Chapter 2
Coherent Optical Communications: Historical Perspectives and Future Directions

Kazuro Kikuchi

Abstract Coherent optical fiber communications were studied extensively in the 1980s mainly because high sensitivity of coherent receivers could elongate the unrepeated transmission distance; however, their research and development have been interrupted for nearly 20 years behind the rapid progress in high-capacity wavelength-division multiplexed (WDM) systems using erbium-doped fiber amplifiers (EDFAs). In 2005, the demonstration of digital carrier phase estimation in coherent receivers has stimulated a widespread interest in coherent optical communications again. This is due to the fact that the digital coherent receiver enables us to employ a variety of spectrally efficient modulation formats such as $M$-ary phase-shift keying (PSK) and quadrature amplitude modulation (QAM) without relying upon a rather complicated optical phase-locked loop. In addition, since the phase information is preserved after detection, we can realize electrical post-processing functions such as compensation for chromatic dispersion and polarization-mode dispersion in the digital domain. These advantages of the born-again coherent receiver have enormous potential for innovating existing optical communication systems. In this chapter, after reviewing the 20-year history of coherent optical communication systems, we describe the principle of operation of coherent detection, the concept of the digital coherent receiver, and its performance evaluation. Finally, challenges for the future are summarized.

2.1 History of Coherent Optical Communications

This section describes the history of coherent optical communications. Technical details will be summarized in Sect. 2.2, and readers should refer to them appropriately.

K. Kikuchi
Department of Electrical Engineering and Information Systems, University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-8656, Japan
e-mail: kikuchi@ginjo.t.u-tokyo.ac.jp
2.1.1 Coherent Optical Communication Systems 20 Years Ago

The research and development in optical fiber communication systems started in the first half of the 1970s. Such systems used intensity modulation of semiconductor lasers, and the optical signal intensity transmitted through an optical fiber was detected by a photodiode, which acted as a square-law detector. This combination of the transmitter and the receiver is called the intensity modulation and direct detection (IMDD) scheme, which has been commonly employed in optical communication systems up to the present date. Such IMDD scheme has a great advantage that the receiver sensitivity is independent of the carrier phase and the state of polarization (SOP) of the incoming signal, which are randomly fluctuating in real systems.

Although the first proposal of coherent optical communications using heterodyne detection was done by DeLange in 1970 [1], it did not attract any attention because the IMDD scheme became mainstream in optical fiber communication systems during the 1970s. On the other hand, in 1980, Okoshi and Kikuchi [2] and Fabre and LeGuén [3] independently demonstrated precise frequency stabilization of semiconductor lasers, which aimed at optical heterodyne detection for optical fiber communications. Figure 2.1 shows the heterodyne receiver proposed in [2] for frequency-division multiplexed (FDM) optical communication systems. Each FDM channel was selected by heterodyne detection with multiple local oscillators (LOs) prepared in the receiver.

Table 2.1 shows the comparison between coherent and IMDD schemes. With coherent receivers, we can restore full information on optical carriers, namely in-phase and quadrature components (or amplitude and phase) of the complex amplitude of the optical electric field and the state of polarization (SOP) of the signal. In exchange for such advantages, coherent receivers are sensitive to the phase and SOP of the incoming signal. To cope with this problem, the configuration of coherent systems becomes much more complicated than that of IMDD systems.

Fig. 2.1 Heterodyne receiver for FDM optical communication systems [2]. LO: local oscillator, PD: photodiode, IFA: intermediate-frequency amplifier, and Det: detector

<table>
<thead>
<tr>
<th>Table 2.1 Comparison between coherent and IMDD schemes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coherent</strong></td>
</tr>
<tr>
<td>Modulation parameters</td>
</tr>
<tr>
<td>Detection method</td>
</tr>
<tr>
<td>Adaptive control</td>
</tr>
</tbody>
</table>
The center frequency drift of semiconductor lasers for a transmitter and a local oscillator could be maintained below 10 MHz as shown in [2, 3]. Even when the frequency drift of the transmitter laser was suppressed, the carrier phase was fluctuating randomly due to large phase noise of semiconductor lasers. Therefore, narrowing spectral linewidths of semiconductor lasers was the crucial issue for realizing stable heterodyne detection. The method of measuring laser linewidths, called the delayed self-heterodyne method, was invented [4], and the spectral property of semiconductor lasers was studied extensively [5–7]. It was found that the linewidth of GaAlAs lasers was typically in the range of 10 MHz. Such a narrow linewidth together with the precisely controlled center frequency accelerated researches of coherent optical communications based on semiconductor lasers.

The SOP dependence of the receiver sensitivity was overcome by the polarization diversity technique [8]. Each polarization component was detected by orthogonally polarized LOs, and post-processing of the outputs could achieve the SOP-independent receiver sensitivity as shown in Fig. 2.2.

Analyses of the receiver sensitivity of various modulation/demodulation schemes were performed [9, 10]. It was found that the shot-noise-limited receiver sensitivity could be achieved by injecting a sufficient LO power into the receiver to combat against circuit noise. In addition, the use of phase information further improved the receiver sensitivity because the symbol distance could be extended on the IQ plane. For example, let us compare the binary phase-shift keying (PSK) modulation with the binary intensity modulation. When the signal peak power is maintained, the symbol distance on the IQ plane in the PSK modulation format is twice as long as that in the intensity modulation scheme, which improves the receiver sensitivity of the PSK modulation format by 6 dB compared with the intensity modulation scheme. The motivation of R&D of coherent optical communications at this stage lay in this high receiver sensitivity that brought forth longer unrepeated transmission distance.

Many demonstrations of heterodyne systems were reported around 1990. In such systems, the frequency-shift keying (FSK) modulation format was most commonly employed because the semiconductor-laser frequency could be easily modulated.

---

**Fig. 2.2** Polarization diversity heterodyne receiver. PBS polarization beam splitter, Demod.: demodulator
by direct bias-current modulation. Such frequency modulation was demodulated by differential detection at the IF stage. Field trial of undersea transmission at 2.5 Gbit/s was reported in [11], where advanced heterodyne technologies, such as CPFSK modulation, polarization diversity, automatic frequency control (AFC) of semiconductor lasers, and differential detection, were introduced as shown in Fig. 2.3.

Homodyne receivers were also investigated in the 1980s. The advantage of homodyne receivers was that the baseband signal was directly obtained, in contrast to heterodyne receivers which require a rather high intermediate frequency. Figure 2.4 is an example of the OPLL-type phase-diversity homodyne receiver for BPSK modulation [12]. The most difficult issue for the OPLL-type homodyne receiver is to recover the carrier phase. The optical signal modulated in the PSK format does not have the carrier component; therefore, to recover the carrier phase, the product between the in-phase component \(\sin\{\theta_s(t) + \theta_n(t)\}\) and the quadrature component \(\cos\{\theta_s(t) + \theta_n(t)\}\) are taken, where \(\theta_s(t)\) and \(\theta_n(t)\) are phase modulation and phase noise, respectively. Such a product eliminates BPSK modulation, leading to the phase error \(\theta_n(t)\) between the transmitter and LO. The phase error is then led to the frequency controlling terminal of LO so that the LO phase tracks the carrier phase. Note that LO acts as a voltage-controlled oscillator (VCO). This type of PLL scheme is well known as the Costas loop. The PLL bandwidth was usually limited below 1 MHz because of large loop delay, and it was difficult to maintain system stability when semiconductor lasers had large phase noise and frequency.
drift. This technical difficulty inherent in OPLL has not been solved perfectly even when we use state-of-the-art distributed-feedback (DFB) semiconductor lasers.

In the 1990s, the invention of erbium-doped fiber amplifiers (EDFAs) made the shot-noise-limited receiver sensitivity of the coherent receiver less significant. This is because the carrier-to-noise ratio (CNR) of the signal transmitted through the amplifier chain is determined from the accumulated amplified spontaneous emission (ASE) rather than the shot noise. In addition, even in unrepeated transmission systems, the EDFA used as a low-noise preamplifier eliminated the need for the coherent receiver with superior sensitivity. The coherent receiver actually had advantages other than the high receiver sensitivity. For example, it could cope with multi-level modulation formats and had post-processing functions such as compensation for group-velocity dispersion (GVD) of the fiber. However, such advantages were neither urgent requirements for the system nor cost-effective solutions in the 1990s.

Technical difficulties in coherent receivers had not been solved by that time. The heterodyne receiver required an intermediate frequency (IF), which should be higher than the signal bit rate. The maximum bit rate of the heterodyne receiver was always less than the half of that the square-law detector could achieve. On the other hand, the homodyne receiver was essentially a baseband receiver; however, the complexity in stable locking of the carrier phase drift had prevented its practical applications.

