Chapter 10

Modulators

Waveguide modulator design and fabrication gives an excellent impression of how rapidly silicon-based photonic components have evolved systematically within the last decades. Very compact and highly transparent waveguides are the base of waveguide modulators. Firstly, waveguides based on silicon nitride/silicon oxide allowed miniaturization up to tenths of microns. Later, silicon/silicon oxide waveguides could be miniaturized to the submicron regime due to their high refractive index contrast. Both waveguide systems have their own merits; silicon waveguides are smaller and more densely packed, but a back-end process makes it easier to fabricate nitride/oxide waveguides. In semiconductor processing, the metallization and insulator layers above the silicon are deposited and structured under more relaxed cleanliness conditions (back end of the line [BEOL]) whereas the processing of the silicon itself requires extremely low contamination levels (front end of the line [FEOL]). Both waveguide systems are transparent in the near-infrared region but offer only a small external influence on the modulation of the light intensity. In the first generation of modulators, this drawback was overcome by a trick composed of three elements. First, the refractive index of one arm of waveguides was modulated by heating (thermo-optic effect)

Silicon-Based Photonics
Erich Kasper and Jinzhong Yu

ISBN 978-981-4303-24-8 (Hardcover), 978-1-315-15651-4 (eBook)
www.jennystanford.com

or by carrier change from diode operation (plasma dispersion effect). The phase of the light wave changes at the end of an arm with a fixed length due to the modulation of the optical length (optical length is geometrical length times refractive index). In a third step, the light from this modulated arm was brought into interference by a Mach–Zehnder interferometer (MZI), with the light from a reference arm translating the phase modulation into intensity modulation. The length of the MZI modulator was rather large (centimeter range) because of the weak matter-light interaction in the transparency regime. The second generation of modulators reduced the lengths dramatically by resonant interference from coupling of light from the waveguide to ring resonators or disc resonators. The third generation of modulators utilizes a direct intensity modulation from electroabsorption effects like the Franz–Keldysh effect (FKE) in bulk semiconductors or the quantum-confined Stark effect (QCSE) in quantum wells (QWs). These electroabsorption modulators (EAMs) operate around the direct bandgap energy reasonably well. The modulation wavelength changes according to the direct bandgap with alloy composition, strain, and QW thickness. Overall, waveguide-integrated Ge and GeSi EAMs [1] have achieved a 3 dB bandwidth greater than 50 GHz, a low energy consumption of ~10 fJ/bit, and a reasonably broad (15–30 nm) operation wavelength range for multichannel wavelength division multiplexing (WDM) applications. The cause of electroabsorption is explained in Fig. 10.1 using the example of a QW. On the left side, one sees the direct-energy gap diagram of a QW with GeSi as the barrier material and Ge as the well material. The carrier wave function and the corresponding lowest energy level are marked for the valence band and the conduction band, respectively.

The bandgap of the QW increases by the quantization energy compared to the bandgap of the bulk well material. With an applied external electric field $\vec{F}$, the QW bandgap decreases but the maxima of the wave functions are spatially separated. The lower bandgap results in a redshift of the absorption, and the spatial mismatch causes a weaker increase in the absorption at higher energies. The electroabsorption also causes a redshift in the indirect band, but this is seldom utilized in devices because of the much lower absorption levels of indirect transitions. In indirect semiconductors with small energy differences between direct and indirect bandgap, like Ge and GeSn, the electroabsorption from the direct bandgap dominates.

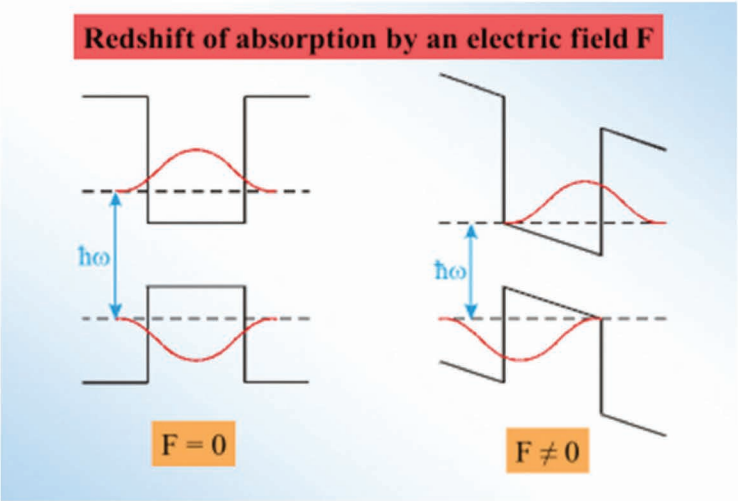

Figure 10.1 Quantum-confined Stark effect.

The modulator is the core device applied in optical cross-connect (OXC) and optical add-drop multiplexing systems; it has great application prospects in the domains of computer and optical communication. Light can be modulated on the basis of the electro-optic effect, the thermo-optic effect, the plasma dispersion effect, FKE, and QCSE, and the modulated light can carry a variety of information, especially information with huge data and high data transmission. In the following parts of the chapter, the main modulation mechanisms in group IV semiconductors and the main electrical and optical structures of silicon modulators are discussed.

10.1 Optical Modulation

There are many parameters to characterize the property of an optical field, such as amplitude (intensity), frequency, phase, and polarization. If some physical methods are used to change one of these parameters regularly with the modulation signal, the light waves can "carry" the information. This is the meaning of optical modulation. The modulation signals often refer to some applied signals, typically electrical, thermal, and optical signals. For the waveguide devices, the most common modulation forms are phase modulation and intensity modulation. Phase modulation is usually

derived from a change in the refractive index of the material involved with the applied signal to change the propagation constant of light in the waveguide. However, the phase change is difficult to detect directly, which is not conducive to the extraction of modulation signals. Therefore, the most commonly used modulation form is intensity modulation, which can be achieved by changing the absorption coefficient of the material involved or converting a phase change based on phase modulation to an intensity change via interferometer or resonator structures. The latter is the often-used way for silicon-based optical waveguide modulators, which can use an MZI or a microring/microdisc resonator to convert phase modulation into intensity modulation.

The parameters of an optical waveguide modulator include modulation depth (MD), extinction ratio (ER), switching time, bandwidth, and insertion loss (IL).

Modulation depth: With the applied signal, P_{max} and P_{min} are the maximum and minimum output power of the modulator. MD is defined as the ratio between the power contrast and the maximum output power:

$$MD = (P_{max} - P_{min})/P_{max} \tag{10.1}$$

A larger MD is better; a waveguide with an MD of 100% has the best modulation.

