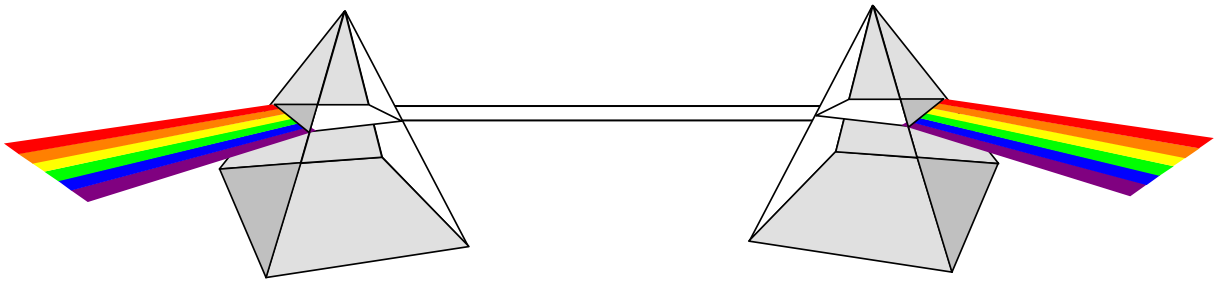
Simulation of Optical Communication Systems

Session 2: External Modulation



Copyright notice

Wired Communications Group (Fachgebiet Leitungsgebundene Übertragungstechnik)
Institute for Communications Engineering (LNT)
Technische Universität München (TUM)
D-80290 München
Germany
Phone: +49 89 23472

Contents

Theoretical Background

As we have seen in the first session of this simulation lab, it is difficult to generate intensity modulation by direct modulation of a semiconductor laser for high-speed communications (e.g. > 5 Gbit/s) due to the limited bandwidth of direct modulation. This problem can be solved if an additional optical component is introduced. The laser is then operated with a constant injection current, simply generating a constant light carrier. This carrier is modulated in a new device, the so-called *external modulator*. The most frequently used external modulator is applying a Mach-Zehnder interferometer structure, and is therefore called Mach-Zehnder Modulator (MZM). This device is investigated in the following.

1. Mach-Zehnder Modulators

1.1. Physical Background and Mathematical Modeling

In a MZM the incoming non-modulated carrier according to $\underline{E}_{in}(t) = \hat{\underline{E}}_{in} \cdot e^{j\omega_c t}$ is first split up into two optical paths (see Fig. 1).

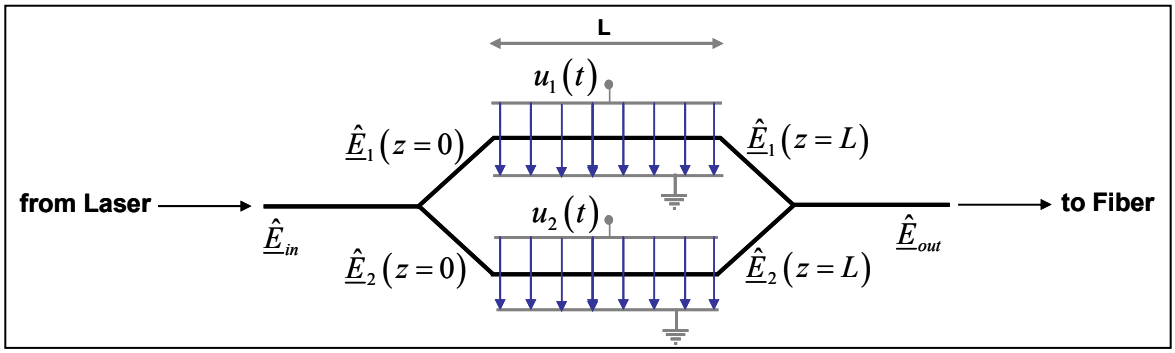


Figure 1: General structure of a Mach-Zehnder Modulator (dual-drive configuration)

The incoming light power is split up in the left Y-branch according to the power splitting ratios κ_1^2 and $1-\kappa_1^2$, i.e. $\hat{\underline{E}}_1(z=0) = \kappa_1 \cdot \hat{\underline{E}}_{in}$, $\hat{\underline{E}}_2(z=0) = \sqrt{1-\kappa_1^2} \cdot \hat{\underline{E}}_{in}$. The waveguide is made of a special material showing a strong electro-optic effect, typically Lithium-Niobate ($LiNbO_3$). Due to this electro-optical effect the refractive index of $LiNbO_3$ changes with an externally applied electrical voltage *u(t)* according to

$$n(t) = n_0 + \varsigma_{EO} \cdot u(t). \qquad (2.1)$$

Here n_0 is the refractive index without external field, ς_{EO} models the strength of the electro-optical effect. As a consequence, the phase velocity of a propagating electrical wave is modified as $c_n = c/n$, and thus the wave number k changes according to $k_n = (\omega/c) \cdot n$, where c is the speed of light in vacuum. Therefore the phase shift of the waves propagating in branch 1 and 2 of the waveguide can be tuned by the applied voltages $u_1(t)$ and $u_2(t)$, respectively, resulting in

