49

Nonlinear Dynamics in Quantum Photonic Structures

49.1 Introduction 662
49.2 Maxwell-Pseudospin Method 667
Pseudospin Equation • Dissipative Dynamics and Master Pseudospin Equation • Vector Maxwell Equations and Coupling to Pseudospin Equations • Coherent Maxwell-Pseudospin Equations in 2D • Numerical Methodology
49.3 Simulation of Nonlinear Dynamics in Optical Waveguides and Semiconductor Microcavities 674
One-Dimensional Model of a Two-Level System • Passive Cavity Properties • Active Nonlinear Optical Waveguide/Cavity Properties • Resonant Nonlinearities and Gain Saturation • Coherent Propagation Effects, Self-Localization and Pattern Formation in 2D Planar Nonlinear Optical Waveguides and Semiconductor Microcavities
49.4 Optical Orientation in Quantum Wells and QDs 691
Circularly Polarized Optical-Pulse Interactions with Two-Level Quantum Systems: Selective Spin Initialization and Readout • Circularly Polarized Optical Pulse Interactions with Degenerate Four-Level Systems
49.5 Modeling Coherent Spin Generation, Manipulation, and Readout in Charged QDs 699
Single-Charged QD Embedded in a Waveguide • Coherent Spin Manipulation through Hot-Charged Exciton States in QDs
49.6 Quantum Stochastic Formalism for Modeling Cavity-Emitter Systems.. 707
Langevin Formalism for Modeling Quantum-Optical Effects • Spontaneous Emission of a QD Ensemble in a Semiconductor Microcavity and Onset of Lasing
49.7 Modeling Optical Rotation in Chiral Carbon Nanotubes with Master Maxwell-Pseudospin Equations......................... 712
Natural Optical Activity • Faraday Rotation and Magneto-Chiral Effects
49.8 Conclusion... 724

Gabriela Slavcheva
and
Mirella Koleva

49.1 Introduction

Quantum optics unites physical optics and the quantum field theory of light. The foundations of the latter were laid by Planck, who succeeded to explain theoretically the black-body radiation spectrum by postulating that the energy of a harmonic oscillator is quantized. In 1927, Dirac developed this theory by solving the long-standing problem of the wave-particle duality and unifying the wave and particle aspects of light by quantization of the electromagnetic radiation. Thus, according to Dirac's single-particle interpretation, the light emitted from a spontaneously emitting source with low photon density is described as single-photon emission events.

A fundamental consequence of light-field quantization is the appearance of vacuum (zero-point energy) fluctuations, or quantum noise, which has no classical analogue and stems from the Heisenberg uncertainty relations. However, some of the greatest successes of quantum theory—such as Planck's black-body spectrum and Einstein's theory of the photoelectric effect—do not actually require a full quantum-mechanical treatment and can instead be explained using semiclassical theory, whereby matter is quantized but light remains a classical wave. Only in the 1960s, motivated by the development of lasers, did researchers start to examine effects which truly required a quantum theory of light. For instance, light quantization is required to explain spontaneous emission, irreversible decay, and the origin of the laser spectral linewidth. The demand to develop further the quantum theory of light in order to explain these effects was met by the pioneering work of Glauber [1,2].

Vacuum quantum fluctuations are necessary to account for phenomena, such as the Lamb shift: a radiative correction to the energy spectrum of hydrogen, as well as for the attractive force between two uncharged conductive plates in a vacuum placed a few nanometres apart, known as the Casimir effect. Some examples of effects which the semiclassical approach fails to explain are quantum beat phenomena in a three-level atom, two-photon interferometry and production of "entangled" states of light, resonance fluorescence (quantum jumps) in the emission of atoms, and the so-called "nonclassical" states of light.

The classification of light states as classical and nonclassical stems from the statistical properties of radiation and, more specifically, from the second-order coherent properties of the light emitter. A crucial step in elucidating these properties was the correlation experiment of Hanbury Brown and Twiss [3], which was designed to measure the coincidence of individual photons at two detectors placed equidistantly from a half-silvered mirror splitting a beam of oncoming low-intensity light. Using the second-order correlation function, $g^{(2)}(\tau)$, quantifying the synchronicity with which photons arrive at the detectors, it was then possible to discern the manner in which different types of light sources emitted light. This method revealed that classical light sources—such as a thermal black-body radiation source or a lamp—exhibited Bose–Einstein statistics manifested by photon "bunching", i.e., an increased likelihood of the photons to arrive at the detectors at the same time, as though they had "bunched together" while traveling from the source to the detectors. On the other hand, a stable coherent light source (e.g., a laser) exhibits Poissonian statistics, which is manifested by a constant unity second-order correlation function, $g^{(2)}(\tau)$, meaning that photons are not intercorrelated and are equally likely to impinge on the detectors with a delay of any length with respect to each other. However, most interestingly of all, nonclassical states of light are described by sub-Poissonian statistics (a diminished likelihood of coincident arrival) called "antibunching": the photons are "diluted" as they travel and eventually emitted one by one. Antibunching is characteristic of resonant fluorescence from single atoms, molecules, or semicondictor nanocrystals and quantum dots (QDs), and is a signature of a single-photon emitter. For perfectly antibunched light, the second-order correlation function vanishes for coincident detections. This is irreconcilable with the classical wave optics and is considered a hallmark of pure quantum behavior.

Photon-number (Fock) states are states of light containing a specific number of photons. The single-photon state is a special case of a Fock state and is a squeezed-light state [4], a quantum state of light resulting from the Heisenberg relations with, e.g., reduced amplitude fluctuations at the expense of enhanced fluctuations of the phase. An ideal amplitude-squeezed source delivers a stream of photons at regular time intervals, which is the precise realization of a single-photon source [5].

Another quantum-mechanical property is entanglement. Quantum systems consisting of interacting subsystems become highly correlated and their individual constituents become "entangled". The global quantum state of an entangled system—for e.g., one consisting of two particles (e.g., a photon pair)—remains correlated even when the two particles are separated by large distances from each other and unable to interact or communicate in any known way. The most famous example of an entangled state is the Einstein–Podolsky–Rosen–Bohm [6] state, whereby two spin 1/2 particles reside in a state with zero total spin angular momentum: $|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle)$. Notably, this state cannot be represented as a simple product of the pure states of the individual particles.

Many of these quantum effects have been observed only in carefully controlled atomic systems; however, with the advent of quantum photonics (the applied device-oriented field of quantum optics), recent advances in semiconductor technology have enabled the fabrication of solid-state devices that emit single photons, opening new avenues for nonclassical light state generation. For instance, the use of radiative cascades from QD excited states is a principal method for generation of quantum entanglement for applications of QDs as polarization-entangled sources. The exchange interaction and the fine structure splitting of the energy levels involved in radiative cascades from QD excited states are of fundamental importance, since they determine the polarization and entanglement of the emitted photons.

Quantum photonics has attracted great interest in recent years. This is due to its potential to revolutionize science and day-to-day life alike through enabling the implementation of faster algorithms, secure transmission of information, vast increase in data storage, and execution of more accurate measurements. One of the major challenges of modeling quantum photonic devices is working out a way of efficient simulation and control of realistic, "open" quantum systems (i.e., allowing for energy to flow in and out of the system) and devices for applications in quantum technologies. For optical quantum information processing to advance beyond demonstration experiments, the development of on-chip capabilities will be essential. Solid-state qubit architectures have emerged as a most promising physical implementation of quantum photonic circuits, taking advantage of the state-of-the-art semiconductor technologies. Reliable, integrated sources of nonclassical light, as well as realization of high-fidelity on-chip readout of solid-state qubits, are a critical requirement for next-generation quantum technologies [7].

Recently, significant progress has been made toward the achievement of high-efficiency solid-state single-photon sources [8–12]. Among these sources, QDs are particularly promising since they can be embedded in devices enabling electrical injection while maintaining efficient light emission. Furthermore, the dot doping—or the number and type of the charge carriers within the dot—can be controlled by applying electrical bias in a Schottky-type structure. A quantum bit (qubit) is a linear superposition of basis states $|0\rangle$ (ground level) and $|1\rangle$ (excited level) in a two-level system, given by: $|\Psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$, with α and β being complex numbers such that $\alpha^2 + \beta^2 = 1$. In a classical computer, the information is encoded in digital bit states 0 and 1. By contrast, in a quantum computer, the state of each bit is permitted to be any quantum-mechanical state of a qubit (manipulated using qubit rotations, also known as gate operations), leading to a massive parallelism and an exponential speed-up of the quantum computations.

The $|0\rangle$ and $|1\rangle$ states of a qubit could be encoded in the spin-up ($|\uparrow\rangle$) and spin-down ($|\downarrow\rangle$) states of a charge carrier. The spin-up and spin-down states of a single electron or hole confined in a charged semiconductor QD are a particularly attractive qubit candidate [13]. This arises from several advantages intrinsic to this quantum system. The first advantage is that the atomic-like electronic structure suppresses coupling of the spin to the solid-state quantum environment and thus impedes the spin decoherence processes. This leads to a very long spin coherence time (denoted T_2)—that is, the time within which the spin loses its phase information—and thus a long-lived quantum state, which is required for performing a large number of spin manipulations. Another important advantage is optical convertibility—i.e., the ability to switch through transfer of angular momentum between a "flying" qubit encoded in a photon and the stationary matter qubit of the spin confined in a QD. This enables the transport of qubits over large distances and facilitates the realization of scalable solid-state qubit architectures for optical quantum computing. Furthermore, this promising platform allows integrating the operations of optical initialization, manipulation, and readout of a spin qubit on a chip. By employing a combination of optical orientation [14] and quantum coherent control [15] techniques, using ultrashort polarized light pulses one can address individual carrier spins in semiconductor QDs, initialize and manipulate them coherently through optically excited states (charged excitons), and detect the resulting spin qubit [16]. For quantum technologies applications, control of both QD position and emission energy is crucial. For instance, deterministic site control of QD location is indispensable for multiqubit operations and scalability. Therefore, new simulation tools are needed to determine the allowed tolerances in the dot positions and emission energies to ensure observation of the desired quantum-optical effects.

Quantum coherent control represents a universal approach for predictable manipulation of the properties of quantum systems, such as atoms and molecules. This technique has most recently been applied to QD solid-state systems. Coherent control in semiconductor nanostructures allows for coherent manipulation of the carrier wavefunctions on a time scale shorter than typical dephasing times. This is a prerequisite for successful implementation of ultrafast optical switching and quantum information processing. Quantum coherent control requires use of ultrashort pulses considerably shorter than the characteristic relaxation times in matter ($\tau \ll T_1, T_2$). This is equivalent to the photon-dipole coupling rate exceeding all dissipative rates in the system, a general condition for achieving the strong-coupling regime. The ultrashort pulses are usually characterized by high field amplitudes and, consequently, lead to nonlinear optical effects, such as coherent pulse propagation and self-induced transparency (SIT) [17]. The phenomenon of SIT is observable above a critical power threshold for a given pulse width: a high-intensity, ultrashort pulse propagating through a medium composed of an ensemble of resonant quantum two-level absorbers, whose relaxation times greatly exceed the pulse duration, travels at a reduced speed and unchanged shape with anomalously low energy loss. The absorbers are driven into the excited state by absorbing ultrashort pulse energy; by reradiating this energy into the pulse, they return to the ground state. Thus, the optical energy is carried through the medium not by the electromagnetic field, but by a coupled light-matter polariton wave. The polariton is a mixed light-matter quasiparticle resulting from the strong coupling of the optical wave to the medium's polarization. As a result, the pulse travels as a solitary wave, known as a SIT-soliton. This soliton is localized both in space and time, in contrast with the well-known nonlinear optics solitons which result from the interplay between nonlinearity and the medium's dispersion and/or diffraction. The condition of SIT is predicted by the remarkable pulse-area theorem (PAT) [18] which establishes a general criterion for stable, ultrashort pulse propagation in attenuating or amplifying media based on an integral quantity—namely, the pulse area. This phenomenon can be preserved to a great extent in solid-state systems—e.g., semiconductors—and has been experimentally demonstrated by picosecond pulse propagation in QD waveguides [19].

When the laser pulse's temporal width becomes comparable with the optical period (i.e., when the pulse envelope contains only a few optical cycles of the carrier wave), a transition to a qualitatively new "extreme nonlinear" regime of strong light-matter interactions is induced in which the electric field itself, rather than intensity envelope, drives the interaction. The majority of the current analytical approaches are based on the standard, slowly varying amplitude approximation (SVEA) and rotating-wave approximation (RWA). However, a number of works have demonstrated the limitations of these approximations, and new phenomena have been predicted on the basis of a non-perturbative approach (see, e.g., [20], and references therein). This requires development of new, non-perturbative theoretical and computational methods beyond the above approximations as shown in Section 49.2, which could account properly for the ultrafast carrier-wave dynamics.

The earlier high-intensity (i.e., high photon density) limit allows to achieve the strong-coupling regime in free space. However, in the opposite, low-intensity (i.e., few-photon) limit, the light-matter coupling of a single photon to a single atom in free space is weak, and despite attempts by several groups [21–23] to improve it, the experimental realization of high-fidelity excited-state preparation remains challenging. In order to improve the performance of single-photon sources, it is necessary to incorporate the emitters in cavities. The cavity helps to select a single mode out of the infinite number of radiation field modes coupled to the emitter. As the emission is preferentially directed into a single mode, owing to the Purcell effect the spontaneous emission lifetime is reduced for emitters located at the antinode of the electric-field intensity in cavities with a small mode volume and high quality factor. Faster emission is desired because it makes the single-photon source more deterministic and reliable.

If the exciton-light coupling in a semiconductor microcavity is weak compared to all dissipative rates in the system, irreversible decay occurs—that is, energy cannot be alternately passed from light to matter. Depending on whether the quantum wells inside the semiconductor microcavity are situated near the node or antinode of the excitatory electric field, either an inhibited or enhanced spontaneous emission takes place. The cavity then operates in the weak coupling regime, whereby the dynamics of the light-matter interaction is dominated by incoherent processes and damping. Conversely, in the strong coupling limit, no irreversible decay takes place. Instead, a sequence of coherent reabsorption and reemission processes occurs and the energy oscillates between the exciton and photon modes, a phenomenon known as optical Rabi oscillations. The spectral signature of strong coupling is an anticrossing between the dot line (i.e., the spectral linewidth of the QD) and the cavity mode. The eigenmodes of the system are mixed entangled light-matter states, giving rise to quasiparticles known as microcavity exciton–polaritons which exhibit bosonic properties. Considerable scientific effort worldwide has been invested in exploration and control of a strongly coupled single QD in a cavity. The latter is considered a solid-state analogue of the atom-cavity system for which many quantum-optical effects have been demonstrated. Some examples include cavity field quantization [24] and photon blockade [25]. Embedding QDs in confined geometries—such as planar [26], micropillar [27], microdisk [28], microsphere [29], nanowire [30], or photonic crystal nanocavities [31]—provides a means for controlling the light-matter interactions. However, unlike isolated atoms, QDs are embedded in a solid-state environment, so confined charge carriers may experience additional interactions with phonons and other neighboring states and charges. These interactions may result in a departure from the behavior of an idealized two-level quantum system, but can also be exploited as additional control measures.

Cavity quantum electrodynamics (CQED), traditionally applied to atom-cavity systems, has gained new momentum after recent reports of quantum-optical experiments in solid-state systems demonstrating nonclassical light behavior, such as nonclassical photon statistics, nonlocality, and quantum entanglement [32,33]. These cutting-edge studies put to the test the very foundations of quantum mechanics and open up new avenues for on-demand generation and exploitation of various nonclassical states of light. So far, close to perfect single-photon interaction with a QD has been achieved in a cavity under the strong coupling regime [34,35].

A major goal of the simulation of quantum-photonic devices is to find an optimum design for combining photonic and solid-state devices on a chip. Integrated optical technologies for quantum computing based on linear optics [36] have been under active development in the recent decade. However, nonlinear optical interactions “on a chip” offer new, unexplored functionalities, and may play a key role in future technologies. In this respect, a major theoretical and computational challenge is the development of new, tractable models of the nonlinear quantum dynamics in realistic devices with complex geometries in multiple dimensions. This chapter reviews the evolution of the original methodology into our fully non-perturbative, mesoscopic, Maxwell’s curl-pseudospin equation-based model of the optically induced spatiotemporal dynamics in quantum-photonic devices. This chapter contains further details about its numerical implementation and some concrete applications.

As we shall show in Section 49.2, our theoretical and computational approach differs substantially from the methods described in Chapter 46 of this book. In that chapter, a number of different approaches for calculation of the spontaneous emission have been employed. In the weak-coupling regime, a classical dipole model is developed whereby the dyadic Green’s function is calculated and the spontaneous emission rate is found from Fermi’s Golden Rule. Analytical expressions for Green’s function are obtainable only for a limited number of device geometries, and more generally one has to resort to numerical methods (e.g., the finite-element method) to compute the electric field. This entails expanding the fields in the eigenmodes in different layers of the specific geometry under consideration. Another approach for modeling cavity-emitter systems is CQED, which is based on the Jaynes–Cummings Hamiltonian, describing the interaction of a two-level system strongly coupled to a single mode of an electromagnetic field. Two approximations are made: (1) for a very low Q cavity, the optical field operators can be replaced with material excitation operators; the dissipation in the system (spontaneous emission rate and dephasing) is taken into account through Lindblad operators. Using the quantum regression theorem that is valid for Markovian processes, an analytical expression for the correlation function is obtained; (2) for a coupled cavity-emitter system, a more complicated master equation is written with an additional dissipation term describing the cavity photon loss. However, for losses weaker than the cavity-emitter coupling, the cavity can be adiabatically eliminated and this results in a master equation describing the emitter only but with modified spontaneous emission rates. By contrast, we use the full master equation for the system density matrix which is valid for all regimes: weak, intermediate, and strong coupling. In addition, in our case the longitudinal and transverse dissipation are taken into account not through Lindblad operators, but by considering the dynamics of population transfer and decoherence between the discrete levels. Another difference between our model presented in Section 49.2 and the model of Gregersen et al. is that in our case, the cavity-emitter (Rabi) coupling is not a model parameter, but is calculated from the initial pulse amplitude, duration, and dipole moment of the specific optical transition, as in the actual experiment. Furthermore, the cavity loss is not a phenomenological parameter either, as it is calculated numerically through the perfectly transparent boundary conditions of the specific cavity geometry. The master equation is solved self-consistently with the vector Maxwell equations for the pulse propagation.

Our approach, predominantly using quantum-optical description based on the Liouville equation of motion for the density matrix, is quite distinctive from the many-body theoretical methods applied in Chapter 48. For instance, the developed by the aforementioned authors quantum-optical semiconductor laser theory beyond the semiclassical approach is based on a correlation-expansion and truncation technique for the quantized light field interacting with carriers in the single-particle picture (electrons and holes) occupying the discrete QD levels. Dynamical equations for the carrier populations in the conduction and valence states of the QD, for the expectation value of the photon number operator of the laser mode and the photon-assisted polarization are derived using Heisenberg operator equations of motion for the creation and annihilation operators. These coupled equations contain carrier–photon and photon–photon correlations, providing access to the second-order correlation function, $g^{(2)}$. The interaction with the environment is accounted through quantum kinetic equations, beyond the Boltzmann equation approach, for the carrier dynamics due to carrier–phonon and carrier-Coulomb interactions. Many-body theory is applied to calculate the corresponding matrix elements. In the former case, effects such as polaron scattering and non-Markovian dynamics are shown for QD systems. This is a very thorough many-body quantized electromagnetic field approach; however, it does not take into account the specific macroscopic boundary conditions of the particular device geometry considered, as the light-matter interaction is described by a single laser mode coupled to the quantum system. An extension for multiple emitters coupled to a single common light mode is proposed, taking advantage of the many-particle picture description of the QD states, such as excitons, trions, multiexcitons, etc., to describe interemitter correlations and radiative coupling. This is somewhat similar to our approach in that it considers many-particle states (e.g., excitons); however, our approach is extended in that it also considers multiple cavity electromagnetic field modes. The cluster expansion approach in Chapter 48 by Gies et al. is in essence a generation of equations of motion for the photon correlations up to a given order of truncation of the expansion and thus providing access to $g^{(2)}$. The dissipation phenomena within this approach—such as the carrier–phonon and carrier–carrier Coulomb scattering—are calculated from a microscopic many-body model (compared to the Lindblad terms and our approach, where they are taken as phenomenological parameters), which is an advantage, as the theory is self-consistent.

In Section 49.2, we will give an overview of the strategy we have used to construct a model for describing coherent light-matter interactions in semiconductor nanostructures.

49.2 Maxwell-Pseudospin Method

49.2.1 Pseudospin Equation

In what follows, we will consider either resonant or near-resonant laser optical field interactions with atom-like discrete-level quantum systems, such as QDs. The optical field we consider is monochromatic and (near-)resonant with a given transition between a pair of atomic levels, thus effectively mimicking a two-level atom. As the two-level atom is conceptually equivalent to a half-spin particle in a magnetic field, the dynamics of the laser-atom interaction is governed by the same Bloch spin vector equation of motion [37] developed for describing nuclear magnetic resonance and the precession of a classical gyromagnet (or its quantum analogue, the spin vector, **S**) in a constant magnetic field, **B**:

$$\frac{\mathrm{d}\mathbf{S}}{\mathrm{d}t} = f\mathbf{B} \times \mathbf{S}. \tag{49.1}$$

Here f guarantees the constant length of the spin vector, thus confining the spin dynamics to the surface of a Bloch sphere. (Note, however, that this equation is valid only for systems with equally spaced energy levels). The formal analogy between a spin and a nonspin (pseudospin) system was pointed out for the first time by Feynman and coworkers [38], showing that when coherent processes are involved in a two-level system, it is sufficient to consider a real three-vector, rather than the complex probability amplitudes in the Schrödinger equation or the complex density-matrix elements [37]. This real-vector representation provides an elegant and intuitive geometrical framework for comprehending the system dynamics in terms of rotations of a real state vector in the Hilbert space. Furthermore, in contrast to the wavefunction formalism, it describes both pure and mixed quantum states. Attempts to extend the formalism analogously for $(\mathrm{N} > 2)$-level quantum systems fail to preserve the simple vector form of the basic equation of motion for the real state pseudospin (coherence, or Bloch) vector. The only known exception to this is the special case of excitation of a spin-J system in constant magnetic field (thus yielding a $(N = 2J + 1)$-level system), due to the equal spacing of the energy levels in this system [39]. Preserving this simple form of the vector equation has a number of advantages: (1) it accounts for the intrinsic symmetry of the underlying Hilbert space of the system and therefore is an exact description, independent of the strength, number, or time dependence of the external forces acting on the system; (2) similar to the two-level system, the dynamical evolution can be illustrated as a rotation in the real physical space of a real coherence vector.

The case of a three-level system is of particular interest, since it provides a useful framework for studying such phenomena as two-photon coherence, resonance Raman scattering, three-level echoes, three-level superradiance, coherent multistep photoionization/photodissociation, coherent population trapping, and electromagnetically induced transparency (EIT). A solution to this problem has been given by Elgin [40] by invoking the invariance of the state vector under rotations of the SU(3) transformation group. The dynamical evolution of a three-level system can be expressed in terms of an eight-dimensional real coherence vector, taking advantage of the group-theory and Gell-Mann's SU(3) generators, which had been developed for the quark triplet in the domain of high-energy physics [41].

A general solution for an N-level quantum system with arbitrary level spacing has been provided by Hioe and Eberly [42]. The starting point is the dynamical evolution of a quantum system described by the Liouville equation for the density operator in the Schrödinger picture

$$i\hbar\frac{\partial\hat{\rho}}{\partial t} = \left[\hat{H}, \hat{\rho}\right], \tag{49.2}$$

where $\hat{H}$ is the total system Hamiltonian:

$$\hat{H} = \hat{H}_0 + \hat{H}_{\mathrm{int}}(t), \tag{49.3}$$

with $\hat{H}_0$ being the unperturbed Hamiltonian of an N-level system: a diagonal matrix with eigenenergies of each level, $\hbar\omega_k$ $(k = 1, \ldots, N)$, along the main diagonal, and $\hat{H}_{int}$ is a time-dependent perturbation (not necessarily small).

Hioe and Eberly showed that the presence of unitary group generators in the time evolution of an N-level quantum system permits to describe the evolution in terms of the rotations of a real coherence vector. Expanding the system Hamiltonian and the density matrix operators in terms of λ-generators of the SU(N) Lie algebra, $\hat{\lambda}_k$, they derived the pseudospin equation of motion in the Heisenberg picture, depicting the time evolution of the real pseudospin vector, **S**, as a generalized rotation in $N^2 - 1$ Hilbert space:

$$\dot{S}_i = f_{ijk}\gamma_j S_k, \quad i, j, k = 1, 2, \ldots, N^2 - 1, \tag{49.4}$$

where the dot stands for the time derivative, γ_j is the torque vector, and $f_{ijk} = \frac{1}{4} i\left[Tr\left(\hat{\lambda}_i\hat{\lambda}_k\hat{\lambda}_j\right) - Tr\left(\hat{\lambda}_i\hat{\lambda}_j\hat{\lambda}_k\right)\right]$, $i, j, k = 1, 2, \ldots, N^2 - 1$, is the fully antisymmetric tensor of the structure constants of the SU(N) group, whose asymmetry guarantees the constant length of the pseudospin vector, $|\mathbf{S}|$, in the (N^2-1)-dimensional Hilbert space. The coherence vector is expressed as the trace of the product of the density and the SU(N) group generators, whereas the torque vector is expressed as the trace of the product of the system Hamiltonian operators and the SU(N) group generators

$$S_j(t) = Tr\left(\hat{\rho}(t).\hat{\lambda}_j\right); \quad \gamma_j(t) = \frac{1}{\hbar} Tr\left(\hat{H}(t).\hat{\lambda}_j\right). \tag{49.5}$$

The SU(N) group generators obey the group property

$$\left[\hat{\lambda}_j, \hat{\lambda}_k\right] = 2if_{jkl}\hat{\lambda}_l \tag{49.6}$$

and orthogonality relations, $Tr\left(\hat{\lambda}_j.\hat{\lambda}_k\right) = \delta_{jk}$, however their choice is not unique.

For the simplest case of a two-level system, $N = 2$, the λ-generators are simply the Pauli matrices $\hat{\sigma}_x$, $\hat{\sigma}_y$, and $\hat{\sigma}_z$. This is obtained from Equation 49.4 when f_{ijk} is replaced by the Levi-Civita symbol, ε_{ijk}. For a dipole-coupling interaction $\hat{H}_{\text{int}} = e\mathbf{E}.\hat{\mathbf{r}}$, where **r** is the local displacement operator and e is the electron charge, and a linearly polarized dipole optical transition (**E** || **r** with an optical selection rule for the z-axis projection of the total angular momentum, $\Delta J_z = 0$) the optical Bloch equations are easily recovered:

$$\dot{\mathbf{S}} = \boldsymbol{\gamma} \times \mathbf{S} \quad \begin{aligned} \dot{S}_1 &= \omega_0 S_2 \\ \dot{S}_2 &= -\omega_0 S_1 + 2\Omega_R S_3 \\ \dot{S}_3 &= -2\Omega_R S_2, \end{aligned} \tag{49.7}$$

where S_i $(i = 1, 2, 3)$ is the real pseudospin three-vector, and the torque vector is given by $\boldsymbol{\gamma} = \left(2\Omega_{\text{R}}, 0, \omega_0\right)$ from Equation 49.5. Here we have defined the Rabi frequency, $\Omega_{\text{R}} = \frac{\wp}{\hbar}E$, with $\wp = \langle i|e\hat{r}|j\rangle$, transition dipole matrix element of $i \rightarrow j$, and ω_0 is the atomic transition resonant frequency. In case of near-resonant excitation, ω_0 should be replaced by $\Delta\omega = \omega_0 - \omega$, where ω is the excitation frequency [43].

In the general case of an electric field vector noncollinear with the local displacement (dipole moment) vector, the optical transition is excited by circularly or elliptically polarized light obeying the selection rule $\Delta J_z = \pm 1$, where the sign "+/–" corresponds to right/left circularly polarized light. The dipole matrix elements are complex vectors yielding complex Rabi frequencies along the x- and y-direction in the plane determined by **E** and **r** (or equivalently, along the dipole moment vector): $\Omega_x = \frac{\wp}{\hbar}E_x$; $\Omega_y = \frac{\wp}{\hbar}E_y$. The torque vector is given by $\boldsymbol{\gamma} = \left(\Omega_x, \Omega_y, \omega_0\right)$ and the resulting pseudospin equations read (see also [37])

$$\dot{\mathbf{S}} = \boldsymbol{\gamma} \times \mathbf{S} \quad \begin{aligned} \dot{S}_1 &= -\omega_0 S_2 - \Omega_y S_3 \\ \dot{S}_2 &= \omega_0 S_1 + \Omega_x S_3 \\ \dot{S}_3 &= \Omega_y S_1 - \Omega_x S_2 \,. \end{aligned} \tag{49.8}$$

The relationship between the real-state coherence vector components and the density-matrix components for the general N-level case is given by Equation 49.5, and for a two-level system reads

$$\begin{aligned} S_1 &= \hat{\rho}_{12} + \hat{\rho}_{21} = 2\mathrm{Re}\left(\hat{\rho}_{12}\right) \\ S_2 &= i\left(\hat{\rho}_{12} - \hat{\rho}_{21}\right) = -2\mathrm{Im}\left(\hat{\rho}_{12}\right) = 2\mathrm{Im}\left(\hat{\rho}_{21}\right) \\ S_3 &= \hat{\rho}_{22} - \hat{\rho}_{21} \,. \end{aligned} \tag{49.9}$$

The physical meaning of the first two components of the pseudospin vector are dispersive (in-phase) and absorptive (in-quadrature) parts of the dipole polarization, and the third component is the population difference between the two levels. Since the level occupation probability is conserved, i.e., $\rho_{11} + \rho_{22} = 1$, it can be easily shown that the length of the coherence vector is unity: $|S|^2 = S_1^2 + S_2^2 + S_3^2 = 1$, and the evolution of the pseudospin is "inscribed" by the pseudospin vector on the surface of the Bloch sphere.

Up until now, we have only considered coherent processes; however, in an open quantum system there is always energy dissipation. For a two-level system, it is easy to define phenomenological longitudinal relaxation time (also known as population relaxation time), T_1, and transverse relaxation time (also known as polarization-decay or dephasing time), T_2, leading to the well-known optical Bloch equations for a two-level system with damping for e.g., linearly polarized transitions:

$$\begin{aligned} \frac{\partial S_1}{\partial t} &= \omega_o S_2 - \frac{1}{T_2} S_1 \\ \frac{\partial S_2}{\partial t} &= -\omega_o S_1 + 2\Omega_x S_3 - \frac{1}{T_2} S_2 \\ \frac{\partial S_3}{\partial t} &= -2\Omega_x S_2 - \frac{1}{T_1}\left(S_3 - S_{30}\right), \end{aligned} \tag{49.10}$$

where $S_{30} = \pm 1$ is the initial population profile; "−" and "+" correspond, respectively, to all of the population in the ground and the excited state. Upon optical excitation, the former describes absorption, while the latter describes optical gain. However, it is not so easy to generalize the dissipation dynamics for a system with an arbitrary number of levels (N-level system), because each pair of levels within this system is characterized by their individual population transfer rates and dephasing rates. We show how to do this generalization next. The time evolution of the two-level quantum system is described by the generalized rotation of the real three-vector, **S**, in the three-dimensional Hilbert space; at all times the tip of **S** is just touching the surface of the Bloch sphere, shown in Figure 49.1.

49.2.2 Dissipative Dynamics and Master Pseudospin Equation

To include energy relaxation and decoherence processes in the system dynamics, we adopt an unconventional approach. We introduce longitudinal relaxation in the Liouville equation for the density matrix by considering first the diagonal population components of the Liouville equation, ρ_{ii}, and taking into account population transfer between all allowed dipole optical transitions in the system with corresponding rates.

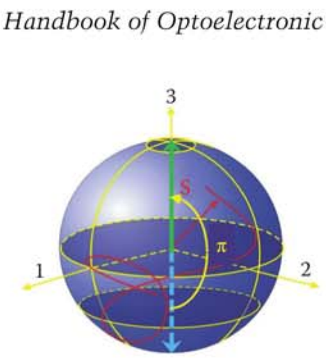


FIGURE 49.1 Geometrical representation of the time evolution of a two-level quantum system by generalized rotation of the real state pseudospin vector, **S**, in the 3D Hilbert space, drawing a trajectory on the $|\mathbf{S}| = const.$ sphere surface. A one-qubit rotation occurs when the vector is rotated through π from state 0 (spin down, dashed arrow) to state 1 (spin up, solid arrow along $\hat{3}$-axis), or vice versa.

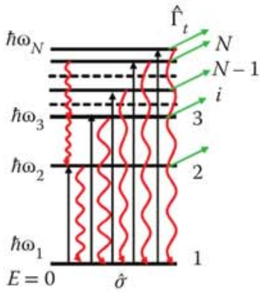


FIGURE 49.2 Schematic of an N-level quantum system with energy zero at the ground state. [Vertical solid arrows, coherent optical excitation; wavy lines, population relaxation $\hat{\sigma}$ (spontaneous emission); tilted arrows, dephasing $\hat{\Gamma}_t$ (decoherence).]

For each Liouville equation we can construct a longitudinal damping matrix of rank N: $\hat{\Gamma}_i$, $i = 1, \ldots, N$, whose elements are the population transfer rates between all pairs of levels. This allows us to introduce a diagonal matrix, $\hat{\sigma} = diag(Tr(\hat{\Gamma}_i.\hat{\rho}))$, which accounts for the longitudinal population relaxation. Similarly, we introduce an off-diagonal matrix, $\hat{\Gamma}_t$, of rank N with zero main diagonal components to account for the dephasing rates of all levels (Figure 49.2). Adding the relaxation and dephasing terms in the Liouville equation thus leads to the following master equation:

$$\frac{\partial\hat{\rho}}{\partial t} = \frac{1}{\hbar}\left[\hat{\rho}, \hat{H}\right] + \hat{\sigma} - \hat{\Gamma}_t.\hat{\rho}. \tag{49.11}$$

Using Equation 49.5, we can now rewrite the density-matrix equations in terms of the pseudospin vector, **S** [44]:

$$\frac{\partial S_j}{\partial t} = \begin{cases} f_{jkl}\Gamma_k S_l + \frac{1}{2}Tr\left(\hat{\sigma}.\hat{\lambda}_j\right) - \frac{1}{T_j}\left(S_j - S_{je}\right), & j = 1, 2, \ldots N\left(\mathrm{N} - 1\right) \\ f_{jkl}\Gamma_k S_l + \frac{1}{2}Tr\left(\hat{\sigma}.\hat{\lambda}_j\right), & j = N\left(N - 1\right) + 1, \ldots, N^2 - 1, \end{cases} \tag{49.12}$$

where we have defined an equilibrium coherence vector, $\mathbf{S_E} = (S_{1e}, S_{2e}, \ldots, S_{(N^2-1)e})$, and $T_j, j = 1, 2\ldots,$ $N^2 - 1$ are phenomenologically introduced nonuniform decay times describing the relaxation of the components of the real state vector towards their equilibrium values, S_{je}. Because of the incoherent processes involved in maintaining the system at a definite level of excitation, the equilibrium system polarization is zero. This means that the subset of $\mathbf{S_E}$ containing the polarization components of the equilibrium coherence vector is zero, i.e., $\mathbf{S}^*_{\mathbf{E}} = (S_{1e}, S_{2e}, ..., S_{N(N-1)e}) = 0$. The second term in the second line of the Equation 49.12 containing the trace of the density matrix through $\hat{\sigma}$ can be rewritten in terms of nonuniform relaxation times, $T_j = T_{N(N-1)+1}, ..., T_{N^2-1}$, describing the relaxation of the population components of the real coherence vector toward their equilibrium values. Therefore, the total equilibrium coherence vector is of the form: $S_E = (0, 0, ..., 0, S_{N(N-1)+1}, ..., S_{N^2-1})$. We shall refer to Equation 49.12 as the master pseudospin equations.

49.2.3 Vector Maxwell Equations and Coupling to Pseudospin Equations

The optical wave propagation is described by the vector Maxwell equations:

$$\begin{aligned}\frac{\partial \mathbf{H}}{\partial t} &= -\frac{1}{\mu}\nabla \times \mathbf{E}\\ \frac{\partial \mathbf{E}}{\partial t} &= \frac{1}{\varepsilon}\nabla \times \mathbf{H} - \frac{1}{\varepsilon}\frac{\partial \mathbf{P}}{\partial t},\end{aligned} \tag{49.13}$$

where $\mathbf{P}$ is the macroscopic medium polarization vector given by

$$\mathbf{P} = -eN_a Tr\,(\hat{\rho}.\hat{\mathbf{r}})\,, \tag{49.14}$$

with N_a signifying the density of the ensemble of resonant dipoles in the medium. Starting with a phenomenologically constructed dipole-coupling interaction Hamiltonian with coherent Rabi frequency terms describing optical transitions between pairs of levels, one can decompose the dipole moment operator, and represent it as a linear combination of λ-generators, whose coefficients are the corresponding Cartesian dipole moment vector components. Then, using Equation 49.14, we can find a relationship between the macroscopic polarization components, $P_{x,y,z}$, and the coherence vector components, S_j. To illustrate this procedure, suppose we have a transverse electromagnetic wave propagating along the z-axis exciting a given $i \rightarrow j\,(i < j)$ transition in an N-level system. Let the wave be circularly polarized in the plane perpendicular to z, with in-plane electric-field components E_x and E_y. Then, the (i, j) term of the interaction Hamiltonian will contain a complex coherent term $\hbar(\Omega_x \pm i\Omega_y)$, where the sign "+/−" depends on the helicity of the pulse, σ^+ or σ^-. (The sign convention is such that "+" and "−" correspond to, respectively, clockwise and counterclockwise rotation of the electric-field vector, when viewed from a direction opposite to that of the propagation of the pulse.) Note that, for the case of a linearly polarized excitation, the Rabi frequency will be real. The in-plane dipole moment vector can be decomposed into two Cartesian components: $e\mathbf{r} = e(\hat{r}_x\vec{e}_x + \hat{r}_y\vec{e}_y)$, where $\vec{e}_x$ and $\vec{e}_y$ are the unit vectors along the x- and y-axes. The coefficients $\hat{r}_x$ and $\hat{r}_y$ are $\hat{\lambda}$-generators; $\hat{r}_x$ contains unity at positions (i, j) and (j, i), whereas $\hat{r}_y$ contains the imaginary number, $-i$, at position (i, j) and i at position (j, i):

$$\mathbf{r} = r_0\left\{\begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot & 0 & 0 & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \cdot & \cdot & 1 & \cdot & 0\\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & \cdot & 1 & 0 & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \cdot & \cdot & 0 & 0 & 0 \end{pmatrix}\vec{e}_x + \begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot & 0 & 0 & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \cdot & \cdot & -i & \cdot & 0\\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & \cdot & i & 0 & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \cdot & \cdot & 0 & 0 & 0 \end{pmatrix}\vec{e}_y\right\}, \tag{49.15}$$

where r_0 is the dipole length scale. The following general formula can be used to obtain the corresponding $\hat{\lambda}$-generators, depending on the number of discrete levels of the system, N:

$$\begin{aligned}\hat{r}_x &= \hat{\lambda}_{j-i+\sum_{d=1}^{i-1}(\mathrm{N}-d)}\\ \hat{r}_y &= \hat{\lambda}_{j-i+\sum_{d=1}^{i-1}(\mathrm{N}-d)+\frac{\mathrm{N(N-1)}}{2}}\end{aligned}. \tag{49.16}$$

In fact, this linear combination reproduces the phenomenologically constructed dipole coupling interaction Hamiltonian, $\hat{H}_{int} = -e\mathbf{r}.\mathbf{E}$, when the dipole moment and the decomposed electric-field vector ($\vec{E} = E_x\vec{e}_x + E_y\vec{e}_y$) are substituted in

$$\hat{H}_{\text{int}} = \hbar\begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot & 0 & 0 & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \cdot & \cdot & \Omega_x - i\Omega_y & \cdot & 0\\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & \cdot & \Omega_x + i\Omega_y & 0 & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \cdot & \cdot & 0 & 0 & 0 \end{pmatrix}. \tag{49.17}$$

The macroscopic polarization vector components, P_x and P_y, are then obtained from Equation 49.14, giving

$$\begin{aligned}P_x &= -\wp N_a Tr(\hat{\rho}.\hat{r}_x)\\ P_y &= -\wp N_a Tr(\hat{\rho}.\hat{r}_y).\end{aligned} \tag{49.18}$$

Now we can write down the vector Maxwell equations for a circularly polarized optical wave exciting the $i \rightarrow j$ transition of an N-level system, and thus inducing dipole polarizations P_x and P_y along the x- and y-directions, respectively:

$$\begin{aligned}\frac{\partial H_x(z,t)}{\partial t} &= \frac{1}{\mu}\frac{\partial E_y(z,t)}{\partial z}\\ \frac{\partial H_y(z,t)}{\partial t} &= -\frac{1}{\mu}\frac{\partial E_x(z,t)}{\partial z}\\ \frac{\partial E_x(z,t)}{\partial t} &= -\frac{1}{\varepsilon}\frac{\partial H_y(z,t)}{\partial z} - \frac{1}{\varepsilon}\frac{\partial P_x(z,t)}{\partial t}\\ \frac{\partial E_y(z,t)}{\partial t} &= \frac{1}{\varepsilon}\frac{\partial H_x(z,t)}{\partial z} - \frac{1}{\varepsilon}\frac{\partial P_y(z,t)}{\partial t}.\end{aligned} \tag{49.19}$$

For the special case of a transverse optical wave linearly polarized along x and propagating in the z-direction, the above system (Equation 49.19) is reduced to

$$\begin{aligned}\frac{\partial H_y(z,t)}{\partial t} &= -\frac{1}{\mu}\frac{\partial E_x(z,t)}{\partial z}\\ \frac{\partial E_x(z,t)}{\partial t} &= -\frac{1}{\varepsilon}\frac{\partial H_y(z,t)}{\partial z} - \frac{1}{\varepsilon}\frac{\partial P_x(z,t)}{\partial t}.\end{aligned} \tag{49.20}$$

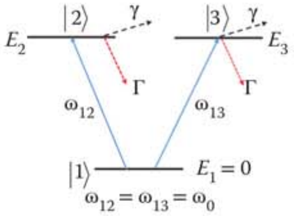


FIGURE 49.3 Energy-level scheme of a V-type three-level system. $\omega_{jk} = (E_j - E_k)/\hbar$, with energy zero chosen at the ground state $|1\rangle$. (Γ, population relaxation rate; γ, dephasing rate.)

We solve Maxwell's curl equations, (Equations 49.19 or 49.20), and the pseudospin equations (Equation 49.12) self-consistently, having established a link between them via the medium's polarization (Equation 49.18). We shall refer to this coupled system of first-order differential equations as the coherent Maxwell's curl-pseudospin model.

49.2.4 Coherent Maxwell-Pseudospin Equations in 2D

Description of nonlinear pulse propagation, interaction, and localization phenomena in multiple dimensions requires dipole coupling to at least a 2D resonant medium, whereby the electric field of a single-photon excitation is coupled to at least two distinct dipole optical transitions. We have previously shown that as a minimum requirement, it is sufficient to consider a degenerate three-level ensemble of dipoles in which two of the allowed dipole optical transitions are excited by each of the two components of the electric-field vector in the waveguide plane [45]. The pseudospin geometrical picture thus allows for adequate modeling of the interaction of an ultrashort laser pulse with a medium in two spatial dimensions. Therefore, we consider a V-type degenerate three-level system of resonant dipoles shown in Figure 49.3, for which a linearly polarized monochromatic electromagnetic wave induces polarizations (dipoles) in two orthogonal directions in the waveguide plane. Two cases arise: a transverse electric (TE) wave, for which $\mathbf{E} = (E_x, 0, 0)$ and $\mathbf{H} = (0, H_y, H_z)$, and transverse magnetic (TM) wave, for which $\mathbf{E} = (0, E_y, E_z)$ and $\mathbf{H} = (H_x, 0, 0)$. The respective vector Maxwell equations are

$$\begin{aligned}
\frac{\partial H_y}{\partial t} &= -\frac{1}{\mu}\frac{\partial E_x}{\partial z} \\
\frac{\partial H_z}{\partial t} &= \frac{1}{\mu}\frac{\partial E_x}{\partial y} \\
\frac{\partial E_x}{\partial t} &= \frac{1}{\varepsilon}\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right) - \frac{1}{\varepsilon}\frac{\partial P_x}{\partial t}
\end{aligned} \tag{49.21}$$

for the TE wave and

$$\begin{aligned}
\frac{\partial H_x}{\partial t} &= -\frac{1}{\mu}\frac{\partial E_z}{\partial y} + \frac{1}{\mu}\frac{\partial E_y}{\partial z} \\
\frac{\partial E_y}{\partial t} &= \frac{1}{\varepsilon}\frac{\partial H_x}{\partial z} - \frac{1}{\varepsilon}\frac{\partial P_y}{\partial t} \\
\frac{\partial E_z}{\partial t} &= -\frac{1}{\varepsilon}\frac{\partial H_x}{\partial y} - \frac{1}{\varepsilon}\frac{\partial P_z}{\partial t}
\end{aligned} \tag{49.22}$$

for the TM wave. We can phenomenologically construct the system Hamiltonian corresponding to each of the cases above:

$$\hat{H}_{TE}(t) = \begin{pmatrix} 0 & \Omega_x & 0 \\ \Omega_x & \omega_0 & 0 \\ 0 & 0 & \omega_0 \end{pmatrix}; \; \hat{H}_{\mathrm{TM}}(t) = \begin{pmatrix} 0 & \Omega_y & \Omega_z \\ \Omega_y & \omega_0 & 0 \\ \Omega_z & 0 & \omega_0 \end{pmatrix}. \tag{49.23}$$

The master pseudospin equations (Equation 49.12) for a three-level system with $i, j, k = 1, 2, \ldots, 8$ are coupled to Maxwell's equations above through the macroscopic polarization, Equation 49.14. It thus becomes possible to derive the following relationships between the polarization components and the pseudospin vector components for the case of a TE wave:

$$P_x = -\wp N_a S_1, \tag{49.24}$$

and for a TM wave:

$$\begin{aligned} P_y &= -\wp N_a S_1 \\ P_z &= -\wp N_a S_3. \end{aligned} \tag{49.25}$$

49.2.5 Numerical Methodology

The Maxwell-pseudospin system is discretized on a Yee grid in 1D (modified Yee grid in 2D, see Figure 49.4) and solved directly in the time domain by the finite-difference time-domain (FDTD) method [46]. At each time step, a solution to all equations is obtained using a predictor-corrector iterative scheme, proven to be very efficient in solving simultaneously a large number of first-order differential equations [47]. The time evolution of a discrete multilevel system under an external perturbation can be viewed as a Goursat-type initial boundary value problem, which is well posed when the time history of the initial electric field is given along, e.g., the left (lower) boundary of the simulation domain in 1D (2D). We impose analytical absorbing boundary conditions at the simulation domain boundaries, based on the Engquist–Majda [48] one-way wave equations discretized by the Mur finite-difference scheme [49] using a two-term Taylor-series approximation [46]. To ensure numerical stability, the time step and the spatial discretization are chosen to satisfy the Courant stability criterion: $\frac{c\Delta t}{\Delta z} \leq 1$ in 1D and $\frac{c\Delta t}{\left(\frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}\right)^{-1/2}} \leq 1$ in 2D [46]. The FDTD numerical technique has numerous advantages over other methods, such as the beam propagation method (BPM). For instance, modeling structures with rapid longitudinal variations (on the scale of the wavelength, λ) of, e.g., refractive index or energy-level population, is impossible to model with BPM, because it relies on the paraxial approximation which in this case is invalid. Moreover, FDTD can handle strong coupling between forward- and backward-propagating fields and can be used to describe general, nonharmonic electromagnetic fields. It can also be used to simulate ultrashort (sub-picosecond) pulses in situations where the pulse bandwidth is significant compared to material resonances.

49.3 Simulation of Nonlinear Dynamics in Optical Waveguides and Semiconductor Microcavities

In this section, we apply the vector Maxwell-pseudospin formalism to the problem of nonlinear optical pulse propagation in planar waveguides and microcavities with embedded quantum systems. We provide a solution in both 1D and 2D. As a first step, we perform validation studies against well-established theoretical and computational methods.

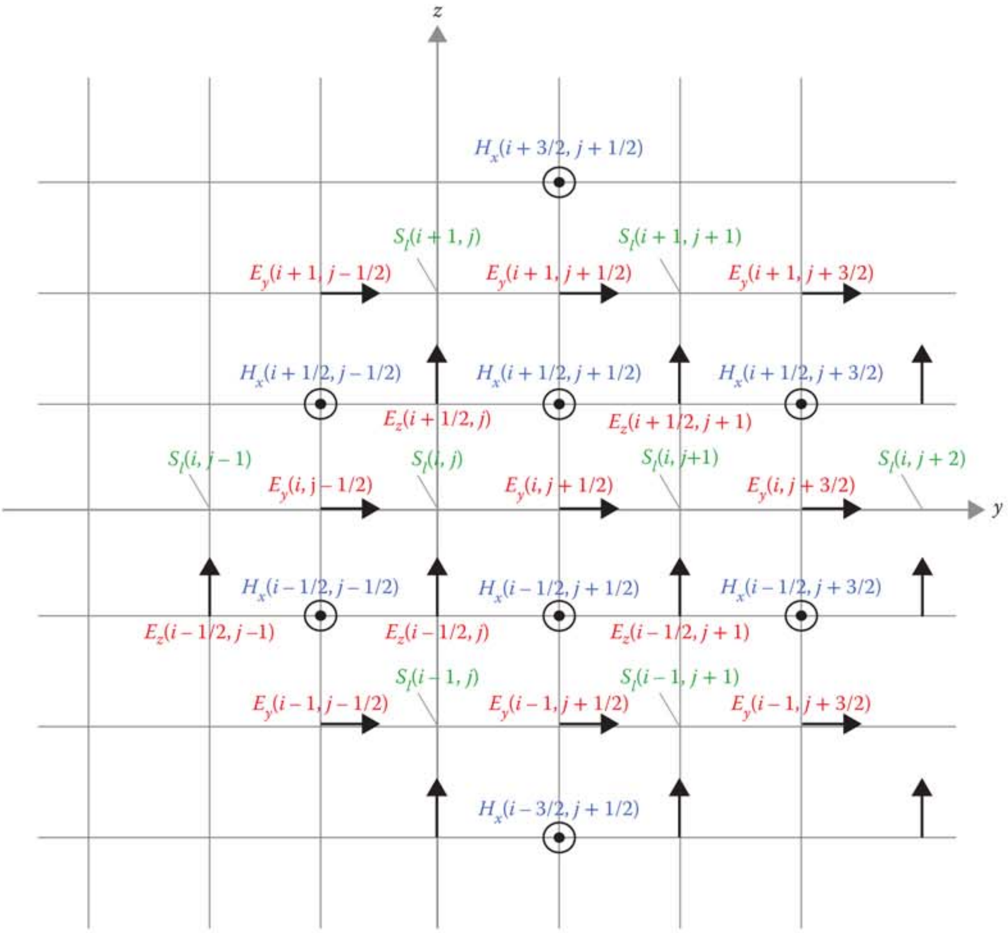


FIGURE 49.4 Modified Yee grid for discretization of the 2D Maxwell-pseudospin system for the purpose of carrying out numerical modeling. The grid is staggered, meaning that the electric and magnetic fields are calculated at every other node. This schematic is for a TM wave: the electric-field vector is in the y–z plane (arrows); the magnetic-field vector is pointing out of the page toward the reader (circle with dot). The electric- and magnetic-field components are spatially separated by $\Delta y/2$ and $\Delta z/2$ in the y–z plane, and temporally by $\Delta t/2$, with the electric-field components situated in the middle of the cells' edges, and magnetic-field components in the center of the cell [46]. The magnetic-field equation is solved at the spatial steps $(i + 1/2)\Delta z$ and $(j + 1/2)\Delta y$ over time step $(k + 1/2)\Delta t$. The electric-field components, E_y and E_z, are solved at the spatial steps $((j + 1/2)\Delta y, i\Delta z)$ and $(j\Delta y, (i + 1/2)\Delta z)$ over time step $k\Delta t$. Quantum-system variables, S_l, are assigned to the empty nodes in the 2D Yee grid, while the electric-field component values at these empty mesh points are estimated by averaging over the electric fields located at the nearest neighbors in the plane.

49.3.1 One-Dimensional Model of a Two-Level System

Consider a two-level quantum system, such as fundamental heavy-hole transitions in a QD. The coherent Maxwell-pseudospin equations describing the interaction of an optical wave propagating along the z-axis and linearly polarized in a transverse direction along x are given by Equations 49.10, 49.20, and 49.24:

$$\frac{\partial H_y}{\partial t} = -\frac{1}{\mu}\frac{\partial E_x}{\partial z}$$

$$
\begin{aligned}
\frac{\partial E_x}{\partial t} &= -\frac{1}{\varepsilon}\frac{\partial H_y}{\partial z} - \frac{\wp N_a}{\varepsilon T_2} S_1 + \frac{\wp N_a \omega_0}{\varepsilon} S_2 \\
\frac{\partial S_1}{\partial t} &= -\frac{1}{T_2} S_1 + \omega_0 S_2 \\
\frac{\partial S_2}{\partial t} &= -\omega_0 S_1 - \frac{1}{T_2} S_2 + 2\frac{\wp}{\hbar} E_x S_3 \\
\frac{\partial S_3}{\partial t} &= -2\frac{\wp}{\hbar} E_x S_2 - \frac{1}{T_1}\left(S_3 - S_{30}\right).
\end{aligned}
\tag{49.26}
$$

We apply an excitation pulse of duration T_p with either a hyperbolic secant or Gaussian envelope and central frequency, ω_0, in resonance with the two-level transition at the boundary $z = 0$ of the simulation domain. Its time dependence is thus

$$
E_x\left(z = 0, t\right) = \begin{cases} E_0 \mathrm{sech}\left(10\frac{t - \frac{T_\mathrm{p}}{2}}{\frac{T_\mathrm{p}}{2}}\right) \sin\left(\omega_0 t\right) \\ E_0 \exp\left[\left(\frac{t - t_0}{t_d}\right)^2\right] \sin\left(\omega_0 t\right), \end{cases}
\tag{49.27}
$$

where E_0 is the initial pulse amplitude, t_0 is the time moment at which the Gaussian pulse is initially centered, and t_d is the characteristic $1/e$ Gaussian decay.

49.3.2 Passive Cavity Properties

To validate our model, we construct a simple, symmetric test Bragg microcavity containing 5 $\lambda/4$ GaAs/$\mathrm{Al_{0.1}Ga_{0.9}As}$/AlAs layers, with λ being 1.4 μm (Figure 49.5f) and compute the passive (cold-)cavity modes and the electric-field standing-wave profile within the cavity. In order to compute the resonant cavity modes, we launch a short broad-band pulse of duration ≤ 100 fs from the left boundary of the structure and sample in time the electric-field evolution at the output on the right. This time trace is Fourier-transformed and normalized with respect to the Fourier transform of the input trace, thus yielding the transmission spectrum of the microcavity. The resonant frequency is extracted from the peak in the stop band of the spectrum (Figure 49.5d). Note that reducing the time step allows to resolve higher-order cavity modes. To obtain the electric-field standing-wave profile within the cavity, a continuous sine wave at the resonance frequency (obtained in the previous step) is pumped from the left boundary, and the time evolution of the intracavity field is monitored. The field amplitude increases in time and eventually reaches a steady value. The results shown in Figure 49.5f are compared with the cavity modes and field distributions computed by the transfer matrix method (TMM).

Using the above FDTD-based methodology, we will subsequently investigate passive cavity properties of more complex cavities. For example, we calculate the resonant cavity modes and standing-wave profile of a Coldren-type microcavity structure designed at $\lambda = 1.3$ μm and consisting of 10 GaAs/$\mathrm{Al_{0.9}Ga_{0.1}As}$ pairs for the bottom distributed Bragg reflector (DBR) mirror, followed by an AlAs supply layer, a tuning $\mathrm{Al_{0.7}Ga_{0.3}As}$ layer, a GaAs $\lambda/2$ cavity and 15 GaAs/$\mathrm{Al_{0.9}Ga_{0.1}As}$ pairs for the top DBR mirror. The simulation results for the resonant cavity modes, namely the transmission spectrum as a function of λ and the electric-field standing-wave profile along the structure, are displayed in Figure 49.6.

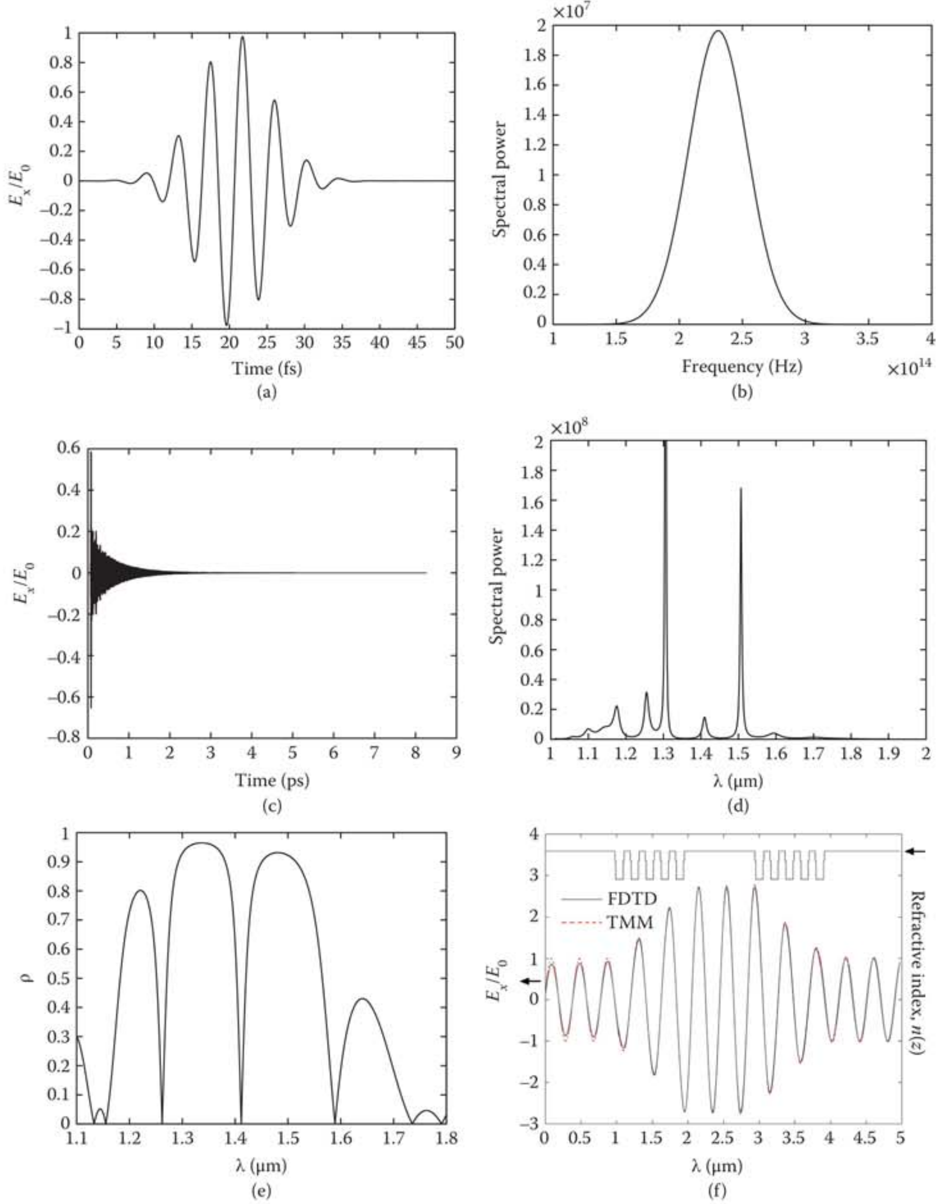


FIGURE 49.5 (a) Initial ultrashort pulse ($T_p \leq 100$ fs). (b) Fourier spectrum of the initial broad-band pulse. (c) Time trace of $E_x(t)$ at the output boundary of the structure. (d) Transmission spectrum as a function of wavelength, λ, computed as the discrete Fast Fourier transform (FFT) of the output time trace in (c), normalized with respect to the input FFT spectrum (b). The resonant cavity wavelength, $\lambda_0 \approx 1.4\,\mu$m, is extracted from the peak in the DBR cavity stop band. (e) Reflectivity spectrum as a function of wavelength computed by the TMM: the resonant cavity wavelength, $\lambda_0 \approx 1.4\,\mu$m is obtained from the dip in the reflectivity spectrum within the DBR cavity stop band. (f) Comparison of the standing-wave profile within the cavity, as obtained by the TMM (dashed curve) and FDTD (solid black curve) methods. The FDTD-inferred instantaneous field standing-wave profile is computed with continuous sine-wave excitation propagating from the left boundary at the cavity resonance frequency equivalent to $\lambda_{\text{pump}} = 1.4116\,\mu$m, obtained from (d) after a sufficiently large number of time steps 2×10^6 (each of duration $\Delta t = 3.33 \times 10^{-3}$ fs) with 31,300 spatial grid points (each spatial step of $\Delta z = 1$ nm), Courant number= 0.75. The scale on the right corresponds to the refractive-index profile of the DBR cavity (top).

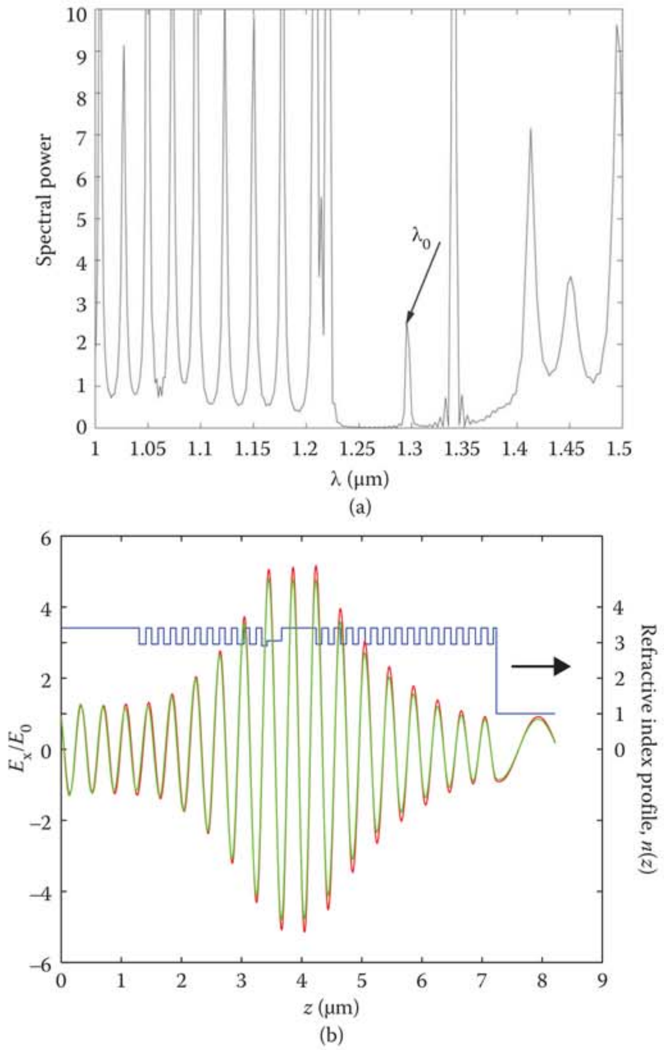


FIGURE 49.6 (a) FFT of the time trace of the signal detected at the output boundary of a Coldren-type microcavity structure, revealing the cavity's (nonnormalized) transmission spectrum. The resonant cavity mode is indicated with λ_0, and corresponds to approximately 1.3 μm. (b) Standing-wave profile after 6 million (dashed curve) and 12 million (solid curve) time steps superimposed on the refractive-index profile of the microcavity (scale on right-hand side). The time step in the simulation was $\Delta t = 3.33 \times 10^{-3}$ fs and the spatial step was $\Delta z = 1$ nm. The amplitude of the standing wave increases as the number of time steps is increased, and eventually reaches a steady maximum value (solid curve). The amplitude is augmented with time due to the build-up of the field caused by multiple reflections at the cavity mirrors.

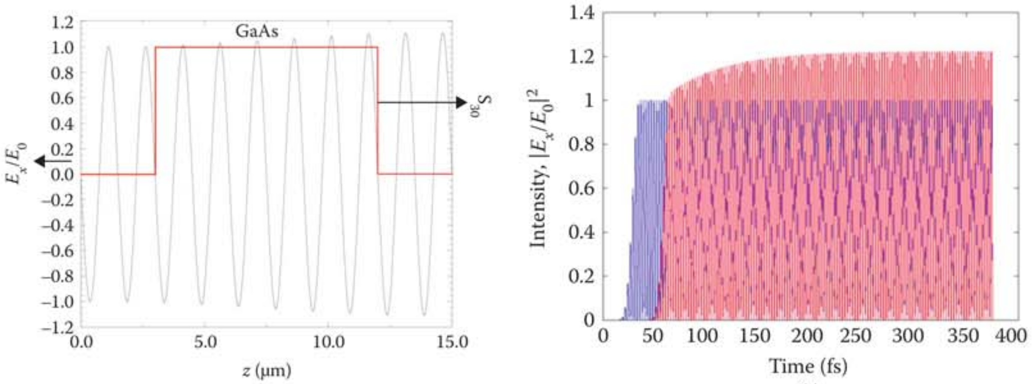
49.3.3 Active Nonlinear Optical Waveguide/Cavity Properties

We have performed gain validation studies of our 1D semiclassical Maxwell-pseudospin model. This was done by establishing the ability of our model to recover the linear small-signal (unsaturated) gain and nonlinear (saturated) gain of a homogeneously broadened (HB) two-level system, and comparing the result against that of the density-matrix theory, which served as a benchmark (see, e.g., [50]). The test structure used for these validation studies was a 15 μm-long simulation domain consisting of a GaAs slab waveguide with a gain region length of $L_g = 9\,\mu\text{m}$, refractive index $n_{\text{GaAs}} = 3.59$, and pumping wavelength of 1.5 μm corresponding to the resonant transition frequency of a two-level system sandwiched between two free-space regions (Figure 49.7a). The GaAs slab was represented by an initially inverted ($S_{30} = 1$) two-level medium with resonant dipole density $N_a = 1 \times 10^{24}\,\text{m}^{-3}$, dipole transition matrix element, $\wp = 1 \times 10^{-29}\,\text{C} \cdot \text{m}$ and population relaxation and dephasing times, $T_1 = 100$ ps and $T_2 = 50$ fs, respectively. We then applied a switch-on pulse: a continuous, sinusoidal electric-field pulse, with pulse carrier frequency ω_0 in resonance with the two-level system, which is smoothly increased in amplitude from 0 up to unit amplitude ($E_0 = 1\ \text{V} \cdot \text{m}^{-1}$) within five periods:

$$E_x(t) = E_0 \sin(\omega_0 t) \begin{cases} 0, & \text{for } t < 0 \\ \left(1 - x^2\right)^4, & \text{for } 0 \le t \le 5T_\text{p} \\ 1, & \text{for } t > 5T_\text{p}, \end{cases} \tag{49.28}$$

with $x = \left(t - 5T_\text{p}\right)/5T_\text{p}$, and T_p being the period of one cycle of the sinusoid. The electric-field amplitude is clearly amplified upon the passage of the sinusoidal pulse across the gain region, thus demonstrating linear small-signal gain (Figure 49.7a). In the time domain, the application of an initial sinusoidal pulse with unit amplitude is amplified at the output end of the structure (Figure 49.7b). Let us compute the complex propagation factor between two points in the gain medium, z_1 and z_2, separated by one dielectric wavelength, $l = z_2 - z_1 = \lambda_0/n$:

$$e^{ik_c(z_2 - z_1)} = \frac{E_x(z_2, \omega)}{E_x(z_1, \omega)}, \tag{49.29}$$



(a) (b)

FIGURE 49.7 (a) Slight amplification of a sinusoidal electric-field switch-on pulse upon its passage through an initially inverted two-level GaAs slab medium of refractive index 3.59 after 1×10^6 time steps (each of duration 1×10^{-4} fs) with Courant stability parameter 0.75. The initial population profile across the structure is shown with a thick solid line (top-hat shape). The spatial discretization in this simulation was 1 Å. (b) Intensity evolution in time of the pulse at the input boundary of the simulated structure (fast-oscillating curve on the left) and the amplified time trace of the pulse at the output boundary (time-delayed curve on the right). The fast oscillations are at the resonant pulse carrier frequency, ω_0.

where $E_x(z_i,\omega)$ $i = 1,2$, are the Fourier transforms of the time traces of the electric field sampled at the two points, and $k_c = \beta + i\gamma$ is the complex wavevector with β-phase shift and γ-gain/absorption coefficient, depending on its sign (+/−). Thus, substituting these back into the left-hand side of Equation 49.29, one can define the amplification and phase factors as $e^{-\gamma.l}$ and $e^{\beta.l}$, respectively. For the amplitude gain coefficient, the steady-state solutions of the density-matrix equations for a two-level system (see [50]) read:

$$\gamma^{ampl}(\omega) = -\frac{\omega_o}{c_o}\frac{N_a\wp^2}{2\varepsilon_o\hbar n_{\mathrm{GaAs}}}g(\omega), \tag{49.30}$$

where n_{GaAs} is the GaAs refractive index, c_0 is the speed of light, ε_0 is the dielectric permeability in vacuum, and the normalized Lorentzian lineshape function is given by: $g(\omega) = \frac{2T_2}{1+(\omega-\omega_o)^2T_2^2}$, with normalization condition: $\int_{-\infty}^{\infty} g(\omega)d\omega = 1$. The "atomic" phase shift of the optical wave induced by the two-level system is given by

$$\Delta\varphi_{\mathrm{at}} = \Delta k.l = \frac{\omega}{c_o}\frac{N_a\wp^2\left(\omega-\omega_o\right)T_2}{4\varepsilon_o\hbar n_{\mathrm{GaAs}}}g(\omega) = \beta l - n_{\mathrm{GaAs}}\frac{\omega}{c_o}l. \tag{49.31}$$

We can now compare the FDTD simulation results obtained with our Maxwell-pseudospin model (Equation 49.26) with the above analytic expressions (Equations 49.30 and 49.31); the results from the two methods are plotted in Figure 49.8. Our model is in remarkably good agreement with the density matrix-calculated gain/absorption coefficients (a) and phase shift (b) in a HB two-level system.

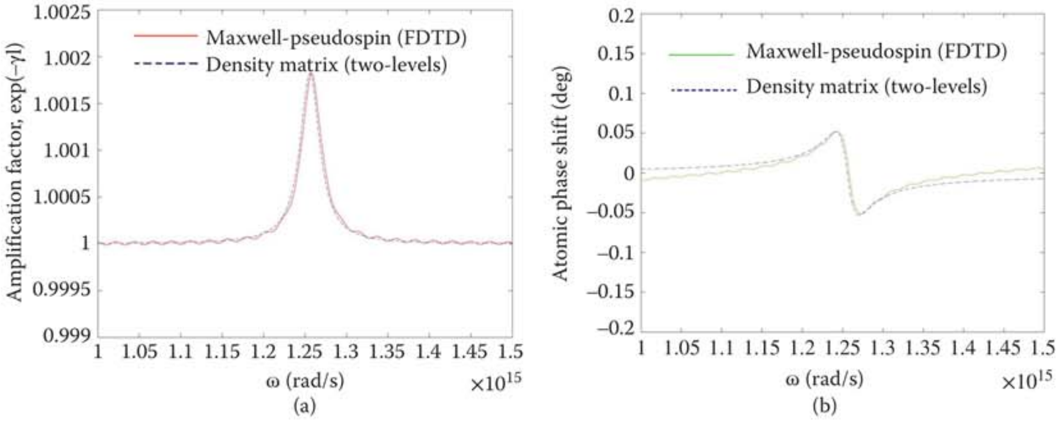


FIGURE 49.8 (a) Amplification spectrum of a switch-on pulse with carrier wave of angular frequency, ω, passing through a GaAs slab waveguide. The spectrum was calculated using two independent methods: by applying the analytical expressions of the density-matrix method to a HB two-level system (Equation 49.30) (dashed curve), and by finding the FDTD solution of the Maxwell-pseudospin set of equations (Equation 49.26) (solid curve). The ω corresponding to the amplification peak is resonant with the optical transition of the two-level quantum system making up the waveguide and is denoted ω_0. Note that the ripple effect observed in the solid curve is an artefact of the FDTD numerical method. (b) ω-dependence of the phase shift (in degrees) of the switch-on pulse emerging from the waveguide. The phase-shift spectrum was inferred independently from the density-matrix method (Equation 49.31) (dashed curve) and the Maxwell-pseudospin method (Equation 49.26) (solid curve). FDTD, finite-difference time-domain.

49.3.4 Resonant Nonlinearities and Gain Saturation

In Section 49.3.3, we employed the analytical steady-state density-matrix equations to perform a successful validation of our Maxwell-pseudospin model against the linear small-signal gain in a HB two-level system. In what follows, we shall show that our model also describes correctly the nonlinear gain dynamics of a HB two-level saturable gain medium. We keep the model structure, source excitation (i.e., input pulse) shape, and material parameters the same as in Section 49.3.3, so that the saturation condition at resonance is satisfied:

$$\frac{\wp^2 E_0^2 T_1 T_2}{\hbar^2} > 1. \tag{49.32}$$

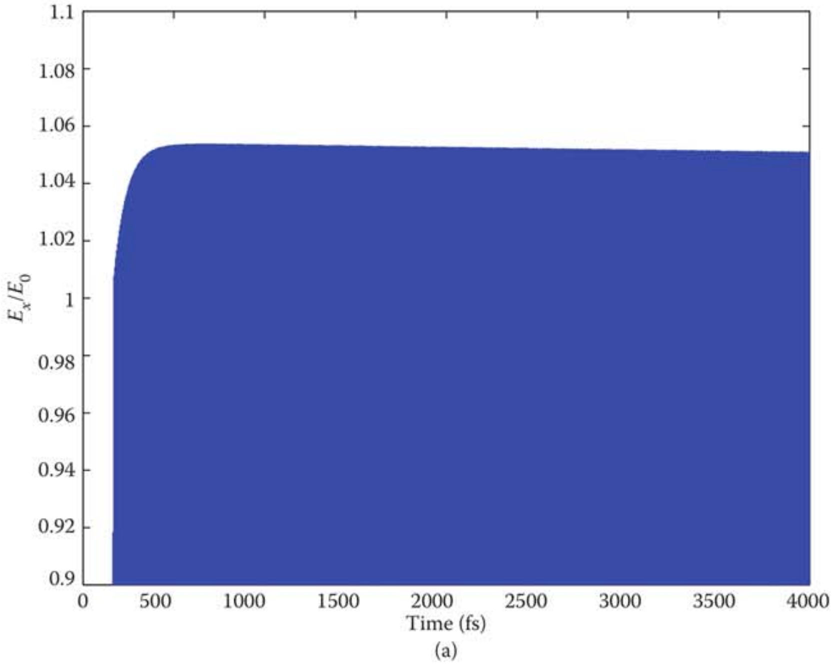
We apply the same source excitation (Equation 49.28) but vary the amplitude, E_0, running separate simulations for different amplitudes, ranging from $E_0 = 5 \times 10^6\ \mathrm{V \cdot m^{-1}}$ to $E_0 = 1 \times 10^8\ \mathrm{V \cdot m^{-1}}$. The nonlinear gain dynamics is shown in Figure 49.9 [51]. The gain peaks and levels off after some time, due to the depletion of the population residing on the upper level. The higher the initial excitation, the faster the gain reaches the steady-state field amplitude value. We proceed to demonstrate quantitative agreement between density-matrix theory of saturation of a two-level system and our Maxwell-pseudospin model. We choose the initial electric-field amplitude $E_0 = 5 \times 10^7\ \mathrm{V \cdot m^{-1}}$, $T_1 = 10$ ps, and $T_2 = 10$ fs, again satisfying the resonant saturation condition (Equation 49.32). The saturation value of the population inversion was calculated in [50] to be

$$S_{3\mathrm{sat}} = \frac{1}{1 + \left(\frac{E_0}{E_{s0}}\right)^2}, \quad E_{s0} = \frac{\hbar}{\wp\sqrt{T_1 T_2}}, \tag{49.33}$$

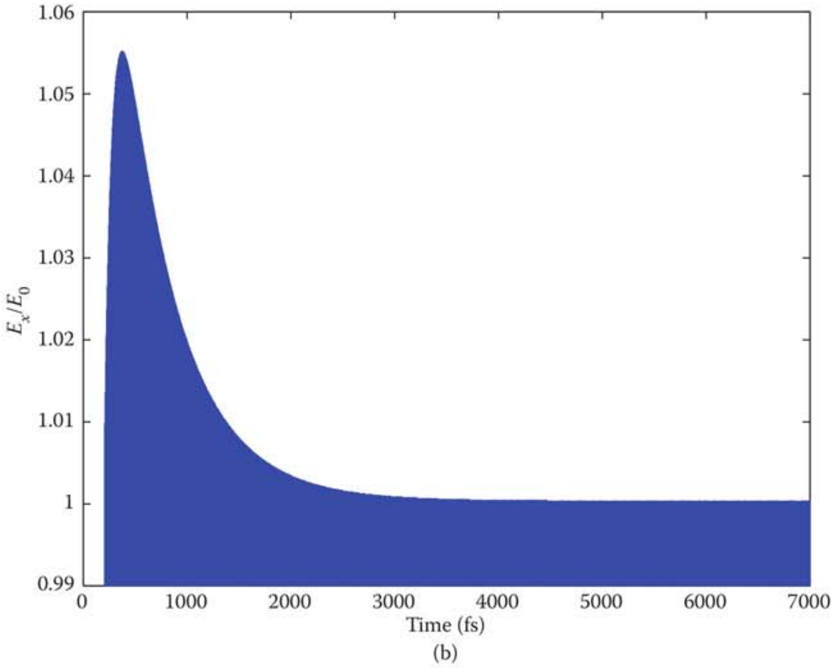
giving a value $S_{3_\mathrm{sat}} = 0.3$ for the chosen parameter set. We can calculate the unsaturated intensity gain coefficient from the maximum intensity using $\gamma_\mathrm{int}^\mathrm{unsat} = \frac{ln(I\mathrm{max}/I_o)}{2L_g}$. By definition, the saturated intensity is the intensity of the signal passing through the laser medium for which the gain coefficient is reduced down to one-half of its original value [52]:

$$\gamma_\mathrm{int}^\mathrm{sat} = \frac{\gamma_\mathrm{int}^\mathrm{unsat}}{1 + I/I_\mathrm{sat}}. \tag{49.34}$$

Given that the normalized saturated field intensity is $I_\mathrm{sat}/I_0 = exp(\frac{\gamma^\mathrm{unsat}}{2}L_g)$ [51], I_sat/I_0 is calculated to be 1.01. In order to make quantitative comparison with our model, we run the code with a large number of time steps (4.8×10^7, of 1×10^{-4} fs duration each) to ensure that the field and population dynamics have reached stationary values. In Figure 49.10a, we plot the final instantaneous electric-field distribution and population profile at subsequent time steps, showing the population relaxing to a steady-state value. In the time domain, the time trace at the waveguide output, (Figure 49.10b), and the population inside the gain region (Figure 49.10c), are plotted. It is easy to check that the steady-state values inferred from our model are in excellent agreement with the theoretically computed values. With this we conclude our corroboration of the ability of our model to reproduce correctly the gain saturation dynamics of a HB two-level quantum system.



(a)



(b)

FIGURE 49.9 Gain saturation dynamics of a HB two-level quantum system excited by a source electric field with different amplitude: (a) $E_0 = 5 \times 10^6\ \mathrm{V \cdot m^{-1}}$; (b) $E_0 = 5 \times 10^7\ \mathrm{V \cdot m^{-1}}$. The time trace of the amplified electric field is sampled at the output of the simulation domain. In both cases, the amplitude of the oscillations (with frequency ω_0) stabilizes after some time. Note that after peaking, the gain in (a) continues to drop approximately linearly at a relatively slow rate over the entire time span shown on the graph, while the gain in (b) reaches a steady-state value after ~4 ps. (G. M. Slavcheva et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2004 IEEE.)

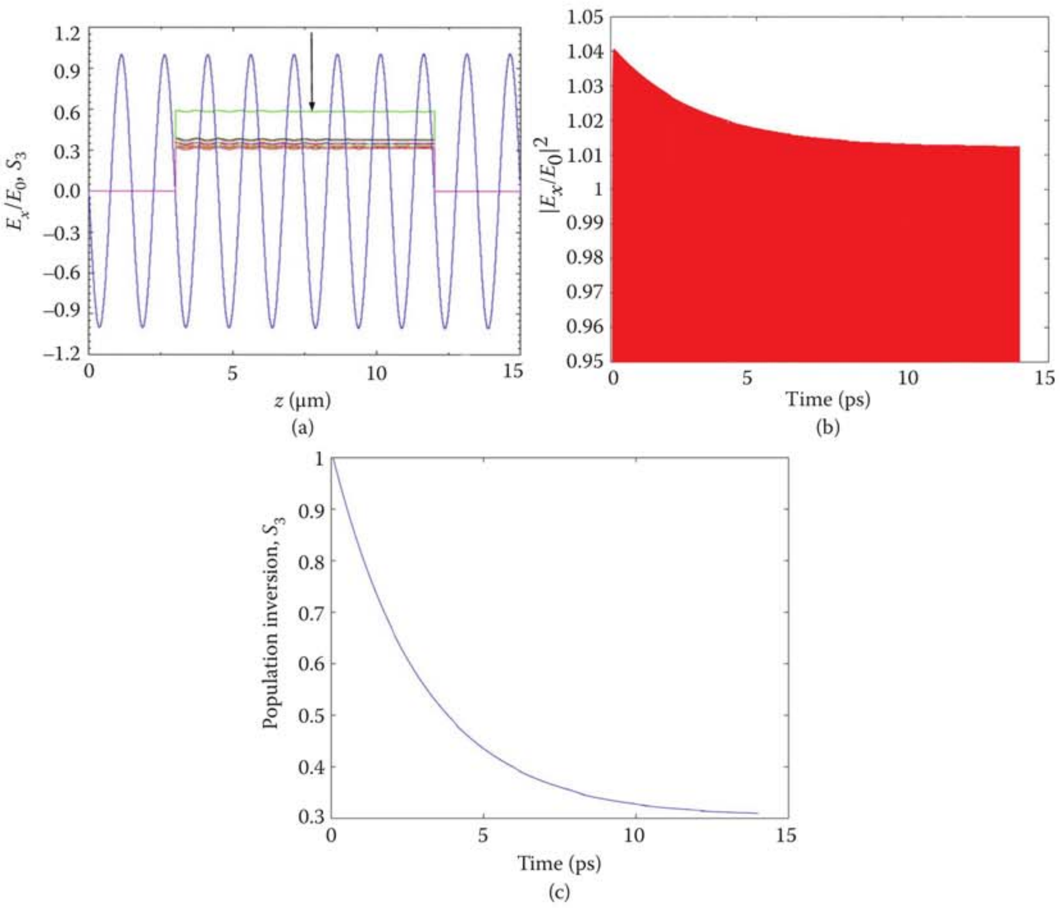


FIGURE 49.10 (a) Instantaneous electric-field amplitude profile across the waveguide (sinusoidal curve), obtained after 4.8×10^7 time steps (corresponding to a total time of ~4.8 ps), and the population inversion profile within the active region at several subsequent time steps (top-hat function with decreasing amplitude). The step in z is $\sim1.8\times10^{-10}$ m and the time step is $\sim1\times10^{-4}$ fs. The population is initially completely inverted: $S_{30} = 1$, but reduces over time (downward-pointing arrow), eventually converging to the theoretically obtained population steady-state value of 0.3. (b) Time evolution of the normalized field intensity showing convergence toward the theoretical value of 1.01. (c) Population relaxation over time, tending toward the theoretically predicted value of 0.3. (G. M. Slavcheva et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2004 IEEE.)

49.3.5 Coherent Propagation Effects, Self-Localization and Pattern Formation in 2D Planar Nonlinear Optical Waveguides and Semiconductor Microcavities

49.3.5.1 2D Nonlinear Optical Waveguides

We employ our 2D Maxwell-pseudospin model (see Section 49.2.4), based on self-consistent solution of the time evolution equations of a degenerate three-level quantum system, (Equation 49.4), for $j, k, l = 1, 2, ..., 8$ and Maxwell's curl equations for a TE-guided mode, Equation 49.21 (Equation 49.22 for a TM equivalent), propagating in a planar parallel-plate mirror optical waveguide, or planar semiconductor microcavity [45], of the type shown in Figure 49.11.