Extinction ratio: The ER corresponds to the MD, and it is a more commonly used parameter that is defined as:

$$ER = 10 \log(P_{max}/P_{min}) \tag{10.2}$$

Switching time: The switching time is an important parameter for modulator performance. Imposing a square-wave signal in the modulator, the output is also a square-wave signal. The difference between the maximum and minimum output power is defined as $\Delta P = P_{max} - P_{min}$. The rise time t_r is the time experienced by the output power to increase from $P_{min} + 10\%\Delta P$ to $P_{min} + 90\%\Delta P$. Similarly, the fall time t_f is the time experienced by the output power to decrease from $P_{max} - 10\%\Delta P$ to $P_{max} - 90\%\Delta P$. Hence, the switching time t_s of a modulator is the larger value of rise time or fall time.

Modulation bandwidth: Modulation bandwidth can be defined as the modulation frequency while the ER reduces to half of the maximum. It is often associated to the switching time t_s:

$$f_{3dB} = 1/t_s \tag{10.3}$$

Insertion loss: It is the total loss of the entire modulator, including the coupling loss, the propagation loss in the waveguide, and the absorption loss in the modulation region. If the input power and output power are P_{in} and P_{out}, respectively, its IL can be written as:

$$IL = 10\log[(P_{in} - P_{out})/P_{in}] \tag{10.4}$$

10.2 Modulation Mechanisms

Let us first introduce the modulation mechanisms by changing the effective refractive index in silicon. There are many physical interactions to modulate light in silicon, such as thermo-optical, electro-optical, and optical-optical (which includes linear and nonlinear optics) effects. A number of mechanisms, such as the electro-optic effect, the thermally induced index change, the carrier-injection-induced index change (carrier plasma effect), and electroabsorption, classify the interaction. In this section, we will first introduce the modulation mechanisms in group-IV semiconductors.

10.2.1 The Electro-optic Effect

Many optical materials demonstrate an optical anisotropic effect with the applied electric field, including the Pockels effect and the Kerr effect. The electro-optic effect refers to the change in the refractive index of the material with the applied electric field $\vec{E}$. The refractive index change caused by the applied electric field can be expressed as

$$n = n_0 + a\vec{E} + b\vec{E}^2 + \ldots, \tag{10.5}$$

where n_0 is the refractive index of the material without the applied electric field. It is the Pockels effect that the refractive index linearly changes with the electric field, and a is the linear electro-optic coefficient. It is the Kerr effect that the refractive index second-orderly changes with the electric field, and b is the quadratic electro-optic coefficient. The Pockels effect exists only in the noncentrosymmetrical crystal, such as gallium arsenide (GaAs) and lithium niobate ($LiNbO_3$); their Pockels coefficients are 1.6×10^{-10} cm/V and 3.08×10^{-9} cm/V, respectively. These electro-optic crystals

are ideal materials for modulators and can easily be used to achieve low-power and high-speed modulators in discrete arrangements.

Monocrystalline silicon is a typical centrosymmetrical crystal with no Pockels effect and a weak Kerr effect [2]. The refractive index in silicon has a change of 10^{-5} with an applied electric field of up to $3*10^5$ V/cm. This change has been able to meet the requirement of the modulator, but the voltage is near the breakdown voltage of lightly doped silicon, which limits the use in low-power applications.

Another electro-optic effect is connected with the Franz–Keldysh effect (FKE) [3, 4]. With the applied electric field, the energy band in silicon tilts so that electron and hole wave functions extend into the bandgap. The transitions of below-the-bandgap energy may occur with the influence of the photon-associated tunneling as shown in Fig. 10.2a. If the energy of the incident light is close to the bandgap, the absorption spectrum will shift to the long wavelength with the applied electric field, which can be used to fabricate electroabsorption modulators (EAMs) by direct intensity modulation.

A change in the refractive index following from a change of absorption coefficient (Kramers–Kroning rule) will dominate indirect bandgaps, which can be used for phase modulation. Figure 10.2b shows the changes in refractive index connected with the FKE. It has been demonstrated that the change in the refractive index caused by the FKE is about 1.3×10^{-5} when wavelength and applied electric field are 1.07 μm and 10^5 V/cm.

Note that the influence of the FKE both for absorption and for refractive index is strong only around the bandgap energy [5]. The indirect bandgap energy of silicon is about 1.12 eV at room temperature. Optical transitions from an indirect bandgap need a momentum transfer from photons, which results in two cut-off energies—1.06 eV and 1.18 eV—for the optical phonon replica (optical phonon energy in Si: about 60 meV).

The band structure of SiGe alloys up to high Ge amounts (80%) is similar to that of Si; only the absolute value of the bandgap energy decreases from 1.12 eV to 0.83 eV. Accordingly, phase modulation based on refractive index changes from FKE may be used with SiGe alloys between about 1.1 eV and 0.8 eV.

The direct bandgap transition of Ge and Ge-rich SiGe alloys dominates the optical absorption behavior of these indirect materials with a small energy difference to the direct transition. For these

materials the direct intensity modulation utilizing the absorption redshift [6] of the FKE is more attractive.

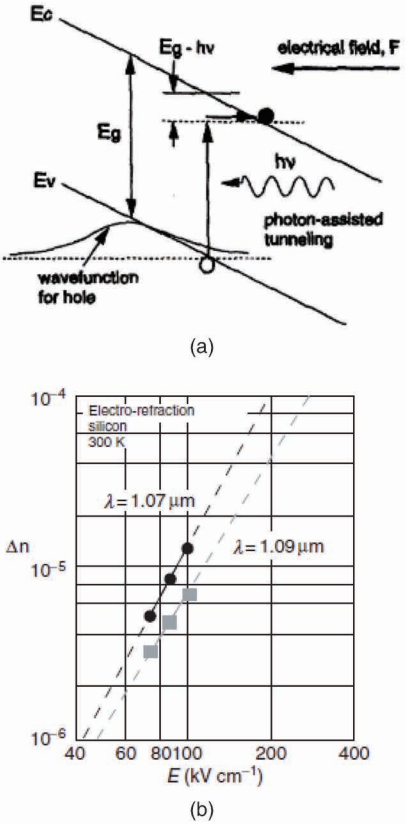

Figure 10.2 (a) the Franz–Keldysh effect causes a tilting energy band in silicon with the applied electric field; (b) the relationship between the changes in the refractive index of silicon and the applied electric field with wavelengths at 1.07 μm and 1.09 μm. © 1987 IEEE. Reprinted, with permission, from Ref. [2].

Alloy layers and quantum wells (QWs) from GeSn extend the wavelength regime of FKE modulators beyond the 1.55 μm band [7].

10.2.2 The Thermo-optic Effect

The material refractive index n changes with the temperature T, which is the thermo-optic effect. The temperature increases lead

to an increase in the effective index in silicon. Silicon has a large thermo-optical coefficient ($\partial n/\partial T \approx 1.86 \times 10^{-4}$ K^{-1}) of 1.55 μm, which is about 2 times and 15 times larger than that of $LiNbO_3$ and SiO_2, respectively. Silicon also has large thermal conductivity (~1.4 × 10^5 W/K). We can fabricate modulators with simple structures by using the thermo-optic effect. However, the response speed of the thermo-optic effect is relatively slow; for the submicron SOI optical modulator, the fastest switching time can just reach microsecond timescales. The thermo-optic effect is very useful for principal investigation of modulators and for modulator fine adjustment of operation at a certain wavelength.