From these reasons, further research and development activities in coherent optical communications have almost been interrupted for nearly 20 years. On the other hand, the EDFA-based IMDD system started to take benefit from wavelength-division multiplexing (WDM) techniques to increase the transmission capacity of a single fiber. The hardware required for WDM networks became widely deployed owing to its simplicity and the relatively low cost associated with optical amplifier repeaters where multiple WDM channels could be amplified simultaneously. The WDM technique marked the beginning of a new era in the history of optical communication systems and brought forth 1,000 times increase in the transmission capacity in the 1990s.

### 2.1.2 Revival of Coherent Optical Communications

With the transmission-capacity increase in WDM systems, coherent technologies have restarted to attract a large interest over the recent years. The motivation lies in finding methods of meeting the ever-increasing bandwidth demand with multi-level modulation formats based on coherent technologies [13]. The first step of the revival of coherent optical communications research was ignited with the quadrature PSK (QPSK) modulation/demodulation experiment featuring optical in-phase and quadrature (IQ) modulation and optical delay detection [14]. In such a scheme, one symbol carries two bits by using the four-point constellation on the complex plane; therefore, we can double the bit rate, while keeping the symbol rate, or maintain the bit rate even with the halved spectral width.
Figure 2.5 shows the comparison of the device structure and the phasor diagram among phase modulation (PM), amplitude modulation (AM), and IQ modulation. Optical amplitude modulation (AM) can be achieved by using phase modulators in a Mach–Zehnder configuration, which are driven in a push–pull mode of operation [15]. Optical IQ modulation, on the other hand, can be realized with Mach–Zehnder-type push–pull modulators in parallel, between which a $\pi/2$-phase shift is given [16]. The IQ components of the optical carrier is modulated independently with the IQ modulator, enabling any kind of modulation formats. IQ modulators integrated on LiNbO$_3$ substrates are now commercially available.

The optical delay detector is composed of an optical one-bit delay line and a double-balanced photodiode as shown in Fig. 2.6. With such a receiver, we can compare the phase of the transmitted signal with that of the previous symbol and restore the data, which is differentially precoded at the transmitter. Using two such receivers in parallel, we can demodulate the signal IQ components separately without the need for LO. The principle of operation of the DQPSK receiver is shown in Fig. 2.7. Let us assume that the previous symbol constitutes the phase reference. Depending on the phase difference $\Delta \theta$ between the current symbol and the previous symbol, we have the following four cases: (1) $\Delta \theta = 0$, (2) $\Delta \theta = \pi/2$, (3) $\Delta \theta = -\pi/2$, and (4) $\Delta \theta = \pi$. Considering the phase shift of $\pm \pi/4$ in Mach–Zehnder arms in Fig. 2.6, we find that the output from the I port is given as the real part of the phasor.
Fig. 2.7 Principle of operation of the DQPSK receiver. From signs of the IQ outputs, we can estimate the phase difference of the QPSK signal indicated by the double line in Fig. 2.7, whereas that from the Q port as the real part of the phasor indicated by the dotted line. Since signs of these outputs from the IQ ports in Fig. 2.6 have the four different patterns, we can demodulate the DQPSK signal. A number of long-distance WDM QPSK transmission experiments have been reported recently, based on IQ modulation and differential IQ demodulation.

The next stage has opened with high-speed digital signal processing (DSP). In the field of radio communications, digital techniques have been widely introduced into transmitters and receivers. Figure 2.8 shows configurations of (a) DSP-based RF transmitter and (b) receiver. At the transmitter, after appropriate DSP, digital data are converted into two-channel analog data through digital-to-analog converters (DACs), which modulate IQ components of the RF carrier. On the other hand, at the receiver, the transmitted RF signal is mixed with LO, and IQ components are demodulated. Such IQ data are converted to the digital domain by using analog-to-digital converters (ADCs), and symbols are decoded through DSP. Digital signal processing at the transmitter and the receiver enables “software-defined radio communications.”

If the RF modulator in Fig. 2.8(a) and the RF mixer in Fig. 2.8(b) are replaced with optical counterparts, which are nothing but the optical IQ modulator and the phase-diversity homodyne receiver, respectively, we can easily imagine the DSP-based optical transmitter and receiver shown in Fig. 2.9. Such an optical transmitter and a receiver were actually reported in [17] and [18–20], respectively.

The recent development of high-speed digital integrated circuits has offered the possibility of treating the electrical signal in a digital signal-processing (DSP) core and retrieving the IQ components of the complex amplitude of the optical carrier from the homodyne-detected signal in a very stable manner. The 20-Gbit/s QPSK signal was demodulated with a phase-diversity homodyne receiver followed by digital carrier phase estimation in [18], although bit-error rate measurements were done still offline. Since the carrier phase is recovered after homodyne detection by means
Fig. 2.8 Configurations of (a) DSP-based RF transmitter and (b) receiver

Fig. 2.9 Configurations of (a) DSP-based optical transmitter and (b) receiver

of DSP, this type of receiver has now commonly been called the “digital coherent receiver.” While an optical phase-locked loop (OPLL) that locks the LO phase to the signal phase is still difficult to achieve due to the loop delay problem, DSP circuits are becoming increasingly faster and provide us with simple and efficient means of estimating the carrier phase. Very fast tracking of the carrier phase improves system stability drastically as compared with the OPLL scheme.
Any kind of multi-level modulation formats can be introduced by using the coherent receiver [21]. While the spectral efficiency of binary modulation formats is limited to 1 bit/s/Hz/polarization, which is called the Nyquist limit, modulation formats with $M$ bits of information per symbol can achieve up to the spectral efficiency of $M$ bit/s/Hz/polarization. Although optical delay detection has been employed to demodulate the quadrature phase-shift keying (QPSK) signal ($M = 2$), further increase in multiplicity is hardly achieved with such a scheme. Figure 2.10 shows constellation maps for BPSK, QPSK, 8PSK, and 16QAM formats. These modulation formats can transmit 1 bit, 2 bits, 3 bits, and 4 bits per symbol, respectively. Each symbol is encoded by using the Gray code scheme.

Another and probably more important advantage of the digital coherent receiver is the post-signal-processing function [22]. The IQ demodulation by the digital coherent receiver is the entirely linear process; therefore, all the information on the complex amplitude of the transmitted optical signal is preserved even after detection, and signal-processing functions acting on the optical carrier, such as optical filtering and dispersion compensation, can be performed at the electrical stage after detection.

The polarization alignment is also made possible after detection by introducing the polarization diversity scheme into the homodyne receiver [23]. The complex amplitude of the horizontal polarization and that of the vertical polarization are simultaneously measured and processed by DSP. Polarization demultiplexing and compensation for polarization-mode dispersion (PMD) have also been demonstrated with the digital coherent receiver [24], where bulky and slow optical-polarization controllers as well as optical delay lines are removed.

Since the digital coherent receiver requires high-speed ADC and DSP, most of the experiments have been done still offline. It means that after transmitted data are stored in a computer, bit errors are analyzed offline. However, very recently, an application-specific integrated circuit (ASIC) designed for the 11.5-Gsymbol/s polarization-multiplexed QPSK signal has been developed [25], and the real-time operation of the digital coherent receiver at the bit rate of 46 Gbit/s has been demonstrated by using such an ASIC [26]. This achievement is really a milestone in the history of the modern coherent optical communications. The combination of coherent detection and DSP is thus expected to become a part of the next generation of
optical communication systems and provide new capabilities that were not possible without the detection of the phase of the optical signal.

### 2.2 Principle of Coherent Optical Detection

This section describes the basic operation principle of coherent optical detection. We show how the coherent receiver measures the complex amplitude of the optical signal with the shot-noise-limited sensitivity and how information on the state of polarization can be extracted by the use of polarization diversity.

#### 2.2.1 Coherent Detection

Figure 2.11 shows the configuration of the coherent optical receiver. The fundamental concept behind coherent detection is to take the product of electric fields of the modulated signal light and the continuous-wave (CW) local oscillator (LO). Let the optical signal incoming from the transmitter be

\[ E_s(t) = A_s(t) \exp(j \omega_s t), \]  \hspace{1cm} (2.1)

where \( A_s(t) \) is the complex amplitude and \( \omega_s \) the angular frequency. Similarly, the electric field of LO prepared at the receiver can be written as

\[ E_{LO}(t) = A_{LO} \exp(j \omega_{LO} t), \]  \hspace{1cm} (2.2)

where \( A_{LO} \) is the constant complex amplitude and \( \omega_{LO} \) the angular frequency of LO. We note here that the complex amplitudes \( A_s \) and \( A_{LO} \) are related to the signal power \( P_s \) and the LO power \( P_{LO} \) by \( P_s = |A_s|^2 / 2 \) and \( P_{LO} = |A_{LO}|^2 / 2 \), respectively.

