

Chapter 7

Applications

7.1 Introduction

The range of applications of PDE optimization is vast and quickly widening to all fields of scientific computing. An outlook of some application problems, for which the design of algorithms with optimal computational complexity is ongoing, is the purpose of this chapter. We discuss PDE optimization problems with representative models of mathematical physics, such as the Navier–Stokes equations, the Schrödinger equation, and the Maxwell equations.

Based on the Navier–Stokes model, we discuss shape design problems where the geometry of the model becomes an optimization variable, and therefore the need arises of defining a “calculus” with respect to this variable. In particular, we focus on the modern powerful approach of nonparametric shape sensitivity analysis that allows us to consider challenging shape optimization problems.

The models described by Schrödinger-type equations are receiving much attention because they are central in nanosciences. At the nanoscale and beyond, we are in the quantum world and the Schrödinger equation becomes the governing model with a large range of applications including quantum optics, quantum chemistry, and design of nanostructures. We discuss quantum control problems where PDE optimization is used to design control potentials to optimally drive quantum systems to perform given tasks.

The Maxwell equations provide one of the most representative models related to inverse problems formulated within the PDE optimization framework. In particular, we consider noninvasive imaging which is very important in medical diagnosis and material probing. In this case, inverse scattering problems are formulated where given incident electromagnetic waves penetrate a medium and based on the scattered waves one determines the presence and the properties of objects hidden in the medium.

7.2 Aerodynamic Shape Design Supported by the Shape Calculus

7.2.1 Overview on Shape Sensitivity Analysis

Shape calculus defines the mathematical framework, when the shape of an object is the unknown quantity and should be determined in accordance with a certain (differentiable) goal. For a gradient-based optimization approach the sensitivity information of an objective function is needed with respect to the shape of the boundary on which it is defined, which is often termed “shape sensitivity analysis.” While these methods have been used quite successfully to derive theoretical results for optimal shapes, such as the optimal supersonic Haak body [160] or the rugby-ball ogive shapes in a Stokes flow [276], most numerical schemes are actually based on smooth ansatz functions such as the very popular Hicks–Henne functions [189] or other curve parameterizations such as b-splines.

However, we consider the nonparametric shape sensitivity approach to numerically conduct large-scale shape optimization. In the following, let the domain under consideration be given by $\Omega \subset \mathbb{R}^3$ compact with sufficiently smooth boundary $\Gamma := \partial\Omega$. We seek to compute the derivative or sensitivity information of general objective functions such as

$$J_1(\Omega) := \int_{\Omega} f \, dA, \quad (7.1)$$

$$J_2(\Omega) := \int_{\Gamma} g \, dS \quad (7.2)$$

with respect to the domain Ω , where f and g are sufficiently smooth functions. To this end, a deformed domain

$$\Omega_t = T_t(\Omega) = \{T_t(x) : x \in \Omega\}$$

is considered. The family of bijective mappings T_t is usually chosen as the perturbation of identity

$$T_t(x) = x + tV(x)$$

or the speed method, where T_t is the solution of the flow equation

$$\frac{\partial x}{\partial t} = V(t, x), \quad x(0) = x_0 \in \Gamma.$$

Here, V is a sufficiently smooth vector field that takes the role of the differentiation direction. For the volume case (7.1), it is easy to see that

$$\begin{aligned} \int_{\Omega_t} f \, dA_t &= \int_{\Omega} f(T_t(x)) \sqrt{\det DT_t^T DT_t(x)} \, dA(x) \\ &= \int_{\Omega} f(T_t(x)) |\det DT_t(x)| \, dA(x), \end{aligned}$$

which can easily be differentiated with respect to t , taking into account that physically

meaningful vectorfields V preserve the orientation:

$$\begin{aligned} dJ_1(\Omega)[V] &= \int_{\Omega} \frac{d}{dt} \Big|_{t=0} f(T_t(x)) \det(DT_t(x)) dA(x) \\ &= \int_{\Omega} \langle \nabla f(x), V(x) \rangle + f(x) \operatorname{div} V(x) dA(x) \\ &= \int_{\Omega} \operatorname{div}(f(x)V(x)) dA(x). \end{aligned} \quad (7.3)$$

A similar computation for the surface objective function results in

$$\begin{aligned} dJ_2(\Omega)[V] &= \frac{d}{dt} \Big|_{t=0} \int_{\Gamma_t} g dS_t \\ &= \int_{\Gamma} \langle \nabla g, V \rangle + g \cdot (\operatorname{div} V - \langle DVn, n \rangle) dS \\ &= \int_{\Gamma} \langle \nabla g, V \rangle + g \operatorname{div}_{\Gamma} V dS, \end{aligned} \quad (7.4)$$

where $\operatorname{div}_{\Gamma}$ is the tangential or surface divergence operator. Thus, the sensitivity equations (7.3) and (7.4) can be interpreted as the directional derivatives in direction V .

7.2.2 The Hadamard Formula

The Hadamard theorem or Hadamard–Zolésio structure theorem [321, 110] states that under some regularity assumptions, the sensitivity equations (7.3) and (7.4) can be expressed as a scalar product of the normal component of the perturbation field V with some shape gradient on the surface of the geometric object to be optimized, i.e.,

$$dJ(\Omega)[V] = dJ(\Gamma)[\langle V, n \rangle n].$$

For the volume case (7.3) applying the divergence theorem results in

$$dJ_1(\Omega)[V] = \int_{\Gamma} \langle V, n \rangle f dS.$$

The surface case, (7.4), requires elements from tangential calculus, especially the tangential Stokes formula. However, one can show that

$$dJ_2(\Omega)[V] = \int_{\Gamma} \langle V, n \rangle \left[\frac{\partial g}{\partial n} + \kappa g \right] dS,$$

where $\kappa := \operatorname{div} n$ is the additive mean curvature. Thus, formulas are created which can be very efficiently evaluated numerically allowing a large-scale shape optimization. The need to compute mesh sensitivity Jacobians is completely circumvented. Instead, some geometric quantities such as the curvature κ must be computed, which can be conveniently done using discrete differential geometry [292].

For the application in computational fluid dynamics and aerodynamic design, the above formula becomes more complex, because the usual objective functions such as fluid

forces in addition to a dependence on the PDE fluid state also depend on the geometry and these quantities must be differentiated accordingly. For a dependence on the normal, one has

$$J(g, \Gamma) := \int_{\Gamma} g(\varphi, D\varphi, n) dS,$$

where $g : \mathbb{R}^d \times \mathbb{R}^{d \times d} \times \mathbb{R}^d \rightarrow \mathbb{R}$, $(\varphi, \zeta, \psi) \mapsto g(\varphi, \zeta, \psi)$ is a sufficiently smooth functional. The shape derivative of the above expression is then given by

$$\begin{aligned} dJ(g, \Gamma)[V] = \int_{\Gamma} \langle V, n \rangle \Big[D_{\varphi} g D\varphi n + D_{\zeta} g D^2 \varphi n \\ + \kappa (g - D_{\psi} g n) + \operatorname{div}_{\Gamma} (D_{\psi} g)^T \Big] dS. \end{aligned}$$

7.2.3 Shape Optimization and the Incompressible Navier–Stokes Equations

The incompressible Navier–Stokes equations are considered in a general setting:

$$\begin{aligned} \min_{(u, p, \Omega)} J(u, p, \Omega) := \int_{\Omega} f(u, Du, p) dA \\ + \int_{\Gamma_0} g(u, D_n u, p, n) dS \end{aligned}$$

subject to

$$\begin{aligned} -\mu \Delta u + \rho u \nabla u + \nabla p &= \rho G \quad \text{in } \Omega, \\ \operatorname{div} u &= 0, \\ u &= u_+ \quad \text{on } \Gamma_+, \\ u &= 0 \quad \text{on } \Gamma_0, \\ pn - \mu \frac{\partial u}{\partial n} &= 0 \quad \text{on } \Gamma_-, \end{aligned}$$

where Γ_+ is the inflow boundary, Γ_- is the outflow boundary, and Γ_0 is the fluid obstacle of which the shape is to be optimized. Here, $f : \mathbb{R}^3 \times \mathbb{R}^{3 \times 3} \times \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}$ are assumed to be continuously differentiable in each argument. In order to keep the notation readable, components of the Jacobian are denoted as follows

$$\begin{aligned} Du &=: [a_{ij}]_{ij} \in \mathbb{R}^{3 \times 3}, \\ D_n u &= Du \cdot n = \frac{\partial u}{\partial n} =: [b_i]_i \in \mathbb{R}^3. \end{aligned}$$

The surface part g is chosen such that there exists a functional $\lambda : \Omega \rightarrow \mathbb{R}^d$ satisfying the following conditions on Γ_0

$$\begin{aligned} \lambda_i &= \frac{1}{\mu} \frac{\partial g}{\partial b_i} \quad i = 1, 2, 3, \\ \langle \lambda, n \rangle &= -\frac{\partial g}{\partial p}. \end{aligned}$$

A formal differentiation of the above leads to the gradient expression [304]

$$\begin{aligned}
 dJ(u, p, \Omega)[V] &= \int_{\Gamma_0} \langle V, n \rangle f(u, Du, p) dS \\
 &+ \int_{\Gamma_0} \langle V, n \rangle [D_{(u,b,p)} g(u, D_n u, p, n) \cdot n + \kappa g(u, D_n u, p, n)] dS \\
 &+ \int_{\Gamma_0} \langle V, n \rangle \left[- \sum_{i=1}^d \left(\frac{\partial g}{\partial u_i} + \mu \frac{\partial \lambda_i}{\partial n} + \sum_{j=1}^d \frac{\partial f}{\partial a_{ij}} n_j \right) \frac{\partial u_i}{\partial n} \right] dS \\
 &+ \int_{\Gamma_0} \langle V, n \rangle [(\operatorname{div}_\Gamma \nabla_n g) - \kappa \langle \nabla_n g, n \rangle] dS,
 \end{aligned}$$

where $\nabla_n g$ denotes the vector consisting of components $\frac{\partial g}{\partial n_i}$ and λ and λ_p solve the adjoint incompressible Navier–Stokes equations

$$\begin{aligned}
 -\mu \Delta \lambda_i - \rho \sum_{j=1}^d \left(\frac{\partial \lambda_j}{\partial x_i} u_j + \frac{\partial \lambda_i}{\partial x_j} u_j \right) - \frac{\partial \lambda_p}{\partial x_i} \\
 = \frac{\partial f}{\partial u_i} - \sum_{j=1}^d \frac{\partial}{\partial x_j} \frac{\partial f}{\partial a_{ij}} \quad \text{in } \Omega, \\
 \operatorname{div} \lambda = \frac{\partial f}{\partial p}
 \end{aligned}$$

with boundary conditions

$$\begin{aligned}
 \lambda &= 0 && \text{on } \Gamma_+, \\
 \lambda_i &= \frac{1}{\mu} \frac{\partial g}{\partial b_i} && \text{on } \Gamma_0, \\
 \langle \lambda, n \rangle &= -\frac{\partial g}{\partial p} && \text{on } \Gamma_0, \\
 \mu \frac{\partial \lambda_i}{\partial n} + \rho \left(\sum_{j=1}^d \lambda_j u_j n_i + \lambda_i u_j n_j \right) + \lambda_p n_i &= 0 && \text{on } \Gamma_-.
 \end{aligned}$$

7.2.4 Shape Hessian Approximation and Operator Symbols

Very efficient optimization methods are usually variants of the SQP or Newton's method, which results in the need of higher-order derivative information. However, shape Hessians are very complex objects and usually need to be approximated in some way. The above Navier–Stokes problem can be made into an excellent problem for studying shape Hessians. By choosing the constant density $\rho = 0$, the nonlinearity vanishes. Coupled with the minimization of the kinetic energy dissipation, one arrives at the following problem, where

the fluid is modeled by the Stokes equations

$$\min_{(u,p,\Omega)} J(u, p, \Omega) := \int_{\Omega} \mu \sum_{i,j=1}^3 \left(\frac{\partial u_i}{\partial x_j} \right)^2 dA$$

subject to

$$\begin{aligned} -\mu \Delta u + \nabla p &= 0 & \text{in } \Omega, \\ \operatorname{div} u &= 0, \\ u &= u_+ & \text{on } \Gamma_+, \\ u &= 0 & \text{on } \Gamma_0, \\ pn - \mu \frac{\partial u}{\partial n} &= 0 & \text{on } \Gamma_-, \\ \operatorname{Vol} &= V_0. \end{aligned}$$

The shape derivative of this problem is given by

$$dJ(u, p, \Omega)[V] = -\mu \int_{\Gamma_0} \langle V, n \rangle \sum_{k=1}^3 \left(\frac{\partial u_k}{\partial n} \right)^2 dS, \quad (7.5)$$

which means the problem is self-adjoint and the Hessian can be studied without the need to consider variations in the adjoint multipliers λ .

Shape gradients can be split in a geometric part, which stems from variations of the objective with respect to the geometry, and a dynamic part, which stems from a variation of the PDE state (u, p) due to a change of the domain on which the PDE is defined. The idea presented here from [305, 302] is to study the symbol of the control to state mapping of the dynamic part only. Considering a sinusoidal perturbation $\tilde{q}(x) = \hat{q}e^{i\omega x}$ of some control q , the pseudodifferential operator nature of the Hessian H can be seen by comparing the input \tilde{q} with the output $H\tilde{q}$. For example, if

$$H\tilde{q} = i\omega\hat{q}e^{i\omega x} = i\omega\tilde{q},$$

then $\Sigma(\omega) := i\omega$ is the symbol of the Hessian, and this corresponds to a classical differential operator of order +1. If, for example, one has

$$H\tilde{q} = -\omega^2\hat{q}e^{i\omega x} = -\omega^2\tilde{q},$$

then $\Sigma(\omega) := -\omega^2$ is the symbol of the Hessian, and this corresponds to a classical differential operator of order +2. However, if

$$H\tilde{q} = |\omega|\hat{q}e^{i\omega x} = |\omega|\tilde{q},$$

then H is a pseudodifferential operator of order +1.

Assuming flow over a flat plate, the domain Ω is considered to be given by

$$\Omega = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 \geq 0\}.$$

Thus, at the boundary $\Gamma = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\}$, the outer normal is given by

$$n = (0, 0, 1)^T,$$

and a complex-valued oscillation of the two-dimensional flat plate is described by

$$x_3 = \alpha(x_1, x_2) := e^{i(\omega_1 x_1 + \omega_2 x_2)},$$

where i is the imaginary unit. Using this setting, the disturbed gradient from (7.5) is given by

$$\begin{aligned} \tilde{G} &:= -2\mu \sum_{k=1}^3 \frac{\partial u_k}{\partial n} \frac{\partial u'_k[\alpha]}{\partial n} \\ &= -2\mu \sum_{k=1}^3 \frac{\partial u_k}{\partial x_3} \frac{\partial u'_k[\alpha]}{\partial x_3} \end{aligned}$$

due to the flat initial domain under consideration. To identify the symbol of the Hessian, the mapping

$$S\alpha := -2\mu \sum_{k=1}^3 \frac{\partial u_k}{\partial x_3} \frac{\partial u'_k[\alpha]}{\partial x_3}$$

must now be characterized by making α appear explicitly in the right-hand side.