10.2.3 The Plasma Dispersion Effect

The plasma dispersion effect means that the changes in the concentration of free carriers lead to the changes in the refractive indexes and absorption coefficients of semiconductor materials. Silicon has a notable plasma dispersion effect (carrier injection or depletion). It has been demonstrated that we can achieve a high speed modulation in this effect. The effect currently is the main operating foundation for silicon optical modulator.

We can derive the expressions for the refractive index and absorption coefficient in silicon as follows, assuming an effective mass model for the bandgap:

$$n = n_0 - \frac{e^2\lambda_0^2}{8\pi^2\varepsilon_0 n_0 c^2}\left[\frac{N_e}{m_e^*} + \frac{N_h}{m_h^*}\right] \tag{10.6}$$

$$\alpha = \frac{e^3\lambda_0^2}{4\pi^2\varepsilon_0 n_0 c^3}\left[\frac{N_e}{(m_e^*)^2\mu_e} + \frac{N_h}{(m_h^*)^2\mu_h}\right] \tag{10.7}$$

Here n and n_0 are the refractive indexes with and without carrier injection, respectively; N is the concentration of the free carriers, m^* is the effective mass, μ is the mobility, c is the velocity of light, and the subscripts e and h represent electron and hole, respectively. The above two equations are the corresponding relation between the refractive index and absorption coefficient in the semiconductor material, namely the free carrier plasma dispersion effect of the material. Going by Eqs. 10.6 and 10.7, if the concentration of the

free carrier increases, the refractive index will decrease but the absorption will be enhanced. The refractive index and the absorption coefficient can be simultaneously changed through a change in the concentration of the free carrier. Therefore, we can make use of the changes in the refractive index (Fig. 10.3) and absorption (Fig. 10.4) to change the phase and intensity of light, achieving light modulation.

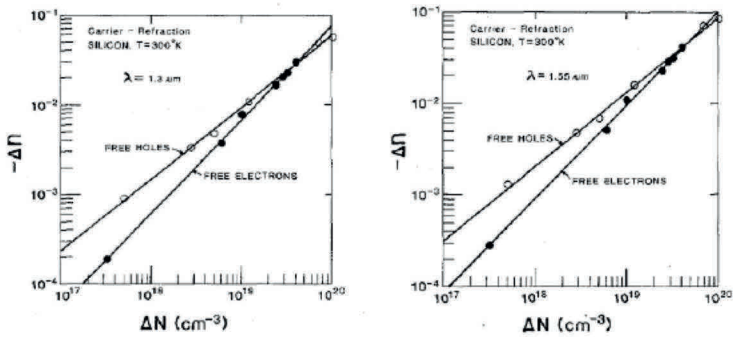

Figure 10.3 Refractive index change with the concentration of free carriers. (a) λ_0 = 1.3 μm and (b) λ_0 = 1.55 μm. © 1987 IEEE. Reprinted, with permission, from Ref. [2].

Figures 10.3 and 10.4 show the variation in the refractive index and absorption coefficient in silicon with a change in the concentration of the free carrier at the wavelengths of 1.3 μm and 1.55 μm. Soref and Bennett [2] used the Kramers–Kroning dispersion relation and experimental data for the absorption spectrum to obtain approximate expressions of the free carrier plasma effect in silicon. The experimental values (Eqs. 10.8 and 10.9) deviate in detail from the simple bandgap model (Eqs. 10.6 and 10.7) because of the real sub-band structure with different effective masses.

$$\lambda_0 = 1.3\,\mu\text{m} \Rightarrow \begin{cases} \Delta n = -\left[6.2\times10^{-22}\Delta N_e + 6.0\times10^{-18}(\Delta N_h)^{0.8}\right] \\ \Delta\alpha = 6.0\times10^{-18}\Delta N_e + 4.0\times10^{-18}\Delta N_h \end{cases} \tag{10.8}$$

$$\lambda_0 = 1.55\,\mu\text{m} \Rightarrow \begin{cases} \Delta n = -\left[8.8\times10^{-22}\Delta N_e + 8.5\times10^{-18}(\Delta N_h)^{0.8}\right] \\ \Delta\alpha = 8.5\times10^{-18}\Delta N_e + 6.0\times10^{-18}\Delta N_h \end{cases} \tag{10.9}$$

Here Δn is the change in the refractive index, $\Delta\alpha$ (cm^{-1}) is the change in the absorption coefficient, ΔN (cm^{-3}) is the change in free carrier concentration, and the subscripts e and h represent electron and hole, respectively. Comparing with the Kerr effect and the FKE, the changes in the refractive index caused by the carrier plasma effect are much larger for indirect bandgap semiconductors. It was the primary choice to achieve silicon electro-optic modulator.

Direct intensity modulation from the electroabsorption effect enlarged the modulator toolbox since heterostructures on Si are available in research and foundry service.

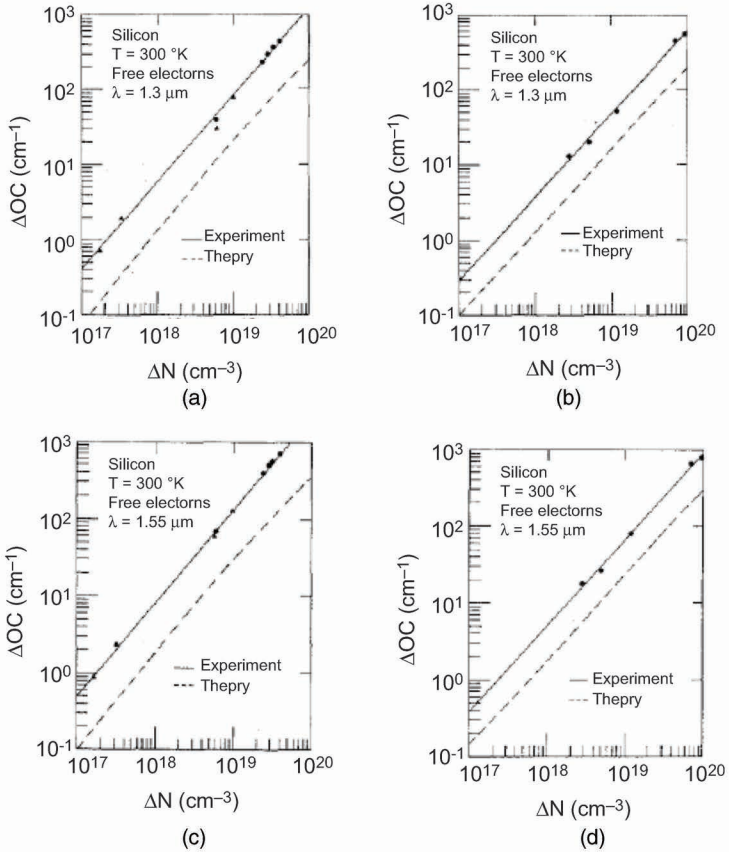

Figure 10.4 The relation between the absorption coefficient and the concentration of the free carrier (electron and hole) in theory and experiment. (a) Electrons, 1.3 μm. (b) Holes, 1.3 μm. (c) Electrons, 1.55 μm. (d) Holes, 1.55 μm. © 1987 IEEE. Reprinted, with permission, from Ref. [2].