Balanced detection is usually introduced into the coherent receiver as a means to suppress the dc component and maximize the signal photocurrent. The concept resides in using a 3-dB optical coupler that adds a 180° phase shift to either the signal field or the LO field between the two output ports. When the signal and LO

---

**Fig. 2.11** Configuration of the coherent receiver that measures the beat between the signal and LO
are co-polarized, the electric fields incident on the upper and lower photodiodes are given as

\[ E_1 = \frac{1}{\sqrt{2}}(E_s + E_{LO}), \]
\[ E_2 = \frac{1}{\sqrt{2}}(E_s - E_{LO}), \]

and the output photocurrents are written as

\[ I_1(t) = R \left[ \text{Re} \left\{ \frac{A_s(t) \exp(j \omega_s t) + A_{LO} \exp(j \omega_{LO} t)}{\sqrt{2}} \right\} \right]^{\text{ms}} \]
\[ = \frac{R}{2} [ P_s(t) + P_{LO} \\
+ 2\sqrt{P_s(t)P_{LO}} \cos \left\{ \omega_{IF} t + \theta_{\text{sig}}(t) - \theta_{LO}(t) \right\}], \quad (2.5) \]

\[ I_2(t) = R \left[ \text{Re} \left\{ \frac{A_s(t) \exp(j \omega_s t) - A_{LO} \exp(j \omega_{LO} t)}{\sqrt{2}} \right\} \right]^{\text{ms}} \]
\[ = \frac{R}{2} [ P_s(t) + P_{LO} \\
- 2\sqrt{P_s(t)P_{LO}} \cos \left\{ \omega_{IF} t + \theta_{\text{sig}}(t) - \theta_{LO}(t) \right\}], \quad (2.6) \]

where “ms” means the mean square with respect to the optical frequencies, “Re” means to take the real part, \( \omega_{IF} \) is known as the intermediate frequency (IF) given by \( \omega_{IF} = |\omega_s - \omega_{LO}| \), and \( \theta_{\text{sig}}(t) \) and \( \theta_{LO}(t) \) are phases of the transmitted signal and LO, respectively. \( R \) is the responsivity of the photodiode given as

\[ R = \frac{e \eta}{\hbar \omega_s}, \quad (2.7) \]

where \( \hbar \) stands for the Planck’s constant, \( e \) the electron charge, and \( \eta \) the quantum efficiency of the photodiode. The balanced detector output is then given as

\[ I(t) = I_1(t) - I_2(t) = 2R \sqrt{P_s(t)P_{LO}} \cos \left\{ \omega_{IF} t + \theta_{\text{sig}}(t) - \theta_{LO}(t) \right\}. \quad (2.8) \]

\( P_{LO} \) is always constant and \( \theta_{LO}(t) \) includes only the phase noise that varies in time.

### 2.2.2 Heterodyne Receivers

Heterodyne detection refers to the case that \( |\omega_{IF}| > \omega_b/2 \), where \( \omega_b \) is the modulation bandwidth of the optical carrier determined by the symbol rate. In such a case,
Eq. (2.8) shows that the electric field of the signal light is down-converted to the IF signal including the amplitude information and the phase information, as shown in Fig. 2.12.

The signal phase is given as $\theta_{\text{sig}}(t) = \theta_s(t) + \theta_{\text{sn}}(t)$, where $\theta_s(t)$ is the phase modulation and $\theta_{\text{sn}}(t)$ the phase noise. The receiver output is given as

$$I(t) = 2R\sqrt{P_s(t)P_{\text{LO}}} \cos \{\omega_{\text{IF}}t + \theta_s(t) + \theta_n(t)\}, \quad (2.9)$$

and we can determine the complex amplitude on $\exp(j\omega_{\text{IF}}t)$ from Eq. (2.9) as

$$I_c(t) = 2R\sqrt{P_s(t)P_{\text{LO}}} \exp j \{\theta_s(t) + \theta_n(t)\}, \quad (2.10)$$

where $\theta_n(t)$ is the total phase noise given as

$$\theta_n(t) = \theta_{\text{sn}}(t) - \theta_{\text{LO}}(t). \quad (2.11)$$

Note that Eq. (2.10) is equivalent to the complex amplitude $A_s(t)$ of the optical signal except for the phase noise increase stemming from LO.

We have three methods of demodulating $I_c(t)$, which are envelope (noncoherent) detection, differential (delay) detection, and synchronous (coherent) detection as shown in Fig. 2.13. In envelope detection, we measure $|I_c(t)|^2$ from Eq. (2.10), which gives us only the information on $P_s(t)$. Differential detection is effective for constant-envelope modulation formats such as $M$-ary PSK. In this scheme, we determine the phase difference between the current symbol and the previous one. In the synchronous detection scheme, although the total phase noise $\theta_n(t) = \theta_{\text{sn}}(t) - \theta_{\text{LO}}(t)$ might vary in time, the electrical phase-locked loop (PLL) can be used to estimate the phase noise and decode the symbol.
2.2.3 Homodyne Receivers

Homodyne detection refers to the case that $\omega_{IF} = 0$. The photodiode current from the homodyne receiver becomes

$$I(t) = 2R\sqrt{P_s(t) P_{LO}} \cos \left\{ \theta_{\text{sig}}(t) - \theta_{\text{LO}}(t) \right\}.$$  \hspace{1cm} (2.12)

Equation (2.12) means that the homodyne receiver measures the inner product between the signal phasor and the LO phasor as shown in Fig. 2.14. In order to decode the symbol correctly, the LO phase $\theta_{\text{LO}}(t)$ must track the transmitter phase noise $\theta_{\text{sn}}(t)$ such that $\theta_{\text{sn}}(t) = 0$. This function is realized by the optical phase-locked loop (OPLL); however, in practice, the implementation of such a loop is not simple and adds to homodyne detection the complexity of the configuration. In addition, Eq. (2.12) only gives the cosine component (in other words, the in-phase component with respect to the LO phase), and the sine component (the quadrature component) cannot be detected. Therefore, this type of homodyne receivers is not able to extract the full information on the signal complex amplitude.

Preparing another LO, whose phase is shifted by $90^\circ$, in the homodyne receiver, we can detect both in-phase and quadrature components of the signal light as shown in Fig. 2.15. This function is achieved by a $90^\circ$ optical hybrid shown in Fig. 2.16. Using such $90^\circ$ optical hybrid, we can obtain four outputs $E_1, E_2, E_3$, and $E_4$ from

Fig. 2.13 Methods of demodulating the IF signal

Fig. 2.14 Phasor diagram of the signal and LO for homodyne detection, which measures the inner product between the signal and LO phasors
Fig. 2.15 Phasor diagram of the signal and LO for phase-diversity homodyne detection. Both of the in-phase and quadrature components are measured at the same time.

Fig. 2.16 Configuration of the phase-diversity homodyne receiver using a 90° optical hybrid.

The two inputs $E_s$ and $E_{LO}$ as

$$E_1 = \frac{1}{2}(E_s + E_{LO}),$$  \hspace{1cm} (2.13)  

$$E_2 = \frac{1}{2}(E_s - E_{LO}),$$  \hspace{1cm} (2.14)  

$$E_3 = \frac{1}{2}(E_s + jE_{LO}),$$  \hspace{1cm} (2.15)  

$$E_4 = \frac{1}{2}(E_s - jE_{LO}).$$  \hspace{1cm} (2.16)  

Output photocurrents from balanced photodetectors are then given as

$$I_I(t) = I_{I1}(t) - I_{I2}(t) = R\sqrt{P_s P_{LO}} \cos\{\theta_{sig}(t) - \theta_{LO}(t)\},$$ \hspace{1cm} (2.17)  

$$I_Q(t) = I_{Q1}(t) - I_{Q2}(t) = R\sqrt{P_s P_{LO}} \sin\{\theta_{sig}(t) - \theta_{LO}(t)\}.$$ \hspace{1cm} (2.18)  

Using Eqs. (2.17) and (2.18), we can restore the complex amplitude as

$$I_c(t) = I_I(t) + jI_Q(t) = R\sqrt{P_s(t) P_{LO}} \exp\{j(\theta_s(t) + \theta_n(t))\},$$ \hspace{1cm} (2.19)  

which is equivalent to the complex amplitude of the optical signal except for the phase noise increase. Note that the complex amplitude is obtained in the baseband in contrast to heterodyne detection.

Equation (2.19) shows that the electric field of the signal light is down-converted to the baseband. As shown in the down-converted spectrum of Fig. 2.17, we need to allow the negative frequency to express the complex amplitude at the baseband, which contains both the in-phase (or cos) and quadrature (or sin) components. In
Fig. 2.17 Spectra of (a) the optical signal and (b) homodyne-detected baseband signal. The conventional homodyne receiver only measures the real part of the optical complex amplitude, whereas the phase-diversity receiver does the complex amplitude itself whose spectrum exists on both the positive and negative frequency sides.

Contrast, since the conventional homodyne receiver only measures the in-phase (or cos) component, the baseband signal exists only in the positive frequency side.

This type of receiver is commonly called the “phase-diversity homodyne receiver” [27] or “intradyne receiver” [28]. Eventually, the phase-diversity homodyne receiver and the heterodyne receiver can similarly restore the full information on the optical complex amplitude, as shown by Eqs. (2.10) and (2.19), respectively. However, since the phase-diversity homodyne receiver generates the baseband signal directly, it is more advantageous over the heterodyne receiver which must deal with a rather high intermediate frequency.