It is easy to show that by setting $S_1 = S_1, S_4 = -S_2, S_7 = S_3$, the TE case is equivalent to a two-level system case (Equation 49.7) and the equations for the remaining coherence vector components are decoupled

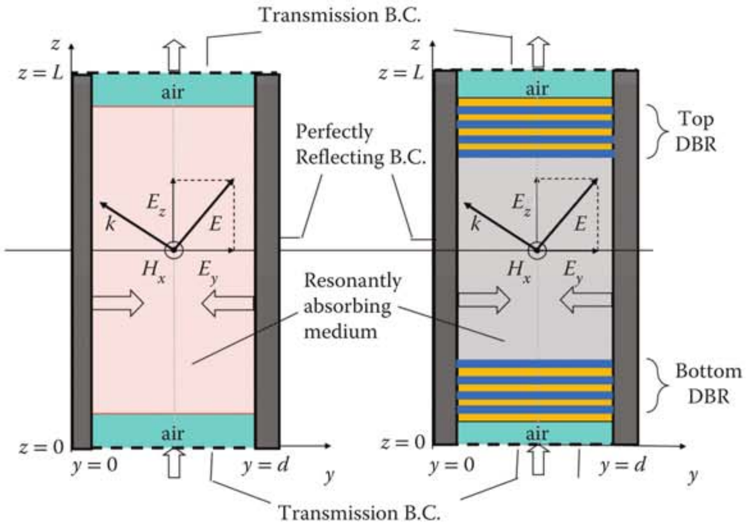


FIGURE 49.11 Planar parallel-mirror waveguide (left panel) and a semiconductor microcavity (right panel) filled with a resonantly absorbing three-level medium. Only TM-guided modes are shown. Perfectly transmitting boundary conditions are imposed on the upper/lower boundaries (upward-pointing arrows) and perfectly reflecting side boundaries are assumed (left/right arrows). The pulse initially propagates through a free-space region (air), then penetrates further, interacting with the degenerate three-level medium, and finally exits through another free-space region. B.C., boundary conditions; DBR, distributed Bragg reflector.

from this system, since there is no coupling to "level 3" in this case. In the FDTD numerical implementation, we use as a source field the TEM (i.e., TM_0) guided mode of a parallel-mirror slab waveguide. The electric field is a plane wave linearly polarized in the transverse x-direction perpendicular to the waveguide plane

$$E_y\left(z=0,y,t\right)=\begin{cases} E_0\mathrm{sech}\,(10\Gamma)\sin\left(\omega_0 t\right) \\ E_0\exp\left[\left(\frac{t-t_0}{t_\mathrm{d}}\right)^2\right]\sin\left(\omega_0 t\right) \end{cases} \quad E_z\left(z=0,y,t\right)=0\,, \tag{49.35}$$

where $\Gamma=\frac{t-\frac{T_\mathrm{p}}{2}}{\frac{T_\mathrm{p}}{2}}$. Here we have allowed for two electric-field envelopes: hyperbolic secant and Gaussian. We should therefore expect that the PAT [18] still holds. The pulse area at a given time is given by

$$\theta\left(z,t\right)=\frac{\wp}{\hbar}\int\limits_{-\infty}^{t}A\left(z,t'\right)dt'\,, \tag{49.36}$$

where the electric-field envelope is as follows:

$$A\left(z,t\right)=E_0\mathrm{sech}\left(\frac{t-\frac{z}{\mathrm{v}}}{\tau}\right)=E_0\mathrm{sech}\,(10\Gamma)\,. \tag{49.37}$$

The area under the entire pulse is given by

$$\theta_{\text{pulse}}(z) = \frac{\wp}{\hbar} \int_{-\infty}^{\infty} A\left(z, t'\right) \mathrm{d}t' \,. \tag{49.38}$$

According to PAT, pulses with areas less than π should be completely absorbed within several absorption lengths; however, pulses with an area of even multiples of π should continue to propagate undistorted as a solitary wave—known as SIT soliton—through the resonantly absorbing medium. We set all resonant dipoles initially to the ground state, i.e., $S_{7e} = -1, S_{8e} = -\frac{1}{\sqrt{3}}$. We then test the predictions of PAT by calculating the pulse amplitude from the PAT, which is given by the following relation for a pulse with a hyperbolic secant-shaped envelope:

$$E_0 = \frac{2\pi\hbar f_0}{\wp \arctan\left(\sinh\left(u\right)\right)\Big|_{-10}^{+10}} \frac{\theta_{\text{pulse}}\left(z=0\right)}{2\pi}, \tag{49.39}$$

where $f_0 = 1/\tau_{\text{p}}$ and $\tau_{\text{p}} = T_{\text{p}}/20$, with $T_{\text{p}} = 100$ fs being the pulse duration chosen for this simulation. If we set the initial pulse area $\theta(z=0)$ to 2π, we get a pulse amplitude corresponding to a soliton pulse of $E_0 = 4.2186 \times 10^9\ \text{V} \cdot \text{m}^{-1}$. The simulated waveguide structure is 150 μm long and 50 μm wide. 3D plots of the soliton field modulus (a) and population relaxation term S_7 (b) distributions in the waveguide *y-z* plane and a slice through the middle of the waveguide's width are displayed in Figure 49.12.

The stable SIT soliton propagation is shown in Figure 49.13a and b: the initial 2π pulse propagates in the resonantly absorbing medium while maintaining its hyperbolic secant shape without any loss. The atoms are driven into the excited state by absorbing ultrashort pulse energy, and by reradiating this energy to the field they return to the ground state. The corresponding population inversion exhibits complete transition from the ground level to the excited state and back to the initial state within one Rabi period. By contrast, for a 0.9π pulse with pulse amplitude $E_0 = 1.18957 \times 10^9\ \text{V} \cdot \text{m}^{-1}$, the initially symmetric pulse is distorted and reshaped during its propagation (Figure 49.13c). The initial equilibrium between the energy absorbed by the leading pulse edge and the energy reemitted by the trailing edge is violated and the pulse becomes increasingly asymmetric. In addition, the system is only partially inverted (Figure 49.13d). The reemitted energy of the continuously growing trailing edge is transmitted through the lower ($z = 0$) boundary. After a sufficiently long time, this pulse should eventually be completely absorbed, in agreement with the PAT.

Let us consider now the more interesting case of a TM_1-guided mode of a parallel-mirror slab waveguide, which cannot be reduced to the two-level system case. We solve self-consistently Maxwell's equations for a TM wave, Equation 49.22, with the pseudospin equations, (Equation 49.4), for a three-level system. As a source field we use the TM_1 slab waveguide mode

$$\begin{aligned}
&E_y\left(z=0, y, t\right) = E_o c_1 \cos\left(\frac{\pi y}{w}\right) \operatorname{sech}\left(10\Gamma\right) \sin\left(\omega_o t\right) = \tilde{E}_y\left(y, t\right) \sin\left(\omega_o t\right) \\
&E_z\left(z=0, y, t\right) = -E_o c_2 \sin\left(\frac{\pi y}{w}\right) \operatorname{sech}\left(10\Gamma\right) \cos\left(\omega_o t\right) = \tilde{E}_z\left(y, t\right) \sin\left(\omega_o t\right) \\
&c_1 = \frac{\sqrt{\frac{\omega_o^2 n^2}{c_o^2} - \left(\frac{\pi}{w}\right)^2}}{\omega_o \varepsilon_o n^2}; \qquad c_2 = \frac{\pi/w}{\omega_o \varepsilon_o n^2},
\end{aligned} \tag{49.40}$$

where *w* is the waveguide width. In order to obtain a stable soliton solution, we generalize the PAT to multiple spatial dimensions, defining the pulse area as the area below the modulus of the field envelope,

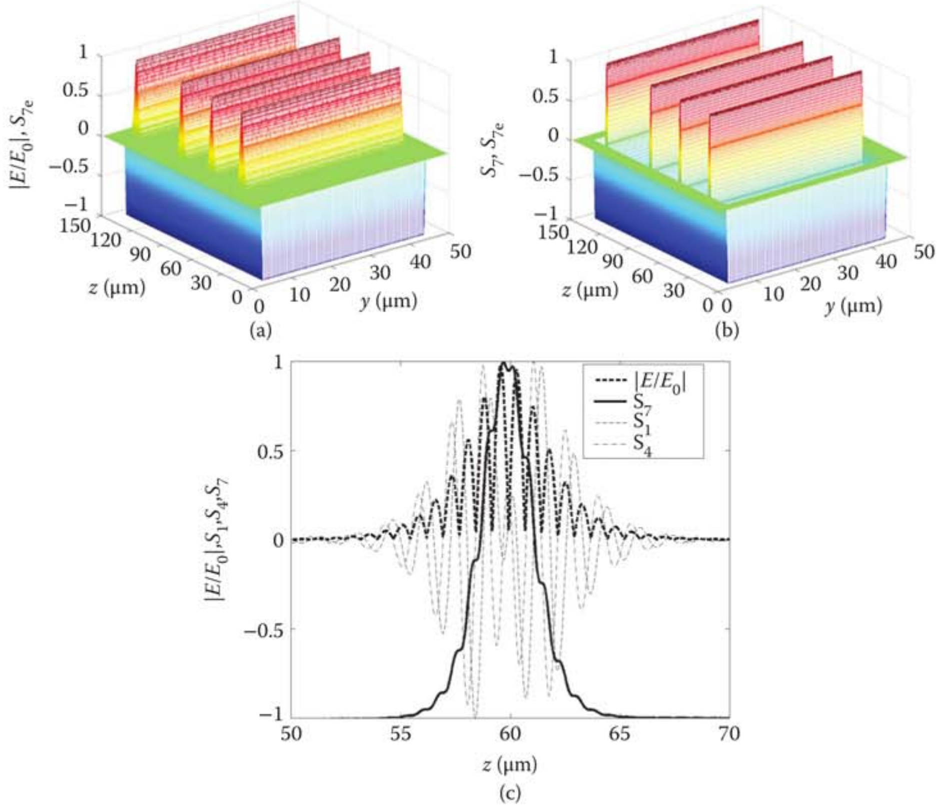


FIGURE 49.12 (a) Superimposed snapshots of the modulus of the electric field for a TEM guided mode propagating in a planar parallel-mirror waveguide at times $t = 150, 250, 350$, and 500 fs after the arrival of the pulse at the left boundary of the waveguide. The initial population profile, S_{7e}, is also shown as an indication of the 2D boundaries of the active medium; the resonant dipoles are initially in the ground state with $S_{7e} = -1$, absorbing the incoming field radiation. (b) Corresponding population inversion term S_7 distribution across the waveguide plane, showing Rabi flops. (c) Slice straight across the middle of the waveguide along the propagation direction, showing the modulus of the electric field, dispersive, and absorptive polarization components, S_1 and S_4, respectively, and the population, S_7, performing a full Rabi flop along the waveguide axis. $\Delta z = 30\,\text{nm}$, $\Delta y = 500\,\text{nm}$, $\Delta t = 9.989 \times 10^{-2}\,\text{fs}$, the pulse wavelength is $\lambda = 1.5\,\mu\text{m}$ and the duration $T_p = 100\,\text{fs}$. (Adapted with permission from G. Slavcheva et al., *Physical Review A*, 66, 063418, 2002. Copyright 2002 by the American Physical Society.)

which depends on the point in the waveguide plane we choose to consider:

$$\theta\left(y, z\right) = \frac{\wp}{\hbar} \int_{-\infty}^{\infty} \left|\tilde{E}\left(y, z, t'\right)\right| \mathrm{d}t' = \frac{\wp}{\hbar} \int_{-\infty}^{\infty} \sqrt{\tilde{E}_y^2 + \tilde{E}_z^2} dt'. \tag{49.41}$$

Assuming that the initial pulse is injected into the active medium at a point $y = y_0$ at the boundary ($z = 0$), stable pulses in the absorbing medium are obtained when

$$\theta\left(y = y_0, z = 0\right) = \frac{\wp}{\hbar} \int_{-\infty}^{\infty} \sqrt{\tilde{E}_y^2\left(y_0, 0, t'\right) + \tilde{E}_z^2\left(y_0, 0, t'\right)} \mathrm{d}t' = 2n\pi, \tag{49.42}$$

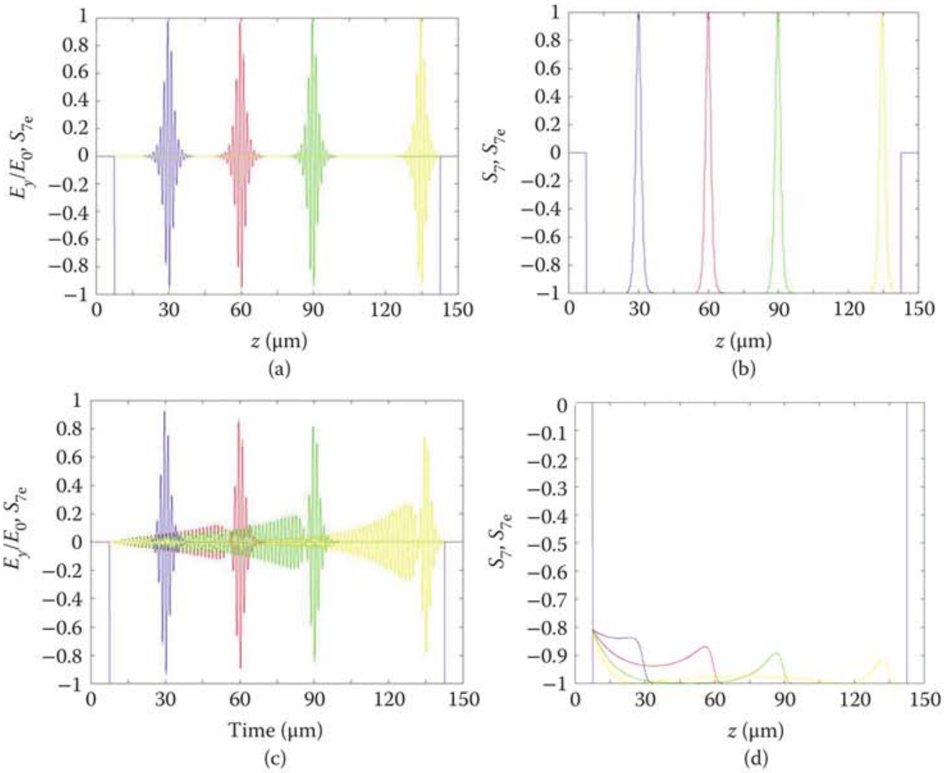


FIGURE 49.13 (a) Time evolution of a 2π initial pulse ($E_0 = 4.2186 \times 10^9\ \mathrm{V \cdot m^{-1}}$) at $t = 150, 250, 350$, and 500 fs after impinging on the left boundary of a planar parallel-mirror waveguide. The soliton propagates undistorted through the resonant medium, the boundaries of which are shown by the initial population profile, $S_{7e} = -1$, i.e., all dipoles are initially in the ground state. (b) The corresponding population inversion term, S_7, exhibits full Rabi flops, whereby the absorbed energy is fully balanced by the emitted energy over one Rabi period. (c) Time evolution of a pulse with initial pulse area 0.9π ($< \pi$) ($E_0 = 1.18957 \times 10^9\ \mathrm{V \cdot m^{-1}}$) at the same simulation times. In this case, the pulse becomes distorted due to imbalance of the absorbed and reemitted energies, and will eventually decay to zero in agreement with the PAT. (d) The resonant system is only partially inverted: there is no population swap between the ground and excited levels. (Adapted with permission from G. Slavcheva et al., *Physical Review A*, 66, 063418, 2002. Copyright 2002 by the American Physical Society.)

with n being an integer. A 2π pulse has an initial pulse amplitude $E_0 = 1.123 \times 10^7\ \mathrm{V \cdot m^{-1}}$. The time evolution of the TM_1 soliton is shown in Figure 49.14.

49.3.5.2 2D Semiconductor Microcavities

Using the 2D model, we numerically investigate conditions of onset of SIT cavity soliton formation in a semiconductor cavity driven by a coherent ultrashort pulse launched from the lower boundary (see Figure 49.11, right panel) [53]. We have designed a number of 2D microcavity structures in order to test the influence of the cavity length on the localization phenomena and pattern formation. A typical 5λ-cavity (where the optical path length between the mirrors of the cavity is equal to an exact multiple of five wavelengths) designed at $\lambda = 1.3$ μm, filled with resonantly absorbing medium sandwiched between a bottom DBR (31.5 pairs of GaAs/AlAs) and a top DBR (24 pairs of $\mathrm{Al_{0.1}Ga_{0.9}As/Al_{0.8}Ga_{0.2}As}$) is shown in Figure 49.15a and b. We apply a driving pulse excitation with duration 100 fs at the left boundary

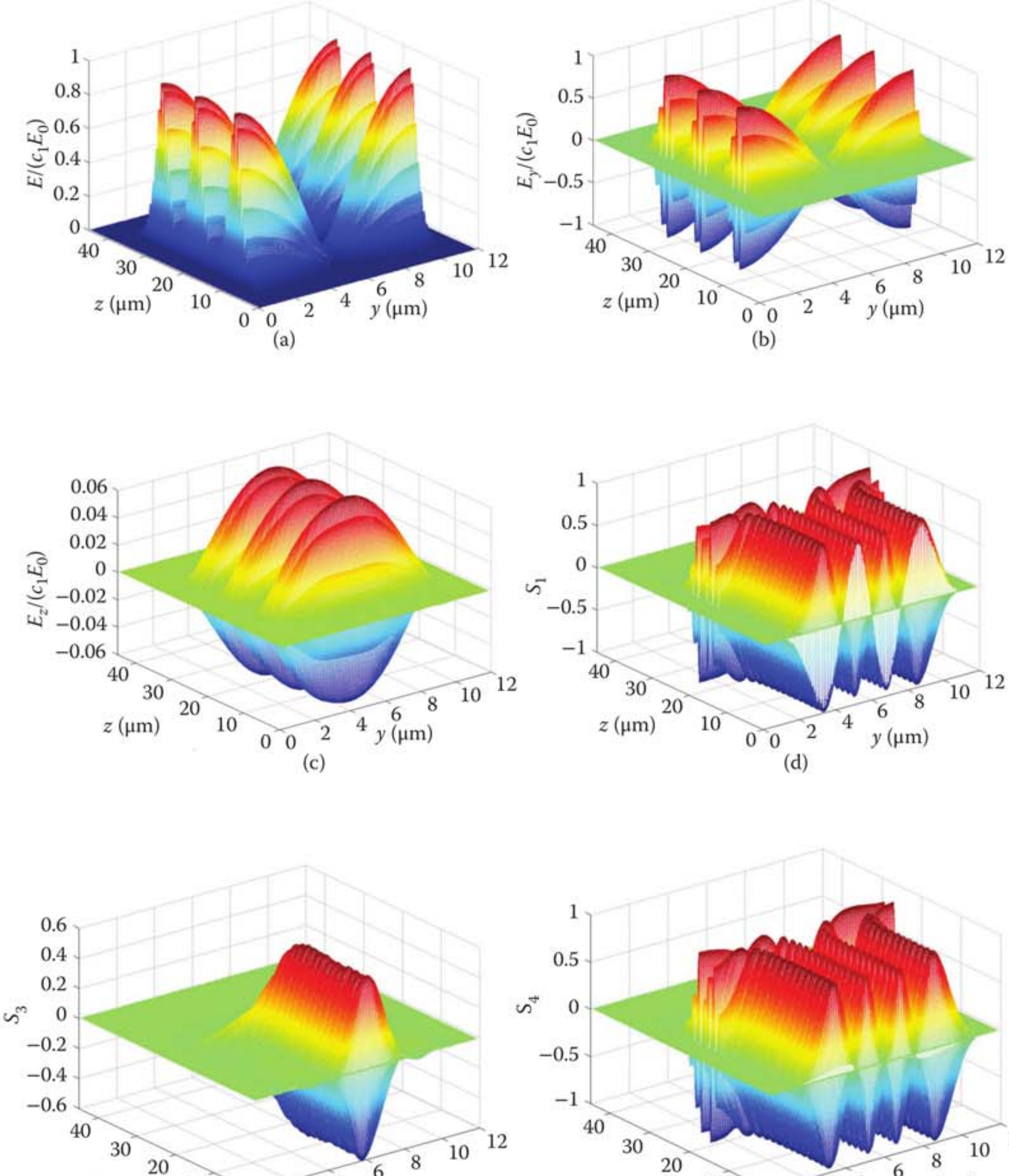


FIGURE 49.14 Time evolution of the (a) modulus, (b) E_y, and (c) E_z electric-field vector component of the TM_1 soliton at times $t = 90, 125$, and 155 fs after its arrival at the left boundary of a planar parallel-mirror waveguide. The 2D pulse travels undistorted in the resonant medium with refractive index, $n = 1$. The waveguide is 50 μm long and 9.9185 μm wide. Discretization grid: $\Delta z = 1.5$ nm; $\Delta y = 119.5$ nm; $\Delta t = 5 \times 10^{-3}$ fs. (d–g) Polarization components of significance (i.e., those corresponding to transitions from one level to another) in a degenerate three-level system. S_1 and S_4: in-phase dispersive components; S_3 and S_6: in-quadrature absorptive components. (*Continued*)

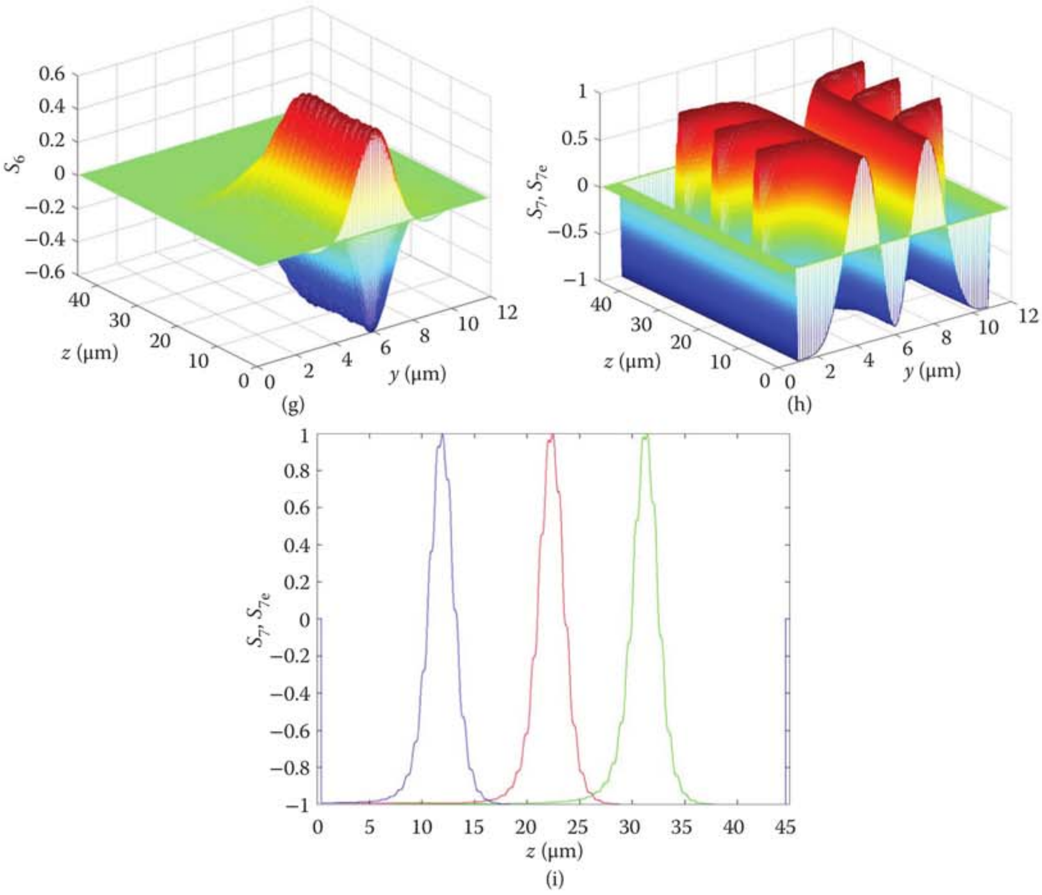


FIGURE 49.14 (*Continued*) (d–g) Polarization components of significance (i.e., those corresponding to transitions from one level to another) in a degenerate three-level system. S_1 and S_4: in-phase dispersive components; S_3 and S_6: in-quadrature absorptive components. (h) Population component S_7 showing population Rabi flops along the waveguide and in a transverse direction (two complete Rabi flops are visible). (i) lengthways cross section of (h) showing complete Rabi flops in the propagation direction, z.

($z = 0$) of the microcavity (Figure 49.15a), which is 12 μm long and $w = 9.92$ μm wide. The dipole number density is $N_a = 10^{24}\,\mathrm{m}^{-3}$ and the dipole matrix element, $\wp = 1 \times 10^{-29}\ \mathrm{C \cdot m}$. The relaxation times are assumed uniform and equal to 100 ps, satisfying the SIT criterion $T_\mathrm{p} \ll T_i,\ i = 1, 2, \ldots, 8$. The initial pulse amplitude $E_0 = 4.8643 \times 10^9\ \mathrm{V \cdot m^{-1}}$ is calculated for a pulse area of 2π, assuming that the PAT still holds. As pointed out in [54–56], this is not a necessary condition for SIT in multilayered media. Even a relatively weak pulse would lead to SIT intracavity patterns after sufficiently long simulation times, due to field enhancement by constructive interference from multiple reflections in the DBR mirrors. The population inversion gradually builds up with time and begins to perform Rabi flopping. In Figure 49.15c through f, the electric-field modulus and the corresponding population inversion profile are plotted for the cavity region. The clearly visible instantaneous roll patterns for the electric field and population inversion, S_7, feature a quasi-stationary standing-wave (i.e., quiescent) SIT soliton. The cavity induces a quasi-stationary population "grating", and the number of the quasi-complete Rabi flops (10) of the population equals the number of half-wavelengths along the cavity length, $L_c = (10\lambda)\,/\,(2n)$. At the

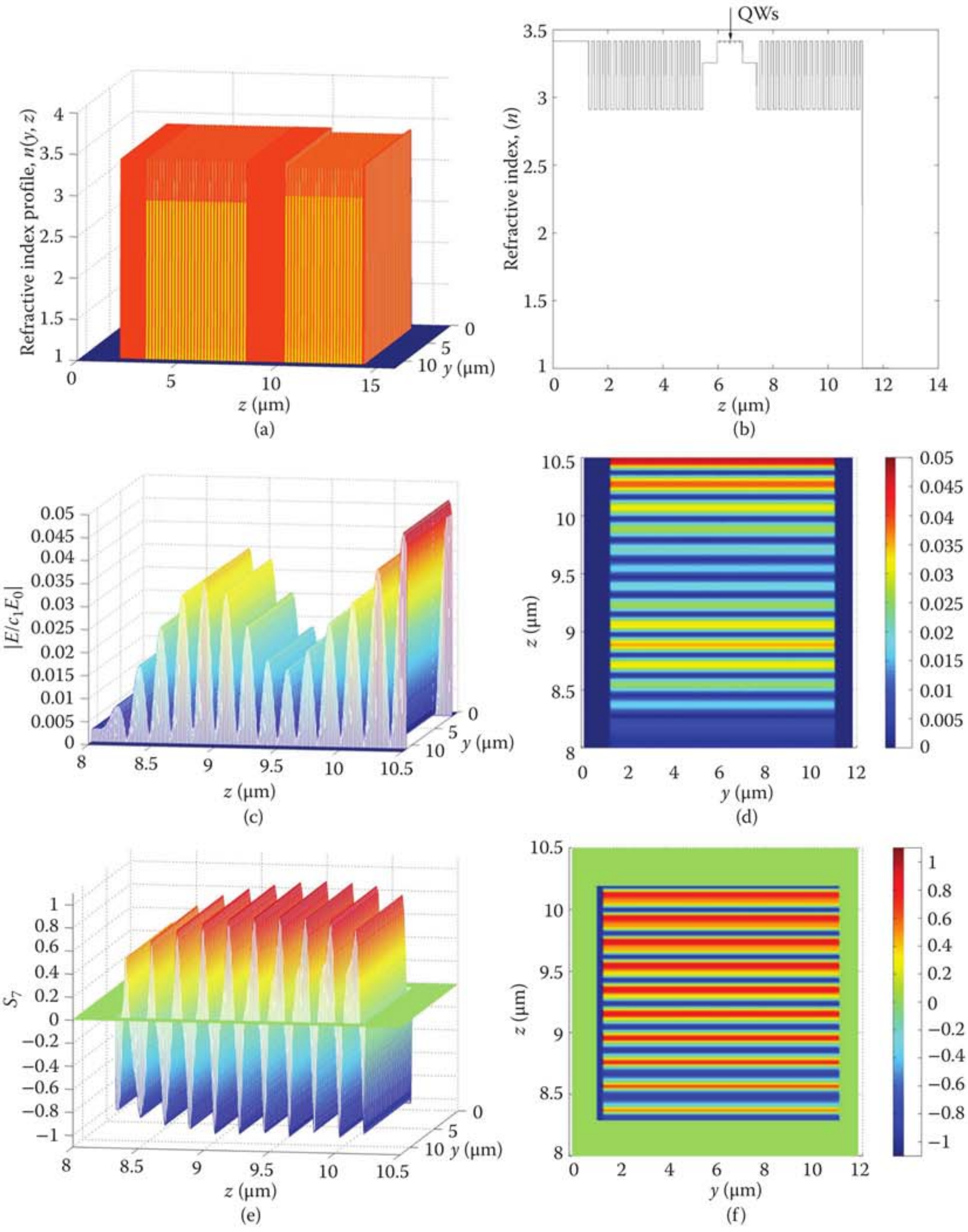


FIGURE 49.15 (a) 2D refractive-index profile across the semiconductor microcavity. (b) Longitudinal slice of (a) at a distance $y_0 = 6$ μm from the left interface, $y = 0$. (c) A snapshot at time $t = 3.6055 \times 10^{-13}$ s of the electric-field modulus distribution across the cavity. (d) Top view of (c) showing the roll patterns formed. (e) Snapshot of the corresponding population inversion profile across the cavity, showing 10 complete Rabi flops. (f) Top view showing the intracavity population "grating" (roll pattern). (G. Slavcheva et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2003 IEEE.) (*Continued*)

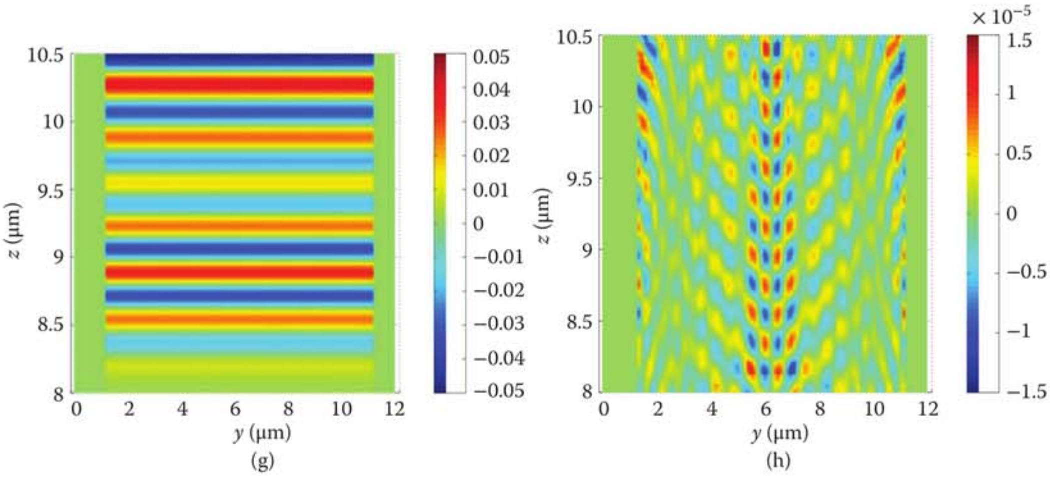


FIGURE 49.15 (*Continued*) (g) Top view of the E_y-component roll pattern. (h) Top view of the 2D pattern of the E_z electric-field component. (G. Slavcheva et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2003 IEEE.)

same time, a roll pattern for the E_y electric-field component, Figure 49.15g, and a more elaborate 2D one, Figure 49.15h, for the E_z field component are observed. Performing simulations with different driving pulse durations on a number of different microcavity designs, we demonstrate numerically the existence of a new type of 2D spatiotemporally localized standing-wave soliton ("light bullet"). The SIT cavity soliton results from the interplay of the resonant nonlinearity and the Bragg reflection interference and dispersion in the microcavity.

49.4 Optical Orientation in Quantum Wells and QDs

Circularly polarized optical excitations of semiconductor nanostructures excite particular interest due to their ability to transfer angular momentum and thus to drive the quantum system to a specific angular-momentum state. The conversion between the flying qubit encoded in photon polarization and the material qubit encoded, for example, in the carrier spin, is known as optical orientation, and this technique has been successfully applied to spin-polarized states in semiconductors [14]. This, in turn, opens up the possibility for faithful transmission of flying qubits between distant locations. If sufficiently short (and therefore high-intensity) optical pulses are used, so that they can interact with the system before it can be affected by its environment, the spin does not have time to decohere (i.e., lose its phase information) and could be coherently manipulated by such light pulses. In this high-intensity, coherent regime, optical Rabi oscillations take place; one-half of each oscillation consists of the population being driven from the ground to the excited state (or vice versa), hence performing a Rabi flop. A Rabi flop corresponds to a one-qubit rotation, whereby the spin is rotated through π; this entails a transition either from spin-down (state 0) to spin-up (state 1) or vice versa (Figure 49.1). The achievement of control and manipulation of the single charge-carrier spin dynamics and the collective many-body dynamics of multiple entangled spins is the key to the physical implementation of quantum computation.

49.4.1 Circularly Polarized Optical-Pulse Interactions with Two-Level Quantum Systems: Selective Spin Initialization and Readout

In bulk semiconductors and in semiconductor structures with reduced dimensionality, dipole optical transitions (including excitonic transitions) between energy levels are only allowable subject to satisfying the angular-momentum optical selection rules, $\Delta J_z = 0, \pm 1$, for linear and circular photon polarization respectively, where J_z is the total angular-momentum projection on the electromagnetic wave propagation direction coinciding with the QD growth axis. We shall consider optical transitions with $\Delta J_z = \pm 1$ that are excited by circularly (or elliptically) polarized light, and will therefore apply the pseudospin equations derived for circularly polarized light in Equation 49.8, coupled to Maxwell's vector equations for an electromagnetic wave propagating along the z-direction, circularly (elliptically) polarized in a plane perpendicular to z [57]:

$$
\begin{aligned}
\frac{\partial H_x(z,t)}{\partial t} &= \frac{1}{\mu}\frac{\partial E_y(z,t)}{\partial z}\\
\frac{\partial H_y(z,t)}{\partial t} &= -\frac{1}{\mu}\frac{\partial E_x(z,t)}{\partial z}\\
\frac{\partial E_x(z,t)}{\partial t} &= -\frac{1}{\varepsilon}\frac{\partial H_y(z,t)}{\partial z} - \frac{1}{\varepsilon}\frac{\partial P_x(z,t)}{\partial t}\\
\frac{\partial E_y(z,t)}{\partial t} &= \frac{1}{\varepsilon}\frac{\partial H_x(z,t)}{\partial z} - \frac{1}{\varepsilon}\frac{\partial P_y(z,t)}{\partial t}.
\end{aligned}
\tag{49.43}
$$

We phenomenologically construct the Hamiltonian corresponding to initially absorbing energy-level schemes (i.e., $S_{30} = -1$), shown in Figure 49.16:

$$
\hat{H} = \hat{H}_0 + \hat{H}_{\text{int}} = \hbar \begin{pmatrix} 0 & -\frac{1}{2}\left(\Omega_x - i\Omega_y\right) \\ -\frac{1}{2}\left(\Omega_x + i\Omega_y\right) & \omega_0 \end{pmatrix}. \tag{49.44}
$$

The Hamiltonian for an initially amplifying two-level system (i.e., $S_{30} = 1$) is the transposed matrix. Here we have introduced a complex Rabi frequency, whose real and imaginary parts correspond to the $E_{x(y)}$ electric-field vector components, $\Omega_x = \frac{\wp}{\hbar}E_x$ and $\Omega_y = \frac{\wp}{\hbar}E_y$.

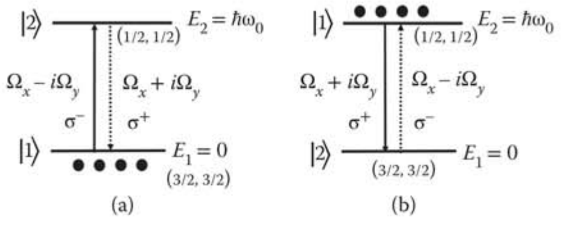


FIGURE 49.16 Energy-level diagram of (a) initially absorbing (i.e., entire charge-carrier population in the ground state) and (b) amplifying (i.e., entire charge-carrier population in the excited state) two-level system, with energy-level separation $\hbar\omega_0$ and allowed dipole optical transitions subject to the condition $\Delta J_z = \pm 1$. States are additionally labeled (J, J_z), where J is the total angular momentum and J_z is its projection along z. (a) A left circularly polarized optical field, σ^-, drives the population into the upper excited level, $|2\rangle$; the population is returned to the ground level by an optical pulse with opposite helicity, σ^+. (b) Having been initially prepared in the excited state, $|1\rangle$, the population in the system is driven by a right circularly polarized pulse, σ^+, to the ground state $|2\rangle$ and returned to the initial excited state by a left circularly polarized pulse, σ^-.

Employing the general methodology in Section 49.2.3, we can easily show that

$$\begin{aligned} P_x &= -\wp N_a S_1, \\ P_y &= -\wp N_a S_2 . \end{aligned} \tag{49.45}$$

We solve numerically by the FDTD method the system of equations, Equations 49.43 and 49.8, coupled through Equation 49.45. The source optical fields for left and right circularly polarized light are each given by two orthogonal, linearly polarized pulses, one time-delayed relative to the other by $\pi/2$, with carrier frequency ω_0 tuned in resonance with the two-level system:

$$\sigma^- \left\{ \begin{array}{l} E_x(z=0,t) = \tilde{E}_x(t)\cos(\omega_o t) \\ E_y(z=0,t) = -\tilde{E}_y(t)\sin(\omega_o t) \end{array} \right. ; \ \sigma^+ \left\{ \begin{array}{l} E_x(z=0,t) = \tilde{E}_x(t)\cos(\omega_o t) \\ E_y(z=0,t) = \tilde{E}_y(t)\sin(\omega_o t), \end{array} \right. \tag{49.46}$$

where the pulse shape is determined by the envelopes:

$$\tilde{E}_{x,y}(t) = \left\{ \begin{array}{l} E_0 \operatorname{sech}(10\Gamma) \\ E_0 \exp\left[-\left(t-t_0\right)^2 / t_d^2\right], \end{array} \right. \tag{49.47}$$

which we choose to be either hyperbolic secant or Gaussian. We consider a 15 μm-long simulation structure filled with either a resonantly absorbing or amplifying ($S_{30} = -1$ or $S_{30} = +1$) two-level medium, assumed to be GaAs (refractive index, $n = 3.3827$). We use a pulse wavelength, $\lambda_0 = 1.5\,\mu\text{m}$ and a pulse duration, $T_\text{p} = 30\,\text{fs}$; the dipole coupling $\wp = 1 \times 10^{-29}\,\text{C} \cdot \text{m}$, and two-level system ensemble density $N_a = 1 \times 10^{24}\,\text{m}^{-3}$ are taken from [47]. We restrict ourselves to the linear regime (initial electric-field amplitude, $E_0 = 1\,\text{V} \cdot \text{m}^{-1}$). The material relaxation times are $T_1 = 100\,\text{ps}$ and $T_2 = 70\,\text{fs}$. We launch a left circularly polarized pulse, σ^-, from the left boundary of the simulation domain and sample the electric field at two locations, z_1 and z_2, along the active medium. The complex propagation factor (see Equation 49.29) over one dielectric wavelength $l = z_2 - z_1 = \lambda/n$ is calculated for the E_x and E_y components, and the results for the gain/absorption coefficient spectra are compared with the steady-state density-matrix theory spectra (see Equation 49.30). A snapshot of the electric-field components, E_x and E_y, at the input and output boundaries are shown in the first column of Figure 49.17. The corresponding Fourier spectra and gain/absorption coefficients are shown in the second and third columns, respectively, of Figure 49.17 for an initially absorbing (Figure 49.17a through f) and initially amplifying (gain) medium (Figure 49.17g through l). The optical wave interacts with the two-level system, driving the population into the upper state (or returning the population to the ground state) by absorbing (or emitting) pulse energy. Therefore, we clearly observe gain/absorption in the Fourier spectra. The gain/absorption coefficients coincide with those for linear polarization. In addition, we have shown that when pumping an absorbing system with a right circularly polarized pulse, the pulse fails to excite the system, and similarly, when pumping an initially inverted system (the whole population residing on the upper level) by a left circularly polarized pulse, the system remains unaffected. Therefore our model correctly describes the excitation using circularly polarized light of dipole optical transitions subject to the optical selection rules, $\Delta J = \pm 1$, in a two-level system. We have thus numerically demonstrated selective excitation of specific spin states by circularly polarized pulses with predefined helicity.

49.4.1.1 SIT and Polarized SIT Soliton Formation

The results in Section 49.4.1 were obtained in the linear (small-signal) regime. Let us now consider the nonlinear coherent regime of ultrashort pulse propagation through resonant media, whose atomic lifetimes greatly exceed the pulse duration (i.e., such that $T_1,\ T_2 \gg T_\text{p}$). The simulation domain is 150 μm long and a two-level resonantly absorbing medium with refractive index $n = 1$ is embedded between two symmetric free-space regions, each with length 7.5 μm. We inject a resonant, left circularly polarized pulse,

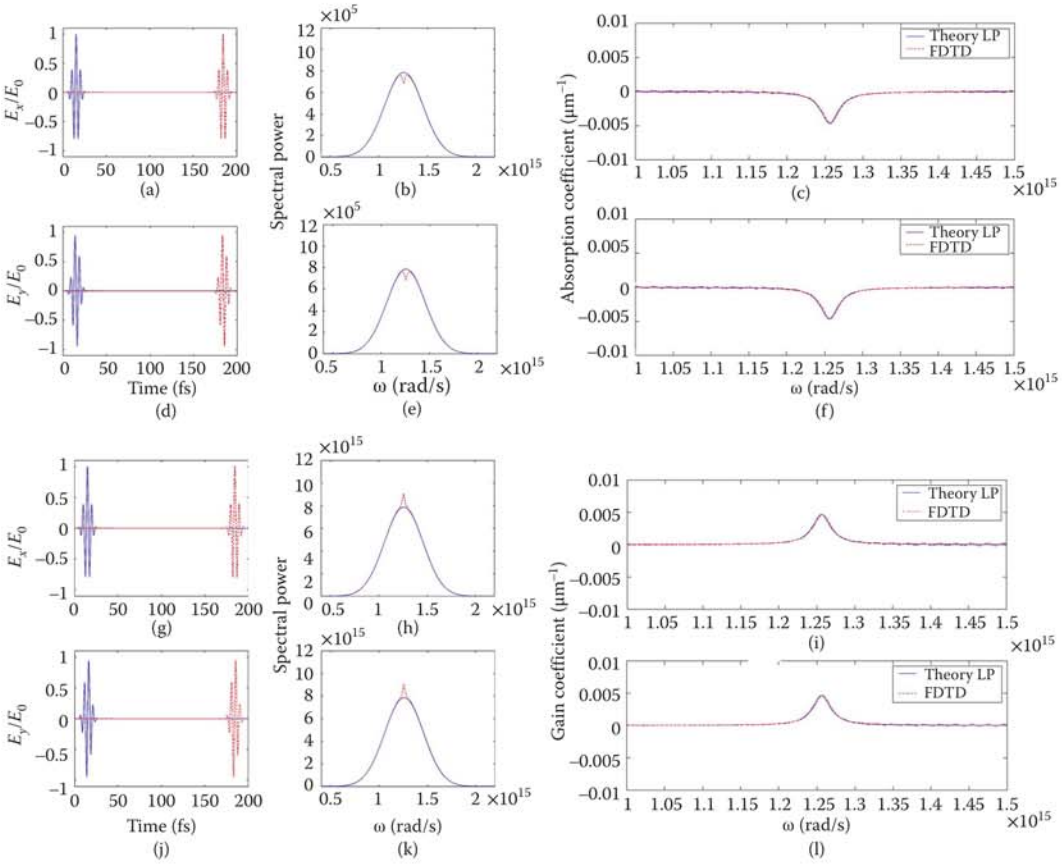


FIGURE 49.17 Time evolution at the input (solid line) and output (dashed line) structure boundaries of the electric-field components, E_x (a) and E_y (d), of a circularly polarized optical pulse with unit amplitude and pulse duration 30 fs for an initially absorbing two-level medium. (b), (e) Fourier spectra corresponding to the E_x and E_y field components, respectively, showing absorption of the pulse (dashed curve dip). (c), (f) Absorption coefficients for E_x and E_y, respectively; in this case, the dashed curve indicates the numerically calculated coefficients for linear polarization, while the solid curve signifies the analytically computed ones (Equation 49.30). (g–l) Respective plots for an initially amplifying (gain) medium; notations are the same as in (a–f). The gain is clearly visible in the Fourier spectra in (h) and (k) (see overshoot of dashed curve). (Adapted with permission from G. Slavcheva and O. Hess. *Physical Review A*, 72, 053804, 2005. Copyright 2005 by the American Physical Society.) (FDTD, finite-difference time-domain.)

σ^-, with hyperbolic secant envelope, since this is the well-known stable soliton solution of the Maxwell-Bloch system (see, e.g., [58]). We choose the electric-field component amplitudes so that the pulse area, according to the PAT, is 2π. This condition implies that $E_0 = 4.2186\times10^9\ \mathrm{V\cdot m^{-1}}$. The time evolution of a 2π pulse is shown in Figure 49.18a. Both circularly polarized pulse components travel undistorted through the resonantly absorbing two-level medium as a solitary wave, and the population inversion is driven through complete Rabi flops by the leading and trailing pulse edges. An expanded view of a snapshot of a circularly polarized SIT soliton at the simulation time $t = 0.2$ ps is shown in Figure 49.18b. By virtue of the PAT, stable soliton solutions in gain (initially inverted, $S_{30} = 1$) medium are obtained for pulse areas which are odd multiples of π. In order to confirm this prediction, we launch a right circularly polarized pulse with amplitude $E_0 = 2.1093\times10^9\ \mathrm{V\cdot m^{-1}}$, which corresponds to a pulse area of π. As expected, a π-pulse completely de-excites the population back to the ground state (Figure 49.18c); the pulse amplitude increases,

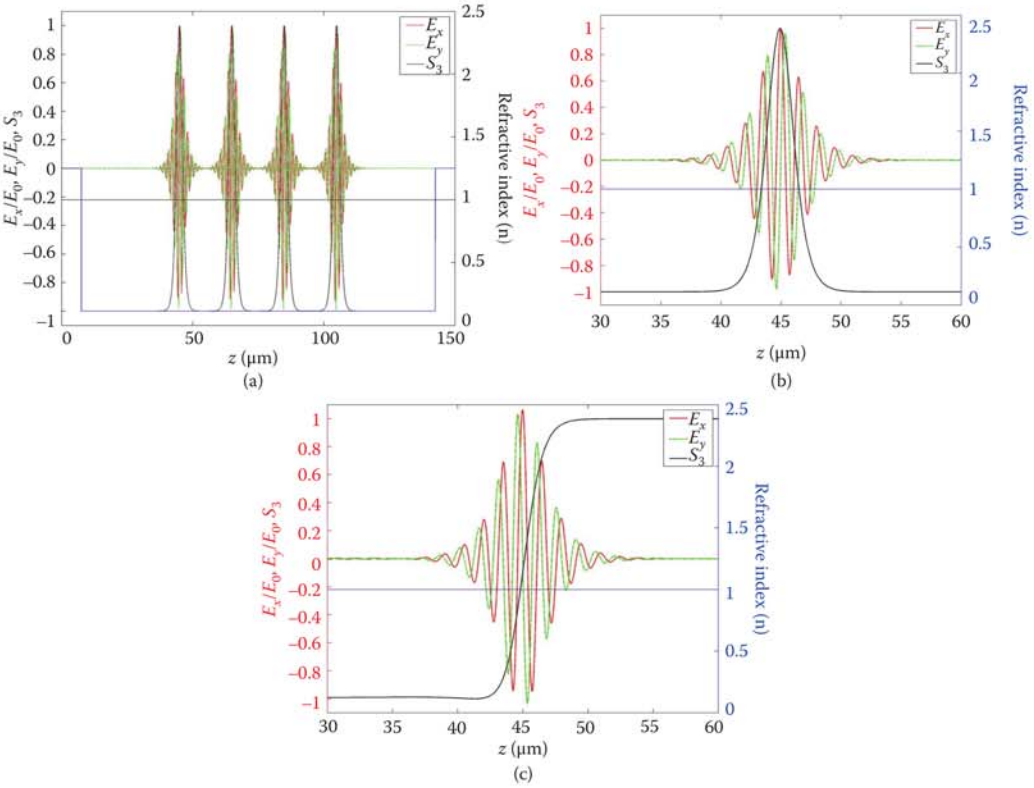


FIGURE 49.18 (a) Time evolution of a 2π, σ^- pulse propagating (from left to right) in a resonantly absorbing medium. The region where the initial charge carrier population profile is $S_{30} = -1$ demarcates the extent of the active medium on the diagram. The normalized electric-field components, E_x and E_y, and the instantaneous population profile, S_3, are plotted at the simulation times $t = 200, 267, 334$ and 400 fs from the moment the pulse is injected at the left boundary of the medium. The pulse travels without loss: it is a SIT soliton. (b) Expanded view of the σ^- SIT soliton in (a) at $t = 267$ fs showing the full Rabi flop of the population inversion. (c) A snapshot of a σ^+ SIT π pulse at $t = 267$ fs that completely inverts the population in the resonantly amplifying medium. (Adapted with permission from G. Slavcheva and O. Hess. *Physical Review A*, 72, 053804, 2005. Copyright 2005 by the American Physical Society.)

while the pulse duration decreases, thereby conserving the pulse area during the pulse propagation in the amplifying medium.

49.4.2 Circularly Polarized Optical Pulse Interactions with Degenerate Four-Level Systems

Let us consider now a more realistic discrete-level quantum system, such as excitonic transitions in bulk or low-dimensional semiconductor systems. For instance, we consider fundamental heavy-hole dipole optical transitions in QDs or quantum wells (QWs), whereby the band structure in the vicinity of the Γ-point ($\mathbf{k} = 0$) in the Brillouin zone (BZ) is isomorphic with a pair of two-level systems corresponding to σ^- and σ^+ heavy-hole excitonic transitions, as shown in Figure 49.19.

Consider a plane electromagnetic wave propagating along z direction, circularly polarized in a plane perpendicular to z, and tuned in resonance with the fundamental heavy-hole transition of the semiconductor system. The resonant nonlinearity is modeled by an ensemble of degenerate four-level systems of the type shown in Figure 49.19, with density N_a. The time-evolution of a four-level system under a dipole-coupling

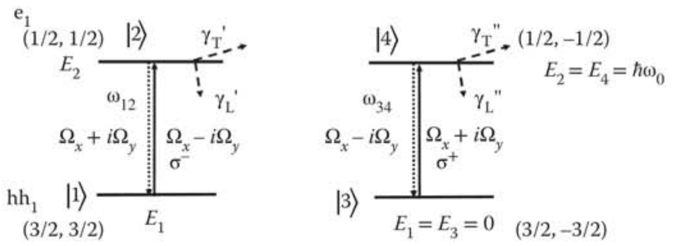


FIGURE 49.19 Energy-level diagram of σ^- and σ^+ heavy-hole transitions in a quantum well near the Γ-point ($\mathbf{k} = 0$) or between s-shell heavy-hole and electron levels in a QD. The energy separation between the upper and lower levels is $\hbar\omega_0$, and the allowed dipole transitions satisfy $\Delta J_z = \pm 1$. Ω_x and Ω_y are, respectively, the real and imaginary part of the Rabi frequencies associated with the coherent transitions. γ_L and γ_T are, respectively, the longitudinal relaxation and transverse dephasing rates of the excited states. e_1 is the bottommost electron energy level in the conduction band, and hh_1 is the topmost heavy-hole energy level in the valence band; together, these energy levels make up the fundamental bandgap.

external perturbation is described in terms of the SU(4) group generators by an equation of motion for a 15-dimensional real state pseudospin vector (see Equation 49.12 for $j, k, l = 1, 2, \ldots, 15$), where the dissipation in the system is accounted for by introducing nonuniform relaxation times, T_j. Without making any assumptions about the initial population redistribution between the levels, the system Hamiltonian is given by [59]

$$\hat{H} = \hbar \begin{pmatrix} 0 & -\frac{1}{2}\left(\Omega_x - i\Omega_y\right) & 0 & 0 \\ -\frac{1}{2}\left(\Omega_x + i\Omega_y\right) & \omega_0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2}\left(\Omega_x + i\Omega_y\right) \\ 0 & 0 & -\frac{1}{2}\left(\Omega_x - i\Omega_y\right) & \omega_0. \end{pmatrix}. \tag{49.48}$$

We investigate the nonlinear coherent propagation regime, where the SIT condition $T_\mathrm{p} \ll T_j$ is satisfied with $T_\mathrm{p} = 100$ fs and $T_j = 100$ ps. A circularly polarized pulse, amplitude-modulated by a hyperbolic secant function and whose carrier-wave frequency is resonant with the optical transition (at $\lambda = 1.5\,\mu\mathrm{m}$), Equation 49.46 and 49.47, is applied at the left boundary $z = 0$ of a 150-μm-long simulation domain, with either resonantly absorbing or amplifying media embedded between two free-space regions of thickness 7.5 μm each. We choose the initial pulse to be a π-pulse, giving $E_0 = 1.093 \times 10^9\ \mathrm{V \cdot m^{-1}}$ according to the PAT. We wish to achieve selective excitation of dipole optical transitions with $\Delta J_z = -1$ or $\Delta J_z = +1$, with a view to manipulate a particular spin state simply by selecting the helicity of the impinging optical pulse. The two leftmost columns of Figure 49.20 show all four possible cases where the initial state population of each two-level system is either entirely in the ground or entirely in the excited state. The two rightmost columns of the same figure also show the outcome of pulse excitation of each respective system. The conclusion that can be drawn is that we can indeed coherently control the spin population of specific spin states in the four-level system by choosing the proper polarization of the injected pulse. We have numerically demonstrated SIT and SIT soliton propagation in a four-level system initially prepared in a state with uniformly distributed populations in either the ground or excited states of the two two-level systems [60]. The ultrashort pulse travels undistorted in the resonant, degenerate four-level medium, Figure 49.21a, and the population is locally driven by the electric field through full Rabi flops simultaneously in both two-level systems, Figure 49.21b (compare with the two-level system soliton, Figure 49.18b).

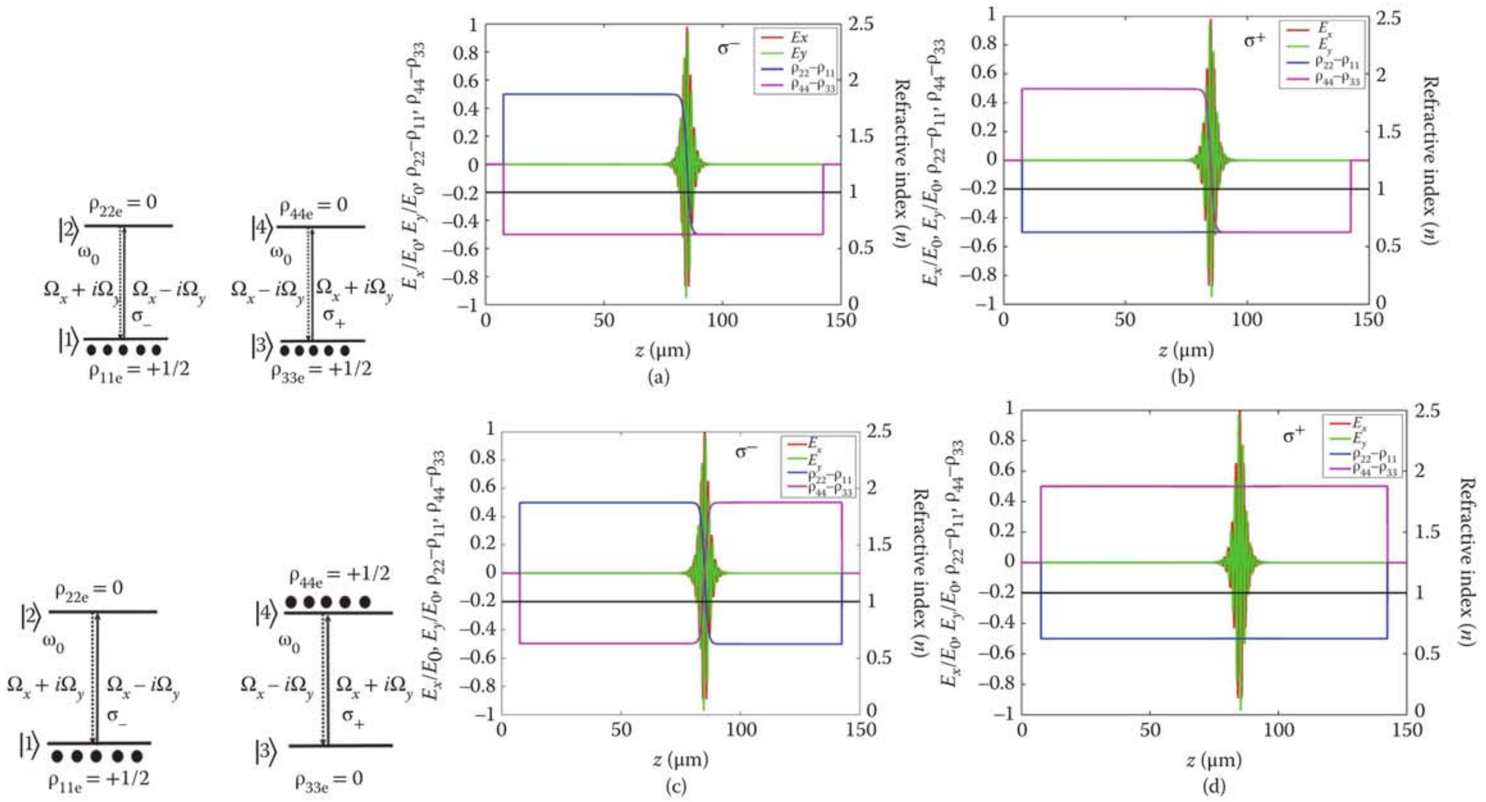


FIGURE 49.20 A snapshot of a circularly polarized pulse at time $t = 333$ fs after injection of the pulse at the medium's left boundary. E_x and E_y electric-field components and the population inversion $\rho_{22} - \rho_{11}$ in the first, and $\rho_{44} - \rho_{33}$ in the second two-level systems are plotted, tracing out the boundaries of the active medium with refractive index, $n = 1$. The leftmost two columns display the initial population distribution between the levels (equal split of the population between the two sets of two-level systems is assumed at the outset). (a) A σ^- π-pulse completely excites the first two-level system, but does not affect the second system. (b) A σ^+ π-pulse completely de-excites the initially inverted second two-level system, but does not affect the first one. (c) A σ^- π-pulse completely excites the first and de-excites the second system. (d) A σ^+ π-pulse does not affect either of the systems: the population in the first remains in the ground state and the population of the second remains in the excited state. (*Continued*)

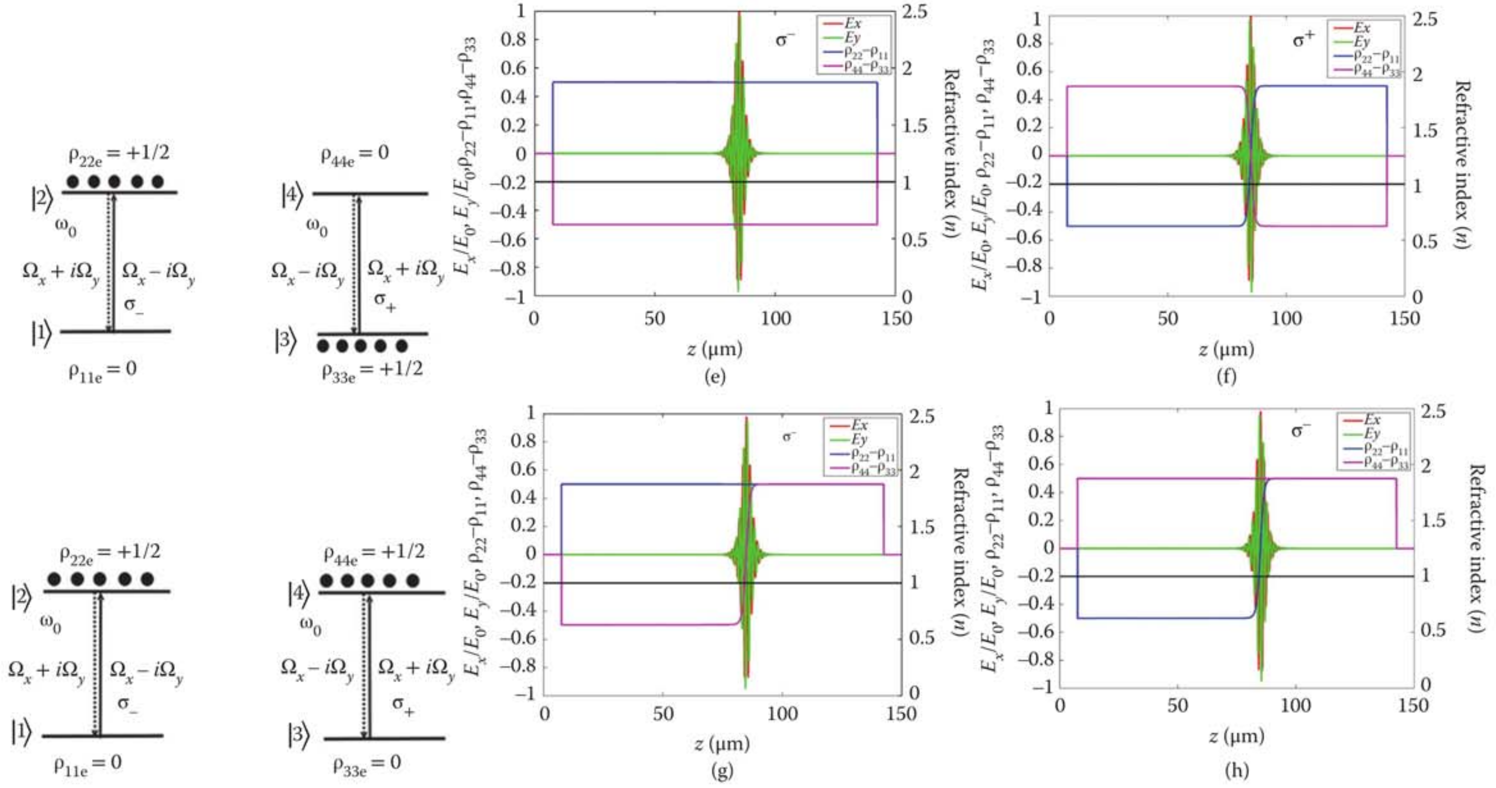
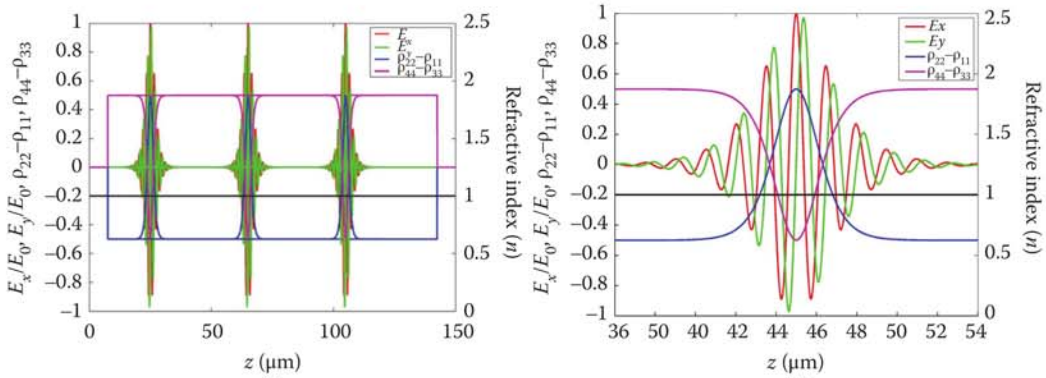


FIGURE 49.20 (*Continued*) (e) A σ^- π-pulse does not change the population in either of the systems. (f) A σ^+ π-pulse inverts simultaneously the population in both systems. (g) A σ^- π-pulse does not affect the first system, but de-excites the second. (h) A σ^+ π-pulse inverts the population in the first system, but does not affect the second system.



(a)

(b)

FIGURE 49.21 (a) Polarized SIT soliton propagation through a four-level medium resonant with the carrier wave of the injected pulse. The diagram shows a superimposition of the soliton at three different times: $t = 133, 266$ and 400 fs from the moment the pulse is injected at the left boundary of the medium. (b) A snapshot of the polarized soliton in the four-level medium, consisting of two pairs of two-level systems, at $t = 200$ fs. The medium is initially prepared in a state with the charge-carrier population residing in the lower-lying level, $|1\rangle$, of the first two-level system and in the upper level, $|4\rangle$, of the second two-level system (see Figure 49.19).

49.5 Modeling Coherent Spin Generation, Manipulation, and Readout in Charged QDs

In Section 49.4, we demonstrated coherent, selective spin excitation dependent on the helicity orientation of an impinging ultrashort optical pulse. In this section, we will show how spin states in low-dimensional semiconductor structures can be generated, manipulated, and read out using ultrashort circularly polarized pulses. Our goal will be to compute the time evolution of a discrete-level system and map the allowed dipole optical transitions that occur in a semiconductor nanostructure (e.g., a single QD or an ensemble of QDs) upon ultrashort, circularly polarized optical excitation. The pulse leaves behind a long-lived coherence, and we shall look for characteristic signatures in the dynamics that would allow us to determine the initial spin state and find useful regimes and schemes for high-fidelity optical manipulation of the spin confined in a QD. One promising approach is to address optically individual carrier spins in semiconductor QDs and manipulate them through optically excited states (charged excitons) by employing the techniques of coherent quantum control and optical orientation.

49.5.1 Single-Charged QD Embedded in a Waveguide

We will study n-doped QDs with an overall negative charge of a single electron. Injection of a single electron into a QD can be achieved either by modulation doping of the barrier region, adjusting the impurity doping within a delta-doped layer to transfer on average one electron per dot to the lowest energy states, or by electrical injection. A large class of self-assembled QDs posses quasi-cylindrical geometry about the quantization axis, z, and the electron and hole single-particle states in the conduction and valence band, respectively, can be approximated by harmonic oscillator potentials [61] (Figure 49.22). The resonant, circularly polarized optical excitation of a charged QD leads to the formation of a three-particle complex—a trion—consisting of two electrons occupying the same lowest conduction-band electron level in a spin singlet state, and a hole occupying the lowest valence-band hole level (Figure 49.22c and d). There is strong evidence from both experiment [62] and theory [63,64] that an intense resonant excitation of the trion transition suppresses the electron spin relaxation due to hyperfine interaction of the electron spins with

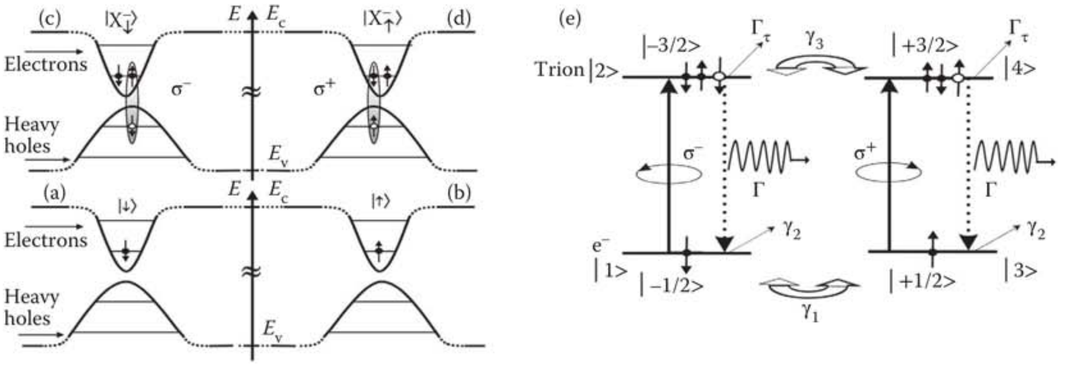


FIGURE 49.22 (a), (b) Schematic of the initial electron states (spin down and spin up) in a singly negatively charged QD and (c), (d) the ground singlet negatively charged exciton, X^-, which forms upon left circularly polarized, σ^-, and right circularly polarized, σ^+, optical excitation of the heavy-hole excitonic transitions. (e) Energy-level diagram of a negatively charged trion (a three-particle complex of two electrons and a hole), equivalent to (a–d); the levels are labeled with the J_z quantum number, and the allowed dipole transitions correspond to $\Delta J_z = \pm 1$ with an energy gap of $\hbar\omega_0$. [Solid upward arrows, coherent optical transitions excited by σ^- or by σ^+; dashed arrows, spontaneous emission rate, Γ; double-headed arrows, transitions due to electron- and hole-spin relaxation, with rates γ_1 and γ_3, respectively; γ_2 and Γ_τ, spin decoherence rates for electrons and holes, respectively.]

the lattice ions' nuclear spins. The latter limits the ability to measure accurately the electron spin orientation at low temperatures. It is obvious that the longer the spin lifetime, the better the chances for optical manipulation and readout of that spin state. We shall investigate both low- and high-intensity nonlinear excitation regimes.

In our treatment of a single quantum object, we shall assume that the ergodic hypothesis holds—i.e., the equivalence of time averages of an observable (in this case, the single-dot polarized, time-resolved photoluminescence TR [PL]) and the ensemble averages at one time over a large number of identical replicas of the dot system. Under such conditions, quantum mechanics allows the prediction of the single-system properties on the basis of macroscopic ensemble averages. Therefore, we shall assume that the time dependence of the optically induced coherent spin generation and subsequent relaxation in a single QD, averaged over a large number of successive measurements, is equivalent to the corresponding spin dynamics of an ensemble of identically prepared QDs. The ensemble of degenerate four-level systems describing the QD resonant nonlinearities is resonantly coupled to an optical wave propagating along the QD growth axis, z, and circularly polarized in a plane perpendicular to z. We use our general methodology to tackle this problem [44] and solve the 1D Maxwell's curl equations (Equation 49.43) coupled to the master pseudospin equations for a four-level system (Equation 49.12) through the induced macroscopic polarization given by

$$\begin{aligned} P_x &= -\wp N_a S_1 \\ P_y &= -\wp N_a S_7. \end{aligned} \tag{49.49}$$

The system under investigation is a GaAs/AlGaAs self-assembled, molecular beam epitaxy (MBE)-grown QD with a height of 5 nm and refractive index, $n_{\mathrm{GaAs}} = 3.63$, sandwiched between two $Al_{0.3}Ga_{0.7}As$ barrier regions, each with a width of 50 nm and refractive index, $n_{\mathrm{AlGaAs}} = 3.46$, at the trion fundamental transition resonance wavelength of 757 nm. The circularly polarized pulse, whose carrier wave is resonant with the trion transition, is described by a hyperbolic secant envelope (see Equation 49.46 and first line of Equation 49.47) and has a duration of 1.3 ps. The simulations are run repeatedly with a different initial pulse amplitude, E_0, which is varied from a low value, $E_0 = 550\,\mathrm{V \cdot m^{-1}}$, through $E_0 = 3 \times 10^6\,\mathrm{V \cdot m^{-1}}$, representing a π-pulse completely inverting the spin population in the system, to $E_0 = 4 \times 10^7\,\mathrm{V \cdot m^{-1}}$, corresponding to a pulse area of 12π and inducing six full Rabi flops. The fundamental trion transition dipole moment is estimated to be $\wp = 4.8 \times 10^{-28}\,\mathrm{C \cdot m}$, and the 3D number density of the resonant four-level dipoles, $N_a = 2.5 \times 10^{24}\,\mathrm{m^{-3}}$, is calculated from the QD surface density to give on average one dot within the microscopic single-dot volume. The trion recombination (spontaneous emission) rate, $\Gamma = 400\,\mathrm{ps^{-1}}$, is taken from [62]; the electron spin-flip rate due to hyperfine interaction with the lattice ions' spins, $\gamma_1 = 0.5\,\mathrm{ns^{-1}}$, has been calculated in [65,66]; the hole spin-flip relaxation is widely assumed to be due to phonon-assisted processes, and an estimate of $\gamma_3 \sim 170\,\mathrm{ps^{-1}}$ [67,68] has been given. The trion-bound electron and hole spin phonon-assisted decoherences are taken as $\gamma_2 = 450\,\mathrm{ps^{-1}}$ and $\Gamma_\tau = 340\,\mathrm{ps^{-1}}$ [69]. We shall assume that the initial spin population resides either in state $|1\rangle$ with spin-down, or in state $|3\rangle$ with spin up. We initially excite the $|1\rangle \rightarrow |2\rangle$ transition by a σ^- pulse. The simulation results for the time evolution of the pulse's electric-field vector components and the populations of all four levels for three initial electric-field amplitudes, E_0, are summarized in Figure 49.23. Excitation of the $|3\rangle \rightarrow |4\rangle$ optical transition is, of course, identical to exciting the $|1\rangle \rightarrow |2\rangle$ transition. The simulated detection of the time-resolved signals shows that there is a sufficiently long time interval—on the order of 400 ps—within which it is possible to determine the spin of the initial states. Indeed, there are two reliable ways for initial spin-state identification: through examining the PL trace profile and through the presence or absence of a second echo pulse after a $\sigma^{+/-}$ π-pulse excitation. The observation of a nonmonotonic PL trace, rather than an exponential decay, is always associated with a spin-up state; in addition, the appearance of a second echo pulse post σ^- π-pulse excitation alludes to a spin up, whereas a repeated echo after a σ^+ π-pulse excitation is indicative of a spin down. The simulations show the onset of the high-intensity Rabi oscillations regime, which suppresses the spin-relaxation processes.

49.5.2 Coherent Spin Manipulation through Hot-Charged Exciton States in QDs

Optical manipulation of the spin of a QD-confined single electron or hole through the resonantly driven trion ground singlet transition (described in the previous section) is considered to be one of the most promising schemes for implementation of spin-based quantum computing, due to the extended spin lifetimes which are limited only by the hyperfine interaction. However, the requirement of a resonantly driven ground trion transition in an inhomogeneously broadened ensemble of charged QDs represents an obstacle for the scalability of architectures based on this quantum system. This problem can be overcome by optical excitation into the excited trion states, taking advantage of a recently discovered effect, herein referred to as "spin filtering". The effect is observed in p-doped QDs under nonresonant, circularly polarized optical excitation, and constitutes an enhanced, photo-induced circular dichroism in the excited trion-state emission, compared to the dichroism exhibited by the ground singlet trion state [70]. The degree of spin polarization is nearly doubled when a high-energy level in the QD is excited resonantly. This is due to an increased spin "injection" efficiency—i.e., the resonant excitation prevents relaxation processes from occurring and the desired spin is maintained. In addition, nonmonotonic dependence of the degree of spin polarization on the optical pulse power is observed, allowing to maximize it by optimizing the pulse characteristics. This spin-filtering effect is promising for realization of high-fidelity schemes for all-optical spin manipulation, since the increased polarization contrast for the two pulse helicities enables highly efficient selective excitation and readout of the spin-up and spin-down populations.

We consider quasi-resonant σ^- or σ^+ circularly polarized excitation of a p-doped QD ensemble into the p-shell in the presence of a resident s-shell hole, and subsequent cascade relaxation to the bright trion ground singlet state, whose decay is detected as PL. The optical excitation creates X^{+*} states, consisting of one electron–hole pair in the s-shell and a resident hole in the p-shell ($1e^1 1h^1 2h^1$), grouped in four degenerate doublets [71]. The exchange interactions can be revealed in the excited exciton spectra, and the energy-level diagram and the spin configurations of the X^{+*} trion are given in Figure 49.24.

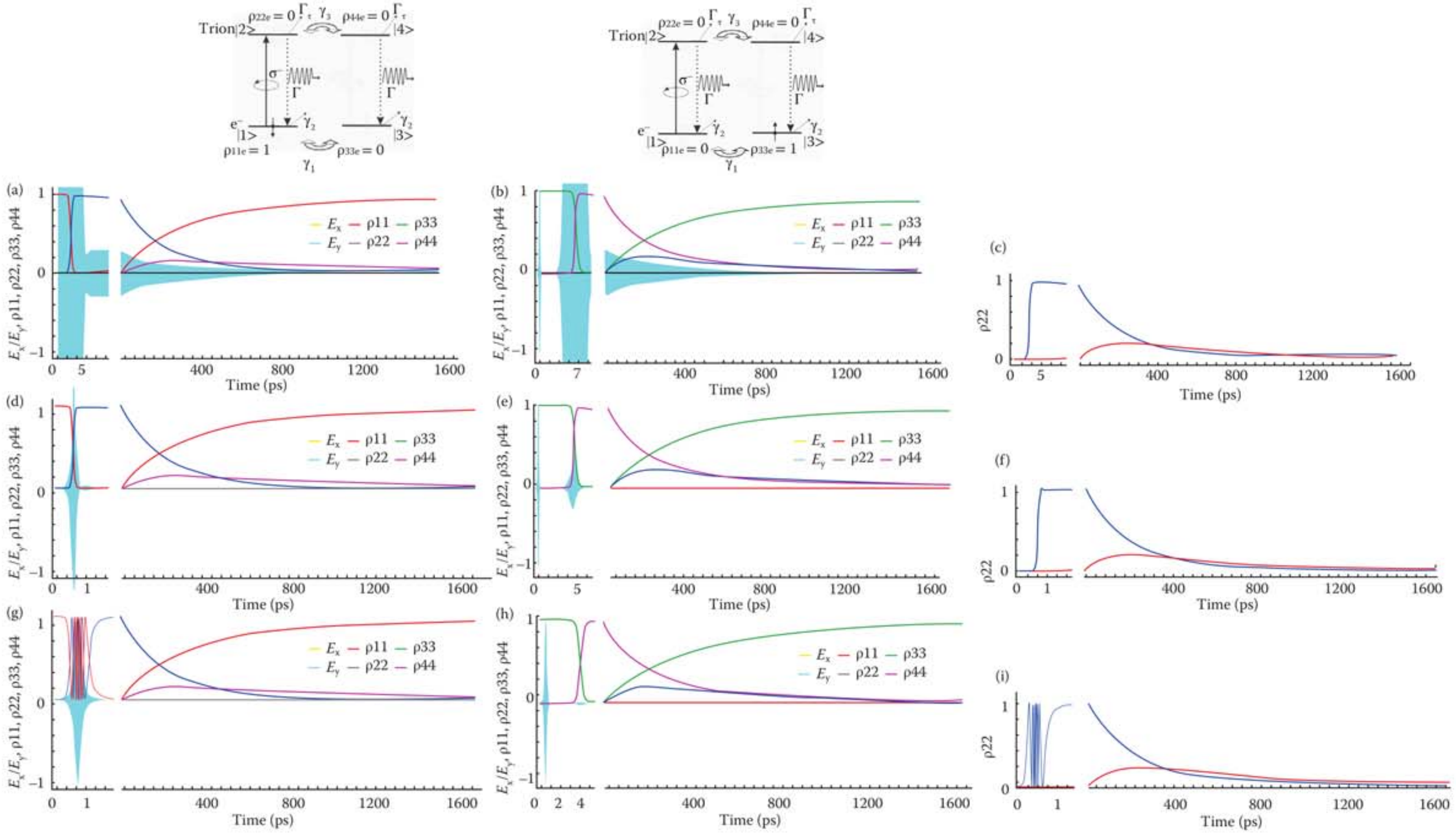


FIGURE 49.23 Top row: initial spin-down (left) and spin-up (right) state. The system is excited by a σ^- polarized pulse. (a–c) Excitation with $E_0 = 550\ \mathrm{V \cdot m^{-1}}$. Time evolution of (a) initial spin-down state $|1\rangle$; (b) initial spin-up state in $|3\rangle$; ρ_{22} (blue curve) represents the trion $|-3/2\rangle$ state spin population proportional to σ^- polarized PL signal detected experimentally. (c, f, i) TRPL signal for $\Downarrow$ (blue curve) and $\Uparrow$ (red curve) exhibiting different dynamics in the short term: ρ_{22} populations from the graphs to their respective left are plotted; (d–f) Excitation with a π-pulse with $E_0 = 3\times10^6\ \mathrm{V \cdot m^{-1}}$. Time-resolved dynamics for (d) spin-down initial state; (e) spin-up initial state. (g–i) Excitation with a 12π-pulse with $E_0 = 4\times10^7\ \mathrm{V \cdot m^{-1}}$. Time-resolved dynamics for (g) spin-down and (h) spin-up initial states. Excitation of initial spin-down state results in six full spin population Rabi flops. (Adapted with permission from G. Slavcheva, *Physical Review B*, 77, 115347, 2008. Copyright 2008 by the American Physical Society.)

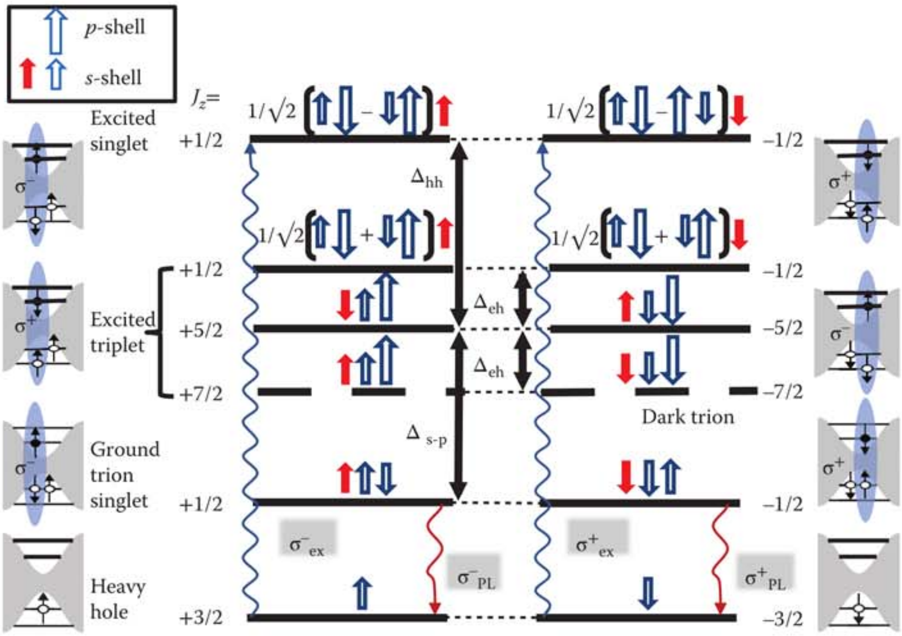


FIGURE 49.24 Energy levels and spin configurations of the hot X^{+*} trion states under σ^- (left six-level system) and σ^+ (right six-level system) optical excitation. The excited singlet is energetically split from the triplet by Δ_{hh}; the excited triplet states are separated from each other by energy Δ_{eh}. The levels are labeled with the total spin projection, J_z, of the electrons and holes. [Small solid arrows, s-shell electron-spin projection; small open arrows, s-shell hole-spin projection; large open arrows, p-shell hole-spin projection; upward wavy arrows, either σ^- or σ^+ excitation; downward wavy arrows, photoluminescence (PL); shaded ellipses, exciton coupling in the radiative states; Δ_{s-p}, energy separation between p-shell and s-shell trion states.]

Spin decoherence mechanisms that need to be considered include the electron–hole anisotropic exchange interaction (AEI), which remains relevant for excited trion states, and the hyperfine interaction between nuclei and either electrons [72] or holes through dipole–dipole interaction [73]. The two degenerate six-level systems for σ^- and σ^+ excitation are coupled through transverse spin decoherence mechanisms, as shown in Figure 49.25.

The system Hamiltonian of a circularly polarized pulse resonantly coupled to an ensemble of six-level resonant absorbers with density N_a is given by

$$\hat{H}^{\mp} = \hbar \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -\frac{1}{2}\left(\Omega_x \mp i\Omega_y\right) \\ 0 & \omega_0 - \Delta_{hh} - \Delta_{sp} & 0 & 0 & 0 & 0 \\ 0 & 0 & \omega_0 - \Delta_{hh} - \Delta_{eh} & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_0 - \Delta_{hh} & 0 & 0 \\ 0 & 0 & 0 & 0 & \omega_0 - \Delta_{hh} + \Delta_{eh} & 0 \\ -\frac{1}{2}\left(\Omega_x \pm i\Omega_y\right) & 0 & 0 & 0 & 0 & \omega_0 \end{pmatrix}, \qquad (49.50)$$

where −/+ correspond to σ^-/σ^+ polarization, $\Omega_x = \wp\frac{E_x}{\hbar}$ and $\Omega_y = \wp\frac{E_y}{\hbar}$ are the time-dependent Rabi frequencies associated with the E_x and E_y electric-field components, and ω_0 and $\wp$ are the resonant transition frequency and the optical dipole matrix element of the ground to excited singlet state transition ($|1\rangle \rightarrow |6\rangle$).

In the presence of relaxation processes (Figure 49.25), under time-dependent external perturbation, e.g., a laser pulse, the time evolution of the six-level quantum system is governed by a master equation,

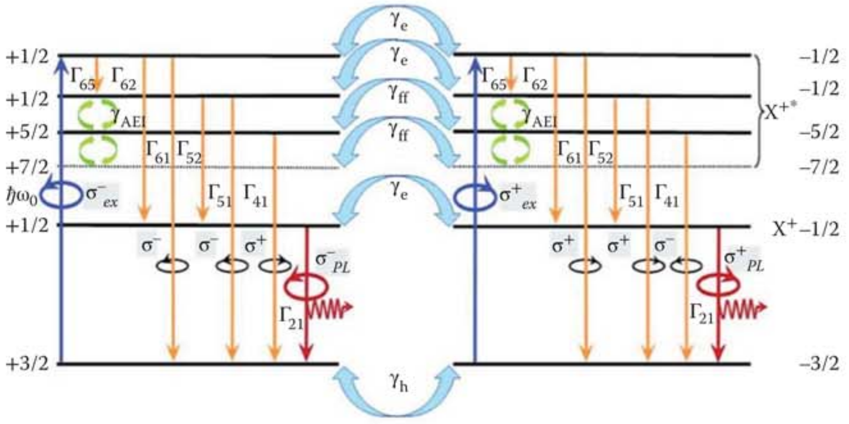


FIGURE 49.25 Discrete-level model of the X^{+*} states in a p-doped QD. The two degenerate six-level systems are coupled via spin decoherence mechanisms represented by double-headed arrows. [γ_h, hole-spin decoherence; γ_e, electron-spin decoherence; γ_{ff}, spin decoherence due to spin flip-flop processes, i.e., anisotropic electron-hole exchange interactions (AEI); ↺ or ↻, circularly polarized transitions; upward arrows, σ_{ex}^- and σ_{ex}^+ resonant pumping; downward arrows, radiative or nonradiative transitions; downward arrows denoted Γ_{21}, detected polarized PL; vertical small curved arrows, spin flip-flop coupling due to AEI.]

Equation 49.12, for the $N^2 - 1 = 35$-dimensional real state pseudospin vector, S_j, coupled to Maxwell's curl equations, Equation 49.43 [74], through the macroscopic polarization induced by the electromagnetic pulse:

$$\begin{aligned} P_x &= -\wp N_a S_5 \\ P_y &= \mp\wp N_a S_{20} \end{aligned} \tag{49.51}$$

where −/+ again correspond to σ^- or σ^+ excitation, respectively.