10.2.4 Strain and Quantization Effects

10.2.4.1 Strain-induced linear electro-optic effect

Monocrystalline silicon has no linear electro-optic effect because it is a centrosymmetrical crystal. However, a significant linear electro-optic effect can be induced in silicon by breaking the crystal symmetry. The symmetry is broken by depositing a straining layer on top of a silicon waveguide [8]. When a straining layer of SiO_2 or Si_3N_4 is deposited on top of an SOI waveguide, it will expand the structure underneath in both directions horizontally. So the crystal symmetry is broken, which leads to a linear electro-optic coefficient in silicon, and the refractive index will have linear changes with the applied electric field. The linear electro-optic coefficient induced by strain in silicon is measured to be 15 pm/V. It is much smaller than the linear electro-optic coefficients induced by the commonly applied electro-optic crystals, such as $LiNO_3$ (360pm/V), but is comparable with those induced by other noncentrosymmetric semiconductors, like GaAs. In addition, the waveguide structure with a large group refractive index can increase the electro-optic coefficient of strained silicon. Photonic crystal waveguide often has a large group refractive index. It has been demonstrated that silicon can obtain the electro-optic coefficient of ~830 pm/V in the photonic crystal waveguide.

High strain values are obtained by lattice-mismatched heterostructures. Strain shifts considerably the bandgap energies [9], which allows an adjustment of operating wavelengths.

10.2.4.2 Quantum-confined Stark effect

The quantum-confined Stark effect (QSCE) is exhibited when an applied electric field is perpendicularly added on the top of QW structures [10]. It is the quantum-confined version of the FKE in bulk materials. The intensity is larger because the thin QW forces the wave function overlap between hole and electron states even under strong electric fields. Thin QW structures made from direct bandgap III–V semiconductors such as GaAs and InP and their alloys exhibit a strong Stark effect [11] mechanism, which allows modulator structures with only several microns of optical path length. The QCSE, at room temperature, in thin germanium QW structures grown on silicon has been reported [12], which causes the absorption spectrum of the QW to red-shift to a long wavelength with the applied electric

field. The QCSE is widely used in the III–V QW structures (such as InGaAs/InAlAs) to fabricate the EAM. However, the QCSE is usually very weak in silicon because of the indirect bandgap. However, a very strong QCSE of the direct transition exists in the GeSi/Ge QW structure. The maximum effective absorption coefficient change in this QW structure is 2800 cm^{-1} at 1438 nm under a 3 V bias. Going by the result, the structure is very promising for small, high-speed, low-power EAMs. Low energy requirement in datacom pushes the use of EAM based on the FKE and the QCSE. All this success [13] indicates that GeSi EAMs are well on the way to system integration in small-footprint, ultra-low-energy, high-speed photonic datalinks.

10.3 Structures for the SOI Waveguide Modulator

It has been widely acknowledged that the plasma dispersion effect and the electroabsorption are the fastest modulation mechanisms for the silicon-based electro-optic modulator. The silicon electro-optic modulator based on the plasma dispersion effect is usually composed of two parts: an electrical structure and an optical structure [14–25]. To change the carrier concentration distribution, many different electrical structures can be used to achieve the injection, accumulation, depletion, and inversion of free carrier, so the refractive index (or absorption coefficient) can be changed correspondingly. There are three commonly used electrical structures: a forward-biased p-i-n diode, a reverse-biased p/n junction, and a metal-oxide-silicon (MOS) capacitor. Optical structures are mainly used for guiding the light and converting phase modulation to intensity modulation. The commonly used optical structures can be classified as interference, total internal reflection, and absorption, where interference also can be classified as the Mach–Zehnder interferometer (MZI), the Fabry–Perot (F-P) resonator, the microring resonator (MRR), and the photonic crystal waveguide.

10.3.1 Electrical Structures

10.3.1.1 Forward-biased p-i-n diode structure

Figure 10.5 shows the forward-biased p-i-n structure in the waveguide modulator: the waveguide core region is the lightly

doped intrinsic region, and the top of rib region or slab region on both sides of the rib region are the heavily doped P^+ region and N^+ region, respectively. The forward-biased p-i-n diode is a simple bipolar electrical unit and also the most commonly used electrical structure. With a forward bias, a large number of nonequilibrium carriers are injected into the intrinsic region. Because of the carrier plasma dispersion effect, the refractive index and absorption coefficient in the rib waveguide are changed simultaneously. The nonequilibrium carriers in the high-impedance lightly doped silicon have a certain diffusion length, so there is a relative large overlap between the varying region of the refractive index and the optical field region, which causes high modulation efficiency.

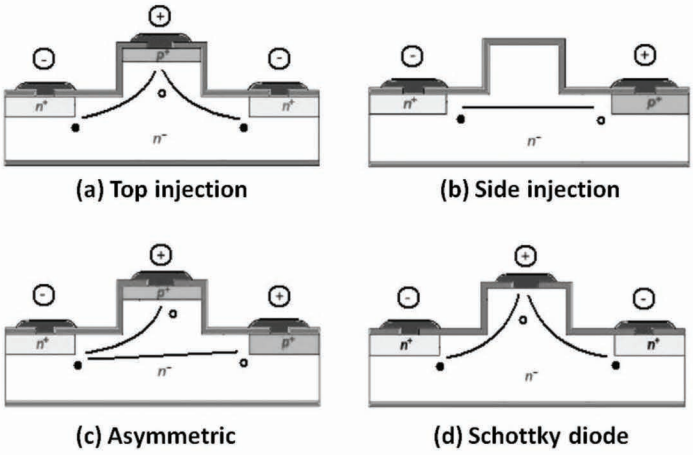

Figure 10.5 Different forward-biased p-i-n structures in the waveguide modulator.

There are four main forms for p-i-n electrical structures in the rib waveguide, as shown in Fig. 10.5, where the p- and n-doped types can be exchanged with each other. The physical factors of the carrier injection, drift, diffusion, absorption, and loss should be considered for design to enhance the modulation efficiency and speed as much as possible. The lateral-injection p-i-n diode structure is the most tolerant to fabrication technology, and it also can enhance the modulation speed by reducing the distance between heavily doped regions on both sides of the rib; hence, the lateral-injection p-i-n diode structure is used often for device design.