Figure 2.18 shows methods of demodulating $I_c(t)$. Similarly to heterodyne detection, we can demodulate $I_c(t)$ by (a) envelope (non-coherent) detection and (b) differential (delay) detection at the baseband. As for synchronous (coherent)
detection, in addition to relying upon (c) conventional OPLL, where the LO phase tracks the signal phase noise, it is possible to estimate the phase noise $\theta_n(t)$ and restore the signal complex amplitude through digital signal processing on the homodyne-detected signal given by Eq. (2.19) (see (c')). This is the basic idea of the “digital coherent receiver,” which has recently been investigated extensively.

### 2.2.4 Homodyne Receiver Employing Phase and Polarization Diversities

It was assumed up to now that the polarization of the incoming signal was always aligned to that of LO. However, in practical systems, the polarization of the incoming signal is unlikely to remain aligned to the state of polarization (SOP) of LO because of random changes on the birefringence of the transmission fiber. One of the most serious problems of the coherent receiver is that the receiver sensitivity is dependent on SOP of the incoming signal. In this subsection, we show that the polarization diversity receiver can cope with the polarization dependence of the receiver sensitivity.

The receiver employing polarization diversity is shown in Fig. 2.19, where two phase-diversity homodyne receivers are combined with the polarization diversity configuration [23]. The incoming signal having an arbitrary SOP is separated into two linear polarization components with a polarization beam splitter (PBS).

Assuming that a single-polarization component of the optical carrier is modulated at the transmitter, let the $x$- and $y$-polarization components after PBS at the receiver be written as

$$
\begin{bmatrix}
E_{sx} \\
E_{sy}
\end{bmatrix} = \begin{bmatrix}
\sqrt{\alpha} A_s e^{j\delta} \\
\sqrt{1 - \alpha} A_s
\end{bmatrix} \exp(j \omega_s t),
$$

(2.20)

![Fig. 2.19 Configuration of the homodyne receiver employing phase and polarization diversities](image)
where \( \alpha \) denotes the power ratio of the two polarization components, and \( \delta \) the phase difference between them. These parameters are dependent on the birefringence of the transmission fiber and time-varying. On the other hand, the \( x \)- and \( y \)-polarization components equally separated from the linearly polarized LO are written as

\[
\begin{bmatrix}
E_{\text{LO},x} \\
E_{\text{LO},y}
\end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix}
A_{\text{LO}} \\
A_{\text{LO}}
\end{bmatrix} \exp(j\omega_{\text{LO}}t). \tag{2.21}
\]

Two 90° optical hybrids in Fig. 2.19 generate electric fields \( E_{1,\ldots,8} \) at double-balanced photodiodes PD1–PD4:

\[
E_{1,2} = \frac{1}{2} \left( E_{sx} \pm E_{\text{LO},x} \right), \tag{2.22}
\]

\[
E_{3,4} = \frac{1}{2} \left( E_{sx} \pm jE_{\text{LO},x} \right), \tag{2.23}
\]

\[
E_{5,6} = \frac{1}{2} \left( E_{sy} \pm E_{\text{LO},y} \right), \tag{2.24}
\]

\[
E_{7,8} = \frac{1}{2} \left( E_{sy} \pm jE_{\text{LO},y} \right). \tag{2.25}
\]

Photocurrents from PD1 to PD4 are then given as

\[
I_{\text{PD}1} = R \sqrt{\frac{\alpha P_s(t) P_{\text{LO}}}{2}} \cos\{\theta_s(t) - \theta_{\text{LO}}(t) + \delta\}, \tag{2.26}
\]

\[
I_{\text{PD}2} = R \sqrt{\frac{\alpha P_s(t) P_{\text{LO}}}{2}} \sin\{\theta_s(t) - \theta_{\text{LO}}(t) + \delta\}, \tag{2.27}
\]

\[
I_{\text{PD}3} = R \sqrt{\frac{(1 - \alpha) P_s(t) P_{\text{LO}}}{2}} \cos\{\theta_s(t) - \theta_{\text{LO}}(t)\}, \tag{2.28}
\]

\[
I_{\text{PD}4} = R \sqrt{\frac{(1 - \alpha) P_s(t) P_{\text{LO}}}{2}} \sin\{\theta_s(t) - \theta_{\text{LO}}(t)\}. \tag{2.29}
\]

From Eqs. (2.26), (2.27), (2.28), and (2.29), we find that the polarization diversity receiver can separately measure complex amplitudes of the two polarization components as
\[ I_{xc}(t) = I_{PD1}(t) + j I_{PD2}(t) \]
\[ = R \sqrt{\frac{\alpha P_s(t) P_{LO}}{2}} \exp j \{ \theta_s(t) + \theta_n(t) + \delta \}, \tag{2.30} \]
\[ I_{yc}(t) = I_{PD3}(t) + j I_{PD4}(t) \]
\[ = R \sqrt{\frac{(1 - \alpha) P_s(t) P_{LO}}{2}} \exp j \{ \theta_s(t) + \theta_n(t) + \delta \}, \tag{2.31} \]

from which we can reconstruct the signal complex amplitude \( A_s(t) \) in a polarization-independent manner.

### 2.2.5 Carrier-to-Noise Ratio

From the complex amplitude at the IF stage given by Eq. (2.10), the carrier-to-noise ratio (CNR) of the heterodyne signal is given as

\[ \gamma_s = \frac{|I_c(t)|^2/2}{2eRP_{LO}B} = \frac{\eta P_s}{hfB}, \tag{2.32} \]

where we assume that the LO shot noise overwhelms the circuit noise with sufficient LO power, and \( B \) is the receiver bandwidth at the IF stage. Noting that the minimum bandwidth at the IF stage is given as

\[ B = \frac{1}{T}, \tag{2.33} \]

where \( T \) is the symbol duration, we find

\[ \gamma_s = \frac{\eta P_s T}{hf} = \eta N_s, \tag{2.34} \]

where \( N_s = P_s T/ hf \) means the number of photons per symbol [29]. This is the shot-noise-limited CNR.

On the other hand, the homodyne phase-diversity receiver can generate the complex amplitude given by Eq. (2.19) at the baseband. Therefore, the mean square of the signal photocurrent is given as \( |I_c(t)|^2 = R^2 P_s P_{LO} \). When reconstructing the complex amplitude, we need to add shot noises due to LO powers of \( P_{LO}/2 \) from the two ports. The receiver bandwidth at the baseband is \( B/2 = 1/2T \); therefore, the total shot-noise current is given as \( eP_{LO}B \). We can thus obtain the CNR \( \gamma_s = \eta P_s / hf B = \eta N_s \), which is the same as the heterodyne receiver. Even when the polarization diversity is introduced, signal processing called “maximal-ratio combining” can maintain CNR [29].
The CNR obtained from coherent detection is the intrinsic CNR of the coherent state of light, which can be understood from the quantum mechanical point of view as follows: Figure 2.20 shows the phasor of light, which is associated with vacuum fluctuations. The average photon energy is $hfN_s$, where $N_s$ denotes the average number of photons, and the noise energy originating from vacuum fluctuation is $hf/2$, which is the half photon energy. Let us consider the case when we measure the phasor of the light by heterodyne detection. When we measure the beat between the signal at the angular frequency of $\omega_s$ and the LO at $\omega_{LO}$, the vacuum fluctuation at $\omega_{image}$ is also merged into the IF band at $\omega_s - \omega_{LO}$. Therefore, the noise energy at IF is $hf$, resulting in the CNR of the IF signal $\gamma_s = N_s$. In the case of phase-diversity homodyne detection, although the image band noise never merges into the detected signal, branching the signal light into the I and Q ports degrades CNR by 3 dB, which gives us the same CNR as heterodyne detection.

By using this picture, we can understand the impact of EDFA on the receiver sensitivity. Figure 2.21 shows the $m$-stage EDFA chain, where the amplifier gain $G$ compensates for the fiber loss $G^{-1}$ periodically. Each amplifier generates noise photons of $(G - 1)n_{sp}$ due to amplified spontaneous emission, where $n_{sp}$ denotes
the spontaneous emission factor. At the output, the accumulated number of noise photons is given as $m(G - 1)n_{sp}$, whereas the number of noise photons due to vacuum fluctuations is one.

Since $m(G - 1)n_{sp} \gg 1$, CNR of the signal transmitted through the amplifier chain is given as

$$\gamma_s = \frac{GN_s}{m(G - 1)n_{sp}} \approx \frac{N_s}{mn_{sp}},$$

which is much lower than the shot-noise limit given by Eq. (2.34); therefore, the shot-noise-limited receiver sensitivity is not necessary in such an amplifier chain. This is the main reason why R&D in coherent optical communications has been interrupted behind the rapid progress in EDFA technologies as discussed in Sect. 2.1.1.