The QD samples modeled are chosen to be identical to the ones used in TRPL experiments, namely MBE grown on a semi-insulating GaAs (001) substrate. The QD layers are sandwiched between two GaAs barriers and the dot areal density is approximately 2×10^{10} cm^{-2} with an average uncapped height of 4 nm, equivalent to a volume dot density, $N_a = 5 \times 10^{22}$ m^{-3}. The simulation domain consists of an InAs QD layer with nominal thickness given by the height of the typical QD (i.e., 4 nm) embedded between two GaAs barrier regions, each with thickness 50 nm (see Figure 49.26).

The circularly polarized Gaussian pulse at the left boundary of the simulation domain, $z = 0$, is modeled using two x and y linearly polarized waves with resonant carrier frequency, ω_0, phase-shifted by $\pi/2$ according to Equation 49.46 and bottom row of Equation 49.47. The system of equations is discretized in space and time, with spatial and temporal steps, $\Delta z = 1$ Å and $\Delta t = 3.33\times10^{-4}$ fs, respectively, and solved numerically directly in the time domain using the FDTD technique.

For comparison with the theory, PL experiments with short (i.e., 50 ps) σ^- or σ^+ optical excitation pulses with pumping wavelength $\lambda_{\text{res}} = 1065$ nm inducing a resonant transition to the excited dot states were carried out on QD ensembles nominally doped with one hole. The polarized PL was detected at the ground singlet X^+ transition, exhibiting a peak at $\lambda_{\text{det}} = 1148$ nm. In our simulations, an optical pulse of duration 50 ps is injected at $z = 0$. The pulse center frequency, ω_0, is tuned in resonance with the energy splitting between the ground heavy-hole level $|1\rangle$ and the excited singlet trion state $|6\rangle$: $\Delta E_{1\to6} = \Delta_{hh} + E_{X^+} + \Delta_{\text{s-p}}$, where $\Delta_{\text{s-p}}$ is the energy separation between the s- and p-shell trion states. The ground trion singlet energy, $E_{X^+} = 1.0815$ eV, is determined from the resonant PL spectra at λ_{det}. $\Delta_{\text{s-p}} \approx 73$ meV is inferred from the PL spectra [70], and $\Delta_{hh} = 12$ meV is taken in agreement with [71]. In order to calculate the Hamiltonian in Equation 49.50, $\Delta_{eh} \approx 0.5$ meV is used, as obtained in [71,75].

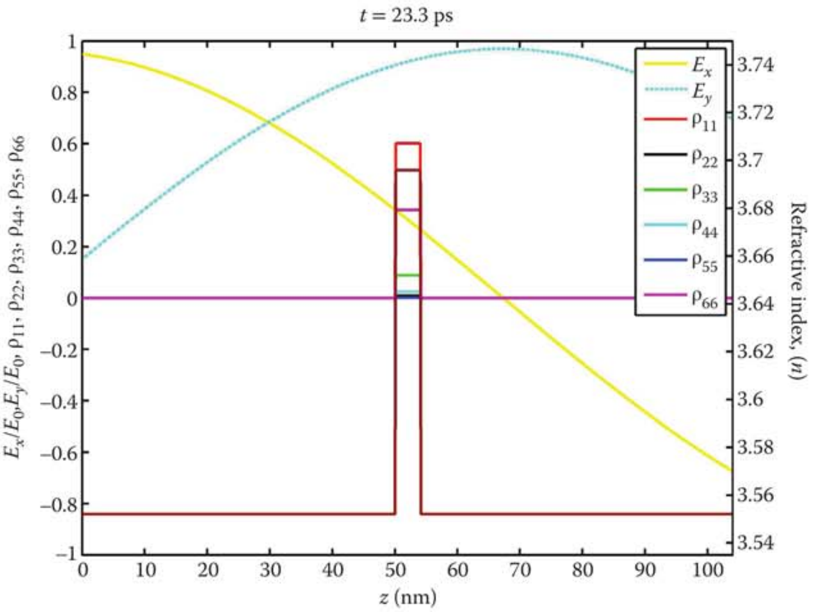


FIGURE 49.26 A snapshot of the spatial distribution of the circularly polarized electric-field components of an excitation pulse interacting with a six-level quantum system. The timing of the snapshot is $t = 23.3$ ps after the source pulse impinges on the left boundary ($z = 0$) of the medium incorporating the quantum system. The charge-carrier spin populations (diagonal density-matrix elements) of all six levels are shown; these indicate the fractional occupancy of each level compared to the total number of charge carriers spread across the different levels. The refractive-index profile of the simulation domain (bottommost curve, right-hand side scale) is produced by an InAs QD layer of thickness 4 nm embedded between two 50-nm-thick GaAs barrier regions.

The pulse area is chosen to be π, so that the pulse completely excites the ground-state spin population to the excited singlet level $|6\rangle$; this corresponds to choosing $E_0 = 2.69 \times 10^5\ \mathrm{V \cdot m^{-1}}$, assuming a dipole matrix element $\wp = 9.83 \times 10^{-29}\ \mathrm{C \cdot m}$. The dipole matrix element's value is comparable to the dipole moment of the X^+ ground singlet transition, and is calculated using the Fermi Golden Rule [50] from the energy separation between the levels and the spontaneous emission rate $\Gamma_{61} \sim 1.2\ \mathrm{ns^{-1}}$, which is close to the experimentally observed value.

The longitudinal relaxation times are taken as follows: the radiative spontaneous decay, $\tau_{21} \approx 1.27$ ns, is experimentally measured from a sample with one hole; spontaneous emission times, $\tau_{41} = 1.35$ ns and $\tau_{51} = \tau_{61} = 1.2$ ns are estimated from the energy-level separation using the Wigner–Weisskopf formula [50]. The nonradiative decay times are taken from theory [76], the transverse spin decoherence rates are obtained from experiment: $\tau_e = 500$ ps [72] and $\tau_h = 14$ ns [73], and a theoretical estimate for $\tau_{ff} = 125$ ps is made on the basis of the Heisenberg uncertainty relations. We should note that there are only two adjustable parameters in our model—namely the largely unknown nonradiative spin relaxation times, τ_{52} and τ_{65}.

The theoretically computed TRPL traces that are displayed in Figure 49.27 clearly exhibit photo-induced circular dichroism in the polarized TRPL emission; this is manifested as a different evolution of PL in time, depending on the pulse helicity used. In contrast to the widely used rate equations model which assumes mono-exponential decays, the dynamics of the quantum system described by the present model is nonlinear and, as a result, the simulated PL decay is not mono-exponential.

The experimentally detected TRPL trace shape strongly depends on detector characteristics: the faster the detector's response and the higher its sensitivity, the better the agreement between theory and

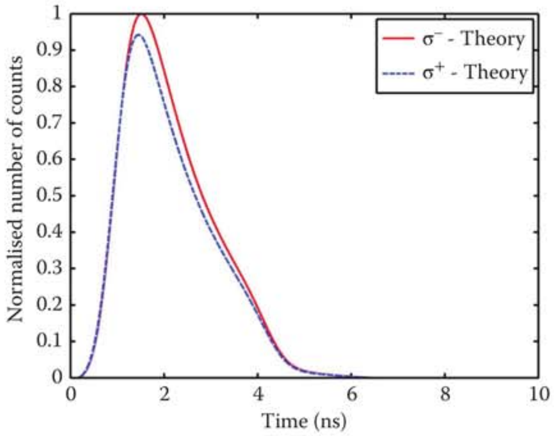


FIGURE 49.27 Theoretically computed polarized TRPL traces for σ^- and σ^+ excitation of a six-level QD. The presence of circular dichroism is clearly visible.

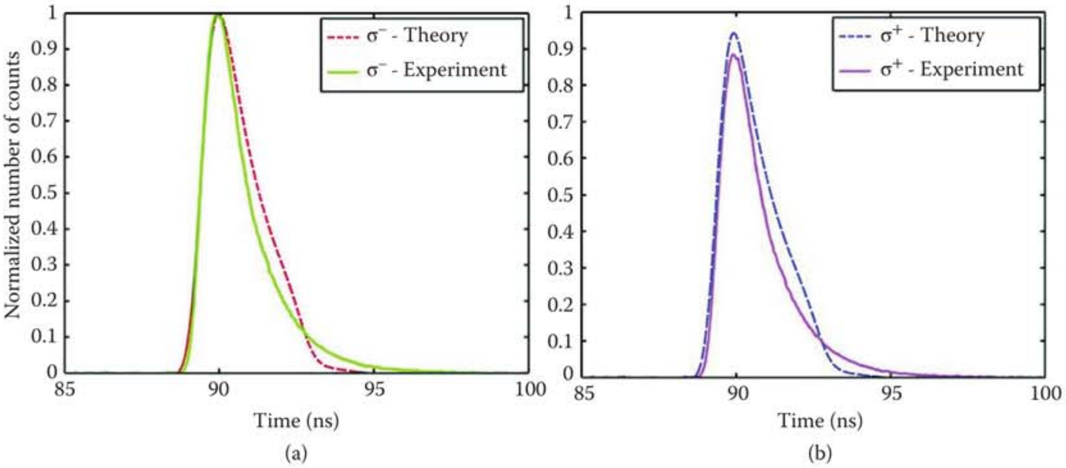


FIGURE 49.28 Comparison between the simulated TRPL trace (after convolution with the detector's response function) and experimentally detected TRPL trace for (a) σ^- excitation and (b) σ^+ excitation of a six-level QD.

experiment. For proper comparison with experiment, we convolve our theoretically calculated polarized traces with the detector's response function and plot them along with the experimental traces in Figure 49.28.

There is excellent agreement between the theoretical and the experimental results. The rising slope and initial decay after the peak in the experimental trace are almost perfectly reproduced by our simulations. The small discrepancy at longer times can be attributed to several factors. Firstly, it is not possible during the course of an experiment to determine precisely which hot trion state is being resonantly excited, so the assumed λ_{res} is only approximate. Secondly, Δ_{s-p} is inferred from the resonant PL spectra and is only a rough estimate, since the excited state does not correspond to a sharp spectral line, but rather a broadened line of finite width; in addition, the inhomogeneous broadening has not been taken into account. The lower peak height for σ^+ excitation can be attributed to the estimated value for the spin flip-flop coupling rate and requires further investigation.

Our theoretical model explains quantitatively the origin of the spin-filtering effect. The experimentally observed enhanced time-resolved circular dichroism in the excited trion state emission is shown to emerge from the dynamical spin flip-flop coupling between the two degenerate sets of excited charged excitonic states. The theory is in very good agreement with experiment, thereby allowing to obtain an estimate for largely unknown intra- and intershell spin-relaxation timescales. The approach also allows to predict optimum pulse parameters, such as power and duration, for efficient control of spin dynamics.

49.6 Quantum Stochastic Formalism for Modeling Cavity-Emitter Systems

Future advances in information systems are expected to exploit the quantum nature of their constituent parts—e.g., in the development of quantum simulators and design of architectures for quantum cryptography. The implementation of these technologies necessitates a fuller understanding of the underlying quantum systems, which in turn will require several fundamental aspects of quantum theory to be tested.

One example of a quantum-optical phenomenon is quantum noise. In contrast to the more familiar classical noise, which can be regarded as spontaneous random fluctuations from a steady state, quantum noise arises from the Heisenberg uncertainty relations, which are fundamental to quantum theory [77]. The concept of quantum noise stems from the statistical interpretation of quantum mechanics and our inability to access any particular one out of the infinite number of degrees of freedom of the electromagnetic field propagating in free space. Quantum noise is closely related to spontaneous emission, irreversible decay and the origin of the spectral line width described by the theory of Weisskopf and Wigner [78], who are considered to have laid the foundations of the quantum noise theory.

At large photon densities, quantum fluctuations are necessary to explain laser linewidth and the threshold properties. By contrast, at low photon densities, light is generated by spontaneous emission described in terms of single-photon emission events within the single-particle Dirac [79] picture. Control over such noise effects will be crucial in the design of devices for quantum information processing.

49.6.1 Langevin Formalism for Modeling Quantum-Optical Effects

Considering a two-level system with a few degrees of freedom immersed in a radiation-field heat bath with an infinite number of degrees of freedom, we exploit the quantum-classical correspondence in the presence of noise and employ the Langevin formalism. The essence of the Langevin approach lies in the realization that the action that the many unknown "bath" (i.e., external) variables have on the system is to modify its deterministic equation of motion, and that this can be accommodated by the inclusion of apparently random terms, known as "Langevin forces."

49.6.2 Spontaneous Emission of a QD Ensemble in a Semiconductor Microcavity and Onset of Lasing

Consider the semiconductor microcavity structures shown in Figure 49.29, designed at $\lambda = 850$ nm and $\lambda = 1.29\,\mu$m. The cavity of the first structure (bottom DBR: 35.5 pairs $\mathrm{AlAs/Al_{0.3}Ga_{0.7}As}$; top DBR: five pairs $\mathrm{AlO/Al_{0.3}Ga_{0.7}As}$) is filled with a resonant two-level medium ($\mathrm{Al_{0.5}Ga_{0.5}As}$), whereas within the second cavity, there are six ($\mathrm{Ga_{0.63}In_{0.37}N_{0.012}As}$) quantum wells filled with the resonant medium. To describe the interaction of an optical wave with these systems, we use the semiclassical 1D Maxwell–Bloch equations in the real vector pseudospin picture, Equation 49.26, for an optical wave linearly polarized along x propagating in the z-direction. In order to model the spontaneous emission of two-level atoms, we add a random electric-field fluctuation, δE_x, at each time step:

$$\left(E_x(z,t) + \delta E_x(z,t)\right) = -\frac{1}{\varepsilon}\frac{\partial H_y}{\partial z} - \frac{N_a\wp}{\varepsilon T_2}S_1 + \frac{N_a\wp\omega_0}{\varepsilon}S_2 . \tag{49.52}$$

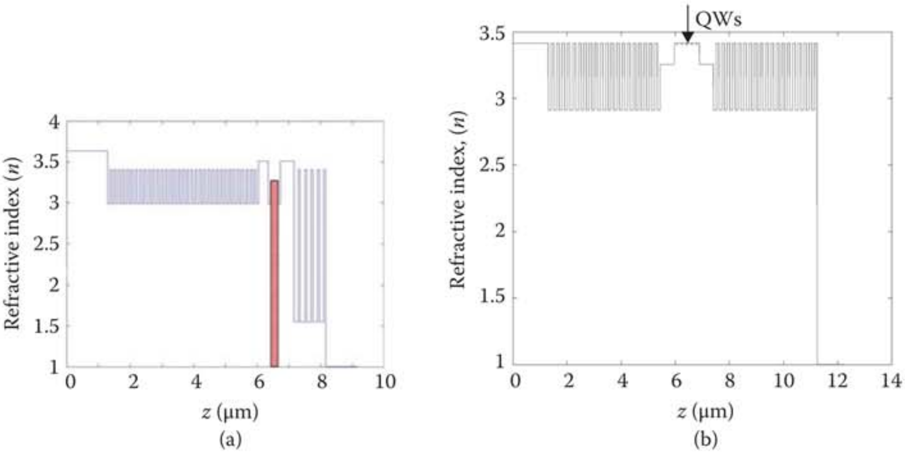


FIGURE 49.29 Refractive-index profile along a semiconductor microcavity (a) filled with gain medium (shaded area) with $n = 3.2736$ at $\lambda = 850$ nm; (b) with 6 GaInNAs-QWs, filled with gain medium of refractive index $n = 3.4$ at $\lambda = 1.29$ μm. (G. Slavcheva et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2004 IEEE.)

We use a pseudorandom number generator based on the Box–Müller method for generating random deviates with a normal (Gaussian) distribution from uniformly distributed in the interval (0, 1) random numbers a and b. At each time step, j, white Gaussian noise with variance $\xi_E = \sigma^2 = 1 \times 10^{-3}\ \mathrm{V^2m^{-2}}$ is implemented according to

$$E_x(z)_j = E_x(z)_j + \sqrt{-2\xi_E \ln(a)}\cos(2\pi b). \tag{49.53}$$

White Gaussian noise has a mean zero and δ-correlated second moment of the random distribution:

$$\begin{aligned}&\langle \delta E_x(z,t)\rangle = 0\\&\left\langle \delta E_x(z,t)\,\delta E_x\left(z,t'\right)\right\rangle = \tilde{\xi}_E\delta\left(t-t'\right) = \xi_E R_{sp}\delta\left(\frac{t-t'}{\Delta t}\right),\end{aligned} \tag{49.54}$$

where an expression for the spontaneous emission rate has been derived:

$$R_{\mathrm{sp}} = \frac{\xi_E(\varepsilon_0\varepsilon)^2}{N_a^2\wp^2T_2^2}. \tag{49.55}$$

Consider the first microcavity: we assume that the resonant medium is initially inverted ($S_{30} = 1$), thus providing gain, and apply solely random white Gaussian noise at each point within the cavity according to Equation 49.53. The spatial distribution of the intracavity electric field as a function of time, and the time evolution of the electric field at the microcavity structure's output boundary are calculated. The parameter set used for this simulation is $\wp = 4.8 \times 10^{-28}\ \mathrm{C\cdot m}$, $n = 3.2736$, $N_a = 1 \times 10^{24}\ \mathrm{m^{-3}}$, initial field amplitude, $E_0 = 700\ \mathrm{V\cdot m^{-1}}$, and relaxation and dephasing times, $T_1 = 10$ ps and $T_2 = 70$ fs, respectively. The cavity field enhancement provided by the DBR mirrors leads to build-up of the electric field within the cavity borne out of the quantum noise. A snapshot of the intracavity electric-field build-up is shown in Figure 49.30, along with the corresponding Rabi flopping of the population dynamics. The amplitude that has arisen solely due to noise gradually builds up towards a standing-wave mode in the cavity.

After this initial amplitude build-up process, the lasing threshold is reached, and coherent oscillations appear at the output boundary. This is the onset of the lasing regime, shown for the unsaturated

$$\Omega_R^2 T_1 T_2 \approx 1.59 \tag{49.56}$$

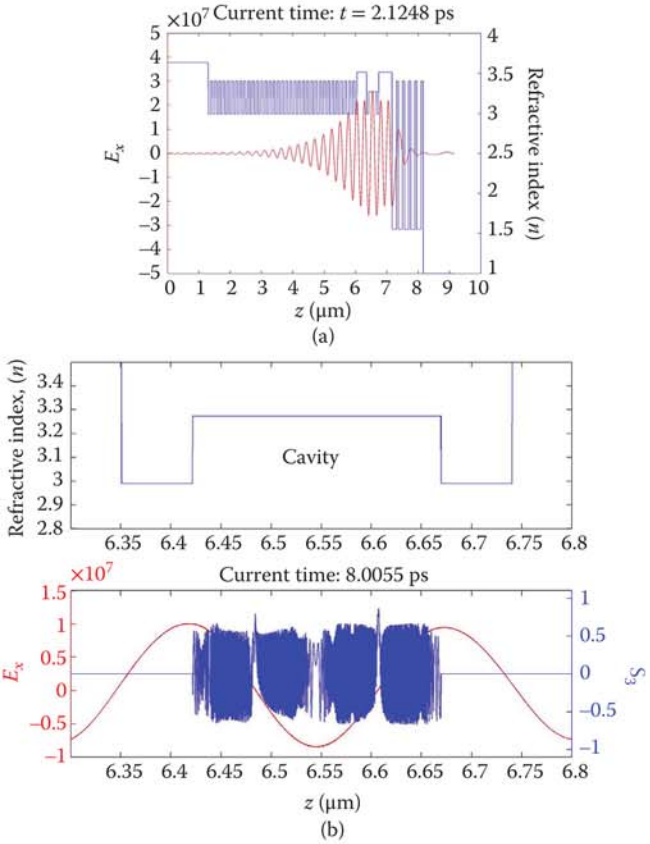


FIGURE 49.30 (a) Stationary electric-field standing-wave profile arisen from quantum noise in the cavity. (b) (Top) Refractive-index profile within the cavity and in its vicinity. (Bottom) A snapshot of the spatial distribution of the intracavity E_x electric field (left axis, slowly-varying sinusoidal curve) and the corresponding population inversion (right axis, fast-oscillating curve) exhibiting fast Rabi oscillations at $t = 8$ ps after the instant when the quantum noise arises. These oscillations result in the fast modulation of the electric-field envelope observed at the output boundary of the structure, which is shown in Figure 49.31a and b.

and saturated

$$\Omega_R^2 T_1 T_2 \approx 16241 \gg 1 \tag{49.57}$$

cases in Figure 49.31 (a,b), respectively. The charge-carrier population performs multiple Rabi flops, oscillating about a steady population inversion value close to zero. The net gain and absorption are therefore similar, resulting in a gain saturation in time.

Let us now consider the structure designed at $\lambda = 1.29\,\mu\text{m}$. The population in all six quantum wells is initially inverted into the upper state, $S_{30} = 1$, representing a nonlinear gain medium, while the barrier regions within the cavity are assumed to be in the ground state, $S_{30} = -1$. Random noise is generated within the cavity with a parameter set corresponding to the unsaturated case above (49.56). In Figure 49.32, a snapshot of the spatial distribution of the electric-field amplitude and the steady-state population inversion are shown. The population inversion relaxes to this value after a sufficiently large number of time steps. The time evolution of the electric field at the output boundary is shown in Figure 49.33, showing continuous build-up of coherent self-sustained oscillations and lasing. The lasing threshold is clearly discernible, and the electric-field envelope exhibits rapid damped relaxation oscillations with a decay rate,

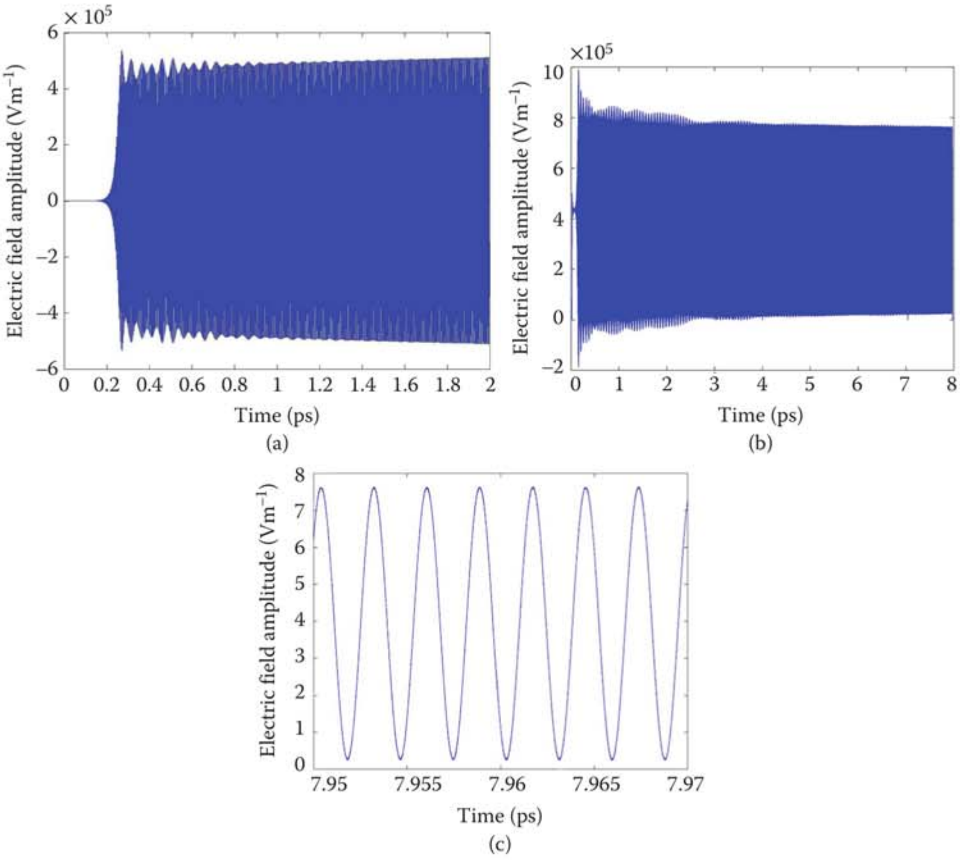
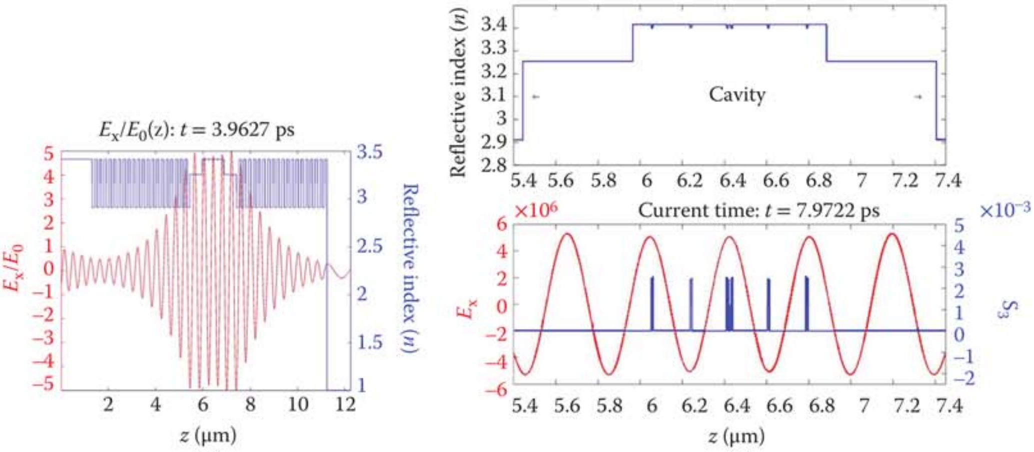


FIGURE 49.31 Time evolution of the electric field at the output boundary of a semiconductor microcavity, showing the noise-induced build-up of coherent oscillations and onset of lasing of the structure displayed in Figure 49.29a. Simulations were run for two different sets of relaxation and dephasing times (T_1 and T_2, respectively). (a) $T_1 = 10$ ps, $T_2 = 70$ fs; (b) $T_1 = 1$ ns, $T_2 = 10$ ps. (c) Expanded view of (b) in the steady-state region, showing single-mode oscillations and lasing at a frequency corresponding to $\lambda = 850$ nm. (G. Slavcheva et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2004 IEEE.)

$\Gamma_R \approx 1.5 \times 10^{11}\ \mathrm{s}^{-1}$. These rapidly varying oscillations are a new feature which is not predicted by the usual rotating-wave (RWA) and slowly-varying envelope (SVEA) approximations. The oscillations eventually reach the steady-state gain saturation value, $S_{3\mathrm{sat}} = 0.0025$, at simulation times exceeding 16 ps.

We perform a Fourier transform of the output electric-field trace in Figure 49.33 and obtain the laser spectral linewidth, shown in Figure 49.34. The obtained linewidth exhibits superfine structure: the main mode is centered at a frequency ω_0 corresponding to $\lambda = 1.289\ \mu\mathrm{m}$, and the wings exhibit satellite peaks at $\omega_0 + n\Omega_{\mathrm{rel}}$. The frequency of the fast relaxation oscillations, $f_{\mathrm{rel}} \approx 2.54$ THz, is determined from the figure.

We showed that an extension within the Langevin framework of our Maxwell-pseudospin model predicts the build-up of coherent, self-sustained oscillations and the onset of lasing, which is a sole consequence of noise being introduced into the cavity. This emphasizes the importance of quantum fluctuations for triggering the lasing regime of cavity operation. The simulations allow to investigate the lasing threshold behavior directly in the time domain and provide an estimate for the coherence time of the laser emission and the laser linewidth. We showed that several important parameters can be extracted from the simulations and that they can be optimized using numerical experiments on a variety of design geometries. Most importantly, the developed quantum stochastic approach is capable of modeling quantum effects, such as



(a) (b)

FIGURE 49.32 (a) Instantaneous profile of the electric-field standing wave borne out of quantum noise in the cavity. (b) (Top) Refractive-index profile implying the cavity boundaries and the location of the quantum wells. (Bottom) Spatial distribution of the stationary electric-field standing wave (left axis, sinusoidal curve) and the corresponding saturated population inversion, $S_{3\text{sat}} = 0.0025$, in the quantum wells within the cavity (right axis, spiky curve). (G. Slavcheva et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2004 IEEE.)

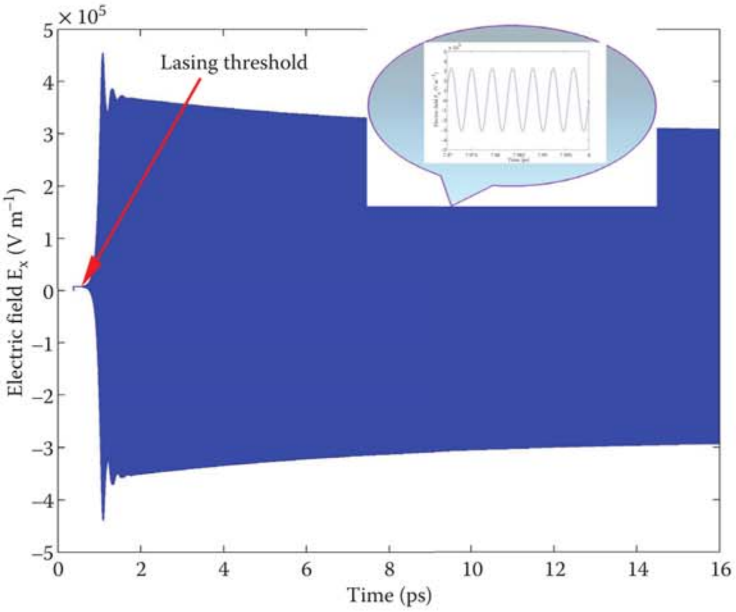


FIGURE 49.33 Build-up in time of coherent, self-sustained oscillations arisen from noise introduced into the cavity of the structure displayed in Figure 49.29b. The lasing threshold is apparent. The decay rate of the electric-field envelope is $\Gamma_R \approx 1.5 \times 10^{11}\ \text{s}^{-1}$. Inset: expanded view of the fast carrier-frequency oscillations with $\lambda = 1.289$ μm; the steady-state region occurs for $t > 16$ ps. (G. Slavcheva et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2004 IEEE.)

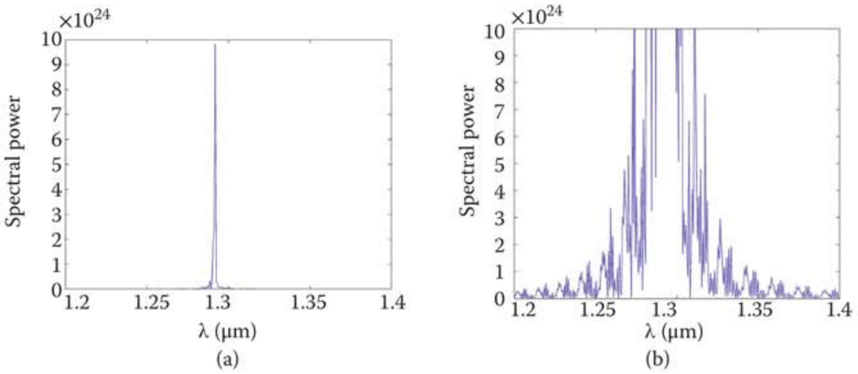


FIGURE 49.34 (a) Laser spectral line obtained by Fourier-transforming the output trace of the electric-field amplitude centered at the lasing frequency, corresponding to $\lambda = 1.289\,\mu\text{m}$. (b) Magnified view of the spectrum in (a) showing relaxation oscillation sidebands. (G. Slavcheva et al., *IEEE Journal of Selected Topics in Quantum Electronics* © 2004 IEEE.)

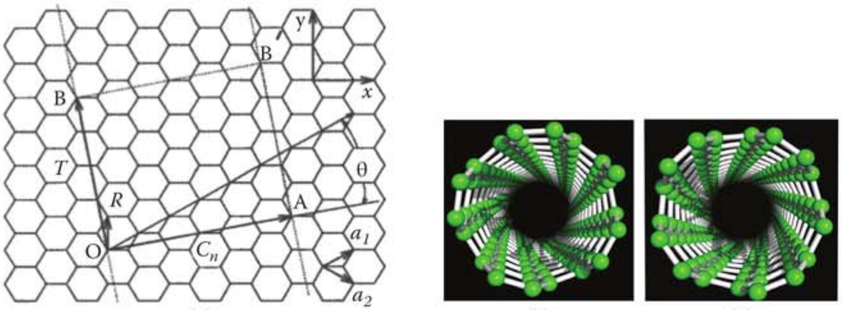
spontaneous emission, and thus can be used to model the system dynamics in the few-photon limit as well as demonstrate and predict quantum-optical effects.

49.7 Modeling Optical Rotation in Chiral Carbon Nanotubes with Master Maxwell-Pseudospin Equations

In this section, we shall employ our vector Maxwell-pseudospin model [44] to investigate how chirality affects the ultrafast nonlinear optical and magneto-optical response of a single chiral CNT [80]. We shall show that a simple discrete-level model of the optically active states near the bandgap edge of a chiral nanotube is sufficient to explain the circular dichroism and birefringence, as well as to quantify the optical rotation per unit length (also known as specific rotatory power), either in the absence or in the presence of an external magnetic field permeating the tube. The model predicts a giant natural gyrotropy for the specific chirality considered in this case, which is comparable with or exceeding that of artificially fabricated helical photonic structures. This is remarkable because the latter are considered to achieve the largest optical rotation of all known materials: orders of magnitude larger than that of crystal birefringent materials or of liquid crystals. Finally, we will give a quantitative estimate of the nonlinear coherent magneto-chiral optical effect in an axial magnetic field.

Chirality is one of the main symmetries of the CNT geometry that determines the optical properties of single-walled CNTs (SWCNTs). SWCNTs are uniquely determined by the chiral vector, $C_h = na_1 + ma_2 \equiv (n, m)$, with $0 < |m| < n$, or equivalently by a pair of integer numbers (n, m) in the planar graphene hexagonal lattice unit-vector basis (Figure 49.35). Two helical forms (enantiomers), e.g., AL $(5, 4)$ and AR $(4, 5)$, can exist. Topologically, the CNT can be viewed as a graphene sheet rolled up along the chiral vector into a cylinder. Chirality serves as a primary classification criterion. Achiral nanotubes, whose mirror image is superimposable, are subdivided into two classes: zigzag ($m = 0$) and armchair ($m = n$) nanotubes. The rest of the nanotubes, whose mirror reflection is not superimposable, are subsumed into the most general class of chiral nanotubes.

There is a fundamental relationship between chirality and optical activity. The interaction of polarized light with chiral materials gives rise to the phenomenon of optical rotation (or optical activity), whereby the polarization plane is rotated continuously during the propagation of the light through the nanotube. In the presence of an external magnetic field, magnetically induced optical activity—also known as Faraday effect—takes place. Both effects manifest themselves as a rotation of the transmitted light; however, the origin of the two effects is fundamentally different. While the natural optical activity is a result of



(a) (b) (c)

FIGURE 49.35 (a) Chiral vector, C_h (C_n on schematic) and chiral angle, $\theta = \angle(a_1, C_h)$, $0° < |\theta| < 30°$, where a_1 and a_2 are the elementary Bravais cell unit vectors. (b) Left-handed AL (5, 4) (with $m > n$; m, n – integers, see text) and (c) right-handed AR (4, 5) (with $m < n$) single-walled CNT molecular structure calculated by the tight-banding method [81]. Shown is a view along the tube axis looking in the negative z-direction. Chirality is determined by the rotation of the so-called armchair chains of carbon atoms either clockwise or counterclockwise when looking in the negative z-direction.

the nonlocal optical response of a medium lacking mirror symmetry, the magnetic optical activity results from time-reversal symmetry breaking by the magnetic field. The two phenomena are linked through the magneto-chiral optical effect which takes place when both symmetries are broken simultaneously [82].

The electronic band structure of an SWCNT is described by the quantization of the wavevector along the tube circumference perpendicular to the tube axis. This results in a discrete spectrum of allowed k-vector states forming subbands in the valence and conduction bands labeled by the quasi-angular momentum quantum number, μ [83]: $k_\perp = \frac{2\pi\mu}{L}$, $\mu = 0, \pm 1, \pm 2, \ldots$, where $L = |C_n|$ is the tube circumference (Figure 49.36a). We should note that due to symmetry in the graphene hexagonal unit cell, there are two equivalent points, K (with $\mu > 0$) and K′ (with $\mu < 0$) in the BZ (Figure 49.36c). The energy dispersion of a CNT can be visualized as intersections of the graphene conical dispersion at the quantized transverse wavevector, $k_\perp$ (Figure 49.36b).

The most widely studied excitation geometry in experiments on individual CNT is the one of linearly polarized excitation perpendicular to the tube axis, shown in Figure 49.37a. In this excitation geometry, the longitudinal E_z electric-field component is dominant, because the transverse E_x component is suppressed by depolarization effects, and linearly polarized intersubband optical transitions from the valence μ subband to the conduction μ subband with $\Delta m = 0$ can be excited (here m signifies the eigenvalue of the quasi-angular momentum operator projection, $\hat{J}_z$). We shall be interested, however, in the unconventional circularly polarized excitation geometry shown in Figure 49.37b, whereby a circularly x–y plane-polarized laser pulse propagates along the tube axis, exciting only one of the two allowed transitions, $\mu \rightarrow \mu \pm 1$, where "$+/-$" corresponds to right and left circularly polarized light, respectively (Figure 49.38b). The molecular structure and the allowed dipole optical intersubband transitions of a (20, 10) chiral CNT are shown in Figure 49.38.

We shall consider aggregates or bundles of aligned SWCNTs grown by chemical-vapor deposition in an electric field [84], or aligned in a polymer matrix [85]. We model the single carbon-nanotube bandgap-edge structure at the K point of the BZ (where the fundamental bandgap opens) by a large ensemble of identical four-level systems, corresponding to the allowed transitions of the two valence subbands closest to the Fermi level for the AL and AR nanotube enantiomers shown in Figure 49.39. We note that the energy-level diagrams for the left- and right-handed helical forms in Figure 49.39 are nonsuperimposable: absorption of σ^+ light excites the $\mu \rightarrow \mu + 1$ transition in AL-handed SWCNT, while absorption of light with the same helicity excites the $\mu \rightarrow \mu - 1$ transition in AR-handed SWCNT. The difference between the optical selection rules for left and right circularly polarized light gives rise to the optical activity [86].

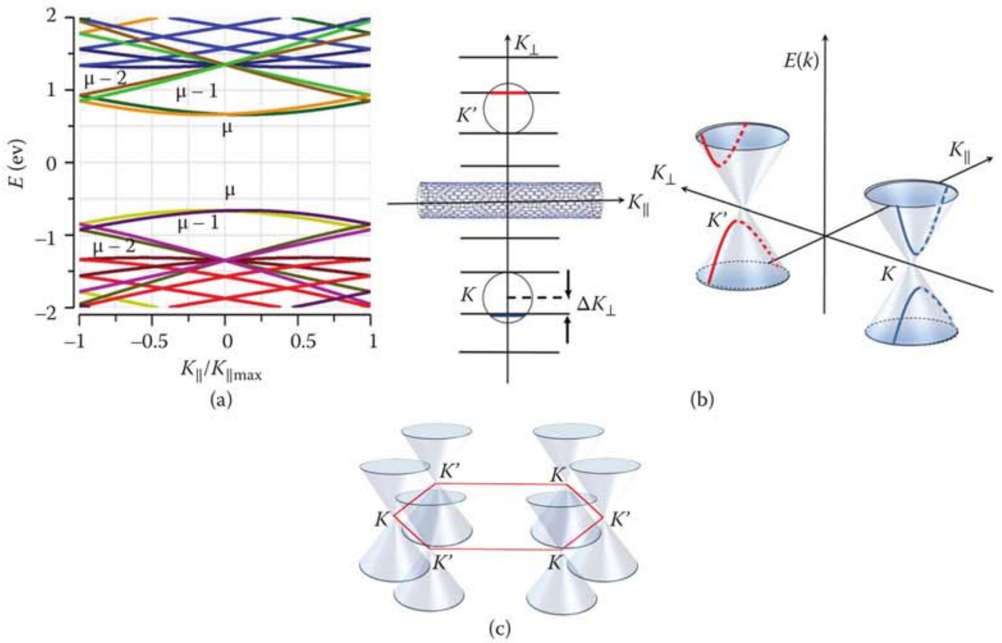


FIGURE 49.36 (a) Energy dispersion as a function of the longitudinal wavevector, $k_{||}$, normalized with respect to the wavevector at the boundary of the first BZ of a (5, 4) SWCNT. $k_{||}$ points along the tube axis. The tube for this chirality is semiconducting—i.e., a bandgap is opened at the center (Γ point) of the BZ. (b) Parabolic cross sections of the graphene linear dispersion, corresponding to points K and K′ in (a) representing the nanotube energy dispersion curves. (c) Unfolded 2D first BZ of graphene, showing the equivalent points K and K′ in the band structure.

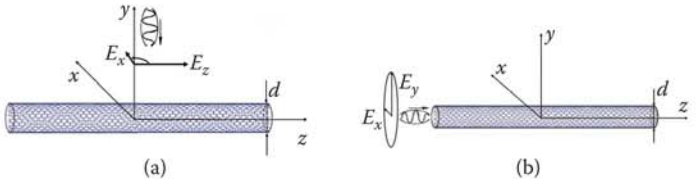


FIGURE 49.37 (a) Linearly polarized excitation perpendicular to the tube axis exciting intersubband optical transitions with $\Delta m = 0$. (b) Circularly polarized excitation along the tube axis exciting $\Delta m = \pm 1$ optical transitions.

49.7.1 Natural Optical Activity

In this section, we compute the natural optical activity rotation angle (rotatory power per unit length) in the absence of an external magnetic field and demonstrate the possibility to manipulate it by engineering the nanotube chirality. We consider the specific case of a chiral, left-handed AL (5, 4) SWCNT, which is chosen for illustration of the general method valid for an arbitrary chirality. The nanotube length is 500 nm and all edge effects will be ignored in our 1D model. We performed tight-binding calculations of the electronic band structure (Figure 49.36a) and determined the fundamental bandgap, $E_g = E_{\mu,\mu} = 1.321$ eV, corresponding to a pumping wavelength of $\lambda = 939$ nm. The tube diameter is 0.611 nm, and the chiral angle, 26.33°. The resonant transition energy for circularly polarized excitations is $E_{\mu,\mu\pm1} = 1.982$ eV, corresponding to a resonant wavelength, $\lambda_0 = 626.5$ nm.

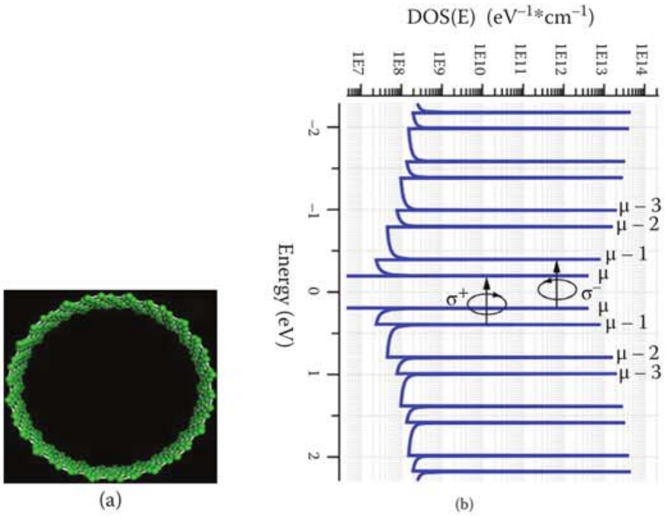


FIGURE 49.38 (a) Molecular structure of an AL (20, 10) left-handed chiral CNT. (b) 1D electron density of states (DOS) plotted as a function of energy at the K point of the BZ ($\mu > 0$). The allowed dipole optical transitions for circularly polarized light are denoted by arrows for left- (σ^-) and right-handed (σ^+) pulse helicity.

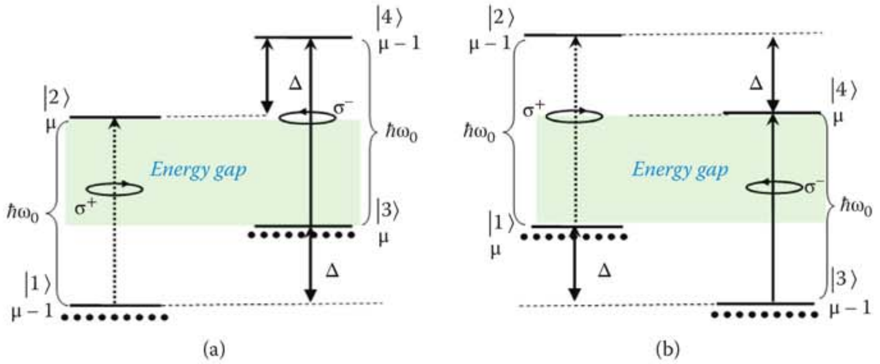
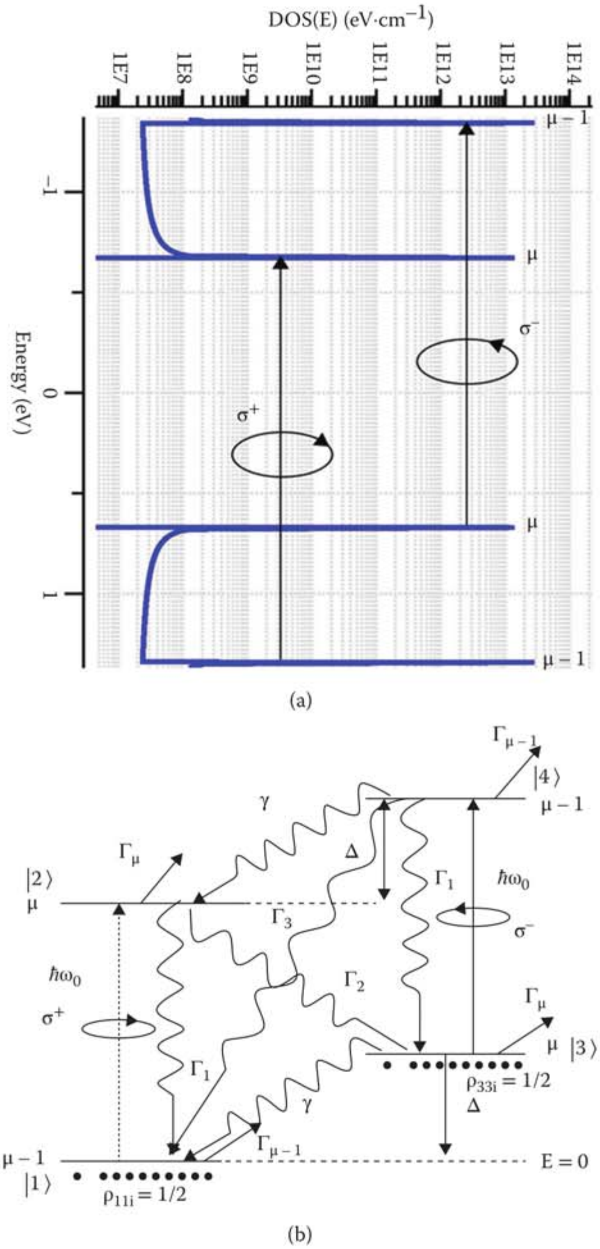


FIGURE 49.39 Energy-level structure at the K (or K') point of the lowest subbands labeled by the subband index μ for an (a) AL (left-handed) and (b) AR (right-handed) SWCNT. The fundamental energy gap is shaded. The dipole optical transitions excited by σ^- and σ^+ circularly polarized light are designated by arrows. Only one of the two transitions is allowed for circularly polarized light, denoted by solid arrow (by contrast, the forbidden transition is indicated by a dashed arrow). Valence-band states below the bandgap are populated. (ω_0, resonant transition frequency; Δ, energy separation between the lowest subband and the second lowest subband near the bandgap.)

We consider a circularly polarized ultrashort pulse of duration 60 fs matching experiments [87], and a pulse area of π, giving an initial electric-field amplitude $E_0 = 6.098 \times 10^8\ \mathrm{V \cdot m^{-1}}$. We inject the pulse from the left boundary of the simulation domain, consisting of the tube embedded between two free-space regions of thickness 50 nm each. The pulse is resonant with the energy $E_{\mu,\mu\pm1}$ of the allowed dipole optical transition and is coupled to an ensemble of identical, HB four-level systems (Figure 49.39a for a left-handed AL SWCNT) used to describe the optical transitions involving the lowest-energy subbands lying nearest to the bandgap edge (Figure 49.40a).



(a)

(b)

FIGURE 49.40 (a) Plot of the 1D density of states (DOS) showing the lowest-energy subbands near the bandgap edge involved in dipole optical transitions excited by circularly polarized light. Only one transition at a time can be excited by each helicity (σ^- or σ^+). (b) Energy-level diagram in the general case of a resonant optical excitation of $\mu - 1 \rightarrow \mu$ and $\mu \rightarrow \mu - 1$ interband transitions by a σ^+ and σ^- polarized pulse, respectively. The resonant transition energy is $\hbar\omega_0$. The initial population of the lowest valence states below the bandgap, ρ_{11i} and ρ_{33i}, is equally distributed between levels $|1\rangle$ and $|3\rangle$ (here, the subscript "i" denotes "initial"). (Δ, the energy separation between the first and the second lowest conduction (or valence) band states; wavy lines denote longitudinal relaxation processes between the levels, associated with population transfer; Γ_1, spontaneous emission (radiative decay) rate of $|2\rangle \rightarrow |1\rangle$ transitions, assumed to be equal to that of $|4\rangle \rightarrow |3\rangle$; Γ_2, spontaneous emission rate for the linearly polarized transition from the conduction μ subband to the valence μ subband; Γ_3, decay rate of the linearly polarized transition from the conduction $\mu - 1$ subband to the valence $\mu - 1$ subband; γ, intraband relaxation rate; Γ_μ and $\Gamma_{\mu-1}$, transverse relaxation (dephasing) rates.)

The corresponding energy-level diagram is shown in Figure 49.40b, with all coherent and spontaneous transitions indicated. Following our coherent vector Maxwell-pseudospin formalism, the system Hamiltonian for a four-level system that applies for either excitation by σ^- or σ^+ is given by

$$\hat{H} = \hbar \begin{pmatrix} 0 & -\frac{1}{2}\left(\Omega_x + i\Omega_y\right) & 0 & 0 \\ -\frac{1}{2}\left(\Omega_x - i\Omega_y\right) & \omega_0 & 0 & 0 \\ 0 & 0 & \Delta & -\frac{1}{2}\left(\Omega_x - i\Omega_y\right) \\ 0 & 0 & -\frac{1}{2}\left(\Omega_x + i\Omega_y\right) & \Delta + \omega_0 \end{pmatrix}, \tag{49.58}$$

where $\Omega_x = \frac{\wp}{\hbar}E_x$ and $\Omega_y = \frac{\wp}{\hbar}E_y$ are the time-dependent Rabi frequencies associated with the E_x and E_y electric-field components, and $\wp$ is the optical dipole matrix element for $\mu \rightarrow \mu \pm 1$ transitions excited by circularly polarized light. This parameter is largely unknown. We have provided an estimate for it,

$$\wp \approx 3 \times 10^{-29}\,\mathrm{C \cdot m}, \tag{49.59}$$

based on an extension of the effective-mass method applied to chirality effects in carbon nanotubes (CNTs) [88], in agreement with the measured radiative lifetime of typical molecular transition in a single CNT [89], which is on the order of $\tau_{\text{spont}} \approx 10\,\text{ns}$.

We consider a four-level resonant medium consisting of a large number of identical aligned CNTs with an average density $N_a = 6.8 \times 10^{24}\,\text{m}^{-3}$, calculated in such a way that the volume of the simulated nanotube with diameter 0.611 nm and length 500 nm contains just a single nanotube on average. A method called the effective medium approximation—used for describing the macroscopic properties of composite materials, such as the dielectric permittivity of a single CNT—has been successfully applied in FDTD modeling of the thermal radiative properties of vertical arrays of multi-walled CNTs [90]. We use two independent theoretical approaches to obtain an estimate for the effective dielectric constant of an isolated SWCNT: the aforementioned effective medium approximation [80] and tight-binding model calculations of the axial component of the imaginary part of the dielectric function tensor [91,92]. This gives a value for the refractive index, $n \approx 2.3$ along the nanotube axis. With dipole matrix element from (49.59), we use the expression for the spontaneous emission rate: $\tau_{\text{spont}} = \frac{3\pi\varepsilon_0\hbar c^3}{n\omega_0^3\wp^2}$, taking into account the energy separation of each transition, to obtain an estimate for the relaxation rates: $\Gamma_1 = 2.907\,\text{ns}^{-1}$, $\Gamma_2 = 9.812\,\text{ns}^{-1}$, and $\Gamma_3 = 1.227\,\text{ns}^{-1}$. We take the experimental value obtained in [87] for the intraband optical transitions, $\gamma = 130\,\text{fs}^{-1}$. Since the dephasing rates for the states involved are largely unknown, we treat them as phenomenological parameters adopting the following values: $\Gamma_\mu = 800\,\text{fs}^{-1}$ and $\Gamma_{\mu-1} = 1.6\,\text{ps}^{-1}$; however, our simulations show that the dynamics is largely insensitive to the choice of these dephasing rates.

We consider separately the cases of an ultrafast resonant optical excitation of the $|1\rangle \rightarrow |2\rangle$ transition with σ^+ helicity, and of the $|3\rangle \rightarrow |4\rangle$ transition with σ^- helicity (Figure 49.40b). Note that in the former case, level $|4\rangle$ does not participate in the relaxation dynamics and the system is effectively a three-level Λ-system, rather than a four-level system. The initial population is assumed to be equally distributed between the lower-lying levels. The ultrashort circularly polarized source pulse is injected into the medium, and the temporal dynamics of the electric-field vector components and population of all four levels are sampled at four different locations along the nanotube z-axis. The spatially resolved temporal dynamics at a point $z = 300\,\text{nm}$ from the left boundary of the simulated structure for σ^+ and σ^- excitations is shown in Figure 49.41.

It would be interesting from an experimental point of view to test whether nanotube chirality could be determined from the ultrafast nonlinear response using ultrashort pulses with either helicity. We compare the Fourier spectra at the output boundary for each helicity of the optical pulse excitation in Figure 49.42. Both transmission spectra exhibit a sharp peak at the resonant wavelength, indicating

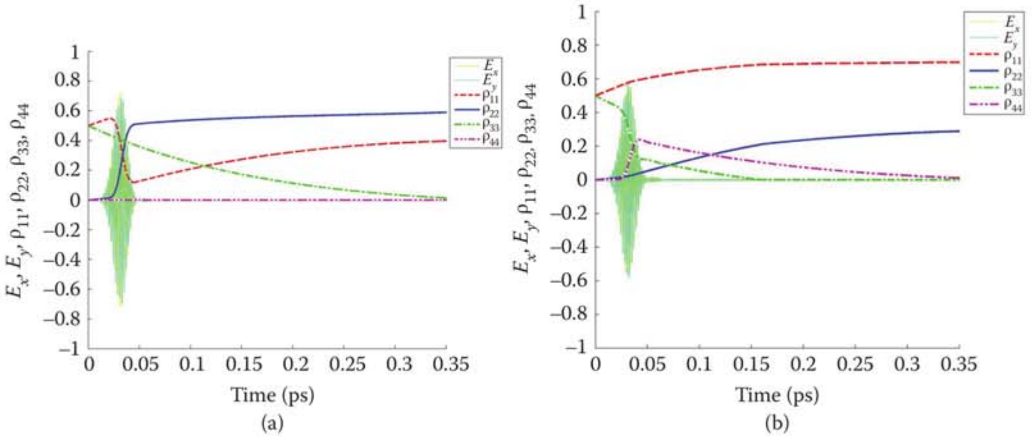


FIGURE 49.41 Time evolution of the electric-field components E_x and E_y, and the populations of all four energy levels of a SWCNT (a) for a circularly polarized σ^+ excitation pulse (note that level $|4\rangle$ is not involved in the dynamics) and (b) for a circularly polarized σ^- excitation pulse. The location is $z = 300$ nm. The left boundary of the SWCNT is at $z = 0$. The SWCNT is 500 nm long and is placed between two free-space regions, each 50 nm long.

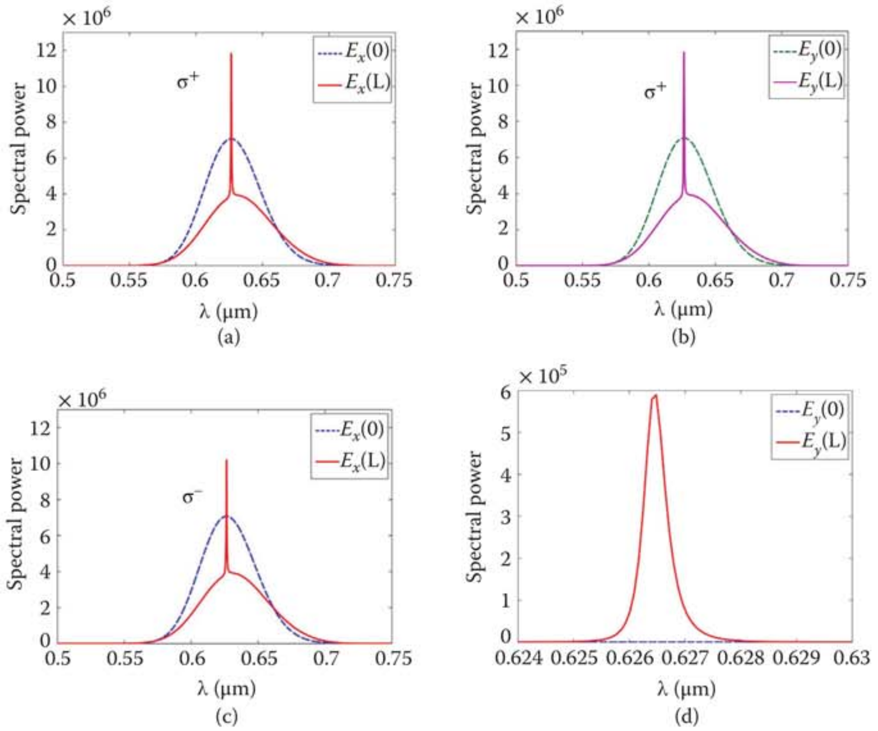


FIGURE 49.42 Transmission Fourier spectra plotted against the wavelength of the E_x and E_y electric-field components at the input, $E_{x,y}(0)$, and output, $E_{x,y}(L)$, of the structure for a (a), (b) σ^+ polarized pulse, and a (c), (d) σ^- polarized pulse. Note that the peak output amplitudes are approximately 20% higher in (a) and (b) compared to in (c) and (d). The simulated structure is as in Figure 49.41.

resonant amplification; however, the peak intensity corresponding to σ^+ excitation is nearly 20% higher. This difference can be exploited in an experiment aiming to determine unambiguously the chirality of a single CNT by first injecting a pulse of one helicity, then the other, and observing the consequences at the output. We show that the difference in the polarization-resolved transmission spectra of a linearly polarized pulse is much more pronounced and allows identification of the precise helical form of the SWCNT.

In order to demonstrate rotation of the electric-field polarization plane during the pulse propagation across the resonant four-level medium, we launch a linearly polarized pulse along x, expressed by:

$$X\begin{cases} E_x(z=0,t) = E_0 e^{-(t-t_0)^2/t_d^2}\cos(\omega_0 t) \\ E_y(z=0,t) = 0. \end{cases} \tag{49.60}$$

The system's temporal dynamics induced by the ultrashort linearly polarized pulse is shown in Figure 49.43a and b on an expanded scale, where the appearance of a second E_y component, and therefore also the appearance of optical rotation of the electric-field vector polarization, is clearly visible. The maximum amplitude of the E_y component continuously increases as the pulse propagates along the nanotube structure. We should therefore expect the maximum optical rotation angle to occur at the output boundary.

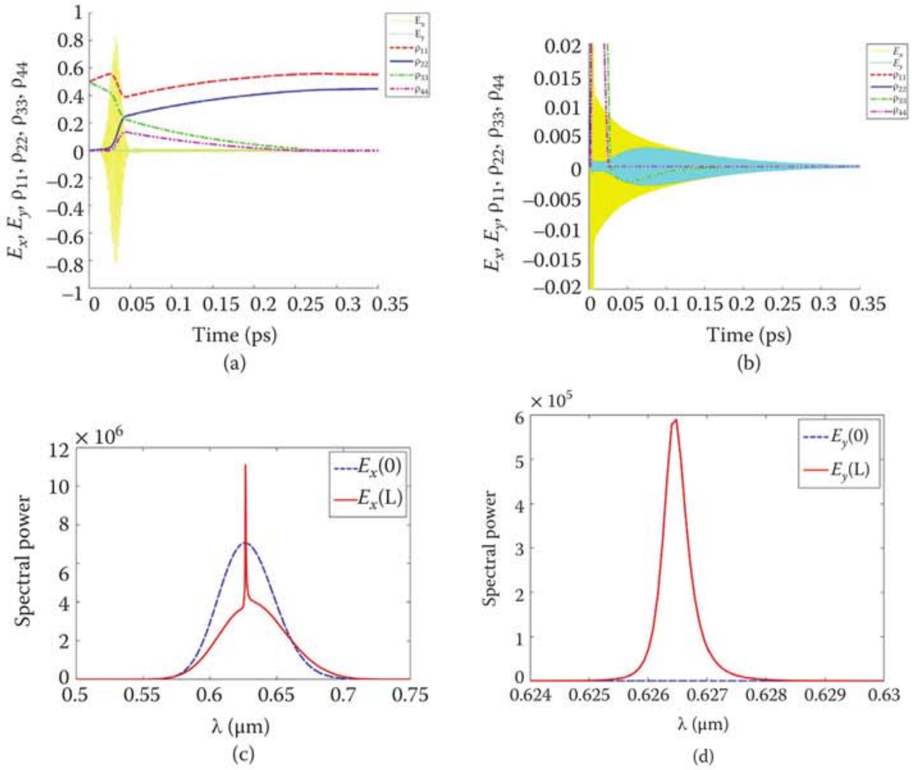
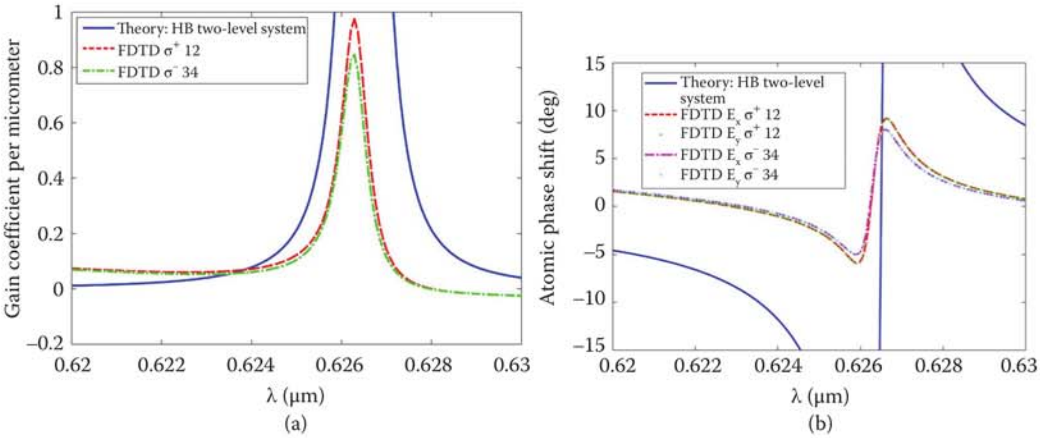


FIGURE 49.43 (a) Time evolution of a linearly polarized pulse with initial value $E_y = 0$, and the populations of all four levels at $z = 550$ nm from the left boundary. The simulated structure is as in Figure 49.41. (b) Magnified view of (a) showing the build-up of the E_y component with time, which is equivalent to a rotation of the electric-field polarization. The transmitted pulse is elliptically polarized and displays different power spectra for (c) the E_x component (sharp resonant peak superimposed on a broader line) and (d) the E_y component (single spectral line).

The polarization-resolved transmission spectra of the E_x and E_y components are quite distinct: while the former is similar to the spectrum obtained for circularly polarized optical excitation, the E_y spectrum exhibits a single spectral line. Conversely, for excitation with light of the opposite helicity (data not shown), the E_x and E_y profiles are interchanged. This difference in the polarization-resolved nonlinear optical response can therefore be used for unambiguous determination of the nanotube chirality.

In order to obtain a quantitative estimate of the natural optical activity in a single chiral CNT, we follow the general methodology developed in Section 49.3.3. We calculate the gain/absorption coefficient from the complex propagation factor, $e^{ik_c(z_2-z_1)}$, for the E_x and E_y electric-field components of a circularly polarized pulse over a distance of one dielectric wavelength, $l = z_2 - z_1 = \lambda_0/n$, where λ_0 is the resonant wavelength. The wavevector is $k_c = \beta + i\gamma$, where β is the phase shift per unit length induced in the optical pulse by the interaction with the resonant medium, and γ is the gain/absorption coefficient. The gain/absorption coefficient allows us to calculate the magnitude of the circular dichroism, whereas the phase shift represents a measure of the rotation angle. A comparison between the spatially resolved gain coefficients for σ^+ and σ^- pulse excitation is shown in Figure 49.44a; the analytically computed gain coefficient of a HB two-level system has also been plotted on the same graph, for reference. The maxima of the gain coefficients for both σ^+ and σ^- occur at the resonant wavelength. The circular dichroism, $\Delta A = G_R - G_L$, is therefore calculated as the average of the difference between the maximal gain coefficients for σ^+ and σ^- excitation (R and L subscripts, respectively) computed at several points along the nanotube. This yields a value for the circular dichroism of $0.083\ \mu\text{m}^{-1}$. By comparison, the absolute value of the circular dichroism of an artificial helicoidal bilayered structure varies in the range of 5–9 dB, which is equivalent to a linear amplitude gain/absorption coefficient of $0.58\text{–}1.04\ \mu\text{m}^{-1}$. The atomic phase shift ($\beta_{R(L),x(y)}l$) and the average rotation angle are plotted in Figure 49.44b. The specific rotatory power per unit length is calculated using $\rho = \pi(n_\text{L} - n_\text{R})/\lambda_0$, where n_L and n_R are the refractive indices for left and right circularly polarized light. Computing this value gives $\approx 2962.24°/\text{mm}$.

Although the circular dichroism is low compared to the artificial chiral photonic structures, the optical rotation, as computed from the phase shift, is enormous compared to other birefringent materials. The



(a) (b)

FIGURE 49.44 (a) Plot against wavelength of the spatially resolved gain coefficient per micron for a σ^+ (dashed curve) and σ^- (dash-dotted curve) circularly polarized ultrashort optical pulse, and the theoretical gain coefficient of a homogeneously broadened (HB) two-level system (solid curve) (Equation 49.30). The simulated structure is as in Figure 49.41. (b) Spectrum of the phase shift, averaged over the entire length of the structure, of the E_x (dashed curve) and E_y (dotted curve) electric-field components of σ^+ excitation, E_x (dash-dotted curve) and E_y (dotted curve) of σ^- optical excitation, and theoretical calculations of this phase shift (Equation 49.31), assuming a HB, resonant two-level system.

specific rotatory power of crystals varies widely from 2.24°/mm for *NaBrO*$_3$ to 522°/mm for *AgGaS*$_2$. Liquid substances exhibit much lower values of specific rotatory power—e.g., $\rho = -0.37°$/mm for turpentine ($T = 10°$, $\lambda = 589.3$ nm); $\rho = 1.18°$/mm for corn syrup, etc. Cholesteric liquid crystals and sculptured thin films exhibit large rotatory power in the visible spectrum: ~1000°/mm and ~6000°/mm, respectively, which is comparable with the calculated rotatory power. We should note that the dependence of the optical rotation angle on chirality opens up the exciting possibility to engineer this angle through manipulating the chirality of the structure.

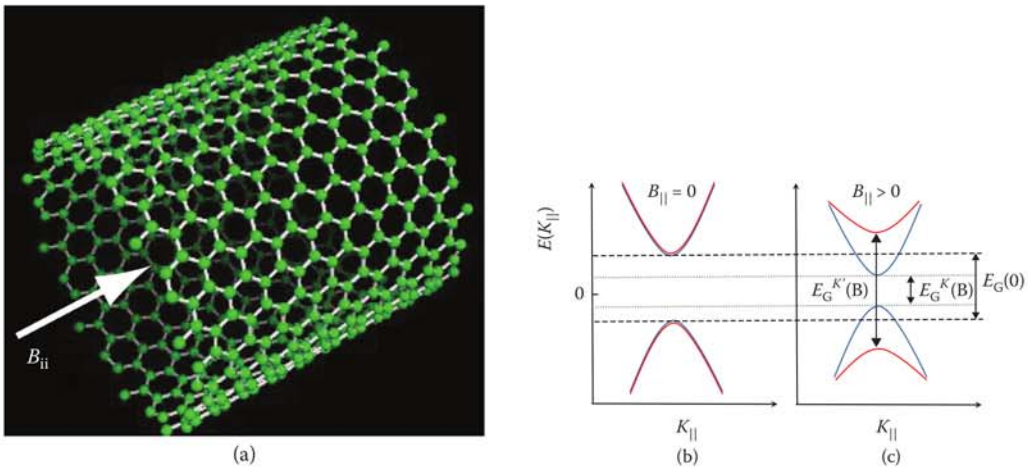
49.7.2 Faraday Rotation and Magneto-Chiral Effects

In this section, we shall develop a theoretical model of the resonant, coherent nonlinear optical activity when a static magnetic field, $B_\parallel$, threads the nanotube (Figure 49.45).

We shall be interested in the Faraday effect, or the rotation of the polarization of a plane-polarized electromagnetic wave propagating in a medium permeated by a static magnetic field, **B**, oriented along the direction of propagation. In the presence of an axial magnetic field, the electronic band structure of a single CNT—and the electronic states near the bandgap edge in particular—change significantly, owing to the combined action of two effects: the spin-B interaction, resulting in Zeeman splitting of the energy levels [93–95], and the appearance of the Aharonov–Bohm phase in the wavefunction [83,96,97]. The two symmetric subbands at the K (or K') point of the BZ are degenerate at $B_\parallel = 0$ (Figure 49.45b). An applied magnetic field along the nanotube axis lifts this degeneracy by shifting the energy levels. As a result, the bandgap of one of the subbands (K') becomes larger while the bandgap of the other subband (K) becomes smaller [94] (Figure 49.45c).

Without loss of generality, we shall consider the electronic states near the bandgap at the K point in the BZ. Therefore, the overall effect will be a bandgap reduction. We consider and calculate separately the contributions to the bandgap from the Zeeman splitting and the Aharonov–Bohm effect.

The effect of the Aharonov–Bohm flux on the bandgap is to induce oscillations in its energy. The energy bandgap thus varies between zero and a fixed value with a period of the flux quantum, $\Phi_0 = h/e$ [83,93],



(a) (b) (c)

FIGURE 49.45 (a) Single (20, 10) chiral nanotube threaded by an axial magnetic field (a so-called Faraday configuration). (b) Magnetic energy bands in the absence of a magnetic field, $B_{\|\|}$; the energy subbands at the point K (K') are degenerate. (c) When an axial magnetic field is applied, the degeneracy is lifted and the energy bandgap of the subband at K' becomes larger, while that of the other subband at K diminishes.

resulting in periodic oscillations of the magneto-optical absorption spectra given by

$$E_G(\Phi) = \begin{cases} 3E_G(0)\left|\frac{1}{3} - \frac{\Phi}{\Phi_0}\right|, & 0 \le \Phi/\Phi_0 \le 1/2 \\ 3E_G(0)\left|\frac{2}{3} - \frac{\Phi}{\Phi_0}\right|, & 1/2 \le \Phi/\Phi_0 \le 1, \end{cases} \tag{49.61}$$

where $E_G(0) = \hbar\omega_0$ is the energy bandgap at zero magnetic field and ω_0 is the resonant transition frequency.

For a magnetic field $B = 8\,\text{T}$ with a flux Φ threading a $(5, 4)$ nanotube with a diameter 0.61145 nm, the ratio $\Phi/\Phi_0 = 0.00057$, and therefore the first of the equations (49.61) above holds. Due to the orbital Aharonov–Bohm effect, this leads to an energy-level shift, or bandgap renormalization of $E_{\text{AB}} = 3.37\,\text{meV}$, corresponding to a resonant angular frequency, $\omega_{\text{AB}} = 5.12 \times 10^{12}\,\text{rad/s}$. At a fixed value of the static magnetic field, the orbital Aharonov–Bohm effect leads to a uniform shift in the energy levels, so the resonant transition frequency ω_0 is replaced by $\omega_0 - \omega_{\text{AB}}$.

In an external magnetic field, the energy levels near the bandgap of Figure 49.40b split and the spin degeneracy is lifted. The resulting energy-level system can be split into two reduced systems of levels, each of which represents a mirror image of the other. The symmetry is broken only by the fact that the allowed optical transitions are different in each case (see Figure 49.46).

The Zeeman splitting, or the spin-B interaction energy, is given by

$$\text{E}_z = \mu_{\text{B}} g_e \sigma B_{\|}\,, \tag{49.62}$$

where $\mu_{\text{B}} = \frac{e\hbar}{2m_e}$ is the Bohr magneton; the electron g-factor, g_e, is taken to be the same as that of pure graphite (≈ 2), $\sigma = \pm 1/2$ is the z-axis projection of the electron spin (spin-up/spin-down state), and m_e is the free electron mass.

We have performed simulations with magnetic fields in the range 2–100 T and found for the considered $(5, 4)$ tube a nonmonotonous dependence of the optical rotation angle on the magnitude of $\mathbf{B}$, with maximum rotation reached for magnetic fields in the range 8–10 T. We therefore selected $B_{\|} = 8\,\text{T}$ to demonstrate the maximum angle of Faraday rotation. Another reason for choosing this value is that it was the maximum magnetic field achievable by the static magnet available in the laboratory of our collaborators at the time when we performed these calculations. For $B_{\|} = 8\,\text{T}$, the Equation (49.62) gives an energy shift of $E_z \approx 0.46\,\text{meV}$, which corresponds to a Zeeman resonant frequency, $\omega_z = 7.026 \times 10^{11}\,\text{rad/s}$.

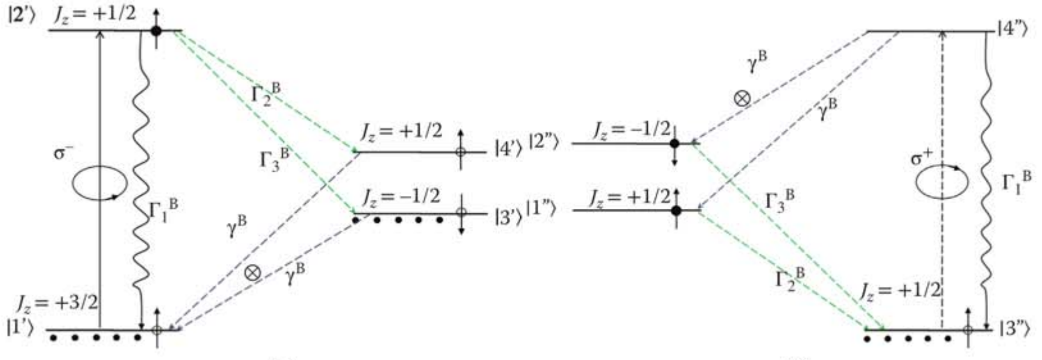


FIGURE 49.46 A four-level system at $B \neq 0$ splits into two nonsuperimposable reduced systems, involving levels energetically closest to the bandgap edge. (a) Left-handed reduced energy-level system excitable by a σ^- optical pulse. (b) Right-handed reduced energy-level system excitable by a σ^+ optical pulse. (Coherent and spontaneous transitions indicated as in Figure 49.40b. Forbidden dipole optical transitions are denoted by ⊗. In fact, there is a X-circle over the dashed lines in Figure 49.46, denoting forbidden transitions between levels).

The system Hamiltonians corresponding to the energy-level schemes displayed in Figure 49.46 are given by

$$\hat{H}' = \hbar \begin{pmatrix} \omega_z & -\frac{1}{2}\left(\Omega_x - i\Omega_y\right) & 0 & 0 \\ -\frac{1}{2}\left(\Omega_x + i\Omega_y\right) & \omega_0 - \omega_z & 0 & 0 \\ 0 & 0 & \Delta - \omega_z & 0 \\ 0 & 0 & 0 & \Delta + \omega_z \end{pmatrix} \tag{49.63}$$

and

$$\hat{H}'' = \hbar \begin{pmatrix} \omega_0 - \omega_z & 0 & 0 & 0 \\ 0 & \omega_0 + \omega_z & 0 & 0 \\ 0 & 0 & \Delta + \omega_z & -\frac{1}{2}\left(\Omega_x + i\Omega_y\right) \\ 0 & 0 & -\frac{1}{2}\left(\Omega_x - i\Omega_y\right) & \Delta + \omega_0 - \omega_z \end{pmatrix}, \tag{49.64}$$

respectively.

We employ our vector Maxwell-pseudospin master equations model for the four-level systems above. The simulation structure is the same as the one described for the $B = 0$ case. Note that due to the presence of a magnetic field, the energy levels $|1\rangle$, $|2\rangle$, $|3\rangle$ and $|4\rangle$ are split into two sets, which we choose to distinguish by adding a $'$ and $''$ to the notation; the first set (denoted by $'$) is accessed only using σ^- light, whereas the second (denoted by $''$) is accessible only via σ^+ light, due to the optical selection rules. We then resonantly excite the $|1'\rangle \to |2'\rangle$ and $|3''\rangle \to |4''\rangle$ transitions using, respectively, a left (σ^-) and right (σ^+) circularly polarized π-pulse of duration 60 fs. The pulse's central frequency is tuned in resonance with the transition frequency, $\omega_0 - \omega_{AB} - 2\omega_z$, and the pulse envelope is a Gaussian function. Owing to the combined effect of the Aharonov-Bohm and Zeeman energy-level shift, the dipole matrix element is modified. An estimate of the optical dipole matrix element in an axial magnetic field can be obtained from the theory developed in [88], taking into account the bandgap reduction at $B = 8\,\mathrm{T}$, giving $\wp = 3.6205 \times 10^{-29}\,\mathrm{C} \cdot \mathrm{m}$. We recalculate the relaxation times using the above mentioned magnetic-field dipole coupling, thus obtaining the following relaxation rates: $\Gamma_1 = 2.914\,\mathrm{ns}^{-1}$, $\Gamma_2 = 9.79\,\mathrm{ns}^{-1}$ and $\Gamma_3 = 9.77\,\mathrm{ns}^{-1}$. The intraband relaxation rate, $\gamma = 130\,\mathrm{fs}^{-1}$, and the dephasing rates, $\Gamma_\mu = 800\,\mathrm{fs}^{-1}$ and $\Gamma_{\mu-1} = 1.6\,\mathrm{ps}^{-1}$, are taken to be the same as for the zero-field case. Note that the initial charge-carrier populations are different for excitation by the σ^- and σ^+ pulses: while in the former case the initial population is assumed to be equally distributed between the valence-band levels $\rho_{1'1'i} = \rho_{3'3'i} = 1/2$ nearest the bandgap edge, in the latter case the whole population is in the single valence-band state $\rho_{4'4'i} = 1$.

Similar to the zero-field case, the time evolution of the electric-field components and populations of all four levels are sampled at several points along the nanotube axis, and the magnetic circular dichroism spectra of the gain/absorption coefficients and phase shifts are calculated and displayed in Figure 49.47. Note that the resonance is shifted toward longer wavelengths due to the bandgap reduction. We should point out that the magnetic circular dichroism spectra at $B = 8\,\mathrm{T}$ are quite distinct from the zero magnetic field ones for natural optical activity (Figure 49.44). While the σ^- polarized pulse is amplified during its propagation along the nanotube, the σ^+ polarized pulse is absorbed, resulting in a much larger net circular dichroism. The average value of the magnetic circular dichroism is $0.706\,\mu\mathrm{m}^{-1}$, an order of magnitude larger than the natural circular dichroism. The different behavior of the calculated gain and absorption spectra under σ^- compared to σ^+ polarized optical pulse excitation is a direct consequence of the different energy-level configuration describing the two cases. While the energy-level system for the σ^- excitation is a four-level system, the one corresponding to the σ^+ excitation is a three-level Λ-system, due to level $|2''\rangle$ being completely decoupled from the rest of the levels as a result of the dipole optical selection rules. The calculated spectra in the latter case are, in fact, a signature of EIT and coherent population trapping

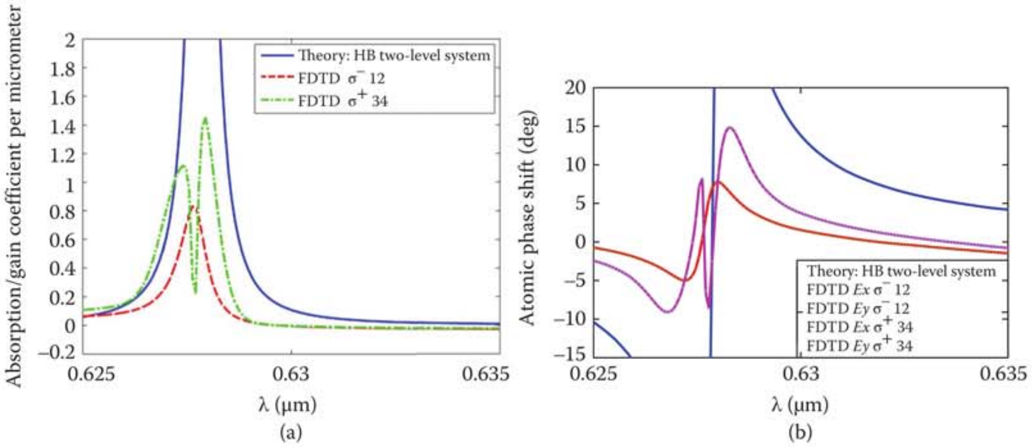


FIGURE 49.47 (a) Spectrum of the gain coefficient per micron, averaged spatially over the entire length of the structure, of the E_x and E_y electric-field vector components of σ^- (dashed curve) and σ^+ (dash-dotted curve) circularly polarized ultrashort optical excitation. The theoretical gain coefficient of a homogeneously broadened (HB), resonant two-level system (solid curve) at $B = 8\,\mathrm{T}$ is also shown. Note that, while the σ^- pulse is amplified during its propagation along the nanotube, the σ^+ pulse is absorbed, resulting in an absorption dip in the spectrum. (b) Spectrum of the phase shift, averaged spatially over the entire length of the structure, at $B = 8\,\mathrm{T}$ of the E_x (dashed curve) and E_y (dots) electric-field components under σ^- excitation, and of E_x (dash-dotted curve) and E_y (dots) under σ^+ circularly polarized ultrashort optical excitation. The phase spectra exhibit double-peaked refraction characteristic of EIT. The theoretical phase shift of a HB resonant two-level system (solid curve), is also given. (FDTD, finite-difference time-domain.)

effects in a three-level system [98], as will be explained subsequently. The absorption at resonance is close to zero, and the spectrum exhibits a dip (Figure 49.47a) similar to the absorption dip characteristic of EIT. The dip is caused by destructive interference between the excitation and emission fields, and leads to the trapping of the charge-carrier population in the ground state (hence, the term "coherent population trapping"). Whereas the phase-shift spectrum for a σ^- pulse excitation is of the type of a two-level atomic phase shift, the spectrum of the phase shift for σ^+ excitation is double peaked (Figure 49.47b). The latter is characteristic of the real part of the dielectric susceptibility in a three-level system exhibiting EIT. The predicted destructive interference in an external axial magnetic field after the passage of the ultrashort pulse is a direct consequence of the timescales of the processes involved in the relaxation dynamics and the phase shift which they introduce in the emission. The calculated average specific rotatory power in a magnetic field $B = 8\,\mathrm{T}$ is $-32580.4°/\mathrm{mm}$, corresponding to an average refractive-index anisotropy $(n_\mathrm{R} - n_\mathrm{L})$ of 6.497 (the purpose of the minus sign in the rotatory power is to denote counterclockwise rotation of the field when viewed from a direction opposite to that of its propagation). This is nearly an order of magnitude greater than the natural optical rotation. We should note that the calculated rotation is a combined effect of the chirality of the nanotube and the magnetic field–induced rotation, and thus can be considered as an estimate for the magneto-chiral effect in a single nanotube. Therefore, our model can be used to study the magnetic field and chirality dependence of the optical rotation angle.

49.8 Conclusion

We have developed a general methodology for description of ultrashort optical-pulse interaction with open, discrete multilevel quantum systems, based on a self-consistent solution of the vector Maxwell's equations and equations for the time evolution of the quantum system in the real-vector representation of the density-matrix geometrical picture. An advantage of this theoretical method is that the coupled vector Maxwell-pseudospin equations can be generated on the fly for a quantum system with an arbitrary number of energy levels using the corresponding unique SU(N) Lie algebra generators. Unlike the widely accepted approach of using simplified two-level systems for description of the system dynamics, our approach allows to model multiple optical transitions and time-dependent gain effects that occur in a realistic quantum system. Our theoretical framework goes beyond the usual rate equations approach, which is unable to model and predict coherent quantum effects that could arise due to couplings and interplay between multiple optical transitions in the system. Most other approaches use the SVEA and RWA which consider only events occurring on a relatively slow timescale—such as the pulse envelope—but do not take account of the pulse's fast carrier wave. By contrast, the proposed approach allows to model rigorously ultrashort optical pulses down to a few optical cycles. Contrary to BPM, no approximations are made to the electromagnetic field equations and their solutions.

We have shown that the model can describe coherent propagation effects: self-induced and EIT, and self-localization effects, such as soliton and pattern formation in cavities. The method can be applied to a wide range of quantum-system geometries by imposing macroscopic boundary conditions and, hence, it can be used to simulate the performance of not only individual idealized devices, but whole interacting systems. We have shown that the model can successfully describe the complex dynamics of an open quantum system, and at the same time provides the capability to observe the feedback mechanisms—including non-Markovian memory effects—as they occur in space and time. In addition, within the Langevin model extension, we were able to model quantum-optical effects such as spontaneous emission. This approach can be used to simulate quantum correlations, nonclassical light generation and manipulation, and few-photon optical nonlinearities.

