41

Nanostructure Solar Cells

41.1 Introduction: Nanostructures in Photovoltaics 441
41.2 Physical Mechanisms of NSSC Device Operation 444
Photogeneration of Electron–Hole Pairs • Charge Separation and Extraction • Recombination
41.3 Modeling of NSSC Device Characteristics 454
Global Detailed Balance Theories for Ideal Systems • Hybrid Models: Semiclassical Transport with Detailed Balance Rates • Fully Microscopic Quantum-Kinetic Models
41.4 Case Study: Single Quantum Well p-i-n SC 461
States • Dynamics • Characteristics
41.5 Summary 465

Urs Aeberhard

41.1 Introduction: Nanostructures in Photovoltaics

The operation of solar cell (SC) devices is governed by a large variety of solid-state properties associated with the interplay of optical, electronic, and vibrational degrees of freedom of the component materials. Light-matter interaction provides the elementary optoelectronic processes of photogeneration and radiative recombination, while coupling of charge carriers to lattice vibrations leads to energy conversion losses due to thermalization, dissipative transport, and nonradiative recombination.

In the presence of spatial inhomogeneities, the optoelectronic material properties change if the characteristic length scale of the spatial variation is reduced below a certain threshold. The critical length scale depends on the physical degree of freedom (optical, electronic, vibrational) and corresponds to the wavelength of the associated wave (electromagnetic [EM] wave, de Broglie wave, sound wave). In this *subwavelength* regime, the physics is governed by the corresponding wave equations (Schrödinger, Maxwell, harmonic lattice dynamics), with the peculiar features such as interference or nonlocality. Along with the states, the interaction of different degrees of freedom is strongly modified in the presence of confinement, mainly due to the effect of localization and symmetry breaking. This tunability of physical properties via variation of material composition, size and shape of structures below a characteristic length scale—termed *nanostructures* (NSs)[†]—provides the basis for the vast field of nanotechnological applications, and the relation between configurational parameters and target NS functionality is a major topic of past and current research. Almost since the advent of semiconductor NS fabrication technologies, such as molecular beam epitaxy (MBE) or metal-organic chemical vapor deposition (MOCVD), but also lithographic processes,

[†] While the term *nano* is used for all of the above degrees of freedom (nanophotonics, -electronics, -phononics) to distinguish the confinement regime from bulk behavior, the characteristic confinement length varies strongly (0.1 μm to 1 cm for photons, 0.5 to 10 nm for electrons, 5 nm to 1 cm for phonons).

the peculiar physical properties of the NS produced in this way have also been utilized for the design of novel SC devices with the potential of increased photovoltaic (PV) conversion efficiency. In later years, new NS fabrication techniques such as nanoimprinting, catalytic vapor deposition, or solution-based processes have enlarged the zoo of NSs and NS architectures proposed for SC with either enhanced efficiency or the potential for substantial cost reduction.

A common way to characterize NSs is according to the dimensionality: The material without confinement, i.e., featuring extended states in all three spatial dimensions, is referred to as *bulk*. For electronic confinement in one, two, and three spatial dimensions, the resulting 2D, 1D, and 0D NSs are called *quantum well* (QW), *quantum wire* (QWR), and *quantum dot* (QD), respectively. At length scales where only optical confinement is present, the corresponding objects are usually called thin films, nanowires, and nanoparticles, respectively. On the other hand, the functionalities of nanostructure solar cell (NSSC) components can be categorized according to the physical degree of freedom that is to be affected: *Optical* functionalities are used for the purpose of light trapping for absorption enhancement via engineering of the EM field strength inside the absorber; the (opto-)*electronic* functionalities, which are related to the tuning of (opto-)electronic properties such as bandgaps, absorption coefficients, mobilities, and carrier relaxation rates via electronic structure engineering; and *vibrational* functionalities, which affect the dissipation of carrier energy due to coupling to lattice vibrations, for instance, via inhibition of suitable phonon modes. As explained earlier, all of the functionalities emerge as the consequence of the nanoscale dimension of the structures that leads to quantization and confinement of the photonic, electronic, and vibrational states and results in physical properties that can deviate considerably from the bulk behavior.

The main applications of NS in photovoltaics are in the so-called second and third generations of PV devices [1], where NS are used in the implementation of SC concepts aiming at either lower cost through strongly reduced material usage and cheap production technology or at increased energy conversion efficiencies through enhanced spectrum utilization and/or reduced thermalization losses. NS with optical functionality have been instrumental in the realization of thin-film SC (second generation) based on low absorption materials, such as thin-film silicon devices [2] or ultrathin absorbers [3]. Relevant representatives of such NS applications for light-trapping include plasmonic metal nanoparticles [4] and nanostructured metal films [5], dielectric scatterers [6] and diffraction gratings [7], engineering of the optical density of states (DOS) via photonic crystals [8], light focusing antenna effects in nanowires and nanorods [9], as well as spectral conversion via QDs [10]. This wide field of optical NS for PV applications is covered by a number of topical reviews [11–14].

The electronic functionality of NS concerns first of all the tuning of optical transitions and nonradiative charge carrier dynamics in high-efficiency SC architectures for concentrator and space applications (third generation). Since the early nineties, multi-QW (MQW—multiple decoupled QW) and QW superlattice (QWSL—periodic and coupled QW) structures have been implemented in different III-V semiconductor materials and have been investigated for application in single-junction SC [15] and multijunction (MJ) PV devices [16], and, more recently, also for hot-carrier SC (HCSC) [17]. The practical relevance of these structures lies mainly in the adjustability of the absorption edge to the wavelength required for efficiency optimization, in the case where suitable bulk materials are not available. Similar bandgap engineering features are provided by multi-QD (MQD) or QD superlattice (QDSL) structures implemented in III-V semiconductors [18–21] or silicon alloys [22], with the advantage of improved strain relaxation mechanisms. However, while QWSCs based on strain-balanced components have enabled efficiencies comparable to the bulk limit [23], severe recombination losses did not allow thus far a similar achievement in QDSC [24]. On the other hand, QD structures were considered to be particularly useful for the implementation of a number of specific third-generation PV concepts: for intermediate band SC (IBSC) due to the possibility to obtain an energetically separated intermediate state or band from the ground state of the QD [25]; for multiple exciton generation (MEG) via impact ionization (providing increased current) due to enhanced electron–electron interaction as a consequence of wave function localization [26]; and for the extraction of hot carriers in HCSC (providing increased voltage) due to inhibited electron–phonon scattering resulting from the low density and spectral sparsity of electronic and vibrational modes [27]. These effects have

been demonstrated in dedicated experimental setups [28–33], but they remain to be shown in working SC devices.

Besides these highly ordered epitaxial QD structures, there is a whole class of QD materials made as colloids from solution, such as, e.g., PbSe and PbS [34], which are arranged into different SC architectures, ranging from QD-sensitized TiO_2 electrodes to hybrid organic–inorganic devices and closely packed QD solids with carrier-selective contacting schemes [35–38]. While present efficiencies are still low (11.3%), these approaches promise low-cost manufacturing, and there is still the prospect of MEG exploitation as observed in the constituent particles [39].

For the sake of completeness, a word is due here on the case of QWRs. Even though they offer a perfect combination of size tunability in transverse dimensions with extended states in transport direction, there are almost no experimental structures including 1D NS with electronic functionality. The reasons are twofold. On the one hand, it is very challenging to produce wires with diameters small enough such as to exhibit electronic confinement. On the other hand, it is not possible to achieve at the same time optical and electronic confinement: At the dimensions suitable for optical confinement, the absorber is electronically bulk-like, and no guided optical modes can exist in the tiny dimensions required for electronic confinement.

As compared to the optical and electronic features emerging at the nanoscale, the vibrational functionalities of NSs have received the least attention so far. Most potential applications consider exploitation of reduced carrier cooling by engineering the phonon modes in NSs such as QWSL [40] and QDSL [22,27]. The issue of confined phonons is discussed also in the context of fast relaxation of hot photogenerated excitons in isolated QD, one of the detrimental factors suppressing MEG, which was thought to be reduced in QD. Recent experiments and simulations emphasize the influence of softened surface modes enabling efficient multiphonon relaxation [41].

All of the above functionalities might be relevant in a single SC device, for instance a HCSC based on a III-V QWSL structure (similar to [40]) with absorption enhancement induced by plasmonic nanoparticles or a cavity effect. In such a case, deviations from bulk physics need to be considered for photons, electrons, and phonons at the same time. Furthermore, while the single functionality of an NS component (e.g., enhancement of interband absorption due to wave function localization) might be beneficial for a certain aspect of device operation (e.g., photogeneration), it may at the same time have a detrimental impact on another aspect (e.g., recombination and transport). Indeed, in experimental implementations of advanced PV concepts, NSs are often found to introduce electronic and optical losses, e.g., due to high defect density at NS interfaces and parasitic absorption into localized states with insufficient carrier extraction efficiency. In the idealized models considered for the determination of efficiency bounds such as the famous Shockley–Queisser (SQ) limit [42], such losses are usually neglected. This explains to some extent the large discrepancy between theoretical ideal case predictions and experimentally achieved results. However, for a realistic assessment of the technological potential of a given NSSC concept, the detrimental aspects associated with the NS need to be considered and quantified. This underlines the importance of advanced modeling approaches reaching beyond bulk semiconductor physics for this category of PV devices, which will be in the focus of the remainder of this chapter.

Space does not permit a detailed discussion of the modeling approaches for all of the NS used in photovoltaics as outlined above. Since nanophotonic aspects of optoelectronic device modeling are covered elsewhere in this book, NS with purely optical functionality—including the important case of the nanowire SC, where the absorber is at the same time its own light-trapping structure—will not be addressed here. The other important class of NSSC that will not be covered comprises the colloidal QD SC architectures. In these devices, the PV processes are strongly influenced by the surface chemistry, such as, e.g., the properties associated with different ligand molecules, but also by strong carrier localization, large exciton binding energies, and the considerable fluctuations in size and position. As a consequence, mobility is usually limited by thermally activated hopping processes [43], resulting in poor transport properties. The disorder-induced smearing of the band edge has also unfavorable impact on the achievable open-circuit voltages, even at the radiative limit and under the assumption of perfect transport [44]. Overall, the situation is

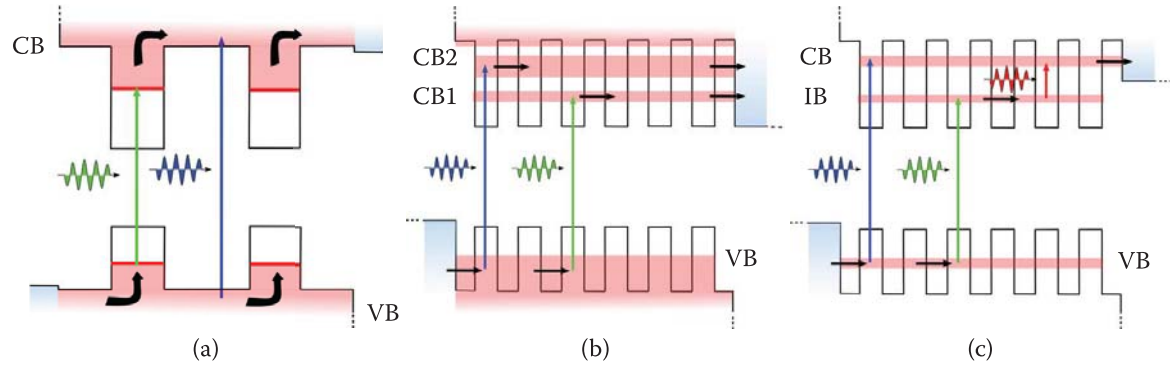


FIGURE 41.1 Common architectures of NSSC: (a) multi-QW or QD structures, where the confined states of distinct NS are not coupled, and photogenerated carriers therefore need first to escape to extended (quasi-)continuum states in order to contribute to current; (b) superlattices of QW or QD, where miniband formation or sequential tunneling enables photocarrier transport and extraction via partially localized NS states; (c) in contrast to (b), the IBSC configuration features noncontacted NS states, which, however, can still be delocalized, but, contrary to (a), should be only radiatively coupled to the contacted continuum states by which carriers are extracted. CB, conduction band; VB, valence band; IB, intermediate band.

closer to the case of organic and hybrid SC, for which a whole range of dedicated simulation approaches exist [45].

The main focus of this chapter, however, will be on the modeling of *regimented* QW and QD architectures as utilized for the implementation of high-efficiency SC concepts using *inorganic* semiconductor materials. Among these structures, three main categories can be distinguished according to the participation of the NS states in the transport of charge carriers. The first category includes devices where the NS states are localized in transport direction and need to be coupled to another set of extended (i.e., bulk-like or unconfined) states to enable charge carrier flow. Typical representatives are the MQW and MQD SC (Figure 41.1a). The second category consists of architectures where both generation-recombination and transport processes are mediated by NS states, which requires a finite degree of wave function delocalization in transport direction related to the coupling of NS, as in QDSL devices with large confinement barriers (Figure 41.1b). Finally, there are structures where coupling-induced delocalization exists between both confined and unconfined states, but only the continuum states are connected to contacts, as in some QD-IBSC implementations (Figure 41.1c).

In Section 41.2, the main physical mechanisms governing the PV device operation are discussed for the different types of NS architectures with regard to their theoretical description and modelization. In Section 41.3, existing methods are reviewed to obtain the device characteristics of NS-based SCs under consideration of the modeling requirements resulting from Section 41.2 and the associated implementation challenges, establishing a simulation hierarchy ranging from global detailed balance approaches for limiting efficiency calculations to microscopic quantum-kinetic theories for the local charge carrier dynamics in nonclassical device regions. In Section 41.4, the simulation of a prototypical NSSC in the form of a single QW photodiode is considered in the light of the approaches discussed in Section 41.3.

41.2 Physical Mechanisms of NSSC Device Operation

In the following, the different stages of the PV energy conversion process in NSSC—from the incident photon flux to the extracted charge current—are considered with focus on the modeling requirements imposed by the presence of NS components.

41.2.1 Photogeneration of Electron–Hole Pairs

The structures that are of interest in NS PV architectures are small compared to the coherence length of sunlight. The spectral *photon flux* $\Phi_\gamma(\mathbf{r}, \hbar\omega)$ at energy $\hbar\omega$, with $\hbar$ the reduced Planck constant and ω the frequency of the light, and position $\mathbf{r}$ inside the absorber of the SC under consideration is therefore well described by the propagation of coherent EM radiation in a dispersive medium, as provided by the solution of (macroscopic) Maxwell's equations. In the most general case, the linear response† of the medium to the transverse electrical field $\boldsymbol{\mathcal{E}}^{\mathrm{t}}$ is characterized by the complex microscopic dielectric tensor $\vec{\varepsilon}(\mathbf{r}, \mathbf{r}', \hbar\omega)$, which considers the electronic excitations in dependence of the energy and polarization of the incident radiation field and of the available electronic states. The corresponding local volume rate $\mathcal{G}$ for the generation of electron–hole pairs by photon absorption is obtained from the dissipated EM energy as follows [46]:

$$\mathcal{G}(\mathbf{r}, \hbar\omega) = \eta_{\mathrm{gen}}(\mathbf{r}, \hbar\omega)(2\hbar\mathcal{V})^{-1} \sum_{\mu,\nu} \int d^3r' \, \Im \left\{ \varepsilon_0 \varepsilon_{\mu\nu}(\mathbf{r}, \mathbf{r}', \hbar\omega) \mathcal{E}^{t*}_{\mu}(\mathbf{r}, \hbar\omega) \mathcal{E}^{t}_{\nu}(\mathbf{r}', \hbar\omega) \right\}, \tag{41.1}$$

where $\Im$ denotes the imaginary part, ε_0 is the vacuum permittivity, η_{gen} is the generation efficiency ($\eta_{\mathrm{gen}} = 1$ means that each photon that is absorbed generates exactly one electron-hole pair), and $\mathcal{V}$ is the absorbing volume. In the above expression, the time-harmonic monochromatic Ansatz $\boldsymbol{\mathcal{E}}^{\mathrm{t}}(\mathbf{r}, t) = \Re\{\boldsymbol{\mathcal{E}}^{\mathrm{t}}(\mathbf{r}, \hbar\omega) \exp(-i\omega t)\}$ was used. While the interaction with the light is inherently nonlocal from the electronic point of view, an averaged local version of the dielectric tensor can be used if the EM field varies only slowly over the length scale of electronic inhomogeneities. For an isotropic medium and unit generation efficiency, Equation 41.1 can then be rewritten as

$$\mathcal{G}(\mathbf{r}, \hbar\omega) = (2\hbar)^{-1} \varepsilon_0 \Im\varepsilon(\mathbf{r}, \hbar\omega) |\boldsymbol{\mathcal{E}}^{\mathrm{t}}(\mathbf{r}, \hbar\omega)|^2 \approx \alpha(\mathbf{r}, \hbar\omega)\Phi(\mathbf{r}, \hbar\omega), \tag{41.2}$$

where

$$\alpha(\mathbf{r}, \hbar\omega) \equiv \frac{\omega}{n_{\mathrm{r}} c_0} \Im\varepsilon(\mathbf{r}, \hbar\omega) \tag{41.3}$$

denotes the (locally defined) *absorption coefficient* that reflects the decay of the optical intensity due to all optical excitations possible at a given photon energy, weighted by their respective strength (c_0: speed of light in vacuum, n_{r}: refractive index of the background).

41.2.1.1 Photon Flux

In the general time-harmonic case, the photon flux is given by the Poynting vector $\mathbf{S}$ of the transverse EM fields ($\boldsymbol{\mathcal{E}}^{\mathrm{t}}, \boldsymbol{\mathcal{B}}^{\mathrm{t}}$):

$$\mathbf{S}(\mathbf{r}, \hbar\omega) = (2\mu_0)^{-1} \left\{ \boldsymbol{\mathcal{E}}^{\mathrm{t}}(\mathbf{r}, \hbar\omega) \times \boldsymbol{\mathcal{B}}^{t}(\mathbf{r}, \hbar\omega) \right\} \equiv \Phi(\mathbf{r}, \hbar\omega)\hbar\omega \, \hat{\mathbf{n}}(\mathbf{r}, \hbar\omega), \tag{41.4}$$

where μ_0 is the vacuum permeability and $\hat{\mathbf{n}}$ is the unit vector in the direction of propagation. The evolution of the flux inside the absorber is related to the EM dissipation due to light-matter coupling via the optical conservation law:

$$\nabla \cdot \mathbf{S}(\mathbf{r}, \hbar\omega) = -\frac{1}{2}\omega\varepsilon_0 \, \Im \left\{ \boldsymbol{\varepsilon}(\mathbf{r}, \hbar\omega) [\boldsymbol{\mathcal{E}}^{\mathrm{t}}(\mathbf{r}, \hbar\omega)]^* \cdot \boldsymbol{\mathcal{E}}^{\mathrm{t}}(\mathbf{r}, \hbar\omega) \right\}, \tag{41.5}$$

† This is appropriate for standard solar illumination conditions, i.e., the AM1.5g spectrum.

where the transverse electrical field is given by Maxwell's equations acquiring the form

$$\mu_0^{-1}\nabla\times\nabla\times\boldsymbol{\mathcal{E}}^{\mathrm{t}}(\mathbf{r},\hbar\omega)=\omega^2\varepsilon_0\boldsymbol{\varepsilon}(\mathbf{r},\hbar\omega)\boldsymbol{\mathcal{E}}^{\mathrm{t}}(\mathbf{r},\hbar\omega), \tag{41.6}$$

$$\nabla\cdot\varepsilon_0\boldsymbol{\varepsilon}(\mathbf{r},\hbar\omega)\boldsymbol{\mathcal{E}}^{\mathrm{t}}(\mathbf{r},\hbar\omega)=0. \tag{41.7}$$

The *absorptance* of the absorbing volume is then given by the ratio $a\equiv P_{\mathrm{abs}}/P_{\mathrm{in}}$ of absorbed and incident EM power, with $P_{\mathrm{abs}}=\int_{\mathcal{V}}d\mathcal{V}(-\nabla\cdot\mathbf{S})$ and $P_{\mathrm{in}}=\int_{\mathcal{S}}d\boldsymbol{S}\cdot\mathbf{S}$, where $\mathcal{S}$ is the surface of the absorbing volume. In the absence of optical confinement—for instance, in optically thick absorbers—the decay of the flux can be written in terms of the absorption coefficient via the *Lambert–Beer* (LB) law, e.g., for propagation along the z-direction:

$$\Phi_{\mathrm{LB}}(z,\hbar\omega)=\Phi_0^{\swarrow}(\hbar\omega)\exp\left(-\int_0^z dz'\,\alpha(z',\hbar\omega)\right), \tag{41.8}$$

with $\Phi_0^{\swarrow}$ denoting the incident photon flux at $z=0$. This defines the absorptance $a_{\mathrm{LB}}(z,\hbar\omega)\equiv 1-\Phi_{\mathrm{LB}}(z,\hbar\omega)/\Phi_0^{\swarrow}(\hbar\omega)$ after an optical path of length z. For planar systems featuring optical elements such as antireflection coatings (ARC) or a back reflector, as frequently used in thin-film PV devices due to the requirement of absorption enhancement via light trapping, the required solution of Maxwell's equations can be performed using a transfer-matrix method (TMM) [47], which relates the coefficients of incident, reflected, and transmitted plane wave components for piecewise constant dielectric function. In the case of 2-D/3-D nanophotonic or plasmonic light trapping structures, Maxwell's equations (Equation 41.7) have to be solved by advanced numerical approaches such as the finite-difference time-domain (FDTD) method or the finite element method (FEM).

41.2.1.2 Absorption Coefficient

Equations 41.2 and 41.3 can be used to determine the absorption coefficient from the complex dielectric function or the linear *susceptibility* $\chi(\omega)=\varepsilon(\omega)-\varepsilon_b$ ($\varepsilon_b\equiv n_r^2$: background dielectric constant) or, equivalently, from the net absorption rate $r_{\mathrm{abs}}=\mathcal{G}/\eta_{\mathrm{gen}}$ (net means after subtraction of the rate for stimulated emission). For weak perturbation, the net volume rate of uncorrelated direct electronic transitions between states v and c by absorption of photons of energy $E_\gamma=\hbar\omega$ is given by *Fermi's Golden Rule* (FGR):

$$r_{v\to c}^{\mathrm{abs}}(E_\gamma)=\frac{2\pi}{\hbar\mathcal{V}}|\langle c|\hat{H}_{e\gamma}|v\rangle|^2\delta(\varepsilon_c-\varepsilon_v-E_\gamma)(f_v-f_c), \tag{41.9}$$

with $\hat{H}_{e\gamma}$ the electron–photon interaction Hamiltonian, $\varepsilon_{c,v}$ the energies of single electron and hole states, and $f_{c,v}$ the occupation of these states. The electron–photon Hamiltonian has the general form

$$\hat{H}_{e\gamma}(\mathbf{r},t)=\frac{e}{2m_0}\left\{\hat{\mathbf{A}}(\mathbf{r},t)\cdot\hat{\mathbf{p}}+\hat{\mathbf{p}}\cdot\hat{\mathbf{A}}(\mathbf{r},t)\right\}+\frac{e^2[A(\mathbf{r},t)]^2}{2m_0}\approx\frac{e}{m_0}\hat{\mathbf{A}}(\mathbf{r},t)\cdot\hat{\mathbf{p}}, \tag{41.10}$$

where $\hat{\mathbf{A}}$ is the EM vector potential, $\hat{\mathbf{p}}=-i\hbar\nabla$ is the momentum operator, and for the last expression (minimal coupling), the assumption of weak vector potential $\hat{\mathbf{A}}$ was made and Coulomb gauge ($\nabla\cdot\mathbf{A}=0$) was used. In the standard *semiclassical* picture of light-matter interaction in semiconductors, the light is treated as a classical wave. Using for the incident field, a monochromatic time-harmonic plane wave expansion:

$$\mathbf{A}(\mathbf{r},t)=\frac{A_0}{2}\boldsymbol{\epsilon}\left\{e^{i(\mathbf{q}\cdot\mathbf{r}-\omega t)}+c.c.\right\} \tag{41.11}$$

of the vector potential, with amplitude $A_0/2$, wave vector $\mathbf{q} = \frac{\omega}{c}\hat{\mathbf{n}}$, and polarization $\boldsymbol{\epsilon}$, together with the minimal coupling form of $H_{e\gamma}$ in the FGR rate (Equation 41.9) yields

$$r^{\text{abs}}_{v\to c}(E_\gamma) = \frac{2\pi}{\hbar V}\frac{e^2A_0^2}{4m_0^2}\mathcal{M}^2_{\text{cv}}\delta(\varepsilon_{cv} - E_\gamma)(f_v - f_c), \tag{41.12}$$

$$\mathcal{M}_{\text{cv}} \equiv |\langle c|e^{i\mathbf{q}\cdot\mathbf{r}}\boldsymbol{\epsilon}\cdot\hat{\mathbf{p}}|v\rangle|, \tag{41.13}$$

where $\varepsilon_{cv} \equiv \varepsilon_c - \varepsilon_v$. If $\mathbf{q}\cdot\mathbf{r} \ll 1$ (i.e., the wavelength of the light is much larger than the characteristic length scale of the absorbing structure), the phase factor $e^{i\mathbf{q}\cdot\mathbf{r}}$ in the matrix element can be neglected, which corresponds to the *dipole approximation*, meaning that contributions of higher order electric and magnetic multipoles are not considered. In this approximation, the coupling via the momentum operator $\hat{\mathbf{p}}$ can be replaced by a coupling via the dipole operator $\hat{\mathbf{d}} = -e\hat{\mathbf{r}}$ ($\hat{\mathbf{r}}$: position operator) due to the relation $(e/m_0)\mathbf{A}\cdot\hat{\mathbf{p}} \approx -\hat{\mathbf{d}}\cdot\boldsymbol{\mathcal{E}}^{\text{t}}$ [48]. The photon flux associated with Equation 41.11 amounts to

$$\Phi_0^{\swarrow}(\hbar\omega) = \frac{n_r c_0 \varepsilon_0 \omega A_0^2}{2\hbar}, \tag{41.14}$$

which, together with Equation 41.12, provides the general FGR absorption coefficient:

$$\alpha^{\text{FGR}}(E_\gamma) = \frac{\pi\hbar}{n_r c_0 \varepsilon_0 E_\gamma \mathcal{V}}\left(\frac{e}{m_0}\right)^2 \sum_{v,c}\mathcal{M}^2_{cv}\delta(\varepsilon_{cv} - E_\gamma)(f_v - f_c), \tag{41.15}$$

where the sum is over all initial and final states.

In Equation 41.15, the occupation of the initial and final states is assumed to be known, and in the case of interband transitions, $f_v = 1$ and $f_c = 0$ are often assumed. However, this is no longer appropriate in the case of transitions between partially filled subbands, as for instance in the case ot the IBSC concept. Thus, in general, the occupation needs to be determined by a microscopic treatment of the charge carrier dynamics under illumination in the framework of the semiconductor Bloch equations [48–50] or nonequilibrium Green's functions [51].

The main quantities determining the generation of electron-hole pairs are, thus, the local value of the transverse EM field on the one hand, and the local value of the dipole or momentum matrix elements (MMEs) and of the density of occupied initial and empty final states of the optical transitions, respectively, on the other hand. While the spatial variation of the transverse EM fields is small over the extent of semiconductor NS in the range of few nanometers, the electronic quantities reflect the impact of symmetry breaking, wave function localization, and energy quantization associated with reduced dimensionality. A prominent example is silicon QDs, where the contribution of direct transitions—absent in bulk—and an increase in oscillator strength due to larger overlap of electron and hole wave functions (in this case equivalent to larger MME) result in a strongly increased absorption [52,53].

Considering Equation 41.12, evaluation of the absorption in semiconductor NSs starts with the computation of the wave functions $\psi_{c,v}$ and energies $\varepsilon_{c,v}$ of the single-particle states $\{|v\rangle, |c\rangle\}$ participating in the optical transitions. There is a vast amount of literature on that topic, which is also covered by a number of chapters in this volume, and will thus not be treated explicitly here. With the electronic structure information, idealized absorption spectra can be obtained from Equation 41.12. However, for a realistic assessment of the PV response, the experimental spectra should be reproduced as closely as possible, which requires consideration of the actual effects that have an impact on the line shape, such as finite fields, excitonic contributions, lifetime broadening due to coupling to phonons and hybridization with continuum states, compositional disorder, and finite-size distributions. Various approaches with increasing degree of sophistication can be found in the literature.

For **QW**, excitons and field effects (i.e., electroabsorption) were included in [54] using a variational ansatz for the exciton wave functions. The consistent consideration of the bound-to-bound state and the bound-to-quasibound state transitions including excitonic and field effects was achieved using an exciton Green's function approach [55]. To describe lineshape broadening beyond phenomenological approaches, microscopic quantum-kinetic theories such as the density matrix formalism, including the different scattering mechanisms (electron–phonon, electron-electron, etc.) can be used [48–50]. Recently, excitonic absorption enhancement for QWs under consideration of the electron–phonon interaction was formulated in a general nonequilibrium Green's function framework [56]. In practice, however, some degree of fitting is used to reproduce the experimental spectra. For QWSC, a semiempirical approach was introduced by Paxman et al. [57], based on earlier work on QW optical modulators [58,59], in which the field-dependent absorption coefficient is expressed as

$$\alpha(E_\gamma, \mathcal{E}^z) = \sum_{c,v} \alpha_{cv} M_{cv}(\mathcal{E}^z) \Big[\sum_l r_{cv}^l \mathcal{L}[E_\gamma, \varepsilon_{cv}(\mathcal{E}^z) - B_{cv}^l] + \int_{\varepsilon_{cv}(\mathcal{E}^z)}^{\infty} dE' \mathcal{L}(E_\gamma, E') \Big], \tag{41.16}$$

where $\mathcal{E}^z$ is the longitudinal electric field in growth direction, $M_{cv} = \int_L dz \psi_c^*(z, \mathcal{E}^z)\psi_v(z, \mathcal{E}^z)$ (L: normalization length) is the overlap matrix element of the field-dependent single-particle electron and hole wave functions, r_{cv}^l (in units of energy) and B_{cv}^l are the relative oscillator strength and binding energy of the lth exciton, respectively, and $\mathcal{L}(E_\gamma, E') = \Gamma/(2\pi[(E_\gamma - E')^2 + \Gamma^2/4])$ is a Lorenzian lineshape function to include homogeneous broadening characterized by Γ. The parameter α_{cv} corresponds to the absorption coefficient at the onset of the continuum absorption, and, in practice, is used together with B and r to fit the experimental spectra. For the absorption at photon energies above the bulk bandgap of the barrier material, the bulk absorption coefficient of the well material is used, neglecting any field effects. A similar semiempirical approach was used in [60], based on [61].

For **QD**, the single-particle absorption spectra can be obtained using similar density matrix formalisms for the linear susceptibility [62] or the equivalent FGR, with straightforward extension from single QD to QDSL [63,64]; in the case of significant excitonic effects, the latter can be treated by means of the configuration interaction (CI) method [65]. While lineshape broadening due to dephasing via electron–phonon interactions can in principle be included in quantum-kinetic approaches, the finite-size distribution of QD ensembles requires additional convolution with some distribution function representing inhomogeneous broadening, commonly of the Gaussian type [66]. In self-assembled QD formed on a wetting layer (WL), the WL states may also contribute to the subgap absorption [67]. For practical device simulation, a common approach is the use of an effective medium absorption coefficient [68]:

$$\bar{\alpha}_{\text{eff}} = w_{\text{QD}}\bar{\alpha}_{\text{QD}} + w_{\text{WL}}\bar{\alpha}_{\text{WL}} + w_{\text{bulk(QD)}}\bar{\alpha}_{\text{bulk(QD)}} + w_{\text{barrier}}\bar{\alpha}_{\text{barrier}}, \tag{41.17}$$

which contains the contribution of the confined states in the QD, the WL, the bulk absorption of the QD material for continuum states, and the barrier absorption, weighted according to the relative volume fractions and a given distribution function reflecting the variations in size and shape. The general expression for the absorption of an ensemble of QD configurations C with distribution f using these simple approaches reads:

$$\bar{\alpha}_{\text{QD}}(E_\gamma) = \int dC f(C) \alpha_C(E_\gamma), \tag{41.18}$$

$$\alpha_C(E_\gamma) = \frac{\pi\hbar}{n_r c_0 \varepsilon_0 E_\gamma \mathcal{V}_{QD}} \left(\frac{e}{m_0}\right)^2 \sum_{c,v} \bar{p}_{cv}^2(C) \frac{\Gamma}{\left(\varepsilon_{cv}(C) - E_\gamma\right)^2 + \Gamma^2/4} (f_v - f_c), \tag{41.19}$$

where $\bar{p}_{cv}$ is the polarization-averaged MME and Γ is the homogeneous broadening. The WL absorption is given by the expression for QW (usually neglecting excitonic absorption), where consideration of thickness fluctuations results in an additional absorption tail. Since the absorbing volume of a QD is not

always well defined, especially in the case of quasibound states, it is common to write the QD contribution absorption in terms of an *absorption cross-section* σ_{QD} per QD and of the QD concentration N_{QD} via $\alpha_{QD} = N_{QD}\sigma_{QD}$.

41.2.2 Charge Separation and Extraction

In general, the extraction of photogenerated charge carriers represents one of the most critical aspects in NS-based PV devices. Indeed, while in some devices based on QW or QD, the tailoring of the absorption via band structure engineering could be successfully demonstrated [69,70], the efficiency of carrier collection has until now remained below the values found for the bulk counterparts, in some cases on a dramatically poor level, especially in devices where transport proceeds via states with increased degree of localization. There are several reasons for the observed performance issue, which are associated with different stages in the extraction process, such as exciton dissociation, carrier escape and capture between localized and extended states, mobility issues associated with scattering, and the virtual absence of true miniband formation in realistic situations, which shall be discussed below.

41.2.2.1 Exciton Dissociation

Since electrons and holes have to be extracted via separate contacts, the dissociation of the photogenerated excitons and the subsequent charge separation need to occur at some point in the device. In bipolar bulk SC, exciton binding energies are usually in the range of a few meV, resulting in thermal dissociation immediately after generation, and the generation can thus safely be described via (Equation 41.2) in terms of noninteracting electron-hole pairs. In isolated NS on the other hand, spatial confinement enhances the electron-hole interaction, and exciton binding energies can amount to multiples of the thermal energy k_BT (k_B: Boltzmann constant, T: temperature), in which case rapid thermal dissociation is no longer possible, and the dissociation process needs to be considered explicitly [53]. Inefficient exciton dissociation can result in a transport behavior dominated by exciton diffusion, similar to the situation in organic SC devices, where a specially designed bulk–heterojunction interface is required for exciton dissociation. As a SC based on exciton diffusion is not likely to reach very high efficiencies, the description of excitons in the simulation of efficient NSSC devices is primarily focused on the excitonic enhancement of optical transitions close to the effective band edge, as described above in the treatment of the absorption coefficient.

41.2.2.2 Charge Carrier Escape and Capture

Charge carriers that are generated in localized states of electronically decoupled NS need to be transferred to extended states in order to contribute to current, which normally amounts to a sequence of scattering events involving confined (2D—QW/0D—QD) subband states followed by scattering from a bound or quasibound state to a (3D) continuum state. While this *escape* process constitutes a loss mechanism in light-emitting devices, which is sought to be suppressed in optimized designs, the opposite applies in NSSC architectures. Hence, understanding the escape process via identification of dominant and limiting mechanisms and design of corresponding models has been an important focus in NSSC research from the beginning. On the other hand, any escape channel enables its inverse capture process, which removes photocarriers from contacted states and increases dark current flow.

In general, while the microscopic processes underlying the carrier dynamics are always related to the fundamental coupling between the electronic, vibrational, and optical degrees of freedom, there are several mechanisms by which photogenerated charge carriers can escape from confined NS states, which are commonly characterized by respective time constants τ^i_{esc}, with total escape rate proportional to $\tau^{-1}_{esc} = \sum_i (\tau^i_{esc})^{-1}$. For **QW** at room temperature—the relevant temperature regime for terrestrial SC operation—and with shallow confinement, the dominant escape mechanism was identified as *thermionic emission* [71], which describes the current due to charge carriers that occupy the part of the energy distribution located above the barrier edge. In the presence of a rectifying built-in field, the escape process is affected by the latter through the modified height (with respect to the center of the QW) and triangular shape of the

barrier, giving rise to escape enhancement due to *tunneling*. For the common assumption of a constant field, the effective barrier for escape in direction of and against the field is given by $V_{B\mp}(\mathcal{E}^z) = \Delta E_b \mp q\mathcal{E}^z L_w/2$, where ΔE_b is the offset of bulk band edges and L_w is the QW width. A general *phenomenological* expression for the corresponding escape time is then given by [71]

$$\tau_{\text{esc}}^{-1} = \frac{1}{L_w}\int_0^\infty \tilde{n}(E)\mathcal{T}(E)v(E)dE, \quad \tilde{n}(E) = \frac{n(E)}{\int_0^\infty n(E)dE}, \tag{41.20}$$

where $\mathcal{T}(E)$ is the probability function for transmission through the barriers, $v(E)$ is the carrier velocity, and $n(E) = g(E)f(E)$ is the spectral carrier density at an energy E measured from the center of the bottom and assuming a given distribution function $f(E)$ for charge carriers in the QW. The actual form of Equation 41.20 depends on the specific description of the spectral quantities ($\mathcal{T}, v, g, f$). The simplest assumption for $\mathcal{T}$ is a step function with unit transmission above the effective barrier edge and zero below, which neglects tunneling contributions. Tunneling can be included on the level of the Wentzel–Kramers–Brillouin (WKB) theory via [72]

$$\mathcal{T}_{\text{WKB}}(E) = \exp\left(-\frac{2}{\hbar}\int_0^{L_B} dz\sqrt{2m_B^*[V_B(z) - E]}\right), \tag{41.21}$$

where m_B^* is the effective mass of carriers in the barrier material, L_B is the width of the barrier, and $V_B(z)$ is the spatial profile of the barrier potential. More accurate approaches for determination of the transmission function are based on analytical or numerical solution of the Schrödinger equation using the exact barrier potential, e.g., by means of an electronic transfer matrix formalism [71]. For free carriers, the velocity is related to the (kinetic) carrier energy via the momentum, e.g., $v_z = \hbar k_z/m_w^* = \sqrt{2E_z/m_w^*}$ for propagation in the z-direction. In basic approximation, the thermal velocity v_{th} is used, corresponding to $E_z = k_B T$. The exact form of the DOS g reflects the transition of dimensionality from 2D to 3D at energies close to the barrier edge and can be captured in advanced descriptions based on scattering states in a real space basis representation, which do not discriminate between bound, quasibound, and continuum states, such as the Green's function approaches [73,74]. In most implementations, however, a combination of 2D square well DOS and 3D bulk DOS is used [71]. Finally, the occupation function f needs to encode the nonequilibrium steady-state carrier distribution under illumination and bias voltage. In general, this is achieved by using Fermi or Boltzmann statistics together with a quasi-Fermi level (QFL) for the carriers in the QW. In the case of Boltzmann statistics, the QFL drops out of the expression for the escape time due to the normalization to the total carrier density [75]. The advantage of the above approach is that it covers any combination of thermionic, direct tunneling, and thermally assisted tunneling escape. The drawback is that it requires assumptions on the carrier distribution in the QW and it does not contain any information on the microscopic mechanisms that lead to the carrier population at a certain energy. Such mechanisms include, e.g., the very fast electron-electron scattering that establishes a hot-carrier quasi-equilibrium at a fs timescale. But even the slower scattering of electrons with longitudinal optical (LO) phonons was shown to produce a hot-carrier population via subband scattering that enables fast thermally assisted carrier escape [76–78]. This latter result, while based on simple application of FGR to electron-LO-phonon scattering between electronic subbands, was confirmed by the state-of-the-art quantum-kinetic simulation of the phonon-assisted tunneling escape [79].

In **QD** devices, in addition to continuum (3D) and confined (0D) QD states, the 2D states of the WL play an important role in the dynamics of carrier transitions. Experimentally, similar thermionic and tunneling escape processes were identified [80–82]. However, thermal escape is much less efficient than in QW, and larger fields are required to enable tunneling escape. In terms of microscopic mechanisms, evidence for both Auger scattering [83] and (multi)phonon absorption [84] was found. While the carrier-carrier scattering depends strongly on carrier density and decreases quickly with QD size due to shrinking Coulomb (Cb)

matrix elements—and is, therefore, less prominent in QW—the phonon processes depend critically on the energetic spacing of the states involved in the transitions, which produces strong resonances in the size dependence of the scattering mechanism. However, the lifetime broadening of QD states due to dynamical processes and delocalization increasing with energy alleviate the impact of the discrete DOS. Finally, in addition to the nonradiative escape processes, QDs offer the possibility of photon-assisted or purely radiative escape via subband transitions as described in Section 41.2.1. In QW, the subband absorption is strongly suppressed for in-plane polarization; hence, special photon management is required to scatter light into large angles, preferentially with coupling to guided modes.

The inverse process to escape is carrier *capture*. In SCs, this process is relevant in MQW and MQD devices, where photogenerated carriers that escaped to the current-carrying continuum states are recaptured into confined states on their way to the contacts. Since on the other hand, this process is instrumental for light-emitting devices, its theoretical description and numerical simulation have received considerable attention and are extensively covered in the literature. For **QW**, traditional models consider the FGR rate for transitions from 3D bulk states to confined states via the emission of LO-phonons [85–95] or carrier-carrier scattering [96,97], in combination with population dynamics from rate equations [98,99] or Monte Carlo simulations [100–104], while more recent approaches are based on quantum-kinetic theory [105]. At the relatively low densities of injected carriers at the operating point of a SC (which is below the bias regime of LED operation), relaxation due to emission of phonons was found to dominate over the effects of carrier-carrier scattering. In the case of **QD** devices, capture from the WL and subband relaxation were evaluated in the FGR framework for Auger processes [106,107], including the effects of line broadening induced by the carrier dynamics [108], and for the emission of single phonons [109,110] and multiple phonons [111], with extension to a polaron picture with finite phonon lifetime [112] indicated due to the strong coupling of QD states to phonons. The mechanisms were also combined in unified theories on semiclassical [113] and quantum-kinetic [114] levels.

41.2.2.3 Transport to and Extraction at Carrier-Selective Contacts

The choice of the picture for charge transport is primarily dictated by the kind of NS device architecture under consideration. In the second and third types of NSSC devices displayed in Figure 41.1, the NS states are involved not only in the absorption, but also in the carrier transport. These states thus need to be to some degree extended, which amounts to the requirement of coupling between individual NS. Indeed, most of the concepts of this kind are based on the formation of minibands in superlattices of QWs or QDs. For that situation, transport may be described as band-like, with a low-field Bloch-type electron *mobility* $\mu = \sigma/(\rho e)$ defined in terms of the conductivity tensor σ, and in the diffusive limit is obtained from the linearized Boltzmann transport equation (BTE) [115]:

$$\mu_{\alpha\beta} = e\left[\sum_{n,\mathbf{K}} \tau_{n\mathbf{K}} v^{\alpha}_{n\mathbf{K}} v^{\beta}_{n\mathbf{K}} \left(-\frac{\partial f^0_{n\mathbf{K}}}{\varepsilon_{n\mathbf{K}}}\right)\right] \Big/ \sum_{n,\mathbf{K}} f^0_{n\mathbf{K}}. \tag{41.22}$$

In Equation 41.22, n is the band index, $\mathbf{K}$ the superlattice wave vector, $\tau_{n\mathbf{K}}$ is the relaxation time, $\varepsilon_{n\mathbf{K}}$ is the miniband dispersion, $f^0_{n\mathbf{K}}$ is the equilibrium Fermi-Dirac electron distribution function, and $v^{\alpha}_{n\mathbf{K}} = \hbar^{-1}\partial_{K_\alpha}\varepsilon_{n\mathbf{K}}$ is the band velocity. The relaxation time depends on the scattering processes that are present in the device and can be computed—in some approximation—from the corresponding FGR formalism [116]:

$$\tau^{-1}_{n\mathbf{K}} = \sum_{n'\mathbf{K}'} \left(\frac{1 - f^0_{n'\mathbf{K}'}}{1 - f^0_{n\mathbf{K}}}\right)\left\{1 - \cos(\theta_{\mathbf{K}\mathbf{K}'})\right\} P^{nn'}_{\mathbf{K}\mathbf{K}'} \tag{41.23}$$

where $\cos(\theta_{\mathbf{KK}'}) = (\mathbf{v}_{n\mathbf{K}} \cdot \mathbf{v}_{n'\mathbf{K}'})/(|\mathbf{v}_{n\mathbf{K}}||\mathbf{v}_{n'\mathbf{K}'}|)$ and $P^{nn'}_{\mathbf{KK}'}$ is the FGR rate for scattering from the occupied state $|n\mathbf{K}\rangle$ to the empty state $|n'\mathbf{K}'\rangle$. However, even in the ideal case of a perfect miniband, the associated Bloch mobility may be critically low, as in the case of silicon QDs in a dielectric matrix material [117]. In realistic situations, due to the presence of built-in fields and any kind of spatial, configurational and compositional disorder, the NS states usually show a high degree of localization such that tunneling between NS is restricted to nearest neighbors. In addition to inducing localization, these effects lead to a misalignment of energy levels in adjacent NS, with the formation of Wannier-Stark ladders in the extreme case of very strong fields [118], in which situation transport is only possible via an inelastic scattering process and is best described by hopping between localized states. In coupled QD systems, due to the sparsity of the DOS, the presence of charge on the NS can lead to pronounced shifts in the energy level structure [53], inhibiting transport (Coulomb blockade regime). In devices where charge transport acquires molecular character, such as in colloidal QD solids, the hopping process involves even a change in configuration and may be phenomenologically described by the Marcus theory [44]. The bulk picture of charge carrier transport in band states does thus in general not provide a proper description of the real situation, which in addition to the remaining coherence and nonlocality effects should include all the localization effects and scattering mechanisms required to overcome energetic misalignment. The modification of the electronic structure due to the presence of NS also affects the scattering processes responsible for the limitation of mobility and for carrier relaxation, such as electron–phonon interaction, which becomes especially relevant in devices where this form of energy dissipation is sought to be suppressed, as in the hot-carrier concept, and in the case where extraction distances are on the order of the mean free path and transport approaches the ballistic regime.

The final step of the PV energy conversion consists in the extraction of the charge carriers at carrier-selective contacts. Carrier selectivity—which is an essential requirement on any PV device architecture [119]—can be described on a phenomenological level via the surface recombination velocities for majority and minority carriers used in the boundary conditions for carrier density/current in macroscopic transport equations. Microscopically, transport across the boundary to the contact depends again on the electronic structure of the interface and on the prevailing scattering processes. Since those can be largely engineered using NS, the latter also play an important role in the implementation of energy-selective [22] and carrier-selective [120] contacts.

41.2.3 Recombination

On the way to or from the contacts, the interaction of optically or electronically injected charge carriers with the EM and vibrational fields can induce different types of *recombination* (the mutual annihilation of an electron and a hole), which are usually classified in *radiative*, *Auger*, and *defect-mediated* processes. Like for the generation process, a proper consideration of NS effects on the recombination processes requires a careful consideration of the wave functions and local DOS together with the occupation of these states in general nonequilibrium conditions.

41.2.3.1 Radiative

Even in ideal, defect-free absorber materials, the principle of *detailed balance* dictates the presence of radiative recombination by spontaneous and stimulated photon emission as the inverse process to carrier generation by photon absorption. Conventionally, the local volume rate of radiative recombination is related to the charge carrier densities and the local optical material constants (α, n_r) as follows [121]:

$$r^{\mathrm{em}}(\mathbf{r}) = \mathcal{B}(\mathbf{r})\left\{\rho_e(\mathbf{r})\rho_h(\mathbf{r}) - \rho_e^0(\mathbf{r})\rho_h^0(\mathbf{r})\right\}, \tag{41.24}$$

$$\mathcal{B}(\mathbf{r}) = \rho_i^{-2}\int dE_\gamma\, \alpha(\mathbf{r}, E_\gamma)\bar{\Phi}_{bb}(E_\gamma), \tag{41.25}$$

where ρ^0 and ρ_i denote equilibrium and intrinsic carrier densities, respectively, and $\bar{\Phi}_{bb} = 4\pi\Phi_{bb}$ is the angle-integrated blackbody radiation flux, with

$$\Phi_{bb}(E_\gamma) = \frac{E_\gamma^2 n_r^2}{4\pi^3\hbar^3 c_0^2}\left\{\exp\left(\frac{E_\gamma}{k_B T}\right) - 1\right\}^{-1}. \tag{41.26}$$

In Equation 41.25, the assumption of emission in an optically homogeneous medium is made, with isotropic photon DOS. In thin-film SCs with nanophotonic light trapping, this is often not appropriate, and the radiative recombination should be modeled starting from the FGR for spontaneous emission under consideration of the actual density of final photon states. For NS, the emission is enhanced as compared to the bulk, due to carrier localization [122] resulting in maximized density product in Equation 41.25 and to larger overlap of electron and hole wave functions in Equation 41.13. A further point to note in the case of NSSC is the dependence of the absorption coefficient on the operating conditions, mainly the bias voltage via the effect of the built-in field on the local electronic structure, which should be reflected in the quantities used in Equation 41.25.

41.2.3.2 Auger

In the regime of large optical injection, relevant for the operation under high optical concentration, the PV performance of a defect-free absorber is limited by the interband Auger recombination mechanism [123], which as an intrinsic effect is unavoidable and needs thus to be considered for a realistic estimate of the limiting energy conversion efficiency. In this process originating in carrier-carrier scattering, electrons recombine with holes by giving up the excess energy to excite either another electron or another hole. In the semiclassical bulk continuum formulation, this rate takes the form [124]

$$r^{\mathrm{Aug}}(\mathbf{r}) = \left[C_e(\mathbf{r})\rho_e(\mathbf{r}) + C_h(\mathbf{r})\rho_h(\mathbf{r})\right]\left\{\rho_e(\mathbf{r})\rho_h(\mathbf{r}) - \rho_e^0(\mathbf{r})\rho_h^0(\mathbf{r})\right\}, \tag{41.27}$$

where $C_{e/h}$ are the Auger coefficients for electron/hole excitation. For NS, some or all of the band indices are replaced by the labels of subbands (QW) or discrete levels (QD). The Auger coefficients can be computed microscopically by application of perturbation theory for the carrier-carrier interaction under consideration of the appropriate dimensionality of the (screened) Cb potential. For QW, FGR-type approaches were used to assess direct [125–129] and phonon-assisted [130] Auger processes, and the latter was also investigated using a Green's function technique [131]. For QD, similar FGR models were developed for colloidal QD [132] and applied to self-assembled III-V single QD [133] and QD arrays [134]. Recently, the complete dynamics of excitons in colloidal QD was also simulated on an *ab initio* level using a combination of density functional theory and molecular dynamics [135]. While in QW, the magnitude of the Auger recombination rate was found to be similar to bulk, it is strongly enhanced in QD due to the effects of wave function localization and associated relaxation of momentum selection rules.

41.2.3.3 Defect Mediated

While radiative and Auger losses are unavoidable and thus present even in structures without imperfections, recombination in real-world devices is usually dominated by defect-mediated processes where the carrier energy is dissipated as heat and is lost for the conversion process. Indeed, there is an inherent increase in surface area associated with the presence of NS, which are thus likely to act as centers of nonradiative recombination. In some cases, it is possible to reduce the NS-related defects via specially engineered strain-balancing techniques [136] or via passivation [137], but in general, the insertion of NS leads to losses in the open-circuit voltage of the SC. With regard to the modeling, the role of lower dimensional NS in the context of defect-mediated recombination has received little attention so far, and models are usually

based on the conventional bulk Shockley–Read–Hall (SRH) theory [138,139], with modification to relate the recombination to interface or surface defects rather than volume defects, e.g., for QW [60]

$$r^{\mathrm{SRH}}(\mathbf{r}) = \frac{\rho_e(\mathbf{r})\rho_h(\mathbf{r}) - \rho_e^0(\mathbf{r})\rho_h^0(\mathbf{r})}{(\rho_e(\mathbf{r}) + \rho_e^t(\mathbf{r}))\tau_e + (\rho_h(\mathbf{r}) + \rho_h^t(\mathbf{r}))\tau_h}, \tag{41.28}$$

where $\tau_s^{-1} = \sigma_s v_{th} N_S / L_w$ $(s = e, h)$ with σ_s the carrier capture cross-section, v_{th} the thermal velocity, N_S the sheet density of interface defects, L_w the QW width, and ρ^t the carrier density for a Fermi level at the trap energy. The capture cross-sections depend on the local electronic structure of the NS and defect states, as well as on the phonon modes coupling to those states in the multiphonon relaxation mechanism [124], and can be obtained based on microscopic theories on phenomenological [140] and fully *ab initio* [141] levels.

41.3 Modeling of NSSC Device Characteristics

The basic function of a SC consists in the conversion of the solar energy flux impinging on its surface into electrical power. A central requirement on any theory suitable for the modeling of SC device operation is thus the ability to predict the PV energy conversion efficiency as defined via the ratio $\eta_{\mathrm{PV}} \equiv P_{el}^{\nearrow} / P_{\gamma}^{\swarrow}$ of generated electrical power to the incident radiative power, with

$$P_{el}^{\nearrow} = \max_V \{I(V) \cdot V\}|_{T_C}, \tag{41.29}$$

$$P_{\gamma}^{\swarrow} = \int dE_\gamma \int_{\mathcal{S}} \int_{\Omega} d\boldsymbol{\Omega} \cdot d\boldsymbol{\mathcal{S}}\, \Phi_{\gamma}^{\swarrow}(\mathbf{r}_{\mathcal{S}}, \Omega, E_\gamma) E_\gamma, \tag{41.30}$$

where V is the terminal voltage (i.e., related to the separation of chemical potentials at the contacts), $I(V)$ is the charge current extracted from the device at this voltage, T_C is the temperature of the SC, E_γ is the photon energy, and $\Phi_{\gamma}^{\swarrow}$ denotes the incident spectral photon flux impinging on a cell of (outer) surface $\mathcal{S}$ under a solid angle of incidence $\Omega = (\theta, \varphi)$. Hence, any PV model needs to provide I as a function of V and $\Phi_{\gamma}^{\swarrow}$ at given T_C.

As outlined in Sections 41.1 and 41.2, depending on the functionality of the NS within the SC device, some or all of the fundamental PV processes are affected by the peculiar physical properties of the NS component, with major implications for the requirements on a suitable simulation approach. The general approach reflects the multiscale and multiphysics scheme depicted in Figure 41.2: Starting from the NS states, the dynamics giving the occupation of these states is evaluated and the resulting generation and recombination rates are used to obtain the device characteristics. Three levels of sophistication that can be identified in the hierarchy of theoretical descriptions of SC device operation providing the current–voltage characteristics $I(V, \Phi_{\gamma}^{\swarrow})$, depending on the explicity of consideration of the various physical mechanisms, are: *detailed balance* theories providing efficiency limits of idealized systems, *semiclassical* transport models with detailed balance generation-recombination rates, and *quantum-kinetic* theories with consistent and fully microscopic treatment of charge carrier dynamics and transport [142]. These three approaches shall now be discussed in some more detail.

41.3.1 Global Detailed Balance Theories for Ideal Systems

On the most basic level, NS properties are considered only in generation and recombination at the radiative limit, while transport is assumed to be ideal. This means that from a device point of view, the system is treated as a black box, and there is no information on geometry or any spatial inhomogeneity in the electronic system that would depart from effective bulk properties. The most prominent examples of this

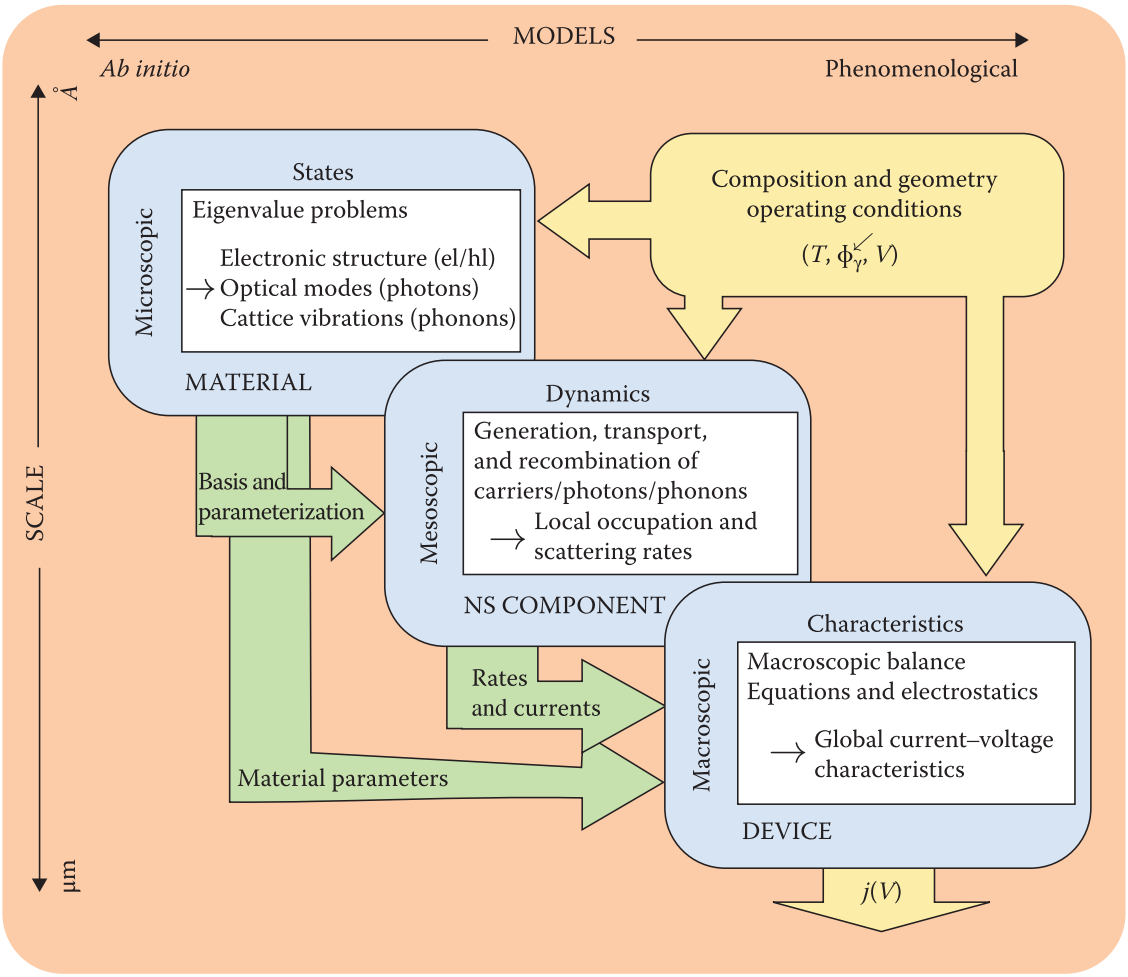


FIGURE 41.2 Schematic representation of the multiscale and multiphysics approach to the simulation of nanostructure-based solar cell devices.

category are the detailed balance models [143] used to obtain an upper bound to the energy conversion efficiency, such as the SQ limit. In this approach, the current–voltage characteristics are obtained by equating the terminal current with the difference of the number of photons per unit time entering and leaving the absorber volume at given terminal bias voltage [143,144]:

$$I(V,\Phi_\gamma^{\swarrow})|_{T_C} = -q\int \mathrm{d}E_\gamma \int_{\mathcal{S}}\int_{\Omega} \mathrm{d}\boldsymbol{\Omega}\cdot\mathrm{d}\boldsymbol{\mathcal{S}}\left\{\Phi_{\mathrm{abs}}(\mathbf{r}_{\mathcal{S}},\Omega,E_\gamma)-\Phi_{\mathrm{em}}(\mathbf{r}_{\mathcal{S}},\Omega,E_\gamma)\right\}|_{V,\Phi_\gamma^{\swarrow},T_C}. \tag{41.31}$$

At the radiative limit, where the recombination is determined by the emission, and assuming unit generation efficiency, the electronic equivalent to the above balance relation is obtained from the integration of the charge continuity equation over the absorber volume $\mathcal{V}$ (e/h: electrons/holes):

$$I(V,\Phi_\gamma^{\swarrow})|_{T_C} = \mp q\int_{\mathcal{V}} d^3\mathbf{r}\left\{\mathcal{G}_{e/h}(\mathbf{r})-\mathcal{R}_{e/h}(\mathbf{r})\right\}|_{V,\Phi_\gamma^{\swarrow},T_C}. \tag{41.32}$$

The absorbed photon flux is derived in terms of the absorptance—or, synonymously, the absorptivity—a of the device, and of the reflection coefficient r for incident light, via

$$\Phi_{\mathrm{abs}}(\mathbf{r}_{\mathcal{S}},\Omega,E_\gamma)=[1-r(\mathbf{r}_{\mathcal{S}},\Omega,E_\gamma)]a(\mathbf{r}_{\mathcal{S}},\Omega',E_\gamma)\left\{\Phi_\gamma^{\swarrow}(\mathbf{r}_{\mathcal{S}},\Omega,E_\gamma)+\Phi_{\gamma 0}(E_\gamma,\Omega,T_{\mathrm{C}})\right\}, \tag{41.33}$$

where Ω' is the solid angle after refraction, related to the former by Fresnel's law [145]. The second term in the curly brackets is the isotropic thermal equilibrium flux and is given by

$$\Phi_{\gamma 0}(E_\gamma, \Omega, T_C) = \Phi_{bb}(E_\gamma, T_C)\cos\theta, \tag{41.34}$$

where Φ_{bb} is the blackbody radiation flux (Equation 41.26) introduced in Section 41.2.3. The generation rate in terms of local photon flux and absorption coefficient is given in Equation 41.2. However, for equivalence with the above result, photon recycling needs to be considered in the evaluation of the local photon flux. At unit generation efficiency and in the absence of nonradiative recombination, the product in front of the incident flux in Equation 41.33 is identified as the *external quantum efficiency* (*EQE*). Under the assumption of perfect transport, i.e., infinite mobility, which amounts to flat QFLs split by the terminal voltage, i.e., $\Delta\mu = qV$, and of an optically isotropic medium, detailed balance then provides the emitted photon flux as a function of the absorptance [144,146]:

$$\Phi_{\mathrm{em}}(\mathbf{r}_S, \Omega, E_\gamma)|_V = [1 - r(\mathbf{r}_S, \Omega, E_\gamma)]a(\mathbf{r}_S, \Omega', E_\gamma)\tilde{\Phi}_{bb}(\Omega, E_\gamma, T_C, V) \tag{41.35}$$

$$\equiv EQE(\mathbf{r}_S, \Omega, E_\gamma)\tilde{\Phi}_{bb}(\Omega, E_\gamma, V), \tag{41.36}$$

where:

$$\tilde{\Phi}_{bb}(\Omega, E_\gamma, T_C, V) = \Phi_{bb}(E_\gamma - qV, T_C)\cos\theta \tag{41.37}$$

$$\approx \Phi_{\gamma 0}(E_\gamma, \Omega, T_C)\mathrm{e}^{qV/(k_B T_C)} \qquad (E_\gamma \gg qV). \tag{41.38}$$

The approximation in Equation 41.38 permits a compact formulation of the current–voltage characteristics:

$$I(V, \Phi_\gamma^{\swarrow})|_{T_C} \approx -q\int dE_\gamma \int_S \int_\Omega d\boldsymbol{\Omega}\cdot d\boldsymbol{S}\, EQE(\mathbf{r}_S, \Omega, E_\gamma)$$
$$\times \left\{\Phi_\gamma^{\swarrow}(\mathbf{r}_S, \Omega, E_\gamma) - \Phi_{\gamma 0}(E_\gamma, \Omega, T_C)\left[\mathrm{e}^{qV/(k_B T_C)} - 1\right]\right\}. \tag{41.39}$$

Finally, the local recombination rate for the determination of the radiative dark current via Equation 41.32 is given by Equation 41.25.

In the SQ limit, the absorptance is a unit step function with threshold energy at the value of the bandgap E_g, which is then the only electronic parameter entering the calculation of the current–voltage characteristics. This simplifies the expressions for the absorbed and emitted photon fluxes considerably. In a first step beyond the SQ analysis, the effect of incomplete absorption with non–abrupt onset as well as the impact of light-trapping structures can be considered by using in Equations 41.33 and 41.35 the proper absorptance computed by optical models for the light propagation as outlined in Section 41.2. It is common to neglect the bias dependence of the photocurrent also in situations where a realistic absorptance is considered, which results in current–voltage characteristics given by the *superposition* of photocurrent I_γ and dark current I_d:

$$I(V, \Phi_\gamma^{\swarrow}) = I_\gamma(\Phi_\gamma^{\swarrow}) + I_\mathrm{d}(V). \tag{41.40}$$

However, this tends to be inappropriate in nanoscale absorbers due to the strong impact of built-in fields and finite state filling on the local absorption coefficient [147,148].

Due to its formal simplicity, the detailed balance analysis of the radiative limit is applied to most novel NS PV devices in a first assessment of the performance potential. On the other hand, the restrictive assumptions made are seldom met in realistic devices, and the upper efficiency bound obtained in this way may therefore be far from a realistic estimate.

Within the above setting of idealized conditions, an alternative starting point for the device characteristics is a modification of the ideal diode equation [149]:

$$I(V,\Phi_\gamma^{\swarrow}) = I_0\left(e^{qV/(k_B T_C)} - 1\right) - I_G(\Phi_\gamma^{\swarrow}) + I_R(V), \tag{41.41}$$

which is formally equivalent to Equation 41.39, but where the generation and recombination currents I_G and I_R are composed of the contributions of the bulk "baseline" cell and of the NS components. The NS contributions are related to those of the baseline using correction factors for quantities such as geometry (fraction of NS material), oscillator strength, and DOS (for radiative transitions). It needs to be noted, however, that such modifications still have to comply with the detailed balance relations between absorption and emission as given above. Furthermore, even in its original formulation devised for the radiative limit, the theory goes beyond that limit in considering in the reverse saturation current I_0 the standard nonradiative minority carrier diffusion contribution of an ideal *p*-(*i*-)*n* diode.

41.3.1.1 Application to QWSC

Soon after the proposal of the QWSC concept, the detailed balance approach was used to show that under the idealized conditions of that theory, the QWSC could not provide an efficiency in excess of that of an ideal gap bulk cell, since both photocurrent generation and radiative recombination are based on the same modified absorptivity [150,151]. An SQ-type detailed balance model was then used to demonstrate that QWSC efficiency could still exceed the bulk limit if variations of the quasi-Fermi level splitting (QFLS) between bulk and QW material were allowed for [152]. This was shown to be equivalent to the detailed balance picture of the intermediate band concept, with the difficulty of the QW featuring suppressed subband absorption and fast population equilibration due the continuous electronic DOS [153]. Indeed, the issue of QFLS variation in QWSC is still a matter of debate and ongoing research. In more recent QWSC developments, the detailed balance approach was used to model tandem configurations and the impact of directional emission in strained QW [23]. Following the original work on the ideal diode QWSC model [149], a number of variations and refinements of this semianalytical approach were devised for MQWSC [154–161] and QWSLSC [159,162–164].

41.3.1.2 Application to QDSC

For QD intermediate band SC (QDIBSC), the original detailed balance IBSC model [165] was adapted to account for the specific absorption characteristics of QD structures [166,167] and was implemented for specific material systems using a Kronig-Penney model for supracrystals of InAsN/GaAsSb QD [168], InGaN/GaN QD [169], and InGaN/InN [170], as well as a $\mathbf{k}\cdot\mathbf{p}$ model for isolated InAs/GaAs QD [171]. The ideal diode theory was applied to theoretically investigate an InAs/GaAs MQD concept with AlGaAs fences to prevent capture of photogenerated carriers [172].

41.3.2 Hybrid Models: Semiclassical Transport with Detailed Balance Rates

The main shortcoming of global detailed balance approaches and ideal diode theories is the lack of a realistic consideration of photocarrier transport, i.e., the assumption of unit collection efficiency, which is often inappropriate in NSSC devices. In order to account for the effects of finite mobility, actual transport equations have to be solved for the charge carriers. The main difficulty in that undertaking resides in the consideration of the contribution of NS states with a higher degree of localization. To obtain the self-consistent occupation of confined NS states (*c*) and bulk host states (*b*), separate but coupled rate equations

need to be formulated for the densities of carriers occupying different types of states. Under the standard assumption of continuum (i.e., no atomistic information) and complete thermalization (i.e., occupation is characterized by a local QFL), the current–voltage characteristics are obtained from the *drift-diffusion* (DD) charge current for carrier species $s \in \{e, h\}$:

$$\mathbf{J}_{\mathrm{s,i}}(\mathbf{r}) = -q\left\{\mu_{\mathrm{s,i}}(\mathbf{r})\rho_{\mathrm{s,i}}(\mathbf{r})\nabla\phi(\mathbf{r}) \mp D_{\mathrm{s,i}}(\mathbf{r})\nabla\rho_{\mathrm{s,i}}(\mathbf{r})\right\}, \qquad i = b, c \tag{41.42}$$

where μ denotes the mobility, $\rho_{e/h}$ denotes the charge carrier densities, D is the diffusion coefficient, ϕ is the electrostatic potential, q is the elementary charge, and the upper (lower) sign is for electrons (holes). In this general formulation, no restriction is made on the charge transport in NS states, which allows for application to all three NS architectures introduced in Section 41.1. The charge carrier densities and the electrostatic potential that determine the charge currents in Equation 41.42 are obtained from the coupled solution of the steady-state continuity equations:

$$\mp\frac{1}{q}\nabla \cdot \mathbf{J}_{s,i}(\mathbf{r}) = \sum_{j\neq i}\left\{\mathcal{G}_{s,j\to i}(\mathbf{r}) - \mathcal{R}_{s,i\to j}(\mathbf{r})\right\}, \tag{41.43}$$

where $\mathcal{G}/\mathcal{R}$ is the carrier (volume) generation/recombination rate due to transitions between (sub)bands and the upper (lower) sign applies again to electrons (holes), and Poisson's equation for the total charge density:

$$\epsilon_0\nabla \cdot \left\{\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})\right\} = -q\left\{\sum_i \rho_{\mathrm{h,i}}(\mathbf{r}) - \sum_j \rho_{\mathrm{e,j}}(\mathbf{r}) + N_{\mathrm{dop}}(\mathbf{r})\right\}, \tag{41.44}$$

where ε is the static dielectric function and $q \cdot N_{\mathrm{dop}}$ is the net charge density due to ionized dopants. The boundary conditions used in the solution of the coupled equations are the same as in the bulk case, reflecting the nature of the physical contact in terms of (Schottky-)barriers and surface recombination. In Equation 41.43, the generation rate $\mathcal{G}_{s,j\to i} = \mathcal{G}_{s,j\to i}[\Phi_\gamma(\mathbf{r}), \alpha_{j\to i}(\mathbf{r}, E_\gamma)]$ $(i = b, c)$ is a functional of the local photon flux and the local absorption coefficient as given by Equation 41.2 and the recombination rate $\mathcal{R}_{s,i\to j} = \mathcal{R}_{s,i\to j}[\rho_{e,i}, \rho_{h,j}, \ldots]$ depends on the carrier densities and on recombination mechanism (radiative, Auger, defect-mediated)–specific parameters. In the case where, due to strong localization as for isolated QW or QD absorbers, direct transport between NS states can be neglected, the current term in Equation 41.43 is absent for these states, and exchange of carriers between localized states always proceeds via scattering to and from extended states. The coupling of the equations for localized and extended states is then provided on the level of a pure rate equation, where the rates are subject to the detailed balance condition and are commonly expressed via the density in the initial state and an associated lifetime depending on the scattering process, e.g.,

$$\mathcal{R}^x_{e,i\to j} = \frac{\rho_{e,i}}{\tau^x_{e,i\to j}}, \qquad \text{x: scattering mechanism.} \tag{41.45}$$

The lifetimes τ (which are local quantities) are either used as fitting parameters to reproduce experimental characteristics, or derived from an appropriate description for the microscopic mechanisms of the scattering process responsible for the coupling as reviewed in Section 41.2, e.g., via FGR, based on the solution of the Schrödinger-Poisson problem for the NS states and energies. For consistency, the macroscopic material parameters based on microscopic information, such as absorption coefficients and mobilities used in the generation term and the current expressions, should be computed on the basis of the same solutions of the microscopic equations for the electronic structure. It needs to be emphasized at this point that the lifetimes themselves still depend on the occupation of the states, e.g., for the radiative recombination, $\tau_e^{\mathrm{em}}(\mathbf{r}) \approx \left[\mathcal{B}(\mathbf{r})\rho_h(\mathbf{r})\right]^{-1}$, where the very small equilibrium term in (Equation 41.24) is neglected.

In the above formulation of the charge current, the local electronic structure does not appear as a parameter that could be determined microscopically. If the Einstein relation between mobility and diffusion coefficient, i.e., $D = \mu k_B T/q$, and the Boltzmann approximation for the carrier distribution function are valid, an equivalent formulation of the charge carrier densities and currents in terms of effective DOS $\mathcal{N}$ and QFL E_F can be used:

$$\rho_{s,i}(\mathbf{r}) = \int dE\, \mathcal{D}_{s,i}(\mathbf{r}, E) f_{s,i}(E) \approx \mathcal{N}_{s,i} \exp\left(\frac{\pm E_{F_{s,i}}(\mathbf{r}) \mp \varepsilon_{s,i}(\mathbf{r})}{k_B T}\right), \tag{41.46}$$

$$\mathcal{N}_{e,i} \equiv \int_{\varepsilon_{e,i}}^{\infty} dE\, \mathcal{D}_{e,i}(\mathbf{r}, E) \exp\left(\frac{\varepsilon_{e,i}(\mathbf{r}) - E}{k_B T}\right), \tag{41.47}$$

$$\mathcal{N}_{h,i} \equiv \int_{-\infty}^{\varepsilon_{h,i}} dE\, \mathcal{D}_{h,i}(\mathbf{r}, E) \exp\left(\frac{E - \varepsilon_{h,i}(\mathbf{r})}{k_B T}\right), \tag{41.48}$$

$$\mathbf{J}_{s,i}(\mathbf{r}) = \pm\, \mu_{s,i}(\mathbf{r}) \rho_{s,i}(\mathbf{r}) \nabla E_{F_{s,i}}(\mathbf{r}), \tag{41.49}$$

with upper (lower) sign applying to electrons (holes) and $\mathcal{D}$ denoting the DOS for carrier species s and band i, with $\varepsilon_{s,i}$ the minimum (maximum) energy associated ("band edge"). This approach, in addition to the implicit consideration of the microscopic DOS for NS states, provides the possibility to study the emergence of distinct QFL for the NS populations, which is instrumental in PV concepts such as the IBSC.

The *hybrid* (macroscopic-microscopic) approach outlined above can be used to reproduce experimental device characteristics with remarkable accuracy. However, being a macroscopic and local model, any situation requiring energy resolution, nonlocality, or coherence cannot be described properly, e.g., nonthermalized carrier distributions or resonant tunneling. To include such processes in a consistent description, transition to a truly microscopic picture of carrier transport is indicated.

41.3.2.1 Applications to QWSC

Semiclassical DD Poisson models with detailed balance rates for generation and recombination were applied to III-V MQWSC [60,173,174] and to Si-SiO_x QWSLSC [175]. Simplifications include rate equations with semianalytical solutions for current contributions [176,177] and the analytical solution of diffusion equation for excess minority carriers in neutral layers [57,154,178].

41.3.2.2 Applications to QDSC

The above mentioned DD Poisson approach with local rates was implemented in a comprehensive fashion for MQDSC [68,179,180] and QD-IBSC [181–183]. Again, there are also approximations in the form of analytical solutions of the diffusion equation [184–186], of simple rate equations for the population of bulk and NS states [187], or a combination of the two [188].

41.3.3 Fully Microscopic Quantum-Kinetic Models

In situations where the NSs are also involved in transport and the latter is not entirely band-like or purely diffusive, a consistent microscopic picture of the PV processes can be formulated by means of nonequilibrium quantum statistical mechanics using the nonequilibrium Green's function formalism (NEGF) [189]. In the NEGF picture, the device current density $\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p$ is given in terms of the charge carrier Green's function components as follows:

$$\mathbf{J}_{n/p}(\mathbf{r}) = \pm \lim_{\mathbf{r}' \to \mathbf{r}} \frac{e\hbar}{m_0} \left(\nabla_{\mathbf{r}} - \nabla_{\mathbf{r}'}\right) \int \frac{dE}{2\pi} G^{<}(\mathbf{r}, \mathbf{r}'; E). \tag{41.50}$$

The Green's functions in Equation 41.50 are determined by the steady-state Keldysh equation:

$$G^{\lessgtr}(\mathbf{r},\mathbf{r}';E) = \int d\mathbf{r}_1 \int d\mathbf{r}_2 G^R(\mathbf{r},\mathbf{r}_1;E)\Sigma^{\lessgtr}(\mathbf{r}_1,\mathbf{r}_2;E)G^A(\mathbf{r}_2,\mathbf{r}';E), \tag{41.51}$$

with retarded and advanced components resulting from the Dyson equation:

$$\int d\mathbf{r}_1 \left[\left\{G_0^{R/A}\right\}^{-1}(\mathbf{r},\mathbf{r}_1;E) - \Sigma^{R/A}(\mathbf{r},\mathbf{r}_1;E)\right] G^{R/A}(\mathbf{r}_1,\mathbf{r}';E) = \delta(\mathbf{r}-\mathbf{r}'). \tag{41.52}$$

While the retarded and advanced Green's functions $G^{R/A}$ are related to the charge carrier DOS, the correlation functions $G^{\lessgtr}$ additionally contain information on the (nonequilibrium) occupation of these states. The noninteracting system is described by G_0. The self-energies Σ, on the other hand, are scattering functions describing the renormalization of the Green's functions due to coupling to the environment in the form of interactions with photons (mediating photogeneration and radiative recombination), phonons (mediating relaxation and indirect transitions), and other carriers (relevant for excitonic and Auger processes). An additional self-energy term describes injection and extraction of carriers as contacts with arbitrary chemical potential, enabling the treatment of an open nonequilibrium system. The computation of the NEGF is self-consistently coupled to the Poisson equation (Equation 41.44) for the mean-field electrostatic potential ϕ—entering the equation for G_0—via the charge carrier density:

$$\rho_{e/h}(\mathbf{r}) = \mp i \int \frac{dE}{2\pi} G^{\lessgtr}(\mathbf{r},\mathbf{r};E). \tag{41.53}$$

For an in-depth review of the general NEGF formalism and its application to NS PV, the reader is referred to [189] and references therein.

In contrast to the semiclassical models presented in Section 41.3.2, the microscopic approach provides insight into the physical mechanisms underlying the PV device operation, as those are properly captured by the formalism. On the downside, the method is extremely demanding from a computational point of view and applications are therefore limited to functional NSSC components of mesoscopic extension. An important direction of future research will thus consist in the combination of quantum-kinetic models for local carrier dynamics in NS regions with semiclassical models for the global characteristics of the extended device.

41.3.3.1 Applications to Nanowire SCs

As discussed in Section 41.1, QWRs are not in the focus of efficient solar energy harvesting systems. However, the low dimensionality allows for an efficient numerical treatment of the system. Indeed, the earliest application of NEGF for PV energy conversion considered quasi-1D carbon nanotube systems with tight-binding (TB) band structure [190–192]. More recently, a $\mathbf{k}\cdot\mathbf{p}$-NEGF approach was formulated to study photocurrents in GaAs nanowires [193].

41.3.3.2 Applications to QWSC

Single and double QW structures in III-V material systems were studied in [194,195] using a simple TB band structure and in [79,196] within an effective mass approach, with focus on the modification of absorption and emission spectra by the complex potential profile and on the escape of carriers from confined to extended, current-carrying states. Photocarrier transport regimes from ballistic extraction to phonon-mediated sequential tunneling in SL absorbers were investigated for Si-SiO_x [197] (including phonon-mediated indirect transitions) and InAlGaAs-InGaAs [198].

41.3.3.3 Applications to QDSC

The main challenge for the implementation of NEGF models for QDSC is the complexity of the electronic structure with respect to 3D spatial variation, which prevents the use of the slab models conveniently applied to planar NSs. A simple approach is to use a coarse-grained model with a basis of localized QD orbitals—e.g., the Wannier functions of a regular QDSL—in the spirit of the TB approximation in atomic systems. This method was used to formulate a NEGF model for Si QDSL absorbers embedded in the SiO_x-SiC matrix material [199], and to study the impact of inter-QD and QD-contact coupling on the PV performance of generic QD arrays [200]. Within such an approach, however, it is not possible to study on a microscopic level, similar to that achieved in MQWSC, the carrier escape and capture processes that are instrumental in the operation of MQDSC.

41.4 Case Study: Single Quantum Well p-i-n SC

In this section, the different elements and levels of sophistication in the simulation of NSSC shall be illustrated by consideration of a prototypical device architecture, for which a single quantum well (SQW) p-i-n photodiode is chosen. The device as displayed in Figure 41.3 consists of a 100-nm thin bulk GaAs baseline SC with heavily doped emitter and base components ($N_A = N_D = 10^{18}$ cm^{-3}) of 20-nm extension. In the center of the intrinsic region, a slab of 10-nm InGaAs is inserted, which forms a type-I QW. For the band offsets, $\Delta E_{C/V} = 0.15/0.10$ eV is chosen. The contacts are designed to be perfectly carrier selective, which corresponds to the assumption of ohmic boundary conditions for majorities and infinite barriers for minorities. The light is assumed to be incident from the p-side, whereas a gold reflector is applied to the *n*-side, which, however, is not considered in the electronic simulation.

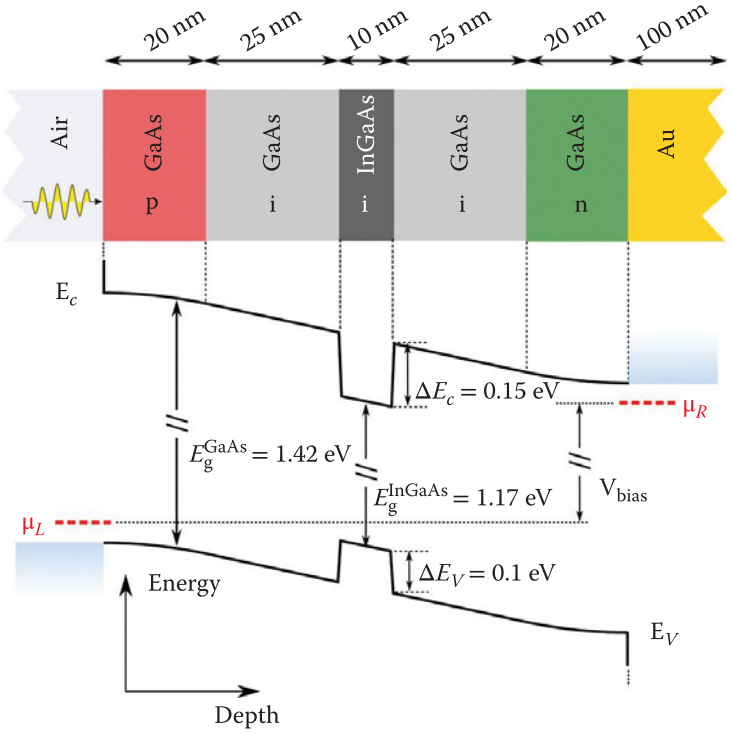


FIGURE 41.3 Single GaAs-InGaAs quantum well p-i-n solar cell architecture used as a test case for the simulation. The contacts are designed to be perfectly selective, i.e., with ohmic boundary conditions for majorities and infinite barriers for minorities. The gold reflector is used only in the optical simulation.

41.4.1 States

In the first step of the simulation, the electronic and optical device states need to be determined. For the description of the electronic structure of the component semiconductor bulk materials, a basic combination of two decoupled single-band effective mass approximation (EMA) models is used, with $m_{e/h}^{*\mathrm{GaAs}} = 0.067/0.1\ m_0$ and $m_{e/h}^{*\mathrm{InGaAs}} = 0.055/0.22\ m_0$. Under the standard thin-film assumption of a transverse continuum due to very large spatial extension perpendicular to the growth direction, the wave functions (WF) can be written as the decomposition

$$\Psi_{n\alpha\mathbf{k}_\parallel}(\mathbf{r}) = \frac{e^{i\mathbf{k}_\parallel \mathbf{r}_\parallel}}{\sqrt{\mathcal{A}}} \psi_{n\alpha\mathbf{k}_\parallel}(z) u_{n0}(\mathbf{r}), \tag{41.54}$$

where n is the bulk Bloch band index, $||$ denotes transverse quantities, $\mathcal{A}$ is the transverse cross-section area, ψ is the envelope function in the growth direction, and u_{n0} is the periodic part of the Bloch function at the Γ point. The 1D EMA Schrödinger equation for the envelope functions of the QW states is solved numerically for closed-system (Dirichlet) boundary conditions and at vanishing field. The wave functions $\psi_{c/v}(z)$ (z: depth) and eigenenergies $\varepsilon_{c/v}$ of the resulting NS states at $\mathbf{k}_\parallel = 0$ are displayed in Figure 41.4a. There are two electron and three hole states with energies below the band edge of the host material, exhibiting the expected symmetry properties and increasing barrier penetration as a function of energy. For comparison, the local density of states (LDOS) at $\mathbf{k}_\parallel = 0$ is computed using the NEGF formalism for identical EMA Hamiltonian and at flat potential (Figure 41.4b), but with open boundary conditions, which, however, does not affect either shape or energies of the confined states. While the solutions of the square well potential are readily available, the situation in the device at operating conditions is quite different, as demonstrated by Figure 41.4c, where the LDOS is shown for a realistic band profile at an operating forward bias voltage

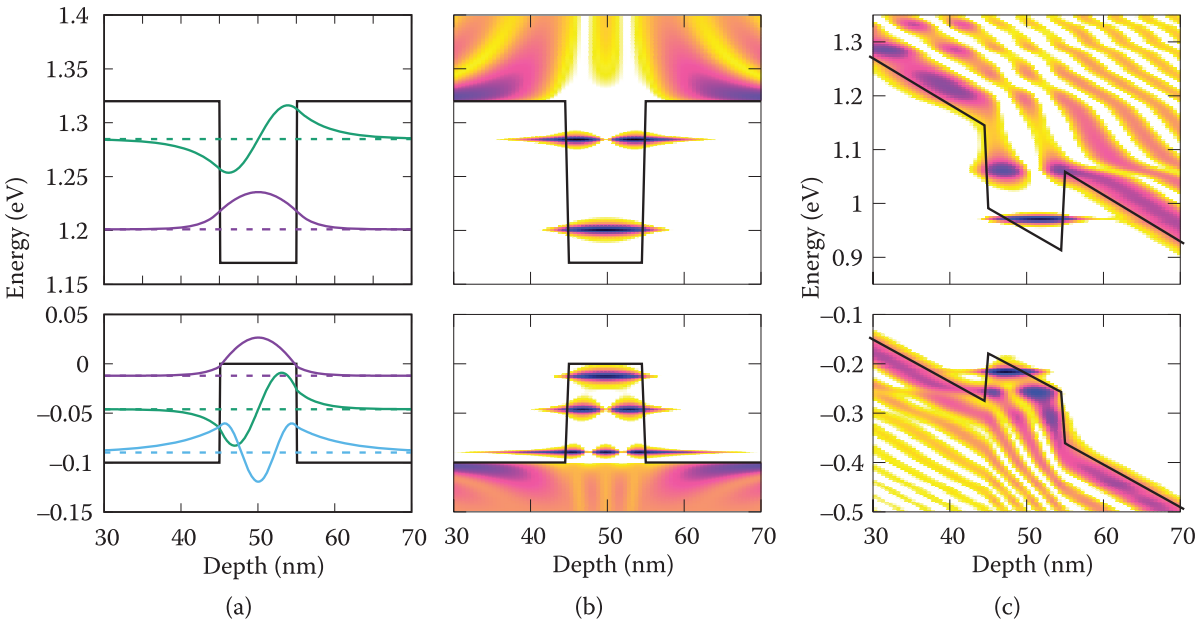


FIGURE 41.4 (a) Wave functions and energies of the eigenstates in the 10-nm InGaAs QW embedded in a GaAs host as obtained by the solution of the 1D single band effective mass Schrödinger equation along the growth direction. (b) One dimensional 1D LDOS ($\mathbf{k}_\parallel = 0$) obtained using the NEGF formalism with the same effective mass Hamiltonian and at flat band conditions. (c) 1D LDOS of the same SQW, but in the p-i-n device configuration close to operating conditions at an applied forward bias voltage $V = 0.8$ V. The most remarkable feature is the unbinding of the higher lying QW states.

of $V = 0.8$ V. Most remarkably, the higher lying QW states lose their confinement in the large built-in field and turn into quasibound resonances.

The optical modes of the device are computed under consideration of the absorbing components and of the reflector using an in-house developed wave propagation solver based on the TMM. As will be shown, this is more appropriate than simple consideration of LB law, since both slab and reflector form an optical cavity, which can lead to nontrivial spatial variation of the optical intensity inside the absorber. Since the states of the quasicontinuum in the QW material are not well described by the chosen approach with closed boundaries, the optical response will be computed only for energies below the host gap. The determination of the optical absorption strength based on the computed electronic states thus forms the next step in the simulation procedure.

41.4.2 Dynamics

The absorption coefficient is evaluated following the FGR procedure outlined in Section 41.2, using the states computed in Section 41.4.1. For energies below the host gap, assuming fully occupied valence bands and empty conduction bands, and neglecting the variation of states and energies with transverse momentum, the absorption coefficient of a square well (SW) potential of width L_w reads

$$\alpha_{\mathrm{SW}}(E_\gamma) = f \cdot E_\gamma^{-1} \sum_{i,j} |M_{c_i v_j}|^2 \Theta\left(E_\gamma - \varepsilon_{c_i v_j}\right), \tag{41.55}$$

with Θ the unit step function and

$$f = \frac{e^2 m_r^* P_{cv}^2}{m_0^2 \hbar n_r c_0 \varepsilon_0 L_w}, \tag{41.56}$$

where m_r^* is the reduced effective mass, $P_{cv} = \sqrt{E_P m_0/6}$ is the bulk MME for Kane energy $E_P = 26.9$ eV, $n_r = 3.7$ is the refractive index of the bulk material, and c_0 and ε_0 are speed of light and permittivity in vacuum, respectively. The overlap matrix elements are

$$M_{c_i v_j} = \int_L dz\, \psi_{c_i}^*(z) \psi_{v_j}(z), \tag{41.57}$$

where L ($> L_w$) is the normalization length of the wave functions, introduce the optical selection rules according to the WF symmetry. As can be verified in Figure 41.5, displaying the above FGR absorption coefficient (full line), only two of the six transitions between the three hole and two electron levels are not suppressed by symmetry. The validity of the selection rules in the absorption coefficient computed via NEGF (dotted line), while evident, is a nontrivial result, which requires careful consideration of the nonlocality in the Green's function picture of electron-photon interaction. In the absence of a symmetric potential, which is not found in the real operating device, the selection rules are strongly relaxed, and the absorption coefficient exhibits a shape that interpolates between 2D and 3D DOS (dashed line).

Based on this absorption coefficient, the emission coefficient $\mathcal{B}$ of the radiative recombination rate introduced in Section 41.2 is determined using the Van Roosbroeck–Shockley (VRS) relation (Equation 41.25). These two quantities are the main ingredients for the assessment of the device characteristics at the radiative limit.

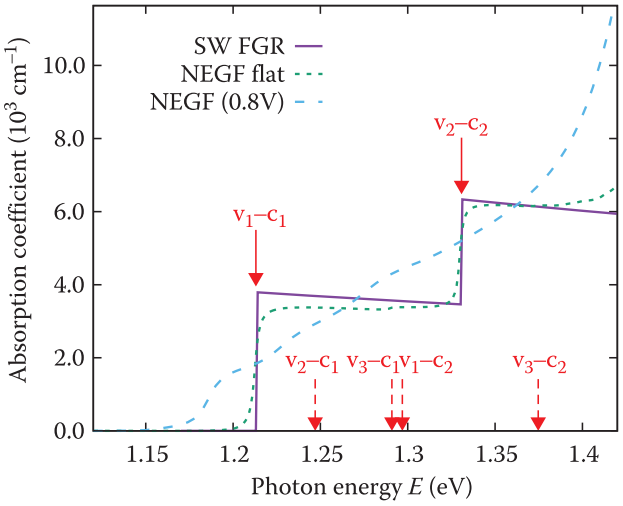


FIGURE 41.5 Absorption coefficient as computed within the Fermi's golden rule approach based on the solution of the single-band effective mass Schrödinger equation for the square well potential (SW FGR, full line), together with the NEGF absorption coefficients (spatially averaged over the QW region) for flat band (dotted) and the real potential at $V = 0.8$ V (dashed). While the selection rules are properly reflected in the NEGF approach, the real spectra feature subband tails and deviate from purely 2-D character due to transitions between quasibound states as a consequence of the large built-in fields present in the intrinsic region of the p-i-n diode.

41.4.3 Characteristics

First, the current–voltage characteristics are evaluated on the level of a detailed balance analysis, i.e., photocurrent and radiative dark current are computed from the absorptance as explained in Section 41.3. To this end, the absorptance is determined both from the LB law (a_{LB}) and from TMM (a_{TMM}) using the previously computed absorption coefficient α_{SW}. The photocurrent density can be evaluated at normal incidence. Dark current, on the other hand, should in principle be evaluated under consideration of the finite opening angle of the loss cone, as determined by the condition of total internal reflection. In the case of emission from high to low refractive index medium, which applies in the situation under consideration here ($n_r^{\mathrm{GaAs}} = 3.6$ to $n_r^{\mathrm{air}} = 1$), the angular dependence may be dropped, i.e., $a(\Omega, E_\gamma) \equiv a(\theta, E_\gamma) \approx a(E_\gamma)$, where the first equality is due to the planar geometry. The angular integration in Equation 41.39 thus provides just a factor of π, and the current–voltage characteristics can be written as

$$J(V) = J_{SC} - J_{0r}\left(e^{qV/(k_B T)} - 1\right), \tag{41.58}$$

with

$$J_{SC} = -q \int \mathrm{d}E_\gamma a(E_\gamma)\Phi_\gamma^{\swarrow}(E_\gamma), \tag{41.59}$$

$$J_{0r} = -q \int \mathrm{d}E_\gamma a(E_\gamma)\pi\Phi_{bb}(E_\gamma). \tag{41.60}$$

For simplicity, a monochromatic incident photon flux at $E_\gamma = 1.3$ eV and intensity of $I_\gamma \equiv \Phi_\gamma^{\swarrow}(E_\gamma)E_\gamma = 0.1$ kW/cm^2 is used for the evaluation of the characteristics. Figure 41.6 displays the JV curves obtained in this way using the absorptances a_{LB} and a_{TMM} based on α_{SW}, which shows that in this unoptimized situation, the reflector does not increase the absorptance of the QW significantly. Comparison with a 1D numerical DD-Poisson-TMM solver (*ASA*-TU Delft) using the same absorption coefficient α_{SW} and corresponding VRS emission coefficient, but otherwise the transport properties of bulk GaAs ($\mu_{n/p} = 400/40$ cm^2/Vs, $\mathcal{N}_{C/V} = 4.35/7.93 \times 10^{17}$ cm^{-3}, $\varepsilon_r = 13.18$, $S_{\mathrm{min/maj}} = 0/10^8$ m/s), reveals—as to be expected—coincidence of the photocurrent with that from a_{TMM}, but also a significant underestimation

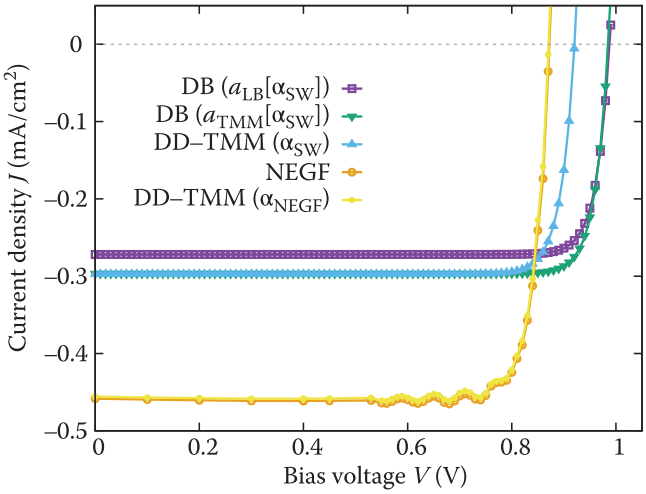


FIGURE 41.6 Current–voltage characteristics as provided by different simulation approaches: detailed balance (DB) currents based on the absorptance from Lambert–Beer's (LB) law or the TMM method, both using the square well potential (SW) absorption coefficient; the semiclassical drift-diffusion-Poisson-TMM model (DD-TMM) for the SW absorption coefficient; the fully microscopic NEGF formalism, revealing a bias-dependent structure in the level of photogeneration. The NEGF characteristics can be reproduced by using in the DD-TMM model the optical data as computed by NEGF, which hints at close to ideal transport in terms of carrier escape from the QW.

of the recombination due to the neglectance of the proper loss cone. The insensitivity to transport is a consequence of the high mobility of GaAs in combination with the short extraction distance and the absence of fast recombination, which results in the flat Fermi levels assumed in the global detailed balance approach. In the full NEGF simulation, on the other hand, both photocurrent and dark current are increased as compared to the characteristics based on the SW absorptance. Remarkably, the NEGF characteristics can be recovered almost completely in the DD-TMM model if local absorption and emission coefficients are extracted from NEGF. This hints at almost perfect transport in terms of unit carrier escape probability (implicitly assumed in the semiclassical model), in consistency with experimental observation for QW structures of similar depth and operated at room temperature. Deviations between the results of the two models are expected in the case of deep wells with presence of fast recombination processes competing with the slow carrier escape: While the NEGF picture considers that a fraction of the photogenerated carriers recombines in the QW prior to escape to the continuum, the semiclassical bulk model assumes the whole generation in bulk states, in which the carriers diffuse quickly from the highly absorbing and—due to detailed balance—emitting QW locations to regions with lower optical coupling strength. That regime, however, is currently not accessible due to the lack of a suitable NEGF picture of nonradiative recombination.

41.5 Summary

The use of NSs in photovoltaics is manifold, and many applications require dedicated models to assess their impact on the PV device performance. There is, however, a common rationale in the consideration of NS in PV device models, according to the device architecture and the NS functionality, which in general amounts to the application of a multiscale and multiphysics scheme. Hence, NS device simulation generally involves many different levels of modeling and simulation, from the determination of NS states to their consideration in the dynamics of carrier generation, transport, and recombination and the final synthesis in the propagation of NS properties to features in the global device characteristics. Reliable consideration of NS effects hence demands a derivation of the key PV material parameters (α, μ, τ) based on a consistent microscopic picture. In the case of ideal transport, an assessment in terms of the optical properties and detailed balance relations between absorption and emission gives already a good estimate of the device

performance. For less ideal systems, the individual loss mechanisms need to be considered in more detail, as provided by semiclassical models on device level, down to the microscopic picture of the fundamental physical mechanisms as delivered by quantum-kinetic approaches in the case where nonclassical effects are relevant, such as transport away from diffusive and band-like regimes.