### 2.3 Digital Signal Processing in Coherent Receivers

This section describes details of digital signal processing for coherent receivers. Outputs from the homodyne receiver comprising phase and polarization diversities are processed by digital signal-processing circuits, restoring the complex amplitude of the signal in a stable manner despite fluctuations of the carrier phase and the signal SOP. Symbol-by-symbol control of such time-varying parameters in the digital domain can greatly enhance the system stability compared with optical control methods.

#### 2.3.1 Basic Concept of the Digital Coherent Receiver

Figure 2.22 shows the basic concept of the digital coherent receiver. First, the incoming signal is detected linearly with the homodyne receiver comprising phase and polarization diversities. Using this receiver, we can obtain full information on the optical carrier, namely the complex amplitude and the state of polarization. Such complex amplitude measured by the receiver is converted to digital data with ADCs and processed by DSP circuits. The progress in the increased performance, speed, and reliability of integrated circuits now makes digital signal processing an attractive approach to recover the optical complex amplitude from the homodyne-detected baseband signal.

The combination of the optical IQ modulator and the IQ demodulator realizes the linear optical communication system as shown in Fig. 2.23. At the transmitter, we define a vector on the complex plane using two voltages driving the IQ modulator. This vector is mapped on the phasor of the optical carrier through the IQ modulator. Such optical IQ modulation is perfectly restored by IQ demodulation, which is
performed by the digital coherent receiver. This is exactly the linear system, where IQ information is preserved even with E/O and O/E conversion processes.

The digital signal-processing circuit is typically composed of the sequence of operations shown in Fig. 2.24 to retrieve the information from the received signal. First, the four-channel ADC asynchronously samples the data and restores the complex amplitudes (see Sect. 2.3.2). Such complex amplitudes are equalized by an fixed equalizer which removes inter-symbol interference (ISI) (see Sect. 2.3.5). Next, the dual-polarization signal is demultiplexed and polarization-mode dispersion (PMD) is compensated for, usually by using the constant-modulus algorithm (CMA) (see Sect. 2.3.4). After the clock is extracted from the interpolated data in the time domain, they are resampled to keep one sample within a symbol interval $T$ (see Sect. 2.3.2). Then the carrier phase is estimated, and the remaining ISI is removed by an adaptive equalizer driven by decoded symbols (see Sect. 2.3.5).
2.3.2 Sampling of the Signal and Clock Extraction

When the PSK signal is encoded on an RZ pulse train (RZ-PSK), we can realize accurate extraction of the clock from the transmitted signal, detecting the intensity of the received signal and using a standard clock recovery circuit. Photocurrents are then sampled and digitized by ADCs at the timing of the extracted clock. In this case, the digital circuit does not need to resample the data, so that it saves on complexity of the digital circuit.

It is also possible to extract the clock from the asynchronously sampled signal as shown in Fig. 2.24. The necessary condition for asynchronous sampling is that the sampling frequency must be equal to or greater than twice the symbol rate. After the clock is extracted from the interpolated data in the time domain, they are resampled to keep one sample within a symbol interval $T$.

2.3.3 Phase Estimation

Since the linewidth of semiconductor DFB lasers used as the transmitter and LO typically ranges from 100 kHz to 10 MHz, the phase noise $\theta_n(t)$ varies much more slowly than the phase modulation $\theta_s(t)$. Therefore, by averaging the carrier phase over many symbol intervals, it is possible to obtain an accurate phase estimate. In the following, assuming the $M$-ary PSK modulation, we explain the phase estimation procedure, but such procedure can easily be extended to quadrature amplitude modulation (QAM).

The phase of the complex amplitude obtained from Eq. (2.21) contains both the phase modulation $\theta_s(i)$ and the phase noise $\theta_n(i)$, where $i$ represents the sample number. The procedure to estimate $\theta_n$ is shown in Fig. 2.25, where the case of QPSK is shown for simplicity. We take the $M$th power of the measured complex amplitude $I_c(i)$, because the phase modulation is removed from $I_c(i)^M$ in the $M$-ary PSK modulation format. Subtracting the phase noise thus estimated from the measured phase, we can restore the phase modulation.
Fig. 2.25 Principle of the $M$th-power phase estimation method. For simplicity, the case of QPSK ($M = 4$) is shown. Taking the $M$th power of the received complex amplitude, we can eliminate the phase modulation and measure the phase noise.

In actual phase estimation, we average $I_c(i)^M$ over $2k + 1$ samples to improve the signal-to-noise ratio of the estimated phase reference. The estimated phase is thus given as

$$
\theta_e(i) = \arg \left( \frac{\sum_{j=-k}^{k} I_c(i + j)^M}{M} \right).
$$

(2.36)

The phase modulation $\theta_s(i)$ is determined by subtracting $\theta_e(i)$ from the measured phase of $\theta(i)$. The phase modulation is then discriminated among $M$ symbols. Figure 2.26 shows the DPS circuit for such phase estimation.

The symbols thus obtained have the phase ambiguity by $2\pi/M$ because we cannot know the absolute phase. It is important to note that the data should be differentially precoded. Differentially decoding the discriminated symbol after symbol discrimination, we can solve the phase ambiguity problem although the bit-error rate is doubled by error multiplication.

The phase estimate $\theta_e(i)$ ranges between $-\pi/M$ and $+\pi/M$. Therefore, if $|\theta_e(i)|$ exceeds $\pi/M$, the phase jump of $2\pi/M$ occurs inevitably as shown in Fig. 2.27. To

Fig. 2.26 DSP circuit for $M$th power phase estimation
cope with this problem, the correction for the phase jump is done by obeying the following rule:

$$\theta_c(i) \leftarrow \theta_c(i) + \frac{2\pi}{M} f(\theta_c(i) - \theta_c(i-1)),$$  \hspace{1cm} (2.37)

where \( f(x) \) is defined as

$$f(x) = \begin{cases} 
+1 & \text{for } x < -\frac{\pi}{M} \\
0 & \text{for } |x| \leq \frac{\pi}{M} \\
-1 & \text{for } x > \frac{\pi}{M} 
\end{cases}$$  \hspace{1cm} (2.38)

This adjustment ensures that the phase estimate follows the trajectory of the physical phase and cycle slips are avoided [30].

### 2.3.4 Polarization Alignment

Throughout this subsection, we regard for simplicity that the measured photocurrent is identical to the optical complex amplitude, neglecting conversion factors from the optical domain to the electrical domain. Let the optical complex amplitude in the \( x \)-polarization state be \( E_x \) and that in the \( y \)-polarization state be \( E_y \) at the receiver.

When the single-polarization signal \( E_{in}(t) \) is transmitted, we can apply the maximal-ratio combining method to align the SOP of the incoming signal [29]. ADCs sample and digitize the four outputs \( I_{PD1}, \ldots, I_{PD4} \) from the phase/polarization diversity receiver shown in Fig. 2.19. The \( x \)- and \( y \)-polarization components of the complex amplitude of the signal are given by Eqs. (2.30) and (2.31) as

$$E_x(i) = \sqrt{2\alpha P_s(i)} \exp{j[\theta_s(i) + \theta_n(i) + \delta]},$$  \hspace{1cm} (2.39)

$$E_y(i) = \sqrt{2(1 - \alpha) P_s(i)} \exp{j[\theta_s(i) + \theta_n(i)}].$$  \hspace{1cm} (2.40)
Figure 2.28 shows the post-processing circuit realizing the maximal-ratio polarization combining process. We define the ratio $r(i)$ as

$$ r(i) = \frac{E_x(i)}{E_y(i)}. \tag{2.41} $$

Polarization parameters $\alpha$ and $\delta$ of the incoming signal vary much more slowly than the phase modulation. Therefore, by averaging $r(i)$ over many symbol intervals, it is possible to obtain accurate values of $\alpha$ and $\delta$: The ratio $r(i)$ averaged over $2\ell + 1$ samples is written as

$$ \bar{r}(i) = \frac{1}{2\ell + 1} \sum_{j=-\ell}^{\ell} r(i), \tag{2.42} $$

and $\alpha$ and $\delta$ are calculated from

$$ \lvert \bar{r}(i) \rvert = \sqrt{\alpha} / \sqrt{\alpha - 1}, \tag{2.43} $$

$$ \text{arg}(\bar{r}(i)) = \delta. \tag{2.44} $$

The signal complex amplitude, independent of SOP of the incoming signal, is then reconstructed by maximal-ratio combining as

$$ E_{in}(i) \propto \bar{r}(i)^* E_x(i) + E_y(i). \tag{2.45} $$

However, the above procedure is not effective when $\alpha \simeq 1$, because $\delta$ contains a large error and the term $\bar{r}^* E_x$ is not determined correctly. In such a case, we should use $E_{in}(i) \propto E_x(i) + E_y(i)/\bar{r}(i)^*$ to reconstruct the signal complex amplitude more accurately. The term $\lvert E_y/\bar{r}^* \rvert$ is much smaller than $\lvert E_x \rvert$, and hence the error in $E_{in}$ is reduced. In summary, the reconstruction formula is given as

$$ E_{in}(i) \propto \begin{cases} \bar{r}(i)^* E_x(i) + E_y(i) & \alpha \leq 0.5 \\ E_x(i) + \frac{E_y(i)}{\bar{r}(i)^*} & \alpha > 0.5 \end{cases}. \tag{2.46} $$
On the other hand, when dual polarizations are transmitted, the constant-modulus algorithm (CMA) has been widely applied to polarization demultiplexing [24, 31]. In what follows, we discuss the principle of operation of polarization demultiplexing.

