elétrico de três níveis, que é então usado para controlar um MZM livre de *chirp* entre seus máximos de transmissão [1]. De forma alternativa, pode ser utilizado um MZM projetado para operar a um quarto da taxa de dados desejada, o qual combinaria as funções de filtro passa-baixas e modulador [2].

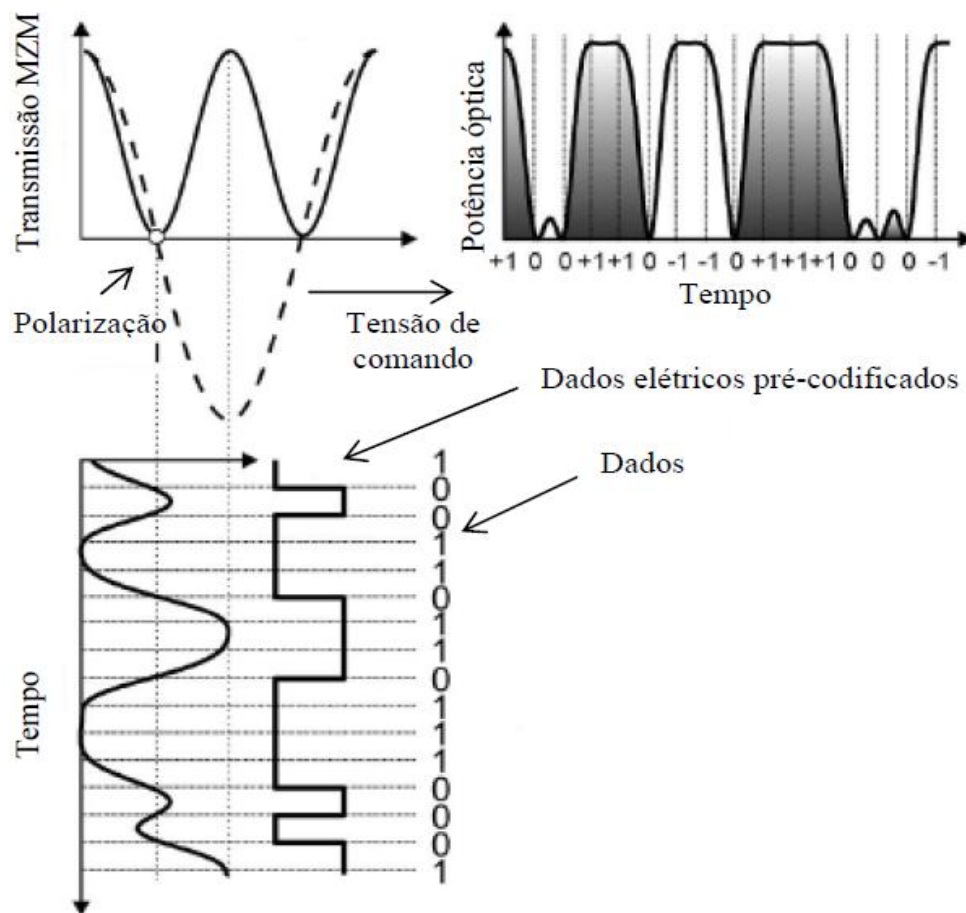


Figura 28 - Visualização das sequências original e pré-codificada de dados e do sinal duobinário de 3 níveis que comanda o MZM entre seus máximos de transmissão. [1]

Entre os transmissores desenvolvidos para a modulação DB, distinguem-se dois tipos: o que produz o DB padrão e o que produz o formato chamado transmissão binária com fase formatada (PSBT – *phase-shaped binary transmission*). No PSBT, os bits “0” contêm uma pequena quantidade de energia. Classicamente, os formatos DB e PSBT são gerados por meio de filtragem elétrica, a qual pode ser implementada por meio de um filtro de atraso e adição (DB) ou por meio de um filtro passa-baixas (PSBT) com características de *roll-off* apropriadas e uma largura de banda de 3 dB de cerca de 25% da taxa de bits. Esses dois modos diferentes de

gerar sinais DB podem ser visualizados na Figura 29. Essa filtragem elétrica também pode ser equivalentemente realizada no domínio óptico. Para isso, basta passar um sinal óptico binário de níveis $\{-1, +1\}$, ou seja, um sinal PSK (*phase shift keying*), por um filtro óptico passa-banda de largura de banda estreita. Com isso, o sinal que sai do filtro óptico já é um sinal DB.

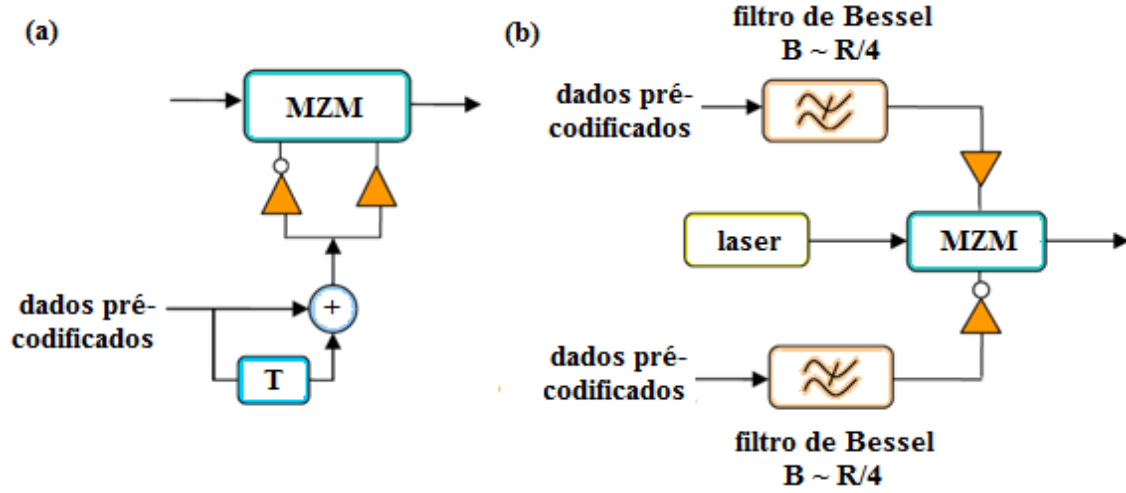


Figura 29 - Circuitos geradores de sinais DB: com filtro de atraso e adição (a) e com filtro passa-baixas (b). B: largura de banda do filtro passa-baixas. R: taxa de dados. [2]

No DB, a pré-codificação dos dados é requerida para que não haja propagação de erro depois da detecção. A pré-codificação pode ser realizada pela seguinte operação:

$$b_k = b_{k-1} \oplus \overline{a_k} \quad (17)$$

Na fórmula acima, onde b_k é o sinal pré-codificado do k-ésimo bit, $\overline{a_k}$ é o k-ésimo bit com sinal invertido e \oplus representa a operação lógica ou-exclusivo (XOR).

Como exemplo, a Tabela 2 mostra a sequência pré-codificada de dados e o sinal DB transmitido para uma sequência aleatória de bits de dados.

Tabela 2 - Sequência de bits, sequência de bits invertidos, dados pré-codificados e sinal duobinário transmitido.

a_k	0	0	1	0	1	1	1	0	0	1	0	1
$\overline{a_k}$	1	1	0	1	0	0	0	1	1	0	1	0
b_{k-1}	1	0	1	1	0	0	0	0	1	0	0	1
b_k	0	1	1	0	0	0	0	1	0	0	1	1
DB	0	0	+1	0	-1	-1	-1	0	0	-1	0	+1

Formatos DB baseados em filtros de atraso e adição resultam numa melhor sensibilidade, enquanto os baseados em filtros passa-baixas oferecem, à custa de uma redução

de cerca de 1 a 2 dB na sensibilidade em relação ao NRZ-OOK, uma maior tolerância à dispersão cromática [2].