$$\hat{\underline{E}}_1(z=L) = \hat{\underline{E}}_1(z=0) \cdot e^{-j\Delta\varphi_1} \quad \text{and} \quad \hat{\underline{E}}_2(z=L) = \hat{\underline{E}}_2(z=0) \cdot e^{-j\Delta\varphi_2}, \qquad (2.2)$$

with $\Delta\varphi_1 = k_n(u_1)\cdot L$ and $\Delta\varphi_2 = k_n(u_2)\cdot L$, respectively. The attenuation of the waveguides is neglected for the sake of simplicity. $\hat{\underline{E}}_1$ and $\hat{\underline{E}}_2$ are then combined at the output of the device according to

$$\hat{\underline{E}}_{out} = \kappa_2 \cdot \hat{\underline{E}}_1(z=L) + \sqrt{1-\kappa_2^2}\cdot \hat{\underline{E}}_2(z=L). \tag{2.3}$$

κ_2^2 and $1-\kappa_2^2$ are the power splitting ratios of the Y-branch on the right-hand side in Fig. 1. This results finally in the output field $\underline{E}_{out}(t)$ according to

$$\underline{E}_{out}(t) = \hat{\underline{E}}_{in}\cdot\left[\kappa_1\kappa_2\cdot e^{-j\Delta\varphi_1} + \sqrt{1-\kappa_1^2}\sqrt{1-\kappa_2^2}\cdot e^{-j\Delta\varphi_2}\right]\cdot e^{j\omega_c t}. \tag{2.4}$$

Under the condition of symmetrical splitting, i.e. $\kappa_1^2 = \kappa_2^2 = 1/2$, we get

$$\underline{E}_{out}(t) = \hat{\underline{E}}_{in}\cdot\frac{1}{2}\left[e^{-j\Delta\varphi_1} + e^{-j\Delta\varphi_2}\right]\cdot e^{j\omega_c t}. \tag{2.5}$$

It can be seen from (2.5) that, depending on the relative phases $\Delta\varphi_1$ and $\Delta\varphi_2$, neither amplitude nor phase of the outcoming field are modified. This becomes more obvious after a simple modification of (2.5), resulting in

$$\begin{aligned}\underline{E}_{out}(t) &= \hat{\underline{E}}_{in}\cdot\frac{1}{2}\left[e^{-j\frac{\Delta\varphi_1-\Delta\varphi_2}{2}} + e^{j\frac{\Delta\varphi_1-\Delta\varphi_2}{2}}\right]\cdot e^{-j\frac{\Delta\varphi_1+\Delta\varphi_2}{2}}\cdot e^{j\omega_c t} = \\ &= \hat{\underline{E}}_{in}\cdot\cos\left(\frac{\Delta\varphi_1-\Delta\varphi_2}{2}\right)\cdot e^{-j\frac{\Delta\varphi_1+\Delta\varphi_2}{2}}\cdot e^{j\omega_c t}.\end{aligned} \tag{2.6}$$

There is a linear relation between the phase shift difference $\Delta\varphi_1$ - $\Delta\varphi_2$ and the voltage difference u_1-u_2. Also, it is important to notice that these quantities are usually time-dependent. This leads us to

$$\Delta\varphi_1(t) - \Delta\varphi_2(t) = \pi\cdot\frac{u_1(t)-u_2(t)}{V_\pi} = \frac{\pi}{V_\pi}\cdot u(t). \tag{2.7}$$

The parameter V_π denotes the voltage difference $u(t) = u_1(t)-u_2(t)$ that generates a relative phase shift of π between upper and lower branch of the MZM. The phase term in (2.6) is now re-written as

$$e^{-j\frac{\Delta\varphi_1(t)+\Delta\varphi_2(t)}{2}} = e^{-j\Theta\cdot\frac{\pi}{V_\pi}\cdot\left(\frac{u_1(t)-u_2(t)}{2}\right)},\ \text{where}\ \Theta = \frac{\Delta\varphi_1(t)+\Delta\varphi_2(t)}{\Delta\varphi_1(t)-\Delta\varphi_2(t)} = \frac{u_1(t)+u_2(t)}{u_1(t)-u_2(t)}. \tag{2.8}$$

Substituting (2.7) and (2.8) into (2.6) finally results in

$$\boxed{\underline{E}_{out}(t) = \hat{\underline{E}}_{in}\cdot\cos\left(\frac{\pi}{V_\pi}\cdot\frac{u_1(t)-u_2(t)}{2}\right)\cdot e^{-j\Theta\cdot\frac{\pi}{V_\pi}\cdot\left(\frac{u_1(t)-u_2(t)}{2}\right)}\cdot e^{j\omega_c t}.} \tag{2.9}$$

The parameter Θ is called the *asymmetry factor* of the dual-drive MZM. To make Θ a time-independent constant, $u_1(t)$ and $u_2(t)$ can no longer be selected independently, but are constrained by

$$u_1(t) = \frac{u(t)}{2} \cdot (\Theta + 1), \quad u_2(t) = \frac{u(t)}{2} \cdot (\Theta - 1), \tag{2.10}$$

with the definition: $-1 \le \Theta \le 1$.

The way the selection of the asymmetry factor Θ acts on u_1 and u_2 is depicted in Figure 2.