Let the Jones matrix of the fiber for transmission be given as

\[
T = \begin{bmatrix}
\sqrt{\alpha} e^{i\delta} & -\sqrt{1 - \alpha} \\
\sqrt{1 - \alpha} & \sqrt{\alpha} e^{-i\delta}
\end{bmatrix},
\]

(2.47)

where \(\alpha\) and \(\delta\) denote the power splitting ratio and the phase difference between the two polarization modes, respectively. When the polarization-multiplexed signal \(E_{in} (t)^T = [E_{in,x} (t), E_{in,y} (t)]^T\), where “\(^T\)” means to take the transposed matrix, is incident on the fiber, the SOP of the output is given as

\[
\begin{bmatrix} E_x (t) \\ E_y (t) \end{bmatrix} = T \begin{bmatrix} E_{in,x} (t) \\ E_{in,y} (t) \end{bmatrix}.
\]

(2.48)

Assuming \(M\)-ary PSK modulation, we normalize the envelope of each input polarization component as

\[
|E_{in,x} (t)|^2 = |E_{in,y} (t)|^2 = 1.
\]

(2.49)

Using the received polarization components \(E_x\) and \(E_y\), we calculate \(E_X\) and \(E_Y\) through DSP as

\[
E_X (t) = r E_x (t) + k E_y (t)
= \left( r \sqrt{\alpha} e^{i\delta} + k \sqrt{1 - \alpha} \right) E_{in,x} (t)
+ \left( -r \sqrt{1 - \alpha} + k \sqrt{\alpha} e^{-i\delta} \right) E_{in,y} (t),
\]

(2.50)

where \(r\) and \(k\) are complex numbers. Here, we define

\[
X \equiv r \sqrt{\alpha} e^{i\delta} + k \sqrt{1 - \alpha},
\]

(2.51)

\[
Y \equiv -r \sqrt{1 - \alpha} + k \sqrt{\alpha} e^{-i\delta}.
\]

(2.52)

Then the power of the \(X\)-component is given as

\[
|E_X (t)|^2 = |XE_{in,x} (t)|^2 + |YE_{in,y} (t)|^2
+ 2 |XY E_{in,x} (t) E_{in,y} (t)| \cos \theta (t),
\]

(2.53)

\[
\theta (t) = \arg \left( \frac{YE_{in,y} (t)}{XE_{in,x} (t)} \right).
\]

(2.54)
Since the phase difference between initial $x$ and $y$ polarization components are time-varying at the symbol rate due to independent PSK modulations, the third term in Eq. (2.53) changes as a function of time. Here, we consider the condition under which $|E_X|^2$ becomes time-independent. As a matter of fact, such condition is given as $Y = 0$ (hereafter Case (I)) or $X = 0$ (hereafter Case (II)). In each case, $(r, k)$ is expressed as

\[
\frac{r_x}{k_x} = \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} e^{-j\delta}, \quad (2.55)
\]

or

\[
\frac{r_y}{k_y} = -\frac{\sqrt{1-\alpha}}{\sqrt{\alpha}} e^{-j\delta}. \quad (2.56)
\]

In Case (I), Eq. (2.55) leads to

\[
E_X = \frac{k_x}{\sqrt{1-\alpha}} E_{in,x}, \quad (2.57)
\]

whereas in Case (II), Eq. (2.56) does

\[
E_Y = \frac{k_y e^{-j\delta}}{\sqrt{\alpha}} E_{in,y}. \quad (2.58)
\]

In Case (I), normalizing $E_X$ such that $|E_X|^2 = 1$, we have

\[
\begin{align*}
  r_x &= \sqrt{\alpha} e^{-j\delta} + j\varphi_x, \quad (2.59) \\
  k_x &= \sqrt{1-\alpha} e^{j\varphi_x}, \quad (2.60)
\end{align*}
\]

where $\varphi_x$ is a real constant. Similarly, in Case (II), we have

\[
\begin{align*}
  r_y &= -\sqrt{1-\alpha} e^{j\varphi_y}, \quad (2.61) \\
  k_y &= \sqrt{\alpha} e^{j\delta} + j\varphi_y, \quad (2.62)
\end{align*}
\]

where $\varphi_y$ is a real constant.

In this way, provided that $r$ and $k$ are controlled so that $|E_X|^2$ does not fluctuate at the symbol rate, we can reach either Case (I) or Case (II). In Case (I), we can restore the initial $x$-polarization as

\[
E_X = E_{in,x} e^{j\varphi_x}. \quad (2.63)
\]
By using $r_x$ and $k_x$ obtained from Eqs. (2.59) and (2.60), respectively, the $y$-polarization is derived from the $Y$ port as

$$E_Y = -k_x^* E_x + r_x^* E_y = E_{in,y} e^{-j\psi_y}.$$  \hspace{1cm} (2.64)

On the other hand, in Case (II), the restored $x$ and $y$ polarizations appear at the $Y$ and $X$ ports, respectively, as

$$E_X = E_{in,x} e^{j\psi_y},$$  \hspace{1cm} (2.65)

$$E_Y = -E_{in,x} e^{-j\psi_y}.$$  \hspace{1cm} (2.66)

The algorithm that controls $(r, k)$ so as to satisfy $|E_X|^2 = 1$ is well known as the constant-modulus algorithm (CMA) or the Godard algorithm [32]. Figure 2.29 shows the DSP circuit for polarization demultiplexing based on CMA. The function of this circuit is expressed as

$$\begin{bmatrix} E_X \\ E_Y \end{bmatrix} = p \begin{bmatrix} E_x \\ E_y \end{bmatrix}.$$  \hspace{1cm} (2.67)

The matrix $p$ is written as

$$p = \begin{bmatrix} p_{xx} & p_{xy} \\ p_{yx} & p_{yy} \end{bmatrix},$$  \hspace{1cm} (2.68)

where elements of the matrix must satisfy the unitary condition:

$$p_{xy} = -p_{yx}^*,$$  \hspace{1cm} (2.69)

$$p_{yy} = p_{xx}^*,$$  \hspace{1cm} (2.70)

$$|p_{xx}|^2 + |p_{xy}|^2 = 1.$$  \hspace{1cm} (2.71)

**Fig. 2.29** DSP circuit for polarization demultiplexing. $p_{yx} = -p_{xy}^*$ and $p_{yy} = p_{xx}^*$ because of unitarity of the matrix $p$.\[\text{Fig. 2.29 DSP circuit for polarization demultiplexing. $p_{yx} = -p_{xy}^*$ and $p_{yy} = p_{xx}^*$ because of unitarity of the matrix $p$.}\]
Following CMA, these matrix elements are updated symbol by symbol as

\begin{align}
p_{xx}(n+1) &= p_{xx}(n) + \mu (1 - |E_X(n)|^2) E_X(n) E_X^*(n), \\
p_{xy}(n+1) &= p_{xy}(n) + \mu (1 - |E_X(n)|^2) E_X(n) E_Y^*(n),
\end{align}

where \( \mu \) is a step-size parameter and \( n \) the number of symbols. By using this algorithm, we can expect that \(|E_X(n)|^2 \to 1\) after the symbol-by-symbol iteration process converges. When \(|E_X(n)|^2 \to 1\), it has already been shown that \( p_{xx}(n) \to r_x \), \( p_{xy}(n) \to k_x \), \( p_{yx}(n) \to -k_x^* \), and \( p_{yy}(n) \to r_x^* \) in Case (I). This result means that the \( x \)-polarization component is restored at the \( X \) port as shown by Eq. (2.63), and the \( y \)-polarization component at the \( Y \) port as shown by Eq. (2.64). On the other hand, in Case (II), we find that the \( x \)-polarization component appears at the \( Y \) port, and the \( y \)-polarization at the \( X \) port as shown by Eqs. (2.65) and (2.66).

When either polarization-mode dispersion (PMD) or polarization-dependent loss (PDL) is involved, unitarity of the matrix \( p \) is not assured. In such a case, the \( Y \) polarization should be aligned independently of the \( X \) component as shown in Fig. 2.30, where \( p_{yy} \) and \( p_{yx} \) are updated by using the following equations:

\begin{align}
p_{yy}(n+1) &= p_{yy}(n) + \mu (1 - |E_Y(n)|^2) E_Y(n) E_Y^*(n), \\
p_{yx}(n+1) &= p_{yx}(n) + \mu (1 - |E_Y(n)|^2) E_Y(n) E_X^*(n).
\end{align}

In this iteration process, we need to carefully choose initial matrix elements for \( p \) in order to prevent the two outputs converging on the same SOP.