The model can describe interactions of matter with circularly/elliptically polarized optical pulses and was successfully applied to the problem of generation, manipulation, and readout of single spins and their dynamics. This success is critical for the development of quantum information technologies—currently a very hot research topic which could potentially revolutionize technological applications. Possible application areas include—but are not limited to—quantum computing and information technologies, such as spin-qubit manipulation in QDs, cold atoms, and NV centers in nanodiamond. CNT and arrays of aligned nanotubes have the potential to serve as functional elements in polarization-sensitive devices and polarization switches, as well as for metamaterial applications in the visible part of the electromagnetic spectrum. Predicted spatio-temporal coherent effects, such as cavity solitons, can be used as integrated low-loss optical switches and for all-optical information processing. Finally, the proposed methodology can be used as a design tool for future quantum-optical devices and for the advancement of quantum-photonic technologies.

References

1. R. J. Glauber. The quantum theory of optical coherence. *Physical Review*, 130:2529, 1963.
2. R. J. Glauber. *Quantum Optics and Electronics*. New York, NY: Gordon Breach, 1965.
3. R. Hanbury Brown and R. Q. Twiss. *Nature*, 177:27, 1956.
4. L. Mandel and S. Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995.
5. B. Lounis and M. Orrit. Single-photon sources. *Report on Progress in Physics*, 68:1129, 2005.
6. A. Einstein, B. Podolsky, and N. Rosen. Can quantum mechanical description of physical reality be considered complete? *Physical Review*, 47:777, 1935.
7. J. L. O'Brien, A. Furusawa, and J. Vuckovic. Photonic quantum technologies. *Nature Photonics*, 3:687, 2009.
8. T. M. Babinec, B. J. M. Hausmann, M. Khan, Y. Zhang, J. R. Maze, Ph. R. Hemmer, and M. LonÄar. A diamond nanowire single-photon source. *Nature Nanotechnology*, 5:195, 2010.
9. J. Claudon, J. Bleuse, N. S. Malik, M. Bazin, P. Jaffrennou, N. Gregersen, C. Sauvan, P. Lalanne, and J.-M. Gerard. A highly efficient single-photon source based on a quantum dot in a photonic nanowire. *Nature Photonics*, 4:174, 2010.

10. D. Dalacu, K. Mnaymneh, V. Sazonova, P. J. Poole, G. C. Aers, J. Lapointe, R. Cheriton, A. J. SpringThorpe, and R. Williams. Deterministic emitter-cavity coupling using a single-site controlled quantum dot. *Physical Review B*, 82:033301, 2010.
11. K. G. Lee, X. W. Chen, H. Eghlidi, P. Kukura, R. Lettow, A. Renn, V. Sandoghdar, and S. Götzinger. A planar dielectric antenna for directional single-photon emission and near-unity collection efficiency. *Nature Photonics*, 5:166, 2011.
12. S. Strauf, N. G. Stoltz, M. T. Rakher, L. A. Coldren, P. M. Petroff, and D. Bouwmeester. High-frequency single-photon source with polarization control. *Nature Photonics*, 1:704, 2007.
13. D. Loss and D. P. diVincenzo. Quantum computation with quantum dots. *Physical Review A*, 57:120, 1998.
14. B. Meier and B. P. Zakharchenya editors. *Optical Orientation*. North-Holland, Amsterdam: Elsevier, 1984.
15. W. S. Warren, H. Rabitz, and M. Dahleh. Coherent control of quantum dynamics: The dream is alive. *Science*, 259:1581, 1993.
16. K. De Greve, D. Press, P. L. McMahon, and Y. Yamamoto. Ultrafast optical control of individual quantum dot spin qubits. *Reports on Progress in Physics*, 76:092501, 2013.
17. M. Sargent III, M. O Scully, and W. E. Lamb Jr. *Laser Physics*. Reading, MA: Addison-Wesley Publishing Company Advanced Book Program, 1974.
18. S. L. McCall and E. L. Hahn. Self-induced transparency by pulsed coherent light. *Physical Review Letters*, 18:908, 1967.
19. S. Schneider, P. Borri, W. Langbein, U. Woggon, J. Förstner, A. Knorr, R. L. Sellin, D. Ouyang, and D. Bimberg. Self-induced transparency in InGaAs quantum-dot waveguides. *Applied Physics Letters*, 83:3668, 2003.
20. S. Hughes. Breakdown of the area theorem: Carrier-wave Rabi flopping of femtosecond optical pulses. *Physical Review Letters*, 81:3363, 1998.
21. D. Pinotsi and A. Imamoglu. Single photon absorption by a single quantum emitter. *Physical Review Letters*, 100:093603, 2008.
22. M. Sondermann, R. Maiwald, H. Konermann, N. Lindlein, U. Peschel, and G. Leuchs. Design of a mode converter for efficient light-atom coupling in free space. *Applied Physics B*, 89:489, 2007.
23. S. J. van Enk. Atoms, dipole waves, and strongly focused light beams. *Physical Review A*, 69:043813, 2004.
24. M. Brune, F. Schmidt-Kaler, A. Maali, J. Dreyer, E. Hagley, J. M. Raimond, and S. Haroche. Quantum Rabi oscillation: A direct test of field quantization in a cavity. *Physical Review Letters*, 76:1800, 1996.
25. K. M. Birnbaum, A. Boca, R. Miller, A. D. Boozer, T. E. Nortup, and H. J. Kimble. Photon blockade in an optical cavity with one trapped atom. *Nature*, 436, 87:87, 2005.
26. R. Young, R. M. Stevenson, P. Atkinson, K. Cooper, D. A. Ritchie, A. J. Shields. *Advances in Solid State Physics*, Vol. 46:55. Berlin: Springer Verlag, 2008.
27. O. Gazzano, S. Michaelis de Vasconcellos, C. Arnold, A. Nowak, E. Galopin, I. Sagnes, L. Lanco, A. Lemaître, and P. Senellart. Bright solid-state sources of indistinguishable single photons. *Nature communications*, 4:1425, 2012.
28. K. Srinivasan and O. Painter. Linear and nonlinear optical spectroscopy of a strongly coupled micro disk-quantum dot system. *Nature*, 450:862, 2007.
29. N. Le Thomas, U. Woggon, O. Schöps, M. V. Artemyev, M. Kazes, and U. Banin. Cavity QED with semiconductor nanocrystals. *Nano Letters*, 6:557, 2006.
30. M. Heiss, Y. Fontana, A. Gustafsson, G. Wust, C. Magen, D. D. O'Regan, J. W. Luo, B. Ketterer, S. Conesa-Boj, A. V. Kuhlmann, J. Houel, E. Russo-Averchi, J. R. Morante, M. Cantoni, N. Mazari, J. Arbiol, A. Zunger, R. J. Warburton, and A. Fontcuberta i Morral. Self-assembled quantum dots in a nanowire system for quantum photonics. *Nature Materials*, 12:439, 2013.

31. H. Kim, R. Bose, T. C. Shen, G. S. Solomon, and E. Waks. A quantum logic gate between a solid-state quantum bit and a photon. *Nature Photonics*, 7:373, 2013.
32. A. Faraon, I. Fushman, D. Englund, N. Stoltz, P. Petroff, and J. Vučković. Coherent generation of non-classical light on a chip via photon-induced tunnelling and blockade. *Nature Physics*, 4:859, 2008.
33. J. Kasprzak, S. Reitzenstein, E. A. Muljarov, C. Kistner, C. Schneider, M. Strauss, S. Hing, A. Forchel, and W. Langbein. Up on the Jaynes-Cummings ladder of a quantum-dot/microcavity system. *Nature Materials*, 9:304, 2010.
34. P. Michler, A. Kiraz, C. Becher, W. V. Schoenfeld, P. M. Petroff, L. Zhang, E. Hu, and A. Imamoglu. A quantum dot single-photon turnstile device. *Science*, 290:2282, 2000.
35. C. Santori, D. Fattal, J. Vuckovic, G. S. Solomon, and Y. Yamamoto. Indistinguishable photons from a single-photon device. *Nature*, 419:594, 2002.
36. E. Knill, R. Laamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:46, 2001.
37. L. Allen and J. H. Eberly. *Optical Resonance and Two-Level Atoms*. New York, NY: Wiley, 1975.
38. R. P. Feynman, F. L. Vernon, and R. W. Hellwarth. Geometrical representation of the Schrödinger equation for solving laser problems. *Journal of Applied Physics*, 28:49, 1957.
39. R. J. Cook and B. W. Shore. Coherent dynamics of N-level atoms and molecules. III. An analytically soluble periodic case. *Physical Review A*, 20:539, 1979.
40. J. N. Elgin. Semiclassical formalism for the treatment of three-level systems. *Physics Letters*, 80A:140, 1980.
41. M. Gell-Mann and Y. Neeman. *The Eightfold Way*. New York, NY: Benjamin, 1964.
42. F. T. Hioe and J. H. Eberly. N-level coherence vectors and higher conservation laws in quantum optics and quantum mechanics. *Physical Review Letters*, 47:838, 1981.
43. S. L. McCall and E. L. Hahn. Self-induced transparency. *Physical Review*, 183:457, 1969.
44. G. Slavcheva. Model for the coherent optical manipulation of a single spin state in a charged quantum dot. *Physical Review B*, 77:115347, 2008.
45. G. Slavcheva, J. M. Arnold, I. Wallace, and R. W. Ziolkowski. Coupled Maxwell-pseudospin equations for investigation of self-induced transparency effects in a degenerate three-level quantum system in two dimensions: Finite-difference time-domain study. *Physical Review A*, 66:063418, 2002.
46. A. Taove. *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House, 2000.
47. R. W. Ziolkowski, J. M. Arnold, and D. M. Gogny. Ultrafast pulse interactions with two-level atoms. *Physical Review*, 52:3082, 1995.
48. B. Engquist and A. Majda. Absorbing boundary conditions for the numerical simulation of waves. *Mathematics of Computation*, 31:629, 1977.
49. G. Mur. Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic-field equations. *IEEE Transactions on Electromagnetic Compatibility*, 23:377, 1981.
50. A. Yariv. *Quantum Electronics*. New York, NY: Wiley, 1989.
51. G. M. Slavcheva, J. M. Arnold, and R.W. Ziolkowski. FDTD simulation of the nonlinear gain dynamics in active optical waveguides and semiconductor microcavities. *IEEE Journal of Selected Topics in Quantum Electronics*, 10(5):1052, 2004.
52. A. E. Siegman. *Lasers*. Mill Valley, CA: University of Science Books, 1986.
53. G. Slavcheva, J. M. Arnold, and R. W. Ziolkowski. Ultrashort pulse lossless propagation through a degenerate three-level medium in nonlinear optical waveguides and semiconductor microcavities. *IEEE Journal of Selected Topics in Quantum Electronics*, 9(3):929, 2003.
54. M. Blaaboer, G. Kurizki, and B. A. Malomed. Spatiotemporally localized solitons in resonantly absorbing Bragg reflectors. *Physical Review E*, 62:R57, 2000.

55. A. Kozhekin and G. Kurizki. Self-induced transparency in Bragg reflectors: Gap solitons near absorption resonances. *Physical Review Letters*, 74:5020, 1995.
56. A. Kozhekin, G. Kurizki, and B. A. Malomed. Standing and moving gap solitons in resonantly absorbing Bragg gratings. *Physical Review Letters*, 81:3647, 1998.
57. G. Slavcheva and O. Hess. Dynamical model of coherent circularly polarized optical pulse interactions with two-level quantum systems. *Physical Review A*, 72:053804, 2005.
58. A. I. Maimistov, A. M. Besharov, O. Elyutin, and Yu. M. Sklyarov. Present state of self-induced transparency theory. *Physics Reports*, 191:1–108, 1990.
59. G. Slavcheva and O. Hess. Spin-dependent dynamics of ultrafast polarised optical pulse propagation in coherent semiconductor quantum systems. *Physica Status Solidi C*, 3:2414, 2006.
60. G. Slavcheva and O. Hess. All-optical coherent control of spin dynamics in semiconductor quantum dots. *Optical and Quantum Electronics*, 31(12–14):973, 2007.
61. A. Wojs, P. Hawrylak, S. Fafard, and L. Jacak. Electronic structure and magneto-optics of self-assembled quantum dots. *Physical Review B*, 54:5604, 1996.
62. A. Greilich, R. Oulton, E. A. Zhukov, I. A. Yugova, D. R. Yakovlev, M. Bayer, A. Shabaev, A. L. Efros, I. A. Merkulov, V. Stavarache, D. Reuter, and A. Wieck. Optical control of spin coherence in singly charged (In,Ga)As/GaAs quantum dots. *Physical Review Letters*, 96:227401, 2006.
63. P. Chen, C. Piermarocchi, L. J. Sham, D. Gammon, and D. G. Steel. Theory of quantum optical control of a single spin in a quantum dot. *Physical Review B*, 69:075320, 2004.
64. A. Shabaev, A. L. Efros, D. Gammon, and I. A. Merkulov. Optical readout and initialization of an electron spin in a single quantum dot. *Physical Review B*, 68:201305(R), 2003.
65. A. V. Khaetskii, D. Loss, and L. Glazman. Electron spin decoherence in quantum dots due to interaction with nuclei. *Physical Review Letters*, 88:186802, 2002.
66. I. A. Merkulov, A. L. Efros, and M. Rosen. Electron spin relaxation by nuclei in semiconductor quantum dots. *Physical Review B*, 65:205309, 2002.
67. D. V. Bulaev and V. Loss. Spin relaxation and decoherence of holes in quantum dots. *Physical Review Letters*, 95:076805, 2005.
68. T. Flissikowski, I. A. Akimov, A. Hundt, and F. Henneberger. Single-hole spin relaxation in a quantum dot. *Physical Review B*, 68:161309(R), 2003.
69. S. Economou, R.-B. Liu, L. J. Sham, and D. G. Steel. Unified theory of consequences of spontaneous emission in a λsystem. *Physical Review B*, 71:195327, 2005.
70. M. Taylor, E. Harbord, P. Spencer, E. Clarke, G. Slavcheva, and R. Murray. Optical spin-filtering effect in charged InAs/GaAs quantum dots. *Applied Physics Letters*, 97:171907, 2010.
71. K. Kavokin. Fine structure of the quantum-dot trion. *Physica Status Solidi: A*, 195:592, 2003.
72. P. F. Braun, X. Marie, L. Lombez, B. Urbaszek, T. Amand, P. Renucci, V. K. Kalevich, K. V. Kavokin, O. Krebs, P. Voisin, and Y. Masumoto. Direct observation of the electron spin relaxation induced by nuclei in quantum dots. *Physical Review Letters*, 94:116601, 2005.
73. B. Eble, C. Testelin, P. Desfonds, F. Bernardot, A. Balocchi, T. Amand, A. Miard, A. Lemaître, X. Marie, and M. Chamarro. Hole-nuclear spin interaction in quantum dots. *Physical Review Letters*, 102:146601, 2009.
74. G. Slavcheva. Model of coherent optical spin manipulation through hot trion states in p-doped InAs/GaAs quantum dots. *ArXiv:1301.7018v1*, 2013.
75. T. Warming, E. Siebert, A. Schliwa, E. Stock, R. Zimmermann, and D. D. Bimberg. Hole-hole and electron-hole exchange interactions in single InAs/GaAs quantum dots. *Physical Review B*, 79:125316, 2009.
76. G. A. Narvaez, G. Bester, and A. Zunger. Carrier relaxation mechanisms in self-assembled (In, Ga) As/GaAs quantum dots: Efficient P→S auger relaxation of electrons. *Physical Review B*, 74:075403, 2006.
77. C. W. Gardiner and P. Zoller. *Quantum Noise*, 3rd ed. Berlin: Springer, 2004.

78. V. Weisskopf and E. Wigner. Berechnung der natürlichen linienbreite auf grund der diracschen lichttheorie. *Zeitschrift für Physik*, 63:54, 1930.
79. P. A. M. Dirac. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society London A*, 114:243, 1927.
80. G. Slavcheva and Ph. Roussignol. Nonlinear coherent magneto-optical response of a single chiral carbon nanotube. *New Journal of Physics*, 12:103004, 2010.
81. L. Yang, M. P. Anantram, J. Han, and J. P. Lu. Band-gap change of carbon nanotubes: Effect of small uniaxial and torsional strain. *Physical Review B*, 60:13874, 1999.
82. G. L. A. Rikken and E. Raupach. Observation of magneto-chiral dichroism. *Nature*, 390:493, 390.
83. A. Ajiki and T. Ando. Electronic states of carbon nanotubes. *Journal of the Physical Society of Japan*, 62:1255, 1993.
84. Y. Zhang, A. Chang, J. Cao, Q. Wang, W. Kim, Y. Li, N. Morris, E. Yenilmez, J. Kong, and H. Daia. Electric-field-directed growth of aligned single-walled carbon nanotubes. *Applied Physics Letters*, 79:3155, 2001.
85. L. Jin, C. Bower, and O. Zhou. Alignment of carbon nanotubes in a polymer matrix by mechanical stretching. *Applied Physics Letters*, 73:1197, 1998.
86. Ge. G. Samsonidze, A. Grüneis, R. Saito, A. Jorio, A. G. Souza Filho, G. Dresselhaus, and M. S. Dresselhaus. Interband optical transitions in left- and right-handed single-wall carbon nanotubes. *Physical Review B*, 69:205402, 2004.
87. J.-S. Lauret, C. Voisin, G. Cassabois, C. Delalande, Ph. Roussignol, O. Jost, and L. Capes. Ultrafast carrier dynamics in single-wall carbon nanotubes. *Physical Review Letters*, 90:057404, 2003.
88. E. L. Ivchenko and B. Spivak. Chirality effects in carbon nanotubes. *Physical Review B*, 66:155404, 2002.
89. F. Wang, G. Dukovic, L. E. Brus, and T. F. Heinz. Time-resolved fluorescence of carbon nanotubes and its implication for radiative lifetimes. *Physical Review Letters*, 92:177401, 2000.
90. H. Bao, X. Ruan, and T. S. Fisher. Optical properties of ordered vertical arrays of multi-walled carbon nanotubes from FDTD simulations. *Optics Express*, 18:6347, 2010.
91. V. N. Popov and L. Henrard. Comparative study of the optical properties of single-walled carbon nanotubes within orthogonal and nonorthogonal tight-binding models. *Physical Review B*, 70:115407, 2004.
92. V. N. Popov. Curvature effects on the structural, electronic and optical properties of isolated single-walled carbon nanotubes within a symmetry-adapted non-orthogonal tight-binding model. *New Journal of Physics*, 6:17, 2004.
93. J. Jiang, J. Ding, and D. Y. Xing. Zeeman effect on the electronic spectral properties of carbon nanotubes in an axial magnetic field. *Physical Review B*, 62:13209, 2000.
94. E. D. Minot, Y. Yaish, V. Sazonova, and P. L. McEuen. Determination of electron orbital magnetic moments in carbon nanotubes. *Nature*, 428:536, 2004.
95. F. L. Shuy, C. P. Chang, R. B. Chen, C. W. Chiu, and M. F. Lin. Magnetoelectronic and optical properties of carbon nanotubes. *Physical Review B*, 67:045405, 2003.
96. H. Ajiki. Magnetic-field effects on the optical spectra of a carbon nanotube. *Physical Review B*, 65:233409, 2002.
97. R. Saito, G. Dresselhaus, and M. S. Dresselhaus. Magnetic energy bands of carbon nanotubes. *Physical Review B*, 50:14698, 1994.
98. K. J. Boller, A. Imamoglu, and S. E. Harris. Observation of electromagnetically induced transparency. *Physical Review Letters*, 66:2593, 1991.

X

Mathematical Methods

50 Drift-Diffusion Models *Patricio Farrell, Nella Rotundo, Duy Hai Doan, Markus Kantner, Jürgen Fuhrmann, and Thomas Koprucki* 733
Introduction • How to Get Started: Numerical Solution of the Stationary 1D Semiconductor Equations • Van Roosbroeck System • Discretizing the van Roosbroeck System • Nonlinear Solvers • Mesh Generation for the Finite Volume Method • Time Discretization • Contact Terminal Currents • An Alternative: The Finite Element Method • Extensions and Outlook

51 Monte Carlo Device Simulations *Katerina Raleva, Abdul R. Shaik, Raghuraj Hathwar, Akash Laturia, Suleman S. Qazi, Robin Daugherty, Dragica Vasileska, and Stephen M. Goodnick* ... 773
Introduction • Bulk Monte Carlo Method Description • Particle-Based Device Simulation • Representative Simulation Results • Outlook

52 Photonics *Frank Schmidt* 807
Landscape of Numerical Methods • Maxwell's Equations in Nondispersive, Chargeless Media • FDTD Method • Fourier Modal Method • Finite Element Method • DG Method

50

Drift-Diffusion Models

50.1 Introduction 733
50.2 How to Get Started: Numerical Solution of the Stationary 1D Semiconductor Equations 734
1D van Roosbroeck Model • Thermodynamic Equilibrium and Local Charge Neutrality • Discretizing the Poisson Problem • Solving the Discrete Nonlinear Poisson Problem • Discretizing the Continuity Equations Using the Scharfetter–Gummel Scheme • Solving the Full Discretized van Roosbroeck System • Why Not Use Central Finite Differences for the Fluxes?
50.3 Van Roosbroeck System 745
The van Roosbroeck System of Equations • Initial and Boundary Conditions • Recombination Processes • Thermodynamic Equilibrium • Non-Boltzmann Statistics and Generalized Einstein Relation • Free Energy and Dissipation Rate • Existence and Uniqueness Results • Maximum Principle
50.4 Discretizing the van Roosbroeck System 753
Finite Volume Method in Higher Dimensions • Scharfetter–Gummel Fluxes and Their Non-Boltzmann Generalizations • Flux Expressions Consistent with the Thermodynamic Equilibrium • Free Energy and Dissipation Rate • Existence, Uniqueness, and Convergence • Maximum Principle
50.5 Nonlinear Solvers 759
Newton's Method • Gummel's Method
50.6 Mesh Generation for the Finite Volume Method 761
Boundary Conforming Delaunay Meshes on Polyhedral Domains • Tensor Product Approaches • Open Problems
50.7 Time Discretization 763
50.8 Contact Terminal Currents 763
50.9 An Alternative: The Finite Element Method 765
50.10 Extensions and Outlook 766
Additional Physical Models • Coupling the van Roosbroeck System to Other Models • Methods for Doping Optimization • Alternative Modeling Approaches for Carrier Transport

Patricio Farrell
Nella Rotundo
Duy Hai Doan
Markus Kantner
Jürgen Fuhrmann
and
Thomas Koprucki

50.1 Introduction

The semiconductor technology is undeniably one of the most important branches in modern industry. Apart from its obvious significance to our daily lives, this technology is an excellent example of a broad collaboration among various disciplines. The development of novel technologies and devices has not only been driven by engineers, but also by physicists, mathematicians, and numerical analysts. On the one hand, rigorous mathematical models allow sophisticated predictions, which might be difficult to observe experimentally. On the other hand, numerical simulations have the potential to optimize device designs without the costly and time-consuming development of prototypes.

In 1950, van Roosbroeck introduced the fundamental semiconductor device equations [1] as a system of three nonlinearly coupled partial differential equations (PDEs). They describe the semiclassical transport of free electrons and holes in a self-consistent electric field using a drift-diffusion approximation. Since then, the so-called *van Roosbroeck system* (frequently also called *drift-diffusion system*) became the standard model to describe the current flow in semiconductor devices at a macroscopic scale. Typical devices modeled by these equations range from diodes and transistors to LEDs, solar cells, and lasers [2,3]. In recent years, the emergence of quantum nanostructures [4–7] and organic semiconductors [8–10] has invigorated the research activities in this area.

In this chapter, we focus on solving the van Roosbroeck system numerically. From a mathematical point of view, the challenge lies in the strong nonlinearities, the drift dominated nature of the underlying physics, the formation of internal and boundary layers as well as the need to accurately mirror qualitative physical properties, such as nonnegativity of carrier densities and consistency with thermodynamical principles. The finite difference scheme invented by Scharfetter and Gummel [11] was the first to deal appropriately with most of these difficulties for charge carrier densities described by Boltzmann statistics. Later on, Fichtner, Rose, and Bank [12] extented it to higher dimensions as a finite volume scheme, providing geometric flexibility. Selberherr combined a lot of these results along with his own contributions in his textbook [13] which has become one of the standard reference works in the field.

However, often the Boltzmann approximation is insufficient and one has to impose Fermi–Dirac statistics or other constitutive laws for the charge carrier densities. For this reason, we present recent attempts to generalize the Scharfetter–Gummel approach while still reflecting the physics correctly.

In order to enable readers not familiar with semiconductor simulations to quickly grasp the main points in the discretization process of the van Roosbroeck system, we start by studying a one-dimensional (1D) problem in Section 50.2. Even though we have to postpone some of the solid-state physics to the third section, we can already introduce many key concepts of the standard finite difference ansatz proposed by Scharfetter and Gummel, which we will interpret as a 1D finite volume scheme.

In Section 50.3, we introduce the full van Roosbroeck system in higher spatial dimensions along with its physical concepts. In Section 50.4, we look at its numerical solution. We present the two-point flux finite volume method for the stationary case, which allows to use the Scharfetter–Gummel scheme in higher dimensions and discuss its generalization to more general carrier statistics. The remaining sections are devoted to special numerical aspects: nonlinear solvers, mesh generation, time stepping, the correct computation of terminal currents, and the finite element method. In the final section, we discuss generalizations of the approach and its embedding into more complex physical models.

50.2 How to Get Started: Numerical Solution of the Stationary 1D Semiconductor Equations

This section serves as a quick introduction to convey the main idea behind the van Roosbroeck system and its discretization. The van Roosbroeck system consists of three equations modeling the charge carrier flow as well as the electrostatic potential distribution in a semiconductor device. However, since the physics behind these equations is rather cumbersome and the focus of this chapter are numerical methods, we omit some of these physical details for now and start with a relatively simple 1D problem. Thus, we can highlight the key elements (and difficulties!) in the discretization of the van Roosbroeck system. This approach may appeal to newcomers to the field of optoelectronics. Experienced readers may jump directly to Section 50.3, where we will dive into the physical meaning of this model.

Here, we wish to model the stationary state of a 1D semiconductor device consisting of a homogeneous material on an interval $\Omega = [0, L]$ with Ohmic contacts at each end. This configuration already covers the simulation of p-n and p-i-n junctions.

50.2.1 1D van Roosbroeck Model

The stationary 1D van Roosbroeck system consists of three nonlinear ordinary differential equations for the unknown electrostatic potential $\psi(x)$, the quasi-Fermi potential for electrons $\varphi_n(x)$ and the quasi-Fermi potential for holes $\varphi_p(x)$. It links Poisson's equation for the electric field to the continuity equations for the carrier densities as follows:

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\varepsilon_s \frac{\mathrm{d}}{\mathrm{d}x}\psi\right) = q\Big(C + p(\psi, \varphi_p) - n(\psi, \varphi_n)\Big), \tag{50.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} j_n = \quad qR(\psi, \varphi_n, \varphi_p), \tag{50.2}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} j_p = -qR(\psi, \varphi_n, \varphi_p). \tag{50.3}$$

The fluxes or current densities are given by

$$j_n = -q\mu_n n(\psi, \varphi_n)\frac{\mathrm{d}}{\mathrm{d}x}\varphi_n \qquad \text{and} \qquad j_p = -q\mu_p p(\psi, \varphi_p)\frac{\mathrm{d}}{\mathrm{d}x}\varphi_p. \tag{50.4}$$

The electron density n and the hole density p are related to the electric potential ψ as well as the quasi-Fermi potentials of electrons and holes via

$$n(\psi, \varphi_n) = N_c \exp\left(\frac{q(\psi - \varphi_n) - E_c}{k_\mathrm{B}T}\right) \quad \text{and} \quad p(\psi, \varphi_p) = N_v \exp\left(\frac{q(\varphi_p - \psi) + E_v}{k_\mathrm{B}T}\right). \tag{50.5}$$

We stress here that in (50.5) we have used the so-called Boltzmann approximation. In general, the exponentials will be replaced by some monotonically increasing *statistical distribution function* $\mathcal{F}$. This function will be discussed in greater detail in Section 50.3.

Next, we explain the constants and functions in the van Roosbroeck system. The elementary charge q and the Boltzmann constant k_B are universal constants. The (absolute) dielectric permittivity $\varepsilon_\mathrm{s} = \varepsilon_0\varepsilon_\mathrm{r}$ is given as the product of the vacuum dielectric permittivity ε_0 and the relative permittivity of the semiconductor ε_r in static (low frequency) limit. The carrier mobilities μ_n and μ_p, the conduction and valence band densities of states N_c and N_v as well as the conduction and valence band-edge energies E_c and E_v are assumed to be constant even though in general they can vary with the material (e.g., due to abrupt or graded heterojunctions). The temperature T is also assumed to be constant; in general it can be space and even time dependent. The doping profile $C = C(x)$ describes material properties and the function $R(\psi, \varphi_n, \varphi_p)$ models the recombination and generation of electron–hole pairs. The electric field $\mathbf{E}$ is determined by the derivative (in higher dimensions the gradient) of the electrostatic potential

$$\mathbf{E} = -\frac{\mathrm{d}}{\mathrm{d}x}\psi\, \mathbf{e}_x,$$

where $\mathbf{e}_x$ denotes the unit normal along the x axis. A more detailed discussion on the physics behind these different quantities will be given in Section 50.3.

In the literature, authors frequently use a different set of unknowns for the van Roosbroeck system by replacing the quasi-Fermi potentials with the carrier densities n and p. However, we prefer to use the quasi-Fermi potentials for several reasons: they allow to easily model abrupt or graded heterostructures and appear naturally in the thermodynamic description of the van Roosbroeck system since the negative gradients of the quasi-Fermi potentials are the driving force of the currents [14]. Even mathematically, they are more beautiful as they make it possible to write the whole van Roosbroeck system in a gradient form [15] consistent with the principles of nonequilbrium thermodynamics [16]. Also carrier densities as opposed to quasi-Fermi potentials vary drastically in magnitude, which may lead to numerical difficulties.

The system of Equations 50.1 through 50.3 needs to be supplied with boundary conditions at $x = 0$ and $x = L$. For applied external voltages U_1 and U_2, we require Dirichlet boundary conditions at the *Ohmic contacts*, i.e.,

$$\psi(0) = \psi_0(0) + U_1, \qquad \psi(L) = \psi_0(L) + U_2, \tag{50.6}$$

$$\varphi_n(0) = U_1, \qquad \varphi_n(L) = U_2, \tag{50.7}$$

$$\varphi_p(0) = U_1, \qquad \varphi_p(L) = U_2. \tag{50.8}$$

We discuss the meaning of the potential ψ_0 in Section 50.2.2.

50.2.2 Thermodynamic Equilibrium and Local Charge Neutrality

The goal is to enforce that the van Roosbroeck system of (50.1) through (50.3) is consistent with the *thermodynamic equilibrium*, which is a physical state defined by vanishing currents, namely

$$j_n = j_p = 0 \qquad \text{implying} \qquad \varphi_0 := \varphi_n = \varphi_p = \text{const.} \tag{50.9}$$

Without loss of generality, we can set $\varphi_0 = 0$. We also discuss the thermodynamic equilibrium in Section 50.3.4. As a consequence, the three differential equations (50.1) through (50.3) reduce to the 1D nonlinear Poisson equation

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\varepsilon_s \frac{\mathrm{d}}{\mathrm{d}x}\psi\right) = q\left(N_v \exp\left(\frac{E_v - q\psi}{k_{\mathrm{B}}T}\right) - N_c \exp\left(\frac{q\psi - E_c}{k_{\mathrm{B}}T}\right) + C\right). \tag{50.10}$$

We supply it with Dirichlet boundary conditions (50.6) for zero voltages $U_1 = U_2 = 0$. The solution of Equation 50.10 with these boundary conditions is the *built-in potential* denoted with ψ_{eq}. In general, we cannot expect to find an analytic expression for ψ_{eq}.

Another important physical concept is *local charge neutrality*, characterized by a vanishing left-hand side in the Poisson equation. Combined with the equilibrium condition (50.9), this leads to

$$0 = q\left(N_v \exp\left(\frac{E_v - q\psi_0}{k_{\mathrm{B}}T}\right) - N_c \exp\left(\frac{q\psi_0 - E_c}{k_{\mathrm{B}}T}\right) + C\right), \tag{50.11}$$

which can be solved for ψ_0 by solving a quadratic equation and omitting the unphysical solution, yielding

$$\psi_0(x) = \frac{E_{\mathrm{c}} + E_{\mathrm{v}}}{2q} - \frac{1}{2}U_{\mathrm{T}} \log\left(\frac{N_c}{N_v}\right) + U_{\mathrm{T}} \operatorname{arcsinh}\left(\frac{C}{2N_{\mathrm{intr}}}\right), \tag{50.12}$$

where the intrinsic carrier density N_{intr} and the *thermal voltage* U_{T} are given by

$$N_{\mathrm{intr}}^2 = N_c N_v \exp\left(-\frac{E_c - E_v}{k_{\mathrm{B}}T}\right) \qquad \text{and} \qquad U_T = \frac{k_{\mathrm{B}}T}{q}. \tag{50.13}$$

The potential ψ_0 serves two purposes. In Section 50.2.4, it will be used to obtain an initial guess for the nonlinear iteration to solve a discretized version of Equation 50.10. On the other hand, modeling (ideal) Ohmic contacts requires local charge neutrality. Hence, the boundary values in (50.6) are obtained by evaluating the function ψ_0 at the contacts $x = 0$ and $x = L$. The voltage difference between the boundary values $\psi_0(0)$ and $\psi_0(L)$ is called *built-in voltage*.

50.2.3 Discretizing the Poisson Problem

Our goal is to find a discrete solution to Equation 50.1. For this, we will use the *finite volume method* which naturally preserves many important physical properties. To see this, the reader may have a quick look in the table of contents and compare the next two sections devoted to the van Roosbroeck system and its numerical solution. It will become apparent that for many aspects the continuous model and its discretized counterpart are closely related.

We discretize the interval using a (possibly) nonuniform mesh of the form

$$0 =: x_1 < \cdots < x_N := L.$$

We associate each node x_k with a *control volume* ω_k defined by the interior subintervals

$$\omega_k = \left[x_{k-1,k}, x_{k,k+1}\right] \quad \text{for} \quad k = 2, \ldots, N-1$$

and the boundary subintervals

$$\omega_1 = \left[x_1, x_{1,2}\right] \quad \text{and} \quad \omega_N = \left[x_{N-1,N}, x_N\right].$$

The boundaries of the kth control volume are given by

$$x_{k-1,k} := \frac{x_{k-1} + x_k}{2} \qquad \text{and} \qquad x_{k,k+1} := \frac{x_k + x_{k+1}}{2}$$

for $k = 2, \ldots, N-1$. We point out that we have generated a disjoint partition of the interval, $\Omega = \bigcup_{k=1}^{N} \omega_k$. This discretization along with its notation is illustrated in Figure 50.1.

We introduce a finite volume discretization for Poisson's equation. Using the fundamental theorem of calculus, we integrate Equation 50.1 over the interior control volumes ω_k and obtain Gauss's law of electrodynamics:

$$-\varepsilon_s \left(\left.\frac{\mathrm{d}\psi}{\mathrm{d}x}\right|_{x_{k,k+1}} - \left.\frac{\mathrm{d}\psi}{\mathrm{d}x}\right|_{x_{k-1,k}} \right) = \int_{\omega_k} q\Big(C + p(\psi, \varphi_p) - n(\psi, \varphi_n)\Big)\, \mathrm{d}x. \tag{50.14}$$

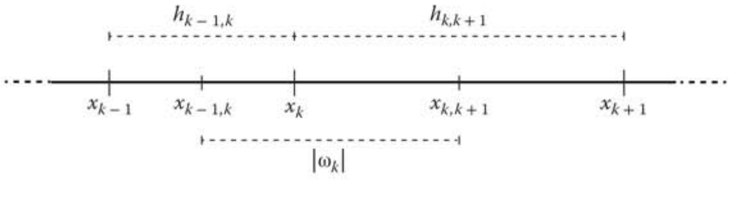


FIGURE 50.1 1D discretization around the kth node along with the corresponding notation.

This equation constitutes a balance law that can be interpreted as follows: the total charge in the kth control volume is given by the difference of the electric displacement $D = -\varepsilon_s \frac{\mathrm{d}}{\mathrm{d}x}\psi$ at its interfaces.

So far, we have derived an integral form of Poisson's equation for every interior control volume. We need approximations of the electric displacements at the boundary of each control volume. For this we employ central finite differences:

$$\left.\frac{\mathrm{d}\psi}{\mathrm{d}x}\right|_{x_{k,k+1}} \approx \frac{\psi_{k+1} - \psi_k}{h_{k,k+1}}, \tag{50.15}$$

with the distance between two neighboring nodes given by

$$h_{k,k+1} := x_{k+1} - x_k.$$

Moreover, the integral over the net charge density is approximated by the rectangle method, evaluating the integrand at node x_k times the size of the control volume $|\omega_k| = x_{k,k+1} - x_{k-1,k}$. That is, the balance equation (50.14) is approximated by

$$-\varepsilon_s\left(\frac{\psi_{k+1} - \psi_k}{h_{k,k+1}} - \frac{\psi_k - \psi_{k-1}}{h_{k-1,k}}\right) = q\Big(C_k + p(\psi_k, \varphi_{p;k}) - n(\psi_k, \varphi_{n;k})\Big)|\omega_k|$$

or, slightly rearranged,

$$\begin{aligned} 0 = -\varepsilon_s\left(\frac{1}{h_{k,k+1}}\psi_{k+1} - \left[\frac{1}{h_{k,k+1}} + \frac{1}{h_{k-1,k}}\right]\psi_k + \frac{1}{h_{k-1,k}}\psi_{k-1}\right) \\ - q\Big(C_k + p(\psi_k, \varphi_{p;k}) - n(\psi_k, \varphi_{n;k})\Big)|\omega_k|, \end{aligned} \tag{50.16}$$

where $k = 2, \ldots, N-1$. The subindices denote evaluation at a given node. So, for example, ψ_k means $\psi(x_k)$ as well as $\varphi_{n;k}$ indicates $\varphi_n(x_k)$. We use this nodal notation frequently within this chapter. To incorporate the Dirichlet boundary conditions (50.6), we simply impose two more equations for the endpoints, namely

$$\psi_1 = \psi_0(0) + U_1 \qquad \text{and} \qquad \psi_N = \psi_0(L) + U_2. \tag{50.17}$$

50.2.4 Solving the Discrete Nonlinear Poisson Problem

Using the finite volume discretization scheme we have just described, we solve the nonlinear Poisson equation (50.16) with Dirichlet boundary conditions (50.17) in thermodynamic equilibrium ($U_1 = U_2 = 0$) to obtain the built-in potential ψ_{eq}. Recalling that there are $N-2$ interior and two boundary nodes, we end up with a system of N variables which can be summarized by the nonlinear discrete system

$$\mathbf{0} = \mathbf{F}_1(\boldsymbol{\psi}_{\mathrm{eq}}),$$

where $\boldsymbol{\psi}_{\mathrm{eq}} = (\psi_{\mathrm{eq}}(x_k))_{k=1}^N$ is the vector of the nodal values. Solving this system is not straightforward due to its inherent nonlinearity. One approach to obtain a solution is Newton's method, which we discuss in more detail in Section 50.5. Newton's method converges quadratically—if one has a starting guess sufficiently close to the solution [17].

50.2.4.1 Solution Procedure: Nonlinear Poisson Problem

1. The easiest way to obtain a starting guess for the solution of Equation 50.16 with Dirichlet boundary conditions (50.17) is to neglect the left-hand side. That is, we use the local charge neutrality condition (50.11) for each node, which we can solve explicitly for ψ_0 via (50.12). With this function, we determine the components (nodal values) of our starting guess for the nonlinear Poisson problem $\boldsymbol{\psi}_\mathbf{0}$ via

$$\boldsymbol{\psi}_\mathbf{0} = (\psi_0(x_k))_{k=1}^N.$$

2. Solving the discrete nonlinear Poisson problem via Newton's method with the starting guess $\boldsymbol{\psi}_\mathbf{0}$ yields the built-in potential vector solution $\boldsymbol{\psi}_{\text{eq}}$.

This solution procedure for the nonlinear Poisson equation for a p-i-n structure is illustrated in Figure 50.2. In the following, we address how to discretize the continuity equations. A task which turns out to be nontrivial.

50.2.5 Discretizing the Continuity Equations Using the Scharfetter–Gummel Scheme

Now, we turn our attention to the continuity equations (50.2) and (50.3). We proceed just as for Poisson's equation: we integrate over the control volumes $\omega_k = [x_{k-1,k}, x_{k,k+1}]$, apply the fundamental theorem of calculus for the left-hand sides and approximate the integral of the recombination term as before. For the interior control volumes we obtain

$$\begin{aligned} 0 &= j_{n;k,k+1} - j_{n;k-1,k} - qR(\psi_k, \varphi_{n;k}, \varphi_{p;k})|\omega_k|, \\ 0 &= j_{p;k,k+1} - j_{p;k-1,k} + qR(\psi_k, \varphi_{n;k}, \varphi_{p;k})|\omega_k|, \end{aligned} \tag{50.18}$$

where $k = 2, \ldots, N-1$. Similar to before, these equations constitute a balance law in integral form which we can interpret as follows: the carrier densities within the control volume ω_k change either due to in- and outflow or by recombination.

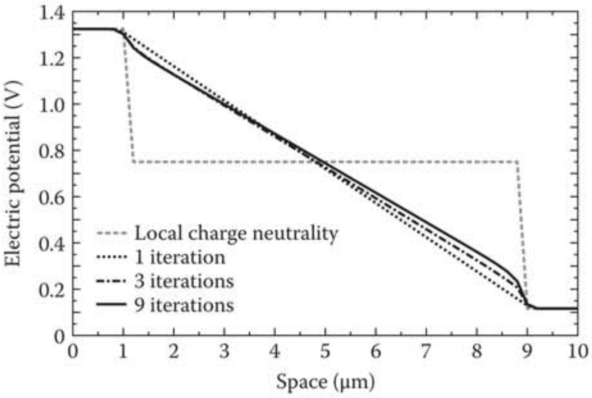


FIGURE 50.2 Convergence of the electric potential toward the thermodynamic equilibrium solution (built-in potential) for a GaAs p-i-n diode on an equidistant mesh with 51 nodes. The locally electroneutral initialization quickly converges to the built-in potential by means of a Newton iteration. The device has total length of 10 μm (1 μm n-doped with $N_\text{D} = 10^{16}\ \text{cm}^{-3}$, 8 μm intrinsic, 1 μm p-doped with $N_\text{A} = 10^{17}\ \text{cm}^{-3}$. Standard parameters for GaAs at $T = 300$ K: $E_v = 0$ eV, $E_c = 1.424$ eV, $N_v = 9.0 \times 10^{18}\ \text{cm}^{-3}$, $N_c = 4.7 \times 10^{17}\ \text{cm}^{-3}$, and $\varepsilon_r = 12.9$.

The fluxes (50.18) need to be expressed in terms of our set of unknowns, namely the nodal values of the electrostatic and the quasi-Fermi potentials. Deriving a suitable approximation for the current expressions $j_{n;k,k+1}$ and $j_{p;k,k+1}$ describing the carrier exchange between neighboring control volumes ω_k and ω_{k+1} is crucial for numerical simulations. Since standard discretization techniques such as central finite differences may lead to unphysical behavior (see Section 50.2.7), the derivation of such an approximation requires particular care.

For the Boltzmann approximation (50.5), a numerically stable scheme was invented in 1969 by Scharfetter and Gummel [11]. In fact, it has been discovered independently several times [18–20]. Its derivation is based on integrating the flux along the interval $[x_k, x_\ell]$ between two neighboring discretization nodes x_k and x_ℓ under the assumption of constant current density and constant electric field. That is, either $\ell = k-1$ or $\ell = k+1$. We illustrate the approach for the electron current density j_n.

In order to obtain a flux approximation, we will make two important assumptions, namely

1. The electric field along each edge is constant. This implies that the derivative of the electric potential $\frac{\mathrm{d}\psi}{\mathrm{d}x}$ along the edge $\left[x_k, x_\ell\right]$ is approximated by $\frac{\psi_\ell - \psi_k}{h_{k,\ell}}$, where again $\psi_\ell = \psi(x_\ell)$ and $\psi_k = \psi(x_k)$. This is consistent with the central difference approximation for the electric field in the Poisson equation, see (50.15).
2. The current density j_n is assumed to be constant along the edge.

The second assumption means that the current fluxes through the boundary from one control volume to another are conserved, which implies that the derivative of the flux along the interval $[x_k, x_\ell]$ vanishes. This means that we assume that there is no recombination along the edge between neighboring nodes. Hence, we obtain for the electron flux (the hole flux follows analogously)

$$\frac{\mathrm{d}}{\mathrm{d}x} j_n = \frac{\mathrm{d}}{\mathrm{d}x}\left(-q\mu_n N_c \exp\left(\frac{q(\psi(x) - \varphi_n(x)) - E_c}{k_\mathrm{B} T}\right) \frac{\mathrm{d}}{\mathrm{d}x}\varphi_n(x)\right) = 0. \tag{50.19}$$

We supply the differential equation with the boundary conditions

$$\varphi_n(x_k) = \varphi_{n;k} \qquad \text{and} \qquad \varphi_n(x_\ell) = \varphi_{n;\ell}. \tag{50.20}$$

The values $\varphi_{n;k}$ and $\varphi_{n;\ell}$ have to be determined. But for now we assume they are given. Equations 50.19 and 50.20 constitute a two-point boundary value problem. Upon integrating once and multiplication by the integrating factor

$$M(x) = \exp\left(-\frac{\psi_\ell - \psi_k}{h_{k,\ell} U_\mathrm{T}} x\right),$$

we obtain for the integration constant $j_{n;k,\ell}$ the equation

$$j_{n;k,\ell} M(x) = q\mu_n N_c U_\mathrm{T} \frac{\mathrm{d}}{\mathrm{d}x}\left(M(x) \exp\left(\frac{q(\psi(x) - \varphi_n(x)) - E_c}{k_\mathrm{B}\mathrm{T}}\right)\right).$$

Integrating once more from x_k to x_ℓ and using the boundary conditons (50.20) yields after some algebraic manipulations the Scharfetter–Gummel expression for the current flux along the edge $[x_k, x_\ell]$ between neighboring control volumes

$$j_{n;k,\ell} = j_{n;k,\ell}(\psi_k, \psi_\ell, \varphi_{n;k}, \varphi_{n;\ell}) = -\frac{q\mu_n U_\mathrm{T}}{h_{k,\ell}} \left[B\left(-\frac{\psi_\ell - \psi_k}{U_\mathrm{T}}\right) N_c \exp\left(\frac{q\left(\psi_k - \varphi_{n;k}\right) - E_c}{k_\mathrm{B} T}\right) \right. \\ \left. - B\left(\frac{\psi_\ell - \psi_k}{U_\mathrm{T}}\right) N_c \exp\left(\frac{q\left(\psi_\ell - \varphi_{n;\ell}\right) - E_c}{k_\mathrm{B}\mathrm{T}}\right) \right], \tag{50.21}$$

where we have introduced the *Bernoulli function*

$$B(x) = \frac{x}{\exp(x) - 1}.$$

This function smoothly interpolates between zero and $-x$, see Figure 50.3. Remembering the relationships (50.5), the flux can be brought into the simpler form

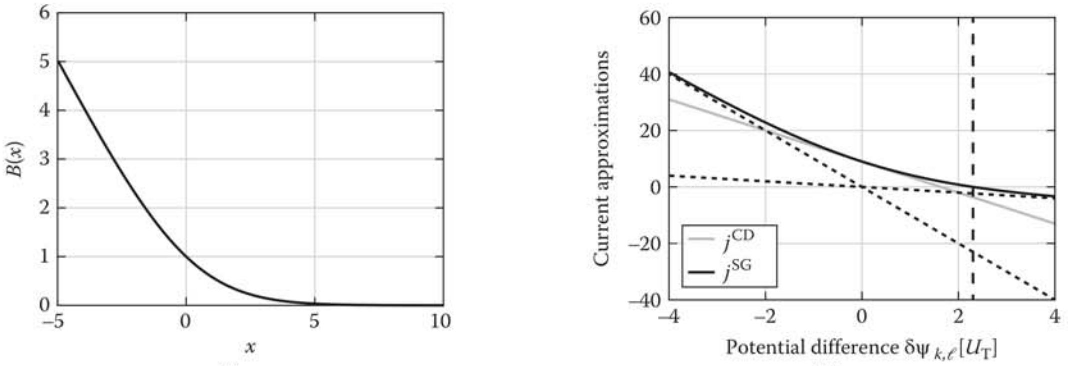
$$j_{n;k,\ell} = -\frac{q\mu_n U_\mathrm{T}}{h_{k,\ell}} \left[B\left(-\frac{\psi_\ell - \psi_k}{U_\mathrm{T}}\right) n_k - B\left(\frac{\psi_\ell - \psi_k}{U_\mathrm{T}}\right) n_\ell \right]. \tag{50.22}$$

An important property of this scheme is its consistency with the thermodynamic equilibrium. That is, similarly to (50.9), we have

$$j_{n;k,\ell} = 0 \quad \text{for} \quad \varphi_{n;k} = \varphi_{n;\ell} = \text{const.} \tag{50.23}$$

For vanishing electric fields corresponding to $\delta\psi_{k,\ell} := \psi_\ell - \psi_k = 0$, the flux $j_{n;k,\ell}$ reduces to a purely diffusive flux

$$j_{n;k,\ell} = -q\left(-D_n \frac{n_\ell - n_k}{h_{k,\ell}}\right).$$



(a) (b)

FIGURE 50.3 On the left the Bernoulli function $B(x)$ is shown. The picture on the right shows different current approximations for $q\mu_n U_\mathrm{T}/h_{k,\ell} = 1$, $n_\ell = 10$, and $n_k = 1$ in terms of the potential difference $\delta\psi_{k,\ell}$. The dashed black line denotes the equilibrium potential difference. The dotted lines show the drift currents, i.e., the asymptotics of the Scharfetter–Gummel flux (50.24) for large positive and for large negative arguments. The central-difference flux is denoted with j^{CD} and the Scharfetter–Gummel flux with j^{SG}.

The diffusion constant D_n is defined via the so-called *Einstein relation* $D_n = \mu_n U_\mathrm{T}$, which we discuss in Section 50.3.5 in more detail. For large electric fields, we obtain asymptotically a drift flux

$$j_{n;k,\ell} = \begin{cases} -q\mu_n n_k \frac{\psi_\ell - \psi_k}{h_{k,\ell}}, & \text{for } \delta\psi_{k,\ell} \gg +U_\mathrm{T}, \\ -q\mu_n n_\ell \frac{\psi_\ell - \psi_k}{h_{k,\ell}}, & \text{for } \delta\psi_{k,\ell} \ll -U_\mathrm{T}. \end{cases} \tag{50.24}$$

These two extreme cases show that the Scharfetter–Gummel scheme naturally encompasses drift and diffusive currents depending on the strength of the local electric field, see Figure 50.3.

Using analogous considerations, we can derive a current expression for the holes

$$j_{p;k,\ell} = \frac{q\mu_p U_\mathrm{T}}{h_{k,\ell}} \left[B\left(\frac{\psi_\ell - \psi_k}{U_\mathrm{T}}\right) p_k - B\left(-\frac{\psi_\ell - \psi_k}{U_\mathrm{T}}\right) p_\ell \right].$$

Since we have an approximation for the fluxes at the control volume boundary, we can derive from the discrete system (50.18) nonlinear equations for the interior nodes in terms of our set of unknowns $(\psi, \varphi_n, \varphi_p)$.

We stress once more that the previous discussion is only valid for the Boltzmann approximation made in (50.5). A discussion on more general distribution functions follows in Sections 50.3.5 and 50.4.2.

50.2.6 Solving the Full Discretized van Roosbroeck System

So far we have described how to derive a set of discrete and nonlinear equations from the 1D van Roosbroeck system ((50.1)–(50.3)), namely

$$\begin{aligned}
0 &= -\varepsilon_s \left(\frac{1}{h_{k,k+1}} \psi_{k+1} - \left[\frac{1}{h_{k,k+1}} + \frac{1}{h_{k-1,k}} \right] \psi_k + \frac{1}{h_{k-1,k}} \psi_{k-1} \right) \\
&\qquad - q\Big(C_k + p(\psi_k, \varphi_{p;k}) - n(\psi_k, \varphi_{n;k}) \Big) |\omega_k|, \\
0 &= j_{n;k,k+1}(\psi_k, \psi_{k+1}, \varphi_{n;k}, \varphi_{n;k+1}) - j_{n;k-1,k}(\psi_{k-1}, \psi_k, \varphi_{n;k-1}, \varphi_{n;k}) - qR(\psi_k, \varphi_{n;k}, \varphi_{p;k}) |\omega_k|, \\
0 &= j_{p;k,k+1}(\psi_k, \psi_{k+1}, \varphi_{p;k}, \varphi_{p;k+1}) - j_{p;k-1,k}(\psi_{k-1}, \psi_k, \varphi_{p;k-1}, \varphi_{p;k}) + qR(\psi_k, \varphi_{n;k}, \varphi_{p;k}) |\omega_k|,
\end{aligned} \tag{50.25}$$

with $k = 2, \ldots, N-1$. For the boundary nodes x_1 and x_N, the equations are given by the boundary conditions ((50.6)–(50.8)):

$$\begin{aligned}
\psi_1 &= \psi(0) = \psi_0(0) + U_1, & \varphi_{n,1} &= \varphi_n(0) = U_1, & \varphi_{p,1} &= \varphi_p(0) = U_1, \\
\psi_N &= \psi(L) = \psi_0(L) + U_2, & \varphi_{n,N} &= \varphi_n(L) = U_2, & \varphi_{p,N} &= \varphi_p(L) = U_2.
\end{aligned}$$

Just as for the Poisson problem we assume there are N control volumes in total. We end up with a system of $3N$ variables which we can summarize by the nonlinear discrete system

$$\mathbf{0} = \mathbf{F}(\boldsymbol{\psi}, \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_p) := \begin{pmatrix} \mathbf{F}_1(\boldsymbol{\psi}, \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_p) \\ \mathbf{F}_2(\boldsymbol{\psi}, \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_p) \\ \mathbf{F}_3(\boldsymbol{\psi}, \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_p) \end{pmatrix},$$

where

$$\boldsymbol{\psi} = (\psi(x_k))_{k=1}^{N}, \qquad \boldsymbol{\varphi}_n = (\varphi_n(x_k))_{k=1}^{N}, \qquad \boldsymbol{\varphi}_p = (\varphi_p(x_k))_{k=1}^{N}.$$

Again, we employ Newton's method to solve this system. We aim to guarantee an appropriate starting guess by starting from thermodynamic equilibrium where the three equations reduce to one. The solution of the full van Roosbroeck system with an applied bias is then obtained via an iteration technique that we state here.

50.2.6.1 Solution Procedure: Full van Roosbroeck System

1. We assume thermodynamic equilibrium and set the constant quasi-Fermi potentials to $\varphi_0 = 0$. With the solution procedure introduced in Section 50.2.4, we obtain a solution vector $\boldsymbol{\psi}_{\text{eq}}$ for the nonlinear Poisson problem with Dirichlet boundary conditions.
2. We denote with $\boldsymbol{\varphi}_{\boldsymbol{n}}^{\mathbf{0}}$ and $\boldsymbol{\varphi}_{\boldsymbol{p}}^{\mathbf{0}}$ the constant vectors, consisting of the equilibrium value for the quasi-Fermi potentials, namely

   $$\boldsymbol{\varphi}_{\boldsymbol{n}}^{\mathbf{0}} = (\varphi_0)_{i=1}^{N} \qquad \text{and} \qquad \boldsymbol{\varphi}_{\boldsymbol{p}}^{\mathbf{0}} = (\varphi_0)_{i=1}^{N}.$$

   We choose zero boundary voltages $U_1 = U_2 = 0$ and remember that we have obtained $\psi_0(0)$ and $\psi_0(L)$ via the charge neutrality condition. Thus we have determined the boundary conditions ((50.6)–(50.8)). Using the equilibrium vector $(\boldsymbol{\psi}_{\text{eq}}, \boldsymbol{\varphi}_{\boldsymbol{n}}^{\mathbf{0}}, \boldsymbol{\varphi}_{\boldsymbol{p}}^{\mathbf{0}})$ as a new starting guess, we can now solve the full discrete system (50.25) with these boundary conditions via Newton's method, yielding a solution $(\boldsymbol{\psi}^{\mathbf{1}}, \boldsymbol{\varphi}_{\boldsymbol{n}}^{\mathbf{1}}, \boldsymbol{\varphi}_{\boldsymbol{p}}^{\mathbf{1}})$. That is, we compute the equilibrium state of the full van Roosbroeck system using the equilibrium of the Poisson equation as a starting guess. We point out that solution and starting guess should agree up to machine precision by construction.
3. We slightly increase U_1 and U_2 to generate a "small" bias $\delta U = U_2 - U_1$. What small here means is not very precise and needs to be determined via trial and error. Now we solve the full discrete system (50.25) via Newton's method with the starting guess $(\boldsymbol{\psi}^{\mathbf{1}}, \boldsymbol{\varphi}_{\boldsymbol{n}}^{\mathbf{1}}, \boldsymbol{\varphi}_{\boldsymbol{p}}^{\mathbf{1}})$ to obtain a new solution $(\boldsymbol{\psi}^{\mathbf{2}}, \boldsymbol{\varphi}_{\boldsymbol{n}}^{\mathbf{2}}, \boldsymbol{\varphi}_{\boldsymbol{p}}^{\mathbf{2}})$.
4. Now we iterate the last step: gradually increase the bias and use the old solution as starting guess for the new one.

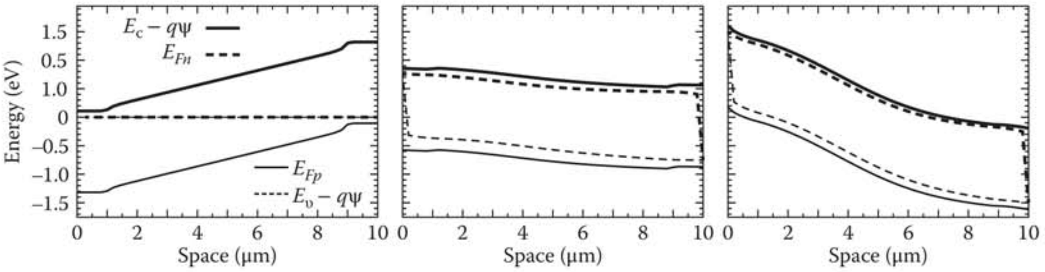
Figures 50.4 and 50.5 show the numerical results for the p-i-n structure, already introduced in Section 50.2.4. They depict the band edge energies, the quasi-Fermi energy levels $E_{F_n} = -q\varphi_n$ and $E_{F_p} = -q\varphi_p$ as well as the carrier densities for different bias values. Additionally, the current voltage characteristics for different n-dopings are shown.

50.2.7 Why Not Use Central Finite Differences for the Fluxes?

Thus far, we have discussed the Scharfetter–Gummel scheme to approximate the flux at the control volume interface. One might ask why not use a simpler flux approximation, for example, central finite differences. This question we address now.

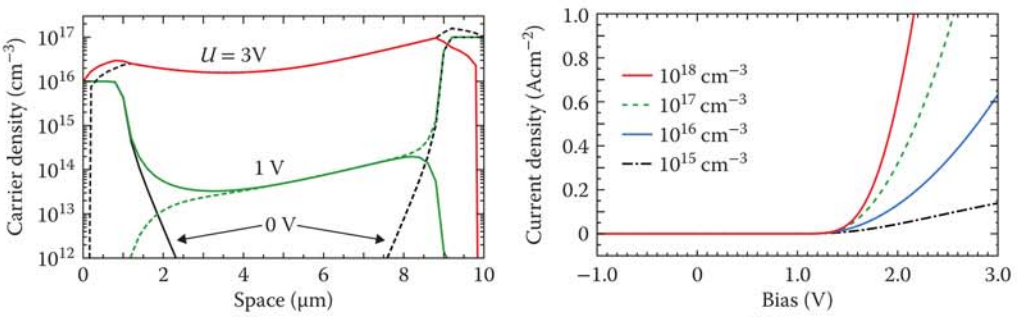
For simplicity we assume a linear electrostatic potential. In this case

$$\frac{\mathrm{d}}{\mathrm{d}x}\psi = \psi(1) - \psi(0) =: \delta\psi.$$



(a) (b) (c)

FIGURE 50.4 Computed band edge energy and quasi-Fermi levels for a GaAs p-i-n diode at different values of the applied voltage. (a) Thermodynamic equilibrium (off-state, 0V), (b) $U = 1.5$ V (flat band), (c) $U = 3.0$ V. Device has total length of 10 μm (1 μm n-doped with $N_D = 10^{16}$ cm^{-3}, 8 μm intrinsic, 1 μm p-doped with $N_A = 10^{17}$ cm^{-3}. The parameters used are standard parameters for GaAs as given in the caption of Figure 50.2 and $\mu_n = 8500$ cm^2 V^{-1} s^{-1}, $\mu_p = 400$ cm^2 V^{-1} s^{-1}, $\tau_n = \tau_p = 1$ ns, $C_n = C_p = 1 \times 10^{-29}$ cm^6 s^{-1}, and $r_{\text{spont}} = 1 \times 10^{-10}$ cm^3 s^{-1}.



(a) (b)

FIGURE 50.5 (a) Carrier densities in the p-i-n-diode at different applied voltages. Electron densities are shown as solid lines, holes are illustrated by dashed lines. One observes how the intrinsic region gets populated by free carriers while increasing the applied bias. (b) Current–voltage curves for the same device with altered values of the donor density N_{D}. Same parameters as in Figure 50.4.

We focus on the continuity equation for electrons on the domain $\Omega = [0, 1]$ which results in a linear differential equation of second order,

$$\frac{\mathrm{d}}{\mathrm{d}x} j_n = qD_n \frac{\mathrm{d}^2}{\mathrm{d}x^2} n - q\mu_n \delta\psi \frac{\mathrm{d}}{\mathrm{d}x} n = 0. \tag{50.26}$$

In this equation, we have introduced the drift-diffusion form of the (electron) flux

$$j_n = q\left(D_n \frac{\mathrm{d}}{\mathrm{d}x} n - n\mu_n \frac{\mathrm{d}}{\mathrm{d}x}\psi\right), \tag{50.27}$$

where we have used again the Einstein relation $D_n = \mu_n U_{\text{T}}$. We omit the physical derivation of the flux here. It is only important to know that this way of characterizing the flux is equivalent to Equation 50.4.

We have already seen how to solve this type of equation numerically by solving *local* two-point boundary value problems for the flux, which yields the Scharfetter–Gummel scheme (50.22). Of course, there are many different ways to discretize the flux (50.27) at a control volume interface $x_{k,\ell}$, needed in (50.18). For example, looking at Equation 50.27 one might consider replacing the derivative of the electron density $\frac{\mathrm{d}}{\mathrm{d}x} n$ with a central finite difference and the density n with an average. That is, at the control volume interface

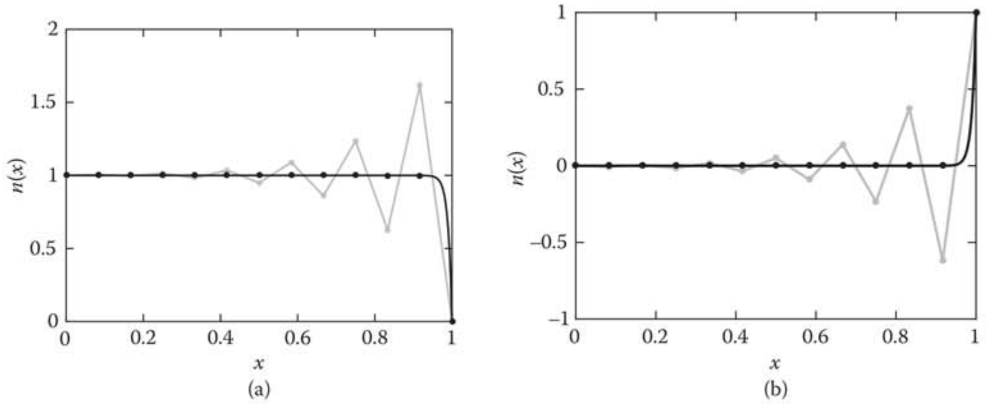


FIGURE 50.6 The exact solution (black) of Equation 50.26 and its central difference approximation (gray) on a uniform mesh with mesh width $h = 1/12$. The boundary conditions (50.28) lead to a numerical solution which violates the maximum principle (left). On the other hand, the boundary conditions (50.29) produce a numerical solution which violates the positivity of the density (right). Here $qD_n = 0.01$ and $q\mu_n\delta\psi = 1$.

we make the approximation

$$j_{n;k,\ell}^{\mathrm{CD}} = qD_n\frac{n_\ell - n_k}{h_{k,\ell}} - q\mu_n\delta\psi\,\frac{n_\ell + n_k}{2} = -\frac{q\mu_n U_{\mathrm{T}}}{h_{k,\ell}}\left(n_k - n_\ell + \frac{\delta\psi_{k,\ell}}{U_{\mathrm{T}}}\,\frac{n_\ell + n_k}{2}\right).$$

This, however, is a very bad idea since the scheme will become unstable, resulting under certain conditions in large oscillations. To show this, we supply Equation 50.26 with two different types of boundary conditions:

$$n(0) = 1 \qquad \text{and} \qquad n(1) = 0 \tag{50.28}$$

as well as

$$n(0) = 0 \qquad \text{and} \qquad n(1) = 1. \tag{50.29}$$

Then, the discretization using central finite differences shows large oscillations for $\mu_n\delta\psi/D_n = 100$, which can be seen in Figure 50.6. For the first set of boundary conditions, we can see that the *maximum principle* is violated, i.e., the solution exceeds both boundary values [21,22]. For the second set of boundary conditions, the numerically computed density becomes negative. Both observations constitute huge violations of basic physical laws and are unacceptable in applications.

The problem here is that the diffusion constant D_n is small compared to the drift constant $\mu_n\delta\psi$. We point out, that in this special 1D case of vanishing recombination the Scharfetter–Gummel scheme, on the other hand, yields *exact* nodal values due to its construction. The Scharfetter–Gummel and the central difference flux are compared in Figure 50.3.

50.3 Van Roosbroeck System

In this section, we make several generalizations to the van Roosbroeck system ((50.1)–(50.3)). In particular, we make the model time dependent and allow higher spatial dimensions. Before stating the equations, we discuss the physical concepts we want to encode in the model.

As in the 1D case, the van Roosbroeck system consists of three equations: the Poisson equation and two continuity equations for the carrier densities. The Poisson equation describes the electric field $\mathbf{E} = -\nabla\psi$ which is generated by a scalar electric potential $\psi(\mathbf{x}, t)$ in the presence of a free charge carrier density. In a (doped) bipolar semiconductor device, this charge density has three ingredients: the density of free (negatively charged) electrons $n(\mathbf{x}, t)$ occupying the conduction band, the density of (positively charged) holes $p(\mathbf{x}, t)$ occupying the valence band and the density of ionized built-in dopants $C(\mathbf{x}) = N_\mathrm{D}^+(\mathbf{x}) - N_\mathrm{A}^-(\mathbf{x})$, where N_D^+ denotes the density of singly ionized donor atoms and N_A^- is the density of singly ionized acceptor atoms.

The continuity equations (in differential form), on the other hand, model the flow of the charge carrier densities due to diffusion and drift governed by the self-consistent electric field, which is generated by the net charge density. Furthermore, the recombination and generation of electron–hole pairs influences the electron and hole densities.

50.3.1 The van Roosbroeck System of Equations

For a bounded spatial domain $\Omega \subset \mathbb{R}^d$ where $d \in \{1, 2, 3\}$, the van Roosbroeck system consists of three coupled nonlinear PDEs of the form

$$-\nabla \cdot \left(\varepsilon_s \nabla\psi\right) = q\left(p - n + C\right), \tag{50.30}$$

$$-q\partial_t n + \nabla \cdot \mathbf{j}_n = qR, \tag{50.31}$$

$$q\partial_t p + \nabla \cdot \mathbf{j}_p = -qR, \tag{50.32}$$

for $\mathbf{x} \in \Omega$ and $t \in [0, T]$. The current densities in Equations 50.31 and 50.32 are given by the usual expressions

$$\mathbf{j}_n = -q\mu_n n\nabla\varphi_n \qquad \text{and} \qquad \mathbf{j}_p = -q\mu_p p\nabla\varphi_p. \tag{50.33}$$

That is, the negative gradients of the quasi-Fermi potentials are the driving forces of the currents [14]. These relationships correspond directly to Equation 50.4.

Within the framework of effective mass approximation [2], the densities of free carriers in a solid are given by

$$n = N_c\mathcal{F}\left(\frac{q(\psi - \varphi_n) - E_c}{k_\mathrm{B}T}\right) \qquad \text{and} \qquad p = N_v\mathcal{F}\left(\frac{E_v - q(\psi - \varphi_p)}{k_\mathrm{B}T}\right). \tag{50.34}$$

Equations 50.34 indicate that the electric potential effectively leads to a bending of the energy band edge levels and thus a nonlinear, self-consistent coupling to the carrier densities is achieved. We assume a globally constant temperature for both carrier species and the crystal lattice.

The function $\mathcal{F}$ describes the occupation of energy states in the semiconductor. Since it plays an important role in the following discussion, we will define and discuss it later in Section 50.3.5. We only point out here that Equations 50.34 is analogous to Equations 50.5 when choosing for $\mathcal{F}$ the exponential function, the so-called *Boltzmann approximation*.

50.3.2 Initial and Boundary Conditions

The system ((50.30)–(50.32)) must be supplemented with initial and boundary conditions. The initial conditions at time $t = 0$ are given by the initial distributions ψ^I, φ_n^I and φ_p^I, i.e.,

$$\psi(\mathbf{x}, 0) = \psi^I(\mathbf{x}), \qquad \varphi_n(\mathbf{x}, 0) = \varphi_n^I(\mathbf{x}), \qquad \varphi_p(\mathbf{x}, 0) = \varphi_p^I(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega.$$

Regarding the boundary conditions, we briefly discuss the most important case where the boundary of the domain Ω can be decomposed into Ohmic contacts, a gate contact and *artificial* interfaces, i.e.,

$$\partial\Omega = \left(\bigcup_{\alpha=1}^{N_O} \Gamma_{O,\alpha}\right) \cup \Gamma_G \cup \Gamma_A.$$

Semiconductor–metal interfaces, such as Ohmic contacts, are modeled by Dirichlet boundary conditions. For any Ohmic contact $\Gamma_{\text{O},\alpha}$ with $\alpha = 1, \dots, N_O$, we set for all $\mathbf{x} \in \Gamma_{O,\alpha}$ and $t \in [0, T]$

$$\psi(\mathbf{x}, t) = \psi_0(\mathbf{x}) + U_\alpha(t), \tag{50.35}$$

$$\varphi_n(\mathbf{x}, t) = U_\alpha(t), \tag{50.36}$$

$$\varphi_p(\mathbf{x}, t) = U_\alpha(t), \tag{50.37}$$

where U_α denotes the corresponding externally applied contact voltage. The value ψ_0 at the boundary is defined by local charge neutrality similar to (50.11), where the exponential is now replaced by $\mathcal{F}$:

$$0 = N_v \mathcal{F}\left(\frac{E_v - q\psi_0(\mathbf{x})}{k_B T}\right) - N_c \mathcal{F}\left(\frac{q\psi_0(\mathbf{x}) - E_c}{k_B T}\right) + C(\mathbf{x}).$$

We just remark that, in general, this equation yields no analytical solution, and therefore its solution needs to be obtained by a nonlinear solver. The boundary conditions for the more advanced nonlinear semiconductor–metal interfaces (Schottky contacts) can be found in [13,23].

Gate contacts are modeled by Robin boundary conditions for the electrostatic potential and homogeneous Neumann boundary conditions for the quasi-Fermi potentials

$$\left.\begin{aligned} \varepsilon_s \nabla\psi(\mathbf{x}, t) \cdot \boldsymbol{\nu} + \tfrac{\varepsilon_{\text{ox}}}{d_{\text{ox}}}(\psi(\mathbf{x}, t) - U_{\text{G}}(t)) &= 0, \\ \mathbf{j}_n(\mathbf{x}, t) \cdot \boldsymbol{\nu} = \mathbf{j}_p(\mathbf{x}, t) \cdot \boldsymbol{\nu} &= 0, \end{aligned}\right. \qquad \text{for all } \mathbf{x} \in \Gamma_{\text{G}} \text{ and } t \in [0, T], \tag{50.38}$$

where ε_{ox} and d_{ox} are the absolute dielectric permittivity and the thickness of the oxide, respectively. The function $U_{\text{G}}(t)$ is the applied voltage at the outside of the insulating gate oxide at Γ_{G}. Here $\boldsymbol{\nu}$ denotes the outer normal vector on the interface.

On the remaining (artificial) interfaces, one typically imposes homogeneous Neumann boundary conditions (natural boundary conditions), namely

$$\nabla\psi(\mathbf{x}, t) \cdot \boldsymbol{\nu} = \mathbf{j}_n(\mathbf{x}, t) \cdot \boldsymbol{\nu} = \mathbf{j}_p(\mathbf{x}, t) \cdot \boldsymbol{\nu} = 0 \qquad \text{for all } \mathbf{x} \in \Gamma_A \text{ and } t \in [0, T]. \tag{50.39}$$

50.3.3 Recombination Processes

The net recombination rate R on the right-hand side of Equations 50.31 and 50.32 describes the radiative and nonradiative generation or recombination of carriers due to thermal excitation and various scattering effects. We assume that the recombination rate $R(n, p)$ is given by the sum of the most common processes, namely the Shockley–Read–Hall recombination R_{SRH}, the spontaneous radiative recombination R_{rad} and the Auger recombination R_{Auger}. All of these rates are of the form

$$R(n, p) = r(n, p)np\left(1 - \exp\left(\frac{q\varphi_n - q\varphi_p}{k_{\mathrm{B}}T}\right)\right), \tag{50.40}$$

where $r(n, p)$ is a model-dependent generation-recombination rate [2,13,24]. In Figure 50.7, one finds the definitions of these rates together with a schematic illustration of the corresponding processes. For Boltzmann statistics, Equation 50.40 is equivalent to the widely used $R(n, p) = r(n, p)(np - N_{\mathrm{intr}}^2)$, where N_{intr} is the intrinsic carrier density defined in (50.13).

50.3.4 Thermodynamic Equilibrium

The thermodynamic equilibrium is characterized by vanishing current fluxes $\mathbf{j}_n(\mathbf{x}, t) = \mathbf{j}_p(\mathbf{x}, t) = \mathbf{0}$. As a consequence, $R(n, p) = 0$ and the quasi-Fermi potentials assume constant values which due to (50.40) are equal:

$$\varphi_0 := \varphi_n = \varphi_p = \mathrm{const}.$$

Without loss of generality, we set $\varphi_0 = 0$. Therefore, the van Roosbroeck system of (50.30)–(50.32) reduces to the nonlinear Poisson equation

$$-\nabla \cdot \varepsilon_s \nabla \psi = q\left(N_v \mathcal{F}\left(\frac{E_v - q\psi}{k_{\mathrm{B}}T}\right) - N_c \mathcal{F}\left(\frac{q\psi - E_c}{k_{\mathrm{B}}T}\right) + C\right). \tag{50.41}$$

We supply it with Dirichlet boundary conditions (50.35) for zero voltages $U_\alpha = 0$ for $\alpha = 1, \ldots, N_O$ and homogeneus Neumann boundary conditions (50.39) elsewhere. The solution of (50.41) with these boundary conditions defines the *built-in potential* denoted with ψ_{eq}.

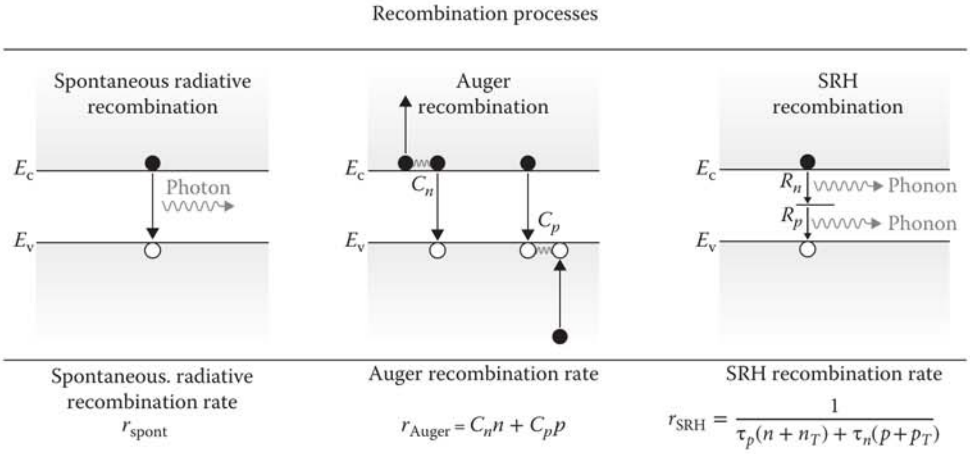


FIGURE 50.7 This figure illustrates three different recombination processes. The carrier life times τ_n, τ_p, the reference carrier densities n_T, p_T, and the coefficients r_{spont}, C_n, C_p are material-dependent parameters.

50.3.5 Non-Boltzmann Statistics and Generalized Einstein Relation

Now, we address the statistical distribution function $\mathcal{F}$ which appears in Equation 50.34. It describes the relationship between the density of the free carriers as well as the electrostatic potential and the quasi-Fermi potentials [2].

Assuming that the electrons in the conduction band are in quasi equilibrium, i.e., they are described by the quasi-Fermi level E_{F_n}, the electron density can be introduced as a convolution integral of the density of states DOS(E) with the Fermi–Dirac distribution function

$$n = \int_{-\infty}^{\infty} \mathrm{DOS}(E) \frac{1}{\exp\left(\frac{E-E_{F_n}}{k_B T}\right)+1} \mathrm{d}E. \tag{50.42}$$

It is possible to express this convolution as a product of the effective density of states N_c and a nondimensionalized statistical distribution function $\mathcal{F}$, i.e.,

$$n(\eta) = N_c \mathcal{F}(\eta). \tag{50.43}$$

We consider organic and inorganic semiconductors. For inorganic, three-dimensional (3D) bulk semiconductors with parabolic bands the statistical distribution function is given by the Fermi–Dirac integral of order 1/2 [2], which can be approximated by the Blakemore [25] or the Boltzmann distribution in the low density limit. For organic semiconductors, it is given by the Gauss–Fermi integral [10], which reduces to a Boltzmann type distribution function (up to some normalizing factor) in the low-density limit or a Blakemore distribution function for vanishing variance σ. The latter corresponds to a δ-shaped density of states [10,26], describing a single electronic transport level. See Figure 50.8 for definitions and a relationship between these functions and Figure 50.9 for a graphical representation.

For inorganic, 3D bulk semiconductors assuming parabolic bands with effective mass m_e^* the effective density of states is given by

$$N_c = 2\left(\frac{m_e^* k_B T}{2\pi\hbar^2}\right)^{3/2},$$

where T is the temperature and $\hbar$ the Planck constant. For organic semiconductors, the effective density of states N_c is given by the density of transport states N_t.

Introducing a general relationship between carrier density as well as quasi-Fermi and electrostatic potential via Equation 50.34 has quite a few implications. The nonlinear Poisson equation (50.41) becomes more complicated. In Section 50.2.4, we have studied how to design useful starting guesses for the Boltzmann approximation. For general distribution functions $\mathcal{F}$, it is no longer possible to obtain an explicit expression for ψ_0, like we have achieved in Equation 50.12. One rather needs to solve a nonlinear equation, for example, via Newton's methods at a given point $\mathbf{x}$. For highly doped regions, the Boltzmann approximation strongly overestimates the carrier densities if $\eta > 0$, see Figure 50.9. This implies that the more realistic Fermi–Dirac distribution leads to a higher built-in voltage. Moreover, from a numerical point of view, the Scharfetter–Gummel method needs to be adjusted. We discuss this in Section 50.4.2.

Finally, the distribution function $\mathcal{F}$ influences the ratio between diffusion and drift. To see this, we write the currents in *drift-diffusion* form. Using the relations for the carrier densities (50.34), we obtain

$$\mathbf{j}_n = -q\mu_n n \nabla\psi + qD_n \nabla n, \qquad \mathbf{j}_p = -q\mu_p p \nabla\psi - qD_p \nabla p.$$

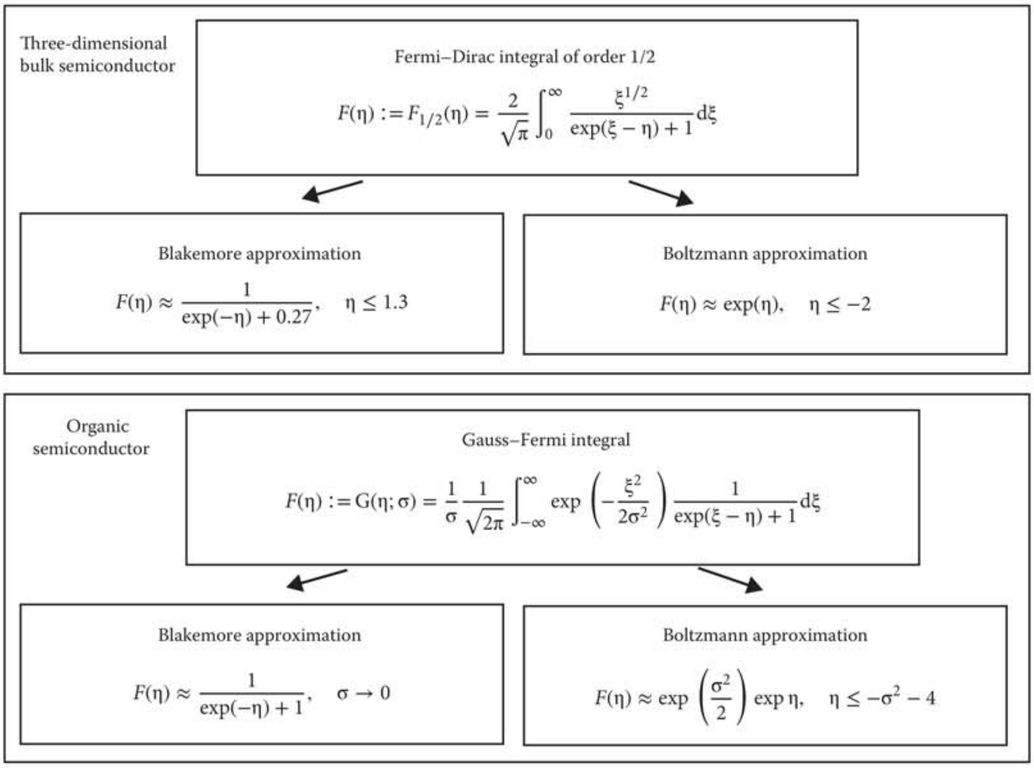
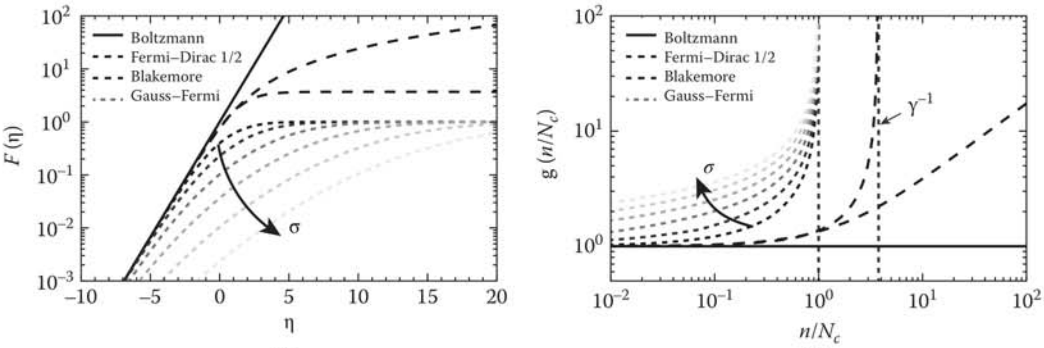


FIGURE 50.8 This figure shows two important choices of the statistical distribution function: for three-dimensional bulk semiconductors and for organic semiconductors [10] described by a Gaussian density of states with variance σ. Both can be approximated using the Blakemore or Boltzmann distribution function.



(a) (b)

FIGURE 50.9 (a) Some frequently considered distribution functions for the density of free carriers. For the Blakemore and the Gauss–Fermi functions one obtains a saturation of the carrier density at high values of η. (b) Illustration of the diffusion enhancement factors for the distribution functions shown in (a). In the case of the Blakemore and the Gauss–Fermi functions the saturation values of the carrier density directly lead to a divergent behavior of the nonlinear diffusion factor.

The diffusion coefficients D_n and D_p are connected to the carrier mobilities by a *generalized* Einstein relation

$$\frac{D_n}{\mu_n} = \frac{k_B T}{q} g\left(\frac{n}{N_c}\right), \qquad \frac{D_p}{\mu_p} = \frac{k_B T}{q} g\left(\frac{p}{N_v}\right) \tag{50.44}$$

with a density-dependent nonlinear factor

$$g(\xi) = \xi \left(\mathcal{F}^{-1}\right)'(\xi),$$

leading in general to a nonlinear diffusion coefficient. For the Boltzmann approximation, we immediately see that $g(\xi) = 1$, which gives the classical Einstein relation $D_n = k_B T \mu_n / q$. For non-exponential distribution function and $\xi \geq 1$, however, we note $g(\xi) \geq 1$, see [27]. For this reason, we call this factor *diffusion enhancement* as proposed in [9]. The dependence of the diffusion enhancement g on the density is depicted in Figure 50.9 for various distribution functions.

50.3.6 Free Energy and Dissipation Rate

We consider consistency with fundamental principles of thermodynamics to be a quality measure for models describing natural processes. So it is of significant interest to study the consistency for both continuous and discretized models. Due to our simplifying assumption of constant temperature, we cannot have energy conservation in the proper sense. The second law of thermodynamics for nonequilibrium processes [16,28] requires nonnegativity of the local entropy production which, multiplied by the temperature T, gives the *dissipation rate*.