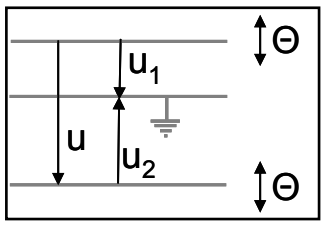


Figure 2: Effect of the asymmetry factor Θ

As we can see from (2.9), if **Θ = 0** is selected, i.e. $u_1(t) = u(t)/2$, $u_2(t) = -u(t)/2$, the modulator generates pure amplitude (or intensity) modulation without any time-dependent modification of the signal phase. If asymmetry is introduced by selecting **Θ ≠ 0**, the intensity modulation is not modified; however, an additional phase modulation is generated. This parasitical, time-dependent phase modulation modifies the instantaneous frequency of the light wave according to

$$\omega(t) = \omega_c - \frac{\Theta}{2} \cdot \frac{\pi}{V_\pi} \cdot \frac{du(t)}{dt}. \tag{2.11}$$

As we know from Session 1 of this course, this frequency chirp, i.e. the deviation $\Delta f(t) = \Delta\omega(t) / 2\pi$ of the instantaneous signal from the carrier frequency, has a negative influence on the propagating signal due to the chromatic dispersion of the fiber. Therefore, strong efforts are made generally to reduce this chirp to negligible values. As we could see above, selecting **Θ = 0**, i.e. $u_2(t) = -u_1(t)$, results in chirp-free intensity modulation. This operation of the dual-drive MZM is called *push-pull mode*.

Another solution is a re-design of the Mach-Zehnder modulator: If the poles of the applied voltage are changed according to Figure 3, the condition $u_2(t) = -u_1(t)$ is met automatically. In this *single-drive configuration* the voltage difference $u(t)$ is directly impressed on the modulator. In this session, we will only be using dual-drive MZMs.

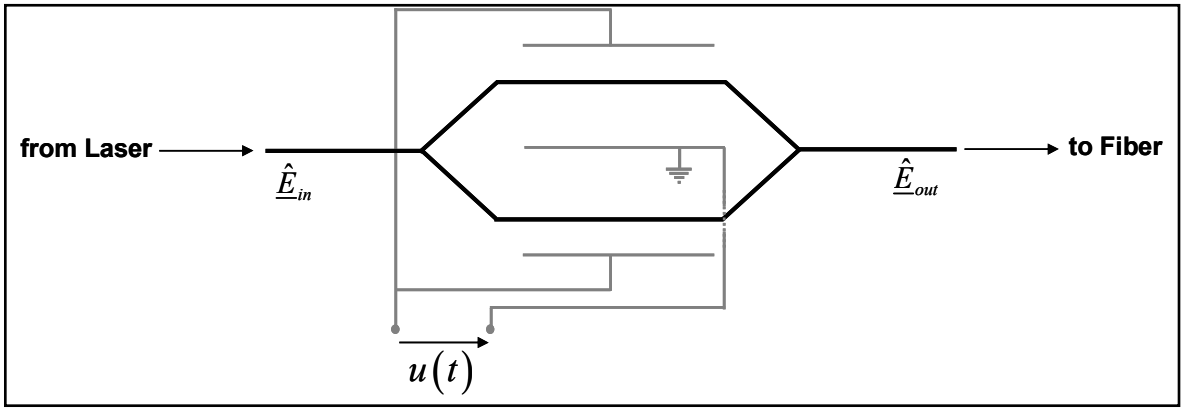


Figure 3: Single-drive Mach-Zehnder Modulator

If the dual-drive MZM is operated in chirp-free mode, the electrical output field is simply given as

$$\underline{E}_{out}(t) = \hat{\underline{E}}_{in} \cdot \cos\left(\frac{\pi}{V_\pi} \cdot \frac{u(t)}{2}\right) \cdot e^{j\omega_c t} = \hat{\underline{E}}_{out}(t) \cdot e^{j\omega_c t}. \qquad (2.12)$$

From (2.12) it can be seen that the amplitude $\hat{\underline{E}}_{out}(t)$ depends on the input voltage *u(t)* according to a cosine function. The power *P(t)* of the modulated carrier is therefore calculated as

$$P(t) \sim \left|\hat{\underline{E}}_{out}(t)\right|^2 = \left|\hat{\underline{E}}_{in}(t)\right|^2 \cdot \left[\cos\left(\frac{\pi}{V_\pi} \cdot \frac{u(t)}{2}\right)\right]^2. \qquad (2.13)$$

The curves relating $\hat{\underline{E}}_{out}$ and *P* with *u* are depicted in Figures 4 and 5.