Devido à sua estreita largura espectral e à alta tolerância à dispersão cromática, o formato duobinário torna-se muito atrativo para a transmissão óptica de altas taxas de bits e sistemas de alta eficiência espectral. Recentemente têm sido demonstrados sistemas com eficiência espectral de 0,8 b/s/Hz a 40 Gb/s por canal. Por isso, atualmente os formatos DB são considerados os mais promissores, no que se refere a custos, para o desenvolvimento de tecnologias de 40 Gb/s sobre a infraestrutura WDM (*wavelength division multiplexing*) de 10 Gb/s por canal [2].

6.7. Inversão alternada de sinal (AMI – *alternate-mark inversion*)

Apesar de ser também classificado como um formato de resposta parcial, o AMI não compartilha da característica de possuir largura de banda limitada, comum a outros formatos de resposta parcial. No formato AMI, ocorrem inversões de fase para cada bit “1” presente na sequência de dados original, independentemente dos bits adjacentes. Em comunicações ópticas, o AMI é geralmente implementado na forma RZ (RZ-AMI) e, assim como em outros formatos RZ, a modulação adicional de fase do AMI ajuda a combater algumas não linearidades da fibra [1]. O espectro óptico e o diagrama de olho do formato RZ-AMI 33%, podem ser visualizados na Figura 30.

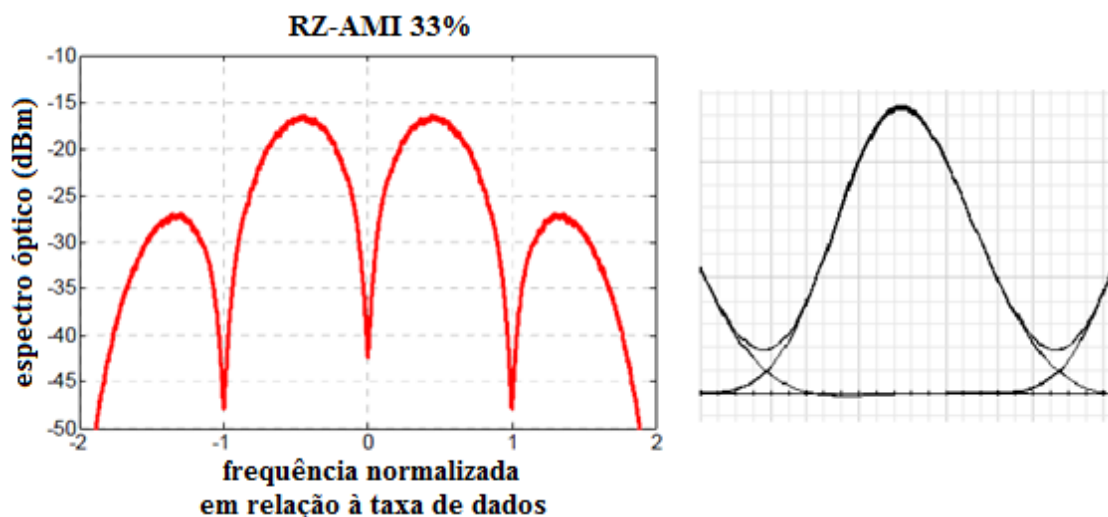


Figura 30 - Espectro óptico e diagrama de olho de um sinal RZ-AMI com ciclo de trabalho de 33%. [2]

Uma das formas de se implementar um transmissor AMI é idêntica à usada para o transmissor DB (ver Figura 29). A única diferença é que, ao invés de utilizar um circuito de atraso e adição (que atua como filtro passa-baixas), é utilizado um circuito de atraso e subtração

(o qual age como um filtro passa-altas). Para obter o RZ-AMI, o modulador de dados é seguido por um modelador de pulso padrão [1].

No AMI, a pré-codificação dos dados também é requerida para que não haja propagação de erro depois da detecção. A pré-codificação pode ser realizada pela seguinte operação:

$$b_k = b_{k-1} \oplus a_k \quad (18)$$

Na fórmula acima, onde b_k é o sinal pré-codificado do k-ésimo bit, a_k é o k-ésimo bit e \oplus representa a operação lógica ou-exclusivo (XOR).

Como exemplo, a Tabela 3 mostra a sequência pré-codificada de dados e o sinal AMI transmitido para uma sequência aleatória de bits de dados.

Tabela 3 - Sequência de bits, dados pré-codificados e sinal AMI transmitido.

a_k	0	0	1	0	1	1	1	0	0	1	0	1
b_{k-1}	1	1	1	0	0	1	0	1	1	1	0	0
b_k	1	1	0	0	1	0	1	1	1	0	0	1
AMI	0	0	-1	0	+1	-1	+1	0	0	-1	0	+1

Outra alternativa de transmissor AMI (vista na Figura 31 (a)) é, primeiro, gerar modulação no formato NRZ binário, podendo-se modular a fase ($\{-1, +1\}$) ou a intensidade ($\{0, +1\}$), como mostrado na parte de cima da Figura 31 (b). Este sinal é então passado através de um filtro de atraso e subtração, implementado por um interferômetro de atraso (DI – *delay interferometer*) com atraso de $\tau \leq T$, onde T é o período do bit. Para agir como um filtro de atraso e subtração, o DI é configurado para provocar interferência destrutiva na sua saída quando não há mudança de fase entre o pulso atrasado (linha tracejada Figura 31) e o pulso atual (linha cheia da Figura 31). Já quando há mudança de fase entre esses pulsos, ou seja, sempre que o sinal de entrada NRZ difere de sua versão atrasada em τ , pulsos ópticos de fase alternada são produzidos na saída do DI (parte de baixo da Figura 31). Sinais RZ-AMI com ciclos de trabalho diferentes podem ser gerados pela variação de τ , sem a necessidade de um modelador de pulso ativo [1].

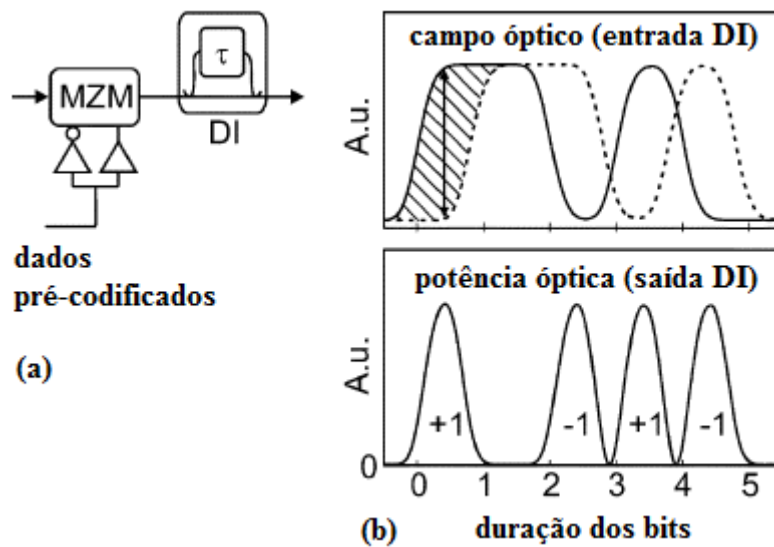


Figura 31- (a) Uma dos tipos de transmissor AMI. (b) Sinal no formato NRZ (linha cheia) interferindo com sua replica atrasada (linha tracejada) em um DI para produzir pulsos de fase alternada. [2]

Ainda existe outra forma de geração do AMI, a qual consiste em passar um sinal DB através de um modelador de pulso CSRZ. As inversões de fase inerentes ao CSRZ convertem o formato DB em AMI. Por isso muitas vezes usa-se o termo “duobinário com portadora suprimida” (DCS – *duobinary-carrier-suppressed*) para se referir ao AMI [1].