10.3.1.2 Reverse-biased p/n junction structure

The waveguide cross section of the modulation region of the electro-optic modulator based on a reverse-biased p/n junction is shown in Fig. 10.6. It is composed of a p/n junction and two electrode layers of p^+ and n^+ without the i-region, like in the above structure. The p/n junction operates with a reverse bias. The medium-doped p-type and n-type silicon in the waveguide core region compose a horizontal or vertical p/n junction. There is a certain carrier distribution in the waveguide without the applied modulation; however, if a reverse bias is imposed, the width of the depletion region will increase and the concentration of carriers around the p/n junction will decrease; then the refractive index in the depletion region will be changed by the changes in the concentrations of carriers.

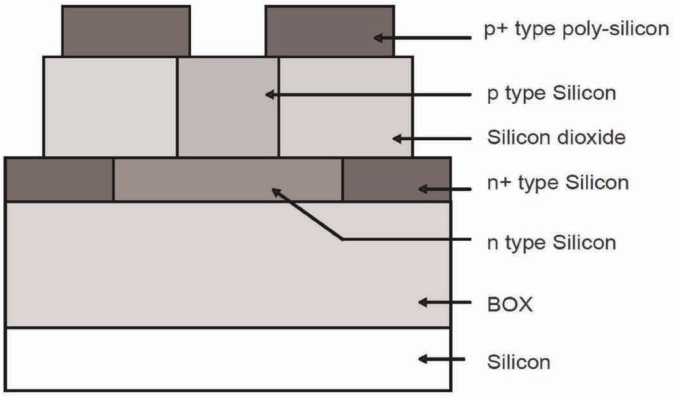

Figure 10.6 Cross-section schematic of a silicon modulator based on a reverse-biased p/n junction structure.

The p/n junction is in the state of carrier depletion with a reverse bias, and carriers mainly drift in the waveguide with the applied electric field, so it has a relative high response speed. Theoretical analysis shows that this modulator can achieve a high-speed modulation of up to several hundred gigahertz. However, its modulation efficiency is relatively low because of limited extension of the depleted region and its small overlap with the optical field. Simultaneously, absorption in the waveguide increases due to a doped

core layer, so the propagation loss of the device correspondingly increases. The modulation efficiency of this structure is difficult to enhance directly, but with strong resonant structures, such as MRRs, it can reduce the amount of refractive index changes in the requirement of electro-optic modulation, which helps achieve better modulation effects. In addition, the modulation speed can be enhanced further.

10.3.1.3 MOS capacitor structure

Figure 10.7 shows the cross-sectional view of a silicon electro-optic modulator based on an MOS capacitor [20]. The first high-speed silicon modulator with a modulation bandwidth exceeding 1 GHz based on this structure was demonstrated. It comprises an n-type doped crystalline silicon slab (the silicon layer of the SOI wafer) and a p-type doped polysilicon rib with a thin gate oxide layer sandwiched between them. With the applied bias, free carriers accumulate on both sides of the gate oxide layer, and then the changes in the refractive index caused by the changes in the carrier concentration achieve the optical modulation in the device.

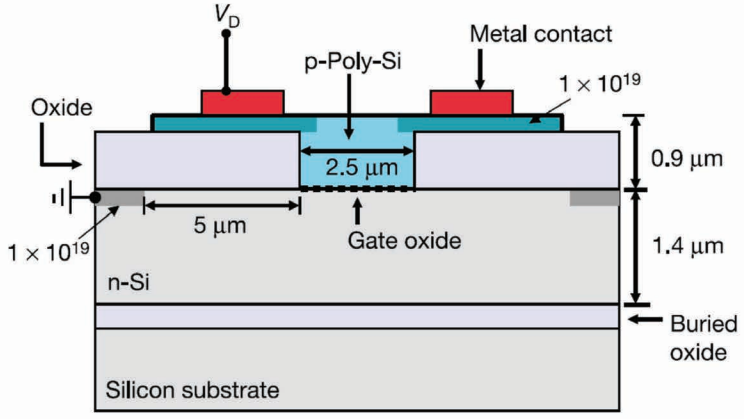

Figure 10.7 The cross section of a silicon modulator based on an MOS capacitor. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, *Nature*, Ref. [20], copyright (2004).

The carriers drift rather than inject or diffuse in the MOS capacity structure waveguide with the applied electric field, so it can obtain

a high-speed modulation in the structure. The modulation speed of the device is mainly dependent on the resistance in the waveguide layer between electrodes and the capacity of the gate oxide layer. Free carriers move in the device controlled by voltage rather than current, so the power consumption and modulation interference by the electro-optic effect are relatively small in the device. However, the region in which carriers accumulate is relative small due to the thin gate oxide layer and the corresponding overlap with optical field is also relatively small, so its total modulation efficiency is not high. The device has strong polarization dependence due to the introduction of the gate oxide layer; the modulation efficiency of TE mode is much larger than the one of TM mode. It is more important that the modulation structure has high requirements for fabrication technology. In order to reduce power consumption and enhance modulation efficiency, the gate oxide layer should be thin as much as possible when ensure it not to be broken down with the applied voltage. The interface between the gate oxide and silicon should be smooth as much as possible to reduce the propagation loss of optical field. The quality of polysilicon on the top of gate oxide is also related to the propagation loss of optical field. Therefore, although the structure can achieve a high speed modulation, it is difficult to fabricate and less tolerant to fabrication technology.

Table 10.1 shows the comparison of the three electrical structures.

Table 10.1 Comparison of the three electrical structures

Electrical structure	Forward biased p-i-n diode	Reverse biased PN junction	MOS capacitor
Modulation efficiency	High	Low	Low
Fabrication	Easy	Difficult	Difficult
Power consumption	Large	Small	Small
Influence of thermal-optic effect	Large	Small	Small
Optical absorption	Weak	Strong	Strong
Modulation speed	Low	High	High

10.3.2 Optical Structures

10.3.2.1 Mach–Zehnder interferometer structure

The MZI structure is composed of the input/output waveguides, two 3 dB couplers (a splitter and a combiner), and two modulation arms (Fig. 10.8). The incident light beam is split into two beams by the splitter. They enter the two modulation arms, propagate a certain distance in the modulation arm, and finally combine into a single beam to provide the output light. A certain optical path difference will be induced through a change in the refractive index of one of the two arms. This optical path difference modulates the magnitude of phase difference between both beams, and the magnitude of phase difference determines the optical field distribution that is generated after the two beams interfere through a combiner. Therefore, it is possible to achieve the modulation or switching function of light by changing the refractive index of one of the two arms. Light propagating in the MZI achieves a conversion from phase modulation to intensity modulation.