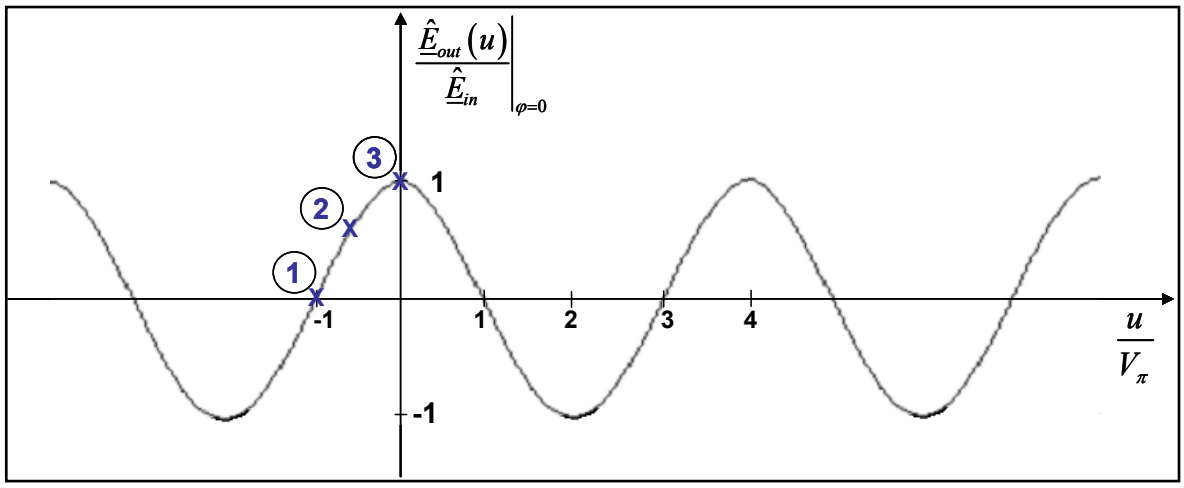


Figure 4: $\hat{\underline{E}}_{out}$ (normalized) as a function of input voltage *u*

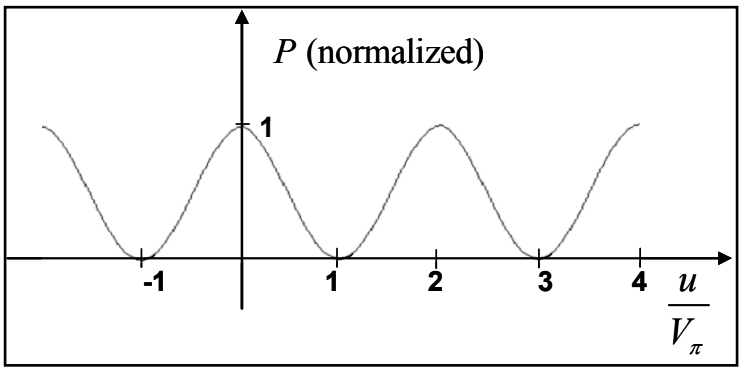


Figure 5: *P* (normalized) as a function of input voltage *u*

The operation point of the MZM is defined by applying a specific bias voltage. The points 1 (*trough*), 2 (*quadrature*) and 3 (*crest*) marked in Figure 4 are used to generate specific optical modulation schemes. Of course any other adequate operation point could be selected due to the periodicity of the curves.

1.2. Pulse shaping and Modulation

Despite their simple working principle, Mach-Zehnder modulators are extremely versatile and can be used to create many different pulse shapes and modulation formats at very high bit rates. The curves depicted in Figures 4 and 5 are the key to creating the various transmitter configurations.

In this lab session, we will concentrate on the pulse shapes *Non-Return-to-Zero* (*NRZ*), *Return-to-Zero* (*RZ*) with various duty cycles and *Carrier-Suppressed Return-to-Zero* (*CSRZ*). As a modulation format, we will consider *On-Off-Keying* (*OOK*). However, if you have some time left at the end of the session, you can try and use Mach-Zehnder Modulators to set up a DPSK transmitter.[1] Lastly, we are going to experiment with a *duobinary* transmitter. Although duobinary is commonly referred to as a modulation scheme, strictly speaking, it is achieved by a combination of a logical precoder (so-called *line coding*) with Amplitude Shift Keying (ASK).[2]

1.2.1. NRZ

A single MZM can be used to produce an NRZ OOK signal. The MZM is biased in the quadrature point, i.e. $u_{Bias} = -V_\pi / 2$. The *driving* or *modulation voltage* u_{mod} has a range of V_π and produces NRZ pulses corresponding to the binary data stream.[3] Table 1 summarizes these parameters. In Figure 6, the modulation voltage is depicted along the time axis (which is drawn shifted by u_{Bias}) for the bit sequence …`010011`….

[1] There is a multitude of other pulse shapes (e.g. Chirped RZ – CRZ pulses) and modulation schemes (e.g. multilevel ASK, DPSK, DQPSK, D8-PSK, as well as combinations of ASK and PSK, etc.), whose transmission properties are in the center of research attention. Some of these advanced transmission systems are already deployed in undersea optical communication links. The transmitter structures covered in this lab session, however, are still by far the most popular in current systems.

[2] Line coding (for which AMI codes are another example) is a technique used to equip the signal with certain desired properties in the time or frequency domain (*spectral shaping*). It must be distinguished from channel coding, where redundancy is added to the signal in order to combat transmission errors at the receiver.

[3] u_{Bias} can be regarded as the constant offset (bias) of *u(t)*. Hence $u(t) = u_{Bias} + u_{\mathrm{mod}}(t)$.