7. Formatos de modulação de fase diferencial

Alguns formatos de modulação já discutidos utilizam o chaveamento *on-off* e, em adição, modulam a fase óptica não para carregar informação mas para tornar o sinal mais robusto à dispersão cromática, à filtragem óptica ou a não linearidades da fibra. Esse é o caso dos formatos AMI, CRZ e CSRZ. Entretanto existem formatos de modulação que utilizam o chaveamento na fase óptica para carregar a informação. Estes são os formatos chaveados por deslocamento de fase (PSK – *phase shift keying*) [2]. Para a demodulação de sinais PSK, elementos que, de alguma forma, convertem fase em intensidade devem ser inseridos antes do fotodetector, haja vista que os receptores de detecção direta são insensíveis à fase óptica. Além disso, como tais receptores não possuem uma referência de fase (são não coerentes), essa referência tem de ser provida pelo próprio sinal: a fase do bit precedente é usada como referência para recuperação do sinal. Por isso, estes formatos, os quais carregam a informação nas transições de fase entre os bits, recebem o nome de “chaveamento por deslocamento de fase diferencial” (DPSK – *differential phase shifted keying*) [2].

Conforme já foi dito, formatos baseados em DPSK não são novos. Eles foram extensivamente estudados no final dos anos 80 e início dos anos 90 para uso principalmente em sistemas a fibra óptica com enlace único que empregavam detecção coerente. Quando os EDFA foram introduzidos, o interesse em sistemas coerentes diminuiu. Assim, por cerca de uma década, sistemas WDM utilizando formatos OOK, em conjunto com amplificadores ópticos, dominaram a pesquisa em sistemas de longa distância. O interesse em sistemas DPSK foi renovado à medida que os sistemas WDM foram impulsionados para níveis de desempenho ainda mais altos [7].

7.1. Modulação por deslocamento diferencial de fase (DPSK - *differential phase shift keying*)

No formato DPSK, a informação é transmitida nos deslocamentos de fase (que podem ser 0 ou π) entre bits adjacentes. Mais especificamente, o bit “1” é representado por uma mudança de fase igual a π , ao passo que o bit “0” não apresenta mudança de fase. Por isso todo *slot* de bit sempre apresenta potência óptica. Assim como ocorre para o OOK, existe tanto o formato NRZ-DPSK, caso em que a potência óptica ocupa todo o *slot* de bit, quanto o RZ-DPSK, no qual a potência óptica não ocupa completamente o *slot* do bit, mas apenas parte dele.

A principal vantagem do DPSK é a redução de 3 dB na relação sinal-ruído requerida para atingir a determinada BER do formato OOK. Isso pode ser percebido pela comparação entre as constelações dos dois formatos, apresentadas na Figura 32, a qual mostra que, no sinal OOK, a potência está presente em apenas um dos símbolos (bit “1”), enquanto, no sinal DPSK, a potência está presente em ambos os símbolos, que estão deslocados de 180° na fase. Portanto,

para que o sinal DPSK tenha a mesma potência óptica média que o OOK, a distância de cada símbolo DPSK em relação à origem do diagrama de constelação deve ser igual à $\frac{1}{\sqrt{2}}$ da do símbolo OOK à origem. Por isso, a distância entre os dois símbolos DPSK é aumentada de $\sqrt{2}$ $\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\right)$ em relação à distância entre os símbolos OOK. Essa distância maior, por sua vez, possibilita ao sinal DPSK aceitar $\sqrt{2}$ mais ruído que o sinal OOK para uma mesma BER, o que significa uma relação sinal-ruído 3 dB menor. Essa menor exigência de relação sinal-ruído para DPSK pode ser usada para aumentar a distância de transmissão, reduzir exigências de potência óptica ou relaxar especificações de componentes [7].

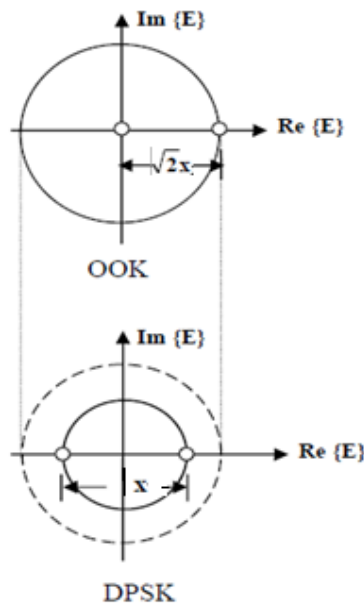


Figura 32 - Constelações dos sinais OOK (superior) e DPSK (inferior). [7]

Outro benefício importante para o DPSK com detecção balanceada é sua maior tolerância a flutuações de potência do sinal no circuito de decisão no receptor, porque o limiar de decisão independe da potência de entrada. O DPSK é também mais robusto à filtragem óptica de banda estreita que o OOK, especialmente quando detecção balanceada é empregada [7]. O DPSK é também mais resistente a alguns efeitos não lineares que o OOK devido aos seguintes fatores:

- nele, a potência óptica está mais uniformemente distribuída que no formato OOK, haja vista que a potência está presente em cada *slot* de bit no DPSK;
- o pico de potência óptica é 3 dB menor no DPSK que no OOK, para a mesma potência óptica média.

O transmissor DPSK pode ser implementado no formato NRZ ou RZ. De forma semelhante aos formatos DB e AMI, o sinal precisa primeiro ser codificado diferencialmente no

transmissor para evitar propagação de erros no receptor. É a sequência de dados pré-codificados que comandará o modulador e gerará ou não a mudança de fase do campo óptico da luz do laser. A modulação de fase pode ser realizada tanto por um modulador de fase quanto por um MZM. A Figura 33 mostra as duas configurações típicas para um transmissor RZ-DPSK. O modulador de fase modula a fase do sinal óptico ao longo do círculo unitário no plano complexo que representa o campo óptico, o que resulta em um sinal de envoltória constante. Com isso, há geração de *chirp* durante as transições dos bits, já que a modulação de fase não ocorre instantaneamente. Um segundo modulador comandado por um sinal senoidal pode ser usado para modelar os pulsos e gerar um sinal RZ-DPSK. Já quando é usado um MZM para a modulação de fase, polariza-se o modulador no seu nulo de transmissão e aplica-se nele um sinal de comando com o dobro da tensão de chaveamento usada na modulação OOK, o que produz uma modulação do sinal ao longo do eixo real do plano complexo. Entretanto, mesmo que a tensão de comando não tenha o valor adequado, sempre se produz deslocamentos de fase exatamente iguais a π em virtude de a fase do campo óptico mudar seu sinal no ponto de nulo da curva de transmissão do MZM. O custo dessa modulação de fase precisa é a presença de um pouco de modulação residual de amplitude (ou seja, diminuição da intensidade óptica) nos instantes de transição de fase. No entanto, visto que a codificação da informação é feita na fase óptica, e não na intensidade, uma modulação exata de fase é mais importante que uma envoltória de intensidade constante; por isso esses desvios de intensidade são de pouca importância, especialmente no caso do RZ-DPSK, que praticamente elimina essa parte do pulso com modulação residual.