References

1. M. A. Green. Third generation photovoltaics: Ultra-high conversion efficiency at low cost. *Prog. Photovolt. Res. Appl.*, 9:123, 2001.
2. J. Müller, B. Rech, J. Springer, and M. Vanecek. TCO and light trapping in silicon thin film solar cells. *Sol. Energy*, 77(6):917–930, 2004.
3. S. Liu, D. Ding, S. R. Johnson, and Y.-H. Zhang. Approaching single-junction theoretical limit using ultra-thin GaAs solar cells with optimal optical designs. In *38th IEEE Photovoltaic Specialists Conference (PVSC)*, pages 002082–002087, 2012.
4. S. Pillai, K. R. Catchpole, T. Trupke, and M. A. Green. Surface plasmon enhanced silicon solar cells. *J. Appl. Phys.*, 101(9):093105, 2007.
5. I. Massiot, C. Colin, N. Péré-Laperne, P. Roca i Cabarrocas, C. Sauvan, P. Lalanne, J.-L. Pelouard, and S. Collin. Nanopatterned front contact for broadband absorption in ultra-thin amorphous silicon solar cells. *Appl. Phys. Lett.*, 101(16):163901, 2012.
6. J. R. Nagel and M. A. Scarpulla. Enhanced absorption in optically thin solar cells by scattering from embedded dielectric nanoparticles. *Opt. Express*, 18(S2):A139–A146, 2010.
7. C. Heine and R. H. Morf. Submicrometer gratings for solar energy applications. *Appl. Opt.*, 34(14):2476–2482, 1995.
8. P. Bermel, C. Luo, L. Zeng, L. C. Kimerling, and J. D. Joannopoulos. Improving thin-film crystalline silicon solar cell efficiencies with photonic crystals. *Opt. Express*, 15(25):16986–17000, 2007.
9. L. Cao, P. Fan, A. P. Vasudev, J. S. White, Z. Yu, W. Cai, J. A. Schuller, S. Fan, and M. L. Brongersma. Semiconductor nanowire optical antenna solar absorbers. *Nano Lett.*, 10(2):439–445, 2010.
10. K. W. J. Barnham, I. I. Ballard, J. G. Connolly, B. G. Kluftinger N. Ekins-Daukes, J. Nelson, C. Rohr, and M. Mazzer. Recent results on quantum well solar cells. *J. Mater. Sci. Mater. Electron.*, 11:531, 2000.
11. K. R. Catchpole and A. Polman. Plasmonic solar cells. *Opt. Express*, 16(26):21793–21800, 2008.
12. H. A. Atwater and A. Polman. Plasmonics for improved photovoltaic devices. *Nat. Mater.*, 9(3):205–213, 2010.
13. K. R. Catchpole, S. Mokkapati, F. Beck, E.-C. Wang, A. McKinley, A. Basch, and J. Lee. Plasmonics and nanophotonics for photovoltaics. *MRS Bull.*, 36:461–467, 2011.
14. S. Mokkapati and K. R. Catchpole. Nanophotonic light trapping in solar cells. *J. Appl. Phys.*, 112(10):101101, 2012.
15. K. W. J. Barnham and G. Duggan. A new approach to high-efficiency multi-band-gap solar cells. *J. Appl. Phys.*, 67:3490–3493, 1990.
16. D. B. Bushnell, K. W. J. Barnham, J. P. Connolly, M. Mazzer, N. J. Ekins-Daukes, J. S. Roberts, G. Hill, R. Airey, and L. Nasi. Effect of barrier composition and well number on the dark current of quantum well solar cells. In *Proceeding of 3rd World Conference on Photovoltaic Energy Conversion*, Osaka, Japan, page 2709, 2003.
17. L. C. Hirst, M. Fuhrer, D. J. Farrell, A. Le Bris, J.-F. Guillemoles, M. J. Y. Tayebjee, R. Clady, T. W. Schmidt, M. Sugiyama, Y. Wang, H. Fujii, and N. J. Ekins-Daukes. InGaAs/GaAsP quantum wells for hot carrier solar cells. *Proc. SPIE.*, 8256:82560X, 2012.
18. R. Oshima, A. Takata, and Y. Okada. Strain-compensated InAs/GaNAs quantum dots for use in high-efficiency solar cells. *Appl. Phys. Lett.*, 93(8):083111, 2008.
19. S. M. Hubbard, C. D. Cress, C. G. Bailey, R. P. Raffaelle, S. G. Bailey, and D. M. Wilt. Effect of strain compensation on quantum dot enhanced GaAs solar cells. *Appl. Phys. Lett.*, 92(12):123512, 2008.

20. Y. Okada, R. Oshima, and A. Takata. Characteristics of InAs/GaNAs strain-compensated quantum dot solar cell. *J. Appl. Phys.*, 106(2):024306, 2009.
21. C. G. Bailey, D. V. Forbes, R. P. Raffaelle, and S. M. Hubbard. Near 1 V open circuit voltage InAs/GaAs quantum dot solar cells. *Appl. Phys. Lett.*, 98(16):163105, 2011.
22. G. Conibeer, M. Green, E.-C. Cho, D. König, Y.-H Cho, T. Fangsuwannarak, G. Scardera, E. Pink, Y. Huang, T. Puzzer, S. Huang, D. Song, C. Flynn, S. Park, X. Hao, and D. Mansfield. Silicon quantum dot nanostructures for tandem photovoltaic cells. *Thin Solid Films*, 516(20):6748–6756, 2008.
23. J. G. J. Adams, B. C. Browne, I. M. Ballard, J. P. Connolly, N. L. A. Chan, A. Ioannides, W. Elder, P. N. Stavrinou, K. W. J. Barnham, and N. J. Ekins-Daukes. Recent results for single-junction and tandem quantum well solar cells. *Prog. Photovolt. Res. Appl.*, 19(7):865–877, 2011.
24. D. Alonso-Álvarez, A. G. Taboada, J. M. Ripalda, B. Alen, Y. González, L. González, J. M. García, F. Briones, A. Martí, A. Luque, A. M. Sánchez, and S. I. Molina. Carrier recombination effects in strain compensated quantum dot stacks embedded in solar cells. *Appl. Phys. Lett.*, 93(12):123114, 2008.
25. A. Martí, L. Cuadra, and A. Luque. Partial filling of a quantum dot intermediate band for solar cells. *IEEE Trans. Electron Dev.*, 48(10):2394–2399, 2001.
26. A. J. Nozik. Quantum dot solar cells. *Physica E*, 14(1–2):PII S1386–9477(02)00374–0, 2002.
27. R. Patterson, M. Kirkengen, B. Puthen Veettil, D. König, M. A. Green, and G. Conibeer. Phonon lifetimes in model quantum dot superlattice systems with applications to the hot carrier solar cell. *Sol. Energy Mater. Sol. Cells*, 94(11):1931–1935, 2010.
28. V. I. Klimov. Detailed-balance power conversion limits of nanocrystal-quantum-dot solar cells in the presence of carrier multiplication. *Appl. Phys. Lett.*, 89(12):123118, 2006.
29. O. E. Semonin, J. M. Luther, S. Choi, H.-Y. Chen, J. Gao, A. J. Nozik, and M. C. Beard. Peak external photocurrent quantum efficiency exceeding 100% via MEG in a quantum dot solar cell. *Science*, 334(6062):1530–1533, 2011.
30. A. Martí, E. Antolín, C. R. Stanley, C. D. Farmer, N. López, P. Diaz, E. Cánovas, P. G. Linares, and A. Luque. Production of photocurrent due to intermediate-to-conduction-band transitions: A demonstration of a key operating principle of the intermediate-band solar cell. *Phys. Rev. Lett.*, 97(24):247701, 2006.
31. P. G. Linares, A. Martí, E. Antolín, and A. Luque. III-V compound semiconductor screening for implementing quantum dot intermediate band solar cells. *J. Appl. Phys.*, 109(1):014313, 2011.
32. L. C. Hirst, R. J. Walters, M. F. Fuehrer, and N. J. Ekins-Daukes. Experimental demonstration of hot-carrier photo-current in an InGaAs quantum well solar cell. *Appl. Phys. Lett.*, 104(23):231115, 2014.
33. J. A. R. Dimmock, S. Day, M. Kauer, K. Smith, and J. Heffernan. Demonstration of a hot-carrier photovoltaic cell. *Prog. Photovolt. Res. Appl.*, 22(2):151–160, 2014.
34. F. W. Wise. Lead salt quantum dots: The limit of strong quantum confinement. *Acc. Chem. Res.*, 33(11):773–780, 2000.
35. A. J. Nozik, M. C. Beard, J. M. Luther, M. Law, R. J. Ellingson, and J. C. Johnson. Semiconductor quantum dots and quantum dot arrays and applications of multiple exciton generation to third-generation photovoltaic solar cells. *Chem. Rev.*, 110(11):6873–6890, 2010.
36. Y. Kanai, Z. Wu, and J. C. Grossman. Charge separation in nanoscale photovoltaic materials: Recent insights from first-principles electronic structure theory. *J. Mater. Chem.*, 20(6):1053, 2010.
37. A. Ruland, C. Schulz-Drost, V. Sgobba, and D. M. Guldi. Enhancing photocurrent efficiencies by resonance energy transfer in CdTe quantum dot multilayers: Towards rainbow solar cells. *Adv. Mater.*, 23(39):4573–7, 2011.
38. I. J. Kramer and E. H. Sargent. The architecture of colloidal quantum dot solar cells: Materials to devices. *Chem. Rev.*, 114(1):863–882, 2014.
39. R. J. Ellingson, M. C. Beard, J. C. Johnson, P. Yu, O. I. Micic, A. J. Nozik, A. Shabaev, and A. L. Efros. Highly efficient multiple exciton generation in colloidal PbSe and PbS quantum dots. *Nano Lett.*, 5(5):865–871, 2005.

40. D. König, Y. Takeda, and B. Puthen-Veettil. Technology-compatible hot carrier solar cell with energy selective hot carrier absorber and carrier-selective contacts. *Appl. Phys. Lett.*, 101(15), 2012.
41. D. Bozyigit, N. Yazdani, M. Yarema, O. Yarema, W. M. M. Lin, S. Volk, K. Vuttivorakulchai, M. Luisier, F. Juranyi, and V. Wood. Soft surfaces of nanomaterials enable strong phonon interactions. *Nature*, 531(7596):618–622, 2016.
42. W. Shockley and H. J. Queisser. Detailed balance limit of efficiency of p-n junction solar cells. *J. Appl. Phys.*, 32(3):510–519, 1961.
43. R. E. Chandler, A. J. Houtepen, J. Nelson, and D. Vanmaekelbergh. Electron transport in quantum dot solids: Monte Carlo simulations of the effects of shell filling, Coulomb repulsions, and site disorder. *Phys. Rev. B*, 75(8):085325, 2007.
44. P. Guyot-Sionnest. Electrical transport in colloidal quantum dot films. *J. Phys. Chem. Lett.*, 3(9):1169–1175, 2012.
45. D. Beljonne and J. Cornil, editors. *Multiscale Modelling of Organic and Hybrid Photovoltaics*. Berlin: Springer, Berlin, Heidelberg, 2014.
46. J. D Jackson. *Classical Electrodynamics*. Hoboken, NJ: John Wiley & Sons, 1998.
47. P. Yeh, A. Yariv, and C. S. Hong. Electromagnetic propagation in periodic stratified media.1. General theory. *J. Opt. Soc. Am.*, 67(4):423–438, 1977.
48. W. Schäfer and M. Wegener. *Semiconductor Optics and Transport Phenomena*. Berlin: Springer, 2002.
49. H. Haug and A. P. Jauho. *Quantum Kinetics in Transport and Optics of Semiconductors*. Berlin: Springer, 1996.
50. H. Haug and S. W. Koch. *Quantum Theory of the Optical and Electronic Properties of Semiconductors*. Singapore: World Scientific, 2004.
51. U. Aeberhard. Quantum-kinetic theory of photocurrent generation via direct and phonon-mediated optical transitions. *Phys. Rev. B*, 84(3):035454, 2011.
52. G. Allan and C. Delerue. Efficient intraband optical transitions in Si nanocrystals. *Phys. Rev. B*, 66(23):233303, 2002.
53. J.-W. Luo, P. Stradins, and A. Zunger. Matrix-embedded silicon quantum dots for photovoltaic applications: A theoretical study of critical factors. *Energy Environ. Sci.*, 4(7):2546, 2011.
54. D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, and C. A. Burrus. Electric field dependence of optical absorption near the band gap of quantum-well structures. *Phys. Rev. B*, 32:1043–1060, 1985.
55. S.-L. Chuang, S. Schmitt-Rink, D. A. B. Miller, and D. S. Chemla. Exciton Green's-function approach to optical absorption in a quantum well with an applied electric field. *Phys. Rev. B*, 43:1500–1509, 1991.
56. U. Aeberhard. Nonequilibrium Green's function theory of coherent excitonic effects in the photocurrent response of semiconductor nanostructures. *Phys. Rev. B*, 86:115317, 2012.
57. M. Paxman, J. Nelson, B. Braun, J. Connolly, K. W. J. Barnham, C. T. Foxon, and J. S. Roberts. Modeling the spectral response of the quantum well solar cell. *J. Appl. Phys.*, 74:614, 1993.
58. P. J. Stevens, M. Whitehead, G. Parry, and K. Woodbridge. Computer modeling of the electric field dependent absorption spectrum of multiple quantum well material. *IEEE J. Quant. Electron.*, 24:2007, 1988.
59. D. Atkinson, G. Parry, and E. J. Austin. Modelling of electroabsorption in coupled quantum wells with applications to voltage optical modulation. *Semicond. Sci. Technol.*, 5:516, 1990.
60. S. M. Ramey and R. Khoie. Modeling of multiple-quantum-well solar cells including capture, escape, and recombination of photoexcited carriers in quantum wells. *IEEE Trans. Electron Dev.*, 50:1179, 2003.
61. G. Lengyel, K. W. Jelley, and R. W. H. Engelmann. A semi-empirical model for electroabsorption in GaAs/AlGaAs multiple quantum well modulator structures. *IEEE J. Quant. Electron.*, 26(2):296–304, 1990.

62. S. Schmitt-Rink, D. A. B. Miller, and D. S. Chemla. Theory of the linear and nonlinear optical properties of semiconductor microcrystallites. *Phys. Rev. B*, 35:8113–8125, 1987.
63. D. L. Nika, E. P. Pokatilov, Q. Shao, and A. A. Balandin. Charge-carrier states and light absorption in ordered quantum dot superlattices. *Phys. Rev. B*, 76(12):125417, 2007.
64. S. Tomić, T. S. Jones, and N. M. Harrison. Absorption characteristics of a quantum dot array induced intermediate band: Implications for solar cell design. *Appl. Phys. Lett.*, 93(26):263105, 2008.
65. Y. Z. Hu, M. Lindberg, and S. W. Koch. Theory of optically excited intrinsic semiconductor quantum dots. *Phys. Rev. B*, 42(3):1713–1723, 1990.
66. W.-Y. Wu, J. N. Schulman, T. Y. Hsu, and U. Efron. Effect of size nonuniformity on the absorption spectrum of a semiconductor quantum dot system. *Appl. Phys. Lett.*, 51(10):710–712, 1987.
67. A. Vasanelli, R. Ferreira, and G. Bastard. Continuous absorption background and decoherence in quantum dots. *Phys. Rev. Lett.*, 89:216804, 2002.
68. A. W. Walker, O. Thériault, J. F. Wheeldon, and K. Hinzer. The effects of absorption and recombination on quantum dot multijunction solar cell efficiency. *IEEE J. Photovolt.*, 3(3):1118–1124, 2013.
69. R. Riölver, B. Berghoff, D. Bätzner, B. Spangenberg, H. Kurz, M. Schmidt, and B. Stegemann. Si/SiO2 multiple quantum wells for all silicon tandem cells: Conductivity and photocurrent measurements. *Thin Solid Films*, 516(20):6763–6766, 2008.
70. M. Zacharias, J. Heitmann, R. Scholz, U. Kahler, M. Schmidt, and J. Bläsing. Size-controlled highly luminescent silicon nanocrystals: A SiO/SiO2 superlattice approach. *Appl. Phys. Lett.*, 80(4):661–663, 2002.
71. J. Nelson, M. Paxman, K. W. J. Barnham, J. S. Roberts, and C. Button. Steady state carrier escape rates from single quantum wells. *IEEE J. Quant. Electron.*, 29:1460–1468, 1993.
72. O. Y. Raisky, W. B. Wang, R. R. Alfano, C. L. Reynolds Jr., and V. Swaminathan. Investigation of photoluminescence and photocurrent in InGaAsP/InP strained multiple quantum well heterostructures. *J. Appl. Phys.*, 81:339, 1997.
73. K. R. Lefebvre and A. F. M. Anwar. Electron and hole escape times in single quantum wells. *J. Appl. Phys.*, 80(6):3595–3597, 1996.
74. K. R. Lefebvre and A. F. M. Anwar. Electron escape time from single quantum wells. *IEEE J. Quant. Electron.*, 33(2):187–191, 1997.
75. H. Schneider and K. V. Klitzing. Thermionic emission and Gaussian transport of holes in a GaAs/Al*x*Ga1−*x*As multiple-quantum-well structure. *Phys. Rev. B*, 38(9):6160, 1988.
76. A. Larsson, P. A. Andrekson, S. T. Eng, and A. Yariv. Tunable superlattice p-i-n photodetectors: characteristics, theory, and application. *IEEE J. Quant. Electron.*, 24(5):787, 1988.
77. D. J. Moss, T. Ido, and H. Sano. Calculation of photogenerated carrier escape rates from GaAs/Al*x*Ga1−*x*As quantum wells. *IEEE J. Quant. Electron.*, 30(4):1015, 1994.
78. C.-Y. Tsai, L. F. Eastman, Y.-H. Lo, and C.-Y. Tsai. Breakdown of thermionic emission theory for quantum wells. *Appl. Phys. Lett.*, 65(4):469–471, 1994.
79. U. Aeberhard. Microscopic theory and numerical simulation of quantum well solar cells. *Proc. SPIE*, 7597:759702, 2010.
80. P. N. Brounkov, A. A. Suvorova, M. V. Maximov, A. F. Tsatsul'nikov, A. E. Zhukov, A. Yu Egorov, A. R. Kovsh, S. G. Konnikov, T. Ihn, S. T. Stoddart, L. Eaves, and P. C. Main. Electron escape from self-assembled InAs/GaAs quantum dot stacks. *Physica B*, 249–251:267–270, 1998.
81. C. M. A. Kapteyn, F. Heinrichsdorff, O. Stier, R. Heitz, M. Grundmann, N. D. Zakharov, D. Bimberg, and P. Werner. Electron escape from InAs quantum dots. *Phys. Rev. B*, 60:14265–14268, 1999.
82. P. W. Fry, J. J. Finley, L. R. Wilson, A. Lemaitre, D. J. Mowbray, M. S. Skolnick, M. Hopkinson, G. Hill, and J. C. Clark. Electric-field-dependent carrier capture and escape in self-assembled InAs/GaAs quantum dots. *Appl. Phys. Lett.*, 77(26):4344–4346, 2000.
83. S. Marcinkevius and R. Leon. Carrier capture and escape in In*x*Ga1−*x*As/GaAs quantum dots: Effects of intermixing. *Phys. Rev. B*, 59:4630–4633, 1999.

84. M. De Giorgi, C. Lingk, G. von Plessen, J. Feldmann, S. De Rinaldis, A. Passaseo, M. De Vittorio, R. Cingolani, and M. Lomascolo. Capture and thermal re-emission of carriers in long-wavelength InGaAs/GaAs quantum dots. *Appl. Phys. Lett.*, 79(24):3968–3970, 2001.
85. J. A. Brum and G. Bastard. Resonant carrier capture by semiconductor quantum wells. *Phys. Rev. B*, 33:1420–1423, 1986.
86. P. W. M. Blom, J. E. M. Haverkort, and J. H. Wolter. Optimization of barrier thickness for efficient carrier capture in graded-index and separate-confinement multiple quantum well lasers. *Appl. Phys. Lett.*, 58(24):2767–2769, 1991.
87. S. C. Kan, D. Vassilovski, T. C. Wu, and K. Y. Lau. Quantum capture and escape in quantum-well lasers—Implications on direct modulation bandwidth limitations. *IEEE Photonic. Technol. Lett.*, 4(5):428–431, 1992.
88. P. Sotirelis, P. von Allmen, and K. Hess. Electron intersubband relaxation in doped quantum wells. *Phys. Rev. B*, 47:12744–12753, 1993.
89. P. Sotirelis and K. Hess. Electron capture in GaAs quantum wells. *Phys. Rev. B*, 49:7543–7547, 1994.
90. B. K. Ridley. Space-charge-mediated capture of electrons and holes in a quantum well. *Phys. Rev. B*, 50:1717–1724, 1994.
91. C.-Y. Tsai, L. F. Eastman, Y.-H. Lo, and C.-Y. Tsai. Carrier capture and escape in multisubband quantum well lasers. *IEEE Photonic. Technol. Lett.*, 6(9):1088–1090, 1994.
92. L. Thibaudeau and B. Vinter. Phonon-assisted carrier capture into a quantum well in an electric field. *Appl. Phys. Lett.*, 65(16):2039–2041, 1994.
93. D. Bradt, Y. M. Sirenko, and V. Mitin. Inelastic and elastic mechanisms of electron capture to a quantum well. *Semicond. Sci. Technol.*, 10(3):260, 1995.
94. K. Kalna and M. Mosko. Electron capture in quantum wells via scattering by electrons, holes, and optical phonons. *Phys. Rev. B*, 54:17730–17737, 1996.
95. M. Mosko and K. Kalna. Carrier capture into a GaAs quantum well with a separate confinement region: Comment on quantum and classical aspects. *Semicond. Sci. Technol.*, 14(9):790, 1999.
96. J. A. Brum and G. Bastard. Direct and indirect carrier capture by semiconductor quantum wells. *Superlattice. Microst.*, 3(1):51–55, 1987.
97. P. W. M. Blom, C. Smit, J. E. M. Haverkort, and J. H. Wolter. Carrier capture into a semiconductor quantum well. *Phys. Rev. B*, 47(4):2072–2081, 1993.
98. C.-Y. Tsai, C.-Y. Tsai, Y.-H. Lo, and L. F. Eastman. Carrier dc and ac capture and escape times in quantum-well lasers. *IEEE Photonic. Technol. Lett.*, 7(6):599–601, 1995.
99. C.-Y. Tsai, C.-Y. Tsai, Y.-H. Lo, R. M. Spencer, and L. F. Eastman. Nonlinear gain coefficients in semiconductor quantum-well lasers: Effects of carrier diffusion, capture, and escape. *IEEE J. Sel. Top. Quant. Electron.*, 1(2):316–330, 1995.
100. Y. Lam and J. Singh. Monte Carlo studies on the well-width dependence of carrier capture time in graded-index separate confinement heterostructure quantum well laser structures. *Appl. Phys. Lett.*, 63(14):1874–1876, 1993.
101. L. Davis, Y. L. Lam, Y. C. Chen, J. Singh, and P. K. Bhattacharya. Carrier capture and relaxation in narrow quantum wells. *IEEE J. Quant. Electron.*, 30(11):2560–2564, 1994.
102. L. F. Register and K. Hess. Simulation of carrier capture in quantum well lasers due to strong inelastic scattering. *Superlattice. Microst.*, 18(3):223–228, 1995.
103. G. C. Crow and R. A. Abram. Monte Carlo simulations of capture into quantum well structures. *Semicond. Sci. Technol.*, 14(1):1, 1999.
104. M. Ryzhii and V. Ryzhii. Monte Carlo modeling of electron transport and capture processes in AlGaAs/GaAs multiple quantum well infrared photodetectors. *Jpn. J. Appl. Phys.*, 38(10R):5922, 1999.
105. J. Hader, J. V. Moloney, and S. W. Koch. Structural dependence of carrier capture time in semiconductor quantum-well lasers. *Appl. Phys. Lett.*, 85(3):369–371, 2004.

106. A. V. Uskov, J. McInerney, F. Adler, H. Schweizer, and M. H. Pilkuhn. Auger carrier capture kinetics in self-assembled quantum dot structures. *Appl. Phys. Lett.*, 72(1):58–60, 1998.
107. I. Magnusdottir, S. Bischoff, A. V. Uskov, and J. Mørk. Geometry dependence of Auger carrier capture rates into cone-shaped self-assembled quantum dots. *Phys. Rev. B*, 67:205326, 2003.
108. A. V. Uskov, I. Magnusdottir, B. Tromborg, J. Mørk, and R. Lang. Line broadening caused by Coulomb carrier-carrier correlations and dynamics of carrier capture and emission in quantum dots. *Appl. Phys. Lett.*, 79(11):1679–1681, 2001.
109. R. Ferreira and G. Bastard. Phonon-assisted capture and intradot Auger relaxation in quantum dots. *Appl. Phys. Lett.*, 74(19):2818–2820, 1999.
110. J. M. Miloszewski, M. S. Wartak, S. G. Wallace, and S. Fafard. Theoretical investigation of carrier capture and escape processes in cylindrical quantum dots. *J. Appl. Phys.*, 114(15):154311, 2013.
111. J.-Z. Zhang and I. Galbraith. Rapid hot-electron capture in self-assembled quantum dots via phonon processes. *Appl. Phys. Lett.*, 89(15):153119, 2006.
112. I. Magnusdottir, A. V. Uskov, R. Ferreira, G. Bastard, J. Mørk, and B. Tromborg. Influence of quasibound states on the carrier capture in quantum dots. *Appl. Phys. Lett.*, 81(23):4318–4320, 2002.
113. T. R. Nielsen, P. Gartner, and F. Jahnke. Many-body theory of carrier capture and relaxation in semiconductor quantum-dot lasers. *Phys. Rev. B*, 69:235314, 2004.
114. J. Seebeck, T. R. Nielsen, P. Gartner, and F. Jahnke. Polarons in semiconductor quantum dots and their role in the quantum kinetics of carrier relaxation. *Phys. Rev. B*, 71:125327, 2005.
115. G. D. Mahan. *Many-Particle Physics*, 2nd ed. New York, NY: Plenum, 1990.
116. T. Gunst, T. Markussen, K. Stokbro, and M. Brandbyge. First-principles method for electron-phonon coupling and electron mobility: Applications to two-dimensional materials. *Phys. Rev. B*, 93:035414, 2016.
117. C.-W. Jiang and M. A. Green. Silicon quantum dot superlattices: Modeling of energy bands, densities of states, and mobilities for silicon tandem solar cell applications. *J. Appl. Phys.*, 99(11):114902, 2006.
118. A. Wacker. Semiconductor superlattices: A model system for nonlinear transport. *Phys. Rep.*, 357:1–111, 2002.
119. P. Würfel. *Physics of Solar Cells: From Basic Principles to Advanced Concepts*. Weinheim: Wiley-VCH, 2009.
120. I. Mora-Sero, L. Bertoluzzi, V. Gonzalez-Pedro, S. Gimenez, F. Fabregat-Santiago, K. W. Kemp, E. H. Sargent, and J. Bisquert. Selective contacts drive charge extraction in quantum dot solids via asymmetry in carrier transfer kinetics. *Nat. Commun.*, 4:2272, 2013.
121. W. Van Roosbroeck and W. Shockley. Photon-radiative recombination of electrons and holes in germanium. *Phys. Rev.*, 94:1558, 1954.
122. E. O. Göbel, H. Jung, J. Kuhl, and K. Ploog. Recombination enhancement due to carrier localization in quantum well structures. *Phys. Rev. Lett.*, 51(17):1588–1591, 1983.
123. A. R. Beattie and P. T. Landsberg. Auger effect in semiconductors. *Proc. R. Soc. London Ser. A*, 249:16, 1959.
124. P. T. Landsberg. *Recombination in Semiconductors.* Cambridge: Cambridge University Press, 1991.
125. L. Chiu and A. Yariv. Auger recombination in quantum-well InGaAsP heterostructure lasers. *IEEE J. Quant. Electron.*, 18(10):1406–1409, 1982.
126. C. Smith, R. A. Abram, and M. G. Burt. Auger recombination in a quantum well heterostructure. *J. Phys. C Solid State Phys.*, 16(5):L171, 1983.
127. N. K. Dutta. Calculation of Auger rates in a quantum well structure and its application to InGaAsP quantum well lasers. *J. Appl. Phys.*, 54(3):1236–1245, 1983.
128. A. Sugimura. Auger recombination effect on threshold current of InGaAsP quantum well lasers. *IEEE J. Quant. Electron.*, 19(6):932–941, 1983.
129. P. K. Basu. Auger recombination rate in quantum well lasers: Modification by electronelectron interaction in quasi two dimensions. *J. Appl. Phys.*, 56(11):3344–3346, 1984.

130. A. Haug. Phonon-assisted auger recombination in quantum well semiconductors. *Appl. Phys. A*, 51:354, 1990.
131. M. Takeshima. Phonon-assisted Auger recombination in a quasi-two-dimensional structure semiconductor. *Phys. Rev. B*, 30(6):3302–3308, 1984.
132. L.-W. Wang, M. Califano, A. Zunger, and A. Franceschetti. Pseudopotential theory of auger processes in CdSe quantum dots. *Phys. Rev. Lett.*, 91:056404, 2003.
133. G. A. Narvaez, G. Bester, and A. Zunger. Carrier relaxation mechanisms in self-assembled (In,Ga)As/GaAs quantum dots: Efficient $P \rightarrow S$ Auger relaxation of electrons. *Phys. Rev. B*, 74:075403, 2006.
134. S. Tomić. Intermediate-band solar cells: Influence of band formation on dynamical processes in InAs/GaAs quantum dot arrays. *Phys. Rev. B*, 82(19):195321, 2010.
135. K. Hyeon-Deuk and O. V. Prezhdo. Time-domain ab initio study of Auger and phonon-assisted auger processes in a semiconductor quantum dot. *Nano Lett.*, 11(4):1845–1850, 2011.
136. N. J. Ekins-Daukes, K. W. J. Barnham, J. P. Connolly, J. S. Roberts, J. C. Clark, G. Hill, and M. Mazzer. Strain-balanced GaAsP/InGaAs quantum well solar cells. *Appl. Phys. Lett.*, 75(26):4195–4197, 1999.
137. K. Ding, U. Aeberhard, O. Astakhov, U. Breuer, M. Beigmohamadi, S. Suckow, B. Berghoff, W. Beyer, F. Finger, R. Carius, and U. Rau. Defect passivation by hydrogen reincorporation for silicon quantum dots in SiC/SiO*x*hetero-superlattice. *J. Non-Cryst. Solids*, 358(17):2145–2149, 2012.
138. W. Shockley and W. T. Read. Statistics of the recombinations of holes and electrons. *Phys. Rev.*, 87:835–842, 1952.
139. R. N. Hall. Electron-hole recombination in germanium. *Phys. Rev.*, 87:387, 1952.
140. A. Schenk. An improved approach to the Shockley-Read-Hall recombination in inhomogeneous fields of space-charge regions. *J. Appl. Phys.*, 71(7):3339–3349, 1992.
141. A. Alkauskas, Q. Yan, and C. G. Van de Walle. First-principles theory of nonradiative carrier capture via multiphonon emission. *Phys. Rev. B*, 90:075202, 2014.
142. U. Aeberhard. Simulation of nanostructure-based high-efficiency solar cells: Challenges, existing approaches, and future directions. *IEEE J. Sel. Top. Quant. Electron*, 19(5):4000411, 2013.
143. G. L. Araújo and A. Martí. Absolute limiting efficiencies for photovoltaic energy conversion. *Sol. Energy Mater. Sol. Cells*, 33:213, 1994.
144. U. Rau. Reciprocity relation between photovoltaic quantum efficiency and electroluminescent emission of solar cells. *Phys. Rev. B*, 76(8):085303, 2007.
145. M. Born and E. Wolf. *Principles of Optics*, 7th ed. Cambridge: Cambridge University Press, 1999.
146. P. Würfel. The chemical potential of radiation. *J. Phys. C Solid State Phys.*, 15:3967, 1982.
147. U. Aeberhard. Simulation of ultra-thin solar cells beyond the limits of the semi-classical bulk picture. *IEEE J. Photovolt.*, 6(3):654–660, 2016.
148. U. Aeberhard. Impact of built-in fields and contact configuration on the characteristics of ultra-thin GaAs solar cells. *Appl. Phys. Lett.*, 109(3), 2016.
149. N. G. Anderson. Ideal theory of quantum well solar cells. *J. Appl. Phys.*, 78:1861, 1995.
150. G. L. Araújo, A. Martí, F. W. Ragay, and J. H. Wolter. Efficiency of multiple quantum well solar cells. In *Proceeding of 12th European Photovoltaic Solar Energy Conference*, Amsterdam, The Netherlands, page 1481, 1994.
151. G. L. Araújo and A. Martí. Electroluminescence coupling in multiple quantum well diodes and solar cells. *Appl. Phys. Lett.*, 66(7):894–895, 1995.
152. S. P. Bremner, R. Corkish, and C. B. Honsberg. Detailed balance efficiency limits with quasi-Fermi level variations. *IEEE Trans. Electron Dev.*, 46(10):1932–1939, 1999.
153. A. Luque, A. Martí, and L Cuadra. Thermodynamic consistency of sub-bandgap absorbing solar cell proposals. *IEEE Trans. Electron Dev.*, 48:2118, 2001.
154. J. P. Connolly, J. Nelson, K. W. J. Barnham, I. Ballard, C. Roberts, J. S. Roberts, and C. T. Foxon. Simulating multiple quantum well solar cells. In *Proceeding of 28th IEEE Photovoltaic Specialists Conference*, Anchorage, Alaska, USA, page 1304, 2000.

155. J. C. Rimada and L. Hernández. Modelling of ideal AlGaAs quantum well solar cells. *Microelectron. J.*, 32(9):719–723, 2001.
156. S. J. Lade and A. Zahedi. A revised ideal model for AlGaAs/GaAs quantum well solar cells. *Microelectron. J.*, 35(5):401–410, 2004.
157. J. C. Rimada, L. Hernandez, J. P. Connolly, and K. W. J. Barnham. Quantum and conversion efficiency calculation of AlGaAs/GaAs multiple quantum well solar cells. *Phys. Status Solidi B*, 242(9):1842–1845, 2005.
158. J. C. Rimada, L. Hernández, J. P. Connolly, and K. W J Barnham. Conversion efficiency enhancement of AlGaAs quantum well solar cells. *Microelectron. J.*, 38(4–5):513–518, 2007.
159. M. Courel, J. C. Rimada, and L. Hernández. GaAs/GaInNAs quantum well and superlattice solar cell. *Appl. Phys. Lett.*, 100(7):073508, 2012.
160. C. I. Cabrera, J. C. Rimada, M. Courel, L. Hernandez, J. P. Connolly, A. Enciso, and D. A. Contreras-Solorio. Modeling multiple quantum well and superlattice solar cells. *Nat. Resour.*, 4(03):235–245, 2013.
161. A. Chatterjee, A. Kumar Biswas, and A. Sinha. An analytical study of the various current components of an AlGaAs/GaAs multiple quantum well solar cell. *Physica E*, 72:128–133, 2015.
162. M. Courel, J. C. Rimada, and L. Hernandez. An approach to high efficiencies using GaAs/GaInNAs multiple quantum well and superlattice solar cell. *J. Appl. Phys.*, 112(5):054511, 2012.
163. A. Alemu and A. Freundlich. Resonant thermotunneling design for high-performance single-junction quantum-well solar cells. *IEEE J. Photovolt.*, 2(3):256–260, 2012.
164. M. Courel, J. C. Rimada, and L. Hernández. AlGaAs/GaAs superlattice solar cells. *Prog. Photovolt. Res. Appl.*, 21(3):276–282, 2013.
165. A. Luque and A. Martí. Increasing the efficiency of ideal solar cells by photon induced transitions at intermediate levels. *Phys. Rev. Lett.*, 78(26):5014–5017, 1997.
166. Z. Wang, T. P. White, and K. R. Catchpole. Plasmonic near-field enhancement for planar ultra-thin photovoltaics. *IEEE Photon. J.*, 5(5):8400608, 2013.
167. W. Hu, M. Maksudur Rahman, M.-Y. Lee, Y. Li, and S. Samukawa. Simulation study of type-II Ge/Si quantum dot for solar cell applications. *J. Appl. Phys.*, 114(12):124509, 2013.
168. Q. Shao, A. A. Balandin, A. I. Fedoseyev, and M. Turowski. Intermediate-band solar cells based on quantum dot supracrystals. *Appl. Phys. Lett.*, 91(16):163503, 2007.
169. Q. Deng, X. Wang, H. Xiao, C. Wang, H. Yin, H. Chen, Q. Hou, D. Lin, J. Li, Z. Wang, and X. Hou. An investigation on $In_xGa_{1-x}N$/GaN multiple quantum well solar cells. *J. Phys. D Appl. Phys.*, 44(26):265103, 2011.
170. W. Wei, Q. Zhang, S. Zhao, and Y. Zhang. Two intermediate bands solar cells of InGaN/InN quantum dot supracrystals. *Appl. Phys. A*, 116(3):1009–1016, 2014.
171. A. Mellor, A. Luque, I. Tobias, and A. Marti. Realistic detailed balance study of the quantum efficiency of quantum dot solar cells. *Adv. Funct. Mater.*, 24(3):339–345, 2014.
172. G. Wei, K.-T. Shiu, N. C. Giebink, and S. R. Forrest. Thermodynamic limits of quantum photovoltaic cell efficiency. *Appl. Phys. Lett.*, 91(22), 2007.
173. J. Nelson, I. Ballard, K. Barnham, J. P. Connolly, J. S. Roberts, and M. Pate. Effect of quantum well location on single quantum well p-i-n photodiode dark currents. *J. Appl. Phys.*, 86:5898, 1999.
174. P. Kailuweit, R. Kellenbenz, S. P. Philipps, W. Guter, A. W. Bett, and F. Dimroth. Numerical simulation and modeling of GaAs quantum-well solar cells. *J. Appl. Phys.*, 107(6):064317, 2010.
175. T. Kirchartz, K. Seino, J.-M. Wagner, U. Rau, and F. Bechstedt. Efficiency limits of Si/SiO2 quantum well solar cells from first-principles calculations. *J. Appl. Phys.*, 105(10):104511, 2009.
176. C.-Y. Tsai and C.-Y. Tsai. Effects of carrier escape and capture processes on quantum well solar cells: A theoretical investigation. *IET Optoelectron.*, 3(6):300–304, 2009.
177. O. Kengradomying, S. Jiang, Q. Wang, N. Vogiatzis, and J. M. Rorison. Modelling escape and capture processes in GaInNAs quantum well solar cells. *Phys. Status Solidi C Curr. Top. Solid State Phys.*, 10(4):585–588, 2013.

178. J. Nelson, K. W. J Barnham, J. P Connolly, G. Haarpaintner, C. Button, and J. Roberts. Quantum well solar cell dark currents. In *Proceeding of 12th EU PVSEC*, Amsterdam, The Netherlands, 1994.
179. M. Gioannini, A. P. Cedola, N. Di Santo, F. Bertazzi, and F. Cappelluti. Simulation of quantum dot solar cells including carrier intersubband dynamics and transport. *IEEE J. Photovolt.*, 3(4):1271–1278, 2013.
180. K. Driscoll, M. F. Bennett, S. J. Polly, D. V. Forbes, and S. M. Hubbard. Effect of quantum dot position and background doping on the performance of quantum dot enhanced GaAs solar cells. *Appl. Phys. Lett.*, 104(2):023119, 2014.
181. K. Yoshida, Y. Okada, and N. Sano. Self-consistent simulation of intermediate band solar cells: Effect of occupation rates on device characteristics. *Appl. Phys. Lett.*, 97(13):133503, 2010.
182. Y.-X. Gu, X.-G. Yang, H.-M. Ji, P.-F. Xu, and T. Yang. Theoretical study of the effects of InAs/GaAs quantum dot layer's position in i-region on current-voltage characteristic in intermediate band solar cells. *Appl. Phys. Lett.*, 101(8):081118, 2012.
183. T. Sogabe, T. Kaizu, Y. Okada, and S. Tomić. Theoretical analysis of GaAs/AlGaAs quantum dots in quantum wire array for intermediate band solar cell. *J. Renew. Sustain. Ener.*, 6(1):011206, 2014.
184. V. Aroutiounian, S. Petrosyan, A. Khachatryan, and K. Touryan. Quantum dot solar cells. *J. Appl. Phys.*, 89(4):2268, 2001.
185. A. Martí, L. Cuadra, and A. Luque. Quasi-drift diffusion model for the quantum dot intermediate band solar cell. *IEEE Trans. Electron Dev.*, 49(9):1632–1639, 2002.
186. C.-C. Lin, M.-H. Tan, C.-P. Tsai, K.-Y. Chuang, and T. S. Lay. Numerical study of quantum-dot-embedded solar cells. *IEEE J. Sel. Top. Quant. Electron.*, 19(5), 2013.
187. V. Aroutiounian, S. Petrosyan, and A. Khachatryan. Studies of the photocurrent in quantum dot solar cells by the application of a new theoretical model. *Sol. Energy Mater. Sol. Cells*, 89(2–3):165–173, 2005.
188. S. M. Willis, J. A. R. Dimmock, F. Tutu, H. Y. Liu, M. G. Peinado, H. E. Assender, A. A. R. Watt, and I. R. Sellers. Defect mediated extraction in InAs/GaAs quantum dot solar cells. *Sol. Energy Mater. Sol. Cells*, 102:142–147, 2012.
189. U. Aeberhard. Theory and simulation of quantum photovoltaic devices based on the non-equilibrium Green's function formalism. *J. Comput. Electron.*, 10:394–413, 2011.
190. D. A. Stewart and F. Leonard. Photocurrents in nanotube junctions. *Phys. Rev. Lett.*, 93:107401, 2004.
191. D. A. Stewart and F. Leonard. Energy conversion efficiency in nanotube optoelectronics. *Nano Lett.*, 5:219, 2005.
192. J. Guo, M. A. Alam, and Y. Yoon. Theoretical investigation on photoconductivity of single intrinsic carbon nanotubes. *Appl. Phys. Lett.*, 88(13):133111, 2006.
193. A. Buin, A. Verma, and S. Saini. Optoelectronic response calculations in the framework of k.p coupled to non-equilibrium Green's functions for 1D systems in the ballistic limit. *J. Appl. Phys.*, 114(3):033111, 2013.
194. U. Aeberhard and R. H. Morf. Microscopic nonequilibrium theory of quantum well solar cells. *Phys. Rev. B*, 77:125343, 2008.
195. U. Aeberhard. Spectral properties of photogenerated carriers in quantum well solar cells. *Sol. Energy Mater. Sol. Cells*, 94(11):1897–1902, 2010.
196. N. Cavassilas, F. Michelini, and M. Bescond. Theoretical comparison of multiple quantum wells and thick-layer designs in InGaN/GaN solar cells. *Appl. Phys. Lett.*, 105(6), 2014.
197. U. Aeberhard. Theory and simulation of photogeneration and transport in Si-SiOx superlattice absorbers. *Nanoscale Res. Lett.*, 6:242, 2011.
198. U. Aeberhard. Simulation of nanostructure-based and ultra-thin film solar cell devices beyond the classical picture. *J. Photon. Energy*, 4(1):042099, 2014.
199. U. Aeberhard. Effective microscopic theory of quantum dot superlattice solar cells. *Opt. Quant. Electron.*, 44:133–140, 2012.
200. A. Berbezier and U. Aeberhard. Impact of nanostructure configuration on the photovoltaic performance of quantum-dot arrays. *Phys. Rev. Applied*, 4:044008, 2015

42

Nanowire Solar Cells: Electro-Optical Performance

42.1 Introduction 475
42.2 Efficiency of Nanostructured Solar Cells 476
Solar Cell Current Density • Thermal Equilibrium • Nonequilibrium • Efficiency Evaluation
42.3 Computational Models 481
Optical Models • Electronic Models
42.4 Light Capture and Emission 483
Absorptivity • Modal Analysis • Shockley–Queisser Efficiency of a Nanowire Array • Purcell Effect
42.5 Electronic Properties 489
Radial versus Axial Junction Arrangement • Surface Recombination
42.6 Summary 492

Bernd Witzigmann

42.1 Introduction

Nanowire arrays show unique optical and electronic features for photovoltaic applications. It has been demonstrated that they exhibit intrinsic optical antireflection properties [1,2]. Moreover, they also can act as optical microconcentrators [3] that focus light to an active region. Nanowires have a large aspect ratio, improving the electronic properties by combining a long optical absorption path along the wire axis and a short lateral carrier extraction [4]. Short carrier extraction paths are beneficial for the carrier collection efficiency. These mechanisms are illustrated in Figure 42.1. On the left, a conventional bulk solar cell collects light in the vertical direction, and the photogenerated carriers move to the respective contacts causing the photo current. A nanostructured solar cell (Figure 42.1a) can act as an optical concentrator in order to focus the incoming light to a small area. Moreover, the carrier extraction path length to the contacts is reduced, which minimizes the probability of carrier loss by unwanted recombination.

Finally, multijunction cells can be realized with strain-relaxed axial heterointerfaces. This allows optimum bandgap combinations [5] and low-defect densities at the same time. As a potential drawback, the large surface to volume ratio can lead to enhanced surface recombination. Epitaxial passivation techniques have been developed in order to suppress surface recombination [6].

In this chapter, a detailed simulation study of the electro-optical characteristics of nanowire array solar cells is presented. An efficiency analysis is derived that treats solar cells with nanostructures on a broader basis using the Shockley–Queisser framework [7]. As addition, the modification of optical density of states (DOS) and wave optics effects are included. Using this framework, the numerical simulation of nanostructured solar cells is discussed. The simulation model consists of the three-dimensional solution of the vectorial Helmholtz equation for solar illumination and a drift-diffusion/Poisson equation system for

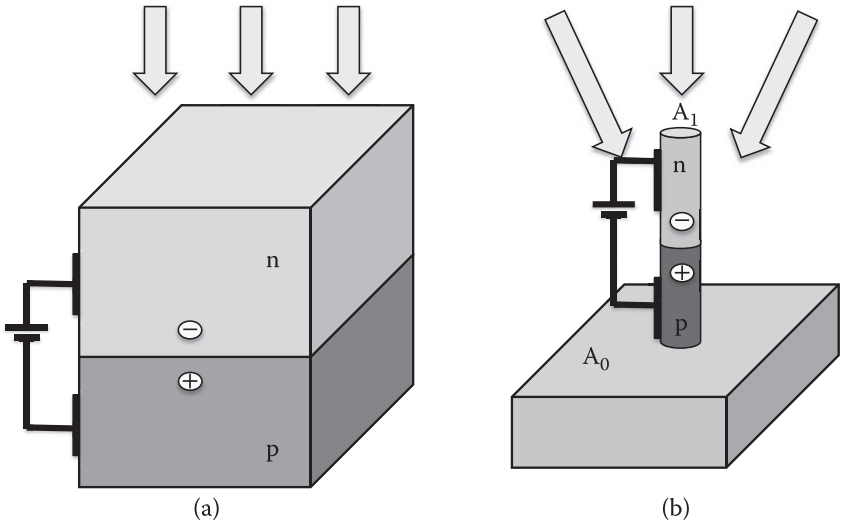


FIGURE 42.1 (a) Bulk solar cell with p-n junction. (b) Solar cell structured as a nanowire.

carrier extraction. In the results section, it is first shown how a nanowire array can act as an electromagnetic broadband absorber. Second, the choice of material composition and dopant placement is discussed with respect to efficiency in single-junction designs. Surface recombination is discussed for both radial and axial p-n junction arrangement.

42.2 Efficiency of Nanostructured Solar Cells

In this section, the efficiency of solar cells that contain nanostructures is discussed. Nanostructures can modify the optical or the electronic properties. As a basic theory, the established derivation by Shockley and Queisser [7] is used, which assumes detailed balance in the radiative limit, i.e., nonradiative recombination processes are not included. The main extension of Shockley and Queisser's classical model is to include the fact that light is viewed as electromagnetic waves and their modes in resonant structures. This leads to a modal dispersion and a strong local variation of the electromagnetic power. Moreover, resonant structures modify the optical DOS, and those enter the absorption as well as the emission properties of a cell, which is based on the Purcell effect [8].

The focus will be put on semiconductor materials, although the derivation can be transferred to other materials in a straightforward manner.

42.2.1 Solar Cell Current Density

The electronic current density of the solar cell can be easily obtained from current continuity in the steady state. It is derived from Maxwell's equations and builds on the mass conservation principle of charged particles:

$$\nabla \vec{J} = q(G - R), \tag{42.1}$$

with $\vec{J}$ being the local current density, G and R representing generation and recombination rates, respectively, and q being the elementary charge. In the integral form, the continuity equation reads

$$\int \vec{J} \mathrm{d}\vec{A} = q \int (G - R)\, \mathrm{d}V. \tag{42.2}$$

The surface area integration on the left hand side is carried out over the contacts (where carriers can leave or enter the solar cell), and the right hand side integration includes the solar cell volume.

42.2.2 Thermal Equilibrium

In thermal equilibrium, all of the emission and absorption processes of the solar cell can be equated, as there is no net current at the contacts. The left hand side of Equation 42.2 is zero, therefore

$$\int G\mathrm{d}V = \int R\mathrm{d}V = \int G^0\mathrm{d}V = \int R^0\mathrm{d}V. \tag{42.3}$$

The local generation rate per volume is

$$G^0 = \int \alpha \mathrm{d}j_{\mathrm{phot}}, \tag{42.4}$$

with α being the absorption coefficient (a material property) and j_{phot} is the photon current density (in units $1/\mathrm{m}^2\mathrm{s}$), which in thermal equilibrium is basically a blackbody radiation at the given reservoir temperature T_0. In order to relate the photon current density to the blackbody radiation (which is an energy density per spectral width), the following relation is used:

$$\mathrm{d}j_{\mathrm{phot}} = \mathrm{d}E_{\mathrm{phot}} \frac{c}{n} \frac{1}{\hbar\omega} \tag{42.5}$$

and

$$dE_{\mathrm{phot}} = D_{\mathrm{opt}} f_{\mathrm{phot}} \hbar\omega \mathrm{d}\hbar\omega. \tag{42.6}$$

The optical DOS D_{opt} depends on the dielectric surrounding of the solar cell and f_{phot} is the photon distribution function. Within an electromagnetic theory, D_{opt} can be calculated as the electromagnetic response of a local dipole excitation [9]:

$$D_{\mathrm{opt}}(\omega, \vec{x}) = \frac{1}{2\pi} \sum_m \int_{\vec{k}} \mathrm{d}\vec{k}\delta(\omega - \omega_m) |\vec{e}_{\mathrm{d}} \vec{E}_{m,\vec{k}}(\vec{r}|^2, \tag{42.7}$$

where $\vec{E}$ is the electric field as obtained from the solution of Maxwell's equations, $\vec{e}_{\mathrm{d}}$ is the dipole unit source, and ω_m the eigenfrequencies of the system.

The photon distribution function in equilibrium is

$$f_{\mathrm{phot}} = \frac{1}{\exp\left(\frac{\hbar\omega}{k_B T_0}\right) - 1}. \tag{42.8}$$

The generation rate, therefore, can be expressed as

$$G^0(\vec{x}) = R^0(\vec{x}) = \int \alpha(\vec{x}, \hbar\omega) \frac{c}{n_r} D^{\mathrm{sc}}_{\mathrm{opt}}(\vec{x}, \hbar\omega) f_{\mathrm{phot}}(T_0) d\hbar\omega, \tag{42.9}$$

which is equal to the recombination in equilibrium.

42.2.3 Nonequilibrium

In the nonequilibrium case, the solar cell is illuminated by solar radiation power, which in our case is represented by an external blackbody source at the temperature T_{s}. This leads to an energy density outside of the solar cell of

$$\mathrm{d}E^{\mathrm{s}}_{\mathrm{phot}} = D^{\mathrm{vac}}_{\mathrm{opt}} f_{\mathrm{phot}}(T_{\mathrm{s}}) \hbar\omega d\hbar\omega. \tag{42.10}$$

The optical DOS here is the vacuum DOS, and the solar temperature T_s enters the Bose–Einstein photon distribution function. In principle, other illumination sources can be used, which then enter the calculation, e.g., as tabulated energy density. The external photon source generates additional electron–hole pairs ΔG inside the solar cell, and therefore

$$G^s = \int \alpha a \frac{c}{n_r^{vac}} D_{opt}^{vac} f_{phot}(T_s) \mathrm{d}\hbar\omega. \tag{42.11}$$

Here, a is the relative absorptivity, the amount of electromagnetic solar power entering the solar cell. It can be calculated using electromagnetic theory by

$$a(\vec{x}) = \frac{\langle \int Re\left\{\nabla \vec{S}\right\} \mathrm{d}V\rangle}{\int Re\left\{\vec{S}\right\} \mathrm{d}\vec{A}}, \tag{42.12}$$

where $\vec{S}$ is the Poynting vector, and the integration in the denominator is performed over a closed surface around the solar cell. This is especially important for nonplanar devices where the solar energy is captured not only from the surface, but from an angle of incidence larger than the half space π. The $\langle\rangle$ symbol in the numerator denotes an averaging over a volume that contains several optical wavelengths so that an inclusion into a drift-diffusion carrier transport framework is possible. The absorptivity $a(\vec{x})$ is the local relative electromagnetic energy entering in the solar cell, and in the ray optics approximation often is expressed as $(1 - r)\exp(-\alpha x)$, where the first factor is the back-reflected power from the air–top surface interface, and the second factor is the diminishing optical intensity inside the device due to internal absorption (Beer–Lambert law).

In the nonequilibrium case, the generation in the cell is increased, which in turn leads to an increased radiative recombination rate. Assuming that the radiative recombination is proportional to the product of electron and hole density, one gets

$$R^s = R^0 \frac{np}{n_i^2} = \frac{N_c N_v}{n_i^2} \exp\left(\frac{\eta_c \eta_v}{k_B T_0}\right) G^0 = G^0 \exp\left(\frac{\Delta E_F}{k_B T_0}\right), \tag{42.13}$$

with the following nomenclature: The intrinsic density is n_i, the effective state densities are N_c or N_v, the reduced Fermi energies are $\eta_c = E_F^n - E^c$ and $\eta_v = -E_F^p + E^v$, and ΔE_F is the quasi-Fermi level separation.

With these quantities, the current densities of the solar cell can be calculated. Using the current continuity equation, one gets

$$\int \vec{J} \mathrm{d}\vec{A} = q \int (G^s - R^s) \mathrm{d}V = q \int \left(G^s - G^0 \exp(\frac{\Delta E_F}{k_B T_0})\right) \mathrm{d}V = I_{gen} - I_{rec}. \tag{42.14}$$

The quasi-Fermi level separation can be approximated by the voltage at the contacts, and after integration of the left hand side, the current of the solar cell becomes

$$I = I_{gen} - I_{sat} \exp\left(\frac{qV}{k_B T_0}\right). \tag{42.15}$$

The currents introduced here are the total generation and recombination currents I_{gen} and I_{rec}, and the reverse saturation current I_{sat}. They can be evaluated from

$$I_{sat} = q \int \int \alpha \frac{c}{n_r^{sc}} D_{opt}^{sc} f_{phot}(T_0) \, \mathrm{d}\hbar\omega \mathrm{d}V \tag{42.16}$$

and

$$I_{\text{gen}} = q \int \int \alpha a \frac{c}{n_r^{\text{vac}}} D_{\text{opt}}^{\text{vac}} f_{\text{phot}}(T_s) \mathrm{d}\hbar\omega \mathrm{d}V. \tag{42.17}$$

The short-circuit current of a solar cell is the current under illumination with zero voltage at the contacts, and therefore accounts to

$$I_{\text{sc}} = I_{\text{gen}} - I_{\text{sat}}. \tag{42.18}$$

With this definition, the current–voltage relationship can be written in the more conventional Shockley form by replacing I_{gen} in Equation 42.15:

$$I = I_{\text{sat}} \left(1 - \exp(\frac{qV}{k_B T_0}) \right) + I_{\text{sc}}. \tag{42.19}$$

The open-circuit voltage is the voltage where the net current flow at the contacts is zero, and it can be evaluated from the relations obtained here as

$$V_{\text{oc}} = \frac{k_B T_0}{q} \log \left(\frac{I_{\text{gen}}}{I_{\text{sat}}} \right) = \frac{k_B T_0}{q} \log \left(1 + \frac{I_{\text{sc}}}{I_{\text{sat}}} \right). \tag{42.20}$$

42.2.4 Efficiency Evaluation

From the current versus voltage characteristics, the electro-optical efficiency of a solar cell can be obtained from

$$\eta_{\text{sc}} = \max \left(\frac{IV}{P_{\text{inc}}} \right), \tag{42.21}$$

where the denominator is the total, spectrally integrated incident solar power. Obtaining a high efficiency translates into maximizing the open-circuit voltage V_{oc} and the short-circuit current I_{sc} as defined in Section 42.2.3, which maximizes the product IV. A maximum short-circuit current is achieved by capturing the incident broadband solar radiation and absorbing the photons in the solar cell. According to Equation 42.17, this requires a large absorption coefficient α, which is a material property, and a good absorptivity a, which is the ability to capture light into the solar cell, an electromagnetic property. In bulk solar cells, the absorptivity a can be maximized by broadband antireflection coatings. If the solar cell thickness is in the order of a few optical wavelengths, waveguiding or even evanescent wave effects can help maximize a. It will be shown in Section 42.2.4.2 how this is applied to nanowire solar cells.

The evaluation of the open-circuit voltage involves a maximum optical generation current I_{gen} and a minimum reverse saturation current I_{sat}, which is a challenge to fulfill. As can be seen from Equation 42.16, the major ingredient to I_{sat} is the optical DOS of the solar cell.

42.2.4.1 Bulk Solar Cell

First, the efficiency evaluation is shown for a bulk solar cell. In principle, all current commercial technologies fall into this category, also the thin-film designs. For a homogeneous medium, the optical DOS is

$$D_{\text{opt}}^{\text{hom}} = \left(\frac{n_r}{c} \right)^3 \frac{1}{4\pi\hbar^3} (\hbar\omega)^2 \Omega, \tag{42.22}$$

where Ω is the solid angle of the radiation emission and absorption, respectively.

For the bulk solar cell, the solid angle of illumination is $\Omega = \pi$, and therefore

$$D_{\text{opt}}^{\text{vac}} = \left(\frac{1}{c}\right)^3 \frac{1}{4\hbar^3}(\hbar\omega)^2. \tag{42.23}$$

The solid angle of emission at any location inside the solar cell is $\Omega = 4\pi$, leading to

$$D_{\text{opt}}^{\text{sc}} = \left(\frac{n_r}{c}\right)^3 \frac{1}{\hbar^3}(\hbar\omega)^2. \tag{42.24}$$

For the absorptivity a, a simple ray optics model can be applied, which leads to

$$a = (1 - r)(1 - \exp(-\alpha d)) \approx 1, \tag{42.25}$$

with r being the reflectivity at the air-semiconductor interface, and d the cell thickness. By design, one can realize a good antireflective coating and sufficient cell thickness, leading to the relation $a \approx 1$. The short-circuit current can therefore be maximized by engineering the cell.

The open-circuit voltage contains the ratio of $\frac{I_{\text{gen}}}{I_{\text{sat}}}$, which is approximately proportional to the ratio $\frac{D_{\text{opt}}^{\text{vac}} n_r^{\text{sc}}}{D_{\text{opt}}^{\text{sc}} n_r^{\text{vac}}}$. This results in a factor $\frac{1}{4n_r^2}$ in the short-circuit current calculation. The physics behind this factor is the light intensity concentration. The factor of 4 in the denominator can be reduced by efficient light management, as it comes from the photon collection from the angle π but photon emission into an angle of 4π [10]. Introducing a dielectric environment that allows emission into an angle π only (by, e.g., photonic crystals or conventional mirrors) increases the open-circuit voltage. Therefore, light management is an important physical mechanism for advanced solar cell design [10,11].

42.2.4.2 Optical Microconcentrators

Waveguides, resonators, or optical antennas can influence the optical absorption behavior of a solar cell. As first case, an optical concentrator that acts as absorber is taken that focuses the incoming solar radiation to the absorber area A_1, whereas the illuminated (large) area is A_0 (for illustration, see Figure 42.1). This leads to an absorptivity $a \propto \frac{A_0}{A_1} = f$, as the absorber concentrates the energy flux density to the area A_1. As a consequence, the short-circuit current is obtained with a much smaller volume of absorption compared to a nonconcentrated case (see integral in Equation 42.18). At the same time, the reverse saturation current (Equation 42.16) is also integrated over a smaller volume compared to a bulk solar cell. The ratio of $\frac{I_{\text{gen}}}{I_{\text{sat}}}$, and therefore the open circuit voltage, increases. However, two assumptions need to be made: First, the reverse saturation current density needs to be independent of the local carrier density, which is, within the Shockley theory, a valid assumption. Second, the optical concentrator modifies the optical DOS of the absorber. This needs to be taken into account in the open-circuit voltage calculation. Therefore, the open-circuit voltage is only increased if the following relation holds:

$$\frac{D_{\text{opt}}^{\text{vac}}}{D_{\text{opt}}^{\text{mc}}} f > \frac{D_{\text{opt}}^{\text{vac}}}{D_{\text{opt}}^{\text{bulk}}}, \tag{42.26}$$

or

$$\frac{D_{\text{opt}}^{\text{bulk}}}{D_{\text{opt}}^{\text{mc}}} f > 1, \tag{42.27}$$

where $D_{\text{opt}}^{\text{bulk}}$ is the optical DOS for the bulk cell and $D_{\text{opt}}^{\text{mc}}$ is the optical DOS for the micro concentrated cell, respectively; their ratio can be interpreted as the Purcell factor. This result shows an interesting design opportunity for high-efficiency nanostructured solar cells: If optical elements are introduced that focus the

solar electromagnetic energy to a small volume, the open-circuit voltage, and therefore the efficiency, can be increased compared to a bulk or thin-film solar cell. However, the Purcell factor, or the ratio of optical DOS, should not be increased at the same rate, as it compensates the concentration effect.

42.2.4.3 Electronic Nanostructures

In the field of lasers and light-emitting diodes, electronic nanostructures such as quantum wells and quantum dots have increased the performance considerably [12]. In solar cells, the application of electronic nanostructures has not been used as a standard feature to improve the cell efficiency. In general, the introduction of nanostructures alters the electronic band structure. It leads to the following consequences:

- The effective bandgap, i.e., the absorption edge, is shifted to the lowest subband energy difference.
- The absorption coefficient increases, as the electronic DOS is higher at the band edge compared to a bulk material, and in addition, the overlap of the envelope wave functions enters the absorption coefficient.
- The absorption coefficient becomes anisotropic due to the anisotropy of the underlying carrier wave functions and DOS.

As a drawback, electronic nanostructures typically have dimensions around 10 nm, which is much smaller than the optical wavelength or a typical absorption path length. Therefore, in order to obtain efficient absorption, multiple nanometer-sized regions need to be stacked in the direction of the light path or efficient focusing of the light onto the absorbing nanostructure needs to be done. Fundamentally, the anisotropy of the absorption and emission can be used to restrict the absorptivity α to the solid angles of the solar radiation, and to reduce the reverse saturation current by a maximum theoretical factor of 4 (see explanation in Section 42.2.4.1) [13]. In addition, the increase of the absorption coefficient leads to smaller active regions, which increases the ratio of short-circuit current to reverse saturation current as outlined in Section 42.2.4.2.

Finally, some advanced concepts have been pursued with electronic nanostructures such as intermediate band solar cells [14] or hot-carrier contacts [15]. The basic idea is to introduce additional energy levels with nanostructures in order to overcome the single bandgap limitation of the semiconductor. Such structures require nonequilibrium carrier transport models and are not discussed here.

42.3 Computational Models

In the previous sections, some fundamental properties of nanostructured solar cells have been described within the semianalytical framework of the Shockley–Queisser model. In general, the incoming light generates electron–hole pairs by absorption in the semiconductor. In Section 42.2, it has been shown that the efficiency of a solar cell depends on the solar optical generation $G^{s}(\vec{x})$, the generation in equilibrium $G^{0}(\vec{x})$, and the recombination under illumination $R^{s}(\vec{x})$ (see Equation 42.14). In principle, those are all local quantities. In nanostructured solar cells, both light intensity and carrier density can be controlled on the scale of tens of nanometers by optical waveguides, plasmonic subwavelength structures, or quantum wells or dots. This introduces a strong local variation of the physical quantities contributing to the cell efficiency. Computational models can capture the local variations, and a correct determination of the efficiency is feasible.

42.3.1 Optical Models

The governing optical equations for solar cells with structures similar or smaller than the optical wavelength are Maxwell's equations. The solar illumination is modeled as a broadband current source, and the resulting electromagnetic field contributes to the local carrier generation rate. There is a multitude of models to solve Maxwell's equations either analytically, semianalytically, or numerically.

A popular analytical solution is the plane wave ansatz, which after some simplifications leads to the transfer matrix method (TMM). Using TMM, light propagation in multilayer structures can be evaluated or the even simpler Beer–Lambert law can be obtained:

$$I(x) = I_0 e^{-\alpha x}, \tag{42.28}$$

with the intensity I.

Most semianalytical methods expand the solution of Maxwell's equations in modes. For nanowire solar cells, which are essentially cylinders, the modes of a cylindrical optical fiber can be used, and the solution is then represented as superposition of the fiber modes [16].

Numerical methods essentially discretize the simulation domain and solve the respective equations only on parts of the domain, e.g., vertices or edges. The finite-difference time-domain (FDTD) method [17] is a popular method that discretizes space and time using explicit finite differences. With sources placed in the simulation domain, the propagation of electromagnetic waves can be computed iteratively in space and time. The FDTD method is straightforward to implement and can be adapted for parallel computing in a natural manner. The original FDTD kernels [17] require rectangular tensor product meshes, which can be restrictive for complex geometries. Moreover, the solutions are only second-order accurate in space and time, which requires many grid points per wavelength for accurate solutions, in particular for optical frequencies. A rather elegant approach to overcome these limitations is the discontinuous Galerkin time-domain method [18]. It combines explicit time stepping with a local finite-element space discretization, and therefore powerful high-order accuracy and flexible meshes can be used. The local elements are connected with constraints for the numerical flux in order to ensure consistency and convergence to the exact solution. For solar cell simulations, which use the broadband solar spectrum as source, the time-domain methods allow the use of spectral broadband sources, with only one simulation run for the entire spectrum. Dispersive materials need special treatment in the time domain, and therefore additional computational resources.

A second class of numerical methods solve Maxwell's equations in the frequency domain. Here, monochromatic harmonic waves replace the time derivatives with an imaginary frequency factor, and any multifrequency solution can be built as superposition in the linear case. Maxwell's equations can be reduced to the Helmholtz equation, which for the electric field $\vec{E}$ reads

$$\left(\Delta - \omega^2 \mu\epsilon\right) \vec{E} = -j\mu\vec{J}, \tag{42.29}$$

where μ and ϵ are the respective permeability and permittivity, ω is the frequency, and $\vec{J}$ is a current source. This equation is valid for homogeneous isotropic materials, and the permittivity is in general a dispersive complex quantity in optics. This equation can be solved by a finite-element method in three dimensions, with sophisticated meshes, and an approximation to the exact solution. In its homogeneous form (without current sources), it yields an eigenvalue problem that give the eigenfrequencies and eigenvectors of a given geometry. For complex permittivities, the eigenvalue problem becomes complex as well, and the imaginary part of the eigenfrequency gives the quality factor of the eigenmode [19,20].

In order to evaluate the efficiency of a solar cell beyond the analytical Shockley–Queisser limit, the framework presented in Section 42.2 as efficiency analysis for nanostructured solar cells requires evaluation of the classical electromagnetic field for the solar cell geometry in the presence of a complex refractive index. This enters both the absorptivity a (see Equation 42.12) and the optical DOS D_{opt} (Equation 42.7).

42.3.2 Electronic Models

Any electronic model describes the dynamics of the charge carriers inside the solar cell, i.e., the photogenerated carriers and the intrinsic carriers under applied bias. In a generalized description, solar cell operation means that the photogenerated carriers thermally relax to the band edges and are subsequently driven to

the contacts creating the photocurrent. In the radiative (i.e., Shockley–Queisser) limit, the carriers can recombine radiatively on their way to the contacts, and of course, the resulting photons can be reabsorbed. This gives an upper limit to the photocurrent. However, carriers can be lost by nonradiative recombination, which decreases the efficiency further. An excellent description of such an efficiency analysis can be found in [21]. In order to understand local features of the carrier transport, such as surface recombination effects, or defect recombination related to specific doping species or materials, a space-resolved carrier transport model is needed. In nanostructures, additional physical properties are potentially relevant, e.g., the thermal relaxation process is different from bulk processes, which can lead to ballistic or semiballistic carrier transport. If these heated carriers can be detected by a contact without thermal relaxation, a higher efficiency is achieved as the thermalization energy loss can be avoided [22].

Therefore, electronic properties of nanostructures require simulations by quantum transport models in order to include effects such as carrier tunneling, phase coherence (and carrier localization), and ballistic transport. These models have been successfully applied to nanowire devices by several groups [23] and established methods are the nonequilibrium Green's functions (NEGF) method or the Wigner function formalism [24]. The NEGF method has been applied to quantum solar cells [25,26]; however, these methods in general are computationally demanding, as they discretize both real and phase space. Therefore, simulation domains of only a few nanometers can be accounted for. In the future, the so called multiscale approaches might be feasible that combine computationally intense quantum transport methods with efficient semiclassical methods; several approaches have been made, with still many open questions [27].

In the following sections, the carrier transport in nanowire solar cells is analyzed. It will turn out that these solar cells have dimensions with diameters larger than 50 nm and lengths more than 1000 nm. In this regime, semiclassical models still have a decent validity, therefore the established and efficient drift-diffusion model will be used. It comprises the self-consistent solution of the current continuity equations for electrons and holes with the Poisson equation for the electrostatic potential.

42.4 Light Capture and Emission

Figure 42.2 shows the arrangement of the nanowire array for a solar cell. The wire diameter is d, the lattice constant is a, and the wire height is L. The first question is how such an array can incouple and absorb light. Solar radiation spans from the ultraviolet to the infrared wavelength range, and the solar cell therefore needs to be a broadband absorber. It has been reported in the literature that an array of nanowires can act as a good broadband absorber [16,28]. For highest efficiency with a single bandgap material, the classical Shockley–Queisser analysis predicts a band gap of the semiconductor material between $E_g \approx 1.3$ eV and 1.4 eV [11].This bandgap optimum can be explained by the balance of having minimum carrier thermalization losses from high-energy photons on the one hand and minimizing the losses due

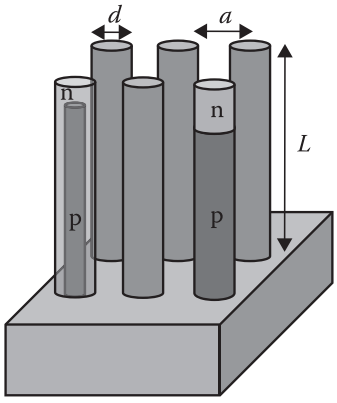


FIGURE 42.2 Illustration of an ordered nanowire array solar cell. The p-n junction can be arranged in a radial (left) or axial (right) manner. The periodicity can be hexagonal or cubic.

to the optical transparency of the semiconductor for low-energy photons. Indium phosphide (InP) with $E_g = 1.34eV$ matches this range very well; therefore, the following analysis will use this material. The findings are applicable to any other material as well.

42.4.1 Absorptivity

Using a three-dimensional finite-element solution of the complex vectorial Maxwell equations, the wire arrangement as shown in Figure 42.2 is illuminated from the top side with a linearly polarized sheet source. The source matches the power spectrum of AM1.5D solar power and is placed at a distance above the wires, which is beyond the evanescent mode decay length. The space between the wires is vacuum, and both the wire and the substrate material is InP, with the complex refractive indices following the dispersion relation given in [12]. The simulation is carried out for a quarter of the wire only, with periodic boundary conditions on the cut planes in order to represent an infinite periodicity of the array . Perfectly matched layer absorbing boundary conditions are placed at the top and the bottom of the simulation domain.

From the electromagnetic fields, the relative local absorptivity $a(\vec{x}, \lambda)$ is calculated according to Equation 42.12, which becomes, after integration of the wire volume, the spectral absorptivity of the entire wire $\tilde{a}(\lambda)$. The optical absorption in the substrate is not included in the integration and does not contribute to the efficiency of the cell. Figure 42.3 shows the spectral absorptivity for different wire periodicity and diameters; the wire height is constant at $L = 2$ μm. A geometric fill factor is defined that describes the volume of active material in a nanostructured cell relative to a bulk material volume. This fill factor can be expressed as

$$f_{nw} = \frac{\pi d^2}{4a^2}. \tag{42.30}$$

The highest fill factor can be achieved with $d = a$ and a fill factor of $f_{nw} \approx 0.79$, i.e., when the nanowires touch. For the simulations shown in Figure 42.3, combinations of periodicity and diameter have been

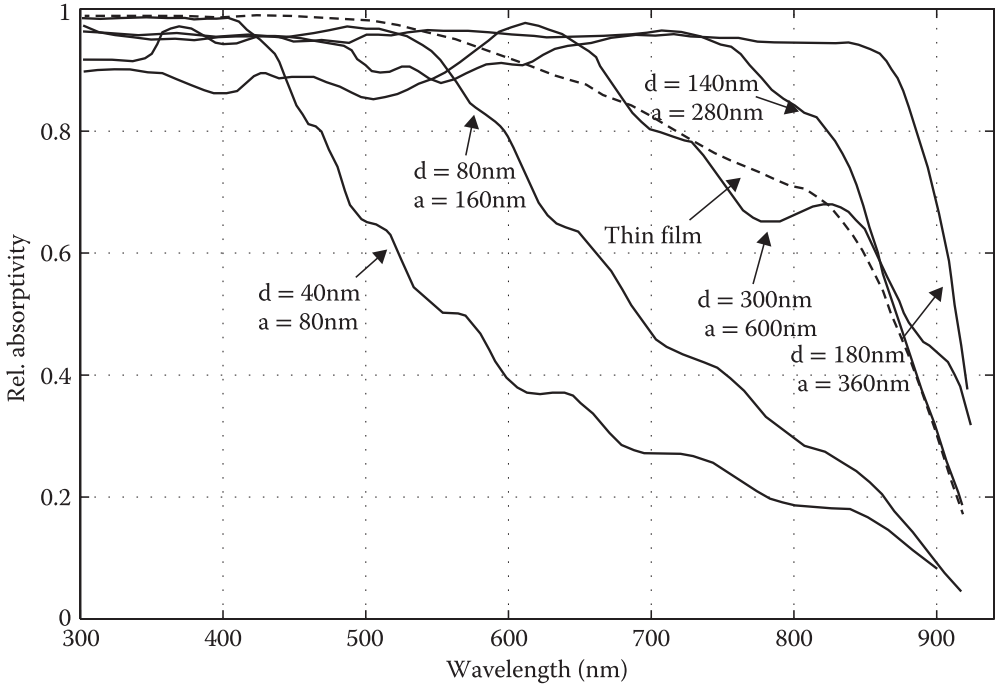


FIGURE 42.3 Spectral relative absorptivity of an InP nanowire array. The wire diameter is labeled d, and the wire distance is a.

chosen so that the material volume fill factor is identical at $f_{nw} = 0.196$. The arrangement with the smallest wire diameter of $d = 40$ nm and distance of $a = 80$ nm shows a substantial reduction of absorptivity above a wavelength of 500 nm. Nanowires with larger diameter and larger array pitch show a better long wavelength absorptivity, and the case of $d = 180$ nm and $a = 360$ nm has an absorptivity of more than 90% up to the bandgap wavelength of InP. The dashed curve shown in Figure 42.3 is the absorptivity of a homogeneous InP film with a thickness of $2\mu m \times 0.196 = 392$ nm, which is the same material volume as the nanowire array. This thin film, however, is covered by an ideal antireflective coating in order to eliminate reflections (which is not added to the wire array case). The decrease of absorptivity at wavelengths larger than 600 nm comes from the insufficient thickness of the InP layer.

42.4.2 Modal Analysis

In order to understand the findings of Section 42.4.1, the spectral behavior of the absorptivity of the nanowire array can be analyzed by a modal description of the electromagnetic field distribution. In [16], it has been shown that the absorptivity obtained from the three-dimensional calculation can be constructed from a separation of the power absorptivity parallel to the nanowire axis according to the Beer–Lambert law. The effective attenuation coefficient is obtained from the two-dimensional mode calculation of a cross-section of the wire array perpendicular to the wire axis. This assumes an infinite extension of the wires along their axes, and the calculation gives the eigenmodes with their complex eigenvalues $n_m^{\text{eff}} = n'_m + in''_m$, with m as the mode index. The real part of the eigenvalue describes the coupling efficiency of the electromagnetic incident TEM wave to the wire array and is inserted in the separation ansatz for the calculation of the reflected power. The imaginary part is the effective absorptivity that enters the Beer–Lambert expression. For an illustration, see Figure 42.4. The optical power at the interface between vacuum and the nanowire array top can be expressed as

$$P_m(z=0) = P_0 \frac{4n_0 n'_m}{(n_0 + n'_m)^2} \frac{\Gamma_m}{\sum_i \Gamma_i}, \tag{42.31}$$

with n_0 being the vacuum refractive index, P_0 the source power, and Γ_m is the overlap integral of the electric field of the incoming TEM mode with the mth mode of the wire array cross-section. The absorbed power in the nanowire material is then

$$P_{\text{abs}}(\lambda) = \sum_i P_i(z=0)\left(1 - \exp\left(\frac{4\pi}{\lambda} n''_i L\right)\right). \tag{42.32}$$

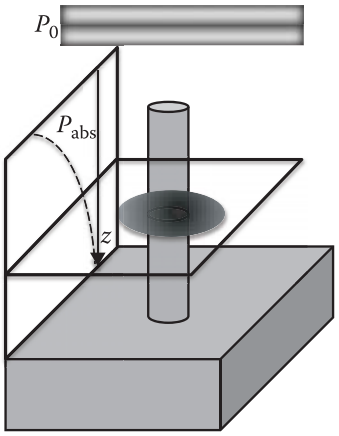


FIGURE 42.4 Illustration of the electromagnetic separation approach for a nanowire array. The incoming plane wave power P_0 is separated into two-dimensional cross-sectional modes and an exponential power decay in z-direction following a Beer–Lambert law.

Using this separation ansatz, the spectral absorptivity curves of the three-dimensional calculation can be reproduced with excellent accuracy, as shown in [28].

This procedure identifies the type of modes that couple the linearly polarized free-space electromagnetic field to the nanowire array. The unit cell of the array consists of the wire itself (which resembles a cylindrical fiber) and the periodic boundary to the neighboring wires that can be viewed as the Neumann boundary condition at the border of the unit cell.

The broadband absorptivity for the nanowire array can be described by only very few modes. For the case of $d = 180$ nm and $a = 360$ nm, only two modes are necessary to create the absorptivity characteristic as shown in Figure 42.3. These modes are plotted in Figure 42.5 with their electric field mode profile and their relative incoupling efficiency. The latter describes the ability of the wire array to act as an antireflective structure.The effect of relative incoupling in Figure 42.5 (bottom figure) is a combination of the confinement factor Γ and the real part of the effective index of the in-plane mode in the nanowire array n' (see Equation 42.31). For incoupling of 90%, the effective index therefore is close to 1 (i.e., the surrounding air), which for a fiber mode, is the regime of weak confinement. Still, most of the electromagnetic power couples into the wire array, and the field distribution as shown in in Figure 42.5 (e.g., mode A at 500 nm) shows a strong field between the wires, which is also absorbed according to the Beer–Lambert law. Between 300 nm and 700 nm, mode A couples close to 90% into the wire, and the electric field analysis shows that the electric field distribution resembles an fiber mode. Finally, the electric field distribution between the wires resembles a fundamental TEM mode, meaning parallel arrangement of the electric field vector. At 750 nm, mode A reaches its cutoff wavelength, and most of the light couples into mode B, which is of type HE_{11}. For further detail, the reader is referred to the optical fiber standard dispersion relations and eigenmodes, which are similar to the case of the sparse wire array.

It is surprising that only two modes are responsible for the broadband absorptivity of the electromagnetic power. This modal description helps understand the characteristics: If the wire spacing is increased from the ideal distance, the overlap integral of the wire modes decreases, and therefore the absorptivity decreases. If the wire spacing is decreased, the effective index of the wire array n' increases, and therefore the relative incoupling decreases. If, at the same distance, the wire diameter is modified, similar effects occur; an increase of the wire diameter moves the fiber mode dispersion relation to the multimode regime, where the effective indices are larger (a standard optical fiber dispersion relation illustrates this). A decrease of the wire diameter will decrease the mode number to a point where only the fundamental mode can be used.

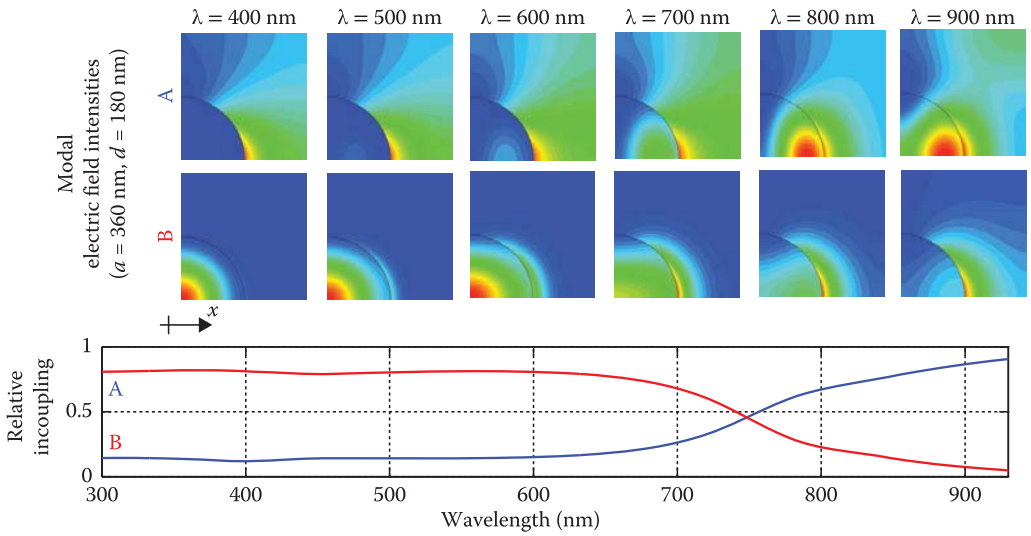


FIGURE 42.5 Top: Electric field distribution in the nanowire array for mode A and mode B, for different wavelengths. Bottom: Relative incoupling P_A/P_0 and P_B/P_0 (see Equation 42.31). (Reprinted from J. Kupec et al., *Optics Express*, 18, 27589–27605, 2010. With permission.)

As a rule of thumb, therefore, the wire distance and diameters should be chosen so that in the dispersion relation of the cross-sectional eigenvalue problem, the bandwidth of operation can span the fundamental and a few higher order modes. Moreover, the wire length should be chosen so that the imaginary part of the eigenvalue ensures full absorption of the optical power according to the Beer–Lambert law.

The modal analysis of the wire array can also be applied for solar radiation under an incident angle different from vertical illumination. For moderate incident angles (below 15°), the results do not change significantly.

42.4.3 Shockley–Queisser Efficiency of a Nanowire Array

Having calculated the relative absorptivity from electromagnetics, the Shockley–Queisser efficiency can subsequently be evaluated from Equation 42.21. Carrier transport effects are not included at this level of description and will be described later. Also, as approximation, the optical DOS in the semiconductor $D_{\mathrm{opt}}^{\mathrm{sc}}$ is taken to be a bulk DOS, for simplicity. This neglects the Purcell effect for the radiative emission, and therefore we get an upper limit for the solar cell efficiency. The impact of the modified optical DOS will be discussed in Section 42.4.4.

Figure 42.6 shows the calculated efficiency of the InP nanowire array with a square wire arrangement and a wire length of $L = 2\,\mu\mathrm{m}$. The efficiency is plotted versus the volumetric fill factor. For small wire diameters of below 100 nm, the efficiency is highest for high fill factors and decreases for larger wire distances. Here, mostly the fundamental HE_{11} mode is responsible for the absorptivity; however, the overlap with the wire is too small for efficient absorption. For larger wire distance, the electromagnetic energy therefore is lost in the substrate. For thicker wires with diameters larger than 100 nm, the most dense array with highest fill factor shows the same efficiency as the thinner wires; however, with increasing wire distance, the efficiency increases up to values of 30% for the diameter of $d = 180$ nm. This increase can be explained by a reduced effective index of the wire array with lower fill factor, and therefore less back reflection from the vacuum–wire array interface. The electromagnetic energy is concentrated to the wires by the modal light capture, and the decrease of efficiency below fill factors of 10% is due to the decreasing coupling factor. As a result, a maximum efficiency of 30% can be reached with a fill factor of approximately 20%, a wire diameter of $d = 180$ nm, and a wire separation of $a = 360$ nm, without an additional anti-reflective coating. Hence, the wire array acts as a microconcentrator element as well as an active absorber at the same time. An InP bulk cell with ideal broadband antireflective coating gives a detailed balance efficiency of $\eta_{\mathrm{ideal}} = 31\%$ for comparison.

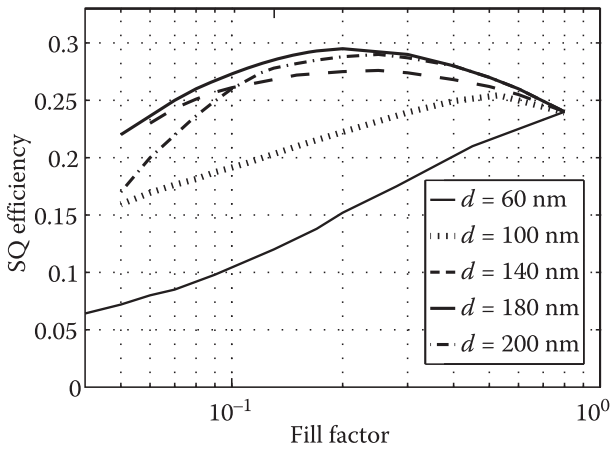


FIGURE 42.6 Efficiency of an InP nanowire array based on the Shockley–Queisser analysis in the radiative limit for different wire diameters *d*. An infinitely thick, perfect antireflective InP bulk cell has an efficiency of approximately 32%.

42.4.4 Purcell Effect

The nanowire array acts as a microconcentrator, and therefore, as discussed in Section 42.2.4.2, the open-circuit voltage can be increased compared to a bulk solar cell, if the concentration factor outweighs the increase in optical DOS. The reciprocity theorem states that any lens or concentrator not only increases the collection angle of the solar cell, but also the emission angle. In addition, the spontaneous emission rate (which in the bulk case is a material property) can be altered if the dielectric surrounding of the emission volume is modified, which changes the recombination rates. For the InP nanowire model array, the question is now how large the Purcell factor is and if it substantially alters the efficiency considerations described up to now.

In order to calculate the optical DOS, dipole sources were placed in the nanowire array, with vertical (z-) and horizontal (xy-) polarization. As position, the cross-section of the nanowire was sampled at a height of $z_0 = 1.9$ μm, where the absorptivity is large. A wavelength sample of $\lambda = 700, 800$, and 900 nm was used, which is close to the spontaneous emission maximum in InP. A description similar to Equation 42.7 has been used in order to calculate the local optical density of states (LDOS) (for a detailed procedure, please see [19]). In Figure 42.7, the LDOS is plotted, normalized to the free-spaced DOS. It basically describes the enhancement or suppression factor for the spontaneous emission rate; a number larger than 1 would mean that spontaneous emission is enhanced. From the calculation, the z-polarized normalized LDOS reaches values around 2–3, close to the wire surface, where the maximum absorption takes place (see Figure 42.5 for comparison). This is a photonic crystal effect, as the z-polarized dipole emits predominantly in the horizontal direction. The xy-LDOS is 0.5–0.8, close to unity. As a consequence, the averaged LDOS over all directions leads to a normalized LDOS very close to unity, which according to Equation 42.27 allows the full benefit of the optical microconcentration effect to enter the open-circuit voltage calculation.

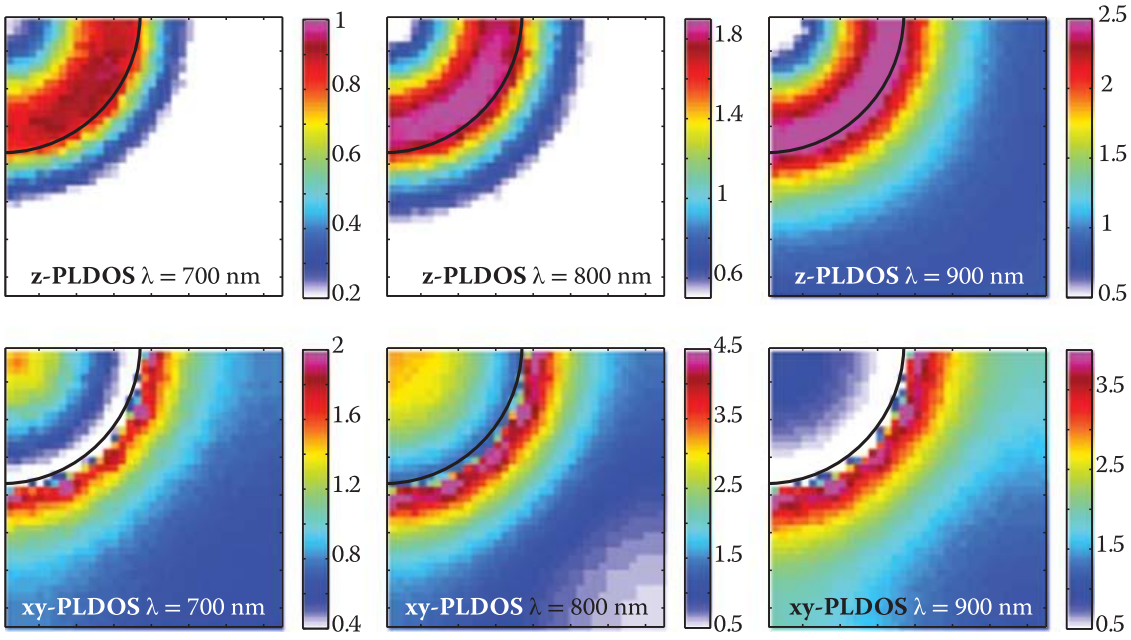


FIGURE 42.7 Local optical density of states (LDOS) relative to the bulk density of states in a nanowire array with diameter of $d = 180$ nm and pitch $a = 360$ nm. The top row is from an emitter polarized parallel to the wire axis, the bottom row for a dipole emitter with horizontal polarization. (Reprinted from J. Kupec et al., *Optics Express*, 18, 27589–27605, 2010. With permission)

42.5 Electronic Properties

42.5.1 Radial versus Axial Junction Arrangement

The nanowire geometry by nature has a high aspect ratio of length to diameter, and the p-n junction can be arranged in radial (i.e., vertical) or axial (i.e., horizontal) manner generating large or small quasi-neutral regions. Both implementations are shown in Figure 42.2 schematically and also have been demonstrated experimentally by epitaxial growth techniques. This p-n junction placement introduces another important aspect to nanostructured solar cell design. Assuming an abrupt junction, and the depletion approximation, the p-n junction can be separated into three distinct regions: the depleted region, and the quasi-neutral p- and n-regions. Photogenerated carriers in the depleted region are separated by the electric field and constitute a local photo current. In the quasi-neutral regions, the carriers diffuse, ideally to the contacts, if a local gradient is present. The cell current at the contacts is reduced if photogenerated carriers recombine on their way to the contact. This can be internal radiative or nonradiative recombination, consisting of spontaneous emission, Auger recombination, and Shockley–Read–Hall recombination. As a general guideline for minimizing the influence of nonradiative mechanisms, the recombination limited carrier lifetime needs to be much larger than the transit time of the carriers to the contacts. This principle is used in the thin-film solar cell technology.

In the analysis in Section 42.2.3, only the radiative recombination was included in the detailed balance limit. The analysis here is also done in the detailed balance limit as well, as it gives an upper boundary for the efficiency.

As example, the InP nanowire geometry is used, with a diameter of $d = 180$ nm and a wire height of 2 μm. The top n-region has a thickness of 100 nm, and the core diameter of the core-shell implementation is $d = 120$ nm. As p-doping, an active concentration of $p = 1{\times}10^{18} cm^{-3}$ and n-doping of $n = 5{\times}10^{18} cm^{-3}$ is used. The junction arrangement of the respective axial and core-shell structure is shown in Figure 42.2. The space and frequency resolved absorptivity profile has been calculated solving the vectorial Helmholtz equation under AM1.5D illumination; an illustration can be seen in Figure 42.8 after integration over frequency.

The wire array fill factor is 20%, which gives a maximum Shockley–Queisser efficiency. After spectral integration of the absorptivity, a local optical generation rate can be evaluated. The electronic simulation has been carried out only for a single wire by solving the continuity equations for holes and electrons coupled to the Poisson equation for the electrostatic potential. This means that any interaction between the

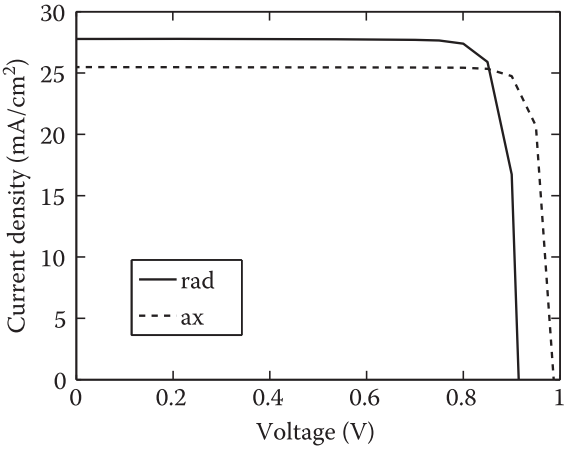


FIGURE 42.8 Current density versus voltage characteristics of axial and radial junction nanowire array solar cells made of InP, illuminated with the AM15.D solar spectrum. (ax, axial; rad, radial.)

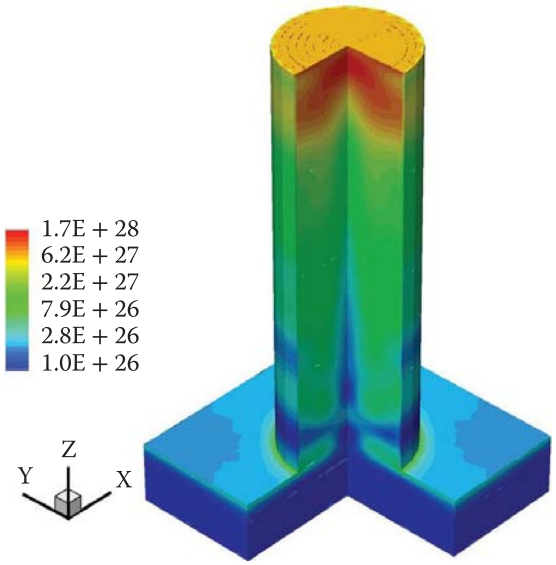


FIGURE 42.9 Optical generation rate profile in a nanowire array (only single nanowire shown) with a solar AM1.5D illumination spectrum (in units 1/s).

wires has been neglected for the carrier transport. Figure 42.9 shows the current versus voltage characteristics for the two designs. The axial p-n junction has a lower short-circuit current than the radial junction; however, it also shows a larger open-circuit voltage. The resulting maximum efficiencies are therefore close: 22% for the axial cell and 21% for the radial cell. In the axial wire, the photocurrent is generated from two effects. For one, the electric field in the depletion zone leads to drift current of the photogenerated carriers. This drift current removes carriers from the depletion zone, and in turn results in a diffusion current of the photogenerated carriers as excess minority carriers toward the junction. For the doping values in this wire, the depletion region is less than 15-nm thick at zero voltage, which would yield almost no photocurrent from direct drift. However, the simulations show that 600 nm of the quasi-neutral p-region contributes to the electron photocurrent by diffusion, and the entire quasi-neutral top n-region of 100 nm contributes to the hole photocurrent. The thickness of these diffusion regions depends critically on the minority carrier lifetime, and therefore impurity densities in the quasi-neutral regions need to be minimized. A more detailed analysis of this effect has been shown in [4]. In our simulation of the axial structure (with only radiative recombination present), a total of 700 nm of the wire height therefore can be viewed as the active region for photocurrent generation [29], which is 35% of the total length of the wire. In contrast, the radial junction contains a depletion region along the entire wire, and the charge separation due to drift and diffusion currents takes place in the entire wire volume. This leads to a higher photocurrent compared to the axial wire. The difference is not as large as the active volume ratio though, as the optical generation rate decreases exponentially from top to bottom, and the high generation top region is an active region for both structures. In [4], it has been found that the efficiency of a radial nanowire solar cell is much less prone to degradation if the minority carrier lifetime decreases, which can be directly explained by the increased depleted volume in the radial cell. It should be noted that the n-doped shell thickness of 30 nm for the radial wire has been chosen since it is larger than the depletion region so that a vertical photocurrent to the n-contact can be established.

The active region analysis explains the increased photocurrent for the radial structure, but it also can explain the reduction in open-circuit voltage compared to the axial device. The open-circuit voltage is given by Equation 42.20, with the fraction $\frac{I_{gen}}{I_{sat}}$ as contribution. Including the above described carrier transport effects, the volume integration for the calculation of the optical generation current I_{gen} and the reverse saturation current I_{sat} should only be done for the active volume. In the case of the axial junction, it is the depletion volume plus the diffusion zones in the quasi-neutral regions, which, in our specific case, account

to the top third of the wire. The radial junction wire collects photocarriers in the entire wire volume, and therefore the integration volume is the entire wire. As stated, this maximizes the optical generation current, and at the same time, also increases the reverse saturation current. Toward the bottom of the radial junction wire, the integration adds the same amount of reverse saturation current as in the upper region; however, the optical generation contribution decreases to a negligible amount. This is due to the exponential decay of the relative absorptivity a (see Equation 42.12). As result, the open-circuit voltage decreases if the wire is too long. This is plotted in Figure 42.10, where the wire array efficiency of a radial junction cell has been calculated for different wire lengths. For wire lengths below 2 μm, the solar spectrum is not fully absorbed within the wire (especially the long wavelengths), which limits the short-circuit current. For lengths above 2 μm, the reverse saturation current increases more strongly than the photocurrent, which decreases the open-circuit voltage. It should be noted that the values for the wire length and resulting efficiencies are specific to the material, dimensions, and doping concentrations that have been chosen for this example; however, the physical principles are valid for the general case of nanowire solar cells. For silicon nanowires, the optical absorption coefficient is smaller, and therefore the optimum wire length would be larger.

42.5.2 Surface Recombination

Up to now, all the calculations have been presented with radiative recombination only, excluding any nonradiative processes. In this section, the impact of surface recombination is analyzed. Nanostructures show a large surface to volume ratio, and any surface effect has a pronounced impact on the characteristic behavior of a device. For solar cells, surface recombination contributes to the nonradiative recombination mechanisms and decreases the cell efficiency, as discussed in the previous section. Here, the goal is to study the impact of surface recombination in the cell efficiency within the framework of the carrier continuity equation and the Poisson equation. The surface recombination rate is calculated via a Shockley–Read–Hall recombination with 2D trap densities [30]:

$$R_s = \frac{n_s p_s - n_i^2}{1/v_p(n_s + n_1) + 1/v_n(p_s + p_1)}, \tag{42.33}$$

where n_i is the intrinsic carrier concentration, v_n and v_p represent the surface recombination velocities, and n_1 and p_1 represent the surface trap densities. This model is used in the simulation procedure. The surface carrier densities are n_s and p_s, respectively.