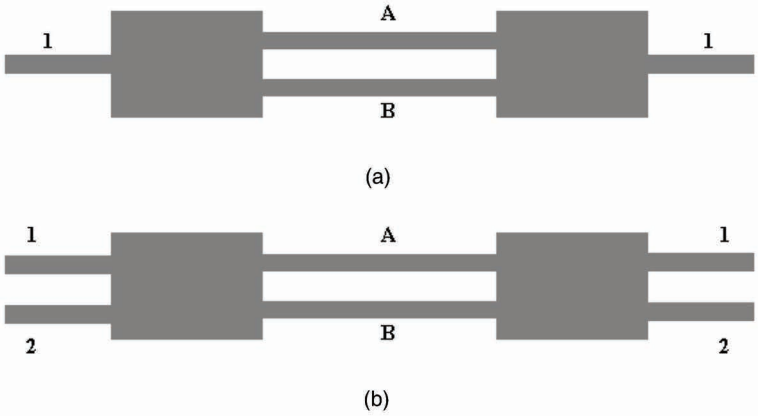

Figure 10.8 Two main MZI configurations: (a) 1 × 1 MZI; (b) 2 × 2 MZI.

The operation principle of the 1 × 1 MZI is relatively simple. Let us take 1 × 1 MZI as an example to derive the operation principle of the MZI. Assuming the splitter in the input port and the combiner in the output port are symmetrical, the splitting ratio of the splitter in the input port is α:(1 – α), the additional loss factor is α, the phase

difference of light in the two output ports is $\Delta\varphi$, the lengths of the two modulation arms are L_A and L_B, the effective refractive indices in the two modulation arm waveguides are n_{effA} and n_{effB}, and the propagation losses are a_A and a_B. After light of power P_{in} enters from the input port and goes through the splitter, the input optical fields in both arms can be expressed as:

$$E_{1A} = \sqrt{\alpha a P_{in}} \exp[j(\omega t + \Delta\varphi_s)] \tag{10.10}$$

$$E_{1B} = \sqrt{\alpha(1-a)P_{in}} \exp[j\omega t] \tag{10.11}$$

At the end of the two arms, the output optical field can be expressed as:

$$E'_{1A} = \sqrt{\alpha a P_{in}} \exp[j(\omega t + \Delta\varphi_s)]\exp(-jk_0 n_{effA} L_A - \alpha_A L_A) \tag{10.12}$$

$$E'_{1B} = \sqrt{\alpha(1-a)P_{in}} \exp[j\omega t]\exp(-jk_0 n_{effB} L_B - \alpha_B L_B) \tag{10.13}$$

The output port is a 2×1 combiner, its output optical field is the interference between the two input fields E'_{1A} and E'_{1B}, and the additional loss factor β, then we can get the output optical field

$$E_{11} = \sqrt{\beta}\{\sqrt{\alpha a P_{in}} \exp[j(\omega t + \Delta\varphi_s)]\exp(-jk_0 n_{effA} L_A - \alpha_A L_A) + \sqrt{\alpha(1-a)P_{in}} \exp[j\omega t]\exp(-jk_0 n_{effB} L_B - \alpha_B L_B)\} \tag{10.14}$$

Therefore, the optical power in the output waveguide is

$$P_{out} = |E_{11}|^2 = \alpha\beta P_{in}\{a\exp(-2\alpha_A L_A) + (1-a)\exp(-2\alpha_B L_B) + 2\sqrt{a(1-a)}\exp(-\alpha_A L_A - \alpha_B L_B)\cos[\Delta\varphi_s - k_0(n_{effA} L_A - n_{effB} L_B)]\} \tag{10.15}$$

From the above analysis, the factors of the beam splitting, the beam combining, the balance between optical fields in both arms, and the phase difference in the MZI all affect the interference and superimposition of the output optical field, and then directly determine the corresponding modulator operation characteristics.

The MZI optical structure can obtain a high ER and a wide optical bandwidth, and the structure is easy to fabricate and has a large fabrication tolerance, but the refractive index change caused by the plasma dispersion effect is often very small. A very long modulation

arm is required to achieve a π phase shift, so the length of the device based on the MZI structure may be very long, up to the millimeter order of magnitude.

10.3.2.2 Fabry–Perot resonator structure

The resonator structure usually can confine the optical field in a small region, so its propagation characteristics are very sensitive to the refractive index changes in the resonator. Based on the principle, even though the refractive index change caused by the plasma dispersion effect is very small, the silicon-based modulators also can be achieved with small dimensions. There are many common used optical resonator structures based on silicon waveguides, including F-P resonators, MRRs, microdisc resonators.

An F-P resonator is a resonator structure widely used in laser. It typically constitutes of two parallel highly reflective mirrors and a medium (or waveguide) between both mirrors, as shown in Fig. 10.9a. When an incident light travels into the left reflective mirror, a part of light is reflected and the other part of light is transmitted. The transmitted light travels through the medium (or waveguide) between both mirrors with a certain phase shift, and then travels into the right mirror; similarly, a part of light is reflected and the other part of light is transmitted. The partial reflected light travels in the medium (or waveguide) in the reverse direction again, and then is reflected by the left mirror. Again and again, there exists the superposition of multiple transmitted lights in the output port. They have a fixed phase difference that is related to light wavelength and optical cavity length. For some specific wavelengths, transmitted light in the output port is in-phase while for other wavelengths, the transmitted light in the output port is out of phase. Therefore, the transmission characteristics of the F-P resonators have resonance peaks. Figure 10.9b shows the transmission spectrum of an F-P resonator based on a silicon waveguide.

If carriers are injected into the waveguide between both reflective mirrors, the refractive index in the waveguide will change, resulting in the change in the optical cavity length of F-P resonators, so the resonance peaks in the transmission spectrum change, as shown in Fig. 10.9b. For a certain fixed wavelength, the transmissivity of total F-P cavity changes, which cause the change in the output intensity, so it can achieve the modulation function. The modulation speed of F-P

modulators is related to not only electrical structures but also optical transmission characteristics. The modulation is achieved mainly by the shift of resonance peaks, if the quality factor (Q) in the F-P cavity is larger, the resonance bandwidth will be narrower, and the refractive index change, required for shifting the same wavelength, is smaller. Therefore, for the same ER, a modulator based on an F-P resonator of high Q can achieve a short switching time.