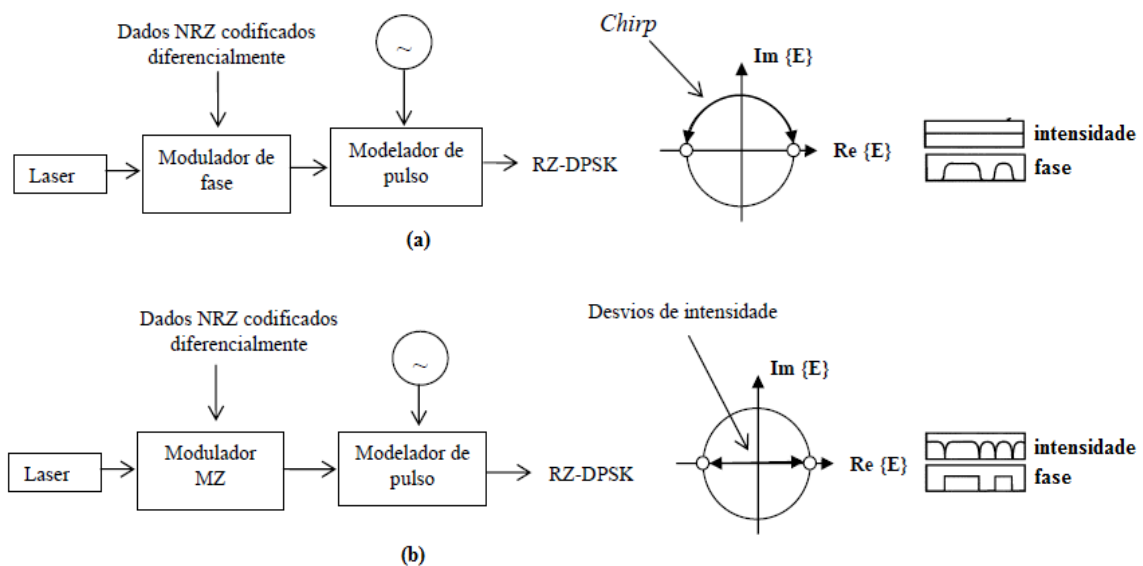


Figura 33 - Duas configurações típicas de um transmissor RZ-DPSK: (a) utilizando um modulador de fase e (b) utilizando um MZM. [2]

A Figura 34 apresenta os diagramas espectral e de olho para o NRZ-DPSK e o RZ-DPSK 33%, ambos gerados por MZM.

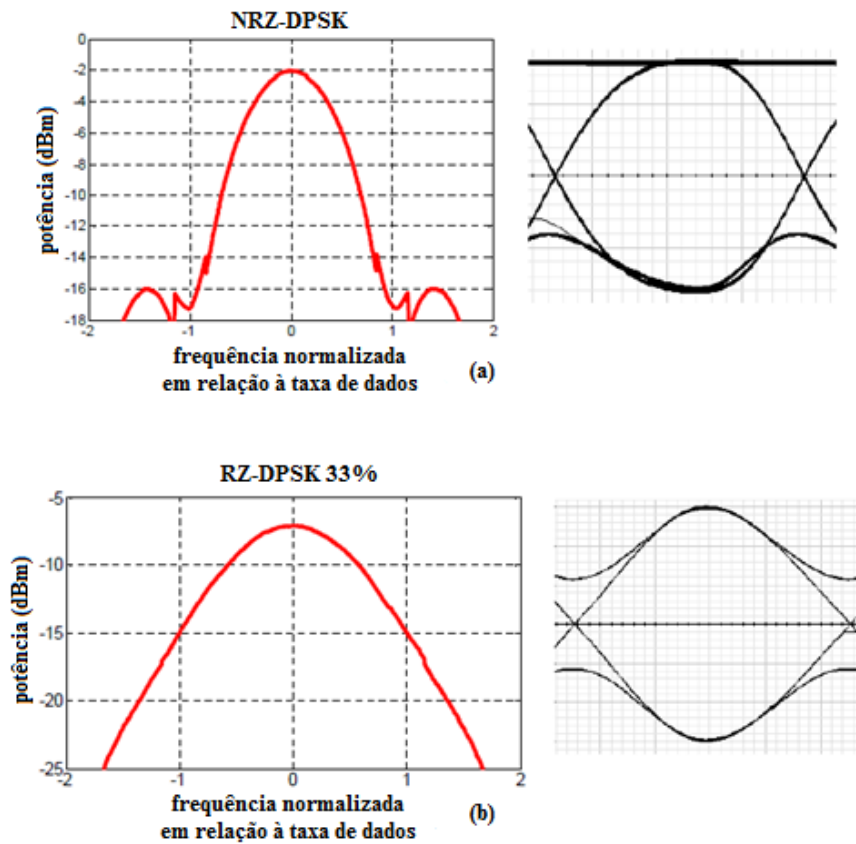


Figura 34 - Diagramas espectral e de olho: (a) NRZ-DPSK e (b) RZ-DPSK 33%. [2]

O sinal DPSK não pode ser demodulado diretamente por detecção direta, dada a insensibilidade desse método à fase óptica. Portanto, para converter a modulação de fase em modulação de intensidade, utiliza-se um interferômetro de atraso (DI – *delay interferometer*), com atraso diferencial de um período de bit, pelo qual o sinal óptico passa antes de ser fotodetectado. O DI divide o sinal óptico modulado em fase em dois caminhos de comprimentos diferentes, atrasa a parte do sinal que percorre o maior caminho e provoca assim, na sua saída, a interferência construtiva ou destrutiva de um pulso com o seu precedente, dependendo da diferença de fase entre eles. Se a interferência dos pulsos for construtiva, haverá presença de potência na saída do DI. Caso contrário, se a interferência for destrutiva, não haverá potência após o DI. Dessa forma, o bit precedente age como uma referência de fase para a demodulação do bit corrente.

Idealmente, uma das portas de saída do DI é ajustada para ter interferência destrutiva sempre que não houver mudança de fase entre os pulsos ópticos interferentes (o que ocorre quando há bit “0” na sequência de dados original) e interferência construtiva quando houver mudança de fase entre os pulsos (o que ocorre quando há bit “1” na sequência de dados

original). Por isso essa porta recebe o nome de porta destrutiva (porta A). A outra porta, devido à conservação de energia, exibe interferência construtiva quando a primeira provoca interferência destrutiva e vice-versa e é, por isso, denominada porta construtiva (porta B). Por essa razão as duas portas de saída do DI apresentarão trens de bits idênticos, mas a porta B terá uma lógica invertida (ou seja, surgem bits “1” onde, na sequência original, apareciam bits “0”). Analisando-se cuidadosamente os sinais demodulados na saída do DI, percebe-se que a porta construtiva carrega modulação DB (age como um filtro de atraso e adição), enquanto que a porta destrutiva apresenta modulação AMI (age como um filtro de atraso e subtração).

Visto que ambas as portas de saída do DI carregam a informação completa (embora a porta construtiva apresente inversão na lógica), ela pode ser detectada através de apenas uma delas (detecção de extremidade única - *single-ended detection*). De modo alternativo, ambas as portas podem ser conectadas a dois fotodiodos para formar um receptor balanceado. A configuração de um receptor DPSK balanceado pode ser visualizada na Figura 35.

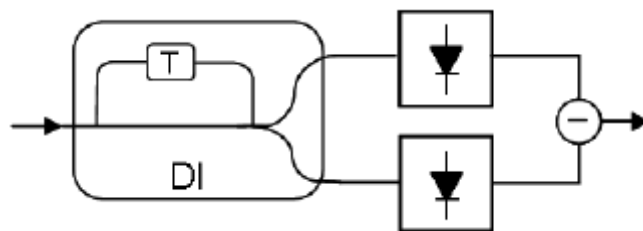


Figura 35 - Receptor DPSK balanceado. [2]

7.2. Modulação por deslocamento diferencial de fase em quadratura (DQPSK - *differential quadrature phase shift keying*)

O DQPSK é o único formato multinível verdadeiro que recebeu atenção até agora em comunicações ópticas. Ele transmite as quatro transições de fase possíveis $\{0, +\pi/2, -\pi/2, \pi\}$ a uma taxa de símbolos igual à metade da taxa efetiva de bits, ou seja, cada símbolo transmitido carrega dois bits [1].