The radial and axial structures have been simulated with constant surface recombination velocity of $v_s = 10^4$ cm/s, which is a realistic value for InP [31], and surface trap densities n_1 and p_1 set to a value for midgap states . Figure 42.11 shows the current–voltage characteristics with (dashed lines) and without surface recombination (solid lines). Turning on surface recombination, the maximum efficiency drops

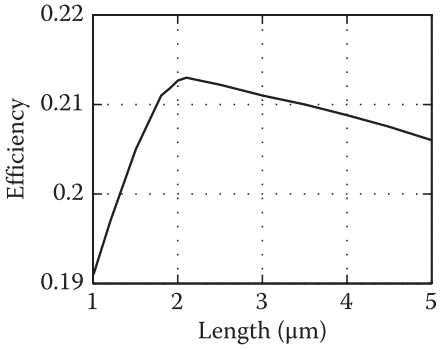


FIGURE 42.10 Cell efficiency of a radial junction nanowire array versus nanowire length.

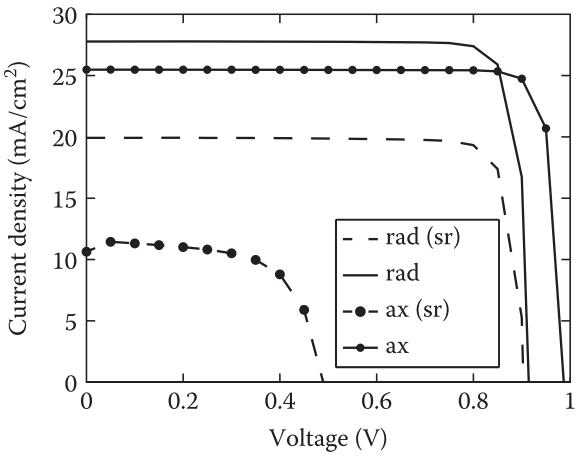


FIGURE 42.11 Current versus voltage relationship of a radial (rad) and axial (ax) junction nanowire array with and without surface recombination (sr).

from 21% to 15.5% for the radial structure and from 22% to 3.5% for the axial structure. It can be shown that the maximum surface recombination rate is achieved for the condition $v_n n_s = v_p p_s$ [32]. Assuming similar recombination velocities for electrons and holes for simplicity, the maximum surface recombination is hence located where electron and hole densities are both high at the same location. For a p-n junction, this occurs in the depletion region. It explains why the axial junction solar cell shows a substantial decrease in efficiency; in this geometry, the depletion region is exposed to the surface. This can be avoided by adding a passivation layer to the surface. For the radial junction, the depletion region thickness is smaller than the outer n-doped shell thickness, which makes the surface a quasi-neutral region. In that respect, the radial junction arrangement is vastly superior and much less prone to degradation from surface recombination effects.

42.6 Summary

Photovoltaic cells have become a fast growing technology for electrical power generation. Conversion efficiency and cost have been major drivers for the development. In that context, the use of nanostructures in solar cells has not lead to a major performance advantage up to now, and most implementations rely on classical p-n junction electronics and classical ray optics. In this chapter, the use of periodic nanowire arrays and their impact on solar cell efficiency has been analyzed. An efficiency analysis is presented that includes the nature of wave optics based on electromagnetic propagation and the optical DOS. The former leads to resonance and interference effects, which in turn causes localization of the electromagnetic energy and spectral dispersion. As main result, the efficiency of a bulk cell with ideal antireflective coating can be achieved with wires of only 20% of material without antireflective coating. In addition, efficient light management improves the solar cell efficiency via a modification of the optical DOS. This leads to a higher open-circuit voltage compared to a classical solar cell.

Acknowledgements

This work has been supported by the EU FP7 grant AMON-RA (grant agreement FP7-214814-1). The author acknowledges the contribution of J. Kupec, S. Yu on the numerical simulation results.

Appendix 42.A: Material Parameters

The InP material parameters used in the simulations are listed here. The refractive index dispersion is shown in Figure 42A.1.

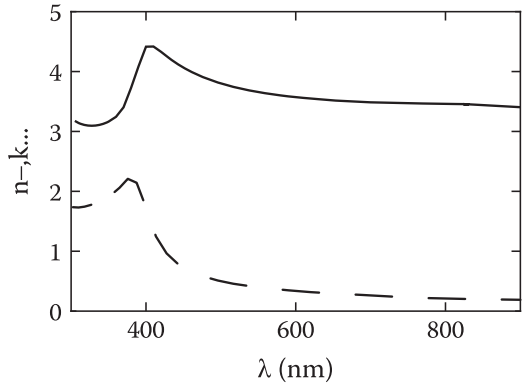


FIGURE 42A.1 InP refractive index n and extinction k versus wavelength (From S. L. Chuang, *Physics of Photonic Devices. Wiley Series in Pure and Applied Optics*, 2009. Copyright Wiley-VCH Verlag GmbH & Co. KGaA).

As mobility model for the electronic simulations, the Arora model has been used:

$$\mu = \mu_{\min} + \frac{\mu_D}{1 + (N_i/N_0)^A}, \tag{42A.1}$$

with N_i being the impurity density, and the remaining parameters have been chosen as [33].

	$\mu_{min}(m^2/V \cdot s)$	$\mu_D(m^2/V \cdot s)$	$N_0(m^{-3})$	A
el	0.040	0.520	3.00e23	0.47
hl	0.001	0.017	4.87e23	0.62

The intrinsic carrier density is $n_i = 1.2 \times 10\ m^{14}m^{-3}$, the bandgap energy is $E_G = 1.34$ eV, and the radiative recombination coefficient is $1.2 \times 10\ m^{-16}m^3/s$.

References

1. L. Tsakalakos, J. Balch, J. Fronheiser, M.-Y. Shih, S. F. Leboeuf, M. Pietrzykowski, P. J. Codella, B. A. Korevaar, O. Sulima, J. Rand, and A. Davuluru. Strong broadband optical absorption in silicon nanowire films. *Instrumentation*, 1:1–10, 2007.
2. E. Garnett and P. Yang. Light trapping in silicon nanowire solar cells. *Nano Letters*, 10:1082–1087, 2010.
3. J. Wallentin, N. Anttu, D. Asoli, M. Huffman, I. Berg, M. H. Magnusson, G. Siefer, P. Fuss-Kailuweit, F. Dimroth, B. Witzig-Mann, H. Q. Xu, L. Samuelson, K. Deppert, and M. T. Borgstrm. InP nanowire array solar cells achieving 13.8% exceeding the ray optics limit. *Science*, 339(6123):1057–1060, 2013.
4. B. Kayes, H. Atwater, and N. Lewis. Comparison of the device physics principles of planar and radial pn junction nanorod solar cells. *Journal of Applied Physics*, 97(11):114302, 2005.
5. C. Boecklin, R. G. Veprek, S. Steiger, and B. Witzigmann. Computational study of an InGaN/GaN nanocolumn light-emitting diode. *Physical Review B*, 81:155306, 2010.
6. G. Mariani, A. C. Scofield, C.-H. Hung, and D. L. Huffaker. GaAs nanopillar-array solar cells employing in situ surface passivation. *Nature Communications*, 4:1497, 2013.

7. W. Shockley and H. J. Queisser. Detailed balance limit of efficiency of pn junction solar cells. *Journal of Applied Physics*, 32(3):510–519, 1961.
8. E. M. Purcell, H. C Torrey, and R. V Pound. Resonance absorption by nuclear magnetic moments in a solid. *Physical Review*, 69(1–2):37–38, 1946.
9. Ivan S. Nikolaev, Willem L. Vos, and A. Femius Koenderink. Accurate calculation of the local density of optical states in inverse-opal photonic crystals. *Journal Optical Society of America B*, 26(5):987–996, 2009.
10. E. Yablonovitch. Statistical ray optics. *Journal of the Optical Society of America*, 72(7):899–907, 1982.
11. J. Nelson. *The Physics of Solar Cells*. London: Imperial College Press, 2003.
12. S. L. Chuang. *Physics of Photonic Devices. Wiley Series in Pure and Applied Optics*, 2nd edition. New York, NY: Wiley, 2009.
13. K. Barnham and G. Duggan. A new approach to high-efficiency multi-band-gap solar cells. *Journal of Applied Physics*, 67:3490–3493, 1990.
14. A. Luque and A. Marti. Increasing the efficiency of ideal solar cells by photon induced transitions at intermediate levels. *Physical Review Letters*, 78:5014–5017, 1997.
15. R. T. Ross and A. J. Nozik. Efficiency of hot-carrier solar energy converters. *Journal of Applied Physics*, 53(5):3813–3818, 1982.
16. J. Kupec and B. Witzigmann. Dispersion, wave propagation and efficiency analysis of nanowire solar cells. *Optics Express*, 17(12):10399–10410, 2009.
17. A. Taove and S. C. Hagness. *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd edition. Norwood, MA: Artech House, 2005.
18. B. Cockburn and C.-W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws v: Multidimensional systems. *Journal of Computational Physics*, 141:199–224, 1998.
19. F. Roemer, B. Witzigmann, O. Chinellato, and P. Arbenz. Investigation of the Purcell effect in photonic crystal cavities with a 3D finite element nanowire solar cells: Electro-optical performance Maxwell solver. *Optical and Quantum Electronics*, 39(4–6):341–352, 2007.
20. F. Roemer and B. Witzigmann. Spectral and spatial properties of the spontaneous emission enhancement in photonic crystal cavities. *Journal Optical Society of America B*, 25(1):31–39, 2008.
21. O. D. Miller, E. Yablonovitch, and S. R. Kurtz. Strong internal and external luminescence as solar cells approach the Shockley–Queisser limit. *IEEE Journal of Photovoltaics*, 2(3):303–311, 2012.
22. D. Knig, K. Casalenuovo, Y. Takeda, G. Conibeer, J. F. Guillemoles, R. Patterson, L. M. Huang, and M.A. Green. Hot carrier solar cells: Principles, materials and design. *Physica E: Low-Dimensional Systems and Nanostructures*, 42(10):2862–2866, 2010.
23. M. Luisier and G. Klimeck. Atomistic full-band simulations of silicon nanowire transistors: Effects of electron-phonon scattering. *Physical Review B*, 80:155430, 2009.
24. W. Frensley. Wigner-function model of resonant-tunneling semiconductor device. *Physical Review B*, 36(3):1570–1580, 1987.
25. U. Aeberhard and R. Morf. Microscopic nonequilibrium theory of quantum well solar cells. *Physical Review B*, 77(12):125343, 2008.
26. U. Aeberhard. Simulation of nanostructure-based and ultra-thin film solar cell devices beyond the classical picture. *Journal of Photonics for Energy*, 4(1):042099, 2014.
27. M. Auf der Maur, G. Penazzi, G. Romano, F. Sacconi, A. Pecchia, and A. Di Carlo. The multiscale paradigm in electronic device simulation. *IEEE Transactions on Electron Devices*, 58(5):1425–1432, 2011.
28. J. Kupec, L. R. Stoop, and B. Witzigmann. Light absorption and emission in nanowire array solar cells. *Optics Express*, 18(26):27589–27605, 2010.
29. S. Yu, J. Kupec, and B. Witzigmann. Efficiency analysis of III-V axial and core-shell nanowire solar cells. *Journal of Computational and Theoretical Nanoscience*, 9(5):1–8, 2012.
30. F. Roemer and B. Witzigmann. Modelling surface effects in nanowire optoelectronics. *Journal of Computational Electronics*, 11(4):431–439, 2012.

31. S. Bothra, S. Tyaqi, S. K. Ghandhi, and J. M. Borrego. Surface recombination velocity and lifetime in InP. *Solid-State Electronics*, 34:47–50, 1991.
32. P. P. Altermatt, A. G. Aberle, J. Zhao, A. Wang, and G. Heiser. A numerical model of p-n junctions bordering on surfaces. *Solar Energy Materials and Solar Cells*, 74:165–174, 2002.
33. M. Sotoodeh, A. H. Khalid, and A. A. Rezazadeh. Empirical low-field mobility model for III–V compounds applicable in device simulation codes. *Journal of Applied Physics*, 87(6), 2000.

43

Thin-Film Solar Cells

43.1 Introduction 497
43.2 Technology Overview 498
Silicon-Based Cells • CdTe-Based Cells • Copper Indium Gallium Selenide–Based Cells • Dye-Sensitized Solar Cells • Organic Solar Cells • Perovskites-Based Cells
43.3 Optical Modeling 503
43.4 Transport Modeling 511
Drift-Diffusion Model • Kinetic Monte Carlo • Example: Transport through 1D Disordered System by kMC
43.5 Specific Modeling Issues 527
Dye Solar Cells • Organic Photovoltaics • Perovskites • Modeling of Morphology and Grain Effects

Matthias Auf der Maur

Tim Albes

and

Alessio Gagliardi

43.1 Introduction

In this chapter, we review simulation models for thin-film solar cells, and especially the specific challenges encountered in modeling of thin-film cells based on emerging technologies. From a technological point of view, the numerous currently available photovoltaic (PV) technologies can be roughly subdivided into two classes. One is based on conventional (mono-)crystalline semiconductor technology, to which belong the first generation of solar cells such as monocrystalline silicon or gallium arsenide p-n junction cells, and the other is based on thin-film technology [1]. The term *thin-film* in this context does not primarily refer to the thickness of the devices, but to the type of production technologies used, although their physical thickness is typically small in the range of few tens of micrometer to well below 1 μm [19,90].

Thin-film technology is characterized by several advantages over conventional technology [4]. First of all, material deposition techniques are often more simple. Depending on the material system, there is a vast diversity in techniques. Many of these are low-temperature methods, when compared to standard crystalline semiconductor technology, and many of them are easily scalable to large-scale production and large device areas. In particular, solar cells based on organic polymers, e.g., can be produced from liquid solutions with processes similar to printing technology, which opens a way to very low-cost large-scale roll-to-roll production [5]. Overall, simpler and lower cost production methods and often cheaper, more abundant, or environmentally unproblematic materials are the main driving factors for development and optimization of thin-film PV concepts. Still, these technologies have to compete with costs of traditional PV, other renewable energy sources, and most of all with that of fossil energy sources in order to penetrate the energy market. In 2014, thin-film technologies accounted for roughly 9% of the global PV module production [6].

The "classical" thin-film solar cell technologies include the second generation cells like amorphous and microcrystalline silicon (a-Si, μc-Si), the chalcogenide alloy copper indium gallium diselenide (CIGS), and cadmium telluride (CdTe), which together count up for almost all of the current thin-film solar cell market.

Although the record efficiency map regularly published by the National Renewable Energy Laboratory (NREL) [7] groups only these technologies under the term "thin-film," we assume a broader interpretation by including and actually focusing on emerging concepts. These concepts are organic-based (bulk) heterojunction cells (organic photovoltaics [OPV]), dye solar cells (DSCs), and Perovskite-based cells. Recent reviews on the state of the art of these emerging technologies, also known as third-generation cells [8], can be found in [9–11].

Modeling of thin-film solar cells requires three essential ingredients, namely a model for the optical absorption, i.e., carrier or charge generation, for the charge separation, and for electronic transport. While the details of the latter depend mostly on the material system and technology of a given cell, the optical absorption is particularly critical for thin-film devices compared to devices employing thick absorbers. This is because they typically consist of several layers of different materials with thicknesses comparable to or smaller than the optical wavelength. Therefore, the optical absorption profile can be significantly determined by interference effects, so that the calculation of the optical absorption requires basically the explicit solution of the full Maxwell equations. Even if the optical absorption in some planar devices with thicker absorber layers, like DSCs, may be calculated with simplified approaches, the inclusion of light management–improving measures like scattering elements or back contact gratings usually requires a scheme based on the solution of Maxwell's equations, too.

Most of the models described in literature are using for the optical modeling constant (but material and wavelength dependent) bulk absorption coefficients, such that the optical absorption can be calculated once, providing the carrier generation profile for the electronic transport simulation. Analytical models for carrier transport can be found under suitable simplifying assumptions, as described e.g., in [1], and such models are very useful to understand basic features of solar cell behavior. Here, however, we are interested in physics-based simulation beyond analytical approximations, which provides more detailed insight into device operation and which, in particular, can be used for device optimization. Moreover, analytical models can typically be formulated only for one-dimensional (1D) problems, while numerical simulations can treat higher dimensions. This allows for simulation-based optimization of the device geometry, e.g., related to contacting schemes, which is hardly possible using only 1D simulations.

The chapter is organized as follows. First, we give a short overview on the basic features of the different thin-film technologies. Then we review the basic physics-based modeling approaches typically used for thin-film solar cells, both for optical absorption and electronic transport. Then we will describe critical details and modeling issues for the different third-generation thin-film concepts. More advanced concepts based on low-dimensional structures, hot carriers or using plasmonic effects, which generally require more complex simulation models, are not covered in this chapter.

43.2 Technology Overview

In the following sections, we shortly review the main characteristics of different thin-film technologies. For more in-depth discussions and further references, we refer, e.g., to [2,12].

43.2.1 Silicon-Based Cells

The currently most important silicon thin-film technologies are based on amorphous (a-Si) and microcrystalline (μc-Si) form. In contrast to crystalline silicon, which forms a regular tetrahedral configuration in a diamond lattice such that every silicon atom has four neighboring atoms, amorphous silicon does not show the same regular structure. The tetrahedral coordination of the atoms is mostly maintained, but random variations in bond angles and random appearance of dangling bonds, i.e., atoms having less than four neighbors, destroys the regular lattice and the associated periodicity.

The dangling bonds in amorphous silicon lead to a large number of defect states inside the bandgap, which adversely influence its electronic properties. In particular, the material can hardly be doped.

However, passivation of at least part of these dangling bonds by hydrogen atoms, known as *hydrogenation*, has been found to be necessary and sufficient to improve considerably material quality and allow it to be doped [13]. The hydrogenated form of a-Si is known as a-Si:H.

μc-Si (or more correctly μc-Si:H) is a more complex material, consisting of amorphous phases and conglomerates of silicon nanocrystals. Their microstructure depends largely on deposition methods and underlying layers, and the presence of grain boundaries between crystalline domains will influence further electronic properties.

As expected, the electronic properties of such disordered materials are worse than those of crystalline silicon. In particular, carrier mobilities are much lower. However, both materials are good enough to build working solar cells.

Both a-Si and μc-Si can be produced by a plasma-enhanced chemical vapor deposition (PECVD) technique introduced around 1970 [13,14]. The advantages of this technique over crystalline silicon deposition methods is that the required substrate temperature is significantly lower (<600 °C) and that it can be employed for large-area deposition. The former allows the use of cheap substrate materials such as plastic, while the latter allows to process large modules at once. Moreover, a-Si and μc-Si have two material intrinsic advantages. First, the loss of long-range order effectively leads to a transition from an indirect bandgap to a direct one, resulting in higher optical absorption coefficient. This allows to use thinner absorber layers without loss of absorption and thus to a cost reduction. Second, because of the different electronic transport properties these cells actually perform better at higher temperatures, contrary to crystalline silicon solar cells.

Typically, a-Si and μc-Si solar cells are built as p-i-n structures as shown in Figure 43.1a rather than pn-junctions, i.e., with an intrinsic layer between the n- and p-doped contact layers, since the carrier diffusion length in doped material are rather short due to the relatively poor transport properties. The transparent conductive oxide (TCO) layer in Figure 43.1a is e.g., indium tin oxide (ITO) or SnO_x. This layer is common to all thin-film solar cell technologies and serves to achieve a sufficiently low front contact resistivity without optical absorption. Cell thickness is usually rather small (around 400 nm for a-Si, few μm for μc-Si) in order to reduce recombination losses. Current best research cell efficiency is around 10% for a-Si:H and 12% for μc-Si single junction cells, compared to 25% of single-crystal silicon cells [15].

One of the most important problems in hydrogenated silicon cells at current is a light-induced performance degradation by prolonged illumination, known as the Wronski–Staebler effect, which is believed to be related to light-induced diffusion of the hydrogen atoms [16].

A recent review on thin-film Si solar technology can be found, e.g., in [17].

43.2.2 CdTe-Based Cells

CdTe is a very suitable material for PV applications. It has a direct bandgap of $E_g \approx 1.45$ eV, which is in the optimum range of values, and it has a high optical absorption coefficient $>10^4$ cm^{-1} over the whole solar spectrum, allowing for thin absorbers.

Both monocrystalline and polycrystalline CdTe cells can be produced. Polycrystalline CdTe can be deposited by a vast range of techniques, including vacuum techniques such as sublimation, evaporation, atomic layer deposition, chemical vapor deposition, but also nonvacuum techniques based on solution processes. Low-temperature processes are especially interesting since they allow for deposition on cheap and flexible substrates and for large-scale roll-to-roll process technology. Deposition is followed by a $CdCl_2$ treatment to improve material quality.

Some disadvantages of CdTe are the inability to tune the bandgap by alloying, and the presence of native defects to an amount that makes it difficult to achieve high-doping levels. The latter is amended by smart contacting schemes. At the front contact, an n-doped CdS buffer layer is introduced before the p-type CdTe absorber layer, which also acts as a buffer layer for CdTe deposition. The back contact is challenging because of the impossibility to achieve high p-doping density and since CdTe has a high electron affinity leading to high Schottky barriers for usual contact metals. Several solutions exist, e.g., suitable combinations of

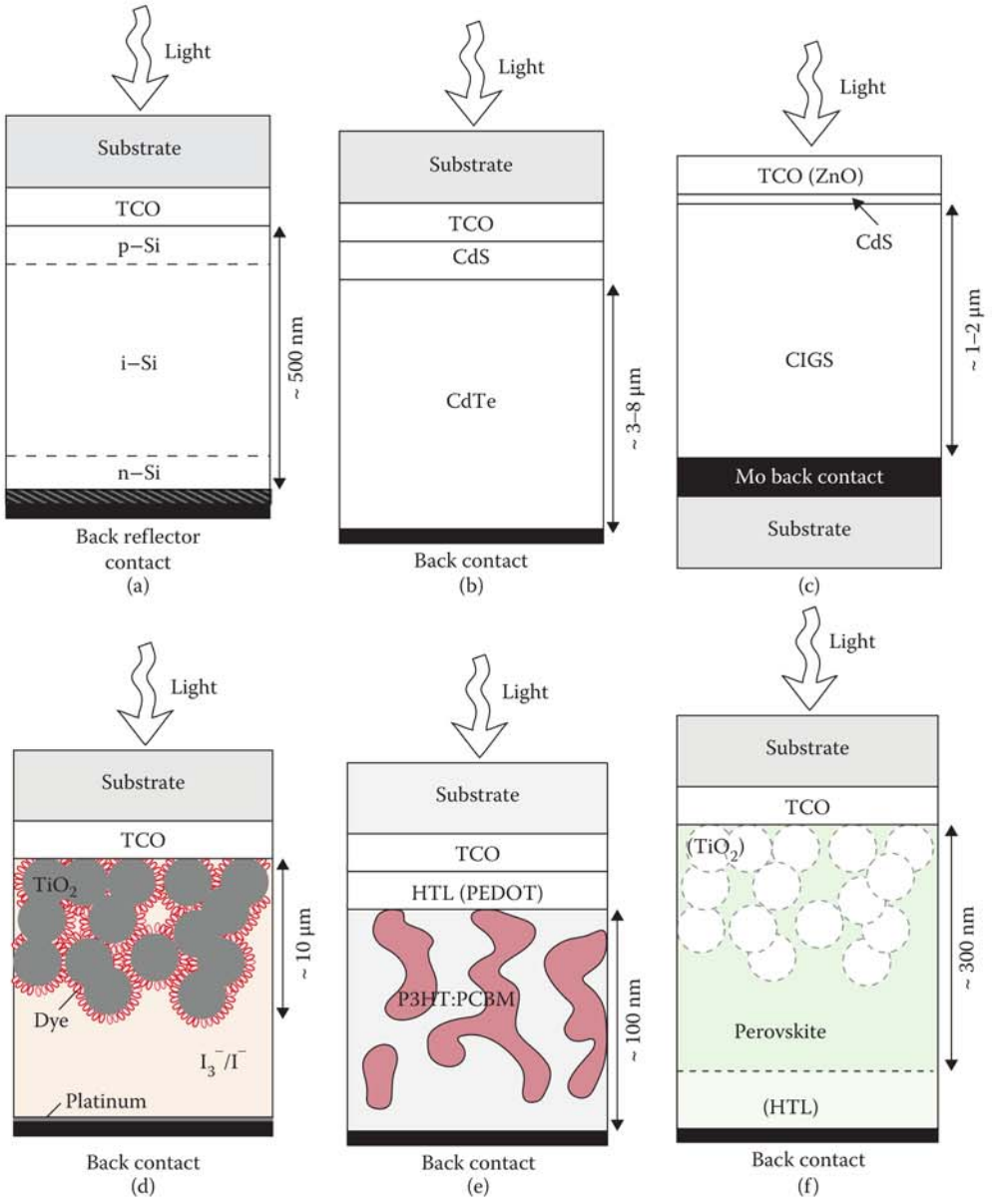


FIGURE 43.1 Typical schematic structures of different thin-film solar cells. (a) Amorphous silicon, with intrinsic absorber layer (i-Si) and doped contact layers (n/p-Si); (b) cadmium telluride (CdTe); (c) copper indium gallium selenide (CIGS); (d) dye solar cell (DSCs): TiO_2 nanoparticles are in the order of 10 nm diameter; (e) organic solar cell: the characteristic dimension of the absorber blend is in the order of 10 nm; and (f) perovskite cell, in some versions including a mesoporous TiO_2 scaffold as used in DSCs. Typical absorber thicknesses are indicated in the figures. TCO is the transparent conducting oxide, like indium tin oxide (ITO) or zinc oxide (ZnO), HTL is the hole transport layer.

buffer layers with metallization or particular surface treatments. Transport and recombination properties in CdTe films depend much on deposition technology and on the properties of the polycrystallinity of the film. The best research cell efficiency reported so far is 21% [15]. A typical CdTe solar cell scheme is shown in Figure 43.1b.

CdTe thin-film solar cells have been commercialized successfully and in 2014 they covered roughly 44% of the market share of all thin-film technologies [6]. It has to be noted that the presence of cadmium in CdTe has induced some doubts about the environmental compatibility of this technology. Recent reviews can be found in [18,19].

43.2.3 Copper Indium Gallium Selenide–Based Cells

Chalcopyrite-based solar cells is the second most important thin-film technology nowadays with a market share of 39% [6]. The most relevant compounds of this material class for PV applications are $CuInSe_2$, $CuInS_2$, and $CuGaSe_2$ having a bandgap of 1.0, 1.5, and 1.7 eV, respectively. These compounds can be easily alloyed, providing an extended range of bandgaps and lattice constants. Typically, the absorber consists of a $Cu(In,Ga)(S,Se)_2$ alloy, giving the name "CIGS," or of $CInS_2$ which has a gap of 1.5 eV. The highest cell efficiency reported in 2015 was around 21%, higher than multicrystalline Si cells [15]. Beside significantly higher efficiency than a-Si thin-film cells, CIGS cells show better stability and have no light-induced degradation.

A typical CIGS cell is shown in Figure 43.1c [20]. The nontransparent back contact is typically molybdenum sputtered onto a suitable substrate. Then follows a polycrystalline p-type CIGS absorber layer of few micrometer thickness. A heterointerface with a very thin n-type CdS is subsequently formed, followed by a TCO-like heavily doped ZnO as front contact and window layer.

Like CdTe and a-Si, CIGS cells can be produced on flexible substrates, allowing for specialized application fields. Reviews can be found, e.g., in [21,22].

43.2.4 Dye-Sensitized Solar Cells

A dye-sensitized solar cell (DSC) is an electrochemical system where the absorption and charge transport components are spatially separated [23]. The standard device is made of a mesoporous oxide, usually titanium dioxide (TiO_2), composed of nanoparticles of 15–20 nm in diameter. The surface of the mesoporous oxide is covered by a monolayer of molecules, the dye, conventionally organometallic complexes like the N719 [24] based on ruthenium. Finally, the oxide is sourrounded by a liquid electrolyte containing a redox pair. The most common one is triiodide/iodide pair (I_3^-/I^-), but many other redox pairs have been tried. In particular, cobalt-based redox pairs [25] have shown interesting properties. The contacts also are made of a glass with a transparent conducting oxide layer like ITO or fluorine-doped tin oxide (FTO) on the TiO_2 side, the anode, while the cathode is a thin layer of platinum deposited on a metallic substrate. In the device, the dye is the active material which absorbs light, while the redox pair in the electrolyte and the TiO_2 are the hole and electron transporting materials, respectively. The mesoporous structure of the TiO_2 is necessary in order to increase the surface for the dye and thus enhance dramatically the absorption efficiency, since necessarily only a monolayer of dye molecules can be used in order to efficiently extract the electrons and regerenate the dye.

The photoconversion in a DSC is far more complex than in conventional PV devices. After the absorption of the photon by the dye, the electron is transferred to the TiO_2 by an extremely fast charge transfer. Then the electron percolates in the mesoporous TiO_2 until it is collected by the ITO or FTO at the anode. The oxidized dye is regenerated by an electrochemical process with the redox pair. In a standard DSC this means that iodide is consumed and triiodide is generated nearby of the dye. The process generates an excess of triiodide with respect to the equilibrium concentration of the I_3^-/I^- couple in the bulk electrolyte. This concentration gradient moves triiodide toward the cathode where it is transformed back into iodide by taking two electrons from the catalyst, the platinum, closing the circuit. The platinum takes electrons directly from the external circuit.

There are several advantages of a DSC: it is a thin-film device, with a thickness in the range of tens of micrometer; it is quite cheap as the most expensive components (platinum and dye) are used in tiny

amounts (few nanometer in thickness and a single molecular layer, respectively); the decoupling of the absorption and transport components strongly reduces recombination processes as electron and hole pairs are immediately separated on different materials at the very beginning of the photoconversion process; and fabrication processes are very cheap.

The main drawback of the device is the liquid component that can easily leak out from microcracks in the encapsulation material or be polluted by the external environment. Moreover, the most common redox pair, triiodide/iodide, is highly corrosive leading to problems in the metals used to make the external contacts and degradation of the dye layer due to desorption of molecules from the surface. All these effects are obviously detrimental for the life span of the device. In order to correct these problems, an alternative DSC has been proposed [26], where the liquid electrolyte for hole transport is substituted by an organic semiconductor (e.g., Spiro-OMeTAD or P3HT [27]). The advantage is that the new hole transporter is more stable and noncorrosive; however, the charge transport in these organic semiconductors is worse leading to higher recombinations at the interface. Moreover, the intermixing between the organic material and the mesoporous TiO_2 is less efficient. DSCs have reached 12% of confirmed efficiency [15]. Recent reviews can be found in [28,29].

43.2.5 Organic Solar Cells

OPV cells are classified as organic because their active layer consists entirely of carbon-based materials, such as polymers or small organic molecules. A major advantage of OPVs is their compatibility with techniques suitable for large-scale and low-cost fabrication, such as roll-to-roll processing. The ability to chemically engineer organic materials allows to create a variety of compounds suitable for OPV applications. Especially due to the high absorption coefficient of specific polymers, very thin layers of 50–300 nm provide sufficient absorption and make OPVs a favorable choice for thin-film solar cells. Additional features, such as mechanical flexibility and low weight, open up the possibility for a variety of applications.

A peculiarity of the organic materials used for OPVs is their low permittivity ($\varepsilon_r \approx 3-4$), which leads to a weak screening of electrostatic interactions. On absorption of a photon, an electron–hole pair is created. Due to the strong Coulomb interaction between the electron and the hole, the pair forms an exciton with binding energies of several hundred millielectron volts (meV). The thermal energy of 25 meV at room temperature is not large enough to split the exciton and separate the charges. Therefore, the concept of a heterojunction composed of a (electron) donor, usually a polymer or small molecule, and an (electron) acceptor was introduced. The most prominent acceptor material is a derivative of the Buckminsterfullerene, [6,6]-phenyl-C61 butyric acid methyl ester (PCBM), because its large electronegativity has shown major improvement in exciton dissociation [30].

The first organic solar cell with efficiency of >1% was developed by Tang [31] and consisted of a planar heterojunction between donor and acceptor materials. Planar architectures have the disadvantage of a low exciton splitting efficiency, since excitons typically have a very small diffusion length and thus only excitons created nearby the junction can diffuse to the interface and become separated before recombination. A boost in device performance was achieved by introducing the concept of a bulk heterojunction (BHJ) [32], which is an intermixed morphology between donor and acceptor materials, as indicated schematically in Figure 43.1c. It provides a trade-off between efficient exciton splitting and good charge transport. The spatial structure of the BHJ is crucial for device performance and is still object of intense study.

The pioneering organic devices nowadays are based on the polymer/fullerene bulk-heterojunction concept with efficiencies of up to 10.8% [15,33]. Although much progress has been achieved in the last decade, higher efficiencies are needed in order for organic solar cells to be commercially competitive with established (inorganic) technologies. Since the experimental investigation of OPVs during operation is difficult, simulations can give valuable insights into the still not fully understood charge generation and transport processes.

More details and reviews on OPV can be found, e.g., in [9,34,35].

43.2.6 Perovskites-Based Cells

Since the first attempts in 2009 and a breakthrough in 2012 regarding efficiency and stability [36], the organometal lead halide perovskite $CH_3NH_3PbX_3$ (with X = Cl, Br, or I) has attracted great attention as absorber material in PV devices. Originally, perovskites have been used as alternative dyes in DSCs, but then perovskite-based solar cells started forming a PV technology on their own. Structurally, these perovskites consist of an ionic cage built from PbX in a tetragonal structure, with an organic methylammonium (CH_3NH_3) molecule at the center [37].

One of the most astonishing properties of perovskite-based solar cells is the fact that perovskite layers can be very easily deposited with simple solution processes, and at the same time easily reaching high photoconversion efficiencies. In fact, recent perovskite cells show efficiencies comparable to the best CdTe and CIGS cells, and thus much higher than all other emerging PV concepts [15]. Such a good performance is due to good absorption properties and good, balanced electron and hole transport properties.

A further useful property of methylammonium lead halide perovskites ($MAPbX_3$) is the ability to tune the bandgap by alloying $MAPbI_3$ with $MAPbBr_3$, which allows to extend the absorption spectra up to 800 nm.

Since perovskite solar cells originally emerged from the DSC technology, their structure often resembles a DSC. Usually, a perovskite layer of few hundred nanometer thickness is deposited on a TiO_2 layer, with or without mesoporous phase, followed by a hole transport layer (HTL), as shown in Figure 43.1f. However, the latter may even be omitted [38].

Substantially, perovskite technology is "dirty," cheap but performing at the same time, reaching efficiencies as high as 20% at cell level [15]. However, it struggles with several open issues. Most importantly, stability is still very poor [39], which is mostly due to the sensitivity of the $MAPbI_3$ to humidity [40]. Second, often the measured current–voltage characteristics show pronounced hysteresis, which is nowadays believed to be due to migration of charged defects, especially lead vacancies, in the perovskite. This poses difficulties for the accurate experimental evaluation of cell efficiencies. Finally, toxicity of lead halide perovskites may be an issue. In fact, lead compounds are toxic, and therefore often banned by law and certainly unpopular in industry. Identification of lead-free perovskites, however, did not yet lead to satisfactory candidates. It seems, that lead-based halides provide best performance.

Recent reviews on perovskite-based solar cells can be found in [37,38,40,41].

43.3 Optical Modeling

The calculation of the optical field distribution in a solar cell is a critical modeling step for any solar cell, as it provides the carrier generation profile, which is the most important factor determining the photocurrent. The generation of carriers is related to the optical absorption of a material for a given wavelength λ and can be written as [42,43]

$$G(\lambda; x) = \alpha(\lambda; x) I(\lambda; x) \tag{43.1}$$

where G is the carrier generation rate per wavelength interval at a given incident free-space wavelength λ at position x, usually given in units of $cm^{-3}s^{-1}nm^{-1}$; α is the optical absorption coefficient at wavelength λ, with units cm^{-1}; and $I(\lambda; x)$ is the local light intensity given as photon flux in $cm^{-2}s^{-1}nm^{-1}$. In Equation 43.1, we assume that every absorbed photon produces a carrier, which may be an unbound electron–hole pair or an exciton. The total generation profile is then obtained by summing the contributions of all relevant wavelengths present in the incident light. The knowledge of the local absorption requires the knowledge of the local wavelength-dependent absorption coefficient, which is usually taken as a material constant, and of the local intensity spectrum. The latter is related to the local strength of the

electromagnetic field, which can be written as [44]

$$I(\lambda; x) = \frac{\lambda \varepsilon_0 n}{2h} |E(\lambda; x)|^2 \tag{43.2}$$

where ϵ_0 is the vacuum permittivity, n the refractive index, h Planck's constant, E the amplitute of the electric field, and λ the free space wavelength. Therefore, the knowledge of the electromagnetic field distribution in a solar cell device allows to calculate the local generation rate G.

While the local electromagnetic field is accessible only via simulation, it is possible to both calculate and measure the absorption spectrum α. The usual way to calculate α is by using Fermi's golden rule and evaluating the optical transition matrix elements between the involved electronic states [1,45]. In this approximation, the transition probability between two states i and f with different energies E_i and E_f is given by

$$P_{i \to f} = \frac{2\pi}{\hbar} |\langle i|H'|f\rangle|^2 \delta(E_f - E_i \pm \hbar\omega), \tag{43.3}$$

where H' is the Hamiltonian describing the light-matter interaction, treated as a perturbation, and $\hbar\omega$ is the photon energy. $\langle i|H'|f\rangle$ in Equation 43.2 is the matrix element coupling initial and final states i and f, and its calculation requires an explicit representation of these states in some suitable basis. The delta function is an expression of the conservation of energy, and the minus sign describes photon absorption, while a positive sign describes photon emission, i.e., radiative recombination. The net rate of photogeneration is then given by the product of transition probability with the number of occupied initial and available final states, and by subtracting emission from absorption. The occupation of the states is given by the carriers' Fermi–Dirac distribution functions f_i and f_f. Dividing the net rate by the incident photon flux leads to the absorption coefficient [45]

$$\alpha(\hbar\omega) = \frac{\pi e^2}{n_r c \varepsilon_0 m_0^2 \omega V} \sum_{\substack{\mathbf{k}_i, \mathbf{k}_f, \\ \sigma_i, \sigma_f}} |\hat{e} \cdot \mathbf{p}_{fi}|^2 \delta(E_f - E_i - \hbar\omega)(f_i - f_f) \tag{43.4}$$

Here, e is the elementary charge, n_r is the relative dielectric constant, ε_0 is the permittivity, m_0 is the electron mass, V is the normalization volume, and the optical matrix element has been substituted by the momentum matrix element applying the dipole approximation [45]. The summation has to be done over the crystal momentum $\mathbf{k}$ and all quantum numbers σ of both initial and final states. For absorption and at not too high optical excitation, one can usually assume all initial states as occupied and all final states as empty, i.e., $f_i \approx 1$ and $f_f \approx 0$.

Radiative recombination is obtained in a similar way; however, in that case only transitions between occupied electron and hole states are relevant. Therefore, the summation can be restricted to states with energies near the band extrema. For absorption, however, all transitions are relevant, and therefore the electronic states have to be known with sufficient accuracy for all energy bands and in the whole reciprocal k-space (or more apropriately in the full first Brillouin zone). This makes calculation of absorption spectra more difficult, so that usually measured spectra are used.

In principle, the local electric field strength at all relevant wavelengths can be obtained by solving Maxwell's equations using one of the standard methods such as finite difference time domain (FDTD) [46], finite elements (FEM) [47], or rigorous coupled-wave analysis (RCWA) [48], but this is not always necessary. For a homogeneous thick absorber, in the absence of coherent effects, the Beer–Lambert approximation can be used [1]. It allows to calculate the local optical intensity simply considering the absorption along the optical path, starting from Equation 43.1. Assume a monochromatic light beam with wavelength λ with intensity I traveling in a material with absorption coefficient α. When it travels an infinitesimal

distance dx, the intensity of the beam reduces by $\alpha I \mathrm{d}x$, since a part of the photons are absorbed. We can therefore deduce a continuity law for the light intensity given by

$$\frac{\mathrm{d}I}{\mathrm{d}x} = -\alpha I \tag{43.5}$$

The solution to this is an exponential function, describing exponential decay of the intensity known as Beer–Lambert law, $I(x) = I_0 \mathrm{e}^{-\alpha x}$, where I_0 is the intensity of the incident beam (after subtraction of reflection losses, i.e., immediately inside the device) and x the distance from the point of incidence. In the case of spatially inhomogeneous absorption, the result is easily generalized to

$$I(\lambda; x) = I_0(\lambda) \mathrm{e}^{-\int_0^x \alpha(\lambda; x)\, \mathrm{d}x} \tag{43.6}$$

Equation 43.6 together with Equation 43.1 are used succesfully for the simulation of solar cells where layer thicknesses are larger than the optical wavelength. However, since this model is based on a ray optics description of light, it will break down whenever effects like interference or diffraction are important. In that case a more rigorous approach is needed to calculate the optical intensity profile.

Under the assumption of a flat, planar layered structure, a more precise model can be formulated, without increasing much the computational efforts. This model is based on a 1D approximation along the optical path of a plane wave, including sequential reflections at material interfaces and transmissions in homogeneous portions of the device, and is known as the *Transfer Matrix Method* (TMM) [49,50]. Today, this is probably the most used model for one-dimensional optical modeling of thin-film devices. The model has been first applied to organic thin-film solar cells in the late 1990s [51–53], and since then it has become a standard tool [54–60].

The basic principle of the TMM is shown schematically in Figure 43.2. To apply the method, we assume that the device can be sliced into layers of homogeneous properties with flat, parallel interfaces. We do not include a possible thick substrate that should be treated with an incoherent method (details on this can be found, e.g., in [44]). Further, we assume a monochromatic optical field written as $E(x)\exp(i\omega t)$, where $E(x)$ is a complex local field amplitude, where x is the coordinate along the stack, with $x = 0$ being the point of incidence. For simplicity, we assume a transverse electromagnetic (TEM) field with orthogonal incidence here, and in the following we drop the harmonic time dependence.

The amplitudes of the electric and magnetic field at each position can be found by using two relations derived from Maxwell's equations. The first relates the fields at the left interface of a layer to that at the right interface by means of the free propagation in a homogeneous medium, while the second relates the fields at the left and the right of each interface by means of continuity conditions. Maxwell's equations in

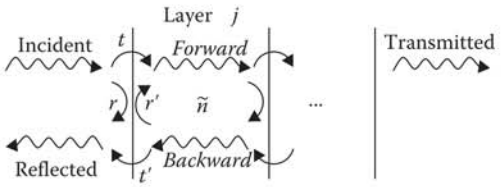


FIGURE 43.2 Scheme of the transfer matrix method. The device is assumed to be sliced in homogeneous layers, and monochromatic light is incident from the left. We decompose the optical field in a forward and backward traveling component. At each interface between two layers, these two components undergo reflection and transmission, r and t, while inside the layer the waves propagate freely with propagation constant according to the complex refractive index $\tilde{n}$.

homogeneous form (without source terms) are given by [49,61]

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \cdot \mathbf{D} = 0 \tag{43.7}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}, \quad \nabla \cdot \mathbf{B} = 0 \tag{43.8}$$

$$\mathbf{B} = \mu \mathbf{H}, \quad \mathbf{D} = \varepsilon \mathbf{E} \tag{43.9}$$

Here, **E**, **H**, **D**, and **B** are the electric field, the magnetic field (or magnetizing field, magnetic field strength), the electric displacement, and the magnetic flux density (or magnetic induction, magnetic field), respectively. The two electric fields and magnetic fields are related by the constitutive relations (Equations 43.7 through 43.9), where ε and μ are the electric permittivity and the magnetic permeability, respectively. The latter is usually equal to the vacuum permeability μ_0 for the materials of interest in this context. Since we are interested in time harmonic fields, Equations 43.7 through 43.9 are conveniently transformed to Fourier space, leading to

$$\nabla \times \mathbf{E} = -i\omega \mathbf{B} \tag{43.10}$$

$$\nabla \times \mathbf{H} = i\omega \mathbf{D} \tag{43.11}$$

while the other expressions remain the same. Note that the fields are now complex quantities. Moreover, the permittivity, which we write as product of vacuum permittivity ε_0 and relative permittivity $\tilde{\varepsilon}$, is frequency dependent (dispersive) and complex in case of lossy, i.e., absorbing materials.

In a homogeneous medium, the solution of Equations 43.7 through 43.9, or equivalently Equations 43.10 and 43.11, is found to be a propagating wave $E(x) = E_0 \exp(\pm ikx)$, where k is the propagation constant (in our case the component along x). The latter is related to the relative permittivity $\tilde{\varepsilon}$, or, equivalently, to the refractive index $\tilde{n} = \sqrt{\tilde{\varepsilon}}$ by [43]

$$k = \frac{2\pi}{\lambda} = \frac{\omega}{c} = \frac{\omega\sqrt{\tilde{\varepsilon}}}{c_0} = \frac{2\pi}{\lambda_0}\tilde{n} \tag{43.12}$$

where c_0 and λ_0 are the vacuum light speed and the free space wavelength, respectively; $\omega = 2\pi f$ is the angular frequency, and we have used the expression for the light speed $c = 1/\sqrt{\varepsilon\mu}$. Note that the refractive index $\tilde{n} = n_0 - i\kappa$ is a complex quantity, where the imaginary part is related to absorption.

It is convenient to decompose the electric field into *forward* and *backward* travelling components E^+ and E^-, as indicated in the figure. This is written as

$$E(x) = E^+(x) + E^-(x) \tag{43.13}$$

where once again all amplitudes are complex quantities.

We can now relate the field amplitudes at the left and right interface of the layer j by

$$E^+_{j;\mathrm{R}} = E^+_{j;\mathrm{L}} \mathrm{e}^{-ik_j t_j} \tag{43.14}$$

$$E^-_{j;\mathrm{R}} = E^-_{j;\mathrm{L}} \mathrm{e}^{ik_j t_j} \tag{43.15}$$

where k_j and t_j are the propagation constant and thickness of layer j, respectively. This can be put into matrix form as

$$\begin{pmatrix} E^+ \\ E^- \end{pmatrix}_{j;\mathrm{R}} = \mathbf{M}_j \begin{pmatrix} E^+ \\ E^- \end{pmatrix}_{j;\mathrm{L}}, \quad \mathbf{M}_j = \begin{pmatrix} \mathrm{e}^{-ik_j t_j} & 0 \\ 0 & \mathrm{e}^{ik_j t_j} \end{pmatrix} \tag{43.16}$$

where we have defined the transmission matrix of layer j as $\mathbf{M}_j$. Comparing Equation 43.16 with Equation 43.6 and considering Equation 43.2, the relation between $\mathrm{Im}(k)$, commonly called *extinction coefficient*, and the absorption coefficient α can be immediately found as

$$\alpha = \frac{4\pi}{\lambda}\kappa \tag{43.17}$$

Based on Maxwell's equations, one can derive the boundary conditions for the in-plane and orthogonal components of the electric and magnetic fields at planar material interfaces (orthogonal to the x-axis) as shown in Figure 43.3, which are well known and given by

$$\mathbf{n}_{lm} \times (\mathbf{E}_m - \mathbf{E}_l) = 0 \tag{43.18}$$

$$\mathbf{n}_{lm} \times (\mathbf{H}_m - \mathbf{H}_l) = 0 \tag{43.19}$$

i.e., the tangential components of **E** and **H** need to be continuous. In the above equations, $\mathbf{n}_{lm}$ denotes the normal on the interface between layers l and m. The orthogonal components of **D** and **B** have to be continuous, which is not relevant in our case of normal incidence.

Furthermore, using Equation 43.10 we can express the magnetic field in terms of the electric field as

$$B^{\pm} = \pm\frac{\tilde{n}}{c_0}E^{\pm} \tag{43.20}$$

which is obtained by evaluating the curl of the electric field and using Equation 43.12. With this we can rewrite Equations 43.18 and 43.19 in terms of the forward and backward traveling electric field components as

$$E_l^+ + E_l^- = E_m^+ + E_m^- \tag{43.21}$$

$$\tilde{n}_l(E_l^+ - E_l^-) = \tilde{n}_m(E_m^+ - E_m^-) \tag{43.22}$$

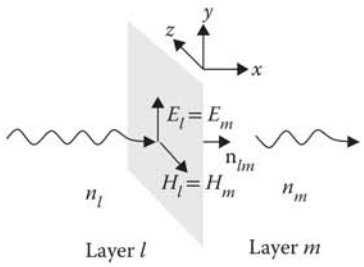


FIGURE 43.3 A material interface and the respective boundary conditions for the in-plane field components.

which can be rearranged such as to provide the transfer matrix of the material interface

$$\begin{pmatrix} E^+ \\ E^- \end{pmatrix}_{l,\mathrm{R}} = \underbrace{\frac{\tilde{n}_l + \tilde{n}_m}{2\tilde{n}_l} \begin{pmatrix} 1 & \frac{\tilde{n}_l - \tilde{n}_m}{\tilde{n}_l + \tilde{n}_m} \\ \frac{\tilde{n}_l - \tilde{n}_m}{\tilde{n}_l + \tilde{n}_m} & 1 \end{pmatrix}}_{\mathbf{I}_{lm}} \begin{pmatrix} E^+ \\ E^- \end{pmatrix}_{m,\mathrm{L}} \tag{43.23}$$

In the expressions for the elements of the transfer matrix $\mathbf{I}_{lm}$, we can identify the reflection and transmission coefficients of the interface between adjacent layers l and $m = l + 1$ with respect to the amplitudes

$$r = r_{lm} = \frac{E^-_{l;\mathrm{R}}}{E^+_{l;\mathrm{R}}} = \frac{n_l - n_m}{n_l + n_m} \tag{43.24}$$

$$t = t_{lm} = \frac{E^+_{m;\mathrm{L}}}{E^+_{l;\mathrm{R}}} = \frac{2n_l}{n_l + n_m} = 1 + r_{lm} \tag{43.25}$$

$$r' = r_{ml} = \frac{E^+_{m;\mathrm{L}}}{E^-_{m;\mathrm{L}}} = -r_{lm} \tag{43.26}$$

$$t' = t_{ml} = \frac{E^-_{l;\mathrm{R}}}{E^-_{m;\mathrm{L}}} = \frac{1 - r^2_{lm}}{t_{lm}} \tag{43.27}$$

so that we can write

$$\mathbf{I}_{lm} = \frac{1}{t_{lm}} \begin{pmatrix} 1 & r_{lm} \\ r_{lm} & 1 \end{pmatrix} \tag{43.28}$$

Combining Equations 43.16 and 43.28 we obtain the transfer matrix for the fields at the left interface between layers l and $m = l + 1$ as

$$\begin{pmatrix} E^+ \\ E^- \end{pmatrix}_l = \mathbf{M}_l \mathbf{I}_{lm} \begin{pmatrix} E^+ \\ E^- \end{pmatrix}_m \tag{43.29}$$

and finally the transfer matrix $\mathbf{T}$ relating the fields to the left and to the right of stack as a product

$$\mathbf{T} = \prod_l \mathbf{M}_l \mathbf{I}_{l,l+1} \tag{43.30}$$

Note that TMM can be implemented in several ways, and that it is inherently numerically unstable for larger layer thicknesses [62]. Stabilization is possible, e.g., by resorting to alternative representations of the transfer matrix [63].

As an example for the importance of interference effects in layered structures, we present a comparison of the local field strength and generation rate between results of a TMM calculation and those of an incoherent calcuation. For the latter, we considered the reflectance and transmittance at each material interface and summed up the intensities of the forward and backward traveling waves. The test structure, similar to the one used in [64], is shown in Figure 43.4a. For simplicity, we assumed total reflection at the aluminum back contact, and we did not consider the glass substrate. The results have been obtained for a wavelength of 500 nm. Figure 43.4b and c presents the field and generation profiles for two different

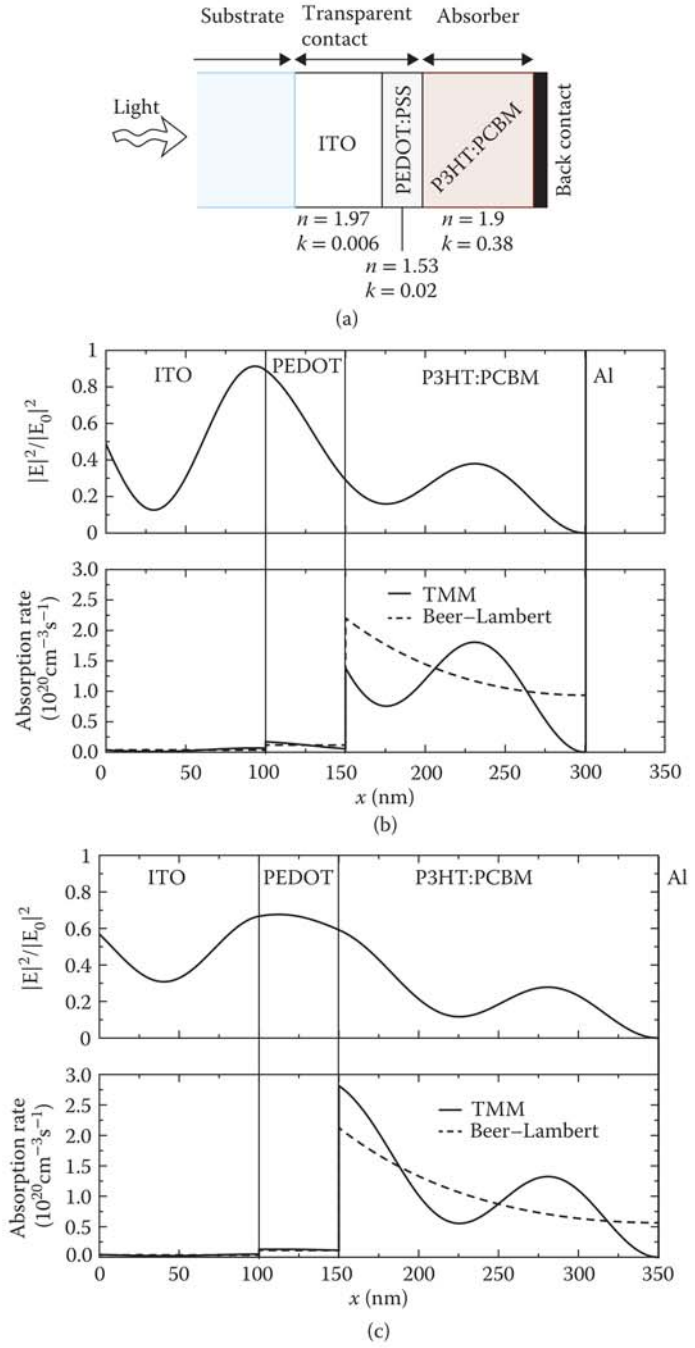


FIGURE 43.4 Structure for comparison of coherent and incoherent optical field calculation. The parameters used in the simulation are given in (a). In (b) and (c) the optical field intensity and absorption profiles for two different absorber thicknesses is shown. The local absorption rate considerably deviates from the Beer–Lambert prediction, in particular for thin layers. For thick layers Beer–Lambert behavior is recovered. (ITO, indium tin oxide; PCBM, [6,6]-phenyl-C61 butyric acid methyl ester;TMM, transfer matrix method.)

absorber thicknesses, showing that the profiles differ significantly from what would be expected from the Beer–Lambert model, and that it depends critically on layer thicknesses. In particular, absorption maxima can be located deeply inside the layer instead of near the interface. When the layers are becoming optically thick, the local absorption rate mediated over a wavelength converges to the result obtained by the Beer–Lambert law.

Figure 43.5 shows another example, borrowed from Reference [53], where the obtained results have been compared with experimental data. In this case, the donor–acceptor heterojunction is formed by the

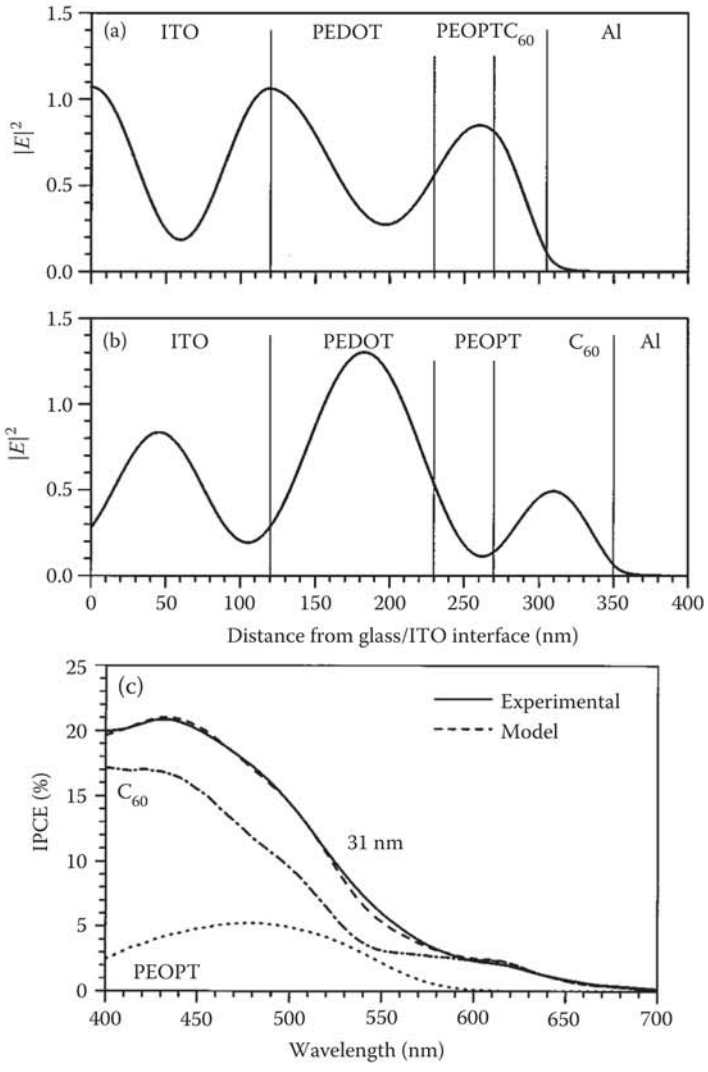


FIGURE 43.5 Field intensity profiles in an organic heterojunction solar cell for two different thicknesses of the C_{60} layer (a, b). The quality and accuracy of the optical simulation is proven by the comparison of experimental and simulated quantum efficiency in (c). Note that the simulation provides the information on where absorption happens, which cannot be obtained from the measurement. For example, in (c) the wavelength resolved contributions to the total absorption in the C_{60} and PEDOT layers is reported for an optimized C_{60} layer thickness. (Reproduced with permission from Leif A.A. Pettersson et al. Modeling photocurrent action spectra of photovoltaic devices based on organic thin films. *Journal of Applied Physics*, 86(1):487—496, 1999 Copyright 1999, American Institute of Physics Publishing LLC.)

PEOPT/C_{60} interface. To achieve maximum efficiency, a peak in field intensity should be obtained near this interface. Figure 43.3 presents the intensity profile at a wavelength of 460 nm for a C_{60} layer thickness of 35 and 80 nm, respectively, indicating that device optimization can benefit substantially from optical simulations. Note that the field profile is optimized in this way for a single wavelength only. Although the field profile cannot be measured directly experimentally, the accuracy of the optical simulation can be assessed by measuring the incident photocurrent efficiency (IPCE) (also denoted as external quantum efficiency, EQE), which is the product of quantum absorptivity and the current collection efficiency at each wavelength and gives the ratio of the number of collected electrons to the number of incident photons [43]. This quantity has been simulated in [53] by combining optical and current transport simulations, and the comparison between measurement and simulation is presented in Figure 43.3 for 31-nm thick C_{60} layer, showing excellent agreement. This example also shows how a simulation can give accurate insight into internal device behavior that is inaccessible by measurement alone, opening the way to efficient device optimization.

While 1D TMM-based modeling is a powerful tool to describe the optical field in a planar layered structure, it is less suitable to treat devices where light management techniques are employed, such as (possibly random) front or back surface texturing. Such approaches are becoming more and more important in order to approach the theoretical limits of light-to-electricity conversion, especially for technologies that are already optimized from an electrical point of view [65]. For organic solar cells, e.g., there is interest in grating structures on the back contact in order to modulate the electric field profile and thus to optimize optical generation and efficiency. Plasmonic effects are also studied for improving solar cell performance [66]. To accurately model texturing, gratings, or plasmonic effects, usually full 3D Maxwell solvers are employed [67,68]. Unfortunately, such calculations are computationally quite expensive, so that approximate models formulated in one dimension are also of interest [69]. Special approaches have been devised also for other cases, like thick layers with incoherent coupling [70].

Also note that from Figure 43.1 it might seem that TMM is not suited for devices based on mesoporous or blended materials, since the materials are not really piecewise homogeneous and there are apparently no planar interfaces. However, since the characteristic dimensions are in the order of at most tens of nanometers, these composite materials can be described with sufficient accuracy as effective materials so that TMM can be applied without particular restrictions.

43.4 Transport Modeling

While optical calculations are important for the simulation of solar cells, taken alone they can only provide upper limits for cell efficiency. For a comprehensive view on cell operation, the transport of the generated carriers and their collection at the electrical contacts has to be taken into account. This section reviews the two carrier transport models that are used most often for solar cell simulations, namely the drift-diffusion model and kinetic Monte Carlo (kMC).

43.4.1 Drift-Diffusion Model

Modeling of electronic transport in thin-film solar cells is mostly based on the drift-diffusion model, which has been the work-horse of physics-based electronic device simulation since decades. A thorough introduction, derivation of the model and implementation hints can be found in different textbooks such as [71], and a discussion of theoretical and mathematical aspects can be found, e.g., in [72]. A more detailed description is provided in Chapter 50.

Since the drift-diffusion model is well known and is described in numerous books and articles, we will provide only a short generic description and concentrate more on the features important for thin-film solar cell simulation.

For a "normal" semiconductor where only electrons and holes are mobile carriers, the drift-diffusion model consists of a set of three continuity equations for the electric displacement $D = -\varepsilon\nabla\varphi$ and the two particle currents J_n and J_p:

$$\nabla(\varepsilon\nabla\varphi) = q(n - p - C) \tag{43.31}$$

$$\frac{\partial n}{\partial t} - \frac{1}{q}\nabla J_n = -R, \qquad J_n = q(D_n\nabla n - \mu_n n\nabla\varphi) \tag{43.32}$$

$$\frac{\partial p}{\partial t} + \frac{1}{q}\nabla J_p = \quad R, \qquad J_p = -q(D_p\nabla p + \mu_p p\nabla\varphi) \tag{43.33}$$

This nonlinear system of partial differential equations is also known as the *basic semiconductor device equations* or the van Roosbroeck system, who was the first to formulate this set of equations in the framework of electronic transport in semiconductors [73]. The first of the above equations (Equation 43.31) is the Poisson equation, relating the electrostatic potential φ with the total charge density $\rho = q(p - n + C)$ where n and p are the electron and hole densities, respectively, while ε is the permittivity. C denotes the net density of fixed charges such as dopants or traps. The other two equations are the continuity equations for the electron and hole currents, where the constitutive equations for the currents are given on the right (which are also known as *drift-diffusion equations*) where q is the electron's charge, $D_{n,p}$ are the electron and hole diffusion coefficients, and $\mu_{n,p}$ are the electron and hole mobilities. R is the net recombination rate, including all processes which lead to destruction or generation of electron–hole pairs.

The three most important recombination models are the trap-assisted (Shockley–Read–Hall [SRH]) recombination, bimolecular (direct, radiative) recombination, and Auger recombination, given by [71,74]

$$R_{\mathrm{SRH}} = \frac{np - n_i^2}{\tau_p(n + n_t) + \tau_n(p + p_t)} \qquad \text{SRH} \tag{43.34}$$

$$R_{\mathrm{rad}} = B(np - n_i^2) \qquad \text{Radiative} \tag{43.35}$$

$$R_{\mathrm{auger}} = (C_n n + C_p p)(np - n_i^2) \qquad \text{Auger} \tag{43.36}$$

Here, n_i^2 is the square of the intrinsic carrier density, given by the product of the equilibrium densities n_0 and p_0. n_t and p_t are parameters depending on the energetic position inside the gap of the trap. Note that Equation 43.34 is valid for a single discrete trap level and in the absence of an electric field, however, it can be generalized to other cases [75,76]. The other parameters in the expressions are mostly material related. The relative importance of the different recombination mechanisms depends on the operating conditions and on material quality. SRH recombination is important at low carrier densities and in presence of defects, while Auger recombination usually becomes dominant only at high carrier densities.

Written in standard form (Equations 43.31 through 43.33), the particle currents are seen to be decomposed into a diffusion term proportional to the gradient of the carrier density and a drift component proportional to the conductivity times the electric field. This is an intuitive description of particle transport, whenever the system is dominated by scattering, that is when the particles undergo Brownian motion with some thermal velocity, and the external perturbation (electric field or density gradient) induces only a small nonzero mean carrier velocity leading to a net current flow.

The system of Equations 43.31 through 43.33, supplemented with suitable boundary conditions, is usually solved numerically, although under certain circumstances analytic solutions exist in one dimensions. For implementation details and numerical methods, we refer to [71] and the vast existing literature on the topic.

The device equations (Equations 43.31 through 43.33) are still the most important model for transport in semiconductor devices, since they can be solved much faster than any other transport model and

therefore represent a good compromise between computational efficiency and physical accuracy. While the equations allow easy interpretation, and can be solved relatively easily, they are based on a number of assumptions that are important when applying them to systems other than crystalline semiconductors. The other important point to note is that they give a good description of device behavior only when good models for the quantities appearing as parameters in Equations 43.31 through 43.33 can be provided. In particular, the models for the recombination and generation processes and for the carrier mobilities, in combination with accurate material-dependent parameters, are important in order to reach satisfactory results with the drift-diffusion model. Most often this implies careful parameter fitting against experimental data, careful checking whether the fitted parameters provide a predictive model or not.

In the following, we discuss some of the basic assumptions underlying the semiconductor device equations (Equation 43.31 through 43.33) and possible consequences when applying them to the materials of interest in this chapter. A discussion of technology-specific issues regarding the basic equations or the parameter models is deferred to Section 43.5.

The standard way to derive the drift-diffusion equations is to start from the semiclassical Boltzmann transport equation [72]. These derivations include implicitly or explicitly assumptions on the distribution function such that the carrier density takes the form of the well-known Boltzmann approximation, i.e.,

$$n = N_\mathrm{c} \mathrm{e}^{\frac{E_\mathrm{F,n} - E_\mathrm{c}}{k_\mathrm{B} T}}, \quad p = N_\mathrm{v} \mathrm{e}^{-\frac{E_\mathrm{F,p} - E_\mathrm{v}}{k_\mathrm{B} T}} \tag{43.37}$$

Here, N_c and N_v are the effective density of states (DOS) in the conduction and valence band, $E_\mathrm{F,n}$ and $E_\mathrm{F,p}$ are the electron and hole quasi Fermi levels, and E_c and E_v are the conduction and valence band edges. The latter can be expressed as $E_\mathrm{c,v} = E_\mathrm{c,v}^{(0)} - q\varphi$, where $E_\mathrm{c,v}^{(0)}$ are material band edges in the absence of an electric field, identifiable with electron affinity and ionization energy, k_B is the Boltzmann constant, and T the temperature.

The diffusion constants and mobilities in Equations 43.31 through 43.33 are not independent, which can be easily seen when substituting Equation 43.27 into Equations 43.31 through 43.33 and solving for equilibrium, i.e., zero current under no external bias nor carrier generation. It is found that the two quantities are related by a so-called Einstein relation [74]:

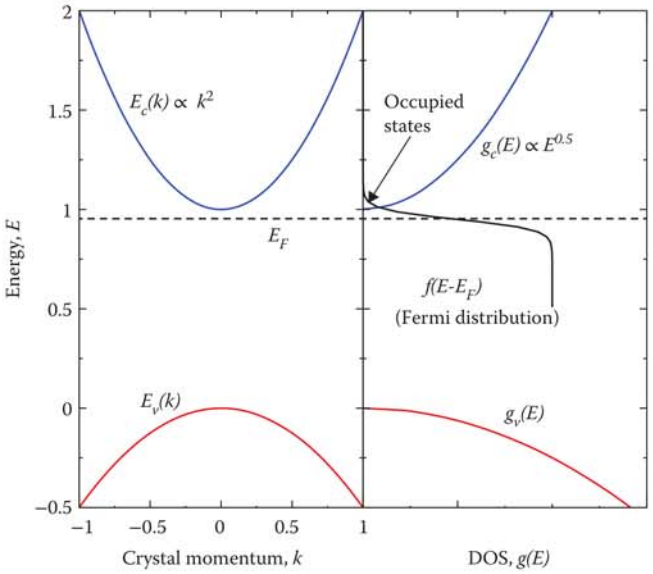
$$D_\mathrm{n} = \frac{k_\mathrm{B} T}{q} \mu_\mathrm{n}, \quad D_\mathrm{p} = \frac{k_\mathrm{B} T}{q} \mu_\mathrm{p} \tag{43.38}$$

which in actual implementations of the drift-diffusion model is usually used to obtain the diffusion coefficients from the mobilities.

The expressions (Equation 43.37) provide an adequate description of the carrier densities for most nondegenerate crystalline semiconductors, i.e., at not too high carrier densities. This is because the same expressions are obtained when calculating the carrier density from the DOS and the Fermi–Dirac distribution, and making the Boltzmann approximation for small densities. Generally, the density, e.g., of electrons is found under the assumption of a local equilibrium distribution by occupying the available states, given by the DOS, with a local quasi Fermi level and integrating over all energies, i.e.,

$$n(\mathbf{r}) = \int g(E) f\left(\frac{E - E_\mathrm{F,n}(\mathbf{r})}{k_\mathrm{B} T}\right) \mathrm{d}E \tag{43.39}$$

where $f(x)$ is the Fermi–Dirac distribution and $g(E)$ the DOS. The DOS of a crystalline semiconductor at the relevant energies, i.e., near the band extrema, is obtained from the energy dispersion, which is approximately parabolic in this range, as shown schematically in Figure 43.6. Details on the calculation of the carrier densities can be found, e.g., in [42,71,74,77].



$$E(k) = E_c^{(0)} + \frac{\hbar^2 k^2}{2m^*}, \quad f(x) = \frac{1}{1+e^x}$$

$$g_c(E) = \frac{1}{\pi^2}\left(\frac{2m^*}{\hbar^2}\right)^{3/2}\sqrt{E}$$

$$n = \frac{1}{(2\pi)^3}\int f\left(\frac{E_c(k)-E_F}{k_B T}\right) dk^3$$

$$= \int_{E_c}^{\infty} g_c(E) f\left(\frac{E-E_F}{k_B T}\right) dE$$

$$= \underbrace{2\left(\frac{m^* kT}{2\pi\hbar^2}\right)^{3/2}}_{N_c} F_{1/2}(E_F/k_B T)$$

$$\approx N_c \cdot e^{(E_F-E_c)/k_B T}$$

FIGURE 43.6 Parabolic approximation to the energy dispersion $E(k)$ of a simple semiconductor and the DOS resulting from this dispersion, obtained by setting $1/(2\pi)^3 \int dk^3 = \int g(E)dE$. On the right, some of the expressions are given, and the calculation of the density is sketched ($m^* = \hbar^2(\partial^2 E(k)/\partial k^2)^{-1}$ is the effective mass, E_F is the quasi Fermi level).

The key point is that the derivations are based on the assumption that the carriers are moving freely in a periodic potential of the underlying crystal lattice, which allows to define the momentum k of the carriers and which leads to the parabolic dispersion valid at least in the vicinity of the band extrema. While this is a good approximation for crystalline semiconductors, it does not hold in many materials of interest in thin-film PV technology. In particular, whenever the periodicity of the crystal structure is lost, the DOS will become substantially different. In an amorphous semiconductor such as a-Si, e.g., the crystalline lattice structure is maintained only locally and only approximately: a-Si is mostly tetrahedrally bonded, but with a statistical distribution of bond angles and with certain number of silicon atoms with three or five neighbors. This modifies the DOS by washing out the band edge and appearance of approximately exponential band tails, known as Urbach tails [78], and appearance of localized states in the gap even in absence of impurity atoms, as shown schematically in Figure 43.7b. Similarly, organic semiconductors are usually characterized by a certain energetic disorder in the molecular energy levels, and the definition of carrier momentum is largely meaningless. In this case, the DOS is completely different from the "standard" semiconductor bulk DOS, and is often well described by a Gaussian distribution with a certain width σ characterizing the statistical scattering of the molecular levels [79], shown in Figure 43.7c.

A consequence of a DOS, which differs from that resulting from a parabolic dispersion, is that the Einstein relation relating the diffusion constants and the mobilities needs to be generalized. This general form can be easily found from Equations 43.31 through 43.33 as

$$D_n = n\left(\frac{\partial n}{\partial \varphi}\right)^{-1}\mu_n, \quad D_p = -p\left(\frac{\partial p}{\partial \varphi}\right)^{-1}\mu_p \tag{43.40}$$

An accurate description of semiconductor transport therefore requires to take the functional form of the DOS into account, although it does not appear explicitly in the basic semiconductor equations. Beside the formal modification of the Einstein relation, the DOS also influences some directly measurable quantities.

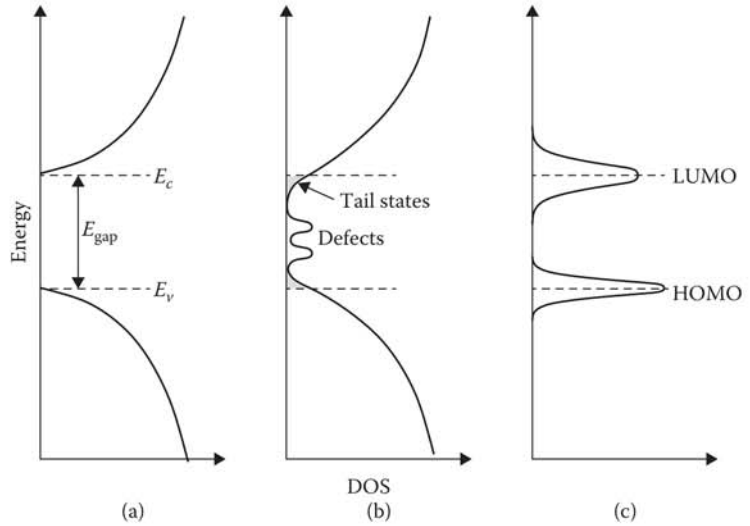


FIGURE 43.7 Sketch of different density of states (DOS): (a) crystalline semiconductor with parabolic band dispersion, (b) amorphous semiconductor with band tails and defect states, (c) disordered semiconductor, e.g., organic, with the DOS formed from energetic and spatial disorder of molecular orbital energies. In the last case, the bands are usually denoted as HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital).

Assume, e.g., a bulk absorber in flat band condition subject to homogeneous optical carrier generation, so that there is no current flow, and assume that only radiative recombination is relevant. In this case, all generated electron–hole pairs have to recombine radiatively according to Equation 43.35, i.e., $G = Bnp$. The splitting of the electron and hole quasi Fermi levels induced by the illumination is a measure for the open-circuit voltage V_{oc} that can be expected. Using Equation 43.37, an analytic formula can be easily found as [80]

$$V_{oc} = \frac{1}{q}(E_{F,n} - E_{F,p}) = \frac{1}{q}E_g + \frac{k_B T}{q}\ln\left(\frac{G}{BN_c N_v}\right) \tag{43.41}$$

and apparently the V_{oc} should vary by a constant rate of $2.3k_B T/q$ per decade in intensity. A different result is obtained if the same quantity is calculated using a different expression for the DOS, e.g., using a Gaussian DOS (GDOS) as typical for organic semiconductors and given by

$$g(E) = \frac{N_0}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(E-E_0)^2}{2\sigma^2}\right) \tag{43.42}$$

Here N_0 is the total density, σ the width (variance) of the Gaussian, and E_0 the center of Gaussian. Using approximations of the formulas for the densities [81], and assuming the same parameters for the electron and hole DOS, we can obtain the open-circuit voltage for sufficiently large carrier densities in this case as

$$V_{oc} = \frac{1}{q}E_g - 2\frac{\sigma\sqrt{2}}{qH(\sigma)}\mathrm{erfc}^{-1}\left(2\sqrt{\frac{G}{BN_0^2}}\right) \tag{43.43}$$

Here, E_g is defined as the difference of the center energies of the electron and hole DOS, and $H(\sigma) = k_B T\sqrt{2}/\sigma \cdot \mathrm{erfc}^{-1}(\exp(-\frac{1}{2}\sigma^2/(k_B T)^2))$. Figure 43.8 compares the derivative of V_{oc} with respect to the logarithm of the generation rate for a parabolic DOS and a GDOS, showing that for typical Gaussian widths

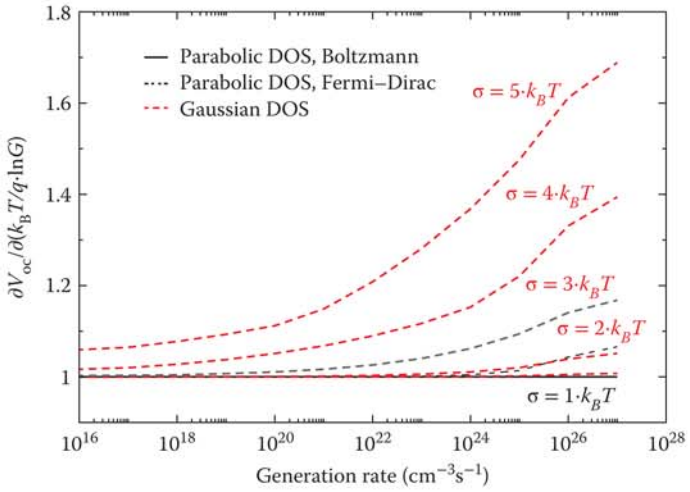


FIGURE 43.8 Variation of the bulk quasi Fermi level splitting due to a homogeneous generation rate. Deviation from the exponential form of the carrier density leads to considerable deviation from the expected rate of change. For the Gaussian density of states (DOS), $N_0 = 10^{21}$ cm^{-3} has been used, for the parabolic case $N_c = N_v = 10^{19}$ cm^{-3}. The recombination rate constant was $B = 10^{-10}$ cm^3·s^{-1}.

the DOS influences the rate even at moderate generation rates. The effect becomes stronger with increasing intensity. The measurable effect of energetic disorder is described, e.g., in [82].

Note that a further modification of the form (Equations 43.31 through 43.33) of the drift-diffusion equations is needed whenever the material parameters in the formula for the carrier densities (i.e., effective DOS or the band edges, see Figure 43.6), or generally the DOS, change spatially. This can be seen easily when evaluating the gradient of the densities, which then leads to additional terms that can be interpreted as effective potentials. These corrections are simple in the case of the standard exponential form (Equation 43.37) of the carrier densities, but they become more complicated in the degenerate case and for other DOS. It is therefore sometimes preferred to write the drift-diffusion model in a different but equivalent form, derived from thermodynamics [61,83]:

$$\nabla(\varepsilon\nabla\varphi) = q(n - p - C) \tag{43.44}$$

$$\frac{\partial n}{\partial t} - \frac{1}{q}\nabla J_n = -R, \qquad J_n = -q\mu_n n\nabla\phi_n \tag{43.45}$$

$$\frac{\partial p}{\partial t} + \frac{1}{q}\nabla J_p = -R, \qquad J_p = -q\mu_p p\nabla\phi_p \tag{43.46}$$

Here, $\phi_{\{n,p\}}$ are the electrochemical potentials of the electrons and holes, respectively, related to the quasi Fermi levels by $E_{F,\{n,p\}} = -q\phi_{\{n,p\}}$. While in Equations 43.31 through 43.33, the dependent variables are the electrostatic potential and the carrier densities, in (Equations 43.44 through 43.46) they are the electrostatic and the electrochemical potentials, and the particle currents are driven by the gradient of the electrochemical potential. The advantage of this formulation with respect to the one in Equation 43.31 through 43.33 is that it is independent on the functional form of the density, and the Einstein relation is not needed explicitly. For the rest, it has the same limitations of applicability as Equations 43.31 through 43.33.

From a technical point of view, note that the decoupled carrier transport equations in Equations 43.44 through 43.46 are quasilinear, with the differential operator being proportional to the carrier density, while in Equations 43.31 through 43.33 the transport equations are linear. This makes the numerical solution procedure more involved.

If we are using the thermodynamic formulation (Equations 43.44 through 43.46) for simulation of solar cells, then we need an adequate model for the DOS or, respectively, the carrier density for all the materials involved, and we need models for the carrier mobilities and the recombination rates. Especially, the mobility is a critical parameter. In a crystalline semiconductor, the mobility is related to the scattering the carriers are subject to. So it depends strongly on temperature and defect or doping densities, and usually mobility increases with decreasing temperature and decreasing defect densities, since the associated scattering rates decrease. Mobility also depends on the electric field, typically by decreasing with increasing field due to a saturation of the mean drift velocity. In fact, the proportionality between drift velocity and electric field, given by $v = \mu_n E$, holds only at low field strength [84]. In noncrystalline materials such as amorphous or organic semiconductors, these dependencies can be radically different. For example, mobility may increase with temperature because carriers can escape more easily local energy minima or overcome small energy barriers. Or it may increase with electric field because the field reduces the effective energy barriers hindering carrier flow. In some cases, such behavior is a manifestation of an underlying microscopic transport that deviates from a diffusive one.

In disordered organic materials, e.g., the deviation from a periodic crystalline structure makes the concept of band-transport inapplicable. Instead, charges are spatially localized and charge transport occurs by transitions between localized states, a process termed as *hopping*. In reality, disordered systems show both disordered and (semi-)crystalline regions so that charge transport is always an interplay between hopping and band conduction. Higher contributions of hopping transport usually lead to lower charge mobilities. Localized states can be imagined as local energy valleys in which charges get confined. Transport can either take place by surmounting the energy barrier to a neighboring localized state by absorbing phonons, or by tunneling through the barrier. The concept of phonon-assisted hopping was introduced by Mott [85]. In 1960, Miller and Abrahams described a hopping model between two localized states consisting of both a contribution due to thermal activation and tunneling. After their model, the hopping process between the states i and j with energy difference $\Delta E = E_j - E_i$ and distance r_{ij} is of the form [86]

$$a_{i\to j} = a_0 \exp\left(-2\gamma r_{ij}\right) \begin{cases} \exp\left(-\frac{\Delta E}{k_B T}\right) & \text{if } \Delta E > 0 \\ 1 & \text{if } \Delta E \leq 0 \end{cases} \tag{43.47}$$

where a_0 is given by a typical phonon frequency in the system and γ is the inverse localization constant. The formula combines the contribution of two different processes to hopping: the first exponential describes a tunneling contribution through a barrier given by the distance r_{ij} of the localized states, and the latter term stands for the thermally activated contribution with the relation between barrier height ΔE_{ij} and the thermal energy $k_B T$.

Another model that takes into account strong coupling between charges and lattice is the polaron model. In organic materials with low permittivities, a charge on a molecule can considerably deform the structure of the molecule. The complex of charge and the distortion of the lattice it induces is known as a quasiparticle called *polaron*. Marcus [87] developed a model for the transition of polarons:

$$a_{i\to j} = \frac{\left|J_{ij}\right|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}}\exp\left(-\frac{(\Delta G_{ij} - \lambda)^2}{4\lambda k_B T}\right) \tag{43.48}$$

where J_{ij} is the transfer integral for electron and hole transfer, ΔG_{ij} is the free energy difference between state, i and j and λ the reorganization energy, which accounts for the reconfiguration of the molecule due to the deformation by the charge.

In order to derive a mobility model for disordered materials, a description for the distribution of available states must be available. The spatial irregularity of these systems leads to fluctuations in the intermolecular interactions. A successful model to represent these energetic fluctuations is the GDOS as given in

Equation 43.42, in which the energy distribution of the localized states is spread by a Gaussian distribution function around the HOMO/LUMO energy levels of the organic material. The GDOS in combination with the Miller–Abrahams hopping was applied by Bässler [79] by Monte Carlo simulations to study the effect of temperature and electric field on the charge mobility. He found a non-Arrhenius temperature dependence $\mu(T) = \mu_0 \exp\left(-(t\hat{\sigma})^2\right)$ with $\hat{\sigma} = \sigma/k_B T$, and, Poole–Frenkel type behavior $\mu(F) = \mu_0 \exp\left(\gamma\sqrt{F}\right)$ for the dependence on the electric field F for a limited range of electric fields. To include the influence of charge carrier densities on the mobility, Pasveer et al. [88] parametrized the results based on a direct solution of the Master equation in a system with a GDOS. They found as expression for the hopping mobility:

$$\mu(T, \rho, F) = \mu(T, \rho) f(T, F) \text{ with} \tag{43.49}$$

$$\mu(T, \rho) = \mu_0 c_1 \exp\left[-c_2 \left(\frac{\sigma}{k_B T}\right)^2\right] \exp\left[\frac{1}{2k_B T}\left(\frac{\sigma^2}{k_B T} - \sigma\right)(2\rho a^3)^\delta\right] \tag{43.50}$$

$$f(T, F) = \exp\left\{0.44\left[\left(\frac{\sigma}{k_B T}\right)^{\frac{3}{2}} - 2.2\right]\left[\sqrt{1 + 0.8\left(\frac{eFa}{\sigma}\right)^2} - 1\right]\right\} \tag{43.51}$$

where ρ is the charge carrier density, either electrons or holes, F is the electric field, $k_B T$ the thermal energy, $a = N_0^{-\frac{1}{3}}$ the distance between localized states, σ the energetic disorder in the GDOS, $c_1 = 1.8 \cdot 10^{-9}$, $c_2 = 0.42$, $\mu_0 = a^2 \nu_0 e/\sigma$, with ν_0 the prefactor in the Miller–Abrahams formula, and

$$\delta = 2\frac{\log(s^2 - s) - \log(\log(4))}{s^2} \text{ with } s = \frac{\sigma}{k_B T} \tag{43.52}$$

This parametrization can be used in drift-diffusion models to account for the hopping nature of transport.

43.4.2 Kinetic Monte Carlo

In this section, we give an overview of the kinetic Monte Carlo (kMC) method [89,90] and how it can be applied to set up a model for organic solar cells.

Monte Carlo methods offer a simple yet powerful tool to solve real world problems by the use of random numbers. In the hierarchy of transport models shown in Figure 43.9, kMC bridges the gap between macroscopic models such as drift-diffusion and molecular dynamics (MD) and quantum mechanical models.

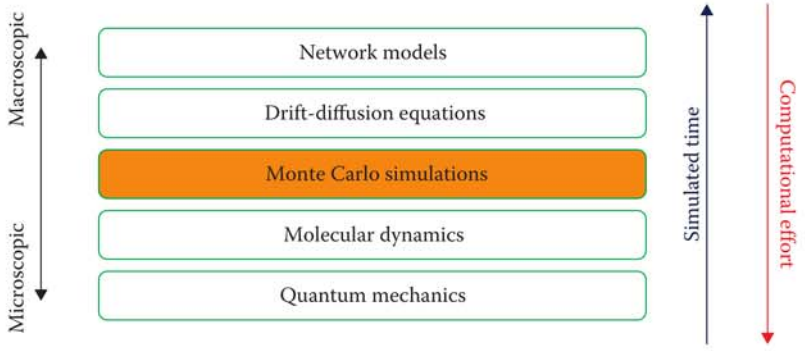


FIGURE 43.9 Hierarchy of transport models. Kinetic Monte Carlo simulations bridge the gap between macroscopic and microscopic modeling techniques.

The idea for the use of random numbers in an algorithm came from Fermi [91] and was further pursued by Ulam and Metropolis in the 1940s studying neutron diffusion in fissionable materials. They developed the most used method of this kind, the Metropolis algorithm [92], during a study about the calculation of state equations for a system of interacting particles described by a hard sphere cutoff potential.

In the following decades, Monte Carlo methods have proven to be a versatile tool and much progress in the development of calculations for a variety of areas, ranging from surface adsorption processes, radiation damage annealing, and statistical physics [93–95], has been achieved. Especially, the development of a Monte Carlo algorithm class that models physical systems as dynamically evolving from state to state by a certain set of transition rates obtained large interest as it allowed to study time-dependent properties of physical systems. This class of algorithms is today commonly known as *kinetic Monte Carlo*. It has received great benefit in the growing computational power of modern processors and provides a flexible tool to simulate dynamical behavior of physical systems, e.g., transport processes in inorganic semiconductor devices [96]. Major contributions to the modern kMC algorithm, as it is described here, have been made by Bortz et al. in 1975 [95] and Gillespie in 1976 [89] and 1977 [97].

Compared to MD simulations, which include the interactions of single atoms with all dynamics of the molecules like vibrations, occurring on very small time scales (fs), kMC characterizes the state of a system by a set of more macroscopic variables which do not include underlying dynamics like fast vibrational modes. These motions on smaller time scales can be ignored as long as the current state can still be assigned to a long-time state. The dynamical evolution of the system is only seen by transitions between such long-time states, called events. Such a system is called an infrequent-event system [90]. Figure 43.10 illustrates this concept. By ignoring the underlying fast motions and only considering the transitions between long-time states, a sufficient amount of processes can be neglected. Thus, much larger overall simulation times of the order of milliseconds to seconds can be reached.

The above mentioned concept can be applied in particular to organic solids, where particles such as electrons, holes and excitons reside in localized states and transport is determined by hopping from one localized state to another. The time scale of these hopping events is much longer than other processes, which do not change the spatial localization of the carriers.

Compared to drift-diffusion, kMC has the advantages of being able to capture the discrete nature of the problem by modeling individual particles at their molecular positions, and of providing a physically more accurate description of hopping transport without having to resort to effective continuous media models. Moreover, it can be much easier and physically more sound to treat blended materials with nanoscale morphology in kMC.

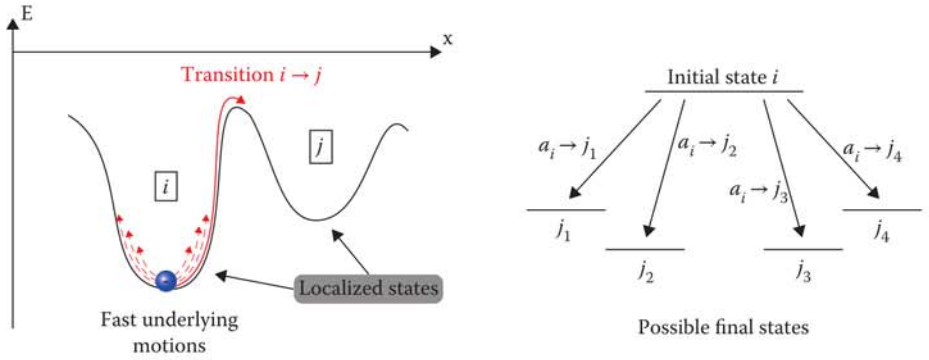


FIGURE 43.10 Transitions from long-time state to long-time state with underlying fast modes in an infrequent-event system (left). State-to-state transitions reduced to the needs for kinetic Monte Carlo (right).

43.4.2.1 kMC Algorithm

The essential concept of the kMC method is the characterization of the considered system by a set of system-specific long-time states. A state is, e.g., defined by the position of all particles in the system at a certain time. Under dynamical evolution, it is possible for the system to change its state by certain system-specific transitions, thus motion of particles from one state to another. The task for an accurate model is to identify all the important processes that take place in the system. By considering processes that happen on long time scales and neglecting those which act on very small time scales, one is able to accomplish a tradeoff between accuracy and sufficiently long simulation times. It is of importance to carefully choose the implemented processes in order not to lose essential parts of the system dynamics.

Let the system be in a certain state i. Depending on the current system setup, a set of several different states j exists to which a transition can take place. This is indicated in Figure 43.10 (right). The transitions from beginning states i to possible end states j can be characterized by rate constants a_{ij} given in units of s^{-1}. Rate constants are a measure for the probability of the corresponding transition from $i \rightarrow j$. An important property must be that the rates are memoryless, meaning that they are only dependent on the setup of the current state and are not dependent on how the system evolved into it. This property is known from the theory of Markov chains [98] and allows, if the set of rate constants is known, to determine the dynamical evolution of the system.

As known from statistical physics [99], stochastical processes characterized by discrete transitions from state to state under temporal evolution can phenomenologically be described by the master equation

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = -\sum_j a_{ij}\, p_i(t) + \sum_j a_{ji}\, p_j(t) \tag{43.53}$$

The probability $p_i(t)$ to find the system in state i at a certain time t changes in a way that is determined by the rate constants a_{ij} and the probabilities of only the initial state $p_i(t)$ and the possible final states $p_j(t)$. Finding solutions of the master equation, be it by analytical or numerical methods, is often complex and not feasible. The kMC method is based on a stochastical framework and provides a numerical approach to obtain the dynamic system evolution based on the time-dependent propagations from state to state. By choosing a transition pathway through a chain of subsequent states, one possible dynamic system evolution is obtained. Such a walk through a pathway of transitions forms a Markov chain. Averaging over a large number of such Markov chains provides an equivalent system behavior as described by the master equation. Therefore, the kMC algorithm is essentially emulating the master equation.

For a detailed derivation of the associated mathematical framework, the reader is referred to the original work of Gillespie [89]. In the following, we present only the equations necessary for a possible implementation of the algorithm. The algorithm as proposed by Gillespie relies on the ability to characterize the system events by a set of transitions

$$\{\mu\} \text{ with } \mu = 1, 2, \ldots, M \tag{43.54}$$

where M is the total number of transitions. μ is equal to a transition from state i to one of the possible states j in that configuration: $i \rightarrow j_\mu$. Then a_μ is the rate value describing process μ and $\{ a_\mu \}$ is the corresponding set of rates. All transitions must describe distinct processes and their number M can vary over time, depending on which events are enabled at a certain step in the system evolution. It is essential to consider all major processes in order to be able to accurately reproduce the behavior of the system. The main task is now to develop a procedure to simulate the time-dependent evolution of the system under consideration of the given transitions. This makes it necessary to derive an algorithm that is able to choose which transition

will be carried out and (for the time dependency) at what time span, all based on the drawing of random numbers.

At any given time, the set of $\{a_\mu\}$ characterizes the possible transitions. They are functions of the current system setup and will change over time, as described later.

Instead of finding a solution to the master equation, the kMC method is based on a function called the probability density function (PDF). The PDF is the suitable function to select a transition and its corresponding time step based on random numbers. The PDF is defined as

$$\begin{aligned} P(\tau,\mu)\,d\tau = {} & \text{probability at time } t \text{ for the next transition} \\ & \text{to occur in time interval } (t+\tau\,,\ t+\tau+d\tau) \\ & \textit{and} \text{ being a transition of type } R_\mu \end{aligned} \tag{43.55}$$

where $P(\tau,\mu)$ is the PDF. This function can be interpreted as follows: it is a joint PDF of the continuous time variable τ $(0 \leq \tau < \infty)$ and the integer variable μ $(\mu = 1, 2, \ldots, M)$ characterizing the transition. At a certain time t, the PDF represents the probability for transition μ to occur within an infinitesimal time span of $d\tau$ after no other event has taken place for a time interval τ. Hence, $P(\tau,\mu)$ can be seen as the independent probabilities to have no transition in the time interval $(t\,,\ t+\tau)$ *and* to have a transition of type R_μ immediately afterward in the infinitesimal interval $(t+\tau\,,\ t+\tau+d\tau)$. In common terms, the transition μ needs time τ before the next transition occurs.

Using the abbreviation

$$a = \sum_{\nu=1}^{M} a_\nu \tag{43.56}$$

as the total transition probability per unit time, the total expression for the PDF is given by

$$P(\tau,\mu) = a_\mu \cdot \exp\left(-\sum_{\nu=1}^{M} a_\nu \tau\right) = a_\mu \exp(-a\tau) \tag{43.57}$$

This function, only dependent on the set of events a_μ, is the starting point to choose a random number pair (τ,μ) that characterizes the next transition and time step in the dynamic evolution of the system. The PDF is properly normalized, meaning that every pair (τ,μ) with $\tau \in (0 \leq \tau < \infty)$ and $\mu = 1, 2, \ldots, M$ has its weighted contribution to the PDF and that it is guaranteed to obtain an associated value for the PDF by randomly choosing one pair:

$$\int_{t=0}^{\infty} \mathrm{d}\tau \sum_{\nu=1}^{M} P(\tau,\nu) = \sum_{\nu=1}^{M} a_\nu \int_{t=0}^{\infty} \mathrm{d}\tau \exp(-a\tau) = 1 \tag{43.58}$$

Now, the fundamental Monte Carlo step can be defined: by choosing a random pair of the variables (τ,μ) according to the PDF $P(\tau,\mu)$ after Equation 43.57, a weighting based on the magnitude of the various transitions and the exponential decrease in time is performed. For the execution of this step, two approaches were suggested by Gillespie: the *direct method* and the *first reaction method*. In this context, the former method is used. The direct method is most efficient in the way it utilizes random numbers to determine a pair of transition and a time step, while the first reaction method can save computational power at the cost of accuracy.

In the direct method, the PDF is divided into two separate PDFs, each of which is itself normalized and only dependent on either the time τ or the transition μ:

$$P(\tau, \mu) = P_1(\tau) \cdot P_2(\mu|\tau) \tag{43.59}$$

Here, the first factor represents the probability density that any of the possible transitions takes place in the time interval $(t+\tau, t+\tau+d\tau)$, irrespective of which one. Thus, $P_1(\tau)$ is the sum over all transitions. With Equation 43.57 follows

$$P_1(\tau) = \sum_{\nu=1}^{M} P(\tau, \nu) = \sum_{\nu=1}^{M} a_\nu \exp(-a\tau) = a \cdot \exp(-a\tau) \tag{43.60}$$

From this and Equation 43.59, $P_2(\mu|\tau)$ is easily derived as

$$P_2(\mu|\tau) = \frac{P(\tau, \mu)}{P_1(\tau)} = \frac{a_\mu}{a} \tag{43.61}$$

and $P_2(\mu|\tau)\, d\tau$ represents the probability that a reaction μ occurs, given the condition that a time step of τ has been chosen before. Both P_1 and P_2 are properly normalized PDFs on their own.

A pair of random numbers (τ, μ) according to the distributions (Equations 43.60 and 43.61) can be produced as follows, using (pseudo-)random number generator with uniform distribution available in all modern programming languages. Every normalized PDF $P(x)$ has an associated cumulative distribution function, defined by

$$F_1(x) = \int_{-\infty}^{x} P_1(x')\, dx' \qquad \text{for a continuous } P_1(x) \tag{43.62}$$

$$F_2(i) = \sum_{i'=1}^{i} P_2(i') \qquad \text{for a discrete } P_2(i) \tag{43.63}$$

Because $\lim_{x \to -\infty} F_1(x) = 0$ and $\lim_{x \to \infty} F_1(x) = 1$ for the continuous case, and $F_2(0) = 0$ and $F_2(M) = 1$ for the discrete case, both functions are monotone and limited. Thus, there exists an inverse function F^{-1} for each of them. It is made sure that every value from the output range has a corresponding value x or i in the domain range. The concept of the inversion method is to draw two uniform random numbers $r_1, r_2 \in (0, 1)$, and calculate $x = F_1^{-1}(r)$ and i from $F_2^{-1}(i)$. This way the obtained values for x and i are distributed according to the PDF from which F is derived.

Using this inversion method and the PDF $P_1(\tau) = a \cdot \exp(-a\tau)$, the corresponding cumulative probability function is[†]

$$F_1(\tau) = \int_0^{\tau} a \cdot \exp(-a\tau')\, d\tau' = 1 - \exp(-a\tau) \tag{43.64}$$

The continuous and exponentially distributed time variable τ, which describes the time passed before one transition (or, for a more vivid interpretation: the time "needed" for one transition), is then

[†] Note that $P_1(t)$ is only defined for $t \geq 0$.