Furthermore, the perturbed states $u'[\alpha]$ and $p'[\alpha]$ are also considered to be oscillatory:

$$\begin{aligned} u'_k[\alpha] &= \hat{u}_k e^{i(\omega_1 x_1 + \omega_2 x_2)} e^{\omega_3 x_3}, \\ p'[\alpha] &= \hat{p} e^{i(\omega_1 x_1 + \omega_2 x_2)} e^{\omega_3 x_3}. \end{aligned}$$

The no-slip boundary condition on $x_3 = 0$ leads to

$$u'_k[\alpha] = \hat{u}_k \alpha e^0,$$

which results in

$$\hat{u}_k = -\frac{\partial u_k}{\partial x_3} \neq 0. \quad (7.6)$$

However, the linearized Stokes PDE must also be solved inside the domain. Applying the Laplace, the gradient, and the divergence operators on the disturbances $u'[\alpha]$ and $p'[\alpha]$ is equivalent to

$$A \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{p} \end{pmatrix} \alpha(x_1, x_2) e^{\omega_3 x_3} = 0,$$

where the matrix A can be thought of as the Stokes equation in the Fourier space. This is noncontradictory to the consequences of the boundary condition (7.6) only if the linearized

system matrix A does not have full rank, which means the determinant of A must vanish:

$$\begin{aligned}\det(A) &= \left[\mu \left(\omega_1^2 + \omega_2^2 - \omega_3^2 \right) \right]^2 \left(\omega_1^2 + \omega_2^2 - \omega_3^2 \right) \stackrel{!}{=} 0 \\ \Rightarrow \omega_3 &= \pm \sqrt{\omega_1^2 + \omega_2^2}.\end{aligned}$$

Hence, it is possible to remove ω_3 from the equations and the local shape derivative of the velocity is given by

$$\begin{aligned}u'_k[\alpha] &= -\frac{\partial u_k}{\partial x_3} \alpha e^{\pm \sqrt{\omega_1^2 + \omega_2^2} x_3} \\ \Rightarrow \frac{\partial u'_k[\alpha]}{\partial x_3} \Big|_{x_3=0} &= \left[-\frac{\partial^2 u_k}{\partial x_3^2} \mp \frac{\partial u_k}{\partial x_3} \sqrt{\omega_1^2 + \omega_2^2} \right] \alpha,\end{aligned}$$

and the operator S is given by

$$S = -2\mu \sum_{k=1}^3 \frac{\partial u_k}{\partial x_3} \left[-\frac{\partial^2 u_k}{\partial x_3^2} \pm \frac{\partial u_k}{\partial x_3} \sqrt{\omega_1^2 + \omega_2^2} \right]. \quad (7.7)$$

In two dimensions, one can assume $\omega_2 \equiv 0$ and the symbol becomes

$$\pm |\omega_1|,$$

which is the symbol of a pseudodifferential operator of order $+1$ closely related to the Dirichlet-to-Neumann map. For a true pseudodifferential operator with symbol $|\omega|$, a discretization using finite differences or surface finite elements is not straightforward. A symbol which can much more easily be approximated is ω^2 , and this corresponds to the tangential Laplace operator, also known as Laplace–Beltrami operator Δ_Γ . Thus, the discrete Hessian is approximated by

$$H_h \approx k \Delta_\Gamma^h + I, \quad (7.8)$$

where $I \in \mathbb{R}^{m \times m}$ is the identity matrix for m surface mesh nodes and k is an appropriately chosen smoothing parameter.

7.2.5 Aerodynamic Design Using Shape Calculus

We conclude this section with the application of the above to a large-scale aerodynamic design problem [306, 307]. The single set-point shape optimization of the very efficient large aircraft (VELA) blended wing-body aircraft is discussed. Rather than using the full Navier–Stokes equations, we employ the compressible Euler equations as a simplified model for the inviscid fluid and the shape is to be optimized such that the pressure or wave drag in a transonic cruise of Mach 0.85 is minimized while the lift is kept constant. Thus, the

optimization problem reads as

$$\begin{aligned}
 & \min_{(\rho, u, E, \Omega)} C_D(\rho, u, E, \Omega) \\
 & \text{subject to} \\
 & \int_{\Omega} \operatorname{div}(\rho u) \, dA = 0, \\
 & \int_{\Omega} \sum_{j=1}^3 \left[\frac{\partial(\rho u_i u_j)}{\partial x_j} \right] + \frac{\partial p}{\partial x_i} \, dA = 0, \\
 & \int_{\Omega} \operatorname{div}(\rho H u) \, dA = 0, \\
 & \langle u, n \rangle = 0 \quad \text{on } \Gamma_0, \\
 & C_L = C_{L_0}, \\
 & \operatorname{Vol} = V_0.
 \end{aligned}$$

Here, the first three constraints are the Euler equations, and the pressure p is linked to the conserved variables

$$U := (\rho, \rho u, \rho E)^T$$

by the perfect gas law

$$p = (\gamma - 1)\rho \left(E - \frac{1}{2} \|u\|^2 \right)$$

and H is the fluid enthalpy. The condition $\langle u, n \rangle = 0$ is the Euler slip boundary condition on the aircraft surface Γ . The farfield boundary conditions are usually more complex, involving characteristics, and are thought of as treated by the flow solver discretely.

Since the gradient-based optimization method needs sensitivity information for both the objective function and the lift constraint, the generalized inviscid fluid forces

$$J(U, \Omega) = \int_{\Gamma_0} \langle p \cdot a, n \rangle \, dS \quad (7.9)$$

are considered, where a is a constant vector describing the rotation of the coordinate system, such that drag, lift, and the appropriate angle of attack are considered simultaneously. The shape derivative of (7.9) is given by [303]

$$\begin{aligned}
 & dJ(U, \Omega)[V] \\
 &= \int_{\Gamma_0} \langle V, n \rangle \left[\frac{\partial(p \cdot a)}{\partial n} n - \lambda U_H \left\langle \frac{\partial u}{\partial n}, n \right\rangle \right. \\
 & \quad \left. + \operatorname{div}_{\Gamma} (p \cdot a - \lambda U_H u) \right] \, dS \quad (7.10)
 \end{aligned}$$

$$\begin{aligned}
 &= \int_{\Gamma_0} \langle V, n \rangle \left[\frac{\partial(p \cdot a)}{\partial n} n - \lambda U_H \left\langle \frac{\partial u}{\partial n}, n \right\rangle + \kappa \langle p \cdot a, n \rangle \right] \\
 & \quad + \langle p \cdot a - \lambda U_H u, dn[V] \rangle \, dS. \quad (7.11)
 \end{aligned}$$

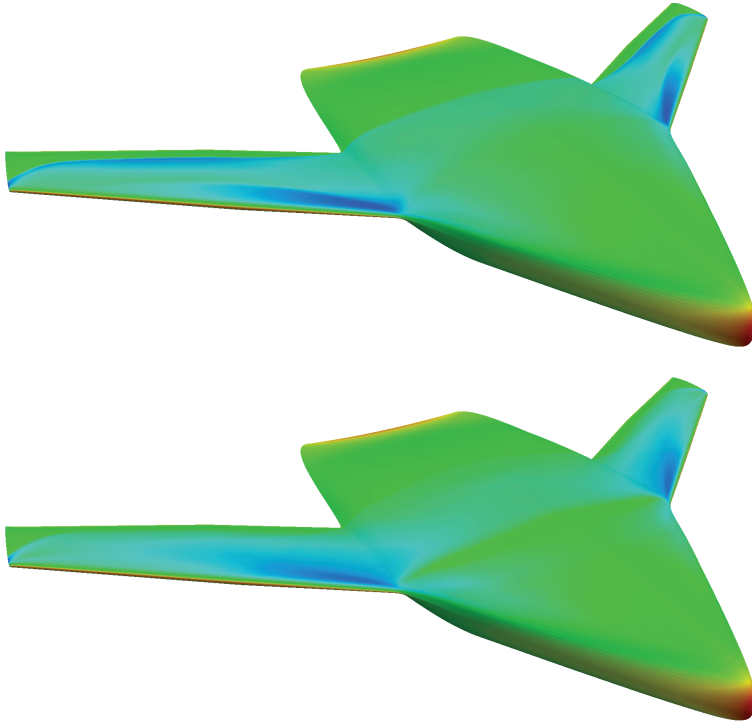


Figure 7.1. Initial and optimized VELA aircraft. Color denotes pressure. Reprinted with permission from C. Schillings, S. Schmidt, and V. Schulz, *Efficient shape optimization for certain and uncertain aerodynamic design*, *Comput. & Fluids*, 46(1) (2011), 78–87.

Additionally, U_H is the vector of conserved variables with the last component replaced by ρH . The adjoint variables λ are given as the solution of [146, 136]

$$-\frac{\partial}{\partial x_1}(A_1^T \lambda) - \frac{\partial}{\partial x_2}(A_2^T \lambda) - \frac{\partial}{\partial x_3}(A_3^T \lambda) = 0 \quad \text{in } \Omega,$$

$$\langle (\lambda_2, \lambda_3, \lambda_4)^T, n \rangle = -\langle a, n \rangle \quad \text{on } \Gamma_0,$$

where the A_i are the Euler flux Jacobians.

The drag minimization problem with constant lift is solved by the application of the one-shot method presented in Section 4.6, where the reduced Hessian approximation B is chosen according to (7.8). The VELA mesh consists of 115,673 surface nodes and 5,859,435 tetrahedra in the volume. Thus, there are approximately $29.3 \cdot 10^6$ unknowns for the fluid, which are computed using the DLR flow solver TAU [131, 140, 317]. The planform of the aircraft is kept constant, such that 113,956 surface nodes of the CFD mesh are the design unknowns for the shape. At cruise condition of Mach 0.85 and 1.8° angle of attack, the initial drag value is $C_D = 4.77 \cdot 10^{-3}$ and the initial lift coefficient is $C_L = 1.787 \cdot 10^{-1}$. The initial and optimized shapes are shown in Figure 7.1. The upper-side transonic shock wave is completely removed, such that the optimized aircraft features a drag value

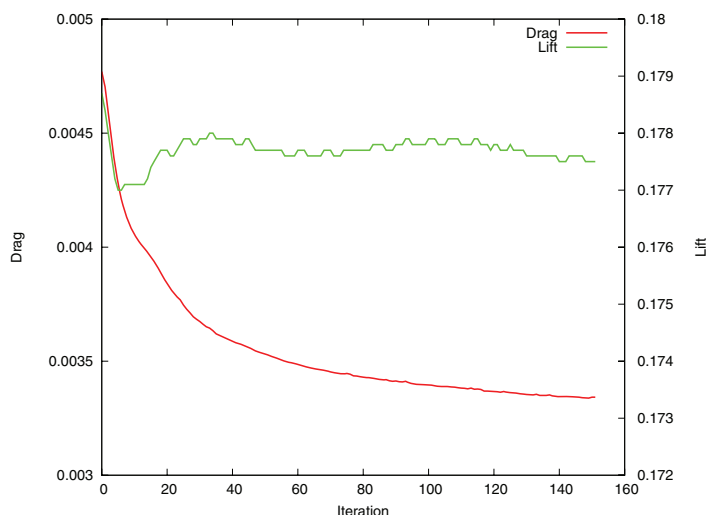


Figure 7.2. Optimization history of the single set-point VELA aircraft. Note that the lift values are stored with four digits only. Reprinted with permission from C. Schillings, S. Schmidt, and V. Schulz, *Efficient shape optimization for certain and uncertain aerodynamic design*, *Comput. & Fluids*, 46(1) (2011), 78–87.

of $C_D = 3.342 \cdot 10^{-3}$, an improvement by 30.06%, and the lift with $C_L = 1.775 \cdot 10^{-1}$ is kept within 0.67% of the original value. The optimization history is shown in Figure 7.2.

7.3 Quantum Control Problems

7.3.1 Introduction to Quantum Control Problems

Control of quantum phenomena is becoming central in a variety of systems with present and perspective applications ranging from quantum optics and quantum chemistry to semiconductor nanostructures, including the emerging fields of quantum computation and quantum communication; see, e.g., [16, 95, 100, 108, 198, 283, 282, 360] and the report of the AMO 2010 study undertaken by the U.S. National Research Council [5]. In most cases, quantum control is aiming at quantum devices where there is the need to manipulate quantum states with the highest possible precision. The successful application of optimal control theory in this field together with the enormous effort towards nanosciences explains the large growing interest of the mathematical and physical communities towards quantum control problems.

In the realm of quantum control problems, we can roughly identify classes of problems ranging from the realization of laser pulses to break a bond in a molecule or to drive a certain chemical reaction [360] to the control of photochemical processes [329] and the control of quantum dots [198]. More generally, control may be required to drive state transitions, maximize observable expectation, and obtain best performance of quantum operators.

In most of the envisioned applications, it is important to define fast control mechanisms that cannot be constructed based on perturbation theory strategies or on a priori parameterized control fields. This fact has motivated the increasing interest in the extension of the optimal control theory framework to the quantum world, resulting in many recent successful results [54, 63, 66, 70, 241, 243, 242, 356]. A pioneering work in this field was done by Peirce, Dahleh, and Rabitz [273], who investigated the optimal dipole control of a diatomic molecule represented by a Morse potential. The focus in this and in other early papers (see the references in [95, 100]) was to validate the ability of the optimal control framework to provide suitable quantum control mechanisms.

However, computational difficulties arising from solving optimality systems with Schrödinger-type equations, the bilinear control structure of the quantum control mechanism, and the resulting nonconvexity of the optimization problem led to research focusing more on finite-dimensional Schrödinger equations; see, e.g., [63, 70, 243]. In this case, the computational schemes of choice have been the monotonic iterative scheme [243, 283] and accelerated versions of the gradient scheme [63, 66]. These schemes perform well for finite-level quantum systems and provide acceptable results [242, 354] when applied to infinite-dimensional systems. More recently, second-order Newton schemes have been proposed [355, 356] by carefully taking into account the variational formulation of quantum optimization problems in complex Hilbert spaces and constructing accurate and stable discretization schemes with time-varying potentials. The development of the MGOPT method to solve infinite-dimensional quantum control problems is discussed in [54].

7.3.2 Finite-Level Quantum Systems

Quantum control problems that are genuinely finite dimensional are problems of controlling spin systems, and in molecular dynamics there are infinite-dimensional systems in interaction with external fields that can be well described by finite-dimensional models.