A forma de transmissor DQPSK mais largamente utilizada consiste de dois MZM em paralelo operando como moduladores de fase [1], os quais são integrados juntos para atingir estabilidade de fase. A Figura 36 detalha a constituição desse tipo de transmissor DQPSK:

- um laser CW (*continuous wave*);
- um divisor que divide igualmente o sinal óptico entre dois caminhos;
- dois moduladores MZM controlados por sinais pré-codificados, gerando assim dois sinais DPSK;

- um alterador de fase para rotacionar o campo óptico em $\pi/2$ de um dos dois sinais DPSK;
- um combinador que gera um único sinal de saída, resultando num sinal de quatro níveis de fase.

Esse tipo de estrutura permite tirar proveito das mudanças exatas de fase produzidas por MZM, independentemente de distorções presentes em seus sinais de comando. Além disso, essa estrutura exige somente sinais eletrônicos de comando binários, que são muito mais fáceis de serem gerados em altas frequências que sinais de vários níveis.

Após o transmissor, existe a opção de adicionar um outro MZM como modelador de pulso para produzir um sinal RZ-DQPSK. Os espectros ópticos para o NRZ-DQPSK e para o RZ-DQPSK 33% são mostrados na Figura 37. Percebe-se que o espectro óptico do DQPSK é idêntico ao do DPSK, porém o DQPSK apresenta o espectro óptico comprimido por um fator de dois.

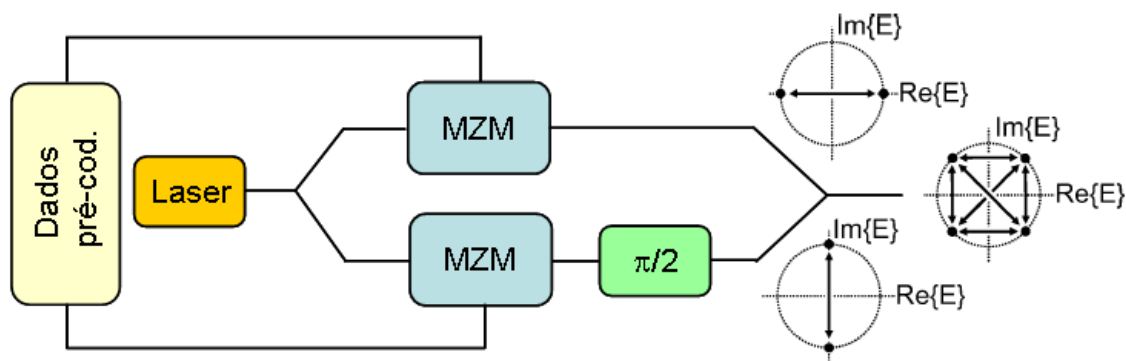


Figura 36 - Diagrama de um transmissor DQPSK utilizando dois MZM em paralelo. [2]

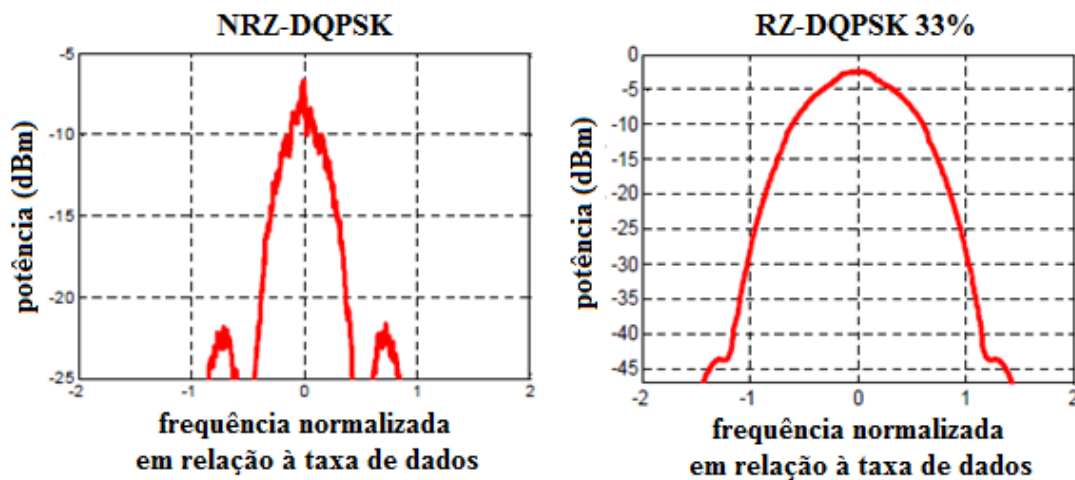


Figura 37 - Diagrama spectral do NRZ-DQPSK e do RZ-DQPSK 33%. [2]

O receptor DQPSK consiste basicamente de dois receptores DPSK balanceados. Esses receptores são um pouco diferentes entre si: cada um possui um braço do DI que atrasa o sinal por uma duração duas vezes maior que a duração do bit (que é a duração do símbolo DQPSK), mas também cada um deles possui um braço que altera a fase do sinal óptico de valores diferentes (um altera a fase de $+\pi/4$ e o outro, de $-\pi/4$). O sinal DQPSK é primeiro dividido em duas partes iguais, as quais são inseridas nos dois receptores DPSK balanceados em paralelo para demodular simultaneamente as duas sequências de dados binários contidas no DQPSK [7]. A configuração do receptor DQPSK é mostrada na Figura 38.

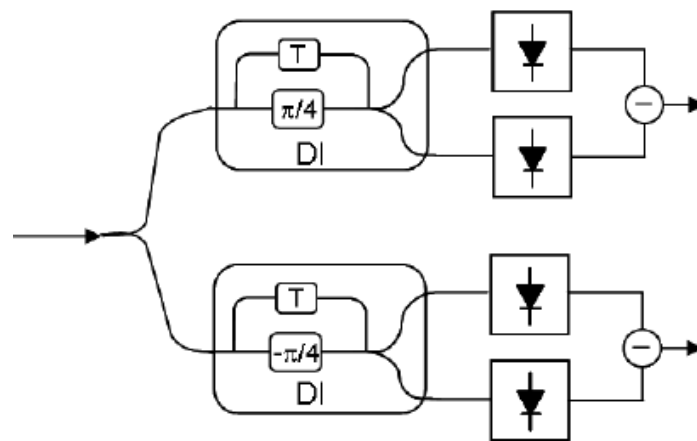


Figura 38 - Configuração de um receptor DQPSK. [2]

A vantagem do DQPSK é que, para a mesma taxa de dados, a taxa de símbolos transmitida é reduzida por um fator de dois. Em consequência, o espectro ocupado pelo DQPSK é reduzido, o que se traduz numa redução por exigência de largura de banda para o transmissor e o receptor [7]. Além disso, o espectro comprimido do DQPSK possibilita alta eficiência espectral em sistemas WDM, bem como uma maior tolerância à dispersão cromática. Sua maior duração de símbolo, comparada à dos símbolos dos formatos OOK (considerando uma mesma taxa de bits), torna-o mais robusto à PMD.

8. Simulações com o Optisystem

Nesta seção são apresentados alguns resultados preliminares de análise sistêmica, obtidos de simulações feitas no software *Optisystem* (versão 7.0), da empresa canadense *Optiwave Systems Inc.* Nessas simulações são visualizados o diagrama de olho e o espectro do sinal, modulado no formato DPSK, na configuração em que o transmissor está ligado diretamente ao receptor (*back-to-back*).