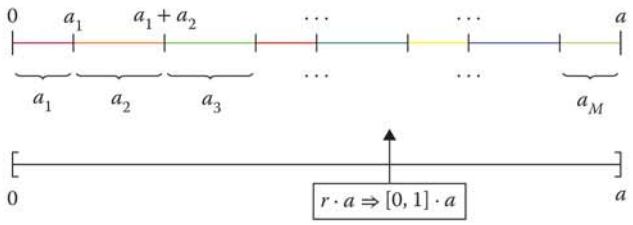


FIGURE 43.11 Drawing a discrete transition by a continuous random variable $r \in [0, 1]$. All enabled rate constants are added to yield the total rate a. Multiplication of a random number in the unit interval with the total rate corresponds to one specific rate process, weighted by its magnitude. Rates are added until the next addition will exceed $r \cdot a$. The last rate in the addition is chosen.

given by the inverse function of F_1. Hence,[†]

$$\tau = F_1^{-1}(r_1) = \frac{1}{a} \ln\left(\frac{1}{r_1}\right) \qquad \text{with } a = \sum_{\nu=1}^{M} a_\nu \text{ and } r_1 \in (0, 1) \tag{43.65}$$

To choose a transition μ from the set of events $\{a_\mu\}$ ($\mu = 1, 2, ..., M$) by a random number r_2, the cumulative probability must satisfy

$$F_2(\mu - 1) < r_2 \leq F_2(\mu) \tag{43.66}$$

That is, $r_2 \in (0, 1)$ corresponds to one plateau in the discrete histogram of F_2. Substituting Equation 43.61 in Equation 43.63 yields an expression for $F_2(\tau|\mu)$:

$$F_2(\tau|\mu) = \sum_{\nu=1}^{\mu} \frac{a_\nu}{a} \tag{43.67}$$

With this, the abovementioned condition reads

$$\sum_{\nu=1}^{\mu-1} \frac{a_\nu}{a} < r_2 \leq \sum_{\nu=1}^{\mu} \frac{a_\nu}{a} \tag{43.68}$$

or, multiplying by a,

$$\sum_{\nu=1}^{\mu-1} a_\nu < r_2 \cdot a \leq \sum_{\nu=1}^{\mu} a_\nu \quad \text{with } a = \sum_{\nu=1}^{M} a_\nu \text{ and } r_2 \in (0, 1) \tag{43.69}$$

This represents a sum of all transition probabilities, up to an index for which the sum is larger than the random variable r_2. This index is set to be the chosen transition μ. The process to derive μ computationally from a random number is illustrated in Figure 43.11.

The essence of the Monte Carlo method is condensed in the two equations (Equations 43.65 and 43.69) to calculate a pair of random numbers (τ, μ), while the rest of the method is fairly straightforward and can be reduced to only the following few, simple steps, shown in a flowchart in Figure 43.12:

[‡] The term $1 - r$ has been replaced by r. Since we are dealing with probabilities, this is valid for uniform random numbers in the unit interval.

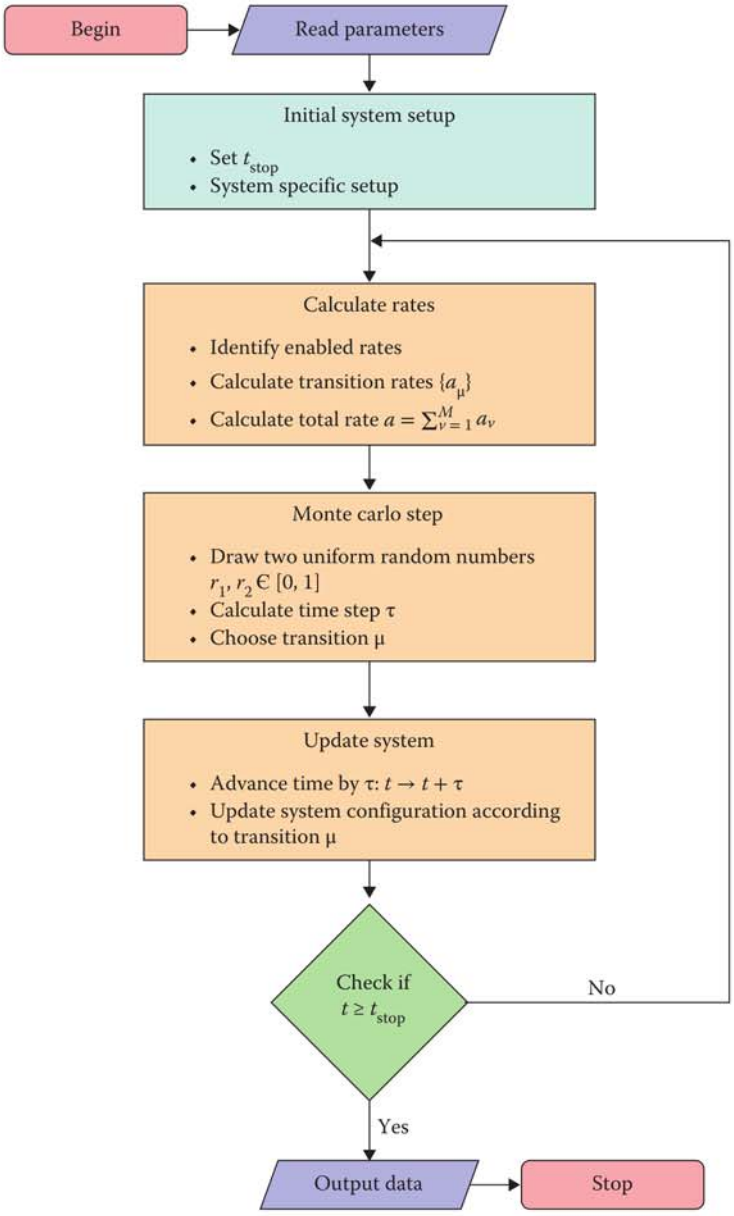


FIGURE 43.12 Flowchart of a general kinetic Monte Carlo algorithm.

0 – Initialization

In the beginning, system-specific parameters are defined in the program. This includes all parameters that must be known in order to set up the system and to calculate the transition rates. Variables to store the occurrence of relevant processes during the simulation are set up here. The simulation time t_{stop} is defined. It determines how long the dynamic evolution should be simulated. The starting time is set to $t = 0$.

1 – Calculation of transition rates

The kMC method is reliant on physical models for all processes that can occur in the system. All physics enters the model in terms of the event rates that are calculated in this step. Depending on the

current state, only some transitions may be activated. Conditions must be met in order to make sure that rates for processes that cannot occur in a certain system state are deactivated. This is realized by setting the value for a deactivated rate $a_\mu = 0$ so that it is not considered in the selection a transition after Equation 43.69. After all rates have been calculated and stored, they are summed up to provide the total rate a.

2 – Monte Carlo step

As described in Section 43.4.2.1, a random pair (τ, μ) is generated according to the PDF $P(\tau, \mu)$ from the list of available (and activated) transitions $\{a_\mu\}$. This accounts for the stochastically correct time step τ for the selection of a transition μ.

3 – Update system

To update the system the time t is advanced by τ. Based on which transition μ has been chosen, the system configuration needs to be updated accordingly. How every update is specifically depends on the physical process.

After updating the system, a check is performed whether the simulation time has reached the specified stopping time t_{stop}. If $t < t_{\text{stop}}$, the simulation will continue by jumping back to step 1 and recalculate the event rates. This way, the system update on the behavior of other processes is reflected. The steps 1–3 represent the so called *Monte Carlo loop*. They will be executed as long as termination condition is not yet fulfilled.

4 – Termination and data output

Through the successive time advance t eventually becomes larger than the stopping time and the Monte Carlo loop is terminated. At last the variables tracking selected events can be evaluated.

Overall, the kMC method offers a straightforward algorithm to simulate the time-dependent evolution of a system. It is only reliant on (1) the assumption that all physical processes can be determined by rate expressions, and (2) that a (pseudo-)random number generator with a sufficiently small correlation between two subsequent numbers and a large period is provided.

43.4.3 Example: Transport through 1D Disordered System by kMC

To illustrate an actual implementation of the kMC algorithm, we provide a short manual on how to model charge transport in a 1D, energetically disordered system, e.g., an organic polymer. We will go through the steps presented in the previous section and in Figure 43.12: the initial system setup, the simulation loop, and the evaluation of quantities.

43.4.3.1 System Setup and Starting Conditions

In the beginning, the available states that the system can occupy must be defined. The setup is illustrated in Figure 43.14. Here, we consider a linear chain of N equidistant nodes indexed by $i = 1, \ldots, N$. The position of each node is quantified by $x_i = i \cdot L$ where L is the fixed distance between all nodes. Then, energy values are assigned to each node accounting for the influence of an applied electric field F, which is assumed to be constant leading to a linear energy drop across the system, and an energetic fluctuation $r(\mu, \sigma)$ drawn from a GDOS distribution with mean μ and standard deviation σ as given in Equation 43.42. This makes up for an energy landscape of the form $E_i = qFx_i + r_{\text{GDOS}}(\mu = 0, \sigma)$. The disorder is picked independenly of the position. A typical configuration for E_i is depicted in Figure 43.14.

Next, initial conditions must be set. The simulation time is set to an initial time $t = 0$. It will be updated during each simulation step. The termination condition is defined by the particle position: if the particle reaches the last node N, the simulation loop stops. We allow one electron of charge $q = -e$ at a time to be in the system. In the beginning it is set to the first node x_1.

A model for the transition rates needs to be defined. We allow the particle to move only to its immediately neighboring sites to the left or to the right. The transitions are described by hopping rates according to the Miller–Abrahams formula (Equation 43.49). There are then only two distinct rates: a_l for hopping to the left and a_r for hopping to the right (Figure 43.13). They depend on the relative energy difference $\Delta E = E_j - E_i$ of the current node i of the particle and the final node $i \pm 1$. Typical values are $a_0 = 10^{11}\ \mathrm{s}^{-1}$, $\gamma = 2\mathrm{nm}$, and $r_{ij} = L = 1$ nm.

43.4.3.2 Simulation Loop

After system setup, the simulation loop is executed until the termination condition is reached. It consists of the following steps:

1. Check the energy landscape to the right and to the left of the current particle position $E_{i\pm1}$.
2. Calculate the hopping rates $a_\mathrm{l}(E_{i-1} - E_i)$ and $a_\mathrm{r}(E_{i+1} - E_i)$.
3. Pick two distinct random numbers r_1, r_2 uniformly from the interval (0, 1).
4. Calculate the total rate $a = a_\mathrm{l} + a_\mathrm{r}$. Then, determine the transition (hopping left or right): If $r_1 \cdot a \leq a_l$, pick a_l. Otherwise, if $a_\mathrm{l} < r_1 \cdot a \leq a_\mathrm{l} + a_\mathrm{r} = a$, pick a_r.
5. Calculate timestep: $\tau = -\frac{r_2}{a}$.
6. Update system according to the transition picked, i.e., change the particle position to the node on the right or on the left.
7. Advance time $t \to t + \tau$.
8. Check termination condition. If still true, go to step 1. Otherwise, exit loop.

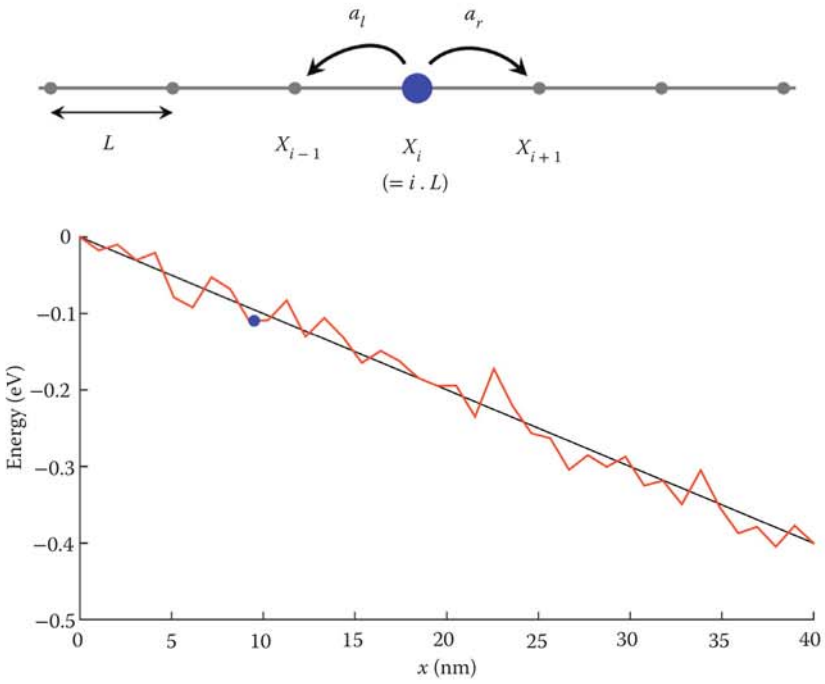


FIGURE 43.13 Setup to model hopping in a 1D system. Top: Spatial discretization of the problem in equidistant nodes. Bottom: Energetic configuration of the nodes. The setup shows the energetic landscape for $N = 40$ nodes comprised of a linear drop of the electric field $F = 10^5$ V/cm and a Gaussian energetic disorder with $\sigma = 30$ meV. The particle is allowed to hop to neighboring nodes under the influence of the energetic configuration.

43.4.3.3 Evaluation

After termination, desired quantities can be evaluated. In this case, since the time that the particle needed to hop through the disordered landscape t_{sim} is now known, the effective velocity $v = \frac{N \cdot L}{t_{sim}}$ can be calculated. From v and the electric field F, the effective transport mobility can be extracted as $\mu = \frac{v}{F}$.

43.4.3.4 Further Info

The general system setup and transition rates can be extended to 2D or 3D (see Figure 43.15). More particle types (holes, excitons) and more effects than hopping, such as injection, recombination, photogeneration of excitons can be introduced to obtain a more realistic model. The principle of choosing transition and timestep stays the same.

43.5 Specific Modeling Issues

In this section, we provide further details on the implementation of models specific to some of the emerging PV technologies, and we cover generic modeling issues.

43.5.1 Dye Solar Cells

DSCs have been extensively simulated using analytical models or numerical ones at the level of drift-diffusion. Most simulations have been performed using 1D models [100–102] and an effective material approximation (see Section 43.5.4 for an extensive description of this approximation). Only few simulations going beyond the effective medium model and trying to include the morphology [103] or to go toward 2D and 3D simulations [104,105] have been reported so far.

The main issues concerning simulations of a DSC are for sure related to the complex morphology of the mesoporous material and the chemical reactions at the interfaces. The latter include both the dye regeneration processes as well as all the recombination events where electrons in the TiO_2 are transferred to reduce ionized dyes or triiodide ions. The main issue in modeling is that all these processes are in general complex electrochemical rates where the charge transfer occurs in several steps. They depend on the effect of the electrolyte, the reorganization energy of the environment, temperature, and voltage [106]. Moreover, different dyes or electrolytes will lead to different kinetics and thus different recombination rates. This leads to recombination rates that formally are completely different from the standard formulas given in Equations 43.34 through 43.36.

Beside that, boundary conditions at the contacts also are very uncommon with respect to more conventional PV devices. While the anode with the ITO or FTO can be approximated by a standard ohmic or Schottky contact, the cathode at the platinum is an electrochemical contact which requires to take into account the formation of an ion accumulation layer. The standard boundary condition is a Butler–Volmer (BV) equation, which is generally a Robin-type boundary condition, since it connects the electron transfer rate with the electrochemical potential at the contact. A BV equation [107] describes the dependency of the electrical current at an electrode j on the electrode potential when both anodic and cathodic reactions are involved:

$$j = j_0 \cdot \left\{ \exp\left[\frac{\alpha_A z e}{k_B T}\eta\right] - \exp\left[\frac{\alpha_C z e}{k_B T}\eta\right] \right\} \tag{43.70}$$

where j_0 is the exchange current density prefactor, $\eta = E - E_{eq}$ the activation overpotential given by the difference in electrode potential and equilibrium potential, α_A and α_C are charge transfer coefficients for anode and cathode, respectively, z is the number of electrons participating in the reaction at the electrode, e

is the elementary charge and $k_B T$ the thermal energy. A Robin-type boundary condition in general is a combination of Dirichlet boundary conditions and Neumann boundary conditions. Imposed on a differential equation of function f, it reads in a general form

$$af + b\frac{\partial f}{\partial n} = g \tag{43.71}$$

where a and b are nonzero constants and g is a function defined on the boundary of the problem. Depending on whether the limiting process is the ionic diffusion of the redox pair or the charge transfer between platinum and triiodide ions, we can have different shapes of the BV equation. Finally, electron transport in the TiO_2 is mediated by trapping and detrapping events, a physical situation captured by the multitrapping model, where the mobility becomes density dependent [108,109]. For solid state DSCs, the simulation issues are more similar to OPVs as the electrolyte is replaced by an organic semiconductor.

43.5.2 Organic Photovoltaics

Modeling of OPV devices incurs several important differences when compared to inorganic devices. First of all, as mentioned already in section 43.4.1, the DOS is fundamentally different as it does not result from electronic states in a periodic potential, but from a both energetically and spatially disordered energy landscape of molecular levels as shown schematically in Figure 43.14. This has to be considered when using or implementing a drift-diffusion based simulation tool, since the standard drift-diffusion model has been derived for carriers with a parabolic DOS. As a further consequence of this energy disorder, transport cannot be described adequately by carriers moving freely in conduction and valence bands, but rather by hopping between energy levels on different molecules [86,110–112], as described in Section 43.4.1. Fortunately, on a macroscopic level, i.e. on a scale much larger than the molecular structure, transport can still be described in terms of drift and diffusion, if a suitable model for the carrier mobilities can be found. Such a model is described in [88], where the mobility has been parametrized based on numerical results of the Master equation.

Due to the different DOSs, the models for transport across interfaces and in particular for injection at the contacts also are different from their counterparts employed in standard drift-diffusion simulations of inorganic semiconductors. The properties of a metal-semiconductor contact, known as Schottky contact in the inorganic case, are controlled by the enetrgetic alignment of the metal work function with the semiconductor band edges, by the semiconductor's Fermi level, presence of interface states, and by the semiconductor DOS. The boundary condition on the particle current e.g. for the electrons at the metal-semiconductor contact can be written as $J_n = qv_R(n - n_0)$, where n_0 is the electron density at the contacts in equilibrium and v_R is a recombination velocity, with units cm/s. The latter quantity depends on the exact

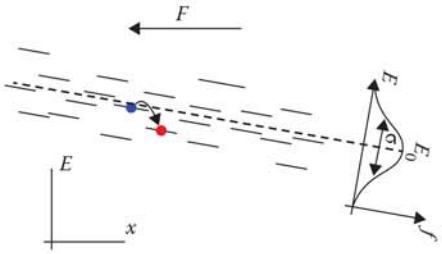


FIGURE 43.14 Schematic representation of the energetic disorder in an organic semiconductor. Each level represents a molecular energy level, which are disordered in energy and in space, following approximately a Gaussian distribution. Transport is driven by hopping between levels, described by energy difference and distance-dependent coefficients that depend also on the external electric field F.

transport mechanism and on the semiconductor DOS, and is therefore very different between inorganic and organic semiconductors [113].

A further difference of OPVs is their low permittivity of $3 \sim 4$ compared to inorganic materials with $\varepsilon_r \approx 10$. This leads to strong local Coulomb interaction which induces the formation of excitons even at room temperature. Since only electrons and holes can be collected at the contacts, the photo-generated excitons need to dissociate first. For this they need to reach the donor/acceptor interface, which is however possible only within their diffusion length of approximately 10 nm. For this reason, OPVs employ a blended donor/acceptor absorber, forming a BHJ. In combination with the hopping transport of charges in spatially and energetically disordered materials, these factors are the key differences to inorganics and make up for most of the challgenges associated with the modeling of OPVs. Since excitonic effects are so important in OPVs, several authors have proposed drift-diffusion-based models including explicitly exciton generation and dissociation [114–116].

The main disadvantage of state-of-the-art drift-diffusion models is the simplification of the BHJ morphology to a 1D effective medium. More details on the background and modeling techniques of the morphology can be found in Section 43.5.4. Since no interfaces exist in the effective medium, all interface effects between the donor/acceptor boundaries, such as exciton splitting and electron–hole recombination, are assumed to be bulk processes and evenly distributed across the photoactive layer. An absorbed photon directly creates and electron–hole pair and exciton diffusion and loss processes cannot be considered. Also, no charge accumulation effects at the donor/acceptor interface can be observed due to the lack of the blend morphology.

Most of these limitations do not apply to kMC simulations. A real 3D blend morphology, explicit treatment of exciton diffusion and splitting, and other interface processes, can easily be implemented. Also, the hopping transport in energetically and spatially disordered materials is naturally representable on a discrete lattice. However, since kMC is a statistical method, it often has to deal with unfeasible simulation times in order to achieve a reasonable accuracy for the desired quantities. This is especially critical when one has to sample very slow processes in presence of very fast events. Specifically, the evaluation of Coulomb interaction is expensive, and since particles are modeled explicitly and not by continuous quantities, the kMC algorithm is very demanding, in particular for a large number of particles. Large particle numbers are induced by a high energetic disorder, small injection barriers ("small barrier problem" in kMC [117]), or large device dimensions/thicknesses and limit the choice of parameters where kMC can be applied to certain ranges. Procedures to parallelize the algorithm and the Coulomb interaction need to be applied [118] to model larger systems and/or get better statistics in the results.

43.5.3 Perovskites

Perovskite, based devices pose two types of issues related to simulation. On the one hand, the fundamental working principle has not been completely unveiled yet, in particular the role of the perovskite in carrier transport. Therefore, it is not clear yet if the perovskite should be used merely as sensitizing absorber similar to the dye in a DSC, or if it should be used in a configuration more similar to an inorganic cell where the absorber acts also as electron and hole transporter. Partly, such questions arise because the development of perovskite PV is driven by the DSC and OPV comunities, with different views on device structure and functioning, and so also the point of view for modeling varies. From a fundamental point of view, being a crystalline material, perovskite seems to be more similar to an inorganic material with well-defined band structure [119], and it would seem that standard simulation models used for inorganic semiconductors should apply.

There are several issues encountered experimentally, where device simulation could give a strong contribution for a better understanding. One open question is whether ferroelectricity and ferroelectric domains exist in perovskite layers and whether they play a role on device functioning. While a fundamental answer to such questions based on density functional theory or MD is not easy, device-level simulation can at least provide inside into the qualitative effects that could be expected in presence of spatially varying

polarization fields. This requires, however, simulation models that can include such effects. In particular, realistic domains have to be generated numerically, and the drift-diffusion model has to include the variable polarization field. Attempts in this direction have been made [120].

Another challenge for modeling is the hysteresis typically observed experimentally when measuring perovskite solar cell current–voltage characteristics [40,121]. Currently, the most probable cause of such hysteresis is believed to be migration of halide ion vacancies [122,123]. However, modeling of this effect in a standard device simulator is not easy, since the simulation model assumes transport of electrons and holes for a given fixed spatial distribution of defects. Therefore, self-consistent modeling of electronic transport and ion migration requires an extension of the standard drift-diffusion model with further mobile species, which is currently not available in standard device simulation software.

43.5.4 Modeling of Morphology and Grain Effects

Many device architectures for OPVs, (solid-state-) DSSCs and perovskite solar cells have regions with complex intermixed morphologies between two different materials where important processes are assumed to be controlled by the shape of the intermixing, especially mobilities through the morphologies and interface processes such as charge recombination and charge generation. For the sake of simplicity and low computational effort, the 3D blend is often simplified and treated as an effective medium, hiding the interface processes and 3D morphology effects in effective models for charge generation and recombination, and the mobility. An effective medium can be defined such that its band edges are given by the lower and higher values of the conduction and valence band edges of the two blended materials, respectively. Parameters like recombination rate constants and mobilities are then usually fitted to reproduce experimental data like current–voltage curves, or experimentally measured mobilities. It is yet to be investigated under which conditions an effective medium treatment is valid and which models are accurate for this description. Alternatively to the effective medium, the mobility can be derived from microscopic simulations considering the 3D blend morphology, like kMC. A comparison between the effective medium approach and a real blend kMC simulation in BHJ OPVs shows, however, that it might be difficult to include all interface effects in a blend into mobility and recombination models within an effective medium [124], especially in low-ε_r materials. A common basis in as many parameters as possible needs to be established for a comparison: the MO levels, the generation profile, and the injection processes, to name a few. By extracting electron and hole mobilities from the 3D simulations, the effect of the blend mobility can effectively be plugged into the effective medium model. It is then tried to obtain an agreement between the two models by tuning the recombination ratio. No general agreement could be obtained, which might indicate that it is not possible to merge all parameters in effective models accurately. It has been observed that, for instance, the charge density distribution within the 3D blend cannot be recovered by an effective model, since local accumulations at an interface are not present in the effective medium where both phases are assumed to be everywhere.

Thus, performing simulations on a realistic 3D blend is important to get a detailed picture about the processes in the blend. For a drift-diffusion model, complex 3D meshes are not easy to generate, increase considerably computation time and are prone to cause problems related to convergence. To avoid these problems for 3D blends, the kMC method is a suitable choice. The kMC algorithm does not have issues with convergence, even for complex morphologies. However, due to the popularity and establishment of drift-diffusion models, more effort should be made to include morphology effects in drift-diffusion-based simulation models. An example in this direction is given, e.g., in [125].

In kMC models, a blend can be created using a spin-exchange algorithm based on the work of Watkins et al. [111] which allows to control the grade of intermixing between two phases. However, the structure is based on a discrete lattice and is not suitable as basis for a DD simulation for which a high-quality finite element mesh is needed. A smoothing algorithm can be applied in order to achieve a mesh suitable for DD simulations. An illustration of a discrete BHJ lattice and a successively smoothed mesh is

depicted in Figure 43.15. Further tests with an implementation of the 3D blend morphology as a mesh for DD simulations need to be performed to increase the common basis of the two models and get more information about where the two models differ and when the effective medium is appropriate.

Apart from nanometer-scale morphology effects, grain boundaries in multicrystalline materials can also affect optoelectronic device properties. Grain dimensions in such materials can range from nanometers to micrometers. A correct description of the effects of the grain boundaries, which can be related to both interface recombination and transport, may require to include the grain interfaces explicitly in the simulation model, especially if there is a limited number of interfaces along the carrier transport path such that an effective medium approximation may become questionable. Grain boundaries in organic materials have been studied related mostly to organic thin-film transistors [126,127]. Similar studies have been performed on polycrystalline silicon [128], and the effect of grain boundaries on carrier collection and recombination has been in CIGS devices has been studied in [129] and [130].

Note that also kMC does not automatically include the effect of grain boundaries, which should be included ad hoc with parameters based on *ab initio* or MD simulations [131].

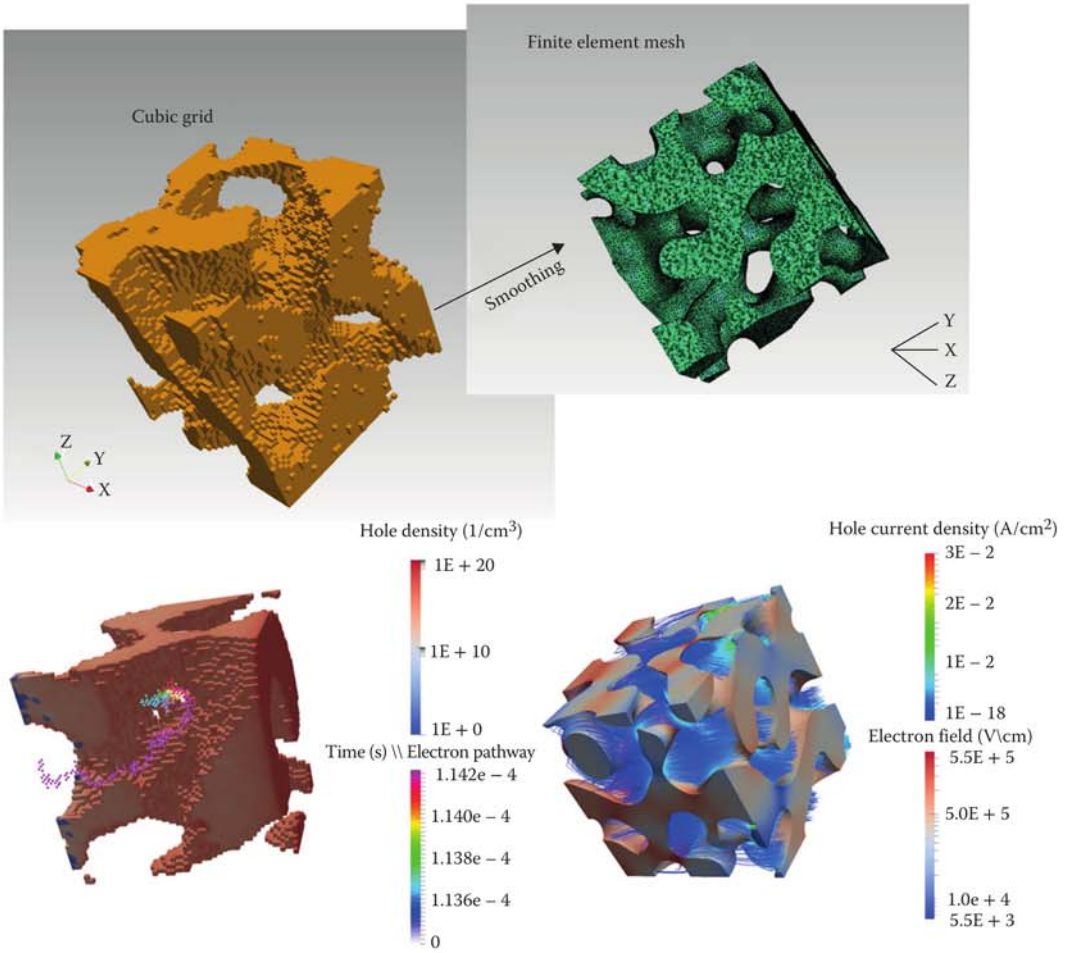


FIGURE 43.15 Smoothing a discrete lattice used for kMC simulations (top left) to obtain a suitable mesh for DD simulations (top right). Exemplary results of the kMC method (bottom left) highlighting the advantage to track single charge carrier trajectories and the DD model to extract, e.g., directed streams such as current densities through the morphology (bottom right).

References

1. J. Nelson. *The Physics of Solar Cells*. London: Imperial College Press, 2003.
2. J. Poortmans and V. Arkhipov, editors. *Thin Film Solar Cells—Fabrication, Characterization and Applications*. Hoboken, NJ: John Wiley & Sons, 2006.
3. K.L. Chopra, P.D. Paulson, and V. Dutta. Thin-film solar cells: An overview. *Progress in Photovoltaics: Research and Applications*, 12(2–3):69–92, 2004.
4. A. Jäger-Waldau. Status and perspectives of thin film photovoltaics. In A. Bosio and A. Romeo, editors, *Thin Film Solar Cells—Current Status and Future Trends*. Hauppauge, NY: Nova Science Publishers, 2011.
5. T.M. Brown, A. Reale, and A. Di Carlo. Organic and hybrid solar cells. In A. Bosio and A. Romeo, editors, *Thin Film Solar Cells—Current Status and Future Trends*. Hauppauge, NY: Nova Science Publishers, 2011.
6. Fraunhofer Institute for Solar Energy (ISE). *Photovoltaics Report*, 2015. Available at http://www.ise.fraunhofer.de/.
7. http://www.nrel.gov/ncpv.
8. M.A. Green. Consolidation of thin-film photovoltaic technology: The coming decade of opportunity. *Progress Photovoltaics: Research and Applications*, 14:383–392, 2006.
9. L. Dou, J. You, Z. Hong, Z. Xu, G. Li, R.A. Street, Y. Yang, L. Dou, J. You, Z. Hong, Z. Xu, G. Li, and Y. Yang. 25th Anniversary Article: A decade of organic/polymeric photovoltaic research. *Advanced Materials*, 25:6642–6671, 2013.
10. W. Cao and J. Xue. Recent progress in organic photovoltaics: Device architecture and optical design. *Energy & Environmental Science*, 7:2123–2144, 2014.
11. S.D. Stranks and H.J. Snaith. Metal-halide perovskites for photovoltaic and lightemitting devices. *Nature Nanotechnology*, 10(5):391–402, 2015.
12. A. Bosio and A. Romeo, editors. *Thin Film Solar Cells—Current Status and Future Trends*. Hauppauge, NY: Nova Science Publishers, 2011.
13. P. Delli Veneri and L.V. Mercaldo. Thin film silicon solar cells. In A. Bosio and A. Romeo, editors, *Thin Film Solar Cells—Current Status and Future Trends*. Hauppauge, NY: Nova Science Publishers, 2011.
14. M. Zeman. Advanced amorphous silicon solar cell technologies. In J. Poortmans and V. Arkhipov, editors, *Thin Film Solar Cells—Fabrication, Characterization and Applications*. Hoboken, NJ: John Wiley & Sons, 2006.
15. M.A. Green, K. Emery, Y. Hishikawa, W. Warta, and E.D. Dunlop. Solar cell efficiency tables (version 46). *Progress in Photovoltaics: Research and Applications*, 23(7):805–812, 2015.
16. H. Fritzsche. Development in understanding and controlling the Staebler-Wronski effect in a-Si:H. *Annual Review of Materials Science*, 31:47–79, 2001.
17. E.A. Schiff, S. Hegedus, and X. Deng. Amorphous silicon-based solar cells. In A. Luque and S. Hegedus, editors, *Handbook of Photovoltaic Science and Engineering*, pp. 487–545. Hoboken, NJ: John Wiley and Sons, 2011.
18. A. Bosio, A. Romeo, D. Menossi, S. Mazzamuto, and N. Romeo. Review: The secondgeneration of CdTe and CuInGaSe2 thin film PV modules. *Crystal Research and Technology*, 46(8):857–864, 2011.
19. S. Girish Kumar and K.S.R. Koteswara Rao. Physics and chemistry of CdTe/CdS thin film heterojunction photovoltaic devices: Fundamental and critical aspects. *Energy and Environmental Science*, 7:45–102, 2014.
20. R. Klenk and M.C. Lux-Steiner. Chalcopyrite based solar cells. In J. Poortmans and V. Arkhipov, editors, *Thin Film Solar Cells—Fabrication, Characterization and Applications*. Hoboken, NJ: John Wiley & Sons, 2006.

21. P. Reinhard, A. Chirila, P. Blosch, F. Pianezzi, S. Nishiwaki, S. Buechelers, and A.N. Tiwari. Review of progress toward 20% efficiency exible cigs solar cells and manufacturing issues of solar modules. *IEEE Journal of Photovoltaics,* 3(1):572–580, 2013.
22. T. Unold and C.A. Kaufmann. *Chalcopyrite Thin-Film Materials and Solar Cells*, Volume 1. Amsterdam: Elsevier, 2012.
23. B. O'Regan and M. Grätzel. A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO2 films. *Nature*, 353:737–740, 1991.
24. K. Kalyanasundaram, editor. *Dye Sensitized Solar Cells*. Lausanne: EPFL Press, 2010.
25. S. Mathew, A. Yella, P. Gao, R. Humphry-Baker, B.F.E. Curchod, N. Ashari-Astani, I. Tavernelli, U. Rothlisberger, M.K. Nazeeruddin, and M. Grtzel. Dye-sensitized solar cells with 13% efficiency achieved through the molecular engineering of porphyrin sensitizers. *Nature Chemistry*, 6(3):242–247, 2014.
26. U. Bach, D. Lupo, P. Comte, J.E. Moser, F. Weissrtel, J. Salbeck, H. Spreitzer, and M. Grtzel. Solid-state dye-sensitized mesoporous TiO_2 solar cells with high photon-to-electron conversion efficiencies. *Nature*, 395(6702):583–585, 1998.
27. F. Matteocci, G. Mincuzzi, F. Giordano, A. Capasso, E. Artuso, C. Barolo, G. Viscardi, T.M. Brown, A. Reale, and A. Di Carlo. Blocking layer optimisation of poly(3-hexylthiopene) based solid state dye sensitized solar cells. *Organic Electronics*, 14(7):1882–1890, 2013.
28. A. Hagfeldt, G. Boschloo, L. Sun, L. Kloo, and H. Pettersson. Dye-sensitized solar cells. *Chemical Reviews*, 110:6595–6663, 2010.
29. A. Hinsch. Status of the dye solar cell technology (DSC) as a guideline for further research. In *28th European PV Solar Energy Conference and Exhibition*, Paris, France, 2013.
30. N.S. Sariciftci, L. Smilowitz, A.J. Heeger, and F. Wudl. Photoinduced electron transfer from a conducting polymer to buckminsterfullerene. *Science*, 258(5087):1474–1476, 1992.
31. C.W. Tang. Two-layer organic photovoltaic cell. *Applied Physics Letters*, 48(2):183–185, 1986.
32. G. Yu, J. Gao, J.C. Hummelen, F. Wudl, and A.J. Heeger. Polymer photovoltaic cells: Enhanced efficiencies via a network of internal donor-acceptor heterojunctions. *Science*, 270(5243):1789–1790, 1995.
33. Y. Liu, J. Zhao, Z. Li, C. Mu, W. Ma, H. Hu, K. Jiang, H. Lin, H. Ade, and H. Yan. Aggregation and morphology control enables multiple cases of high-efficiency polymer solar cells. *Nature Communications*, 5:5293, 2014.
34. C. Deibel and V. Dyakonov. Polymer-fullerene bulk heterojunction solar cells. *Reports on Progress in Physics*, 73(9):096401, 2010.
35. G. Li, R. Zhu, and Y. Yang. Polymer solar cells. *Nature Photonics*, 6(3):153–161, 2012.
36. H.-S. Kim, C.-R. Lee, J.-H. Im, K.-B. Lee, T. Moehl, A. Marchioro, S.-J. Moon, R. Humphry-Baker, J.-H. Yum, J.E. Moser, M. Grtzel, and N.-G. Park. Lead iodide perovskite sensitized all-solid-state submicron thin film mesoscopic solar cell with efficiency exceeding 9%. *Scientific Reports*, 2:591, 2012.
37. H.S. Jung and N.-G. Park. to Perovskite solar cells: From materials devices. *Small*, 11(1):10–25, 2015.
38. P. Gao, M. Grätzel, and M.K. Nazeeruddin. Organohalide lead perovskites for photovoltaic applications. *Energy and Environmental Science*, 7(8):2448–2463, 2014.
39. D. Wang, M. Wright, N.K. Elumalai, and A. Uddin. Stability of perovskite solar cells. *Solar Energy Materials and Solar Cells*, 147:255–275, 2016.
40. M. Grätzel. The light and shade of perovskite solar cells. *Nature Materials*, 13(9):838–842, 2014.
41. P.P. Boix, K. Nonomura, N. Mathews, and S.G. Mhaisalkar. Current progress and future perspectives for organic/inorganic perovskite solar cells. *Materials Today*, 17(1):16–23, 2014.
42. P. Würfel. *Physics of Solar Cells*. Weinheim: Wiley-VCH Verlag, 2009.
43. G.P. Smestad. *Optoelectronics of Solar Cells*. SPIE Press, 2002.

44. N.-K. Persson and O. Inganäs. Simulations of optical processes in organic photovoltaic devices. In S.-S. Sung and N.S. Sariciftci, editors, *Organic Photovoltaics—Mechanisms, Materials and Devices*. Hoboken, NJ: Taylor & Francis, 2005.
45. S.L. Chuang. *Physics of Optoelectronic Devices*. *Wiley Series in Pure and Applied Optics*, 1st edition. Hoboken, NJ: John Wiley & Sons, 1995.
46. A. Taove and S.C. Hagness. *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House Publishers, 2005.
47. J. Jin. *The Finite Element Method in Electromagnetics*. Hoboken, NJ: Wiley, 2002.
48. I. Semenikhin, M. Zanuccoli, M. Benzi, V. Vyurkov, E. Sangiorgi, and C. Fiegna. Computational efficient RCWA method for simulation of thin film solar cells. *Optical and Quantum Electronics*, 44(3):149–154, 2012.
49. M. Born and E. Wolf. *Principles of Optics*. New York, NY: Cambridge University Press, 1999.
50. B. Harbecke. Coherent and incoherent reflection and transmission of multilayer structures. *Applied Physics B Photophysics and Laser Chemistry*, 39(3):165–170, 1986.
51. J.J.M. Halls, K. Pichler, R.H. Friend, S.C. Moratti, and A.B. Holmes. Exciton diffusion and dissociation in a poly(*p*-phenylenevinylene)/C_{60} heterojunction photovoltaic cell. *Applied Physics Letters*, 68(22):3120–3122, 1996.
52. J.J.M. Halls. *Photoconductive Properties of Conjugated Polymers*. PhD dissertation, Cambridge: St. John's College, 1997.
53. L.A.A. Pettersson, L.S. Roman, and O. Inganäs. Modeling photocurrent action spectra of photovoltaic devices based on organic thin films. *Journal of Applied Physics*, 86(1):487–496, 1999.
54. L.A.A. Pettersson, L.S. Roman, and O. Inganäs. Quantum efficiency of exciton-to-charge generation in organic photovoltaic devices. *Journal of Applied Physics*, 89(10):5564–5569, 2001.
55. N.-K. Persson, M. Schubert, and O. Inganäs. Optical modelling of a layered photovoltaic device with a polyuorene derivative/fullerene as the active layer. *Solar Energy Materials and Solar Cells*, 83(23):169–186, 2004.
56. H. Hoppe, N. Arnold, N.S. Sariciftci, and D. Meissner. Modeling the optical absorption within conjugated polymer/fullerene-based bulk-heterojunction organic solar cells. *Solar Energy Materials and Solar Cells*, 80(1):105–113, 2003.
57. D.P. Gruber, G. Meinhardt, and W. Papousek. Spatial distribution of light absorption in organic photovoltaic devices. *Solar Energy*, 79(6):697–704, 2005.
58. F. Monestier, J.-J. Simon, P. Torchio, L. Escoubas, F Flory, S. Bailly, R. de Bettignies, S. Guillerez, and C. Defranoux. Modeling the short-circuit current density of polymer solar cells based on p3HT:PCBM blend. *Solar Energy Materials and Solar Cells*, 91(5):405–410, 2007.
59. R. Husermann, E. Knapp, M. Moos, N.A. Reinke, T. Flatz, and B. Ruhstaller. Coupled optoelectronic simulation of organic bulk-heterojunction solar cells: Parameter extraction and sensitivity analysis. *Journal of Applied Physics*, 106(10):104507, 2009.
60. G.F. Burkhard, E.T. Hoke, and M.D. McGehee. Accounting for interference, scattering, and electrode absorption to make accurate internal quantum efficiency measurements in organic and other thin solar cells. *Advanced Materials*, 22(30):3293–3297, 2010.
61. L.D. Landau, E.M. Lifshitz, and L.P. Pitaevskii. *Electrodynamics of Continuous Media*. Amsterdam: Elsevier, 1993.
62. H.-Y.D. Yang. A spectral recursive transformation method for electromagnetic waves in generalized anisotropic layered media. *IEEE Transactions on Antennas and Propagation,* 45(3):520–526, 1997.
63. J. Ning and E.L. Tan. Hybrid matrix method for stable analysis of electromagnetic waves in stratified bianisotropic media. *IEEE Microwave and Wireless Components Letters,* 18(10):653–655, 2008.
64. A.H. Fallahpour, A. Gagliardi, F. Santoni, D. Gentilini, A. Zampetti, M. Auf der Maur, and A. Di Carlo. Modeling and simulation of energetically disordered organic solar cells. *Journal of Applied Physics*, 116(18), 2014.

65. M. Zeman, O. Isabella, K. Jger, R. Santbergen, S. Solntsev, M. Topic, and J. Krc. Advanced light management approaches for thin-film silicon solar cells. *Energy Procedia*, 15:189–199, 2012.
66. S. In, D.R. Mason, H. Lee, M. Jung, C. Lee, and N. Park. Enhanced light trapping and power conversion efficiency in ultrathin plasmonic organic solar cells: A coupled optical-electrical multiphysics study on the effect of nanoparticle geometry. *ACS Photonics*, 2(1):78–85, 2015.
67. M. Peters, B. Bläsi, S.W. Glunz, A.G. Aberle, J. Luther, and C. Battaglia. Optical simulation of silicon thin-film solar cells. *Energy Procedia*, 15:212–219, 2012.
68. A.H. Fallahpour, G. Ulisse, M. Auf der Maur, A. Di Carlo, and F. Brunetti. 3-D simulation and optimization of organic solar cell with periodic back contact grating electrode. *IEEE Journal of Photovoltaics*, 5(2):591–596, 2015.
69. M. Zeman, O. Isabella, S. Solntsev, and K. Jäger. Modelling of thin-film silicon solar cells. *Solar Energy Materials and Solar Cells*, 119:94–111, 2013.
70. N. Tucher, J. Eisenlohr, P. Kiefel, O. Höhn, H. Hauser, M. Peters, C. Müller, J.C. Goldschmidt, and B. Bläsi. 3D optical simulation formalism optos for textured silicon solar cells. *Optics Express*, 23(24):A1720–A1734, 2015.
71. S. Selberherr. *Analysis and Simulation of Semiconductor Devices*, 1st edition. New York, NY: Springer-Verlag Wien, 1984.
72. P.A. Markowich, C.A. Ringhofer, and C. Schmeiser. *Semiconductor Equations*, 1st edition. New York, NY: Springer-Verlag Wien, 1990.
73. W.V. van Roosbroeck. Theory of the flow of electrons and holes in germanium and other semiconductors. *Bell System Technical Journal*, 29:560–607, 1955.
74. S.M. Sze. *Semiconductor Devices: Physics and Technology*. New York, NY: John Wiley & Sons, 1985.
75. A. Schenk. An improved approach to the Shockley-Read-Hall recombination in inhomogeneous fields of spacecharge regions. *Journal of Applied Physics*, 71(7):3339–3349, 1992.
76. T. Goudon, V. Miljanovic, and C. Schmeiser. On the Shockley-Read-Hall model: Generation-recombination in semiconductors. *SIAM Journal on Applied Mathematics*, 67(4):1183–1201, 2007.
77. E. O'Reilly. *Quantum Theory of Solids*. London: Taylor & Francis, 2002.
78. S. John, C. Soukoulis, M.H. Cohen, and E.N. Economou. Theory of electron band tails and the urbach optical-absorption edge. *Physical Review Letters*, 57:1777–1780, 1986.
79. H. Bässler. Charge transport in disordered organic photoconductors: A Monte Carlo simulation study. *Physica Status Solidi (B)*, 175(1):15–56, 1993.
80. L.J.A. Koster, V.D. Mihailetchi, R. Ramaker, and P.W.M. Blom. Light intensity dependence of open-circuit voltage of polymer: Fullerene solar cells. *Applied Physics Letters*, 86(12):123509, 2005.
81. G. Paasch and S. Scheinert. Charge carrier density of organics with Gaussian density of states: Analytical approximation for the Gauss-Fermi integral. *Journal of Applied Physics*, 107(10):104501, 2010.
82. A. Zampetti, A.H. Fallahpour, M. Dianetti, L. Salamandra, F. Santoni, A. Gagliardi, M. Auf der Maur, F. Brunetti, A. Reale, T.M. Brown, and A. Di Carlo. Influence of the interface material layers and semiconductor energetic disorder on the open circuit voltage in polymer solar cells. *Journal of Polymer Science Part B: Polymer Physics*, 53(10):690–699, 2015.
83. G.K. Wachutka. Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling. *IEEE Transactions on Computer-Aided Design*, 11:1141–1149, 1990.
84. N.A. Zakhleniuk. Nonequilibrium drift-diffusion transport in semiconductors in presence of strong inhomogeneous electric fields. *Applied Physics Letters*, 89(25):252112, 2006.
85. N.F. Mott. On the transition to metallic conduction in semiconductors. *Canadian Journal of Physics*, 34(12A):1356–1368, 1956.
86. A. Miller and E. Abrahams. Impurity conduction at low concentrations. *Physical Review*, 120(3):745, 1960.
87. R.A. Marcus. Chemical and electrochemical electron-transfer theory. *Annual Review of Physical Chemistry*, 15(1):155–196, 1964.

88. W.F. Pasveer, J. Cottaar, C. Tanase, R. Coehoorn, P.A. Bobbert, P.W.M. Blom, D.M. De Leeuw, and M.A.J. Michels. Unified description of charge-carrier mobilities in disordered semiconducting polymers. *Physical Review Letters*, 94(20):206601, 2005.
89. D.T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976.
90. A.F. Voter. Introduction to the kinetic Monte Carlo Method. *Radiation Effects*, 235:1–23, 2005.
91. H.L. Anderson. Scientific uses of the maniac. *Journal of Statistical Physics*, 43(5–6):731–748, 1986.
92. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087, 1953.
93. J.R. Beeler (Jr). Displacement spikes in cubic metals. i. α-Iron, copper, and tungsten. *Physical Review*, 150(2):470, 1966.
94. E.S. Hood, B.H. Toby, and W.H. Weinberg. Precursor-mediated molecular chemisorption and thermal desorption: The interrelationships among energetics, kinetics, and adsorbate lattice structure. *Physical Review Letters*, 55(22):2437, 1985.
95. A.B. Bortz, M.H. Kalos, and J.L. Lebowitz. A new algorithm for Monte Carlo simulation of Ising spin systems. *Journal of Computational Physics*, 17(1):10–18, 1975.
96. C. Jacoboni and P. Lugli. *The Monte Carlo Method for Semiconductor Device Simulation. Computational Microelectronics*. Vienna: Springer, 1989.
97. D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
98. J.R. Norris. *Markov Chains*. Number 2008. Cambridge: Cambridge University Press, 1998.
99. F. Schwabl. *Statistical Mechanics. Advanced Texts in Physics*. Berlin: Springer, 2006.
100. J. Ferber, R. Stangl, and J. Luther. An electrical model of the dye sensitized solar cell. *Solar Energy Materials and Solar Cells*, 53:29–54, 1998.
101. P.R.F. Barnes and B.C. O'Regan. Electron recombination kinetics and the analysis of collection efficiency and diffusion length measurements in dye sensitized solar cells. *The Journal of Physical Chemistry C*, 114(44):19134–19140, 2010.
102. A. Gagliardi, S. Mastroianni, D. Gentilini, F. Giordano, A. Reale, T.M. Brown, and A.D. Carlo. Multiscale modeling of dye solar cells and comparison with experimental data. *IEEE Journal of Selected Topics in Quantum Electronics,* 16(6):1611–1618, 2010.
103. A. Gagliardi, M. Auf der Maur, D. Gentilini, F. Di Fonzo, A. Abrusci, H.J. Snaith, G. Divitini, C. Ducati, and A. Di Carlo. The real TiO_2/HTM interface of solid-state dye solar cells: Role of trapped states from a multiscale modelling perspective. *Nanoscale*, 7(3):1136–1144, 2015.
104. A. Gagliardi, M. Auf der Maur, D. Gentilini, and A. Di Carlo. Simulation of dye solar cells: Through and beyond one dimension. *Journal of Computational Electronics*, 10(4):424–436, 2011.
105. A. Gagliardi, M. Auf der Maur, and A. Di Carlo. Theoretical investigation of a dye solar cell wrapped around an optical fiber. *IEEE Journal of Quantum Electronics*, 47(9):1214–1221, 2011.
106. J. Nissfolk, K. Fredin, A. Hagfeldt, and G. Boschloo. Recombination and transport processes in dye-sensitized solar cells investigated under working conditions. *The Journal of Physical Chemistry B*, 110(36):17715–17718, 2006.
107. U. Landau. Current distribution in electrochemical cells: Analytical and numerical modeling. In *Modern Aspects of Electrochemistry No. 44: Modelling and Numerical Simulations II*, pp. 451–501. New York, NY: Springer New York, 2009.
108. J. Bisquert and V.S. Vikhrenko. Interpretation of the time constants measured by kinetic techniques in nanostructured semiconductor electrodes and dye-sensitized solar cells. *Journal of Physical Chemistry B*, 108(7):2313–2322, 2004.
109. D. Gentilini, A. Gagliardi, M. Auf der Maur, L. Vesce, D. D'Ercole, T.M. Brown, A. Reale, and A. Di Carlo. Correlation between cell performance and physical transport parameters in dye solar cells. *Journal of Physical Chemistry C*, 116(1):1151–1157, 2012.

110. S.D. Baranovskii. Theoretical description of charge transport in disordered organic semiconductors. *Physica Status Solidi (B)*, 251(3):487–525, 2014.
111. P.K. Watkins, A.B. Walker, and G.L.B. Verschoor. Dynamical Monte Carlo modelling of organic solar cells: The dependence of internal quantum efficiency on morphology. *Nano Letters*, 5(9):1814–1818, 2005.
112. M. Casalegno, G. Raos, and R. Po. Methodological assessment of kinetic Monte Carlo simulations of organic photovoltaic devices: The treatment of electrostatic interactions. *Journal of Chemical Physics*, 132(9):094705, 2010.
113. F. Santoni, A. Gagliardi, M. Auf der Maur, and A. Di Carlo. The relevance of correct injection model to simulate electrical properties of organic semiconductors. *Organic Electronics*, 15(7):1557–1570, 2014.
114. I. Kamohara, M. Townsend, and B. Cottle. Simulation of heterojunction organic thin film devices and exciton diffusion analysis in stacked-hetero device. *Journal of Applied Physics*, 97(1):014501, 2005.
115. Il-S. Park, S.-R. Park, D.-Y. Shin, J.-S. Oh, W.-J. Song, and J.-H. Yoon. Modeling and simulation of electronic and excitonic emission properties in organic host/guest systems. *Organic Electronics*, 11(2):218–226, 2010.
116. B. Ruhstaller, T. Beierlein, H. Riel, S. Karg, J.C. Scott, and W. Riess. Simulating electronic and optical processes in multilayer organic light-emitting devices. *IEEE Journal of Selected Topics in Quantum Electronics*, 9(3):723–731, 2003.
117. D. Kipp and V. Ganesan. A kinetic Monte Carlo model with improved charge injection model for the photocurrent characteristics of organic solar cells. *Journal of Applied Physics*, 113(23):234502, 2013.
118. N.J. van der Kaap and L.J.A. Koster. Massively parallel kinetic monte carlo simulations of charge carrier transport in organic semiconductors. *Journal of Computational Physics*, 307:321–332, 2016.
119. L. Pedesseau, J.-M. Jancu, A. Rolland, E. Deleporte, C. Katan, and J. Even. Electronic properties of 2D and 3D hybrid organic/inorganic perovskites for optoelectronic and photovoltaic applications. *Optical and Quantum Electronics*, 46(10):1225–1232, 2014.
120. A. Pecchia, D. Gentilini, D. Rossi, M. Auf der Maur, and A. Di Carlo. Role of ferroelectric nanodomains in the transport properties of perovskite solar cells. *Nano Letters*, 16(2):988–992, 2016.
121. D. Liu, M.K. Gangishetty, and T.L. Kelly. Effect of CH3NH3PbI3 thickness on device efficiency in planar heterojunction perovskite solar cells. *Journal of Materials Chemistry A*, 2:19873–19881, 2014.
122. S. Meloni, T. Moehl, W. Tress, M. Franckeviius, M. Saliba, Y.H. Lee, P. Gao, M.K. Nazeeruddin, S.M. Zakeeruddin, U. Rothlisberger, and M. Graetzel. Ionic polarization-induced current-voltage hysteresis in $CH_3NH_3PbX_3$ perovskite solar cells. *Nature Communications*, 7, 2016, Article number: 10334. doi:10.1038/ncomms10334.
123. J.M. Azpiroz, E. Mosconi, J. Bisquert, and F. De Angelis. Defect migration in methylammonium lead iodide and its role in perovskite solar cell operation. *Energy and Environmental Science*, 8(7):2118–2127, 2015.
124. T. Albes, P. Lugli, and A. Gagliardi. Investigation of the blend morphology in bulk-heterojunction organic solar cells. *IEEE Transactions on Nanotechnology*, 15(2):281–288, 2016.
125. H.K. Kodali and B. Ganapathysubramanian. A computational framework to investigate charge transport in heterogeneous organic photovoltaic devices. *Computer Methods in Applied Mechanics and Engineering*, 247:113–129, 2012.
126. A. Bolognesi, M. Berliocchi, M. Manenti, A. Di Carlo, P. Lugli, K. Lmimouni, and C. Dufour. Effects of grain boundaries, field-dependent mobility, and interface trap states on the electrical characteristics of pentacene TFT. *IEEE Transactions on Electron Devices*, 51(12):1997–2003, 2004.
127. Y. Hu, L. Wang, Q. Qi, D. Li, and C. Jiang. Charge transport model based on single-layered grains and grain boundaries for polycrystalline pentacene thin-film transistors. *Journal of Physical Chemistry C*, 115(47):23568–23573, 2011.

128. M. Kimura, S. Inoue, T. Shimoda, and T. Sameshima. Device simulation of carrier transport through grain boundaries in lightly doped polysilicon films and dependence on dopant density. *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers*, 40(9A):5237–5243, 2001.
129. W.K. Metzger and M. Gloeckler. The impact of charged grain boundaries on thin-film solar cells and characterization. *Journal of Applied Physics*, 98(6):063701, 2005.
130. K. Taretto and U. Rau. Numerical simulation of carrier collection and recombination at grain boundaries in Cu(In,Ga)Se_2 solar cells. *Journal of Applied Physics*, 103(9), 2008.
131. H. Kobayashi and Y. Tokita. Modeling of hole transport across grain boundaries in organic semiconductors for mesoscale simulations. *Applied Physics Express*, 8(5):051602, 2015.

44

Electroluminescent Refrigerators*

Kuan-Chen Lee

and

Shun-Tung Yen

44.1 Introduction 541
44.2 Theoretical Methods 542
Fundamentals of EL Refrigeration • Device Structure and Main Equations • Recombination and Generation Processes • Boundary Conditions • Output Quantities and Models for Steady-State Analysis
44.3 Results and Discussion 549
Influence of Device Structures on EL Refrigeration • Influence of Optical Processes on EL Refrigeration
44.4 Summary 557

44.1 Introduction

Solid-state optical refrigeration, which is featured by carrying the internal thermal energy away through light emission, has progressed rapidly after the first demonstration of photoluminescent (PL) refrigeration in 1995 (Epstein et al., 1995). To date, it has been shown that an Yb-doped $LiYF_4$ cooler can reach a temperature of 114 K and give a cooling power of 750 mW at room temperature (Melgaard et al., 2014). Further improving this kind of rare-earth-doped coolers to an even lower cooling temperature will be a difficult task. This is because the cooling process involves the thermalization of dopants and becomes inefficient when the thermal energy is comparable to the energy difference between discrete dopant levels in the ground-state manifold (Sheik-Bahae and Epstein, 2007). Different to the doped coolers, a semiconductor luminescent refrigerator is not restricted by such limitation and is expected to have lower operation temperature and higher cooling power. Moreover, it can be directly integrated with other semiconductor devices. These attractive features have stimulated extensive research on semiconductor luminescent refrigeration (Sheik-Bahae and Epstein, 2007, 2009). However, the required external efficiency for cooling is as high as nearly unity such that the realization of semiconductor PL refrigeration was not achieved until recently by Zhang et al. (2013) on CdS mircorod.

Semiconductor luminescent refrigeration using electrical pumping scheme, on the other hand, receives growing attention because of its looser requirement on the external efficiency compared to semiconductor PL refrigeration. Despite several works on the feasibility of electroluminescent (EL) refrigeration several decades ago (Berdahl, 1985; Dousmanis et al., 1964), the experimental verification of EL cooling has been reported just recently. In 2012, Santhanam et al. (2012) biased a light-emitting diode (LED) under several tens of microvolts and showed that the output emission power can exceed twice of the applied electrical

* This chapter is reproduced with permission from Yen and Lee (2010), and Lee and Yen (2012).

input power. This positive difference of the output over the input powers implies that the LED is capable of extracting thermal energy from the environments, as a refrigerator, to keep heat flow balanced. However, the cooling power is smaller than nanowatts so that the cooling phenomenon is not directly observed. It has been known for a while that the cooling power density of an EL cooler can be significantly high if the cooler is biased at a voltage of $\sim E_g/q$, where E_g is the bandgap energy of the active layer and q the elementary charge. Extensive studies have been made to investigate the cooling feasibility and capability of EL refrigeration in this bias region. Mal'shukov and Chao studied the influence of the Auger recombination on the cooling capability of a GaAs-based refrigerator. They showed that a net cooling power density of several watts per square centimeter is achievable (Mal'shukov and Chao, 2001). Wang et al. performed a detailed self-consistent calculation with various recombination mechanisms accounted. Their analysis showed an impressive cooling efficiency of 35% for a cooler of p-GaAs active layer (Wang et al., 2005). Oksanen and Tulkki (2010) proposed a thermal heat pump, which consists of an EL cooler integrated directly with an absorbing diode, to deal with the problem of low light extraction efficiency encountered in semiconductor luminescent refrigeration. Other theoretical works, such as on heat transfer and cooling temperature (Han et al., 2007; Heikkilä et al., 2010; Lin et al., 2012; Piprek and Li, 2016), the design of refrigerator structure (Oksanen and Tulkki, 2010; Yen and Lee, 2010), the influence of the active materials (Yu et al., 2007), light extraction efficiency (Yen and Lee, 2010), and the photon recycling effect (Heikkilä et al., 2009; Lee and Yen, 2012), has also been conducted to provide a deep insight into EL refrigeration. Experimental works on the feasibility (Xue et al., 2015) and the application (Liu et al., 2013) of EL refrigeration are also being conducted. Although the high-cooling-power EL refrigeration has not been achieved, a thorough understanding of its operation mechanism still plays a necessary role in its realization.

In the following sections, we begin with the fundamental concept of EL refrigeration. The self-consistent calculation generally used for analyzing EL refrigeration and the operation principle of EL refrigeration is introduced in Section 44.2. Calculated results together with the analysis of various structural and optical effects on EL refrigeration are given in Section 44.3. Finally, a conclusion is given in Section 44.4.

44.2 Theoretical Methods

44.2.1 Fundamentals of EL Refrigeration

In EL refrigeration, electrons and holes are mediated as refrigerants and are intended to be injected into the active region for radiative recombination, which generates photons with thermal energy carried away. Recently, studies indicate that the extracted thermal energy comes from the thermalization of carriers during transport, similar to the Peltier effect in thermoelectric (TE) cooling (Santhanam et al., 2012; Oksanen and Tulkki, 2015; Xue et al., 2015). This process can be seen during the cooling cycle of EL refrigeration as illustrated in Figure 44.1a. At terminal contacts, carriers are injected into the device under an external bias V and flow all the way to the active layer. In the process of carrier transport, local cooling occurs in the place where the external bias is not large enough for carriers to overcome the potential barriers. This situation can be seen, for instance, in the cladding layer near the active layer. In such regions, only carriers of sufficiently high energy can surmount the barrier and move forward as shown in the inset circle. This causes high-energy empty states that are then occupied by carriers through thermalization, i.e., drawing phonons from the lattice matrix. The redistribution of carriers progresses rapidly and hence the quasi-thermal equilibrium is achieved as a result of much short phonon relaxation time. In the active layer, the injected electron–hole pairs recombine through various mechanisms. A heating energy of qV is generated if the electron–hole pair recombines nonradiatively. Contrary, for each photon generated from radiative recombination escaping out of the device, the internal energy of the whole system is reduced by $\hbar\omega - qV$. Here, $\hbar\omega$ is the photon energy, which comes from the combination of the thermal energy from the lattice matrix and the electric potential energy from the external bias. These three processes, i.e., electrical excitation from external bias, thermalization during transport, and relaxation through radiative recombination, complete the cooling cycle of EL refrigeration. Figure 44.1b shows the system view of an ideal EL cooler

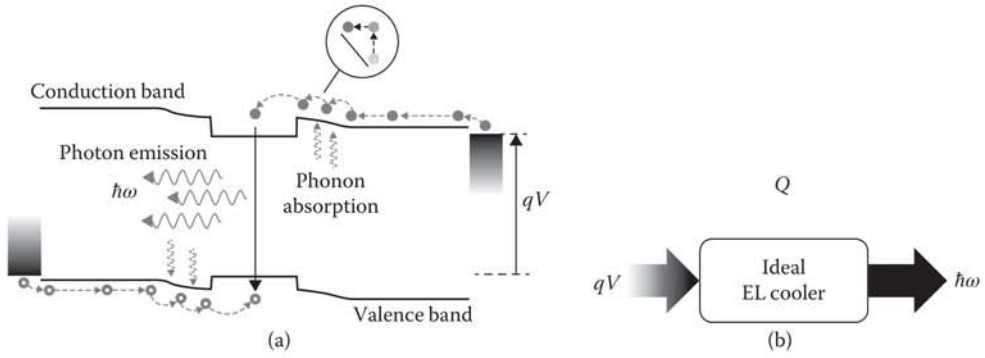


FIGURE 44.1 (a) The illustration of cooling cycling for EL refrigeration. The inset circle shows the thermalization process of carriers, which happens in the cladding layer near the active region. This is similar to the Peltier effect in thermoelectric cooling and attributed to EL refrigeration. (b) The system view of an ideal EL cooler. Net cooling occurs when $Q > 0$; otherwise, the device generates heat.

for a single cooling cycle, in which the parasitic effects are omitted. In this figure, an energy flow per cycle $Q = \hbar\omega - qV$ is required to be drawn from the ambient environment for compensating the internal energy loss of the system. If $Q > 0$, heat flows into the device and net cooling occurs; otherwise, the device generates heat. In this chapter, the cooling power, one of the most important parameters for coolers, is obtained in a similar way as Q described earlier, for which the input and the output powers are calculated from a self-consistent calculation based on one-dimensional drift-diffusion model with various recombination mechanisms accounted. In such a picture, the cooler is assumed to be in direct contact with a heat reservoir and the thermal resistance inside the cooler is neglected so that the temperature of the whole device is kept at a constant value of 300 K. In reality, a temperature gradient between the cooler and its environment exists due to a finite thermal conductance. The temperature distribution may also be nonuniform inside the cooler because of a nonuniform distribution of TE heat exchange (Yu et al., 2007). To account these effects, the heat balance equation between the cooler and its environment and the heat transfer equation inside the cooler have to be considered. Interested readers can refer to the works of Ashcroft and Mermin (1976), Han et al. (2007), Heikkilä et al. (2010), Lin et al. (2012), and Piprek and Li (2016) for more information. The geometric effects, such as current spreading and the detailed optical processes, are not accounted either. The consideration of these effects relies on the detailed information about the device geometry and a full three-dimensional simulation, which is beyond the scope of this chapter. The goal of this chapter is to provide a comprehensive view on how various structural and optical components affect the performance of an EL cooler without losing accuracy.

44.2.2 Device Structure and Main Equations

The device structure used in this work is composed of seven layers, as shown in Figure 44.2. They are, in order, a 100-nm p-doped $Al_{0.25}Ga_{0.75}As$ cladding layer (with an acceptor concentration of $N_a = 1 \times 10^{18}$ cm^{-3}), a 50-nm p-doped $Al_xGa_{1-x}As$ carrier blocking layer ($N_a = 1 \times 10^{18}$ cm^{-3}), a 50-nm undoped $Al_{0.25}Ga_{0.75}As$ spacer, an undoped GaAs active layer of thickness $L = 100$ nm, a 50-nm undoped $Al_{0.25}Ga_{0.75}As$ spacer layer, a 50-nm n-doped $Al_xGa_{1-x}As$ carrier blocking layer (with a donor concentration of $N_d = 1 \times 10^{18}$ cm^{-3}), and a 100-nm n-doped $Al_{0.25}Ga_{0.75}As$ cladding layer ($N_d = 1 \times 10^{18}$ cm^{-3}). Figure 44.2 also demonstrates the calculated energy band profiles, including the band edges of the conduction band E_c and the valence band E_v, the electron quasi-Fermi level E_{Fc}, and the hole quasi-Fermi level E_{Fv}, along the growth direction of z under a forward bias of $V = 1.4$ V. The Al content x is 0.4 in the carrier blocking layer. For the drift-diffusion model, the central problem is to find out the electrostatic potential Φ

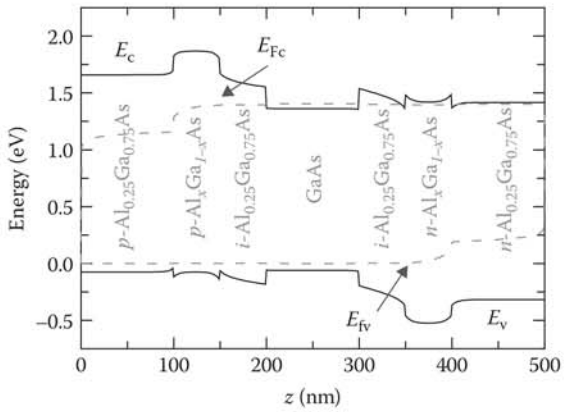


FIGURE 44.2 The energy band profiles, E_c and E_v, and the quasi-Fermi level profiles, E_{Fc} and E_{Fv}, along the growth direction z of the conduction and the valence bands, respectively, for the heterostructure under consideration. The profiles are obtained under a forward bias of 1.4 V.

and the quasi-Fermi levels for electrons E_{Fc}, and for holes E_{Fv} by solving three coupled equations (Horio and Yanai, 1990; Yen and Lee, 2010). These are the one-dimensional Poisson equation in the steady state:

$$\frac{\mathrm{d}}{\mathrm{d}z}\epsilon\frac{\mathrm{d}}{\mathrm{d}z}\Phi = -q\left(p - n + N_\mathrm{d}^+ - N_\mathrm{a}^-\right), \tag{44.1}$$

and the one-dimensional continuity equations for the electron and the hole currents:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}J_\mathrm{n} &= q(R-G),\\ \frac{\mathrm{d}}{\mathrm{d}z}J_\mathrm{p} &= -q(R-G),\end{aligned} \tag{44.2}$$

where ϵ is the electric permittivity; p, n, N_d^+, and N_a^- are the concentrations of the holes, electrons, ionized donors, and ionized acceptors, respectively; J_n (J_p) is the current densities due to the electron (hole) flow; R is the electron–hole recombination rate minus the intrinsic generation rate due to thermal excitation; and G is the electron–hole generation rate due to the photon recycling effect to be described. It is noted that all of the aforementioned p, n, N_d^+, N_a^-, J_n, J_p, R, and G are functions of Φ, E_{Fc}, and E_{Fv}. The first six ones can be expressed by the following formulas:

$$n = N_\mathrm{c}F_{1/2}\left(\eta_\mathrm{n}\right), \eta_\mathrm{n} \equiv \left(E_{Fc} - E_\mathrm{c}\right)/k_\mathrm{B}T, \tag{44.3}$$

$$p = N_\mathrm{v}F_{1/2}\left(\eta_\mathrm{p}\right), \eta_\mathrm{p} \equiv \left(E_\mathrm{v} - E_{Fv}\right)/k_\mathrm{B}T, \tag{44.4}$$

$$N_\mathrm{d}^+ = \frac{N_\mathrm{d}}{\left\{1 + 2\exp\left[\left(E_{Fc} - E_\mathrm{d}\right)/k_\mathrm{B}T\right]\right\}}, \tag{44.5}$$

$$N_\mathrm{a}^- = \frac{N_\mathrm{a}}{\left\{1 + 4\exp\left[\left(E_\mathrm{a} - E_{Fv}\right)/k_\mathrm{B}T\right]\right\}}, \tag{44.6}$$

$$J_\mathrm{n} = -q\mu_\mathrm{n}n\mathrm{d}\Phi/\mathrm{d}z + qD_\mathrm{n}\mathrm{d}n/\mathrm{d}z, \tag{44.7}$$

$$J_\mathrm{p} = -q\mu_\mathrm{p}p\mathrm{d}\Phi/\mathrm{d}z - qD_\mathrm{p}\mathrm{d}p/\mathrm{d}z, \tag{44.8}$$

where N_c (N_v) is the effective density of states for the conduction (valence) band, F_j the Fermi–Dirac integral of order j, k_B the Boltzmann constant, T the temperature, E_d (E_a) the donor (acceptor) level, μ_n (μ_p) the electron (hole) mobility, and D_n (D_p) the diffusivity for the electrons and the holes.

44.2.3 Recombination and Generation Processes

For the recombination processes in the device, the Shockley–Read–Hall (SRH), the radiative, and the Auger recombinations are considered in the high-quality GaAs active layer, while only the SRH recombination process is assumed to be dominant in the other layers containing aluminum. Therefore, the recombination rate is

$$R = \begin{cases} R_{SRH} + R_{rad} + R_{Aug} & z \text{ in the active layer,} \\ R_{SRH} & \text{otherwise,} \end{cases} \tag{44.9}$$

where R_{SRH}, R_{rad}, and R_{Aug} are the recombination rates due to the SRH, the radiative, and the Auger recombination processes, respectively. In the calculation of R_{rad}, we consider the spontaneous emission contributed from the transition from the conduction band (c) to both the heavy-hole (hh) and the light-hole (lh) bands and express R_{rad} as

$$R_{rad} = \int r(\hbar\omega)\, d\hbar\omega, \; r = r_{hh} + r_{lh}, \tag{44.10}$$

where $\hbar$ is the reduced Planck constant, ω the optical angular frequency, and r the total spontaneous emission rate that consists of r_{hh} and r_{lh} attributed to the transitions of c→hh and c→lh, respectively. With the parabolic band approximation, r_{hh} can be expressed as (Chuang, 1995)

$$r_{hh}(\hbar\omega) = C_{hh} \frac{\hbar\omega\sqrt{\hbar\omega - E_g}\,\Theta(\hbar\omega - E_g)}{\left[1 + e^{-\eta_n} \exp\left(\frac{m_r}{m_c}\frac{\hbar\omega - E_g}{k_B T}\right)\right]\left[1 + e^{-\eta_p} \exp\left(\frac{m_r}{m_{hh}}\frac{\hbar\omega - E_g}{k_B T}\right)\right]}, \tag{44.11}$$

where

$$C_{hh} = \frac{q^2 n_r m_r^{3/2} E_p}{3\sqrt{2}\pi^3 \epsilon_0 m_0 \hbar^5 c^3}. \tag{44.12}$$

ϵ_0, m_0, and c are the electric permittivity, the electron mass, and the light speed, *in vacuo*, respectively; n_r is the refractive index; E_p is the optical matrix parameter; $m_r = m_c m_{hh}/(m_c + m_{hh})$ is the reduced mass; and Θ is the unit step function. For r_{lh}, it is simply obtained by replacing the role of m_{hh} in Equations 44.11 and 44.12 with the effective mass of the light hole band, m_{lh}.

For the Auger recombination rate, the CHHS and CHCC processes are considered. Therefore, R_{Aug} is

$$R_{Aug} = \left(C_p p + C_n n\right)\left(np - n_i^2\right), \tag{44.13}$$

where C_p (C_n) is the Auger coefficient corresponding to the CHHS (CHCC) process and n_i the intrinsic carrier density. For the SRH recombination, we use

$$R_{SRH} = \frac{pn - n_i^2}{\tau_n\left(p + n_i\right) + \tau_p\left(n + n_i\right)}, \tag{44.14}$$

where τ_n (τ_p) is the minority carrier lifetime when $p \gg n, n_i$ ($n \gg p, n_i$).

For the generation rate G, the carriers generated by photon recycling are accounted in the active layer. Photon recycling is a process involving the absorption of spontaneously emitted photons by the active material itself and hence strongly related to the distribution of optical field in the device. A detailed consideration of the recycling process relies on the information about the device geometry, which is out of the scope for the present work. In this work, the generation rate G is directly connected to R_{rad} by

$$G = \frac{1}{L}\eta_{\mathrm{pr}}\left(1-\eta_{\mathrm{xp}}\right)\int\limits_{z\in \text{ active layer}} R_{\mathrm{rad}}dz, \tag{44.15}$$

where η_{xp} is the light extraction efficiency, the probability for the spontaneously generated photons emitting out of the device, and η_{pr} is the photon recycling efficiency, the ratio of trapped photons eventually being absorbed by the active layer. Therefore, Equation 44.15 implies that the photon recycling process generates carriers uniformly in the active layer, which results from the absorption of a portion of spontaneously generated photons trapped in the device. The assumption of a uniform carrier generation rate is justifiable for an active layer thickness no longer than the inverse of the absorption coefficient.

44.2.4 Boundary Conditions

To obtain an unambiguous distribution of Φ, E_{Fc}, and E_{Fv}, the boundary conditions for Equations 44.1 and 44.2 have to be specified. For the boundary condition of Φ, it is simply assumed that the electrostatic potential Φ and the displacement field $-\epsilon d\Phi/dz$ are continuous at interfaces. The boundary conditions of E_{Fc} and E_{Fv} are, however, complicated and classified into two cases according to the magnitude of band offsets. For the case of a band offset smaller than $2k_{\mathrm{B}}T$, the quasi-Fermi levels are determined from the continuity of drift-diffusion currents at the both sides of the interface under the assumption of a continuous quasi-Fermi level. For the other case of a larger band offset, the drift-diffusion model is no longer applicable for carriers' transport between layers. In this case, the thermionic process has to be taken into account, which allows quasi-Fermi levels to split at the interface. The thermionic emission current density due to the electron flow at the interface z_{i} with Fermi–Dirac distribution accounted can be expressed as†

$$J_{\mathrm{n}}\left(z_{\mathrm{i}}\right) = A_{\mathrm{n}}\left(z_{\mathrm{i}}^{+}\right)T^{2}F_{1}\left[\xi_{\mathrm{n}}\left(z_{\mathrm{i}}^{+}\right)\right] - A_{\mathrm{n}}\left(z_{\mathrm{i}}^{-}\right)T^{2}F_{1}\left[\xi_{\mathrm{n}}\left(z_{\mathrm{i}}^{-}\right)\right], \tag{44.16}$$

where $z_{\mathrm{i}}^{\pm} = \lim_{\varepsilon\to 0}\left(z_{\mathrm{i}}\pm\varepsilon\right)$, $A_{\mathrm{n}}\left(z_{\mathrm{i}}^{\pm}\right)$ is the effective Richardson constant of electrons at $z_{\mathrm{i}}^{\pm}$, and $\xi_{\mathrm{n}}\left(z_{\mathrm{i}}^{\pm}\right)$ is the energy difference, normalized by $k_{\mathrm{B}}T$, from the quasi-Fermi levels at $z_{\mathrm{i}}^{\pm}$ to the larger one of the conduction band edges at the interface. By matching the drift-diffusion current to the thermionic current at interfaces, we can write the boundary condition for E_{Fc} as

$$\begin{aligned} &J_{\mathrm{n}}\left(z_{\mathrm{i}}\right) = J_{\mathrm{n}}\left(z_{\mathrm{i}}^{\pm}\right), \quad \text{for}\left|E_{\mathrm{c}}\left(z_{\mathrm{i}}^{+}\right) - E_{\mathrm{c}}\left(z_{\mathrm{i}}^{-}\right)\right| \geq 2k_{\mathrm{B}}T,\\ &J_{\mathrm{n}}\left(z_{\mathrm{i}}^{+}\right) = J_{\mathrm{n}}\left(z_{\mathrm{i}}^{-}\right), \quad \text{otherwise,} \end{aligned} \tag{44.17}$$

with $J_{\mathrm{n}}\left(z_{\mathrm{i}}^{\pm}\right)$ being the drift-diffusion current at $z_{\mathrm{i}}^{\pm}$. The boundary condition for E_{Fv} is determined by the continuity of J_{p} at the interface, similar to that for E_{Fc} just described above.

At the end of surfaces on which the metal contacts are to be made, the Fermi levels are assumed to be pinned at a fixed position from the band edges such that the carrier densities equal to the values of those

† It is noted that, at a low temperature, the contribution of the tunneling current through an abrupt heterojunction may not be neglected. In such a situation, Equation 44.16 has to be modified. Interested readers can refer to Yang et al. (1993) for detailed expression.

in thermal equilibrium. The Fermi level positions at the surface are thus determined by the charge neutral condition, $p + N_d^+ = n + N_A^-$, with $E_{Fc} = E_{Fv}$.

With the boundary conditions being specified, the differential Equations 44.1 and 44.2 can be solved steadily by the Newton–Raphson iteration scheme for each applied voltage V. The external electrical work qV done on each of injected carriers gives the Fermi level difference between two surfaces $E_{Fc}|_{z \text{ at cathode}} - E_{Fc}|_{z \text{ at anode}} = qV$. Parameters used in this work are listed in Table 44.1, except for mobility, which is obtained from an empirical model (Sotoodeh et al., 2000). The band offset ratio ($\Delta E_c/\Delta E_v$) is set as 0.6:0.4 for GaAs/$Al_xGa_{1-x}As$ interfaces. The binding energies of donors and acceptors are set at k_BT for simplicity.

44.2.5 Output Quantities and Models for Steady-State Analysis

Once the solutions Φ, E_{Fc}, and E_{Fv} are obtained, the quantities of interest, including various current components, power densities, and efficiencies, can be calculated and used for further analysis. Figure 44.3 shows the working principle of an EL refrigerator. In the steady state, electrons and holes are injected

TABLE 44.1 Material Parameters of GaAs and AlGaAs at 300 K Used for Calculation

Parameter	Symbol	Unit	GaAs	$Al_xGa_{1-x}As$ ($x < 0.45$)
Bandgap	E_g	eV	$1.424 + 1.247x$	
Effective mass	m_c	m_0	$0.063 + 0.083x$	
	m_{hh}		$0.51 + 0.25x$	
	m_{lh}		$0.082 + 0.068x$	
Effective density of states	N_c	cm^{-3}	$2.5 \times 10^{19}(m_c/m_0)^{3/2}$	
	N_v		$2.5 \times 10^{19}[(m_{hh}/m_0)^{3/2} + (m_{lh}/m_0)^{3/2}]$	
Effective Richardson constant	A_n	$A/cm^2/K^2$	$120 \times (m_c/m_0)$	
	A_p		$120 \times (m_{hh} + m_{lh})/m_0$	
Electric permittivity	ϵ	ϵ_0	$12.9 - 2.84x$	
Refractive index	n_r	–	3.6	–
Optical matrix parameter	E_p	eV	25.7	–
SRH recombination lifetime	τ_n	s	1.3×10^{-6}	1×10^{-8}
	τ_p		1.2×10^{-6}	1×10^{-8}
Auger coefficient	C_n	cm^6/s	1.0×10^{-31}	–
	C_p		1.2×10^{-30}	–

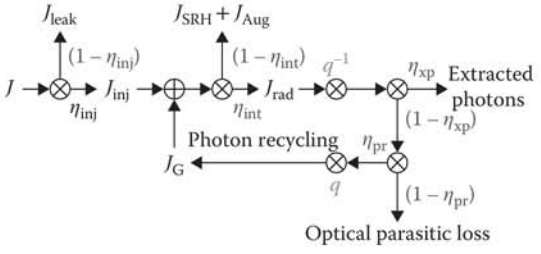


FIGURE 44.3 A diagram illustrating of the operation of an electroluminescent refrigerator with photon recycling as a feedback. It indicates the relationships among various currents. Symbol + stands for an addition operation that sums its inputs, indicated by inward arrows with the tips pointing at the symbol, and gives its output, indicated by an outward arrow with the tip pointing away from the symbol. Symbol × stands for multiplication operations, which give their outputs, indicated by outward arrows, by multiplying their inputs, indicated by inward arrows, by the factors along with the individual outward arrows.

from the electrodes under a forward bias, leading to a total current J in the device. Most of the electrons and holes are intended to be injected into the active region, recombine therein, and result in a major current component, called the injection current J_{inj}. The injection efficiency η_{inj} is defined as the ratio J_{inj}/J. The other part of the small current, called the leakage current J_{leak}, arises from the SRH recombination outside the active region $J'_{\text{SRH}} = \int_{z\notin\text{active layer}} R_{\text{SRH}}\text{d}z$ and the recombination current at the device surfaces $J_{\text{sr}} = J_{\text{n}}\big|_{z\text{ at anode}} + J_{\text{p}}\big|_{z\text{ at cathode}}$. In this figure, we see that there are two means of injecting carriers into the active region. One is by the external electric injection and the other is by the photon recycling effect, which result in two effective current components J_{inj} and J_{G}, respectively. As a result, the sum $J_{\text{inj}} + J_{\text{G}}$ accounts for the total effective injection rate of electrons and holes to the active region. In the steady-state condition, the injection rate is balanced by the recombination rate in the active layer. Hence,

$$J_{\text{inj}} + J_{\text{G}} = J_{\text{SRH}} + J_{\text{rad}} + J_{\text{Aug}}, \tag{44.18}$$

where J_{SRH}, J_{rad}, and J_{Aug} are the current components in the active region due to the SRH, the radiative, and the Auger recombinations, respectively, with $J_i = q\int_{z\in\text{active region}} R_i\text{d}z$, for i = SRH, rad, and Aug. The internal quantum efficiency η_{int} is naturally defined as the probability for electron–hole pairs recombining through the radiative recombination:

$$\eta_{\text{int}} = \frac{J_{\text{rad}}}{J_{\text{SRH}} + J_{\text{rad}} + J_{\text{Aug}}}. \tag{44.19}$$

Owing to the radiative recombination, photons are generated with a rate J_{rad}/q. Since only a ratio η_{xp} of generated photons can escape out of the device, the emission flux is $\eta_{\text{xp}}J_{\text{rad}}/q$ and the corresponding emission power is:

$$P_{\text{rad}} = \eta_{\text{xp}}\int\int r\hbar\omega\text{d}\hbar\omega\text{d}z \equiv \eta_{\text{xp}}\langle\hbar\omega\rangle J_{\text{rad}}/q, \tag{44.20}$$

where $\langle\hbar\omega\rangle$ is the average photon energy. In addition to these emitted photons, the others are trapped in the device, either being parasitically absorbed, causing heat, or being recycled, reincarnating into electron–hole pairs in the active region. Figure 44.3 shows that the photon recycling effect is a feedback process, supplying a portion of electron–hole pairs required for recombinations. From Figure 44.3, we can simply express J_{G} as $J_{\text{G}} = \eta_{\text{pr}}\left(1-\eta_{\text{xp}}\right)J_{\text{rad}}$, which is equivalent to Equation 44.15 with the equality $J_{\text{G}} = qGL$.

In conventional optoelectronic devices, the external quantum efficiency η_{ext} is defined as the ratio of the photon flux emitted out of the devices to the carrier flux injected into the devices, i.e., $\eta_{\text{ext}} = \eta_{\text{xp}}J_{\text{rad}}/J$. Using Equation 44.18 and the definitions of various efficiencies, η_{ext} with photon recycling effect accounted can be expressed in terms of the other efficiencies:

$$\eta_{\text{ext}} = \frac{\eta_{\text{xp}}\eta_{\text{int}}\eta_{\text{inj}}}{1-\eta_{\text{int}}\eta_{\text{pr}}\left(1-\eta_{\text{xp}}\right)}. \tag{44.21}$$

In the absence of photon recycling ($\eta_{\text{pr}} = 0$), η_{ext} returns to the conventional form of $\eta_{\text{ext}} = \eta_{\text{xp}}\eta_{\text{int}}\eta_{\text{inj}}$.

For the purpose of refrigeration, the cooling power and the cooling efficiency are two important parameters of interests. Following the concepts shown in Figure 44.1b, we define the cooling power as the power difference of the output and the input powers, which is required to keep the cooler operating steadily at 300 K:

$$P_{\text{c}} = P_{\text{rad}} - JV = \eta_{\text{c}}JV, \tag{44.22}$$

where $\eta_c \equiv P_c/JV$ is known as the cooling efficiency. The notion of EL refrigeration is to set the devices in the cooling mode, which requires a positive cooling power or cooling efficiency. Using Equation 44.20 and the definition of η_{ext}, one can rewrite the cooling efficiency as

$$\eta_c = \frac{\eta_{ext}}{\eta_{ext,cr}} - 1, \tag{44.23}$$

where

$$\eta_{ext,cr} \equiv \frac{qV}{\langle \hbar\omega \rangle}, \tag{44.24}$$

is the critical external efficiency for the devices to act in the cooling mode, which is consistent with the result from Figure 44.1b if the external efficiency of the cooler is considered. From Equation 44.21, the critical extraction efficiency for cooling is:

$$\eta_{xp,cr} = \frac{1 - \eta_{pr}\eta_{int}}{\eta_{int}\left(\eta_{inj}/\eta_{ext,cr} - \eta_{pr}\right)}, \tag{44.25}$$

for $\eta_{inj}/\eta_{ext,cr} > \eta_{pr}$.

44.3 Results and Discussion

In this section, we present the results simulated by the aforementioned self-consistent calculation to investigate various structural and optical effects on the cooling feasibility and capability of GaAs/AlGaAs EL refrigerators. In the discussion of the structural effects, we have assumed an unity light extraction efficiency ($\eta_{xp} = 1$).

44.3.1 Influence of Device Structures on EL Refrigeration

44.3.1.1 Leakage Current and Carrier Blocking Layer

We first study the cooling feasibility of a simple double heterostructure (DH) EL cooler by plotting the internal, the external, and the critical efficiencies shown in Figure 44.4 as functions of the applied voltage

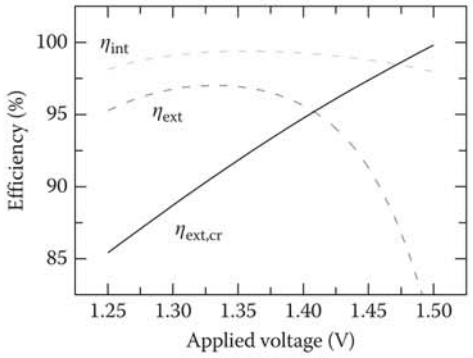


FIGURE 44.4 The internal efficiency η_{int}, the external efficiency η_{ext}, and the critical external efficiency $\eta_{ext,cr}$ as functions of the applied voltage V for the heterostructure with an active layer of width $L = 0.1$ μm and carrier blocking layers of Al composition $x = 0.25$, which has no effect on blocking carriers.

V for $\eta_{xp} = 1$. The Al content x in the blocking layer is 0.25, which is identical with that in the cladding and the spacer layers so that there is no blocking effect. As can be seen, while the critical external efficiency $\eta_{ext,cr}$ increases linearly with the applied voltage V, the internal efficiency η_{int} first increases to a maximum at $V = 1.36$ V and then decreases. The variation of η_{int} on V reveals a more important role played by the SRH and the Auger recombinations at a lower and a higher voltage, respectively. Similar to η_{int}, the external efficiency η_{ext} reaches to a maximum at $V = 1.33$ V and then droops more severely than η_{int} as V further increases, crossing with $\eta_{ext,cr}$ at the critical voltage $V_{cr} = 1.41$ V, where $\eta_{ext} = \eta_{ext,cr}$. We see that the DH EL cooler can act in the cooling mode for $V < 1.41$ V.

It is not difficult to find that the deterioration of η_{ext} is caused by a rapid reduction in the injection efficiency η_{inj} since $\eta_{inj} = \eta_{ext}/\eta_{int}$ for $\eta_{xp} = 1$. This implies that V_{cr} can be improved if the leakage current can be alleviated. In our calculation, it is shown that the strong surface recombination dominates in the leakage current as a result of a high minority carrier density near the terminal contacts. This situation becomes worse in particular for the devices of shorter cladding layers as in this work. The high minority carrier density can be reduced significantly by carrier blocking layers. In Figure 44.5a, we show the electron density in the center of the p-cladding layer as a function of V for $x = 0.25$, 0.3, 0.35, and 0.4. As can be seen, the electron density decreases monotonically with increasing x. This indicates that the energy barrier provided by the blocking layer can effectively block the minority carriers, preventing them from flowing over the carrier blocking layer. This leads to a lower minority carrier density in the cladding layer and hence a lower surface recombination current. The low surface recombination current is confirmed by the injection efficiency as plotted in Figure 44.5b. Evidently, the injection efficiency is improved significantly by increasing the barrier height when $x \leq 0.3$. The improvement however becomes insignificant as x increases from 0.35 to 0.4, in contrast to the result from Figure 44.5a This is because the leakage current across the blocking layer is almost eliminated when $x = 0.35$. In this case, $\eta_{inj} \approx 1$ and the residue of the leakage current is mostly due to the SRH recombination in the space layer. Contrarily, for $x \leq 0.3$, the injection efficiency decreases rapidly at high bias as a result of J_{sr}.

In Figure 44.6, we show the cooling power as a function of the applied voltage for EL cooler of various Al content in the blocking layer. As expected, the critical voltage increases from 1.41 V for $x = 0.25$ to 1.47 V for $x = 0.35$ and 0.4. Moreover, the cooling power also increases from 7.5 to 20 W/cm^2 for $x = 0.25$ to 0.35. The improvement of cooling power comes from an improved η_{inj} and hence η_{ext} as indicated by Equations 44.22 and 44.23, i.e., $P_c = JV\left(\eta_{ext}/\eta_{ext,cr} - 1\right)$. Further increasing the Al content to 0.4 does not help in improving the cooling power, which is consistent with the variation of the injection efficiency as shown in Figure 44.5b.

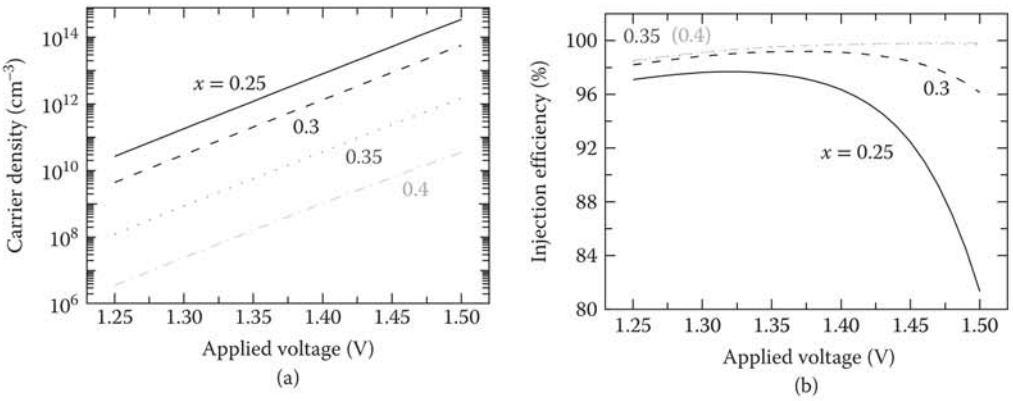


FIGURE 44.5 (a) The minority carrier (electron) density in the center of the *p*-cladding layer and (b) the injection efficiency as functions of the applied voltage V for various aluminum content $x = 0.25$, 0.3, 0.35, and 0.4.

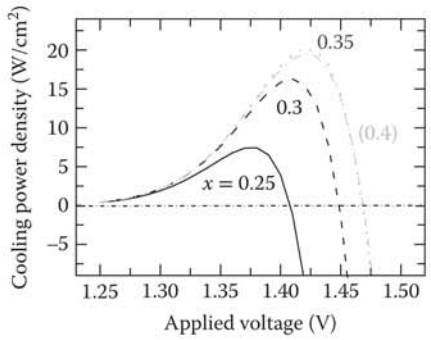


FIGURE 44.6 The ideal cooling power density P_c^0 as a function of the applied voltage V for various Al compositions of $x = 0.25$, 0.3, 0.35, and 0.4 in the carrier blocking layer. The extraction efficiency is set at unity.

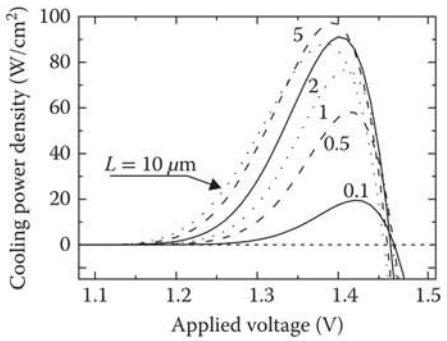


FIGURE 44.7 The ideal cooling power density P_c^0 as a function of V for $L = 0.1$, 0.5, 1, 2, 5, and 10 μm.

44.3.1.2 Active Layer Thickness

The cooling power density of 20 W/cm^2 seems to be the limiting value for the device of active layer thickness $L = 0.1$ μm. To further enhance it, one can increase the active layer thickness. Intuitively, a wider active layer can contain more electron–hole pairs to recombine therein, leading to an enhanced J_{rad} and hence more effective cooling cycles. Figure 44.7 shows the cooling power density P_c as a function of V for the structures with the same Al content ($x = 0.4$) in the carrier blocking layer but various active layer thickness $L = 0.1$, 0.5, 1, 2, 5, and 10 μm; here, $\eta_{xp} = 1$. As can be seen, the maximum cooling power $P_{c,max}$ can be improved about five times up to as high as 97 W/cm^2 by increasing L from 0.1 to 5 μm. However, this improvement is only prominent for $L \leq 5$ μm. Further increasing the active layer thickness does not enhance but degrades the maximum cooling power due to the remarked nonuniformity of carrier distribution, as L is comparable to or greater than the carrier diffusion length. From our calculation, we find that the average carrier density in the active layer is lower for a larger L; this difference in the average carrier density between different Ls is particularly remarkable when V is high. This renders the cooling power for $L = 10$ μm larger than that for $L = 5$ μm at a low V but smaller at a high V.

In Figure 44.8, we show the dependences of various efficiencies on the active layer thickness by plotting the injection efficiency η_{inj} and the internal efficiency η_{int} in Figure 44.8a, the product $\eta_{inj}\eta_{int}$ and the critical external efficiency $\eta_{ext,cr}$ in Figure 44.8b, and the cooling efficiency η_c in Figure 44.8c as functions of the applied voltage V for the active layer thickness $L = 0.1$, 1, and 10 μm and $\eta_{xp} = 1$. The Al content of the blocking layer is $x = 0.4$. Since J_{inj} is enhanced by increasing L, the η_{inj} is also enhanced. For $L \geq 1$ μm, the injection efficiency is nearly 100%, almost independent of the applied voltage. As been described, the internal efficiency η_{int} reaches to a maximum at about 1.37 V. It declines more profoundly as V decreases

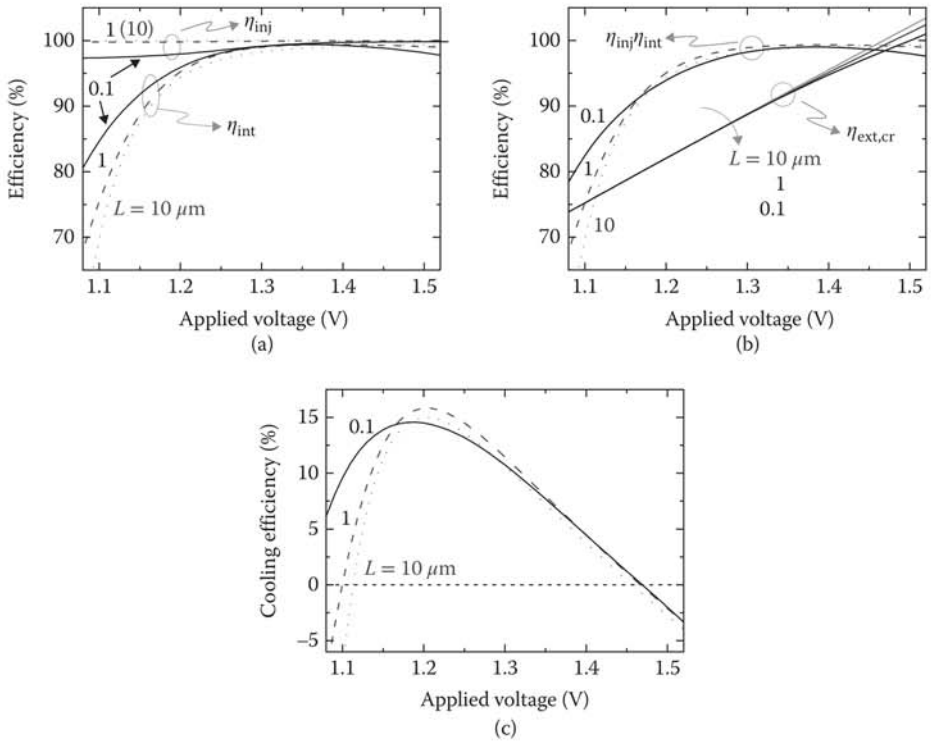


FIGURE 44.8 (a) The injection efficiency η_{inj} and the internal efficiency η_{int}, (b) the production of $\eta_{inj}\eta_{int}$ and the critical extraction efficiency $\eta_{ext,cr}$, and (c) the ideal cooling efficiency η_c^0 as functions of the applied voltage for $L = 0.1$, 1, and 10 μm.

than increases, implying a more important SRH recombination at low V than the Auger recombination at high V. Such a lowering of η_{int} at low V is more prominent for a larger L because of a lower carrier density in the active layer. On the other side, at high V, the Auger recombination is less important for a larger L due also to a lower carrier density. The ratio of $\eta_{ext,cr}/(\eta_{inj}\eta_{int})$ is the critical extraction efficiency $\eta_{xp,cr}$, which gives the design margin for η_{xp}. For instance, at $V = 1.2$ V, $\eta_{inj}\eta_{int} = 94\%$ and $\eta_{ext,cr} = 82\%$, as can be found in Figure 44.8b, give $\eta_{xp,cr} = 87\%$. Evidently, the margin becomes tighter when bias V deviates from 1.2 V. If an EL cooler is improperly biased at a voltage where $\eta_{xp,cr} \geq 1$, the ideal cooling efficiency will become negative (Figure 44.8c). In this case, it is impossible for the device to act in the cooling mode. At $V = 1.39$ V where the maximum cooling power occurs (see Figure 44.7), $\eta_{xp,cr} = 96\%$ for $L = 10$ μm. In such a bias condition, the corresponding cooling efficiency is only about 4%. The cooling efficiency can be improved by biasing the cooler at $V = 1.2$ V at which it can be as high as 15% at the cost of the cooling power density.