Table 1: Voltage Parameters of a MZM producing OOK NRZ modulation

u_{Bias}	$-V_\pi / 2$ (quadrature)
u_{mod} function	NRZ
u_{mod} frequency	(bit rate)
u_{mod} range	$\left[\frac{-V_\pi}{2};\frac{V_\pi}{2}\right]$

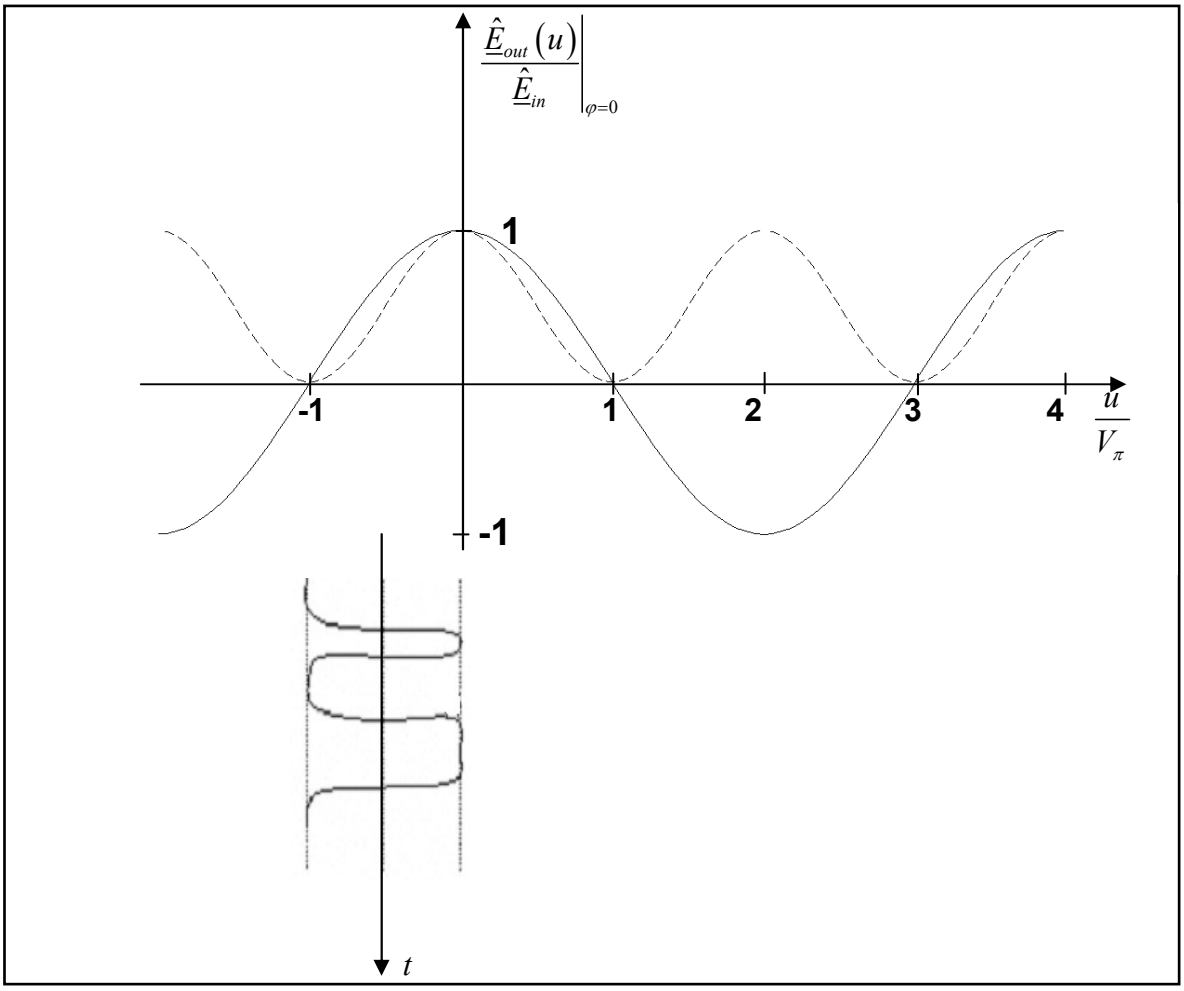


Figure 6: Generation of OOK NRZ modulation (*E* (solid) and *P* (dashed) normalized)

1.2.2. RZ

Compared to the simpler NRZ pulses, *Return-to-Zero* (RZ) pulses have several advantages:

- In long sequences of *ones*, RZ pulses still allow the receiver to recover the signal clock.
- For the same pulse energy, the peak power is higher in RZ than in NRZ pulses, resulting in a higher SNR at the sampling instant.
- Depending on the system configuration, RZ pulses can yield a better system performance than NRZ pulses. It is, however, difficult to make general statements about the performance gain of RZ over NRZ, as this gain depends on many system parameters such as signal power, filter bandwidths, fiber nonlinearities, dispersion

map and data rate. As a rough rule, it can be said that RZ plays out its advantages with increasing data rates and signal powers (i.e. higher fiber nonlinearities) [5].

These advantages, however, come at the cost of an increased transmitter complexity. To generate a signal with RZ pulses, a setup employing two MZMs is used (Figure 7). The first MZM, called the *pulse carver* (PC), is generating a train of identical pulses, which serve as an input to the second MZM, the *data* or *gating modulator* (DM). The purpose of the DM, which is configured as an NRZ OOK modulator as introduced in the previous section, is simply to block or let pass single pulses depending on the data to be transmitted.