Inicialmente foi necessária uma etapa de aprendizado de manipulação do software, que possui tutoriais para autoaprendizado. Em seguida, foi montado o transmissor NRZ-DPSK esquematizado na Figura 39, que é constituído de um gerador de sequência de bits pseudo-aleatórios, um pré-codificador duobinário (o qual transmite um bit que é resultado da operação ou-exclusivo entre o bit de dado a ser enviado e o bit anteriormente transmitido), um gerador de pulsos NRZ, um laser operando no modo *continuous wave* e um MZM. O laser foi colocado para operar em 1552,52 nm (ou 193,1 THz). O MZM foi polarizado no nulo de transmissão de sua função de transferência e controlado por sinais simétricos ($V_1(t) = -V_2(t)$) de amplitude igual a $V_\pi/2$, sendo V_π a tensão de chaveamento do MZM.

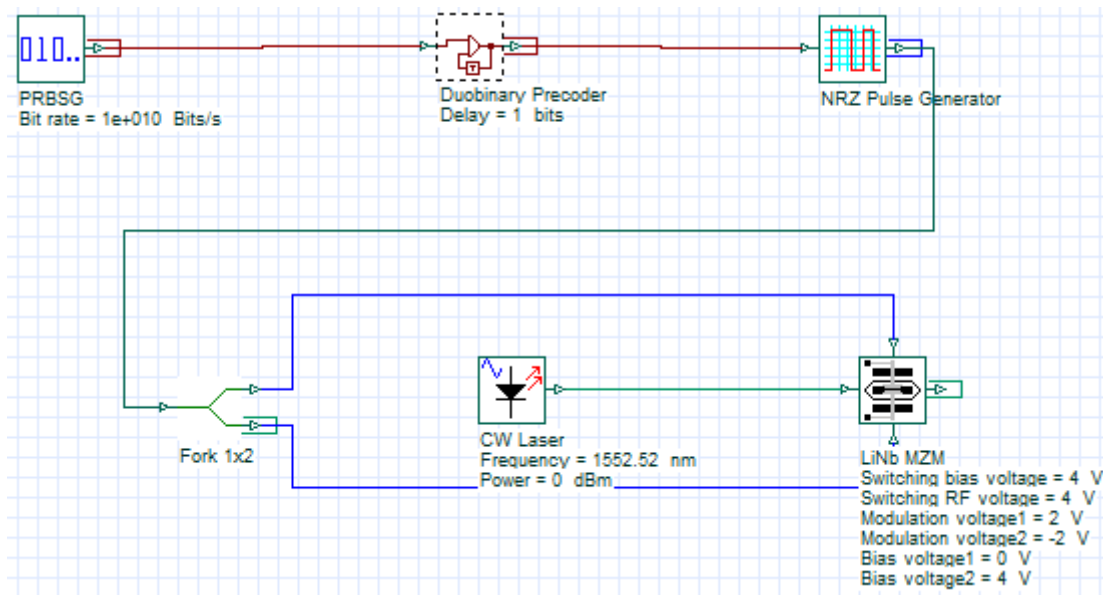


Figura 39 - Transmissor NRZ-DPSK.

Após montado o transmissor, foi montado o receptor visto na Figura 40, composto por um interferômetro de atraso igual ao período dos bits produzidos, dois fotodetectores (um para decodificar o sinal DB e o outro para decodificar o sinal AMI), um subtrator elétrico para fazer a combinação dos dois sinais anteriores e, finalmente, um filtro de Bessel passa-baixas com frequência de corte ajustada para 75% do valor da taxa de dados.

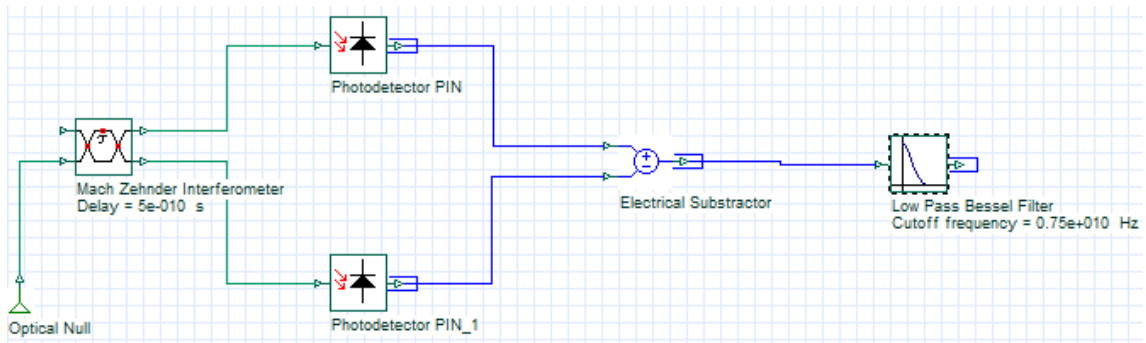


Figura 40 - Receptor NRZ-DPSK.

A seguir analisou-se o diagrama de olho do sinal DPSK para dois valores distintos de taxa de dados: 10 Gb/s (Figura 41 (a)) e 40 Gb/s (Figura 41 (b)). Observa-se, na figura, que houve uma pequena diminuição da abertura do olho no diagrama de olho quando a taxa foi quadruplicada. Vale ressaltar que o sinal que está fechando a abertura nos dois diagramas de olho é apenas um sinal espúrio que, sempre que se executa a simulação, aparece no *slot* do primeiro bit enviado, independentemente de esse bit ser “0” ou “1”.

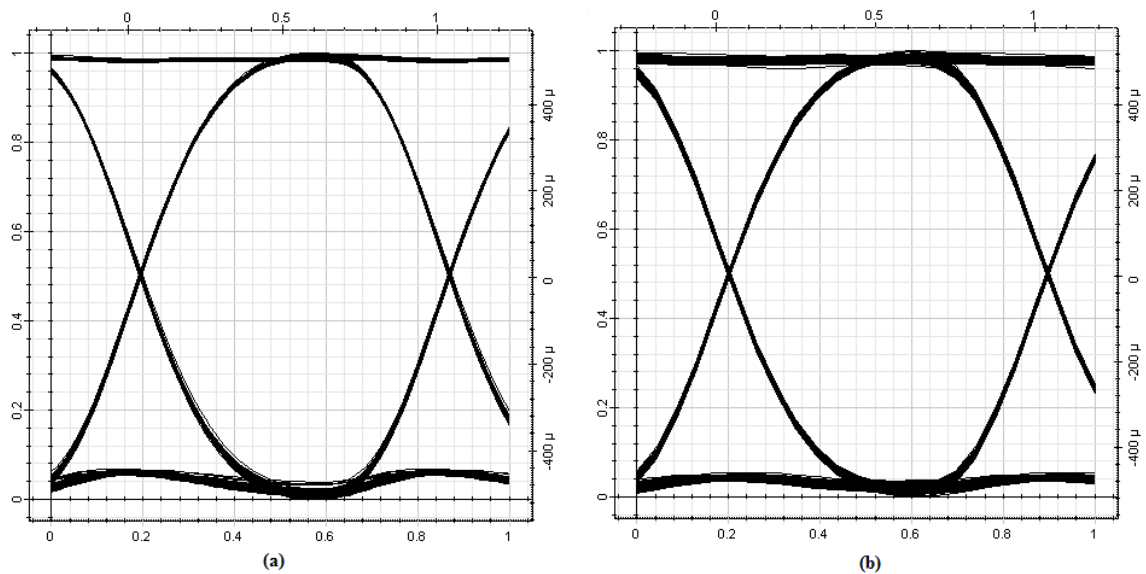


Figura 41 - Diagramas de olho de dois sinais DPSK: (a) um a 10 Gb/s e (b) outro a 40 Gb/s.