44.3.1.3 Doping in Active Layer

In Figure 44.9, we show the influences of the doping types and concentrations on EL refrigeration by plotting the ideal cooling power as a function of the applied voltage V for the active regions which are undoped, doped with $N_d = 1 \times 10^{18}$ cm^{-3}, with $N_a = 1 \times 10^{18}$ cm^{-3}, and with $N_a = 5 \times 10^{18}$ cm^{-3}. Here, the thickness of the active layer is $L = 1$ μm and the Al content of the blocking layers is $x = 0.4$. We find that intentionally doping into the active region does not improve the cooling power. The degradation caused by p-type doping is mainly due to the enhanced Auger recombination. The n-type doping suppresses the radiative recombination and also the Auger recombination, leading to a slightly enhanced cooling power at high V but a reduced cooling power at low V.

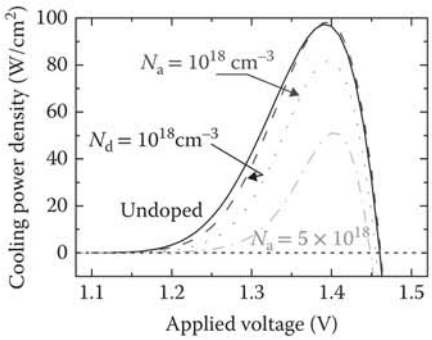


FIGURE 44.9 The ideal cooling power density P_c^0 as a function of the applied voltage for active layers with different doping types and concentrations, including undoped, doped with $N_d = 1 \times 10^{18}$ cm^{-3}, with $N_a = 1 \times 10^{18}$ cm^{-3}, and with $N_a = 5 \times 10^{18}$ cm^{-3}. Here, the active layer thickness is $L = 1$ μm, the Al content of the blocking layer is $x = 0.4$, and the light extraction efficiency is set at unity.

44.3.2 Influence of Optical Processes on EL Refrigeration

44.3.2.1 Influence of Light Extraction

Light extraction efficiency influences significantly the performance of an EL refrigerator in the absence of photon recycling. This is because the photons trapped within the device turn out to be totally absorbed and transformed into thermal energy. In this subsection, we study the influence of the light extraction efficiency on EL refrigeration for the worst case of zero photon recycling efficiency ($\eta_{pr} = 0$). Here, the active layer thickness L is set at 1 μm and the Al content $x = 0.4$ is used for the carrier blocking layer. Figure 44.10a shows the curves of the cooling power P_c as a function of the applied voltage V for five values of extraction efficiency $\eta_{xp} = 0.8$, 0.85, 0.9, 0.95, and 1. The curves exhibit a peak for $\eta_{xp} = 0.9$, 0.95, and 1, but no peak for $\eta_{xp} = 0.8$ and 0.85. As expected, the height of the peak diminishes drastically from 79 to 2 W/cm^2 as η_{xp} reduces slightly from unity to 0.9. The voltage at which the curve peaks moves from 1.41 V for $\eta_{xp} = 1$ to 1.28 V for $\eta_{xp} = 0.9$. The cooling power falls rapidly from the peak value to negative as the voltage increases.

These behaviors can be understood from Equation 44.23, i.e., $\eta_c = \eta_{ext}/\eta_{ext,cr} - 1$, with the help of the relationships between the curves of the corresponding external efficiencies η_{ext} and the critical external efficiencies $\eta_{ext,cr}$ shown in Figure 44.10b. The curves of $\eta_{ext,cr}$ for the different η_{xp} values merge into a single line, since $\eta_{ext,cr} = qV/\langle\hbar\omega\rangle$ is nearly proportional to V and independent of η_{xp}. The curves of η_{ext} for the different η_{xp} values behave similarly and separate almost equidistantly from their neighbors. As V increases, the η_{ext} value increases and then almost reaches saturation at a level $\eta_{ext} \approx \eta_{xp}$, with a slight decline in the high-voltage region. These features can be explained by the simple relation $\eta_{ext} = \eta_{xp}\eta_{int}\eta_{inj}$, where the injection efficiency η_{inj} is almost unity for the device containing carrier blocking layers. Therefore, the variation of η_{ext} with V for various η_{xp} simply follows the variation of the internal efficiency η_{int} multiplied by a scaling factor of η_{xp}. In Figure 44.10b, we see that there exist operation regions of refrigeration for $\eta_{xp} = 0.9$, 0.95, and 1, but no operation region for $\eta_{xp} = 0.8$ and 0.85, where the operation region of refrigeration is a range of voltage in which the device operates in the cooling modes ($\eta_c > 0$ or $\eta_{ext} > \eta_{ext,cr}$). This explains the feature shown in Figure 44.10a that only the three curves for $\eta_{xp} = 0.9$, 0.95, and 1 exhibit a peak of positive P_c. Figure 44.10c shows the cooling efficiency η_c as a function of V for $\eta_{xp} = 0.85$, 0.9, 0.95, and 1. We see that the peaks for η_c-V curves fall at $V = 1.2$ V. The cooling efficiencies decline when the operation voltage deviates from 1.2 V. Different to the peak voltage for η_c-V curves, the peaks of P_c-Vcurves in Figure 44.10a occur at voltages near the critical voltages, for which the cooling efficiency is small. This is why a slight increase of V may cause the drastic reduction of P_c from the peak value and turn the operation of the device from the cooling mode to the heating mode.

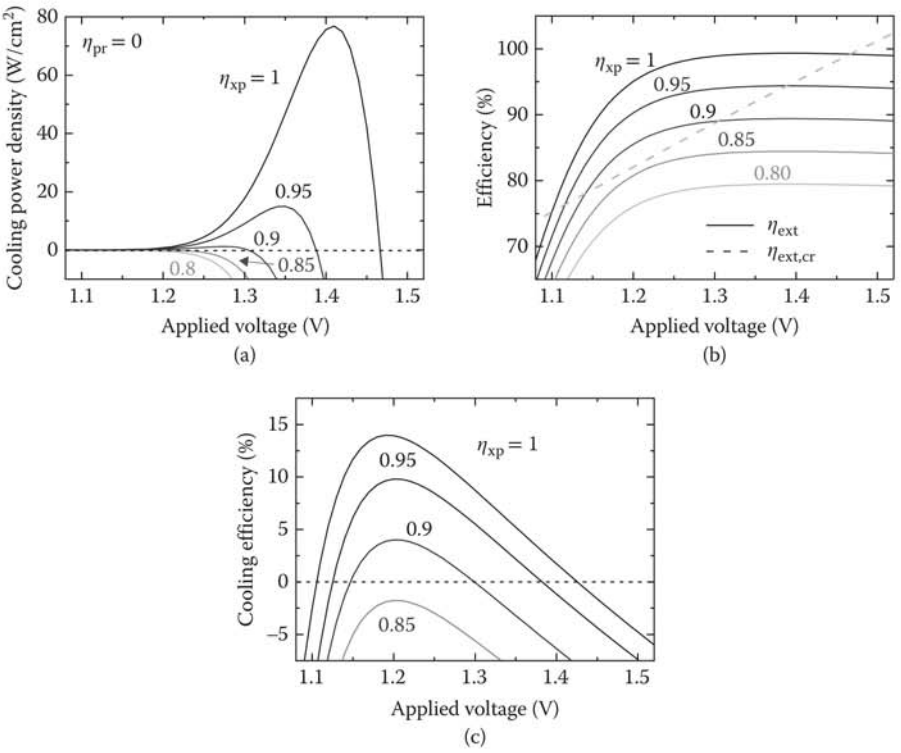


FIGURE 44.10 (a) The cooling power density P_c, (b) the external quantum efficiency η_{ext} and the critical extraction efficiency $\eta_{ext,cr}$, and (c) the cooling efficiency η_c as functions of the applied voltage for various extraction efficiency $\eta_{xp} = 0.8, 0.85, 0.9, 0.95$, and 1. The photon recycling efficiency is $\eta_{pr} = 0$.

44.3.2.2 Influence of Photon Recycling

Actually, most of the trapped photons can be reabsorbed by the high-quality GaAs active region and reincarnated into electron–hole pairs rather than into the thermal energy if the device structure is well designed. The photon recycling, as will be described, alleviates considerably the requirement of extraction efficiency for EL refrigeration. In this subsection, we demonstrate the photon recycling effect by examining another extreme case of unity recycling efficiency $\eta_{pr} = 1$, i.e., all of the trapped photons are recycled, turning into electron–hole pairs in the active region.

Figure 44.11a shows the cooling power density P_c as a function of V for five values of the extraction efficiency $\eta_{xp} = 0.2$, 0.4, 0.6, 0.8, and 1. All the curves in the figure behave as an asymmetrical peak of positive P_c, even for the low extraction efficiency of 0.2. The height of the peak now diminishes gently as η_{xp} reduces. These peaks occur almost at the same voltage ($V = 1.41$ V) except for $\eta_{xp} = 0.2$, for which the curve peaks at 1.39 V. Similarly, these behaviors can also be understood from Equation 44.23 and the relations between the corresponding η_{ext} and $\eta_{ext,cr}$ shown in Figure 44.11b. Comparing with those in Figure 44.10b, the external efficiencies η_{ext} shown in Figure 44.11b still vary with V in the manner that they first rise, reach a maximum, and then slightly decline. The difference is that η_{ext} is now much less sensitive to η_{xp}, especially in the voltage range where the internal efficiency η_{int} is near unity, as plotted in Figure 44.11c together with η_{inj} for various η_{xp}. In the voltage range between 1.3 and 1.5 V, where η_{int} is close to unity, the external efficiency η_{ext} can be approximated by

$$\eta_{ext} \approx \eta_{inj}\left[1 - \frac{1}{\eta_{xp}}\left(1 - \eta_{int}\right)\right]. \tag{44.26}$$

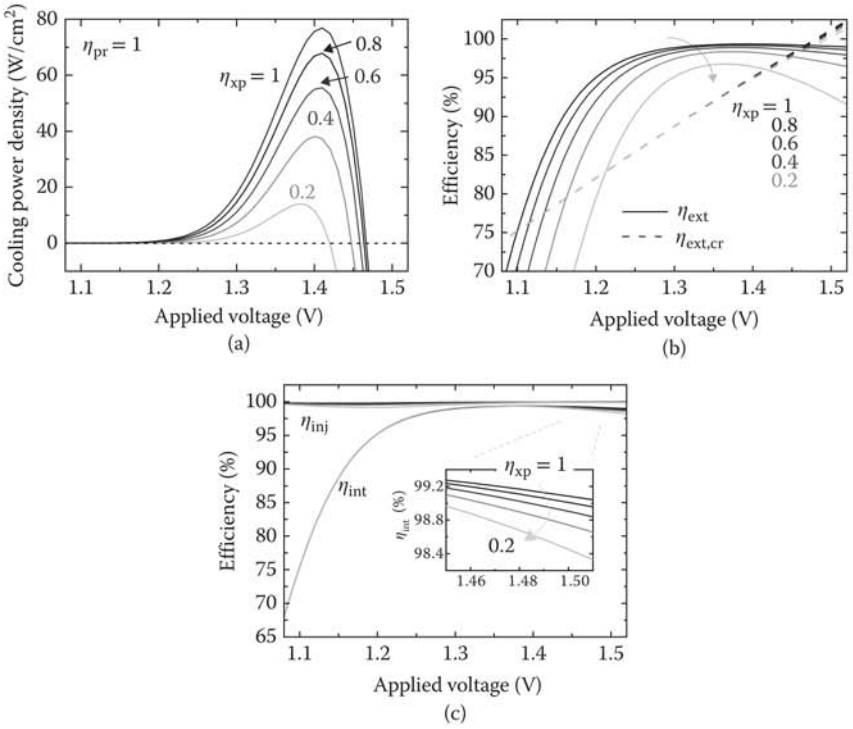


FIGURE 44.11 (a) The cooling power density P_c, (b) the external quantum efficiency η_{ext} and the critical external efficiency $\eta_{ext,cr}$, and (c) the injection efficiency η_{inj} and the internal efficiency η_{int} as functions of V for various extraction efficiency $\eta_{xp} = 0.2$, 0.4, 0.6, 0.8, and 1. The inset of (c) shows η_{int} at high V. The photon recycling efficiency is set at unity.

Equation 44.26 explains that, as $1 - \eta_{int}$ is small, η_{ext} is insensitive to η_{xp} and has a value close to η_{inj}, which is nearly unity.

In principle, the variation of η_{ext} basically follows the variation of η_{int} modified by photon recycling. This is because the only important loss is the nonradiative recombination in the active region when we set $\eta_{pr} = 1$ and $\eta_{inj} \approx 1$. For $\eta_{xp} = 1$, all the generated photons are emitted out of the device and there is no photon recycling, regardless of the value of η_{pr}, as can also be seen from Equation 44.21. The photon recycling effect manifests when η_{xp} deviates from unity. It causes an effective feedback current J_G and boosts the carrier concentration in the active region. The increasing carrier concentration degrades η_{int} and hence η_{ext} in the high-voltage region. This effect is particularly important for a low η_{xp} (see Figure 44.11b and the inset in Figure 44.11c). The lines of $\eta_{ext,cr}$ shown in Figure 44.11b do not merge as completely as those for the case of $\eta_{pr} = 0$ shown in Figure 44.10b. This is a consequence of the photon recycling effect, which causes a higher average photon energy $\langle \hbar\omega \rangle$ and, hence, a lower $\eta_{ext,cr}$ for a lower η_{xp}, due to a higher carrier concentration. The condition that $P_c = 0$ occurs at a critical voltage of about 1.47 V for $\eta_{xp} = 0.4$, 0.6, 0.8, and 1. The voltages are close to each other because of the crowding of the η_{ext}-V curves in the high-voltage region, which cross the $\eta_{ext,cr}$-V curves in the vicinity of 1.47 V. This implies that the P_c-V curves peak almost at the same voltage. The voltages for the peaks of the η_c-V curves, which can be estimated from Figure 44.11b, move from 1.2 to 1.3 V as η_{xp} decreases from 1 to 0.2.

Under the condition that $\eta_{pr} = 1$ and $\eta_{inj} \approx 1$, the critical extraction efficiency $\eta_{xp,cr}$ for the device to operate in the cooling mode can be written as

$$\eta_{xp,cr} \approx \frac{(1/\eta_{int}) - 1}{(1/\eta_{ext,cr}) - 1}. \tag{44.27}$$

The above expression implies that $\eta_{xp,cr}$ can be very small when η_{int} approaches unity, while $\eta_{ext,cr}$ does not. An EL refrigerator having a very low η_{xp} could operate in the cooling mode if there were no optical parasitic loss ($\eta_{pr} = 1$) and no current leakage ($\eta_{inj} = 1$).

The aforementioned discussion reveals that a high internal efficiency is essential to the photon recycling process. This implies that the quality of the materials making up the device is critical for refrigeration. If the material quality is not as high as in the study and gives $\eta_{int} < \eta_{ext,cr}$, it is impossible for the device to operate in the cooling mode as a result of the inequality $\eta_{ext} < \eta_{int}$ (see Equation 44.21), regardless of the values of the other efficiencies.

We now know that the internal efficiency η_{int} is an important factor in determining η_{ext} and, hence, to the performance of the device. It is influenced by photon recycling through the feedback generation of electron–hole pairs in the active region. To further gain insight into this effect, we set $\eta_{pr} = 1$ and investigate the average carrier density in the active region (Figure 44.12) as a function of η_{xp} when the device is biased at $V = 1.2$ and 1.4 V. As shown in Figure 44.12a, the average carrier density is low ($\approx 3 \times 10^{16}$ cm^{-3}) and nearly independent of η_{xp} for the device biased at 1.2 V, but is more than 5×10^{17} cm^{-3} and increases as η_{xp} reduces for the device biased at 1.4 V. This difference arises from the difference in the electric potential energy of the electrons (holes) in the active region relative to that in the *n*-type (*p*-type) cladding layer, from which the electrons (holes) are injected. This situation can be seen in Figure 44.12b, which shows the conduction band profiles at the *n*-doping side as functions of z for $V = 1.2$ and 1.4 V. When the device is biased at a low voltage, such as $V = 1.2$ V, the electric potential energy of the electrons is higher in the active layer than in the *n*-type cladding layer. This drop in electric potential energy causes the excess carriers, due to optical generation, to leak out of the active region, leading to an effective current J_G opposite the injection current J_{inj}. Consequently, accumulation of carriers in the active region is negligible, and the carrier density is basically independent of the photon recycling (and, hence, η_{int}). The drop in electric potential energy diminishes as the bias increases and, eventually, the situation reverses. At $V = 1.4$ V, the electric potential energy of the electrons is lower in the active region than in the n-type cladding layer. The potential profile now forms a potential well for the carriers and prohibits most of the excess carriers in the active region from leaking out. As a result, the optically generated carriers are mostly accumulated within the active region, recombined therein, and contribute an additional current, known as J_G. The carrier accumulation depends on the photon recycling and becomes particularly important for low η_{xp}, as shown in Figure 44.12a.

The variation of recombination current densities (J_{SRH}, J_{rad}, and J_{Aug}) on the extraction efficiency also follows the variation of carrier density such that they are invariant on η_{xp} at $V = 1.2$ V but increase considerably at $V = 1.4$ V for a decreasing η_{xp} (Lee and Yen, 2012). The feedback current J_G, which compensates J_{inj} for the total recombination in the active region, increases with decreasing η_{xp} for both

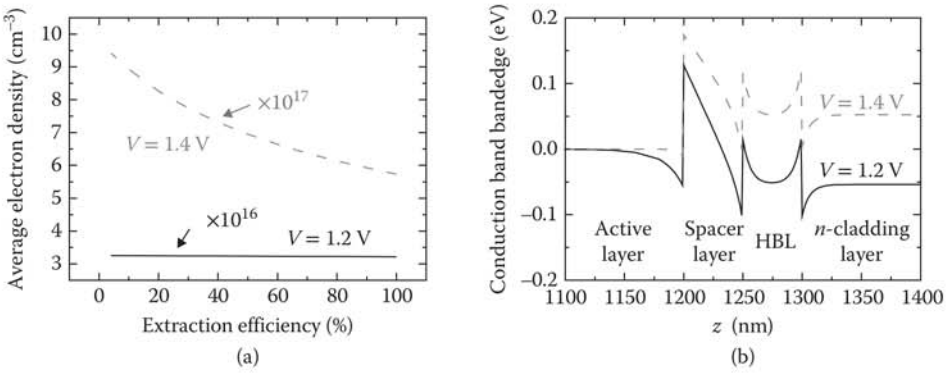


FIGURE 44.12 (a) The average electron densities in the active region for $V = 1.2$ and 1.4 V as functions of the extraction efficiency and (b) the conduction band profiles near the *n*-doping side as functions of z for $V = 1.2$ and 1.4 V. (HBL, hole-blocking layer.)

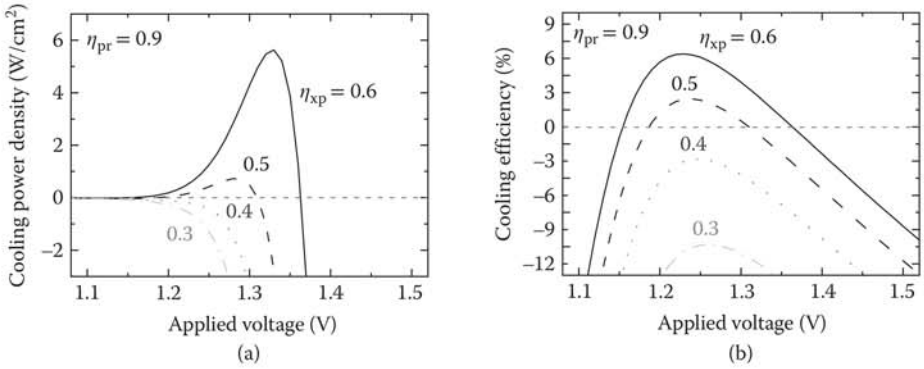


FIGURE 44.13 (a) The cooling power density P_c and (b) the cooling efficiency η_c as functions of applied voltage for various light extraction efficiencies $\eta_{xp} = 0.3$, 0.4, 0.5, and 0.6. The recycling efficiency is set at a practical value of 0.9.

applied voltages because $J_G = \eta_{pr}(1 - \eta_{xp})J_{rad}$. This leads to a significant reduction in J_{inj} and hence improves external efficiency for a low η_{xp}.

For a more practical case, we show the cooling power P_c in Figure 44.13a and the cooling efficiency η_c in Figure 44.13b as functions of the applied voltage V. Since $L = 1$ μm, η_{xp} is at most 60%. We therefore consider $\eta_{xp} = 0.3$, 0.4, 0.5, and 0.6 and set $\eta_{pr} = 0.9$, which is potentially achievable for a high-quality GaAs material. The resulting P_c-V relationship changes with η_{xp} to a degree milder than for $\eta_{pr} = 0$ (Figure 44.10a) but more severe than for $\eta_{pr} = 1$ (Figure 44.11a). The peak cooling power is now 5.6 W/cm^2 at $V = 1.33$ V for $\eta_{xp} = 0.6$, which is only 10% for the case of $\eta_{pr} = 1$ (Figure 44.11a). At the voltage of 1.33 V, the cooling efficiency η_c is small (2%) for $\eta_{xp} = 0.6$ and falls negative as η_{xp} reduces to 0.5, associated with the severe degradation of P_c from 5.6 to −3 W/cm^2. The situation can be cured by biasing the device at a lower voltage for an improved η_c. For example, at $V = 1.24$ V, η_c is 6.4% for $\eta_{xp} = 0.6$ and remains positive (about 2%) for $\eta_{xp} = 0.5$. The device can operate in the cooling mode for $\eta_{xp} > 0.45$. However, the cooling power is sacrificed. It is now only 0.9 W/cm^2 for $\eta_{xp} = 0.6$ and 0.3 W/cm^2 for $\eta_{xp} = 0.5$.

In the end of this chapter, we remark on the choices of active materials and operation temperatures for EL refrigeration. From Equation 44.24, the critical external efficiency of an EL cooler operating at E_g/q for maximum cooling power can be approximated by $\eta_{ext,cr} \approx E_g/\left(E_g + 2k_BT\right)$. Evidently, the external efficiency required for cooling is alleviated with decreasing bandgap energy and an increasing operating temperature. Indeed, recent research has shown that operating a 450-nm InGaN LED at 615 K is capable of producing a negative Peltier heating power inside the device (Xue et al., 2015). Moreover, theoretical study also confirms that elevating the operating temperature can improve the cooling power and efficiency for a GaAs EL cooler (Heikkilä et al., 2010). These improvements, as can be seen from the approximation of $\eta_{ext,cr}$, are strongly related to the ratio $E_g/2k_BT$. A smaller value of $E_g/2k_BT$ implies a fewer number of effective cooling cycles required for compensating a single parasitic loss. This means that the EL refrigeration will be more likely achieved for a cooler with an active material of smaller bandgap energy and operating in a higher temperature if the internal efficiency is high.

44.4 Summary

We discussed the influences of device structures and optical processes on the feasibility and capability of GaAs-based EL refrigerators by self-consistent calculation. For EL devices with short cladding layers, the leakage current, which arises from the minority carriers flowing toward the opposite electrodes, deteriorates the injection efficiency significantly at high bias. The degradation of injection efficiency can be cured

by employing carrier blocking layers of $Al_xGa_{1-x}As$ ($x \geq 0.35$), leading to an injection efficiency of nearly unity. The maximum cooling power density generally increases with the active layer thickness because of an increasing total radiative recombination rate. A GaAs active layer with thickness of 5 μm can give a high internal efficiency of nearly 100% and a limiting cooling power density of 97 W/cm^2. Further increasing the active layer thickness causes a nonuniform distribution of carriers, reducing the radiative power and also the maximum cooling power. Doping in the active layer does not help significantly in improving the cooling power density. Contrarily, a heavily *p*-doped active layer has much lower cooling power density than an undoped one owing to the enhanced Auger recombination.

The cooling capability depends strongly on the external efficiency, and hence the extraction efficiency, due to a small ratio of the extracted thermal energy to the work done on each injected electron–hole pair. For a GaAs EL cooler to act in the cooling mode, an external efficiency of 82% is needed at 1.2 V. This corresponds to an extraction efficiency of 87% if the trapped photons are all absorbed parasitically. This harsh requirement of extraction efficiency can be alleviated by photon recycling process provided that the internal efficiency is sufficiently high. The photon recycling behaves as an additional means of generating carriers, improving the external efficiency by reducing the driving currents. Therefore, for an ideal case of unity recycling efficiency, an EL device can act in the cooling mode even if the extraction efficiency is only 20%. For a more practical case of 90% recycling efficiency, our calculation shows that a maximum cooling power of 5.6 W/cm^2 and a maximum cooling efficiency of 6.4% can be achieved for an EL cooler of 1-μm GaAs active layer. The limiting extraction efficiency for this cooler is found to be 45%, which is achievable for the state-of-the-art technology. This reveals a good possibility of experimentally realizing EL refrigeration for high-cooling-power application.

References

Ashcroft NW, Mermin ND (1976) *Solid State Physics*. Fort Worth, TX: Harcourt Brace College.

Berdahl P (1985) Radiant refrigeration by semiconductor diodes. *J. Appl. Phys.* 58:1369–1374.

Chuang SL (1995) *Physics of Optoelectronic Devices*. New York, NY: Wiley.

Dousmanis GC, Mueller CW, Nelson H, Petzinger KG (1964) Evidence of refrigerating action by means of photon emission in semiconductor diodes. *Phys. Rev.* 133:A316–A318.

Epstein RI, Buchwald MI, Edwards BC, Gosnell TR, Mungan CE (1995) Observation of laser-induced fluorescent cooling of a solid. *Nature* 377:500–503.

Han P, Jin KJ, Ren SR, Zhou YL, Lu HB (2007) Numerical analysis of optothermionic refrigeration in semiconductor triple-well structure. *J. Appl. Phys.* 102:114501.

Heikkilä O, Oksanen J, Tulkki J (2009) Ultimate limit and temperature dependency of light-emitting diode efficiency. *J. Appl. Phys.* 105:093119.

Heikkilä O, Oksanen J, Tulkki J (2010) The challenge of unity wall plug efficiency: The effects of internal heating on the efficiency of light emitting diodes. *J. Appl. Phys.* 107:033105.

Horio K, Yanai H (1990) Numeral modeling of heterojunctions including the thermionic emission mechanism at heterojunction interface. *IEEE Tran. Electron. Devices* 37:1093–1098.

Lee KC, Yen ST (2012) Photon recycling effect on electroluminescent refrigeration. *J. Appl. Phys.* 111:014511.

Lin FR, Lee KC, Yen ST (2012) The ultimate cooling temperature of semiconductor electroluminescent refrigeration. In *the 12th International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD 2012)*, Shanghai, China, ThPD2.

Liu X, Zhao G, Zhang Y, Deppe DG (2013) Semiconductor laser monolithically pumped with a light emitting diode operating in the thermoelectrophotonic regime. *Appl. Phys. Lett.* 102:081116.

Mal'shukov AG, Chao KA (2001) Opto-thermionic refrigeration in semiconductor heterostructures. *Phys. Rev. Lett.* 86:5570–5573.

Melgaard S, Seletskiy D, Polyak V, Asmerom Y, Shiek-Bahae M (2014) Identification of parasitic loss in Yb:YLF and prospects for optical refrigeration down to 80 K. *Opt. Express* 22:7756–7764.

Oksanen J, Tulkki J (2010) Thermophotonic heat pump—A theoretical model and numerical simulations. *J. Appl. Phys.* 107:093106.

Oksanen J, Tulkki J (2015) LEDs feed on waste heat. *Nat. Photonics* 9:782–783.

Piprek J, Li Z-M (2016) Electroluminescent cooling mechanisms in InGaN/GaN light-emitting diodes. *Opt. Quant. Electron*. 48:472.

Santhanam P, Gray DJ Jr., Ram RJ (2012) Thermoelectrically pumped light-emitting diodes operating above unity efficiency. *Phys. Rev. Lett.* 108:097403.

Sheik-Bahae M, Epstein RI (2007) Optical refrigeration. *Nat. Photonics* 4:693–699.

Sheik-Bahae M, Epstein RI (2009) Laser cooling of solids. *Laser Photon. Rev.* 3:67–83.

Sotoodeh M, Khalid AH, Rezazadeh AA (2000) Empirical low-field mobility model for III-V compounds applicable in device simulation codes. *J. Appl. Phys.* 87:2890–2900.

Wang JB, Ding D, Yu SQ, Johnson SR, Zhang YH (2005) Electroluminescence cooling in semiconductors. In 2005 *Conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science and Photonic Applications Systems Technologies, QELS'05*, Baltimore, MD, 655–657.

Xue J, Zhao Y, Oh SH, Herrington WF, Speck JS, Denbaars ST, Nakamura S, Ram RJ (2015) Thermally enhanced blue light-emitting diode. *Appl. Phys. Lett*. 107:121109.

Yang K, East JR, Haddad GI (1993) Numerical modeling of abrupt heterojunctions using a thermionic-field emission boundary condition. *Solid State Electron.* 36:321–330

Yen ST, Lee KC (2010) Analysis of heterostructures for electroluminescent refrigeration and light emitting without heat generation. *J. Appl. Phys.* 107:054513.

Yu SQ, Wang JB, Ding D, Johnson SR, Valsileska D, Zhang YH (2007) Impact of electronic density of states on electroluminescence refrigeration. *Solid State Electron.* 51:1387–1390.

Zhang J, Li D, Chen R, Xiong Q (2013) Laser cooling of a semiconductor by 40 Kelvin. *Nature* 493:504–508.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

45

Photonic Crystal Laser Diodes

Maciej Dems

45.1 Photonic Crystals Basics 561
Optical Periodic Structures • Photonic Bandgap • Photonic Crystals in Lasers
45.2 Fully Vectorial Method for Modeling VCSELs with Subwavelength Structures 565
Plane-Wave Admittance Method • Plane-Wave Reflection Transfer
45.3 VCSELs with Photonic Crystals for Light Confinement..... 570
45.4 Subwavelength Gratings in VCSELs..... 574
High-Contrast Subwavelength Grating Theory • Modeling of Subwavelength Grating Reflectivity • Cavities with Subwavelength Grating Mirrors • Manufacturing Sensitivity of High-Contrast Grating Mirrors
45.5 Summary..... 582

45.1 Photonic Crystals Basics

45.1.1 Optical Periodic Structures

Photonic crystals are periodic structures designed to affect electromagnetic waves in a similar manner to how solid-state crystals affect electrons. The simplest, one-dimensional form of a photonic crystal is a periodic multilayer film (Bragg mirror). Such films were already being studied 120 years ago, in 1887, by Rayleigh [1]. They continued to be the subject of intensive research, resulting in the development of thin-film optics: the science and technology of fabricating dielectric multilayer mirrors, filters, polarizers, antireflection coatings, and so on. However, it took a whole century before the technology was expanded into two and three dimensions by Yablonovitch [2] and John [3] in 1987. Since then, the term *photonic crystal* has become widely used and the subject is studied all over the world.

The physics of photonic crystals has much in common with solid-state physics, as both concern the solution of wave equations in a periodic medium. Hence, concepts known from solid-state physics can also be used in photonics. In particular, for any regular lattice, it is possible to define a *reciprocal lattice* in the Fourier space, i.e., the space of the wavevectors. Such reciprocal lattices are an important tool for the analysis of light propagation in photonic crystals. Because of the Bloch theorem, any eigenmode in periodic medium can be represented as

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_{\mathbf{k}}(\mathbf{r}) \exp(-i\,\mathbf{k}\cdot\mathbf{r}), \tag{45.1}$$

where $\mathbf{k}$ is a wavevector and $\mathbf{E}_{\mathbf{k}}$ is a periodic function with the same periodicity as the refractive index, i.e.,

$$n_{\mathrm{R}}(\mathbf{r}+\mathbf{a}_i) = n_{\mathrm{R}}(\mathbf{r}), \tag{45.2}$$

$$\mathbf{E}_{\mathbf{k}}(\mathbf{r}+\mathbf{a}_i) = \mathbf{E}_{\mathbf{k}}(\mathbf{r}). \tag{45.3}$$

Vectors $\mathbf{a}_i$ are referred to as the base vectors of the periodic lattice and their number is equal to the number of dimensions (in Equations 45.2 and 45.3).

A reciprocal lattice has vectors $\mathbf{b}_j$ defined in the space of wavevectors. They fulfill the condition

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij},$$

where δ_{ij} is Kronecker's delta. For three 3D lattices, $\mathbf{b}_j$ can be expressed directly as

$$\mathbf{b}_1 = 2\pi\frac{\mathbf{a}_2 \times \mathbf{a}_3}{a_1 a_2 a_3}, \qquad \mathbf{b}_2 = 2\pi\frac{\mathbf{a}_3 \times \mathbf{a}_1}{a_1 a_2 a_3}, \text{ and} \qquad \mathbf{b}_3 = 2\pi\frac{\mathbf{a}_1 \times \mathbf{a}_2}{a_1 a_2 a_3},$$

and for 2D lattices as

$$\mathbf{b}_1 = 2\pi\frac{\mathbf{a}_2 \times \hat{\mathbf{n}}}{a_1 a_2}, \qquad \mathbf{b}_2 = 2\pi\frac{\hat{\mathbf{n}} \times \mathbf{a}_1}{a_1 a_2},$$

where $\hat{\mathbf{n}}$ is the unit vector perpendicular to the lattice plane. These vectors define the periodicity of the wavevector $\mathbf{k}$ from Equation 45.1. This periodicity enables the modal analysis of PCs with the vectors $\mathbf{k}$ located within a limited area, called the first Brillouin zone [4].

In the case of two-dimensional photonic crystals, two lattice configurations are particularly common. These are the square lattice and the hexagonal lattice, shown in Figure 45.1a and c. Their reciprocal lattices have the same properties as corresponding physical lattices and are depicted in Figure 45.1b and d. Each of these lattices possesses several high symmetry points, which are important for determination of the photonic band structure. The Γ point is always defined for wavevector $\mathbf{k} = \mathbf{0}$. Others are located at the edges of the first Brillouin zone. For the square lattice, these are X for $\mathbf{k} = \frac{\pi}{a}\hat{\mathbf{x}}$ and M for $\mathbf{k} = \frac{\pi}{a}\hat{\mathbf{x}} + \frac{\pi}{a}\hat{\mathbf{y}}$[†]. In the case of the hexagonal lattice, the high symmetry points are M ($\mathbf{k} = \frac{\pi}{a}\hat{\mathbf{x}} + \frac{\pi}{a\sqrt{3}}\hat{\mathbf{y}}$) and K ($\mathbf{k} = \frac{4\pi}{3a}\hat{\mathbf{x}}$). These positions are given for the points marked in Figure 45.1, although there are other points of each kind on the other edges and corners. In practice, this changes little, as the eigenmodes may be subjected to any symmetry transformation that does not change the distribution of the refractive index. Hence, the triangles made by the adjacent symmetry points are the *irreducible zone*, into which all the possible wavevectors can

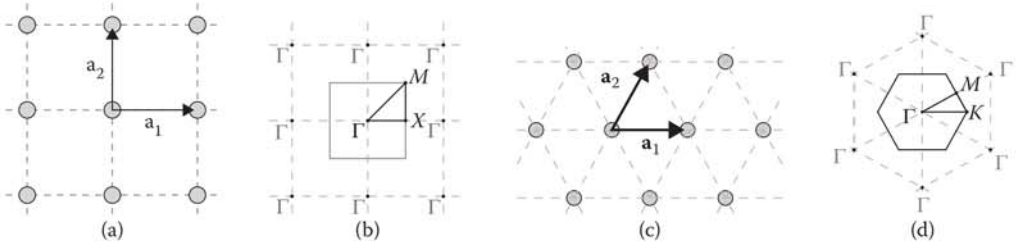


FIGURE 45.1 Arrangements of periodic cells and base vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ of two-dimensional photonic crystals for (a) rectangular and (c) triangular lattices and reciprocal lattices of (b) rectangular and (d) triangular two-dimensional photonic crystals. The important high-symmetry points in the middle and at the edges of the first Brillouin zone (marked with solid lines) are shown. The periodicity of this lattice allows each point outside of the first Brillouin zone to be mapped within it. All the lattice nodes correspond to the Γ point. The triangles formed by the high-symmetry points and shown in the figures are irreducible zones. Each point outside this zone can be mapped to one inside it by periodicity and symmetry transformations.

[†] $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are unit tensors of the Cartesian coordinate system.

be transformed using periodicity and symmetry transformations. As a consequence, analysis of the Bloch modes of a photonic crystal can be limited to this irreducible zone. In order to determine the bandgap, it is even sufficient to compute the frequencies for the wavevectors at the edges of that zone.

45.1.2 Photonic Bandgap

One of the most important properties of photonic crystals is the *photonic bandgap* (PBG), i.e., a range of frequencies in which light cannot propagate through the crystal [5]. If a light wave of such frequency were to propagate into the photonic crystal, it would be almost 100% reflected. This effect is well known in one-dimensional periodic stacks and is commonly used in distributed Bragg reflectors (DBRs, see Chapter 34). Figure 45.2a shows the reflectivity spectrum of a typical quarter-wavelength GaAs/AlAs DBR. The clearly visible high-reflectivity range between 927 nm and 1039 nm is the bandgap. The same effect in two dimensions can be seen in Figure 45.2b, which shows the optical bands as a horizontal wavevector to normalized frequency relation and the density of optical states within an infinite crystal. The bandgap is the region with zero density of states. In this frequency range, there is no band for any wavevector.

Two- and three-dimensional photonic crystals can be considered a multi-dimensional equivalent of DBRs. Hence, one of their major applications is to confine light. This is what they are most commonly used for in vertical-cavity surface-emitting lasers (VCSELs).

In order to confine light using a photonic crystal, one must create a *defect*: a localized perturbation of the periodic structure. Such perturbations can be either linear, in which case we have a waveguide, or isolated, in which case it makes a cavity. Examples of two-dimensional photonic crystal waveguides are shown in Figure 45.3. Because of the reflective properties of the photonic crystal within the bandgap, the waveguides can be bended with sharp angles and still effectively transmit light. In lasers, such waveguides can be coupled with edge-emitting lasers. In the case of surface-emitting lasers, the light is confined within a two-dimensional photonic crystal cavity. The most common configurations of such cavities are shown in Figure 45.4.

45.1.3 Photonic Crystals in Lasers

Photonic crystals are seldom used in edge-emitting lasers. One design of such a laser is presented in [6]. It is a W1 or W3 linear waveguide filled with an active material. Such lasers are reported to be intrinsically

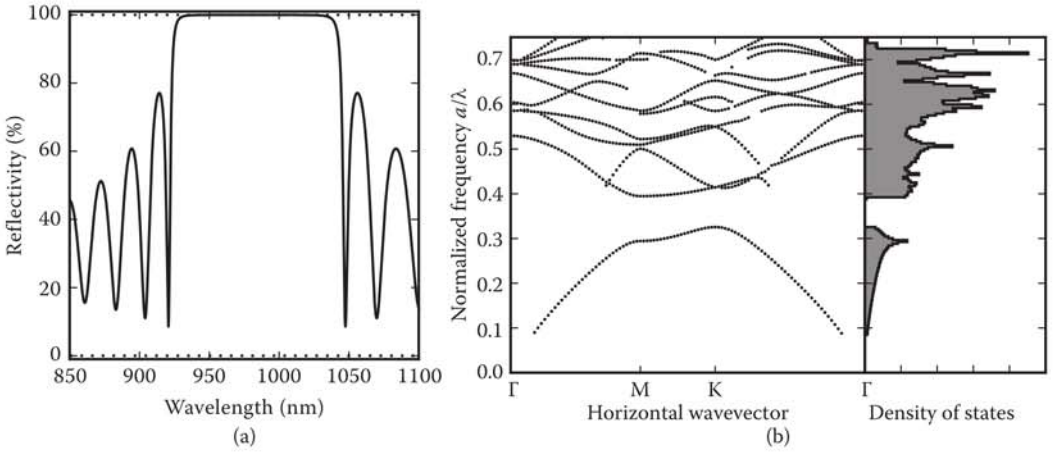


FIGURE 45.2 (a) Reflectivity of a quarter-wavelength one-dimensional stack and (b) band diagram and density of states (right histogram) of a two-dimensional photonic crystal. A photonic bandgap is visible in both cases as the frequency/wavelength region with very high reflectivity or vanishing density of states.

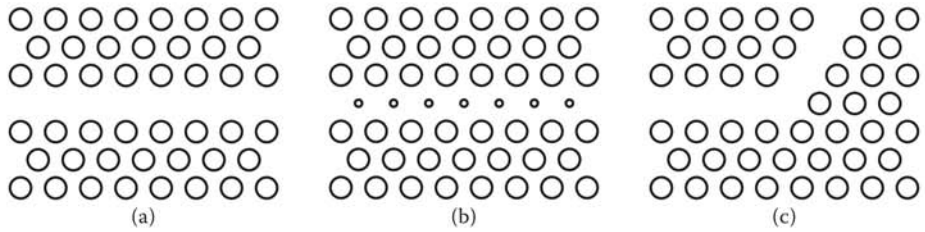


FIGURE 45.3 Examples of photonic crystal waveguides: (a) simple W1 waveguide with one row of missing holes, (b) linear waveguide with smaller holes, and (c) bent waveguide.

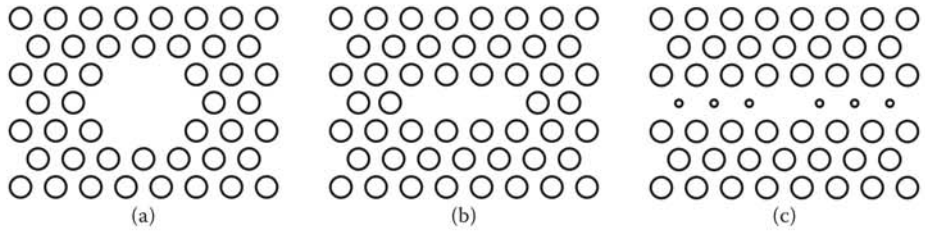


FIGURE 45.4 Examples of photonic crystal cavities: (a) cavity formed by removing seven holes, (b) L3 cavity with shifted edge holes, and (c) birefringent cavity formed by one missing hole and a row of smaller holes.

single-mode, in contrast to classical first-order DFB lasers. Depending on the configuration, the emitting mode has its wavevector either at the Γ or K point of the hexagonal photonic crystal lattice. The emission spectrum has two peaks, corresponding to the edges of the so-called mini-stop-band, which appears when two modes cross (as is the case at both Γ and K points). If emission takes place at frequencies that do not fall within the bandgap of the photonic crystal, only one of these modes is confined to the waveguide. The other spreads infinitely into the photonic crystal, and therefore suffers significantly larger losses. This is the main cause of single-mode operation.

In contrast to edge-emitting lasers, photonic crystals are much more commonly used in surface-emitting lasers. The first experimental designs were published at the end of the nineties. Similarly to photonic-crystal edge-emitting lasers (PC-EELs), such lasers were at first pumped with an optical beam and their design was not very complicated. The simplest surface-emitting photonic crystal laser is a PC membrane with a defect inside. The greatest advantage of such a design is its simplicity and the ease of fabricating a photonic crystal with a small lattice constant. For a typical hexagonal lattice of air holes, etched in a semiconductor material, the pitch must be around $0.3\,\lambda$ if the frequency of the lasing mode c/λ is to be inside the bandgap. Because of the Purcell effect, the small size of the cavity helps to achieve a high spontaneous emission rate and a high Q-factor that can reach up to 10^4 [7].

In most designs of photonic crystal surface-emitting lasers [8–12], the light is not confined by the bandgap, but through the effective index of the defect area, which is higher than that of the surrounding photonic crystal. The role of the photonic crystal is to provide single-mode emission on either a higher order [9] or a fundamental mode [12]. Single-mode emission is one of the most promising applications of photonic crystals in the area of VCSEL design. This is especially true for high-power lasers [13], where classical designs fail because of the necessity of using broad area cavities.

The origin of a single-mode emission in PC-VCSELs can be attributed to one of the two factors. The first is a significant increase of modal losses for the higher-order modes, caused by light scattering. By carefully choosing the etching depth of the holes in the top DBR and their diameter, one can maximize this difference while keeping optical losses for the fundamental mode reasonably low [14,15]. The other factor is a cutoff similar to that found in so-called *endlessly single-mode* fibers [16], i.e., fibers that remain

single-mode for any particular wavelength. The origin of this phenomena lies in the dependence of the PC effective index on the wavelength. The number of modes in a classical fiber depends on the parameter V_{eff}, which is defined as

$$V_{\text{eff}} = \frac{2\pi r}{\lambda}\left(n_{\text{core}}^2 - n_{\text{clad}}^2\right)^{1/2}, \tag{45.4}$$

where r is the radius of the core, λ is the wavelength, and n_{core} and n_{clad} are the effective indexes of the fiber core and the cladding, respectively. The fiber is single-mode if $V_{\text{eff}} < 2.405$. It is possible to design a photonic crystal fiber in which the V_{eff} does not exceed 2.405 even for $\lambda \to 0$. Because the VCSEL cavity can be approximated as a waveguide (although many phenomena are disregarded in such an approximation), Equation 45.4 can be used in this case. In practice, this means that single-mode operation is achieved if the diameters and depths of the holes in the photonic crystal are sufficiently small.

45.2 Fully Vectorial Method for Modeling VCSELs with Subwavelength Structures

Unlike classical oxide-confined VCSELs, discussed in Chapter 34, VCSELs incorporating subwavelength structures cannot be modeled using either the effective frequency method or any other approximate scalar method. The reason for this is the fact that such methods do not adequately consider light scattering from the subwavelength structures. Hence, in order to correctly model the optical properties of such devices, one needs to apply more exact vector methods, such as the finite-difference time-domain method [17], plane-wave expansion [18], the finite element method (FEM) [19], the coupled mode model (CMM) [20], the plane-wave admittance method (PWRT) [21], or plane-wave reflection transfer [22]. The greatest drawback of all these models is the increased computation time and computer memory usage. This cost is, however, necessary to bear, as only the rigorous solution of Maxwell equations provided by these methods allows light scattering from subwavelength structures to be included in the model.

In this chapter, we illustrate the fundamental optical properties of surface-emitting lasers incorporating subwavelength periodic structures, called photonic crystals. First, we need to sketch the principles of the rigorous vector methods used to obtain the results presented later in the chapter. Instead of describing the details of all popular vector models (some of which are discussed elsewhere in this book), we focus on the two that have proved to be the most useful for our purposes: the plane-wave admittance method and PWRT.

45.2.1 Plane-Wave Admittance Method

The plane-wave admittance method is an extension of the vector approach presented in Chapter 34. It uses the same procedure—admittance transfer—to find the eigenmodes of a resonant cavity. However, in this method, the electromagnetic field is expanded in the plane-wave Fourier basis. This basis is particularly useful for periodic structures, as it has natural periodic boundary conditions. Hence, modeling of infinite two-dimensional photonic crystals is straightforward and requires little effort. However, to analyze localized modes in lasers and correctly consider emitted radiation, one needs to introduce absorbing perfectly matched layers at the edges of the computational domain.

As the exact computation procedure was described in Chapter 34, it need not be repeated here. Instead, this section presents a derivation of the matrix form of Maxwell equations using the plane-wave expansion approach. Such matrices can be entered straightforwardly into computer memory for effective computations.

To start the derivation, consider an anisotropic linear dielectric or semiconductor material without any free charges ($\rho = 0$) or currents ($\mathbf{j} = 0$). In such a material, the time-independent Maxwell equations read

$$\nabla \times \mathbf{E} = -i\omega\mu\mu_0\mathbf{H}, \tag{45.5}$$

$$\nabla \times \mathbf{H} = \quad i\omega\varepsilon\varepsilon_0\mathbf{E}, \tag{45.6}$$

where $i^2 = -1$, ω is the angular frequency of light, and physical observables are the real parts of $\mathbf{E}$ and $\mathbf{H}$.

Now consider an anisotropic material in Cartesian coordinates. Assuming that both permittivity and permeability of the material can be represented as diagonal tensors, the Equations 45.5 and 45.6 can be expanded as

$$-\partial_z E_y + \partial_y E_z = -i\omega\mu_x\mu_0 H_x, \tag{45.7}$$

$$\partial_x E_z - \partial_z E_x = -i\omega\mu_y\mu_0 H_y, \tag{45.8}$$

$$-\partial_y E_x + \partial_x E_y = -i\omega\mu_z\mu_0 H_z, \tag{45.9}$$

$$-\partial_z H_y + \partial_y H_z = i\omega\varepsilon_x\varepsilon_0 E_x, \tag{45.10}$$

$$\partial_x H_z - \partial_z H_x = i\omega\varepsilon_y\varepsilon_0 E_y, \tag{45.11}$$

$$-\partial_y H_x + \partial_x H_y = i\omega\varepsilon_z\varepsilon_0 E_z, \tag{45.12}$$

where ∂_ξ means a partial derivative in the ξ direction. These equations can be transformed into the form already introduced in Chapter 34:

$$\partial_z \begin{bmatrix} -E_y \\ E_x \end{bmatrix} = -i\frac{\eta_0}{k_0} \begin{bmatrix} \partial_y\varepsilon_z^{-1}\partial_y + \mu_x k_0^2 & -\partial_y\varepsilon_z^{-1}\partial_x \\ -\partial_x\varepsilon_z^{-1}\partial_y & \partial_x\varepsilon_z^{-1}\partial_x + \mu_y k_0^2 \end{bmatrix} \begin{bmatrix} H_x \\ H_y \end{bmatrix}, \tag{45.13}$$

$$\partial_z \begin{bmatrix} H_x \\ H_y \end{bmatrix} = -i\frac{1}{\eta_0 k_0} \begin{bmatrix} \partial_x\mu_z^{-1}\partial_x + \varepsilon_y k_0^2 & \partial_x\mu_z^{-1}\partial_y \\ \partial_y\mu_z^{-1}\partial_x & \partial_y\mu_z^{-1}\partial_y + \varepsilon_x k_0^2 \end{bmatrix} \begin{bmatrix} -E_y \\ E_x \end{bmatrix}, \tag{45.14}$$

where $k_0 = \omega/c = \omega\left(\mu_0\varepsilon_0\right)^{1/2}$.

Now, consider an electromagnetic field in a structure exhibiting two-dimensional periodicity, i.e.,

$$\varepsilon(\mathbf{r}) = \varepsilon(\mathbf{r} + \mathbf{a}_i) \qquad (i = 1, 2), \tag{45.15}$$

$$\mu(\mathbf{r}) = \mu(\mathbf{r} + \mathbf{a}_i) \qquad (i = 1, 2), \tag{45.16}$$

with $\mathbf{a}_1$ and $\mathbf{a}_2$ being elementary lattice vectors in the xy plane. Because of their periodicity, the solutions of the Maxwell equations obey the Bloch theorem, i.e., both the electric and magnetic fields can be represented as

$$\mathbf{E}(\mathbf{r}) = \bar{\mathbf{E}}(\mathbf{r})\exp(-i\,\mathbf{k}_{xy}\cdot\mathbf{r}), \tag{45.17}$$

$$\mathbf{H}(\mathbf{r}) = \bar{\mathbf{H}}(\mathbf{r})\exp(-i\,\mathbf{k}_{xy}\cdot\mathbf{r}), \tag{45.18}$$

where $\bar{\mathbf{E}}(\mathbf{r})$, $\bar{\mathbf{H}}(\mathbf{r})$ are periodic functions and $\mathbf{k}_{xy}$ is a projection of the mode wavevector into the xy plane. Furthermore, the functions $\bar{\mathbf{E}}(\mathbf{r})$ and $\bar{\mathbf{H}}(\mathbf{r})$ can be expanded into the two-dimensional

Fourier (plane-wave) series as

$$\bar{\mathbf{E}}(\mathbf{r}) = \mathbf{E}^{\mathbf{g}} \exp(i\,\mathbf{g}\cdot\mathbf{r}) = \mathbf{E}^{\mathbf{g}} \left|\varphi_{\mathbf{g}}\right\rangle, \tag{45.19}$$

$$\bar{\mathbf{H}}(\mathbf{r}) = \mathbf{H}^{\mathbf{g}} \exp(i\,\mathbf{g}\cdot\mathbf{r}) = \mathbf{H}^{\mathbf{g}} \left|\varphi_{\mathbf{g}}\right\rangle, \tag{45.20}$$

where the Einstein summation convention[†] is used and $\mathbf{g}$ is the reciprocal lattice vector:

$$\mathbf{g} = l_1\mathbf{b}_1 + l_2\mathbf{b}_2, \tag{45.21}$$

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}, \tag{45.22}$$

with l_1 and l_2 being arbitrary integers and δ_{ij} the Kronecker's delta. As we perform Fourier expansion only in the xy plane, the coefficients $\mathbf{E}^{\mathbf{g}}$ and $\mathbf{H}^{\mathbf{g}}$ are functions of z.

Using the Bloch theorem together with Equations 45.19 and 45.20, we can represent both fields as

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{\mathbf{g}} \left|\varphi_{\mathbf{g}-\mathbf{k}_{xy}}\right\rangle, \tag{45.23}$$

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}^{\mathbf{g}} \left|\varphi_{\mathbf{g}-\mathbf{k}_{xy}}\right\rangle. \tag{45.24}$$

Now consider Equations 45.13 and 45.14. As the Fourier basis is orthonormal, i.e.,

$$\langle\varphi_{\mathbf{g}}|\varphi_{\mathbf{g}'}\rangle = \delta_{\mathbf{g}\mathbf{g}'}, \tag{45.25}$$

we can introduce Equations 45.23 and 45.24 into Equations 45.13 and 45.14 and left-multiply them by $\langle\varphi_{\mathbf{g}-\mathbf{k}_{xy}}|$ to obtain

$$\partial_z \begin{bmatrix} E_y^{\mathbf{g}} \\ E_x^{\mathbf{g}} \end{bmatrix} = -i\frac{\eta_0}{k_0} \left\langle \varphi_{\mathbf{g}-\mathbf{k}_{xy}} \right| \begin{matrix} \partial_y\varepsilon_z^{-1}\partial_y + \mu_x k_0^2 & -\partial_y\varepsilon_z^{-1}\partial_x \\ -\partial_x\varepsilon_z^{-1}\partial_y & \partial_x\varepsilon_z^{-1}\partial_x + \mu_y k_0^2 \end{matrix} \left| \varphi_{\mathbf{g}'-\mathbf{k}_{xy}} \right\rangle \begin{bmatrix} H_x^{\mathbf{g}'} \\ H_y^{\mathbf{g}'} \end{bmatrix}, \tag{45.26}$$

$$\partial_z \begin{bmatrix} H_x^{\mathbf{g}'} \\ H_y^{\mathbf{g}'} \end{bmatrix} = -\frac{i}{\eta_0 k_0} \left\langle \varphi_{\mathbf{g}-\mathbf{k}_{xy}} \right| \begin{matrix} \partial_x\mu_z^{-1}\partial_x + \varepsilon_y k_0^2 & \partial_x\mu_z^{-1}\partial_y \\ \partial_y\mu_z^{-1}\partial_x & \partial_y\mu_z^{-1}\partial_y + \varepsilon_x k_0^2 \end{matrix} \left| \varphi_{\mathbf{g}'-\mathbf{k}_{xy}} \right\rangle \begin{bmatrix} E_y^{\mathbf{g}} \\ E_x^{\mathbf{g}} \end{bmatrix}, \tag{45.27}$$

where, in order to simplify notation, we represent $E_y(\mathbf{r})$ as

$$E_y(\mathbf{r}) = -E_y^{\mathbf{g}} \left|\varphi_{\mathbf{g}-\mathbf{k}_{xy}}\right\rangle \tag{45.28}$$

instead of directly using Equation 45.23.

The matrices in the above equations can be computed easily when both permittivity and permeability are expanded in the Fourier basis. For this purpose, the analyzed layer is assumed to be invariant in the

[†] By this convention, if an index is repeated twice in any product, we perform summation over this index. We use this convention in this whole chapter, unless explicitly stated otherwise.

z-direction. Then, we can write

$$\varepsilon_i(\mathbf{r}) = \varepsilon_i^{\mathbf{g}} \left|\varphi_{\mathbf{g}}\right\rangle \quad (i = x, y), \tag{45.29}$$

$$\mu_i(\mathbf{r}) = \mu_i^{\mathbf{g}} \left|\varphi_{\mathbf{g}}\right\rangle \quad (i = x, y), \tag{45.30}$$

$$\varepsilon_z^{-1}(\mathbf{r}) = \kappa^{\mathbf{g}} \left|\varphi_{\mathbf{g}}\right\rangle, \tag{45.31}$$

$$\mu_z^{-1}(\mathbf{r}) = \gamma^{\mathbf{g}} \left|\varphi_{\mathbf{g}}\right\rangle. \tag{45.32}$$

Substituting this into Equations 45.26 and 45.27, we finally obtain the matrix form of the equations in the plane-wave basis:

$$\partial_z \begin{bmatrix} E_y^{\mathbf{g}} \\ E_x^{\mathbf{g}} \end{bmatrix} = -i\frac{\eta_0}{k_0} \begin{bmatrix} -(g_y - k_y)(g'_y - k_y)\kappa^{\mathbf{g}-\mathbf{g}'} + k_0^2\mu_x^{\mathbf{g}-\mathbf{g}'} & (g_y - k_y)(g'_x - k_x)\kappa^{\mathbf{g}-\mathbf{g}'} \\ (g_x - k_x)(g'_y - k_y)\kappa^{\mathbf{g}-\mathbf{g}'} & -(g_x - k_x)(g'_x - k_x)\kappa^{\mathbf{g}-\mathbf{g}'} + k_0^2\mu_y^{\mathbf{g}-\mathbf{g}'} \end{bmatrix} \begin{bmatrix} H_x^{\mathbf{g}'} \\ H_y^{\mathbf{g}'} \end{bmatrix}, \tag{45.33}$$

$$\partial_z \begin{bmatrix} H_x^{\mathbf{g}} \\ H_y^{\mathbf{g}} \end{bmatrix} = -\frac{i}{\eta_0 k_0} \begin{bmatrix} -(g_x - k_x)(g'_x - k_x)\gamma^{\mathbf{g}-\mathbf{g}'} + k_0^2\varepsilon_y^{\mathbf{g}-\mathbf{g}'} & -(g'_x - k_x)(g'_y - k_y)\gamma^{\mathbf{g}-\mathbf{g}'} \\ -(g_y - k_y)(g'_x - k_x)\gamma^{\mathbf{g}-\mathbf{g}'} & -(g_y - k_y)(g'_y - k_y)\gamma^{\mathbf{g}-\mathbf{g}'} + k_0^2\varepsilon_x^{\mathbf{g}-\mathbf{g}'} \end{bmatrix} \begin{bmatrix} E_y^{\mathbf{g}'} \\ E_x^{\mathbf{g}'} \end{bmatrix}, \tag{45.34}$$

where k_x, k_y, g_x, and g_y are the corresponding components of the mode wavevector and the reciprocal lattice vector $\mathbf{g}$, respectively. These equations can be introduced into computer memory for further computations by truncating the Fourier basis $\{\mathbf{g}\}$ at some point. The introduced error will depend on two factors—the distribution of the electromagnetic field and the distribution of the material parameters. The former, as a result of simulation, cannot be influenced directly, although the latter can be convoluted with a Gaussian window to improve convergence.

The vertical components of electric and magnetic fields can also be expanded in the Fourier basis as

$$E_z^{\mathbf{g}} = \frac{\eta_0}{k_0}\kappa^{\mathbf{g}-\mathbf{g}'}\left[-(g'_y - k_y)H_x^{\mathbf{g}'} + (g'_x - k_x)H_y^{\mathbf{g}'}\right], \tag{45.35}$$

$$H_z^{\mathbf{g}} = \frac{1}{k_0\eta_0}\gamma^{\mathbf{g}-\mathbf{g}'}\left[(g'_x - k_x)E_y^{\mathbf{g}'} + (g'_y - k_y)E_x^{\mathbf{g}'}\right]. \tag{45.36}$$

Once Equations 45.33 and 45.34 have been derived and represented in their numerical forms, we can use them to determine the eigenmodes of the analyzed structure. For this purpose, we first show their solution for a single z-invariant layer and then present a numerically stable technique for combining fields in all layers and deriving an eigenmode condition for the whole structure. The procedure described below was originally used with the method of lines [23] and is also applied to the plane-wave admittance method.

Further in this section, we will use the short representation of the vectors and matrices introduced in Equations 45.33 and 45.34. In particular, we will name

$$\hat{\mathbf{E}} = \begin{bmatrix} E_y^{\mathbf{g}} \\ E_x^{\mathbf{g}} \end{bmatrix},$$

$$\hat{\mathbf{H}} = \begin{bmatrix} H_x^{\mathbf{g}} \\ H_y^{\mathbf{g}} \end{bmatrix},$$

$$\mathbf{R}_\mathbf{E} = \frac{1}{\eta_0 k_0}\begin{bmatrix} -(g_x - k_x)(g'_x - k_x)\gamma^{\mathbf{g}-\mathbf{g}'} + k_0^2\varepsilon_y^{\mathbf{g}-\mathbf{g}'} & -(g'_x - k_x)(g'_y - k_y)\gamma^{\mathbf{g}-\mathbf{g}'} \\ -(g_y - k_y)(g'_x - k_x)\gamma^{\mathbf{g}-\mathbf{g}'} & -(g_y - k_y)(g'_y - k_y)\gamma^{\mathbf{g}-\mathbf{g}'} + k_0^2\varepsilon_x^{\mathbf{g}-\mathbf{g}'} \end{bmatrix},$$

$$\mathbf{R}_\mathbf{H} = \frac{\eta_0}{k_0}\begin{bmatrix} -(g_y - k_y)(g'_y - k_y)\kappa^{\mathbf{g}-\mathbf{g}'} + k_0^2\mu_x^{\mathbf{g}-\mathbf{g}'} & (g_y - k_y)(g'_x - k_x)\kappa^{\mathbf{g}-\mathbf{g}'} \\ (g_x - k_x)(g'_y - k_y)\kappa^{\mathbf{g}-\mathbf{g}'} & -(g_x - k_x)(g'_x - k_x)\kappa^{\mathbf{g}-\mathbf{g}'} + k_0^2\mu_y^{\mathbf{g}-\mathbf{g}'} \end{bmatrix}.$$

It is worth noting that the whole procedure described below can be applied with any correct representation of the vectors and matrices above. Thus, it is applicable not only to plane-wave expansion, but also to any orthogonal or non-orthogonal basis. In particular, the representation of $\hat{\mathbf{E}}$, $\hat{\mathbf{H}}$, $\mathbf{R}_\mathbf{E}$, and $\mathbf{R}_\mathbf{H}$ by finite-difference approximation is used in the method of lines.

Take Equations 45.33 and 45.34 and represent them in their short form:

$$\partial_z \hat{\mathbf{E}} = -i\mathbf{R}_\mathbf{H}\hat{\mathbf{H}}, \tag{45.37}$$

$$\partial_z \hat{\mathbf{H}} = -i\mathbf{R}_\mathbf{E}\hat{\mathbf{E}}. \tag{45.38}$$

By taking the z-derivative of one of these equations and substituting the other into it, one can write two second-order equations with the fields decoupled:

$$\partial_z^2 \hat{\mathbf{E}} = -\mathbf{Q}_\mathbf{E}\hat{\mathbf{E}}, \tag{45.39}$$

$$\partial_z^2 \hat{\mathbf{H}} = -\mathbf{Q}_\mathbf{H}\hat{\mathbf{H}}, \tag{45.40}$$

where $\mathbf{Q}_\mathbf{E} = \mathbf{R}_\mathbf{H}\mathbf{R}_\mathbf{E}$ and $\mathbf{Q}_\mathbf{H} = \mathbf{R}_\mathbf{E}\mathbf{R}_\mathbf{H}$.

The eigenmodes can be found using the procedure described in Chapter 34 by constructing and solving equation $\mathbf{Y}\tilde{\mathbf{E}} = 0$.

45.2.2 Plane-Wave Reflection Transfer

PWRT differs from the plane-wave admittance method in the way the diagonalized Equations 45.33 and 45.34 are solved. In this method, the electric field in each layer uniform in z-direction is chosen to have the form

$$\tilde{\mathbf{E}} = \exp(-i\boldsymbol{\Gamma}z)\tilde{\mathbf{F}} + \exp(i\boldsymbol{\Gamma}z)\tilde{\mathbf{B}}, \tag{45.41}$$

where $\tilde{\mathbf{F}}$ denotes the amplitude of the forward propagating wave and $\tilde{\mathbf{B}}$ the amplitude of the backward propagating wave. The reflection transfer means an iterative determination of the reflection matrix $\mathbf{R}$, which is defined in each layer as a relation between $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{B}}$ such that

$$\tilde{\mathbf{F}} = \mathbf{R}\,\tilde{\mathbf{B}}. \tag{45.42}$$

We assume that in the first layer, which has infinite thickness, there is no in-going field (we will introduce any optional incident field in the last layer). Thus, $\tilde{\mathbf{F}} \equiv 0$ and so $\mathbf{R} = \mathbf{0}$. The reflection matrices in the subsequent layers can be determined in an iterative process using the requirement that the electric and magnetic fields are continuous at layer interfaces.

By careful selection of units in the magnetic field, it is possible to show that in the diagonalized domain it takes the form of

$$\tilde{\mathbf{H}} = \exp(-i\boldsymbol{\Gamma}z)\tilde{\mathbf{F}} - \exp(i\boldsymbol{\Gamma}z)\tilde{\mathbf{B}}, \tag{45.43}$$

which can be derived from Equation 45.41. The requirement of field continuity between the n-th and $(n+1)$th layers can be expressed as

$$\mathbf{T_E}^n \left\{ \boldsymbol{\varphi}^n \tilde{\mathbf{F}}^n + \left(\boldsymbol{\varphi}^n\right)^{-1} \tilde{\mathbf{B}}^n \right\} = \mathbf{T_E}^{n+1} \left\{ \tilde{\mathbf{F}}^{n+1} + \tilde{\mathbf{B}}^{n+1} \right\}, \tag{45.44}$$

$$\mathbf{T_H}^n \left\{ \boldsymbol{\varphi}^n \tilde{\mathbf{F}}^n - \left(\boldsymbol{\varphi}^n\right)^{-1} \tilde{\mathbf{B}}^n \right\} = \mathbf{T_H}^{n+1} \left\{ \tilde{\mathbf{F}}^{n+1} - \tilde{\mathbf{B}}^{n+1} \right\}, \tag{45.45}$$

where $\boldsymbol{\varphi}^n = \exp\left(-i\boldsymbol{\Gamma} d_n\right)$ and d_n is the thickness of nth layer. Using the relation (Equation 45.42), we obtain

$$\tilde{\mathbf{F}}^{n+1} + \tilde{\mathbf{B}}^{n+1} = \mathbb{E}^n \left(\boldsymbol{\varphi}^n\right)^{-1} \tilde{\mathbf{B}}^n, \tag{45.46}$$

$$\tilde{\mathbf{F}}^{n+1} - \tilde{\mathbf{B}}^{n+1} = \mathbb{H}^n \left(\boldsymbol{\varphi}^n\right)^{-1} \tilde{\mathbf{B}}^n, \tag{45.47}$$

where

$$\mathbb{E}^n = \left(\mathbf{T_E}^{n+1}\right)^{-1} \mathbf{T_E}^n \left\{ \boldsymbol{\varphi}^n \mathbf{R}^n \boldsymbol{\varphi}^n + \mathbf{I} \right\}, \tag{45.48}$$

$$\mathbb{H}^n = \left(\mathbf{T_H}^{n+1}\right)^{-1} \mathbf{T_H}^n \left\{ \boldsymbol{\varphi}^n \mathbf{R}^n \boldsymbol{\varphi}^n - \mathbf{I} \right\}. \tag{45.49}$$

This, after simple algebra, gives the reflection matrix for the $(n+1)$th layer in the form

$$\mathbf{R}^{n+1} = \left[\mathbb{E}^n + \mathbb{H}^n\right] \left[\mathbb{E}^n - \mathbb{H}^n\right]^{-1}. \tag{45.50}$$

In contrast to plane-wave admittance transfer, we need not assume the electric field at the boundaries is equal to zero and, hence, there is no need to use absorbing perfectly matched layers as vertical boundary conditions. Instead, at the edge of the structure, we can assume the $\mathbf{R}^0$ matrix to be equal to 0. This means that in the outermost layer, there is no forward propagating field. This field corresponds to the wave coming into the structure. As the analyzed device is a laser resonant cavity, it only emits light; hence, $\tilde{\mathbf{F}}^0 = \mathbf{0}$ and $\tilde{\mathbf{B}}^0 \neq \mathbf{0}$. Then, similarly to PWAM, we can perform iterative progressions of the reflection matrix from both top and bottom boundaries, up to a selected layer. In this, the following relations hold:

$$\tilde{\mathbf{F}}^{\text{top}} = \tilde{\mathbf{B}}^{\text{bottom}}, \tag{45.51}$$

$$\tilde{\mathbf{B}}^{\text{top}} = \tilde{\mathbf{F}}^{\text{bottom}}. \tag{45.52}$$

Using this condition, it is possible to construct a matching matrix $\mathbf{M}$ such that

$$\mathbf{M}\tilde{\mathbf{F}}^{\text{bottom}} = \left[\mathbf{R}^{\text{top}} - \left(\mathbf{R}^{\text{bottom}}\right)^{-1}\right] \tilde{\mathbf{F}}^{\text{bottom}} = \mathbf{0}.$$

In other words, this matrix must be singular, which is the condition for finding the eigenmodes.

45.3 VCSELs with Photonic Crystals for Light Confinement

In VCSELs, photonic crystals are used to provide effective light confinement. By etching a regular lattice of holes in either the DBRs or the VCSEL cavity (or both), together with intentional lattice defects, it is possible to mold the optical mode distribution. In particular, one can laterally confine the fundamental laser mode within a photonic crystal defect without adding any more lowindex oxidation layers. There are several advantages of this approach: First, it allows the fabrication of VCSELs with well-defined cavities in

material systems where selective oxidation is either impossible or difficult to achieve. Second, nonuniform patterns in the periodic lattice can be tuned to interact differently with fundamental and high-order modes, providing increased mode selectivity and allowing single-mode VCSEL operation at higher currents (see Section 45.1.3).

To illustrate the exclusive optical properties of PC-VCSELs, consider a basic optical-only VCSEL structure made of GaAs, as shown in Figure 45.5a. It contains 24 GaAs/AlAs pairs of top DBRs, 29 pairs of GaAs/AlAs bottom DBRs, and a very short single-wavelength GaAs cavity. The exact layer structure is listed in Table 45.1. The photonic crystal is based on a triangular lattice (of pitch L) and takes the form of three hexagons surrounding a small cavity in the form of a missing PC hole. Holes with diameter d are etched either in the top DBRs (with varying etching depths), throughout the whole structure, or only inside the cavity.

We now investigate the impact of photonic crystal parameters on the properties of the fundamental mode in the VCSEL in question. Due to its Gaussian-like shape, this mode is the most preferred in most applications. Furthermore, VCSELs with incorporated photonic crystals are known for their high mode selectivity [24], which makes it possible to tune them into a single-mode source with high power emission.

Our results are depicted in Figures 45.6 through 45.9. They express typical tendencies observed in PC-VCSELs. Figure 45.6 shows that increasing the size of the holes results in a decrease in the resonant wavelength, which can be explained by two factors. The first is the decrease in the area of the photonic crystal defect, which forms the cavity (corresponding to a decrease in the aperture of a classical VCSEL).

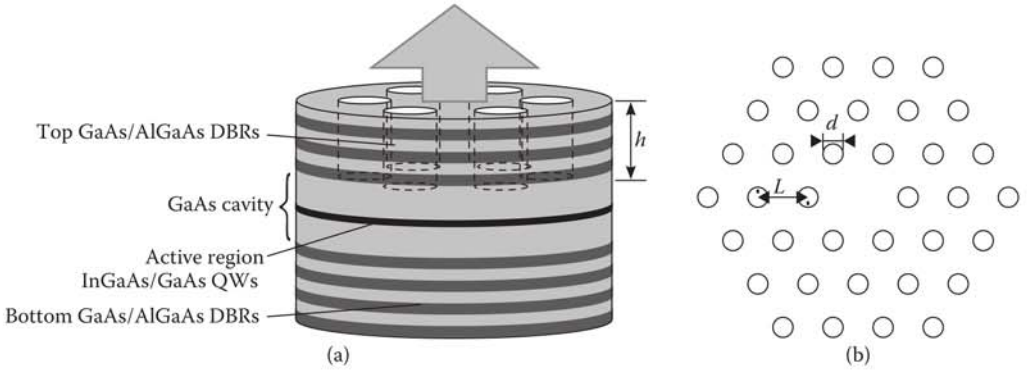


FIGURE 45.5 Analyzed VCSEL structure: (a) schematic three-dimensional view, (b) PC structure. (DBRs, distributed Bragg reflectors; QWs, quantum wells.)

TABLE 45.1 Details of the Layer Structure of the Analyzed VCSEL.

	Thickness [nm]	Material	Refractive index
		Air	1.00
Top DBR	69.40	GaAs	3.53
24 pairs	79.55	AlGaAs	3.08
	121.71	GaAs	3.53
Cavity	$3 \times 8.00 + 2 \times 5.00$	QW/GaAs	$\begin{cases} 3.56 + j n_g & \text{for } r < a \\ 3.56 - 0.01j & \text{for } r \geq a \end{cases}$ / 3.53
	121.71	GaAs	3.53
Bottom DBR	79.55	AlGaAs	3.08
29 pairs	69.40	GaAs	3.53
Substrate		GaAs	3.53

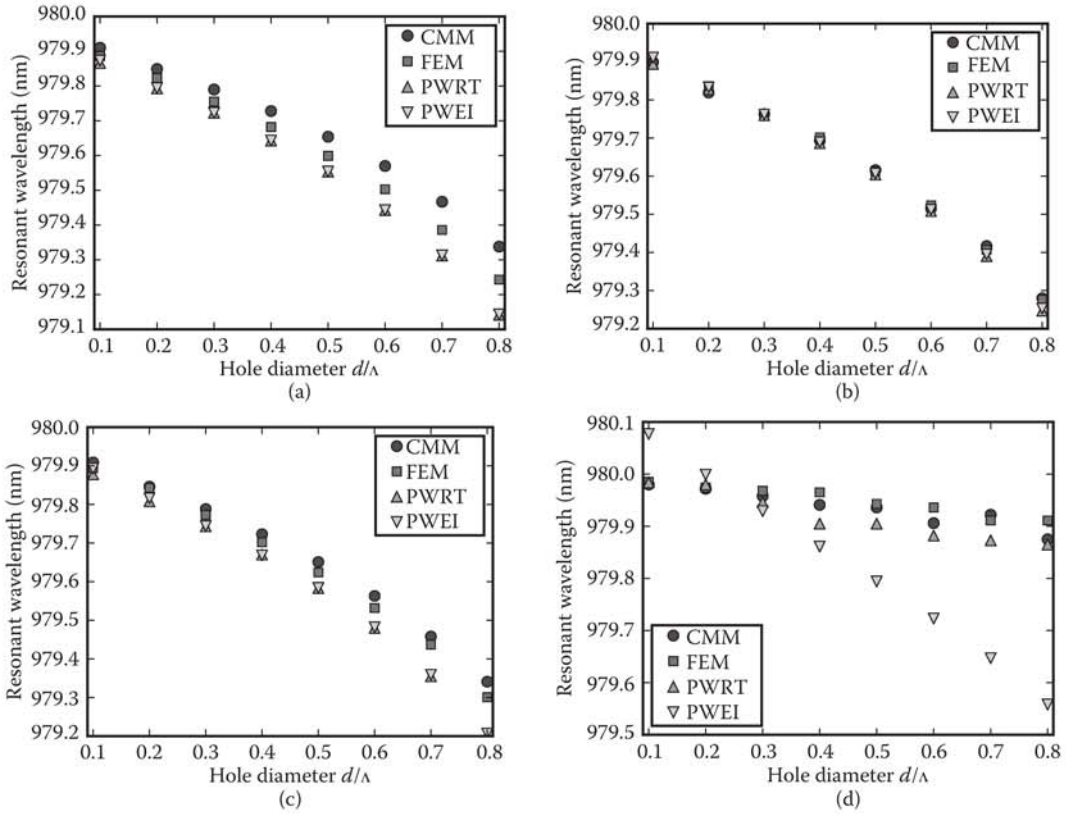


FIGURE 45.6 Resonant wavelengths of PC-VCSELs with photonic crystals etched in different parts of the structure ((a) whole structure, (b) cavity only, (c) 24 pairs of the top DBR, (d) 10 pairs of the top DBR) as a function of hole diameters d. The photonic crystal pitch Λ is 4 μm in all the plots.

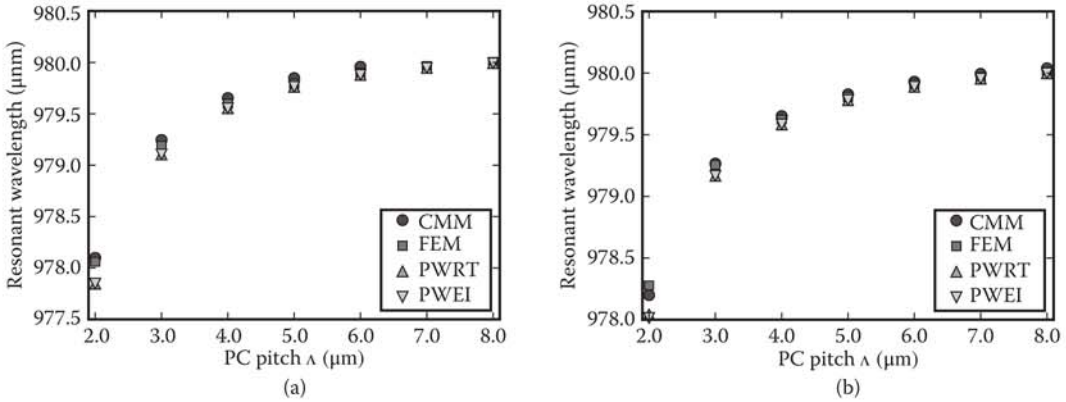


FIGURE 45.7 Resonant wavelengths of PC-VCSELs with photonic crystals etched in different parts of the structure ((a) whole structure, (b) cavity only, (c) 24 pairs of the top DBR, (d) 10 pairs of the top DBR) as a function of the photonic crystal pitch Λ. The diameter of the etched holes d is proportional to the pitch and equals 0.5 Λ.

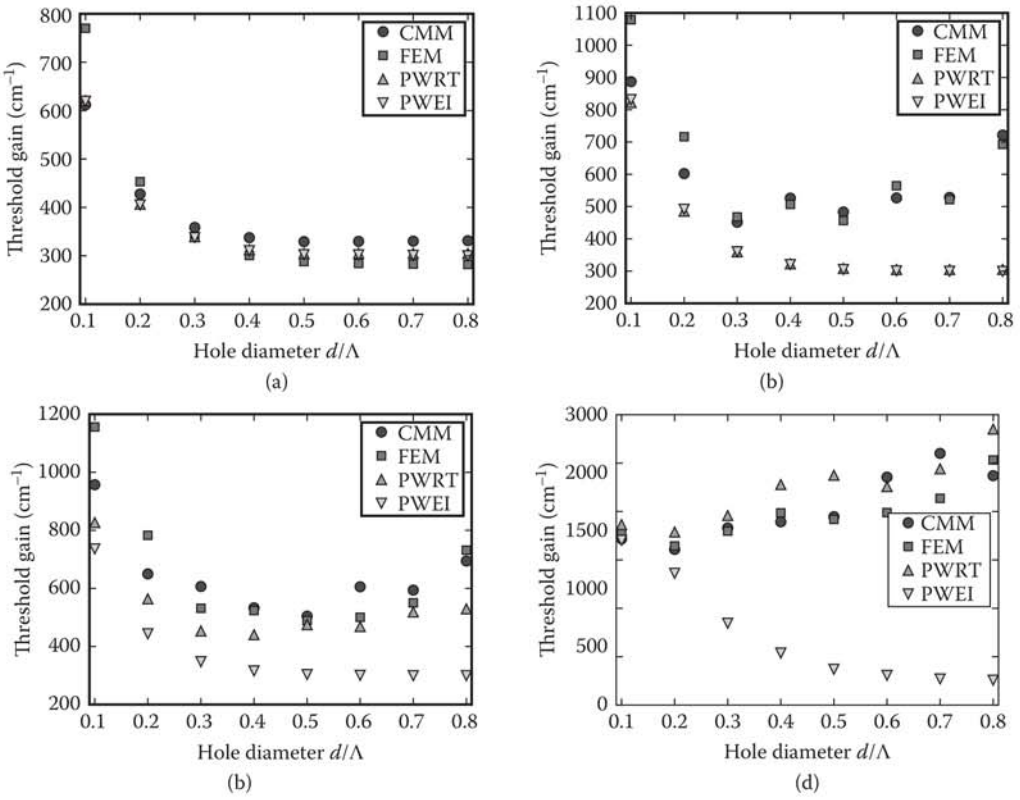


FIGURE 45.8 Threshold gains of PC-VCSELs with photonic crystals etched in different parts of the structure ((a) whole structure, (b) cavity only, (c) 24 pairs of the top DBR, (d) 10 pairs of the top DBR) as a function of the hole diameters d. The photonic crystal pitch Λ is 4 μm in all the plots.

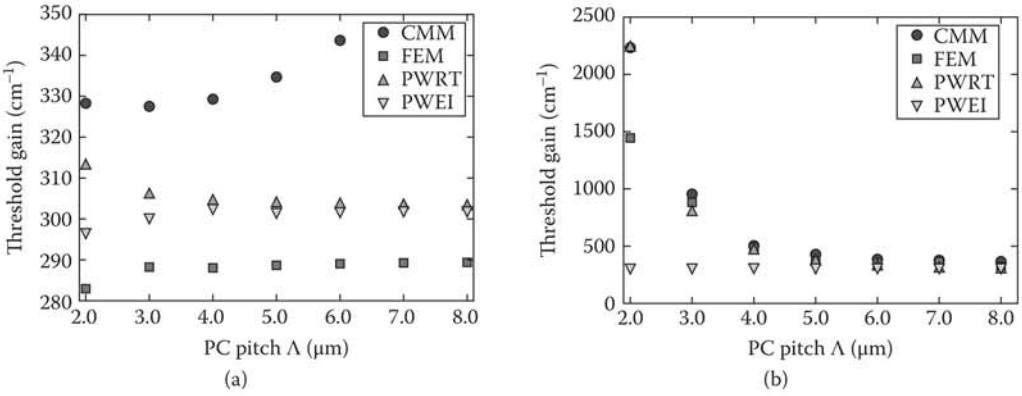


FIGURE 45.9 Threshold gains of PC-VCSELs with photonic crystals etched in different parts of the structure ((a) whole structure, (b) cavity only, (c) 24 pairs of the top DBR, (d) 10 pairs of the top DBR) as a function of the photonic crystal pitch Λ. The diameter of the etched holes d is proportional to the pitch and equals 0.5 Λ.

The other factor is the increase of the effective refractive index contrast between the cavity and the photonic crystal surrounding it. Larger holes mean a larger average air-to-semiconductor ratio, which results in a lower effective index outside the cavity.

The opposite tendency is shown in Figure 45.7. Increasing the photonic crystal pitch results in a rapid increase in the resonant wavelength, as the cavity area increases in proportion to Λ. Due to the constant d/Λ ratio, the effective index contrast remains unchanged. The change is fastest for smaller pitches, where the light confinement is high. With increasing Λ, it increases asymptotically to 980 nm, which is the resonant wavelength of the presented VCSEL with no lateral confinement at all.

Much more interesting is the behavior of the threshold gain. In each case, it is large for small holes (Figure 45.8) because of the weak confinement and large overlap of the mode with the absorbing areas of the active region outside the gain aperture. As the size of the holes increases, the mode is increasingly better confined and therefore the threshold decreases. However, with shallow etching (Figure 45.8b through d), scattering starts to play an important role, and with larger holes the threshold starts increasing again. On the other hand, the photonic crystal pitch does not influence the threshold gain significantly (Figure 45.9), except when it drops too strongly (below 4 μm) in structures with holes etched only in the DBR (Figure 45.9b), in which case light scattering starts to play an important role.

In each case, the same optical properties were computed using four different numerical methods: PWRT presented above, the CMM [20], the FEM [25], and the plane-wave effective index (PWEI) [26]. These methods have already been rigorously compared [27]. Here, we show them to emphasize that the three-dimensional structure of a PC-VCSEL is complex to analyze numerically, even nowadays. Practically all of the methods have shown behavior qualitatively different to the others in at least one case. This shows how important it is to calibrate numerical models and to choose the right one, which will yield correct results for a particular case.

In most circumstances, however, numerical analysis allows some conclusions to be made regarding the optical properties of PC-VCSELs. The results depicted in Figures 45.6 and 45.8 show computed resonant wavelengths and threshold gains as functions of the hole diameters when the photonic crystal pitch Λ equals 4 μm. A pitch of 4 μm is a typical value and has been kept fixed for all etchings. Figures 45.7 and 45.9 show the results obtained from different sizes of the photonic crystal lattice constant, with holes etched through the whole cavity and all layers of the top DBR.

In most cases, the PC-VCSEL resonant wavelength is predicted consistently by all numerical methods. The only exception is with large, shallow etched holes, as can be seen in Figure 45.6d. The differences between the vector methods (CMM, FEM, and PWRT) and the scalar method (PWEI), which gives a too small wavelength, are clearly noticeable. The origin of these discrepancies is the significant light scattering at the bottoms of the holes, which is correctly considered by the vector methods, but neglected by the scalar method.

Unsurprisingly, greater differences are observable in the threshold gain data. The methods show good agreement only in case of the $\Lambda = 4$ μm wholly etched structure. All other situations require more detailed analysis. Figure 45.8c shows the threshold gains for holes etched only in the top DBR. Due to their complicated arrangement, significant scattering occurs at their edges, which is reflected in an increase in the threshold gain for $d/\Lambda > 0.4$. This effect is not predicted by PWEI, which—as a variation of the effective index method—is incapable of properly considering scattering losses. PWEI even more significantly underestimates the threshold gain for shallow holes (Figure 45.6d), where scattering is much stronger due to the weaker mode confinement.

45.4 Subwavelength Gratings in VCSELs

In the first part of this chapter, we described the two-dimensional photonic crystals used to provide light confinement of the optical mode. However, periodic subwavelength structures can offer much more. In what follows, we will look at specific one-dimensional photonic crystals: subwavelength gratings with high

refractive index contrast (referred to as high-contrast gratings or HCGs). Such structures may effectively replace DBR mirrors in VCSELs, as they can provide very high reflectivity, while being significantly thinner than a multilayer DBR stack.

Because the light emitted by most VCSELs is linearly polarized, one-dimensional linear gratings can be efficiently applied. Furthermore, they ensure stable polarization of the emitted light, which otherwise might suddenly change at higher output powers [28]. For the same reason, circular gratings are much less suitable as mirror replacements for VCSELs, as they cannot provide uniform reflectivity for linearly polarized light.

In order to investigate the properties of VCSELs with subwavelength gratings, we first need to understand the behavior of grating mirrors. For this reason, in the next subsection, we will outline the basic physical principles behind the high reflectivity of HCGs. Next, we will study their reflectivity in various configurations. After that, we will move on to the determination of resonant modes in VCSEL cavities with subwavelength grating.

45.4.1 High-Contrast Subwavelength Grating Theory

Despite their very slight thicknesses—usually only a few hundreds of nanometers—HCGs can be effective mirrors with reflectivity comparable or exceeding that of thick multilayered DBRs. This behavior can be explained in two ways: either as an effect of the Fano resonance between the incident wave and the excited mode propagating inside the grating [29] or as an effect of destructive interference between transmitted and leaky modes [30]. Below, we briefly summarize the latter explanation.

Consider a grating suspended in air. It can be seen as a periodic array of waveguides, propagating light in a direction perpendicular to the grating plane (Figure 45.10). Assume an incident plane wave propagating in the z-direction and polarized parallel to the grating bars (with the nonzero E_y component). At the incidence side of the grating (region 1 in Figure 45.10), the electric field can be described as the single zero-order incident plane wave and multiple orders of the reflected wave:

$$E_1(x,z) = E_I \exp\left(-ik_0 z\right) + \sum_n E_R^n \cos\left(2n\pi x/L\right) \exp\left(+i\sqrt{k_0^2 - (2n\pi/L)^2}\, z\right), \tag{45.53}$$

where E_I is the amplitude of the incident field, E_R^n is the horizontal distribution of the reflected modes of different orders, $k_0 = \omega/c$ is the normalized frequency, and L is the grating period. Inside the grating, the field is the sum of forward and backward propagating waveguide array modes:

$$E_2(x,z) = \sum_m \left[E_F^m(x) \exp\left(-i\beta_m z\right) + E_B^m(x) \exp\left(+i\beta_m z\right)\right], \tag{45.54}$$

where β_m is the propagation constant of the mth mode and E_F^m and E_B^m are amplitudes of the forward and backward propagating modes, respectively. Finally, in the last section, the field consists only of the

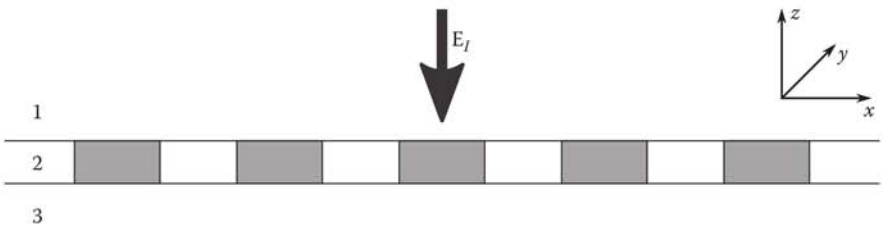


FIGURE 45.10 Grating as a periodic waveguide array.

transmitted modes with amplitudes E_T^n:

$$E_3(x,z) = \sum_n E_T^n \cos(2n\pi x/L) \exp\left(-i\sqrt{k_0^2 - (2n\pi/L)^2}\, z\right).$$

For subwavelength gratings $L < \lambda = 2\pi/k_0$, only the zero orders of both the reflected and transmitted modes propagate and the other modes are evanescent. Although they do not transport any energy, they are very important for ensuring proper boundary conditions for Maxwell equations. By applying these boundary conditions to both grating boundaries, it is possible to determine the amplitudes of both the reflected and transmitted modes (E_R^n and E_T^n) and the modes inside the grating (E_F^m and E_B^m). As only one transmitted mode (E_T^0) propagates energy, it is possible to tune the grating parameters in such a way that $E_T^0 = 0$. This corresponds to the zero energy transmission equivalent, as there is no gain inside the grating, up to 100% reflectivity.

For gratings operating in a near-subwavelength regime, the periodic waveguide array supports only two even modes with propagation constants β_0 and β_2 [31]. In such a case, E_T^0 can be derived analytically as

$$E_T^0 = \frac{1}{L}\int_0^L \left\{E_F^0(x) + E_B^0(x)\right\} dx + \frac{1}{L}\int_0^L \left\{E_F^2(x) + E_B^2(x)\right\} dx.$$

In other words, E_T^0 is a sum of the spatial averages of both waveguide array modes. If these averages are antimatched, i.e.,

$$\int_0^L \left\{E_F^0(x) + E_B^0(x)\right\} dx = -\int_0^L \left\{E_F^2(x) + E_B^2(x)\right\} dx,$$

there is no energy transmission through the grating. This corresponds to the destructive interference of these modes at the far edge of the grating.

For a fuller derivation and discussion of the modal properties of subwavelength gratings, we refer the reader to [31], where the interaction between gratings and light waves is described in detail.

45.4.2 Modeling of Subwavelength Grating Reflectivity

The reflectivity of subwavelength HCGs can be studied using the PWRT method presented at the beginning of this chapter. This method is mathematically equivalent to the popular rigorous coupled-wave analysis (RCWA) [32]. Both methods rely on expansion of the electromagnetic field into a complex Fourier series and both require the analyzed structure to be split into a set of layers, each of which is uniform in the surface normal direction (see Figure 45.11). PWRT, which was used to obtain all the results presented in this chapter, is applied in the same way as for cavity analysis, as described in Section 45.2.2 with the last step omitted. Instead, after determining the reflection matrix $\bar{\mathbf{R}}$, this is multiplied by a predefined incidence vector $\mathbf{F}$ to calculate the vector of reflected diffraction orders $\mathbf{B}$:

$$\mathbf{B} = \bar{\mathbf{R}}\,\mathbf{F}. \tag{45.55}$$

In most cases, the incident vector has only one nonzero component, which corresponds to the incident plane wave of the zero-order. It is possible to construct more complex incidence vectors, e.g., one forming a Gaussian beam. However, for subwavelength gratings the reflection of the zero-order plane wave is dominant and the results for a Gaussian beam do not differ significantly from those for a plane wave.

Figure 45.11 shows a typical single-period cell of a subwavelength grating. It is comprised of parallel bars made of a high-refractive-index material (usually Si or GaAs), surrounded on both sides by low-index

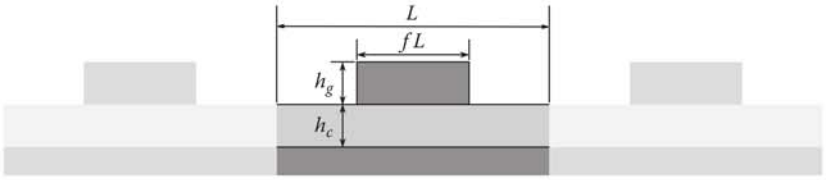


FIGURE 45.11 Schematic diagram of the computational domain for grating reflection computation. The structure is separated into a number of layers uniform in the perpendicular direction and only one grating period is analyzed.

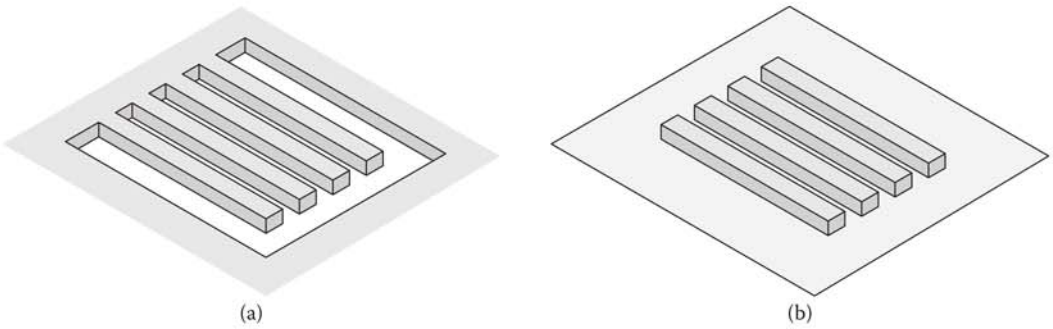


FIGURE 45.12 (a) High-contrast grating membrane suspended in air and (b) similar grating located on low-index cladding (e.g., SiO_2). The former structure is more difficult to fabricate; however, it allows the membrane to be shifted using piezoelectric forces for VCSEL cavity tuning.

layers. These layers can be either the air on both sides (in which case we speak of grating membranes; see Figure 45.12a) or one of the layers may be some kind of low-index solid (usually an oxide, such as amorphous SiO_2). The former are significantly more difficult to fabricate, as the process involves etching a sacrificial layer first and then dissolving it. The resulting grating is suspended in the air. A common problem is then buckling of the stripes, so some additional stress-relaxing etching may be necessary. The advantage of this solution is the fact that the membrane can be placed on a piezoelectric element and then moved freely, allowing for cavity tuning. On the other hand, the latter (Figure 45.12b) is usually much easier to fabricate and the bars are permanently attached to the cladding. The lack of tunability is compensated by the durable mount and no problems with buckling. If the refractive index of the cladding is low enough, there are no significant differences in the optical properties of either type of grating.

Although the gratings are very thin compared to multilayered DBRs, they can provide very high reflectivity, which in an ideal situation can even reach 100%. Due to their strongly anisotropic geometry, the gratings' reflective properties are qualitatively different, depending on the incident light polarization. For an ideal uniform grating, we can distinguish two distinct polarizations: one with an electrical field vector parallel to the grating bars (transverse electric, TE) and the other perpendicular to it (transverse magnetic, TM). Depending on the grating design, any of these polarizations can be strongly reflected, while the other is reflected only moderately. This effect can be used to design VCSELs with strong anisotropy and fixed polarization of the emitted light.

Sample reflectivities computed for a Si membrane suspended in air are shown in Figure 45.13a and b. These figures show reflectivity for both light polarizations with two different gratings. Their exact dimensions are summarized in Table 45.2. Both version were designed to provide high reflectivity at a wavelength of 1500 nm. As can be seen in Figure 45.13, both TE and TM polarization can be strongly reflected, with almost zero transmission. If two such peaks coincide, we observe a wide reflectivity bandwidth. This is clearly visible in Figure 45.13a, where the wide reflectivity area has a small dip in the middle, indicating that it consists of two independent peaks.

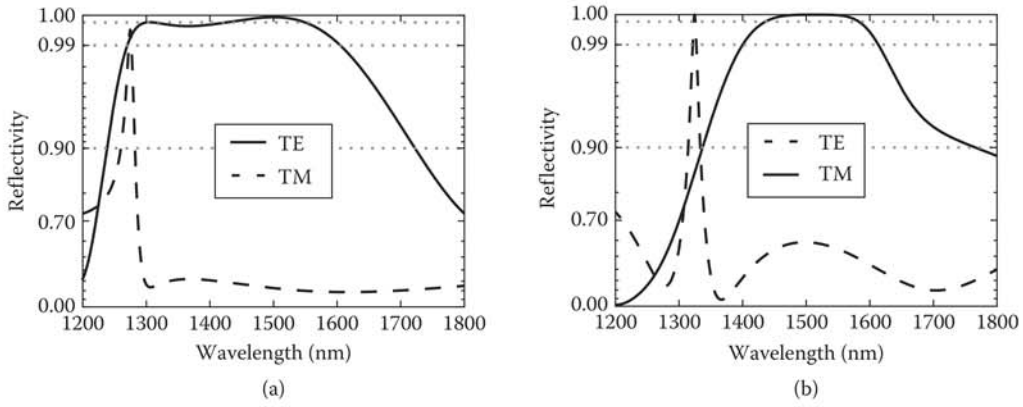


FIGURE 45.13 Silicon-on-silica grating reflectivity for TE and TM polarization. The dimensions of the gratings are optimized to provide high reflectivity around 1500 nm for (a) TE and (b) TM polarization. The exact grating dimensions are summarized in Table 45.2. The dotted lines show reflectivity levels important for laser operation: 90%, 99%, and 99.8%. Notice the non-linear scale in the plots.