We show that a finite-level quantum system is modeled by a Schrödinger equation for an n -component wave function $\psi : [0, T] \rightarrow \mathbb{C}^n$ as follows

$$i\dot{\psi}(t) = H(u(t))\psi(t), \quad \psi(0) = \psi_0, \quad (7.12)$$

for $t \in (0, T]$, and $T > 0$ is a given terminal time. The Hamiltonian matrix $H \in \mathbb{C}^{n \times n}$ depends on the external control field $u : [0, T] \rightarrow \mathbb{C}$ and $\psi_0 \in \mathbb{C}^n$ is a given initial condition with $|\psi_0|^2 = 1$. The Hamiltonian $H = H_0 + H_c(u)$ has two constitutive components: the constant free Hamiltonian $H_0 \in \mathbb{C}^{n \times n}$ describing the unperturbed (uncontrolled) system and the control Hamiltonian $H_c \in \mathbb{C}^{n \times n}$ modeling the coupling of the quantum state to an external control field, typically a laser (electromagnetic) field. In particular, a dipole control mechanism can be written as $H_c(u) = H_d u$, where H_d is a Hermitian dipole matrix. Less investigated is the case of dipole and polarization control terms, $H_c(u) = H_d u + H_p u^2$; see, e.g., [106].

The choice of T is a modeling issue motivated by physical considerations. Based on the quantum indeterminacy principle $\Delta\lambda \Delta t \geq \hbar$ (where \hbar is the Planck constant that we set equal to one) we can state the following: For small T a highly energetic ($\Delta\lambda \gg 1$) optimal control results, thus involving many energy levels. On the other hand, for large T additional decoherence channels become important, which should be avoided in order to have an efficient control process; see, e.g., [198].

Strictly speaking, the wavefunction description given in (7.12) is appropriate for an isolated quantum system and in that case the governing Hamiltonian H is Hermitian. For a more realistic nonisolated system with environment couplings and subject to control, a more general density-matrix description would be required. Alternatively, a non-Hermitian Hamiltonian as in [66] can be constructed to account for environment losses.

Consider the Hermitian case that is very important in applications and that is also the most studied. Notice that being H Hermitian the evolution of ψ is unitary. This means that (7.12) defines an evolution in the unit sphere of \mathbb{C}^n . Further, we assume that the spectrum of H_0 is nondegenerate, with real eigenvalues $\lambda_1, \dots, \lambda_n$, called energy levels, and normalized eigenvectors ϕ_1, \dots, ϕ_n , called eigenstates, representing an orthonormal basis.

Once the eigenvalues and eigenvectors are known, the time evolution of a free state is computed as follows. If $\psi_0 = \sum_{j=1}^n c_0^j \phi_j$, then

$$\psi(t) = \sum_{j=1}^n c^j(t) \phi_j \quad \text{with} \quad c^j(t) = c_0^j \exp(-\lambda_j t),$$

where $|\psi(t)|^2 = \sum_{j=1}^n |c^j(t)|^2 = 1$. In this framework, $|c^j(t)|^2$ is the probability that making a measurement of energy of the system at time t , we obtain λ_j as its value.

Now, denote with $\langle \cdot, \cdot \rangle$ the complex scalar product in \mathbb{C}^n and assume the system at time $t = 0$ is in an ϕ_i -eigenstate in the sense that $|\langle \psi(0), \phi_i \rangle|^2 = 1$. A typical state transition problem is to design control functions u_1, \dots, u_m such that at time T the system is in another prescribed ϕ_j -eigenstate such that $|\langle \psi(T), \phi_j \rangle|^2 = 1$. Here, we have $H_c(u) = \sum_{k=1}^m H_k u_k$. Notice that because of the complex absolute function $|\cdot|$, the value of $|\langle \psi, \phi_j \rangle|^2 = 1$ is invariant under a phase shift $\psi \rightarrow \psi e^{i\theta}$.

For finite-level quantum systems with n energy levels, the state space is the unit sphere $S^{2n-1} \subset \mathbb{C}^n$. These problems are right-invariant control problems on the Lie group of special unitary matrices $SU(n)$, and the quantum controllability problem of proving that for every couple of points in $SU(n)$ one can find controls steering the system from one point to the other is well understood. Indeed the system is controllable if and only if [108, 284] we have

$$\text{Lie}\{iH_0, iH_1, \dots, iH_m\} = su(n),$$

where $su(n)$ denotes the subalgebra of zero trace matrices in $u(n)$, the Lie algebra of skew-Hermitian $n \times n$ matrices considered as a Lie algebra over the real field.

However, verifying the above condition is not always straightforward, and alternative necessary and sufficient conditions for controllability have been formulated; see, e.g., [283, 301]. On the other hand, much less is known on the controllability of infinite-dimensional (infinite-level) quantum systems; see [20, 27, 103, 342] and the references given therein.

In the following, we focus on the problem of determining an optimal control field $u \in L^2(0, T; \mathbb{C})$, such that (7.12) is fulfilled and a number of optimality criteria are met. The most frequent objectives encountered in quantum control problems require that the control sequence drive the system at a time T close to a desired target configuration $\psi_d \in \mathbb{C}^n$. In addition, limited laser resources and smoothness of the shape of the laser field are accounted for through a minimization of appropriate norms of the control function.

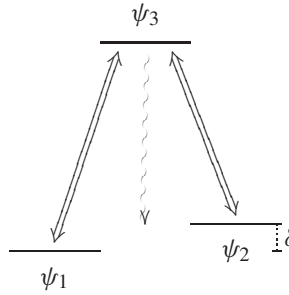


Figure 7.3. Prototypical Λ -type three-level scheme: ψ_1 and ψ_2 are long-lived states, whereas ψ_3 is a short-lived state which is optically coupled to both ψ_1 and ψ_2 ; wiggled line indicates relaxation and dephasing of state ψ_3 . Reprinted with permission from A. Borzi, J. Salomon, and S. Volkwein, *Formulation and numerical solution of finite-level quantum optimal control problems*, *J. Comput. Appl. Math.*, 216(1) (2008), 170–197.

Further, one can require suppressing population of intermediate states which suffer strong environment losses. All these requirements are realized in a cost functional as

$$\begin{aligned}
 J(\psi, u) &= \frac{1}{2} |\psi(T) - \psi_d|_{\mathbb{C}^n}^2 + \frac{\gamma}{2} \|u\|_{L^2(0,T;\mathbb{C})}^2 + \frac{\mu}{2} \|\dot{u}\|_{L^2(0,T;\mathbb{C})}^2 \\
 &\quad + \frac{1}{2} \sum_{j \in I} \alpha_j \|\psi_j\|_{L^2(0,T;\mathbb{C})}^2,
 \end{aligned} \tag{7.13}$$

where the constants $\gamma, \mu \geq 0$, $\gamma + \mu > 0$, are regularization parameters, which allow us to vary the relative importance of the objectives represented by the various terms. The goal of the first term of the cost functional is to track the state ψ close to a given terminal state at $t = T$. The second and third terms are for the regularization of the problem so that existence of at least one optimal control is ensured. In the last term of (7.13), which penalizes the occupation of certain states ψ_j , the set $I \subset \{1, \dots, n\}$ denotes a subset of possible state indices and $\alpha_j \geq 0$, $j \in I$, are weighting factors.

In [63, 66] a quantum control problem with (7.12)–(7.13) is considered. The optimality system is discretized by implicit second-order CN schemes and solved using cascadic NCG schemes, as discussed in Chapter 4, and monotonic schemes. Numerical results show that this optimization problem has very flat minima and the resulting optimal control is sensitive to the tolerance on the norm of the reduced gradient. In the following, we give details of this application.

Consider a three-level quantum system whose configuration is represented by $\psi = (\psi_1, \psi_2, \psi_3) \in H^1(0, T; \mathbb{C}^3)$ which consists of two long-lived states ψ_1 and ψ_2 , which are energetically separated by some amount δ , and a state ψ_3 , which has a finite lifetime because of environment coupling (wiggled line); see Figure 7.3. Such Λ -type configurations have a long-standing history in quantum optics, and more recently, similar configurations have received increasing attention also in semiconductor quantum dots; see, e.g., [198] and the references given therein.

Table 7.1. Optimization results depending on optimization parameters, $tol_{abs} = 10^{-7}$.

γ	μ	α	$ \psi(T) - \psi_d _{\mathbb{C}^3}$	J
10^{-7}	10^{-7}	0.05	$8.6 \cdot 10^{-4}$	$2.37 \cdot 10^{-3}$
10^{-7}	10^{-9}	0.05	$3.7 \cdot 10^{-4}$	$5.46 \cdot 10^{-4}$
10^{-7}	0	0.05	$6.9 \cdot 10^{-5}$	$1.41 \cdot 10^{-4}$
10^{-7}	0	0	$1.2 \cdot 10^{-3}$	$2.33 \cdot 10^{-6}$
10^{-4}	10^{-4}	0.05	$3.3 \cdot 10^{-2}$	$6.52 \cdot 10^{-2}$
10^{-4}	10^{-6}	0.05	$4.4 \cdot 10^{-3}$	$9.03 \cdot 10^{-3}$
10^{-4}	0	0.05	$2.7 \cdot 10^{-3}$	$5.68 \cdot 10^{-3}$
10^{-4}	0	0	$8.3 \cdot 10^{-3}$	$3.34 \cdot 10^{-4}$

Time evolution of this finite-level quantum system is governed by the following non-Hermitian Hamiltonian

$$H_0 = \frac{1}{2} \begin{pmatrix} -\delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & -i\gamma_0 \end{pmatrix}, \tag{7.14}$$

where the term $-i\gamma_0$ accounts for environmental losses (e.g., spontaneous photon emissions). The external control field is assumed complex, $u(t) \in \mathbb{C}$, which corresponds to two real control fields. The coupling of this field to the system is modeled as follows

$$H_c(u) = -\frac{1}{2} \begin{pmatrix} 0 & 0 & \mu_1 u \\ 0 & 0 & \mu_2 u \\ \mu_1 u^* & \mu_2 u^* & 0 \end{pmatrix}, \tag{7.15}$$

where μ_1 and μ_2 describe the coupling strengths of states ψ_1 and ψ_2 to the interconnecting state ψ_3 (e.g., optical dipole matrix elements). Typical initial and final states are given by

$$\psi_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \psi_d = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

respectively.

Now, we discuss the effect of different choices of values of the optimization parameters using the results reported in Table 7.1 and Figure 7.4. These results are obtained with a cascadic NCG which appears to be the most robust and efficient scheme for this class of problems.

In Table 7.1, we see that smaller values of $|\psi(T) - \psi_d|_{\mathbb{C}^3}$ are attained for smaller γ . We remark that $\gamma = 10^{-7}$ is quite small, and that makes the problem quite stiff and ill-conditioned. The NCG algorithm appears to be robust with respect to changes of γ . We also can see the effect of the regularization parameter μ . As μ increases, $|\psi(T) - \psi_d|_{\mathbb{C}^3}$ increases, demonstrating that the additional smoothness of the control function (slightly) reduces the capability of tracking. Apparently, larger μ makes the problem behaving better, resulting in less computational effort. Concerning the parameter $\alpha_3 = \alpha$ we obtain better

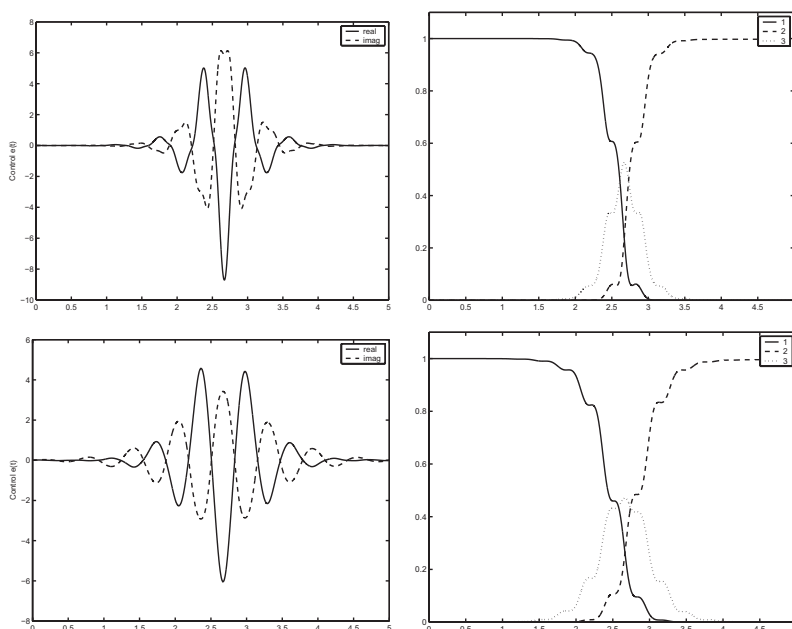


Figure 7.4. Optimal solutions for $\gamma = 10^{-4}$ and $\mu = 0$ (top) and $\mu = 10^{-6}$ (bottom): $\alpha_3 = 0.01$. Left: the control field; right: the wave function. Reprinted with permission from A. Borzi, J. Salomon, and S. Volkwein, *Formulation and numerical solution of finite-level quantum optimal control problems*, *J. Comput. Appl. Math.*, 216(1) (2008), 170–197.

tracking for nonzero α . This is expected since we have $|\psi_d|_{\mathbb{C}^3} = 1$, whereas $|\psi(T)|_{\mathbb{C}^3} < 1$ whenever $\gamma_0 > 0$, because of dissipation. By taking $\alpha > 0$ dissipation is reduced and therefore better $|\psi(T) - \psi_d|_{\mathbb{C}^3}$ is possibly achieved.

7.3.3 Infinite-Dimensional Quantum Systems

Optimization with infinite-dimensional quantum systems is one of the most challenging research fields in computational sciences. In fact, the complexity of multiparticle quantum systems and their foreseen applications outmatch our present computational capabilities [294] and our understanding of physical processes at nanoscales [5].

A fundamental model to start with the investigation of optimization problems of infinite dimensional quantum systems is the control problem of one-particle state transition in a confining potential. The term infinite dimensional refers to the fact that the constraint is a PDE and the free Hamiltonian is unbounded. We have $H_0 = -\Delta + V_0$, where the minus Laplacian represents the kinetic energy operator of the particle and V_0 the confining potential.

In quantum mechanics, the quantum state of a particle is described by a wavefunction $\psi : \Omega \times (0, T) \rightarrow \mathbb{C}$ that is governed by the following time-dependent Schrödinger equation

(TDSE) [101]

$$i \partial_t \psi(x, t) = \{-\Delta + V(x, t)\} \psi(x, t), \quad (7.16)$$

where we choose the scaling of the Planck constant $\hbar = 1$ and the mass $m = 1/2$. We consider the potential $V(x, t)$ consisting of a stationary part $V_0(x)$ and a time varying control part.