For the sake of readability, we discuss these concepts for Boltzmann statistics [29] here. See [30] and [31] for more general statistics. For a triple (ψ, n, p) and a thermodynamic equilibrium solution $(\psi_{eq}, n_{eq}, p_{eq}) := (\psi_{eq}, n(\psi_{eq}), p(\psi_{eq}))$ of the full van Roosbroeck system the *free energy* is defined as

$$\begin{aligned} \mathbb{F}(\psi, n, p) = &\int_\Omega \left(n \log \frac{n}{n_{eq}} - n + p \log \frac{p}{p_{eq}} - p + n_{eq} + p_{eq} \right) d\mathbf{x} \\ &+ \int_\Omega \frac{\varepsilon_s}{2} \left| \nabla \left(\psi - \psi_{eq} \right) \right|^2 d\mathbf{x} + \frac{1}{2} \int_{\Gamma_G} \frac{\varepsilon_{ox}}{d_{ox}} \left(\psi - \psi_{eq} \right)^2 ds. \end{aligned} \tag{50.45}$$

For a transient solution $(\psi(t), n(t), p(t))$ of ((50.30)–(50.32)), the function $\mathbb{L}(t) = \mathbb{F}(\psi(t), n(t), p(t))$ decays exponentially as t tends to infinity and one has

$$\mathbb{L}(t) = \mathbb{L}(0) - \int_0^t \mathbb{D}(\tau) d\tau,$$

where $\mathbb{D}(t)$ is the nonnegative dissipation rate [29]

$$\mathbb{D}(t) = \int_\Omega \left(n\mu_n |\nabla \varphi_n|^2 + p\mu_p |\nabla \varphi_p|^2 + r(n,p)(np - 1)\log(np) \right) d\mathbf{x} \geq 0 \tag{50.46}$$

depending on t via the time evolution of n and p. This result confirms the consistency with the second law of thermodynamics. Incidentally, the function $\mathbb{L}$ is a Lyapunov function, allowing in certain situations to prove the global stability of the thermodynamical equilibrium. Furthermore, it can be used as a tool to prove uniqueness of solutions to the system ((50.30)–(50.32)). This leads us naturally to the following subsection.

50.3.7 Existence and Uniqueness Results

When modeling complex physical phenomena, it is often necessary to simplify the underlying physical model. Therefore, a sound mathematical investigation is necessary to assess the implications of such simplifications. Take, for example, the electrostatic potential of a device. Physically, we know that such a potential exists. Hence, a sound model should guarantee its existence—also from a mathematical point of view. If the existence cannot be shown, then the model does not represent the physical world accurately. Apart from existence, uniqueness of the solution is often desirable as well as continuous dependence on the initial data.

The mathematical technique used to prove existence and uniqueness depends on the device geometry and the model. The first existence result on the van Roosbroeck system was shown by Mock [32]. Since then several results have been obtained by Gajewski and Gröger, we refer the interested reader to [33,34]. In [34], the key tool to show the existence and uniqueness of the time-dependent system is based on finding a Lyapunov function. Moreover, Gajewski and Gröger presented the first result which considered Fermi–Dirac statistics instead of Boltzmann statistics. There are other important existence results. We would like to mention explicitly the results studied by Markowich see, for example, [35], Jüngel [36], and Jerome [37].

We would like to show the reader that these results are not of purely analytical interest but can also be used to design numerical methods. Gummel's method [38], for example (see Section 50.5), is based on the same fixed point iteration technique used to obtain the existence result which we discuss next. The connection to numerics can be found in [37]. We would like to give the reader an idea regarding the proof of a standard result without going into the tricky details. We follow ideas from Markowich's textbook [39].

We consider the stationary van Roosbroeck system with Boltzmann statistics. Using a scaling that can be found in [39, Section 2.4] and making a change of variables, we can rewrite the steady-state van Roosbroeck system in the following way:

$$-\lambda^2 \Delta\psi = \delta^2 e^{-\psi} v - \delta^2 e^{\psi} u + C, \tag{50.47}$$

$$\nabla \cdot (\mu_n e^{\psi} \nabla u) = \tilde{R}, \tag{50.48}$$

$$\nabla \cdot (\mu_p e^{-\psi} \nabla v) = \tilde{R}. \tag{50.49}$$

The densities n and p are related to the so-called *Slotboom variables* u and v via $n = \delta^2 e^{\psi} u$ and $p = \delta^2 e^{-\psi} v$. In Equations 50.47 through 50.49, the parameter λ denotes the normalized characteristic Debye length of the device, δ^2 indicates the scaled intrinsic carrier density and $\tilde{R}$ is the scaled recombination term. It is of the form $\tilde{R} = c(\psi, u, v)(uv - 1)$ with $c(\psi, u, v) > 0$. The chosen scaling and change of variables make it possible to rewrite the boundary conditions. We consider only Ohmic contacts and homogeneous Neumann boundaries.

In order to prove the existence of solutions for the system ((50.47)–(50.49)) equipped with suitable boundary conditions, we need to make several technical assumptions. We refer the interested reader to [39] where these assumptions are well explained as well as physically and mathematically justified. Here, we would rather like to focus on the key idea of the proof which is exploited numerically, omitting mathematical detail. The existence proof is based on an iteration scheme which considers the Poisson equation (50.47) decoupled from the continuity equations (50.48) and (50.49). The proof consists of the following steps:

1. We fix some $u_0, v_0 > 0$, and consider the semilinear elliptic problem

   $$-\lambda^2 \Delta\psi = \delta^2 e^{-\psi} v_0 - \delta^2 e^{\psi} u_0 + C, \text{ in } \Omega, \qquad \text{(plus BC for } \psi \text{ on } \partial\Omega). \tag{50.50}$$

   Using standard analytical results for semilinear elliptic equations (namely the Leray–Schauder fixed point theorem and the maximum principle [40]) we prove that there exists a unique solution of the problem (50.50) that we denote with ψ_1.

2. We insert the solution ψ_1 of Equation 50.50 into the decoupled linear elliptic equations

$$\nabla \cdot (\mu_n e^{\psi_1} \nabla u) = c(\psi_1, u_0, v_0)(uv_0 - 1), \text{ in } \Omega, \qquad \text{(plus BC for } u \text{ on } \partial\Omega), \tag{50.51}$$

$$\nabla \cdot (\mu_p e^{-\psi_1} \nabla v) = c(\psi_1, u_0, v_0)(u_0 v - 1), \text{ in } \Omega, \qquad \text{(plus BC for } v \text{ on } \partial\Omega). \tag{50.52}$$

   We point out that in the right-hand side of Equation 50.51 the rate c depends only on ψ_1 determined from the previous step and on the fixed u_0, v_0. The other term, $uv_0 - 1$, depends linearly on the unknown u. Analogous considerations are valid for Equation 50.52. Physically, this means that in the factor describing relaxation to equilibrium, $uv - 1$, we freeze the hole density, more precisely the Slotboom variable v_0, in the continuity equation for the electrons and vice versa. Thanks to the nonnegativity of $c(\psi_1, u, v)$ there exist unique solutions u_1 and v_1 to Equations 50.51 and 50.52. This is ensured by standard results for linear elliptic equations (see, e.g., [40]).
3. Based on the first two steps, we can now define a map H which maps (u_0, v_0) onto (u_1, v_1). This map is known as *Gummel map*. It is possible to prove that this map has a fixed point (u^*, v^*) which determines a (weak) solution (ψ^*, u^*, v^*) of the coupled system ((50.47)–(50.49)). The electrostatic potential ψ^* that solves Poisson's equation (50.47) can be determined using the first step by substituting (u_0, v_0) with (u^*, v^*).

Finally, we discuss the uniqueness of solutions to the van Roosbroeck system. It is well known that for the steady-state system, the uniqueness of the solution cannot be shown without additional assumptions, for example, on the applied voltage. This is no surprise, as in fact, some semiconductor devices (e.g., thyristors [41]) are designed to have multiple steady states. However, for a device in thermal equilibrium (see Section 50.3.4), we have $u = v = 1$. In this case, there is a unique function which satisfies the Poisson equation (50.47). Hence, the solution $(\psi_e, 1, 1)$ is the unique equilibrium solution of the system ((50.47)–(50.49)), see [32]. For sufficiently small bias voltages, the uniqueness of the solution is shown in [39] under some specific assumptions on the recombination rates. All recombination mechanisms appearing in Figure 50.3.3 fulfill these assumptions. However, impact ionization rates, for example, are excluded.

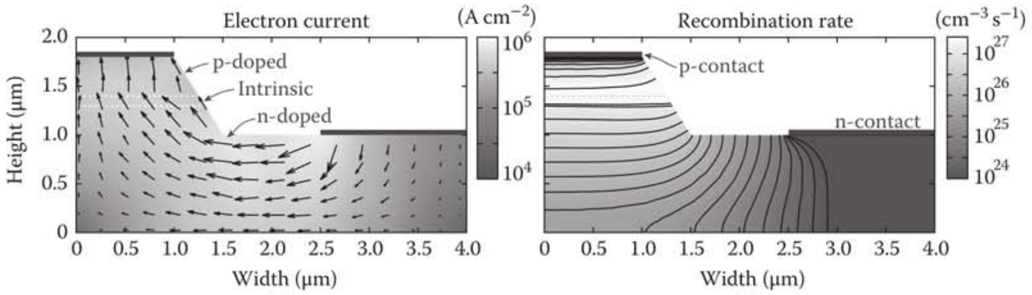
50.3.8 Maximum Principle

An important mathematical tool, often used for proofs similar to the ones in the previous section, is the maximum principle. Intuitively, the maximum principle states that if the domain is bounded and the right-hand side of the elliptic equation is positive, then the maximum of the solution is attained on the boundary of the domain [40].

50.4 Discretizing the van Roosbroeck System

In this section, we introduce a method for discretizing the van Roosbroeck system which is close to the physicist's approach to derive PDEs based on a subdivision of the computational domain into *representative elementary volumes* or *control volumes*. The *two-point flux* finite volume method described here can be seen as a straightforward generalization, preserving the properties of the 1D Scharfetter–Gummel scheme in higher dimensions. Figure 50.10 shows a 2D simulation. The 2D variant of this approach was introduced as *box method* in [12]. Historically, it goes back to [42]. The 3D variant of this method was probably first investigated in [29,43].

The method has two main ingredients: a geometry-based approach to obtain a system describing communicating control volumes and a consistent description of the fluxes between two adjacent control volumes. These are discussed in the following two subsections. We finish this section by describing various properties of the finite volume scheme.



(a) (b)

FIGURE 50.10 Exemplary results of a 2D simulation. The example is inspired by a ridge-waveguide laser with Ohmic contacts on the top of a mesa and at the side. For simplicity, we restrict to a homogeneous material which is here GaAs with the parameters as given in Figure 50.2. Within the mesa, an intrinsic domain is enclosed by a p-doped top-layer ($N_A = 2 \times 10^{18}\,\mathrm{cm}^{-3}$) and the n-doped substrate ($N_D = 2 \times 10^{18}\,\mathrm{cm}^{-3}$). The boundaries of the doped domains are shown in the pictures as dashed lines. (a) At an applied voltage of 1.5 V (flat band conditions) a significant electron current flow can be observed. The absolute value of the current density is shown by the gray scale, the arrows indicate the direction of particle flow (in the case of electrons the current density vector points in the opposite direction). (b) Plot of the total recombination rate at the same bias. The maximal recombination rate is observed in the vicinity of the intrinsic domain.

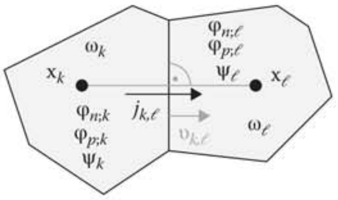


FIGURE 50.11 Two adjacent control volumes ω_k and ω_ℓ with corresponding data.

50.4.1 Finite Volume Method in Higher Dimensions

We partition the domain into N nonintersecting, convex polyhedral control volumes ω_k such that $\Omega = \bigcup_{k=1}^{N} \omega_k$. We associate with each control volume a node $\mathbf{x}_k \in \omega_k$. If the control volume intersects with the boundary of our domain, we demand that the node lies on the boundary, that means $\mathbf{x}_k \in \partial\Omega \cap \omega_k$. We assume that the partition is *admissible* [44], i.e., the *edge* $\overline{\mathbf{x}_k\mathbf{x}_\ell}$ with length $h_{k,\ell}$ is orthogonal to $\partial\omega_k \cap \partial\omega_\ell$. Thus, the normal vectors to $\partial\omega_k$ can be calculated by $\boldsymbol{\nu}_{k,\ell} = (\mathbf{x}_\ell - \mathbf{x}_k)/\|\mathbf{x}_\ell - \mathbf{x}_k\|$. See Figure 50.11 for details.

The set of all nodes and control volumes is called the *mesh* of the domain Ω. For a discussion of constructive ways to obtain such mesh partitions, see Section 50.6. In order to keep the equations simple, we introduce two abbreviations,

$$\eta_n\left(\psi, \varphi_n\right) = \frac{q\left(\psi - \varphi_n\right) - E_c}{k_B T} \quad \text{and} \quad \eta_p\left(\psi, \varphi_p\right) = \frac{E_v - q\left(\psi - \varphi_n\right)}{k_B T},$$

and assume that we have only a single gate contact with gate voltage U_G. We integrate the van Roosbroeck system ((50.30)–(50.32)) over ω_k and apply Gauss's divergence theorem, resulting in the integral equations

$$-\int_{\partial\omega_k} \varepsilon_s \nabla\psi \cdot \boldsymbol{\nu}\,\mathrm{d}s + \int_{\partial\omega_k \cap \Gamma_G} \frac{\varepsilon_{\mathrm{ox}}}{d_{\mathrm{ox}}}(\psi - U_G)\,\mathrm{d}s = q \int_{\omega_k} \left(C - N_c \mathcal{F}\left(\eta_n\left(\psi, \varphi_n\right)\right) + N_v \mathcal{F}\left(\eta_p\left(\psi, \varphi_p\right)\right)\right) \mathrm{d}\mathbf{x},$$

$$-q\frac{\partial}{\partial t}\int\limits_{\omega_k} N_c\mathcal{F}\left(\eta_n\left(\psi,\varphi_n\right)\right)\mathrm{d}\mathbf{x} + \int\limits_{\partial\omega_k} \mathbf{j}_n\cdot\boldsymbol{\nu}\mathrm{d}s = q\int\limits_{\omega_k} R d\mathbf{x},$$

$$q\frac{\partial}{\partial t}\int\limits_{\omega_k} N_v\mathcal{F}\left(\eta_p\left(\psi,\varphi_p\right)\right)\mathrm{d}\mathbf{x} + \int\limits_{\partial\omega_k} \mathbf{j}_p\cdot\boldsymbol{\nu}\mathrm{d}s = -q\int\limits_{\omega_k} R\mathrm{d}\mathbf{x},$$

for $k = 1, \dots, N$. Here, $\boldsymbol{\nu}$ is outward-pointing unit normal to the control volume ω_k; see Figure 50.11. These equations represent an integral form of the van Roosbroeck system on every control volume. In particular, the first equation is Gauss's law of electrodynamics. The other two equations are balance laws for the carrier densities. The densities in each control volume change only due to in- and outflow through the boundary or recombination.

Next, the surface integrals are split into the sum of integrals over the planar interfaces between the control volume ω_k and its neighbors. Employing one point quadrature rules for the surface and volume integrals, we deduce the finite volume scheme

$$\begin{aligned}\sum_{\omega_\ell\in\mathcal{N}(\omega_k)} |\partial\omega_k\cap\partial\omega_\ell| D_{k,\ell} &= q|\omega_k|\left(C_k - N_c\mathcal{F}\left(\eta_n\left(\psi_k,\varphi_{n;k}\right)\right) + N_v\mathcal{F}\left(\eta_p\left(\psi_k,\varphi_{p;k}\right)\right)\right)\\ &\quad + |\partial\omega_k\cap\Gamma_G|\frac{\varepsilon_{\mathrm{ox}}}{d_{\mathrm{ox}}}(U_G-\psi_k),\end{aligned} \tag{50.53}$$

$$-q|\omega_k|N_c\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}\left(\eta_n\left(\psi_k,\varphi_{n;k}\right)\right) + \sum_{\omega_\ell\in\mathcal{N}(\omega_k)} |\partial\omega_k\cap\partial\omega_\ell| j_{n;k,\ell} = q|\omega_k|R_k, \tag{50.54}$$

$$q|\omega_k|N_v\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}\left(\eta_p\left(\psi_k,\varphi_{p;k}\right)\right) + \sum_{\omega_\ell\in\mathcal{N}(\omega_k)} |\partial\omega_k\cap\partial\omega_\ell| j_{p;k,\ell} = -q|\omega_k|R_k. \tag{50.55}$$

In the above mentioned formulas, $\mathcal{N}\left(\omega_k\right)$ denotes the set of all control volumes neigboring ω_k. In 2D, the measure $|\partial\omega_k\cap\partial\omega_\ell|$ corresponds to the length of the boundary line segment and in 3D to the area of the intersection of the boundary surfaces. The measure $|\omega_k|$ is in 2D given by the area and in 3D by the volume of the control volume ω_k. The unknowns ψ_k, $\varphi_{n;k}$, and $\varphi_{p;k}$ are approximations of the electric potential as well as the quasi-Fermi potentials for electrons and holes evaluated at node $\mathbf{x}_k$ like already introduced for the 1D case in Section 50.2. Accordingly, R_k is defined as

$$R_k = R\left(N_c\mathcal{F}\left(\eta_n\left(\psi_k,\varphi_{n;k}\right)\right), N_v\mathcal{F}\left(\eta_p\left(\psi_k,\varphi_{p;k}\right)\right)\right).$$

The doping is defined by the integral average

$$C_k = \frac{1}{|\omega_k|}\int_{\omega_k} C(\mathbf{x})\mathrm{d}\mathbf{x},$$

which can be estimated by its nodal value $C(\mathbf{x}_k)$. The numerical fluxes $D_{k,\ell}$, $j_{n;k,\ell}$ and $j_{p;k,\ell}$ approximate $-\varepsilon\nabla\psi\cdot\boldsymbol{\nu}_{k\ell}$, $\mathbf{j}_n\cdot\boldsymbol{\nu}_{k\ell}$ and $\mathbf{j}_p\cdot\boldsymbol{\nu}_{k\ell}$, respectively, on the interfaces between two adjacent control volumes ω_k and ω_ℓ, see Figure 50.11. As in the 1D case, we assume that these fluxes can be expressed as nonlinear functions depending on the values $\psi_k, \varphi_{n;k}, \varphi_{p;k}$ and $\psi_\ell, \varphi_{n;\ell}, \varphi_{p;\ell}$.

The electric displacement flux is approximated by

$$D_{k,\ell} = -\varepsilon_s\frac{\psi_\ell-\psi_k}{h_{k,\ell}},$$

where

$$h_{k,\ell} = \|\mathbf{x}_{\ell} - \mathbf{x}_k\|$$

is the edge length, compare with Equation 50.15.

For the Boltzmann approximation, the numerical charge carrier fluxes can be approximated via the Scharfetter–Gummel expression (50.21). In theory, any classical technique such as central differences or upwinding could be used to discretize the numerical fluxes. However, the former suffers from instability issues (as seen in Section 50.2.7) and neither scheme is consistent with the thermodynamic equilibrium. We will explain what this means in Section 50.4.3. In the next section, we discuss thermodynamically consistent schemes for non-Boltzmann statistics.

The finite volume scheme ((50.53)–(50.55)) yields a nonlinear system of $3N$ equations depending on $3N$ variables. We can directly substitute Dirichlet boundary values in the equations. In practice, however, this way of handling Dirichlet values is very technical to implement. Exploiting floating point arithmetic, the Dirichlet penalty method [45] provides a reasonable alternative. Physically, it replaces the Dirichlet boundary conditions by gate boundary conditions with very high oxide permittivity.

50.4.2 Scharfetter–Gummel Fluxes and Their Non-Boltzmann Generalizations

In 1D, we have already discussed the highly effective flux discretization scheme proposed by Scharfetter and Gummel for Boltzmann statistics [11]. They derive the numerical flux by locally solving a linear two-point boundary value problem. The 1D idea immediately carries over to higher dimensions if we insert the Scharfetter–Gummel (50.21) into the discrete system ((50.53)–(50.55)). However, if the Boltzmann approximation is not valid anymore, we can no longer derive the flux like Scharfetter and Gummel suggested.

This motivated the work in [46] where the Scharfetter–Gummel idea was generalized to a large class of nonlinear convection-diffusion problems, allowing to define consistent numerical fluxes from nonlinear two-point boundary value problems. Unfortunately, these *generalized Scharfetter–Gummel schemes* cannot be expressed by closed formulas. Sometimes, however, the local fluxes can be obtained iteratively [27]. It is also possible to approximate the two-point boundary value problems by simpler ones (e.g., by freezing some coefficients). This leads to *modified Scharfetter–Gummel schemes*. We address some of these schemes now.

For the sake of readability, we provide the formulas for electrons only. The formulas for holes follow similarly.

50.4.2.1 Generalized Scharfetter–Gummel Schemes

By the construction of our mesh, it suffices to study the 1D flux j_n along the edge $\overline{\mathbf{x}_k\mathbf{x}_\ell}$. For general distribution functions $\mathcal{F}$, we want to solve the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}x} j_n = \frac{\mathrm{d}}{\mathrm{d}x}\Big(- q\mu_n N_c \mathcal{F}\left(\eta_n\left(\psi, \varphi_n\right)\right) \nabla\varphi_n \Big) = 0$$

on the interval $[0, h_{k,\ell}]$ with boundary conditions

$$\varphi_n\left(0\right) = \varphi_{n;k} \quad \text{and} \quad \varphi_n\left(h_{k,\ell}\right) = \varphi_{n;\ell},$$

where $\varphi_{n;k}$ and $\varphi_{n,\ell}$ are the values of the quasi-Fermi potentials at nodes $\mathbf{x}_k$ and $\mathbf{x}_\ell$, respectively, see Figure 50.11. We assume now that the flux is constant between both nodes and denote it with $j_{n;k,\ell}$. Integrating twice leads to an integral equation [27] for the unknown current, namely

$$\int_{\eta_{n;k}}^{\eta_{n;\ell}} \left(\frac{j_{n;k,l}/j_0}{\mathcal{F}(\eta)} + \frac{\psi_\ell - \psi_k}{U_\mathrm{T}} \right)^{-1} d\eta = 1, \tag{50.56}$$

where $j_0 = q\mu_n N_c \frac{U_\mathrm{T}}{h_{k,\ell}}$ and the limits are given by $\eta_{n;k} = \eta_n\left(\psi_k, \varphi_{n;k}\right)$ and $\eta_{n;\ell} = \eta_n\left(\psi_\ell, \varphi_{n;\ell}\right)$. For strictly monotonously increasing $\mathcal{F}(\eta)$ this equation has always a unique solution [47].

For the Boltzmann approximation $\mathcal{F}(\eta) = \exp(\eta)$, this integral equation can be solved analytically and yields the classical Scharfetter–Gummel expression for the flux (50.21). In [27], it was shown that for the Blakemore distribution function $\mathcal{F}(\eta) = \frac{1}{\exp(-\eta)+\gamma}$, the integral equation yields a fixed point equation

$$\hat{j}_{n;k,\ell} = B\left(\gamma\hat{j}_{n;k,\ell} + \frac{\psi_\ell - \psi_k}{U_\mathrm{T}}\right) e^{\eta_{n;\ell}} - B\left(-\left[\gamma\hat{j}_{n;k,\ell} + \frac{\psi_\ell - \psi_k}{U_\mathrm{T}}\right]\right) e^{\eta_{n;k}} \tag{50.57}$$

for the nondimensionalized edge current $\hat{j}_{n;k,\ell} = j_{n;k,l}/j_0$. The right-hand side is a Scharfetter–Gummel expression where the argument of the Bernoulli function is shifted by $\gamma\hat{j}_{n;k,\ell}$. Hence, this reduces to the classical Scharfetter–Gummel scheme for $\gamma = 0$.

Since the Bernoulli function is strictly decreasing, this fixed point equation possesses a unique solution $\hat{j}_{n;k,\ell}$. If we want to use the flux given by (50.57) in the discrete system ((50.53)–(50.55)) , we need to solve for the fluxes twice (once for electrons and once for holes) on each discretization edge $\overline{\mathbf{x}_k\mathbf{x}_\ell}$. A few Newton steps are sufficient to solve this equation iteratively.

Even though (50.57) is restricted to the Blakemore approximation, it provides a useful scheme in the context of organic semiconductors. There it arises naturally as a model for materials with δ-shaped density of states [10,26], describing a single electronic transport level, see Figures 50.8 and 50.9. This has been described in Section 50.3.5.

Unfortunately, for a general statistical distribution function, no corresponding equation has been derived so far. Therefore, in [47] it was proposed to use piecewise approximations of $\mathcal{F}$ by Blakemore type or rational approximations of Padé type in order to obtain piecewise integrable expression from the local boundary value problem.

50.4.2.2 Modified Scharfetter–Gummel Schemes

Bessemoulin-Chatard [48] derived a finite volume scheme for convection-diffusion problems by averaging the nonlinear diffusion term appropriately. This idea was generalized to more general distribution functions in [49], introducing a logarithmic average of the nonlinear diffusion enhancement

$$g_{n;k,\ell} = \frac{\eta_{n;k} - \eta_{n;\ell}}{\log\mathcal{F}\left(\eta_{n;k}\right) - \log\mathcal{F}\left(\eta_{n;\ell}\right)}$$

along the discretization edge. Using the generalized Einstein relation (50.44), one immediately observes that the diffusion enhancement g can be seen as a modification factor of the thermal voltage U_T. Replacing U_T in the Scharfetter–Gummel expression (50.21) by $U_\mathrm{T}^* = U_\mathrm{T} g_{n;k,\ell}$, we deduce the following modified Scharfetter–Gummel scheme

$$j_{n;k,\ell} = -\frac{q\mu_n U_\mathrm{T}}{h_{k,\ell}} g_{n;k,\ell} \left(N_c\mathcal{F}\left(\eta_{n;k}\right) B\left(-\frac{\psi_\ell - \psi_k}{U_\mathrm{T} g_{n;k,\ell}}\right) - N_c\mathcal{F}\left(\eta_{n;\ell}\right) B\left(\frac{\psi_\ell - \psi_k}{U_\mathrm{T} g_{n;k,\ell}}\right)\right), \tag{50.58}$$

approximating the current along the edge.

Previously, we have replaced the thermal voltage by a suitable average along the edge. Now, we want to approximate $\mathcal{F}(\eta)$ along the edge by an exponential (Boltzmann approximation) and adjusting the effective density of states N_c accordingly, i.e.,

$$N_c\mathcal{F}(\eta) \approx N^*_{c;k,\ell}\exp(\eta).$$

This choice makes it possible to use the original Scharfetter–Gummel flux (50.21) by replacing N_c with $N^*_{c;k,\ell}$. One choice for the modified density of states is

$$N^*_{c;k,\ell}(\eta^*) = N_c\frac{\mathcal{F}(\eta^*)}{\exp(\eta^*)},$$

where $\eta^* \in [\eta_k, \eta_\ell]$, assuming $\eta_k \leq \eta_\ell$. In practice, we might consider taking the geometric average between $N_c(\eta_k)$ and $N_c(\eta_\ell)$, which leads to another modified Scharfetter–Gummel scheme:

$$j_{n;k,\ell} = -\frac{q\mu_n U_\mathrm{T}}{h_{k,\ell}}N_c\sqrt{\frac{\mathcal{F}\left(\eta_{n;k}\right)\mathcal{F}\left(\eta_{n;\ell}\right)}{\exp(\eta_{n;k})\exp(\eta_{n;\ell})}}\left(B\left(-\frac{\psi_\ell-\psi_k}{U_\mathrm{T}}\right)e^{\eta_{n;k}} - B\left(\frac{\psi_\ell-\psi_k}{U_\mathrm{T}}\right)e^{\eta_{n;\ell}}\right). \tag{50.59}$$

The idea behind this scheme was introduced in [50] for the numerical solution of the generalized Nernst–Planck system which is similar to the van Roosbroeck system ((50.30)–(50.33)). A variant of this scheme for Fermi–Dirac statistics is described in [14,51] and numerically implemented in the semiconductor device simulation package WIAS-TeSCA [52].

50.4.3 Flux Expressions Consistent with the Thermodynamic Equilibrium

As introduced in Section 50.3.4, the currents of holes and electrons vanish in the thermodynamic equilibrium when no bias is applied. If a numerical scheme for zero bias boundary conditions results into vanishing numerical fluxes, we call it *consistent with the thermodynamical equilibrium*. In practice, one can examine this consistency by checking whether $j_{n;k,\ell} = j_{p;k,\ell} = 0$ if $\varphi_{n;k} = \varphi_{n;\ell}$ and $\varphi_{p;k} = \varphi_{p;\ell}$. Violating this property causes unphysical dissipation (spurious Joule heating) in the steady state attained for zero bias boundary conditions, which by definition is supposed to be the thermodynamic equilibrium [48].

All schemes introduced in Section 50.4.2 are consistent with the thermodynamical equilibrium. Indeed, let us assume that the quasi-Fermi potentials φ_n and φ_p between two adjacent control volumes ω_k and ω_ℓ are equal. The consistency of the classical Scharfetter–Gummel (50.21) and the scheme using a modified density of states (50.59) are obvious since the Bernoulli function satisfies $B(-x) = \exp(x)B(x)$.

Due to the definition from the solution of the two-point boundary value problem, the generalized Scharfetter–Gummel scheme defined in Equation 50.56 and its specialization for the Blakemore approximation (50.57) are consistent with the thermodynamic equilibrium. And finally, the logarithmic average of the diffusion enhancement is the only possible average which guarantees consistency with thermodynamics in the scheme (50.58). For details see [27,47,49].

50.4.4 Free Energy and Dissipation Rate

For Boltzmann statistics, it is rather straightforward to define discrete analoga of the free energy (50.45) and the dissipation rate (50.46). The positivity of the discrete dissipation rate was shown in [29], and the exponential decay of the free energy to its equilibrium value was proven in [53]. An overview of the entropy method for finite volume schemes has been given in [54]. First results on more general statistics functions in this respect have been obtained in [55].

These pioneering works strongly indicate that the chosen discretization approach results in discrete models, which are consistent with the structural assumptions of nonequilibrium thermodynamics. A full account of these issues in the context of general statistics functions remains an open research topic.

50.4.5 Existence, Uniqueness, and Convergence

There are very few existence proofs for the solutions of the discretized system ((50.53)–(50.55)). For the Boltzmann approximation, Gärtner [56] proved that the discretized steady-state system has a solution, which becomes unique if a small bias is applied. A similar result for Fermi statistics has been obtained in [47].

A convergence theory for the finite volume scheme for the full discrete system ((50.53)–(50.55)) and general flux functions is still missing. However, practical experience and a number of results make its convergence plausible.

For example, in one space dimension, second-order convergence in the discrete maximum norm for the Scharfetter–Gummel scheme has been shown in [19]. Under the assumption that second derivatives of the continuous solution exist, in [57] for moderately sized drift terms and two-dimensional (2D), square grids, first-order convergence for the simple upwind scheme (see e.g., [21]), and second-order convergence for the exponential fitting scheme in the L_2-norm has been shown. Reinterpretations of the finite volume Scharfetter–Gummel scheme as a nonstandard finite element method allowed to obtain convergence estimates for Scharfetter–Gummel schemes on Delaunay grids (see Section 50.6) [58,59]. For a general approach to the convergence theory of finite volume schemes, see [44]. In [46], weak convergence (no order estimate) for a generalization of the Scharfetter–Gummel scheme to nonlinear convection-diffusion problems has been shown. A convergence proof for a variant of the van Roosbroeck system discretized with the simple upwind scheme was given in [60].

50.4.6 Maximum Principle

When applied to the drift-diffusion formulation, compared to various variants of the stabilized finite element method, the two-point flux finite volume scheme is outstanding in the sense that it guarantees positivity of densities and absense of unphysical oscillations [61,62]. It allows to carry over the discussions of Section 50.2.7 to the higher dimensional case.

50.5 Nonlinear Solvers

In Section 50.2, we have already briefly mentioned that we need to solve a nonlinear discrete system of the form

$$\mathbf{0} = \mathbf{F}(\boldsymbol{\psi}, \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_p) := \begin{pmatrix} \mathbf{F}_1(\boldsymbol{\psi}, \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_p) \\ \mathbf{F}_2(\boldsymbol{\psi}, \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_p) \\ \mathbf{F}_3(\boldsymbol{\psi}, \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_p) \end{pmatrix}.$$

Here, we discuss two ways to do this: Newton's method and *Gummel's iteration method* [38], which is based on a blockwise decoupling of the discrete system.

50.5.1 Newton's Method

Assuming that a good starting guess has been provided, for example, by following the ideas in Section 50.2.6, Newton's method constructs a new $(k+1)$th iterate from the kth one by solving the linear system

$$\begin{pmatrix} \frac{\partial \mathbf{F}_1}{\partial \boldsymbol{\psi}} & \frac{\partial \mathbf{F}_1}{\partial \boldsymbol{\varphi}_n} & \frac{\partial \mathbf{F}_1}{\partial \boldsymbol{\varphi}_p} \\ \frac{\partial \mathbf{F}_2}{\partial \boldsymbol{\psi}} & \frac{\partial \mathbf{F}_2}{\partial \boldsymbol{\varphi}_n} & \frac{\partial \mathbf{F}_2}{\partial \boldsymbol{\varphi}_p} \\ \frac{\partial \mathbf{F}_3}{\partial \boldsymbol{\psi}} & \frac{\partial \mathbf{F}_3}{\partial \boldsymbol{\varphi}_n} & \frac{\partial \mathbf{F}_3}{\partial \boldsymbol{\varphi}_p} \end{pmatrix}^k \begin{pmatrix} \boldsymbol{\delta\psi}^k \\ \boldsymbol{\delta\varphi}_n^k \\ \boldsymbol{\delta\varphi}_p^k \end{pmatrix} = - \begin{pmatrix} \mathbf{F}_1(\boldsymbol{\psi}^k, \boldsymbol{\varphi}_n^k, \boldsymbol{\varphi}_p^k) \\ \mathbf{F}_2(\boldsymbol{\psi}^k, \boldsymbol{\varphi}_n^k, \boldsymbol{\varphi}_p^k) \\ \mathbf{F}_3(\boldsymbol{\psi}^k, \boldsymbol{\varphi}_n^k, \boldsymbol{\varphi}_p^k) \end{pmatrix} \tag{50.60}$$

for the update vector

$$\begin{pmatrix} \boldsymbol{\delta\psi}^k \\ \boldsymbol{\delta\varphi}_n^k \\ \boldsymbol{\delta\varphi}_p^k \end{pmatrix} := \begin{pmatrix} \boldsymbol{\psi}^{k+1} - \boldsymbol{\psi}^k \\ \boldsymbol{\varphi}_n^{k+1} - \boldsymbol{\varphi}_n^k \\ \boldsymbol{\varphi}_p^{k+1} - \boldsymbol{\varphi}_p^k \end{pmatrix}.$$

Then the new iterate is obtained by adding the previous vector to the update. The advantage of Newton's method is that it converges quadratically if the starting guess is sufficiently close to the solution [17]. This allows to obtain highly accurate discrete solutions at low additional cost. The major drawback is that it might converge very slowly or even fail to converge if the starting guess is too far from the actual solution. Damping—multiplying the update with a factor less than 1—is known to increase the convergence region. Another remedy is parameter embedding where one slowly changes a parameter, always using the old solution as a new starting guess. We have already employed this embedding technique in Section 50.2.6 when gradually increasing the bias voltage in the solution procedure.

An important advantage of finite element and finite volume discretizations is the fact that they create only next-neighbor couplings in the discretized systems. The resulting linearized systems are therefore sparse, i.e., the maximum number of nonzero elements in a matrix row is bounded by a constant independent of the number of discretization cells, making it possible to use highly economic storage schemes. For 2D applications, one can usually solve the resulting sparse linear system (50.60) via sparse direct solvers, such as PARDISO [63–66] or UMFPACK [67,68]. Direct solvers calculate a representation of the system matrix as a product of easily solvable lower and upper triangular matrices, which unfortunately are not anymore sparse, increasing memory consumption and computational time especially in large 3D applications. This fill-in phenomenon is avoided by preconditioned Krylov subspace methods [69], which compared to the direct solvers need significant effort in tuning and adaptation to the problem at hand.

50.5.2 Gummel's Method

Gummel [13,38] suggested decoupling the three equations in the van Roosbroeck system. He devised an iterative method at the continuous level, which for Boltzmann statistics and drift-diffusion form leads to alternating between solving linear differential equations for the electric potential as well as the charge carrier densities. Suppose, one already knows the iterate (ψ^k, n^k, p^k). In order to obtain the electric potential at the new level, ψ^{k+1}, one solves

$$0 = -\frac{1}{q}\nabla \cdot \left(\varepsilon_s \nabla \psi^{k+1}\right) + \left(n^k + p^k\right)\left(\frac{\psi^{k+1} - \psi^k}{U_\mathrm{T}}\right) + n^k - p^k - C.$$

This formulation is motivated by linearizing a fixed point problem that one obtains from the Poisson equation [13]. Once one has solved for the electrostatic potential at the new level, one can successively solve the continuity equations from

$$
\begin{aligned}
0 &= \nabla \cdot \mathbf{j}_n(\psi^{k+1}, n^{k+1}) - qR(\psi^{k+1}, n^{k+1}, p^k) \\
0 &= \nabla \cdot \mathbf{j}_p(\psi^{k+1}, p^{k+1}) + qR(\psi^{k+1}, n^{k+1}, p^{k+1}).
\end{aligned}
$$

Gummel's method is known to have a larger convergence region and a slower asymptotic convergence rate compared to Newtons's method. Depending on the software environment chosen, it also may be easier to implement. Note that standard existence and uniqueness proofs for the van Roosbroeck system rely on a similar decoupling strategy, see Section 50.3.7.

50.6 Mesh Generation for the Finite Volume Method

While appearing to be just a technical footnote compared to the questions of modeling, analysis and discretization, mesh generation in reality is a hard problem which deserves special attention. We focus on methods to construct admissible subdivisions of the computational domain Ω as used in Section 50.4.

50.6.1 Boundary Conforming Delaunay Meshes on Polyhedral Domains

Assume a partition (triangulation in 2D, tetrahedralization in 3D) of the polyhedral domain $\Omega = \bigcup_{k=1}^{N_\Sigma} \Sigma_k$ into nonoverlapping simplices Σ_k as it is commonly used for finite element methods [70]. We require that this simplicial partition of Ω has the *boundary conforming Delaunay property* [71]. For a triangulation of a 2D domain this property is equivalent to:

(1) For any two triangles with a common edge, the sum of their respective angles opposite to that edge is less or equal to 180°.
(2) For any triangle sharing an edge with $\partial\Omega$, its angle opposite to that edge is less or equal to 90°.

For a given vertex $x_i \in X$, where X is the set of vertices, the *Voronoi cell* around X is the set

$$
V_i = \left\{ \mathbf{x} \in \mathbb{R}^d : \|\mathbf{x} - \mathbf{x}_i\| < \|\mathbf{x} - \mathbf{x}_j\| \text{ for all } \mathbf{x}_j \in X \text{ with } \mathbf{x}_j \neq \mathbf{x}_i \right\}.
$$

The *Voronoi diagram*—the set of Voronoi cells for all vertices in X—is dual to the *Delaunay triangulation* of the point set X in the sense that for each edge $\overline{\mathbf{x}_i\mathbf{x}_j}$ in the Delaunay triangulation, $\partial V_i \cap \partial V_j \neq \emptyset$. If the simplicial partition is boundary conforming Delaunay, the restricted Voronoi cells $\omega_i = \Omega \cap V_i$ are well defined and can be obtained by joining the circumcenters of the simplices adjacent to the vertex $\mathbf{x}_i$. These restricted Voronoi cells provide an admissible control volume partition as required in Section 50.4. We note that in order to implement the finite volume method as described in Section 50.4, there is no need for an explicit construction of the control volumina ω_i as geometrical objects. Given the simplical partition, it is sufficient to base the calculations on the simplicial contributions $s_{ij}^k = |\partial\omega_i \cap \partial\omega_j \cap \Sigma_k|$ and $|\omega_i \cap \Sigma_K|$, and to use a simplex-based assembly loop as common for finite elements [29]. Figure 50.12 shows (boundary conforming) Delaunay triangulations.

There are several efficient algorithms to construct Delaunay triangulations for a given point set X [73,74]. These are good starting points of devising meshing algorithms in general. Many different problems complicate the generation of a mesh. For example, the boundary conforming Delaunay property is rather difficult to achieve, in particular it requires the careful insertion of additional points on the boundary. In 3D, slivers (very flat tetrahedra) must be avoided. And finally, it may be very complicated to fulfill additional requirements like constraints on the minimum angle or the local element size. Though there are still unsolved problems, the triangle ([72], 2D, free for noncommercial use) and TetGen ([75], 3D, open source) mesh generators help to create boundary conforming Delaunay meshes based on algorithms which are proven to

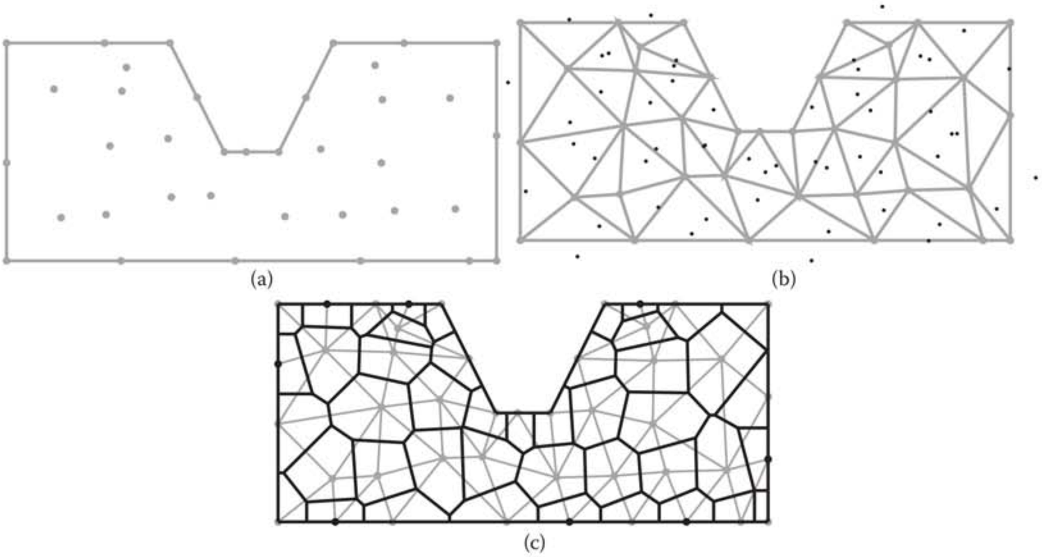


FIGURE 50.12 (a) Piecewise linear description of computational domain with given point cloud. (b) Delaunay triangulation of domain and triangle circumcenters. As some boundary triangles have angles larger than 90° angles opposite to the boundary, their circumcenters (small dots) are outside of the domain. (c) Boundary conforming Delaunay triangulation with automatically inserted additional points at the boundary (large dots), and Voronoi and restricted Voronoi cells (black). Created using triangle [72].

deliver meshes with the desired properties in finite time for a broad class of geometries given by a piecewise linear description of the boundary.

Other mesh generation approaches, in particular the advancing front [76] and the octree method [77], are similarly widespread. However, while popular in finite element community, their design makes it much harder to use them as a starting point for generating boundary conforming Delaunay meshes.

50.6.2 Tensor Product Approaches

In simplified geometrical situations, rectangular and cuboid mesh structures are a straightforward way to generate admissible finite volume partitions of the computational domain [12]. Extrusion of a 2D boundary conforming Delaunay base mesh into a 3D mesh consisting of prisms (and optional subsequent subdivision of these prisms into tetrahedra) provides another method to create admissible finite volume partitions [78].

50.6.3 Open Problems

Charge densities and potentials in semiconductors exhibit interior and boundary layers, i.e., thin regions charaterized by rapid gradients in one direction and slow gradients in the orthogonal direction. This may happen in space charge regions in the vicinity of gate contacts and at p-n junctions. These layers and other possible "hotspots" due to recombination, for example, call for local mesh adaptation. Element size control via the mesh generators using *a priori* knowledge is often used in this situation [72,75]. The development of reliable a posteriori mesh refinement criteria is an open problem in the context of the strongly coupled van Roosbroeck system. A physically motivated heuristic approach based on the equidistribution of the dissipation rate was suggested in [29] and investigated in [79].

Mesh generation tools mostly fail to create boundary conforming Delaunay meshes which use thin, anisotropic simplices to resolve interior and boundary layers with an optimal number of points. A partial remedy in some situations can be provided by node offsetting [80], a technique still under development. In certain cases, tensor product approaches [78] give reasonable results.

50.7 Time Discretization

In order to discretize the time derivative in the van Roosbroeck system ((50.30)–(50.32)), we define a series of discrete time values $0 = t^0, t^1, \ldots, t^M = T$ with step lengths $\tau^m = t^{m+1} - t^m$. We describe the temporal semi-discretization of the (spatially) continuous problem. For practical purposes it needs to be combined with the space discretization approach described in Section 50.4. The implict Euler time discretization scheme assumes given values $\psi^m, \varphi_n^m, \varphi_p^m, n^m, p^m$ of the variables at time t^m. Values at the subsequent time t^{m+1} are calculated by solving the nonlinear system of equations

$$-\nabla \cdot (\varepsilon_s \nabla \psi^{m+1}) = q(C - n^{m+1} + p^{m+1}), \tag{50.61}$$

$$-q\frac{n^{m+1} - n^m}{\tau^m} - q\nabla \cdot \mu_n n^{m+1} \nabla \varphi_n^{m+1} = qR(n^{m+1}, p^{m+1}), \tag{50.62}$$

$$q\frac{p^{m+1} - p^m}{\tau^m} - q\nabla \cdot \mu_p p^{m+1} \nabla \varphi_p^{m+1} = -qR(n^{m+1}, p^{m+1}), \tag{50.63}$$

where

$$\begin{aligned} n^{m+1} &= N_c \mathcal{F}(\eta_n(\psi^{m+1}, \varphi_n^{m+1})), \\ p^{m+1} &= N_v \mathcal{F}(\eta_p(\psi^{m+1}, \varphi_p^{m+1})). \end{aligned}$$

As a consequence, the nonlinear system ((50.61)–(50.63)), which is structurally similar to Equations 50.53 through 50.55, has to be solved for each time step by iterative methods described in Section 50.5. An initial guess for these methods can be obtained from the old time step t^m or a linear extrapolation involving data from several previous time steps [81].

It is advisable to choose the size of the time steps τ^m according to multiple criteria, including the convergence radius of Newton's method, the free energy, and local changes of the potentials.

At first glance, the computational cost of having to solve a nonlinear system for each time step appears to be rather high. However, this scheme is stable independent of the time step size and allows to carry over all advantages of the finite volume scheme from the stationary to the transient case. The transient discrete solution converges to the stationary one, in particular, for the corresponding boundary condition to the thermodynamic equilibrium solution. The free energy of the system decays during this approach to the equilibrium. In several cases, it has been proven that this decay is exponential [55,82]. By preserving the physics, the implicit Euler scheme fits well to the thermodynamically motivated discretization schemes. The use of other schemes implies giving up some of these physical properties. If the resulting deviations from thermodynamic properties can be kept under control, higher order schemes (e.g., backward differentiation formula (BDF) methods [83]) might help to reduce computational costs due to a higher temporal convergence order.

Due to the perceived complexity and high computational costs of solving coupled nonlinear PDE systems, linear implicit methods, mixing linearization, and time discretizations are alternatives. Most prominent is the scheme devised in [84].

50.8 Contact Terminal Currents

In simulations, one is usually interested in IV-curves, i.e., the dependency of terminal currents on applied voltages. Therefore, calculating terminal currents accurately is crucial to a successful postprocessing of the simulated field data. The *total current density* is given by

$$\mathbf{j} = \mathbf{j}_d + \mathbf{j}_n + \mathbf{j}_p,$$

where $\mathbf{j}_d = -\varepsilon_s \nabla(\partial_t \psi) = \partial_t \mathbf{D}$ is the *displacement current density*. Taking the time derivative of the Poisson equation (50.30) yields

$$\begin{aligned} -\nabla \cdot \varepsilon_s \nabla(\partial_t \psi) = q(\partial_t n - \partial_t p) &= -\nabla \cdot \mathbf{j}_n - qR - \nabla \cdot \mathbf{j}_p + qR \\ \nabla \cdot \partial_t \mathbf{D} &= -\nabla \cdot \mathbf{j}_n - \nabla \cdot \mathbf{j}_p, \end{aligned}$$

resulting in $\nabla \cdot \mathbf{j} = 0$, which physically implies charge conservation.

Given a set of contacts $\Gamma_\alpha \subset \partial\Omega$ for $\alpha = 1, \dots, N_C$, the terminal current through contact α is defined as

$$I_\alpha = \int_{\Gamma_\alpha} \mathbf{j} \cdot \boldsymbol{\nu} \, \mathrm{d}s.$$

The correct calculation of this integral presupposes a regularity of the solution and its derivative which is not supported by analytical theory. Therefore, in [51,52] a test function–based technique was proposed, which uses the weak formulation of the problem. For a physical motivation of this approach and an interpretation as a generalized Shockley–Ramo theorem we refer to [85]. For any α, let the function $T_\alpha : \Omega \to \mathbb{R}$ solve the boundary value problem

$$-\nabla^2 T_\alpha = 0 \qquad \text{in } \Omega, \tag{50.64}$$

$$\nabla T_\alpha \cdot \boldsymbol{\nu} = 0 \qquad \text{on } \partial\Omega \setminus \bigcup_{\beta=1}^{N_C} \Gamma_\beta, \tag{50.65}$$

$$T_\alpha = \delta_{\alpha\beta} \qquad \text{on } \Gamma_\beta \text{ (for } \beta = 1, \dots, N_C), \tag{50.66}$$

where $\delta_{\alpha\beta}$ denotes the Kronecker delta. As a consequence,

$$\begin{aligned} I_\alpha = \int_{\Gamma_\alpha} \mathbf{j} \cdot \boldsymbol{\nu} \, \mathrm{d}s = \int_{\partial\Omega} T_\alpha \mathbf{j} \cdot \boldsymbol{\nu} \, \mathrm{d}s &= \int_\Omega \nabla \cdot (T_\alpha \mathbf{j}) \, \mathrm{d}\mathbf{x} \\ &= \int_\Omega \nabla T_\alpha \cdot \mathbf{j} \, \mathrm{d}\mathbf{x} + \int_\Omega T_\alpha \nabla \cdot \mathbf{j} \, \mathrm{d}\mathbf{x} \\ &= \int_\Omega \nabla T_\alpha \cdot (\mathbf{j}_n + \mathbf{j}_p) \, \mathrm{d}\mathbf{x} + \partial_t \int_\Omega \nabla T_\alpha \cdot \mathbf{D} \, \mathrm{d}\mathbf{x}. \end{aligned}$$

Given the semi-discretization in time from Section 50.7, one arrives at

$$I_\alpha^{m+1} = \int_\Omega \nabla T_\alpha \cdot (\mathbf{j}_n^{m+1} + \mathbf{j}_p^{m+1}) \, \mathrm{d}\mathbf{x} + \frac{1}{\tau^m} \int_\Omega \nabla T_\alpha \cdot \left(\mathbf{D}^{m+1} - \mathbf{D}^m\right) \, \mathrm{d}\mathbf{x}.$$

Like in Section 50.4, we approximate the system (50.64)–(50.66) by a finite volume discretization. Using partial integration and the notation introduced in Section 50.4 we come up with an expression for the total current through the α-th contact of the form

$$I_\alpha^{m+1} \approx \sum_{\substack{\omega_\ell \in \mathcal{N}(\omega_k) \\ k<\ell}} |\partial\omega_k \cap \partial\omega_\ell| \left(j_{n;k,\ell}^{m+1} + j_{p;k,\ell}^{m+1} + \frac{D_{k,\ell}^{m+1} - D_{k,\ell}^{m}}{\tau^m} \right) (T_{\alpha;k} - T_{\alpha;\ell}). \tag{50.67}$$

50.9 An Alternative: The Finite Element Method

We have put special emphasis on the finite volume method. However, it is worth pointing out that there are other approaches. The finite difference method that has been used to solve the van Roosbroeck system [13] on tensor product meshes is equivalent to the previously introduced finite volume method.

The most popular ansatz to deal with unstructured meshes is the finite element method [70]. It starts with a *weak formulation* of the steady-state van Roosbroeck problem ((50.30)–(50.32)). Each equation of the system is multiplied with a so-called test function and integrated in space. Integrating by parts, each equation of the steady-state system can be restated as the following problem:

$$\text{find } u \in U \text{ such that } a(u, v) = f(v) \text{ for all } v \in V,$$

where U is a function space containing the ansatz functions u and V is a function space containing the test functions v. The map $a: U \times V \to \mathbb{R}$ has the structure

$$a(u, v) = \int_\Omega \alpha(u)\nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x} - \int_{\partial\Omega} \nabla u \cdot \mathbf{n}\, v\, ds \tag{50.68}$$

and the functional $f: V \to \mathbb{R}$ is continuous. The idea of a finite element method is to approximate this continuous problem with a discrete version. We approximate the infinite dimensional spaces U and V with finite dimensional subspaces U_h and V_h spanned by basis functions with localized support. Then we obtain a finite dimensional version of Equation 50.68, namely

$$\text{find } u_h \in U_h \text{ such that } a_h(u_h, v_h) = f(v_h) \text{ for all } v_h \in V_h.$$

As it is sufficient to fulfill this condition for finitely many localized basis functions in V_h, we obtain a sparse (containing many zeros) nonlinear system of equations for the basis coefficients of u_h, assuming we are able to calculate the integrals. One usually approximates these integrals by quadrature rules.

Auf der Maur [86] suggested to write the stationary van Roosbroeck system in this framework, using the quasi-Fermi potential formulation of the flux. In this case one has

$$a^{\psi}(\psi, v) = \int_\Omega \varepsilon_s \nabla\psi \cdot \nabla v \,\mathrm{d}\mathbf{x} + \int_{\Gamma_G} \frac{\varepsilon_{\mathrm{ox}}}{d_{\mathrm{ox}}}(\psi - U_{\mathrm{G}}) v \,\mathrm{d}s, \qquad f^{\psi}(v) = q\int_\Omega (p - n + C) v \,\mathrm{d}\mathbf{x}, \tag{50.69}$$

$$a^{\varphi_n}(\varphi_n, v) = \int_\Omega \mu_n n \nabla\varphi_n \cdot \nabla v \,\mathrm{d}\mathbf{x}, \qquad f^{\varphi_n}(v) = \int_\Omega R v \,\mathrm{d}\mathbf{x}, \tag{50.70}$$

$$a^{\varphi_p}(\varphi_p, v) = \int_\Omega \mu_p p \nabla\varphi_p \cdot \nabla v \,\mathrm{d}\mathbf{x}, \qquad f^{\varphi_p}(v) = -\int_\Omega R v \,\mathrm{d}\mathbf{x}. \tag{50.71}$$

At the Dirichlet parts of the boundary, namely at the Ohmic contacts $\Gamma_{\mathrm{O},\alpha}$ introduced in Section 50.3, we assume that test functions v vanish and the ansatz functions denoted by ψ, φ_n, and φ_p fulfill the Dirchlet boundary conditions ((50.35)–(50.37)). Note, that the gate boundary conditions (50.38) have been incorparated into the boundary part of the quadratic form (50.69).

We point out that for Poisson's equation $\alpha^{\psi}(\psi) = \varepsilon_s$ and for the continuity equations we have $\alpha^{\varphi_n}(\varphi_n) = \mu_n n$ and $\alpha^{\varphi_p}(\varphi_p) = \mu_p p$, respectively. Due to the formulation of the problem with quasi-Fermi potentials as unknowns, this means that $\alpha^{\varphi_n}(\varphi_n)$ and $\alpha^{\varphi_p}(\varphi_p)$ introduce nonlinearities in a^{φ_n} and a^{φ_p}. The integrals on the left-hand side of Equations 50.69 through 50.71 have no analytical expressions and are approximated using quadrature rules [70].

Essential for this finite element scheme is the underlying potential-based formulation of the flux in terms of φ_n and φ_p. If one attempts to apply the finite element method using the density-based, drift-diffusion formulation of the flux, one encounters a lot of technical difficulties and one has to stabilize the finite element method [61,87]. A convergence proof of the finite element method for a similar system in a somewhat different context (and without recombination) is available in [88].

50.10 Extensions and Outlook

In this chapter, we discussed the most important numerical solution techniques of the van Roosbroeck system, describing the transport of electrons and holes in semiconductors in a self-consistent electric field. This approach provides the core functionalities for a solver, which can be enhanced by models describing additional physical phenomena and more complex device structures including, for example, heterostructures. We mention a number of additional extensions and point to the corresponding literature.

50.10.1 Additional Physical Models

Many devices require a more accurate description of the carrier mobilities. For example, models of ionized impurity scattering, high-field drift velocity saturation and similar effects [13,89] introduce dependencies of the mobilities on carrier densities and the electric field strength, resulting in even more nonlinear couplings in the system.

The carrier transport in organic semiconductors is governed by hopping transport between the energy states of neighboring molecules. The van Roosbroeck system can be used to describe this effectively [8,9]. This may require the use of statistical distribution functions that reflect the distribution of the energy transport levels such as Blakemore for δ-shaped densities of states or the Gauss–Fermi integral [10]. In this context, the correct treatmeant of the diffusion enhancement is vital. One also has to account for the nonlinear mobilities [8,90] related to the energetic disorder characteristic for organic semiconductors.

Spin-polarized drift-diffusion models have been proposed in [91] for the description of spintronic devices. They generalize the van Roosbroeck system, introducing spin-resolved densities for electron and holes and additional mechanisms describing the spin relaxation. They have been recently studied also from a numerical point of view in [92–94].

50.10.2 Coupling the van Roosbroeck System to Other Models

When modeling lasers one has to couple the van Roosbroeck system to equations for the optical modes. This introduces additional recombination processes, describing stimulated emission. Additionally, one has to consider balance equations for the photon number. A comprehensive model (also including heating effects) along with a numerical solution strategy is given in [14].

In semiconductor devices with embedded nanostructures (such as quantum dots and quantum wells), one has to couple the drift-diffusion equations to equations describing the dynamics of the carriers localized in the nanostructures [4,5,7]. A multiscale approach for coupling atomistic with continuum drift-diffusion models is presented in [95].

If heating effects become important, the van Roosbroeck system needs to be extended by an energy transport model for the heat flow in the device. For a thermodynamically consistent extension of the van Roosbroeck system to account for this effect we refer to [14,31].

50.10.3 Methods for Doping Optimization

An important task when designing semiconductor devices is finding a suitable doping profile. This is an analytical and numerical challenge. Electric properties of the device can be improved by optimizing the doping profiles using suitable objective functionals [96–98]. Recently, the approach introduced in [96] has been extented for the optimization of doping profiles in lasers [99].

50.10.4 Alternative Modeling Approaches for Carrier Transport

The van Roosbroeck system uses implicitly the assumption that the carrier ensemble is locally in quasi equilibrium. When this assumption is not met, for example, for hot electrons, one has to consider alternative approaches. The most common ones are hydrodynamic models [100] and approximations of the Boltzmann transport equation which are derived from a spherical harmonics expansion [101]. The book [102] gives a mathematically oriented overview of these topics.

Acknowledgments

The authors are grateful for innumerable discussions with K. Gärtner and H. Gajewski, who helped to shape their still incomplete knowledge on the topic. This work has been supported by ERC-2010-AdG no. 267802 "Analysis of Multiscale Systems Driven by Functionals" (N.R.), by the Deutsche Forschungsgemeinschaft DFG within CRC 787 "Semiconductor Nanophotonics" (T.K., N.R., M.K) and partially funded in the framework of the project "Macroscopic Modeling of Transport and Reaction Processes in Magnesium-Air-Batteries" (grant 03EK3027D) under the research initiative "Energy storage" of the German Federal government.

References

1. W. van Roosbroeck. Theory of the flow of electrons and holes in germanium and other semiconductors. *Bell Syst. Tech. J.*, 29(4):560–607, 1950.
2. S. M. Sze and K. K. Ng. *Physics of Semiconductor Devices*, 3rd edition. Wiley, 2006. ISBN: 978-0-471-14323-9.
3. J. Piprek. *Optoelectronic Devices: Advanced Simulation and Analysis*. Berlin: Springer, 2005.
4. M. Grupen and K. Hess. Simulation of carrier transport and nonlinearities in quantum-well laser diodes. *IEEE J. Quant. Electron.*, 34(1):120–140, 1998.
5. S. Steiger, R. G. Veprek, and B. Witzigmann. Unified simulation of transport and luminescence in optoelectronic nanostructure. *J. Comput. Electron.*, 7(4):509–520, 2008.
6. T. Koprucki, H.-C. Kaiser, and J. Fuhrmann. Electronic states in semiconductor nanostructures and upscaling to semi-classical models. In A. Mielke, editor, *Analysis, Modeling and Simulation of Multiscale Problems*, pp. 365–394. Heidelberg, Berlin: Springer, 2006.
7. T. Koprucki, A. Wilms, A. Knorr, and U. Bandelow. Modeling of quantum dot lasers with microscopic treatment of Coulomb effects. *Opt. Quant. Electron.*, 42:777–783, 2011.
8. R. Coehoorn, W. F. Pasveer, P. A. Bobbert, and M. A. J. Michels. Charge-carrier concentration dependence of the hopping mobility in organic materials with Gaussian disorder. *Phys. Rev. B.*, 72(15):155206, 2005.
9. S. L. M. van Mensfoort and R. Coehoorn. Effect of Gaussian disorder on the voltage dependence of the current density in sandwich-type devices based on organic semiconductors. *Phys. Rev. B.*, 78(8):085207, 2008.
10. G. Paasch and S. Scheinert. Charge carrier density of organics with Gaussian density of states: Analytical approximation for the Gauss-Fermi integral. *J. Appl. Phys.*, 107(10):104501, 2010.

11. D. L. Scharfetter and H. K. Gummel. Large-signal analysis of a silicon Read diode oscillator. *IEEE Trans. Electron. Dev.*, 16(1):64–77, 1969.
12. W. Fichtner, D. J. Rose, and R. E. Bank. Semiconductor device simulation. *SIAM J. Sci. Stat. Comput.*, 4(3):391–415, 1983.
13. S. Selberherr. *Analysis and Simulation of Semiconductor Devices*. Berlin: Springer, 1984.
14. U. Bandelow, H. Gajewski, and R. Hünlich. Fabry–Perot lasers: Thermodynamics-based modeling. In J. Piprek, editor, *Optoelectronic Devices*. Berlin: Springer, 2005.
15. A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift diffusion systems. *Nonlinearity*, 24(4):1329, 2011.
16. S. R. de Groot and P. Mazur. *Non-Equilibrium Thermodynamics*. North Holland: Dover, 1962.
17. P. Deuflhard *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, volume 35. Berlin: Springer, 2011.
18. D. N. Allen and R. V. Southwell. Relaxation methods applied to determine the motion, in two dimensions, of a viscous fluid past a fixed cylinder. *Quart. J. Mech. Appl. Math.*, 8:129–145, 1955.
19. A. M. Il'in. A difference scheme for a differential equation with a small parameter multiplying the second derivative. *Mat. Zametki.*, 6:237–248, 1969.
20. J. S. Chang and G. Cooper. A practical difference scheme for Fokker-Planck equations. *J. Comput. Phys.*, 6(1):1–16, 1970.
21. H.-G. Roos, M. Stynes, and L. Tobiska. *Robust Numerical Methods for Singularly Perturbed Differential Equations*, volume 24. Berlin: Springer, 2008.
22. K. W. Morton. *Numerical Solution of Convection-Diffusion Problems: Applied Mathematics*. London: Taylor & Francis, 1996.
23. D. Schröder. *Modelling of Interface Carrier Transport for Device Simulation*. Wien: Springer, 1994.
24. S. Y. Karpov. *Visible Light-Emitting Diodes*, Chapter 14. In J. Piprek, editor, *Nitride Semiconductor Devices: Principles and Simulation*, pp. 303–325. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2007.
25. J. S. Blakemore. The parameters of partially degenerate semiconductors. *Proc. Phys. Soc. London A.*, 65:460–461, 1952.
26. M. Gruber, E. Zojer, F. Schürrer, and K. Zojer. Impact of materials versus geometric parameters on the contact resistance in organic thin-film transistors. *Adv. Funct. Mater.*, 23(23):2941–2952, 2013.
27. T. Koprucki and K. Gärtner. Discretization scheme for drift-diffusion equations with strong diffusion enhancement. *Opt. Quant. Electron.*, 45(7):791–796, 2013.
28. R. Haase. *Thermodynamics of Irreversible Processes*. Berlin: Addison-Wesley, 1968.
29. H. Gajewski and K. Gärtner. On the discretization of van Roosbroeck's equations with magnetic field. *Z. Angew. Math. Mech.*, 76(5):247–264, 1996.
30. H. Gajewski and K. Gröger. Semiconductor equations for variable mobilities based on Boltzmann statistics or Fermi-Dirac statistics. *Math. Nachr.*, 140(1):7–36, 1989.
31. G. Albinus, H. Gajewski, and R. Hünlich. Thermodynamic design of energy models of semiconductor devices. *Nonlinearity*, 15(2):367, 2002.
32. M. S. Mock. On equations describing steady-state carrier distributions in a semiconductor device. *Comm. Pure Appl. Math.*, 25(25):781–792, 1972.
33. H. Gajewski and K. Gröger. On the basic equations for carrier transport in semiconductors. *J. Math. Anal. Appl.*, 113:12–35, 1986.
34. H. Gajewski. On existence, uniqueness and asymptotic behavior of solutions of the basic equations for carrier transport in semiconductors. *Z. Angew. Math. Mech.*, 65(2):101–108, 1985.
35. P. A. Markowich. A nonlinear eigenvalue problem modelling the avalanche effect in semiconductor diodes. *SIAM J. Math. Anal.*, 6:1268–1283, 1985.
36. A. Jüngel. On the existence and uniqueness of transient solutions of a degenerate nonlinear drift-diffusion model for semiconductors. *Math. Models Methods Appl. Sci.*, 04:677, 1994.

37. J. W. Jerome. *Analysis of Charge Transport. A Mathematical Study of Semiconductor Devices*. Berlin Heidelberg: Springer, 1996.
38. H. K. Gummel. A self-consistent iterative scheme for one-dimensional steady state transistor calculations. *IEEE Trans. Electron Dev.*, 11(10):455–465, 1964.
39. P. A. Markowich. *The Stationary Semiconductor Device Equations*. Berlin: Springer-Verlag, 1986.
40. D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Berlin: Springer, 2001. Reprint of the 1998 edition.
41. I. Rubinstein. *Electro-Diffusion of Ions*, volume 11. Philadelphia, PA: SIAM, 1990.
42. R. H. Macneal. An asymmetrical finite difference network. *Quart. Math. Appl.*, 11:295–310, 1953.
43. P. Fleischmann and S. Selberherr. Three-dimensional Delaunay mesh generation using a modified advancing front approach. In *Proceedings of IMR97*, pp. 267–278, 1997.
44. R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In *Handbook of Numerical Analysis*, P.G. Ciarlet and J.L. Lions (Eds.), 7:713–1018, 2000.
45. J. W. Barrett and C. M. Elliott. Finite element approximation of the Dirichlet problem using the boundary penalty method. *Numer. Math.*, 49(4):343–366, 1986.
46. R. Eymard, J. Fuhrmann, and K. Gärtner. A finite volume scheme for nonlinear parabolic equations derived from one-dimensional local Dirichlet problems. *Numer. Math.*, 102(3):463–495, 2006.
47. K. Gärtner. Existence of bounded discrete steady state solutions of the van Roosbroeck system with monotone Fermi-Dirac statistic functions. *J. Comput. Electron.*, 14(3):773–787, 2015.
48. M. Bessemoulin-Chatard. A finite volume scheme for convection–diffusion equations with nonlinear diffusion derived from the Scharfetter–Gummel scheme. *Numer. Math.*, 121(4):637–670, 2012.
49. T. Koprucki, N. Rotundo, P. Farrell, D. H. Doan, and J. Fuhrmann. On thermodynamic consistency of a Scharfetter–Gummel scheme based on a modified thermal voltage for drift-diffusion equations with diffusion enhancement. *Opt. Quant. Electron.*, 47(6):1327–1332, 2015.
50. J. Fuhrmann. Comparison and numerical treatment of generalised Nernst-Planck models. *Comput. Phys. Commun.*, 196:166–178, 2015.
51. H. Gajewski. *Analysis und Numerik von Ladungstransport in Halbleitern*. WIAS Report No. 6. Berlin: WIAS, 1993. ISSN 0942-9077.
52. H. Gajewski, M. Liero, R. Nürnberg, and H. Stephan. *WIAS-TeSCA User Manual 1.2*. WIAS Technical Report No. 14. Berlin: WIAS, 2016. ISSN: WIAS 1618-7776.
53. A. Glitzky and K. Gärtner. Energy estimates for continuous and discretized electro-reaction–diffusion systems. *Nonlinear Anal.*, 70(2):788–805, 2009.
54. C. Chainais-Hillairet. Entropy method and asymptotic behaviours of finite volume schemes. In J. Fuhrmann, M. Ohlberger, and C. Rohde, editors, *Finite Volumes for Complex Applications VII – Methods and Theoretical Aspects*, pp. 17–35. Berlin: Springer, 2014.
55. M. Bessemoulin-Chatard and C. Chainais-Hillairet. Exponential decay of a finite volume scheme to the thermal equilibrium for drift–diffusion systems. *J. Numer. Math.*, 2016. (Online first, doi: 10.1515/jnma-2016-0007)
    M. Bessemoulin-Chatard and C. Chainais-Hillairet. Exponential decay of a finite volume scheme to the thermal equilibrium for drift–diffusion systems. *arXiv preprintarXiv:1601.00813*, 2016.
56. K. Gärtner. Existence of bounded discrete steady-state solutions of the van Roosbroeck system on boundary conforming delaunay grids. *SIAM J. Sci. Comput.*, 31(2):1347–1362, 2009.
57. R. D. Lazarov, I. D. Mishev, and P. S. Vassilevski. Finite volume methods for convection diffusion problems. *SIAM J. Numer. Anal.*, 33(1):31–55, 1996.
58. J. J. H. Miller and S. Wang. An analysis of the Scharfetter-Gummel box method for the stationary semiconductor device equations. *RAIRO-Math. Num.*, 28(2):123–140, 1994.
59. J. Xu and L. Zikatanov. A monotone finite element scheme for convection-diffusion equations. *Math. Comput. Am. Math. Soc.*, 68(228):1429–1446, 1999.

60. C. Chainais-Hillairet, J.-G. Liu, and Y.-J. Peng. Finite volume scheme for multidimensional drift-diffusion equations and convergence analysis. *ESAIM: Math. Model. Numer. Anal. Math. Num.*, 37(2):319–338, 2003.
61. R. E. Bank, W. M. Coughran Jr, and L. C. Cowsar. The finite volume Scharfetter-Gummel method for steady convection diffusion equations. *Comput. Vis. Sci.*, 1(3):123–136, 1998.
62. M. Augustin, A. Caiazzo, A. Fiebach, J. Fuhrmann, V. John, A. Linke, and R. Umla. An assessment of discretizations for convection-dominated convection-diffusion equations. *Comp. Meth. Appl. Mech. Eng.*, 200:3395–3409, 2011.
63. A. Kuzmin, M. Luisier, and O. Schenk. Fast methods for computing selected elements of the greens function in massively parallel nanoelectronic device simulations. In F. Wolf, B. Mohr, and D. Mey, editors, *Euro-Par 2013 Parallel Processing*, volume 8097, *Lecture Notes in Computer Science*, pp. 533–544. Berlin: Springer, 2013.
64. O. Schenk, M. Bollhöfer, and R. A. Römer. On large-scale diagonalization techniques for the Anderson model of localization. *SIAM Rev.*, 50(1):91–112, 2008.
65. O. Schenk, A. Wächter, and M. Hagemann. Matching-based preprocessing algorithms to the solution of saddle-point problems in large-scale nonconvex interior-point optimization. *Comput. Optim. Appl.*, 36(2–3):321–341, 2007.
66. O. Schenk. PARDISO version 5.0.0. http://www.pardiso-project.org. Accessed on February 22, 2016.
67. T. A. Davis. *Direct Methods for Sparse Linear Systems (Fundamentals of Algorithms 2)*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2006.
68. T. Davis. SuiteSparse version 4.5.1. http://faculty.cse.tamu.edu/davis/suitesparse.html. Accessed on February 22, 2016.
69. Y. Saad. *Iterative Methods for Sparse Linear Systems*. Philadelphia, PA: SIAM, 2003.
70. P. Ciarlet. *The Finite Element Method for Elliptic Problems*. Philadelphia, PA: SIAM, 2002.
71. H. Si, K. Gärtner, and J. Fuhrmann. Boundary conforming Delaunay mesh generation. *Comput. Math. Math. Phys.*, 50(1):38–53, 2010.
72. J. Shewchuk. Triangle: A two-dimensional quality mesh generator and Delaunay triangulator. http://www.cs.cmu.edu/~quake/triangle.html. Accessed on December 1, 2015.
73. H. Edelsbrunner. *Algorithms in Combinatorial Geometry*. Heidelberg: Springer-Verlag, 1987.
74. S.-W. Cheng, T. K. Dey, and J. R. Shewchuk. *Delaunay Mesh Generation*. Boca Raton, FL: Chapman & Hall/CRC, 1st edition, 2012.
75. H. Si. TetGen version 1.5. http://tetgen.org/. Accessed on 2015-12-01.
76. J. Schöberl. Netgen an advancing front 2d/3d-mesh generator based on abstract rules. *Comput. Vis. Sci.*, 1(1):41–52, 1997.
77. M. S. Shephard and M. K. Georges. Automatic three-dimensional mesh generation by the finite octree technique. *Int. J. Numer. Meth. Eng.*, 32(4):709–749, 1991.
78. K. Gärtner and R. Richter. DEPFET sensor design using an experimental 3d device simulator. *Nucl. Instrum. Meth. A.*, 568(1):12–17, 2006.
79. B. Schmithüsen, K. Gärtner, and W. Fichtner. Grid adaptation for device simulation according to the dissipation rate. In *Simulation of Semiconductor Processes and Devices 1998*, pp. 197–200. Berlin: Springer, 1998.
80. J. Krause, N. Strecker, and W. Fichtner. Boundary-sensitive mesh generation using an offsetting technique. *Int. J. Numer. Meth. Eng.*, 49(1–2):51–59, 2000.
81. A. Fiebach, A. Glitzky, and A. Linke. Uniform global bounds for solutions of an implicit voronoi finite volume method for reaction–diffusion problems. *Numer. Math.*, 128(1):31–72, 2014.
82. A. Glitzky. Exponential decay of the free energy for discretized electro-reaction–diffusion systems. *Nonlinearity*, 21(9):1989, 2008.
83. E. Hairer, S. P. Norsett, and G. Wanner. *Solving Ordinary Differential Equations: Nonstiff Problems. v. 2: Stiff and Differential-Algebraic Problems*. Berlin: Springer Verlag, 2010.

84. M. S. Mock. A time-dependent numerical model of the insulated-gate field-effect transistor. *Solid-State Electron.*, 24(10):959–966, 1981.
85. P. D. Yoder, K. Gärtner, U. Krumbein, and W. Fichtner. Optimized terminal current calculation for Monte Carlo device simulation. *IEEE Trans. Comput-Aided Design Integr. Circuits Syst.*, 16(10):1082–1087, 1997.
86. M. Auf Der Maur. *A Multiscale Simulation Environment for Electronic and Optoelectronic Devices*. PhD thesis, University of Tor Vergata, 2008.
87. V. John and P. Knobloch. On spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part I–A review. *Comput. Methods Appl. Mech. Eng.*, 196(17):2197–2215, 2007.
88. A. Prohl and M. Schmuck. Convergent discretizations for the Nernst-Planck-Poisson system. *Numer. Math.*, 111(4):591–630, 2009.
89. V. Palankovski and R. Quay. *Analysis and Simulation of Heterostructure Devices*. Computational Microelectronics. Vienna: Springer Science & Business Media, 2004.
90. H. Bässler. Charge transport in disordered organic photoconductors a Monte Carlo simulation study. *Phys. Stat. Sol. (B).*, 175(1):15–56, 1993.
91. J. Fabian, S. C. Erwin, and I. Zutic. Bipolar spintronics: Fundamentals and applications. *IBM J. Res. Dev.*, 50(1):121–139, 2006.
92. A. Glitzky and K. Gärtner. Existence of bounded steady state solutions to spin-polarized drift-diffusion systems. *SIAM J. Math. Anal.*, 41(6):2489–2513, 2010.
93. A. Jüngel, C. Negulescu, and P. Shpartko. Bounded weak solutions to a matrix drift diffusion model for spin-coherent electron transport in semiconductors. *Math. Models Methods Appl. Sci.*, 25(5):929–958, 2015.
94. C. Chainais-Hillairet, A. Jüngel, and P. Shpartko. A finite-volume scheme for a spinorial matrix drift-diffusion model for semiconductors. *Numer. Methods for Partial Differ. Equ.*, 32(3):819–846, 2016.
95. M. Auf der Maur, A. Pecchia, G. Penazzi, F. Sacconi, and A. Carlo. Coupling atomistic and continuous media models for electronic device simulation. *J. Comput. Electron.*, 12(4):553–562, 2013.
96. M. Hinze and R. Pinnau. An optimal control approach to semiconductor design. *Math. Models Methods Appl. Sci.*, 12(1):89–107, 2002.
97. M. Hinze and R. Pinnau. Second-order approach to optimal semiconductor design. *J. Optim. Theory. Appl.*, 133(2):179–199, 2007.
98. M. Burger and R. Pinnau. Fast optimal design of semiconductor devices. *SIAM J. Appl. Math.*, 64(1):108–126, 2003.
99. D. Peschka, N. Rotundo, and M. Thomas. Towards doping optimization of semiconductor lasers. To appear in. *J. Comput. Theor. Trans.*, 45(5):410–423, 2016.
100. T. Grasser, T.-W. Tang, H. Kosina, and S. Selberherr. A review of hydrodynamic and energy-transport models for semiconductor device simulation. *Proc. IEEE*, 91(2):251–274, 2003.
101. S.-M. Hong and C. Jungemann. A fully coupled scheme for a Boltzmann-Poisson equation solver based on a spherical harmonics expansion. *J. Comput. Electron.*, 8(3):225–241, 2009.
102. A. Jüngel. *Transport Equations for Semiconductors*, volume 773. Berlin: Springer, 2009.

51

Monte Carlo Device Simulations

Katerina Raleva
Abdul R. Shaik
Raghuraj Hathwar
Akash Laturia
Suleman S. Qazi
Robin Daugherty
Dragica Vasileska
and
Stephen M. Goodnick

51.1 Introduction 773
Transistor Scaling • More Moore Technology • Importance of Modeling and Simulation
51.2 Bulk Monte Carlo Method Description 779
51.3 Particle-Based Device Simulation 791
Calculation of the Current • Ohmic Contacts • Time-Step • PM Coupling
51.4 Representative Simulation Results 796
Bulk Monte Carlo Simulations of Different Materials • Modeling Thermal Effects in SOI Devices Using Particle-Based Device Simulator
51.5 Outlook 803

51.1 Introduction

The semiconductor device–based electronics industry has been the largest industry in the world with global sales of over a trillion dollars since 1998. The revolution in the semiconductor industry, a large portion of the electronics industry, began in 1947 with the fabrication of bipolar devices on slabs of polycrystalline germanium (Ge). Single-crystalline materials were later proposed and introduced that made possible the fabrication of grown junction transistors. Migration to silicon (Si)–based devices was initially hindered by the stability of the Si/SiO$_2$ materials system, necessitating a new generation of crystal pullers with improved environmental controls to prevent SiO$_2$ formation. Later, the stability and low interface-state density of the Si/SiO$_2$ materials system provided passivation of junctions and eventually, the migration from bipolar devices to field-effect devices in 1960. By 1968, both complementary metal–oxide–semiconductor (CMOS) devices and poly-Si gate technology that allowed self-alignment of the gate to the source/drain of the device had been developed. These innovations permitted a significant reduction in power dissipation and a reduction of the overlap capacitance, improving frequency performance and resulting in the essential components of the modern CMOS device.

Since the invention of the bipolar transistor in 1947, the number and variety of semiconductor devices have increased tremendously as advanced technology, new materials, and broadened comprehension have been applied to the creation and innovation of new devices. Dr. William Shockley suggested in the fifties the use of semiconductors of different bandgap for the fabrication of heterostructure devices. Professor Herbert Kroemer's contributions to heterostructures—from heterostructures bipolar transistors to lasers—culminated in a Nobel Prize in Physics in 2000 and have paved the way for novel heterostructure devices including those in silicon. The unique properties of semiconductor materials have enabled the development of a wide variety of ingenious devices that have literally changed our world. To date, there are about

TABLE 51.1 Major Semiconductor Device Discoveries

- 1874: Metal-semiconductor contact
- 1907: Light-emitting diode
- 1947: Bipolar junction transistor (BJT)
- 1954: Solar cell
- 1957: Heterojunction bipolar transistor (HBT)
- 1958: Tunnel diode
- 1959: Integrated circuits
- 1960: Field-effect transistors (FETs)
- 1962: Semiconductor lasers
- 1966: Metal-semiconductor FET (MESFET)
- 1967: Nonvolatile semiconductor memory
- 1974: Resonant tunneling diode (RTD)
- 1980: Modulation (MOD) FET
- 1994: Room-temperature single-electron memory cell (SEMC)
- 2001: 15 nm MOSFET
- 2006: 3 nm FinFET
- 2012: Single atom transistor in Si technology

60 major devices, with over 100 device variations related to them. A list of most of the basic semiconductor devices discovered and used over the past century with the date of their introduction is shown in Table 51.1.

The metal–oxide–semiconductor field-effect transistor (MOSFET) and related integrated circuits (ICs) now constitute about 90% of the semiconductor device market. Combining Si with the elegance of the field-effect transistor (FET) structure has allowed simultaneously making devices smaller, faster, and cheaper—the mantra that has driven the modern semiconductor microelectronics industry. Indeed, the single factor driving continuous device improvement has been the semiconductor industry's relentless effort to reduce the cost per function on a chip. The way this is done is to put more devices on a chip while either reducing manufacturing costs or holding them constant. This leads to three methods of reducing the cost per function. The first is *transistor scaling*, which involves reducing the transistor size in accordance with some goal, i.e., keeping the electric field constant from one generation to the next. With smaller transistors, more can fit into a given area than in previous generations. The second method is *circuit cleverness*, which is associated with the physical layout of the transistors with respect to each other. If the transistors can be packed into a tighter space, then more devices can fit into a given area than before. The third method is to make the die larger. More devices can be fabricated on a *larger die*. All the while, the semiconductor industry is constantly looking for technological breakthroughs to decrease the manufacturing cost. All of this effort serves to reduce the cost per function on a chip.

51.1.1 Transistor Scaling

Device engineers are most concerned with the method of scaling introduced in Section 51.1. The semiconductor industry has been so successful in providing continued system performance improvement year after year that the Semiconductor Industry Association (SIA) has been publishing roadmaps for semiconductor technology since 1992. These roadmaps represent a consensus outlook of industry trends, taking history as a guide. SIA roadmaps [1] incorporate participation from the global semiconductor industry, including the United States, Europe, Japan, Korea, and Taiwan. They basically affirm the desire of the industry to continue with *Moore's law* [2], which is often stated as the doubling of transistor performance and quadrupling of the number of devices on a chip every 3 years. The phenomenal progress signified by Moore's law has been achieved through scaling of the MOSFET from larger to smaller physical dimensions [3].

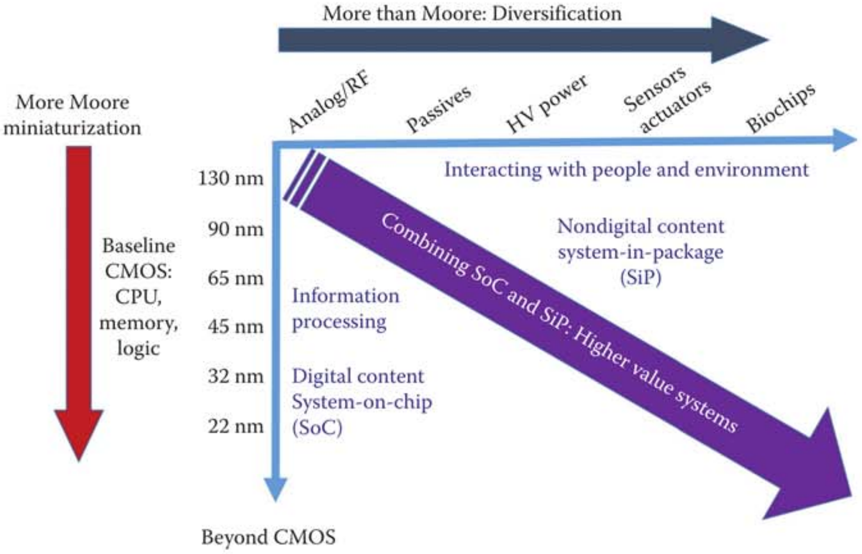


FIGURE 51.1 Schematic description of More Moore, More than Moore, and beyond complementary metal–oxide–semiconductor (CMOS). (RF, radio frequency; HV, high-voltage. [Reprinted from ITRS Roadmap for Semiconductors. Available at: www.itrs.org. With permission.])

The *More Moore* domain is internationally defined as an attempt to further develop advanced CMOS technologies and reduce the associated cost per function by introducing strain and silicon-on-insulator (SOI) technology (see Figure 51.1). Almost 70% of the total semiconductor components market is directly impacted by advanced CMOS miniaturization achieved in the More Moore domain. This 70% comprises three component groups of similar size, namely microprocessors, mass memories, and digital logic. As an example, in recent years, technology has advanced to fabricate ICs at 14-nm gate length commercially [4]. Fab industry giants such as Intel, TSMC, Samsung, and Global Foundries have plans to fabricate ICs at 10 nm by 2017. Samsung has already fabricated and tested 128 Mb SRAM in 10 nm [5] and is hoping to commercialize the process by the end of 2016. This aggressive scaling of technology is possible because of the advent of FinFETs and fully-depleted silicon-on-insulator (FD-SOI) device technology [6].