TABLE 45.2 Dimensions for Gratings with Reflectivities Shown in Figure 45.13a and b.

	(a)	(b)
Grating period L [μm]	1.00	0.64
Grating fill-factor f	0.30	0.64
Grating thickness h_g [μm]	0.22	0.42
Air-gap/cladding thickness h_c [μm]	0.83	0.83

Note: The meaning of the parameters is explained in Figure 45.11. The refractive index of grating bars and substrate is 3.48 (silicon) and of the cladding is 1.47 (silica).

Due to the finite thickness of the grating, it can rather be considered as an array of cavities for the two waveguide array modes. The negative interference described in Section 45.4.1 corresponds to the resonance thickness of these cavities. If we plot grating reflectivity as a function of wavelength and grating thickness (Figure 45.14a and b), we can notice two overlapping periodic patterns, corresponding to the resonance conditions of both waveguide array modes (with propagation constants β_0 and β_2), giving a characteristic chessboard pattern.

Similar calculations can be performed for an HCG located on cladding (Figure 45.12b). With carefully chosen dimensions, it can reflect both TE and TM polarizations equally well. A sample reflectivity spectrum and reflectivity map for TM light polarization is shown in Figure 45.15. The grating parameters are as follows: $L = 0.70\,\mu\text{m}$, $h_g = 0.46\,\mu\text{m}$, $f = 0.75$, and $h_c = 0.83\,\mu\text{m}$. The cladding material is silica with a refractive index of 1.47. Such gratings are usually easier to manufacture and show the same patterns as air membranes.

45.4.3 Cavities with Subwavelength Grating Mirrors

The grating mirrors presented in Section 45.4.2 can be used to replace multi-layered Bragg mirrors as reflectors in VCSELs. Typical structures for such lasers—both with an HCG membrane suspended in air and an HCG grating located on a low-index cladding—are presented in Figure 45.16. For manufacturing feasibility, the bottom mirrors are classical DBRs and the top ones are replaced with HCGs. In the case of the

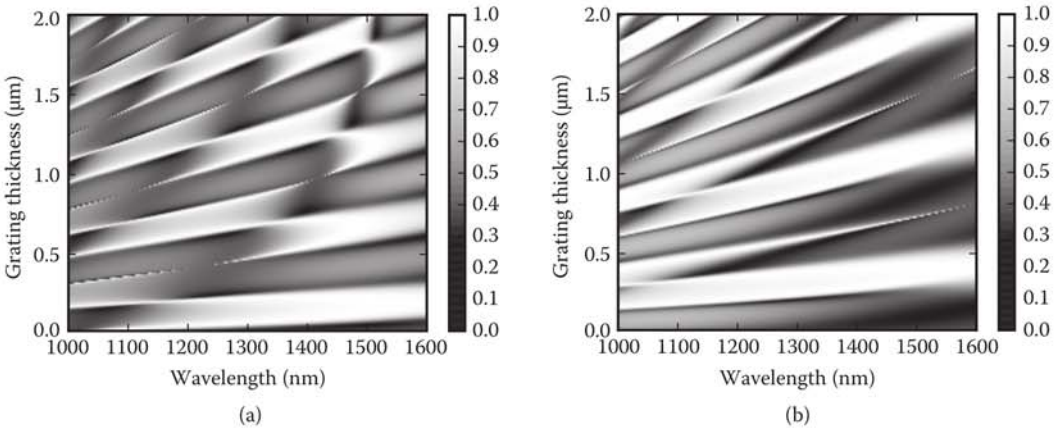


FIGURE 45.14 Reflectivity maps for gratings optimized for (a) TE and (b) TM polarizations as a function of the wavelength and grating thickness.

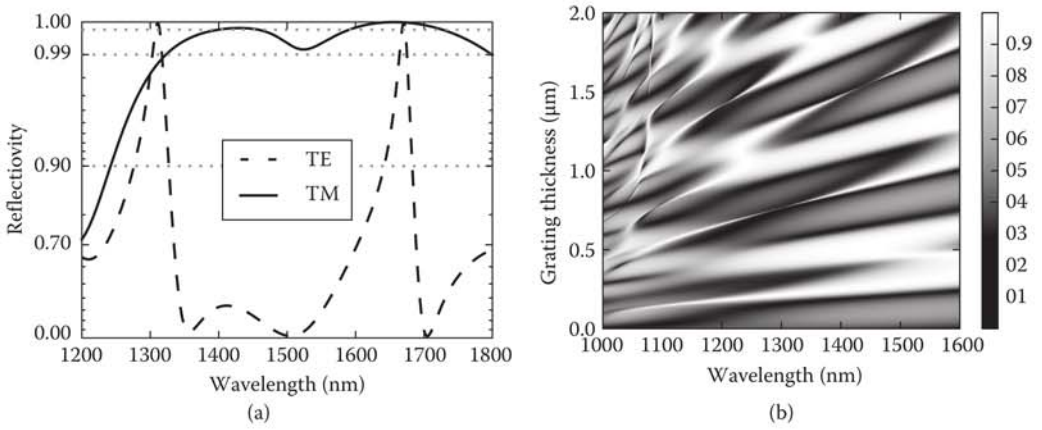


FIGURE 45.15 (a) Reflectivity spectrum and (b) TM reflectivity map for various thicknesses of a grating located on dielectric cladding (Figure 45.12b).

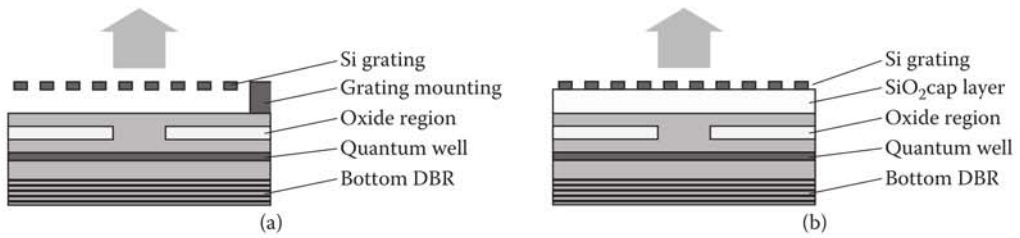


FIGURE 45.16 Typical structure of an HCG-VCSEL with (a) grating membrane suspended in air and (b) grating layer located on a low-index cladding.

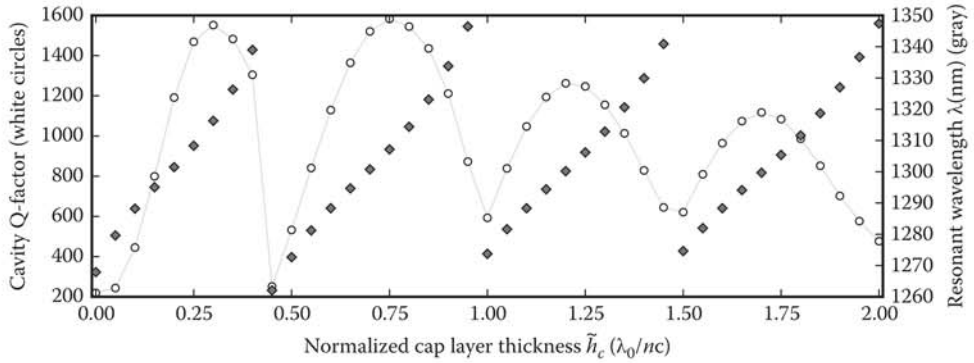


FIGURE 45.17 Cavity Q-factor (light circles) and resonant wavelength (gray diamonds) as functions of the normalized cap layer thickness.

design shown in Figure 45.16a, the membrane can be shifted by piezoelectric forces, effectively changing the cavity length and enabling the emitted wavelength to be tuned. Such a tunable HCG VCSEL has been demonstrated by [33].

In this section, we will analyze an HCG-VCSEL with the grating located on a low-index cladding layer (Figure 45.16b). It consists of a 2-λ GaAs/AlGaAs cavity† with a quantum well in one anti-node and an oxide aperture in the second. The bottom mirror is constructed from 29 pairs of GaAs/AlGaAs quarter-wavelength DBRs. At the top, there is an Si HCG separated by a SiO_2 cap layer that provides low index on both sides of the HCG. The cavity and the bottom DBR are designed for a wavelength of 1.3 μm and the HCG parameters are as follows (see Figure 45.11 for the meaning of the symbols): $L = 0.567\,\mu m$, $h_g = 0.373\,\mu m$, $f = 0.65$, and h_c varies. We are considering the cold-cavity properties of the laser, thus there is no gain in the quantum well. The oxide aperture is 4 μm and there are a total of seven grating periods within the aperture window. As the computations have to be fully three-dimensional, 891 plane waves have been used to represent the lateral electromagnetic field distribution in each layer: 11 in the direction parallel to the grating and 81 in the direction perpendicular to the grating.

Figure 45.17 shows the computed resonant wavelength of the emitted light and the cavity Q-factor‡ as functions of cap layer thickness normalized to the target cavity wavelength λ_0 of 1.3 μm. This normalized thickness is computed as $\tilde{h_c} = h_c n_c/\lambda_0$. As can be seen, both the resonant wavelength and the cavity Q-factor depend periodically on $\tilde{h_c}$, with a period of $\lambda/2$ and the highest Q-factor corresponding to the quarter-wavelength cap thickness (i.e., $\tilde{h_c} = 1/4 + k/2$ with k being an arbitrary integer positive number). However, as the cap thickness increases over $\tilde{h_c} = 1.0$, the consequent maxima correspond to a lower and lower Q-factor. This can be attributed to the fact that the finite size of the grating causes an increase in scattering losses, as the divergent beam is not fully back-reflected to the active region.

45.4.4 Manufacturing Sensitivity of High-Contrast Grating Mirrors

The results in Figure 45.17 show that HCGs can be effectively used to replace DBR mirrors in VCSELs. However, the question arises whether they are similarly robust. The need to etch a precise lithographic pattern can make them prone to manufacturing errors. We therefore investigated the reflectivity of a grating mirror with an imperfect periodic grating. We performed simulations of multiple grating periods in which grating bar thicknesses and positions were randomly varied with a specified standard deviation σ. Due

† 2-λ means that the optical cavity length is exactly two wavelengths.

‡ Q-factor is a dimensionless rate of energy loss relative to the stored energy of the resonator. It can be determined from the equation $Q = \frac{1}{2}\mathrm{Re}(\tilde{\omega})/\mathrm{Im}(\tilde{\omega})$, where $\tilde{\omega}$ is the complex eigen-frequency of the resonant mode in the cavity.

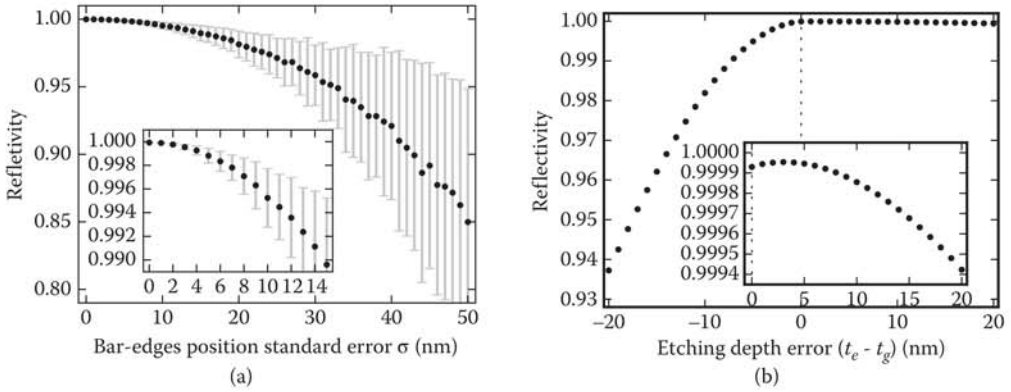


FIGURE 45.18 (a) Computed reflectivity for an HCG mirror with the position of the bar edges varied randomly using different standard deviations. The vertical bars show the uncertainties of the results. The inset shows the magnified part of the full figure. (b) Dependence of reflectivity on the difference between etching depth (h_e) and the thickness of the Si layer (h_g, see Figure 45.11). Negative values mean too shallow holes while positive values designate holes that are too deep. The latter are shown magnified in the inset.

to the randomness introduced, we repeated our calculations 1000 times for each value of σ. The results are presented in Figure 45.18a. Their uncertainties (shown with gray vertical lines) are computed using Student's *t*-distribution for a confidence interval of 90%.

As can be seen, even a small variation causes a significant drop in reflectivity. As the precision of typical photo-lithography is around 10 nm, mirror reflectivity can drop by 0.5% for $\sigma = 10$ nm. Larger manufacturing errors can make the mirrors unusable in VCSELs, as reflectivity can drop well below 90%. The largest allowable error, for which the mirrors reflect more than 99% of the incident light in at least half of the cases, is $\sigma = 15$ nm. The stronger requirement of reflectivity exceeding 99.8% is fulfilled if the error is not greater than $\sigma = 8$ nm. This clearly shows that the main physical phenomenon behind the properties of HCG mirrors is Bragg reflection, in which perfect periodicity plays a crucial role. This contrasts with most modern designs for photonic crystal VCSELs, in which the most important factor is the reduced effective index of the photonic crystal area [8].

Another test we have performed is an analysis of the precision of etching depth. In the ideal case, the Si bars should rest fully separated on a SiO_2 layer. In other words, the spaces between the bars should be etched to a depth exactly equal to the thickness of the Si layer. In reality, this may not always be the case, as the etching can be slightly deeper or slightly shallower. However, locally one can assume uniform etching conditions, whereby the depths of adjacent etched areas should be the same.

In order to investigate the impact of etching depth precision on the performance of HCG mirrors, we computed their reflectivity for various etching depths h_e (h_e is the depth of etching spaces between the bars; in an ideal situation it should be equal to the thickness of the high-index grating layer h_g, see Figure 45.11). Incident light is the same as in the previous case, i.e., the polarization is perpendicular to the bars and the wavelength is 1.3 μm. The results are shown in Figure 45.18b. As one can see, too shallow etching causes a strong decrease in reflectivity. On the other hand, too deep etching does not influence the results significantly. Furthermore, etching 3 nm of the SiO_2 layer can increase reflectivity slightly, although this effect is so weak that it is negligible in practice. The main conclusion is that it is better to etch the space around the bars a little more than desired in order to provide optimum mirror reflectivity even in the case of manufacturing imprecisions.

Similar analysis can be performed for HCG VCSELs, as described in Section 45.4.3. As the reflectivity of HCGs drops with manufacturing imprecisions, so should the VCSEL cavity Q-factor. A numerical analysis of this effect is shown in Figure 45.19. The computation procedure is identical to that used for generating Figure 45.18. However, each point was averaged over 100 simulations instead of 1000 (due to the larger

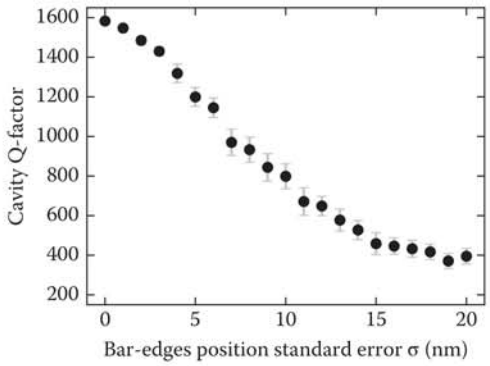


FIGURE 45.19 HCG-VCSEL cavity Q-factor with an imperfect mirror. The edges of the grating bars are varied randomly using different standard deviations. The vertical bars show the uncertainties of the results.

numerical effort necessary to determine the proper VCSEL modes). The resonant wavelength is obviously unaffected, as it is determined by the cavity itself and not the quality of the mirror. We can notice that even a 10 nm error can lead to a twofold reduction in the Q-factor. This must be taken into consideration in designs for devices that are to be produced on an industrial scale. Furthermore, the analysis presented does not consider the effects of the surface roughness of the bars, which to some extent is inevitable in the lithography process.

45.5 Summary

Periodic structures have been used in laser diodes for a long time. Their most basic form are one-dimensional distributed Bragg reflectors, commonly used in VCSELs. However, two- or three-dimensional photonic crystals allow further tuning of laser properties, e.g., to achieve a single-mode light emission. One-dimensional subwavelength high-contrast gratings can make efficient broadband laser mirrors with reduced size and wavelength tuning ability.

The downside of such structures is the increased laser complexity, which requires more advanced numerical methods for the device optical modeling. In particular, only vectorial methods like the finite-difference time-domain method, the plane-wave expansion, the finite element method, the coupled mode model, the plane-wave admittance method, or the plane-wave reflection transfer can yield proper optical characteristics of PC-VCSELs.

References

1. Strutt, J. W. (1887). On the maintenance of vibrations by forces of double frequency, and on the propagation of waves through a medium endowed with a periodic structure. *Phil. Mag.* 27, 145–159.
2. Yablonovitch, E. (1987). Inhibited spontaneous emission in solid-state physics and electronics. *Phys. Rev. Lett.* 58, 2059–2062.
3. John, S. (1987). Strong localization of photons in certain disordered dielectric superlattices. *Phys. Rev. Lett.* 58, 2486–2489.
4. Kittel, C. (2004). *Introduction to Solid State Physics* (8th ed.). Hoboken, NJ: Wiley.
5. Joannopoulos, J. D., R. D. Meade, and J. N. Winn (1995). *Photonic Crystals: Molding the Flow of Light*. Princeton, NY: Princeton University Press.
6. Checoury, X., P. Boucaud, J.-M. Lourtioz, F. Pommereau, C. Cuisin, E. Derouin, O. Drisse, L. Legouezigou, O. L. Legouezigou, F. Lelarge, F. Poingt, G.-H. Duan, S. Bonnefont, D. Mulin, J.

Valentin, O. Gauthier-Lafaye, F. Lozes-Dupuy, and A. Talneau (2005). Distributed feedback-like laser emission in photonic crystal waveguides on InP substrate. *IEEE J. Sel. Top. Quant. Electron.* 11, 1180–1186.

7. Painter, O., J. Vuckovic, and A. Scherer (1999). Defect modes of a two-dimensional photonic crystal in an optically thin dielectric slab. *J. Opt. Soc. Am. B.* 16, 275–285.
8. Danner, A. J., J. J. Raftery Jr., L. P. O., and K. D. Choquette (2006). Single mode photonic crystal vertical cavity lasers. *Appl. Phys. Lett.* 88, 091114.
9. Danner, A. J., J. J. Raftery Jr., N. Yokouchi, and K. D. Choquette (2004). Transverse modes of photonic crystal vertical-cavity lasers. *Appl. Phys. Lett.* 84, 1031–1033.
10. Leisher, P. O., A. J. Danner, and K. D. Choquette (2006). Single-mode 1.3 μm photonic crystal vertical-cavity surface-emitting laser. *IEEE Photon. Technol. Lett.* 18, 2156–2158.
11. Song, D. S., S. H. Kim, H. G. Park, C. K. Kim, and Y. H. Lee (2002). Single-fundamental mode photonic-crystal vertical-cavity surface-emitting lasers. *Appl. Phys. Lett.* 80, 3901–3903.
12. Yokouchi, N., A. J. Danner, and K. D. Choquette (2003). Two-dimensional photonic crystal confined vertical-cavity surface-emitting lasers. *IEEE J. Sel. Top. Quant. Electron.* 9, 1439–1445.
13. Liu, H., M. Yan, P. Shum, H. Ghafouri-Shiraz, and D. Liu (2004). Design and analysis of anti-resonant reflecting photonic crystal VCSEL lasers. *Opt. Express.* 12, 4269–4274.
14. Czyszanowski, T., M. Dems, and K. Panajotov (2007a). Impact of the hole depth on the modal behaviour of long wavelength photonic crystal VCSELs. *J. Phys. D: Appl. Phys.* 40, 2732–2735.
15. Czyszanowski, T., M. Dems, R. P. Sarzała, W. Nakwaski, and K. Panajotov (2011). Precise lateral mode control in photonic crystal vertical-cavity surface-emitting lasers. *IEEE J. Quant. Electron.* 47, 1291–1296.
16. Birks, T. A., J. C. Knight, and P. S. J. Russell (1997). Endlessly single-mode photonic crystal fiber. *Opt. Lett.* 22, 961–963.
17. Taflove, A. and S. C. Hagness (2000). *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (2nd ed.). Boston, MA: Artec House.
18. Johnson, S. G. and J. D. Joannopoulos (2001). Block iterative frequency-domain methods for Maxwell's equations in a planewave basis. *Opt. Express.* 8, 173–190.
19. Nyakas, P. (2007). Full-vectorial three-dimensional finite element optical simulation of vertical-cavity surface-emitting lasers. *IEEE J. Lightwave Techn.* 25, 2427–2434.
20. Bava, G. P., P. Debernardi, and L. Fratta (2001). Three-dimensional model for vectorial fields in vertical-cavity surface-emitting lasers. *Phys Rev. A.* 63, 23816.
21. Dems, M., R. Kotynski, and K. Panajotov (2005). Plane-wave admittance method—A novel approach for determining the electromagnetic modes in photonic structures. *Opt. Express.* 13, 3196–3207.
22. Dems, M. (2011). Modelling of high-contrast grating mirrors. The impact of imperfections on their performance in VCSELs. *Opto-Electr. Rev.* 19, 340–345.
23. Conradi, O., S. F. Helfert, and R. Pregla (2001). Comprehensive modeling of vertical-cavity laser-diodes by the method of lines. *IEEE J. Quant. Electron.* 37, 928–935.
24. Czyszanowski, T., M. Dems, and K. Panajotov (2007b). Single mode condition and modes discrimination in photonic-crystal 1.3 μm AlInGaAs/InP VCSEL. *Opt. Express.* 15, 5604–5609.
25. Gedney, S. D. (1996). An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices. *IEEE T. Antenn. Propag.* 44, 1630–1639.
26. Dems, M. (2009). Semi-vectorial method based on effective index for VCSEL analysis. *J. Opt. Soc. Am. B.* 26, 792–796.
27. Dems, M., I.-S. Chung, P. Nyakas, S. Bischoff, and K. Panajotov (2010). Numerical methods for modeling photonic-crystal VCSELs. *Opt. Express.* 18, 16042–16054.
28. Panajotov, K., B. Ryvkin, J. Danckaert, M. Peeters, H. Thienpont, and I. Veretennicoff (1998). Polarization switching in VCSEL's due to thermal lensing. *IEEE Photon. Technol. Lett.* 10, 6–8.
29. Fan, S., W. Suh, and J. Joannopoulos (2003). Temporal coupled-mode theory for the Fano resonance in optical resonators. *J. Opt. Soc. Am. A.* 20 (3), 569–572.

30. Karagodsky, V., F. G. Sedgwick, and C. J. Chang-Hasnain (2010). Theoretical analysis of subwavelength high contrast grating reflectors. *Opt. Express*. 18, 16973–16988.
31. Chang-Hasnain, C. J. and W. Yang (2012). High-contrast gratings for integrated optoelectronics. *Adv. Opt. Photon*. 4, 379–440.
32. Rosenblatt, D., A. Sharon, and A. A. Friesem (1997). Resonant grating waveguide structures. *IEEE J. Quant. Electron*. 33, 2038–2059.
33. Huang, M. C. Y., Y. Zhou, and C. J. Chang-Hasnain (2008). A nanoelectromechanical tunable laser. *Nat. Photon*. 2, 180–184.

46

Single-Photon Sources

Niels Gregersen
Dara P. S. McCutcheon
and
Jesper Mørk

46.1 Introduction 585
Characteristics • Single-Photon Generation Schemes • Photon Extraction Strategies
46.2 Modeling of Spontaneous Emission 590
Classical Dipole • Spontaneous Emission • Simulation Techniques
46.3 Cavity QED Approach 594
Micropillar
46.4 Waveguide QED Approach 596
Photonic Nanowire • Photonic Crystal Waveguide
46.5 Planar Dielectric Antenna Approach 600
46.6 Decoherence Effects 601
Quantum Emitter Master Equation • Emitter-Cavity Master Equation • Spectral Wandering • Coupling to Longitudinal Acoustic Phonons
46.7 Summary 604

46.1 Introduction

The single-photon source (SPS) is a solid-state device capable of emitting single photons on demand. It is a key component within the emerging field of quantum information processing, where the information is encoded on single photons. In these applications, the quantum mechanical properties of the single photon are exploited, e.g., to enable secure communication where undetected eavesdropping is impossible or to perform quantum calculations taking advantage of massive parallelization.

This tutorial chapter reviews the characteristic figure of merits of the deterministic SPS based on a quantum emitter in a semiconductor solid-state system. The theory and the modeling methods used to describe single-photon emission are discussed. Photon extraction strategies based on cavity and waveguide quantum electrodynamics (QED) are presented, and finally we discuss decoherence mechanisms due to the solid-state environment.

46.1.1 Characteristics

For use in quantum information technology, the SPS must meet three basic performance requirements: pure single-photon emission, indistinguishability, and high efficiency, as discussed separately in the three following subsections.

46.1.1.1 Pure Single-Photon Emission

The most fundamental characteristic of the SPS is that it should emit only one photon at a time. This ability is tested in the so-called Hanbury, Brown, and Twiss (HBT) experiment illustrated in Figure 46.1.

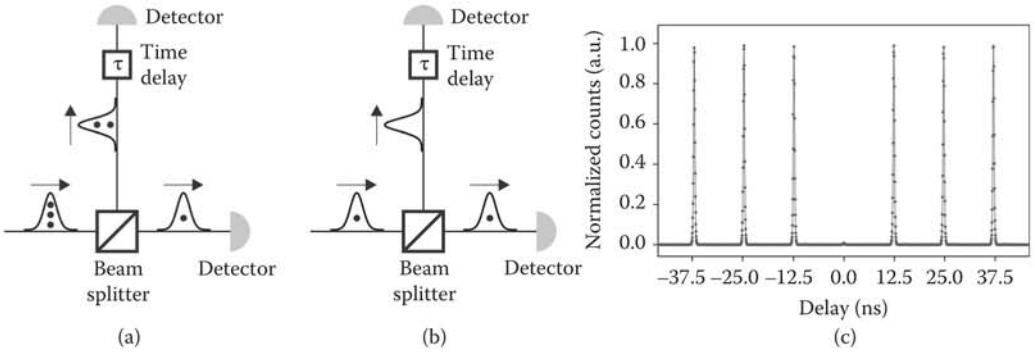


FIGURE 46.1 The Hanbury, Brown, and Twiss experiment illustrated for input pulses containing (a) multiple photons and (b) single photons. (c) The $g^{(2)}(\tau)$ autocorrelation function for a micropillar SPS. (Reprinted with permission from Ding, X. et al., *Physical Review Letters*, 116, 020401, 2016. Copyright 2016 by the American Physical Society.)

Here an optical pulse ideally containing only one photon is directed toward a beam splitter and is either reflected or transmitted toward two detectors. A variable time delay τ is implemented in one of the two output paths, and it is adjusted such that the photons from either of the two ports of the beam splitter arrive simultaneously at the two detectors for zero time delay. The single-photon emission ability is now characterized by measuring the autocorrelation function $g^{(2)}(\tau)$. The experiment is carried out by sending in light pulses from the SPS and by measuring the number of simultaneous detection events in the two detectors as function of time delay.

If the input pulse contains multiple photons as depicted in Figure 46.1a, some photons may be reflected while others are transmitted leading to coincidence clicks in the two detectors at zero time delay such that $g^{(2)}(\tau = 0) \neq 0$. However, if the pulse contains only a single photon as illustrated in Figure 46.1b, this photon will be either transmitted or reflected and no coincidence events are measured. An example of such a HBT measurement is given in Figure 46.1c (Ding et al., 2016), where we observe coincidence counts as function of delay for a SPS generating pulses every 12 ns. The peaks in the measurement correspond to integer values of the pulse time separation, where photons from different pulses may lead to coincidence clicks. However, at zero time delay, the correlation counts are strongly suppressed corresponding to an ideal single-photon emission with $g^{(2)}(\tau = 0) = 0$.

46.1.1.2 Indistinguishability

A second requirement for the SPS is that the emitted photons should be quantum mechanically indistinguishable, meaning that the emitted photons should have the same wavelength, phase variation across the pulse, polarization, optical mode profile, etc. The single-photon indistinguishability is characterized in the Hong–Ou–Mandel (HOM) experiment sketched in Figure 46.2 (Hong et al., 1987).

In this experiment, two single-photon pulses are inbound on a 50–50 beam splitter, where the two pulses are either transmitted or reflected into the output paths. A variable time delay τ is implemented in one of the two output paths adjusted such that the photons reach the detectors simultaneously at zero time delay. The experiment then takes place by sending in single-photon pulses along the input paths and by measuring the correlation function $g^{(2)}(\tau)$ describing the simultaneous detection events for this setup as function of time delay.

The impinging photons may leave the beam splitter either in separate output paths or together as illustrated in Figure 46.2a and b, respectively. In the case where the input photons are distinguishable, the output photons may leave the beam splitter along separate paths leading to $g^{(2)}(\tau = 0) \neq 0$. However, when the input photons are completely indistinguishable and arrive simultaneously at the beam splitter, they always exit the beam splitter along the same output path as depicted in Figure 46.2b, resulting in a $g^{(2)}(\tau = 0) = 0$. This phenomenon is known as the HOM effect and can be understood from a quantum-optical consideration of the probabilities of the four possible output states sketched in Figure 46.2c. The

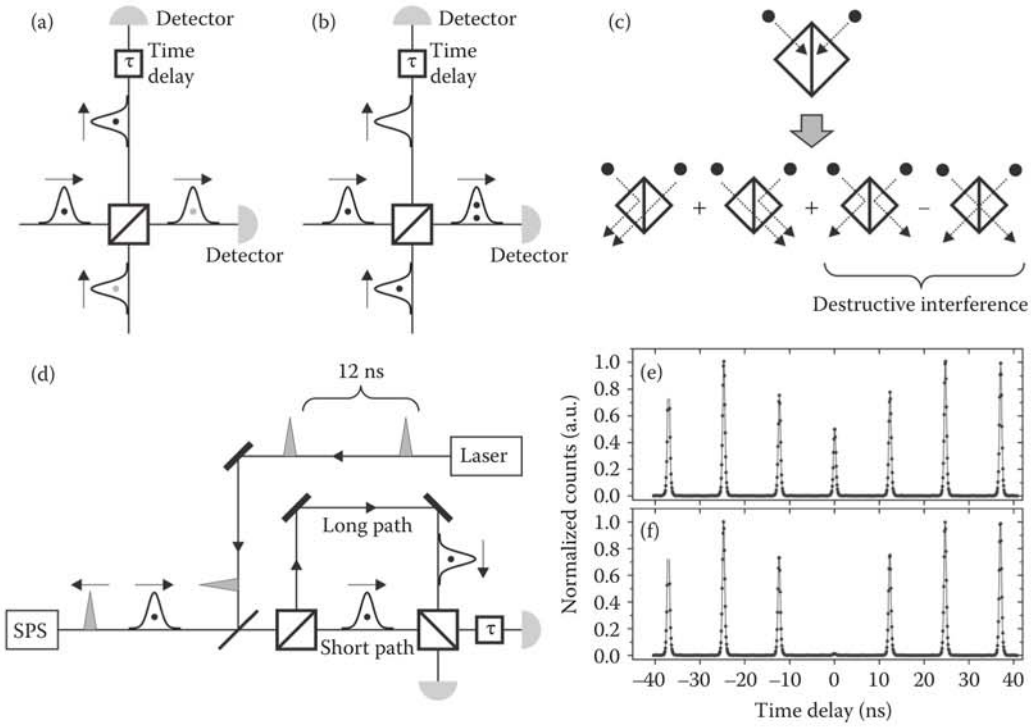


FIGURE 46.2 Outcome of the Hong–Ou–Mandel experiment for (a) distinguishable and (b) indistinguishable photons. (c) Various possible output states. Sketch of a typical experimental setup. Measured coincidence counts in (e) cross and (f) parallel polarization. (Reprinted figure with permission from Ding, X. et al., *Physical Review Letters*, 116, 020401, 2016. Copyright 2016 by the American Physical Society.)

probability amplitude of the photons leaving the beam splitter along separate output paths consists of two contributions corresponding to the photons being both either reflected or transmitted at the beam splitter, and for indistinguishable photons these two contributions add up destructively (Gerry and Knight, 2008).

In most HOM experiments, the impinging photons are generated by the same source using the setup sketched in Figure 46.2d. A laser generates classical pulses at a typical repetition rate of 80 MHz. The pulses are transmitted to the SPS, and each laser pulse triggers a single-photon emission event leading to a train of single photons with 12.5-ns time separation inbound on the first beam splitter. Two subsequent photons in the train can propagate along the short and the long paths in various configurations. The long and short path lengths are adjusted such that a first photon taking the long path and a second photon taking the short path will arrive simultaneously at the beam splitter. Typical HOM measurements for indistinguishable photons having orthogonal or parallel polarization are illustrated in Figure 46.2e and f, respectively. For orthogonal polarizations, the photons are distinguishable and $g^{(2)}(\tau = 0) \neq 0$, whereas for parallel polarizations the HOM effect with $g^{(2)}(\tau = 0) = 0$ is observed. Again, $g^{(2)}(\tau)$ has peaks for time delays equal to an integer times the pulse separation. These peaks correspond to coincidence events for photon pairs that do not arrive simultaneously at the beam splitter and are thus not subject to the HOM effect. The indistinguishability is then quantified as one minus the normalized area of the $\tau = 0$ peak and should ideally equal unity.

46.1.1.3 Efficiency

The third important characteristic of the SPS is its efficiency or brightness (Gregersen et al., 2013). The ideal SPS should be deterministic and provide single photons on demand. It should act as a "single-photon gun" such that every trigger releases exactly one photon into the collection optics. The efficiency ε is then

defined as the number of photons detected by the collection optics per trigger. Within the QED design approaches discussed in Sections 46.3 and 46.4, the efficiency is characterized using two governing parameters as illustrated in Figure 46.3. The first is the spontaneous emission factor β describing the fraction of light coupled from the emitter into the optical mode of interest. The β factor can be written as

$$\beta = \frac{\Gamma_M}{\Gamma_M + \Gamma_{Rad}}, \tag{46.1}$$

where Γ_M is the spontaneous emission rate into the optical mode and Γ_{Rad} is the emission rate into background radiation modes. One may also add a term Γ_{NonRad} to the denominator in the definition (46.1) to take into account nonradiative recombination channels of the quantum emitter. Whereas Γ_M and Γ_{Rad} can be modified by tailoring the photonic environment, the parameter Γ_{NonRad} represents inherent properties of the emitter that generally cannot easily be controlled.

The second governing parameter is the outcoupling efficiency γ describing the fraction of light transmitted from the optical mode to the first objective lens. In this simple picture, the efficiency is given by $\varepsilon = \beta\gamma$. In a multiphoton interference experiment involving j photons, the total success probability of the experiment will scale as ε^j. For scalable optical quantum information technologies involving a large number of photons, a single-photon generation efficiency close to unity is thus required.

46.1.2 Single-Photon Generation Schemes

A major workhorse for generating single photons has been the spontaneous parametric down-conversion (SPDC) process (Brendel et al., 1999). Let us recall that in a nonlinear isotropic material, the polarization **P** is related to the electric field **E** as

$$\mathbf{P} = \varepsilon_0 \left(\chi^{(1)}\mathbf{E} + \chi^{(2)}\mathbf{E}^2 + \chi^{(3)}\mathbf{E}^3 + \ldots\right), \tag{46.2}$$

where $\chi^{(n)}$ is the nth-order susceptibility. In the SPDC scheme, a crystal featuring a $\chi^{(2)}$ nonlinearity such as lithium niobate is illuminated with a strong laser pump pulse with an energy $\hbar\omega_p$ as sketched in Figure 46.4. Provided that proper phase-matching conditions are met, the $\chi^{(2)}$ nonlinearity then enables conversion of a pump photon to two single photons, the signal and the idler photons of energies $\hbar\omega_s$ and $\hbar\omega_i$, respectively, such that $\hbar\omega_p = \hbar\omega_s + \hbar\omega_i$. The advantage of the SPDC source is its simplicity and the high

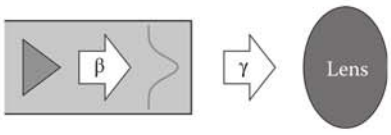


FIGURE 46.3 Efficiency model for the quantum electrodynamics approach. The single-photon emitter is represented by the triangle.

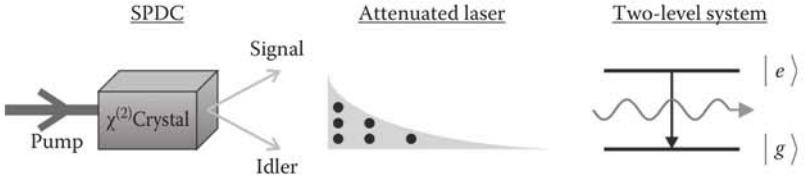


FIGURE 46.4 Major single-photon generation schemes.

degree of indistinguishability of the signal and idler photons. Its main drawback is that the downconversion process is inherently probabilistic, leading to efficiencies below 0.01. While the efficiency may be improved by cranking up the pump power, this occurs at the cost of significant increase of the probability of multiphoton emission. The probabilistic nature thus limits SPDC to experiments involving only a few photons.

Another commercially important source is the attenuated laser SPS (Buttler et al., 2000). A pulse generated using a classical laser propagates through an absorbing material and is attenuated to the single-photon level. The advantage of the attenuated laser is its simplicity, and similar to the SPDC process its main drawback is the probabilistic nature of the attenuation leading to either low single-photon generation efficiency or high probability of multiphoton events. Additionally, the generated photons are not indistinguishable, preventing their use in many quantum optics and quantum information experiments. However, the attenuated laser remains an important device in commercial products for quantum key distribution, where clean single-photon emission and indistinguishability are not required.

The quantum-dot emitter (Shields, 2007) has recently emerged as a solution, overcoming the probabilistic limitations of the SPDC and attenuated laser strategies. It consists of a two-level system as shown in Figure 46.4, where an electron can occupy either a ground state $|g\rangle$ or an excited state $|e\rangle$ characterized by energies E_g and E_e, respectively. The system is initially in the ground state and is then excited using an optical or electrical pulse, after which the electron returns to the ground state by the spontaneous emission of a single photon of energy $\hbar\omega_0 = E_e - E_g$. A main advantage of the two-level emitter is that the upper level can contain only one excitation at a time, thereby ensuring the single-photon nature of the emitted light, combined with the deterministic nature of the spontaneous emission event and the possibility of structuring the medium to realize highly directional emission, potentially allowing for unity efficiency.

Any solid-state system featuring discretized energy levels can in principle be employed as a SPS. The InAs quantum dot (QD) grown in a GaAs host material as shown in Figure 46.5a represents a mature platform for a deterministic SPS. The QD consists of a pyramidal-shaped island of InAs atoms surrounded by GaAs. Due to the lower bandgap of InAs, the electron wavefunction will be fully confined in all three dimensions leading to discretized energy levels in the conduction and valence bands depicted in Figure 46.5b. The emission of InAs QD ensembles features strong inhomogeneous broadening with typical emission wavelength between 900 and 950 nm. The lowest energy *s* state in the conduction band is typically used as the excited state. Due to the spin degree of freedom, the *s* state supports a biexciton leading to two-photon emission. However, the InAs QD features a fine-structure splitting arising from the exciton binding energy of a few nanometers allowing for the selection of the excitonic line using spectral filtering thus enabling single-photon emission at cryogenic temperatures. The indistinguishability of the emitted photons depends greatly on the charge environment and the excitation scheme employed, as discussed in detail in Section 46.6.

Other two-level systems including diamond nitrogen-vacancy centers (Babinec et al., 2010) and organic molecules (Siyushev et al., 2014) have been pursued for single-photon generation. These systems feature the advantages of room temperature operation and very long spin coherence times. However, the maturity

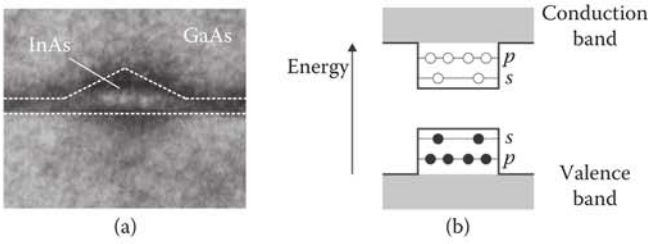


FIGURE 46.5 Scanning electron micrograph (a) and energy level diagram (b) of an InAs quantum dot.

FIGURE 46.6 Major photon extraction strategies. (Reprinted from Gregersen, N et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2013 IEEE and Reprinted by permission from Macmillan Publishers Ltd: Claudon, J et al., *Nature Photonics* 4(3): 174–177, copyright 2010.)

of III–V semiconductor cleanroom fabrication has made the InAs QD the leading platform for highly efficient SPSs, and for this reason we focus in this tutorial on the InAs QD.

46.1.3 Photon Extraction Strategies

While the self-assembled InAs QD features an internal quantum efficiency η of almost unity, a major challenge in single-photon engineering lies in the efficient extraction of the emitted photon (Gregersen et al., 2013). Figure 46.6 summarizes the main SPS photon extraction strategies. The simplest approach consists of a planar sample with a QD in a bulk material. The main advantages of this approach are its simplicity as no additional structuring is needed after QD growth as well as the absence of fluctuating surface charges near the QD as discussed further in Section 46.6. The disadvantage is the poor extraction efficiency. The symmetry of the system leads to light emission in all directions limiting the efficiency to only a few percent.

The first strategy proposed to improve the efficiency and control the light emission was based on cavity QED, where the QD was placed in a microcavity (Gérard, 2003). Here the Purcell effect ensures a preferential single-photon emission from the QD into the cavity mode. This preferential coupling requires a careful spectral alignment between the QD emission line and the cavity line as discussed in detail in Section 46.3. This approach has led to efficiencies of 0.8 (Gazzano et al., 2013) as well as an indistinguishability of 0.99 from the micropillar cavity SPS (Ding et al., 2016; Somaschi et al., 2016), representing the current state of the art in highly efficient sources of indistinguishable photons.

A second design strategy is based on waveguide QED, where the QD is positioned inside a waveguide (Claudon et al., 2013). The preferential coupling to the waveguide mode is ensured by a dielectric screening effect (Bleuse et al., 2011). The main advantage of this nonresonant approach is that no spectral alignment is required. While an efficiency of 0.75 has been demonstrated for the photonic nanowire SPS, the demonstration of highly indistinguishable photons from a waveguide-based SPS is still lacking.

46.2 Modeling of Spontaneous Emission

A major challenge in engineering deterministic SPSs lies in optimizing the photon extraction efficiency and the indistinguishability to achieve values close to unity. Proper modeling of the light emission from a two-level system generally requires a QED description of the spontaneous emission process. However, SPSs generally operate in the weak coupling regime, and here it turns out that the spontaneous emission rate can be calculated in a purely classical manner (Novotny and Hecht, 2012).

46.2.1 Classical Dipole

We first consider the electromagnetic field generated by a classical current distribution $\mathbf{j}$ emitting light at the angular frequency ω_0. We thus work in the frequency domain with harmonic time dependence such that $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r})e^{-i\omega_0 t}$ and $\mathbf{j}(\mathbf{r}, t) = \mathbf{j}(\mathbf{r})e^{-i\omega_0 t}$. Furthermore, we assume nonmagnetic materials and the

absence of free charges for simplicity. The classical Maxwell's equations are then combined to yield the second-order wave equation for the electric field:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - \varepsilon_r(\mathbf{r}) k_0^2 \mathbf{E}(\mathbf{r}) = i\omega_0 \mu_0 \mathbf{j}(\mathbf{r}), \tag{46.3}$$

where $k_0 = \omega_0/c$ is the free-space wavenumber and ε_r is the dielectric constant.

Equation 46.3 can be solved in an elegant manner by introducing the dyadic Green's function $\overline{\overline{G}}$ (Novotny and Hecht, 2012). It is defined as the solution to

$$\nabla \times \nabla \times \overline{\overline{G}}(\mathbf{r}, \mathbf{r}') - \varepsilon_r(\mathbf{r}) k_0^2 \overline{\overline{G}}(\mathbf{r}, \mathbf{r}') = \overline{\overline{I}} \delta\left(\mathbf{r} - \mathbf{r}'\right), \tag{46.4}$$

where $\overline{\overline{I}}$ is the unit dyad. $\overline{\overline{G}}$ and $\overline{\overline{I}}$ are both 3×3 tensors and we stress that $\overline{\overline{G}}(\mathbf{r}, \mathbf{r}')$ depends on both $\mathbf{r}$ and $\mathbf{r}'$. Now, using this definition of the dyadic Green's function, we note that

$$\mathbf{E}(\mathbf{r}) = i\omega_0 \mu_0 \int_V \overline{\overline{G}}(\mathbf{r}, \mathbf{r}') \mathbf{j}(\mathbf{r}') \, \mathrm{d}\mathbf{r}' \tag{46.5}$$

is a solution to Equation 46.3. The Green's function $\overline{\overline{G}}(\mathbf{r}, \mathbf{r}')$ can thus be interpreted physically as the field generated at the position $\mathbf{r}$ by a current at the position $\mathbf{r}'$. Once the Green's function has been calculated, the entire field profile can be determined from Equation 46.5.

The standard procedure for computing the power P emitted by the current distribution is to integrate the time average of the normal component of the Poynting vector over the surface S of volume V enclosing the current distribution as

$$P = \frac{1}{2} \int_S Re\left(\mathbf{E} \times \mathbf{H}^*\right) \cdot \mathbf{n}_S \mathrm{d}S, \tag{46.6}$$

where $\mathbf{n}_S$ is a unit vector normal to the surface S. Now, the average energy dissipation rate W inside the volume is given by

$$W = -\frac{1}{2} \int_V Re\left(\mathbf{j}^* \cdot \mathbf{E}\right) \mathrm{d}\mathbf{r}, \tag{46.7}$$

and according to Poynting's theorem, we have $P = W$.

We now consider the current distribution generated by a point dipole $\mathbf{d}$ positioned at $\mathbf{r}_0$, as illustrated in Figure 46.7, such that

$$\mathbf{j}(\mathbf{r}) = -i\omega_0 \mathbf{d} \delta\left(\mathbf{r} - \mathbf{r}_0\right). \tag{46.8}$$

By inserting Equation 46.8 into Equation 46.5 and subsequently Equation 46.5 into Equation 46.7, we obtain an expression for the power P in terms of the dipole moment given by

$$P = \frac{\omega_0^3 \mu_0 \left|\mathbf{d}\right|^2}{2} \mathrm{Im}\left(\mathbf{n}_\mathrm{d} \cdot \overline{\overline{G}}(\mathbf{r}_0, \mathbf{r}_0) \cdot \mathbf{n}_\mathrm{d}\right), \tag{46.9}$$

where $\mathbf{n}_\mathrm{d} = \mathbf{d}/\left|\mathbf{d}\right|$ is the orientation of the dipole moment. The power P emitted by the point dipole can thus be obtained by evaluating the imaginary part of the Green's function at the position of the dipole.

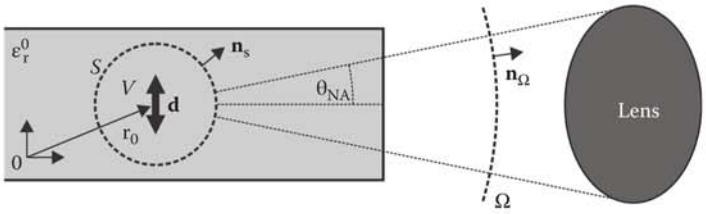


FIGURE 46.7 Light emission from a classical dipole **d** placed at the position $\mathbf{r}_0$ inside a volume V in a material of dielectric constant ε_r^0. The volume is limited by the surface S having normal unit vector $\mathbf{n}_S$. The entire structure is enclosed in a sphere Ω with normal unit vector $\mathbf{n}_\Omega$.

46.2.2 Spontaneous Emission

In the weak coupling regime, the total spontaneous emission rate Γ of a two-level system at the position $\mathbf{r}_0$, with transition energy $\hbar\omega_0$ and with dipole moment **d,** is derived from Fermi's golden rule as

$$\Gamma = \frac{\pi\omega_0}{\hbar\varepsilon_0} |\mathbf{d}|^2 \rho_L \left(\mathbf{n}_\mathrm{d}, \mathbf{r}_0, \omega_0\right). \tag{46.10}$$

Here, $\rho_L \left(\mathbf{n}_\mathrm{d}, \mathbf{r}_0, \omega_0\right)$ is the local density of states (LDOS) defined as

$$\rho_L \left(\mathbf{n}_\mathrm{d}, \mathbf{r}_0, \omega_0\right) = \int \left|\mathbf{n}_\mathrm{d} \cdot \widetilde{\mathbf{E}}_\alpha(\mathbf{r}_0)\right|^2 \delta(\omega_0 - \omega_\alpha) d\alpha, \tag{46.11}$$

where $\widetilde{\mathbf{E}}_\alpha$ are the so-called normal modes solutions to the wave equation in the absence of free currents and charges:

$$\nabla \times \nabla \times \widetilde{\mathbf{E}}_\alpha(\mathbf{r}) - \varepsilon(\mathbf{r}) \left(\frac{\omega_\alpha}{c}\right)^2 \widetilde{\mathbf{E}}_\alpha(\mathbf{r}) = 0, \tag{46.12}$$

and the index α is a short notation for all appropriate discrete and continuous parameters describing a mode. The normal modes are normalized such that

$$\int \varepsilon_r(\mathbf{r}) \widetilde{\mathbf{E}}_\alpha^*(\mathbf{r}) \cdot \widetilde{\mathbf{E}}_{\alpha'}(\mathbf{r}) \mathrm{d}\mathbf{r} = \delta(\alpha - \alpha'). \tag{46.13}$$

We now write the dyadic Green's function as an expansion in normal functions. Referring the reader interested in the derivation to Novotny and Hecht (2012), we simply state the result:

$$Im \left(\overline{\overline{G}}(\mathbf{r}_0, \mathbf{r}_0, \omega_0)\right) = \frac{\pi c^2}{2\omega_0} \int \widetilde{\mathbf{E}}_\alpha^*(\mathbf{r}_0) \widetilde{\mathbf{E}}_\alpha(\mathbf{r}_0) \delta(\omega_0 - \omega_\alpha) d\alpha, \tag{46.14}$$

where the integrand contains the outer product of the normal modes. Comparing this expression to Equation 46.11, we observe that the imaginary part of the Green's function is closely related to the LDOS. Indeed, outer multiplications of the dyad with the dipole unit vector and using the definition (46.11) lead to

$$\rho_L \left(\mathbf{n}_\mathrm{d}, \mathbf{r}_0, \omega_0\right) = \frac{2\omega_0}{\pi c^2} \mathrm{Im} \left(\mathbf{n}_\mathrm{d} \cdot \overline{\overline{G}}(\mathbf{r}_0, \mathbf{r}_0, \omega_0) \cdot \mathbf{n}_\mathrm{d}\right), \tag{46.15}$$

and we observe that, except for physical constants, the LDOS is identical to the average power emitted by a classical dipole as given in Equation 46.9. This means that we obtain the very important relationship

$$\frac{\Gamma}{\Gamma_{\text{Bulk}}} = \frac{P}{P_{\text{Bulk}}}, \tag{46.16}$$

which states that the spontaneous emission rate, Γ, of a two-level system embedded in a medium with dielectric constant ε_r^0 normalized to its value, Γ_{Bulk}, in a bulk material of dielectric constant ε_r^0 is equal to the power dissipation, P, from a classical dipole normalized to the rate, P_{Bulk}, in the bulk material.

46.2.3 Simulation Techniques

Analytic expressions for the Green's function are only available in a limited number of geometries and generally we rely on numerical simulation techniques to compute the electric field. When using spatial discretization methods such as the finite-difference time-domain technique and the finite-elements method (Lavrinenko et al., 2014), the field is expanded on a grid as sketched in Figure 46.8a and either Maxwell's equations or a vectorial wave equation are directly solved to give the optical field.

Alternatively, one may use a modal method (Lavrinenko et al., 2014). Here, the geometry is sliced into layers uniform along a propagation axis as illustrated in Figure 46.8b and the eigenmodes of each layer are then computed. The optical field is then expanded on the eigenmodes, and the fields at each side of a layer interface are connected using the scattering matrix formalism. This allows for direct access to scattering coefficients and the possibility of studying the various parts of the SPS individually.

As we have seen, the QD can be modeled as a classical point dipole placed inside the geometry under study. The electric field profile generated by the dipole is then computed, and the classical power P can then be obtained from Equation 46.7. We note that, for the current source defined by Equation 46.8, the real part of the electric field diverges at the position of the dipole, and an evaluation of the power P by integrating over a small sphere using Equation 46.6 is thus often preferred for numerical stability. The power collected by the lens of numerical aperture NA can be evaluated by computing the electromagnetic field on the surface of a large sphere Ω using a standard near-field to far-field transformation (Balanis, 1989) and integrating the Poynting vector over the unit solid angle subject to the condition $\theta < \theta_{\text{NA}}$, where $\theta_{\text{NA}} = \text{asin}(\text{NA})$. The efficiency ε is then given as

$$\varepsilon = \frac{\frac{1}{2}\int\limits_{\theta<\theta_{\text{NA}}} \text{Re}\,(\mathbf{E}\times\mathbf{H}^*)\cdot\mathbf{n}_\Omega d\Omega}{P}. \tag{46.17}$$

While Equation 46.17 is exact, this direct procedure does not give complete insight into the governing physics. For this reason, the efficiency in the QED SPS designs is instead treated using an elements-splitting

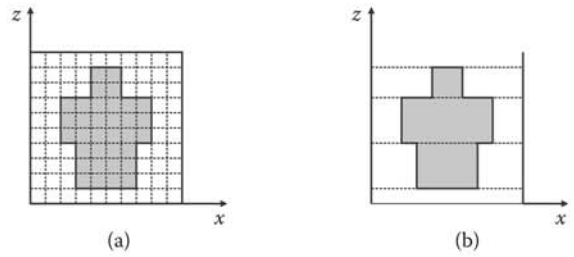


FIGURE 46.8 Discretization schemes for (a) the finite-difference time-domain technique and for (b) the modal method.

approach where β, Γ_M, $\Gamma_{Rad,}$ and γ are analyzed and optimized separately and the efficiency is computed using $\varepsilon = \beta\gamma$. The rates Γ_M and Γ_{Rad} are available from Equation 46.11 as

$$\Gamma_M = \frac{\pi\omega_0}{\hbar\varepsilon_0} |\mathbf{d}|^2 \left|\mathbf{n}_d \cdot \widetilde{\mathbf{E}}_M(\mathbf{r}_0)\right|^2 \tag{46.18}$$

$$\Gamma_{Rad} = \frac{\pi\omega_0}{\hbar\varepsilon_0} |\mathbf{d}|^2 \int\limits_{\alpha \neq M} \left|\mathbf{n}_d \cdot \widetilde{\mathbf{E}}_\alpha(\mathbf{r}_0)\right|^2 \delta(\omega_0 - \omega_\alpha) d\alpha, \tag{46.19}$$

and we observe that their evaluation requires access to the electric field profiles of the optical mode of interest as well as the radiation modes.

When performing numerical simulations of SPS designs, the advantages of the spatial discretization methods are their maturity, their availability in commercial user-friendly software packages (Lumerical, n.d.; JCMWave, n.d.), and their ability to exploit parallelization as required in the treatment of large geometries. Their major weakness is that they do not give direct access to the mode profiles $\widetilde{\mathbf{E}}_\alpha(\mathbf{r})$. While the total power P can be obtained from Equation 46.6 or Equation 46.7, the individual rates Γ_M and Γ_{Rad} are not directly available and β thus cannot be directly evaluated. On the other hand, in the modal expansion methods, the mode profiles $\widetilde{\mathbf{E}}_\alpha(\mathbf{r})$ are directly available, and this represents a major advantage. However, the drawbacks of the modal method are the lack of user-friendly software packages as well as its difficulty in handling large 3-D geometries due to the lack of eigenvalue solvers capable of exploiting parallelization.

46.3 Cavity QED Approach

Inspection of Equation 46.1 reveals two options for optimizing β. The first is to increase Γ_M using cavity QED effects in the weak coupling regime. Here, the emitter is placed inside a cavity with quality factor Q, mode volume V, and resonance wavelength $\lambda_r = 2\pi c/\omega_r$ as illustrated in Figure 46.9a. If the emitter is placed at a maximum of the optical field profile and the emitter and cavity lines are spectrally aligned such that $\omega_r = \omega_0$, the spontaneous emission rate Γ_M into the cavity mode is enhanced by the factor F_P known as the Purcell factor (Purcell, 1946). It describes the rate Γ_M relative to the rate Γ_{Bulk} for an emitter in a bulk medium and is given by (Gérard, 2003)

$$F_P = \frac{\Gamma_M}{\Gamma_{Bulk}} = \frac{3}{4\pi^2} \frac{Q}{V} \left(\frac{\lambda_r}{n}\right)^3, \tag{46.20}$$

where n is the refractive index of the material enclosing the emitter. The mode volume is computed as†

$$V = \frac{\int \varepsilon_r(\mathbf{r}) \left|\widetilde{\mathbf{E}}_M(\mathbf{r})\right|^2 d\mathbf{r}}{\varepsilon_r(\mathbf{r}_{max}) \left|\widetilde{\mathbf{E}}_M(\mathbf{r}_{max})\right|^2}, \tag{46.21}$$

where the denominator is evaluated at the position $\mathbf{r}_{max}$ of a maximum of the electric field intensity profile.

The expression (Equation 46.20) holds provided the emitter line is narrow compared to the cavity linewidth. If this is not the case, an integration over the electronic density-of-states should be performed,

† When integrated over all space, the integral in Equation 46.21 diverges. To remedy this, it is possible to define a generalized mode volume (Kristensen et al., 2012) that converges in an unambiguous way. However, in high-Q cavities, one can usually employ Equation 46.21 and limit the integration to the device geometry with negligible error.

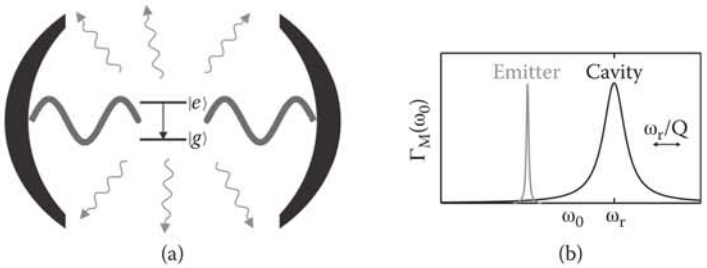


FIGURE 46.9 (a) Cavity-enhanced spontaneous emission. (b) Cavity emission rate Γ_M as a function of frequency.

which generally leads to a reduction in the enhancement. The β factor can be written in terms of the Purcell factor as

$$\beta = \frac{F_P}{F_P + (\Gamma_{Rad}/\Gamma_{Bulk})} \approx \frac{F_P}{F_P + 1}. \tag{46.22}$$

For micropillars with diameters in the few micrometer range, the spontaneous emission into radiation modes is approximately equal to the rate in a bulk material, and we can make the approximation $\Gamma_{Rad} = \Gamma_{Bulk}$ leading to the second equality in Equation 46.22.

We observe from Equations 46.20 through 46.22 that a high-efficiency cavity-based SPS requires a large Q factor combined with a low mode volume, and the optical design challenge for the cavity-based SPS is to maximize the ratio Q/V while maintaining a high outcoupling coefficient γ.

The expression (46.20) assumes perfect spectral alignment between the emitter and the cavity line. If they are detuned, a correction factor given by a Lorentzian lineshape function is introduced, such that (Lorke et al., 2013)

$$\frac{\Gamma_M(\omega_0)}{\Gamma_{Bulk}} = \frac{3}{4\pi^2}\frac{Q}{V}\left(\frac{\lambda_r}{n}\right)^3 \frac{(\delta\omega_r)^2}{(\delta\omega_r)^2 + (\omega_0 - \omega_r)^2}, \tag{46.23}$$

where $\delta\omega_r = \omega_r/Q$ is the cavity linewidth. The cavity emission rate Γ_M is plotted in Figure 46.9b as function of frequency and we observe that for an emitter-cavity misalignment of only a few linewidths $\delta\omega_r$, the enhancement is lost. For the cavity design approach to work, it is thus absolutely necessary to control the spectral alignment.

46.3.1 Micropillar

The AlGaAs microcavity pillar geometry (Reitzenstein and Forchel, 2010) illustrated in Figure 46.10a is one of the most popular platforms for constructing a high-efficiency SPS. It is a vertical structure consisting of a GaAs cavity with an embedded InAs QD surrounded by GaAs/AlAs distributed Bragg reflectors (DBRs). The optical cavity mode profile is confined vertically by the DBRs and laterally by total internal reflection. The emitted light is collected by an objective lens from above.

The thicknesses of the cavity and the DBR layers should be chosen as λ_d/n_{eff} and $\lambda_d/(4n_{eff})$, respectively, where λ_d is the design wavelength and n_{eff} is the effective index of the fundamental mode. These thicknesses ensure that the resonance is centered on the DBR stopband for maximum reflection (Coldren et al., 2012).

In the ideal structure, light escapes only through the bottom and top DBRs. The light emitted into the substrate is lost and the reflectivity of the bottom DBR should thus be significantly higher than that of the top DBR to obtain a high efficiency. The transmission γ to the lens for a micropillar of infinite diameter is presented in Figure 46.10b, and we observe that $\gamma > 0.99$ requires approximately 18 extra layer pairs in the bottom DBR.

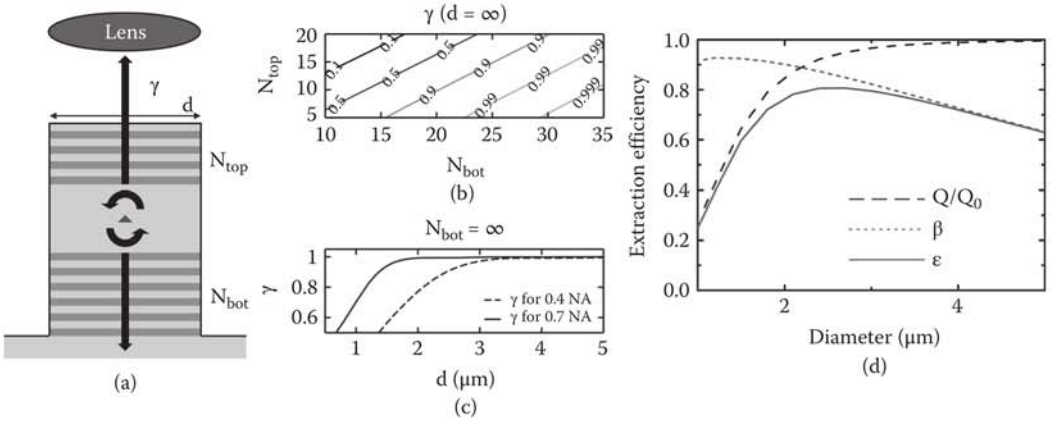


FIGURE 46.10 (a) The micropillar geometry. Transmission γ as function of DBR layer pairs (b) and as function of the pillar diameter (c). (d) Extraction efficiency as function of the pillar diameter. (Reprinted from Gregersen, N. et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2013 IEEE and reprinted by permission from Macmillan Publishers Ltd: Gazzano, O. et al., *Nature Communications* 4: 1425, copyright 2013.)

After the numbers of layer pairs are fixed, the pillar diameter should be chosen. The transmission γ as function of the pillar diameter is shown in Figure 46.10c for a perfectly reflecting bottom DBR. The far-field emission pattern is related to the Fourier transform of the near-field profile. For small diameters, the far-field emission pattern is broad and the transmission is low. For larger diameters, the beam divergence is reduced and the transmission improves. We observe that a pillar diameter above ~2 μm leads to near-unity transmission for a 0.7 NA lens.

In the optimization of the extraction efficiency, two additional mechanisms should be considered. While the ideal structure emits light only along the vertical axis, fabricated pillar structures suffer from sidewall imperfections leading to in-plane scattering of light. The amount of light lost can be estimated by measuring the Q factor and comparing with the quality factor Q_0 of a planar sample without microstructuring. The light lost due to imperfections is given by $1 - Q/Q_0$ and the predicted efficiency should then be corrected by the fraction Q/Q_0 such that $\varepsilon = \beta\gamma Q/Q_0$. The dashed curve in Figure 46.10d shows a fit to experimentally measured Q/Q_0 ratios for the pillar geometry described in Gazzano et al. (2013). As the diameter increases, the relative overlap of the mode profile with the imperfections decreases, and Q thus increases with diameter toward the planar value.

While a large pillar diameter optimizes γ and Q/Q_0 toward unity, it also increases the mode volume V leading to a reduced F_P. When choosing the diameter, it is thus necessary to consider the magnitude of β as function of diameter as given by Equation 46.22. The calculated ratio $\beta = F_P/(F_P + 1)$ is displayed in Figure 46.10d as the dotted curve and it decreases with the diameter due to the increase in V. The product $\varepsilon = \beta Q/Q_0$ is shown in the figure as the full curve, and we observe that a maximum of ~ 0.8 is obtained for $d \sim 2.5$ μm. This estimation of the efficiency assumes a transmission γ of unity, which is a good approximation when using a 0.7 NA lens.

This micropillar SPS reported in Gazzano et al. (2013) indeed features an efficiency of 0.79. The simultaneous spectral and spatial alignment was achieved using an *in situ* fabrication method pioneered by the group.

46.4 Waveguide QED Approach

The second option for optimizing β as defined by Equation 46.1 is to decrease Γ_{Rad} and this represents the waveguide QED approach. Two main geometries taking advantage of this approach are the photonic nanowire presented in Section 46.4.1 and the photonic crystal waveguide discussed in Section 46.4.2.

Assuming that the QD is positioned at a maximum of the guided mode, the spontaneous emission rate to the guided mode in an infinite uniform waveguide is given by (Claudon et al., 2013)

$$\frac{\Gamma_{\mathrm{M}}}{\Gamma_{\mathrm{Bulk}}} = \frac{3}{4\pi A}\left(\frac{\lambda}{n}\right)^2 \frac{n_{\mathrm{g}}}{n}. \tag{46.24}$$

Here, n_{g} is the group index and

$$A = \frac{\int \varepsilon_r(\mathbf{r}_\perp)\left|\mathbf{E}(\mathbf{r}_\perp)\right|^2 \mathrm{d}\mathbf{r}_\perp}{\varepsilon_r(\mathbf{r}_{\max})\left|\mathbf{E}(\mathbf{r}_{\max})\right|^2}, \tag{46.25}$$

where the integration is over the lateral plane normal to the waveguide axis. In the case of the periodic waveguide of periodicity a, the rate is given by

$$\frac{\Gamma_{\mathrm{M}}}{\Gamma_{\mathrm{Bulk}}} = \frac{3a}{4\pi V}\left(\frac{\lambda}{n}\right)^2 \frac{n_{\mathrm{g}}}{n}, \tag{46.26}$$

where the integration (46.21) of the mode volume V is limited to a single periodic cell.

The waveguide QED approach relies on significant suppression of the spontaneous emission into radiation modes, and this rate Γ_{Rad} should be quantified. The suppression occurs for waveguide diameters on the order of λ/n, and Γ_{Rad} can no longer be approximated to its value in a bulk material. Instead, Γ_{Rad} should be computed using full 3-D numerical simulations. Whereas the calculation of Γ_{M} is usually straightforward, a precise computation of Γ_{Rad} is generally quite demanding because the radiation modes propagate away from the device and perfect absorbing boundary conditions are required to avoid artificial reflections from the boundary walls. It is often determined indirectly by first computing the total spontaneous emission rate Γ_{T}, which for most methods is more accessible, and then subtracting Γ_{M} such that $\Gamma_{\mathrm{Rad}} = \Gamma_{\mathrm{T}} - \Gamma_{\mathrm{M}}$.

46.4.1 Photonic Nanowire

The photonic nanowire (Claudon et al., 2013) is a vertical GaAs cylinder with an embedded QD sitting on a substrate. The geometry is illustrated in Figure 46.11a. Light is guided in the nanowire by total internal reflection at the nanowire-air interface. Single photons emitted by the QD are predominantly coupled to the fundamental HE_{11} waveguide mode as will be discussed in the following. The emission is again collected from an objective lens above the nanowire.

Similar to the micropillar SPS, a bottom mirror is required to avoid loss of light into the substrate. Furthermore, the small nanowire diameter leads to a highly divergent far-field emission pattern, and a top taper strategy is required to reduce this divergence. The reader is referred to (Claudon et al., 2013; Gregersen et al., 2013) for more details on the mirror and the taper strategies. In this tutorial, we will simply study spontaneous emission in the infinite nanowire of Figure 46.11b.

The physical origin of the strong suppression of the spontaneous emission into radiation modes can be understood from Figure 46.11c and d. We first consider a radiation mode defined from the illumination of an infinite photonic nanowire by a plane wave $E_x(\mathbf{r}) = E_0 \exp(ik_0 y)$ propagating along the y axis. The corresponding field intensity profile is shown in Figure 46.11c for a nanowire with diameter $d << \lambda$. Outside the nanowire we observe some field enhancement; however, inside the nanowire the field is strongly suppressed. This field suppression is due to a dielectric screening effect (Bleuse et al., 2011) illustrated in Figure 46.11d, which displays the refractive index profile $n(x)$ and the normalized electric field component $E_x(x)/E_0$ in along the $y = 0$ axis. Now, at the nanowire-air interface, continuity is dictated by Maxwell's equations for the component $D_x(x) = n(x)^2 E_x(x)$ of the displacement field normal to the boundary.

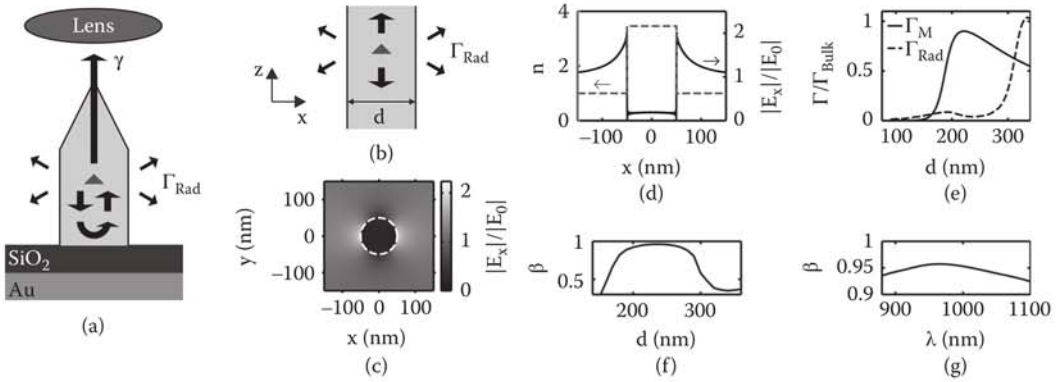


FIGURE 46.11 (a) The full photonic nanowire SPS geometry. (b) The infinite nanowire geometry. (c) The normalized E_x field amplitude of a plane wave propagating along the y axis and scattered by a nanowire of diameter $d = 100$ nm. The nanowire boundary is represented by the dashed white line. (d) Normalized E_x field amplitude (solid) and refractive index profile (dashed) at the position $y = 0$. (e) Spontaneous emission rates into the fundamental waveguide mode (solid) and into radiation modes (dashed) as function of nanowire diameter d. Spontaneous emission β factor (f) as function of d and (g) as function of λ for $d = 240$ nm. Other parameters: $\lambda = 950$ nm and $n_{GaAs} = 3.46$ unless otherwise noted. (Reprinted from Gregersen, N. et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2013 IEEE.)

Since the refractive index of GaAs is large, around 3.5, the E_x component inside the nanowire must be correspondingly suppressed. Indeed, an analytic expression for the spontaneous emission inside the nanowire is available in the electrostatic limit, where $d << \lambda$, and is given by

$$\frac{\Gamma}{\Gamma_{Bulk}} = \frac{\Gamma_{Rad}}{\Gamma_{Bulk}} = \frac{4}{n\left(n^2+1\right)^2}. \tag{46.27}$$

We observe that for GaAs the suppression is on the order of 150. The expression (46.27) is valid for a dipole oriented in the plane normal to the axis, which is the case for InAs QDs. However, for a dipole oriented along the z axis, the relevant boundary condition for the radiation modes at the nanowire-air interface is simply continuity of the E_z component. The mechanism ensuring suppression of the electric field is no longer present, and Γ_{Rad} thus remains substantial even when $d << \lambda$.

The normalized spontaneous emission rates Γ_M and Γ_{Rad} are shown in Figure 46.11e for finite nanowire diameters for an on-axis radially oriented dipole. For $d < 150$ nm, both rates are low. As d increases toward 200 nm, the nanowire becomes large enough to support a well-confined fundamental mode and Γ_M increases toward unity. However, for $d < 300$ nm, Γ_{Rad} remains below 0.1 Γ_{Bulk} due to the dielectric screening effect discussed above. The corresponding β factor is presented in Figure 46.11f, and we observe a maximum β of ~0.95 for a diameter of ~240 nm.

Importantly, this screening effect is not a resonant effect. The β factor as a function of wavelength is presented in Figure 46.11g, and we observe that $\beta > 0.93$ is obtained in a 100-nm range. The waveguide QED is thus broadband, and no spectral alignment between an emitter line and a narrow cavity is required. This represents a significant advantage in the fabrication and a major asset of the waveguide QED approach.

46.4.2 Photonic Crystal Waveguide

For integrated quantum photonics, in-plane emission of the emitted light into a planar waveguide is desired. A platform allowing for the efficient in-plane emission of single photons is the photonic crystal (PhC) membrane with an embedded QD (Yao, Manga Rao, and Hughes, 2010) as illustrated in Figure 46.12a. By

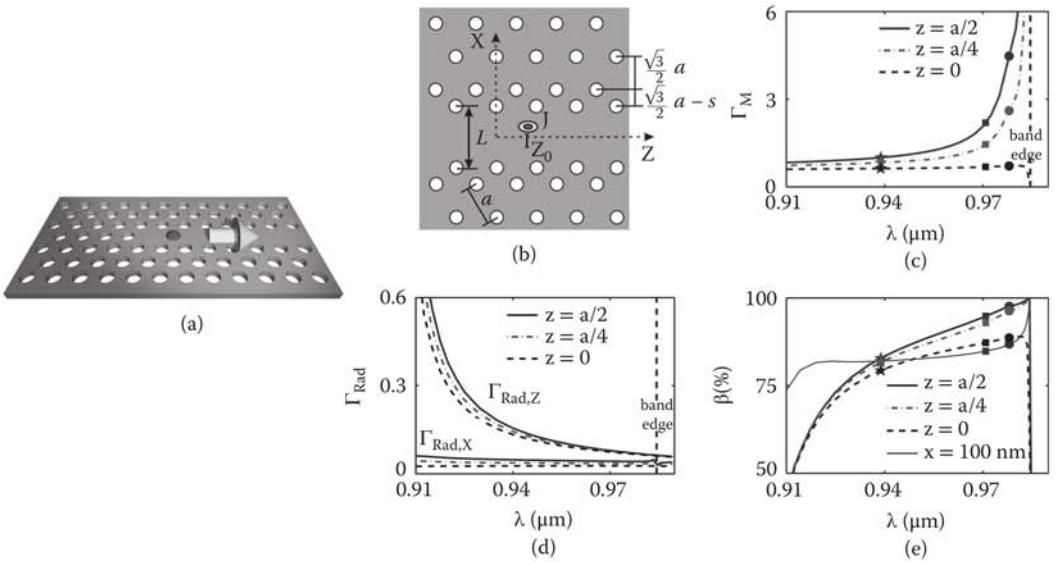


FIGURE 46.12 (a) The PhC waveguide SPS. (b) Geometry of the infinite periodic PhC waveguide. (c) Normalized SE decay rates Γ_M as function of λ. Normalized SE decay rates (d) into the radiation modes for the two in-plane dipole moment orientations X and Z. (e) Associated β factors. The thin solid curve is obtained for an off-axis dipole location ($x = 100$ nm and $z = a/2$). Circles, squares, and stars are reference marks locating the fundamental Bloch-mode group velocities equal to $c/20$, $c/10$, and $c/5$, respectively, where c is the speed of light. (Reprinted with permission from Lecamp, G. et al., *Physical Review Letters*, 99, 023902, 2007. Copyright 2007 by the American Physical Society.)

removing a row of holes, a waveguide can be formed. The waveguide confines light in the plane using the PhC bandgap and vertically using total internal reflection.

In the following, we study the spontaneous emission of the infinite periodic PhC waveguide illustrated in Figure 46.12b. In this tutorial, let us simply consider a dipole positioned in the center of the waveguide $x = 0$ at the position $z = a/2$ with hole displacement $s = 0$. The corresponding results are displayed in Figure 46.12c through 46.12e as full curves. The spontaneous emission rate Γ_M into the fundamental PhC waveguide mode for a dipole oriented along the x axis is shown in Figure 46.12c. For short λ, the rate is similar to that in a bulk material; however, near the band edge Γ_M diverges. This strong enhancement is due to the slow-light effect occurring near the band edge, where the group refractive index diverges (Baba, 2008).

The emission into radiation modes is illustrated in Figure 46.12d for dipoles aligned along the x and the z axes. $\Gamma_{Rad,X}$ is strongly suppressed in the entire wavelength regime due to the same dielectric screening effect discussed for the photonic nanowire. As for the photonic nanowire, this suppression does not take place for the dipole oriented along the z axis and $\Gamma_{Rad,Z}$ is small only near the band edge.

The average spontaneous emission β factor for dipoles oriented along the x and z axes can be written as

$$\beta = \frac{\Gamma_M}{\Gamma_M + \Gamma_{Rad,X} + \Gamma_{Rad,Z}} \tag{46.28}$$

and is depicted in Figure 46.12e. We observe that β approaches unity near the band edge thanks to the suppression of Γ_{Rad} combined with the enhanced Γ_M. The β factor remains larger than 0.75 over a 50 nm wavelength range, an effect predominantly due to the broadband suppression of the spontaneous emission into radiation modes. By displacing the QD x position, the range can be increased as illustrated by the thin solid curve for $x = 100$ nm in Figure 46.12e.

46.5 Planar Dielectric Antenna Approach

A final SPS design discussed in this tutorial is the planar dielectric antenna (PDA) illustrated in Figure 46.13a. This design approach (Chen, Götzinger, and Sandoghdar, 2011) differs conceptually from the previous QED strategies. Whereas the quantum emitter in the QED approaches couples to a specific optical cavity or waveguide mode with the coupling coefficient β, the quantum emitter in the PDA design instead couples to a continuous range of radiation modes and the collection efficiency ε is computed by integrating the contributions from this entire range.

We consider the geometry illustrated in the inset of Figure 46.13b, where a quantum emitter with horizontal electric dipole (HED) or vertical electric dipole (VED) moment orientation is placed in a material of refractive index n_2. In this geometry, light will be collected by an objective lens placed below the emitter.

Without any precautions, light is emitted in all directions due to symmetry, and it is necessary to implement a top mirror to reflect light propagating upward back toward the bottom. While a gold mirror can provide a high broadband reflectivity, coupling to surface plasmons reduces the reflectivity for certain angles of incidence (Chen, Götzinger, and Sandoghdar, 2011). However, the high reflectivity can be restored by introducing an intermediate low-index layer of index $n_3 < n_2$.

With the introduction of the metal mirror, the light is now propagating downward with a highly divergent far-field pattern. To reduce this divergence, a solid immersion lens of index $n_1 > n_2$ is introduced and the downward propagating field is now limited to a cone of polar angle $\theta = \sin^{-1}(n_2/n_1)$. The power density and the collection efficiency as function of the polar angle θ is shown in Figure 46.13b for the two dipole orientations. We observe that a collection efficiency above 0.99 is obtained for an angle $\theta = 55°$. The wavelength dependence on the efficiency is presented in Figure 46.13c, and we observe that $\varepsilon > 0.99$ is maintained over a wavelength range of several hundred nanometers for both emitter orientations. The inset of Figure 46.13c shows the influence of the spacer thickness s on the collection efficiency. As s decreases, the increased coupling of the vertical components of the electric field generated by a VED with surface plasmons reduces the efficiency. However, the field generated by the HED has much weaker vertical components; the coupling to surface plasmons is reduced and the efficiency for this dipole orientation thus remains high.

Whereas the PDA design features record-high broadband efficiency, the condition $n_1 > n_2$ imposes constraints on the choice of material for the medium 1 if the emitter is, e.g., a QD in GaAs with index $n_2 = 3.5$. This can be remedied by introducing an intermediate low-index buffer layer between media 1 and 2. However, an important currently unresolved issue for the PDA design remains, namely the highly non-Gaussian far-field emission pattern illustrated in Figure 46.13b, which makes efficient coupling of the emitted photons to the Gaussian mode profile of a single-mode fiber a challenge.

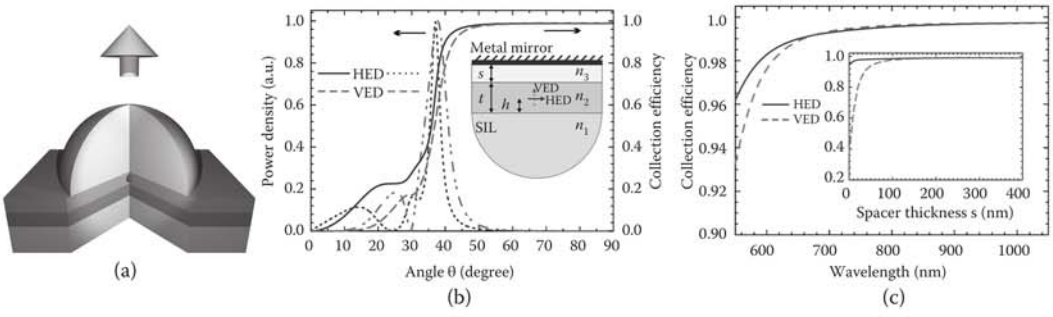


FIGURE 46.13 (a) Planar dielectric antenna. (b) Power density (left axis) and collection efficiency (right axis) versus emission/collection angle θ at a wavelength of 637 nm. The inset shows the schematics of the device with vertically and horizontally oriented dipole emitters. (c) Collection efficiency as a function of wavelength. The inset plots the collection efficiency versus spacer thickness s at 637 nm. (Reprinted from Chen, X,-W. et al., *Optics Letters* 36, 3545–3547, 2011. With permission of Optical Society of America.)

46.6 Decoherence Effects

As discussed above, as well as producing single photons and having a high extraction efficiency, an ideal SPS would produce photons that are quantum mechanically indistinguishable. With reference to Figure 46.2, when two indistinguishable photons are incident on a 50–50 beam splitter they emerge in the same output arm, which provides the fundamental quantum interference necessary for quantum information applications. For photons that are not completely indistinguishable, there remains a nonzero probability that a single photon emerges in both output arms. Based on these arguments, the indistinguishability of a SPS can be quantified by Kaer et al. (2013):

$$I = \frac{\int_0^\infty \int_0^\infty \left| \left\langle \hat{A}^\dagger(t+\tau)\hat{A}(t) \right\rangle \right|^2 dt d\tau}{\int_0^\infty \int_0^\infty \left\langle \hat{A}^\dagger(t)\hat{A}(t) \right\rangle \left\langle \hat{A}^\dagger(t+\tau)\hat{A}(t+\tau) \right\rangle dt d\tau}, \tag{46.29}$$

where $I = 1$ for perfectly indistinguishable photons and $I = 0$ for distinguishable photons. The operators $\hat{A}^\dagger$ and $\hat{A}$ are the creation and annihilation operators for the electric field incident on the beam splitter. Further details of the derivation leading to Equation 46.29 can be found in Unsleber et al. (2015). The challenge in modeling the indistinguishability of photons emitted from a semiconductor-based SPS lies in (a) relating the correlation function and expectation values in Equation 46.29 to operators pertaining to the quantum emitter, and (b) ensuring that the model of the emitter itself captures all important processes that affect the photon indistinguishability.

46.6.1 Quantum Emitter Master Equation

To gain some insight, we first consider a QD emitting in a bulk material or a very low-Q cavity. In this case, provided the emitter is sufficiently far from the beam splitter optics, the field operators in Equation 46.29 can be replaced with operators describing the emitter. We write

$$\hat{A}(t) \sim \hat{\sigma}(t), \tag{46.30}$$

where $\hat{\sigma} = |g\rangle\langle e|$ annihilates an excitation of the emitter (Kiraz et al., 2004). The proportionality constants that we have omitted are independent of time, and with reference to Equation 46.29 we see that they will cancel when the indistinguishability is calculated.

Our task now is to calculate the correlation function in the numerator of Equation 46.29, which can be achieved with help from the quantum regression theorem. To begin, we must first write down a master equation describing the dynamics of the quantum emitter, which takes the form of an equation of motion for the density operator $\hat{\rho}$ (Carmichael, 1999). This can be done to various levels of rigor, but in order to capture the most important qualitative features it suffices to write

$$\frac{d}{dt}\hat{\rho} = \Gamma_{\text{Bulk}} L_{\hat{\sigma}}(\hat{\rho}) + 2\gamma^* L_{\hat{\sigma}^\dagger\hat{\sigma}}(\hat{\rho}), \tag{46.31}$$

where the Lindblad operators are defined to satisfy $L_{\hat{O}}(\hat{\rho}) = \hat{O}\hat{\rho}\hat{O}^\dagger - \left(\hat{O}^\dagger\hat{O}\hat{\rho} + \hat{\rho}\hat{O}^\dagger\hat{O}\right)/2$ and Γ_{Bulk} is the bulk spontaneous emission rate. The first term on the right side of Equation 46.31 describes the decay of the excited state while the second term describes the decoherence of the two-level system. The quantity γ^* is the dephasing or decoherence rate, which we will discuss further in the following sections. Equation 46.31

allows us to establish equations of the motion for expectation values of the emitter. The dipole expectation value, e.g., obeys

$$\frac{d}{dt}\left\langle \hat{\sigma} \right\rangle = -\frac{1}{2}\left(\Gamma_{\text{Bulk}} + 2\gamma^*\right)\left\langle \hat{\sigma} \right\rangle. \tag{46.32}$$

Armed with Equation 46.31, we can now invoke the quantum regression theorem, which states that the correlation function in Equation 46.29 obeys the same equation of motion as the dipole operator in Equation 46.32 but with respect to τ, namely $\frac{d}{d\tau}\left\langle \hat{\sigma}^\dagger(t+\tau)\hat{\sigma}(t) \right\rangle = -\frac{1}{2}\left(\Gamma_{\text{Bulk}} + 2\gamma^*\right)\left\langle \hat{\sigma}^\dagger(t+\tau)\hat{\sigma}(t) \right\rangle$. The initial condition is found by setting $\tau = 0$, from which we find $\left\langle \hat{\sigma}^\dagger(t)\hat{\sigma}(t) \right\rangle = e^{-\Gamma_{\text{Bulk}}t}$ if we assume the emitter is initially in its excited state at time $t = 0$. This allows us to find the correlation function, and inserting this solution into Equation 46.29 we find that the indistinguishability is given by Kaer et al. (2013):

$$I = \frac{\Gamma_{\text{Bulk}}}{\Gamma_{\text{Bulk}} + 2\gamma^*}. \tag{46.33}$$

This expression demonstrates that for our simplistic model, the indistinguishability depends solely on the ratio, $\Gamma_{\text{Bulk}}/\gamma^*$ that we should seek to maximize.

46.6.2 Emitter-Cavity Master Equation

So far we have considered a QD emitting into a bulk medium. However, using similar techniques to those above it is also possible to calculate the indistinguishability of photons emitted from a coupled emitter-cavity system. The calculation is performed by replacing the master equation in Equation 46.31 with the more complicated expression (Kaer et al., 2013)

$$\frac{d}{dt}\hat{\rho} = -i\left[\hat{H}, \hat{\rho}\right] + \Gamma_{\text{Bulk}} L_{\hat{\sigma}}(\hat{\rho}) + 2\gamma^* L_{\hat{\sigma}^\dagger\hat{\sigma}}(\hat{\rho}) + \kappa L_{\hat{a}}(\hat{\rho}). \tag{46.34}$$

Here, $\hat{H} = g\left(\hat{\sigma}\hat{a}^\dagger + \hat{a}\hat{\sigma}^\dagger\right)$ describes the interaction between the emitter and the cavity with g the emitter-cavity coupling strength and $\hat{a}$ the annihilation operator of the cavity mode. The last term in Equation 46.34 describes the outcoupling of light from the cavity with κ the cavity mode decay rate. Broadly speaking, provided the losses in the system are weaker than the emitter-cavity coupling strength, the cavity mode can be adiabatically eliminated from this equation of motion. For details of this procedure, we refer the reader to Kaer et al. (2013). This results in a master equation describing the emitter only, which is equal to Equation 46.31 but with Γ_{Bulk} replaced with $\Gamma = \Gamma_{\text{Bulk}} + \Gamma_{\text{M}}$ with $\Gamma_{\text{M}} = \Gamma_{\text{Bulk}} F_{\text{P}}$ the emission rate into the cavity, and the Purcell factor now given by

$$F_{\text{P}} = \frac{4g^2}{\kappa\Gamma_{\text{Bulk}}}. \tag{46.35}$$

We note that the Purcell factor given here is in fact equivalent to that given in Equation 46.20. The apparent difference is a matter of notation, with Equation 46.35 being preferred in the present context since it makes explicit the dependence on the emitter-cavity coupling strength g. With this modification, the indistinguishability measure for a weakly coupled emitter-cavity system becomes

$$I = \frac{\left(F_{\text{P}} + 1\right)\Gamma_{\text{Bulk}}}{\left(F_{\text{P}} + 1\right)\Gamma_{\text{Bulk}} + 2\gamma^*}. \tag{46.36}$$

This expression encompasses the two main strategies used to increase photon indistinguishability. First, as one might expect, eliminating sources of decoherence, and thus reducing γ^*, serves to increase I. Methods to achieve this are briefly discussed in the following section. Second, placing the emitter in a photonic structure such as a micropillar can give rise to a Purcell enhancement of the spontaneous emission rate,

which will also increase the indistinguishability. In either case, the strategy amounts to extracting the single photon on a timescale faster than the decoherence rate. We note, however, that these simple arguments only apply as long as the dephasing rate γ^* and the Purcell factor are independent parameters. As we will see below, dephasing caused by coupling to phonons can increase as the Purcell factor is increased, meaning that increasing photon indistinguishability by increasing the spontaneous emission rate may not always be advisable.

46.6.3 Spectral Wandering

Having introduced the basic theory behind photon indistinguishability, we shall now discuss two main mechanisms of dephasing in semiconductor emitters. We begin with spectral wandering, which refers to the process by which the spectral emission line of a QD moves (wanders) over time (Berthelot et al., 2006). In such a case, successively emitted photons may have different frequencies and therefore cannot be considered indistinguishable. Many processes in semiconductor systems can give rise to spectral wandering, with the amplitude and frequency distribution being dependent on the nature and timescale of the process. The most common process in charge-neutral QDs is spectral diffusion caused by the temporary trapping of charges in nearby sites, which induces a change of the electrostatic environment of the QD.

In order to model spectral diffusion, it is typically sufficient to include a phenomenological dephasing term in the master equation describing the QD degrees of freedom as in Equation 46.31 above. As we have seen, dephasing of this form can be combated by increasing the Purcell effect. Recalling also that the efficiency of a source increases with the Purcell effect, we conclude that in the case of simple phenomenological dephasing, such as that caused by spectral wandering, an increased Purcell factor is always advantageous.

46.6.4 Coupling to Longitudinal Acoustic Phonons

Coupling to phonons, present in any semiconductor sample, represents an additional source of decoherence. It has been demonstrated that the dominant phonon-mediated dephasing mechanism is due to coupling of the QD exciton to longitudinal acoustic phonons via the deformation potential (Ramsay et al., 2010). Phonons deform the QD structure and give rise to fluctuations in the exciton energy, which in turn cause a dephasing of superpositions of the ground and excited state. As such phonons are sensitive to the QD dressed state energy difference, which itself depends on the strength of any QD-cavity coupling g. Therefore, in contrast to spectral wandering, phonons cannot be accurately included by the addition of a simple phenomenological dephasing rate that is independent of other parameters.

For these reasons, correctly incorporating phonon coupling into the description of a QD-based SPS is particularly challenging. One approach is to include the exciton-phonon coupling at a Hamiltonian level, where one essentially takes a large but finite number of harmonic oscillators to model the phonon environment. The dynamics describing the complete system can then be solved numerically, giving results that are exact to within a set numerical tolerance (Kaer et al., 2012, 2013). This method, however, requires a large amount of computing power and can obscure the underlying physical processes. Another less numerically demanding method involves deriving a time-local master equation describing the emitter-cavity system using projection operators. This results in more tractable expression from which the indistinguishability can be calculated, though is valid only to second order in the exciton–phonon coupling strength. A comparison of these methods and a simple phenomenological approach can be found in Kaer et al. (2013).

Here we will adopt a simple model that manages to capture the salient qualitative features while being simple to implement. We replace the constant dephasing rate γ^* in Equation 46.34 with the expression (Ramsay et al., 2010)

$$\gamma_{\mathrm{ph}} = \frac{\pi}{4} J(2g) \coth\left(\frac{g}{k_{\mathrm{B}}T}\right), \tag{46.37}$$

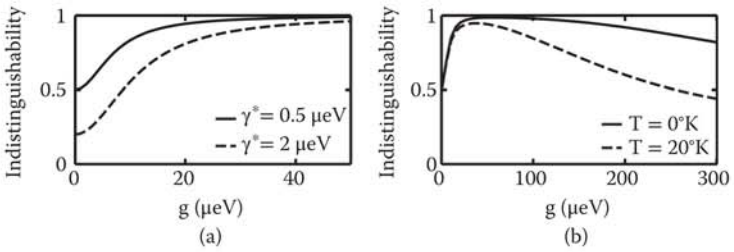


FIGURE 46.14 (a) Photon indistinguishability as a function of emitter-cavity coupling strength for a simple constant dephasing rate of $\gamma^* = 0.5\,\mu eV$ (solid) and $\gamma^* = 2\,\mu eV$ (dashed). (b) A similar plot where the dephasing rate is composed of a constant value of $\gamma^* = 0.5\,\mu eV$ plus a phonon-induced contribution given by Equation 46.37 for temperatures of $T = 0$ K (solid) and $T = 20$ K (dashed). Other parameters are $\Gamma_{Bulk} = 1\,\mu eV$, $\kappa = 100\,\mu eV$, $\eta = 0.032\ ps^2$ and $\omega_c = 0.95\,\mu eV$.

where T is the sample temperature and the function $J(\omega)$ is the exciton-phonon spectral density. For typical GaAs QDs, the spectral density takes the form

$$J(\omega) = \eta\omega^3 e^{-\left(\frac{\omega}{\omega_c}\right)^2}, \tag{46.38}$$

where η is a quantity describing the overall strength of the interaction and can be determined from material parameters, while ω_c is a characteristic cutoff frequency that is inversely proportional to the linear dimension of the QD size (Ramsay et al., 2010). Using Equation 46.37 in Equation 46.34 and again invoking the quantum regression theorem, Equation 46.29 can be evaluated to explore the indistinguishability of photons emitted from a coupled QD-cavity system subject to dephasing caused by phonons.

One final consideration is the correct operator to insert into Equation 46.29. Since in the majority of this chapter we are interested in the emission via the cavity mode, if this mode cannot be adiabatically eliminated it is more appropriate to use the cavity operator $\hat{a}$ rather than the emitter operator in Equation 46.29. In doing so, we are explicitly considering the indistinguishability of photons emitted into the cavity mode of interest.

In Figure 46.14a, we show the photon indistinguishability as a function of the emitter-cavity coupling strength g for a constant dephasing rate of $\gamma^* = 0.5\,\mu eV$ (solid) and $\gamma^* = 2\,\mu eV$ (dashed). For $g = 0$, the indistinguishability is given by its value in bulk (Equation 46.33). As g is increased, the coupling of the emitter to the cavity gives rise to an enhancement of the Purcell effect that accelerates the spontaneous emission, thus increasing the indistinguishability. Figure 46.14b shows a similar plot where we now include a constant dephasing rate $\gamma^* = 0.5\,\mu eV$ and also a phonon contribution given by Equation 46.37. For small values of g, we again observe an increase in I due to the Purcell enhancement. However, as g is increased further, the phonon contribution to the dephasing increases, which serves to reduce the indistinguishability. These plots demonstrate the subtle interplay between the beneficial effects of the Purcell enhancement and the detrimental effects of increased coupling to phonons. The reader is referred to (Kaer et al., 2012, 2013; Kaer and Mørk, 2014) for a more extensive discussion of these effects.

46.7 Summary

In this chapter, we have presented the key figures of merit for the single-photon source. For applications in quantum information processing, clean single-photon emission, efficient photon extraction, and indistinguishable photon emission are required. Sources based on spontaneous parametric down-conversion or attenuation are simple to implement but suffer inherently from low efficiency. We have presented the two-level system with focus on the semiconductor quantum dot as an attractive platform for realizing

a deterministic highly efficient single-photon source. Design strategies for controlling the light emission and ensuring a near-unity coupling efficiency based on cavity and waveguide quantum electrodynamics approaches as well as the planar dielectric antenna have been presented. We have discussed the theory and major numerical simulation techniques for modeling the light emission. Finally, we have discussed the influence of several decoherence processes arising from the solid-state environment including simple pure dephasing mechanisms such as charge fluctuations and the more advanced phonon-induced decoherence process due to interaction with longitudinal acoustic phonons.

While, in recent years, tremendous progress has been witnessed towards achieving efficient and pure emission of indistinguishable photons, the simultaneous realization of near-unity efficiency and indistinguishability remains elusive. Fundamental challenges persist both on the design level and on the fabrication level making single-photon sources an interesting research topic most likely for many years to come.

References

Baba, T. 2008. Slow light in photonic crystals. *Nature Photonics* 2 (8): 465–473. doi:10.1038/nphoton.2008.146.

Babinec, T. M., B. J. M. Hausmann, M. Khan, Y. Zhang, J. R. Maze, P. R. Hemmer, and M. Lončar. 2010. A diamond nanowire single-photon source. *Nature Nanotechnology* 5 (3): 195–199. doi:10.1038/nnano.2010.6.

Balanis, C. A. 1989. *Advanced Engineering Electromagnetics*. 1st ed. Hoboken, NJ: Wiley.

Berthelot, A., I. Favero, G. Cassabois, C. Voisin, C. Delalande, P. Roussignol, R. Ferreira, and J-M. Gérard. 2006. Unconventional motional narrowing in the optical spectrum of a semiconductor quantum dot. *Nature Physics* 2 (11): 759–764. doi:10.1038/nphys433. http://arxiv.org/abs/cond-mat/0610346/nhttp://www.nature.com/doifinder/10.1038/nphys433.

Bleuse, J., J. Claudon, M. Creasey, N. S. Malik, J-M. Gérard, I. Maksymov, J-P. Hugonin, and P. Lalanne. 2011. Inhibition, enhancement, and control of spontaneous emission in photonic nanowires. *Physical Review Letters* 106 (10): 103601. doi:10.1103/PhysRevLett.106.103601.

Brendel, J., N. Gisin, W. Tittel, and H. Zbinden. 1999. Pulsed energy-time entangled twin-photon source for quantum communication. *Physical Review Letters* 82 (12): 2594–2597. doi:10.1103/PhysRevLett.82.2594. http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.82.2594.