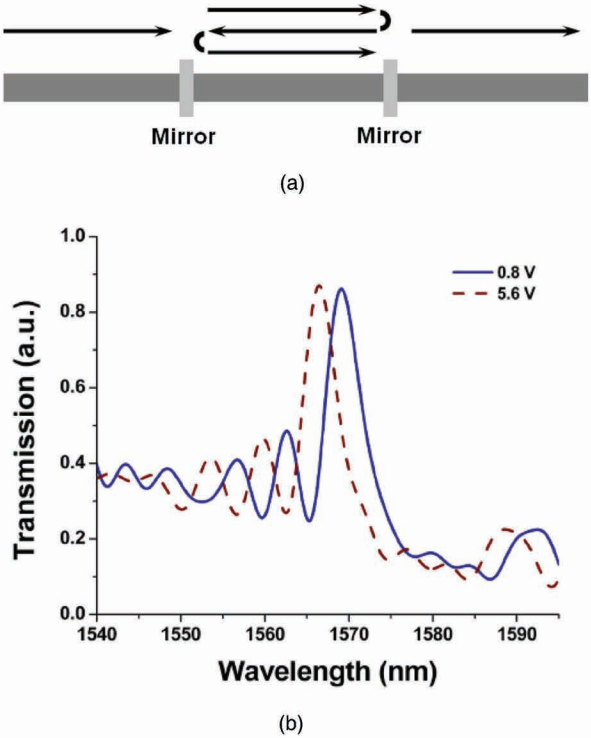

Figure 10.9 (a) The schematic of an F-P resonator; (b) the transmission spectrum of an F-P resonator based on silicon waveguide. Reprinted with permission from Ref. [14] © The Optical Society.

10.3.2.3 Microring resonator structure

Marcatili [26] first proposed the optical MRR in 1969. Recently, MRR has been attracted wide attention in silicon-based photonic devices by researchers. An MRR is an important optical structure with wavelength selectivity, so it can be used to control the light

transmission path; and the optical field intensity in the resonant cavity is very high, so it also can be used for many nonlinear optics experiments. An MRR is an ideal basic building block for very-large-scale-integrated (VLSI) optical circuits, which has been widely used in fabricating a serial of passive or active optical functional devices, such as add-drop filters, multiplexers/de-multiplexers, lasers, modulators, switches, gates, delay lines, sensors, etc.

An MRR comprises a ring or racetrack waveguide and one or two straight waveguides coupled with the ring or racetrack waveguide, as shown in Fig. 10.10a. Unlike the F-P resonator, the transmitted light and reflected light export from different waveguides: the transmitted light travels through the through port and the reflected light travels through the drop port; hence, with MRRs, it is easier to achieve optical switches with multi-input and multi-output. Light can be coupled into the ring via an evanescent field through the input waveguide. The ring will be resonant for light whose phase change after each full round trip in the ring is an integer multiple of 2π. The light coupled into the ring from the input waveguide will be in phase with the original light in the ring and constructive interference, on the other hand, the light coupled back into the input bus will be reverse-phase with the original light in the input bus and destructive interference. The light in the ring can achieve a very high intensity and export through the drop port by coupling into the output bus. While the phase change after each full round trip in the ring is not an integer multiple of 2π, light coupled into the ring will be very small and most of them transmit directly through the through port, that is, the MRR is transparent to the nonresonant light.

An MRR is functionally the same as the F-P resonator, however, comparing with the F-P resonator, the quality factor Q of an MRR is much higher, so it is more conducive to achieve a high speed electro-optic modulator. Its output spectrum is similar to the F-P with resonance peaks. The resonance peaks can be shifted by changing the refractive index in the ring waveguide via electrical structures, and then the modulation in the MRR structure is achieved. Figure 10.10b shows the output spectrum in the through port of an MRR where an electrical structure of forward biased p-i-n diode is embedded to inject carriers. The solid line represents the output spectrum in the through port without the injection of carriers, and the dotted lines represent the output spectrums in the through port with different

applied biases. It can be clearly seen that the output optical power in the through port increases, while the applied bias increases for some specific wavelengths.

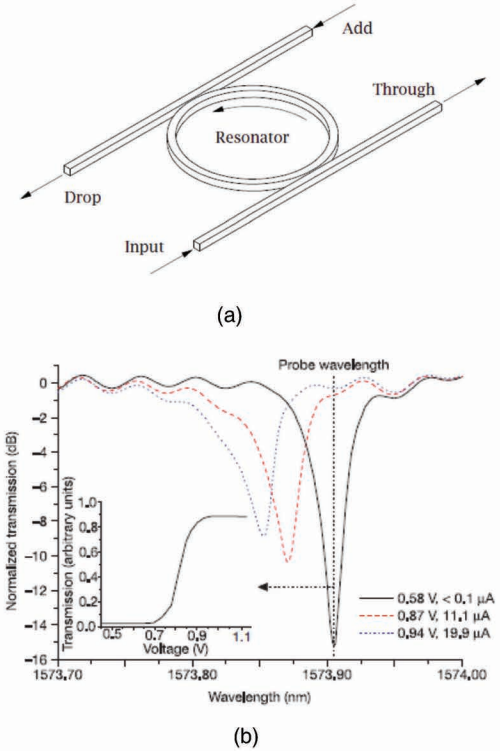

Figure 10.10 (a) The schematic of a microring structure; (b) the output spectrum of a ring resonator. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, Ref. [17], copyright (2005).

If a disc waveguide is used to replace the ring waveguide in the MRR, the microdisc resonator is generated. A microdisc resonator can achieve a higher Q, and its optical response speed is much higher than electrical response speed, so it is expected to achieve a higher-frequency modulator.

Both the F-P resonator and the microring (microdisc) resonator can greatly reduce the size of modulators to micrometer order of magnitude, which is a great advantage for large scale optoelectronic

integration. However, resonator modulators can only operate in the specific wavelength range with a narrow wavelength band, which requires a carefully consideration on the operating wavelength of device for the design. In addition, because the refractive index change is very sensitive in the resonator structure, some parasitic effects have an impact on the performance of electro-optic modulators. For instance, the device based on the electrical structure of p-i-n diode generates heat due to the injection of current. However, the thermo-optic effect and plasma dispersion effect is competing with each other, and their refractive index changes is just opposite. If the device design is not reasonable, the plasma dispersion effect may mask the electro-optic effect.

10.4 Summary

Two technical directions turned out to be dominating research and applications of modulators in Si-based photonic systems. (i) First, Si waveguide modulators based on phase modulation from the plasma dispersion effect matured utilizing different CMOS technologies [27] for resonant interference (microring, microdisc) in small-footprint devices. (ii) Heterostructures (Si/Ge and GeSn/Ge) that are necessary for detectors and light sources paved the way for recent progress in EAMs, which give direct intensity modulation from Franz–Keldish and QCSE [1, 13]. System integration of Ge/GeSi EAM has gained momentum driven by the energy efficiency requirements in datacom. Recent highlights are high speed greater than 50 GHz, low energy consumption ~10 fJ/bit, and a reasonably broad (15–30 nm) operation wavelength range for multichannel wavelength division multiplexing (WDM) applications. Research groups [28, 29] demonstrated a 16-channel GeSi EAM-photodiode array coupled to multicore fibers to achieve 896 Gbps for short-reach optical links, and they demonstrated a 100 Gbps nonreturn to zero (NRZ) on-off key (OOK) transmission format using GeSi EAM. The direct intensity modulation by EAM is especially integration friendly because the electrical and optical structures for detection, light emission, and modulation are identical. The electrical structure is given by a diode [30] that is operated for detection at zero bias or at small reverse bias. For light emission the diode is driven in the forward direction, and for modulation the diode is switched between reverse bias

and zero bias. The optical structure needs coupling between the waveguide and the active device that is done by butt coupling or evanescent coupling, same for all three devices.