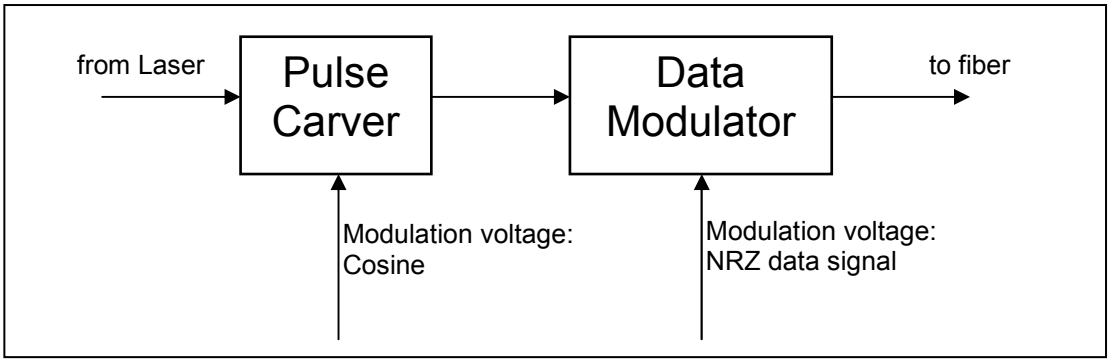


Figure 7: Setup of RZ transmitter

The pulse carver is driven by a cosine function. Depending on the bias voltage and the modulation voltage's range and frequency, different types of RZ pulses can be generated. RZ pulses are characterized by their *duty cycle*, which is the pulse's full width at half intensity relative to the bit slot duration. *RZ33* denotes an RZ pulse with 33% duty cycle; *RZ50* has 50% duty cycle. In contrast to these two, *RZ67* pulses have an alternating field sign. They are covered in the next section.

RZ50 Pulse Generation

To produce RZ pulses with 50% duty cycle, the PC is driven with the parameters given in Table 2.

Table 2: Voltage Parameters of a MZM producing RZ 50 pulses

u_{Bias}	$\pm V_\pi / 2$ (quadrature)
u_{mod} function	Cosine
u_{mod} frequency	Bit frequency (e.g. 10 GHz for a 10 Gbit/s signal)
u_{mod} range	$\left[\frac{-V_\pi}{2};\frac{V_\pi}{2}\right]$

Question

What are A and B in $u_{\mathrm{mod}} = A \cdot \cos\left(B \cdot t\right)$ if we want to produce RZ50 pulses for a 10Gbit/s signal? Sketch the modulation voltage versus time in Figure 8.

$$u(t) = u_{bias} + u_{\mathrm{mod}}(t), \text{ where } u_{bias} = V_\pi / 2 \text{ and } u_{\mathrm{mod}}(t) = \underbrace{V_\pi / 2}_{A} \cdot \cos\left(\underbrace{2\pi \cdot 10\ \mathrm{GHz}}_{B} \cdot t \right)$$