Buttler, W. T., R. J. Hughes, S. K. Lamoreaux, G. L. Morgan, J. E. Nordholt, and C. G. Peterson. 2000. Daylight quantum key distribution over 1.6 Km. *Physical Review Letters* 84 (24): 5652–5655.

Carmichael, H. J. 1999. *Statistical Methods in Quantum Optics*. 1st ed. Berlin-Heidelberg: Springer.

Chen, X-W., S. Götzinger, and V. Sandoghdar. 2011. 99% efficiency in collecting photons from a single emitter. *Optics Letters* 36 (18): 3545–3547. doi:10.1364/OL.36.003545. http://ol.osa.org/abstract.cfm?URI=ol-36-18-3545.

Claudon, J., J. Bleuse, N. S. Malik, M. Bazin, P. Jaffrennou, N. Gregersen, C. Sauvan, P. Lalanne, and J-M. Gérard. 2010. A highly efficient single-photon source based on a quantum dot in a photonic nanowire. *Nature Photonics* 4 (3): 174–177. doi:10.1038/nphoton.2009.287. http://dx.doi.org/10.1038/nphoton.2009.287.

Claudon, J., N. Gregersen, P. Lalanne, and J-M. Gérard. 2013. Harnessing light with photonic nanowires: Fundamentals and applications to quantum optics. *ChemPhysChem* 14 (11): 2393–2402. doi:10.1002/cphc.201300033.

Coldren, L. A., S. W. Corzine, and M. L. Mashanovitch. 2012. *Diode Lasers and Photonic Integrated Circuits*. 2nd ed. Hoboken, NJ: Wiley.

Ding, X., Y. He, Z.-C. Duan, N. Gregersen, M.-C. Chen, S. Unsleber, S. Maier, et al. 2016. On-demand single photons with high extraction efficiency and near-unity indistinguishability from a resonantly driven quantum dot in a micropillar. *Physical Review Letters* 116 (2): 020401. doi:10.1103/PhysRevLett.116.020401. http://arxiv.org/abs/1601.00284.

Gazzano, O., S. Michaelis de Vasconcellos, C. Arnold, A. Nowak, E. Galopin, I. Sagnes, L. Lanco, A. Lemaître, and P. Senellart. 2013. Bright solid-state sources of indistinguishable single photons. *Nature Communications* 4: 1425. doi:10.1038/ncomms2434. http://www.ncbi.nlm.nih.gov/pubmed/23385570.

Gérard, J.-M. 2003. Solid-state cavity-quantum electrodynamics with self-assembled quantum dots. In *Single Quantum Dots, Topics in Applied Physics*, vol. 90, edited by P. Michler, 269–315. Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-540-39180-77. http://www.springerlink.com/index/4nn94hfbm1gpmyxq.pdf/nhttp://www.springerlink.com/content/4NN94HFBM1GPMYXQ.

Gerry, C. C. and P. L. Knight. 2008. *Introductory Quantum Optics*. New York, NY: Cambridge University Press.

Gregersen, N., P. Kaer, and J. Mørk. 2013. Modeling and design of high-efficiency single-photon sources. *IEEE Journal of Selected Topics in Quantum Electronics* 19 (5): 9000516. doi:10.1109/JSTQE.2013.2255265. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6493378.

Hong, C. K., Z. Y. Ou, and L. Mandel. 1987. Measurement of subpicosecond time intervals between two photons by interference. *Physical Review Letters* 59 (18): 2044–2046. doi:10.1103/PhysRevLett.59.2044. http://journals.aps.org.globalproxy.cvt.dk/prl/abstract/10.1103/PhysRevLett.59.2044.

JCMWave. n.d. JCMSuite. http://www.jcmwave.com.

Kaer, P., N. Gregersen, and J. Mørk. 2013. The role of phonon scattering in the indistinguishability of photons emitted from semiconductor cavity QED systems. *New Journal of Physics* 15: 035027. doi:10.1088/1367-2630/15/3/035027.

Kaer, P. and J. Mørk. 2014. Decoherence in semiconductor cavity QED systems due to phonon couplings. *Physical Review B* 90 (3): 035312. doi:10.1103/PhysRevB.90.035312. http://journals.aps.org/prb/abstract/10.1103/PhysRevB.90.035312.

Kaer, P., T. R. Nielsen, P. Lodahl, A.-P. Jauho, and J. Mørk. 2012. Microscopic theory of phonon-induced effects on semiconductor quantum dot decay dynamics in cavity QED. *Physical Review B* 86 (8): 085302. doi:10.1103/PhysRevB.86.085302.

Kiraz, A., M. Atatüre, and A. Imamoğlu. 2004. Quantum-dot single-photon sources: Prospects for applications in linear optics quantum-information processing. *Physical Review A* 69 (3): 032305. doi:10.1103/PhysRevA.69.032305. http://journals.aps.org/pra/abstract/10.1103/PhysRevA.69.032305.

Kristensen, P. T., C. Van Vlack, and S. Hughes. 2012. Generalized effective mode volume for leaky optical cavities. *Optics Letters* 37 (10): 1649–1651. doi:10.1364/OL.37.001649.

Lavrinenko, A. V., J. Lægsgaard, N. Gregersen, F. Schmidt, and T. Søndergaard. 2014. *Numerical Methods in Photonics*. Boca Raton, FL: CRC Press. http://www.crcpress.com/product/isbn/9781466563889.

Lecamp, G., P. Lalanne, and J. P. Hugonin. 2007. Very large spontaneous-emission β factors in photonic-crystal waveguides. *Physical Review Letters* 99 (2): 023902. doi:10.1103/PhysRevLett.99.023902.

Lorke, M., T. Suhr, N. Gregersen, and J. Mørk. 2013. Theory of nanolaser devices: Rate equation analysis versus microscopic theory. *Physical Review B* 87 (20): 205310. doi:10.1103/PhysRevB.87.205310.

Lumerical. n.d. FDTD Solutions. http://www.lumerical.com/.

Novotny, L. and B. Hecht. 2012. *Principles of Nano-Optics*. 2nd ed. New York, NY: Cambridge University Press.

Purcell, E. M. 1946. Spontaneous emission probabilities at radio frequencies. *Physical Review* 69: 681.

Ramsay, A. J., A. V. Gopal, E. M. Gauger, A. Nazir, B. W. Lovett, A. M. Fox, and M. S. Skolnick. 2010. Damping of exciton rabi rotations by acoustic phonons in optically excited InGaAs/-GaAs quantum dots. *Physical Review Letters* 104 (1): 017402. doi:10.1103/PhysRevLett.104.017402. http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.104.017402.

Reitzenstein, S. and A. Forchel. 2010. Quantum dot micropillars. *Journal of Physics D: Applied Physics* 43 (3): 033001. doi:10.1088/0022-3727/43/3/033001.

Shields, A. J. 2007. Semiconductor quantum light sources. *Nature Photonics* 1 (4): 215–223. doi:10.1038/nphoton.2007.46. http://www.nature.com/nphoton/journal/v1/n4/full/nphoton.2007.46.html.

Siyushev, P., G. Stein, J. Wrachtrup, and I. Gerhardt. 2014. Molecular photons interfaced with alkali atoms. *Nature* 509 (7498): 66–70. doi:10.1038/nature13191. http://www.ncbi.nlm.nih.gov/pubmed/24784217.

Somaschi, N., V. Giesz, L. De Santis, J. C. Loredo, M. P. Almeida, G. Hornecker, S. L. Portalupi, et al. 2016. Near-optimal single-photon sources in the solid state. *Nature Photonics* 10 (5): 340–345. doi:10.1038/nphoton.2016.23. http://dx.doi.org/10.1038/nphoton.2016.23.

Unsleber, S., D. P. S. McCutcheon, M. Dambach, M. Lermer, N. Gregersen, S. Höfling, J. Mørk, C. Schneider, and M. Kamp. 2015. Two-photon interference from a quantum dot microcavity: Persistent pure dephasing and suppression of time jitter. *Physical Review B* 91 (7): 075413. doi:10.1103/PhysRevB.91.075413.

Yao, P., V. S. C. Manga Rao, and S. Hughes. 2010. On-chip single photon sources using planar photonic crystals and single quantum dots. *Laser & Photonics Reviews* 4 (4): 499–516. doi:10.1002/lpor.200810081. http://doi.wiley.com/10.1002/lpor.200810081.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

47

Nanoplasmonic Lasers and Spasers

A Freddie Page

and

Ortwin Hess

47.1 Introduction 609
47.2 Models for Metals and Gain 610
Drude–Lorentz Model • Multilevel Maxwell–Bloch Theory
47.3 Transfer Matrix Method 613
47.4 Evolutionary Algorithm 616
47.5 Finite-Difference Time-Domain Method 618
47.6 Loss Compensation 620
47.7 Lasing Dynamics 623
47.8 Stopped-Light Lasing 623
47.9 Summary 624

47.1 Introduction

Semiconductor-based electronics has, in recent years, enjoyed a miniaturization trend down to structure sizes on nanometer scales. However, corresponding photonic devices, such as, in particular, lasers, have not been able to follow suit with a similarly dramatic pace [1]. This is because the generation of photons in small volumes appeared challenging due to the diffraction limit, which apparently seemed to set the minimum wavelength of light to be approximately inversely proportional to its frequency. Considering typical semiconductor band gaps with associated frequencies, this would imply that we would have a lower limit of optical devices and lasers of typically around 1 μm. Yet the availability of nanoscale lasers would be a true starting point for on-chip optical circuits, with applications in telecommunications, data storage, optical computing and sensing. So, can we shrink lasers to be much smaller than the wavelength of the light they emit, particularly down to the nanoscale?

Let us recall that lasers operate on the basis of two features: feedback and gain. In general, a cavity provides feedback by supporting resonant lasing modes (closed loops where light constructively interferes with itself on the round trip). On each loop, the gain medium amplifies the light, and when the amplification is greater than the threshold of energy lost (to dissipation in the material or by escaping the cavity), the system can enter a lasing regime. The physical size of a laser is thus dependent on the cavity supporting lasing modes and this, in turn, is limited by diffraction. Metallic cavities, on the other hand, allow for the circumvention of this diffraction limit by confining light in the subwavelength surface plasmon modes [2,3]. On a metal-insulator interface, light couples to the oscillations of surface charges as *surface plasmon polariton* (SPP) modes. But how does this help? There are two primary features of SPPs that make them useful in nano-optics applications: First, SPPs break the diffraction limit, i.e., they have wavevectors larger than what is generally allowed for a photon of a given frequency. Second, SPPs have or are associated with large field enhancements at the material interface. This combination of shorter wavelengths and higher field densities then significantly enhances the interaction of light with the electronic systems of active media

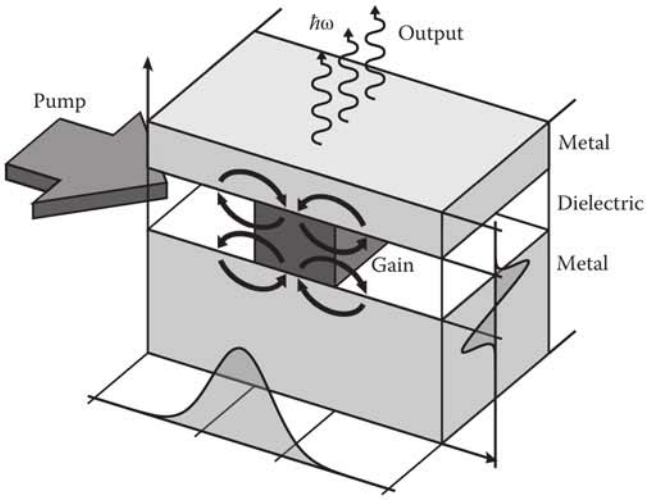


FIGURE 47.1 Nanoplasmonic laser illustration. Light is trapped in a metallic structure such that it forms a closed path (black arrows) that forms in overlap with a gain medium (dark box). In this setup, the gain is pumped from the side and emits from the top. A field strength profile is also shown superimposed on the structure.

[4,5], particularly those interactions that are usually weak. Having subwavelength confinement is, however, frequently traded off to a certain degree with not insignificant dissipative material losses in the metallic components, which requires a corresponding round-trip gain in order to be compensated [6].

Lasers that operate below the diffraction limit on nanometer scales by using metallic cavities are hence called nanoplasmonic lasers. A schematic illustration is given in Figure 47.1. Here there are two options for the light emission: Generated light can be channeled into photonic modes, as in a conventional laser, but can also generate surface plasmons, in which case the laser is more aptly described as a *spaser* for "surface plasmon amplification by stimulated emission of radiation" [7]. Indeed, the possibility to make nanoplasmonic lasers has been explored in various configurations such as, nanorod [8], double-fishnet [9], whispering gallery [10], and stopped light lasers [11].

This chapter will outline a range of techniques that enable the design, optimization, and simulation of nanoplasmonic lasers while tracking the example of a stopped light laser to outline concepts.

47.2 Models for Metals and Gain

47.2.1 Drude–Lorentz Model

In the following, we will chart our theoretical basis for describing the optical response of both metals (which provide subwavelength confinement) and semiconductor gain media (which provides the amplification in a nanoplasmonic laser), involving a combination of a Drude–Lorentz model and a multilevel Maxwell–Bloch theory, respectively, in the harmonic oscillator form. The Drude–Lorentz model describes the oscillations of charges in response to an applied electric field, while a Maxwell–Bloch approach is used to describe the behavior of dipole transitions such as in a four-level system model of a semiconductor.

In its most general terms, the Drude–Lorentz model relates the time evolution of the polarization of a material, **P**, to a driving electric field, **E**, by a second-order differential equation:

$$\frac{\partial^2 \mathbf{P}}{\partial t^2} + \gamma_L \frac{\partial \mathbf{P}}{\partial t} + \omega_L^2 \mathbf{P} = \omega_p^2 \varepsilon_0 \mathbf{E}\,, \tag{47.1}$$

where γ_L is a damping term, ω_L is the restoring force frequency, and ω_p is the plasma frequency—the coupling strength to the driving **E** field. It is a causal response that captures the behavior of most resonance phenomena. In the frequency domain, it is expressed as

$$\tilde{\mathbf{P}} = \varepsilon_0 \frac{\omega_p^2}{\omega_L^2 - \omega(\omega + i\gamma_L)} \tilde{\mathbf{E}} \,, \tag{47.2}$$

where the Drude–Lorentz susceptibility can be extracted:

$$\chi_L(\omega) = \frac{\omega_p^2}{\omega_L^2 - \omega(\omega + i\gamma_L)} \,. \tag{47.3}$$

For $\gamma_L > 0$, this obeys the Kramers–Kronig relations relating real and imaginary parts as required by causality.

The Drude–Lorentz model describes noble metals well, and one may explain the origins of each of the terms with a simple model for a free electron gas. For each electron, let the displacement from its nucleus be given as $\mathbf{x}(t)$. A Newtonian force equation can be built assuming the electron acts as a free particle under the Lorentz force law of an external field **E**, i.e.,

$$m^* \frac{\partial^2 \mathbf{x}}{\partial t^2} = -e\mathbf{E} \,, \tag{47.4}$$

where m^* is the effective mass of an electron and e is its charge. Additional forces can be included, such as a collision term proportional to the velocity and inversely to an average collision time $F_{\text{col}} = -\frac{\partial \mathbf{x}}{\partial t}/\tau$. In this free electron picture, there is no restoring force as for metals, electrons are not bound to their lattice sites. The macroscopic polarization can be related to these electron displacements, as each displacement induces an electronic dipole moment, and these can be averaged to become the polarization by multiplication by the electron number density, $\mathbf{P} = -en\mathbf{x}$. A comparison of terms allows for the associating:

$$\omega_p^2 \to \frac{ne^2}{m^*} \,, \; \gamma \to \tau^{-1} \,, \; \omega_L \to 0 \,. \tag{47.5}$$

This is the Drude model, and with an additional static background permittivity ε_{bg}, it describes metals well.

Real materials typically have multiple resonances, which can be incorporated by adding multiple Drude–Lorentz resonances:

$$\chi = \sum_i \frac{\omega_{pi}^2}{\omega_{Li}^2 - \omega(\omega + i\gamma_{Li})} \,. \tag{47.6}$$

These can have a variety of origins, i.e., in addition to the plasma mode discussed, rotational modes, electronic transitions, etc., and indeed the magnetic response, χ_m, can also be described as a sum of Drude–Lorentz resonances.

47.2.2 Multilevel Maxwell–Bloch Theory

A Drude–Lorentz (harmonic oscillator) type system can also model the electronic transitions between two levels of a semiconductor system, with an energy difference $\hbar\omega_e$. In this picture, there is, however, a field-dependent occupation density of upper and lower levels. Light interacts with this *two-level system* by stimulated emission and absorption processes. When the lower level is more occupied than the upper

level, photons are absorbed and electromagnetic energy is dissipated within the system; however, when the reverse is true, photons can be emitted, stimulated by other photons, adding gain to the system. This dissipative/amplifying response is accompanied by a corresponding refractive response. The strength of the resonance is proportional to the *inversion density* of the two levels, i.e., how much more the higher level is occupied than the lower one, $\Delta N = N_2 - N_1$, such that a layer with emitters embedded may be represented by

$$\varepsilon = \varepsilon_{bg} + \frac{-\Delta N \omega_{pe}^2}{\omega_e^2 - \omega(\omega + i\gamma_e)}, \tag{47.7}$$

$$\omega_{pe}^2 = \sqrt{\varepsilon_{bg}}\gamma_e \sigma_e c, \tag{47.8}$$

where ε_{bg} is the permittivity of the layer hosting the resonance, σ_e is the emission cross-section, and γ_e is the width of the resonance [6]. Note that for negative inversion, the emitter becomes an absorber as there is a higher density of emitters in the lower state. In time, the inversion varies as light is emitted or absorbed. The auxiliary equations to take this into account are

$$\frac{\partial N_1}{\partial t} = -\frac{1}{\hbar\omega_e}\left(\frac{\partial \mathbf{P}_e}{\partial t} + \frac{\gamma_e}{2}\mathbf{P}_e\right)\cdot \mathbf{E} + \frac{N_2}{\tau_{21}}, \tag{47.9}$$

$$\frac{\partial N_2}{\partial t} = +\frac{1}{\hbar\omega_e}\left(\frac{\partial \mathbf{P}_e}{\partial t} + \frac{\gamma_e}{2}\mathbf{P}_e\right)\cdot \mathbf{E} - \frac{N_2}{\tau_{21}}, \tag{47.10}$$

with the first terms representing stimulated emission, and the final ones to account for non-radiative recombination at a rate τ_{21}^{-1}.

The two-level system is appropriate for modeling a single electronic transition; however, there is no way in which a two-level system can reach a state of inversion by relying solely on the processes of spontaneous emission and absorption—the best that can be achieved is equal occupation of the levels. A *four-level system*, on the other hand, can be constructed such that there can be a dynamically maintained population inversion between two of its levels, from where stimulated emission can occur. The construction contains two two-level systems, labeled e and a for emission and absorption; e with energy levels 1 and 2 and energy gap $\hbar\omega_e$, in-between a with energy levels 0 and 3 and greater energy gap $\hbar\omega_a$, as depicted in Figure 47.2. The two upper levels, 2 and 3, are coupled by a fast non-radiative relaxation channel, with rate τ_{32}^{-1}, as are the two lower levels, 0 and 1 (τ_{10}^{-1}). This has the effect of rapidly depleting the first and third levels shortly

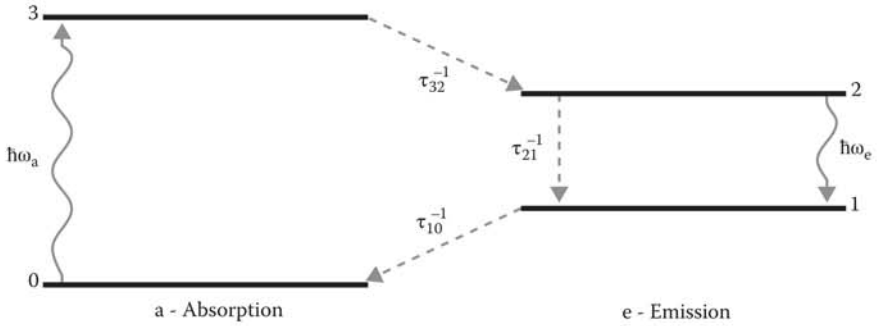


FIGURE 47.2 Schematic of the four-level system. A four-level system comprises two two-level systems, a and e, with active transitions ($0 \leftrightarrow 3$) and ($1 \leftrightarrow 2$), respectively, which are coupled by nonradioactive transition rates τ_{32}^{-1} and τ_{10}^{-1}. In this scheme, the emission subsystem permits radiative transitions at an energy of $\hbar\omega$ and has a slow nonradiative recombination rate τ_{21}^{-1}.

after they become occupied. Between the levels of the emission two-level subsystem, there is additionally a (slow) nonradiative channel (τ_{21}^{-1}).

The key point of a four-level system is that electrons that are pumped from levels 0 and 3, will quickly decay to level 2, leading to an inversion density of level 2 over level 1, which is available for stimulated emission. This allows the optical pumping between levels 0 and 3 to generate inversion between levels 1 and 2.

Four-level systems are incorporated into time-domain simulations using the differential equation for the polarization given in Equation 47.1 for each transition:

$$\frac{\partial^2 \mathbf{P}_\mathrm{a}}{\partial t^2} + \gamma_\mathrm{a} \frac{\partial \mathbf{P}_\mathrm{a}}{\partial t} + \omega_\mathrm{a}^2 \mathbf{P}_\mathrm{a} = -\Delta N \omega_\mathrm{pa}^2 \varepsilon_0 \mathbf{E}, \tag{47.11}$$

$$\frac{\partial^2 \mathbf{P}_\mathrm{e}}{\partial t^2} + \gamma_\mathrm{e} \frac{\partial \mathbf{P}_\mathrm{e}}{\partial t} + \omega_\mathrm{e}^2 \mathbf{P}_\mathrm{e} = -\Delta N \omega_\mathrm{pe}^2 \varepsilon_0 \mathbf{E}\,. \tag{47.12}$$

Here the polarization has been split off into a two parts $\mathbf{P}_\mathrm{a}$ and $\mathbf{P}_\mathrm{e}$ that are connected with the radiative resonances, which are added to the total polarization. The corresponding level occupation densities update with the auxiliary equations [6]

$$\frac{\partial N_3}{\partial t} = \left(\frac{\partial \mathbf{P}_\mathrm{a}}{\partial t} + \frac{\gamma_\mathrm{a}}{2} \mathbf{P}_\mathrm{a} \right) \cdot \mathbf{E} - \frac{N_3}{\tau_{32}}, \tag{47.13}$$

$$\frac{\partial N_2}{\partial t} = \frac{N_3}{\tau_{32}} + \frac{1}{\hbar \omega_\mathrm{e}} \left(\frac{\partial \mathbf{P}_\mathrm{e}}{\partial t} + \frac{\gamma_\mathrm{e}}{2} \mathbf{P}_\mathrm{e} \right) \cdot \mathbf{E} - \frac{N_2}{\tau_{21}}, \tag{47.14}$$

$$\frac{\partial N_1}{\partial t} = \frac{N_2}{\tau_{21}} - \frac{1}{\hbar \omega_\mathrm{e}} \left(\frac{\partial \mathbf{P}_\mathrm{e}}{\partial t} + \frac{\gamma_\mathrm{e}}{2} \mathbf{P}_\mathrm{e} \right) \cdot \mathbf{E} - \frac{N_1}{\tau_{10}}, \tag{47.15}$$

$$\frac{\partial N_0}{\partial t} = \frac{N_1}{\tau_{10}} - \left(\frac{\partial \mathbf{P}_\mathrm{a}}{\partial t} + \frac{\gamma_\mathrm{a}}{2} \mathbf{P}_\mathrm{a} \right) \cdot \mathbf{E}\,, \tag{47.16}$$

which encode both the stimulated terms and nonradiative recombinations.

47.3 Transfer Matrix Method

The geometry of the perhaps simplest types of nanoplasmonic lasers is planar waveguide structures. Here, planar layers of heterogeneous materials are stacked to form a waveguide in the center. The modes of such structures depend strongly on the width and material of each layer and can be calculated using the transfer matrix method (TMM), a linear, frequency-domain analysis that shall be used to design and optimize structures before simulation in time domain.

Isotropic homogeneous materials support transverse plane wave solutions in bulk. These waves have the dispersion relation

$$\varepsilon \mu \frac{\omega^2}{c^2} - k^2 = 0\,, \tag{47.17}$$

and the relative amplitudes of the electric and magnetic fields set by

$$E = \sqrt{\frac{\mu}{\varepsilon}} Z_0 H\,, \tag{47.18}$$

where Z_0 is the impedance of free space. When two materials meet at a planar interface, the solutions in each space must be matched together, and the prescription here is that the E and H field components parallel to the interface are continuous. Each interface couples a forward traveling and a backward traveling wave in each layer together.

Let us consider a stack with an axis parallel to the z-direction and waves traveling in a plane perpendicular to the y-direction. There are two independent polarization solutions, *transverse magnetic* (TM) where H is perpendicular to the plane of incidence (parallel to y) with E parallel to the plane, and *transverse electric* (TE) where E is perpendicular to the plane of incidence. TM equations shall be considered herein as this is the character of the SPP solutions. The corresponding TE equations can be recovered by making the substitutions: $\varepsilon \to \mu$, $\mu \to \varepsilon$, $\mathbf{E} \to -Z_0\mathbf{H}$, $Z_0\mathbf{H} \to \mathbf{E}$.

The system can be solved by performing a change of basis between forward and backward propagating waves and electric and magnetic field components, i.e.,

$$\begin{pmatrix} E_{xi} \\ Z_0 H_{yi} \end{pmatrix} = \underbrace{\begin{pmatrix} u_i & -u_i \\ 1 & 1 \end{pmatrix}}_{\mathbf{U}} \begin{pmatrix} A_i^+ \\ A_i^- \end{pmatrix}, \tag{47.19}$$

and translating the forward and backward field components within a layer:

$$\left.\begin{pmatrix} A_i^+ \\ A_i^- \end{pmatrix}\right|_z = \underbrace{\begin{pmatrix} e^{ik_{zi}(z-z')} & 0 \\ 0 & e^{-ik_{zi}(z-z')} \end{pmatrix}}_{\mathbf{T}} \left.\begin{pmatrix} A_i^+ \\ A_i^- \end{pmatrix}\right|_{z'}, \tag{47.20}$$

with $u_i = ck_{zi}/(\varepsilon_{xi}\omega)$. These are the matrices of the TMM. The matrices, $\mathbf{U}$ and $\mathbf{T}$, are assigned such that $\mathbf{U}$ converts from the basis of forward and backward traveling waves to electric and magnetic field components, and $\mathbf{T}$ propagates forward and backward traveling waves from a point z' to a point z.

For a stack of N planar slabs on a substrate as in Figure 47.3, each with its own material properties, we demonstrate how the TMM can determine the bound modes. Let the top interface of the ith slab be at z_i, vector components with that label, e.g., $A_i^{\pm}$, be measured from there, i.e., just below the top interface in that

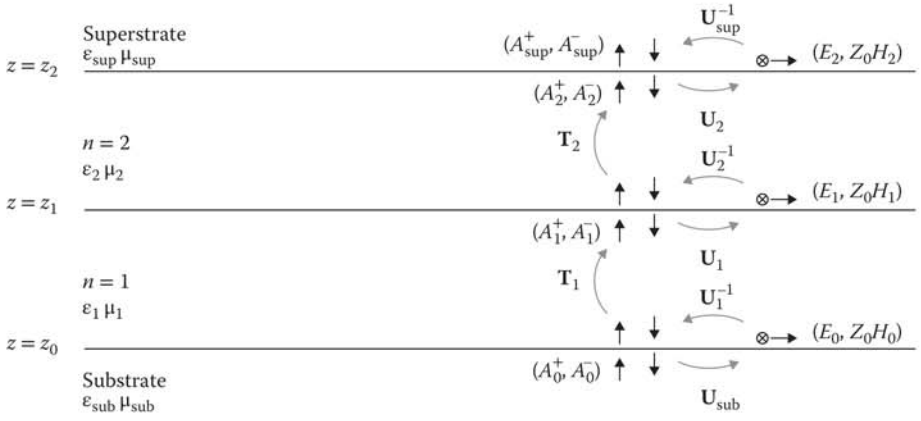


FIGURE 47.3 Schematic of the transfer matrix method. Fields (black arrows) are operated on by matrix transformations (gray arrows). The forward and backward components at the top of the substrate layer (A_0^+, A_0^-) are transferred through the stack to forward and backward components at the bottom of the superstrate layer (A_{sup}^+, A_{sup}^-). Fields move between bases of forward and backward waves ($\uparrow\downarrow$) and electric and magnetic components ($\otimes\rightarrow$).

layer. Starting in the forward and backward basis, one may use the transfer matrices to propagate through the stack from the one layer to the next:

$$\begin{pmatrix} A_{i+1}^{+} \\ A_{i+1}^{-} \end{pmatrix} = \mathbf{T}_{i+1}\mathbf{U}_{i+1}^{-1}\mathbf{U}_{i}\begin{pmatrix} A_{i}^{+} \\ A_{i}^{-} \end{pmatrix} . \tag{47.21}$$

The exception to this being the superstrate, the $(N+1)$th layer, which by definition does not have a top interface. As such, $E_{\text{sup}}^{\pm}$ is defined specially as the forward and backward components just above the top interface of the Nth layer, i.e.,

$$\begin{pmatrix} A_{\text{sup}}^{+} \\ A_{\text{sup}}^{-} \end{pmatrix} = \mathbf{U}_{\text{sup}}^{-1}\mathbf{U}_{N}\begin{pmatrix} A_{N}^{+} \\ A_{N}^{-} \end{pmatrix} . \tag{47.22}$$

The full transfer matrix $\mathcal{M}$ will map forward and backward traveling waves in the substrate to the superstrate:

$$\begin{pmatrix} A_{\text{sup}}^{+} \\ A_{\text{sup}}^{-} \end{pmatrix} = \mathcal{M}\begin{pmatrix} A_{\text{sub}}^{+} \\ A_{\text{sub}}^{-} \end{pmatrix} , \tag{47.23}$$

and is given as the product of each layer hop:

$$\mathcal{M} = \mathbf{U}_{\text{sup}}^{-1}\mathbf{U}_{N}\cdots\mathbf{T}_{1}\mathbf{U}_{1}^{-1}\mathbf{U}_{\text{sub}} . \tag{47.24}$$

Terms with similar indices can be grouped to give a matrix for each layer:

$$\mathbf{M}_{i} = \mathbf{U}_{i}\mathbf{T}_{i}\mathbf{U}_{i}^{-1} , \tag{47.25}$$

which maps electric and magnetic fields at the bottom of the layer through to the top. As such, the transfer matrix for the whole stack is

$$\mathcal{M} = \mathbf{U}_{\text{sup}}^{-1}\left(\prod_{i=N}^{1}\mathbf{M}_{i}\right)\mathbf{U}_{\text{sub}} , \tag{47.26}$$

noting the reverse order on the product.

The eigenmodes of the system that have non-zero fields for a zero driving field, i.e.,

$$\begin{pmatrix} A_{\text{sup}}^{+} \\ 0 \end{pmatrix} = \mathcal{M}\begin{pmatrix} 0 \\ A_{\text{sub}}^{-} \end{pmatrix} , \tag{47.27}$$

with the prescription that $\operatorname{Im} k_z > 0$ in both the superstrate and substrate such that fields on both sides decay away to infinity. This yields the condition that $\mathcal{M}_{22} = 0$ for bound modes. For dispersive media, this is satisfied for either one of the frequency ω or wavevector q being a complex number, with the other remaining real. In general, analytical dispersion relation solutions do not exist, but rather solutions for $\mathcal{M}_{22}(\mathbf{q}, \omega) = 0$ are found implicitly, using a complex root-finding algorithm.

To illustrate the method, let's consider the SPP solution. Consider a metal substrate with a planar interface to a dielectric layer. In this case, the transfer matrix is given as

$$\mathcal{M} = \mathbf{U}_{\text{sup}}^{-1}\mathbf{U}_{\text{sub}} \tag{47.28}$$

$$= \frac{1}{2}\begin{pmatrix} \frac{\varepsilon_{\text{bg}}\omega}{ck_{z,\text{air}}} & 1 \\ -\frac{\varepsilon_{\text{bg}}\omega}{ck_{z,\text{air}}} & 1 \end{pmatrix}\begin{pmatrix} \frac{ck_{z,\text{metal}}}{\varepsilon(\omega)\omega} & -\frac{ck_{z,\text{air}}}{\varepsilon(\omega)\omega} \\ 1 & 1 \end{pmatrix}, \tag{47.29}$$

which gives the SPP condition:

$$\mathcal{M}_{22} = \frac{1}{2}\left(\frac{\varepsilon_{\text{bg}}\omega}{ck_{z,\text{air}}}\frac{ck_{z,\text{metal}}}{\varepsilon(\omega)\omega} + 1\right) = 0, \tag{47.30}$$

$$\frac{\varepsilon_{\text{bg}}}{\varepsilon(\omega)} + \frac{k_{z,\text{air}}}{k_{z,\text{metal}}} = 0\,. \tag{47.31}$$

One can see that for this equation to have solutions, given $\varepsilon_{\text{bg}} > 0$ and $\operatorname{Im} k_z > 0$, it must hold that $\operatorname{Re}\varepsilon(\omega) < 0$, which is characteristic of metals for frequencies below their plasma frequency.

For a lossless Drude metal, i.e., $\varepsilon(\omega) = \varepsilon_\infty - (\omega_{\text{p}}/\omega)^2$, an analytical solution exists:

$$q^2 = \varepsilon_{\text{bg}}\left(\frac{\omega}{c}\right)^2 \frac{\varepsilon_\infty\omega^2 - (\varepsilon_{\text{bg}} + \varepsilon_\infty)\omega_{\text{sp}}^2}{(\varepsilon_{\text{bg}} + \varepsilon_\infty)(\omega^2 - \omega_{\text{sp}}^2)}\,, \tag{47.32}$$

where $\omega_{\text{sp}} = \omega_{\text{p}}/\sqrt{\varepsilon_{\text{bg}} + \varepsilon_\infty}$ is the surface plasmon frequency.

47.4 Evolutionary Algorithm

For a planar stratified structure whose properties can be determined by a TMM, one may wish to optimize some aspect of its response to electromagnetic radiation. This may be in the dispersion relation of bound modes or in the transmission and reflection spectra. For example, to prescribe the frequency of a bound mode, or a reflection peak; or set whether there is only a single mode or many modes within a frequency range; or control the amplitude and phase response to an incoming wavepacket. These properties are entirely determined by the composition of the structure; however, it is often not obvious how changes in the structure lead to changes in the response.

The structures to be considered can be described by a finite number of numeric or discrete parameters, i.e., for each layer: the thickness of the layer, the material model to use, and the parameters of that model. This is well suited to optimization using an *evolutionary algorithm* (EA) [12].

In brief, an EA operates by storing multiple copies of these parameter sets, perturbing the values of the parameters randomly, ordering the sets by how the system they describe performs against a fitness metric, and keeping only the best variants to the next round.

More fully, in the language of an EA, this would be to say that a *population* of *chromosomes* is kept, and on each iteration of the algorithm, new chromosomes are generated from *mutations* of the current ones; the next-generation population is filled by *selection* of these chromosomes by a *fitness function*. The goal of the EA is to find a chromosome that extremizes the fitness function globally, and as such an EA is most suitable when the evaluation of the fitness function can be performed relatively quickly as a large number of members of the state space must be evaluated and compared.

A chromosome is simply the set of parameters that describes a system, i.e., a list of numbers or discrete option values. They are atomic entities, independent of other chromosomes. Each chromosome may have its fitness function evaluated to determine how suitable its corresponding physical system is to a problem.

Each iteration of the EA holds a set of chromosomes, ordered by their fitness function, known as a *population*. The first-generation population is either seeded with randomly generated structures, or structures that have been tuned by hand, perhaps by approximative theoretical methods. Populations of subsequent generations are constructed by first taking a small quantile of the fittest chromosomes from the previous iteration, with the remaining places filled with *mutations* of chromosomes selected from the previous set. The progress of the algorithm can be tracked by examining the fittest chromosome at the end of each iteration. At the end of each iteration, a *terminating condition* is evaluated, which determines if the algorithm should stop, returning the fittest chromosome, e.g., if the best structure in the population is within a defined tolerance of the target output, or if a set number of iterations have elapsed. If the terminating condition is not met, then the algorithm continues for another iteration. The operation of the EA is described as a flowchart in Figure 47.4.

The mutation process generates a new chromosome by randomly perturbing the values of parameters of a parent chromosome. It is by mutation that the state space of chromosomes is mapped.

For the layered structures considered here, the mutation process will perturb both the properties of the layers themselves and the relationship between layers. Numeric parameters, such as layer thicknesses and those used in material models, are perturbed by a Gaussian random variable, and the discrete parameters, such as the layer model, can be randomly selected from a list. Typically, only one or few parameters will be changed in each mutation. The parameters of the layers themselves do not sit in isolation. The number of layers in these structures is variable and mutation allows for adding or removing of layers or indeed swapping the order of layers and shifting the boundary between adjacent layers.

Whereas the EA is widely applicable to problems in nano-optics, the fitness function is not general and has to be constructed for each problem based on the type and kind of characteristics that are to be selected for. As an example, the EA can be applied to the case of stopped-light lasing. Here structures are desired that support a bound mode with two points of zero group velocity (ZGV) in their dispersion. An SL structure is deemed better if the ratio of the separations of the wavevectors and frequencies of the ZGV points, the band index $|c/v_b| = (k_2 - k_1)/(\omega_2 - \omega_1)$, is higher. This ensures that a wide range of wavevectors can be held at a single frequency, which reduces the group velocity for the whole band and the wave dispersion that will occur for a wavepacket made from Fourier components in the band. Figure 47.5 shows an example of

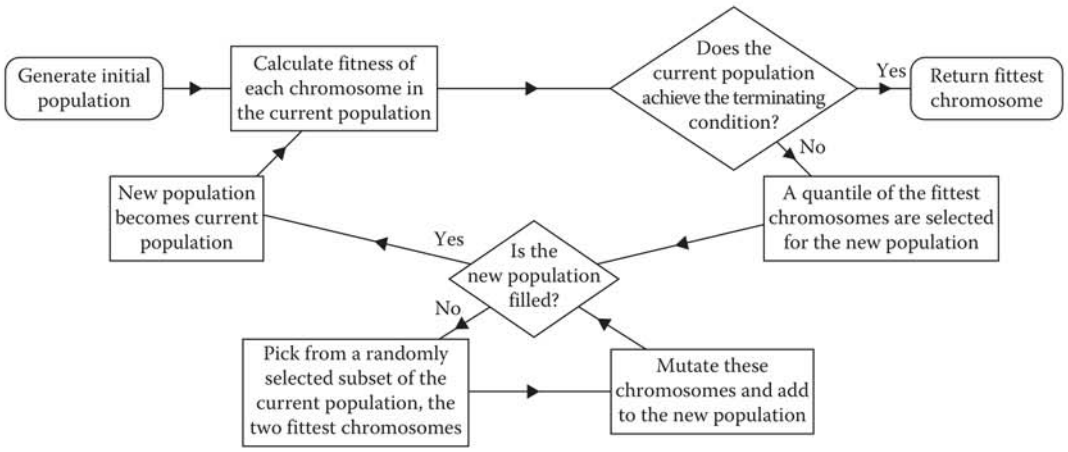


FIGURE 47.4 Evolutionary algorithm flowchart.

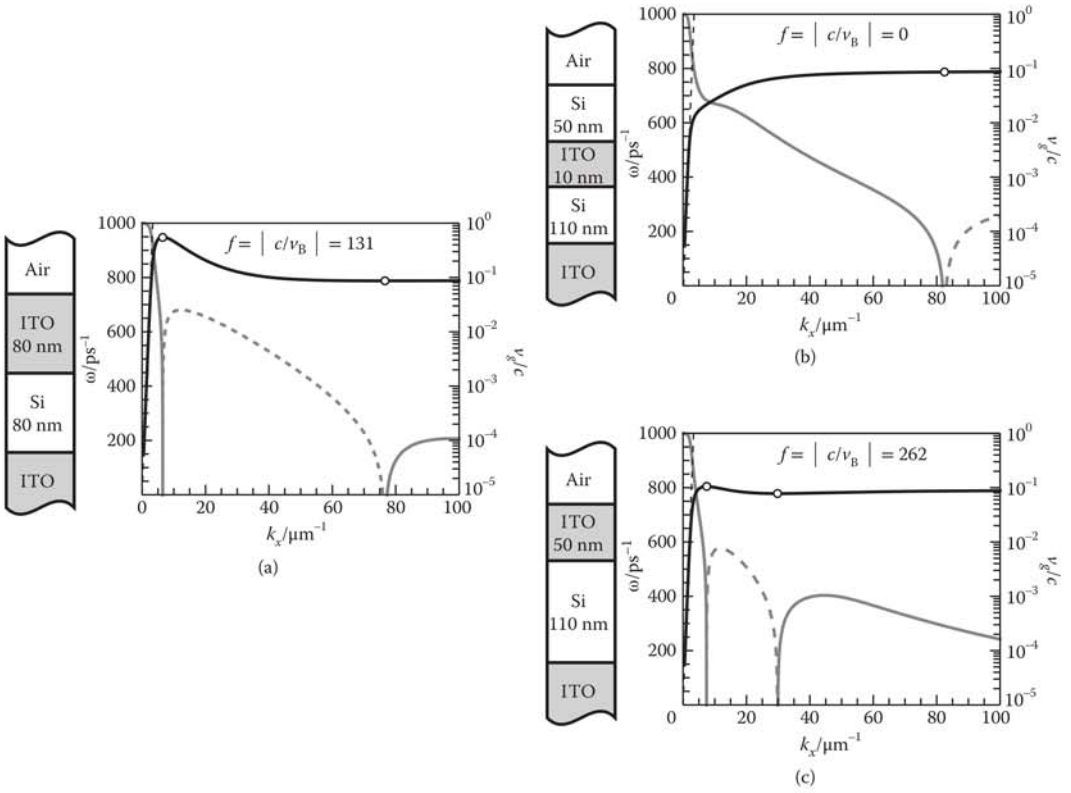


FIGURE 47.5 Diagrams of planar plasmonic nanostructures and corresponding dispersion relations (black solid curve). The group velocity is also shown (gray solid, positive; gray dashed, negative) as well as zero group velocity points (circles). An evolutionary algorithm, using *band index* as a metric, given a. as a starting chromosome may produce b. and c. as child chromosomes. b. has a metric of zero and is thrown out, whereas c. may stay to the next generation as it has a higher metric than its parent.

potential SL structures, their dispersion relations, and group velocity. Each structure is labeled by its band index, and for a parent structure, two possible child structures are presented.

47.5 Finite-Difference Time-Domain Method

When describing processes such as gain, a frequency-domain analysis can only bring insight to a certain level as it has limitations to describe nonlinearities. When plasmons or phonons are emitted, electrons are demoted from higher energetic states to lower ones. This depletes the population inversion, reducing the available gain over time, leading to a nonlinear dependence in the field strength. In addition, spatially resolved effects, such as spatial hole burning, cannot be considered in TMM that assumes uniformity in the direction of propagation, whereas the level of inversion can vary both in this direction and perpendicular to the stacking. Depending on the mode formed, some regions may host higher field densities that can deplete the local gain. Thus a dynamic, spatially resolved, time-domain simulation is required to capture all the aspects of a nanoplasmonic lasing device. Here we will discuss and use the *finite-difference time-domain* (FDTD) method [13,14].

In FDTD, the goal is to solve Maxwell's equations for successive instants in time on a discretized lattice. This is approached by focusing on the curl equations and making the time derivatives of the fields the

subject. Let us assume that there are no external sources:

$$\frac{\partial \mathbf{E}}{\partial t} = \frac{1}{\varepsilon_0}\left(\nabla \times \mathbf{H} - \frac{\partial \mathbf{P}}{\partial t}\right) \tag{47.33}$$

$$\frac{\partial \mathbf{H}}{\partial t} = -\frac{1}{\mu_0}\left(\nabla \times \mathbf{E} - \frac{\partial \mathbf{M}}{\partial t}\right), \tag{47.34}$$

where **P** and **M** are the polarization and magnetization, respectively. Now, space and time are discretized as $(t, x, y, z) \to (n\Delta_t, i\Delta_x, j\Delta_x, k\Delta_x)$ and derivatives are turned into finite differences, i.e.,

$$\frac{\partial \mathbf{E}}{\partial t}(\mathbf{x}, t) \to \frac{\mathbf{E}^{n+1}_{i,j,k} - \mathbf{E}^{n}_{i,j,k}}{\Delta_t}. \tag{47.35}$$

Strictly, the above is a central difference and approximates the time derivative at a time $(n+1/2)\Delta_t$. Therefore, the corresponding right-hand side terms must be evaluated at this time. The curl operator is handled similarly:

$$\nabla \times \mathbf{E}(\mathbf{x}, t)|_x \to \frac{E_z|^n_{i,j+1,k} - E_z|^n_{i,j,k}}{\Delta_x} - \frac{E_y|^n_{i,j,k+1} - E_y|^n_{i,j,k}}{\Delta_x}. \tag{47.36}$$

In order to make best use of this central difference, the lattice is arranged as a Yee grid [13], where fields are stored separately by component and **E** and **P** fields are specified on integer time coordinates, whereas **H** and **M** are specified on half-integer time coordinates. The lattice coordinates are more complicated; for **E**, components are stored at half-integer coordinates in their own direction and at integer coordinates otherwise. This is the opposite for the **H** field, as illustrated in Figure 47.6. From here, one can write the *update equations*:

$$\begin{aligned} E_x|^{n+1}_{i+\frac{1}{2},j,k} &= E_x|^{n}_{i+\frac{1}{2},j,k} - \varepsilon_0^{-1} P_x|^{n+1}_{i+\frac{1}{2},j,k} + \varepsilon_0^{-1} P_x|^{n}_{i+\frac{1}{2},j,k} \\ &+ \frac{c\Delta_t}{\Delta_x}\left(Z_0 H_z|^{n+\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2},k} - Z_0 H_z|^{n+\frac{1}{2}}_{i+\frac{1}{2},j-\frac{1}{2},k}\right) \\ &- \frac{c\Delta_t}{\Delta_x}\left(Z_0 H_y|^{n+\frac{1}{2}}_{i+\frac{1}{2},j,k+\frac{1}{2}} - Z_0 H_y|^{n+\frac{1}{2}}_{i+\frac{1}{2},j,k-\frac{1}{2}}\right) \end{aligned} \tag{47.37}$$

$$\begin{aligned} Z_0 H_x|^{n+\frac{1}{2}}_{i,j+\frac{1}{2},k+\frac{1}{2}} &= Z_0 H_x|^{n-\frac{1}{2}}_{i,j+\frac{1}{2},k+\frac{1}{2}} + cM_x|^{n+\frac{1}{2}}_{i,j+\frac{1}{2},k+\frac{1}{2}} - cM_x|^{n-\frac{1}{2}}_{i,j+\frac{1}{2},k+\frac{1}{2}} \\ &+ \frac{c\Delta_t}{\Delta_x}\left(E_z|^{n}_{i,j+1,k+\frac{1}{2}} - E_z|^{n}_{i,j,k+\frac{1}{2}}\right) \\ &- \frac{c\Delta_t}{\Delta_x}\left(E_y|^{n}_{i,j+\frac{1}{2},k+1} - E_y|^{n}_{i,j+\frac{1}{2},k}\right), \end{aligned} \tag{47.38}$$

with corresponding equations for the omitted components.

With the updated equations of the primary fields determined, it now remains to be described how the auxiliary fields are updated. In general, the constitutive equations are functionals of all past values of all fields, though this would prove tricky to implement directly. In practice, material responses can be cast in terms of a finite difference if they can be expressed in terms of a temporal differential equation. This lends the Drude–Lorentz model well to inclusion within FDTD by converting its auxiliary equations into finite differences.

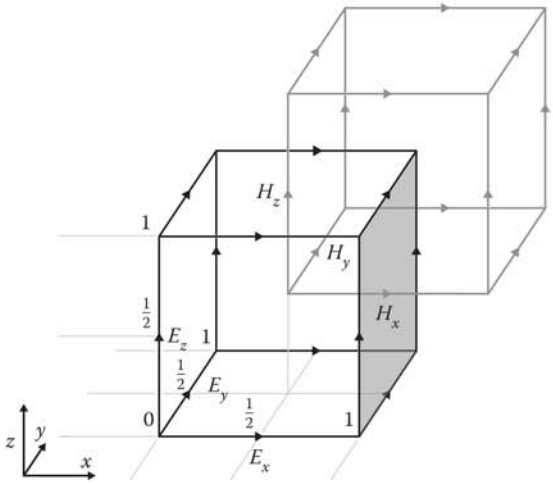


FIGURE 47.6 Schematic of FDTD Yee cell. Electric (black) and magnetic (gray) field grids. They are shifted half a step from each other, forming an interlocking link. Each field line passes through the center of a square formed from lines of the other field, e.g., H_x here passes through the center of the highlighted E-field face, where the finite difference $\nabla \times \mathbf{E}$ is defined.

The FDTD algorithm proceeds by iterating over, and updating in order, the **P** field, **E** field, **M** field, and **H** field over all space for each timestep.

Of course, there are limitations to FDTD. Its staggered grid structure means that material interfaces are not sharply defined, and any curvature on an interface will also experience a staircasing effect that can lead to spurious hotspots especially in plasmonic structures where geometric effects play a key role. The grid is also regular, which is to say it cannot be adaptively refined in areas of interest, rather all points must be simulated to the same resolution. This adds to the numerical load, as the smallest wavelength component of the fields, as well as the smallest geometrical features must be provided with sufficient resolution. This is again especially relevant in plasmonics, where field confinement is a feature.

Because of its nature as a full-wave time-domain solver, FDTD can often be the final say on a theoretical investigation. Short of doing a real-world experiment, FDTD numerical simulations find their utility as a heavy-duty tool. That is, the Maxwell's equations are solved at every discretized point in space, so it is often slower and more cumbersome, but can as a result of the inherent thoroughness produce results for questions that other techniques are unable to.

47.6 Loss Compensation

The high field enhancements afforded by plasmon modes supported by negative permittivity in metals must be traded off against Ohmic losses in the optical frequencies. Loss compensation is the incorporation of gain elements into a system in order to offset the loss in a particular frequency range while retaining desired properties from the lossy structure.

The properties one might wish to retain, e.g., negative effective index [6], stopped light points [11], or hyperbolic dispersion [15], etc., are frequency dependent. Causality will influence the position and width of the frequency range where losses can be compensated and desirable properties are to be found simultaneously [16,17], and adding gain to a system will necessarily change the refractive properties in accordance with the Kramers–Kronig relations.

The induced change in the mode structure by the addition of gain can be analyzed in a frequency-domain analysis, modeling gain as a four-level system in frequency domain with constant electron occupation

densities. Using the TMM, this would mean replacing a part or whole of a dielectric layer with an equivalent Drude–Lorentz model with properties that map to the dielectric layer for equilibrium carrier occupation densities.

To illustrate this, the SL structure as optimized previously is modified by replacing the dielectric layer with a quantum emitter. The emission frequency is set to match either one of the ZGV points, and the other parameters are set representing the inclusion of realistic laser dye molecules [18]. The resulting dispersion and loss relations are shown in Figure 47.7 for inversion densities ΔN varying between 0 and N, with a fixed emitter density, and excitation about ZGV1 and ZGV2. The first point to note is that, even on full inversion, the addition of gain does not significantly change the dispersion, with the maximum shift in frequency being around 0.6%. The presence of ZGV points is preserved, though they may drift slightly, i.e., ZGV2 moves slightly right with emission about ZGV1. There is no change at the frequency that is being excited because the permittivity change of a Lorentzian is zero at the resonant frequency. Adding gain has the side effect of making the structure a slightly better SL structure as the band velocity is marginally reduced.

The key change though is in the loss. As the inversion density increases, the loss decreases for wavevectors where the dispersion is within the gain width. Initially the plasmonic loss is reduced as the inversion increases, and then for an inversion of around $\Delta N/N \approx 0.4$, some wavevectors become undamped, and even eventually experience gain. Plasmons sitting in these modes will grow exponentially in amplitude while small enough to remain in the small-signal gain regime. As ZGV2 is flatter than ZGV1, when the

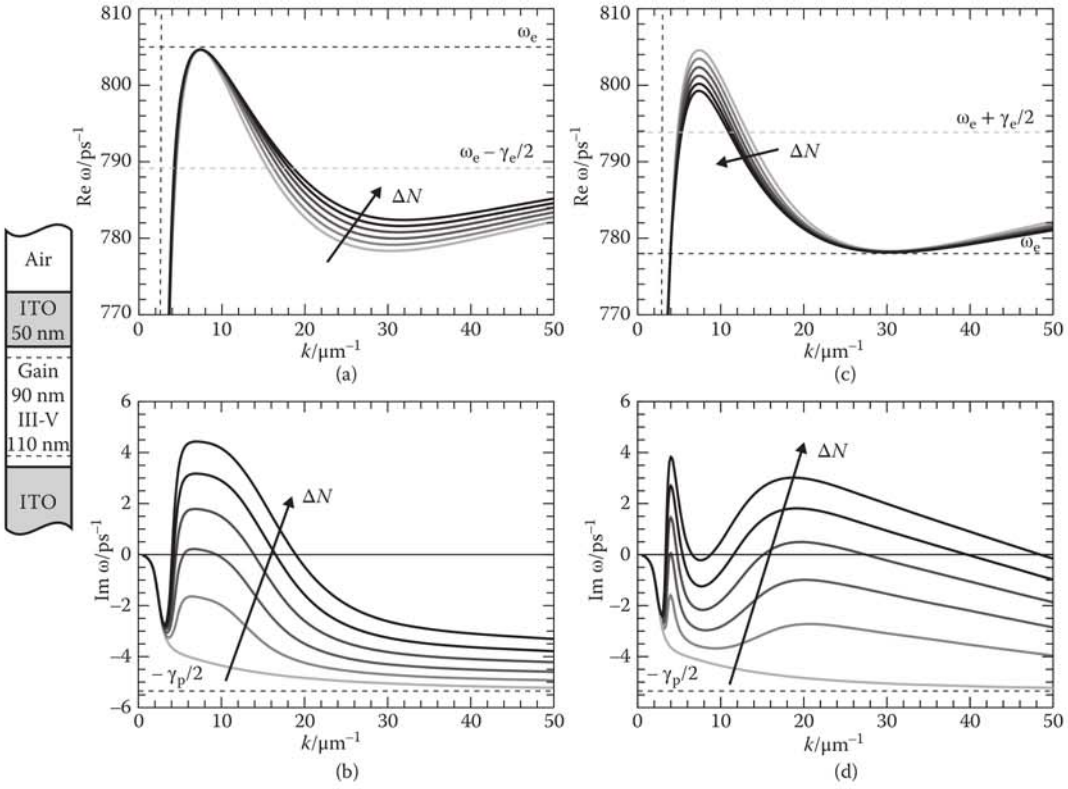


FIGURE 47.7 Perturbation of dispersion relation and loss with Lorentzian gain. The change of mode shape plotted as Lorentzian gain is introduced to the system. For subfigures on the left, emission is about ZGV1, and on the right about ZGV2. a and b show a zoom in of how the dispersion relation changes with increasing inversion density, while c and d show the corresponding loss/gain of the mode.

emission is about this point, a wider range of frequencies fall within the gain width, leading to a larger range of k values that can become undamped.

An analysis of loss compensation can also be performed in the time domain. This has been done in Reference [6] where the structure was a negative refractive index double-fishnet metamaterial. Pump-probe spectroscopy was performed in FDTD in order to retrieve effective parameters using realistic parameters of rhodamine 800 dye molecules as a four-level system. It was shown that the imaginary part of the effective refractive index, which encodes the loss, transitioned from positive to zero as pumping increased while retaining a negative real part of the effective index between wavelengths of 706 and 716 nm. The results are shown in Figure 47.8.

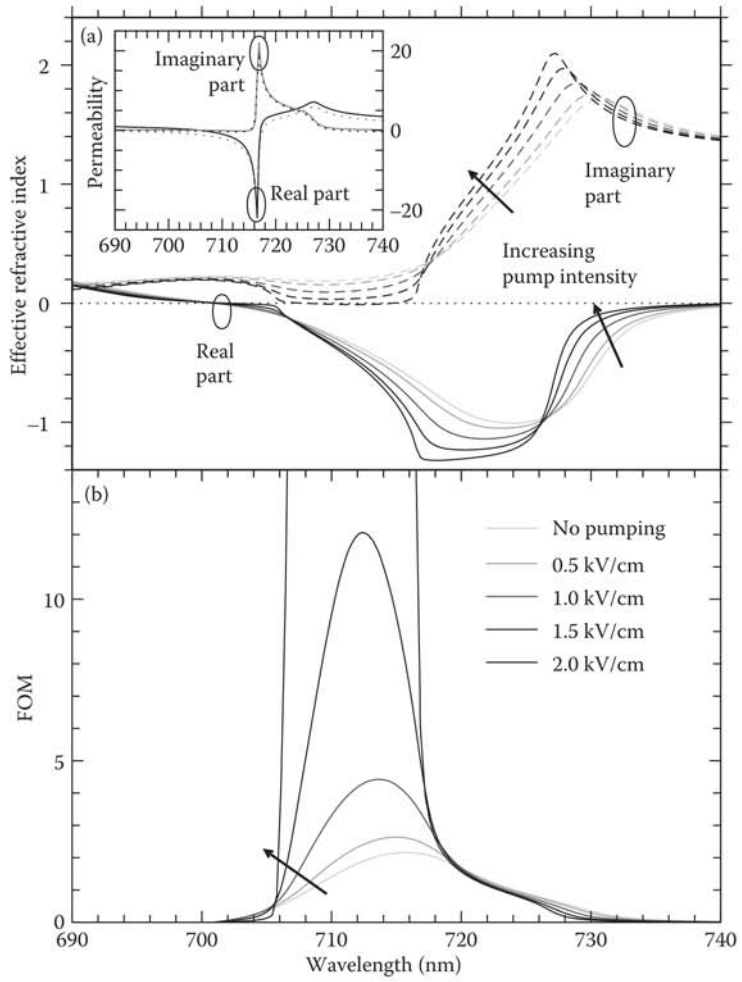


FIGURE 47.8 (a) Real and imaginary part of the retrieved effective refractive indices of the double-fishnet structure for different pump amplitudes. The peak electric field amplitude of the pump increases in steps of 0.5 kV/cm from no pumping (lightest) to a maximum of 2.0 kV/cm (darkest). The inset shows the real and imaginary part of the effective permeability (black and gray line, respectively) and the result of the Kramers–Kronig relation (dotted lines) for the highest peak electric field amplitude of 2.0 kV/cm . (b) The figures of merit (FOM) for the same pumping amplitudes. (Reproduced from Sebastian Wuestner et al., *Phys. Rev. Lett.*, 105,127401, 2010.)

47.7 Lasing Dynamics

In order to transition from loss compensation to lasing, photons emitted by the gain medium must feedback and stimulate on average more emission than photons that are lost on the round trip. Photons can be lost by escaping the cavity to radiation modes, absorbed incoherently by dissipative channels, or absorption by the gain medium.

Initially, for a cold cavity with no lasing mode field built up, an inverted gain system may spontaneously emit into cavity modes. This spontaneous emission may trigger further stimulated emission in a process known as *amplified spontaneous emission* (ASE), which can be added to the semiclassical FDTD description by including spatially resolved dynamic Langevin noise to the system, which accounts for the dissipative reservoirs feeding back stochastically on the system. The noise couples to the four-level system and induces incoherent transitions, which then become amplified, allowing for the triggering of the lasing regime [8].

As a device enters a lasing regime, characteristic relaxation oscillations can be seen where inversion initially builds up and spontaneous emission events are induced. The fields then are amplified as they stimulate further emission, growing the field in the mode coherently. From here on the emitted fields start to grow exponentially, and when they are of sufficient strength will deplete the inversion density. This reduction of inversion feeds back by decreasing the available gain, leading to a decrease in field energy as the energy in the field is lost to dissipative processes in the metal layers. The decrease in energy density allows for the inversion to rebuild. This continues in an oscillatory manner with the energy density lagging behind the inversion by 90°. The amplitude of the inversion and energy oscillations decrease with each cycle until a stable steady value for both is reached. In contrast, the control structure, albeit entering a regime of relaxation oscillations, is more erratic and the oscillations do not settle to a steady state. Instead, the oscillations continue with a factor of 4 between the peak energy density and the trough. An example of these relaxation oscillations in shown for the SL laser set in Figure 47.9 IN contrast to a "control structure" where the waveguide is not of high enough quality for the device to enter the lasing regime. Relaxation oscillations are accompanied by a characteristic decrease of the width of the power spectrum, from a wide peak in ASE to a sharp one in the lasing regime. This observation is usually taken as a crucial evidence for lasing in experimental setups.

47.8 Stopped-Light Lasing

Stopped-light lasing differs from traditional nanoplasmonic lasers in that there is no cavity with predefined modes. Whereas in the structures presented, light is trapped in a waveguide mode, it is free to propagate in one direction, albeit at ZGV. This means that there is a continuum of planewave modes for the gain to emit into, which get selected on the basis of the overlap with the gain media. The feedback in the system can be understood by considering the spatially resolved, cycle-averaged Poynting vector, energy density, and inversion that are plotted in Figure 47.10 for a structure with a gain width of 1000 nm. The energy is concentrated on metal-dielectric interfaces, strongest on the lower interface, and is localized around the gain medium despite there being no cavity along the horizontal direction.

The Poynting vector shows how energy circulates in the structure. There are four energy flux vortices, one in each corner of the structure; as described in the figure caption, these have energy moving out of the gain region into the metal layers and into it in dielectric layers. Considering the x component of the Poynting flux Figure 47.10a, the forward and backward flows are in exact balance. The balanced counter-propagation of energy in the negative-permittivity (metal) layers against the dielectric layer is the basis of the (subwavelength) feedback in the system.

The inversion is shown in Figure 47.10d; there are areas of spatial hole burning where the energy density is highest. The spatial modulation seen is explained when considering the field profile and its formation.

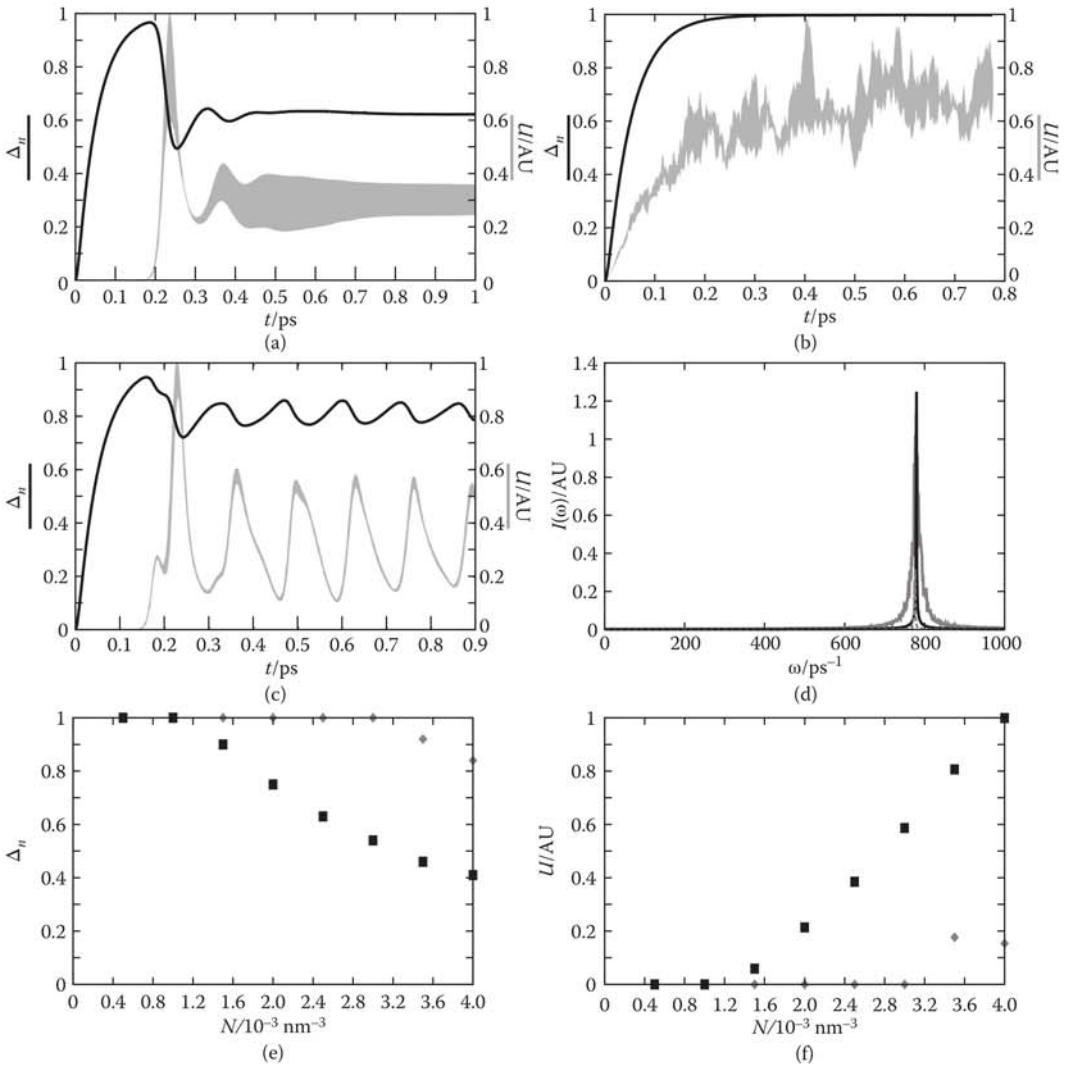


FIGURE 47.9 Lasing onset with increasing gain density. a, b, c. Energy density and mean inversion for (a) the stopped light structure with a gain density $N = 0.002\,\text{nm}^{-3}$, (b) the stopped light structure with a gain density $N = 0.001\,\text{nm}^{-3}$, and (c) the control structure with a gain density $N = 0.004\,\text{nm}^{-3}$. (d) Power spectra for SL structure at $N = 0.002\,\text{nm}^{-3}$ (black), SL structure at $N = 0.001\,\text{nm}^{-3}$ (×5000 dashed gray), and control structure at $N = 0.004\,\text{nm}^{-3}$ (solid gray). (e) Steady-state inversion of SL (black) and control (gray) structures with varying gain density. (f) Steady-state cycle-averaged energy density with varying gain density.

47.9 Summary

In this chapter, we have explored techniques to design, optimize, and simulate a nanoplasmonic laser. We started by defining harmonic oscillator type Drude–Lorentz and Maxwell–Bloch models, a universal type of models for metals, semiconductors, and gain media. From a simple model of forces in time domain, a causal frequency-domain response function was derived. From here, the four-level system was introduced as two dipole transitions in the frequency domain and as a dynamic system with self-consistently evolving

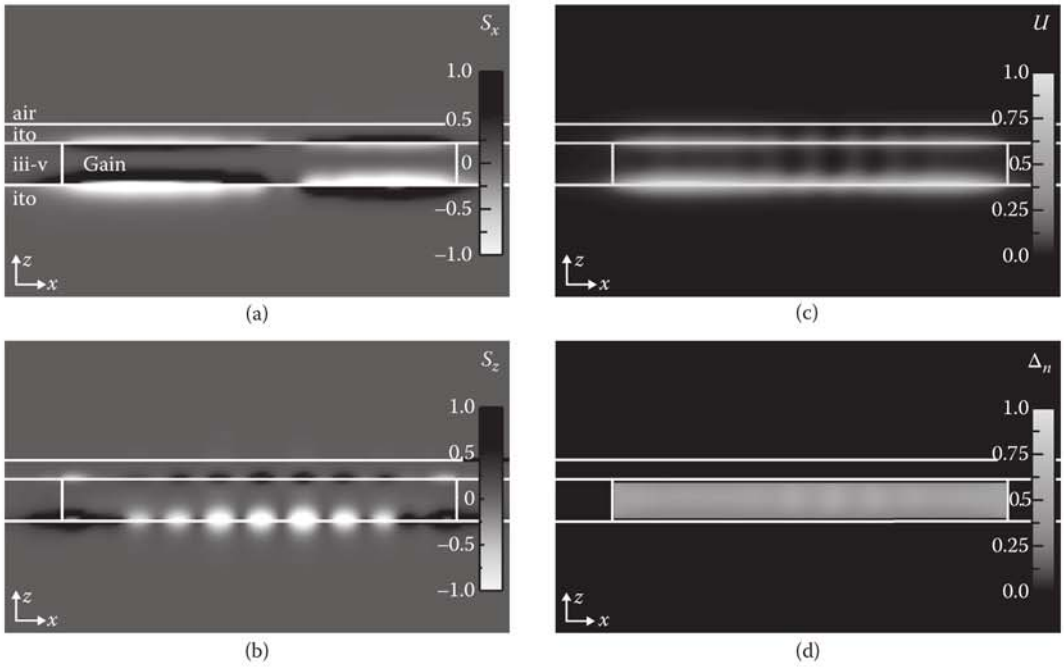


FIGURE 47.10 Energy flux and density of the lasing mode. The lasing mode of a stopped-light laser structure with a width 1000 nm. (a,b) The x and z components of the cycle-averaged Poynting vector, respectively, with positive values in black and negative in white. Focusing on the bottom right lobe of the flux in panel a, energy flows into the gain region in the dielectric layers and out of it in the metal layer. For the same lobe in panel b, energy moves downward out of the gain region toward the center, and upward into the region at the edge. The counter-clockwise energy vortex in this corner is mirrored in the other corners. (c) The cycle-averaged energy density and (d) inversion plots. The inversion is the complement of the energy density, being highest in areas with low energy density and depleted where the mode sits.

electron occupation levels in the time domain. The TMM was next introduced as an exact frequency-domain tool to analyze planar structures. EA are suited to optimize systems that can be characterized by a range of numeric or discrete parameters and perform characterized by a fitness metric. The EA approach was introduced to and demonstrated in the case of the TMM and applied to designing stopped light laser structures. Next, FDTD, the time-domain numerical method, was discussed and used to simulate active hybrid metallo-dielectric structures with nonlinearities and time-dependant dispersion such as a lasing system. This allowed for the simulation of lasing dynamics where loss compensation, ASE, and relaxation oscillations could be captured. Finally, some specifics of the recently established stopped light laser principle were explored as an example system of a designed and optimized nanoplasmonic laser.

References

1. Jennifer A. Dionne, Luke A. Sweatlock, Matthew T. Sheldon, A. Paul Alivisatos, and Harry A. Atwater. Silicon-based plasmonics for on-chip photonics. *IEEE Journal of Selected Topics in Quantum Electronics*, 16(1):295–306, 2010.
2. Anatoly V. Zayats and Igor I. Smolyaninov. Near-field photonics: Surface plasmon polaritons and localized surface plasmons. *Journal of Optics A: Pure and Applied Optics*, 5(4):S16, 2003.
3. Dmitri K. Gramotnev and Sergey I. Bozhevolnyi. Plasmonics beyond the diffraction limit. *Nature Photonics*, 4(2):83–91, 2010.

4. Kevin F. MacDonald, Zsolt L. Samson, Mark I. Stockman, and Nikolay I. Zheludev. Ultrafast active plasmonics. *Nature Photonics*, 3(1):55–58, 2009.
5. Ortwin Hess, John B. Pendry, Stefan A. Maier, Rupert Oulton, Joachim M. Hamm, and Kosmas L. Tsakmakidis. Active nanoplasmonic metamaterials. *Nature Materials*, 11(7):573–584, 2012.
6. Sebastian Wuestner, Andreas Pusch, Kosmas L. Tsakmakidis, Joachim M. Hamm, and Ortwin Hess. Overcoming losses with gain in a negative refractive index meta-material. *Physical Review Letters*, 105(12):127401, 2010.
7. David J. Bergman and Mark I. Stockman. Surface plasmon amplification by stimulated emission of radiation: Quantum generation of coherent surface plasmons in nanosystems. *Physical Review Letters*, 90:27402, 2003.
8. Andreas Pusch, Sebastian Wuestner, Joachim M. Hamm, Kosmas L. Tsakmakidis, and Ortwin Hess. Coherent amplification and noise in gain-enhanced nanoplasmonic metamaterials: A Maxwell-Bloch Langevin approach. *ACS Nano*, 6(3):2420–2431, 2012.
9. Sebastian Wuestner, Joachim M. Hamm, Andreas Pusch, Fabian Renn, Kosmas L. Tsakmakidis, and Ortwin Hess. Control and dynamic competition of bright and dark lasing states in active nanoplasmonic metamaterials. *Physical Review B*, 85(20):201406, 2012.
10. Bumki Min, Eric Ostby, Volker Sorger, Erick Ulin-Avila, Lan Yang, Xiang Zhang, and Kerry Vahala. High-Q surface-plasmon-polariton whispering-gallery microcavity. *Nature*, 457(7228):455–458, 2009.
11. Tim W. Pickering, Joachim M. Hamm, A. Freddie Page, Sebastian Wuestner, and Ortwin Hess. Cavity-free plasmonic nanolasing enabled by dispersionless stopped light. *Nature Communications*, 5:4972, 2014.
12. William M. Spears. *Evolutionary Algorithms: The Role of Mutation and Recombination. Natural Computing Series*. Heidelberg: Springer Berlin Heidelberg, 2000.
13. Kane Yee. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14:302–307, 1966.
14. Allen Taflove. *Computational Electrodynamics: The Finite-Difference Time-Domain Method. Antennas and Propagation Library*. Norwood, MA: Artech House, 1995.
15. Xingjie Ni, Satoshi Ishii, Mark D. Thoreson, Vladimir M. Shalaev, Seunghoon Han, Sangyoon Lee, and Alexander V. Kildishev. Loss-compensated and active hyperbolic metamaterials. *Optics Express*, 19(25):25242–25254, 2011.
16. Johannes Skaar. Fresnel equations and the refractive index of active media. *Physical Review E*, 73(2):26605, 2006.
17. Paul Kinsler and Martin McCall. Causality-based criteria for a negative refractive index must be used with care. *Physical Review Letters*, 101(16):1–4, 2008.
18. P. Sperber, W. Spangler, Bernd Meier, and Alfons Penzkofer. Experimental and theoretical investigation of tunable picosecond pulse generation in longitudinally pumped dye laser generators and amplifiers. *Optical and Quantum Electronics*, 20(5):395–431, 1988.

48

Quantum-Dot Nanolasers

Christopher Gies
Michael Lorke
Frank Jahnke
and
Weng W. Chow

48.1 Going Nano: A New Era of Small Lasers 627
48.2 Connecting to Earlier Work: Laser Rate Equations and the β Factor 628
48.3 Key Aspects of Nanolasers 631
48.4 Essential Physics I: Quantum-Optical Semiconductor Laser Theory 635
Dynamical Laser Equations • Pump Process for Continuous-Wave-Excited Lasers
48.5 Essential Physics II: Carrier Kinetics in QD Systems 639
Carrier–Phonon Interaction • Coulomb Interaction
48.6 Connecting to Experiments: Transition to Lasing and the β Factor 646
48.7 Interemitter Coupling Effects and Superradiance in Nanolasers 647
48.8 Nonresonant Coupling and Lasing from Multiexciton States in Few-QD Systems 651
48.9 Outlook 653

48.1 Going Nano: A New Era of Small Lasers

The impact of semiconductor lasers on our daily lives is immense. They perform work every time someone gathers information for the internet, makes a telephone call, prints an article or pays for an item at a store. There are over 2.5 billion VCSELs† in use at our homes or offices, and roughly 100 million are produced each year to keep up with demand. VCSEL technology is now considered conventional, and research has progressed to smaller devices, modulated at higher speeds and producing output with greater spectral stability. This chapter focuses on two of the advances. One is the reduction of optical cavity volume by as much as two orders of magnitude, making a transition from microlasers (such as VCSELs) to nanolasers. Two is a change in active medium from quantum wells (QWs) to quantum dots (QDs), bringing about, e.g., devices operating with very few (tens) of emitters. These developments ushered in a new era for semiconductor device physics, one where quantum optical and many-body electron interaction effects dominate. The underlying motivation is the control of spontaneous emission [1]. The typical laser mitigates the randomness (noise) caused by spontaneous emission by overwhelming it with stimulated emission. Here, we are speaking of actually quieting the spontaneous emission noise, both spatially and temporally. In the former, we use nanocavities to inhibit spontaneous emission in undesirable directions. With the latter, we use very

† Vertical cavity surface emitting lasers, cf. Chapter 34.

few QDs (ideally only one) to control the timing of photon emission, thereby improving photon statistics beyond the limit described by the Poisson distribution. Devices incorporating the two advances are being fabricated and experiments are being performed. The results are both promising and exciting from the device and physics aspects, respectively. They also lead to many new questions and renewed interest in some old ones: What is lasing and where is the threshold? Why is there not a phonon bottleneck? What is the homogeneous width of a QD transition [2–4]? What is the inhomogeneous broadening in my samples? Is there really thresholdless lasing [5–7]? Are QD lasers better than QW lasers [8]? Why do QD lasers not show modulation speeds as advertised? This chapter describes a theoretical framework capable of addressing all the abovementioned questions. The building blocks come from quantum electrodynamics (QED), which is the quantum theory of the interaction of light with matter, and carrier interactions described by many-body theory.

The fundamental concepts of lasing are often discussed in terms of rate equations for material excitations and cavity photons that may well be familiar to the reader. We begin this chapter with an excursion to show profit and limitations of laser rate equations for understanding the behavior of nanolasers, and at the same time establish a connection to more sophisticated semiconductor laser models. In Section 48.3, we give a short overview of the key aspects in modeling nanolasers and the characterization of their emission properties. In Section 48.4, we introduce a microscopic laser model for solid-state emitters that goes beyond the rate equations and that provides access to the quantum-statistical properties of the emission. A central aspect of device performance is determined by the efficiency of carrier scattering processes following excitation. Section 48.5 provides an overview of the underlying mechanisms of carrier–carrier and carrier-phonon interaction and explains methods to incorporate the dynamics into semiconductor laser models within different levels of approximation. The underlying formalism allows one to address a wide variety of effects present in semiconductor nanolasers, and we provide a few examples in Sections 48.6 and 48.7, where QD lasing and radiative coupling between individual emitters are discussed. Section 48.8 gives an outlook toward systems with only a few QDs inside the cavity, where the individual electronic many-particle properties of each emitter matter, and where device properties are dominated by quantum and correlation effects.

48.2 Connecting to Earlier Work: Laser Rate Equations and the β Factor

Lasing operation takes place when pumping of the active material provides sufficient gain to compensate the losses. It is thus determined by the interplay of the excitation and emission dynamics of the gain material, taking into account the feedback provided by the radiation field inside the resonator. A basic understanding of lasing action can be obtained from laser rate equations in the form of two coupled equations for the gain (in this case, the number of excitons N in the QD ensemble) and the intracavity photon number n:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} N &= P - \gamma_l (n+1) N - \gamma_{\mathrm{nl}} N, \\ \frac{\mathrm{d}}{\mathrm{d}t} n &= -\kappa n + \gamma_l (n+1) N. \end{aligned} \tag{48.1}$$

Here, P and κ are the pump and photon loss rates, γ_l the spontaneous emission rate into the laser mode, γ_{nl} the rate of radiative recombination into other modes, and $\gamma_{\mathrm{sp}} = \gamma_l + \gamma_{\mathrm{nl}}$ the total spontaneous emission rate. The bracket $(n+1)$ contains both the stimulated and spontaneous emission contributions. The ratio

$$\beta = \frac{\gamma_l}{\gamma_{\mathrm{sp}}} \tag{48.2}$$

defines the β-factor illustrated in Figure 48.1. It is a key characteristic quantity of any laser device as it defines device performance in the spontaneous-emission regime (i.e., below threshold) and, more importantly, the threshold current. The rate equations (Equation 48.1) are often expressed in terms of β:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} N &= P - \beta\gamma_{\mathrm{sp}}\,(n+1)N - (1-\beta)\gamma_{\mathrm{sp}}\,N, \\ \frac{\mathrm{d}}{\mathrm{d}t} n &= -\kappa n + \beta\gamma_{\mathrm{sp}}\,(n+1)N. \end{aligned} \tag{48.3}$$

In conventional lasers, $\beta \approx 10^{-5}$ [9], so that the majority of spontaneously emitted photons are lost. Once the system starts lasing at sufficiently high excitation, stimulated emission completely overwhelms spontaneous emission, so that the fraction of photons lost into nonlasing modes becomes insignificant. This behavior can be directly traced by accessing both emission channels separately. In micropillar lasers, the fundamental mode often used as laser mode has a far-field emission in vertical direction of the pillar, whereas losses occur laterally through the pillar side walls. The input–output curve (photon number versus pump rate) collected separately from both directions, together with a schematic of the experiment, is shown in Figure 48.2 [10]. As long as spontaneous emission is the dominating emission process, both emission intensities increase linearly with pump with a fixed ratio determined by the β factor. At higher excitation powers, stimulated emission from the laser mode sets in and photon emission from the laser mode increases nonlinearly and more strongly than the emission into loss channels. The resulting increase of the normalized ratio r of both emission channels therefore gives direct account of the onset of stimulated emission. Note that the S shape of the intensity curve for the lateral emission is due to stray light from the laser mode that is collected in lateral direction. Ideally, a clamped intensity is expected.

When continuous excitation is considered, the rate equations are solved in the steady state. Input–output curves obtained from the rate equations (Equations 48.1) for different values of β are shown in Figure 48.3. For $\beta \ll 1$, a sudden increase in emission intensity is visible that allows one to identify the lasing threshold. With increasing coupling efficiency β, the threshold shifts to lower pump rates, offering the prospect of devices with reduced threshold currents. At the same time, the intensity jump decreases until the "thresholdless" case is obtained for $\beta = 1$. In this regime an identification of the transition into lasing is impossible from the emission intensity alone. A characterization of the threshold requires going beyond the rate equation approximation to study statistical and coherence properties of the emission. One of the attractiveness of nanolasers is that they are capable of operating in a regime close to $\beta = 1$ due to their small effective mode volumes and large cavity lifetimes that cause strong funneling of spontaneously emitted photons into a single resonator mode [1]. At the same time, the photonic density of states is reduced away from the cavity-mode frequency, which further suppresses emission into nonlasing modes and thereby also

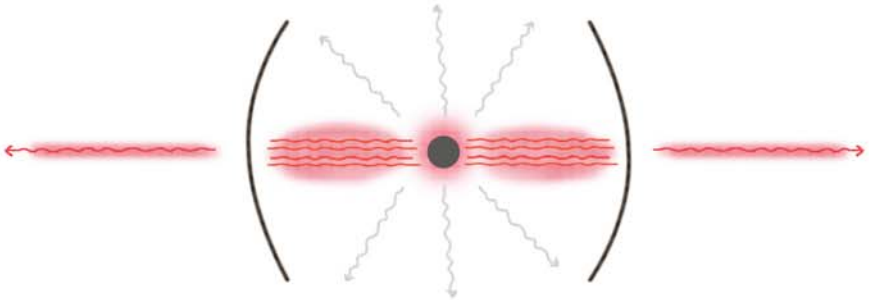


FIGURE 48.1 Spontaneous emission of the emitter into the laser mode (black lines) and into nonlasing modes (gray lines). The ratio of the spontaneous emission into the laser mode and the total spontaneous emission is described by the β factor.

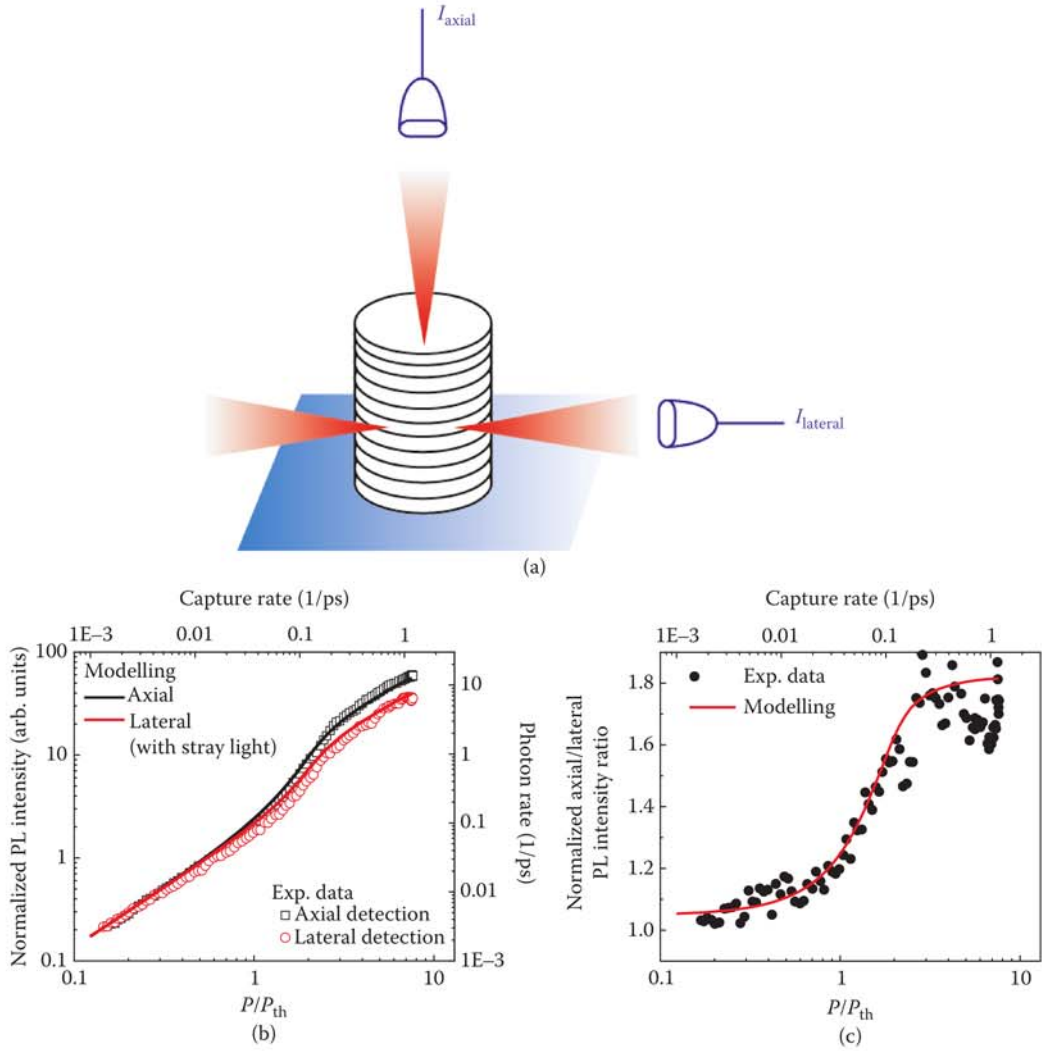


FIGURE 48.2 (a) Sketch of an experiment for detecting laser-mode emission in axial direction, and leaky mode emission in lateral direction separately. (b) Input–output characteristics of a quantum dot (QD) micropillar laser. Emission intensity into the laser mode (axial direction) and into nonlasing modes (lateral direction) are shown separately. (c) Their ratio r increases at the onset of stimulated emission, providing a direct visualization of the transition into the lasing regime. The results have been obtained from the semiconductor laser theory introduced in Section 48.4. Note that in lateral emission some scattered light from the axial emission is collected, which is responsible for the unclamped intensity behavior. [PL, photoluminescence. (Figures b and c adapted from Musiał, A. et al., *Physical Review B*, 91, 205310, 2015.)]

increases the β factor. While cavities operate at a specific narrow frequency window, structures that cause a broadband enhancement of spontaneous emission without using a cavity, such as photonic trumpets [11], are based on the very same concept.

The rate equations (Equation 48.1) have further drawbacks for modeling QD-based nanolasers. Most importantly, they do not provide information on laser field coherence and photon correlations. The rate equations are obtained by neglecting all correlations between light and matter degrees of freedom. Due to the enhanced interaction inside the microresonator, photon, and carrier-photon correlations have been demonstrated to play an important role in nanolasers, and this is where new approaches are required

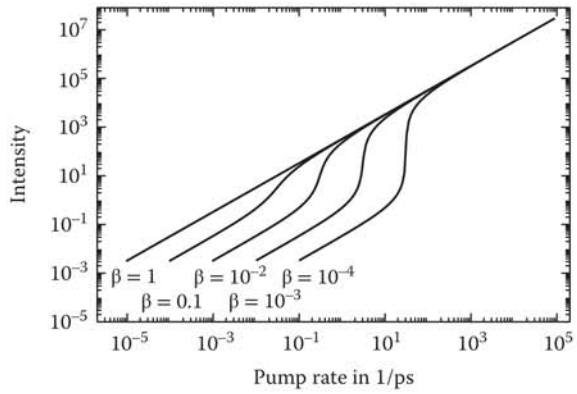


FIGURE 48.3 Laser input–output characteristics as obtained from rate equations. Shown are steady-state results at each pump value, parameterized by the β factor and the cavity loss rate. With increasing β, the threshold shifts to lower pumping and the laser threshold becomes increasingly hard to identify, and additional indicators beyond the reach of the rate-equation description are needed.

to understand and control these properties for new devices. In Section 48.4, we introduce a systematic approach to include semiconductor and correlation effects in extended laser models and discuss effects in nanolasers that set these systems apart from conventional laser devices.

48.3 Key Aspects of Nanolasers

Nanolasers use resonators that confine light in a volume of the light wavelength cube or smaller. The free spectral range, i.e., the energetic distance between resonator modes, grows with decreasing size of the cavity mode volume. Consequently, there is increased potential that the material gain overlaps only with one, or at most, a very few modes [12] of the resonator, as illustrated in the top right panel of Figure 48.4. This sets them apart from other lasers, such as edge-emitting lasers and the typical VCSELs [13,14], where optical modes lie more densely and emission is typically multimode (left panel). A nano-resonator may be a disk or pillar, employing distributed-Bragg reflectors [15], total internal reflection at semiconductor-air interfaces [16], or photonic crystal cavities [17].

It is a great advantage of small resonators that spontaneous emission enhancement $\propto Q/V$ can be used to strongly enhance the light–matter interaction and to favor emission into the single mode and thereby attain high values of the β factor that can approach unity. Here, V is the effective mode volume, and $Q = E/\Delta E$, with $E = \hbar\omega$ the mode energy and ΔE the mode linewidth, is the quality factor of the mode. The spontaneous-emission enhancement is direly needed to make up for the limited amount of gain material that can be placed inside the cavity. This so-called Purcell enhancement [18] can be directly observed in the temporal decay of the emission signal following short excitation. This is shown in Figure 48.5 [19] for micropillar nanolasers of different diameter and, therefore, different effective mode volume V.

As mentioned in Section 48.2, additional indicators are required to identify lasing in high-β devices due to the vanishing kink in the input–output curve. The most established method is to identify the statistical fluctuations of the emitted light. In the thermal regime below threshold, photons prefer to come in bunches. Lasing, on the other hand, originates from a coherent state of the light field, and photons arrive randomly in time. A transition from bunching to close to Poissonian emission is therefore a strong indicator for crossing the threshold to lasing. State-of-the-art streak-camera measurements give a direct visual account of this behavior: Figure 48.6 [20] shows temporally resolved single photon detection events for a system that has been excited by a short laser pulse. Counting statistics can be used to identify single photons, photon pairs, or bunches of more photons as indicated by the boxes. In practice, measurements are repeated thousands

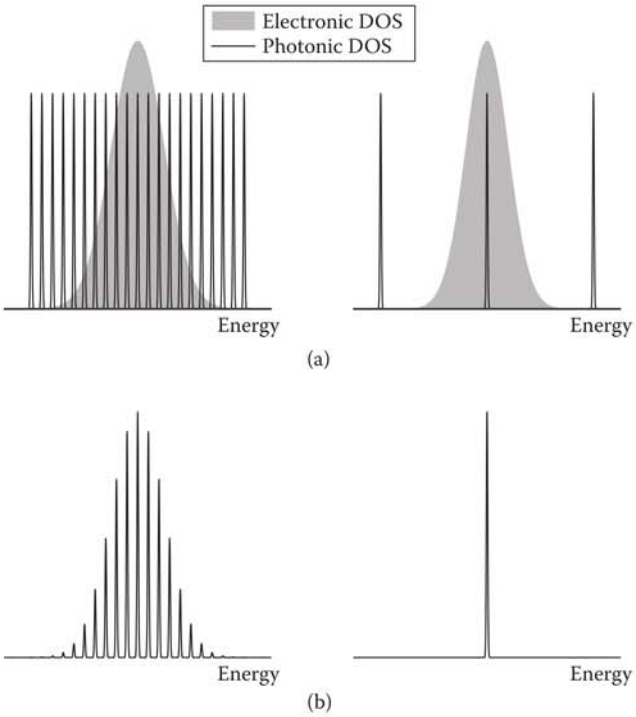


FIGURE 48.4 (a) Illustration of the photonic (black) and electronic (shaded) density of states of a typical laser (left) and a nanolaser (right) with the latter having a much larger free spectral range (energetic spacing between photon modes). The QD gain in the right panel is inhomogeneously broadened due to variations in composition or dimension, see Section 48.8. (b) The resulting spontaneous emission spectra are determined by the overlap of the photonic and electronic density of states. On the other hand, with the proper pillar, disk or photonic crystal design, a nanolaser can be single mode both below and above the lasing threshold. (DOS, density of states.)

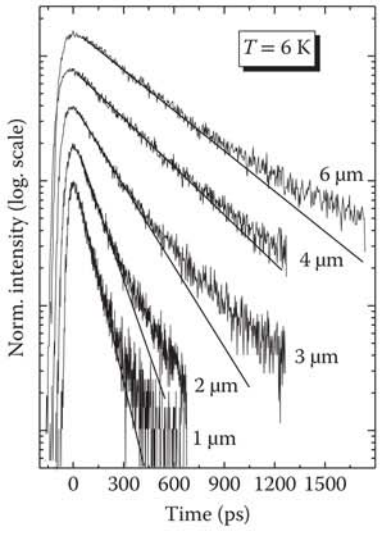


FIGURE 48.5 Time-resolved photoluminescence spectra following a *weak* excitation pulse. A decrease of the spontaneous emission time is observed with decreasing diameter of the micropillar cavities, reflecting the Purcell-enhancement $\propto Q/V$. The weak excitation ensures that the effect is not related to stimulated emission. (Reprinted from Schwab, M. et al., *Physical Review B*, 74, 045323, 2006. With permission.)

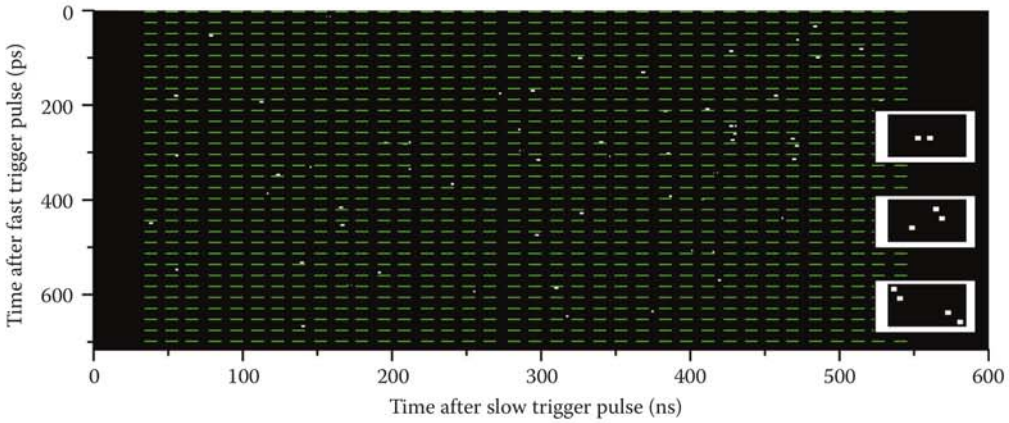


FIGURE 48.6 Streak-camera measurements provide impressive account of statistical fluctuations of the emitted light on the level of single photons. Single- and multiphoton emission events can be detected within a time-bin size of 2 ps (see magnification in the boxes) and provide insight into the quantum-mechanical state of the light field. To perform counting statistics in order to construct correlation functions like the ones shown in Figure 48.7, ten thousands of such measurements are performed. (Reprinted from Aßmann, M. et al., *Science* 325, 297–300, 2009. With permission).

of times. From the collected data, photon correlation functions $g^{(n)}$ can be obtained. They are of particular relevance, as they allow one to characterize the state of the quantum-mechanical light field as given by the distribution function p_n, which describes the *photon statistics* in terms of the probability of finding the system in the photon number state $|n\rangle$. Photon correlation functions are connected to the moments of the photon statistics:

$$\langle n^i \rangle = \sum_{k=0}^{\infty} k^i p_k. \tag{48.4}$$

As an example, the second-order photon-correlation function (sometimes referred to as photon autocorrelation function) is given by

$$g^{(2)} = \frac{\langle n^2 \rangle - \langle n \rangle}{\langle n \rangle^2}. \tag{48.5}$$

Correlation functions of order n take on a value of $n!$ if the light field is in a thermal state, and at all orders they possess a value of 1 for a perfectly coherent state [21]. For a micropillar QD nanolaser, this has been experimentally observed for $g^{(2)}, g^{(3)}$ and $g^{(4)}$ as function of pump rate, shown in Figure 48.7.

More commonly used than a streak camera is a Hanbury Brown and Twiss type measurement [22] schematically shown in Figure 48.8, where the bunching behavior of photons is quantified by the simultaneous registration of two photons at the two detector arms, providing access to the second-order correlation function $g^{(2)}$. For detectors, avalanche photodiodes (APDs) are used. Their time resolution is typically in the range of hundreds of picoseconds. Whether this resolution is sufficient depends on the type of measurement performed and the coherence time τ_{coh} of the emission signal. On the timescale of the latter, $g^{(2)}(\tau)$ goes to unity with respect to the delay time τ between two detection events, and the actual $g^{(2)}(0)$ at $\tau = 0$ can no longer be resolved if the detector's time resolution exceeds the coherence time, leading to the effect that $g^{(2)}(0)$ seems to take on a value of 1 even for thermal light [23].

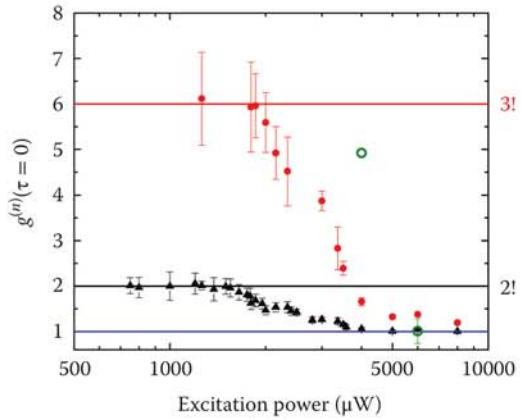


FIGURE 48.7 Photon autocorrelation functions $g^{(2)}(0)$ (triangles), $g^{(3)}(0)$ (dots), and $g^{(4)}(0)$ (open circles) as function of excitation power obtained from streak-camera measurements like the ones shown in the previous figure. For a typical LED to laser transition, all correlation functions show a transition from their thermal values $g^{(n)} = n!$ to that of a coherent field, $g^{(n)} = 1$. (Courtesy of Marc Aßmann, TU Dortmund.)