![Fig. 2.30 DSP circuit controlling \( X \) polarization and \( Y \) polarization independently. This circuit should be used when unitarity of the matrix \( p \) is not assured](image)
2.3.5 *Equalization of Inter-symbol Interference*

When the optical signal is transmitted through a fiber link, linear impairments stemming from group-velocity dispersion (GVD) of the fiber, tight optical filtering, and electrical bandwidth limitation induce ISI, which degrades the BER performance. Since the restored complex amplitude contains the information on the amplitude and phase of the optical signal, ISI can be equalized by using the transversal filter shown in Fig. 2.31, where the tap spacing is equal to the symbol interval $T$. The spacing of $T/2$ is also employed commonly. The transversal filter is also called the finite-impulse response (FIR) filter. When we represent the received complex amplitude as $x(n)$ where $n$ is the sample number, the output from the transversal filter is given as

$$y(n) = \sum_{i=0}^{k} c_i x(n - i), \quad (2.76)$$

where $c_i$ are complex tap weights and $k$ denotes the number of taps. Theoretically, with a sufficient number of taps, this filter can compensate for any kind of linear impairments, because the filter can generate an arbitrary transfer function.

Tap coefficients can be prefixed by using a given transfer function or adaptively controlled. Two kinds of adaptive control methods are explained in the following. We first define the tap-coefficient vector and the input-signal vector as

$$c(n) = \begin{bmatrix} c_0(n) \\ c_1(n) \\ \vdots \\ c_{k-1}(n) \\ c_k(n) \end{bmatrix}, \quad (2.77)$$

![Fig. 2.31](image) Configuration of the transversal filter for ISI equalization. Delayed replicas are summed up with complex tap weights $c_i$. 
Then the filter output is given as
\[
 y(n) = c(n)^T x(n) = \sum_{i=0}^{k} c_i(n) x(n-i). \tag{2.79}
\]

One of the methods of adaptively controlling tap coefficients is the blind equalization based on CMA discussed in Sect. 2.3.4. In this method, the tap coefficients are updated by
\[
c(n+1) = c(n) + \mu \left\{ 1 - |y(n)|^2 \right\} y(n) x(n)^*, \tag{2.80}
\]
where we do not need decoded symbols for updating tap coefficients. The principle is based on the fact that when ISI is suppressed, the envelope of the PSK signal becomes constant. Equation (2.80) controls tap coefficients so that \( |y(n)|^2 \rightarrow 1 \), and ISI is suppressed in such a case.

Figure 2.32 shows another adaptive-equalization circuit based on the least-mean-square (LMS) algorithm [33]. The tap coefficients are updated by using the LMS algorithm as follows:
\[
c(n+1) = c(n) + \mu \epsilon(n) x(n)^*. \tag{2.81}
\]

In the training mode, we transmit a training signal having a fixed pattern in advance, and the error \( \epsilon(n) \) is given as the difference between the training signal and the filter output \( y(n) \). The tap coefficients are updated until the error tends to zero. Once the tap coefficients have been fixed in the training mode, the operation mode of the
filter is switched to the tracking mode, where the error $\varepsilon(n)$ is given as the difference between the decoded symbol $d(n)$ and the filter output $y(n)$. In the tracking mode, the tap coefficients are adaptively updated so that $|\varepsilon(n)|$ is suppressed even if sources of transmission impairments are time-varying.

If we need to equalize PMD, we should modify the polarization-demultiplexing circuit shown in Fig. 2.29 such that each matrix element becomes a multi-tap transversal filter instead of a one-tap configuration. We define input-signal vectors and tap-coefficient vectors as

$$E_k(n) = \begin{bmatrix} E_k(n) \\ E_k(n-1) \\ \vdots \\ E_k(n-m) \end{bmatrix}, \quad (2.82)$$

$$p_{ij}(n) = \begin{bmatrix} p_{ij,0}(n) \\ p_{ij,1}(n) \\ \vdots \\ p_{ij,m}(n) \end{bmatrix}. \quad (2.83)$$

where $i$, $j$, and $k$ are any one of $x$ and $y$ and $m$ is the number of taps. The tap coefficients can be updated by CMA as

$$p_{xx}(n+1) = p_{xx}(n) + \mu(1 - |E_X(n)|^2)E_X(n)E_X^*(n), \quad (2.84)$$

$$p_{xy}(n+1) = p_{xy}(n) + \mu(1 - |E_X(n)|^2)E_X(n)E_Y^*(n), \quad (2.85)$$

$$p_{yx}(n+1) = p_{yx}(n) + \mu(1 - |E_Y(n)|^2)E_Y(n)E_X^*(n), \quad (2.86)$$

$$p_{yy}(n+1) = p_{yy}(n) + \mu(1 - |E_Y(n)|^2)E_Y(n)E_Y^*(n). \quad (2.87)$$

Updating tap coefficients by Eqs. (2.84), (2.85), (2.86), and (2.87), we can compensate for PMD because the envelope fluctuation is suppressed. At the same time, we can demultiplex dual polarizations as shown in the one-tap case (see Sect. 2.3.4). It is also possible to introduce the least-mean-square (LMS) algorithm instead of CMA.

### 2.4 Performance of the Digital Coherent Receiver

This section deals with evaluation of basic characteristics of the digital coherent receiver. After showing the optical circuit for the homodyne receiver comprising phase and polarization diversities, we discuss the receiver sensitivity limit, the polarization dependence of the receiver sensitivity, and the phase noise tolerance.
2.4.1 Optical Circuit for the Homodyne Receiver Comprising Phase and Polarization Diversities

The optical 90° hybrid can be implemented into the phase-diversity homodyne receiver by using free-space optical components as shown in Fig. 2.33 [19, 20]. Orthogonal states of polarization for the local oscillator (LO) and the incoming signal create the 90° hybrid necessary for phase diversity. With the λ/4 waveplate (QWP), the polarization of the LO becomes circular, while the signal remains linearly polarized and its polarization angle is 45° with respect to principal axes of polarization beam splitters (PBSs). After passing through the half mirror (HM), the polarization beam splitters separate the two polarization components of the LO and signal while two balanced photodiodes PD1 and PD2 detect the beat between the LO and signal in each polarization. When the circularly polarized LO is split with a PBS, the phase difference between the split beams is 90°. On the other hand, there is no phase difference between the split signal beams, because the signal is linearly polarized. Photocurrents from PD 1 and 2 are then given by Eqs. (2.17) and (2.18).

The optical 90° hybrid using planar lightwave circuits (PLC), where the 90° phase shift is created by an optical path-length difference, has also been demonstrated [34]. This is a promising approach for integrating an LO and PDs on a PLC platform.

The optical circuit for the homodyne phase/polarization diversity receiver composed of free-space optical components is shown in Fig. 2.34. In this receiver, two

![Fig. 2.33 Optical circuit for the homodyne phase-diversity receiver. QWP: quarter-wave plate PBS: polarization beam splitter HM: polarization-independent half mirror, Coll.: collimator](image)

![Fig. 2.34 Optical circuit for the homodyne receiver employing phase and polarization diversities. See Fig. 2.33 for abbreviations](image)
homodyne phase-diversity receivers are combined with the polarization diversity
configuration. The incoming signal having an arbitrary SOP is separated into two
polarization components with a polarization beam splitter (PBS). We refer to the
x- and y-polarizations with respect to the principal axes of the PBS. On the other
hand, the local oscillator (LO) is split into two paths with a half mirror (HM), after
its SOP is made circular by a quarter-wave plate (QWP). Right- and left-hand sides
of the receiver then constitute phase-diversity receivers for x- and y-polarizations,
respectively. Photocurrents from PD 1, 2, 3, and 4 are then given by Eqs. (2.26),
(2.27), (2.28), and (2.29), respectively.

2.4.2 Receiver Sensitivity

The back-to-back BER of the BPSK signal is measured to access the sensitivity of
the homodyne-phase-diversity receiver [29]. Data were precoded at the pulse pat-
tern generator (PPG) such that differential decoding of the transmitted data resulted
in a pseudo-random binary sequence (PRBS) (2^7–1). Lasers used as a transmitter
and an LO were 1.55-μm distributed-feedback (DFB) semiconductor lasers, whose
linewidths were about 150 kHz. The laser temperature and bias current were care-
fully maintained via feedback control to keep frequency drifts of the lasers below
10 MHz. The transmitter laser output was modulated through a LiNbO₃ push–pull
modulator to generate a BPSK signal at a bit rate of 10 Gbit/s. The received power
was adjusted with an attenuator and monitored with a power meter. The received sig-
nal was amplified with an erbium-doped fiber amplifier (EDFA) to −10 dBm before
it was detected with the coherent receiver. The maximum LO power was 10 dBm,
which was determined from the allowable power of the photodiodes. The signals
I_{PD1} and I_{PD2} were simultaneously sampled at a rate of 20 Gsample/s with ADCs.
The 3-dB bandwidth of ADCs was 8 GHz. The BER measurement was performed
offline. The collected samples were resampled to keep only one point per symbol
and combined to form a 100-ksymbol-long stream. During phase estimation, we
used the averaging span k = 10. The SOP of the signal was controlled manually so
as to maximize the receiver sensitivity.