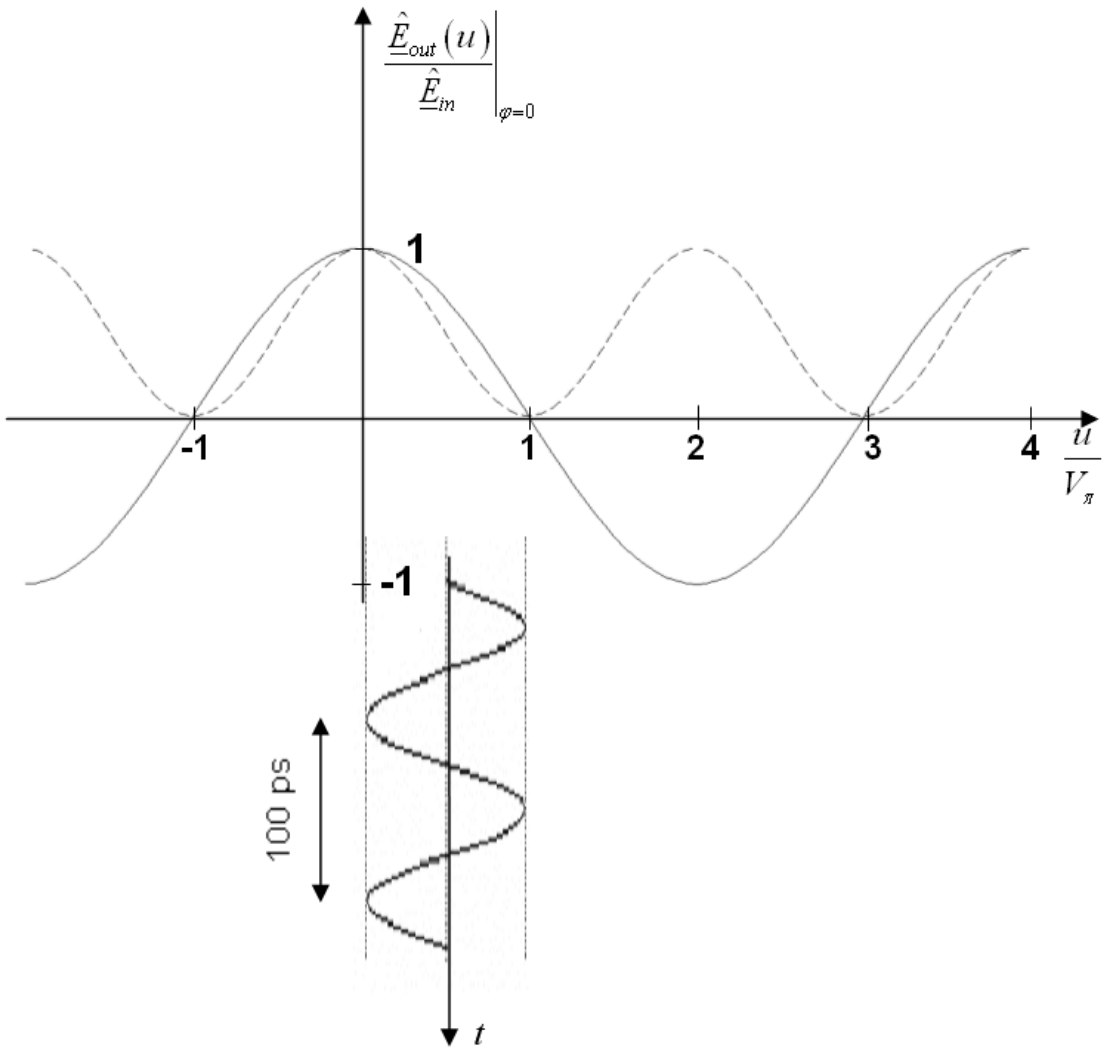


Figure 8: Generation of RZ50 pulses (*E* (solid) and *P* (dashed) normalized)

RZ33 Pulse Generation

RZ33 pulses are even narrower in the time domain than RZ50 pulses and accordingly have a broader spectrum. They can be generated using the parameters given in Table 3.

Table 3: Voltage Parameters of a MZM producing RZ 33 pulses

u_{Bias}	0 (crest)
u_{mod} function	Cosine
u_{mod} frequency	Half the bit frequency (e.g. 5 GHz for a 10 Gbit/s signal)
u_{mod} range	$[-V_\pi; V_\pi]$

Question

Calculate an expression for the pulse train function (i.e. the output power of the PC as a function of time) by inserting $u(t) = u_{Bias} + u_{\mathrm{mod}}(t)$ as given in Table 3 into Equation (2.13). Show that the duty cycle of these pulses is 33% (=1/3).

$u(t) = V_\pi \cdot \cos\left(2\pi \cdot B/2 \cdot t\right)$, with *B*: Bit frequency, *1/B*: Bit duration

Substituting into (2.13) yields

$$P_{\mathrm{normalized}} = \cos^2\left(\frac{\pi}{V_\pi} \cdot \frac{u(t)}{2}\right) = \cos^2\left(\frac{\pi}{2} \cdot \cos\left(2\pi \frac{B}{2} t\right)\right) \overset{!}{=} \frac{1}{2}$$

$$\Rightarrow \frac{\pi}{2} \cdot \cos\left(2\pi \frac{B}{2} t_0\right) = \arccos\frac{1}{\sqrt{2}} = \frac{\pi}{2}$$

$$\Rightarrow \cos\left(\pi \cdot B \cdot t_0\right) = \frac{1}{2}$$

$$\Rightarrow \pi \cdot B \cdot t_0 = \frac{\pi}{3}$$

$$\Rightarrow t_0 = \frac{1}{3} \cdot \frac{1}{B}$$

At t_0, the pulse has reached half the maximum intensity. Because the pulse is symmetrical, $(1/B - T_{\mathrm{FWHM}}) = 2 \cdot t_0$.

Hence, $T_{\mathrm{FWHM}} = 1/B - 2 \cdot t_0 = \frac{1}{3} \cdot \frac{1}{B}$.

The duty cycle is 33%.

1.2.3. CSRZ

At V_π, the power P is minimum and the electrical field changes its sign. When the bias voltage is set to this point, a pulse carver driven by a cosine function will produce pulses that have alternating field signs. Therefore, the generated signal is DC-free (in the electric field domain). In the frequency spectrum, this property corresponds to an absence of discrete spectral components at the carrier frequency. For that reason, these pulses are called *Carrier-Suppressed RZ* (CSRZ). The parameters used to create CSRZ pulses are given in Table 4. The duty cycle of CSRZ pulses is 67%, so they are also called *RZ67*.

Table 4: Voltage Parameters of a MZM producing CSRZ pulses

u_{Bias}	V_π (trough)
u_{mod} function	Cosine
u_{mod} frequency	Half the bit frequency (e.g. 5 GHz for a 10 Gbit/s signal)
u_{mod} range	$[-V_\pi; V_\pi]$

CSRZ pulses have the advantage of being quite robust against ISI caused by dispersion. To understand why, imagine two neighboring pulses spreading in time such that they overlap partly. Adjacent CSRZ pulses have contrary signs, so that ideally, they completely cancel. In the lab, this can be observed from a larger eye opening compared to other RZ pulses.