Os diagramas espectrais dos sinais elétricos já decodificados também podem ser conferidos na Figura 42. Em cada sinal, nesse diagrama, podem-se observar dois lóbulos, cada qual com largura de banda aproximadamente igual à taxa de dados, como era de esperar.

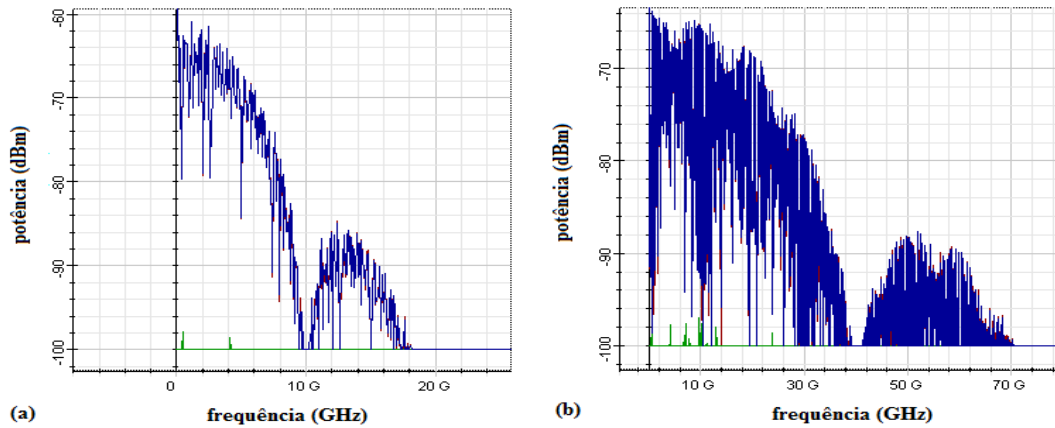


Figura 42 - Diagramas espectrais de dois sinais DPSK: (a) um a 10 Gb/s e (b) outro a 40 Gb/s.

As Figura 43 (taxa de 10 Gb/s) e Figura 44 (taxa de 40 Gb/s) comparam o sinal da sequência de bits enviados e o sinal óptico produzido pelo MZM. Notam-se as transições de fase que ocorrem sempre que um bit “1” é enviado, que são os “vales” de intensidade representando a modulação residual de intensidade do MZM. Outra característica de formatos de modulação de fase que também pode ser percebida nessas duas figuras é a presença potência óptica em todos os *slots* de bit.

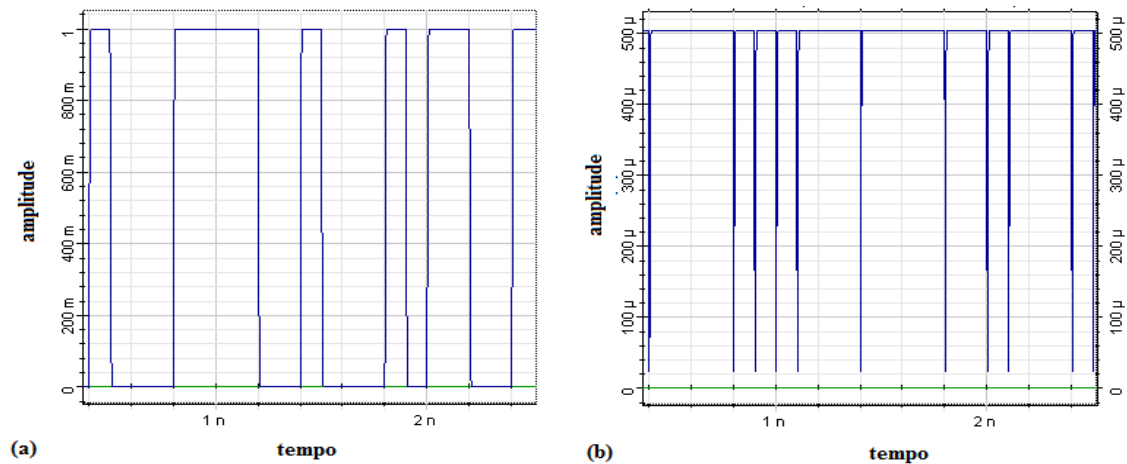


Figura 43 – Sequência de bits enviados a 10 Gb/s (a) e sinal óptico produzido pelo MZM (b).

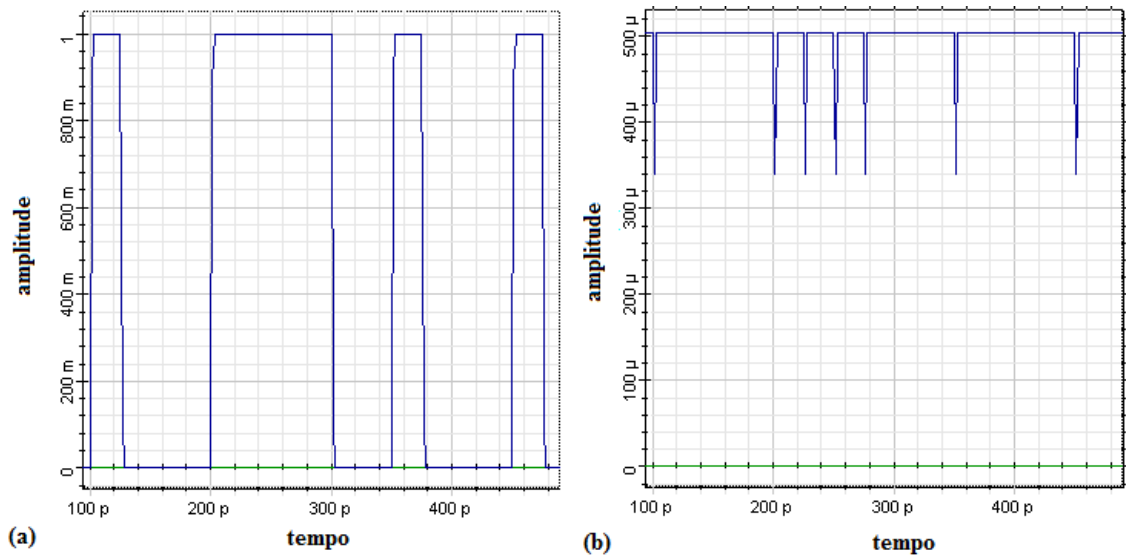


Figura 44 - Sequência de bits enviados a 40 Gb/s (a) e sinal óptico produzido pelo MZM (b).

Finalmente vê-se, nas Figura 45 (taxa de 10 Gb/s) e Figura 46 (taxa de 40 Gb/s), a diferença entre as formas de onda correspondentes ao sinal de dados ainda não codificados (Figura 45 (a) e Figura 46 (a)) e o sinal que é demodulado no receptor (Figura 45 (b) e Figura 46 (b)). Percebe-se que houve degradação da forma de onda, mas não a ponto de impossibilitar a transmissão da informação. Como dito inicialmente, os resultados apresentados são preliminares. As configurações das paletas de simulação, vistas nas Figura 39 e Figura 40, precisam ser aprimoradas para que as formas de onda, vistas através dos visualizadores (no tempo e na frequência), sejam mais bem representadas. Por exemplo, o filtro no receptor deveria ser otimizado de forma a reduzir o ruído observado na Figura 42 sem, contudo, alterar as formas de onda de saída, nas Figura 45 (b) e Figura 46 (b).