Notice that quantum systems are defined in unbounded spaces while the presence of a confinement potential results in wavefunctions ψ whose support is localized in a bounded region. Therefore, with Ω we represent a spatial domain that is large enough to represent ψ during evolution. In this case, periodic or homogeneous boundary conditions may be appropriate; see [8] for more sophisticated boundary conditions to solve numerically the TDSE on unbounded domains.

The next representative infinite-dimensional quantum system discussed in this section is the Bose–Einstein condensate (BEC) model that describes the state of matter formed by a cloud of bosons (e.g., helium-4, rubidium) cooled to temperatures very near to absolute zero. Under such conditions, the atoms with magnetic spin collapse into the lowest quantum state sharing the same wavefunction. This wavefunction describes the mean-field dynamics of a coherent BEC. It is modeled by the following Gross–Pitaevskii equation [71, 109]

$$i \partial_t \psi(x, t) = \left\{ -\Delta + V(x, u(t)) + g |\psi(x, t)|^2 \right\} \psi(x, t), \quad (7.17)$$

where $x \in \Omega$ and $t \in (0, T)$, with g a coupling constant related to the scattering length of the atoms, density, and transversal confinement. In (7.17), the confining potential $V(x, u(t))$ is produced by magnetic microtraps whose variation is described by a control function $u : (0, T) \rightarrow \mathbb{R}$.

Among models with higher complexity, we find multiparticle Schrödinger-type equations that allow us to describe quantum systems like molecules. In particular, we mention the time-dependent Kohn–Sham equation [126, 223] and the n -particle Schrödinger equation; see, e.g., [368] and the references given therein. For these two classes of systems, optimization and control are topics of future research efforts.

In the following, we discuss in detail the solution of optimal control problems governed by (7.16) and (7.17). By doing this, we address some numerical optimization issues which arise in quantum control problems. In particular, we encounter the formulation of optimization operators in complex Hilbert spaces. We start discussing a control problem with (7.16) and a dipole control mechanism.

A Dipole Quantum Control Problem

In the quantum mechanical framework, a dynamically stable system like an atom or a molecule exists in the presence of a stationary confining potential with a “well” envelope [101]. The states of this system are defined based on the following eigenproblem

$$\{-\Delta + V_0(x) - \lambda_j\} \phi_j(x) = 0, \quad j = 1, 2, \dots \quad (7.18)$$

The eigenfunctions ϕ_j represent the stationary states and the eigenvalues λ_j represent the corresponding energy. Like in the finite-dimensional case, the time evolution of these states is formally given by $\psi_j(x, t) = \phi_j(x) \exp(-i\lambda_j t)$. A representative stationary potential with various applications in semiconductor nanostructures, which is meaningful in one

spatial dimension, with $\Omega = (-\ell/2, \ell/2)$, is the following infinite barrier well potential

$$V_0(x) = 0 \quad \text{for } x \in \Omega \quad \text{and} \quad V_0(\pm\ell/2) = +\infty.$$

The infinite barrier condition is equivalent to homogeneous Dirichlet boundary conditions for the wavefunction, and thus we have $\lambda_j = \frac{j^2\pi^2}{\ell^2}$ and $\phi_j(x) = \cos(j\pi x/\ell)$.

To discuss some mathematical properties of the present framework, we give some definitions. We write $\operatorname{Re}(z)$ and $\operatorname{Im}(z)$ for the real and imaginary part of a complex $z \in \mathbb{C}$. Moreover, z^* stands for the complex conjugate of z and $|z|_{\mathbb{C}} = \sqrt{z^*z}$ for its absolute value. We define $\mathcal{H} = L^2(\Omega; \mathbb{C})$, the Hilbert space endowed with the inner product

$$(\varphi, \psi)_{\mathcal{H}} = \int_{\Omega} \varphi(x)^* \psi(x) dx \quad \text{for } \varphi, \psi \in \mathcal{H}$$

and the induced norm $\|\varphi\|_{\mathcal{H}}$ for $\varphi \in \mathcal{H}$. The Hilbert space $\mathcal{V} = H_0^1(\Omega; \mathbb{C})$ is given by

$$\mathcal{V} = \left\{ \varphi \in \mathcal{H} \mid \|\varphi\|_{\mathcal{V}} = \left(\int_{\Omega} |\nabla \varphi(x)|_{\mathbb{C}}^2 dx \right)^{1/2} < \infty, \quad \varphi = 0 \text{ on } \partial\Omega \right\},$$

supplied with the inner product $(\varphi, \psi)_{\mathcal{V}} = (\nabla \varphi, \nabla \psi)_{\mathcal{H}}$ for $\varphi, \psi \in \mathcal{V}$ and the induced norm $\|\varphi\|_{\mathcal{V}}$. We also need the Hilbert space

$$\mathcal{W} = L^2(0, T; H_0^1(\Omega; \mathbb{C}) \cap H^2(\Omega; \mathbb{C})) \cap H^1(0, T; L^2(\Omega; \mathbb{C})).$$

For more details on the above Lebesgue and Sobolev spaces see, e.g., [124].

Now, consider the TDSE with an initial state of the quantum system given by $\psi_0(x) \in \mathcal{V}$ at $t = 0$. One recognizes that the Schrödinger evolution operator is time-reversible (nondissipative) and therefore ψ cannot have better regularity than ψ_0 [71]. We also see that with a time-varying potential there is no energy conservation. In fact, we want to change the energy of the system. However, we have mass conservation as stated by the following proposition; see, e.g., [101].

Proposition 7.1. *Let $V(x, t) = V_0(x) + u(t)x$ and $\|\psi_0(\cdot)\|_{\mathcal{H}} = 1$; then we have $\|\psi(\cdot, t)\|_{\mathcal{H}} = 1$ for all $t \in [0, T]$.*

An important result of perturbation theory is that in a long time horizon a time-harmonic control $u(t)$ is able to induce transition between two eigenstates $i \rightarrow j$ if its frequency equals the difference of energy of the two states; see, e.g., Chapter 15 of [101]. Therefore, it is relatively easy to control state transitions for long time intervals. However, the problem becomes very difficult if short time intervals are considered. Here short means that $T \approx 2\pi/\omega$, where $\omega = \lambda_j - \lambda_i$, and in this case the perturbation theory is inapplicable and the control $u(t)$ deviates greatly from an harmonic function. In this case, an optimal control approach is unavoidable.

To formulate the optimal control problem, we have to decide in which functional space the control is sought. From the previous discussion, it appears that the control space $\mathcal{U} = H_0^1(0, T; \mathbb{R})$ is the most appropriate for dipole controls, as it means that the laser pulse cannot change instantaneously and it accommodates sinusoidal functions for long time controls. In the optimal control framework, this choice means that the objective of the

optimization has a regularization term of the form $\|u\|_{\mathcal{U}}^2$, where this norm is induced by the following inner product

$$(u, v)_{\mathcal{U}} = \int_0^T (u(t)v(t) + \alpha \dot{u}(t)\dot{v}(t)) dt \quad \text{for } u, v \in \mathcal{U}, \quad (7.19)$$

with $0 < \alpha \ll 1$. Notice that with this norm the control is continuous since $H_0^1(0, T)$ is compactly embedded in $C^0([0, T])$ in one dimension. Use of smaller values of α allows for controls with larger rates of change. We have that the control is zero at the beginning and at the end of the time interval which is the maximum time window for the laser pulse.

The present control problem requires finding a control $u \in \mathcal{U}$ such that a quantum system initially in the state ψ_0 evolves with (7.16) to a state $\psi(\cdot, T)$ that is as close as possible to a desired target configuration ψ_d . This aim is formulated by requiring us to minimize the objective given by the cost functional

$$\min J(\psi, u) := \frac{1}{2} \left(1 - \|P\psi(\cdot, T)\|_{\mathcal{H}}^2 \right) + \frac{\gamma}{2} \|u\|_{\mathcal{U}}^2, \quad (7.20)$$

under the constraint given by the TDSE, including the initial condition, denoted as

$$c(\psi, u) := \{i \partial_t - H(u)\} \psi = 0, \quad (7.21)$$

where $H(u) = -\Delta + V_0(x) + u(t)x$ and we introduce the projector $P\psi = (\psi_d, \psi)_{\mathcal{H}} \psi_d$. The goal of the first term of the cost functional is to track the given terminal state ψ_d up to a global phase $e^{i\varphi T}$ which cannot be specified.

A control suitable for fast quantum state transition can be obtained in the optimal control formulation given by (7.20) and (7.21). To characterize the solution to this problem, we introduce the following Lagrangian

$$L(\psi, u, p) = J(\psi, u) + Re \int_0^T \int_{\Omega} p^*(x, t) c(\psi, u)(x, t) dx dt, \quad (7.22)$$

where p is the Lagrange multiplier. In [356] it is proved that any minima of (7.20) and (7.21) correspond to an extremal point of the Lagrangian; see also [235, 239]. Therefore, taking the Fréchet derivatives of $L(\psi, u, p)$ with respect to the optimization variables gives the following first-order optimality system that characterizes the optimal solution. We have

$$\begin{aligned} \{i \partial_t + \Delta - V_0(x) - u(t)x\} \psi(x, t) &= 0, \\ \{i \partial_t + \Delta - V_0(x) - u(t)x\} p(x, t) &= 0, \\ -\gamma u(t) + \gamma \alpha \ddot{u}(t) + Re \int_{\Omega} p^*(x, t) x \psi(x, t) dx &= 0. \end{aligned} \quad (7.23)$$

This system consists of the state equation, the adjoint equation, and the optimality condition, respectively, with homogeneous Dirichlet boundary conditions, and initial and terminal conditions given by

$$\begin{aligned} \psi(x, 0) &= \psi_0(x), \\ p(x, T) &= i (\psi_d(\cdot), \psi(\cdot, T))_{\mathcal{H}} \psi_d(x), \\ u(0) &= 0, \quad u(T) = 0. \end{aligned} \quad (7.24)$$

In [356] it is proved that there exists at least one solution to (7.23)–(7.24). In this reference it is also proved that if γ is sufficiently large and the projection $\|P\psi(T)\|_{\mathcal{H}}^2$ is sufficiently close to 1, then the second-order sufficient optimality condition holds and the optimization problem is locally strictly convex.

Notice that the control is a function of time only, and the state and adjoint variables can be seen as implicit functions of the control. Therefore the dimensionality of the optimization problem can be reduced significantly, introducing a reduced cost functional $\hat{J}(u) = J(\psi(u), u)$. The corresponding gradient is given by

$$(\nabla \hat{J}(u))(t) = \gamma u(t) - \gamma \alpha \ddot{u}(t) - \operatorname{Re} \int_{\Omega} p^*(x, t) x \psi(x, t) dx. \quad (7.25)$$

Therefore we have that $\nabla \hat{J}(u) \in H^{-1}(0, T; \mathbb{R})$, which is problematic with a gradient-based approach because the gradient is not in the same space of the solution and thus it does not provide an update to the control along the descent direction. It has been shown [354] that this problem can be solved by using the Riesz representation of the gradient in the $H_0^1(0, T)$ space as a means of Sobolev smoothing.

We remark that a main difficulty in the analysis of quantum control problems is that they may admit multiple solutions (as most bilinear control problems). This is the case of the present control problem. In fact, we have the following proposition [356].

Proposition 7.2. *Let the initial and target states be eigenfunctions and the stationary potential be symmetric. Then the reduced cost functional does not have a unique minimizer. In particular, if $u^*(t)$ is a minimizer, then so is $-u^*(t)$ and consequently $\hat{J}(u)$ is nonconvex independently of the values of the regularization parameters.*

The Discrete Dipole Quantum Control Problem and Numerical Results

We recall the introductory discussion in Chapter 3 and pursue the approach of discretize-before-optimize in order to avoid any discrepancy between the directional derivative and its approximation, and to guarantee a symmetric Hessian; see [16, 99]. In our experience, gradient inconsistency is usually not negligible in the case of hyperbolic- and Schrödinger-type equations with bilinear controls.

Regarding discretization of the Schrödinger equation, the time-discretization schemes discussed in Chapter 3 are not appropriate since they do preserve unitarity. Furthermore, in the case of time-varying Hamiltonian, also the CN scheme is not norm-preserving. For this reason, in [356] a modified Crank–Nicolson (MCN) method is proposed which is proved to be second-order accurate and unconditionally stable also in the case of time-varying potentials. Let N_t be the number of time steps $t_k = k \delta t$, $k = 0, 1, \dots, N_t$, where the size $\delta t = \frac{T}{N_t}$, and let N_x be the number of grid points of the Ω discretization including the boundary points. The TDSE discretized by the MCN scheme results in the following

$$\psi_k - \psi_{k-1} = -\frac{i\delta t}{4}[H(t_k) + H(t_{k-1})][\psi_k + \psi_{k-1}]. \quad (7.26)$$

Let $\mathbf{y}_k \in \mathbb{C}^{N_x-2}$ be a vector which contains the values of ψ on the interior grid points at the k th time step. Let $N_d = (N_x - 2)(N_t - 1)$ be the total number degrees of freedom of the state variable, which excludes the initial and boundary points.

Based on [355], we discuss a formulation in complex spaces of the discrete equality constraint (7.26) and of the optimization operators involved. It results in a more compact notation than the real-imaginary matrix notation used in [356]. We introduce the following matrices

$$\mathbf{A}_1 = \mathbf{I} + \frac{i\delta t}{2}\mathbf{H}_0, \quad \mathbf{A}_2 = \mathbf{I} - \frac{i\delta t}{2}\mathbf{H}_0, \quad \mathbf{B} = \frac{i\delta t}{4}\mathbf{X}, \quad (7.27)$$

where \mathbf{H}_0 is the discretization of the stationary Hamiltonian $H_0 = -\Delta + V_0(x)$, using linear finite elements on a uniform grid, and \mathbf{X} is the discretized position operator. The k th discrete equality constraint may now be written as follows

$$\begin{aligned} \mathbf{c}_k(\mathbf{y}, \mathbf{u}) = & [\mathbf{A}_1 - (u_k + u_{k-1})\mathbf{B}]\mathbf{y}_k \\ & - [\mathbf{A}_2 + (u_k + u_{k-1})\mathbf{B}]\mathbf{y}_{k-1}. \end{aligned} \quad (7.28)$$

To take variations using the Wirtinger calculus [346], we write the cost in terms of the discretized wavefunction and its complex conjugate

$$J(\mathbf{y}, \bar{\mathbf{y}}, \mathbf{u}) = \frac{1}{2}(1 - \bar{\mathbf{y}}^\top \mathbf{P}\mathbf{y}) + \frac{\gamma}{2}\mathbf{u}^\top \mathbf{W}\mathbf{u}, \quad (7.29)$$

where $^\top$ means transpose, and $\mathbf{P} \in \mathbb{R}^{N_x \times N_d}$ is the discrete projection operator, defined by

$$\mathbf{P}\mathbf{y} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ (\mathbf{y}_d^\top \mathbf{y}_{N_t})\mathbf{y}_d \end{pmatrix}, \quad (7.30)$$

where \mathbf{W} is the finite element approximation to the Helmholtz operator $I - \alpha \partial_t^2$.