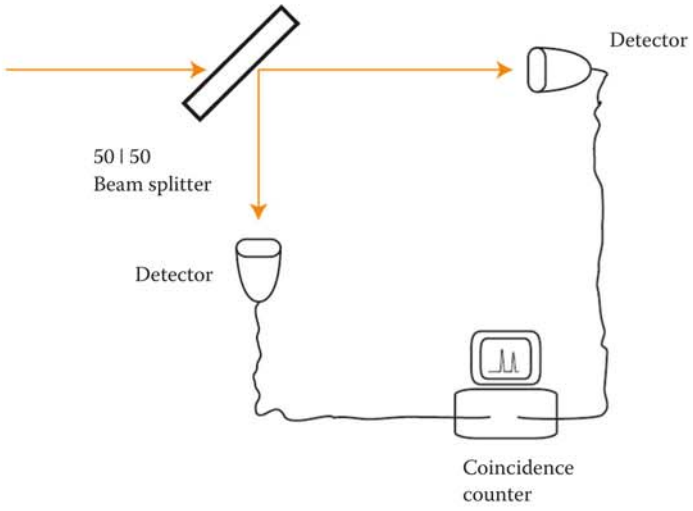


FIGURE 48.8 Illustration of a Hanbury Brown and Twiss measurement. A beam splitter distributes the emission between two detectors. Equal-time detection of two or more photons provides access to the second-order photon autocorrelation function $g^{(2)}(\tau = 0)$. A delay line in one arm can be used to measure $g^{(2)}(\tau \neq 0)$.

Another accessible indicator for coherent emission is a linewidth narrowing, or equivalently, an increase of the coherence time. Experimentally, the linewidth can be determined from spectra or Michelson interferometry measurements [21]. In the steady state $t = t_{ss}$, the coherence time τ_{coh} is related to the first-order coherence of the emission

$$\tau_{coh} = \int_{-\infty}^{\infty} d\tau \, |g^{(1)}(t_{ss}, \tau)|^2, \tag{48.6}$$

which requires the calculation of the two-time quantity in the integral by means of the quantum-regression theorem [24,25]. We will not go into more detail here and refer the reader to [26].

The criteria for lasing in nanolaser devices are a subject of much discussion [5,27]. Ideally, a combination of several criteria is used to identify the lasing regime and the threshold. These include the following:

1. Coherent emission approaching Poissonian photon statistics p_n.
2. A value close to 1 in the second-order photon correlation function $g^{(2)}(0)$. Note that this is a weaker criterion than 1., as $g^{(2)}(0)$ is only linked to the first two moments of the full photon distribution function p_n, c.f. Equation 48.5.
3. An increase in the coherence time at the onset of lasing, when phase-coherent stimulated emission begins to dominate over spontaneous emission.
4. Cavity-photon-induced feedback of the gain material providing gain instead of absorption.
5. A mean intracavity photon number above 1.
6. An S-shaped nonlinear region in the input–output characteristics at threshold.
7. Carrier dynamics developing a hole-burning effect in the carrier-population functions at the energy of the laser transition, cf. Figure 48.13.

All of these criteria have their justifications, but on their own, neither may suffice to conclusively identify lasing. For example, coherent emission (1) is a prerequisite of lasing, but it may be realized without amplification (4 and 5) or originate from an external coherent drive. An S-shaped jump (6) in the input–output characteristic can originate from a variety of effects, such as a transition between multiexciton configurations feeding the laser mode [28], background contributions [29], or spectral wandering. On its own it is not a clear indicator for lasing. Moreover, lasing can take place in the absence of a nonlinear emission regime, e.g., if spontaneous emission losses are small ($\beta \to 1$), or saturation effects set in before the S shape can fully develop, which is particularly relevant under pulsed excitation [30].

48.4 Essential Physics I: Quantum-Optical Semiconductor Laser Theory

Different approaches can be taken in formulating a laser theory, including semiclassical methods, where the radiation field is classical and the system dynamics under the influence of the light–matter interaction is described by Maxwell–Bloch equations [31,32]. In the following, we introduce an extended laser model that accounts for semiconductor-specific effects, such as carrier scattering and dephasing, and that provides access to the statistical properties of the emission. The model is obtained in the framework of a correlation-expansion technique for the quantized light field interacting with carriers occupying the discrete QD conduction- and valence states. The method is versatile, as it allows one to include various aspects of semiconductor and correlation effects on a required level in a systematic and consistent manner. At the same time, the rate equations are contained as limiting case, as has been shown in [33].

In contrast to semiclassical approaches, using a quantized light field introduces the possibility to investigate emission properties beyond the classical regime. In fact, nonclassical effects, such as antibunched light emission around the threshold region, have been demonstrated in QD nanolasers [34]. In semiclassical models, the electromagnetic field drives a coherent polarization $\langle v_\nu^\dagger c_\nu \rangle$ across the bandgap of the semiconductor medium, as illustrated in Figure 48.9. The operators c_ν ($c_\nu^\dagger$) annihilate (create) a conduction-band carrier in the state $|\nu\rangle$, the operators v_ν ($v_\nu^\dagger$) are the equivalent for valence carriers. The resulting dynamics is familiar from atomic two-level models, where a resonant field drives Rabi oscillations. In a fully quantum-mechanical approach, the electromagnetic field is described by modes that are occupied by field quanta, the photons. If the modes are energetically well separated, like in a nanolaser, where the cavity possesses a large free spectral range, it suffices to consider the interaction of the active medium with photons in a single mode q. The coherent polarization $\langle v_\nu^\dagger c_\nu \rangle$ is then replaced by a photon-assisted polarization $\langle b_q^\dagger v_\nu^\dagger c_\nu \rangle$ that describes a carrier deexcitation (or excitation via $\langle b_q c_\nu^\dagger v_\nu \rangle$) by simultaneous emission (absorption)

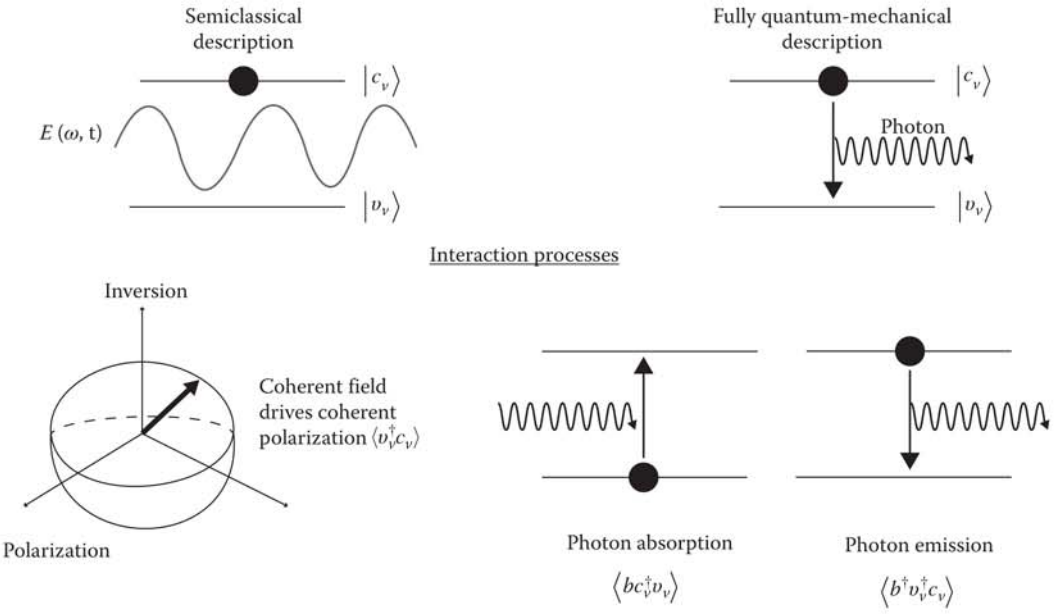


FIGURE 48.9 Illustration of the fundamental light–matter interaction in a semiclassical (left) and fully quantum-mechanical description (right). Semiclassically, matter degrees of freedom are described quantum-mechanically (e.g., by the states $|c_\nu\rangle$ and $|v_\nu\rangle$). The interaction with a classical coherent field $E(\omega, t)$ drives a coherent polarization $\langle v_\nu^\dagger c_\nu\rangle$. The polarization and population dynamics of often described in terms of the rotation of a vector on a unit sphere (Bloch sphere), as illustrated in the bottom left. By quantizing the light field, field quanta (photons) are introduced, and the fundamental interaction processes become the electronic (de)excitations via the simultaneous absorption or emission of a photon (bottom right). The semiclassical coherent polarization is replaced by the photon-assisted polarization $\langle b^\dagger v_\nu^\dagger c_\nu\rangle$.

of a photon into the mode q; the Bose operators b_q ($b_q^\dagger$) annihilate (create) a photon in mode q. The quantum-mechanical and semiclassical pictures are schematically compared in Figure 48.9.

The fully quantum-mechanical laser model is obtained by calculating the dynamical equations for carrier populations $f_\nu^c = \langle c_\nu^\dagger c_\nu\rangle$, $f_\nu^v = \langle v_\nu^\dagger v_\nu\rangle$ in the conduction and valence states of the QD, and the expectation value of the photon number operator in the laser mode q, for which we drop the index q in the following. These equations are obtained from Heisenberg's equations of motion for operators $A(t)$, $\dot{A} = i[H, A]/\hbar$. In the Hamiltonian, we consider the free contributions of QD single-particle states (first and second term in Equation 48.8) and the electromagnetic field (third term in Equation (48.8)), and the light–matter interaction,

$$H = H^{\text{free}} + H^{\text{LM}}, \tag{48.7}$$

with

$$H^{\text{free}} = \sum_\nu \varepsilon_\nu^c c_\nu^\dagger c_\nu + \sum_\nu \varepsilon_\nu^v v_\nu^\dagger v_\nu + \hbar\omega\left(b^\dagger b + \frac{1}{2}\right) \tag{48.8}$$

and

$$H_{\text{LM}} = -i\sum_{\alpha\nu}\left(g_{\alpha\nu}\, c_\alpha^\dagger v_\nu b + g_{\alpha\nu}\, v_\alpha^\dagger c_\nu b\right) + \text{h.c.} \tag{48.9}$$

Here, $\hbar\omega$ is the mode energy, and $g_{\alpha\nu}$ is the light–matter coupling strength that is determined by the overlap of the wave functions of the QD single-particle states $|\alpha\rangle$ and $|\nu\rangle$ and the mode function of the confined electromagnetic field [33]. The sums run over the single-particle states of the individual QD emitters, and the index is to be understood to contain both, QD label and single-particle state. The interaction Hamiltonian contains the elementary processes of electron–hole recombination and generation via emission and absorption of a photon, including the ones schematically shown in the bottom right of Figure 48.9. Note that the counter-rotating terms in Equation (48.9) are often neglected in the rotating wave approximation.

48.4.1 Dynamical Laser Equations

For the dynamical evolution of the photon number $\langle b^\dagger b\rangle$ in the given mode and the electron and hole populations $f_\nu^e = \langle c_\nu^\dagger c_\nu\rangle$, $f_\nu^h = 1 - \langle v_\nu^\dagger v_\nu\rangle$, the contribution of the light–matter interaction H_{LM} in the Heisenberg equations of motion leads to [33]

$$\left(\hbar\frac{\mathrm{d}}{\mathrm{d}t} + \kappa\right)\langle b^\dagger b\rangle = 2\,\mathrm{Re}\sum_{\nu'}|g_{\nu'}|^2\,\langle b^\dagger v_{\nu'}^\dagger c_{\nu'}\rangle, \tag{48.10}$$

$$\hbar\frac{\mathrm{d}}{\mathrm{d}t}f_\nu^{e,h} = -2\,\mathrm{Re}\,|g_\nu|^2\,\langle b^\dagger v_\nu^\dagger c_\nu\rangle. \tag{48.11}$$

Here, we have scaled $\langle b^\dagger v_\nu^\dagger c_\nu\rangle \to g_{q\nu}\langle b^\dagger v_\nu^\dagger c_\nu\rangle$ to have the modulus of the coupling matrix elements appear and have further assumed that g is diagonal in the equal-envelope approximation. The finite lifetime of the cavity mode is introduced by considering a complex mode energy, $\tilde{\omega} = \omega - i\kappa$, where κ is directly connected to the Q-factor of the laser mode, $Q = \hbar\omega/2\kappa$. The dynamics of the photon number is determined by the abovementioned photon-assisted polarization $\langle b^\dagger v_\nu^\dagger c_\nu\rangle$ that describes the expectation value for a correlated event, where a photon in the mode q is created in connection with an interband transition of an electron from a conduction to a valence state. The sum over ν involves all possible interband transitions from various QDs.

An important component of the carrier dynamics is the recombination into nonlasing modes giving rise to the β factor discussed in Section 48.2. For the fraction $1-\beta$ of carriers emitting into nonlasing modes, a term proportional to the spontaneous emission rate can be added to the carrier population equations to account for the losses [33]:

$$\frac{\mathrm{d}}{\mathrm{d}t}f_\nu^{e,h}\Big|_{\mathrm{nl}} = \gamma_{\mathrm{sp}}(1-\beta)f_\nu^e f_\nu^h. \tag{48.12}$$

In contrast to the rate equations (Equation 48.3), where this contribution is typically accounted for by the β coefficient, the spontaneous emission losses in a semiconductor are determined by both electron and hole populations. The dynamical equation for the cavity-photon-assisted polarization is given by

$$\left(\hbar\frac{\mathrm{d}}{\mathrm{d}t} + \kappa + \Gamma + i\left(\tilde{\varepsilon}_\nu^e + \tilde{\varepsilon}_\nu^h - \hbar\omega\right)\right)\langle b^\dagger v_\nu^\dagger c_\nu\rangle =$$
$$f_\nu^e f_\nu^h - \left(1 - f_\nu^e - f_\nu^h\right)\langle b^\dagger b\rangle + \frac{1}{g_\nu}\sum_\alpha g_\alpha C^{\mathrm{x}}_{\alpha\nu\nu\alpha} + \delta\langle b^\dagger b c_\nu^\dagger c_\nu\rangle - \delta\langle b^\dagger b v_\nu^\dagger v_\nu\rangle. \tag{48.13}$$

The free evolution of $\langle b^\dagger v_\nu^\dagger c_\nu\rangle$ is determined by the detuning of the QD transitions from the cavity laser mode. This oscillatory term drops out if the transition is in perfect resonance with the mode. Only accounting for identical resonant QDs is a common and well-justified practice in ensembles of many (>100 emitters), as the summation over all QDs becomes a mere prefactor of the QD number, which is a significant simplification. In a semiconductor, the source term of spontaneous emission is described by an expectation value of four carrier operators $\langle c_\alpha^\dagger v_\alpha v_\nu^\dagger c_\nu\rangle$, see [35]. The Hartree–Fock factorization of

this source term leads to $f_\nu^e f_\nu^h$, which appears as the first term on the right hand side of Equation 48.13. Corrections to this factorization are included in $C^{\mathrm{x}}_{\alpha'\nu\nu'\alpha} = \delta\langle c^\dagger_{\alpha'} v^\dagger_\nu c_{\nu'} v_\alpha\rangle$. Dephasing is what destroys a coherent polarization. In semiconductors, dephasing arises from carrier scattering processes, excitation of carriers, and from the interaction with lattice vibrations that are responsible for the homogeneous linewidth. These contributions are here summarized in a phenomenological constant Γ.

The correlation functions $\delta\langle b^\dagger_q b_q c^\dagger_\nu c_\nu\rangle$ and $\delta\langle b^\dagger_q b_q v^\dagger_\nu v_\nu\rangle$ introduce carrier-photon correlations that are neglected on the level of rate equations. Their time evolution is given by [33]

$$\left(\hbar\frac{\mathrm{d}}{\mathrm{d}t} + 2\kappa\right)\delta\langle b^\dagger b c^\dagger_\nu c_\nu\rangle = -2\left|g_\nu\right|^2 \mathrm{Re}\Big[\delta\langle b^\dagger b^\dagger b v^\dagger_\nu c_\nu\rangle + \left(\langle b^\dagger b\rangle + f^e_\nu\right)\langle b^\dagger v^\dagger_\nu c_\nu\rangle\Big], \tag{48.14}$$

$$\left(\hbar\frac{\mathrm{d}}{\mathrm{d}t} + 2\kappa\right)\delta\langle b^\dagger b v^\dagger_\nu v_\nu\rangle = 2\left|g_\nu\right|^2 \mathrm{Re}\Big[\delta\langle b^\dagger b^\dagger b v^\dagger_\nu c_\nu\rangle - +\left(\langle b^\dagger b\rangle + f^h_\nu\right)\langle b^\dagger v^\dagger_\nu c_\nu\rangle\Big] \tag{48.15}$$

which couples to

$$\left(\hbar\frac{\mathrm{d}}{\mathrm{d}t} + 3\kappa + \Gamma + i\left(\tilde{\varepsilon}^e_\nu + \tilde{\varepsilon}^h_\nu - \hbar\omega\right)\right)\delta\langle b^\dagger b^\dagger b v^\dagger_\nu c_\nu\rangle = -2\left|g_\nu\right|^2\langle b^\dagger v^\dagger_\nu c_\nu\rangle^2 + 2\left[\langle b^\dagger b\rangle + f^h_\nu - f^e_\nu\right]\times$$
$$\left[\delta\langle b^\dagger b c^\dagger_\nu c_\nu\rangle - \delta\langle b^\dagger b v^\dagger_\nu v_\nu\rangle\right] - \left(1 - f^e_\nu - f^h_\nu\right)\delta\langle b^\dagger b^\dagger b b\rangle \tag{48.16}$$

and

$$\left(\hbar\frac{\mathrm{d}}{\mathrm{d}t} + 4\kappa\right)\delta\langle b^\dagger b^\dagger b b\rangle = 4\left|g_\nu\right|^2\sum_{\nu'}\delta\langle b^\dagger b^\dagger b v^\dagger_{\nu'} c_{\nu'}\rangle. \tag{48.17}$$

On this level of approximation, we obtain a closed system of coupled equations that go beyond the rate equations by containing carrier-photon and photon–photon correlation functions that provide access to the second-order photon-correlation function

$$g^{(2)}(\tau = 0) = 2 + \frac{\delta\langle b^\dagger b^\dagger b b\rangle}{\langle b^\dagger b\rangle^2}. \tag{48.18}$$

48.4.2 Pump Process for Continuous-Wave-Excited Lasers

The earlier equations describe the light–matter interaction of a semiconductor system in terms of carrier populations and the polarization-analogue for the quantized light field, i.e., the photon-assisted polarization. An important component is yet missing, namely the interaction of charge carriers with the environment that allows for some sort of excitation mechanism. Pumping can be either done by current injection or by optical excitation. In both cases, electrons and holes are typically created in the continuum states of the wetting layer (WL) or barrier material, from where they are captured into the discrete QD states. This incoherent capture is one example of carrier scattering processes that, at this stage, are still missing from our description. While the quantitative modeling of these many-body processes is discussed in detail in Section 48.5, for continuous-wave (CW) excitation it is often sufficient to consider only the dynamics of carriers in the localized QD states. In its most simple form, two localized states are used for electrons and holes each, which closely resembles a four-level laser scheme. This choice is physically motivated by the electronic structure of typical QDs, which, depending on their size, consists of several confined states. We refer to the lowest (highest) and second-lowest (second highest) electron (hole) state as *s*- and *p*-states, respectively, as shown in Figure 48.10. In this reduced level scheme, the *p*-states are used for carrier pumping at rate P, which approximates the combined effect of carrier excitation in higher

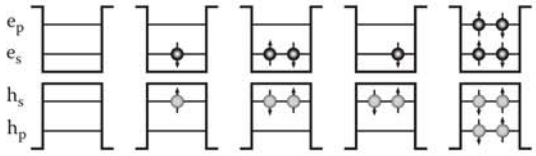


FIGURE 48.10 In a QD with several confined single-particle states for electrons and holes (the s- and p-states are shown), a multitude of many-particle configurations can be realized by occupying these states with electrons and holes. These many-particle configurations contain the neutral, charged, and multiexciton states, and dark states that differ in energy due to the Coulomb interaction between carriers.

states and the successive capture into the QD states:

$$\frac{\mathrm{d}}{\mathrm{d}t} f_{\mathrm{p}}^{\mathrm{e}}\Big|_{\mathrm{Pump}} = P(1 - f_{\mathrm{p}}^{\mathrm{e}} - f_{\mathrm{p}}^{\mathrm{h}}) \tag{48.19}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} f_{\mathrm{p}}^{\mathrm{h}}\Big|_{\mathrm{Pump}} = P(1 - f_{p}^{\mathrm{e}} - f_{\mathrm{p}}^{\mathrm{h}}). \tag{48.20}$$

The blocking factors in the brackets ensure that the occupation remains bounded by unity. From their generation in the QD *p*-states, a phenomenological set of equations distributes carriers into the QD s-states:

$$\frac{\mathrm{d}}{\mathrm{d}t} f_{\mathrm{s}}^{\mathrm{e}}\Big|_{\mathrm{scatt}} = \gamma_{\mathrm{p}\to\mathrm{s}}^{r,e} f_{\mathrm{p}}^{\mathrm{e}}(1 - f_{\mathrm{s}}^{\mathrm{e}}) - \gamma_{s\to p}^{r,e} f_{\mathrm{s}}^{\mathrm{e}}(1 - f_{\mathrm{p}}^{\mathrm{e}}) \tag{48.21}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} f_{\mathrm{s}}^{\mathrm{h}}\Big|_{\mathrm{scatt}} = \gamma_{\mathrm{p}\to\mathrm{s}}^{r,h} f_{\mathrm{p}}^{\mathrm{h}}(1 - f_{\mathrm{s}}^{\mathrm{h}}) - \gamma_{s\to p}^{r,h} f_{\mathrm{s}}^{\mathrm{h}}(1 - f_{\mathrm{p}}^{\mathrm{h}}) \tag{48.22}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} f_{\mathrm{p}}^{\mathrm{e}}\Big|_{\mathrm{scatt}} = -\gamma_{\mathrm{p}\to\mathrm{s}}^{\mathrm{r,e}} f_{\mathrm{p}}^{\mathrm{e}}(1 - f_{\mathrm{s}}^{\mathrm{e}}) + \gamma_{s\to p}^{\mathrm{r,e}} f_{\mathrm{s}}^{\mathrm{e}}(1 - f_{\mathrm{p}}^{\mathrm{e}}) \tag{48.23}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} f_{\mathrm{p}}^{\mathrm{h}}\Big|_{\mathrm{scatt}} = -\gamma_{\mathrm{p}\to\mathrm{s}}^{\mathrm{r,h}} f_{\mathrm{p}}^{\mathrm{h}}(1 - f_{\mathrm{s}}^{\mathrm{h}}) + \gamma_{\mathrm{s}\to\mathrm{p}}^{\mathrm{r,h}} f_{\mathrm{s}}^{\mathrm{h}}(1 - f_{\mathrm{p}}^{\mathrm{h}}). \tag{48.24}$$

Equations 48.21 through 48.24 are motivated in the more general concepts of carrier kinetics in the following section. Together, Equations 48.10–48.24 serve as an accessible and complete model for QD-based lasers, and we discuss exemplary results in Section 48.6.

48.5 Essential Physics II: Carrier Kinetics in QD Systems

As we overviewed in Section 48.2, rate equations can provide important insights for the theoretical understanding of lasers. They can be used to describe input–output curves [36–39], and, combined with Langevin approaches [40,41] even the photon statistics and intensity noise.

An important reason why carrier dynamics is essential for the understanding of laser operation is the excitation mechanism. To see this, let us assume that carrier populations in lasers are given by Fermi–Dirac distribution functions (often a reasonable approximation). This implies that states with below-average kinetic energy are Pauli-blocked for the excitation mechanism. This is illustrated in Figure 48.11, where states with above-average kinetic energy are available for the excitation, while highly occupied states (shaded region, applied to 77K) are not. Therefore, it is important to understand how carriers, injected at the green-shaded area, end up in the spectral window from where they recombine. Accountable for this are two mechanisms that we discuss in this section. The carrier–carrier Coulomb interaction redistributes carriers between different states and leads to an equilibration towards a Fermi–Dirac distribution [42–46]. But as every carrier scattered downward in energy is partnered by a scattering partner that is scattered upward (right sketch in Figure 48.12), the Coulomb scattering is not able to dissipate energy.

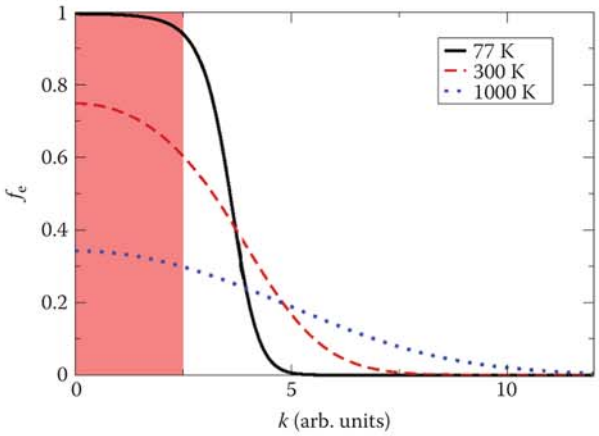


FIGURE 48.11 Fermi–Dirac distribution for electrons in a QW at different temperatures and a carrier density of $5 \times 10^{11}\text{cm}^{-2}$, the shaded background indicates the availability for carrier injection for the low-T distribution.

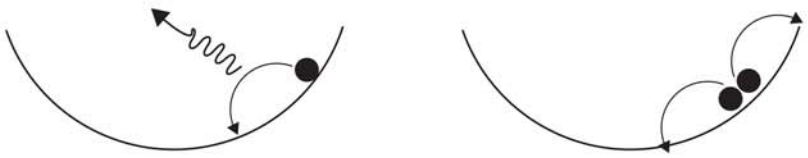


FIGURE 48.12 Conceptual sketch of carrier scattering processes. Left: carrier–phonon interaction, right: carrier–carrier Coulomb interaction.

This leads to a hot Fermi–Dirac function such as depicted in Figure 48.11, that has significantly different properties regarding inversion and gain. The cooling toward the lattice temperature is provided by the carrier-phonon interaction [47–51], that provides for the cooling by taking up the energy of the hot carriers and distributing it toward crystal lattice (left sketch in Figure 48.12). In a laser under CW-excitation conditions the energetic difference from excitation to emission energy also leads to heating and often the carrier-phonon interaction is not strong enough to provide complete cooling towards the lattice temperature. Therefore, a dynamical equilibrium is reached that may consist of Fermi–Dirac distributions with several 100K above the lattice temperature [52,53].

As already outlined in Section 48.4, it is generally necessary to set up equations of motion not for the total carrier density N but for individual carrier populations of different single particle states $f_\alpha^c = \langle c_\alpha^\dagger c_\alpha \rangle$ to describe these effects. Such so-called "kinetic equations" together with the laser equations, Equations 48.10 through 48.20, enable us to treat excitation kinetics and turn-on dynamics of semiconductor QD-based laser systems and will be the topic of this section.

The question of what timescales and by which processes the carriers approach (quasi-) equilibrium conditions (if at all) influences the design of laser devices, as the answer to these questions determines, e.g., the dynamical response and turn-on delay [53–57]. While rate equations can also describe the turn-on dynamics for conventional laser devices, where carrier dynamics takes place on a fs to ps timescale, whereas the laser turn-on happens on the order of 10 ns, the situation is very much different in nanocavities with large Purcell enhancement, i.e., the enhancement of the spontaneous emission rate due to the presence of the cavity, where the turn-on delay can be on the order of 10 ps [53,58] and thus on the same timescale as the carrier scattering. The use of population function both for QD and WL states also allows one to describe effects such as spectral hole burning, or the difference between optical and injection pumping, that are unavailable to a theoretical model that tracks only the dynamics of the total carrier density N.

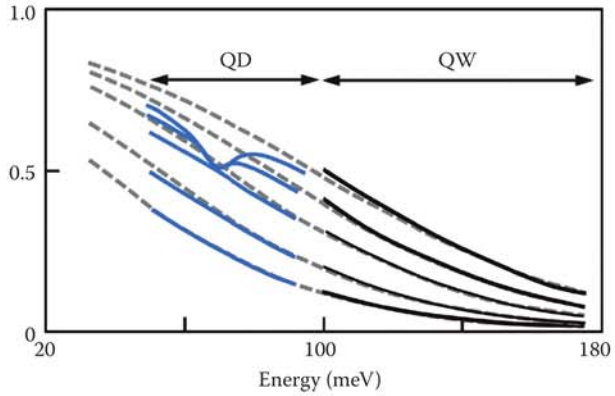


FIGURE 48.13 Carrier populations in a QD+WL system under laser operation for increasing pump intensity. The gray lines describe the population of QD carriers (subject to inhomogeneous broadening), the black lines show the populations of the WL states. For high pump rates the "hole" burnt into the population function by laser operation is clearly visible. The shaded-dashed lines are corresponding Fermi–Dirac functions. (From Chow, W. W. et al., *Light: Science & Applications*, 3, e201, 2014. With permission.)

An example of such carrier population functions are shown in Figure 48.13. The seemingly continuous features of the discrete QD single-particle state populations originate from the large number of emitters in the inhomogeneously broadened ensemble, cf. Section 48.8.

The simplest equation for such state-dependent carrier dynamics is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} f_\alpha = (1 - f_\alpha)\Gamma_\alpha^{\mathrm{in}} - f_\alpha \Gamma_\alpha^{\mathrm{out}}, \tag{48.25}$$

where $\Gamma_\alpha^{\mathrm{in}}$ and $\Gamma_\alpha^{\mathrm{out}}$ give the in- (out-) scattering rates into (out of) the state α. A more detailed version of Eq. (48.25) has the structure

$$\frac{\mathrm{d}}{\mathrm{d}t} f_\alpha = \sum_\beta f_\beta (1 - f_\alpha)\gamma_{\beta\to\alpha} - f_\alpha (1 - f_\beta)\gamma_{\alpha\to\beta} \tag{48.26}$$

and describes that the population of state α is enlarged by scattering from β to α and reduced by scattering from α to β. The Pauli-blocking terms $f_\beta(1 - f_\alpha)$ and $f_\alpha(1 - f_\beta)$ occur as scattering has to be proportional to the occupancy f of the initial and to the non-occupancy $(1 - f)$ of the final state. Under continuous excitation (CW laser), it may suffice to restrict the description to the s-shell and p-shell states of the QD and to use phenomenological scattering rates $\gamma_{\beta\to\alpha}^{e,h}$, in which case we recover Equation 48.21 through 48.24 in Section 48.4.

An often used approximation that is generally of relevance if not only QD states are discussed but also the continuous states of the WL or barrier are taken into consideration is the so-called relaxation time approximation

$$\frac{d}{dt} f_\alpha = \frac{f_\alpha - F_\alpha(N, T)}{\tau}, \tag{48.27}$$

which describes the relaxation of a nonequilibrium carrier population f_α toward a Fermi function $F_\alpha(N, T)$ at temperature T and carrier density N on a time-scale τ. While this approximation often works well, it also necessitates to determine the time-scale τ and the carrier temperature T. The last point is not trivial as in laser devices the carriers typically posses a higher temperature than the lattice [52,53]. To answer these

questions and also to investigate situations where the relaxation time approximation (Equation 48.27) is not applicable we have to take the microscopic origin of the scattering process into account. In the following, we discuss the dominant physical mechanisms behind the carrier dynamics, which are carrier-phonon scattering and carrier–carrier Coulomb scattering.

48.5.1 Carrier–Phonon Interaction

As discussed earlier, carrier temperature is an important factor for laser operation as it controls the population inversion. Hence the optical gain depends strongly on the temperature of the carrier system [59,60]. As the carrier–carrier Coulomb interaction cannot dissipate energy, the source for thermalization of the carriers toward the lattice temperature is the interaction with phonons. To investigate this interaction mechanism, we first have to determine which phonon modes are responsible for the carrier dynamics. A main difference between different phonon branches is whether they are longitudinal or transversal in nature. One can show [61] that due to their ability to form polarizations, mainly longitudinal phonons couple to the carrier system and that for most situations the interaction with transversal phonons is negligible. This leaves us with two phonon branches that are most important for the carrier-phonon interaction. These are the longitudinal optical (LO) phonons and longitudinal acoustic (LA) phonons. The speed of propagation of an acoustic phonon, which is also the speed of sound in the lattice, is given by the slope of the acoustic dispersion relation, $\frac{\partial\omega_k}{\partial k}$. At low values of k, the dispersion relation is almost linear, and the speed of sound is approximately independent of the phonon wavenumber.

Even though we are interested in laser structures, where the active material is composed of nanostructures (QDs), it is a good approximation to take the phonon modes and interaction potentials of the surrounding bulk material. This is the case, as we are mostly interested in long-wavelength phonons where the phonon wavelength is much larger than the size of the nanostructure we are investigating.

The interaction of carriers with LO phonons can be described by the Fröhlich coupling [61]. As this coupling mechanism is Coulombic in nature, the corresponding interaction matrix elements drop off rapidly with increasing momentum transfer $\mathbf{q}$. Therefore a often used approximation consists of treating the LO phonons as dispersionless with constant frequency $\hbar\omega_{\mathbf{q}} = \hbar\omega_{\mathrm{LO}}$ [62]. For acoustic phonons, the main interaction mechanisms are the deformation potential coupling and the piezoelectric coupling [62]. While these play an important role in the line broadening of QD lines in the low temperature and low carrier-density regime, they do not provide an efficient source for carrier scattering.

The interaction matrix elements for the interaction of carriers with LO-phonons are [60,63]

$$|M_{\alpha\beta}|^2 = \frac{M_{\mathrm{LO}}^2}{e^2/\varepsilon_0} V_{\alpha\beta\alpha\beta}, \tag{48.28}$$

containing Coulomb matrix elements $V_{\alpha\beta\alpha\beta}$ (see Equation 48.33 and Reference [64]), the elementary charge e and the vacuum permittivity ε_0. The prefactor $M_{\mathrm{LO}} = 4\pi\alpha\frac{\hbar}{\sqrt{2m}}(\hbar\omega_{\mathrm{LO}})^{\frac{3}{2}}$ includes the polar coupling strength α and the reduced mass m. The matrix elements are given here in a way that explicitly shows the Coulombic nature of the carrier-LO-phonon coupling.

A first approach toward carrier scattering is given by the Boltzmann equation [40,61]. For carrier-phonon scattering, this is given by Equation 48.26 with

$$\gamma_{\beta\to\alpha} = \sum_q M_{\beta\alpha}(q)\Big((1+n_q)\delta(\epsilon_\beta - \epsilon_\alpha - \hbar\omega_{\mathrm{LO}}) + n_q\delta(\epsilon_\beta - \epsilon_\alpha + \hbar\omega_{\mathrm{LO}})\Big). \tag{48.29}$$

The terms proportional to n_q and $(1 + n_q)$ describe phonon absorption and emission processes. In QD systems, additional effects exist that are not described within the framework of a Boltzmann equation, but are closely connected to carrier scattering, as they stem from the same microscopic interaction mechanism. These are quasi-particles and non-Markovian effects. The main reason is that the carrier-phonon interaction leads to the formation of a new quasi-particle. The quasi-particle obtained by dressing the carriers with the carrier-phonon interaction—the polaron—describes the lattice distortion accompanying the electron in its motion. In the framework of Green's functions (GFs), polarons are described using the retarded GF Kadanoff–Baym equation [63]:

$$\left[i\hbar\frac{\partial}{\partial\tau} - e_\alpha^a\right] G_\alpha^{a,R}(\tau) = \delta(\tau) + \int d\tau' \;\; \Sigma_\alpha^{a,R}(\tau - \tau')\, G_\alpha^{a,R}(\tau'). \tag{48.30}$$

The corresponding retarded selfenergy in random-phase approximation (RPA) [62] is given by

$$\Sigma_\alpha^{a,R}(\tau) \;=\; i\hbar \sum_\beta |M_{\alpha\beta}|^2 G_\beta^{a,R}(\tau)\, d^<(-\tau), \tag{48.31}$$

where the phonon propagators $d^{\gtrless}$ contain the phonon frequency and the phonon population [63]. The spectral function $\hat{G}_\alpha(\hbar\omega) = 2\mathrm{Im}(G_\alpha^{a,R}(\hbar\omega))$, which follows from the retarded GF, can be seen as a generalization of the single particle energy, i.e., a free particle has a spectral function of the form $\hat{G}_\alpha(\hbar\omega) = \delta(\epsilon_\alpha - \hbar\omega)$. In contrast, the Polaronic spectral functions but also posses a finite spectral width that is connected to the lifetime of the respective quasi-particle.

Using the quasi-particles as described by the polaron retarded GF, we can formulate a quantum-kinetic equation for polaron scattering:

$$\frac{\partial f_\alpha^a(t)}{\partial t} = 2\,\mathrm{Re}\sum_\beta \int_{-\infty}^{t} dt' \; |M_{\alpha\beta}|^2 G_\beta^{a,R}(t,t') \left[G_\alpha^{a,R}(t,t')\right]^* \\ * \left\{ \left[f_\beta^a(t')(1 - f_\alpha^a(t'))\right] d^>(t',t) - \left[f_\alpha^a(t')(1 - f_\beta^a(t'))\right] d^<(t',t) \right\}, \tag{48.32}$$

which contains scattering by phonon emission and absorption processes just as in the Boltzmann equation. Effects included here that go beyond the Boltzmann treatment are a weakening of the exact energy conservation due to the polaron properties (finite spectral width) and due to non-Markovian effects described by the explicit dependence of the time change of f_α^a at time t on all earlier times.

Results from such a quantum-kinetic treatment are shown in Figure 48.14. The QD level spacing is chosen to be 10% higher than the LO-phonon energy of 36meV. Therefore, a Boltzmann scattering equation predicts a complete inhibition of scattering due to the δ-functions in Equation (48.29). In contrast, the quantum-kinetic treatment shows that, due to quasi-particle effects (dotted line), scattering channels open up. The solid line shows the combined influence of quasi-particles and non-Markovian effects, that are caused by the time integrals in Equation (48.32). The inhibition of scattering as given by a Boltzmann equation, was predicted as a "phonon-bottleneck" early on [65,66], however, fast-carrier scattering was observed e.g., in Reference [67], showing the relevance of polaronic effects. For details see References [63,68,69].

48.5.2 Coulomb Interaction

The most important carrier scattering mechanism in the regime of high excitation densities, relevant for laser operation, is the carrier–carrier Coulomb interaction. It leads to a fast redistribution of carriers from

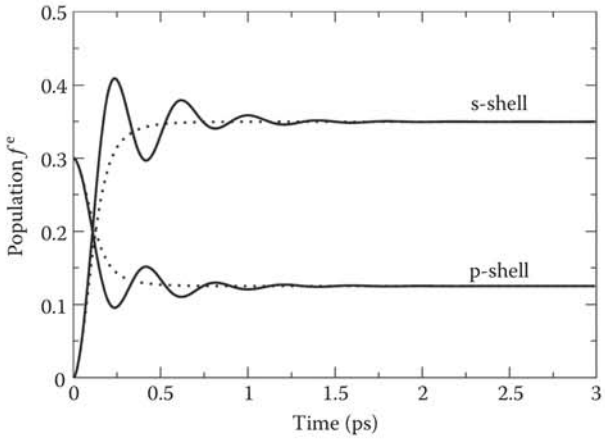


FIGURE 48.14 Redistribution of carrier occupations due to scattering of electrons from p-shell to s-shell in a QD via the carrier phonon interaction. In a Boltzmann-equation approach the occupations would remain constant due to a complete lack of scattering processes as described in the text (Reprinted from Seebeck, J. et al., *Physical Review B*, 71, 125327, 2005. With permission).

the spectral window where excitation takes place toward the spectral window from where they recombine. While Coulomb scattering cannot dissipate energy and thermalize the system (see Figure 48.12), it leads to equilibration toward a hot Fermi–Dirac function.

To determine the strength of the Coulomb interaction in nanostructures, we need to determine matrix elements of the Coulomb interaction in the single-particle basis given by, e.g., tight-binding descriptions of QDs. Using such single-particle wavefunctions, we can construct the Coulomb matrix elements

$$V_{\alpha\beta\gamma\delta} = \frac{1}{A} \sum_{\mathbf{q}} V_{\mathbf{q}} \left\langle \alpha \left| e^{-i\mathbf{q}\cdot\mathbf{r}} \right| \delta \right\rangle \left\langle \beta \left| e^{+i\mathbf{q}\cdot\mathbf{r}} \right| \gamma \right\rangle, \tag{48.33}$$

consisting of overlap integrals between single-particle wavefunctions and the Coulomb potential $V_{\mathbf{q}}$ [64]. A closer analysis of these Matrix elements also points towards a source for the efficiency of Coulomb scattering in nanostructures. Even though the available phase space of scattering partners is reduced significantly in QDs, compared to QW or bulk systems, this is partially balanced by enhanced interaction matrix elements due to the strong localization of the QD wave functions.

Like for the carrier-phonon scattering a first description of carrier–carrier scattering can be given by a Boltzmann-type equation

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f_\alpha(t) = & \frac{2\pi}{\hbar} \sum_{\beta\gamma\delta} \left(\left| W_{\alpha\gamma\delta\beta} \right|^2 - W_{\alpha\gamma\delta\beta} W^*_{\alpha\gamma\beta\delta} \right) \delta(\epsilon_\alpha - \epsilon_\beta + \epsilon_\gamma - \epsilon_\delta) \\ & \Big[f_\beta(t)(1 - f_\alpha(t)) f_\delta(t)(1 - f_\gamma(t)) - f_\alpha(t)(1 - f_\beta(t)) f_\gamma(t)(1 - f_\delta(t)) \Big], \end{aligned} \tag{48.34}$$

which contains screened interaction matrix elements $W_{\alpha\gamma\delta\beta}$ and an energy conserving δ-function. A feature of the carrier–carrier Coulomb scattering in QDs is that, while most scattering channels become extremely inefficient at low carrier density due to the population factors in Equation (48.34), some some scattering channels like electron–hole scattering are possible at very low carrier densities [70].

Results for the carrier–carrier Coulomb scattering in a QD laser under optical cw excitation are presented in Figure 48.15, where the carrier population functions is shown as a function of energy for different times. Discrete symbols denote the QD populations for s-shell and p-shell and the lines the continuous

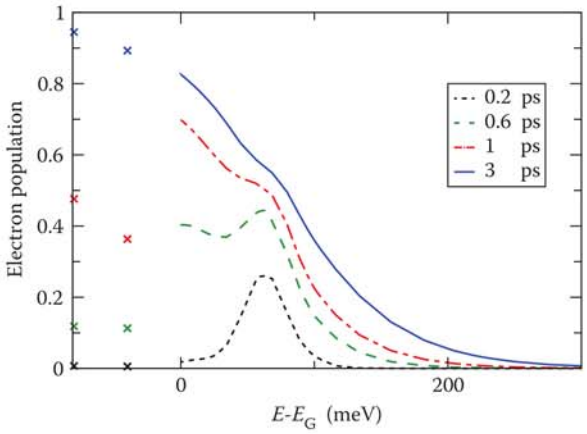


FIGURE 48.15 Carrier populations as a function of energy for electrons in a QD system under laser operation for different times. Symbols denote QD populations, lines WL populations. (Reprinted from Lorke, M. *Applied Physics Letters*, 99, 151110, 2011. With permission.)

WL states that are optically excited. After the initial Gaussian population profile has been excited by the pump, an ultra-fast redistribution is observed, that immediately starts to populate the QD states. It should be noted that, contrary to an often assumed situation [56,71], the timescales between the relaxation dynamics in the WL and the dynamics of the QD populations are not decoupled. On one hand, this is due to the efficiency of the initial in-scattering into the QD, as the QD states are initially empty. On the other hand, the relaxation within the WL is significantly slower than in a pure QW, as the capture into the QD states is most efficient for the WL states with low quasi-momenta. Therefore, these states are constantly depleted during the early stage of the kinetics, which slows down the relaxation of the WL distribution itself toward a quasi-equilibrium distribution [53] Furthermore, after the initial relaxation that lasts about 3 ps, a heating of the carrier population is observed on a timescale of 100 ps, as the pumping into the WL injects carriers at a higher energy than the emission energy.

These results also show (compared to those of the preceding section) that Coulomb and phonon scattering act on the same timescale in QD systems.

As for the phonons, quantum-kinetic effects lead to quasi-particle formation and non-Markovian effects. The corresponding quantum-kinetic equations can be written as

$$\frac{\partial f_\alpha^a(t)}{\partial t} = \frac{2}{\hbar}\mathrm{Re}\sum_{\beta\gamma\delta}\int_{-\infty}^{t} \mathrm{d}t' \Big(\left|W_{\alpha\gamma\delta\beta}\right|^2 - W_{\alpha\gamma\delta\beta}W^*_{\alpha\gamma\beta\delta}\Big) G_\beta^{a,\mathrm{R}}(t,t') \left[G_\alpha^{a,\mathrm{R}}(t,t')\right]^* * G_\delta^{a,\mathrm{R}}(t,t') \left[G_\gamma^{a,\mathrm{R}}(t,t')\right]^*$$
$$* \Big[f_\beta(t')(1-f_\alpha(t'))f_\delta(t')(1-f_\gamma(t')) - f_\alpha(t')(1-f_\beta(t'))f_\gamma(t')(1-f_\delta(t')) \Big]. \quad (48.35)$$

As for the Boltzmann-type equation for carrier scattering, this equation contains direct and exchange scattering, proportional to $\left|W_{\alpha\gamma\delta\beta}\right|^2$ and $W_{\alpha\gamma\delta\beta}W^*_{\alpha\gamma\beta\delta}$, respectively. The energy conserving δ-function is generalized by the product of retarded GFs just as for the carrier-phonon interaction and via the explicit time dependence on earlier times memory effects are included. While these equations have not been evaluated in detail due to their numerical complexity, it is expected that the included quasi-particle and non-Markovian effects would influence several properties of nanolasers. On one hand, large signal modulation and turn-on delay are expected to be influenced by non-Markovian effects. On the other hand, quasi-particle effects can, as for the carrier-phonon interaction, alter the scattering rates

influencing also the modulation response and threshold current. In the photon-assisted polarization, Equation 48.13 exist analogous expressions to the scattering contributions given in Equation 48.35 that lead to excitation-enhanced dephasing

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle b^\dagger \nu_\alpha^\dagger c_\alpha\rangle(t) = \int_{-\infty}^{t} \mathrm{d}t'(-\Gamma_\alpha^{\mathrm{DD}}(t,t'))\langle b^\dagger \nu_\alpha^\dagger c_\alpha\rangle(t') + \sum_\beta \Gamma_{\alpha\beta}^{\mathrm{OD}}(t,t')\langle b^\dagger \nu_\beta^\dagger c_\beta\rangle(t'), \tag{48.36}$$

where Γ^{DD} and Γ^{OD} consist of interaction matrix elements W, retarded GFs $G^{a,\mathrm{R}}$ and population functions f_α in a similar way as they appear in Equation 48.35 for the carrier scattering. We have investigated the influence of quasi-particle and non-Markovian effects on the quasi-classical analogues of these quantities in detail to study optical gain spectra of QD systems [8,60,72–75] and we expect an influence of these contributions on the turn-on behavior and the modulation dynamics [38].

48.6 Connecting to Experiments: Transition to Lasing and the β Factor

As first example, we discuss characteristic input–output curves and statistical properties of a Purcell-enhanced QD nanolaser device. In Figure 48.16 results from the semiconductor laser model are shown for various values of the β factor and fixed number of QD emitters. For $\beta = 1$, all spontaneous emission is

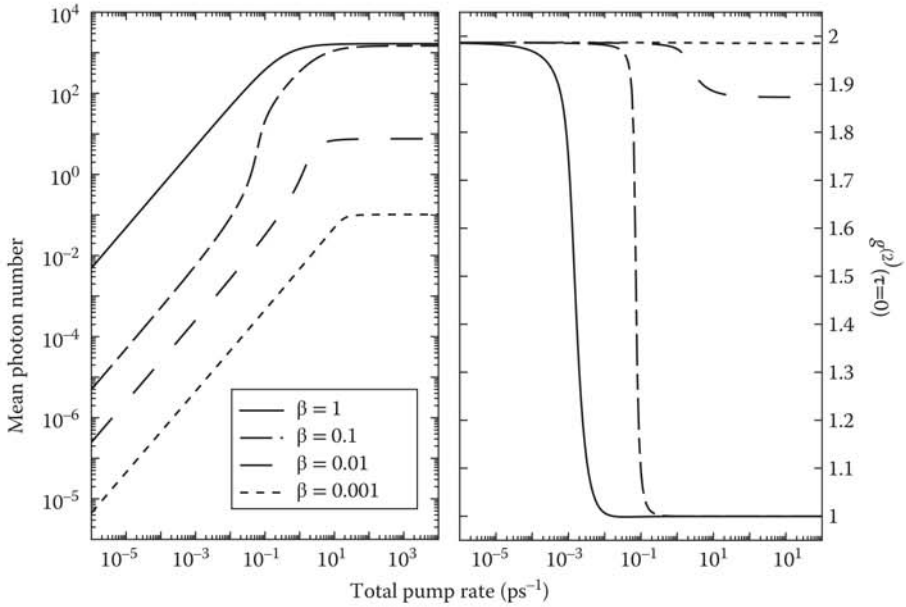


FIGURE 48.16 Input–output curves (left) and autocorrelation function $g^{(2)}(0)$ (right) for a nanolaser with 300 QD emitters. Curves are shown for different values of the β factor. For $\beta = 1$, a visible threshold is absent in the input–output curve, whereas $g^{(2)}(0)$ shows a clear transition from thermal to coherent emission. Note that the jump from below to above threshold is larger than in the results obtained from rate equations (Figure 48.3) and does here no longer scale with $1/\beta$. For $\beta = 0.1$, a soft kink appears in the input–output curve. For smaller values of β, the limited number of QD emitters cannot provide sufficient gain to reach lasing before saturation is reached. A cavity-Q factor of 32,000 and a cavity-enhanced spontaneous emission rate of $\tau_{\mathrm{sp}} = 50$ ps has been used.

directed into the laser mode. In contrast to the rate-equation result that we have discussed in the context of Figure 48.3, saturation is visible at strong pumping. This effect is important in QD nanolasers, since the small size of the resonator naturally limits the amount of gain material that can be brought into the cavity. A monotonous increase of the output intensity with pumping, as given by the rate equations (Equation 48.1), would be unphysical. As a consequence, for smaller values of β, the gain can become insufficient to reach lasing before saturation sets in. In the example in Figure 48.16, this is visible for $\beta = 0.01$, where the threshold is reached, but the emission remains largely thermal with $g^{(2)}(0) \approx 1.9$, and for $\beta = 0.001$, where the QD emission saturates long before the laser threshold at a low mean intracavity photon number of 0.1. Access to the photon autocorrelation function $g^{(2)}(0)$ provides an important tool in judging the emission properties far beyond what can be inferred from the emission intensity alone and is a significant advancement over the rate equations.

Deeper insight into cavity-QED effects in nanolasers can be obtained from the emission and correlation *dynamics*. The top panel in Figure 48.17 shows time resolved the emission pulse (solid curves) and autocorrelation function $g^{(2)}(t, \tau = 0)$ (symbols) following picosecond-short optical excitation of the nanolaser. Also here the high time resolution has been realized by using a streak-camera setup. Following a sufficiently strong excitation pulse, the statistical properties of the emission pulse change as function of time from thermal to coherent and back to thermal. Coherent emission is reached around the peak region of the emission pulse and forms a plateau when the excitation level is above the threshold region. The data agrees well with time-dependent results of the laser model (bottom panel), which are readily available from the time evolution of the coupled differential equations presented in Section 48.4. In contrast to conventional measurements that integrate over the emission pulse, here it is possible to track the degree of coherence during the pulse and to provide an understanding of the coherence it contains. In their combination, state-of-the art measurements and microscopic models can reveal new aspects of solid-state light sources, reflecting the emission dynamics and the dynamics of quantum-mechanical correlations down to the few-photon level. The explicit treatment of dynamical properties of nanolasers, such as carrier dynamics, turn-on delay, and modulation response requires to go beyond the simple approach in Equations 48.21 through 48.24 for the carrier scattering and use the more involved methods of Section 48.5 due to the sensitivity of these quantities [55].

48.7 Interemitter Coupling Effects and Superradiance in Nanolasers

When several emitters couple to a common light field, the interaction introduces correlations between them. The eigenstates of the interacting system are then no longer given by those of the individual systems, but by so-called Dicke states of the collective system. In 1954, Dicke has shown [76] that from certain collective states, emission can be enhanced or suppressed, which is referred to as superradiance and subradiance. In particular, for the Dicke state of the half inverted system of N two-level emitters, the recombination rate is proportional to $N(N-1)$ [77]. This quadratic behavior with emitter number has become one of the hallmarks of superradiance. Since its discovery by Dicke, superradiance has been extensively and continuously studied in a variety of systems. Most commonly, however, superradiance is associated with temporal modifications of the emission in spatially extended systems, such as clouds of atoms [78] or, when it comes to semiconductor systems, excitons in extended systems [79].

In an ensemble of QDs that are embedded in a microcavity, depending on the inhomogeneous broadening of the QDs and the cavity Q factor, a fraction of the QDs overlaps both spectrally and spatially with the mode. The common light field these resonant solid-state emitters are subjected to provides a mechanism to form inter-emitter correlations that can lead to sub- and superradiant effects in QD nanolasers. Laser models typically assume that emitters act individually and neglect such effects. In the following, we outline how to use the formalism introduced in Section 48.4 to formulate a laser theory that includes inter-emitter coupling in an approximative way and give a quantitative discussion of its effects. As it turns out, their

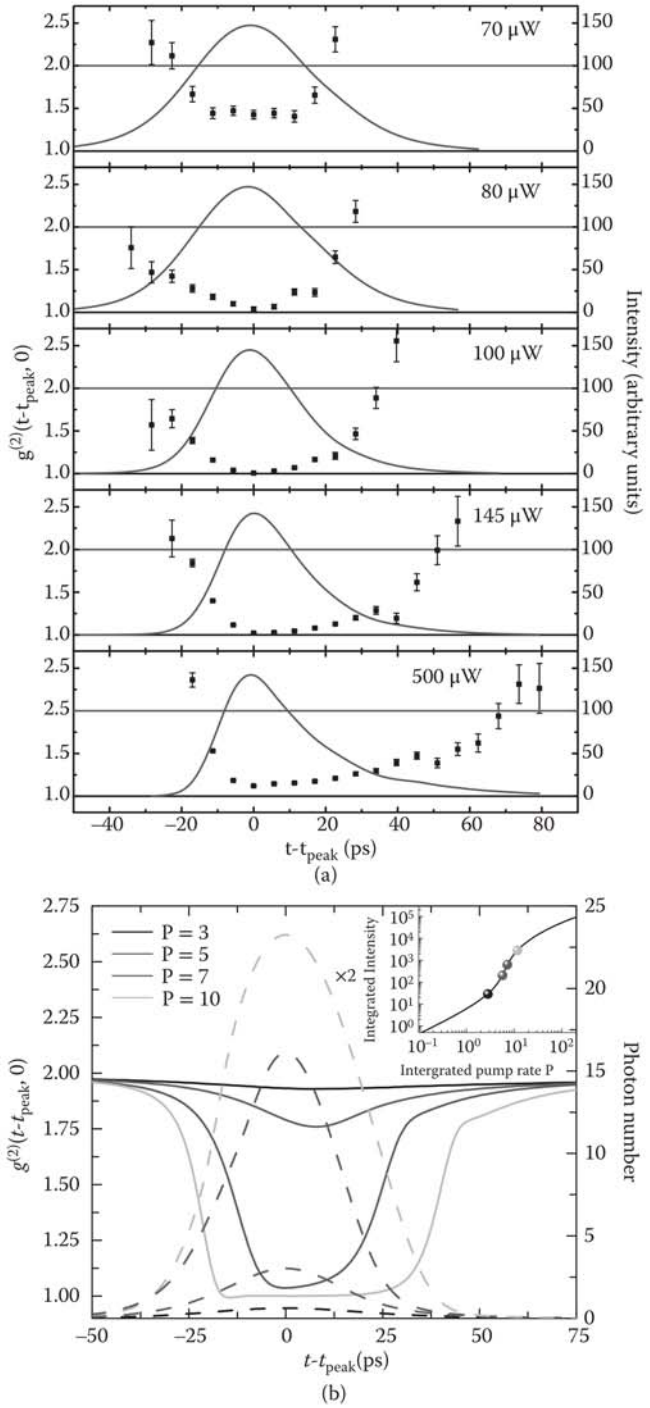


FIGURE 48.17 Time-resolved measurements (a) and theoretical results (b) for emission intensity and photon autocorrelation function following picos-second short optical excitation of a QD-micropillar laser. (Reprinted from Aßmann, M. et al., *Physical Review B* 81, 165314, 2010. With permission).

impact on both stationary and dynamical laser properties can be quite significant. A detailed description of the formalism is found in [80].

To this end, we introduce a formulation for operators that act on the electronic QD states in terms of the many-particle configurations $|i\rangle$, where i refers to configurations like ground state, exciton, trion, biexciton, various dark states, etc. These configuration states, some of which are shown in Figure 48.10, form the basis of the single-QD subspaces of the Hilbert space. Operators Q_{ij}^{α} are introduced that describe either transitions from configuration j to configuration i for $i \neq j$, or the probability that configuration i is realized for $i = j$ in QD α. This approach has advantages over descriptions that use single-particle operators c_ν and v_ν, like we did in Section 48.4. In doing so, many-electron configurations of one emitter are addressed by a single operator Q rather than many single-particle creation and annihilation operators. It further simplifies the systematic approximation scheme used to truncate the hierarchy of equations of motion with respect to correlations involving *different* QDs. Interemitter correlations are described by a hierarchy of expectation values $\langle Q_{ij}^{\alpha} Q_{kl}^{\beta} \rangle$, $\langle Q_{ij}^{\alpha} Q_{kl}^{\beta} Q_{mn}^{\gamma} \rangle$, ... that involve transitions in QDs α and β, or QDs α, β and γ, etc. The cluster expansion approach facilitates a truncation at a certain order of interemitter coupling, e.g. on the level of pair-correlations between emitters given by $\langle Q_{ij}^{\alpha} Q_{kl}^{\beta} \rangle$ and illustrated in Figure 48.18. All quantum-mechanical expectation values containing electronic QD operators acting on more than two emitters are then neglected. On this level, which is the lowest order that contains interemitter coupling effects, a system of N identical QDs with four confined states each is described by about 300 coupled differential equations, also including photon correlations up to the fourth order providing access to the autocorrelation function $g^{(2)}(0)$. The generation of the equations of motion can be assisted by using computer algebra [80]. While the formalism may appear to involve a great deal of effort, it is highly efficient in the numerical evaluation, and computations for many emitters can typically be performed on a single workstation.

The impact of interemitter coupling on the steady-state properties of a continuously driven QD nanolaser can be seen in Figure 48.19. In the top panel, input–output curves are shown that have been obtained from two separate calculations, one including and one omitting QD–QD correlations. C_F in the bottom panel visualizes the difference in the output intensity I, i.e.,

$$C_F = \frac{I_{\text{rad. coupled QDs}}}{I_{\text{independent QDs}}} - 1. \tag{48.37}$$

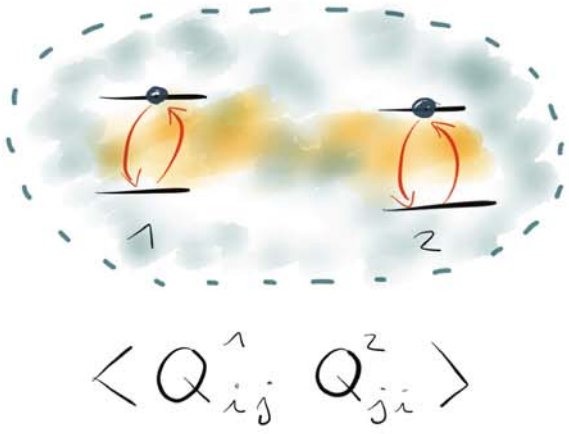


FIGURE 48.18 Illustration of photon-mediated inter-emitter dipole correlations between two emitters.

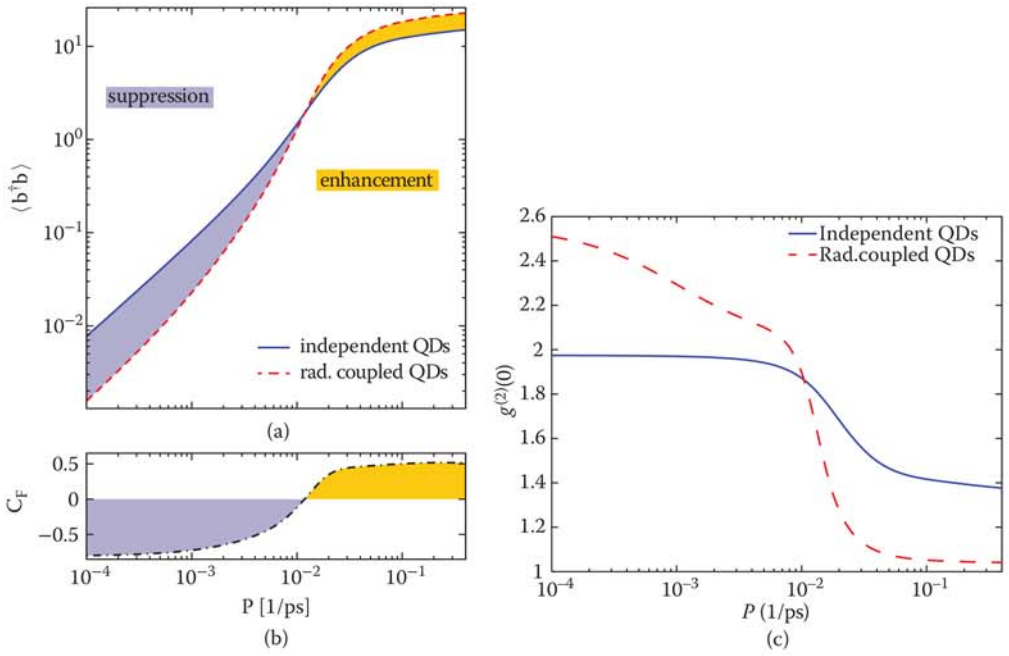


FIGURE 48.19 (a) Input–output curve for a QD nanolaser with typical specifications: 100 resonant emitters, cavity loss rate $\kappa = 0.05/\text{ps}$ (corresponding to $Q = 20,000$), carrier relaxation rate $\gamma_r = 0.05/\text{ps}$, and spontaneous losses into leaky modes $\gamma_{\text{spont}} = 0.01/\text{ps}$. (b) The cooperativity factor visualizes the difference caused by the radiative coupling. (Adapted from Leymann, H. A. M. et al., *Physical Review Applied*, 4, 044018, 2015. With permission.) (c) Corresponding autocorrelation functions. At low excitation, superthermal bunching is a signature of the subradiant emission regime. At high excitation, the threshold to lasing is not fully crossed if emitters act individually and radiative coupling is neglected. Dipole–dipole correlations between distant emitters create additional coherence in the superradiant regime.

The possibility to "switch" the coupling on an off is a particular advantage of the equation-of-motion method and allows to directly assess the impact of the radiative coupling. The following effects are observed:

1. In the low-excitation regime, the inter–emitter coupling creates a subradiant state with a reduced spontaneous emission rate and reduced photon output. Once the laser threshold is crossed, a superradiant state is formed which enhances the laser output *in addition* to stimulated emission provided by photons in the cavity. In sum, the effect on the input–output curve over the whole excitation regime is quite significant. In particular, in the given example, which has been calculated for a typical microcavity laser with 100 QD emitters as active medium, one would underestimate the true β factor by one order of magnitude when ignoring the radiative coupling
2. The input/output characteristics exhibit a different and untypical slope below threshold.
3. In the subradiant regime (low emission intensity), dipole correlations between pairs of QDs cause an increased probability to emit photons synchronously, leading to super-thermal bunching behavior in the statistical properties of the emission, as shown in the right panel of Figure 48.19.
4. Due to inter-emitter coupling, fewer QDs are required to reach the lasing threshold. In small systems, where achieving sufficient gain due to the limited number of emitters in the cavity has always been an issue, the presence of radiative coupling can explain why lasing is more easily reached than expected from conventional laser theories that do not include QD–QD correlations.

5. Coherent emission with $g^{(2)}(0) = 1$ is achieved at lower intracavity mean photon number in the presence of interemitter coupling. This may be interpreted as an increase in the "coherence per photon" due to the alignment of dipole correlations between emitters.

Interestingly, these theoretical predictions are hard to verify experimentally. A strong indicator for the presence of a dipole-correlated phase is the super-thermal photon bunching [81], which has recently been observed in time-resolved studies using a streak-camera setup in combination with microscopic theory [82]. In the same study, highly accelerated spontaneous emission and subradiant photon trapping have been demonstrated. In their sum, the observation of these criteria give a convincing account of the presence of strong inter-emitter coupling effects in a typical QD nanolaser.

48.8 Nonresonant Coupling and Lasing from Multiexciton States in Few-QD Systems

An inherent property of self-assembled QDs that are grown in the so-called *Stranski–Krastanov* mode [83] is inhomogeneous broadening. It refers to inhomogeneity in size, shape and material composition from one QD to the next, causing differences in the confinement potential and with it the electronic single-particle states. While the typical linewidth of QD transitions is tens to hundreds of μeV depending on temperature [2,84,85], the line of an ensemble of self-assembled QDs consists of the spectral lines of the individual emitters, which for many emitters then appears as a single line with a typical broadening of 20–50 meV. An illustrative demonstration of this behavior is given in Figure 48.20 [86]. Inhomogeneous broadening is generally seen as a weak point of self-assembled QDs, as it takes away some of the advantage of the narrow and well-defined transition energies that is characteristic for the single emitter. On the other hand, in cavity-QED it simplifies the task of creating spectral overlap between emitters and a cavity mode (cf. Figure 48.4) and compensates for spectral wandering, since the ensemble broadening is large compared to the cavity linewidth (typically 50–500 μeV). The influence of detuning of emitters in the ensemble and a cavity mode is greatly reduced by nonresonant coupling mechanisms discussed later. The influence of inhomogeneous broadening on QD laser properties has been studied in the framework of a QD laser theory in [27].

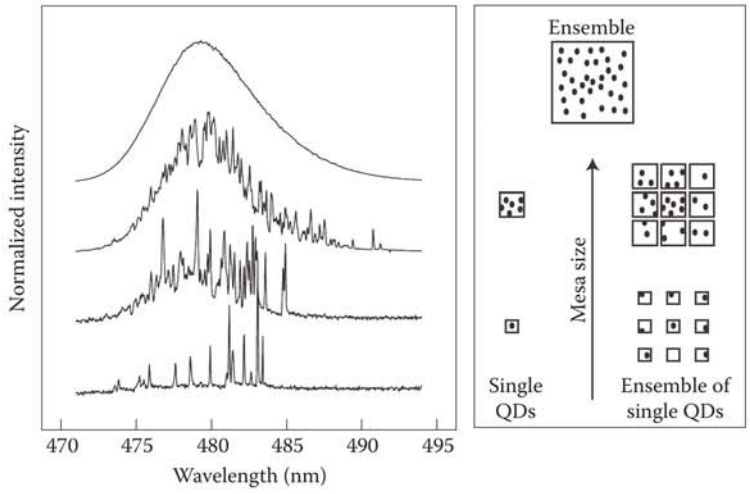


FIGURE 48.20 Experimental results illustrating inhomogeneous broadening in a quantum dot (QD) ensemble. With decreasing mesa size, the QD ensemble consists of fewer and fewer emitters (illustrated right). The broad inhomogenous emission spectrum for many emitters (top left) is revealed to consist of discrete lines as the number of emitters is reduced (bottom left). (Reprinted from Scheibner, M. et al., *Nature Physics*, 3, 106–110, 2007. With permission.)

In this section, we look into what happens if the resonator contains only a few emitters, and the gain spectrum consists of a multitude of separate sharp lines. Such a situation is depicted in the left panel of Figure 48.21 [87]. Then, the emission properties are determined by the interplay of various many-particle configurations of each QD emitter, and their contribution to the photon production depends on the relative spectral position to the cavity mode. While at first it seems contradictory that lasing can be achieved in such a system, strong emission enhancement can drive even a single emitter close to the regime of lasing. At the same time, nonresonant coupling mechanisms haven been identified to provide means for a detuned QD to emit photons into the cavity mode. Nonresonant cavity feeding plays a significant role in nanolasers with ensembles of emitters, but a quantification of its effect is nearly impossible to obtain in such a system. The few-emitter limit, however, offers the unique possibility to study non-resonant mode coupling in a highly controllable environment. In fact, the underlying physical process has long been elusive. Intensive research on this topic has identified three mechanisms that are responsible for the effect:

1. Accoustic phonons can bridge small energy gaps in the range of 1 meV between a detuned emitter resonance and the cavity mode [88–93]. The efficiency is determined by the imbalance between phonon-assisted cavity-photon emission, and the reverse process. Low temperature favors this asymmetry, because low phonon population makes the emission of a phonon more likely than its absorption. Therefore, phonon-mediated nonresonant coupling is more efficient for coupling blue-detuned emitter transitions to the mode [94].
2. At higher carrier densities, interactions with the quasi-continuous WL states provide means for Auger-like scattering processes that allow to bridge larger energy gaps of up to 10 meV [95–99]. While their efficiency is small in comparison to the ultrafast intraband relaxation [53,64,100–102], Auger-coupling is of central importance in the high-excitation regime of nanolasers.
3. QDs possess a rich electronic structure that gives rise to a multitude of multiexciton states, as illustrated in Figure 48.10. The Coulomb interaction energetically separates transitions from different multi-exciton states. This energetic separation, and the possibility of interband recombinations from higher confined states (e.g., the p shell) covers an even larger energetic window, as has been shown

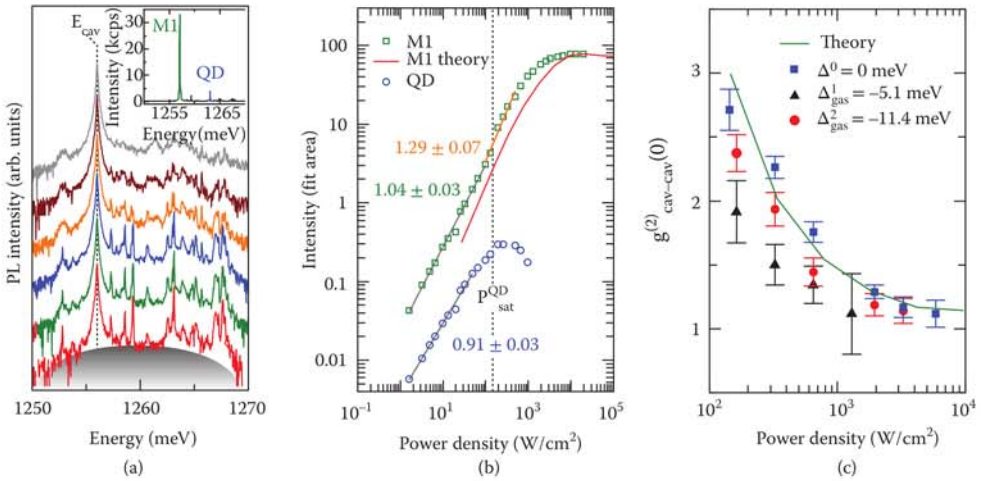


FIGURE 48.21 **(a)** Spectra of a few-quantum dot (QD) photonic-crystal cavity laser for excitation energies ranging from 0.14 to 5.9 kW/cm^2. **(b)** Input–output curves at the cavity-mode energy (squares) and the QD exciton (circles), the latter showing saturation while the cavity-mode emission continues to increase. **(c)** Second-order photon correlation function at the cavity mode for three different detunings of the cavity-mode energy. Irrespective of the detuning between QD transitions and mode, lasing is achieved at sufficiently high excitation. PL, photoluminescence. (Reprinted from Lichtmannecker, S. et al., *Scienti. Rep.*, 7, 2017).

in [99,103,104]. Thus, even if the QD exciton line, which dominates the spectrum at low excitation, is detuned from the cavity mode, a higher transition, e.g., from the biexciton, may be in perfect resonance with the mode and lead to efficient photon production at higher excitation.

There are several approaches to include the non-resonant coupling of detuned emitters in a laser theory, which superficially all have the same effect of producing additional photons in the resonator mode. In their origin they differ fundamentally, which is reflected in the saturation behavior and in the statistical properties of the emission. On the one hand, a "direct" cavity feeding is often used in the literature [105–108], where photons are directly generated in the cavity mode by an inverse cavity loss term. Such a feeding mechanism, acting together with the cavity loss, gives rise to a thermal photonic state with a temperature defined by the cavity-pump-to-loss ratio. This overlooks the fact that the background photons are generated through exciting an active medium, and hence are expected to have a coherent component [89]. On the other hand, detuned emitters can be explicitly included in the calculation, and their coupling to a detuned cavity mode via phonon or Auger coupling is facilitated by an effective system-bath coupling (Lindblad) term.This approach has the advantage that properties of the gain medium are preserved. The resulting cavity feeding is strongly non-linear with pump rate and exhibits saturation if the number of detuned emitters is limited. Furthermore, the overall emission can become coherent if the background contributions drive the laser above the threshold [89].

As an example, the left part of Figure 48.21 [109] shows the spectrum of a few-QD photonic-crystal nanocavity laser. Only few discrete QD lines are found in the vicinity of the mode (M1). The input–output characteristics is that of a nanolaser with a slight kink, hinting at a β-factor close to unity (middle panel), and coherent emission with $g^{(2)}(0) = 1$ being reached (right panel). Remarkably, the performance of the device is insensitive to tuning of the cavity-mode energy, which can be realized by gas deposition on the photonic crystal cavity. Evidence for this is found in the autocorrelation function in the right panel, which shows a clear transition to lasing with $g^{(2)}(0) = 1$ *independently* of the exact cavity position. In a system with few discrete emission channels this provides strong evidence for the importance of nonresonant coupling effects. Similar results have been reported in [110], to which the authors referred to as "self-tuning gain" effect. Moreover, in the spontaneous emission regime the coupling of various multiexcitonic emission channels from different emitters inside the cavity leads to super-thermal photon bunching with $g^{(2)}(0) > 2$. A density-matrix theory including quantum-mechanical correlations between these transitions, shown as green curve in the right panel, is in good agreement with the measured data. Neither rate equations, nor conventional laser models can unravel the intricate quantum-mechanical nature that governs the laser properties of few-emitter cavity-QED systems.

48.9 Outlook

With present-day nanolaser devices, we are closer than ever to reaching fundamental limits of light–matter interaction in semiconductor systems. Electronic excitations can be generated and used on the level of single excitons [111,112]. Brought into interaction with a single mode of microresonator, semiconductor nanolasers operate at the transition to the regime of quantum optics, where cavity-QED effects, such as photon blockade, photon-antibunching, and vacuum Rabi oscillations can coexist with lasing in different excitation regimes of the same device [106,113,114]. This offers fascinating prospects both from a fundamental and an applied point of view.

Acknowledgments

The authors would like to thank Joachim Piprek for the kind invitation to write this chapter. We also thank Andreas Beuthner for his help in creating the schematic figures. Christopher Gies and Frank Jahnke acknowledge financial support from the German Science Foundation (DFG).

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. WWC thanks the hospitality of the Technical University Berlin and travel support from the German Research Foundation via the collaborative research center 787.

References

1. Vahala, K. J. (2003). Optical microcavities. *Nature 424* (6950), 839–846.
2. Borri, P., W. Langbein, S. Schneider, U. Woggon, R. L. Sellin, D. Ouyang, and D. Bimberg (2002). Relaxation and dephasing of multiexcitons in semiconductor quantum dots. *Physical Review Letters 89* (18), 187401.
3. Borri, P., W. Langbein, U. Woggon, V. Stavarache, D. Reuter, and A. D. Wieck (2005). Exciton dephasing via phonon interactions in InAs quantum dots: Dependence on quantum confinement. *Physical Review B 71* (11), 115328.
4. Florian, M., A. Steinhoff, C. Gies, and F. Jahnke (2016). Scattering-induced dephasing of many-particle transitions in semiconductor quantum dots. *Applied Physics B 122* (1), 1–7.
5. Blood, P. (2013). The laser diode: 50 years on. *IEEE Journal of Selected Topics in Quantum Electronics 19* (4), 1503201–1503201.
6. Ning, C. (2013). What is laser threshold? *IEEE Journal of Selected Topics in Quantum Electronics 19* (4), 1503604–1503604.
7. Rice, P. R. and H. J. Carmichael (1994). Photon statistics of a cavity-QED laser: A comment on the laser–phase-transition analogy. *Physical Review A 50* (5), 4318–4329.
8. Chow, W. W., M. Lorke, and F. Jahnke (2011). Will quantum dots replace quantum wells as the active medium of choice in future semiconductor lasers? *IEEE Journal of Selected Topics in Quantum Electronics 17*, 1349–1355.
9. Yokoyama, H. and S. D. Brorson (1989). Rate equation analysis of microcavity lasers. *Journal of Applied Physical 66*, 4801.
10. Musiał, A., C. Hopfmann, T. Heindel, C. Gies, M. Florian, H. A. M. Leymann, A. Foerster, C. Schneider, F. Jahnke, S. Höfling, M. Kamp, and S. Reitzenstein (2015). Correlations between axial and lateral emission of coupled quantum dot micropillar cavities. *Physical Review B 91* (20), 205310.
11. Munsch, M., N. S. Malik, E. Dupuy, A. Delga, J. Bleuse, J.-M. Gérard, J. Claudon, N. Gregersen, and J. Mørk (2013). Dielectric GaAs antenna ensuring an efficient broadband coupling between an InAs quantum dot and a Gaussian optical beam. *Physical Review Letters 110* (17), 177402.
12. Leymann, H. A. M., C. Hopfmann, F. Albert, A. Foerster, M. Khanbekyan, C. Schneider, S. Höfling, A. Forchel, M. Kamp, J. Wiersig, and S. Reitzenstein (2013). Intensity fluctuations in bimodal micropillar lasers enhanced by quantum-dot gain competition. *Physical Review A 87* (5), 053819.
13. Agrawal, G. P. and N. K. Dutta (1993). *Semiconductor Laser*. Norwell, MA: Kluwer.
14. Chuang, S. L. (1995). *Physics of Optoelectronic Devices*. New York, NY: Wiley.
15. Lohmeyer, H., K. Sebald, C. Kruse, R. Kröger, J. Gutowski, D. Hommel, J. Wiersig, N. Baer, and F. Jahnke (2006). Confined optical modes in monolithic II-VI pillar microcavities. *Applied Physics Letters 88* (5), 051101.
16. McCall, S. L., A. F. J. Levi, R. E. Slusher, S. J. Pearton, and R. A. Logan (1992). Whispering gallery mode microdisk lasers. *Applied Physics Letters 60* (3), 289–291.
17. Joannopoulos, J. D., P. R. Villeneuve, and S. Fan (1997). Photonic crystals: Putting a new twist on light. *Nature 386* (6621), 143–149.
18. Purcell, E. M. (1946). Spontaneous emission probability at radio frequencies. *Physical Review 69*, 681.

19. Schwab, M., H. Kurtze, T. Auer, T. Berstermann, M. Bayer, J. Wiersig, N. Baer, C. Gies, F. Jahnke, J. P. Reithmaier, A. Forchel, M. Benyoucef, and P. Michler (2006). Radiative emission dynamics of quantum dots in a single cavity micropillar. *Physical Review B 74* (4), 045323.
20. Aßmann, M., F. Veit, M. Bayer, M. v. d. Poel, and J. M. Hvam (2009). Higher-order photon bunching in a semiconductor microcavity. *Science 325* (5938), 297–300. PMID: 19608912.
21. Loudon, R. (2000). *The Quantum Theory of Light*. Oxford, UK: Oxford University Press.
22. Hanbury Brown, R. and R. Q. Twiss (1956). Correlation between photons in two coherent beams of light. *Nature 177*, 27–29.
23. Ulrich, S. M., C. Gies, J. Wiersig, S. Reitzenstein, C. Hofmann, A. Löffer, A. Forchel, F. Jahnke, and P. Michler (2007). Photon statistics of semiconductor microcavity lasers. *Physical Review Letters 98*, 043906.
24. Gardiner, C. and P. Zoller (2004). *Quantum Noise A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*. Heidelberg, Berlin, New York: Springer Verlag.
25. Gies, C., M. Florian, P. Gartner, and F. Jahnke (2012). Modelling single quantum dots in microcavities. In F. Jahnke (Ed.), *Quantum Optics with Semiconductor Nanostructures*. Philadelphia, PA: Woodhead Publishing Limited.
26. Ates, S., C. Gies, S. M. Ulrich, J. Wiersig, S. Reitzenstein, A. Löffler, A. Forchel, F. Jahnke, and P. Michler (2008). Influence of the spontaneous optical emission factor β on the first-order coherence of a semiconductor microcavity laser. *Physical Review B 78* (15), 155319.
27. Chow, W. W., F. Jahnke, and C. Gies (2014). Emission properties of nanolasers during the transition to lasing. *Light: Science & Applications 3* (8), e201.
28. Ritter, S., P. Gartner, C. Gies, and F. Jahnke (2010). Emission properties and photon statistics of a single quantum dot laser. *Optics Express 18* (10), 9909–9921.
29. Reitzenstein, S., C. Böckler, A. Bazhenov, A. Gorbunov, A. Löffer, M. Kamp, V. D. Kulakovskii, and A. Forchel (2008). Single quantum dot controlled lasing effects in high-q micropillar cavities. *Optics Express 16* (7), 4848–4857.
30. Gies, C., J. Wiersig, and F. Jahnke (2008). Output characteristics of pulsed and continuous-wave-excited quantum-dot microcavity lasers. *Physical Review Letters 101* (6), 067401.
31. Chow, W. and S. Koch (2005). Theory of semiconductor quantum-dot laser dynamics. *IEEE Journal of Quantum Electronics 41* (4), 495–505.
32. Kolarczik, M., N. Owschimikow, J. Korn, B. Lingnau, Y. Kaptan, D. Bimberg, E. Schöll, K. Lüdge, and U. Woggon (2013). Quantum coherence induces pulse shape modification in a semiconductor optical amplifier at room temperature. *Nature Communications 4*, 2953.
33. Gies, C., J. Wiersig, M. Lorke, and F. Jahnke (2007). Semiconductor model for quantum-dot-based microcavity lasers. *Physical Review A 75* (1), 013803.
34. Wiersig, J., C. Gies, F. Jahnke, M. Aßmann, T. Bestermann, M. Bayer, C. Kistner, S. Reitzenstein, S. H. A. Forchel, C. Kruse, J. Kalden, and D. Hommel (2009). Direct observation of correlations between individual photon emission events of a microcavity laser. *Nature 460*, 245.
35. Baer, N., C. Gies, J. Wiersig, and F. Jahnke (2006). Luminescence of a semiconductor quantum dot system. *European Physical Journal B 50*, 411.
36. Gregersen, N., T. Suhr, M. Lorke, and J. Mørk (2012). Quantum-dot nano-cavity lasers with Purcell-enhanced stimulated emission. *Applied Physics Letters 100* (13), 131107.
37. Lermer, M., N. Gregersen, M. Lorke, E. Schild, P. Gold, J. Mørk, C. Schneider, A. Forchel, S. Reitzenstein, S. Hing, and M. Kamp (2013). High beta lasing in micropillar cavities with adiabatic layer design. *Applied Physics Letters 102* (5), 052114.
38. Lorke, M., T. Suhr, N. Gregersen, and J. Mørk (2013). Theory of nanolaser devices: Rate equation analysis versus microscopic theory. *Physical Review B 87*, 205310.

39. Suhr, T., N. Gregersen, M. Lorke, and J. Mørk (2011). Modulation response of quantum dot nanolightemitting-diodes exploiting Purcell-enhanced spontaneous emission. *Applied Physics Letters 98* (21), 211109.
40. Chow, W. and S. Koch (1999). *Semiconductor-Laser Fundamentals* (1st ed.). Berlin: Springer-Verlag.
41. Coldren, L. and S. Corzine (1995). *Diode Lasers and Photonic Integrated Circuits*. New York, NY: Wiley.
42. Binder, R., H. Köhler, M. Bonitz, and N. Kwong (1997). Green's function description of momentum orientation relaxation of photoexcited electron plasmas in semiconductors. *Physical Review B 55*, 5110.
43. Bonitz, M., D. Kremp, D. Scott, R. Binder, W. Kraeft, and H. Köhler (1996). Numerical analysis of nonmarkovian effects in charge-carrier scattering: One-time kinetic equations. *Journal of Physics: Condensed Mattter 8*, 6057.
44. ElSayed, K., S. Schuster, H. Haug, G. Herzel, and K. Henneberger (1994). Subpicosecond plasmon response: Buildup of screening. *Physical Review B 49*, 7337.
45. Henneberger, K. (1988). Resonant laser excitation and electron-hole kinetics of a semiconductor: I. nonequilibrium green's function treatment and fundamental equations. *Physica A 150*, 419.
46. Jahnke, F. and S. W. Koch (1995). Many-body theory for semiconductor microcavity lasers. *Physical Review A 52*, 1712.
47. Bányai, L., P. Gartner, and H. Haug (1998). Self-consistent RPA retarded polaron Green function for quantum kinetics. *European Physical Journal B 1*, 209.
48. Gartner, P., L. Bányai, and H. Haug (1999). Two-time electron-LO-phonon quantum kinetics and the generalized Kadanoff-Baym approximation. *Physical Review B 60*, 14234.
49. Gartner, P., L. Bányai, and H. Haug (2002). Self-consistent RPA for the intermediate-coupling polaron. *Physical Review B 66*, 75205.
50. Hartmann, M. and W. Schäfer (1992). Real time approach to relaxation and dephasing processes in semiconductors. *Physical Status Solidi (B) 173*, 165.
51. Haug, H. (1992). Interband quantum kinetics with LO-phonon scattering in a laser-pulse-excited semiconductor I. Theory. *Physical Status Solidi (B) 173*, 139.
52. Jahnke, F. and S. W. Koch (1993). Theory of carrier heating through injection pumping and lasing in semiconductor microcavity lasers. *Optics Letters 18* (17), 1438–1440.
53. Lorke, M., T. R. Nielsen, and J. Mørk (2011). Switch-on dynamics of nanocavity laser devices. *Applied Physics Letters 99* (15), 151110.
54. Lingnau, B., K. Lüdge, E. Schöll, and W. W. Chow (2010). Many-body and nonequilibrium effects on relaxation oscillations in a quantum-dot microcavity laser. *Applied Physics Letters 97*, 111102.
55. Lorke, M., T. R. Nielsen, and J. Mørk (2010). Influence of carrier dynamics on the modulation bandwidth of quantum-dot based nanocavity devices. *Applied Physics Letters 97* (21), 211106.
56. Luedge, K., M. J. P. Bormann, E. Malic, P. Hoevel, M. Kuntz, D. Bimberg, A. Knorr, and E. Schoell (2008). Turn-on dynamics and modulation response in semiconductor quantum dot lasers. *Physical Review B 78*, 035316.
57. Luedge, K. and E. Scholl (2009). Quantum-dot lasers—Desynchronized nonlinear dynamics of electrons and holes. *IEEE Journal of Quantum Electronics 45* (11), 1396–1403.
58. Aßmann, M., F. Veit, M. Bayer, C. Gies, F. Jahnke, S. Reitzenstein, S. Höfling, L. Worschech, and A. Forchel (2010). Ultrafast tracking of second-order photon correlations in the emission of quantum-dot microresonator lasers. *Physical Review B 81* (16), 165314.
59. Lorke, M., W. Chow, T. R. Nielsen, J. Seebeck, P. Gartner, and F. Jahnke (2006). Anomaly in the excitation dependence of optical gain in semiconductor quantum dots. *Physical Review B 74*, 035334.
60. Lorke, M., T. Nielsen, J. Seebeck, P. Gartner, and F. Jahnke (2006). Influence of carrier-carrier and carrier-phonon correlations on optical absorption and gain in quantum-dot systems. *Physical Review B 73*, 85324.

61. Schäfer, W. and M. Wegener (2002). *Semiconductor Optics and Transport Phenomena* (1st ed.). Berlin: Springer-Verlag.
62. Mahan, G. (1990). *Many-Particle Physics*. New York, NY: Plenum Press.
63. Seebeck, J., T. Nielsen, P. Gartner, and F. Jahnke (2005). Polarons in semiconductor quantum dots and their role in the quantum kinetics of carrier relaxation. *Physical Review B 71*, 125327.
64. Nielsen, T. R., P. Gartner, and F. Jahnke (2004). Many-body theory of carrier capture and relaxation in semiconductor quantum-dot lasers. *Physical Review B 69* (23), 235314.
65. Benisty, H. (1995). Reduced electron-phonon relaxation rates in quantum-box systems: Theoretical analysis. *Physical Review B 51*, 13281.
66. Inoshita, T. and H. Sakaki (1996). Electron-phonon interaction and the so-called phonon bottleneck effect in semiconductor quantum dots. *Physica B 227*, 373.
67. Kurtze, H., J. Seebeck, P. Gartner, D. R. Yakovlev, D. Reuter, A. D. Wieck, M. Bayer, and F. Jahnke (2009). Carrier relaxation dynamics in self-assembled semiconductor quantum dots. *Physical Review B 80*, 235319.
68. Schuh, K., F. Jahnke, and M. Lorke (2011). Rapid adiabatic passage in quantum dots: Influence of scattering and dephasing. *Applied Physics Letters 99* (1), 011105.
69. Seebeck, J., T. Nielsen, P. Gartner, and F. Jahnke (2006). Quantum kinetic theory of phonon-assisted carrier transitions in nitride-based quantum-dot systems. *European Physical Journal B 49*, 167.
70. Seebeck, J., M. Lorke, P. Gartner, and F. Jahnke (2009). Carrier-carrier and carrier-phonon scattering in the low-density and low-temperature regime for resonantly pumped semiconductor quantum dots. *Physica Status Solidi (C) 6*, 488.
71. Uskov, A. V., C. Meuer, H. Schmeckebier, and D. Bimberg (2011). Auger capture induced carrier heating in quantum dot lasers and amplifiers. *Applied Physics Express 4* (2), 022202.
72. Goldmann, E., M. Lorke, T. Frauenheim, and F. Jahnke (2014). Negative differential gain in quantum dot systems: Interplay of structural properties and many-body effects. *Applied Physics Letters 104* (24), 242108.
73. Lorke, M., W. Chow, and F. Jahnke (2007). Reduction of optical gain in semiconductor quantum dots. *Proceedings of SPIE 6468*, 646818.
74. Lorke, M., F. Jahnke, and W. Chow (2007). Excitation dependence of gain and carrier induced refractive index changes in quantum dots. *Applied Physical Letters 90*, 51112.
75. Lorke, M., J. Seebeck, T. Nielsen, P. Gartner, and F. Jahnke (2006). Excitation dependence of the homogeneous linewidths in quantum dots. *Physical Status Solidi (C) 3*, 2393–2396.
76. Dicke, R. H. (1954). Coherence in spontaneous radiation processes. *Physical Review 93*, 99–110.
77. Mandel, L. and E. Wolf (1995). *Optical Coherence and Quantum Optics*. New York, NY: Cambridge University Press.
78. Bohnet, J. G., Z. Chen, J. M. Weiner, D. Meiser, M. J. Holland, and J. K. Thompson (2012). A steady-state superradiant laser with less than one intracavity photon. *Nature 484* (7392), 78–81.
79. Timothy Noe Ii, G., J.-H. Kim, J. Lee, Y. Wang, A. K. Wójcik, S. A. McGill, D. H. Reitze, A. A. Belyanin, and J. Kono (2012). Giant superfluorescent bursts from a semiconductor magneto-plasma. *Nature Physics 8* (3), 219–224.
80. Leymann, H. A. M., A. Foerster, F. Jahnke, J. Wiersig, and C. Gies (2015). Sub—And superradiance in nanolasers. *Physical Review Applied 4* (4), 044018.
81. Auffèves, A., D. Gerace, S. Portolan, A. Drezet, and M. F. Santos (2011). Few emitters in a cavity: From cooperative emission to individualization. *New Journal of Physics 13* (9), 093020.
82. Jahnke, F., C. Gies, M. Aßmann, M. Bayer, H.A.M. Leymann, A. Foerster, J. Wiersig, C. Schneider, M. Kamp, and S. Höfling (2016). Giant photon bunching, superradiant pulse emission and excitation trapping in quantum-dot nanolasers. *Nature Communications 7*, 11540.
83. Bimberg, D., M. Grundmann, and N. N. Ledentsov (1998). *Quantum Dot Heterostructures*. Hoboken, NJ: John Wiley & Sons.

84. Bayer, M. and A. Forchel (2002). Temperature dependence of the exciton homogeneous linewidth in In0.60Ga0.40As/GaAs self-assembled quantum dots. *Physical Review B 65* (4), 041308.
85. Moody, G., R. Singh, H. Li, I. A. Akimov, M. Bayer, D. Reuter, A. D. Wieck, and S. T. Cundiff (2013). Fifth-order nonlinear optical response of excitonic states in an InAs quantum dot ensemble measured with two-dimensional spectroscopy. *Physical Review B 87* (4), 045313.
86. Scheibner, M., T. Schmidt, L. Worschech, A. Forchel, G. Bacher, T. Passow, and D. Hommel (2007). Superradiance of quantum dots. *Nature Physics 3* (2), 106–110.
87. Lichtmannecker, S., M. Florian, T. Reichert, M. Blauth, M. Bichler, et al. (2017). A few-emitter solid-state multi-exciton laser. *Scientific Reports 7* (7420). doi: 10.1038/s41598-017-07097-9
88. Ates, S., S. M. Ulrich, A. Ulhaq, S. Reitzenstein, A. Löffler, S. Höfling, A. Forchel, and P. Michler (2009). Non-resonant dot—cavity coupling and its potential for resonant single-quantum-dot spectroscopy. *Nature Photonics 3* (12), 724–728.
89. Florian, M., P. Gartner, C. Gies, and F. Jahnke (2013). Phonon-mediated off-resonant coupling effects in semiconductor quantum-dot lasers. *New Journal of Physics 15* (3), 035019.
90. Hohenester, U., A. Laucht, M. Kaniber, N. Hauke, A. Neumann, A. Mohtashami, M. Seliger, M. Bichler, and J. J. Finley (2009). Phonon-assisted transitions from quantum dot excitons to cavity photons. *Physical Review B 80* (20), 201311.
91. Majumdar, A., E. D. Kim, Y. Gong, M. Bajcsy, and J. Vučković (2011). Phonon mediated off-resonant quantum dot–cavity coupling under resonant excitation of the quantum dot. *Physical Review B 84* (8), 085309.
92. Press, D., S. Götzinger, S. Reitzenstein, C. Hofmann, A. Löffler, M. Kamp, A. Forchel, and Y. Yamamoto (2007). Photon antibunching from a single quantum-dot-microcavity system in the strong coupling regime. *Physical Review Letters 98* (11), 117402.
93. Tarel, G. and V. Savona (2009). Emission spectrum of a quantum dot embedded in a nanocavity. *Physica Status Solidi (C) 6* (4), 902–905.
94. Roy, C. and S. Hughes (2011). Influence of electron–acoustic-phonon scattering on intensity power broadening in a coherently driven quantum-dot–cavity system. *Physical Review X 1* (2), 021009.
95. Chauvin, N., C. Zinoni, M. Francardi, A. Gerardino, L. Balet, B. Alloing, L. H. Li, and A. Fiore (2009). Controlling the charge environment of single quantum dots in a photonic-crystal cavity. *Physical Review B 80* (24), 241306.
96. Florian, M., P. Gartner, A. Steinhoff, C. Gies, and F. Jahnke (2014). Coulomb-assisted cavity feeding in nonresonant optical emission from a quantum dot. *Physical Review B 89* (16), 161302(R).
97. Karrai, K., R. J. Warburton, C. Schulhauser, A. Högele, B. Urbaszek, E. J. McGhee, A. O. Govorov, J. M. Garcia, B. D. Gerardot, and P. M. Petroff (2004). Hybridization of electronic states in quantum dots through photon emission. *Nature 427* (6970), 135–138.
98. Settnes, M., P. Kaer, A. Moelbjerg, and J. Mork (2013). Auger processes mediating the nonresonant optical emission from a semiconductor quantum dot embedded inside an optical cavity. *Physical Review Letters 111* (6), 067403.
99. Winger, M., T. Volz, G. Tarel, S. Portolan, A. Badolato, K. J. Hennessy, E. L. Hu, A. Beveratos, J. Finley, V. Savona, and A. Imamoğlu (2009). Explanation of photon correlations in the far-off-resonance optical emission from a quantum-dot–cavity system. *Physical Review Letters 103* (20), 207403.
100. Bockelmann, U. and T. Egeler (1992). Electron relaxation in quantum dots by means of auger processes. *Physical Review B 46*, 15574.
101. Efros, A. L., V. A. Kharchenko, and M. Rosen (1995). Breaking of the phonon bottleneck in nanometer quantum dots role of auger-like processes. *Solid State Communications 93*, 281.
102. Uskov, A. V., F. Adler, H. Schweizer, and M. H. Pilkuhn (1997). Auger carrier relaxation in self-assembled quantum dots by collisions with two-dimensional carriers. *Journal of Applied Physics 81*, 7895.

103. Dekel, E., D. Gershoni, E. Ehrenfreund, D. Spektor, J. M. Garcia, and P. M. Petroff (1998). Multiexciton spectroscopy of a single self-assembled quantum dot. *Physical Review Letters 80* (22), 4991–4994.
104. Laucht, A., M. Kaniber, A. Mohtashami, N. Hauke, M. Bichler, and J. J. Finley (2010). Temporal monitoring of nonresonant feeding of semiconductor nanocavity modes by quantum dot multiexciton transitions. *Physical Review B 81*, 241302.
105. del Valle, E., F. P. Laussy, and C. Tejedor (2009). Luminescence spectra of quantum dots in microcavities. II. Fermions. *Physical Review B 79*, 235326.
106. Nomura, M., N. Kumagai, S. Iwamoto, Y. Ota, and Y. Arakawa (2010). Laser oscillation in a strongly coupled single-quantum-dot–nanocavity system. *Nature Physics 6*, 279–283.
107. Valle, E. d. and F. P. Laussy (2011). Effective cavity pumping from weakly coupled quantum dots. *Superlattices and Microstructures 49* (3), 241–245.
108. Yao, P., P. K. Pathak, E. Illes, S. Hughes, S. Münch, S. Reitzenstein, P. Franeck, A. Löffer, T. Heindel, S. Höfling, L. Worschech, and A. Forchel (2010). Nonlinear photoluminescence spectra from a quantum-dot cavity system: Interplay of pump-induced stimulated emission and anharmonic cavity QED. *Physical Review B 81* (3), 033309.
109. Lichtmannecker, S., M. Florian, T. Reichert, M. Blauth, M. Bichler, F. Jahnke, J. J. Finley, C. Gies, and M. Kaniber (2016). A few-emitter solid-state multi-exciton laser. *ArXiv:1602.03998 [condmat]*.
110. Strauf, S., K. Hennessy, M. T. Rakher, Y.-S. Choi, A. Badolato, L. C. Andreani, E. L. Hu, P. M. Petroff, and D. Bouwmeester (2006). Self-tuned quantum dot gain in photonic crystal lasers. *Physical Review Letters 96*, 127404.
111. Mlynek, J. A., A. A. Abdumalikov, C. Eichler, and A. Wallraff (2014). Observation of Dicke superradiance for two artificial atoms in a cavity with high decay rate. *Nature Communications 5*, 5186.
112. Müller, M., S. Bounouar, K. D. Jöns, M. Glässl, and P. Michler (2014). On-demand generation of indistinguishable polarization-entangled photon pairs. *Nature Photonics 8* (3), 224–228.
113. Muñoz, C. S., E. del Valle, A. G. Tudela, K. Müller, S. Lichtmannecker, M. Kaniber, C. Tejedor, J. J. Finley, and F. P. Laussy (2014). Emitters of n-photon bundles. *Nature Photonics 8* (7), 550–555.
114. Gies, C. et al., Phys. Rev. A **96**, 023806 (2017).

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com