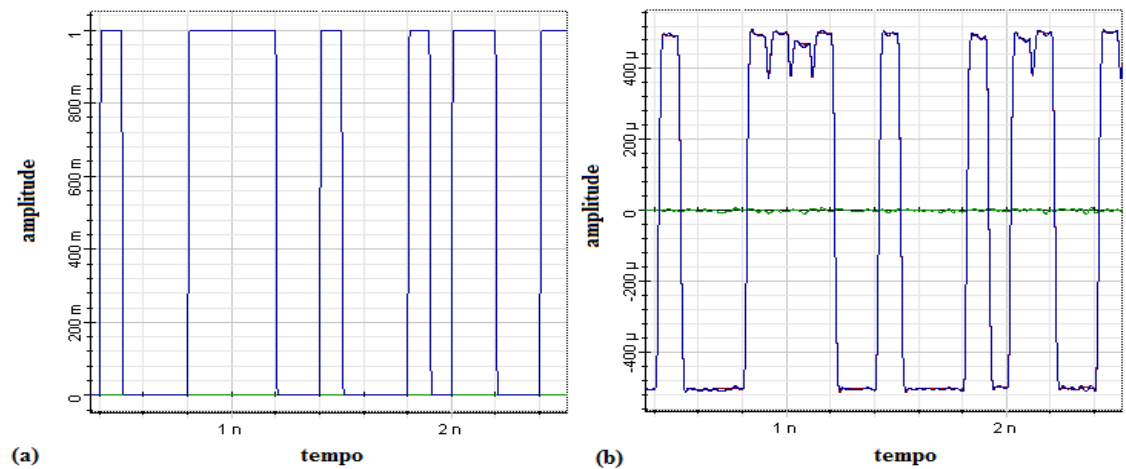


Figura 45 - Formas de onda de entrada (a) e saída (b) do sinal DPSK a 10 Gb/s.

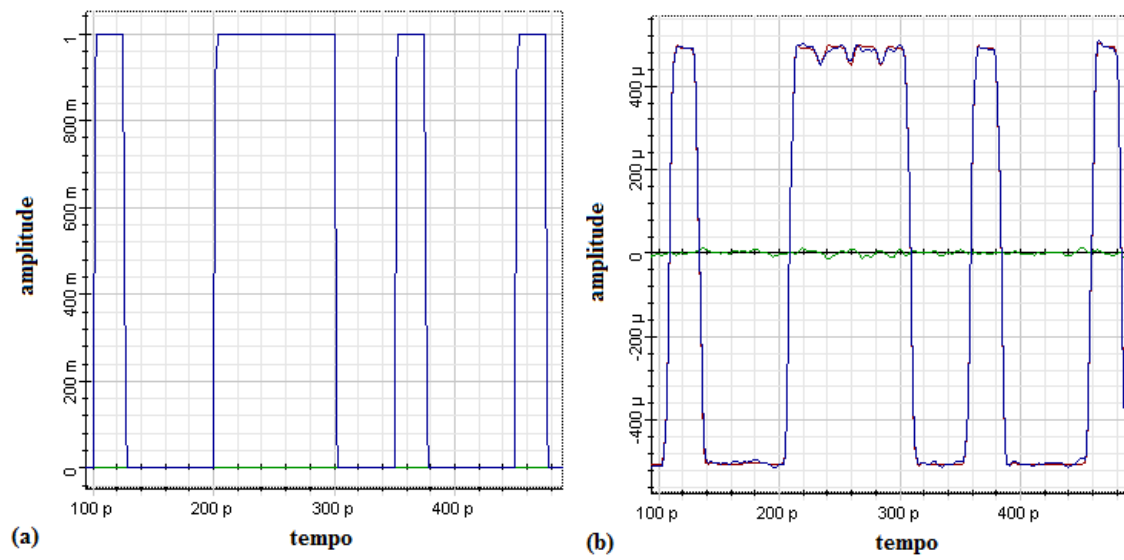


Figura 46 - Formas de onda de entrada (a) e saída (b) do sinal DPSK a 40 Gb/s.

9. Conclusão

Neste trabalho foram abordados os subsistemas que compõem um sistema de comunicação óptica e os principais conceitos envolvidos na transmissão, propagação e recepção do sinal óptico. Foram descritos os principais fatores que degradam o sinal durante sua propagação em fibras ópticas. Conforme mostrado, a transmissão nos sistemas de nova geração impõe complexos desafios porque os efeitos físicos que degradam o sinal durante a propagação na fibra são muito mais críticos para taxas de transmissão a partir de 40 Gb/s por canal do que são para os da geração atual de 10 Gb/s [2].

Também foram expostos os formatos de modulação que mais estão sendo discutidos pela comunidade científica para aumentar a capacidade de transmissão de dados e atender a crescente demanda por banda do mercado de telecomunicações. Foram descritas as tecnologias e estruturas utilizadas na geração de cada formato e apresentou-se a classificação, as vantagens e as desvantagens de cada um deles. Por exemplo, percebe-se que existe uma relação de compromisso entre a tolerância à dispersão cromática e aos efeitos não-lineares. Os formatos de modulação que têm espectro mais largo são pouco tolerantes à dispersão, mas bem resistentes a efeitos não lineares. Por outro lado, os tipos de modulação com estreita largura espectral são mais tolerantes à dispersão, mas pouco resistentes aos efeitos não lineares. Os formatos modulados na fase mostram ser tolerantes a ambos os efeitos.

Utilizando o software comercial *Optisystem* (versão 7.0), da empresa canadense *Optiwave Systems Inc.*, foram feitas simulações preliminares para o formato DPSK, na configuração *back-to-back*, que mostraram o diagrama de olho, a forma de onda e o espectro óptico do sinal DPSK transmitido a 10 Gb/s e a 40 Gb/s. Para uma análise sistêmica mais completa, ajustes nas paletas de simulação bem como simulações com enlaces de fibra deverão ser realizados.

Os formatos avançados de modulação são ferramentas essenciais para habilitar os sistemas de nova geração, com taxas de transmissão a partir de 40 Gb/s por canal. Um dos grandes desafios a ser superado é a adequação dos sistemas dessa nova geração à infraestrutura atualmente existente, que foi instalada e otimizada para trabalhar com transmissão de taxas de até 10 Gb/s por canal.

Referências bibliográficas

- [1] P. J. Winzer, R. -J. Essiambre, “Advanced Optical Modulation Formats”, *Proceedings of the IEEE*, Vol. 94, N° 5, pp. 952-985, 2006.
- [2] C. R. da Silveira, “Estudo de formatos especiais de modulação digital para comunicações ópticas”, dissertação de mestrado, Universidade de São Paulo, Departamento de Engenharia Elétrica da Escola de Engenharia de São Carlos, 2009.
- [3] G. P. Agrawal, *Fiber-Optic Communication Systems*, John Wiley & Sons, 2002.
- [4] Nick Massa, *Fiber Optic Telecommunication*, Springfield Technical Community College, Springfield, Massachusetts.
- [5] I. P. Kaminow, Tingye Li, A. E. Willner, *Optical fiber telecommunications V*, 2008.
- [6] Notas de aula do Prof. Dr. Ben-Hur Viana Borges, Departamento de Engenharia Elétrica da Escola de Engenharia de São Carlos, USP.
- [7] A. H. Gnauck, and P. J. Winzer, “Optical Phase-Shift-Keyed Transmission”, *Journal of Lightwave Technology*, Vol. 23, N° 1, 2005.
- [8] P. J. Winzer, C. Dorrer, R.-J. Essiambre, I. Kang, “Chirped Return-to-Zero Modulation by Imbalanced Pulse Carver Driving Signals”, *IEEE Photonics Technology Letters*, Vol. 16, N° 5, maio de 2004.
- [9] Notas de aula do Prof. Dr. Murilo Araújo Romero e do Prof. José Carlos Sartori, Departamento de Engenharia Elétrica da Escola de Engenharia de São Carlos, USP.