More than Moore (MtM) refers to a set of technologies that enable nondigital micro-/ nanoelectronic functions. MtM radio frequency (RF), high-voltage power, solid-state lighting (SSL), medical ultrasound, biochips and microfluidics, energy scavenging, electronic imaging, sensors, and actuators on CMOS platforms are some of the applications. They are based on, or derived from, Si technology but do not necessarily scale with Moore's Law. MtM devices typically provide conversion of nondigital as well as nonelectronic information, such as mechanical, thermal, acoustic, chemical, optical, and biomedical functions to digital data, and vice versa.

51.1.2 More Moore Technology

For digital circuits, a figure of merit for MOSFETs for unloaded circuits is CV/I, where C is the gate capacitance, V the voltage swing, and I the current drive of the MOSFET. For loaded circuits, the current drive of the MOSFET is of paramount importance. Keeping in mind both the CV/I metric and the benefits of a large current drive, one notes that device performance may be improved [7] by (1) inducing a larger charge density for a given gate voltage drive; (2) enhancing the carrier transport by improving the mobility, saturation velocity, or ballistic transport; (3) ensuring device scalability to achieve a shorter channel length; and (4) reducing parasitic capacitances and resistances. For capitalizing these opportunities, the proposed technology options generally fall into two categories: new materials and new device structures. In many cases, the introduction of a new material requires the use of a new device structure, or vice versa. To fabricate devices beyond current scaling limits, IC companies have simultaneously pushed the planar, bulk Si CMOS design while incorporating alternative gate stack materials (high-k dielectric [8] and metal gates), band engineering methods (using strained Si [9] or SiGe [10]), and alternative transistor structures. The concept of a band-engineered transistor is to enhance the mobility of electrons and/or holes in the channel by modifying the band structure of Si in the channel in a way such that the physical structure of the transistor remains substantially unchanged. This enhanced mobility increases the transistor transconductance (g_m) and on-drive current (I_{on}). A SiGe layer or a strained Si on a relaxed SiGe layer is used as an enhanced-mobility channel layer. It has already been demonstrated experimentally that at $T = 300$ K (room temperature), effective hole enhancement of about 50% can be achieved using the SiGe technology [10]. Intel has adopted strained Si technology for the first time for its 90-nm process [11]. This resulted in nearly 20% performance improvement per technology generation, with only a few additional process steps. The challenge in identifying suitable high-k dielectrics and metal gates for both conventional p-channel MOS (PMOS) and n-channel MOS (NMOS) transistors has led to early adoption of alternative transistor designs. These include primarily partially-depleted (PD) SOI and FD-SOI devices. Today, there is extensive research in double-gate (DG) structures and FinFET transistors [12], which have better electrostatic integrity and, theoretically, have better transport properties than single-gated FETs. A FinFET is a form of a DG transistor having surface conduction channels on two opposite vertical surfaces and having current flow in the horizontal direction. The channel length is given by the horizontal separation between source and drain and is usually determined by a lithographic step combined with a side-wall spacer etch process. Many innovative structures, involving structural challenges such as fabrication on nanometer-scale fins and nanometer-scale planarization over an entire wafer are currently being implemented. Some new and revolutionary technology such as carbon nanotubes or molecular transistors might be on the horizon, but it is not clear, in view of the predicted future capabilities of CMOS, whether they will be competitive. In conclusion, the semiconductor industry is approaching the end of an era of scaling gains by rote shrinkage of device dimensions and entering a postscaling era, a new phase of CMOS evolution in which innovation is demanded simply to compete. The trends in benefits to density, performance, and power will be continued through such innovations. Rather than coming to a close, a new era of CMOS technology is just beginning.

51.1.3 Importance of Modeling and Simulation

As semiconductor devices are scaled into nanoscale regimes, first, velocity saturation starts to limit the carrier mobility due to pronounced intervalley scattering; when the device dimensions are scaled below 200 nm, velocity overshoot (velocity of the carriers in the channel is larger than the steady-state saturation velocity) starts to dominate the device behavior leading to larger on-state currents. Alongside with the developments in the semiconductor nanotechnology, in recent years, there has been significant progress in physically based modeling of semiconductor devices. There are two issues that make simulation important. Product cycles are getting shorter with each generation, and the demand for production wafers shadows development efforts in the factory. Consider the product cycle issue first. In order for companies to maintain their competitive edge, products have to be taken from design to production in less than 18 months. As a result, the development phase of the cycle is getting shorter. Contrast this requirement with the fact that it takes 2–3 months to run a wafer lot through a factory, depending on its complexity. The specifications for experiments run through the factory must be near the final solution. While simulations may not be completely predictive, they provide a good initial guess. This can ultimately reduce the number of iterations during the device development phase. The second issue that reinforces the need for simulation is the production pressures that factories face. To meet customer demand, development factories are making way for production space. It is also expensive to run experiments through a production facility. Such resources could have otherwise been used to produce sellable products. Again, device simulation can be used to decrease the number of experiments run through a factory. Device simulation can be used as a tool to guide manufacturing down the right path, thereby decreasing the development time and costs. Besides offering the possibility to test hypothetical devices that have not (or could not) yet been manufactured, device simulation offers unique insight into device behavior by allowing the observation of phenomena that cannot be measured on real devices. It is related to, but usually separate from, process simulation, which deals with various physical processes such as material growth, oxidation, impurity diffusion, etching, and metal deposition inherent in device fabrication leading to ICs. Device simulation is distinct from another important aspect of computer-aided design (CAD), device modeling, which deals with compact behavioral models for devices and subcircuits relevant for circuit simulation in commercial packages such as simulation program with integrated circuit emphasis (SPICE). Device simulation can provide the parameters that are used to generate the compact behavioral models, and when coupled with process simulation and circuit simulation, it provides a hierarchical approach to technology computer-aided design (TCAD).

The main components of semiconductor device simulation at any level of approximation are illustrated in Figure 51.2 [13]. There are two main kernels, which must be solved self-consistently with one another, the *transport equations* governing charge flow, and the *fields* driving charge flow. Both are coupled strongly to one another, and hence must be solved simultaneously. The fields arise from external sources, as well as the charge and current densities which act as sources for the time-varying electric and magnetic fields obtained from the solution of Maxwell's equations. Under appropriate conditions, only the quasi-static electric fields arising from the solution of Poisson's equation are necessary. The fields, in turn, are driving forces for charge transport as illustrated in Figure 51.3 for the various levels of approximation within a hierarchical structure ranging from compact modeling at the top to an exact quantum-mechanical description at the bottom. For devices for which gradual channel approximation cannot be used due to the two-dimensional (2D) nature of the electrostatic potential and the electric fields driving the carriers from source to drain, drift-diffusion models have been exploited. These models are valid, in general, for large devices in which the fields are not that high so that there is no degradation of the mobility due to the electric field. The validity of the drift-diffusion models can be extended to take into account the velocity saturation effect with the introduction of field-dependent mobility and diffusion coefficients. When velocity overshoot becomes important (for length scales less than 200 nm—see Figure 51.4), the drift-diffusion model is no longer valid and the hydrodynamic model must be used.

The hydrodynamic model has been the workhorse for technology development and several high-end commercial device simulators have appeared including Silvaco, Synopsys, Crosslight, etc. The advantages of the hydrodynamic model are that it allows quick simulation runs, but the problem is that the amount of the velocity overshoot depends upon the choice of the energy relaxation time. The smaller the device,

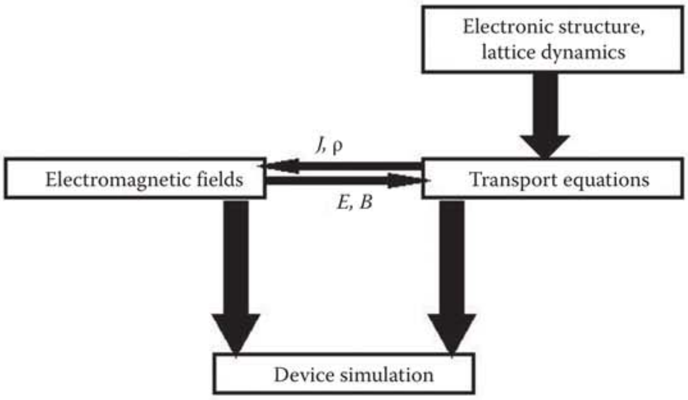
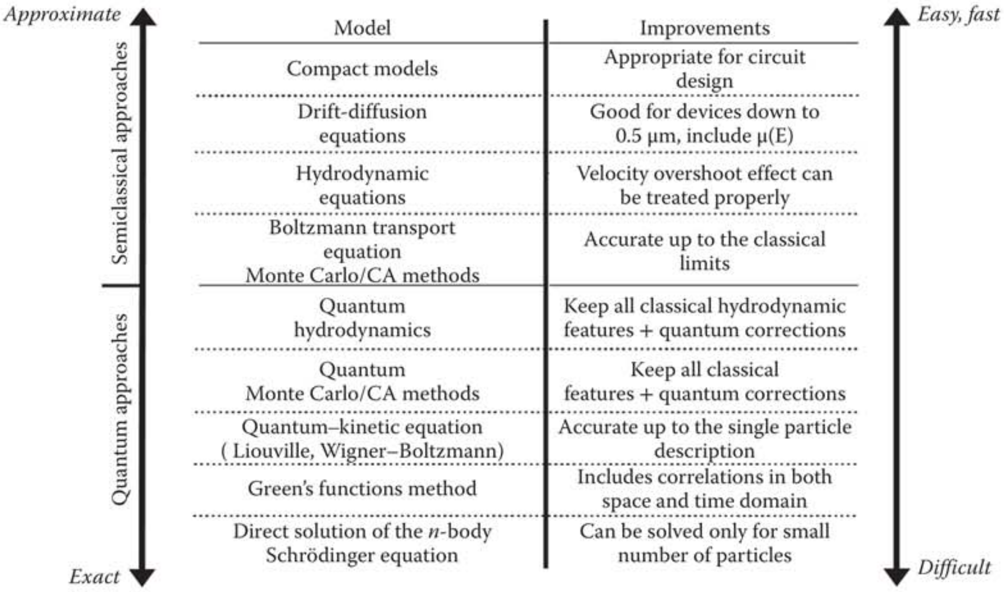


FIGURE 51.2 A schematic description of the building blocks of device simulator.



	Model	Improvements
Semiclassical approaches	Compact models	Appropriate for circuit design
	Drift-diffusion equations	Good for devices down to 0.5 μm, include μ(E)
	Hydrodynamic equations	Velocity overshoot effect can be treated properly
	Boltzmann transport equation Monte Carlo/CA methods	Accurate up to the classical limits
Quantum approaches	Quantum hydrodynamics	Keep all classical hydrodynamic features + quantum corrections
	Quantum Monte Carlo/CA methods	Keep all classical features + quantum corrections
	Quantum–kinetic equation (Liouville, Wigner–Boltzmann)	Accurate up to the single particle description
	Green's functions method	Includes correlations in both space and time domain
	Direct solution of the *n*-body Schrödinger equation	Can be solved only for small number of particles

FIGURE 51.3 Illustration of the hierarchy of transport models.

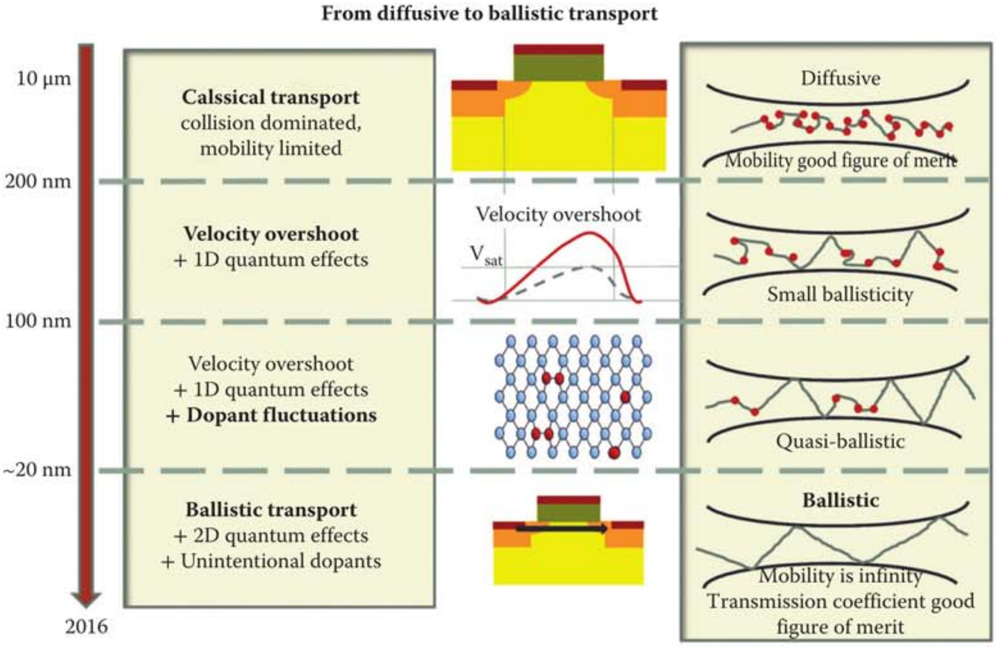


FIGURE 51.4 From diffusive to ballistic transport when scaling semiconductor devices into nanometer scale.

the larger the deviation when using the same set of energy relaxation times. A standard way in calculating the energy relaxation times is to use bulk Monte Carlo simulations. However, the energy relaxation times are material-, device geometry-, and doping-dependent parameters, so their determination ahead of time is not possible.

To avoid the problem of the proper choice of the energy relaxation times in hydrodynamic models, a direct solution of the Boltzmann transport equation (BTE)

$$\frac{\partial f}{\partial t} + \frac{1}{\hbar}\nabla_k E(k) \cdot \nabla_r f + \left(-\frac{e\mathbf{E}}{\hbar}\right) \cdot \nabla_k f = \left(\frac{\partial f}{\partial t}\right)_{\mathbf{coll}} \tag{51.1}$$

using the Monte Carlo method is the best method of choice for nanoscale devices in which position and momentum can still be treated as independent variables. In Equation 51.1, $f(r, k, t)$ is the semiclassical distribution function, $E(k)$ the dispersion relation for the electrons, and $\mathbf{E}$ the electric field. When, for example, collisional broadening of the states is important, quantum transport approaches must be pursued.

This chapter focuses on Monte Carlo methods for the solution of the BTE and is organized as follows. Monte Carlo solution of the BTE for homogeneous systems is given in Section 51.2. Particle-based device simulator is described in Section 51.3. Representative simulation results for bulk material and for semiconductor devices are given in Section 51.4. Conclusions and future directions of research are presented in Section 51.5.

51.2 Bulk Monte Carlo Method Description

The ensemble Monte Carlo (EMC) techniques have been used for well over 40 years as a numerical method to simulate nonequilibrium transport in semiconductor materials and devices and have been the subject of numerous books and reviews [14–16]. In the applications to transport problem, a random walk is generated using the random number–generating algorithms common to modern computers to simulate the stochastic motion of particles subject to collision processes. This process of random walk generation is part of a very general technique used to evaluate integral equations and is connected to the general random sampling technique used in the evaluation of multidimensional integrals [17].

The basic technique as applied to transport problems is to simulate the free particle motion (referred to as the *free flight*) terminated by instantaneous random *scattering events*. The Monte Carlo algorithm consists of generating random free-flight times for each particle, choosing the type of scattering occurring at the end of the free flight, changing the final energy and momentum of the particle after scattering, and then repeating the procedure for the next free flight. Sampling the particle motion at various times throughout the simulation allows for the statistical estimation of physically interesting quantities such as the single particle distribution function, the average drift velocity in the presence of an applied electric field, the average energy of the particles, etc. By simulating an *ensemble* of particles, representative of the physical system of interest, the nonstationary time-dependent evolution of the electron and hole distributions under the influence of a time-dependent driving force may be simulated. This particle-based picture, in which the particle motion is decomposed into free flights terminated by instantaneous collisions, is basically the same approximate picture underlying the derivation of the semiclassical BTE. In fact, it may be shown that the one-particle distribution function obtained from the random walk Monte Carlo technique satisfies the BTE for a homogeneous system in the long-time limit [18]. This semiclassical picture breaks down when quantum-mechanical effects become pronounced, and one cannot unambiguously describe the instantaneous position and momentum of a particle. In the following, we describe first the standard Monte Carlo algorithm used to simulate charge transport in semiconductors. We then discuss how this basic model for charge transport within the BTE is self-consistently solved with the appropriate field equations to perform particle-based device simulation (Section 51.3).

In the bulk Monte Carlo method [19], particle motion is assumed to consist of free flights terminated by instantaneous scattering events, which change the momentum and energy of the particle after scattering. So, the first task is to generate free flights of random time duration for each particle. To simulate this process, the probability density, $P(t)$, is required, in which $P(t)\mathrm{d}t$ is the joint probability that a particle will arrive at time t without scattering after a previous collision occurring at time $t = 0$, and then suffer a collision in a time interval dt around time t. The probability of scattering in the time interval dt around t may be written as $\Gamma[\mathbf{k}(t)]\mathrm{d}t$, where $\Gamma[\mathbf{k}(t)]$ is the scattering rate of an electron or hole of wavevector **k**. The scattering rate, $\Gamma[\mathbf{k}(t)]$, represents the sum of the contributions from each individual scattering mechanism, which are usually calculated quantum mechanically using perturbation theory [20]. The implicit dependence of $\Gamma[\mathbf{k}(t)]$ on time reflects the change in **k** due to acceleration by internal and external fields. For electrons subject to time-independent electric and magnetic fields, the time evolution of **k** between collisions is represented as

$$\mathbf{k}(t) = \mathbf{k}(0) - \frac{e(\mathbf{E} + \mathbf{v} \times \mathbf{B})t}{\hbar}, \tag{51.2}$$

where **E** is the electric field, **v** the electron velocity, and **B** the magnetic flux density. In terms of the scattering rate, $\Gamma[\mathbf{k}(t)]$, the probability that a particle has not suffered a collision after a time t is given by $\exp\left(-\int_0^t \Gamma\left[k(t')\right]\mathrm{d}t'\right)$. Thus, the probability of scattering in the time interval dt after a free flight of time $t$may be written as the joint probability:

$$P(t)\mathrm{d}t = \Gamma[k(t)]\exp\left[-\int_0^t \Gamma\left[k\left(t'\right)\right]\mathrm{d}t'\right]\mathrm{d}t. \tag{51.3}$$

Random flight times may be generated according to the probability density $P(t)$ above using, for example, the pseudo-random number generator implicit on most modern computers, which generate uniformly distributed random numbers in the range [0, 1]. Using a direct method, random flight times sampled from $P(t)$ may be generated according to

$$r = \int_0^{t_r} P(t)\mathrm{d}t, \tag{51.4}$$

where r is a uniformly distributed random number and t_r is the desired free-flight time. Integrating Equation 51.4 with $P(t)$ given by Equation 51.3 above yields

$$r = 1 - \exp\left[-\int_0^{t_r} \Gamma\left[k(t')\right]dt'\right]. \tag{51.5}$$

Since $1 - r$ is statistically the same as r, Equation 51.5 may be simplified to

$$-\ln r = \int_0^{t_r} \Gamma\left[\mathbf{k}\left(t'\right)\right]\mathrm{d}t'. \tag{51.6}$$

Equation 51.6 is the fundamental equation used to generate the random free-flight time after each scattering event, resulting in a random walk process related to the underlying particle distribution function. If there is no external driving field leading to a change of **k** between scattering events (e.g., in ultrafast photoexcitation experiments with no applied bias), the time dependence vanishes, and the integral is trivially evaluated. In the general case, where this simplification is not possible, it is expedient to introduce the so-called self-scattering method [21] in which one introduces a fictitious scattering mechanism whose scattering rate always adjusts itself in such a way that the total (self-scattering plus real scattering) rate is energy-independent:

$$\Gamma = \Gamma\left[\mathbf{k}(t')\right] + \Gamma_{\text{self}}\left[\mathbf{k}(t')\right], \tag{51.7}$$

where $\Gamma_{\text{self}}[\mathbf{k}(t')]$ is the self-scattering rate. The self-scattering mechanism itself is defined such that the final state before and after scattering is identical. Hence, it has no effect on the free-flight trajectory of a particle when selected as the terminating scattering mechanism, yet results in the simplification of Equation 51.6 such that the free flight is given by

$$t_r = -\frac{1}{\Gamma}\ln r. \tag{51.8}$$

The constant total rate (including self-scattering) Γ must be chosen at the start of the simulation interval (there may be multiple such intervals throughout an entire simulation) so that it is larger than the maximum scattering encountered during the same time interval. In the simplest case, a single value is chosen at the beginning of the entire simulation (constant gamma method), checking to ensure that the real rate never exceeds this value during the simulation. Other schemes may be chosen that are more computationally efficient, and which modify the choice of Γ at fixed time increments [22].

The algorithm described above determines the random free-flight times during which the particle dynamics is treated semiclassically. For the scattering process itself, we need the type of scattering (i.e., impurity, acoustic phonon, photon emission, etc.) which terminates the free flight, and the final energy and momentum of the particle(s) after scattering. The type of scattering that terminates the free flight is chosen using a uniform random number between 0 and Γ, and using this pointer to select among the relative total scattering rates of all processes including self-scattering at the final energy and momentum of the particle:

$$\Gamma = \Gamma_{\text{self}}[n,\mathbf{k}] + \Gamma_1[n,\mathbf{k}] + \Gamma_2[n,\mathbf{k}] + \cdots \Gamma_N[n,\mathbf{k}], \tag{51.9}$$

with n the band index of the particle (or sub-band in the case of reduced-dimensionality systems), and $\mathbf{k}$ the wavevector at the end of the free flight. This process is illustrated schematically in Figure 51.5.

Once the type of scattering terminating the free flight is selected, the final energy and momentum (as well as band or subband) of the particle due to this type of scattering must be selected. For elastic scattering processes such as ionized impurity scattering, the energy before and after scattering is the same. For the interaction between electrons and the vibrational modes of the lattice described as quasi-particles known as phonons, electrons exchange finite amounts of energy with the lattice in terms of emission and absorption of phonons. For determining the final momentum after scattering, the scattering rate, $\Gamma_j[n,\mathbf{k};m,\mathbf{k}']$ of the jth scattering mechanism is needed, where n and m are the initial and final band indices, and $\mathbf{k}$ and $\mathbf{k}'$ are the particle wavevectors before and after scattering. Defining a spherical coordinate system

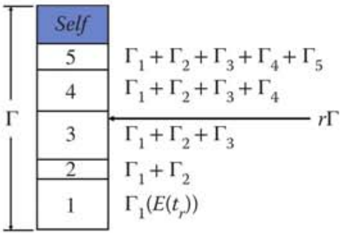


FIGURE 51.5 Selection of the type of scattering terminating a free flight in the Monte Carlo algorithm.

around the initial wavevector **k**, the final wavevector $\mathbf{k}'$ is specified by $|\mathbf{k}'|$ (which depends on conservation of energy) as well as the azimuthal and polar angles, ϕ and θ around **k**. Typically, the scattering rate, $\Gamma_j[n,\mathbf{k};m,\mathbf{k}']$, only depends on the angle θ between **k** and $\mathbf{k}'$. Therefore, ϕ may be chosen using a uniform random number between 0 and 2π (i.e., $2\pi r$), while θ is chosen according to the angular dependence for scattering arising from $\Gamma_j[n,\mathbf{k};m,\mathbf{k}']$. If the probability for scattering into a certain angle $P(\theta)\mathrm{d}\theta$ is integrable, then random angles satisfying this probability density may be generated from a uniform distribution between 0 and 1 through a prescribed recipe. Otherwise, a rejection technique may be used to select random angles according to $P(\theta)$. Common scattering mechanisms that contribute to transport are summarized in Figure 51.6. The corresponding scattering rate expressions for general nonparabolic bands are summarized in Table 51.2 [23–25].

A general Monte Carlo code is then developed as follows: First, a subroutine is typically called that contains all material and scattering rates parameters for the scattering mechanisms included in the theoretical model. After the material and run parameters are read, in the first step of the Monte Carlo simulation procedure, it is necessary to construct scattering tables for the Γ, L, and X valleys (for GaAs as a prototypical example) that initialize a series of events that are summarized in Figure 51.7. Then, for each energy, the cumulative scattering rates for each valley are stored in separate look-up tables and renormalized according to the maximum scattering rate (including self-scattering) that occurs over the range of energies stored.

Having constructed the scattering table and after renormalizing the table, examples of which are given for GaAs in Figures 51.8 and 51.9 for the Γ, L, and X valleys, the next step is to initialize the carrier's wavevector and energy and the initial free-flight time. This is accomplished by calling the initialization subroutine. Energy and wavevector histograms of the initial carrier energy and the components of the wavevector along the x-, y-, and z-axes are shown in Figure 51.10. For good statistics, the number of particles simulated is typically taken to be 10,000, and one can see the statistical fluctuation of these average quantities associated with the finite number of particles. Notice that the variance is inversely proportional to the square-root of the number of particles. Also note that the initial y-component for the wavevector is symmetric around the y-axis, which means that the average wavevector along the y-axis is zero, which should be expected since the electric field along the y-component is zero at $t = 0$. Identical distributions have been obtained for the x- and for the z-components of the wavevector. Also note that the energy distribution has the Maxwell–Boltzmann form, as it should be expected. One can also estimate from this graph that the average energy of the carriers is on the order of $(3/2)k_\mathrm{B}T$ (k_B is the Boltzmann constant and T the temperature).

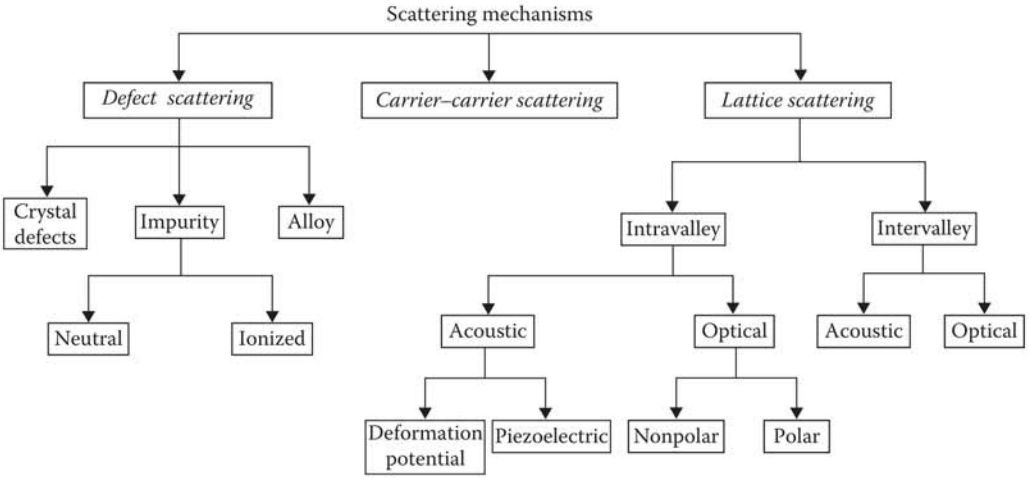


FIGURE 51.6 Scattering mechanisms in a typical semiconductor.

TABLE 51.2 Scattering Rate Expressions Γ(k) = W(k) = W(E) for Electrons for General Nonparabolic Energy Bands

---

1. *Ionized impurity scattering:* Scattering from ionized Coulomb charges in the Brooks–Herring approach.

$W(E) = \left(\frac{\sqrt{2}e^4 N_I m_d^{\frac{3}{2}}}{\pi\varepsilon_s^2 ?^4}\right) \times \left(\sqrt{E(1+\alpha E)} \times (1+2\alpha E)\right) \times \left(\frac{1}{q_D^2\left(q_D^2 + \frac{8m_d E(1+\alpha E)}{\hbar^2}\right)}\right)$ $q_D = \sqrt{\frac{e^2(n+p)}{\varepsilon_s K_B T_L}}$, where n and p are the electron and hole concentrations, respectively, k_B is the Boltzmann constant, m_d is the density of states mass, T_L is the lattice temperature, ε_s is the permittivity of the material, E is the energy, and α is the nonparabolicity factor.

2. *Alloy disorder scattering* (e.g., $Al_xGa_{1-x}As$): Scattering due to the presence of a random alloy.

$W(E) = \left(\frac{x(1-x)a^3}{\pi}\right) \times \left(\frac{D_{\text{alloy}}^2 d}{\hbar^4}\right) \times m_d\sqrt{2m_d E(1+\alpha E)} \times (1+2\alpha E)$,

where x is the mole fraction, d the lattice disorder ($0 \le d \le 1$), a the lattice constant, and D_{alloy} the alloy disorder scattering potential.

3. *Dislocation scattering* (e.g., GaN): Scattering due to crystal defects/dislocations (Coulomb-type interaction).

$W(E) = \left(\frac{N_{\text{dis}} m_d e^4}{4\hbar^3 \varepsilon^2 c^2}\right) \times \left(\frac{\lambda^4}{\left(1+\frac{8\lambda^2 m_d E(1+\alpha E)}{\hbar^2}\right)^{\frac{3}{2}}}\right) \times \left(1 + \frac{4\lambda^2 m_d E(1+\alpha E)}{\hbar^2}\right) \times (1+2\alpha E)$ $\lambda = \sqrt{\frac{\varepsilon K_B T_L}{e^2 n'}}$, where n' is the effective screening concentration and N_{dis} the line dislocation density.

4. *Acoustic deformation potential scattering:* Describes fluctuations in the band structure due to deformation of the unit cell because of the strain due to the different amounts of displacement of the atoms (in the same direction).

$W(E) = \left(\frac{2\pi D_{ac}^2 K_B T_L}{\hbar C_l}\right) \times \left(\frac{(2m_d)^{\frac{3}{2}}\sqrt{E(1+\alpha E)}}{4\pi^2\hbar^3}\right) \times (1+2\alpha E)$,

where D_{ac} is the acoustic deformation potential constant and C_l is the sound velocity.

5. *Piezoelectric scattering:* Dipole-type interaction in zinc-blende materials in which there is charge transfer between the two atoms in the unit cell. The optical modes of vibration (vibration of the atoms in opposite direction) directly affect the magnitude of the dipole moment, which in turn leads to scattering potential.

$W(\mathrm{E}) = \left(\frac{m_d^{\frac{1}{2}} \mathrm{K_B T_L}}{4\sqrt{2}\pi\rho \mathrm{V}_s^2 \hbar^2}\right) \times \left(\frac{1+2\alpha E}{\sqrt{E(1+\alpha E)}}\right) \times \left(\frac{e e_{pz}}{\varepsilon_\infty}\right)^2 \times \ln\left(1 + \frac{8m_d E(1+\alpha E)}{\hbar^2 q_D^2}\right)$ $q_D = \sqrt{\frac{e^2(n+p)}{\varepsilon_s K_B T_L}}$, where ρ is the crystal density, v_s sound velocity, e_{pz} the piezoelectric constant, and ε_∞ the high-frequency dielectric constant.

6. *Polar optical phonon (POP) scattering:* In polar materials (where there is transfer of charge from one atom to the other, which in turn, leads to a dipole moment), optical modes of vibration cause direct modification of the dipole moment, which then leads to polar optical phonon scattering which is a Coulomb-type interaction.

$W(E) = \left(\frac{\sqrt{m_d}e^2 W_{\mathrm{LO}}}{4\sqrt{2}\pi\hbar e_p}\right) \times \left(N_0 + \frac{1}{2} \mp \frac{1}{2}\right) \times \left(\frac{1+2\alpha E_k'}{\gamma_k}\right) \times F\left(E_k, E_k'\right)$

$N_0 = \frac{1}{\frac{\hbar_{\mathrm{LO}} W}{e^{K_B T_{L-1}}}}$ $\varepsilon_p = \frac{1}{\frac{1}{\varepsilon_{\text{high}}} - \frac{1}{\varepsilon_{\text{low}}}}$ $F(E_k, E_k') = \ln\left(\mathrm{mod}\left(\frac{\sqrt{\gamma_k}+\sqrt{\gamma_{k'}}}{\sqrt{\gamma_k}-\sqrt{\gamma_{k'}}}\right)\right)$

$\gamma_K = E_k(1+\alpha)E_k$ $E_k = E_k \pm \hbar W_{\mathrm{LO}}$,

where W_{LO} is the optical phonon energy and the indices high/low stand for high frequency/static dielectric constants.

7. *Intervalley phonon scattering:* Scattering between equivalent or nonequivalent valleys.

$W(E) = \left(\frac{\pi D_{ij}^2 Z_j}{\rho W_{ij}}\right) \times \left(n(W_{ij})\frac{1}{2} \mp \frac{1}{2}\right) \times \left(\frac{(2m_d)^{\frac{3}{2}}\sqrt{E_f(1+\alpha E_f)}}{4\pi^2\hbar^3}\right) \times (1+2\gamma E_f)$ $E_f = F \pm \hbar w_{ij} - \Delta E_{ij}$,

where D_{ij} is the coupling constant, Z_j the number of final valleys, ρ the crystal density, and w_{ij} the energy of the phonon involved in the process.

Sources: M. Lundstrom, Fundamentals of Carrier Transport, Cambridge University Press, London, 2000; B. K. Ridley, *Quantum Processes in Semiconductors, Oxford University Press, 1982; D. Vasileska nanoHUB.org page: nanohub.org/groups/dragica_vasileska/semiclassical* For more details on the derivation of these formulas, please refer to Refs. [26–28].

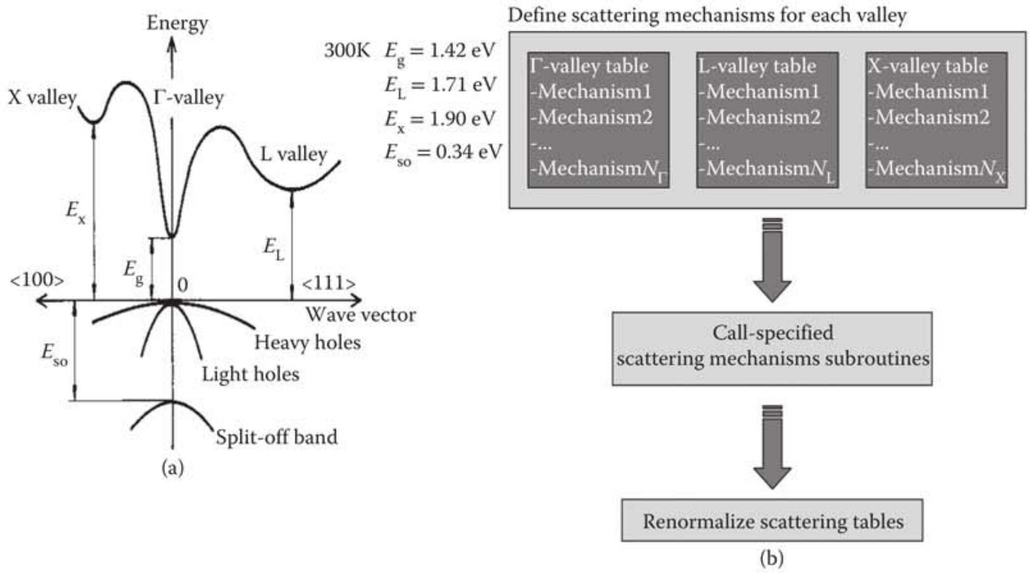


FIGURE 51.7 (a) Schematics of a bandstructure for GaAs. (b) Procedure for the creation of the scattering tables for electron transport.

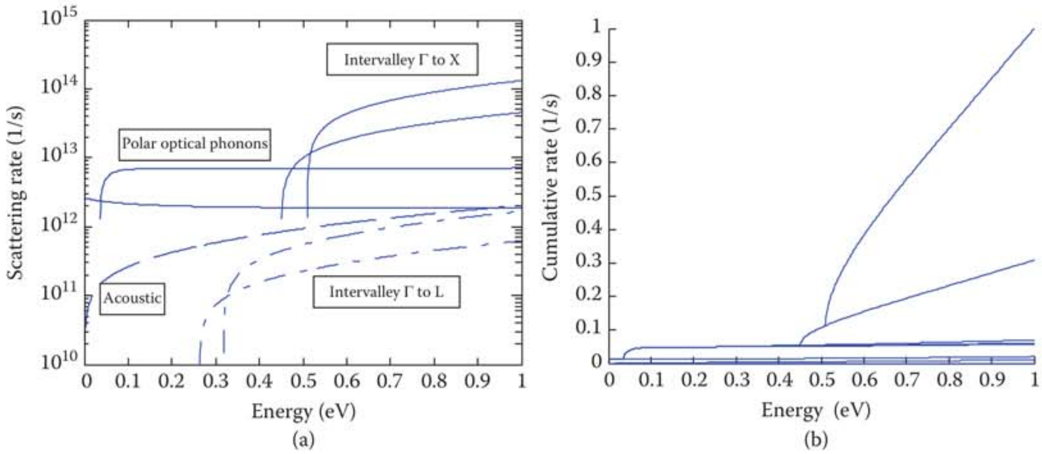


FIGURE 51.8 (a) Scattering rates for the Γ valley. For simplicity, we have omitted Coulomb scattering in these calculations. In (a) the dashed line corresponds to the acoustic phonon scattering rate, solid lines correspond to polar optical phonon scattering (absorption and emission), and the dashed-dotted line corresponds to intervalley scattering from Γ to L valleys. Since the L valley is along the [111] direction, there are eight equivalent directions, and since these valleys are shared, there are a total of four equivalent L valleys. The dotted line corresponds to scattering from Γ to X valleys. The X valleys are in the [100] direction and since there are six equivalent [100] directions and the valleys are shared between Brillouin zones, there are three equivalent X valleys. (b) Normalized cumulative scattering table for the Γ valley. Everything above the top line up to $\Gamma = 1$ is self-scattering so it is advisable when checking the scattering mechanisms to first check whether the scattering mechanism chosen is self-scattering or not. This is in particular important for energies below 0.5 eV for this particular scattering table when the Γ to X intervalley scattering (absorption and emission) takes over.

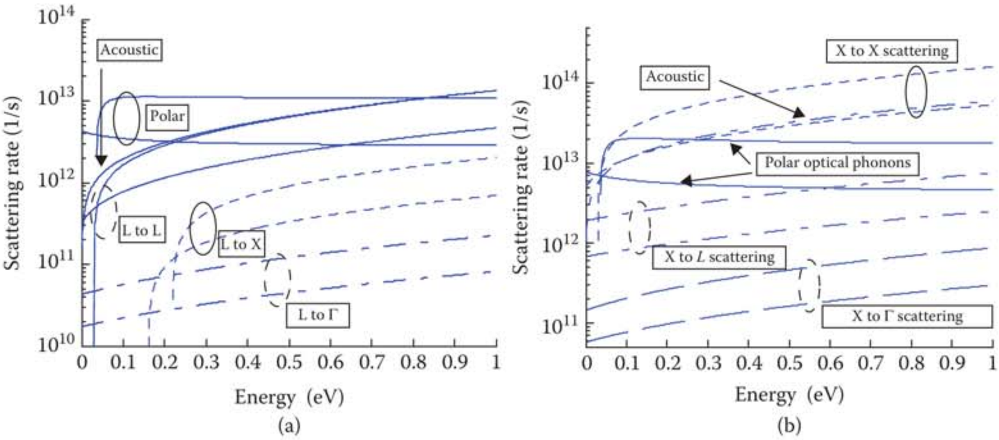


FIGURE 51.9 Scattering rates for the L (a) and X (b) valleys used to create the corresponding normalized scattering tables (not shown here).

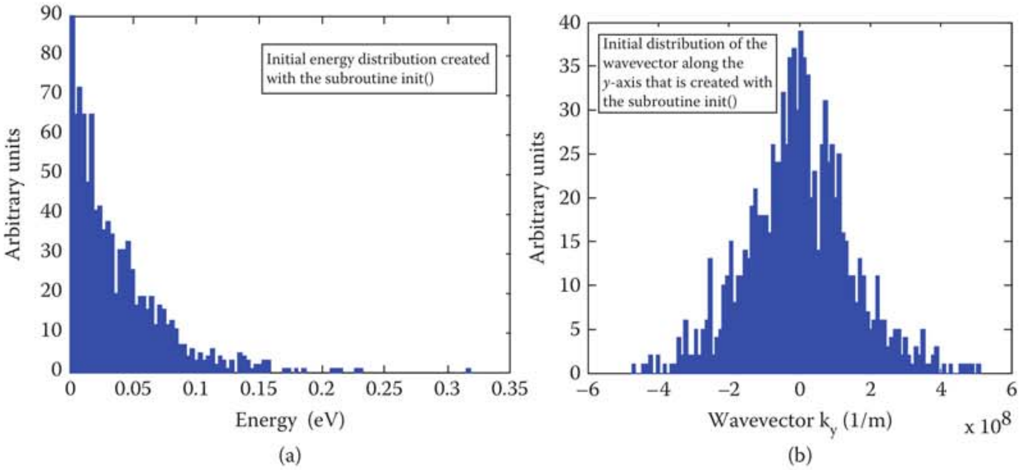


FIGURE 51.10 Initial carrier distribution for an ensemble of 10,000 particles. (a) Energy distribution. (b) Distribution of wavevector k_y.

When the initialization process is finished, the main free-flight-scatter procedure takes place until the completion of the simulation time. There are two components in this process: first, the carriers accelerate freely due to the electric field, accomplished by calling the **drift()** subroutine (which implements Equation 51.2), and then their free-flights are interrupted by random scattering events that are managed by the **scatter_carrier()** subroutine. The flowchart for performing the free-flight-scatter process within one time-step Δt is shown diagrammatically in Figure 51.11.

In the **scatter_carrier()** subroutine from Figure 51.11, first the scattering mechanism terminating the free flight is chosen, to which certain attributes are associated such as the change in energy after scattering. For inelastic scattering processes, we have the change in energy due to emission or absorption of phonons, for example. Also, the nature of the scattering process is identified as isotropic or anisotropic. Note that when performing acoustic phonon and intervalley scattering for GaAs, both of which are isotropic scattering processes, no coordinate system transformation is needed to determine the final wavevector after scattering. Because polar optical phonon and Coulomb scattering mechanisms are anisotropic, it is necessary to do a rotation of the coordinate system, scatter the carrier in the rotated system, and then perform

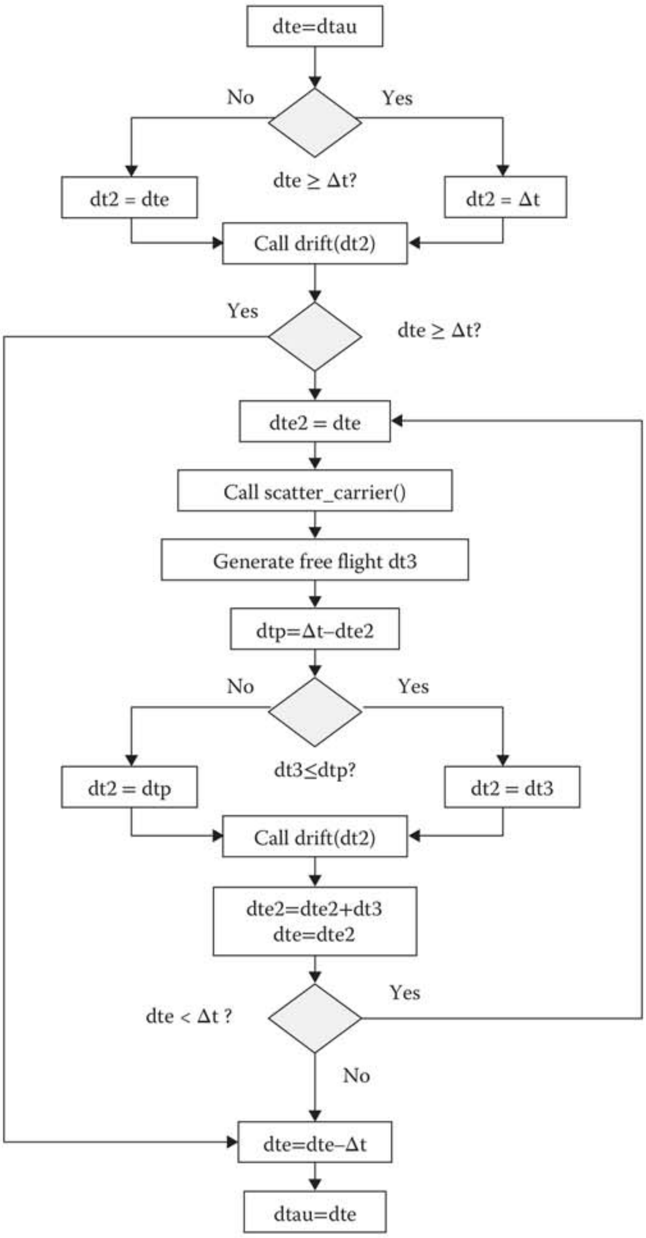


FIGURE 51.11 Free-flight-scatter procedure within one time-step. Time variable dtau is the remaining free-flight time from previous time-step Δt.

inverse coordinate transformation. This procedure is needed because it is much easier to determine final carrier momentum in the rotated coordinate system in which the initial wavevector k is aligned with the z-axis. For this case, one can calculate that the final polar angle of scattering with polar optical phonons for parabolic bands in the rotated coordinate system as

$$\cos\theta = \frac{(1+\xi) - (1+2\xi)^r}{\xi}, \quad \xi = \frac{2\sqrt{E_k\left(E_k \pm \hbar\omega_0\right)}}{\left(\sqrt{E_k} - \sqrt{E_k \pm \hbar\omega_0}\right)^2}, \tag{51.10}$$

where E_k is the carrier energy, $\hbar\omega_0$ the polar optical phonon energy, and r a random number uniformly distributed between 0 and 1. The final angle for scattering with ionized impurities (Coulomb scattering) and for parabolic bands is

$$\cos\theta = 1 - \frac{2r}{1 + 4k^2 L_D^2 (1 - r)}, \tag{51.11}$$

where k is the carrier wavevector and L_D is the Debye screening length. The azimuthal angle for both scattering processes is simply calculated using $\varphi = 2\pi r$. The importance of properly calculating the polar angle θ after scattering to describe small angles of deflection in the case of Coulomb or polar optical phonon scattering is illustrated in Figure 51.13. The histogram of the polar angle (from 0 to $\pi = 3.141592654$) after scattering for electron–polar optical phonon scattering is presented in Figure 51.12, from where we can clearly see the preference for small angle deflections that are characteristic for any Coulomb-type interaction (polar optical phonon is in fact electron–dipole interaction). Graphical representation of the determination of the final angle after scattering for both isotropic and anisotropic scattering processes is shown in Figure 51.13.

The direct technique described above can be applied when the integrals describing $\cos\theta$ can be analytically calculated. For most cases of interest, the integral cannot be easily inverted. In these cases, a rejection technique may be employed. The procedure for the rejection technique goes as follows:

- Choose a maximum value C, such that $C > f(x)$ for all x in the interval (a, b).
- Choose pairs of random numbers, one between a and b $\left(x_1 = a + r_1(b - a)\right)$ and another one$f_1 = r_1' C$ between 0 and C, where r_1 and r_1' are random numbers uniformly distributed between 0 and 1.
- If $f_1 \leq f(x_1)$, then the number x_1 is accepted as a suitable value, otherwise it is rejected.

The three steps described above are schematically shown in Figure 51.14. For $x = x_1$, $r_1 C$ is larger than $f(x_1)$ and in this case, if this represents the final polar angle for scattering, this angle is rejected and a new sequence of two random numbers is generated to determine x_2 and $r_2 C$. In this second case, $f(x_2) > r_2 C$ and the polar angle $\theta = x_2$ is selected (for polar angle selection $a = 0$ and $b = \pi$).

After the simulation is completed, typical results to check are the velocity–time, the energy–time, and the valley occupation versus time characteristics, such as those shown in Figure 51.15, where the velocity–time characteristics for applied electric fields ranging from 0.5 to 7 kV/cm, with an electric field increment

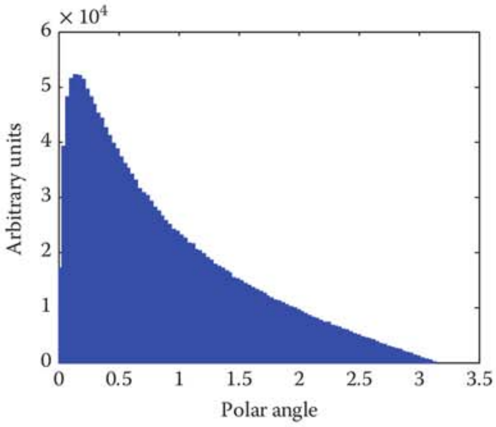
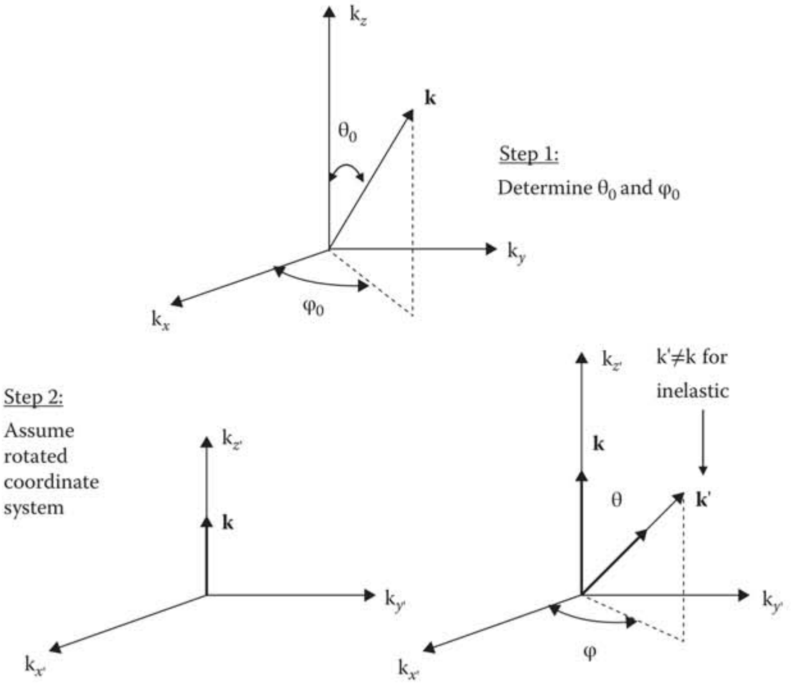


FIGURE 51.12 Histogram of the polar angle for electron–polar optical phonon scattering.

1. Isotropic scattering processes

$$\cos\theta = 1 - 2r,\ \varphi = 2\pi r$$

2. Anisotropic scattering processes (Coulomb, POP)



Step 3:

perform scattering

$$\cos\theta = \frac{(1+\xi) - (1+2\xi)^r}{\xi},\quad \xi = \frac{2\sqrt{E_k(E_k \pm \hbar\omega_0)}}{\left(\sqrt{E_k} - \sqrt{E_k \pm \hbar\omega_0}\right)^2} \text{POP}$$

$$\cos\theta = 1 - \frac{2r}{1 + 4k^2L^2{}_D(1-r)} \text{ Coulomb}$$

$\varphi = 2\pi r$ for both

Step 4:

$k_{xp} = k'\sin\theta\cos\varphi,\ k_{yp} = k'\sin\theta^*\sin\varphi,\ k_{zp} = k'\cos\theta$

Return back to the original coordinate system:
$k_x = k_{xp}\cos\varphi_0\cos\theta_0 - k_{yp}\sin\varphi_0 + k_{zp}\cos\varphi_0\sin\theta_0$
$k_y = k_{xp}\sin\varphi_0\cos\theta_0 + k_{yp}\cos\varphi_0 + k_{zp}\sin\varphi_0\sin\theta_0$
$k_z = -k_{xp}\sin\theta_0 + k_{zp}\cos\theta_0$

FIGURE 51.13 Description of final angle selection for isotropic and anisotropic scattering processes using the direct technique.

of 0.5 kV/cm, are shown. These clearly demonstrate that after a transient phase, the system reaches a stationary steady state, after which time we can start taking averages for calculating steady-state quantities.

From the results shown in Figure 51.15, one can see that steady state is achieved for larger time intervals when the electric field value is increased and the carriers are still sitting in the Γ valley. Afterward, time needed to get to steady state decreases. This trend is related to the valley repopulation and movement of the

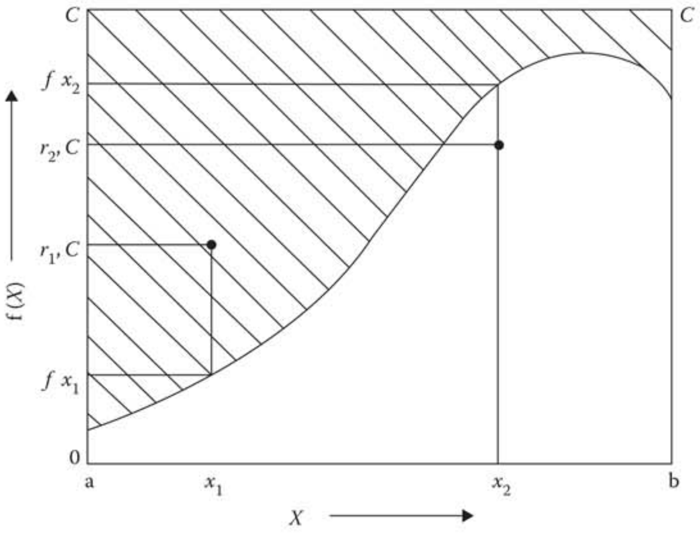


FIGURE 51.14 Schematic description of the rejection technique.

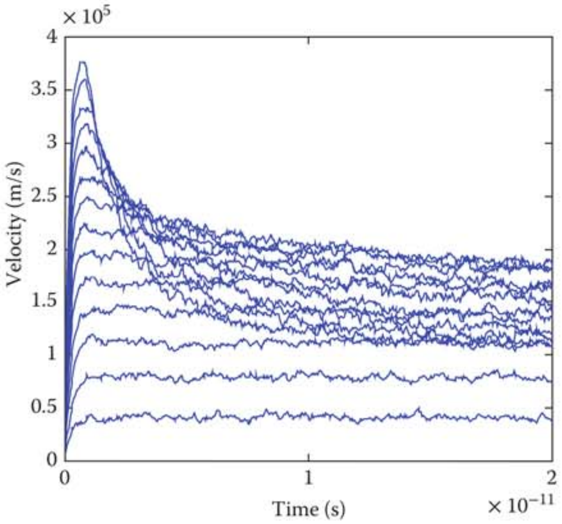


FIGURE 51.15 Time evolution of the drift velocity for electric field strengths ranging between 0.5 and 7 kV/cm, in 0.5 kV/cm increments.

carriers from Γ, into the X and finally into the L valley. The steady-state velocity-field and valley population versus electric field characteristics are shown in Figures 51.16 and 51.17, respectively. One can clearly see on the velocity–field characteristics that a low field mobility of about 8000 cm^2/V-s is correctly reproduced for GaAs without the use of any adjustable parameters.

At this point, it is advisable to check the energy and wavevector histograms (Figure 51.18) to ensure that the energy range chosen in the scattering tables is correct or not for the particular maximum electric field strength being considered, which gives the worst case scenario. Since, as already noted, we apply the electric field in the *y*-direction, for comparative purposes, we plot the histograms of the *x*-component of the wavevector, *y*-component of the wavevector, and the histogram of the final carrier energy distribution

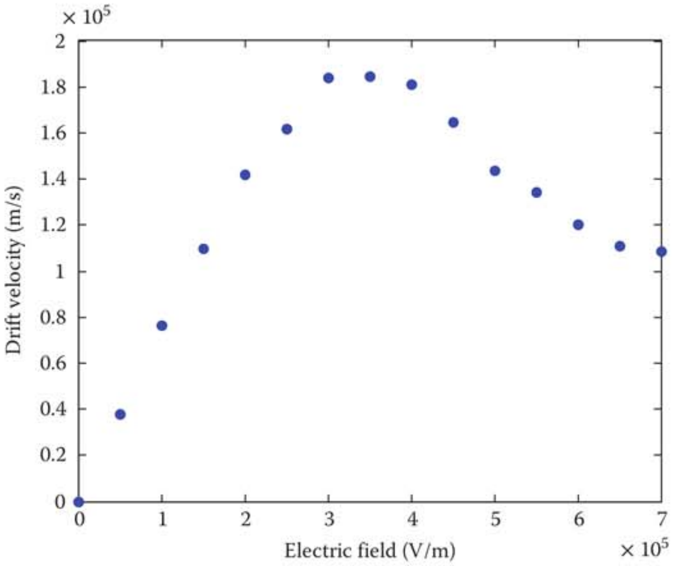


FIGURE 51.16 Steady-state drift velocity versus electric field.

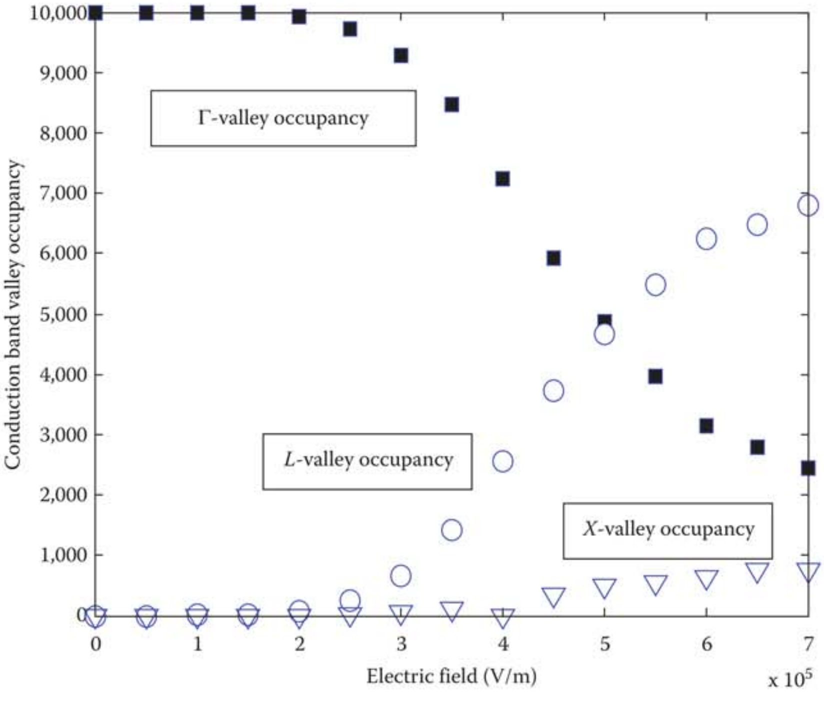


FIGURE 51.17 Different valley occupancy versus electric field.

for which a drifted Maxwellian form is evident. Since there is no field applied in the x-direction, we see that the average wavevector in the x-direction is 0. Due to the application of the field in the y-direction, there is a finite positive shift in the y-component of the velocity, which is yet another signature for the displaced Maxwellian form of the energy distribution in the bottom histogram.

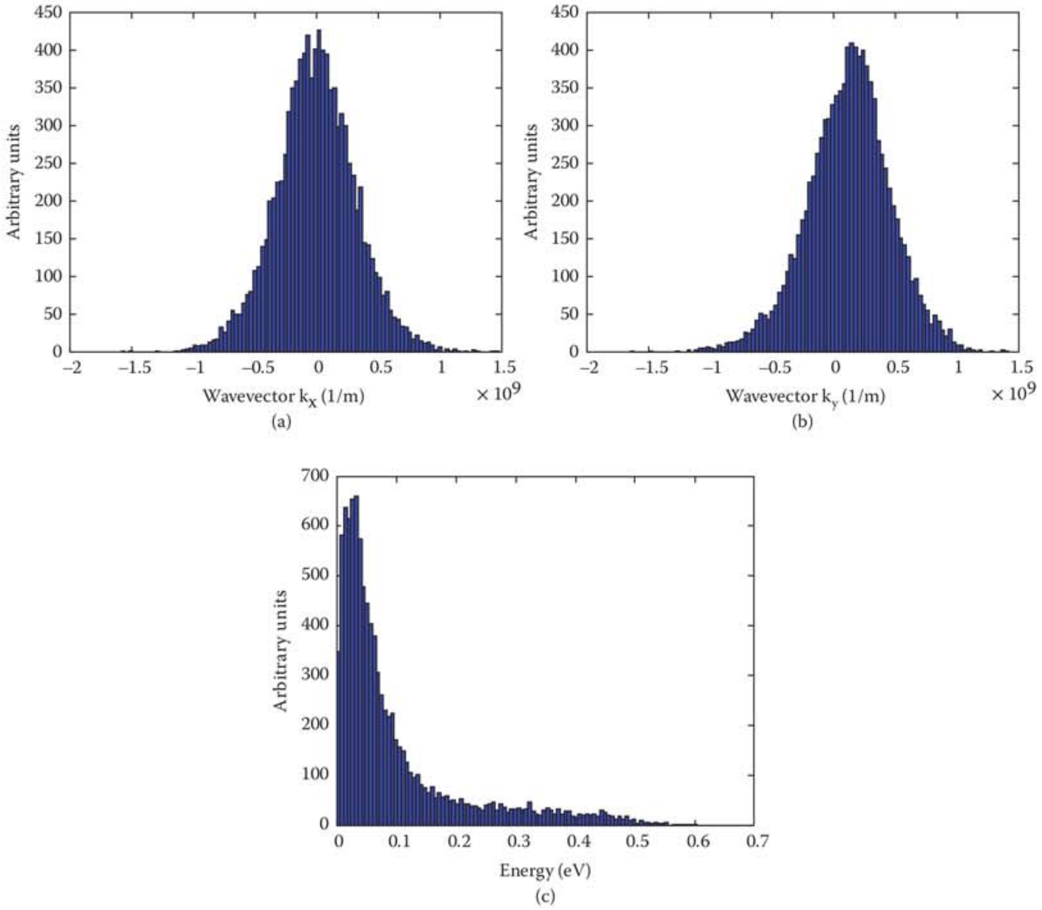


FIGURE 51.18 (a) Histogram of the x-component of the wavevector. (b) Histogram of the y-component of the wavevector. (c) Histogram of the carrier energy. Applied electric field is 7 kV/cm.

51.3 Particle-Based Device Simulation

In Section 51.2, we introduced the numerical solution of the BTE using Monte Carlo method. Within a device, both the transport kernel and the field solver are coupled to each other. In Poisson's equation, the field associated with the potential is the driving force accelerating particles in the Monte Carlo phase, for example, while the distribution of mobile (both electrons and holes) and fixed charges (e.g., donors and acceptors) provides the source of the electric field. Below, we give a brief description of the Monte Carlo particle-based device simulator development.

Within the particle-based EMC method, with its time-marching algorithm, Poisson's equation may be decoupled from the BTE over a suitably small time step (typically less than the inverse plasma frequency corresponding to the highest carrier density in the device). Over this time interval, carriers accelerate according to the frozen field profile from the previous time-step solution of Poisson's equation, and then Poisson's equation is solved at the end of the time interval with the frozen configuration of charges arising from the Monte Carlo phase [6,26]. Note that Poisson's equation is solved on a mesh, whereas the solution of charge motion using EMC occurs over a continuous range of coordinate space in terms of the particle position. Therefore, a particle-mesh (PM) coupling is needed for both the charge assignment and the force interpolation. The PM coupling is broken into four steps: (1) assign particle charge to the mesh, (2) solve the Poisson equation on the mesh, (3) calculate the mesh-defined forces, and (4) interpolate to find forces on the particle.

The motion in real space of particles under the influence of electric fields is somewhat more complicated due to the band structure. The velocity of a particle in real space is related to the E-$\mathbf{k}$ dispersion relation defining the bandstructure as

$$\begin{aligned} \mathbf{v}(t) &= \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{1}{\hbar}\nabla_k E(\mathbf{k}(t)) \\ \frac{\mathrm{d}\mathbf{k}}{\mathrm{d}t} &= \frac{q\mathbf{E}(\mathbf{r})}{\hbar} \end{aligned}, \tag{51.12}$$

where the rate of change of the crystal momentum is related to the local electric field acting on the particle through the acceleration theorem expressed by the second equation. In turn, the change in crystal momentum, $\mathbf{k}(t)$, is related to the velocity through the gradient of E with respect to $\mathbf{k}$. If one has to use the full bandstructure of the semiconductor, then integration of these equations to find $\mathbf{r}(t)$ is only possible numerically, using, for example, a Runge–Kutta algorithm [27].

To simulate the steady-state behavior of a device, the system must be initialized in some initial condition, with the desired potentials applied to the contacts, and then the simulation proceeds in a time stepping manner until steady state is reached. This process may take several picoseconds of simulation time, and consequently several thousand time-steps based on the usual time increments required for stability. Clearly, the closer the initial state of the system is to the steady-state solution, the quicker the convergence. If one is, for example, simulating the first bias point for a transistor simulation, and has no *a priori* knowledge of the solution, a common starting point for the initial guess is to start out with charge neutrality, i.e., to assign particles randomly according to the doping profile in the device and based on the super-particle charge assignment of the particles, so that initially the system is charge neutral on the average. (For 2D device simulation, one should keep in mind that each particle actually represents a rod of charge into the third dimension.) Subsequent simulations at the same device at different bias conditions can use the steady-state solution at the previous bias point as a good initial guess. After assigning charges randomly in the device structure, charge is then assigned to each mesh point using the nearest-grid-point (NGP), cloud-in-cell (CIC), or nearest-element-cell (NEC) PM coupling methods, and Poisson's equation is solved. The forces are then interpolated on the grid, and particles are accelerated over the next time-step. A flowchart of a typical Monte Carlo device simulation is shown in Figure 51.19.

As the simulation evolves, charge will flow in and out of the contacts, and depletion regions internal to the device will form until steady state is reached. The charge passing through the contacts at each time step can be tabulated, and a plot of the cumulative charge as a function of time gives the steady-state current. Figure 51.20 shows the particle distribution in three dimension of a metal-semiconductor field-effect transistor (MESFET), where the dots indicate the individual simulated particles for two different gate biases. Here, the heavily doped MESFET region (shown by the inner box) is surrounded by semi-insulating GaAs forming the rest of the simulation domain. Figure 51.20a corresponds to no net gate bias (i.e., the gate is positively biased to overcome the built-in potential of the Schottky contact), while Figure 51.20b corresponds to a net negative bias applied to the gate, such that the channel is close to pinch-off. One can see the evident depletion of carriers under the gate under the latter conditions.

51.3.1 Calculation of the Current

The device output current can be determined using two different yet consistent methods. First, by keeping track of the charges entering and exiting each terminal/contact, the net number of charges over a period of the simulation can be used to calculate the terminal current. The net charge crossing a terminal boundary

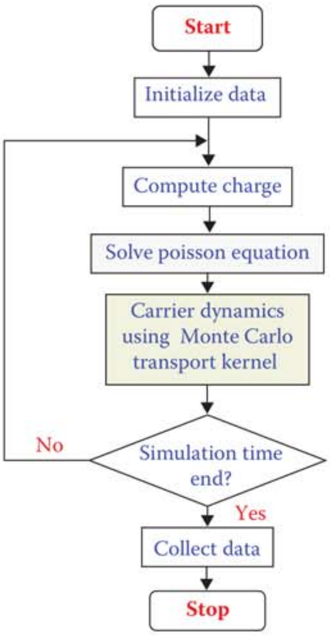


FIGURE 51.19 Flowchart of a typical particle-based device simulation.

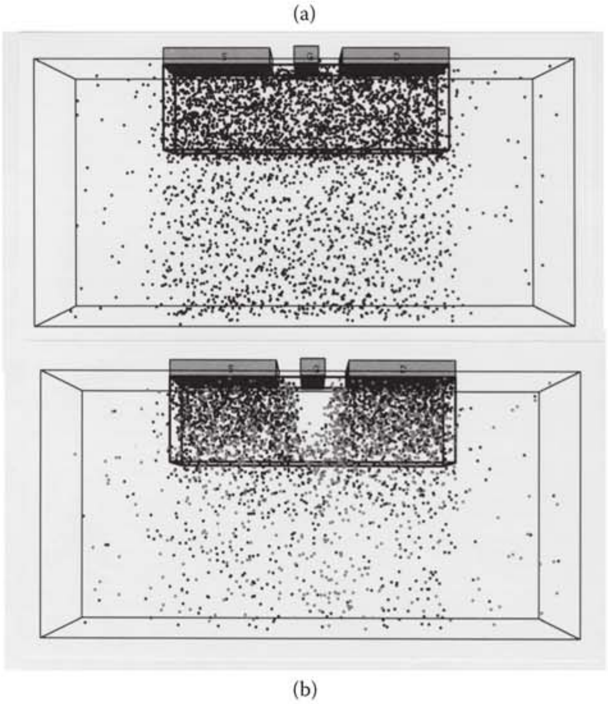


FIGURE 51.20 Example of the particle distribution in an MESFET structure simulated in three dimension using an EMC approach. (a) The device with zero gate voltage applied and (b) the device with a negative gate voltage applied, close to pinch-off.

is determined by

$$Q(t) = e(n_{\text{abs}}(t) - n_{\text{injec}}(t)) + \varepsilon \int E_y(x, t)\text{d}y, \tag{51.13}$$

where n_{abs} is the number of particles that are absorbed by the contact (exit), n_{injec} the number of particles that have been injected at the contact, and E_y the vertical field at the contact. The second term in Equation 51.13 on the right-hand side is used to account for the displacement current due to the changing field at the contact. Equation 51.13 assumes the contact is at the top of the device and that the fields in the x- and z- directions are negligible. The charge e in Equation 51.13 should be multiplied by the particle charge if it is not unity. The slope of $Q(t)$versus time gives a measure of the terminal current. In steady state, the current can be found by

$$I = \frac{\text{d}Q(t)}{\text{d}t} = \frac{e(n_{\text{net}})}{\Delta t}, \tag{51.14}$$

where $n_{\text{net}} = n_{\text{abs}} - n_{\text{injec}}$ is the net number of particles exiting the contact over a fixed period of time Δt.

In a second method, the sum of the electron velocities in a portion of the channel region of the device is used to calculate the current. The electron current density through a cross-section of the device is given by

$$J = env_{\text{d}}, \tag{51.15}$$

where v_{d} is the average electron drift velocity and n the carrier concentration. If there are a total of N particles in a differential volume, $\text{d}V = \text{d}L \cdot \text{d}A$,the current found by integrating Equation 51.15 over the cross-sectional area, $\text{d}A$, is

$$I = \frac{eNv_{\text{d}}}{\text{d}L}, \text{ or } I = \frac{e}{\text{d}L} \sum_{i=1}^{N} v_x(i), \tag{51.16}$$

where $v_x(i)$ is the velocity along the channel of the ith electron. The device is divided into several sections along the x-axis, and the number of electrons and their corresponding velocity is added for each section after each free flight. The total x-velocity in each section is then averaged over several time-steps to determine the current for that section. Total device current can be determined from the average of several sections, which gives a much smoother result compared to counting terminal charges. By breaking the device into sections, individual currents can be compared to verify that there is conservation of particles (constant current) throughout the device. In addition, sections near the source and drain regions may have a high y-component in their velocity and should be excluded from the current calculations. Finally, by using several sections in the channel, the average energy and velocity of electrons along the channel can be observed to ensure the proper physical characteristics. The two methods for the calculation of the current are performed, as illustrated in Figure 51.17, on a 50-nm channel length MOSFET device.

Extrapolating the slope of the curve shown in Figure 51.21a, that represents the cumulative electron charge that enters/exits the source/drain contact, leads to source/drain current of 0.5205/0.5193 mA/μm. When compared with the results shown in Figure 51.21b, it is evident that both current measurement techniques discussed in this section give current values with relative error less than 1%.

51.3.2 Ohmic Contacts

Another issue that has to be addressed in particle-based simulations is the real space boundary conditions for the particle part of the simulation. Reflecting boundary conditions are usually imposed at the artificial

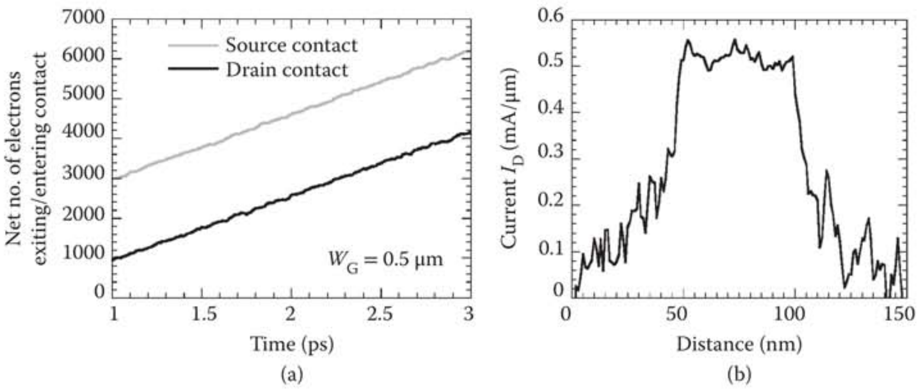


FIGURE 51.21 (a) Net charge entering/exiting the source/drain contact. (b) Average current along the channel. The gate length of the device being modeled equals 50 nm. $V_G = 1.4$ V and $V_D = 1$ V in these simulations. Width of the gate is $W = 0.5$ μm.

boundaries. As far as the ohmic contacts are concerned, they require more careful consideration because electrons crossing the source and drain contact regions contribute to the corresponding terminal current. Commonly employed models for the contacts include the following [28]:

- Electrons are injected at the opposite contact with the same energy and wavevector **k**. If the source and drain contacts are in the same plane, as in the case of MOSFET simulations, the sign of **k**, normal to the contact will change. This is an unphysical model, however [29].
- Electrons are injected at the opposite contact with a wavevector randomly selected based upon a thermal distribution. This is also an unphysical model.
- Contact regions are considered to be in thermal equilibrium. The total number of electrons in a small region near the contact are kept constant, with the number of electrons equal to the number of dopant ions in the region. This is a very good model most commonly employed in actual device simulations.
- Another method uses "reservoirs" of electrons adjacent to the contacts. Electrons naturally diffuse into the contacts from the reservoirs, which are not treated as part of the device during the solution of Poisson's equation. This approach gives results similar to the velocity weighted Maxwellian but at the expense of increased computational time due to the extra electrons simulated. It is an excellent model employed in few most sophisticated particle-based simulators.

There are also several possibilities for the choice of the distribution function—Maxwellian, displaced Maxwellian, and velocity-weighted Maxwellian [10].

51.3.3 Time-Step

As in the case of solving the drift-diffusion, hydrodynamic, or full Maxwell's equations, for a stable Monte Carlo device simulation, one has to choose the appropriate time step, Δt, and the spatial mesh size (Δx, Δy, and/or Δz). The time-step and the mesh size may correlate to each other in connection with the numerical stability. For example, as discussed in the context of solving drift-diffusion simulations, the time-step Δt must be related to the plasma frequency:

$$\omega_p = \frac{e^2 n}{\varepsilon_s m^*}, \tag{51.17}$$

where n is the carrier density. From the viewpoint of the stability criterion, Δt must be much smaller than the inverse plasma frequency. The highest carrier density specified in the device model is used to estimate

Δt. If the material is a multivalley semiconductor, the smallest effective mass to be experienced by the carriers must be used in Equation 51.17 as well. In the case of GaAs, with the doping of 5×10^{17} cm^{-3}, $\omega_{\mathrm{p}} \cong 5 \times 10^{13}$; hence, Δt must be smaller than 0.02 ps.

The mesh size for the spatial resolution of the potential is dictated by the charge variations. Hence, one has to choose the mesh size to be smaller than the smallest wavelength of the charge variations. The smallest distance is approximately equal to the Debye length, given as

$$\lambda_{\mathrm{D}} = \sqrt{\frac{\varepsilon_{\mathrm{s}} k_{\mathrm{B}} T}{e^2 n}}. \tag{51.18}$$

The highest carrier density specified in the model should be used to estimate λ_{D} from the stability criterion. The mesh size must be chosen to be smaller than the value given by Equation 51.18. In the case of GaAs, with the doping density of 5×10^{17} cm^{-3}, $\lambda_{\mathrm{D}} \cong 6$ nm.

Based on the discussion above, the time step (Δt) and the mesh size (Δx, Δy, and/or Δz) can be specified separately. However, the Δt chosen must be checked again by calculating the distance $l_{\max}$, defined as

$$l_{\max} = \mathbf{v}_{\max} \times \Delta t, \tag{51.19}$$

where $\mathbf{v}_{\max}$ is the maximum carrier velocity that can be approximated by the maximum group velocity of the electrons in the semiconductor (on the order of 10^8 cm/s). Therefore, the distance $l_{\max}$ is regarded as the maximum distance the carriers can propagate during Δt. The time step chosen must be small enough so that $l_{\max}$ is smaller than the spatial mesh size chosen using Equation 51.19. This is because large Δt chosen may cause substantial change in the charge distribution, while the field distribution in the simulation is only updated every Δt.

51.3.4 PM Coupling

As mentioned earlier, the position of charge as described by the EMC algorithm is continuous, whereas Poisson's equation is solved on a mesh, hence the charge associated with the individual particles must be mapped onto the field mesh in some manner. The charge assignment and force interpolation schemes usually employed in self-consistent Monte Carlo device simulations are the NGP, CIC, and NEC schemes [30]. In the NGP scheme, the particle position is mapped into the charge density at the closest grid point to a given particle. This has the advantage of simplicity, but leads to a noisy charge distribution, which may exacerbate numerical instability. Alternately, within the CIC scheme, a finite volume is associated with each particle spanning several cells in the mesh, and a fractional portion of the charge per particle is assigned to grid points according to the relative volume of the "cloud" occupying the cell corresponding to the grid point. This method has the advantage of smoothing the charge distribution due to the discrete charges of the particle-based method, but may result in an artificial "self-force" acting on the particle, particularly if an inhomogeneous mesh is used. For the case of inhomogeneous mesh or spatially varying dielectric permittivity, the NEC scheme is the best choice. Within the NEC scheme, the charge is moved at the center of the cell and then CIC-like charge assignment to the node points is performed. The PM coupling sequence is presented in Figure 51.22.

51.4 Representative Simulation Results

51.4.1 Bulk Monte Carlo Simulations of Different Materials

In most semiconductors, in order to properly simulate high field transport, it is necessary to consider more than one conduction band valley. To calculate the drift velocity along any direction, the effective mass along

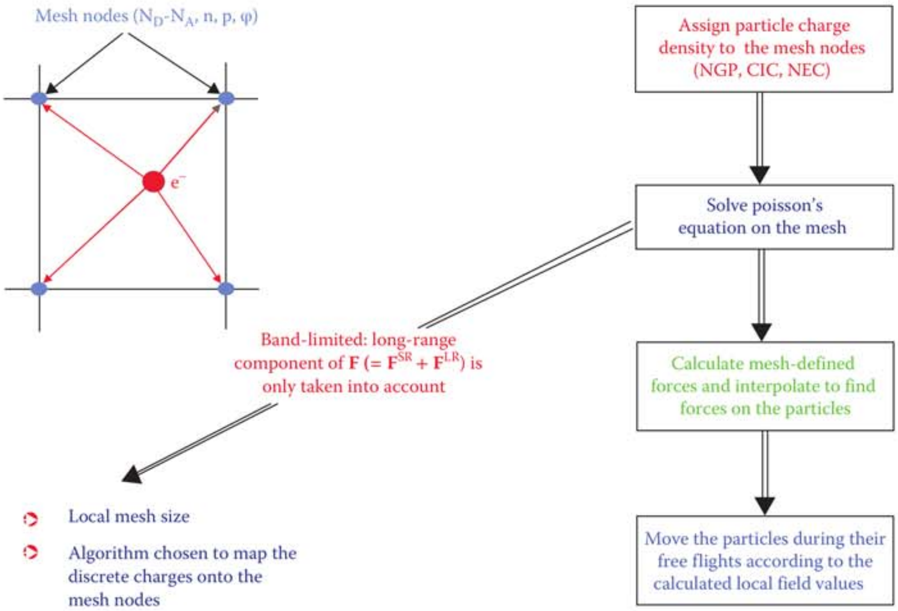


FIGURE 51.22 Particle-mesh coupling sequence. (SR, short range; LR, long range.)

that particular direction is required. In GaAs, for example, a subvalley of the L valley lies along the [111] direction. We know the effective masses along the transverse and longitudinal directions of the subvalley but we do not know the effective mass along the [100] direction. This makes it hard to calculate the drift velocity along the [100] direction.

Assume the Monte Carlo simulation is run on the x, y, z- coordinate system where the x-direction is [100], y-direction is [010], and z-direction is [001]. Let the three perpendicular directions that describe the subvalley be $[a_1, b_1, c_1]$, $[a_2, b_2, c_2]$, and $[a_3, b_3, c_3]$. The electrons in the Monte Carlo simulation will be drifted according to the x, y, z coordinate system so there will be k_x the wavevector along [100], k_y the wavevector along [010], and k_z the wavevector along [001]. The drift velocity of the electron is first calculated along the three directions that describe the subvalley and along which we know the effective masses.

$$v_d, [[a]_1 b_1 c_1] = \frac{\hbar k_{[[a]_1 b_1 c_1]}}{m_1 (1 + 2\alpha E)} v_d, [[a]_1 b_1 c_1] = \frac{\hbar k_{[[a]_1 b_1 c_1]}}{m_1 (1 + 2\alpha E)} \tag{51.20}$$

Here m_1 is the effective mass of the electron along $[a_1, b_1, c_1]$. The final expression of drift velocity along that direction is then calculated by calculating $k_{[[a]_1 b_1 c_1]} k_{[[a]_1 b_1 c_1]}$ using a simple transformation of coordinates to give [31]

$$k_{[[a]_1 b_1 c_1]} = \frac{k_x a_1}{\sqrt{\left(a_1^2 + b_1^2 + c_1^2\right)}} + \frac{k_y b_1}{\sqrt{\left(a_1^2 + b_1^2 + c_1^2\right)}} + \frac{k_z a_1}{\sqrt{\left(a_1^2 + b_1^2 + c_1^2\right)}}. \tag{51.21}$$

Using similar methods we get v_d, $\left[[a]_2 b_2 c_2\right]$ and v_d, $\left[[a]_3 b_3 c_3\right]$. The coordinates system is then transformed once again back to the x,y,z coordinate system to get the drift velocities along x,y, and z.

$$v_x = \frac{a_1 v_{\mathrm{d},[[a]_1 b_1 c_1]}}{\sqrt{\left(a_1^2 + b_1^2 + c_1^2\right)}} + \frac{a_2 v_{\mathrm{d},[[a]_2 b_2 c_2]}}{\sqrt{\left(a_2^2 + b_2^2 + c_2^2\right)}} + \frac{a_3 v_{\mathrm{d},[[a]_3 b_3 c_3]}}{\sqrt{\left(a_3^2 + b_3^2 + c_3^2\right)}}$$

$$v_y = \frac{b_1 v_{\mathrm{d},[[a]_1 b_1 c_1]}}{\sqrt{\left(a_1^2 + b_1^2 + c_1^2\right)}} + \frac{b_2 v_{\mathrm{d},[[a]_2 b_2 c_2]}}{\sqrt{\left(a_2^2 + b_2^2 + c_2^2\right)}} + \frac{b_3 v_{\mathrm{d},[[a]_3 b_3 c_3]}}{\sqrt{\left(a_3^2 + b_3^2 + c_3^2\right)}} \tag{51.22}$$

$$v_x = \frac{c_1 v_{\mathrm{d},[[a]_1 b_1 c_1]}}{\sqrt{\left(a_1^2 + b_1^2 + c_1^2\right)}} + \frac{c_2 v_{\mathrm{d},[[a]_2 b_2 c_2]}}{\sqrt{\left(a_2^2 + b_2^2 + c_2^2\right)}} + \frac{c_3 v_{\mathrm{d},[[a]_3 b_3 c_3]}}{\sqrt{\left(a_3^2 + b_3^2 + c_3^2\right)}}$$

Also, as the three directions are mutually perpendicular we have

$$\begin{aligned} a_1 a_2 + b_1 b_2 + c_1 c_2 &= 0 \\ a_1 a_3 + b_1 b_3 + c_1 c_3 &= 0 \\ a_3 a_2 + b_3 b_2 + c_3 c_2 &= 0. \end{aligned} \tag{51.23}$$

For N electrons in the simulation, an average drift velocity is then calculated as

$$\langle v_x \rangle = \frac{1}{N} \sum_{i=1}^{N} v_{x,i}, \quad \left\langle v_y \right\rangle = \frac{1}{N} \sum_{i=1}^{N} v_{y,i}, \quad \langle v_z \rangle = \frac{1}{N} \sum_{i=1}^{N} v_{z,i}, \tag{51.24}$$

where each drift velocity depends on the subvalley the electron is currently in. The mobility is calculated as the slope of the velocity versus field curve for low fields. The range of "low fields" varies from material to material, but in all cases, it is the range of fields for which the velocity linearly varies with the field.

The results presented in Figure 51.23 shows the difference in velocity of electrons for different materials. For GaN, the peak velocity of 3×10^7 cm/s occurs at around 150 kV/cm so that in the electric field range shown in Figure 51.19 it is clear that for electric fields below 30 kV/cm, GaN carriers in the bulk GaN materials are still in the low-field regime.

As the drift velocities saturate at much smaller electric fields in Si and Ge, they are plotted on a separate graph (see Figure 51.24) [32].

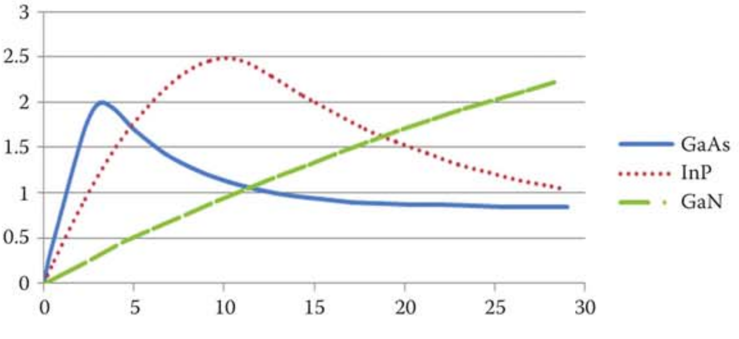


FIGURE 51.23 Drift velocity (10^7 cm/s) versus electric field (kV/cm).

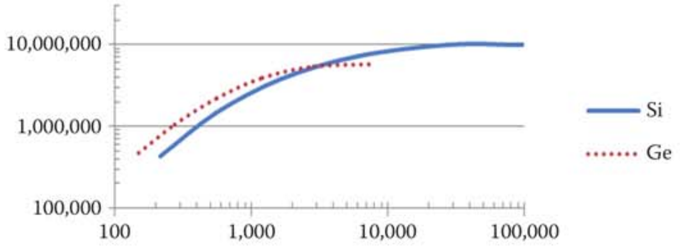


FIGURE 51.24 Electron drift velocity (cm/s) versus electric field (V/cm) in Si and Ge along the principal crystallographic directions.