Acknowledgments

We thank Xuejun Xu and Kang Xiong for the collection of basic information and figure drawings.

References

1. X. Wang and J. Liu (2018). Emerging technologies in Si active photonics, *J. Semicond.*, **39**, 061001-1–061001-29.
2. R. A. Soref and B. R. Bennett (1987). Electrooptical effects in silicon, *IEEE J. Quantum Electron.*, **23**, 123–129.
3. W. Franz (1958). Einflus eines elektrischen Feldes auf eine optische Absorptionskante, *Z. Naturforsch.*, **13a**, 484–489.
4. L. V. Keldysh (1957). Behaviour of non-metallic crystals in strong electric fields, *JETP*, **33**, 994–1003.
5. A. Frova, P. Handler, F. A. Germano and D. E. Aspnes (1966). Electroabsorption effects at the band edges of silicon and germanium, *Phys. Rev.*, **145**, 575–583.
6. M. Schmid, M. Kaschel, M. Gollhofer, M. Oehme, J. Werner, E. Kasper and J. Schulze (2012). Franz-Keldysh effect of germanium-on-silicon p-i-n diodes within a wide temperature range, *Thin Solid Films*, **525**, 110–114.
7. M. Oehme, K. Kostecki, M. Schmid, M. Kaschel, M. Gollhofer, K. Ye, D. Widmann, R. Koerner, S. Bechler, E. Kasper and J. Schulze (2014). Franz-Keldysh effect in GeSn pin photodetectors, *Appl. Phys. Lett.*, **104**, 161115.
8. R. S. Jacobsen, K. N. Andersen, P. I. Borel, J. Fage-Pedersen, L. H. Frandsen, O. Hansen, M. Kristensen, A. V. Lavrinenko, G. Moulin, H. Ou, C. Peucheret, B. Zsigri and A. Bjarklev (2006). Strained silicon as a new electro-optic material, *Nature*, **441**, 199–202.
9. Y. Ishikawa, K. Wada, D. D. Cannon, J. Liu, H.-C. Luan and L. C. Kimerling (2003). Strain-induced band gap shrinkage in Ge grown on Si substrate, *Appl. Phys. Lett.*, **82**, 13.

10. D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood and C. A. Burrus (1984). Quantum-confined Stark effect, *Phys. Rev. Lett.*, **53**, 2173.

11. J. Stark (1913). Splitting of spectral lines of hydrogen in an electric field, *Scientific talks*, *Nobel Prize in Physics 1919*.

12. Y.-H. Kuo, Y. K. Lee, Y. Ge, S. Ren, J. E. Roth, T. I. Kamins, D. A. B. Miller and J. S. Harris (2005). Strong quantum-confined Stark effect in germanium quantum-well structures on silicon, *Nature*, **437**, 1334–1337.

13. J F. Liu (2016). *Photonics and Electronics with Germanium* (Willey-VCH Verlag).

14. B. Schmidt, Q. Xu, J. Shakya, S. Manipatruni and M. Lipson (2007). Compact electro-optic modulator on silicon-on-insulator substrates using cavities with ultra-small modal volumes, *Opt. Express*, **15**, 3140–3148.

15. X. Tu (2008). Study on the silicon-on-insulator optical waveguide electro-optical modulator, PhD thesis, Institute of Semiconductors, Chinese Academy of Sciences.

16. A. R. M. Zain, N. P. Johnson, M. Sorel and R. M. De La Rue (2008). Ultra high quality factor one dimensional photonic crystal/photonic wire micro-cavities in silicon-on-insulator (SOI), *Opt. Express*, **16**, 12084–12089.

17. Q. Xu, B. Schmidt, S. Pradhan and M. Lipson (2005). Micrometre-scale silicon electro-optic modulator, *Nature*, **435**, 325–328.

18. L. Friedman, R. A. Soref and J. P. Lorenzo (1988). Silicon double-injection electrooptic modulator with junction gate control, *J. Appl. Phys.*, **63**, 1831–1839.

19. A. Cutolo, M. Iodice, P. Spirito and L. Zeni (1997). Silicon electro-optic modulator based on a three terminal device integrated on a low-loss single-mode SOI waveguide, *J. Lightwave Technol.*, **15**, 505–518.

20. A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu and M. Paniccia (2004). High-speed silicon optical modulator based on a metaloxide-semiconductor capacitor, *Nature*, **427**, 615–618.

21. Q. F. Xu, S. Manipatruni, B. Schmidt, J. Shakya and M. Lipson (2007). 12.5 Gbit/s carrier-injection-based silicon micro-ring modulators, *Opt. Express*, **15**, 430–436.

22. L. Liao, D. Samara-Rubio, M. Morse, A. Liu, D. Hodge, D. Rubin, U. D. Keil and T. Franck (2005). High speed silicon Mach-Zehnder modulator, *Opt. Express*, **13**, 3129–3135.

23. F. Y. Gardes, G. T. Reed, N. G. Emerson and C. E. Png (2005). A sub-micron depletion-type photonic modulator in silicon on insulator, *Opt. Express*, **13**, 8845–8854.

24. A. Liu, L. Liao, D. Rubin, H. Nguyen, B. Ciftcioglu, Y. Chetrit, N. Izhaky and M. Paniccia (2007). High-speed optical modulation based on carrier depletion in a silicon waveguide, *Opt. Express*, **15**, 660–668.

25. W. M. J. Green, M. J. Rooks, L. Sekaric and Y. A. Vlasov (2007). Ultra-compact, low RF power, 10 Gb/s Silicon Mach-Zehnder modulator, *Opt. Express*, **15**, 17106–17113.

26. E. A. J. Marcatili (1969). Dielectric rectangular waveguide and directional coupler for integrated optics, *Bell Syst. Tech. J.*, **48**, 2071–2102.

27. G.-S. Jeong, W. Bae and D.-K. Jeong (2017). Review of CMOS integrated circuit technologies for high-speed photo-detection, *Sensors*, **17**, 1962.

28. P. De Heyn, V. I. Kopp, S. A. Srinivasan, et al. (2017). Ultra-dense 16×56 Gb/s NRZ GeSi EAM-PD arrays coupled to multicore fiber for short-reach 896Gb/s optical links, *Optical Fiber Communication Conference*, Th1B.7.

29. J. Verbist, M. Verplaetse, S. A. Srivinasan, et al. (2017). First real-time 100-Gb/s NRZ-OOK transmission over 2 km with a silicon photonic electro-absorption modulator. *Optical Fiber Communication Conference*, Th5C.4.

30. M. Oehme, E. Kasper and J. Schulze (2013). GeSn heterojunction diode: detector and emitter in one device, *ECS J. Solid State Sci. Technol.*, **2**, 76–78.