We consider the discrete Lagrangian, with Lagrange multipliers $\mathbf{p}_1, \dots, \mathbf{p}_{N_t}$ and $\bar{\mathbf{p}}_1, \dots, \bar{\mathbf{p}}_{N_t}$:

$$\begin{aligned} L(\mathbf{y}, \bar{\mathbf{y}}, \mathbf{u}, \mathbf{p}, \bar{\mathbf{p}}) = & J(\mathbf{y}, \bar{\mathbf{y}}, \mathbf{u}) \\ & + \sum_{k=1}^{N_t} \mathbf{p}_k^\top \mathbf{c}_k(\mathbf{y}, \mathbf{u}) + \bar{\mathbf{p}}_k^\top \bar{\mathbf{c}}_k(\mathbf{y}, \mathbf{u}). \end{aligned} \quad (7.31)$$

Here, the variables are understood to be column vectors. The components of the derivatives of the Lagrange function are row vectors given as follows

$$L_{\mathbf{y}} = -\frac{1}{2}\bar{\mathbf{y}}^\top \mathbf{P} + \mathbf{p}^\top \mathbf{c}_{\mathbf{y}}, \quad (7.32)$$

$$L_{\bar{\mathbf{y}}} = -\frac{1}{2}\mathbf{y}^\top \mathbf{P} + \bar{\mathbf{p}}^\top \bar{\mathbf{c}}_{\bar{\mathbf{y}}}, \quad (7.33)$$

$$L_{\mathbf{u}} = \gamma \mathbf{u}^\top \mathbf{W} + \mathbf{p}^\top \mathbf{c}_{\mathbf{u}} + \bar{\mathbf{p}}^\top \bar{\mathbf{c}}_{\mathbf{u}}, \quad (7.34)$$

$$L_{\mathbf{p}} = \mathbf{c}^\top, \quad (7.35)$$

$$L_{\bar{\mathbf{p}}} = \bar{\mathbf{c}}^\top. \quad (7.36)$$

With this notation, the discrete equality constraint $\mathbf{c}^\top \in \mathbb{R}^{1 \times N_d}$ is a row vector given by $\mathbf{c}^\top = (\mathbf{c}_1^\top \quad \mathbf{c}_2^\top \quad \dots \quad \mathbf{c}_{N_t}^\top)$.

We collect the components (7.32)–(7.36) to define the full gradient

$$\nabla L = \begin{pmatrix} L_{\mathbf{y}}, & L_{\bar{\mathbf{y}}}, & L_{\mathbf{u}}, & L_{\mathbf{p}}, & L_{\bar{\mathbf{p}}} \end{pmatrix}.$$

The Hessian is defined as the Jacobian of the Hermitian conjugate of this gradient, which we denote with $\nabla^2 L = \nabla(\nabla L)^*$. This approach results in the following Hermitian Hessian

$$\nabla^2 L = \begin{pmatrix} L_{yy} & 0 & L_{yu} & 0 & \mathbf{c}_y^* \\ 0 & L_{\bar{y}\bar{y}} & L_{\bar{y}u} & \mathbf{c}_y^\top & 0 \\ L_{yu}^* & L_{\bar{y}u}^\top & L_{uu} & \mathbf{c}_u^\top & \mathbf{c}_u^* \\ 0 & \bar{\mathbf{c}}_y & \bar{\mathbf{c}}_u & 0 & 0 \\ \mathbf{c}_y & 0 & \mathbf{c}_u & 0 & 0 \end{pmatrix}. \quad (7.37)$$

To make the notion clear, $\mathbf{c}_y \in \mathbb{C}^{N_d \times N_d}$ is the part of the Jacobian of the equality constraint obtained by differentiating with respect to the state \mathbf{y} , \mathbf{c}_y^\top is its transpose without complex conjugation, \mathbf{c}_y^* is the Hermitian conjugate, and $\bar{\mathbf{c}}_y$ is its complex conjugate without transposition. These operators are given in expanded matrix form in the following

$$\begin{aligned} \mathbf{c}_u &= - \begin{pmatrix} \mathbf{B}(\mathbf{y}_1 + \mathbf{y}_0) & & & & \\ \mathbf{B}(\mathbf{y}_2 + \mathbf{y}_1) & \mathbf{B}(\mathbf{y}_2 + \mathbf{y}_1) & & & \\ & \mathbf{B}(\mathbf{y}_3 + \mathbf{y}_2) & \mathbf{B}(\mathbf{y}_3 + \mathbf{y}_2) & & \\ & & \ddots & \ddots & \\ & & & \mathbf{B}(\mathbf{y}_{N_t} + \mathbf{y}_{N_t-1}) \end{pmatrix}, \\ \mathbf{c}_y &= \begin{pmatrix} \mathbf{A}_1 - (u_1 + u_0)\mathbf{B} & & & & \\ -\mathbf{A}_2 - (u_2 + u_1)\mathbf{B} & \mathbf{A}_1 - (u_1 + u_0)\mathbf{B} & & & \\ & \ddots & \ddots & \ddots & \\ & & -\mathbf{A}_2 - (u_{N_t} + u_{N_t-1})\mathbf{B} & \mathbf{A}_1 - (u_{N_t} + u_{N_t-1})\mathbf{B} \end{pmatrix}, \\ L_{yu} &= - \begin{pmatrix} \mathbf{B}^*(\bar{\mathbf{p}}_1 + \bar{\mathbf{p}}_2) & \mathbf{B}^*\bar{\mathbf{p}}_2 & & & \\ \mathbf{B}^*\bar{\mathbf{p}}_2 & \mathbf{B}^*(\bar{\mathbf{p}}_2 + \bar{\mathbf{p}}_3) & \mathbf{B}^*\bar{\mathbf{p}}_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \mathbf{B}^*\bar{\mathbf{p}}_{N_t-1} & \mathbf{B}^*(\bar{\mathbf{p}}_{N_t-1} + \bar{\mathbf{p}}_{N_t}) & \\ & & & \mathbf{B}^*\bar{\mathbf{p}}_{N_t} \end{pmatrix}, \\ L_{yy} = L_{\bar{y}\bar{y}} &= - \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & 0 & 0 \\ 0 & \cdots & 0 & -\frac{1}{2}\mathbf{P} \end{pmatrix}, \\ L_{uu} &= \gamma \mathbf{W}. \end{aligned}$$

It is important to recognize that the complex differentiation and transposition operators do not commute, so the order of operations is critical and this notation is employed to make clear that in each case differentiation occurs first and then conjugation second.

Using the fact that $\mathbf{c} = 0$ and $L_{\mathbf{y}} = 0$ by construction, the full KKT Hessian and gradient equations allow us to compute the differential change in the state and adjoint variables due to a differential change in the control

$$\delta \mathbf{y} = -\mathbf{c}_y^{-1} \mathbf{c}_u \delta \mathbf{u}, \quad (7.38)$$

$$\delta \mathbf{p} = -\mathbf{c}_y^{-\top} (L_{\bar{y}\bar{y}} \delta \bar{\mathbf{y}} + L_{\bar{y}u} \delta \mathbf{u}). \quad (7.39)$$

Using the Schur reduction on the KKT system allows us to write the Newton method in terms of the reduced cost functional as follows

$$\nabla^2 \hat{J}(\mathbf{u}) \delta \mathbf{u} = -\nabla \hat{J}^*(\mathbf{u}),$$

where $\nabla \hat{J}^*(\mathbf{u}) = L_{\mathbf{u}}^* = L_{\mathbf{u}}^\top$, and the application of the reduced Hessian on $\delta \mathbf{u}$ is given by

$$\nabla^2 \hat{J}(\mathbf{u}) \delta \mathbf{u} = \mathbf{L}_{\mathbf{uu}} \delta \mathbf{u} + 2 \operatorname{Re}[\mathbf{c}_{\mathbf{u}}^\top \delta \mathbf{p} + \mathbf{L}_{\mathbf{yu}}^\top \delta \mathbf{y}]. \quad (7.40)$$

Based on this framework, a globalized Krylov–Newton scheme and the corresponding code are presented in [355]. This methodology is discussed in detail in Section 4.2. In the following, we report results obtained with the Krylov–Newton scheme.

In Figure 7.5, optimal controls for different state transitions are depicted and in Table 7.2, we report the decrease of the norm of the gradient for the method of steepest descent and for the Krylov–Newton method for successive iterations. From this table, typical Newton’s second-order convergence can be seen. The Krylov–Newton method attains a gradient which is zero to machine precision in just 8 steps. For the test of Table 7.2, 400 time steps and 200 spatial grid points were used with $T = 1$, $\alpha = 10^{-2}$, $\gamma = 10^{-3}$, $\psi_0(x) = \cos(\pi x/2)$, $\psi_d(x) = \frac{1}{\sqrt{10}}[\cos(\pi x/2) + 3\cos(\pi x)]$.

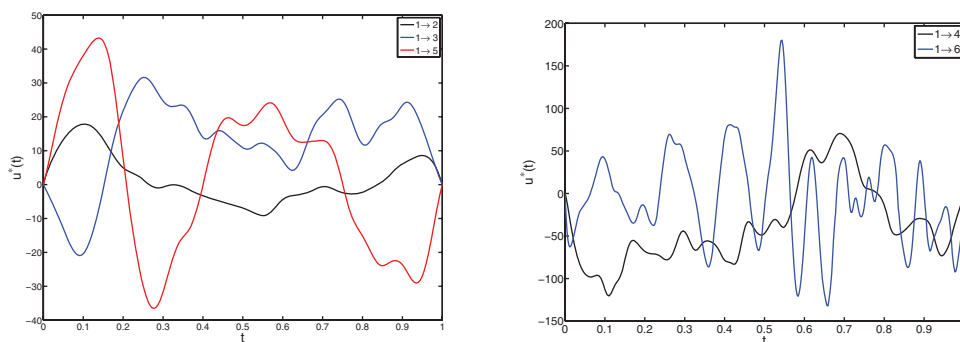


Figure 7.5. Left: Optimal controls for transitions from the first state to the second, the third, and the fifth states. Right: Optimal controls for transitions from the first state to the fourth and the sixth states. This figure first appeared in G. von Winckel, A. Borzi, and S. Volkwein, *A globalized Newton method for the accurate solution of a dipole quantum control problem*, *SIAM J. Sci. Comput.*, 31(6) (2009), 4176–4203.

Optimal Control of BECs

Quantum optimal control of transport of BECs in magnetic microtraps was proposed [199] to solve the problem for a trapped-atom interferometer setup where a dilute BEC should be split from a single to a double-well ground state. We recall that a coherent BEC is described by the following

$$i \partial_t \psi(x, t) = \left\{ -\Delta + V(x, u(t)) + g |\psi(x, t)|^2 \right\} \psi(x, t), \quad (7.41)$$

where $x \in \Omega$ and $t \in [0, T]$. We measure mass in units of the atom mass and length in units of micrometers. We assume that the quantum state wavefunction $\psi(x, t)$ is normalized to

Table 7.2. Convergence of the steepest descent scheme and of the Krylov–Newton scheme with respect to $\|\nabla \hat{J}(u)\|$.

Iteration	$\ \nabla \hat{J}_{SD}(u)\ $	$\ \nabla \hat{J}_{KN}(u)\ $
1	1.8615×10^{-4}	1.8615×10^{-4}
2	6.5263×10^{-5}	6.5263×10^{-5}
3	6.0031×10^{-5}	2.4732×10^{-5}
4	2.3535×10^{-5}	1.5557×10^{-5}
5	2.8106×10^{-5}	1.2316×10^{-6}
6	1.5703×10^{-5}	1.0977×10^{-8}
7	1.7062×10^{-5}	3.5480×10^{-13}
8	1.0322×10^{-5}	2.0009×10^{-17}
9	1.3312×10^{-5}	0
50	1.9114×10^{-7}	0

one, $\int_{\Omega} |\psi(x, t)|^2 dx = 1$, $t \geq 0$, and therefore g in (7.41) incorporates the number of atoms N_A . The fact that V is a confinement potential results in a wavefunction ψ whose support is localized in a bounded region. Therefore, with Ω we represent a spatial domain that is large enough to represent the support of ψ during evolution and we use periodic boundary conditions.

The magnetic control potential $V_u(x, t) = V(x, u(t))$ is modulated by a control function $u(t)$, and the initial and final potential configurations are given; therefore we require that $u(t)$ take initial and final values of zero and one, respectively. These two extremal values correspond to the case where the potential V_u is convex and to the case where it has a double-well structure. Furthermore we require that V_u be spatially symmetric with respect to the origin of coordinates.

Suppose that initially the system is in the ground state ψ_0 for the potential $V(x, 0)$. Upon varying $u(t)$ in the time interval $t \in [0, T]$ from zero to one, the system will pass through a sequence of states and will end up in the final state $\psi(T)$. Our purpose is to determine an optimal control time evolution that allows us to channel the system from the initial state ψ_0 at time zero to a desired state ψ_d at final time T and ψ_d represents the ground state for the potential $V(x, 1)$ at time T . The ground state for a given potential $V(x, u)$ is defined as the stationary state $\phi(x)$ with $\int_{\Omega} |\phi(x)|^2 dx = 1$ that minimizes the following energy [22]

$$E_u(\phi) = \frac{1}{2} \int_{\Omega} \left(|\nabla \phi(x)|^2 + V_u(x) |\phi(x)|^2 + \frac{g}{2} |\phi(x)|^4 \right) dx. \quad (7.42)$$

See Figure 7.6 for stationary solutions and corresponding initial target states.

We choose the control potential proposed in [231] to create condensates of trapped atoms coupled with a radio frequency fields. We have that

$$V(x, u) = -\frac{u^2 d^2}{8c} x^2 + \frac{1}{c} x^4, \quad (7.43)$$

where $c = 40$ and d is a parameter corresponding to twice the distance of the two minima in the double-well potential.