51.4.2 Modeling Thermal Effects in SOI Devices Using Particle-Based Device Simulator

There are many applications of the Monte Carlo approach to device simulations that can be found in the literature. Here, we show the example of Monte Carlo device simulation to self-heating which requires coupling of the Monte Carlo device solver self-consistently with a solver for the phonon bath. Self-heating (despite the fact that supply voltages are reduced when reducing the size of the channel of nanoscale transistors) is still a problem, since there are sufficiently large electric fields in the smallest devices that can accelerate the carriers to average energies of 0.5–0.8 eV. These hot carriers then interact with the lattice and transfer their energy to the phonon bath (both acoustic and optical), thus heating the lattice and increasing the lattice vibrations (scattering). This leads to a negative feedback on the carrier drift velocity or mobility (if transport is scattering dominated). Thus, we may conclude that when transistors are in the ON-state, there is a strong interaction of the electrons with the lattice on a very short time scale (in particular with the optical phonons). The optical phonons in a longer time scale decay into acoustic phonons through a harmonic multiple phonon processes, thus removing the heat from the hot spot. Therefore, to properly treat the operation of nanoscale devices it is necessary to, at least, consider the electrons within the BTE picture and, in the lowest approximation, via energy balance equations including the acoustic and the optical phonon baths separately. This is what the Arizona State University (ASU) group has implemented couple of years ago [33]. Our ultimate goal is to have coupled the developed phonon BTE solver [34] with the electron BTE solver in a self-consistent manner.

51.4.2.1 Theoretical Model and Computation Details

A schematic description of our 2D/3D electrothermal simulator is shown in Figure 51.25. It is important to note that the EMC transport kernel for the electrons is in the loop for the solution of the energy balance equations for the acoustic and optical phonon temperatures. The criterion for convergence of the whole self-consistent loop (measured in terms of Gummel cycles) is convergence in the current to the third significant digit.

Between iterations, there are variables that are being exchanged between the electron and phonon solvers. Namely, after running the EMC for the electrons in the device for about 5–10 ps for 2D device analysis and 10 ps for the 3D nanowire simulations, the electron transport solver passes to the phonon energy balance solver the information about the spatial variation of the electron density, spatial variation of the average drift velocity, and spatial variation of the average electron temperature, where an assumption is made that the drift energy is much smaller than the thermal energy. These variables are then used in the solution of the energy balance equations for the acoustic and optical phonons to get updated values for the spatial variation of the optical phonon and acoustic phonon temperatures, which are then used as inputs, in the choice of the scattering table for the electrons (proper scattering table is chosen with the acoustic and optical phonon temperatures at the NGP). Then, when the proper table is selected, based on the energy of the electron and a selection of a random number, the scattering mechanism is selected.

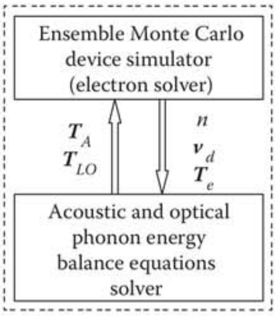


FIGURE 51.25 Exchange of variables between the two kernels. The electron density (n), electron drift velocity (v_d), and electron temperature (T_e), obtained from the electron solver, are input variables for the phonon kernel. The output variables of the phonon solver are acoustic and optical phonon temperature profiles (T_A and T_{LO}, respectively). They enter in the beginning of the free-flight-scattering phase through phonon temperature-dependent scattering tables.

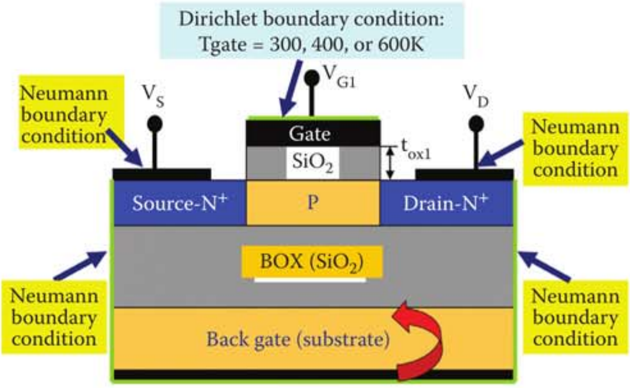


FIGURE 51.26 Schematic of the simulated device structure which illustrates the boundary conditions used in some of the simulations presented. Note also that the boundary conditions at the bottom of the substrate can be mapped as a boundary condition at the bottom of the BOX.

The question of the proper boundary conditions for the electronic part of the problem is rather clear and has been discussed in many papers and texts in the literature [35]. The problem in properly specifying the phonon boundary conditions is the selection of acoustic and optical phonon temperatures either at the artificial boundaries or at the contacts. Typically, acoustic phonon temperature is equated with the lattice temperature. To better understand the choice of boundary conditions we have considered in various works [15], we want to point out that lattice temperature is analogous to electrostatic potential and heat flux is analogous to current. As it is well known, when considering electrical behavior of the device, at least one node within the structure has to have Dirichlet boundary conditions specified (specify the potential for Poisson). Analogously, for the lattice temperature, we need at least one node that is a thermal contact and whose temperature is set to 300 K. As illustrated in Figure 51.26 for the case of FD-SOI devices, the bottom of the BOX is assumed to be at thermal equilibrium. Also, the gate contact is assumed to be at thermal equilibrium, but not necessarily at 300 K. For example, in one set of simulations, we vary the temperature on the gate to be 300, 400, and 600 K. Now comes the question: What happens to the source and drain contacts and the side artificial boundaries? Specifying Dirichlet boundary conditions at the ohmic contacts is not accurate from a standpoint that there is current flowing through the contacts, and since the contacts have finite resistance, there will be Joule heating (so the problem becomes unconstrained). The best solution to this, as we have done in several studies, is to extend the metal contact to become part of the simulation

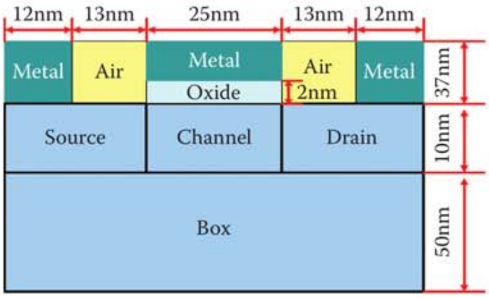


FIGURE 51.27 Device structure with extended domain. The geometrical dimensions are for the simulated 25-nm channel length FD-SOI nMOSFET.

domain and to apply at the very ends of the contact isothermal boundary conditions (see Figure 51.27). With the top and bottom specified, we have vertical transport of heat through the structure.

The next question is, is there lateral transfer of heat across the artificial boundaries? Here we can consider two cases: case when the neighboring device is ON, in which case Neumann boundary conditions are appropriate (derivative of the temperature), and case when the neighboring device is OFF, in which case Dirichlet boundary conditions are appropriate. In the case when Neumann boundary conditions are applied across the artificial boundaries, the heat transport remains vertical, but for the case when Dirichlet boundary conditions are applied across the boundary, the heat transport has both vertical and horizontal component. The case of Neumann boundary conditions can happen in analog circuits such as current mirror in which both transistors are simultaneously ON, whereas the case of Dirichlet boundary conditions occurs in digital circuits in which most of the time the transistors are OFF.

Yet another issue that deserves attention in getting physically correct results is the proper choice of the thermal conductivity for thin Si slabs and for nanowires. Asheghi and coworkers [36] and Li Shi and coworkers [37] have demonstrated via experimental measurements that thermal conductivities of thin Si films and Si nanowires, respectively, strongly depend of the geometry as for smaller geometries phonon boundary scattering can reduce the thermal conductivity of the Si film or nanowire by a factor of 10 or more from its bulk value. Moreover, thermal conductivity is temperature-dependent quantity. We made extensive efforts, using the theory of Sondhaimer for conductivity of metals [38], to arrive at an empirical formula that simultaneously described the thickness and temperature dependence of the thermal conductivity [39]. Our empirical expression perfectly matched the experimental data of Asheghi and coworkers [16]. In the model, it was assumed that phonon boundary scattering was perfectly diffusive.

51.4.2.2 Simulation Results

To validate our premise that self-heating effects become smaller as we scale devices into the nanometer regime (due to more pronounced velocity overshoot effects and thinner BOX layer that leads to smaller thermal resistance), in Table 51.3 we present the parameters of the devices being simulated. Simulation results for the lattice temperature profile in the Si film for devices with channel lengths from 25 to 180 nm are shown in Figure 51.28. We clearly see that the hot spot is in the channel for 180-nm channel length device and moves into the drain contact for 25-nm channel length device. The current degradation for isothermal temperatures on the gate of 300, 400, and 600 K are shown in Figure 51.29. These are the worst case scenario results as we use Neumann boundary conditions at the side boundaries.

In obtaining the results presented in Figures 51.28 and 51.29, constant thermal conductivity value of 13 W/m-K was used in the simulations [40]. The use of constant thermal conductivity model with value of 13 W/m-K leads to underestimation (25 nm channel length) or overestimation (180 nm channel length) of the average maximum temperature of the lattice in the channel region of the device.

TABLE 51.3 Parameters for Various Simulated Device Technology Nodes (Constant Field Scaling)

L(nm)	t_{ox} (nm)	t_{si} (nm)	t_{box} (nm)	N_{ch} (cm^{-3})	$V_{GS} = V_{DS}$ (V)	I_D (mA/ µm)
25	2	10	50	1×10^{18}	1.2	1.82
45	2	18	60	1×10^{18}	1.2	1.41
60	2	24	80	1×10^{18}	1.2	1.14
80	2	32	100	1×10^{17}	1.5	1.78
90	2	36	120	1×10^{17}	1.5	1.67
100	2	40	140	1×10^{17}	1.5	1.57
120	3	48	160	1×10^{17}	1.8	1.37
140	3	56	180	1×10^{17}	1.8	1.23
180	3	72	200	1×10^{17}	1.8	1.03

Note: *Parameters of the simulated structure given in Figure 51.26 are L, gate length;* t_{ox}, *gate oxide thickness;* t_{Si}, *active Si-layer thickness;* t_{box}, *BOX thickness;* N_{ch}, *channel doping concentration;* I_D, *isothermal current value (300 K).*

51.4.2.3 Summary of Results

In this section, we have presented modeling of self-heating effects in FD-SOI devices. Larger velocity overshoot and smaller BOX thickness lead to smaller degradation due to self-heating effects in 25-nm channel length FD-SOI devices. The amount of self-heating significantly depends upon the magnitude of the thermal conductivity used. Bulk values are inadequate and proper thickness and temperature dependence of

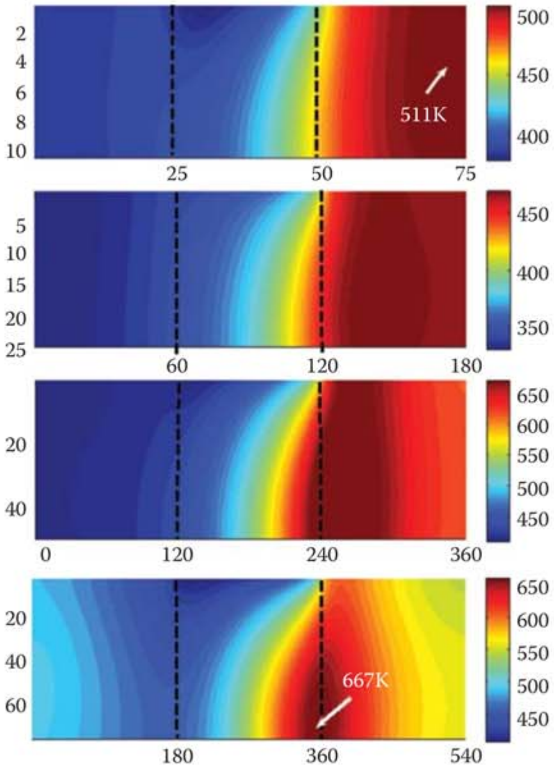


FIGURE 51.28 Lattice temperature profile in the active Si-layer for the FD-SOI devices from Table 51.2 ranging from 25 nm (top) to 180 nm (bottom) channel length. Shown here is only the thin Si film of the SOI device, namely the active area of the device.

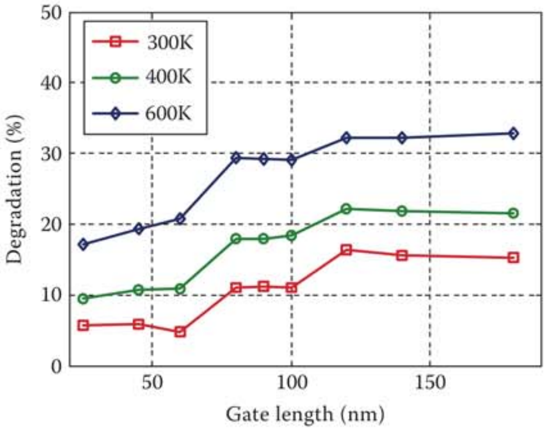


FIGURE 51.29 Current degradation (abs(I_{th}−I_o)/I_o, where I_{th} is the on current when the self-heating model is incorporated and I_o is current in the isothermal case) versus technology generation for the FD-SOI devices from Table 51.2 ranging from 25- to 180-nm channel length. Different curves correspond to different temperatures on the gate electrode (bottom set of results: T_{gate} = 300 K, middle set of results: T_{gate} = 400 K and top set of results: T_{gate} = 600 K).

the thermal conductivity must be taken into consideration. The choice of the BOX material makes significant impact on the amount of observed self-heating. Both diamond and Si are good materials to be used as BOX. Diamond, when compared to aluminum nitride (AlN) is a better heat spreader.

51.5 Outlook

The drift-diffusion solvers are applicable in situations in which the bias conditions and the device geometry are such that electric fields are relatively low and velocity saturation model is applicable [41]. Situations in which drift-diffusion models are applicable are the Si-based power MOSFET devices, bipolar junction transistors, light-emitting diodes (that are used more and more in solid-state lightning), and crystalline solar cells, just to name a few. In some of these devices, such as power transistors and light-emitting diodes (LEDs) used for solid-state lighting, it is of paramount importance to incorporate self-heating models within the drift-diffusion framework [42]. The accuracy of simple heating models in conjunction with drift-diffusion models is to some degree questionable so that it is in many circumstances justifiable to use particle-based device simulators with more exact self-heating models [43].

On the other hand, the hydrodynamic models do not suffer from the limitations of the drift-diffusion approaches and the incorporation of the additional energy balance equation allows one to include velocity overshoot in the model. Velocity overshoot and nonstationary transport are key features of conventional MOSFET devices with gate lengths of 200 nm and below [44]. However, the magnitude of the velocity overshoot observed via simulations depends strongly upon the choice of the energy relaxation time, mostly in sub-100-nm channel length devices. This, in turn, affects the magnitude of the drain current. The reason for such drastic differences in the results when different energy relaxation times are used is due to the fact that the energy relaxation time is a material-dependent as well as device geometry-dependent parameter. So, to calculate better estimates for the velocity overshoot, higher moments of the BTE are needed [45]. These, in turn involve parameters that are more and more ambiguous on the expense of increased computational cost and when the computational cost of hydrodynamic models exceeds one of particle-based device simulators, there is no point in using moment methods. In these circumstances, the direct solution of the BTE via the Monte Carlo method becomes a method of choice. Thus, we might conclude that it is advisable to use particle-based device simulators when nanoscale devices are concerned [33].

But how far down in the scaling can we go? Particle-based device simulators capture on an equal footing both ballistic and diffusive transport, so if the ballisticity factor in the device increases [46], there is no problem that ballistic transport is effectively captured with particle-based device simulators [47]. Quantum-mechanical size quantization effects can also be captured by solving in slices the corresponding one-dimensional (1D) or 2D Schrödinger equation if one is concerned with conventional or FD-SOI MOSFETs, or nanowire transistors, respectively. What cannot be captured with particle-based device simulators is if there are local strains and stresses in the ultra-nanoscale devices, but that can also be cured via coupling of Monte Carlo device simulators with atomistic models [48,49] for bandstructure calculation.

In summary, Monte Carlo device simulators are a powerful tool for modeling devices ranging from the nanoscale regime to the microscale regime. What cannot be modeled with particle-based device simulators are resonant tunneling diodes in which quantum interference effects dominate the device behavior [50]. Efforts have been made along this direction as well [51], but the inclusion of the quantum-mechanical phase alongside with well-defined particle trajectory still remains open field of research.

References

1. ITRS International Technology Roadmap for Semiconductors. Available at: http://www.itrs.org.
2. G. Moore. Progress in digital integrated electronics. *IEDM Tech. Digest*, 11–13, 1975.
3. R. Dennard, F. H. Gaensslen, H. N. Yu, L. Rideout, E. Bassous, and A. R. LeBlanc. Design of Ion-Implanted MOSFET's with very small physical dimensions. *IEEE J. Solid State Circ.*, Vol. 9 (5), 256–268, 1974.
4. E. Fayneh, M. Yuffe, E. Knoll, M. Zelikson, M. Abozaed, Y. Talker, Z. Shmuely, and S. A. Rahme. 14 nm 6th-generation core processor SoC with low power consumption and improved performance. *IEEE International Solid-State Circuits Conference*, San Francisco, pp. 72–74, 2016.
5. T. Song, W. Rim, S. Park, Y. Kim, J. Jung, G. Yang, S. Baek, J. Choi, B. Kwon, Y. Lee, S. Kim, G. Kim, H. S. Won, J. H. Ku, S. S. Paak, E. Jung, S. S. Park, and K. Kim. A 10 nm FinFET 128 Mb SRAM with assist adjustment system for power, performance, and area optimization. *IEEE International Solid-State Circuits Conference*, San Francisco, pp. 306–309, 2016.
6. H.-Y. Chen, F.-C. Chen, Y.-L. Chan, K.-N. Yang, F.-L. Yang, and C. Hu. Method of fabricating a necked FINFET device. US 7122412 B2, 2006.
7. L. Geppert. A quantum leap for photonics. *IEEE Spectr.*, 28–33, April 9, 2004.
8. W. Zhu, J. P. Han, and T. P. Ma. Mobility measurement and degradation mechanisms of MOSFETs made with ultrathin high-k dielectrics. *IEEE Trans. Electron Dev.*, Vol. 51, 98–105, 2004.
9. J. Welser, J. L. Hoyt, and J. F. Gibbons. Temperature and scaling behavior of strained-Si N-MOSFET's. *IEDM Tech. Dig.*, 1000–1002, 1992.
10. R. Oberhuber, G. Zandler, and P. Vogl. Subband structure and mobility of two-dimensional holes in strained Si/SiGe MOSFET's. *Phys. Rev. B.*, Vol. 58, 9941–9948, 1998.
11. R. Chau, J. Kavalieros, B. Doyle, A. Murthy, N. Paulsen, D. Lionberger, D. Barlage, R. Arghavani, B. Roberts, and M. Doczy. A 50 nm depleted-substrate CMOS transistor (DST). *IEDM Techn. Dig.*, 29.1.1–29.1.4, 2001.
12. Y.-K. Choi, K. Asano, N. Lindert, V. Subramanian, T.-J King., J. Bokor, and C. Hu. Ultrathin-body SOI MOSFET for deep-sub-tenth micron era. *IEEE Electron Dev. Lett.*, Vol. 21 (5), 254–255, 2000.
13. D. Vasileska and S. M. Goodnick. Computational electronics. *Mater. Sci. Eng. R Rev. J.*, Vol. R38, 181–236, 2002.
14. C. Jacoboni and L. Reggiani. The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials. *Rev. Mod. Phys.*, Vol. 55, 645–705, 1983.
15. C. Jacoboni and P. Lugli. *The Monte Carlo Method for Semiconductor Device Simulation*. Vienna, VA: Springer-Verlag, 1989.
16. K. Hess. *Monte Carlo Device Simulation: Full Band and Beyond*. Boston, MA: Kluwer Academic Publishing, 1991.
17. M. H. Kalos and P. A. Whitlock. *Monte Carlo Methods*. New York, NY: Wiley, 1986.

18. D. K. Ferry. *Semiconductors*. New York, NY: Macmillan, 1991.
19. D. Vasileska, S. M. Goodnick, and G. Klimeck. *Computational Electronics: Semiclassical and Quantum Transport Modeling*. New York, NY: Taylor & Francis, 2010.
20. D. K. Ferry. *Quantum Mechanics: An Introduction for Device Physicists and Electrical Engineers*. New York, NY: Taylor & Francis, 2001.
21. H. D. Rees. Calculation of distribution functions by exploiting the stability of the steady state. *J. Phys. Chem. Solids*, Vol. 30, 643, 1969.
22. R. M. Yorston. Free-flight time generation in the Monte Carlo simulation of carrier transport in semiconductors. *J. Comp. Phys*., Vol. 64, 177, 1986.
23. M. Lundstrom. *Fundamentals of Carrier Transport*. London: Cambridge University Press, 2000.
24. B. K. Ridley. *Quantum Processes in Semiconductors*. Oxford: Oxford University Press, 1982.
25. D. Vasileska. https://nanohub.org/groups/dragica_vasileska/semiclassical.
26. K. Tomizawa. *Numerical Simulation of Submicron Semiconductor Devices*. Boston, MA: Artech House, 1993.
27. M. V. Fischetti and S. E. Laux. Band structure, deformation potentials and carrier mobility in strained Si, Ge and SiGe alloys. *J. Appl. Phys*., Vol. 80, 2234–2252, 1996.
28. T. Gonzalez and D. Pardo. Physical models of ohmic contact for Monte Carlo device simulation. *Solid State Electron*., Vol. 39, 555–562, 1996.
29. P. A. Blakey, S. S. Cherensky, and P. Sumer. *Physics of Submicron Structures*. New York, NY: Plenum Press, 1984.
30. S. E. Laux. On particle-mesh coupling in Monte Carlo device simulation. *IEEE Trans. Comp. Aided Des. Int. Circ. Sys*., Vol. 15, 1266, 1996.
31. R. Hathwar. Generalized Monte Carlo tool for investigating low-field and high field properties of materials using non-parabolic band structure model. MS Thesis, Arizona State University, June 2011.
32. S. S. Qazi. Electrical and thermal transport in alternative device technologies. MS Thesis, Arizona State University, November 2013.
33. K. Raleva, D. Vasileska, S. M. Goodnick, and M. Nedjalkov. Modeling thermal effects in nanodevices. *IEEE Trans. Electron Dev.*, Vol. 55 (6), 1306, 2008.
34. A. R. Shaik. Multi-scale study of heat transfer using Monte Carlo technique for phonon transport. MS Thesis, Arizona State University, April 2016.
35. D. Vasileska, K. Raleva, and S. M. Goodnick. Self-heating effects in nanoscale FD SOI devices: The role of the substrate, boundary conditions at various interfaces, and the dielectric material type for the box. *IEEE Trans. Electron Dev.*, Vol. 56, 3064, 2009.
36. W. Liu and M. Asheghi. Phonon-boundary scattering in ultrathin single-crystal silicon layers. *Appl. Phys. Lett.*, Vol. 84, 3819, 2004.
37. D. Li, Y. Wu, P. Kim, L. Shi, P. Yang, and A. Majumdar. Thermal conductivity of individual silicon nanowires. *Appl. Phys. Lett*., Vol. 83, 2934, 2003.
38. E. H. Sondheimer. The mean free path of electrons in metals. *Adv. Phys.*, Vol. 1, 1, 1952, reprinted in *Adv. Phys.*, 50, 499, 2001.
39. D. Vasileska, K. Raleva, and S. M. Goodnick. Electrothermal studies of FD SOI devices that utilize a new theoretical model for the temperature and thickness dependence of the thermal conductivity. *IEEE Transactions on Electron Devices*, Vol. 57, 726–728, 2010.
40. K. Raleva and D. Vasileska. The importance of thermal conductivity modeling for simulations of self-heating effects in FD SOI devices. *J. Comput. Elec*., Vol. 12 (4), 601–610, 2013.
41. D. Vasileska and S. M. Goodnick, Editors. *Nano-Electronic Devices: Semiclassical and Quantum Transport Modeling.* Vienna, VA: Springer Verlag, June 2011.
42. *www.silvaco.com*.
43. D. Vasileska. Modeling thermal effects in nano-devices. *Microelectron. Eng.*, Vol. 109 (9), 163–167 2013.

44. T. Grasser, T. W. Tang, H. Kosina, and S. Selberherr. A review of hydrodynamic and energy-transport models for semiconductor device simulation. *Proc. IEEE*, Vol. 91 (2), 251–274, 2003.
45. T. Grasser, H. Kosina, C. Heitzinger, and S. Selberherr. Characterization of the hot electron distribution function using six moments. *J. Appl. Phys,* Vol. 91, 3869, 2002.
46. M. Lundstrom. Elementary scattering theory of the Si MOSFET. *IEEE Electron Dev. Lett.,* Vol. 18, 361–363, 1997.
47. M. V. Fischetti and S. E. Laux. Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects. *Physical Review B*, Vol. 38(14), 9721, 1988.
48. T. Frauenheim, G. Seifert, M. Elstner, T. Niehaus, C. Köhler, M. Amkreutz et al. Atomistic simulations of complex materials: Ground-state and excited-state properties. *J. Phys. Condens. Matter*, Vol. 14 (11), 3015, 2002.
49. G. Klimeck, S. S. Ahmed, H. Bae, N. Kharche, S. Clark, B. Haley, S. Lee et al. Atomistic simulation of realistically sized nanodevices using NEMO 3-D—Part 1: models and benchmarks. *IEEE Trans. Electron Dev.,* Vol. 54 (9), 2079–2089, 2007.
50. R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic. Single and multiband modeling of quantum electron transport through layered semiconductor devices. *J. Appl. Phys.*, Vol. 81 (12), 7845–7869, 1997.
51. L. Shifren and D. K. Ferry. Wigner function quantum Monte Carlo. *Physica B: Condens. Matter*, Vol. 314 (1), 72–75, 2002.

52

Photonics

52.1 Landscape of Numerical Methods 808
Time-Harmonic versus Time-Dependent Modeling • Exact versus Approximate Models • Order of Approximation and Flexibility of Meshes • Computing Aspects: CPU-Time versus Memory • Special Adaptation to Photonics
52.2 Maxwell's Equations in Nondispersive, Chargeless Media 809
Scattering, Mode Propagation, and Resonance Problems • Simplification: 2D Geometries and 2D Fields • Transformation Rules • Perfectly Matched Layers
52.3 FDTD Method 817
Historical Notes • Concept: Discrete Differential Operators on Staggered Grids • Spatial Discretization of Maxwell's Equations • Time Evolution • Dispersion, Stability, and Courant–Friedrichs–Lewy Condition • Discrete Div-Conditions • Convergence
52.4 Fourier Modal Method 823
Historical Notes • Concept: Segmentwise Waveguide Treatment • Transfer-Matrix Approach • Scattering-Matrix Approach • Layered Media Stack as Model Case • FMM in 2D • Convergence
52.5 Finite Element Method 833
Historical Notes • Concept: Patchwise Tangential Continuous, Polynomial Approximation on Unstructured Grids • 1D: Weak Form and Variational Formulation • 1D: Linear Finite Elements, Local Matrices, and Assembly Process • 2D and 3D: Weak Form and Variational Formulation • 2D and 3D: Linear Finite Elements, Local Matrices, and Assembly Process • Convergence
52.6 DG Method.. 844
Historical Notes • Concept: Combine Adaptive Spatial FEM Discretization with the Time Treatment in Finite Volume Methods • Scalar Advection Equation as Model Equation: The Finite Volume Method in 1D • Maxwell's Equations in Conservation Form • Concept of DG-Algorithms for 3D Maxwell's Equations • Convergence

Frank Schmidt

This chapter deals with numerical simulation in photonics. It provides an overview of the most widely used methods, explains their concepts, and gives some details, as long as it is necessary for a qualitative understanding. It is far from being complete but tries to describe the landscape of simulation tools. Parts of this review are closely related to the book *Numerical Methods in Photonics* [1], which itself is a comprehensive representation of the field covered by many very detailed monographs. The methods we describe in the following are the *finite-difference time-domain FDTD method*, closely related to the *finite integration technique* (FIT), the *Fourier modal method* (FMM), the *finite element method* (FEM), and the *discontinuous galerkin* (DG) method. We omit other very useful methods often specialized to certain applications, in particular:

- Beam propagation method (BPM), for simulation of paraxial beam propagation [2]
- Integral methods such as boundary integral methods, for simulation problems in grating design [3]
- Plane wave expansion (PWE) for simulations of band structures [4]

and refer the reader to the given monographs.

52.1 Landscape of Numerical Methods

There are many aspects to consider if one has to choose a numerical method to solve a particular problem in photonics. We now discuss a few.

52.1.1 Time-Harmonic versus Time-Dependent Modeling

In the most general case, one wants to solve time-dependent Maxwell's equations with general, i.e., nonlocal, nonlinear, dispersive media. Both FDTD and DG methods in time provide such general solutions. On the other hand, especially in the case of high accuracy requirements, strict conditions on the maximal stepsizes in time may cause a high amount of central processing unit (CPU)-time consumption. Therefore, in case of linear and nondispersive problems, it is often useful to consider time-harmonic problems, and, if needed, to combine time-harmonic solutions via Fourier transform with different frequencies to obtain a time-dependent behavior.

52.1.2 Exact versus Approximate Models

It is common to approximate Maxwell's equations by structurally simpler equations. Especially, the simplification in terms of polarizations have a long tradition, where the full vectorial equations are replaced by semivectorial or scalar approximations. The scalar Helmholtz equation, e.g., is well investigated and a number of efficient numerical approaches exists, hence at first glance it seems reasonable to try Helmholtz solvers instead of Maxwell solvers. But the full vectorial Maxwell solvers available today do not have significant disadvantages with respect to speed or memory consumption, thus it is obsolete in many cases to use such semivectorial or scalar approximations. The case is different for special situations such as paraxial field propagations along distinguished propagation directions. As long as reflections can be neglected, in the case of waveguide tapers for example, so-called BPM are useful. They map a time-harmonic three-dimensional (3D) scattering problem approximately to a time-like evolution problem with 2D cross sections, which results in a substantial reduction of the computational effort.

52.1.3 Order of Approximation and Flexibility of Meshes

All methods described in the following offer, at least in principle, high-order approximations in space and time, thus resulting in high-order convergence. Nevertheless, this is not always used in practice. FDTD methods, e. g., are typically linked to a second-order discretization in time and space on staggered grids. Therefore, once geometrical flexibility is needed, one would typically, instead of enhancing FDTD, directly apply DG. Implementation of DG is more involved than of FDTD, but offers the flexibility of FEM meshes by design and yields the time evolution with costs comparable to FDTD.

52.1.4 Computing Aspects: CPU-Time versus Memory

The time-dependent solvers FDTD and DG can be treated by explicit time integrators. Thus, the evolution procedure results in much less memory consumption than that is typical for FMM or FEM, where either full or large sparse matrices have to be inverted. Hence, on memory-restricted systems, FDTD or DG are advantageous. The price to pay is longer computing times.

52.1.5 Special Adaptation to Photonics

FDTD, DG, and FEM are general methods for solving electromagnetic problems; they are not specially tailored to photonics. Since mode propagation is often the central physical effect, numerical methods based on the interaction of waveguide or cavity modes are often very effective. The dominant class of methods here are the FMMs. Subdividing the structure into waveguide sections and relating waveguide modes to each other result in algorithms which often offer surprisingly fast convergence in the range of lower, but in practice sufficient accuracy.

52.2 Maxwell's Equations in Nondispersive, Chargeless Media

In the most general case, the numerical methods of photonics solve time-dependent Maxwell's equations with general media, i.e., with nonlocal, nonlinear, dispersive media. For a concise and clear presentation of the different numerical methods, we use only a subset of the general equations. Our discussions are based on

$$\begin{aligned}\nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}\\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}\\ \nabla \cdot \mathbf{D} &= 0\\ \nabla \cdot \mathbf{B} &= 0\end{aligned} \tag{52.1}$$

completed by the corresponding constitutive equations

$$\mathbf{D} = \varepsilon \mathbf{E} \qquad \mathbf{B} = \mu \mathbf{H} \qquad \mathbf{J} = \sigma \mathbf{E}. \tag{52.2}$$

These equations describe the fields in linear, nondispersive, and chargeless media, where the magnetic field **H**, the electric field **E**, the electric displacement **D**, and the magnetic flux density **B** are vectorial quantities defined for the entire space and positive times, $\mathbf{H}, \mathbf{E}, \mathbf{D}, \mathbf{B} : \mathbb{R}^3 \times \mathbb{R}^+ \longrightarrow \mathbb{R}^3$, and the permittivity ε, the permeability μ, and the conductivity σ are scalar, piecewise continuous functions in space, $\varepsilon, \mu, \sigma : \mathbb{R}^3 \longrightarrow \mathbb{R}$.

The field equations combined with the constitutive equations yield the system, on which our representation of the FDTD method and the DG method are based:

$$\begin{pmatrix} -\sigma & \nabla\times \\ \nabla\times & \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \frac{\partial}{\partial t}\left[\begin{pmatrix} \epsilon & \\ & -\mu \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}\right]. \tag{52.3}$$

The other methods, FEM, FMM, etc., are based on the time-harmonic form of Equation 52.3. The concept is to use a Fourier transform of the fields in Equation 52.3 for a single frequency ω. Assuming the single frequency dependence to be $\exp(-i\omega t)$, the Fourier transformed fields become complex but live in space only, $\mathbf{H}, \mathbf{E} : \mathbb{R}^3 \longrightarrow \mathbb{C}^3$. We do not introduce a special notation for the time-harmonic fields, since the meaning becomes clear from the context. In the frequency domain, Equation 52.3 reads

$$\begin{pmatrix} & \nabla\times \\ \nabla\times & \end{pmatrix}\begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = -i\omega\left[\begin{pmatrix} \epsilon - \frac{\sigma}{i\omega} & \\ & -\mu \end{pmatrix}\begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}\right]. \tag{52.4}$$

52.2.1 Scattering, Mode Propagation, and Resonance Problems

Maxwell's equations take different forms depending on the different modeling situations. In the case of time-dependent problems, these are the different types of linear and nonlinear scattering problems. In the case of time-harmonic problems, where mostly but not exclusively, linear problems are considered, we have scattering, mode propagation, and resonance problems. Time-dependent solutions to Equation 52.3 including also nonlocal, dispersive, or nonlinear materials are the most general classical solutions to Maxwell's equations—the corresponding computational effort is usually relatively large. Solutions to the time-harmonic system (52.4) are often much cheaper.

52.2.1.1 Scattering on Unbounded Domain

Figure 52.1 shows the situation of a time-harmonic scattering problem, the situation for time-dependent scattering is very similar. There is a source field $\mathbf{E}_{\text{source}}$ traveling through the unbounded space, being itself a solution to Maxwell's equations (Equation 52.4). If no scatterer would be present, this would be the only field we have. But if it hits an object, scattered fields $\mathbf{E}_{\text{scattered}}$ (or in short $\mathbf{E}_s$) are generated. We need to compute the field inside a computational domain Ω, which we also call the interior domain, in contrast to the exterior domain surrounding Ω. Further, we denote the boundary of the computational domain with $\partial\Omega$.

Let the source field $\mathbf{E}_{\text{source}}(\mathbf{r})$, with $\mathbf{r}$ in the exterior domain, $\mathbf{r} \in \mathbb{R}^3 \setminus \bar{\Omega}$ be given. Then, the general scattering problem in 3D is defined by the following conditions:

1. Interior problem. The field on the computational domain obeys Equation 52.4.
2. Exterior problem. The fields in the exterior domain are a superposition of the source and the scattered fields $\mathbf{E}_{\text{source}}$ and $\mathbf{E}_s$. The source field and the scattered field must also obey Maxwell's equations.
3. Continuity of the tangential data and their normal derivative along the boundary of the computational domain.
4. In the case of time-harmonic fields, the Silver–Müller radiation condition for the scattered field

$$\lim_{r\to\infty} r\left(\left(\nabla \times \mathbf{E}_s\right) \times \mathbf{r}^0 - i\,|\mathbf{k}|\,\mathbf{E}_s\right) = 0 \tag{52.5}$$

holds true uniformly in all directions $\mathbf{r}$. Here, $\mathbf{r}^0$ denotes the radial vector of unit length.

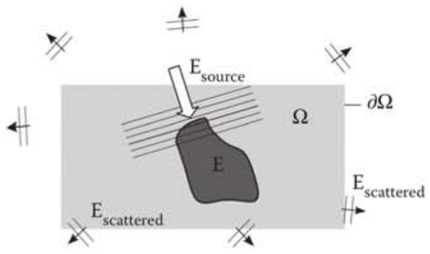


FIGURE 52.1 Scattering problem.

Typically, the radiation condition is not directly applied. But the methods that treat the exterior part of the problem ensure that the scattered fields satisfy the condition. The most prominent and successful method here is the perfectly matched layer (PML) method, see the Section 52.2.4.

52.2.1.2 Resonance Problems

In resonance problems, we consider Equation 52.4 as an eigenvalue problem with the unknown eigenvalue ω and the unknown resonance mode, the eigenvector, $(\mathbf{E},\mathbf{H})^T$. The special challenge lies in the treatment of the exterior domain. Either the boundary conditions on the boundary of the computational domain are explicitly known, e. g., as a magnetic or electric wall, or methods from scattering problems are applied.

52.2.1.3 Mode Propagation Problems

Again, we consider Equation 52.4 as an eigenvalue problem. This time, we specialize the problem further to z-invariant problems: $\varepsilon = \varepsilon(x, y)$. We take a fixed frequency ω and look for solutions of the type $\mathbf{E}(x, y, z) = \mathbf{E}(x, y)e^{ik_z z}$ with a (unknown) phase velocity k_z. Two common forms of the eigenvalue problem for k_z are the following: Since any z-dependence is given by factors $\exp(ik_z z)$, one can introduce a specialized nabla-operator $\nabla_{k_z} = \mathbf{e}_x\partial_x + \mathbf{e}_y\partial_y + ik_z\mathbf{e}_z$ and Equation 52.4 reads

$$\begin{pmatrix} & \nabla_{k_z}\times \\ \nabla_{k_z}\times & \end{pmatrix}\begin{pmatrix}\mathbf{E}\\ \mathbf{H}\end{pmatrix} = -i\omega\left[\begin{pmatrix}\epsilon-\frac{\sigma}{i\omega} & \\ & -\mu\end{pmatrix}\begin{pmatrix}\mathbf{E}\\ \mathbf{H}\end{pmatrix}\right].$$

Here, the eigenvector consists of both $\mathbf{E}$ and $\mathbf{H}$, the eigenvalue k_z appears in linear form. Alternately, eliminating, e. g., the magnetic field, and using the decompositions $\mathbf{E} = \mathbf{E}_\perp + \mathbf{e}_z E_z$ and $\nabla = \nabla_\perp + i\mathbf{e}_z k_z$, one derives from Equation 52.4 the modal eigenvalue problem

$$\begin{bmatrix} \frac{1}{\varepsilon}\nabla_\perp\times\frac{1}{\mu}\nabla_\perp\times \;+\frac{k_z^2}{\varepsilon\mu} & \frac{ik_z}{\varepsilon\mu}\nabla_\perp \\ i\frac{k_z}{\varepsilon}\nabla_\perp\frac{1}{\mu}\cdot & -\frac{1}{\varepsilon}\nabla_\perp\cdot\frac{1}{\mu}\nabla_\perp \end{bmatrix}\begin{bmatrix}\mathbf{E}_\perp\\ E_z\end{bmatrix} = \omega^2\begin{bmatrix}\mathbf{E}_\perp\\ E_z\end{bmatrix} \tag{52.6}$$

The eigenvector consists just of the $\mathbf{E}$-field, but this is a quadratic eigenvalue problem in k_z. The same procedure for the magnetic field yields

$$\begin{bmatrix} \frac{1}{\mu}\nabla_\perp\times\frac{1}{\varepsilon}\nabla_\perp\times \;+\frac{k_z^2}{\varepsilon\mu} & \frac{ik_z}{\varepsilon\mu}\nabla_\perp \\ i\frac{k_z}{\mu}\nabla_\perp\frac{1}{\varepsilon}\cdot & -\frac{1}{\mu}\nabla_\perp\cdot\frac{1}{\varepsilon}\nabla_\perp \end{bmatrix}\begin{bmatrix}\mathbf{H}_\perp\\ H_z\end{bmatrix} = \omega^2\begin{bmatrix}\mathbf{H}_\perp\\ H_z\end{bmatrix}. \tag{52.7}$$

52.2.2 Simplification: 2D Geometries and 2D Fields

In a 2D geometry, the permittivity ε depends only on two coordinates, say the x- and y-coordinates: $\varepsilon = \varepsilon(x, y)$, just as in the mode propagation case in Section 52.2.1. In the mode propagation case, all fields depend on a common phase factor $\exp(ik_z z)$. If we set $k_z = 0$, i. e., we consider only fields with transversal position dependence $\mathbf{E} = \mathbf{E}(x, y)$, we obtain from (Equations 52.6 and 52.7) the decoupled 2D equations

$$\begin{aligned} \nabla_\perp\times\nabla_\perp\times\mathbf{E}_\perp(x, y) &= \varepsilon(x, y)\mu\omega^2\mathbf{E}_\perp(x, y) \\ -\nabla_\perp\cdot\nabla_\perp E_z(x, y) &= \varepsilon(x, y)\mu\omega^2 E_z(x, y). \end{aligned} \tag{52.8}$$

If we additionally assume uniformity in y-direction, we find the eigenvalue problem for waveguides with a 1D cross section

$$\partial_x^2 E_z + \varepsilon(x)\mu\omega^2 E_z = k_y^2 E_z. \tag{52.9}$$

The same applies to the magnetic field:

$$\begin{aligned}\nabla_\perp \times \frac{1}{\varepsilon(x,y)} \nabla_\perp \times \mathbf{H}_\perp &= \mu\omega^2 \mathbf{H}_\perp(x,y)\\ \varepsilon(x)\partial_x \frac{1}{\varepsilon(x)} \partial_x H_z + \varepsilon(x)\mu\omega^2 H_z &= k_y^2 H_z. \end{aligned}\tag{52.10}$$

These equations are simpler than the original ones and they build the basis of many (approximate) solvers. They are simpler in the operator structure and they are linear in the eigenvalue k_y^2.

Note that, conventionally, the propagation direction of waveguide modes along the axis of waveguides is chosen to be the z-axis. To meet this convention, one has to interchange the y- and z-indices in Equations 52.9) and 52.10.

As a consequence of decoupling, we can split the entire field globally into a transversal electric (TE or s-polarization) and a transversal magnetic (TM or p-polarization) field:

$$\begin{pmatrix} \mathbf{E}(x,y) \\ \mathbf{H}(x,y) \end{pmatrix} = \underbrace{\begin{pmatrix} E_z(x,y)\mathbf{e}_z \\ H_x(x,y)\mathbf{e}_x + H_y(x,y)\mathbf{e}_y \end{pmatrix}}_{\text{s-polarization (TE)}} + \underbrace{\begin{pmatrix} E_x(x,y)\mathbf{e}_x + E_y(x,y)\mathbf{e}_y \\ H_z(x,y)\mathbf{e}_z \end{pmatrix}}_{\text{p-polarization (TM)}}. \tag{52.11}$$

52.2.3 Transformation Rules

Transformation techniques often yield considerable simplifications, e.g.,

- Exploiting symmetries by adapting coordinate systems (cylindrical, spherical coordinate systems, etc.)
- Using precomputations on reference elements as in FEM, cf. Section 52.5
- Handling of PML, cf. Section 52.2.4, by a complex coordinate stretching

The transformation rules describe the change of the equations when we change a local coordinate system. The PML, seen as a complex scaling of coordinates, can be treated as change of a coordinate system. The computations in FEM simplify if they are done in the reference coordinate systems and then mapped to the physical systems.

Geometric quantities and field quantities transform differently. For an easy *algebraic* notation, it is useful to write geometric quantities in row vectors and field quantities in column vectors of their coefficients. Expressions like $\mathbf{dl} \cdot \mathbf{E}$ with $\mathbf{dl} = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}z)$ (a row vector of path components) and $\mathbf{E} = (E_x, E_y, E_z)^T$ (a column vector of field components) have then the right algebraic form and can be programed directly as they appear in the formula.

52.2.3.1 Mapping of Geometric Quantities

Let an arbitrary domain mapping $Q : \widehat{K} \to K$ be given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = Q\left(\widehat{x}, \widehat{y}, \widehat{z}\right), \tag{52.12}$$

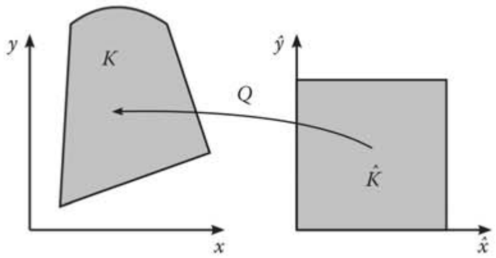


FIGURE 52.2 Example of a domain mapping. A square is mapped to a curvilinear quadrilateral. Quantities belonging to the reference domain are marked with a hat.

see Figure 52.2. This way we map *points* from the reference domain to *points* in the original domain, where the points are defined by the coefficients $\widehat{x}, \widehat{y}, \widehat{z}$ in the reference system and by the coefficients x, y, z in the original system.

We study how infinitesimally small path, face, and volume elements $\mathbf{dl}$, $\mathrm{d}\mathbf{A}$, $\mathrm{d}V$, transform. In 3D, we describe $\mathbf{dl}$ and $\mathrm{d}\mathbf{A}$ by three-coefficient row vectors. A linearization of the mapping at point $\left(\widehat{x}, \widehat{y}, \widehat{z}\right)$ yields

$$\begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{pmatrix} = J\left(\widehat{x}, \widehat{y}, \widehat{z}\right) \begin{pmatrix} \mathrm{d}\widehat{x} \\ \mathrm{d}\widehat{y} \\ \mathrm{d}\widehat{z} \end{pmatrix}, \qquad J\left(\widehat{x}, \widehat{y}, \widehat{z}\right) := \begin{pmatrix} \partial_{\widehat{x}} Q_1 & \partial_{\widehat{y}} Q_1 & \partial_{\widehat{z}} Q_1 \\ \partial_{\widehat{x}} Q_2 & \partial_{\widehat{y}} Q_2 & \partial_{\widehat{z}} Q_2 \\ \partial_{\widehat{x}} Q_3 & \partial_{\widehat{y}} Q_3 & \partial_{\widehat{z}} Q_3 \end{pmatrix}. \tag{52.13}$$

Defining the path element as the row vector $\mathbf{dl} = \left(\mathrm{d}x, \mathrm{d}y, \mathrm{d}z\right)$, we obtain the path element transformation

$$\mathbf{dl} = \mathrm{d}\widehat{\mathbf{l}} J^T. \tag{52.14}$$

A similar computation for a face element $\mathrm{d}\mathbf{A} = \left(\mathrm{d}A_x, \mathrm{d}A_y, \mathrm{d}A_z\right)$ spanned by two path elements gives, with $|J|$ the determinant of J, the transformation rule

$$\mathrm{d}\mathbf{A} = \mathrm{d}\widehat{\mathbf{A}} J^{-1} \, |J| \tag{52.15}$$

and for a volume element $\mathrm{d}V$ spanned by three path elements we obtain

$$\mathrm{d}V = \mathrm{d}\widehat{V} \, |J| \, . \tag{52.16}$$

52.2.3.2 Transformation of Grad, Curl, Div

We use the nabla notation describing a column vector

$$\mathbf{grad}\, v(x, y, z) = \nabla_{xyz} v(x, y, z) = \begin{pmatrix} \partial_x v(x, y, z) \\ \partial_y v(x, y, z) \\ \partial_z v(x, y, z) \end{pmatrix}.$$

Let an arbitrary function scalar $v(x, y, z)$ be given which lives on K and allows for the computation of the gradient. We compute the scalar quantity $\mathbf{dl} \cdot \nabla_{xyz} v$. This is a physical quantity with the meaning of a voltage (e.g., $\mathbf{E} \cdot \mathbf{ds}$) or a current ($\mathbf{H} \cdot \mathbf{ds}$). This quantity must not change if we change the coordinate system, hence, we require

$$\mathbf{dl} \cdot \nabla_{xyz} v(x, y, z) = \mathrm{d}\widehat{\mathbf{l}} \cdot \nabla_{\widehat{x}\widehat{y}\widehat{z}} \widehat{v}\left(\widehat{x}, \widehat{y}, \widehat{z}\right),$$

TABLE 52.1 Transformation of Geometric and Field Quantities

	Geometric Quantity	Field Quantity	Nabla Operation
Segment	$\mathrm{d}\mathbf{l} = \mathrm{d}\widehat{\mathbf{l}}J^{T}$	$\mathbf{E} = J^{-T}\widehat{\mathbf{E}}$ $\mathbf{H} = J^{-T}\widehat{\mathbf{H}}$	$\nabla_{xyz}v = J^{-T}\nabla_{\widehat{x}\widehat{y}\widehat{z}}\widehat{v}$
Face	$\mathrm{d}\mathbf{A} = \mathrm{d}\widehat{\mathbf{A}}\,\lvert J\rvert\, J^{-1}$	$\mathbf{D} = \frac{J}{\lvert J\rvert}\widehat{\mathbf{D}}$ $\mathbf{B} = \frac{J}{\lvert J\rvert}\widehat{\mathbf{B}}$	$\nabla_{xyz} \times \mathbf{v} = \frac{J}{\lvert J\rvert}\nabla_{\widehat{x}\widehat{y}\widehat{z}} \times \widehat{\mathbf{v}}$
Volume	$\mathrm{d}V = \mathrm{d}\widehat{V}\,\lvert J\rvert$	$\rho = \frac{1}{\lvert J\rvert}\widehat{\rho}$	$\nabla_{xyz} \cdot \mathbf{v} = \frac{1}{\lvert J\rvert}\nabla_{\widehat{x}\widehat{y}\widehat{z}} \cdot \widehat{\mathbf{v}}$

where $\widehat{v}\left(\widehat{x},\widehat{y},\widehat{z}\right)$ in the reference domain follows from $v\left(\widehat{x},\widehat{y},\widehat{z}\right)$ in the original domain by $v(x,y,z) = v\left(Q\left(\widehat{x},\widehat{y},\widehat{z}\right)\right) = (v\circ Q)\left(\widehat{x},\widehat{y},\widehat{z}\right)$, hence $\widehat{v} = v\circ Q$, based on the coordinate transform Equation 52.12. It follows from Equation 52.14 that

$$\mathrm{d}\widehat{\mathbf{l}}J^{T} \cdot \nabla_{xyz}v(x,y,z) = \mathrm{d}\widehat{\mathbf{l}} \cdot \nabla_{\widehat{x}\widehat{y}\widehat{z}}\widehat{v}\left(\widehat{x},\widehat{y},\widehat{z}\right).$$

Since this should hold true for all functions v, we find the mapping of the gradient

$$\nabla_{xyz} = J^{-T}\nabla_{\widehat{x}\widehat{y}\widehat{z}}. \tag{52.17}$$

The coefficient of the nabla operator transform like the coefficients of the field strength. The same way one computes the transformations rules for **curl** and **div**, as collected in Table 52.1.

52.2.3.3 Transformation of μ and ε

Rewriting Maxwell's **curl** equations from the physical to the reference domain we get

$$\frac{J}{|J|}\nabla_{\widehat{x}\widehat{y}\widehat{z}} \times \widehat{\mathbf{H}} = i\omega\varepsilon J^{-T}\widehat{\mathbf{E}}$$
$$\frac{J}{|J|}\nabla_{\widehat{x}\widehat{y}\widehat{z}} \times \widehat{\mathbf{E}} = -i\omega\mu J^{-T}\widehat{\mathbf{H}},$$

which shows that the entire transformation can be traced back to a mapping of the material quantities

$$\widehat{\varepsilon} : = |J|\, J^{-1}\varepsilon J^{-T} \tag{52.18}$$
$$\widehat{\mu} : = |J|\, J^{-1}\mu J^{-T}. \tag{52.19}$$

52.2.4 Perfectly Matched Layers

A standard problem in the simulation of scattering problems is the treatment of scattered fields outside the domain of interest Ω_{int}, cf. Figure 52.3. The predominantly used method is the PML technique introduced by Berenger [5]. The PML treatment of Berenger relies on a split field formulation of Maxwell's equations in the PML region. This has been shown in [6] to be only weakly stable. In contrast, the uniaxial PML (UPML) [7] does not require such a splitting of Maxwell's fields and maintains the strong well-posedness of the original Maxwell system. Both formulations are equivalent to an elegant and general approach: the stretched-coordinate PML [8,9]. We describe briefly only the latter one. It has among others two remarkable properties

1. It can be used in a similar way in higher space dimensions.
2. It can be plugged into existing codes provided that the codes allow for sufficient general materials and field representations.

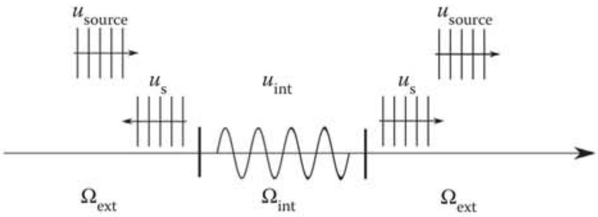


FIGURE 52.3 Sketch of a 1D time-harmonic scattering problem. The source fields are given, the interior field u_{int} and the scattered field u_S has to be computed.

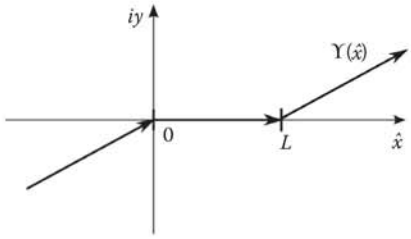


FIGURE 52.4 Complex PML path. Within the interior domain $(0, L)$ the x-coordinate is real but becomes complex outside.

52.2.4.1 PML in 1D

Let us consider the right exterior domain $x > L$, where L marks the right boundary point of the situation depicted in Figure 52.3. The general solution of the scattered field in the scalar 1D version of Equation 52.8 is

$$u_s = c_- e^{-ikx} + c_+ e^{ikx}, \quad x > L.$$

We have to discriminate between the incoming version of the scattered field $c_- \exp(-ikx)$, which we need to drop and the outgoing scattered field $c_+ \exp(ikx)$. Here, x is the real space coordinate. Now, we allow x to be a path in the complex plane, not just on the real axis. We parameterize path x with the real parameter $\widehat{x}$ to derive a path in the complex plane; now $\widehat{x}$ takes the role of the reference coordinate as in Section 52.2.3:

$$\begin{aligned} x(\widehat{x}) &= L + (1 + i\sigma)\widehat{x} \\ &\sigma > 0, \text{ real, fixed} \\ &\widehat{x} > 0, \text{ real reference coordinate;} \end{aligned} \tag{52.20}$$

compare Figure 52.4, which yields the following mapping:

$$\begin{aligned} c_- \exp(-ikx) &\mapsto c_- \exp(-ikL) \exp(-ik\widehat{x}) \exp(\sigma k\widehat{x}) \\ c_+ \exp(ikx) &\mapsto c_+ \exp(ikL) \exp(ik\widehat{x}) \exp(-\sigma k\widehat{x}). \end{aligned}$$

The first expression, the incoming field, blows up for $\widehat{x} \to \infty$, due to the factor $\exp(\sigma k\widehat{x})$, whereas the second one decreases exponentially with $\exp(-\sigma k\widehat{x})$. So for $\widehat{x}$ large enough, we have a clear discrimination:

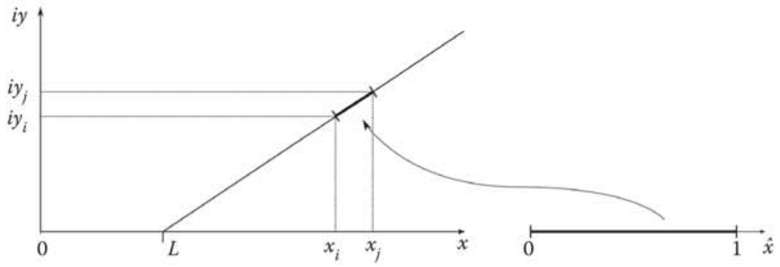


FIGURE 52.5 Transformation from the unit interval to the complex-valued PML interval.

nearly the entire field value comes from the left traveling scattered field. Hence, a proper formulation of Sommerfeld's radiation condition in terms of the complexified path $x(\widehat{x})$ is

$$u\left(x\left(\widehat{x}\right)\right) \longrightarrow 0 \text{ as } \widehat{x} \longrightarrow \infty.$$

Let us stay in the right exterior domain and introduce the complex distance variable x:

$$x(\widehat{x}) = \widehat{x} + i\sigma\widehat{x}.$$

We consider the complexification of the spatial variable within the Helmholtz equation:

$$\partial_x^2 u + k^2 u = 0 \qquad \Rightarrow \qquad \frac{1}{(1+i\sigma)^2}\frac{\partial^2 u}{\partial \widehat{x}^2} + k^2 u = 0.$$

This step is decisive. If we are allowed to generalize the real derivative to a complex derivative, i. e., if the derivative is independent of the direction in the complex plane, we can use the linear path transformation, see Figure 52.5, (or others) to obtain the last equation.

This transformation is what we in fact exploit within the algorithmic realization. We define the complexification of the real x-coordinate to a complex coordinate x by means of the real reference coordinate $\widehat{x}$

$$x = \begin{cases} \widehat{x}\,(1+i\sigma), & \widehat{x} < 0 \\ \widehat{x}, & 0 \le \widehat{x} < L \\ L + \left(\widehat{x} - L\right)(1+i\sigma), & \widehat{x} \ge L. \end{cases} \tag{52.21}$$

52.2.4.2 PML in 2D and 3D

The (tensor product type) generalization of Equation 52.20 to 2D is obviously the following (σ_x, σ_y are real, positive constants):

$$x = \begin{cases} \widehat{x}\left(1+i\sigma_x\right), & \widehat{x} < 0 \\ \widehat{x}, & 0 \le \widehat{x} < L\,, \\ L + \left(\widehat{x} - L\right)\left(1+i\sigma_x\right), & \widehat{x} \ge L \end{cases} \quad y = \begin{cases} \widehat{y}\left(1+i\sigma_y\right), & \widehat{y} < 0 \\ \widehat{y}, & 0 \le \widehat{y} < L \\ L + \left(\widehat{y} - L\right)\left(1+i\sigma_y\right), & \widehat{y} \ge L. \end{cases} \tag{52.22}$$

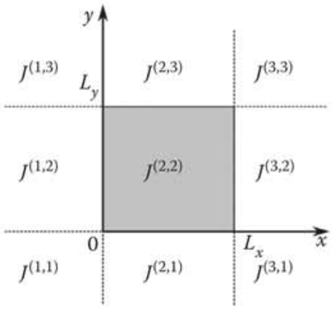


FIGURE 52.6 PML-regions in 2D lexicographically enumerated.

The Jacobian of the mapping $(\hat{x}, \hat{y}) \mapsto (x, y)$ depends on the position, see Figure 52.6, and evaluates to, e.g.,

$$J^{(2,2)} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \qquad J^{(3,2)} = \begin{pmatrix} 1 + i\sigma_x & \\ & 1 \end{pmatrix} \qquad J^{(3,3)} = \begin{pmatrix} 1 + i\sigma_x & \\ & 1 + i\sigma_y \end{pmatrix}.$$

The realization of the PML now consists in using the transformed permittivity (Equation 52.18) and permeability (Equation 52.19), where the Jacobians are computed from the complex stretching of the space variables (Equation 52.22). We proceed now as follows. We consider a subdomain of the exterior, say $\Omega^{(2,1)} \subset \mathbb{R}^2$ from Figure 52.6 and the complex solution $u^{(2,1)} : \Omega^{(2,1)} \to \mathbb{C}$ living there. Then, we use the mapping from the real to the complex coordinates to define a complex-valued spatial domain: $\Omega^{(2,1)} \to \Omega^{c,(2,1)} \subset \mathbb{C}^2$. On the complexified domain $\Omega^{c,(2,1)}$, we have a complex extension of $u^{(2,1)}$ which we denote by $u^{c,(2,1)}$. The solution $u^{c,(2,1)}$ is different to the one on the real space $\Omega^{(2,1)}$, it is the complex continuation of the original function. The key point is that our original function $u^{(2,1)}(x, y)$ has a complex extension by its nature, and we point to its values by introducing complex instead of only real spatial coordinates.

52.3 FDTD Method

The FDTD method is the most widely used method for the simulation of electromagnetic field propagation. It combines algorithmic simplicity with a deep universality. Nonlinear, nonlocal, and dispersive material responses can be treated. The reason for this generality lies in the fact that the structure of the FDTD scheme is very close to the structure of Maxwell's equations and has several algorithmic features:

- It is simple to implement and fast.
- High-order discretizations are feasible.
- It is explicit in time.
- There is a comprehensive theory.

Besides these pros, there are also severe drawbacks:

- The simple local approximation based on Cartesian staggered grids restricts geometric flexibility unless the grid is prohibitively fine (staircasing problem).
- The inherent second-order accuracy in time and space of the original method limits long-time computations on large domains.

52.3.1 Historical Notes

The initial and most important step was done by Yee in 1966 [10]. He used a so-called staggered mesh discretization in space and time, see the following subsections, and replaced the differential operators **grad**, **curl**, and **div** by finite-difference approximations to obtain a discrete approximation of second-order accuracy. Since then, a huge number of contributions improved the method and extended the fields of applications. An interesting view on the derivation of the FDTD equations has been given by Weiland in 1977 [11] starting from an integral formulation of Maxwell's equations. This approach is called FIT and yields the same set of equations. We use this approach to motivate the finite difference equations.

A number of efforts have been aimed at addressing the above-mentioned problems. The staircasing problem has been treated in many papers, e.g., [12]. High-order methods have been proposed, e.g., in [13, 14].The *generalized finite difference method* introduced by Bossavit [15] in 2001 takes ideas from differential geometry and topology and stimulates new interpretations of classical finite element schemes in terms of finite differences. Clemens and Weiland [16] follow a similar spirit in the framework of FIT to extend classical schemes.

A comprehensive overview of FDTD methods and further improvements along with detailed applications is given in [17] which is the standard reference book on FDTD methods. A thoroughly different approach to time-dependent problems is given by the class of DG methods which by its inherent construction offer high-order approximations in space and time as well as flexible meshes, cf. Section 52.6.

52.3.2 Concept: Discrete Differential Operators on Staggered Grids

Figure 52.7 explains the basic construction principle. We consider two entangled loops. One, say the light gray loop, is associated with the electric field, the other one with the magnetic field (left part of Figure 52.7). These loops will be used later for line integrals. The straight edges of the loops themselves are the edges of a cube, see the middle part of Figure 52.7. Then, by adding cubes related to the electric and magnetic fields, the complete spatial Yee mesh is generated. It can be considered as composed from two independent Cartesian meshes where the "electric" mesh and the "magnetic" mesh are shifted compared to each other by half of an elementary edge length of a cube.

52.3.2.1 Degrees of Freedom

A central question in any numerical scheme is the proper definition of the degrees of freedom (DoF). Typically, the stability of a method essentially relies on this. In FDTD methods, it is common to project the fields to the edges and to take the values at the midpoints of the edges as DoF. For example, we project $\mathbf{E}$ to the ith edge $\mathbf{e}_i$, $\mathbf{E} \cdot \mathbf{e}_i$, take the value at the midpoint, and denote it as E_i. We motivate the DoFs in the following by approximations of integral expressions (e.g., by the midpoint rule) as it is done in the FIT. This offers additional insight into the structure of the equations. The same results are obtained if the differential operators are approximated by finite difference stencils.

52.3.2.2 Discrete Gradient

For a direct discretization of Maxwell's equations, it is not necessary to consider the discrete approximation of the gradient but it completes the insight into the structure of the subsequent discretization process. Consider an edge $\mathbf{e}$ of length h connecting two neighboring vertices v_1, v_2 of the Yee mesh. The fundamental theorem of calculus

$$\int_{v_1}^{v_2} \mathbf{grad}\, \phi \, d\mathbf{e} = \phi(v_2) - \phi(v_1),$$

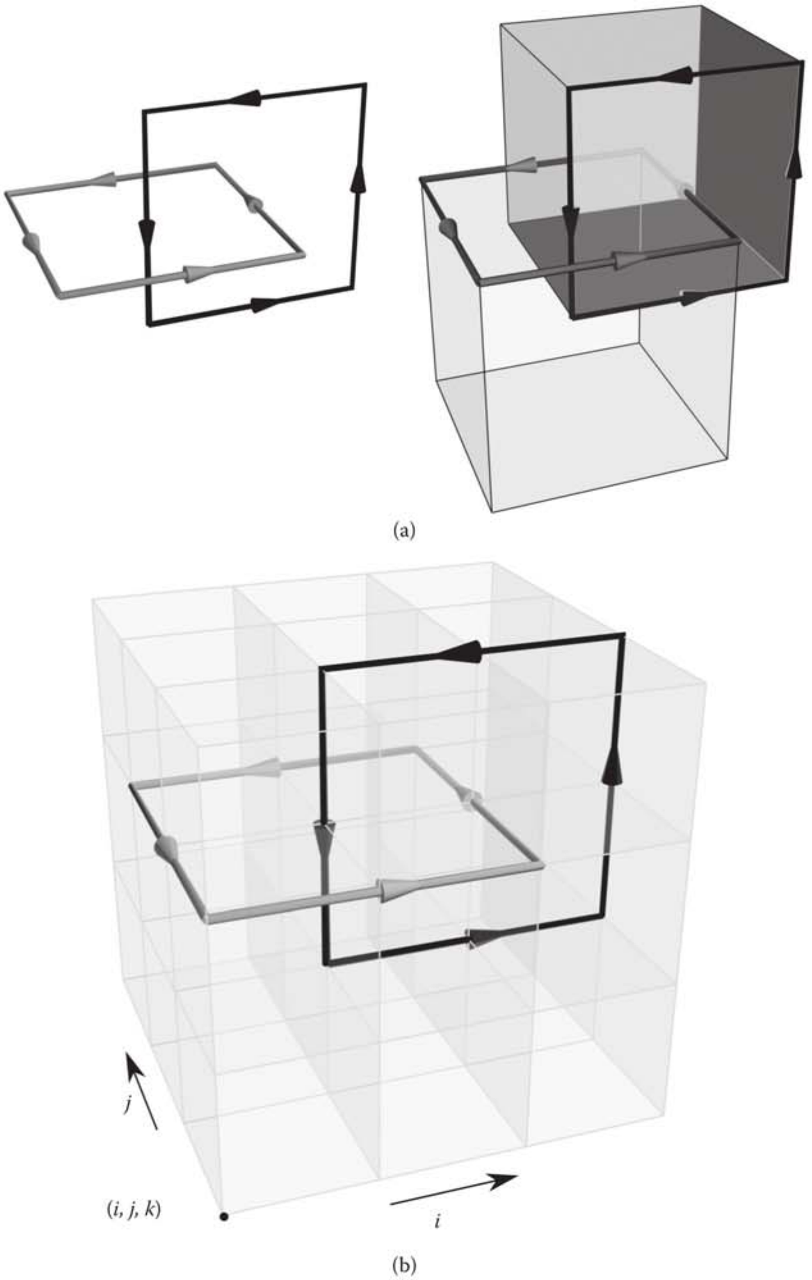


FIGURE 52.7 Schematic interaction of electric and magnetic fields and its realization with staggered grids.

motivates the definition of the discrete gradient operator $\mathbf{grad}_\mathrm{h}$:

$$\mathbf{grad}_\mathrm{h}\, \phi\big|_{\mathbf{e}} \,:=\, \frac{\mathbf{e}^0}{h}\left(\phi(v_2) - \phi(v_1)\right) \text{ forall edges e},$$

where $\mathbf{e}^0$ is the normalization of $\mathbf{e}$ to unit length. The $\mathbf{grad}_\mathrm{h}$ -operator maps quantities defined on *vertices* to quantities defined on *edges*.

52.3.2.3 Discrete Curl

The definition of the discrete **curl** -operator, $\mathbf{curl}_\mathrm{h}$, relies on the Stokes theorem

$$\int_{\mathbf{f}} \mathbf{curl}\,\mathbf{E}\ \mathrm{d}\mathbf{f} = \int_{\partial\mathbf{f}} \mathbf{E}\ \mathrm{d}\mathbf{e}$$

with **f** a plane, rectangular face of a cube and $\partial\mathbf{f}$ its boundary consisting of four edges $\mathbf{e}_i$. The discrete version reads

$$\mathbf{curl}_\mathrm{h}\,\mathbf{E}\big|_{\mathbf{f}} \ : \ = \frac{\mathbf{f}^0}{h_1 h_2} \sum_{i=1}^{4} \iota_{e_i,f}\,\left|\mathbf{e}_i\right| E_i \text{ forall faces } \mathbf{f}.$$

The iota-symbol $\iota_{e_i,f}$ denotes the incidence number of edge $\mathbf{e}_i$ with respect to the orientation of the face **f**. It holds $\iota_{e_i,f} = 1$, if edge $\mathbf{e}_i$ shares the same direction as that induced by the face vector **f**, otherwise $\iota_{e_i,f} = -1$. The $\mathbf{curl}_\mathrm{h}$ -operator maps quantities defined on *edges* to quantities defined on *faces*. For the light gray loop of Figure 52.7, e.g., starting with the edge in front of the figure and assuming the edge lengths of the rectangle are all the same $h_1 = h_2 = h$, this is

$$\mathbf{curl}_\mathrm{h}\,\mathbf{E}\big|_{\mathbf{f}} \ : \ = \frac{\mathbf{f}^0}{h}\left(E_1 + E_2 - E_3 - E_4\right).$$

52.3.2.4 Discrete Divergence

The definition of the discrete **div** -operator, $\mathbf{div}_\mathrm{h}$, relies on the Gauss theorem

$$\int_{V} \mathbf{div}\,\mathbf{D}\ dV = \int_{\partial V} \mathbf{D}\ \mathrm{d}\mathbf{f}.$$

Here, V is the volume of the cube, ∂V its boundary consisting of 6 faces, and **f** is the outward pointing surface vector. The discrete version reads

$$\mathbf{div}_\mathrm{h}\,\mathbf{D}\big|_{c} = \frac{1}{h_1 h_2 h_3} \sum_{i=1}^{6} \iota_{e_f,c}\,\left|\mathbf{f}_i\right| D_i \text{ for all cells } c.$$

The iota-symbol $\iota_{f_i,c}$ denotes the incidence number of face $\mathbf{f}_i$ with respect to the orientation of cell c, defined positive by the outward orientation of all surface vectors. Hence, it holds $\iota_{f_i,c} = 1$, if face $\mathbf{f}_i$ points outward c, otherwise $\iota_{f_i,c} = -1$. The $\mathbf{div}_\mathrm{h}$ -operator maps quantities defined on *faces* to quantities defined on *cells*.

52.3.2.5 Exact Sequences

As a direct consequence of the fact that the Yee mesh is orientable, it follows

$$\begin{aligned} \mathbf{curl}_\mathrm{h} \cdot \mathbf{grad}_\mathrm{h} &= 0 \\ \mathbf{div}_\mathrm{h} \cdot \mathbf{curl}_\mathrm{h} &= 0. \end{aligned} \tag{52.23}$$

Hence, the discrete operators share the same properties as the continuous ones. For the Yee mesh, the property follows easily from a direct computation.

52.3.3 Spatial Discretization of Maxwell's Equations

Let us denote by $\mathbf{E}_\mathrm{h}$ the column vector of all DoFs associated with the primal mesh and by $\mathbf{H}_\mathrm{h}$ the column vector of all DoFs associated with the dual mesh. Then, the spatial discrete version of Equation 52.3, setting $\sigma = 0$ for simplicity, reads

$$\varepsilon\partial_t \mathbf{E}_\mathrm{h}\big|_{\mathbf{e}_\mathrm{p}} = \mathbf{curl}_\mathrm{h}\, \mathbf{H}_\mathrm{h}\big|_{\mathbf{f}_\mathrm{d}} \qquad \text{for all faces } \mathbf{f}_\mathrm{d} \text{ of the dual mesh,} \tag{52.24}$$

$$\mu\partial_t \mathbf{H}_\mathrm{h}\big|_{\mathbf{e}_\mathrm{d}} = -\,\mathbf{curl}_\mathrm{h}\, \mathbf{E}_\mathrm{h}\big|_{\mathbf{f}_\mathrm{p}} \qquad \text{for all inner faces } \mathbf{f}_\mathrm{p} \text{ of the primal mesh.} \tag{52.25}$$

Here, $\mathbf{e}_\mathrm{p}$ denotes the edge of the primal mesh that penetrates face $\mathbf{f}_\mathrm{d}$ of the dual mesh and $\mathbf{e}_\mathrm{d}$ denotes the edge of the dual mesh going through $\mathbf{f}_\mathrm{p}$. For implementations, it is convenient to consider both the primal and the dual meshes as submeshes of a common Cartesian mesh.

52.3.4 Time Evolution

The time evolution of the semidiscretized equations (Equations 52.24 and 52.25) can be computed by a number of different time integrators. Frequently used is the so-called leapfrog scheme. It fits perfectly to the spatial discretization and offers the advantage of being an explicit method.

52.3.4.1 Leapfrog Scheme in Time

We introduce a uniform mesh in time

$$0 = t_0 < t_{\frac{1}{2}} < t_1 < \ldots < t_i < t_{i+\frac{1}{2}} < t_{i+1} \ldots < t^N = T,$$

with $t_{i+1} - t_i = \Delta t$ for all time steps and indicate the discrete time by a superscript at the quantities to use, e. g., the discrete electric field vector at time $t_{n+\frac{1}{2}}$ is $\mathbf{E}_\mathrm{h}^{n+\frac{1}{2}}$. Using a central finite difference approximation of the time derivative in Equation 52.24, we obtain its fully discretized version

$$\mathbf{E}_\mathrm{h}\big|_{\mathbf{e}_\mathrm{p}}^{n+\frac{1}{2}} = \mathbf{E}_\mathrm{h}\big|_{\mathbf{e}_\mathrm{p}}^{n-\frac{1}{2}} + \frac{\Delta t}{\varepsilon}\,\mathbf{curl}_\mathrm{h}\, \mathbf{H}_\mathrm{h}\big|_{\mathbf{f}_\mathrm{d}}^{n} \qquad \text{for all faces } \mathbf{f}_\mathrm{d} \text{ of the dual mesh.} \tag{52.26}$$

The time derivative is centered around n, hence it is of second-order accuracy. Using the updated electric field of Equation 52.26, Equation 52.25 is updated accordingly:

$$\mathbf{H}_\mathrm{h}\big|_{\mathbf{e}_\mathrm{d}}^{n+1} = \mathbf{H}_\mathrm{h}\big|_{\mathbf{e}_\mathrm{d}}^{n} - \frac{\Delta t}{\mu}\,\mathbf{curl}_\mathrm{h}\, \mathbf{E}_\mathrm{h}\big|_{\mathbf{f}_\mathrm{p}}^{n+\frac{1}{2}} \qquad \text{for all inner faces } \mathbf{f}_\mathrm{p} \text{ of the primal mesh.} \tag{52.27}$$

Equations 52.26 and 52.27 are typical FDTD update equations.

52.3.5 Dispersion, Stability, and Courant–Friedrichs–Lewy Condition

Let us consider a monochromatic plane wave of frequency ω in a homogeneous medium, with speed of light $c = 1/\sqrt{\varepsilon\mu}$, ε and μ independent of position, and a wavevector $\mathbf{k} = (k_x, k_y, k_z)$. Then the plane wave

$$\mathbf{E}e^{-i\omega t + i\mathbf{kr}}$$

with a constant vector **E** is a solution of Maxwell's equations without sources, if the dispersion relation

$$\omega^2 = \mathbf{k}^2 c^2 \tag{52.28}$$

and orthogonality condition

$$\mathbf{kE} = 0$$

hold true. The first one follows from the **curl**-**curl** equations, the second one from the divergence condition. For the Yee mesh, there is a discrete counterpart. The discrete plane wave satisfies Equations 52.26 and 52.27 takes the natural form

$$\mathbf{E}_{\mathrm{h}} e^{-i\omega t_j + i\mathbf{k}\mathbf{r}_{\mathrm{h}}}.$$

The corresponding discrete dispersion relation, however, contains the stepsizes in time, Δt, and space, h_x, h_y, h_z,

$$\frac{1}{\Delta t^2}\sin^2\frac{\omega\Delta t}{2} = c^2\left(\left(\frac{1}{h_x^2}\sin^2\frac{k_x h_x}{2}\right) + \left(\frac{1}{h_y^2}\sin^2\frac{k_y h_y}{2}\right) + \left(\frac{1}{h_z^2}\sin^2\frac{k_z h_z}{2}\right)\right). \tag{52.29}$$

For a derivation see, e. g.,[17] or [1]. Clearly, this discrete dispersion relation approaches Equation 52.28 as $\Delta t, h_x, h_y, h_z \to 0$ as it must be. But the sin-function of the left-hand side of Equation 52.29 has a further consequence. Let k_x, k_y, k_z be real, hence the plane wave is spatially bounded everywhere as in the case of the the continuous plane wave. Multiplying with Δt^2 from the left-hand side of Equation 52.29 might result in a real number larger than 1, if Δt is large enough. Then ω onthe left-hand side becomes imaginary, and since the expression involves $\sin^2$, both positive and negative imaginary frequencies occur. The positive one will cause an exponential increase of the discrete plane wave in time, hence the solution becomes unstable. Consequently, for stability we must require

$$c^2\Delta t^2\left(\frac{1}{h_x^2} + \frac{1}{h_x^2} + \frac{1}{h_z^2}\right) \le 1$$

or, for uniform meshes, $h = h_x = h_y = h_z$,

$$\Delta t \le \frac{h}{c\sqrt{3}}.$$

These inequalities are the Courant–Friedrichs–Lewy (CFL) conditions for the wave equation.

52.3.6 Discrete Div-Conditions

It is a remarkable property that the Yee scheme preserves the **div**-conditions of Equation 52.1 for homogeneous media in discrete form [18]. Computing the discrete divergence of the updated Equation 52.26

$$\mathbf{div}_{\mathrm{h}}\,\varepsilon\mathbf{E}_{\mathrm{h}}\big|_{\mathbf{e}_{\mathrm{p}}}^{n+\frac{1}{2}} = \mathbf{div}_{\mathrm{h}}\,\varepsilon\mathbf{E}_{\mathrm{h}}\big|_{\mathbf{e}_{\mathrm{p}}}^{n-\frac{1}{2}} + \Delta t\,\mathbf{div}_{\mathrm{h}}\;\mathbf{curl}_{\mathrm{h}}\,\mathbf{H}_{\mathrm{h}}\big|_{\mathbf{f}_{\mathrm{d}}}^{n} \quad \text{for all cells of the dual mesh.}$$

The last term vanishes due to property (Equation 52.23), hence we have the convergence

$$\mathbf{div}_{\mathrm{h}}\, \varepsilon \mathbf{E}_{\mathrm{h}}\big|_{\mathbf{e}_{\mathrm{p}}}^{n+\frac{1}{2}} = \mathbf{div}_{\mathrm{h}}\, \varepsilon \mathbf{E}_{\mathrm{h}}\big|_{\mathbf{e}_{\mathrm{p}}}^{n-\frac{1}{2}} = \ldots = \mathbf{div}_{\mathrm{h}}\, \varepsilon \mathbf{E}_{\mathrm{h}}\big|_{\mathbf{e}_{\mathrm{p}}}^{-\frac{1}{2}} \quad \text{for all cells of the dual mesh.}$$

The divergence attributed to a cell is preserved over all time steps. For the magnetic field, we obtain the corresponding result the same way:

$$\mathbf{div}_{\mathrm{h}}\, \mu \mathbf{H}_{\mathrm{h}}\big|_{\mathbf{e}_{\mathrm{d}}}^{n+1} = \mathbf{div}_{\mathrm{h}}\, \mu \mathbf{H}_{\mathrm{h}}\big|_{\mathbf{e}_{\mathrm{d}}}^{n} = \ldots = \mathbf{div}_{\mathrm{h}}\, \mu \mathbf{H}_{\mathrm{h}}\big|_{\mathbf{e}_{\mathrm{d}}}^{0} \quad \text{for all cells of the primal mesh.}$$

If the initial data at time step $n = -1/2$ and $n = 0$ are divergence-free,

$$\mathbf{div}_{\mathrm{h}}\, \varepsilon \mathbf{E}_{\mathrm{h}}^{-\frac{1}{2}} = 0 \qquad \mathbf{div}_{\mathrm{h}}\, \mu \mathbf{H}_{\mathrm{h}}^{0} = 0,$$

this property is preserved for all following time steps.

52.3.7 Convergence

Let, as before, $\mathbf{E}_{\mathrm{h}}$ be the vector of field values attributed to the edges of the primal mesh with values computed at the midpoint of edges. Let us denote the vector of the exact field values, projected to the edges, by $\mathbf{E}$; hence $\mathbf{E}_{\mathrm{h}}$ approximates $\mathbf{E}$ pointwise. We define the following, mesh-dependent discrete norm

$$\|\mathbf{E} - \mathbf{E}_h\|^2 \,:\, = h^3 \left(\mathbf{E}_{\mathrm{h}} - \mathbf{E}\right)^2,$$

and the analog for $\mathbf{H}$ with values attributed to the centroids of the cell faces.The following holds true. Let $\mathbf{E}$ and $\mathbf{H}$ be three times continuously differentiable in Ω. Then, there exists a constant C independent of h such that

$$\|\mathbf{E} - \mathbf{E}_{\mathbf{h}}\| + \|\mathbf{H} - \mathbf{H}_{\mathbf{h}}\| \leq Ch^2.$$

FDTD converges with second order. This classical result extends also to nonuniform meshes as shown by Monk and Süli [19].

52.4 Fourier Modal Method

This summary of FMMs follows Chapter 6 of [1]; for more details see that textbook and the references given therein.

52.4.1 Historical Notes

The FMM is a family of methods that belong to the class of spectral methods for the frequency domain and has been well established for more than three decades now. The term spectral method means that the field approximation is realized with functions which have a global support. In contrast, in FDTD methods, we consider only the pointwise-defined functions, and in FEM methods, we have basis functions with a very small, local support.

The origins of the FMM dates back to the 1960s and were devoted to grating analyses [20,21]. In 1981, Moharam and Gaylord established in their seminal paper [22] a computational framework they called rigorous coupled wave analysis (RCWA). Other synonyms or variations of the method include the PWE, the eigenmode expansion (EME), and eigenmode expansion technique (EET). A typical observation in RCWA methods has been that the convergence of TM fields is notably worse than that of TE fields. In 1996, Lalanne and Granet and coworkers [23,24] found a solution to overcome this TM convergence problem and in the same year Li published his factorization rules [25] formalizing the results. Currently, different proposals are being investigated to establish a fast convergent 3D FMM, e. g., [26].

52.4.2 Concept: Segmentwise Waveguide Treatment

The family of FMMs are characterized by the following building blocks:

1. The geometry is subdivided into a sequence of waveguide-type sections, see Figure 52.8b and c.
2. The field within each waveguide section is represented by up and down propagating modes as well as evanescent modes.
3. The transversal fields of the waveguide-type eigenmodes are computed in Fourier representation.
4. The interaction of the individual sections is realized by posing proper continuity conditions at the interfaces and a scattering-matrix formalism.

52.4.3 Transfer-Matrix Approach

We consider the fields in two subsequent layers i and $i+1$, cf. Figure 52.9. At the moment, we do not focus on the special structure of the electromagnetic field except that we can decompose it into upward ($\mathbf{E}^+$) and downward ($\mathbf{E}^-$) propagating parts with respect to the z-direction. Let $\mathbf{E}_i^\pm(x,z)$ be arbitrary normalized solutions of Maxwell's equations split into propagation directions and $a_i^\pm \in \mathbb{C}$ their weights. It holds that

$$\begin{array}{rcccc} \text{Field in layer } i & : & a_i^+\mathbf{E}_i^+(x,z) & + & a_i^-\mathbf{E}_i^-(x,z) \\ \Downarrow \text{ implies} & & & & \\ \text{Field in layer } i+1 & : & a_{i+1}^+\mathbf{E}_{i+1}^+(x,z) & + & a_{i+1}^-\mathbf{E}_{i+1}^-(x,z). \end{array} \tag{52.30}$$

Given the complete field in one layer, Maxwell's equations including continuity conditions enforce the complete field in the other layer. This can be expressed by a transfer-matrix relation of the coefficients

$$\mathbf{A}_{i+1}\begin{pmatrix} a_{i+1}^+ \\ a_{i+1}^- \end{pmatrix} = \mathbf{A}_i\begin{pmatrix} a_i^+ \\ a_i^- \end{pmatrix}, \tag{52.31}$$

where $\mathbf{A}_i$, $\mathbf{A}_{i+1}$ are complex 2×2 matrices. We introduced this symmetric form since the count sequence of layers (top-down or bottom-up) is arbitrary. The typical application of the 1D transfer-matrix method is, given n layers of finite thickness with indices 0 and $n+1$ referring to substrate and superstrate, the

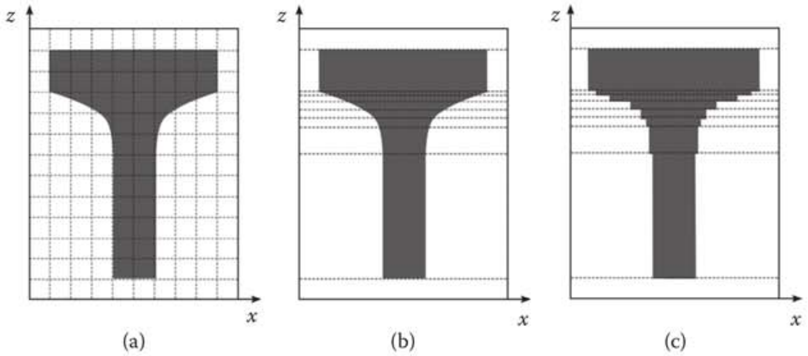


FIGURE 52.8 (a) A schematic geometry and a Cartesian mesh as used in FDTD methods. (b) Slicing of the geometry into segments along z. (c) Waveguide-like approximation of the geometry used in Fourier modal methods.