Figure 2.35 shows back-to-back BERs measured as a function of the received
power with and without optical preamplification. The power was measured before
the preamplifier in the case with preamplification. The LO power is also changed
when we do not use optical preamplification. The dotted curve is the theoretical
one, which is calculated from the formula

\[ P_e = \frac{1}{2} \text{erfc} \left( \sqrt{\gamma_s} \right), \quad (2.88) \]

where \text{erfc}(\ast) means the complimentary error function and we assume the ideal
CNR given as

\[ \gamma_s = N_s = \frac{P_i T}{h \omega_s}. \quad (2.89) \]
Without optical preamplification, the increase in the LO power improves the receiver sensitivity; however, 10-dB power penalty still remains from the theoretical limit even at $P_{\text{LO}} = 10$ dBm, because we cannot reach the shot-noise limit due to the relatively large circuit noise. On the other hand, with optical preamplification, power penalty is as small as 3 dB, which stems from the spontaneous emission factor larger than 1 and ISI due to bandwidth limitation of the receiver. Since the theoretical power penalty due to preamplification is 0 dB when the spontaneous emission factor of the amplifier $n_{sp} = 1$, the use of the preamplifier is very effective if the LO power is not sufficient. Refer to [29] for sensitivity analyses of coherent receivers with optical preamplification.

### 2.4.3 Polarization Sensitivity

We evaluate the performance of the homodyne phase/polarization diversity receiver in the single-polarization 10-Gbit/s BPSK system [29]. During digital signal processing for polarization alignment and carrier phase estimation, we used $\ell = 10$ and $k = 10$, respectively.

Figure 2.36 shows back-to-back BERs as a function of the received power before the preamplifier. While we change polarization-parameter values of $\alpha(=0, 0.25, 0.5, 0.75, 1)$, BER curves are almost independent of $\alpha$. The power penalty is negligible compared with the case where we do not employ polarization diversity but align the signal SOP optimally ($\alpha = 0, 1$). Thus, we find that the receiver is insensitive to the polarization fluctuation of the incoming signal.

### 2.4.4 Phase Noise Tolerance

The requirement for the laser linewidth in $M$-ary PSK systems becomes much more stringent as $M$ increases. To evaluate the bit-error rate (BER) performance of the
Fig. 2.36 Bit-error rate curves when polarization diversity is introduced. BER curves are almost insensitive to $\alpha$. The power penalty due to polarization diversity is also negligibly small.

Homodyne phase-diversity receiver for the 8PSK case, we conduct a computer simulation where the phase noise variance $\sigma_p^2$ and the phase estimation span $k$ are used as parameters [21]. We assume that the laser phase noise accumulated during a symbol interval $T$ has a Gaussian distribution with the variance $\sigma_p^2 = 2\pi(\delta f_s + \delta f_{LO})T$, where $\delta f_s$ and $\delta f_{LO}$ denote 3-dB linewidths of the transmitter and LO, respectively, and that signals are contaminated by additive white Gaussian noise.

Figure 2.37 shows the simulated BER performance of 8PSK systems calculated as a function of the SNR per bit when $\sigma_p = 0, 1.4 \times 10^{-2}$ and $4.3 \times 10^{-2}$, together with the ideal performance assuming perfect carrier recovery. For comparison, simulated BERs of the differential demodulation scheme are also shown in Fig. 2.37. For 10-Gsymbol/s systems, $\sigma_p$ of $1.4 \times 10^{-2}$ gives a linewidth of $\delta f = 150$ kHz for either transmitter or LO laser, and $\sigma_p$ of $4.3 \times 10^{-2}$ gives 1.5 MHz. Without phase noise, the BER performance of the phase estimation method approaches the ideal performance when $k$ becomes larger; therefore, our phase estimation method is very

\[ \delta f = 0 \quad \delta f = 150 \text{ kHz} \quad \delta f = 1.50 \text{ MHz} \]

Fig. 2.37 Simulation results on BER performance in 8PSK systems for three cases, i.e., the ideal homodyne scheme without phase noise, the differential detection, and the homodyne detection with phase estimation ($k = 2, 5, 10, 25, 100$), as a function of the SN ratio per bit.
effective for obtaining a noiseless accurate phase reference. In contrast, differential demodulation incurs a nearly 3-dB penalty in the SNR, because the reference phase has the same amount of noise as the target phase. As far as the carrier phase is relatively constant over duration of \((2k + 1) T\), the phase estimation provides excellent performance. When \(\sigma_p = 1.4 \times 10^{-2}\), the phase estimation is still efficient and the SNR penalty is less than 1 dB at \(\text{BER} = 10^{-6}\) for \(k = 5, 10, \) and 25. However, the phase noise increase gradually neutralizes the advantage of phase estimation compared to differential demodulation. This is because the phase difference between the first and last samples of \((2k + 1)\) samples becomes significant. For example, in the case of \(\sigma_p = 4.3 \times 10^{-2}\), the error floor appears even when \(k = 5\), and the merit of phase estimation is small. Also note that independently of \(\sigma_p\), the BER performance when \(k = 0\) is identical to that of differential detection. These results indicate that if the laser linewidth is as low as 150 kHz, the acceptable performance in 8PSK systems is obtained by the optimal choice of \(k = 10\).

### 2.4.5 Coherent Demodulation of Multi-level Encoded Signals

We demonstrate coherent demodulation of optical multi-level coded signals. By using distributed-feedback (DFB) semiconductor lasers with linewidths of 150 kHz as a transmitter and a local oscillator, BPSK, QPSK, 8PSK, and 16QAM signals are successfully demodulated at the symbol rate of 10 Gsymbol/s [35].

Figure 2.38(a) shows back-to-back BERs measured as a function of the received power for 10-Gsymbol/s BPSK, QPSK, 8PSK, and 16QAM. A single polarization component is transmitted. At the receiver, the averaging span of phase estimation is optimized for each modulation format. We introduce an adaptive equalizer for suppressing ISI stemming from the receiver-bandwidth limitation. Figure 2.38(b)

![Fig. 2.38 BER performance in BPSK, QPSK, 8PSK, and 16QAM. (a) Measured results, (b) theoretical ones in the shot-noise-limited condition](image-url)
shows theoretical BER curves in the shot-noise-limited condition. With increase in the level of modulation, we find that the power penalty from the theoretical sensitivity increases. This is because the BER performance becomes more sensitive against ISI as the level of modulation increases.

### 2.5 Challenges for the Future

The progress of digital coherent technologies is very rapid. The real-time operation of 46-Gbit/s dual-polarization QPSK receivers has been demonstrated, and the next target is operation at 100 Gbit/s, which aims at transmission of 100-Gbit/s ether signals. As to offline experiments, spectrally efficient transmission has been demonstrated by using higher-level modulation formats such as 8QAM [36] and 16QAM [37]. However, the following is the technical problems that must be tackled and accomplished before the practical coherent optical communication system is realized in the future.

1. Hybrid integration of planar lightwave circuits (PLCs) for phase and polarization diversities, double-balanced photodiodes, and a local oscillator is an important technical task, which enables cost reduction of the coherent receiver and improves system stability.
2. A tunable local oscillator with a narrow linewidth is still the key component of the high-performance coherent receiver.
3. High-speed operation of the coherent receiver relies on the development of high-speed ADC and DSP. The processing speed to cope with \( > 25 \text{-Gsymbol/s} \) systems is desirable for 100-Gbit/s ether applications.
4. More flexible signal processing such as advanced forward error correction (FEC) should be available in the DSP core.
5. For long-distance transmission of multi-level coded optical signals, fiber nonlinearity ultimately limits the system performance. Post-compensation for fiber nonlinearity [38] such as self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM) is an important problem.

The combination of coherent detection and DSP provides new capabilities that were not possible without detection of the phase of the optical signal. We believe that the born-again coherent optical technology will renovate existing optical communication systems in the near future.

### References

31. K. Kikuchi, LEOS Summer Topicals, TuC1.1, Acapulco, Mexico (21–23 July 2008)
33. S. Haykin, Adaptive Filter Theory (Prentice Hall, Englewood, 2001)
High Spectral Density Optical Communication Technologies
Nakazawa, M.; Kikuchi, K.; Miyazaki, T. (Eds.)
2010, X, 338 p. 287 illus., 13 illus. in color., Hardcover
ISBN: 978-3-642-10418-3