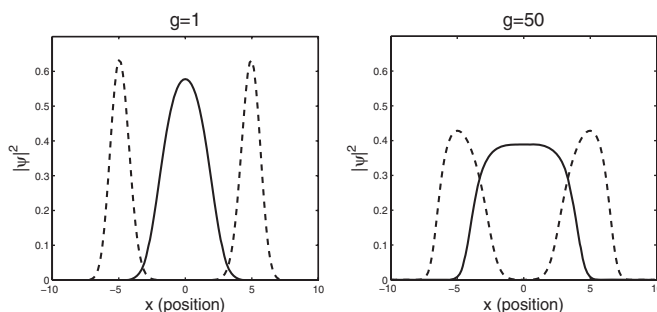


Figure 7.6. Stationary solutions of the Gross–Pitaevskii equation for nonlinearity strengths $g = 1$ and $g = 50$. The initial state ψ_0 is shown with the solid line and the target state ψ_d is represented by the dashed line. Reprinted with permission from G. von Winckel and A. Borzi, *Computational techniques for a quantum control problem with H^1 -cost*, *Inverse Problems*, 24 (2008), 034007.

For our purpose, in order to define a well-defined control problem, we consider the cost functional [199]

$$J(\psi, u) = \frac{1}{2}(1 - |(\psi_d, \psi(T))|^2) + \frac{\gamma}{2} \int_0^T (\dot{u}(t))^2 dt \quad (7.44)$$

with $(u, v) = \int_{\Omega} u(x)^* v(x) dx$ the usual inner product in complex spaces and $\|u\| = (u, u)^{1/2}$. The second term in the cost functional represents an H^1 cost and aims at penalizing fast varying confinement potentials that are more difficult to realize in real experiments.

The control problem under consideration is therefore to minimize $J(\psi, u)$ subject to the condition that $\psi(x, t)$ fulfills the BEC equation (7.41) with given initial conditions. As in the dipole control problem, to solve this problem we introduce a Lagrange function where $p(x, t)$ is the Lagrange multiplier, and we obtain the following optimality system characterizing the solution to the BEC optimal control problem. We have [54]

$$i \frac{\partial \psi}{\partial t} = (-\Delta + V_u + g|\psi|^2) \psi, \quad (7.45)$$

$$i \frac{\partial p}{\partial t} = (-\Delta + V_u + 2g|\psi|^2) p + g \psi^2 p^*, \quad (7.46)$$

$$\gamma \frac{d^2 u}{dt^2} = -\operatorname{Re} \left(\psi, \frac{\partial V_u}{\partial u} p \right), \quad (7.47)$$

which has to be solved together with the initial and terminal conditions

$$\psi(0) = \psi_0, \quad (7.48)$$

$$ip(T) = -(\psi_d, \psi(T)) \psi_d, \quad (7.49)$$

$$u(0) = 0, \quad u(T) = 1. \quad (7.50)$$

Because of H^1 regularization we have a natural setting to impose the required Dirichlet boundary conditions on the control function, $u(0) = 0$ and $u(T) = 1$. See [54] for a discussion concerning second-order optimality conditions.

We have that (7.45) is uniquely solvable for every $u \in H^1(0, T; \mathbb{R})$ such that V_u is a symmetric double-well potential; see [296]. Thus, it is meaningful to introduce the reduced cost functional $\hat{J} : H^1(0, T; \mathbb{R}) \rightarrow \mathbb{R}$ given by $\hat{J}(u) = J(\psi(u), u)$, where $\psi(u)$ denotes the unique solution to (7.45) for given u . One can show that the (L^2) gradient of \hat{J} with respect to u is given by

$$\nabla \hat{J}(u) = -\gamma \frac{d^2 u}{dt^2} - \operatorname{Re} \left(\psi, \frac{\partial V_u}{\partial u} p \right), \quad (7.51)$$

where ψ and p solve the state and the adjoint equations with given u .

Now, the Taylor series of $\hat{J}(u)$ in a Hilbert space is

$$\hat{J}(u + \epsilon \varphi) = \hat{J}(u) + \epsilon (\nabla \hat{J}(u), \varphi)_X + \frac{\epsilon^2}{2} ([\nabla^2 \hat{J}(u)] \varphi, \varphi)_X + O(\epsilon^3).$$

The actual gradient will depend on the choice of which inner product space is used. If we choose the space $X = L^2(0, T; \mathbb{R})$, we have the gradient given by (7.51).

In the case of the $X = H^1(0, T; \mathbb{R})$ formulation, we can determine the formula for the gradient because the Taylor series must be identical term-by-term regardless of the choice of X . Since we know the gradient in L^2 given by (7.51), we can determine the gradient in H^1 by requiring

$$(\nabla \hat{J}_{H^1}(u), \varphi)_{H^1} = (\nabla \hat{J}_{L^2}(u), \varphi)_{L^2}.$$

Using the definition of the H^1 inner product and integrating by parts shows that the H^1 gradient must satisfy the one-dimensional Poisson equation with homogeneous Dirichlet conditions. That is,

$$\frac{d^2}{dt^2} [\nabla \hat{J}(u)] = \gamma \frac{d^2 u}{dt^2} + \operatorname{Re} \left(\psi, \frac{\partial V_u}{\partial u} p \right), \quad (7.52)$$

with $[\nabla \hat{J}(u)](0) = 0$ and $[\nabla \hat{J}(u)](T) = 0$. Two important differences between the L^2 and H^1 formulations are immediately apparent. First, the L^2 gradient does not vanish at $t = T$, and second, the H^1 gradient possesses the same degree of smoothness as u , while the L^2 gradient does not.

The discussion on which optimization space is used is very important also for the numerical implementation. As pointed out in [354], not only is the definition of the gradient also space dependent but also the construction of the optimization schemes. In [354], optimization methods in L^2 and H^1 spaces are discussed with a focus on the nonlinear conjugate gradient scheme of Hager and Zhang [181] and on the BFGS algorithm. These schemes are illustrated in Chapter 4. We remark that our BEC control problem has been successfully solved in [54] by applying the MGOPT optimization strategy.

For the discretization of the BEC control problem, we use an unconditionally stable explicit second-order norm-preserving time-splitting spectral scheme (TSSP); see [23, 54, 354] for more details. We take $\Omega = (-\ell/2, \ell/2) \subset \mathbb{R}$, where ℓ is large enough so that the support of the state and adjoint variables is well within the domain. The spatial mesh size is $h = \ell/N$ with $N = 128$. The time interval $(0, T)$ is divided in M subintervals; thus the time step size is given by $\delta t = T/M$. The function ψ is periodic in Ω in the sense that $\psi(-\ell/2^+) = \psi(\ell/2^-)$.

Some of the results of numerical experiments are summarized in Figure 7.7, which shows a series of optimal control curves corresponding to different time horizons, obtained

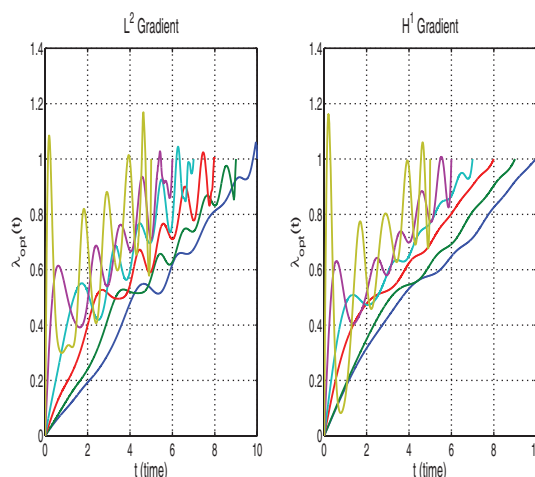


Figure 7.7. The optimal control function in the L^2 and H^1 settings for decreasing time horizons. As T becomes smaller, u becomes a more oscillating function; here $\gamma = 10^{-4}$. Reprinted with permission from G. von Winckel and A. Borzi, *Computational techniques for a quantum control problem with H^1 -cost*, *Inverse Problems*, 24 (2008), 034007.

by minimizing in the L^2 and H^1 spaces. We see that, as T becomes shorter, a more oscillatory optimal control results. However, the solutions obtained with H^1 space minimization are less oscillating. Indeed we see that with the same setting, the controls obtained in the two functional space settings are different.

The regularization parameter γ plays an important role in the L^2 -based optimization, as it forces the solution to have a finite H^1 norm, while this will always be the case by construction with the H^1 -based formulation. In the L^2 space, the limit as $\gamma \rightarrow 0$ small gives a control function $u(t)$ which has a steep gradient at $t = T$. On the other hand, since the H^1 -based scheme produces $u \in H^1$ at every optimization step, the precise value of the regularization term is not essential to obtain a smooth control. Rather in this case the γ term serves only to penalize highly oscillatory solutions. In the limit as $\gamma \rightarrow 0$, the H^1 -based solution approaches a smooth curve which gives a small cost functional.

Additional results are reported in Tables 7.3 through 7.6, which allow a more detailed discussion. These tables show the computational performance of the HZ-NCG and BFGS schemes using the L^2 and H^1 spaces minimization, respectively. With the nonlinearity strength $g = 10$ and final time $T = 6$, this is a difficult optimization problem, the optimal control being far away from the original linear function. The CPU usage in both spaces is comparable. We see that, at convergence and similar computational effort, the H^1 -based approach produces a smaller objective and substantially smaller norm of the gradient on the finest mesh than the L^2 -based approach.

The advantage of the H^1 -based minimization on the L^2 -based scheme is more evident in Tables 7.5 and 7.6, which show the performance of the BFGS scheme in the L^2 and H^1 spaces, respectively. The L^2 -based BFGS method requires progressively more iterations to reach the minimum, whereas the H^1 approach requires fewer iterations on the

Table 7.3. Results with L^2 -based HZ-NCG minimization with $g = 10$ and $T = 6$.

M	\hat{f}	$\ \nabla \hat{f}\ _{L^2}$	Iterations	CPU time (sec)
400	2.9676×10^{-3}	4.4182×10^{-2}	81	1.6271×10^2
800	1.6492×10^{-3}	2.113×10^{-2}	110	4.1789×10^2
1600	1.8103×10^{-3}	1.1400×10^{-2}	155	1.1932×10^3
3200	3.4755×10^{-3}	1.13815×10^{-1}	76	1.1985×10^3

Table 7.4. Results with H^1 -based HZ-NCG minimization with $g = 10$ and $T = 6$.

M	\hat{f}	$\ \nabla \hat{f}\ _{L^2}$	Iterations	CPU time (sec)
400	3.0295×10^{-2}	2.7397×10^{-1}	24	3.3911×10^1
800	2.2888×10^{-3}	2.9977×10^{-2}	217	7.9421×10^2
1600	1.1995×10^{-3}	1.8310×10^{-2}	153	1.1986×10^3
3200	2.4388×10^{-3}	4.5315×10^{-3}	74	1.1883×10^3

Table 7.5. Results with L^2 -based BFGS minimization with $g = 10$ and $T = 6$.

M	\hat{f}	$\ \nabla \hat{f}\ _{L^2}$	Iterations	CPU time (sec)
400	1.8922×10^{-3}	5.8348×10^{-2}	25	6.1990×10^1
800	1.2472×10^{-3}	9.7978×10^{-2}	34	1.7211×10^2
1600	3.5128×10^{-4}	1.3702×10^{-2}	105	9.6031×10^2
3200	1.1394×10^{-4}	5.8215×10^{-3}	500	8.1292×10^3

Table 7.6. Results with H^1 -based BFGS minimization with $g = 10$ and $T = 6$.

M	\hat{f}	$\ \nabla \hat{f}\ _{L^2}$	Iterations	CPU time (sec)
400	1.6605×10^{-2}	1.4288×10^{-1}	15	3.8407×10^1
800	5.5963×10^{-4}	4.5284×10^{-2}	62	2.8107×10^2
1600	2.9634×10^{-4}	1.0733×10^{-2}	30	3.6334×10^2
3200	1.0562×10^{-4}	3.6378×10^{-3}	37	9.6153×10^2

finest mesh, similar to NCG. Compared to the previous two tables, the BFGS scheme provides at least one order of magnitude better results in terms of the value of the objective and concerning CPU times. See [354] for additional results.

Figure 7.8 shows results for a wavefunction splitting for an optimized and a linear [199] control function [54]. One can see that the wavefunction becomes split along the time evolution and is transported in the respective minima. Excited vibrational states that originate during the initial splitting process can be seen plotting the absolute value of the wavefunction profiles at final time. These are evident in the case of a linearly varying u where the split is not complete and part of the population remains localized between the two wells.

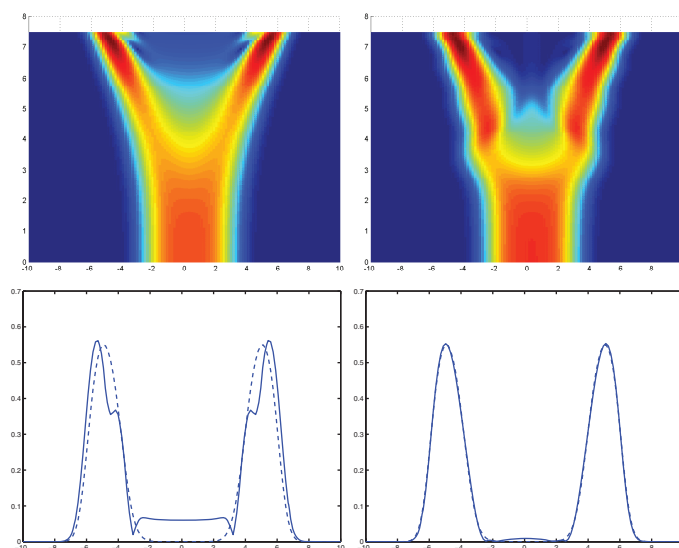


Figure 7.8. The function $|\psi(x,t)|$ on the space-time domain (top) for the linear (left) and optimized (right) controls. The corresponding profiles at final time (bottom, continuous line) compared to the desired state (dashed line). Mesh 128×1250 ; $\gamma = 10^{-4}$. This figure first appeared in A. Borzi and U. Hohenester, *Multigrid optimization schemes for solving Bose–Einstein condensate control problems*, *SIAM J. Sci. Comput.*, 30(1) (2008), 441–462.

7.4 Electromagnetic Inverse Scattering

7.4.1 Introduction

In this section, we discuss electromagnetic inverse scattering problems in the PDE-based optimization framework. These problems arise in many different practical application fields and are among the most representative classes of inverse problems. We provide details of the formulation of the inverse scattering problem as a PDE optimization problem and discuss some aspects of its numerical implementation.

Consider an incident electromagnetic wave which propagates in a known background medium where there are objects with electromagnetic properties different from the background. Because of this difference, these objects scatter the electromagnetic waves, and with this scenario, we can distinguish two kinds of scattering problems. The first category is that of direct scattering problems where the incident waves, the background medium, and the characteristics of the objects are given and computation of the scattered field generated by the interaction of the incident waves with the objects is required. The inverse scattering problems represent the second category of problems where the incident waves, the scattered waves, and the medium are given and we have the problem of deducing the features of the unknown objects. These problems are very important in applications concerning noninvasive imaging. In fact, because electromagnetic waves can penetrate various media where objects may be placed, the solution of inverse scattering problems allows us to determine the presence and the properties of these objects hidden in the medium.