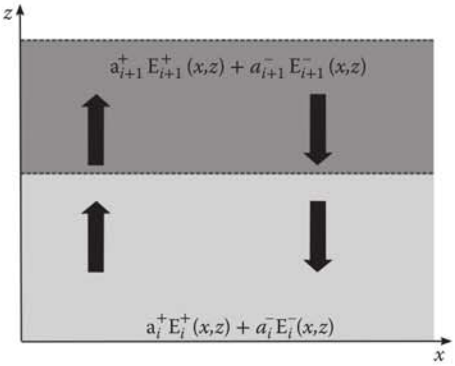


FIGURE 52.9 Superposition of up and down propagating fields in two subsequent layers.

update procedure

$$\begin{pmatrix} a_{n+1}^{+} \\ a_{n+1}^{-} \end{pmatrix} = \mathbf{T}_{n,n+1} \cdot \ldots \cdot \mathbf{T}_{i,i+1} \cdot \ldots \cdot \mathbf{T}_{0,1} \begin{pmatrix} a_{0}^{+} \\ a_{0}^{-} \end{pmatrix}.$$

This procedure may become unstable due to the exponentials involved in $T_{i,i+1}$.

52.4.4 Scattering-Matrix Approach

The scattering-matrix approach uses merely the same representation as the transfer-matrix approach, except that the computation of the coefficients $a_i^{\pm}$ is reordered (Figure 52.10). The scattering-matrix notation corresponding to Equation 52.31 is

$$\begin{pmatrix} a_{i}^{-} \\ a_{i+1}^{+} \end{pmatrix} = \mathbf{S}_{i,i+1} \begin{pmatrix} a_{i}^{+} \\ a_{i+1}^{-} \end{pmatrix} \tag{52.32}$$

with the obvious computation of the coefficients of $\mathbf{S}_{i,i+1}$ from $\mathbf{A}_i$ and $\mathbf{A}_{i+1}$. The notion scattering follows from the fact that we can consider the coefficients $\left(a_i^{+}, a_{i+1}^{-}\right)$ as fields *incident* to the layer separating boundary and the coefficients $\left(a_i^{-}, a_{i+1}^{+}\right)$ as fields *scattered* away from the boundary, respectively. The scattering matrix for an n-layer system is derived as follows. We rewrite Equation 52.31 as

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_i\left(d_i\right) & -\mathbf{A}_{i+1}(0) \end{pmatrix} \begin{pmatrix} a_{i}^{+} \\ a_{i}^{-} \\ a_{i+1}^{+} \\ a_{i+1}^{-} \end{pmatrix}.$$

Here, the top coefficient a_i^{+} of the column vector and the bottom coefficient a_{i+1}^{-} are the incident coefficients. We repeat this for the next layer, increasing the index i by 1

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_i\left(d_i\right) & -\mathbf{A}_{i+1}(0) & \\ & \mathbf{A}_{i+1}\left(d_{i+1}\right) & -\mathbf{A}_{i+2}(0) \end{pmatrix} \begin{pmatrix} a_{i}^{+} \\ a_{i}^{-} \\ a_{i+1}^{+} \\ a_{i+1}^{-} \\ a_{i+2}^{+} \\ a_{i+2}^{-} \end{pmatrix}.$$

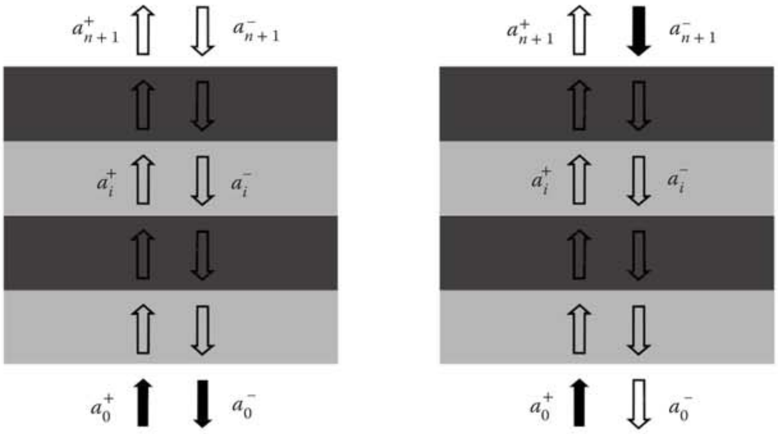


FIGURE 52.10 (a) Transfer-matrix approach. The data a_0^+, a_0^- are given (filled arrows), the others are computed. (b) Scattering-matrix approach. The data a_0^+, a_{n+1}^- are given, the others are computed.

Again, the top and bottom coefficients of the column vector are the incident coefficients of the layer system. Moving these coefficients to the left-hand side, deleting the first and last column of the system matrix, and inverting it yields the scattering-matrix formulation with scattering matrix **S**:

$$\mathbf{S}\begin{pmatrix} a_0^+ \\ 0 \\ \vdots \\ 0 \\ a_{n+1}^- \end{pmatrix} = \begin{pmatrix} a_0^- \\ a_1^+ \\ a_1^- \\ \vdots \\ \\ a_{n+1}^+ \end{pmatrix}. \tag{52.33}$$

52.4.5 Layered Media Stack as Model Case

A layered media stack is a 1D problem, if we choose the z-axis normal to the faces of the layers. To discuss the structure of FMM algorithms, we construct the standard transfer-matrix algorithm in scattering-matrix formulation from the viewpoint of the FMM. To find the field representation in layer i, we have to solve Maxwell's equations in layer i, i. e., either the equation for TE or TM. Let us consider the TE case. The modal ansatz

$$\text{layer } i: \quad E_y(x,z) = E(x)e^{ik_z z} = E_0 e^{ik_x x + ik_z z}$$

satisfies Equation 52.9 if $n_i^2 k_0^2 - k_x^2 = k_{z,i}^2$. Hence, given k_x, we have

$$\text{layer } i: \quad E_{y,i}(x,z) = a_i^+ e^{ik_x x + ik_{z,i} z} + a_i^- e^{ik_x x - ik_{z,i} z}, \tag{52.34}$$

which is the first line in Equation 52.30. Equally, we have for layer $i+1$

$$\text{layer } i+1: \quad E_{y,i+1}(x,z) = a_{i+1}^+ e^{ik_x x + ik_{z,i+1} z} + a_{i+1}^- e^{ik_x x - ik_{z,i+1} z}.$$

We compute matrices $\mathbf{A}_i$ and $\mathbf{A}_{i+1}$. Let the boundary between the two layers be placed at $z = 0$. The tangential electric and magnetic fields must be continuous. It follows for the electric field

$$a_i^+ e^{ik_x x} + a_i^- e^{ik_x x} = a_{i+1}^+ e^{ik_x x} + a_{i+1}^- e^{ik_x x} \quad \text{for all } x \in \mathbb{R},$$

hence

$$a_i^+ + a_i^- = a_{i+1}^+ + a_{i+1}^-. \tag{52.35}$$

Given the electric field in each layer, we compute the tangential component of the magnetic field

$$\begin{array}{rrcll} \text{layer } i: & \omega\mu H_{z,i}(x,z) & = & a_i^+ k_{z,i} e^{ik_x x + ik_{z,i} z} & - \quad a_i^- k_{z,i} e^{ik_x x - ik_{z,i} z} \\ \text{layer } i+1: & \omega\mu H_{z,i+1}(x,z) & = & a_{i+1}^+ k_{z,i+1} e^{ik_x x + ik_{z,i+1} z} & - \quad a_{i+1}^- k_{z,i+1} e^{ik_x x - ik_{z,i+1} z} \end{array}$$

and obtain the continuity condition

$$k_{z,i} a_i^+ - k_{z,i} a_i^- = k_{z,i+1} a_{i+1}^+ - k_{z,i+1} a_{i+1}^-. \tag{52.36}$$

Consequently, taking Equations 52.35 and 52.36,

$$\begin{pmatrix} 1 & 1 \\ k_{z,i+1} & -k_{z,i+1} \end{pmatrix} \begin{pmatrix} a_{i+1}^+ \\ a_{i+1}^- \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ k_{z,i} & -k_{z,i} \end{pmatrix} \begin{pmatrix} a_i^+ \\ a_i^- \end{pmatrix}. \tag{52.37}$$

Repeating this calculation with a finite layer thickness d_i between the two boundaries, we obtain

$$A_i(d_i) = \begin{pmatrix} 1 & 1 \\ k_{z,i} & -k_{z,i} \end{pmatrix} \begin{pmatrix} e^{ik_{z,i} d_i} & 0 \\ 0 & -e^{ik_{z,i} d_i} \end{pmatrix}. \tag{52.38}$$

Collecting the steps, we obtain the classical transfer-matrix algorithm in scattering formulation Algorithm 52.1. In 1D it is not necessary to save the eigenvectors since they are the same for all layers. But in 2D and 3D, the eigenvectors are different in each layer, so we introduced the step here for full parallelism with Algorithm 52.1.

Algorithm 52.1: Transfer-Matrix Algorithm in Scattering Formulation

```
Require: $a_0^+$ and $a_{n+1}^-$ to define source fields.
  for all layers do
    Solve each layer's eigenvalue problem (Equation 52.9).
    Save: eigenvectors $\exp(ik_x x)$ and eigenvalues $k_{z,i}$
    for later field representation (Equation 52.34).
    Compute the matrices $A_i$ by (Equation 52.38).
  end for
  Assemble the system (Equation 52.33).
  Solve.
  return coefficients $a_0^-, \dots, a_{n+1}^+$.
```

52.4.6 FMM in 2D

We generalize the foregoing, where ε has been constant in each layer. Now, let ε vary with x, each layer may have a different dependence $\varepsilon_i(x)$. The eigenvalue problems are Equations 52.9 and 52.10 with layerwise z-independent permittivities. In RCWA and FMM methods, these eigenvalue equations are solved based on a Fourier representation of the fields and the permittivities. This implies essentially two properties:

- Advantage: There is no lateral mesh needed. The nonlocal (or spectral) basis functions of fields and permittivities are simply constructed which results in a straightforward implementation.
- Disadvantage: The convergence of RCWA and FMM depend on the convergence properties of the Fourier representation which are not advantageous for nonsmooth functions.

52.4.6.1 Fourier Series Representation of Functions and Products of Functions

We consider a function $f(x)$ which is periodic with L: $f(x+L) = f(x)$. Hence, $f(x)$ has a Fourier series representation

$$f(x) = \sum_{m=-\infty}^{\infty} f_m e^{im\frac{2\pi}{L}x}. \tag{52.39}$$

We denote its Fourier transform by $\widehat{f}$

$$\widehat{f}(k) = \frac{1}{L}\int_0^L f(x)e^{-ikx}\,dx \text{ and compute } f_m = \widehat{f}\left(\frac{2\pi}{L}m\right), \quad m \in \mathbb{Z}.$$

We consider a second function $g(x)$, also periodic with interval L, and study the product $h(x) = f(x)g(x)$, which is also L-periodic and consequently has a Fourier representation

$$h(x) = \sum_{m=-\infty}^{\infty} h_m e^{im\frac{2\pi}{L}x}.$$

In terms of the Fourier coefficients of f and g, this is

$$\begin{aligned} h(x) &= \left(\sum_{m=-\infty}^{\infty} f_m e^{im\frac{2\pi}{L}x}\right)\left(\sum_{n=-\infty}^{\infty} g_n e^{in\frac{2\pi}{L}x}\right) = \sum_{m=-\infty}^{\infty}\left(\sum_{n=-\infty}^{\infty} f_{m-n}g_n\right)e^{im\frac{2\pi}{L}x} \\ h_m &= \sum_{n=-\infty}^{\infty} f_{m-n}g_n. \end{aligned} \tag{52.40}$$

We write the last equation in matrix-vector form

$$\begin{aligned} \mathbf{h} &= \mathbf{Fg} \\ \mathbf{F} &= \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & f_0 & f_{-1} & f_{-2} & f_{-3} & \cdots \\ \cdots & f_1 & f_0 & f_{-1} & f_{-2} & \cdots \\ \cdots & f_2 & f_1 & f_0 & f_{-1} & \cdots \\ \cdots & f_3 & f_2 & f_1 & f_0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}, \end{aligned}$$

where the convolution matrix $\mathbf{F}$ is called the Toeplitz matrix.

52.4.6.2 Fourier Series Convergence in Dependence of the Smoothness of the Functions

In electromagnetics, we have to deal with piecewise smooth functions which have different global smoothness properties. Roughly, we have three classes:

Discontinuous: Functions that are piecewise smooth, but may have jumps or singularities. Example: Normal field components in a 3D setting with material discontinuities.

Continuous: Functions that are continuous but not differentiable at positions, where different materials meet. Example: H_y in TM fields as solution of, e.g., Equation 52.10.

Smooth: Functions that are continuous *and* the first derivative exists everywhere, even at positions, where different materials meet. Example: E_y in s-polarization, e.g., Equation 52.9.

52.4.6.3 Truncation

For an algorithmic application, the infinite series has to be truncated. The truncated Fourier series (Equation 52.39) then reads

$$f^M(x) = \sum_{m=-M}^{M} f_m e^{im\frac{2\pi}{L}x}, \quad f_m = \frac{1}{L}\int_0^L f(x)e^{-ik_m x}\,dx, \quad k_m = \frac{2\pi}{L}m \tag{52.41}$$

and the symmetric truncation of (Equation 52.40) yields

$$h^M(x) = \sum_{m=-M}^{M}\left(\sum_{n=-M}^{M} f_{m-n}g_n\right)e^{im\frac{2\pi}{L}x}, \quad h_m^M = \sum_{n=-M}^{M} f_{m-n}g_n.$$

Not only does the number of terms of the convolution sum change, but the individual coefficients h_n^M change too. The truncation changes the problem from an infinite to a finite one. The finite dimensional problem will differ from the exact, infinite dimensional one, as always in numerics, and the quality of this difference depends on the nature of the applied series truncation. The truncated matrix-vector form reads

$$\mathbf{h}^M = \mathbf{F}^M\mathbf{g}^M \qquad \text{“direct rule”} \tag{52.42}$$

$$\mathbf{F}^M = \left(f_{m-n}\right)_{-M\le m,n\le M} \tag{52.43}$$

with F^M the finite dimensional Toeplitz matrix and $f_{m-n} = 0$ for $|m-n| > M$. As shorthand for $\mathbf{F}^M$ generated by coefficients f_m we also write $[\![f]\!] := \mathbf{F}^M$.

52.4.6.4 Li's Factorization Rules

It has turned out that the originally observed slow convergence rate of TM problems is related to the smoothness properties of the products of field and permittivity and a special, polarization-dependent treatment can be used to remove the difficulty [23–25]. To this end, we introduce the discontinuity-types shown in Figure 52.11:

1. Type: I The product $f(x)g(x)$ results in $h(x)$ which exhibits also jump discontinuities.
2. Type: II The product $f(x)g(x)$ yields a continuous function $h(x)$. The discontinuities are removed in $h(x)$.

For a formulation of the FMM algorithm, taking into account the accelerated TM technique, we need additionally, the finite Fourier series representation of the inverse function $f^{\dagger}$:

$$f^{\dagger}(x) := \frac{1}{f(x)}$$

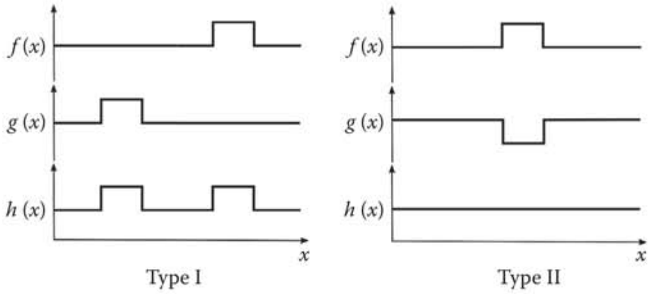


FIGURE 52.11 Two different types of jump discontinuities. Type I: The product $f(x)g(x)$ results in $h(x)$ which exhibits also jump discontinuities. Type II: The product $f(x)g(x)$ yields a continuous function $h(x)$.

together with its matrix-vector representation as counterpart to Equation 52.42

$$\mathbf{h}^M = \left(\mathbf{F}^{M\dagger}\right)^{-1}\mathbf{g}^M \qquad \text{“inverse rule.”} \tag{52.44}$$

The factorization rules are as follows:

1. Type: I If the discontinuities are not removable (Type I), apply the direct rule (Equation 52.42) for fastest convergence.
2. Type: II If the discontinuities are removable (Type II), apply the inverse rule (Equation 52.44) for fastest convergence.

52.4.6.5 Eigenvalue Problems in Fourier Space and Application of the Factorization Rules

The eigenvalue problems Equations 52.9 and 52.10 with ε independent of z hold true in each layer. We consider the TE equation Equation 52.9. Its continuous Fourier transform reads

$$\partial_x^2 E_y + \varepsilon(x)\mu\omega^2 E_y = k_z^2 E_y \quad \overset{\text{Fourier transform}}{\Longrightarrow} \quad (ik_x)^2\widehat{E}_y + \mu\omega^2\widehat{\varepsilon} * \widehat{E}_y = k_z^2\widehat{E}_y. \tag{52.45}$$

Restricting $E_y(x)$ to functions periodic over L (discretization in k-space), truncating the number of Fourier basis functions, and using the Toeplitz matrix to compute the convolution we obtain the eigenvalue equation in Fourier space

$$(i\mathbf{K}_x)^2\widehat{\mathbf{E}}_y + \mu\omega^2 \left[\!\left[\widehat{\varepsilon}\right]\!\right] \widehat{\mathbf{E}}_y = k_z^2\widehat{\mathbf{E}}_y. \tag{52.46}$$

Equation 52.46 yields the generalization of the plane waves (TE) used in Equation 52.34 for the derivation of the transfer-matrix algorithm. The same procedure applies to the TM case, where strict consideration of the factorization rules is important. We compute

$$\begin{aligned} &\varepsilon(x)\partial_x\frac{1}{\varepsilon(x)}\partial_x H_y + \varepsilon(x)\mu\omega^2 H_y = k_z^2 H_y. \\ \overset{\text{Fourier transform}}{\Longrightarrow} \quad &\widehat{\varepsilon} * (ik_x)\widehat{\left(\frac{1}{\varepsilon}\right)} * (ik_x)\widehat{H}_y + \mu\omega^2\widehat{\varepsilon} * \widehat{H}_y = k_z^2\widehat{H}_y. \end{aligned} \tag{52.47}$$

For proper discretization based on the factorization rules, we rewrite Equation 52.47, factoring out $\widehat{\varepsilon}$,

$$\widehat{\varepsilon} * \left((ik_x)\widehat{\left(\frac{1}{\varepsilon}\right)} * (ik_x) + \mu\omega^2 \right) \widehat{H}_y = k_z^2 \widehat{H}_y. \tag{52.48}$$

First, $\partial_x H_y$ and $1/\varepsilon$ are discontinuous, their product is continuous, hence factorization rule II applies. Second, the large parenthesis on the left-hand side of Equation 52.48 must be discontinuous since ε is discontinuous. The product, however, is continuous, since the right-hand side is continuous. Hence, the factorization rule type II applies again:

$$\left[\!\left[\widehat{\left(\frac{1}{\varepsilon}\right)}\right]\!\right]^{-1} \left(i\mathbf{K}_x \left[\!\left[\widehat{(\varepsilon)}\right]\!\right]^{-1} (i\mathbf{K}_x) + \mu\omega^2 \right) \widehat{\mathbf{H}}_y = k_z^2 \widehat{\mathbf{H}}_y. \tag{52.49}$$

52.4.6.6 Continuity Condition

We derive the continuity condition Equation 52.31 for the 2D case. In two subsequent waveguide sections, the multimode generalization of Equation 52.30 holds:

$$\begin{array}{rcccc} \text{field in section } i & : & \sum_m a_{i,m}^{+} \mathbf{E}_{i,m}^{+}(x,z) & + & \sum_m a_{i,m}^{-} \mathbf{E}_{i,m}^{-}(x,z) \\ \Downarrow \text{ implies} & & & & \\ \text{field in section } i+1 & : & \sum_m a_{i+1,m}^{+} \mathbf{E}_{i+1,m}^{+}(x,z) & + & \sum_m a_{i+1,m}^{-} \mathbf{E}_{i+1,m}^{-}(x,z). \end{array} \tag{52.50}$$

The generalization from plane waves (Equation 52.34) to waveguide modes reads

$$\text{section } i: \quad E_{y,i}(x,z) = \sum_m a_{i,m}^{+} \mathbf{E}_{i,m}(x) e^{ik_{z,i,m}z} + \sum_m a_{i,m}^{-} \mathbf{E}_{i,m}(x) e^{-ik_{z,i,m}z}. \tag{52.51}$$

Each mth normalized mode in section i has a Fouriermode representation $\mathbf{E}_{i,m} = \sum_l b_{i,m,l} \exp ik_{x,l}x$. Here, the weights $b_{i,m,l}$ (ith section, mth waveguide mode, lth Fouriermode of this waveguide mode) are, in case of TE fields, the coefficients of the eigenvectors (Equation 52.46), i.e., $b_{i,m,l} = \left(\widehat{\mathbf{E}}_y\right)_l$, if $\widehat{E}_y$ is the mth waveguide mode in the ith section. Hence,

$$E_{y,i}(x,z) = \sum_m a_{i,m}^{+} \sum_l b_{i,m,l} e^{ik_{x,i,m,l}x} e^{ik_{z,i,m}z} + \sum_m a_{i,m}^{-} \sum_l b_{i,m,l} e^{ik_{x,i,m,l}x} e^{-ik_{z,i,m}z}.$$

Setting the interface position between sections i and $i + 1$ to $z = 0$ and requiring continuity of E_y yields

$$\sum_{m=-M}^{M} b_{i,m,l} \left(a_{i,m}^{+} + a_{i,m}^{-} \right) = \sum_{m=-M}^{M} b_{i+1,m,l} \left(a_{i+1,m}^{+} + a_{i+1,m}^{-} \right), \qquad -M \le l \le M.$$

In the same way, one computes the 2D counterpart to Equation 52.36

$$\sum_{m=-M}^{M} b_{i,m,l} \left(k_{z,i,m} a_{i,m}^{+} - k_{z,i,m} a_{i,m}^{-} \right) = \sum_{m=-M}^{M} b_{i+1,m,l} \left(k_{z,i+1,m} a_{i+1,m}^{+} - k_{z,i+1,m} a_{i+1,m}^{-} \right)$$
$$-M \le l \le M$$

Introducing the matrix

$$\mathbf{B}_i = \left(b_{i,l,m}\right)^T = \begin{pmatrix} \widehat{\mathbf{E}}_{y,-M}, & \cdots & ,\widehat{\mathbf{E}}_{y,m}, & \cdots & ,\widehat{\mathbf{E}}_{y,M} \end{pmatrix}^T$$

allows writing the two continuity conditions in compact block matrix form:

$$\begin{pmatrix} \mathbf{B}_i & \\ & \mathbf{B}_i \end{pmatrix} \begin{pmatrix} \mathbf{I}_\mathrm{d} & \mathbf{I}_\mathrm{d} \\ \operatorname{diag} \mathbf{k}_{z,i} & -\operatorname{diag} \mathbf{k}_{z,i} \end{pmatrix} \begin{pmatrix} \mathbf{a}_i^+ \\ \mathbf{a}_i^- \end{pmatrix}$$
$$= \begin{pmatrix} \mathbf{B}_{i+1} & \\ & \mathbf{B}_{i+1} \end{pmatrix} \begin{pmatrix} \mathbf{I}_\mathrm{d} & \mathbf{I}_\mathrm{d} \\ \operatorname{diag} \mathbf{k}_{z,i+1} & -\operatorname{diag} \mathbf{k}_{z,i+1} \end{pmatrix} \begin{pmatrix} \mathbf{a}_{i+1}^+ \\ \mathbf{a}_{i+1}^- \end{pmatrix}$$

In this form, the similarity to the continuity condition Equation 52.37 of the transfer-matrix algorithm becomes obvious. For compact algorithmic notation, we define

$$\mathbf{A}_i := \begin{pmatrix} \mathbf{B}_i & \\ & \mathbf{B}_i \end{pmatrix} \begin{pmatrix} \mathbf{I}_\mathrm{d} & \mathbf{I}_\mathrm{d} \\ \operatorname{diag} \mathbf{k}_{z,i} & -\operatorname{diag} \mathbf{k}_{z,i} \end{pmatrix} \tag{52.52}$$

and have exactly as in the 1D case

$$\mathbf{A}_i \begin{pmatrix} \mathbf{a}_i^+ \\ \mathbf{a}_i^- \end{pmatrix} = \mathbf{A}_{i+1} \begin{pmatrix} \mathbf{a}_{i+1}^+ \\ \mathbf{a}_{i+1}^- \end{pmatrix}.$$

Collecting the steps, we obtain the FMM algorithm in 2D, Algorithm 52.2.

Algorithm 52.2: Fourier Modal Method in 2D

```
Require: $\mathbf{a}_0^+$ and $\mathbf{a}_{n+1}^-$ to define source fields.
  for all sections $i$ do
    Solve each sections's eigenvalue problem, either (Equation 52.46), TE, or (Equation 52.49), TM.
    Save: eigenvectors $\widehat{\mathbf{E}}_{y,i,m}$, eigenvalues $k_{z,i,m}$
    for later field representation (Equation 52.51).
    Compute the matrices $A_i$ by (Equation 52.52).
  end for
  Assemble the system (Equation 52.33).
  Solve.
  return vectors $\mathbf{a}_0^-, \ldots, \mathbf{a}_{n+1}^+$.
```

52.4.7 Convergence

The two significant approximation errors in FMM are the geometrical approximation by staircasing and the trunction in the number of Fourier modes. These problems have been addressed in numerous papers (see also the introduction). Especially, the representation of permittivities, which are not continuous in lateral direction, by global Fourier modes poses a convergence problem. Since in this case, the corresponding

Fourier coefficients decay only as $1/n$, where n is the number of the Fouriermode, the spatial approximation is worse than the $1/n^2$ we know, e. g., from the FDTD method. On the other hand, physics very often behaves in a way that only a few modes are sufficient to describe reasonably wave propagation and scattering. In these cases, FMM has been proven to be a very efficient method.

52.5 Finite Element Method

Several excellent books treating the finite element analysis of Maxwell's equations are available; the books of Peter Monk [27] and Leszek Demkowicz [28] are very recommendable. The summary of FEM here follows Chapter 8 of [1].

52.5.1 Historical Notes

The deep understanding of the relation between the scalar Helmholtz case, which enforces the use of classical, nodal based Lagrange shape functions, and the vectorial Maxwell case relies greatly on the work of Whitney and Nédélec [29,30], Bosssavit [31], and Webb [32]. Many important contributions of other authors have added to the field, which makes the FEM today a very flexible, effective, and well-studied method for electrodynamical simulations. An early survey of finite element studies with focus on nano-optical applications has been given in [33].

52.5.2 Concept: Patchwise Tangential Continuous, Polynomial Approximation on Unstructured Grids

The FEM is based on a *variational formulation* of Maxwell's equations that involves integral expressions on the computational domain. The FEM does not use any approximation to Maxwell's equations itself, except the discretization of the geometry, but approximates the solution space in which one tries to find a reasonable approximation to the exact solution. This is in contrast to the finite difference method which approximates Maxwell's equations directly via an approximation of the differential operators by finite difference stencils. In some respects, the FEM is comparable to FMMs: one takes a number of trial functions (the Fourier modes in case of FMM) and asks in which way a superposition of them gives a reasonable approximation to the exact solutions.

The solution space of FEM, however, is obtained by subdividing the computational domain into small patches and by providing a number of polynomials on each patch for the approximation of the solution. The patches together with the local polynomials defined on them are called *finite elements*. The most common examples of finite elements are triangles and rectangles in 2D and tetrahedrons and cubes in 3D together with constant, linear, quadratic, and cubic polynomials. These locally defined polynomial spaces have to be pieced together to ensure tangential continuity of the electric and magnetic field across the boundaries of neighboring patches. Once these local approximations with proper continuity conditions have been defined, these are inserted into the variational equation. This results in a linear system whose solution is a piecewise polynomial approximation of the exact solution. Hence, the two basic ingredients in each FEM for Maxwell's equations are

1. A variational formulation of Maxwell's equations.
2. A suitable construction of finite elements based on polynomials defined on local geometric patches to transform the variational formulation into a discrete, algebraic problem.

This concept has a number of remarkable properties:

- Complex geometrical shapes can be treated without geometrical approximations, e.g., rounding can be well and easily approximated, cf. Figure 52.12.

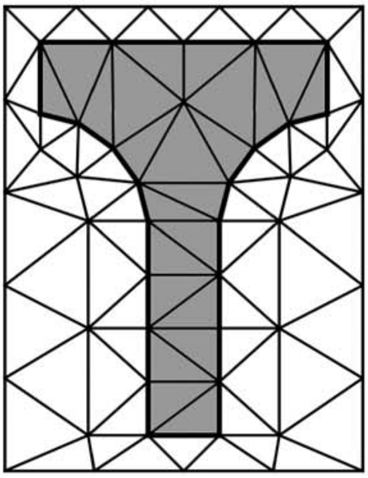


FIGURE 52.12 Sketch of a 2D triangulation for FEM, cf. also Figure 52.8 with staircase approximation. The triangles are shaped and placed such that the geometry is well covered by relatively few triangles.

- The finite element mesh can easily be adapted to the behavior of the solution, e.g., to singularities at corners.
- High-order approximations are available and ensure fast convergence.

52.5.3 1D: Weak Form and Variational Formulation

For the introduction of the finite element concept, we start with the most simple approximation of the time-harmonic scattering problem—the scalar Helmholtz equation in 1D, defined on the whole real axis, see Figure 52.3. The exterior domain should have a spatially invariant k. We let our computational domain Ω be the interval $(0, L)$. The source field is a plane wave coming from the left. A possible problem formulation reads: Find $u(x)$, $x \in \Omega = (0, L)$ such that

$$\partial_x^2 u(x) + k^2 u(x) = 0, \qquad x \in \mathbb{R} \tag{52.53}$$

$$u_{\text{source}}(x) = e^{ikx}, \qquad x \in \mathbb{R} \setminus \Omega. \tag{52.54}$$

Let us choose the interval $(0, L)$. We multiply Equation 52.53 with the complex conjugate of a so-called test function $v(x)$ and integrate over the domain Ω to arrive at

$$\int\limits_0^L v^*(x) \left(\partial_x^2 u(x) + k^2 u(x)\right) dx = 0 \tag{52.55}$$

where nothing has been specified so far about $v(x)$, except that we must be allowed to compute the integral. This formulation cannot be equivalent to Equation 52.53, since very special choices of v would equate the integral to zero even for functions, which do not satisfy Equation 52.53. Hence, in order to have the chance to get something equivalent to Equation 52.53, Equation 52.55 must hold true for many functions $v(x)$. We call the class of functions from which we take $v(x)$, the *test space*. In contrast, the space of functions containing the desired solution $u(x)$ is called the *trial* or *ansatz space*. In general, both spaces can be different. It is subject to the convergence theory of finite elements to define precisely the test and trial spaces and to conclude convergence rates. For our purpose here, it is sufficient to define these spaces in a way that makes the construction of the numerical algorithms reasonable. In the first attempt, we require $\int_0^L v^* v \, dx < \infty$, which is the class of square-integrable functions, the $L^2(0, L)$-functions. We will make the choice of the space more precise below, but only in a framework that introduces usual FEM notions and gives a first look at the topic. A comprehensive treatment can be found, e. g., in [27]. The functions $u(x)$ should be twice differentiable since the second derivative is contained in Equation 52.53. Hence, we seek them in the class $C^2(0, L)$. So in a first attempt to construct a variational equation, we formulate find a function $u(x) \in C^2(0, L)$ such that Equation 52.55 holds true for all $v \in L^2(0, L)$.

52.5.3.1 Weak Form

We perform an integration by parts on the left-hand side of Equation 52.55:

$$\int_0^L v^*(x) \left(\partial_x^2 u(x) + k^2 u(x)\right) dx = \int_0^L \left(-\partial_x v^*(x) \partial_x u(x) + k^2 v^*(x) u(x)\right) dx + \left. v^*(x) \partial_x u(x) \right|_0^L . \tag{52.56}$$

It is convenient to abbreviate the negative of the integral without boundary terms by

$$a_{\text{int}}(v, u) := \int_0^L \left(\partial_x v^*(x) \partial_x u(x) - k^2 v^*(x) u(x)\right) dx \tag{52.57}$$

to rewrite (Equation 52.56) to

$$\int_{\Omega_{\text{int}}} v^*(x) \left(\partial_x^2 u(x) + k^2 u(x)\right) dx =: -a_{\text{int}}(v, u) + \sum_{x \in \partial\Omega_{\text{int}}} v^*(x) \partial_n u(x). \tag{52.58}$$

The term Equation 52.57 is called a *sesquilinear form* because it is antilinear in its first argument (the v, due to complex conjugation) and linear in its second argument (the u). Hence, $a_{\text{int}}(v, u)$ maps two complex functions via integration to a complex number. We may repeat this for the scattered part of the exterior field

$$a_{\text{int}}(v, u) - \sum_{x \in \partial\Omega_{\text{int}}} v^*(x) \partial_n u(x) = 0 \quad \text{in } \Omega_{\text{int}} \tag{52.59}$$

$$a_{\text{ext}}(v, u_s) - \sum_{x \in \partial\Omega_{\text{int}}} v^*(x) \partial_n u_s(x) = 0 \quad \text{in } \mathbb{R} \setminus \Omega_{\text{int}}, \tag{52.60}$$

where we ensure (see discussion below) that the boundary term at infinity vanishes and only the boundary $\partial\Omega_{\text{int}}$ contributes to the equation. Using on this boundary, $u = u_{\text{source}} + u_s$ and introducing an auxiliary function g with

$$\begin{aligned} u_s &= \underbrace{u_s + g}_{w} - g \\ g(\partial\Omega_{\text{int}}) &= u_{\text{source}}(\partial\Omega_{\text{int}}) \\ g(\infty) &= 0 \end{aligned} \tag{52.61}$$

we obtain

$$a_{\mathrm{int}}(v,u) - \sum_{x \in \partial\Omega_{\mathrm{int}}} v^*(x)\partial_n\left(u_{\mathrm{source}} + u_{\mathrm{s}}\right) = 0 \quad \text{in } \Omega_{\mathrm{int}} \tag{52.62}$$

$$a_{\mathrm{ext}}(v, w-g) - \sum_{x \in \partial\Omega_{\mathrm{int}}} v^*(x)\partial_n u_{\mathrm{s}}(x) = 0 \quad \text{in } \mathbb{R} \setminus \Omega_{\mathrm{int}}. \tag{52.63}$$

Adding these two equations, taking into account the different normal directions, we end up with

$$a_{\mathrm{int}}(v,u) + a_{\mathrm{ext}}(v,w) = \sum_{x \in \partial\Omega_{\mathrm{int}}} v^*(x)\partial_n u_{\mathrm{source}} + a_{\mathrm{ext}}(v,g). \tag{52.64}$$

Next, using the function u, which lives only on the interior domain, and the function w, which lives only on the exterior domain, we define a joint, continuous function, again called u. This gives rise to the variational formulation of the scattering problem.

52.5.3.2 Variational Formulation

Let the source field data u_{source} and $\partial_n u_{\mathrm{source}}$ at the boundary $\partial\Omega$ of the computational domain Ω be given. The variational form of the scattering problem reads

$$\text{Find } u \in V: \quad a(v,u) = \sum_{x \in \partial\Omega_{\mathrm{int}}} v^*(x)\partial_n u_{\mathrm{source}} + a_{\mathrm{ext}}(v,g) \quad \text{for all } v \in V, \tag{52.65}$$

with $a(\cdot,\cdot) = a_{\mathrm{int}}(\cdot,\cdot) + a_{\mathrm{ext}}(\cdot,\cdot)$, with a_{int} and a_{ext} given by Equation 52.59, Equation 52.60 and g defined by Equation 52.61. Note that in Equation 52.60 we neglected the contribution from the outer boundary $\partial\Omega$ext. The reason is that the expression is established for the scattered field u_{s} and we perform the integration over the complex extended exterior domain as described in Section 52.2.4. The complex extension leads to an exponential decay of u_{s}, hence the contribution of $\partial\Omega$ext vanishes.

52.5.4 1D: Linear Finite Elements, Local Matrices, and Assembly Process

The discretization of a given variational form to get the FEM equations is a straight forward procedure: just restrict the infinite dimensional test and trial space V to a finite dimensional space V_h:

$$\begin{array}{rcll} \text{Find } u \in V: \ a(v,u) & = & b(v) & \text{for all } v \in V. \\ & \downarrow & & \\ \text{Find } u_h \in V_h: \ a(v_h,u_h) & = & b(v_h) & \text{for all } v_h \in V_h. \end{array} \tag{52.66}$$

The space V_h is not uniquely predetermined. The different types of Galerkin-type methods differ in the construction of the elements of V_h. Let V_h be spanned by N linearly independent functions $v_j(x)$, $j = 1, \ldots, N$. We represent the discrete approximation of the exact solution u by $u_h(x) = \sum_{j=1}^{N} u_j v_j(x)$, insert this into Equation 52.66:

$$a\left(v_h, \sum_{j=1}^{N} u_j v_j(x)\right) = b(v_h) \tag{52.67}$$

and get the linear system

$$i = 1, \dots, N: \qquad \sum_{j=1}^{N} a_{ij} u_j = b_i \qquad \text{with } a_{ij} = a\left(v_i, v_j\right) \quad \text{and } b_i = b\left(v_i\right). \tag{52.68}$$

52.5.4.1 Linear Finite Elements

We want to construct the matrix $\left(a_{ij}\right)$ as a sparse matrix, i.e., a matrix with predominantly zero entries, which is a very useful numerical property. To this end we construct test functions v_j that exist only on small patches. The most common finite elements are the *piecewise linear "hat"*-functions shown in Figure 52.13a:

$$v_i(x) = \begin{cases} \frac{x - x_{i-1}}{h_{i-1}} & \text{for} \quad x \in \left[x_{i-1}, x_i\right] \\ -\frac{x - x_{i+1}}{h_i} & \text{for} \quad x \in \left[x_i, x_{i+1}\right] \\ 0 & \text{elsewhere,} \end{cases} \tag{52.69}$$

where we introduced the lengths of the subdomains with $h_{i-1} = x_i - x_{i-1}$ and $h_i = x_{i+1} - x_i$.

52.5.4.2 Assembly Process: Patchwise Computation of $A = \left(a_{ij}\right)$

We start from Equation 52.67 with the summation moved to front of $a(\cdot, \cdot)$:

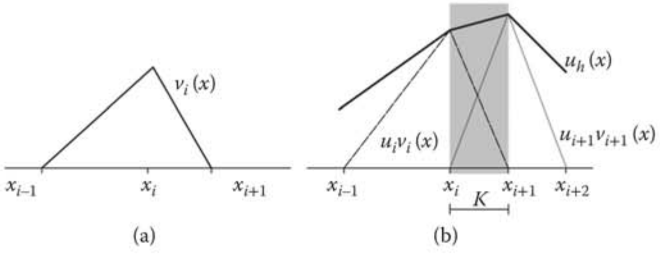
$$i = 1, \dots, N \qquad \sum_{j=1}^{N} a\left(v_i(x), v_j(x)\right) u_j = b(v_i).$$

Next we denote $a(\cdot, \cdot)$ restricted to patch K by $a(\cdot, \cdot)|_K = a_k(\cdot, \cdot)$ and rewrite the above in terms of a sum over patches:

$$i = 1, \dots, N \qquad \sum_{j=1}^{N} \sum_{k} a_k\left(v_i(x), v_j(x)\right) u_j = b(v_i).$$

Interchanging the order of summation we have

$$i = 1, \dots, N \qquad \sum_{k} \sum_{j=1}^{N} a_k\left(v_i(x), v_j(x)\right) u_j = b(v_i).$$



(a) (b)

FIGURE 52.13 (a) Piecewise linear test function $v_i(x)$. (b) Piecewise linear function u_h composed from trial functions $v_i(x)$ and $v_{i+1}(x)$. Only trial functions v_i and v_{i+1} have support on the interval K, but no other trial function.

In principle, this is a summation over k different matrices $A_k = a_k\left(v_i(x), v_j(x)\right)$, each of dimension $N \times N$ multiplied with a vector $u = (u)_i$ of dimension N, hence

$$\left(\sum_k A_k\right) u = b(v_i).$$

The matrices A_k have full dimension. However, since they are computed only over a single patch K they have entries different from zero only if both v_i and v_j have support over K, hence most entries but a few are zero. Suppose we have a set of (few) trial functions $v_{i_1}, v_{i_2}, \dots, v_{i_n}$ with index set $\{i_1, i_2, \dots, i_n\}$ which contribute to A_k assigned to patch K. We count these indices as follows:

$$\begin{array}{ccc} \text{local DOFs} & \rightarrow & \text{global DOFs} \\ (1, 2, \dots, n) & \rightarrow & \left(i_1, i_2, \dots, i_n\right) \end{array}$$

and store the connectivity data, using MATLAB ® notation, in an element connectivity vector

$$c = [\mathrm{i}_1, \mathrm{i}_2, \dots, \mathrm{i}_n]. \tag{52.70}$$

We have such a mapping for each patch connecting the *local* DOFs to the *global* ones. With the help of the local DOFs we define dense, so-called *local* matrices $A_k^{\mathrm{loc}} \in \mathbb{C}^{n \times n}$ of small dimension n and corresponding *local* vectors $b_k^{\mathrm{loc}} \in \mathbb{C}^n$:

$$\begin{array}{rcl} A_k^{\mathrm{loc}}(l, m) & := & A_k\left(i_l, i_m\right) \quad \text{for all indices with non-vanishing values of } A_k, \text{ and} \\ b_k^{\mathrm{loc}}(l) & := & b_k(i_l). \end{array}$$

The process of distributing the entries of the local matrices to the global matrices is called index scattering. The MATLAB ® statement

$$A(c, c) = A^{\mathrm{loc}}$$

scatters the entries of the local matrix A^{loc} to their global counterparts.

52.5.4.3 Computation of Local Matrices

For a convenient computation of the local vectors and matrices, it is useful to decompose the form $a(v, u)$, e.g., Equation 52.57, into the so-called *stiffness* and *mass* parts $s(v, u)$, $m(v, u)$, respectively:

$$a(v, u) = s(v, u) - m(v, u) \quad \text{with}$$

$$s(v, u) := \int_0^L \partial_x v^* \partial_x u \quad \text{and} \quad m(v, u) := \int_0^L k^2 v^* u.$$

The terms stiffness and mass come from mechanical systems and have no meaning in electrodynamics. But, as we will see, both parts have very different mathematical properties.

Let us compute the stiffness matrix based on the hat-functions. We consider the patch K from Figure 52.13 which gives us the local⟶global assignment of indices

$$(1, 2) \longrightarrow (i, i + 1).$$

Using Equation 52.69, we get

$$s^{\text{loc}} = \begin{pmatrix} s\left(v_i, v_i\right)\big|_K & s\left(v_i, v_{i+1}\right)\big|_K \\ s\left(v_{i+1}, v_i\right)\big|_K & s\left(v_{i+1}, v_{i+1}\right)\big|_K \end{pmatrix}.$$

We compute directly

$$s^{\text{loc}} = \frac{1}{h_i}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad \text{and the in same way} \quad m^{\text{loc}} = h_i \begin{pmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} \end{pmatrix}.$$

The structure of a finite element algorithm is given in Algorithm 52.3.

52.5.5 2D and 3D: Weak Form and Variational Formulation

The Maxwell scattering problem with PML is derived in complete analogy to Section 52.5.3. Let the source field data $\mathbf{E}_{\text{source}}$ and $\nabla \times \mathbf{E}_{\text{source}}$ at the boundary $\partial\Omega$ of the computational domain Ω be given. The variational form of the scattering problem reads:

$$\text{Find } \mathbf{E} \in V: \quad a(\mathbf{v}, \mathbf{E}) = \int_{\partial\Omega} d\mathbf{s} \cdot \left(\mathbf{v}^* \times \left[\left(\mu^c\right)^{-1} \nabla \times \mathbf{E}_{\text{source}}\right]\right) + a_{\text{ext}}\left(\mathbf{v}, \mathbf{g}\right) \quad \text{for all } \mathbf{v} \in V, \tag{52.71}$$

with

$$\begin{aligned} a(\cdot,\cdot) &= a_{\text{int}}(\cdot,\cdot) + a_{\text{ext}}(\cdot,\cdot) \\ a_{\text{int}}(\mathbf{v}, \mathbf{E}) &= \int_{\Omega} dx\, dy\, \left(\nabla \times \mathbf{v}^*\right) \cdot \mu^{-1} \left(\nabla \times \mathbf{E}\right) - \omega^2 \int_{\Omega} dx\, dy\, \mathbf{v}^* \varepsilon \mathbf{E} \\ a_{\text{ext}}(\mathbf{v}, \mathbf{g}) &= \int_{\Omega_{\text{ext}}} dx\, dy\, \left(\nabla \times \mathbf{v}^*\right) \cdot \left(\mu^c\right)^{-1} \left(\nabla \times \mathbf{g}\right) - \omega^2 \int_{\Omega_{\text{ext}}} dx\, dy\, \mathbf{v}^* \varepsilon^c \mathbf{E} \end{aligned}$$

and $\mathbf{g}$ is an arbitrary tangentially continuous function with

$$\begin{aligned} \mathbf{g}(\partial\Omega) &= \mathbf{E}_{\text{source}}(\partial\Omega) \\ \mathbf{g}(\infty) &= 0. \end{aligned}$$

The transformed permeabilities and permittivities are

$$\left(\mu^c\right)^{-1} = \frac{1}{|J|} J^T \mu^{-1} J, \qquad \varepsilon^c = |J|\, J^{-1} \varepsilon J^{-T},$$

with J the Jacobian of the complex extension. Here, we let Ω_{ext} be the original, real exterior domain and the PML-complexification is hidden in the transformed materials μ^c and ε^c, see Equations 52.18 and 52.19.

52.5.6 2D and 3D: Linear Finite Elements, Local Matrices, and Assembly Process

The basic scheme of deriving the numerical algorithm is exactly the same as in 1D, see Section 52.5.4, and also Algorithm 52.3:

1. Construct a mesh that describes the geometry plus surrounding domain. The meshes in 2D are triangular or quadrilateral meshes, in 3D these are meshes based on tetrahedrons, bricks, prisms, and pyramids. Typically, special mesh generators are employed. The construction of high-quality meshes is often a nontrivial task and a research field in its own right.
2. On each patch of the mesh, say on a triangle, a polynomial space is defined which is the basis of the construction of so-called shape functions. The local space of shape functions must be capable of approximating the solutions on the patch. Given a polynomial space, different families of shape functions can be constructed which all span the same space. A local matrix and a local right-hand side vector is attributed to each patch. The computation of these involves mapping from reference elements.
3. The shape functions of neighboring patches are glued together to build the discrete test and trial functions v_i. These functions must be tangentially continuous across the patch boundaries to be a global basis function of the solution space for Maxwell's equations. One DoF is attributed to each v_i.
4. Having obtained all test and trial functions v_i, the assembly process is carried out as in 1D, see Equations 52.67 and 52.68. The scattering of local to global matrices depends on the neighborhood relation of patches.

Algorithm 52.3: Basic Finite Element Algorithm

```
construct a tessellation $\bar{\Omega} = \cup_i \bar{K}_i$
**for all** patches $K_i$ **do**
    map $\widehat{K} \to K_i$
    compute local matrices $\mathbf{A}_i^{\mathrm{loc}}$
    compute local vectors $\mathbf{b}_i^{\mathrm{loc}}$
    scatter the local matrices and vectors to global ones
    update global matrices $\mathbf{A}$ and right-hand side $\mathbf{b}$
**end for**
solve $\mathbf{A}\mathbf{u}_h = \mathbf{b}$
**return** $\mathbf{u}_h$
```

52.5.6.1 Finite Elements

Figure 52.14 recalls the construction of the global test and trial functions from the local shape functions. The process is the same in higher dimensions, but there are many more possibilities for how to do it.

We illustrate these possibilities by means of some examples in 2D. First, we consider a mesh built from quadrilaterals, see Figure 52.15a. For a technical simplification of finite element algorithms, each quadrilateral is considered to be a map from a unit square, see Figure 52.16. The test and trial functions are constructed over the unit square and then mapped to their original positions. The lowest order trial functions are linear functions, in the case of a unit square, we use simply tensor products of 1D linear functions, see Section 52.5.4.

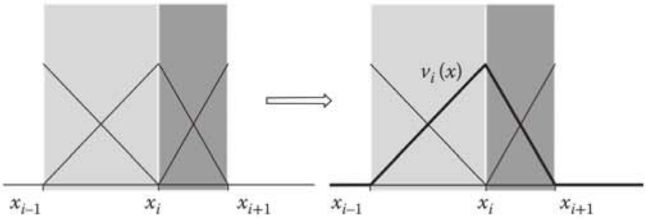


FIGURE 52.14 Construction of a global trial function v_i. Two neighboring intervals with two linear shape functions, each have the common boundary point x_i. One of the shape functions of each interval is taken and glued to the other and continued to a global, continuous trial function with local support.

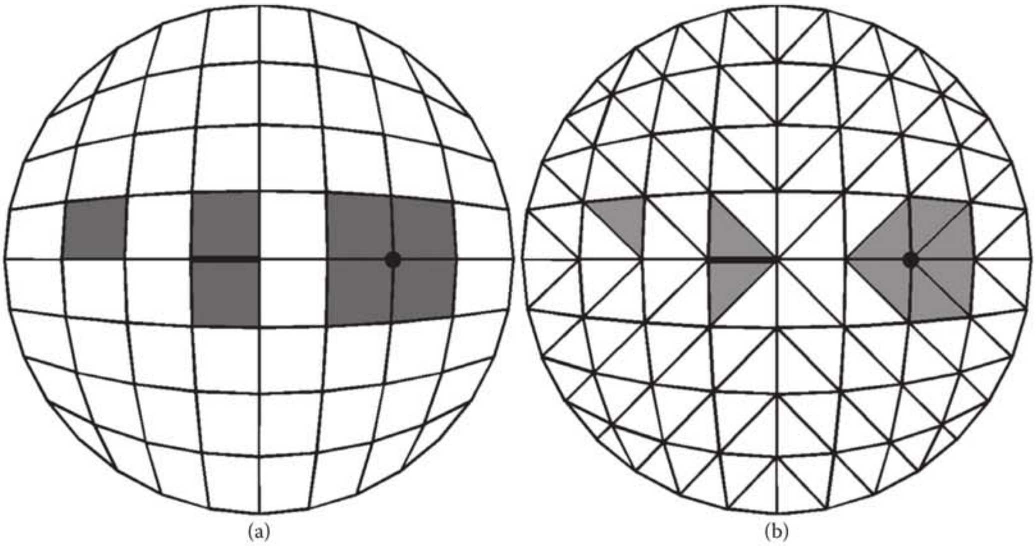


FIGURE 52.15 Structured mesh based on quadrilaterals (a) and a triangular mesh with the same vertices (b). The gray shaded areas show the geometric possibilities of constructing trial functions based on the interior of a single patch, neighboring patches with a common edge and neighboring patches with a common vertex.

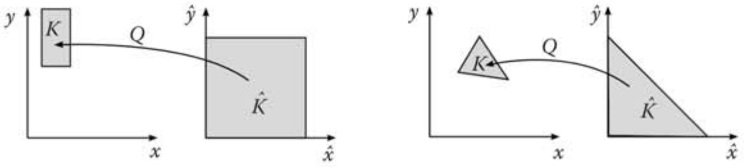
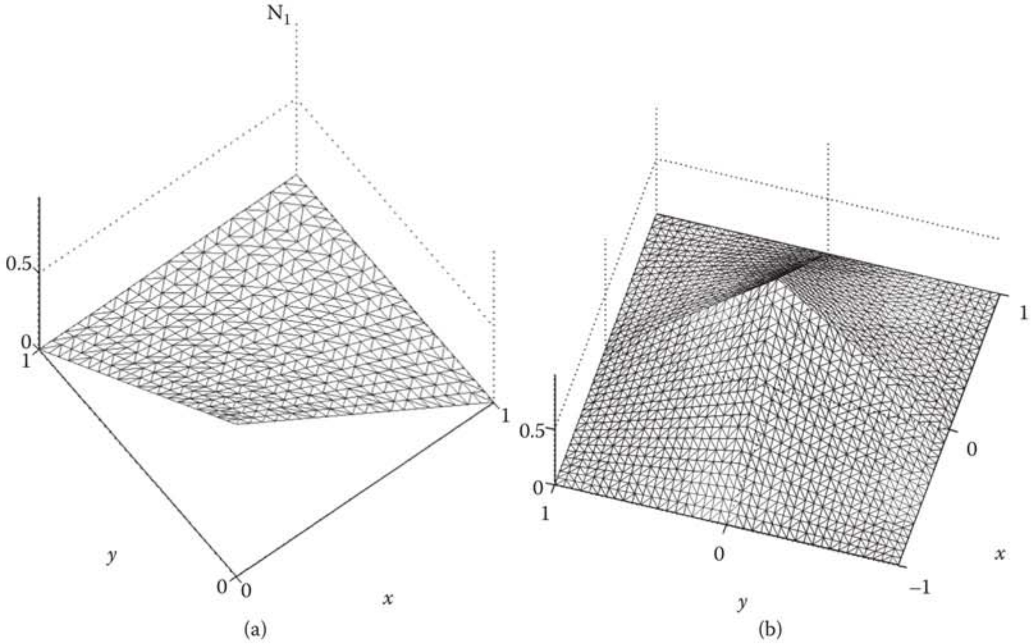


FIGURE 52.16 Mappings from reference elements to desired shapes.

52.5.6.2 Scalar Helmholtz Equation

Scalar trial functions v_i are used to approximate the solution of Helmholtz equations (Equation 52.8). In the case of a quadrilateral mesh with a unit square as reference element, the tensor product of the 1D shape functions gives four bilinear shape functions:

$$\{N_i : i = 1, \ldots, 4\} = \{(1-\hat{x})(1-\hat{y}), \hat{x}(1-\hat{y}), \hat{x}\hat{y}, (1-\hat{x})\hat{y}\}.$$



(a) (b)

FIGURE 52.17 Bilinear shape function N_1 on a unit square and the composed trialfunction.

Figure 52.17a shows the shape function N_1, which is associated to the vertex $(0, 0)$ of the reference coordinate system $\widehat{K}$. In the spirit of the 1D construction (Figure 52.14), neighboring shape functions will be connected to continuous, global trial functions v_i as shown in Figure 52.17b. The common vertex is the defining feature of the resulting global trial function. The DoF attributed to this function is usually taken as its nodal value at the vertex position.

52.5.6.3 Maxwell's Equations

In the case of Maxwell's equations, the fields are generally no longer continuous. They are tangentially continuous but may have jumps or even singularities in normal field components at material discontinuities. Hence, nodal values, at least along material discontinuities, do not represent the physical situation properly. Therefore, the lowest order vectorial shape functions for Maxwell'equations are so-called edge functions. They are constructed in a way that they allow for an easy extension to tangential continuous, global trial functions. The four lowest order edge functions attributed to the unit square are shown in Figure 52.18.

52.5.6.4 Computation of Local Matrices and Assembly Process

This is very similar to the procedure discussed for the 1D Helmholtz problem Section 52.5.4. According to the Maxwell scattering formulation (Equation 52.71), we have to evaluate the variational formulation (Equation 52.71) on a finite dimensional test and trial space V_h. The computation of the local matrices and vectors goes the same way as in Section 52.5.4 with

$$a_{i,j} = a\left(\mathbf{v}_i, \mathbf{v}_j\right) = \underbrace{\int\limits_{\Omega_{\text{cd}}} dx\, dy\, \left(\nabla \times \mathbf{v}_i^*\right) \cdot \mu^{-1} \left(\nabla \times \mathbf{v_j}\right)}_{\text{stiffness part } s_{ij}} - \omega^2 \underbrace{\int\limits_{\Omega_{\text{cd}}} dx\, dy\, \mathbf{v}_i^* \varepsilon \mathbf{v}_j}_{\text{mass part } m_{ij}}$$

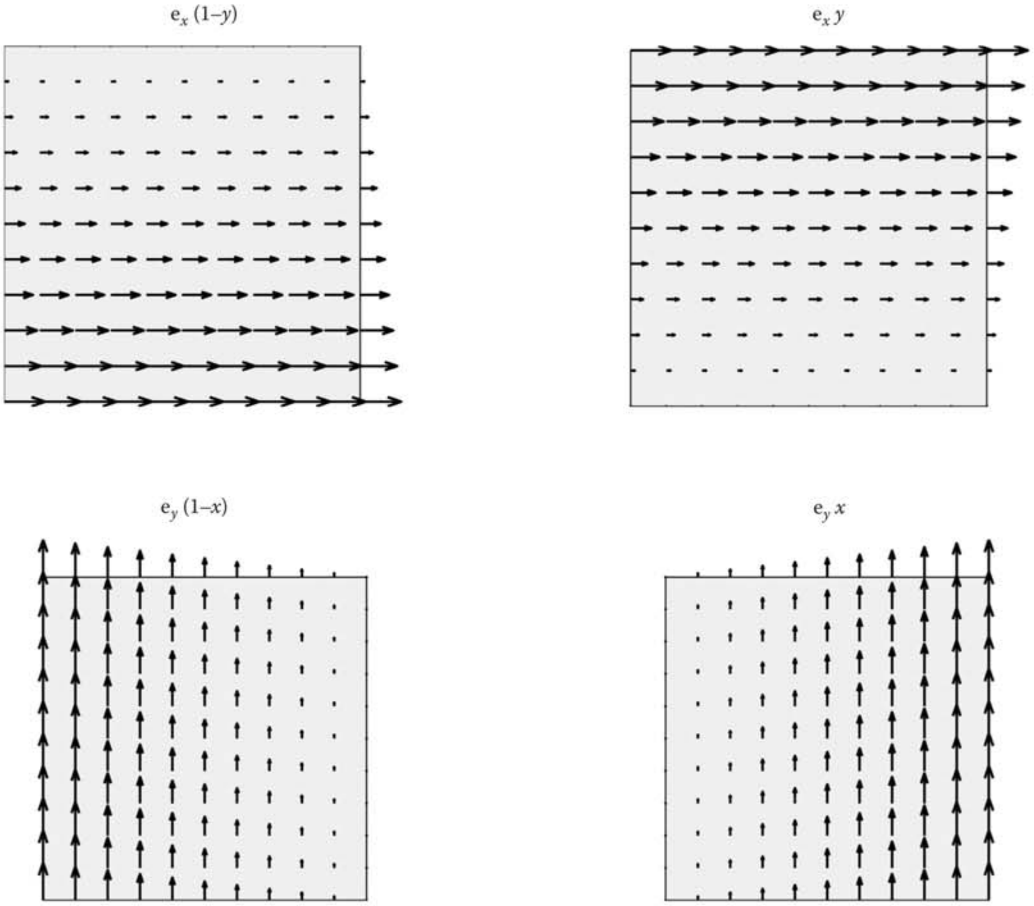


FIGURE 52.18 Vectorial, first-order edge elements of the unit square.

and

$$(b)_i = \int_{\partial\Omega} d\mathbf{f}\cdot\left(\mathbf{v}_i^*\times\left[\mu^{-1}\nabla\times\mathbf{E}_{\text{source}}\right]\right) + a_{\text{ext}}\left(\mathbf{v}_i,\mathbf{g}\right). \tag{52.72}$$

52.5.7 Convergence

The convergence analysis for edge element discretizations of time-harmonic scattering problems yields very useful insights into the behavior of numerical codes. The analysis itself is a challenging task since even in the lossless case the variational equations are neither elliptic nor definite. As a typical example we refer to Monk [34], who investigated a lossless scattering problem with sources on a bounded domain with electric boundary conditions. The error estimation uses the **curl** -norm

$$\|\mathbf{E}-\mathbf{E}_h\|^2_{\mathbf{curl}} := \|\mathbf{E}-\mathbf{E}_h\|^2 + \|\nabla\times(\mathbf{E}-\mathbf{E}_h)\|^2,$$

where $\|\cdot\|$ is the standard L^2-norm, h denotes again a typical meshsize and as subscript it indicates discrete approximation. Depending on the properties of the computational domain and the smoothness of the exact solution there are constants C and δ independent of h and a constant $h_0 > 0$ independent of $\mathbf{E}$ and $\mathbf{E}_h$ such that for all $0 < h < h_0$

$$\|\mathbf{E} - \mathbf{E}_h\|_{\mathbf{curl}} \leq \frac{1}{1 - Ch^{1/2+\delta}} \inf_{\mathbf{v}_h \in V_h} \|\mathbf{E} - \mathbf{v}_h\|_{\mathbf{curl}}.$$

This result shows that for h less than a threshold value h_0, the FEM scheme converges to the exact solution, but in the worst case without a polynomial rate. However, if the solution is smooth quantified as $\mathbf{E} \in H^s(\mathbf{curl}\,; \Omega)$ and s is a smoothness parameter, then, we obtain in fact the desired polynomial convergence if we use high-order finite elements

$$\|\mathbf{E} - \mathbf{E}_h\|_{\mathbf{curl}} \leq Ch^s.$$

52.6 DG Method

The FDTD method has become the method of choice in physics and engineering when time-dependent Maxwell's equations have to be solved. The reasons are its simplicity, robustness and sufficient accuracy for most practical problems. On the other hand, it has severe limitations. These are mainly the difficultie incorporating complicated geometry due to its link to staggered grids unless the grid is prohibitively fine, the inherent second-order accuracy in time and space. Both topics have been addressed successfully in the frame work of FDTM methods, see the section about FDTD. However, DG methods present a successful alternative with the two main features geometrical flexibility and high-order approximations both in time and space inherently anchored in the method from the beginning.

52.6.1 Historical Notes

As any other method described in this chapter, the DG method has a long history with many contributions from many sides. Reed and Hill [35] applied a first form of the DG method to the neutron transport equation in 1973.

Since then, it has undergone a fast development and finds numerous applications to e.g., the Euler equations of gas dynamics, the shallow water equations, the equations of magneto-hydrodynamics, the compressible Navier–Stokes equations, and Maxwell's equations only to mention a few; see e.g., [36] for an overview and analysis of many of these applications.

The combination of the DG method for spatial discretization and of an explicit Runge–Kutta scheme for time-integration has been introduced by Cockburn and Shu in several papers, [37] maybe being one of the latest. [38] gives an introduction to the DG method and contains a more complete list of references on the (RKDG) method; it can be downloaded online. We also mention [39–41], which is a review paper about RKDG methods, as references on the DG method. [42] gives a review of the DG method applied to nanophotonics.

52.6.2 Concept: Combine Adaptive Spatial FEM Discretization with the Time Treatment in Finite Volume Methods

The FEM treats successfully complicated geometries in a natural way. There is no basic obstacle in combining local high-order approximations on unstructured meshes composed from different elements like tetrahedrons, prisms, bricks, and pyramids. The difficulties here lie on the side of implementations and is linked to the complexity of the problem: nontrivial geometries have to be described, decomposed into geometric patches, a detailed and efficient bookkeeping has to be realized for fast traversing. A direct application of FEM to time-dependent problems are possible but has a severe drawback: due to the occurrence of the mass matrix in conjunction with the time-derivative, an inversion of the mass matrix is needed. This can be done in several ways; nevertheless, it causes a problem not present in FDTD methods. The reason for the necessity to invert a mass matrix lies in the construction of finite elements. Typically, neighboring elements couple to each other in a way that the resulting system matrices are sparse but yield a global coupling of DoF which does not break down to structures which are cheaply invertible.

Finite volume methods, on the other hand, offer explicit time-integration schemes like those in FDTD methods. The key idea is to localize most of all spatial discretization into cells which are treated completely independent from each other. The cells are connected to each other by an explicit kind of continuity statement. This connection has to satisfy natural requirements like continuity or convergence properties but is not derived from first principles. There are many choices which can be adapted to given situations. The most intriguing feature is that by construction, each cell "sees" its own time evolution which results in an explicit scheme. The key idea of DG methods now is to combine this kind of time-treatment with a high-order spatial discretization localized to individual cell which are connected only by explicitly forced interface conditions on the cell boundaries.

52.6.3 Scalar Advection Equation as Model Equation: The Finite Volume Method in 1D

Consider the scalar advection equation

$$\begin{aligned}\partial_t u + \partial_x f(u) &= 0\\ f(u) &= au, \quad a \text{ is a real constant,}\\ u(x,0) &= u_0(x), \quad x \in \mathbb{R},\end{aligned} \tag{52.73}$$

where $\partial_x f(u)$ is the 1D model for divergence in higher dimensions $\nabla \cdot \mathbf{F}(\mathbf{u})$. The solution to Equation 52.73 is $u(x,t) = u_0(x - at)$. If $a > 0$, this is a right traveling wave, if $a < 0$, this is a left traveling wave. In order to analyze the stability of the FVM described below, we consider energy conservation. We multiply the wave equation (Equation 52.73) with u and integrate over a subdomain $\Omega_i = (x_i, x_{i+1})$:

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}u^2 + \frac{d}{dx}\frac{1}{2}au^2 &= 0\\ \frac{d}{dt}\int\limits_{x_i}^{x_{i+1}} \frac{1}{2}u^2\,dx + \int\limits_{x_i}^{x_{i+1}} \frac{d}{dx}\frac{1}{2}au^2\,dx &= 0\\ \frac{d}{dt}\,\|u\|^2_{\Omega_i} &= -a\left(u^2\left(x_{i+1}\right) - u^2\left(x_i\right)\right).\end{aligned} \tag{52.74}$$

52.6.3.1 Direction of Information Flow

From Equation 52.74 we deduce the direction of information flow and stability under perturbation. Let $a > 0$, the wave is right propagating. At a fixed time, we consider the function $u(x)$ on the interval (x_i, x_{i+1}). Since (Equation 52.73) is linear, we can subtract a constant such that we obtain a new $u(x)$ with $u(x_i) = 0$. When we consider this value as a given zero boundary condition at the left side, we read from (Equation 52.74) $\frac{d}{dt}\,\|u\|^2_{\Omega_i} \leq 0$ independently from the right value. This right value might be the result of a computation, hence it might be contaminated by an error. Independent of contamination, there is no increase in the norm of u, hence a perturbation does not lead to instability. If we had chosen the right boundary to be the place for the exact boundary condition, the opposite would be the case. There would always be an increase in $\|u\|^2_{\Omega_i}$ in time.

52.6.3.2 Link to Finite Volume Methods

The whole computational domain Ω is decomposed into small control volumes $\bar{\Omega} = \cup\bar{\Omega}_i$, $\Omega_i \cap \Omega_j = 0$ for $i \neq j$. We consider one subdomain Ω_i and use the conservation form of Equation 52.73 applying Gauss's theorem:

$$\partial_t \int_{\Omega_i} u\, dV + \int_{\partial\Omega_i} f dS = 0, \tag{52.75}$$

where in higher dimensions the surface integral takes the form $\int_S \mathbf{f} \cdot \mathbf{n}\, dS$ and $\mathbf{n}$ is the outwardly directed unit normal vector. Next, let the surface $\partial\Omega_i$ of cell Ω_i be subdivided into surfaces S_k which connect cell i to neighboring cells. Let $|\Omega_i|$ be the volume of Ω_i. We introduce the cell average u_i as an unknown target quantity

$$\partial_t u_i + \frac{1}{|\Omega_i|} \sum_k \int_{S_k} f\, dS = 0 \tag{52.76}$$

$$u_i = \frac{1}{|\Omega_i|} \int_{\Omega_i} u\, dV. \tag{52.77}$$

On the 1D mesh this results in the numerical scheme:

$$\partial_t u_i + \frac{1}{|\Omega_i|} \left(f_{i+1/2} - f_{i-1/2}\right) = 0; \tag{52.78}$$

see Figure 52.19. As in the FDTD method, the solution is based on a combination of quantities defined in primal meshes (the function u_h at integer positions) and on dual meshes (the discrete flux f_h at half counting positions). The main question is how do we define the values of $f = au$ at positions $x_{i-1/2}$ and $x_{i+1/2}$?

The answer depends, as discussed above, on the direction of wave propagation. It is natural to select the points in the solution at current time that are "upwind" of the solution at the position i for increasing time to maintain the causal connection. Depending then on the direction in which the solution is translated, we have different numerical schemes.

52.6.3.3 Finite Volume Method with First-Order Upwind Scheme

If $a > 0$, there is propagation in the positive x-direction. It holds $u_{i-1/2} \approx u_{i-1}$ and $u_{i+1/2} \approx u_i$, hence $f_{i+1/2} = au_i, f_{i-1/2} = au_{i-1}$, and the scheme (Equation 52.78) would be

$$\partial_t u_i + \frac{a}{\Delta x} \left(u_i - u_{i-1}\right) = 0.$$

In contrast, if $a < 0$, there is propagation in the negative x-direction. It holds $u_{i-1/2} \approx u_i$ and $u_{i+1/2} \approx u_{i+1}$, hence $f_{i+1/2} = au_{i+1}, f_{i-1/2} = au_i$, and the scheme (Equation 52.78) would be

$$\partial_t u_i + \frac{a}{\Delta x} \left(u_{i+1} - u_i\right) = 0.$$

Let us fix position $x_{i+1/2}$. The approximation of the flux just used is called the numerical flux, f^{num},

$$f^{\text{num}}_{i+1/2}\left(u_i, u_{i+1}\right) = \begin{cases} au_i & \text{if } a > 0 \\ au_{i+1} & \text{if } a < 0. \end{cases} \tag{52.79}$$

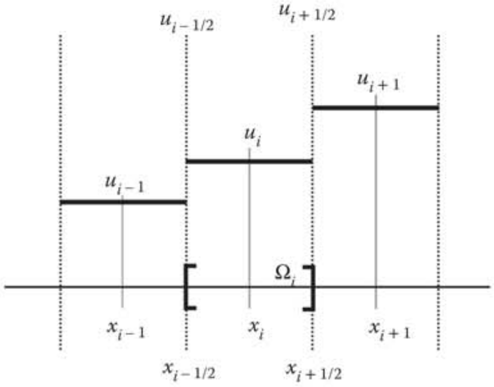


FIGURE 52.19 Vertex-centered finite volume method in 1D.

This can be compactly denoted as

$$f^{\text{num}}_{i+1/2}\left(u_i, u_{i+1}\right) = a\underbrace{\frac{1}{2}\left(u_i + u_{i+1}\right)}_{:=\{u\}_{i+1/2}} + \frac{|a|}{2}\underbrace{\left(u_i - u_{i+1}\right)}_{:=[\![u]\!]_{i+1/2}}$$

$$= a\{\!\{u\}\!\}_{i+1/2} + \frac{|a|}{2}[\![u]\!]_{i+1/2}\,.$$

We describe the typical finite volume discretization and give explicitly the 1D scheme:

1. Subdivide the domain into finite control volumes Ω_i. In one space dimension this is $\Omega_i = \left(x_{i-1/2}, x_{i+1/2}\right)$.
2. The conservation law is computed independently on each individual control volume Ω_i (52.76)

$$\partial_t u_i = -\frac{1}{|\Omega_i|}\sum_k \int_{S_k} f dS$$

with u_i the approximation to the average (Equation 52.77)

$$u_i = \frac{1}{|\Omega_i|}\int_{\Omega_i} u\, dV.$$

The resulting upwind scheme is, with flux (Equation 52.79),

$$\begin{aligned}\partial_t u_i &= -\frac{1}{|\Omega_i|}\left(f_{i+1/2} - f_{i-1/2}\right)\\ &= -\frac{1}{|\Omega_i|}\left(f^{\text{num}}_{i+1/2}\left(u_i, u_{i+1}\right) - f^{\text{num}}_{i+1/2}\left(u_{i-1}, u_i\right)\right).\end{aligned} \qquad (52.80)$$

3. Solve the semidiscrete system (Equation 52.80) with a suitable time-integrator. There are many possibilities, the most often used scheme in conjunction with time-dependent Maxwell's equations are low-storage Runge–Kutta methods. The resulting scheme after discretization in time must be a so-called monotonic scheme. Monotonic numerical schemes are ones which, given an initial distribution which is monotonic, produce a monotonic distribution after advection. A consequence of this property is that monotonic schemes neither create new extrema in the solution nor amplify existing extrema.

52.6.3.4 DG Method in 1D

The goal is to replace the low-order (piecewise constant) solution by a high-order approximation. We consider cell Ω_i. Exactly as in the FEM, we multiply (Equation 52.73) with a test function $v \in V$:

$$\int_{\Omega_i} v\partial_t u \, dx + \int_{\Omega_i} v\partial_x f(u) \, dx = 0.$$

An integration by parts gives

$$\int_{\Omega_i} v\partial_t u \, dx - \int_{\Omega_i} \partial_x v f(u) \, dx = - \int_{\partial\Omega_i} v f(u) \, ds.$$

Again as in FEM, we obtain the numerical method in *weak form* if we restrict the general space V to a finite dimensional space (of possibly high order) $V_h \subset V$:

$$\int_{\Omega_i} v_h \partial_t u_h \, dx - \int_{\Omega_i} \partial_x v_h f(u_h) \, dx = - \int_{\partial\Omega_i} v_h f^{\text{num}}(u_h) \, ds \quad \text{for all } v_h \in V_h. \tag{52.81}$$

The corresponding strong version results, if we undo the partial integration:

$$\int_{\Omega_i} v_h \partial_t u_h \, dx + \int_{\Omega_i} v_h \, \partial_x f(u_h) \, dx = \int_{\partial\Omega_i} v_h \left(f\left(u_h\right) - f^{\text{num}}(u_h) \right) \, ds \quad \text{for all } v_h \in V_h. \tag{52.82}$$

52.6.4 Maxwell's Equations in Conservation Form

We generalize the above step by step and present the framework of DG for Maxwell's equation's in 3D. The matrix-valued flux function $\mathbf{F} := \left(\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3\right)$ is composed of three column vectors of length n, $\mathbf{F}_i : \Omega \subset \mathbb{R}^3 \to \mathbb{R}^n$, with $i = 1, 2, 3$, and $\mathbf{F}_i = \left(F_{1i}, \ldots, F_{ni}\right)^T$. The divergence $\nabla \cdot \mathbf{F}$ of this matrix-valued field $\mathbf{F} : \Omega \subset \mathbb{R}^3 \to \mathbb{R}^{n\times 3}$ is defined as

$$\nabla \cdot \mathbf{F} := \begin{pmatrix} \partial_{x_1} F_{11} & +\partial_{x_2} F_{12} + & \partial_{x_3} F_{13} \\ & \vdots & \\ \partial_{x_1} F_{n1} & +\partial_{x_2} F_{n2} + & \partial_{x_3} F_{n3} \end{pmatrix}.$$

The generalization of the 1D advection equation (Equation 52.73) reads

$$\begin{aligned} \mathbf{Q}\partial_t u + \nabla \cdot \mathbf{F}(\mathbf{u}) &= 0 \\ u(x, 0) = u_0(x), &\quad x \in \mathbb{R}^3. \end{aligned} \tag{52.83}$$

Next, we bring Maxwell's equations into this conservation form. From

$$\begin{pmatrix} \varepsilon & \\ & \mu \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \begin{pmatrix} & \nabla\times \\ -\nabla\times & \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}$$

we read

$$\mathbf{Q} = \begin{pmatrix} \varepsilon & \\ & \mu \end{pmatrix} \qquad \mathbf{F}_i = \begin{pmatrix} -\mathbf{e}_i \times \mathbf{H} \\ \mathbf{e}_i \times \mathbf{E} \end{pmatrix}$$

with $\mathbf{e}_i$ being one of the three Cartesian unit vectors. Hence, there is in fact a conservation form (Equation 52.83) of Maxwell's equations.

52.6.5 Concept of DG-Algorithms for 3D Maxwell's Equations

The 3D algorithm follows the 1D concept. The only missing link is the upwind flux for the situation of Maxwell's equations. Following the same strategy as in 1D, Hesthaven and Warburg [43] derive the following expression:

$$\mathbf{n} \cdot \left(\mathbf{F} - \mathbf{F}^{\mathrm{num}}\right) = \begin{pmatrix} \bar{Z}^{-1}\mathbf{n} \times \left(-Z^{+} [\![\mathbf{H}]\!] + \mathbf{n} \times [\![\mathbf{E}]\!]\right) \\ \bar{Y}^{-1}\mathbf{n} \times \left(-Y^{+} [\![\mathbf{E}]\!] + \mathbf{n} \times [\![\mathbf{H}]\!]\right) \end{pmatrix}. \tag{52.84}$$

Here Z and Y are the wave impedance and admittance, respectively,

$$Z^{\pm} = \sqrt{\frac{\mu^{\pm}}{\varepsilon^{\pm}}}, \quad Y^{\pm} = \left(Z^{\pm}\right)^{-1}, \quad \bar{Z} = \frac{Z^{+} + Z^{-}}{2}, \quad \bar{Y} = \frac{Y^{+} + Y^{-}}{2}.$$

The superscript "−" denotes the limit from the interior to the cell boundary, and the superscript "+" the limit from the neighboring cell to the boundary.

Algorithm 52.4: Basic Discontinuous Galerkin Algorithm

```
1. Discretization in space
construct a tessellation $\bar{\Omega} = \cup_i \bar{\Omega}_i$
for all patches $\Omega_i$ do
  map $\widehat{K} \to K_i$
  compute local matrices on local volumes
  assemble them (independently) to a global matrix
  compute local matrices on interior interfaces
  connect them to the global matrix
  compute local matrices on boundary interfaces
  connect them to build up the right-hand side
end for
2. Time evolution
use a time integrator (e. g., Runge–Kutta, leap-frog) to solve the semidiscrete, block-diagonal system in time
return vector of field values and fluxes in time
```

52.6.6 Convergence

We consider Maxwell's equations with homogeneous, isotropic materials and a convex domain $\Omega \subset \mathbb{R}^3$ and arbitrary smooth initial data. This is a simplified model case which has no singularities. In this case optimal L^2-convergence for a fixed time interval is obtained [44]:

$$\left\|\mathbf{u} - \mathbf{u}_h\right\| \leq Ch^{p+1}.$$

This is the same as we would have, under the same conditions, for finite elements.

Symbol Definition

$\mathbf{r}^0$ Vector of unit length.

Ω Computational domain.

$\partial\Omega$ Boundary of Ω.

$\mathbf{E}_\perp$ $\mathbf{e}_x E_x + \mathbf{e}_y E_y$

$\nabla_\perp$ $\mathbf{e}_x \partial_x + \mathbf{e}_y \partial_y$

∇_{k_z} $\mathbf{e}_x \partial_x + \mathbf{e}_y \partial_y + i k_z \mathbf{e}_z$

$\iota_{e,f}$ incidence number edge-face

c speed of light

$\mathbf{k}$ wavevector

$\mathbf{r}$ position

$\llbracket f \rrbracket$ Töplitz matrix generated by f

$\llbracket \cdot \rrbracket$ jump across a face

$\{\!\{ \cdot \}\!\}$ average across a face

References

1. Andrei V Lavrinenko, Jesper Lagsgaard, Niels Gregersen, Frank Schmidt, and Thomas Søndergaard. *Numerical Methods in Photonics*. Boca Raton, FL: CRC Press, 2014.
2. Junji Yamauchi. *Propagating Beam Analysis of Optical Waveguides*. Electronic and Electrical Engineering Research Studies. Optoelectronics and Microwaves Series. Exeter, UK: Research Studies Press, 2003.
3. Jean-Claude Nédélec. *Acoustic and Electromagnetic Equations: Integral Representations for Harmonic Problems*, Vol. 144. Springer Science & Business Media, 2013.
4. Richard A Norton and Robert Scheichl. Planewave expansion methods for photonic crystal fibres. *Applied Numerical Mathematics*, 63:88–104, 2013.
5. Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2):185–200, 1994.
6. Saul Abarbanel and David Gottlieb. On the construction and analysis of absorbing layers in CEM. *Applied Numerical Mathematics*, 27(4):331–340, 1998.
7. Stephen D Gedney. An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices. *IEEE Transactions on Antennas and Propagation*, 44(12):1630–1639, 1996.
8. Weng Cho Chew and William H Weedon. A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates. *Microwave and Optical Technology Letters*, 7(13):599–604, 1994.
9. Fernando L Teixeira and Weng C Chew. General closed-form PML constitutive tensors to match arbitrary bianisotropic and dispersive linear media. *IEEE Microwave and Guided Wave Letters,* 8(6):223–225, 1998.

10. Kane S Yee. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas Propagation*, 14(3):302–307, 1966.
11. Thomas Weiland. A discretization model for the solution of Maxwell's equations for six component fields. *Archiv Elektronik und Uebertragungstechnik*, 31:116–120, 1977.
12. Adi Ditkowski, K Dridi, and Jan S Hesthaven. Convergent Cartesian grid methods for Maxwell's equations in complex geometries. *Journal of Computational Physics*, 170(1):39–80, 2001.
13. Jan S Hesthaven. High-order accurate methods in time-domain computational electromagnetics: A review. *Advances in Imaging and Electron Physics*, 127:59–125, 2003.
14. Allen Taflove. *Advances in Computational Electrodynamics, the Finite-Difference time Domain*. Boston: Artech House, 1998.
15. Alain Bossavit. 'Generalized finite differences' in computational electromagnetics. *Progress in Electromagnetics Research*, 32:45–64, 2001.
16. Markus Clemens and Thomas Weiland. Discrete electromagnetism with the finite integration technique. *Progress in Electromagnetics Research*, 32:65–87, 2001.
17. Allen Taflove and Susan C Hagness. *Computational Electrodynamics: The Finite Difference Time-Domain Method*, 3rd edition. Boston, MA: Artech House Publishers, 2005.
18. Anders Bondeson, Thomas Rylander, and Par Ingelstrm. *Computational Electromagnetics*, Vol. 51. New York, NY: Springer Verlag, 2005.
19. Peter Monk and Endre Süli. A convergence analysis of Yee's scheme on nonuniform grids. *SIAM Journal on Numerical Analysis*, 31(2):393–412, 1994.
20. Theodor Tamir, HC Wang, and AA Oliner. Wave propagation in sinusoidally stratified dielectric media. *IEEE Transactions on Microwave Theory and Techniques*, 12(3):323–335, 1964.
21. C Yeh, KF Casey, and ZA Kaprielian. Transverse magnetic wave propagation in sinusoidally stratified dielectric media. *IEEE Transactions on Microwave Theory and Techniques*, 13(3):297–302, 1965.
22. MG Moharam and TK Gaylord. Rigorous coupled-wave analysis of planar-grating diffraction. *Journal of the Optical Society of America*, 71(7):811–818, 1981.
23. Philippe Lalanne and G Michael Morris. Highly improved convergence of the coupled wave method for tm polarization. *JOSA A*, 13(4):779–784, 1996.
24. G Granet and B Guizal. Efficient implementation of the coupled-wave method for metallic lamellar gratings in tm polarization. *Journal of the Optical Society of America A*, 13(5):1019–1023, 1996.
25. Lifeng Li. Use of Fourier series in the analysis of discontinuous periodic structures. *Journal of the Optical Society of America A*, 13(9):1870–1876, 1996.
26. Jens Küchenmeister. Three-dimensional adaptive coordinate transformations for the Fourier modal method. *Optics Express*, 22(2):1342–1349, 2014.
27. Peter Monk. *Finite element Methods for Maxwell's Equations*. Oxford University Press, 2003.
28. Leszek Demkowicz. *Computing with hp-Adaptive Finite Elements: One and Two Dimensional Elliptic and Maxwell Problems*, Vol. 1. CRC Press, 2006.
29. Jean-Claude Nédélec. Mixed finite elements in R^3. *Numerische Mathematik*, 35(3):315–341, 1980.
30. Jean-Claude Nédélec. A new family of mixed finite elements in R^3. *Numerische Mathematik*, 50:57–81, 1986.
31. Alain Bossavit. Whitney forms: A class of finite elements for three-dimensional computations in electromagnetism. *IEE Proceedings A (Physical Science, Measurement and Instrumentation, Management and Education, Reviews)*, 135(8):493–500, 1988.
32. Jon P Webb. Hierarchal vector basis functions of arbitrary order for triangular and tetrahedral finite elements. *IEEE Transactions on Antennas and Propagation*, 47(8):1244–1253, 1999.
33. Jan Pomplun, Sven Burger, Lin Zschiedrich, and Frank Schmidt. Adaptive finite element method for simulation of optical nano structures. *Physica Status Solidi (B)*, 244(10):3419–3434, 2007.
34. Peter Monk. A simple proof of convergence for an edge element discretization of Maxwell's equations. In *Computational Electromagnetics*, pp. 127–141. Springer, 2003.

35. WH Reed and TR Hill. Triangular mesh methods for the neutron transport equation. Report LA-UR-73-479. Los Alamos, NM: Los Alamos Scientific Laboratory, 1973.
36. Bernardo Cockburn, George E Karniadakis, and Chi-Wang Shu. *The Development of Discontinuous Galerkin Methods*. Lecture notes in computational science and engineering. Berlin, Heidelberg: Springer-Verlag, 2000.
37. Bernardo Cockburn and Chi-Wang Shu. The Runge–Kutta discontinuous Galerkin method for conservation laws V: Multidimensional systems. *Journal of Computational Physics*, 141(2):199–224, 1998.
38. Bernardo Cockburn. An introduction to the discontinuous Galerkin method for convection-dominated problems. In *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations*, pp. 151–268. Berlin Heidelberg: Springer, 1998.
39. Bernardo Cockburn, George E Karniadakis, and Chi-Wang Shu. *Discontinuous Galerkin Methods: Theory, Computation and Applications*. Berlin Heidelberg: Springer, 2000.
40. Bernardo Cockburn and Chi-Wang Shu. Runge–Kutta discontinuous Galerkin methods for convection-dominated problems. *Journal of Scientific Computing*, 16(3):173–261, 2001.
41. Jan S Hesthaven and Tim Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer Science & Business Media, 2007.
42. Kurt Busch, Michael Koenig, and Jens Niegemann. Discontinuous Galerkin methods in nanophotonics. *Laser & Photonics Reviews*, 5(6):773–809, 2011.
43. Jan S Hesthaven and Timothy Warburton. Nodal high-order methods on unstructured grids: I. Time-domain solution of Maxwell's equations. *Journal of Computational Physics*, 181(1):186–221, 2002.
44. Marcus J Grote, Anna Schneebeli, and Dominik Schtzau. Interior penalty discontinuous Galerkin method for Maxwell's equations: Optimal l2-norm error estimates. *IMA journal of Numerical Analysis*, 28(3):440–468, 2008.