1.2.4. Duobinary

Duobinary modulation is, strictly speaking, not a modulation format, but rather 3-level Amplitude Shift Keying (ASK) with a logical precoder. The values transmitted are $P = 0$ (i.e. $u = V_\pi$) for a logical zero, and $P_{normalized} = 1$ (i.e. $u = 0$ or $u = 2V_\pi$) for a logical one. The data modulator is therefore driven by a three-level ($u = 0, V_\pi, 2V_\pi$) electrical NRZ signal. The transformation of the binary data sequence to a ternary sequence is achieved by a logical precoder, whose coding rule is as follows:

- A logical *zero* is always coded as `0`.
- A logical *one* is coded as `+1` or `-1`. If an odd number of *zeros* is transmitted, the sign of the following *one* is changed.

The block diagram of a duobinary coder is depicted in Figure 9.

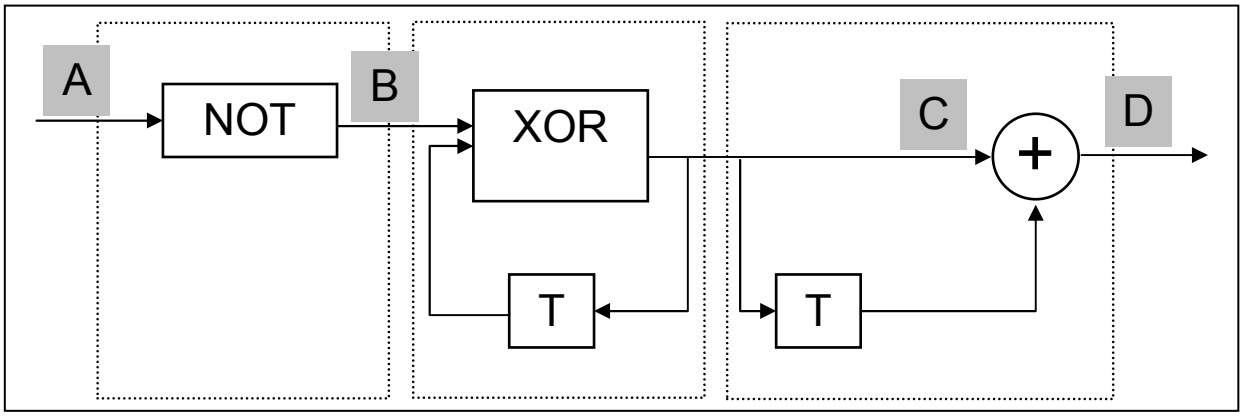


Figure 9: Block diagram of duobinary precoder

For the system in Figure 9 to make sense, two conventions are important:

- The output of the XOR element is bipolar, i.e. a `1` is output as `+1`, a `0` as `−1`.
- The output of the total system will have an absolute amplitude of 2, i.e. to obtain the duobinary code described above, we would have to divide by 2.

The block diagram in Figure 9, however, illustrates the way duobinary signals are generated in *OptiSystem*. The first dotted box corresponds to the OptiSystem element *Binary NOT*, the second to the element *Duobinary Precoder*, and the third box corresponds to the *Duobinary Pulse Generator*.
The example given in Table 5 will help clarify the principle.

Table 5: Example of duobinary line coding

A	0	1	0	1	1	1	1	0	0	1	1	0	0	0	1	1	0	1
B	1	0	1	0	0	0	0	1	1	0	0	1	1	1	0	0	1	0
C	1	1	-1	-1	-1	-1	-1	1	-1	-1	-1	1	-1	1	1	1	-1	-1
D	0	+2	0	-2	-2	-2	-2	0	0	-2	-2	0	0	0	+2	+2	0	-2

The main advantage of duobinary modulation is its dispersion tolerance. In particular, isolated *zeros* (such as in …101…) cause the sign of the *ones* to change. Because of the opposed field signs, the neighboring pulses ideally cancel when leaking into the *zero* bit slot.
Duobinary coding is usually used with NRZ pulses. Therefore, only one MZM is required at the transmitter. Because of the broad NRZ pulses used, duobinary signals have a narrow spectrum suitable for WDM systems with a small channel separation, which is another advantage.

Question

What bias voltage u_{Bias} do you choose for duobinary modulation? Which values does u_{mod} take (compare with Figure 6)?

$u_{\mathrm{Bias}} = V_\pi$ (trough)

$u_{\mathrm{mod}} \in \{-V_\pi, 0, V_\pi\}$

Bibliography

[1] N. Hanik; Optical Communication Systems, Lecture Notes at Technische Universität München (TUM), 2008
[2] G.P. Agrawal; Fiber-Optic Communication Systems, Third Edition, John Wiley & Sons, New York, 2002
[3] E. Voges, K. Petermann (eds.); Optische Kommunikationstechnik, Springer, Berlin, 2002
[4] I. Kaminov, T. Li (eds.), Optical Fiber Telecommunications IV, Academic Press, San Diego, 2002
[5] P.J. Winzer, R.-J. Essiambre, Advanced Modulation Formats, Proceedings of the IEEE, Vol. 94, No. 5, pp. 952-985, May 2006