We focus on electromagnetic inverse scattering problems in the space-time domain [31], trying to determine the electromagnetic properties of objects in the medium, and the geometric structure and position of these scatterers. We consider the time-domain formulation that naturally accommodates data corresponding to a large spectrum of possible frequencies of the incident field [159, 161, 163, 164, 246, 269, 287, 335]. Notice that the Maxwell equations could be considered in the frequency domain with monochromatic fields and one would need to process the data in the frequency domain with particular techniques; see, e.g., [298]. In the discussion that follows, a two-dimensional spatial domain with a model given by the Maxwell curl equations is considered. In particular, we consider a two-dimensional scattering scenario and the governing Maxwell curl equations including initial conditions, and radiation conditions. The objective includes observation of the electric and magnetic fields' data misfit and regularization terms of zero- and first-order Tikhonov type [105] for the electromagnetic properties that are sought. To characterize the solution of the optimization problem, the first-order optimality system is discussed, including the formulation of the radiation conditions for the adjoint equations. The implementation of radiation conditions is important to avoid unphysical wave reflections at the boundary of the computation domain. However, see [161, 164] for simplified boundary conditions.

To solve the optimality system, the finite difference time domain (FDTD) method is illustrated in the one-dimensional case, and we report some details concerning numerical dispersion, stability, and accuracy of FDTD solutions. Notice that the scientific literature of discretization schemes for Maxwell problems is vast; see, e.g., [257] and the references given therein. Further, we discuss the perfectly matched layer scheme to implement free propagation in the free space of the scattered waves. This section is completed with numerical results with an inversion problem.

7.4.2 The Scattering Problem in the Time Domain

We formulate a time-domain electromagnetic inverse scattering problem in the framework of optimization with PDE constraints. We start introducing a general two-dimensional scattering scenario and the governing Maxwell curl equations with initial and radiation boundary conditions. Our objective is to minimize a functional of electromagnetic field data misfit, including a regularization functional for the electromagnetic properties that are sought. To characterize the solution of the resulting optimization problem, we derive the corresponding first-order optimality system.

Consider the two-dimensional scenario illustrated in Figure 7.9.

This scenario consists of a domain V which can be unbounded. If it is bounded, we denote with ∂V its boundary. The domain V represents also a free space background medium, characterized by the electric permittivity ε_0 and the magnetic permeability μ_0 . In a subdomain $D \subset V$, called scatterer domain, we imagine that we have a spatial distribution of the electromagnetic properties $\varepsilon(\vec{r})$, $\mu(\vec{r})$, and $\sigma(\vec{r})$, which are different from those of the background medium. Outside the domain D , we have electromagnetic sources, i.e., an antenna that generates an electromagnetic field that propagates in the domain V and also in the domain D where the scattering occurs.

In the direct scattering problem, we know the sources, the background medium, and also the distribution of electromagnetic properties in the scatterer domain D . The purpose is to calculate the fields $(\vec{E}(\vec{r}, t), \vec{H}(\vec{r}, t))$ in V by solving the following Maxwell curl

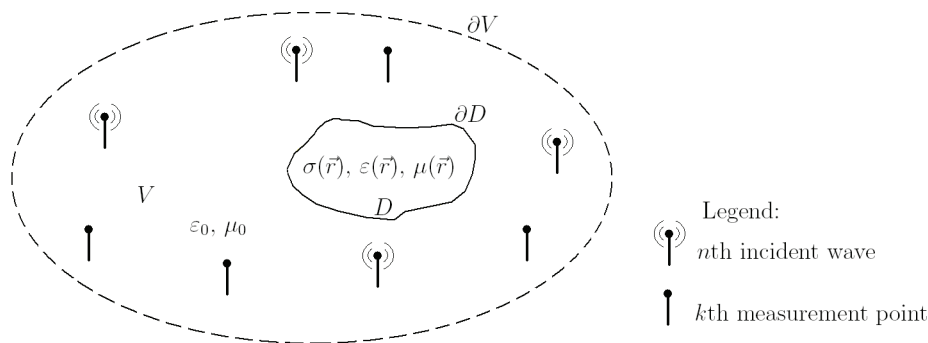


Figure 7.9. A scenario of direct and inverse scattering problems.

equations

$$\begin{cases} \nabla \times \vec{E} + \mu \frac{\partial \vec{H}}{\partial t} = 0, \\ \nabla \times \vec{H} - \epsilon \frac{\partial \vec{E}}{\partial t} - \sigma \vec{E} - J = 0, \end{cases} \quad (7.53)$$

where we have used the compact notation (\vec{E}, \vec{H}) instead of $(\vec{E}(\vec{r}, t), \vec{H}(\vec{r}, t))$. This set of equation needs in addition to specify the initial conditions and the boundary conditions that the evolving field must satisfy. The initial conditions set the values of the electromagnetic field at the starting time, $t = 0$. The boundary conditions prescribe the behavior of the fields on the boundary of the domain V for all times. Given the system (7.53) and the initial and boundary conditions, we can solve the wave equations for the fields.

In an inverse scattering problem, we have antennas and receivers placed in certain positions in V , external to the scatterer domain D , which generate and receive and record the fields during the time interval $[0, T]$. With this data, we apply our inversion procedure to determine the properties $\mathbf{p} = (\epsilon(\vec{r}), \mu(\vec{r}), \sigma(\vec{r}))$ in D , knowing the sources and the background medium. Since the presence and positioning of the scatterers in D are unknown, a procedure that is defined on all the subdomain D must be applied. In this procedure, the domain D is illuminated by N different incident waves and, for every incidence wave, we measure the fields in K different measurement points during the time interval $[0, T]$.

To solve the inverse scattering problem, one defines a functional that contains the discrepancies between the measured fields and the calculated fields, and the aim is at minimizing the misfit by varying the unknown electromagnetic properties which define the optimization variables. Because the ill-posedness of this problem, a regularization term is required and the following functional results

$$J(\mathbf{p}, \vec{E}, \vec{H}) = \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^K \int_0^T \left(\|\vec{E}_{nk} - \vec{E}_{nk}^m\|^2 + \eta_0^2 \|\vec{H}_{nk} - \vec{H}_{nk}^m\|^2 \right) dt + R(\mathbf{p}), \quad (7.54)$$

where \vec{E}_{nk} and \vec{H}_{nk} are the calculated fields, whereas \vec{E}_{nk}^m and \vec{H}_{nk}^m are the measured fields. These four fields are evaluated in the k th measurement points for the n th incident wave.

Here, $R(\mathbf{p})$ represents the regularization term given by the following

$$R(\mathbf{p}) = \frac{\beta}{2} \int_D \left(|\varepsilon_0^{-1} \nabla \varepsilon|^2 + |\nabla \sigma|^2 + |\mu_0^{-1} \nabla \mu|^2 \right) dv + \frac{\gamma}{2} \int_D \left(|\varepsilon - \bar{\varepsilon}|^2 + |\sigma - \bar{\sigma}|^2 + |\mu - \bar{\mu}|^2 \right) dv. \quad (7.55)$$

In this regularization functional we distinguish two different terms. The first term represents a first-order Tikhonov regularization scheme, as in [161, 287]. A regularization term similar to (7.55) is considered in [164]. Minimizing the functional J with this type of regularization means that we search the model's parameters \mathbf{p} in a smooth space with spatially slow varying functions. A result of this choice is that we obtain the edges of the reconstructed objects smeared out by diffusion. The other regularization term represents a zero-order Tikhonov regularization, and it corresponds to the requirement that the ε sought be as close as possible to $\bar{\varepsilon}$, and similarly σ be as close as possible to $\bar{\sigma}$, and μ be as close as possible to $\bar{\mu}$. This regularization term can be used even if we do not have a priori information with known $\bar{\varepsilon}$, $\bar{\mu}$, $\bar{\sigma}$. In this case we choose these values to be zero and the resulting regularization term penalizes large parameters values.

The objective defined by (7.54) is to be minimized under the constraints given by the Maxwell curl equations for each incidence field. We have

$$\begin{cases} \nabla \times \vec{E}_n + \mu \frac{\partial \vec{H}_n}{\partial t} = 0, \\ \nabla \times \vec{H}_n - \varepsilon \frac{\partial \vec{E}_n}{\partial t} - \sigma \vec{E}_n - \vec{J}_n = 0, \end{cases} \quad (7.56)$$

$$\vec{E}_n|_{t=0} = 0, \quad \vec{H}_n|_{t=0} = 0, \quad (7.57)$$

$$\lim_{r \rightarrow \infty} r \left[\nabla \times \begin{pmatrix} \vec{E}_n \\ \vec{H}_n \end{pmatrix} + \frac{1}{c_0} \hat{n}_{V_\infty} \times \frac{\partial}{\partial t} \begin{pmatrix} \vec{E}_n \\ \vec{H}_n \end{pmatrix} \right] = 0, \quad (7.58)$$

$$\Sigma(\vec{E}_n, \vec{H}_n, \vec{J}_n) = f(\vec{r}, t). \quad (7.59)$$

The system (7.56) is constituted by the Maxwell curl equations and the initial conditions (7.57), and (7.59) represents the (hard and soft) sources. The limit (7.58) represents the Sommerfeld radiation condition at infinity, where $k_0 = \omega \sqrt{\mu_0 \varepsilon_0} = 2\pi f \sqrt{\mu_0 \varepsilon_0}$ and \hat{n}_{V_∞} is the outward-pointing vector normal to ∂V , the surface boundary of V_∞ . This constraint means that, for every estimated \mathbf{p} and for the n th incident wave, the calculated fields have to satisfy the Maxwell curl equations.

To solve this infinite-dimensional constrained optimization problem in the framework of PDE optimization, we introduce the following Lagrange functional

$$\begin{aligned} L(\mathbf{p}, \vec{E}, \vec{H}, \vec{e}, \vec{h}) &= J(\mathbf{p}, \vec{E}, \vec{H}) \\ &+ \sum_{n=1}^N \int_0^T \int_V \left(\vec{e}_n \cdot \left(\nabla \times \vec{H}_n - \varepsilon \frac{\partial \vec{E}_n}{\partial t} - \sigma \vec{E}_n - \vec{J}_n \right) \right) dt dv \\ &+ \sum_{n=1}^N \int_0^T \int_V \left(\vec{h}_n \cdot \left(\nabla \times \vec{E}_n + \mu \frac{\partial \vec{H}_n}{\partial t} \right) \right) dt dv, \end{aligned} \quad (7.60)$$

where \vec{e}_n and \vec{h}_n represent the Lagrange multipliers.

The characterization of the solution of this constrained minimization problem results in the optimality conditions of the first order that are formally given by the conditions for extrema of the Lagrange functional. For this purpose, we compute the derivatives of the Lagrange functional (7.60) with respect to $\vec{E}, \vec{H}, \vec{e}, \varepsilon, \mu, \sigma$, with the inner product given by

$$(\vec{\varphi}, \vec{\vartheta}) = \sum_{n=1}^N \int_0^T \int_V (\vec{\varphi} \cdot \vec{\vartheta}) \, dv dt.$$

We also use the L^2 -norm in space which is denoted as follows

$$||\vec{\varphi}||^2 = \int_V (\vec{\varphi} \cdot \vec{\varphi}) \, dv. \quad (7.61)$$

We have that the functional derivative with respect to the Lagrange multipliers gives the Maxwell curl equations. Further, the derivatives with respect to the fields give the adjoint equations and the derivatives with respect to the electromagnetic properties provide the optimality conditions; see [246] for details. The following optimality system is obtained:

- The direct problem

$$\begin{cases} \nabla \times \vec{E}_n + \mu \frac{\partial \vec{H}_n}{\partial t} = 0, \\ \nabla \times \vec{H}_n - \varepsilon \frac{\partial \vec{E}_n}{\partial t} - \sigma \vec{E}_n - J_n = 0, \end{cases}$$

$$\vec{E}_n|_{t=0} = 0, \quad \vec{H}_n|_{t=0} = 0,$$

$$\lim_{r \rightarrow \infty} r \left[\hat{n}_{V_\infty} \times \left(\nabla \times \begin{pmatrix} \vec{E}_n \\ \vec{H}_n \end{pmatrix} \right) - j k_0 \begin{pmatrix} \vec{E}_n \\ \vec{H}_n \end{pmatrix} \right] = 0,$$

$$\Sigma(\vec{E}_n, \vec{H}_n, \vec{J}_n) = f(\vec{r}, t);$$

- The adjoint problem

$$\begin{cases} \nabla \times \vec{e}_n - \mu \frac{\partial \vec{h}_n}{\partial t} + \eta_0^2 \sum_{k=1}^K (\vec{H}_n - \vec{H}_n^m) \delta(\vec{r} - \vec{r}_k) = 0, \\ \nabla \times \vec{h}_n + \varepsilon \frac{\partial \vec{e}_n}{\partial t} - \sigma \vec{e}_n + \sum_{k=1}^K (\vec{E}_n - \vec{E}_n^m) \delta(\vec{r} - \vec{r}_k) = 0, \end{cases}$$

$$\vec{e}_n|_{t=T} = 0, \quad \vec{h}_n|_{t=T} = 0,$$

$$\lim_{r \rightarrow \infty} r \left[\hat{n}_{V_\infty} \times \left(\nabla \times \begin{pmatrix} \vec{e}_n \\ \vec{h}_n \end{pmatrix} \right) + j k_0 \begin{pmatrix} \vec{e}_n \\ \vec{h}_n \end{pmatrix} \right] = 0;$$

- The optimality conditions with respect to μ , ε , and σ , respectively

$$\begin{cases} \sum_{n=1}^N \int_0^T \vec{h}_n \cdot \frac{\partial \vec{H}_n}{\partial t} dt - \beta \mu_0^{-2} \nabla^2 \mu + \gamma(\mu - \bar{\mu}) = 0 \\ - \left(\sum_{n=1}^N \int_0^T \vec{e}_n \cdot \frac{\partial \vec{E}_n}{\partial t} dt + \beta \varepsilon_0^{-2} \nabla^2 \varepsilon - \gamma(\varepsilon - \bar{\varepsilon}) \right) = 0 \\ - \left(\sum_{n=1}^N \int_0^T \vec{e}_n \cdot \vec{E}_n dt + \beta \nabla^2 \sigma - \gamma(\sigma - \bar{\sigma}) \right) = 0. \end{cases}$$

Notice that the Lagrange multipliers represent ingoing waves whereas the fields \vec{E} and \vec{H} are outgoing waves. The Lagrange multipliers have the same nature as electromagnetic field variables. The sources that drive these fields are the discrepancies between the measured and calculated fields at the measurement points.

The residuals of the optimality conditions provide the minus gradient of the objective functional with respect to the optimization variables, and therefore they represent the descent direction that is required in any gradient-based optimization procedure as those discussed in Chapter 4. We present results obtained with a steepest descent scheme; see [161, 164] concerning the application of quasi-Newton schemes to electromagnetic inverse problems.

7.4.3 The Maxwell Curl Equations

In order to illustrate some approximation aspects of the Maxwell curl equations and of the corresponding adjoint equations, we discuss the FDTD discretization of these equations with different spatial dimensions as follows.

Consider the following four Maxwell equations

$$\begin{cases} \nabla \cdot \vec{B} = 0 & \text{Gauss's law for the magnetic field,} \\ \nabla \cdot \vec{D} = 0 & \text{Gauss's law for the electric field,} \\ \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} - \vec{M} & \text{Faraday's law,} \\ \nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{J} & \text{Ampere's law.} \end{cases} \quad (7.62)$$

In a linear, isotropic, and nondispersive medium, we can write

$$\vec{D} = \varepsilon \vec{E} = \varepsilon_0 \varepsilon_r \vec{E}, \quad \vec{B} = \mu \vec{H} = \mu_0 \mu_r \vec{H}$$

and

$$\vec{J} = \vec{J}_s + \sigma \vec{E}, \quad \vec{M} = \vec{M}_s + \sigma^* \vec{H}.$$

Thus, the equations in (7.62) become

$$\begin{cases} \frac{\partial \vec{H}}{\partial t} = -\frac{1}{\mu} \nabla \times \vec{E} - \frac{1}{\mu} (\vec{M}_s + \sigma^* \vec{H}), \\ \frac{\partial \vec{E}}{\partial t} = \frac{1}{\varepsilon} \nabla \times \vec{H} - \frac{1}{\varepsilon} (\vec{J}_s + \sigma \vec{E}). \end{cases} \quad (7.63)$$

In Cartesian coordinates, the equations (7.63) have the following components

$$\begin{cases} \frac{\partial H_x}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} - (M_{sx} + \sigma^* H_x) \right), \\ \frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} - (M_{sy} + \sigma^* H_y) \right), \\ \frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - (M_{sz} + \sigma^* H_z) \right), \end{cases} \quad (7.64)$$

$$\begin{cases} \frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - (J_{sx} + \sigma E_x) \right), \\ \frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} - (J_{sy} + \sigma E_y) \right), \\ \frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - (J_{sz} + \sigma E_z) \right). \end{cases} \quad (7.65)$$

These six coupled PDEs are the fundamental equations in the three-dimensional case.

The two-dimensional case is obtained setting $\partial/\partial z = 0$ in (7.64) and in (7.65). Then, we can group the six equations in the following two independent set of equations

$$\begin{cases} \frac{\partial H_x}{\partial t} = \frac{1}{\mu} \left(-\frac{\partial E_z}{\partial y} - (M_{sx} + \sigma^* H_x) \right), \\ \frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_z}{\partial x} - (M_{sy} + \sigma^* H_y) \right), \\ \frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - (J_{sz} + \sigma E_z) \right) \end{cases} \quad (7.66)$$

and

$$\begin{cases} \frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_z}{\partial y} - (J_{sx} + \sigma E_x) \right), \\ \frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon} \left(-\frac{\partial H_z}{\partial x} - (J_{sy} + \sigma E_y) \right), \\ \frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - (M_{sz} + \sigma^* H_z) \right). \end{cases} \quad (7.67)$$

The set in (7.66) is called the transverse magnetic mode (TM_z), and the set in (7.67) is called the transverse electric mode (TE_z).

For completeness, we write the equations for the one-dimensional case. We impose $\partial/\partial y = 0$ in both modes and obtain

$$\begin{cases} \frac{\partial H_x}{\partial t} = -\frac{1}{\mu} (M_{sx} + \sigma^* H_x), \\ \frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_z}{\partial x} - (M_{sy} + \sigma^* H_y) \right), \\ \frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_y}{\partial x} - (J_{sz} + \sigma E_z) \right) \end{cases}$$

and

$$\begin{cases} \frac{\partial E_x}{\partial t} = -\frac{1}{\varepsilon} (J_{sx} + \sigma E_x), \\ \frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon} \left(-\frac{\partial H_z}{\partial x} - (J_{sy} + \sigma E_y) \right), \\ \frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left(-\frac{\partial E_y}{\partial x} - (M_{sz} + \sigma^* H_z) \right). \end{cases}$$

Now, we pose M_{sx} and J_{sx} zero for every t and also H_x zero in $t = 0$. Consequently, we have $E_x = 0$ and $H_x = 0$ for every t . Then, we obtain the following two sets of equations

$$\begin{cases} \frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_z}{\partial x} - (M_{sy} + \sigma^* H_y) \right), \\ \frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_y}{\partial x} - (J_{sz} + \sigma E_z) \right) \end{cases} \quad (7.68)$$

and

$$\begin{cases} \frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon} \left(-\frac{\partial H_z}{\partial x} - (J_{sy} + \sigma E_y) \right), \\ \frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left(-\frac{\partial E_y}{\partial x} - (M_{sz} + \sigma^* H_z) \right). \end{cases} \quad (7.69)$$

The system (7.68) is called TEM mode x -directed and z -polarized. The system (7.69) is called TEM mode x -directed and y -polarized.

7.4.4 The FDTD Discretization

To evaluate the functional gradient and therefore define an optimization step, we need to solve the direct Maxwell problem and the adjoint Maxwell equations. This solution procedure can be formulated in the framework of FDTD methods; see [120, 333] for a detailed introduction to this methodology. We remark that the Maxwell equations are first-order PDEs. In this case, it is known that stable finite difference schemes are obtained in a staggered grid. It was Yee [366] who first proposed such a scheme defined on a staggered grid that is now called the Yee grid. The Yee algorithm is based upon the discretization of the Maxwell curl equations in time and space domain with central difference formula on a staggered grid. The scheme does not enforce the Gauss laws, because they are a consequence of the two Maxwell curl equations. The electric and magnetic fields are sampled on different discrete spatial points and at different discrete temporal points.

To illustrate the Yee scheme in the three-dimensional case, consider a cubic spatial domain in \mathbb{R}^3 , divided in N_x , N_y , and N_z cells in the x , y , and z directions, respectively. In this lattice we can identify a spatial point as $(i, j, k) = (i \Delta x, j \Delta y, k \Delta z)$, where i , j , k are integers and Δx , Δy , and Δz are the spatial increments on the x , y , and z axes, respectively. The generic single cell in the position (i, j, k) is called the Yee cell and is illustrated in Figure 7.10.

Also the time is discretized, with an index n and time increment Δt , uniform over the observation interval. A function u in a three-dimensional grid on the point (i, j, k) and at the n th discretized time instant is individuated by

$$u(i \Delta x, j \Delta y, k \Delta z, n \Delta t) = u_{ijk}^n.$$

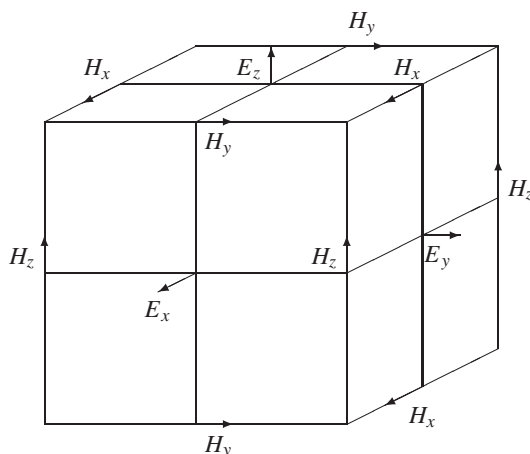


Figure 7.10. *The Yee cell.*

The sampling instants of the electric and magnetic fields are not the same. The magnetic field is sampled at the instants $0, \Delta t, 2\Delta t, \dots, n\Delta t, \dots$, and the electric field is sampled at $1/2\Delta t, (1 + 1/2)\Delta t, \dots, (n + 1/2)\Delta t, \dots$. Second-order accurate centered differences are used for the first-order derivatives [333].

Consider the one-dimensional case, where we have increments only in the x and the t axis. We start with the discretization of (7.68), more precisely with the first equation. We imagine evaluating the equation in the point $(i, n + 1/2)$. We have

$$\frac{H_{y|i}^{n+1} - H_{y|i}^n}{\Delta t} = \frac{1}{\mu_i} \left(\frac{E_{z|i+1/2}^{n+1/2} - E_{z|i-1/2}^{n+1/2}}{\Delta x} - M_{s y|i}^{n+1/2} - \sigma_i^* H_{y|i}^{n+1/2} \right).$$

With the hypothesis that the field H is evaluated every time step Δt , we do not have its value in $(n + 1/2)\Delta t$. Because of this, we use a semi-implicit approximation as follows

$$H_{y|i}^{n+1/2} = \frac{H_{y|i}^n + H_{y|i}^{n+1}}{2}.$$

Substituting this and manipulating the equation, we obtain

$$H_{y|i}^{n+1} = \frac{1 - \frac{\Delta t \sigma_i^*}{2\mu_i}}{1 + \frac{\Delta t \sigma_i^*}{2\mu_i}} H_{y|i}^n + \frac{\frac{\Delta t}{\mu_i \Delta x}}{1 + \frac{\Delta t \sigma_i^*}{2\mu_i}} \left(E_{z|i+1/2}^{n+1/2} - E_{z|i-1/2}^{n+1/2} - \Delta x M_{s y|i}^{n+1/2} \right).$$

In the same manner, we discretize the second equation of (7.68) on $(i - 1/2, n)$ as follows

$$\frac{E_{z|i-1/2}^{n+1/2} - E_{z|i-1/2}^{n-1/2}}{\Delta t} = \frac{1}{\varepsilon_{i-1/2}} \left(\frac{H_{y|i}^n - H_{y|i-1}^n}{\Delta x} - J_{s z|i-1/2}^n - \sigma_{i-1/2} E_{z|i-1/2}^n \right).$$

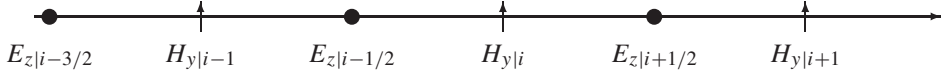


Figure 7.11. Spatial arrangement for E_z and H_y in the one-dimensional case.

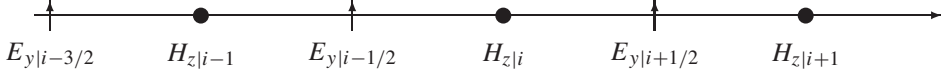


Figure 7.12. Spatial arrangement for E_y and H_z in the one-dimensional case.

Using the semi-implicit approximation on E , since its value in $n\Delta t$ is not stored, we have

$$E_{z|i-1/2}^{n+1/2} = \frac{1 - \frac{\Delta t \sigma_{i-1/2}}{2\varepsilon_{i-1/2}}}{1 + \frac{\Delta t \sigma_{i-1/2}}{2\varepsilon_{i-1/2}}} E_{z|i-1/2}^{n-1/2} + \frac{\frac{\Delta t}{\varepsilon_{i-1/2} \Delta x}}{1 + \frac{\Delta t \sigma_{i-1/2}}{2\varepsilon_{i-1/2}}} \left(H_{y|i}^n - H_{y|i-1}^n - \Delta x J_{sz|i-1/2}^n \right).$$

The same reasoning is applied to the set of equations (7.69). Therefore we obtain the following discretized set of equations, and their spatial arrangements are given in Figures 7.11 and 7.12

$$\begin{cases} H_{y|i}^{n+1} = \frac{1 - \frac{\Delta t \sigma_i^*}{2\mu_i}}{1 + \frac{\Delta t \sigma_i^*}{2\mu_i}} H_{y|i}^n + \frac{\frac{\Delta t}{\mu_i \Delta x}}{1 + \frac{\Delta t \sigma_i^*}{2\mu_i}} \left(E_{z|i+1/2}^{n+1/2} - E_{z|i-1/2}^{n+1/2} - \Delta x M_{sy|i}^{n+1/2} \right), \\ E_{z|i-1/2}^{n+1/2} = \frac{1 - \frac{\Delta t \sigma_{i-1/2}}{2\varepsilon_{i-1/2}}}{1 + \frac{\Delta t \sigma_{i-1/2}}{2\varepsilon_{i-1/2}}} E_{z|i-1/2}^{n-1/2} + \frac{\frac{\Delta t}{\varepsilon_{i-1/2} \Delta x}}{1 + \frac{\Delta t \sigma_{i-1/2}}{2\varepsilon_{i-1/2}}} \left(H_{y|i}^n - H_{y|i-1}^n - \Delta x J_{sz|i-1/2}^n \right), \\ E_{y|i-1/2}^{n+1/2} = \frac{1 - \frac{\Delta t \sigma_{i-1/2}}{2\varepsilon_{i-1/2}}}{1 + \frac{\Delta t \sigma_{i-1/2}}{2\varepsilon_{i-1/2}}} E_{y|i-1/2}^{n-1/2} + \frac{\frac{\Delta t}{\varepsilon_{i-1/2} \Delta x}}{1 + \frac{\Delta t \sigma_{i-1/2}}{2\varepsilon_{i-1/2}}} \left(H_{z|i-1}^n - H_{z|i}^n - \Delta x J_{sy|i-1/2}^n \right), \\ H_{z|i}^{n+1} = \frac{1 - \frac{\Delta t \sigma_i^*}{2\mu_i}}{1 + \frac{\Delta t \sigma_i^*}{2\mu_i}} H_{z|i}^n + \frac{\frac{\Delta t}{\mu_i \Delta x}}{1 + \frac{\Delta t \sigma_i^*}{2\mu_i}} \left(E_{y|i-1/2}^{n+1/2} - E_{y|i+1/2}^{n+1/2} - \Delta x M_{sz|i}^{n+1/2} \right). \end{cases}$$

Typically, in the implementation of the FDTD algorithm it is not needed to logically offset the fields. Thus, when considering the $i = 1$ position in the arrays which contain the values of the fields E_z and H_y , we have the values $E_{z|1/2}$ and $H_{y|1}$.

While we do not write the equations for the two-dimensional case and refer the reader to [246, 287] for details, we would like to show the spatial arrangements of the electromagnetic fields in the Yee cell as depicted in Figure 7.10. The TE_z mode is obtained if we consider the fields in the plane $k + 1/2$; the TM_z mode is obtained in the cut plane k . Thus, in the bidimensional lattice, the two modes are as shown in the Figures 7.13 and 7.14.

The FDTD scheme may result in nonphysical dispersion of the waves in the grid, while the need of a stable scheme requires us to satisfy a relation between the time step and the space increment, thus influencing the numerical accuracy of the solution. Concerning dispersion, one obtains waves with phase velocities different from the phase velocities

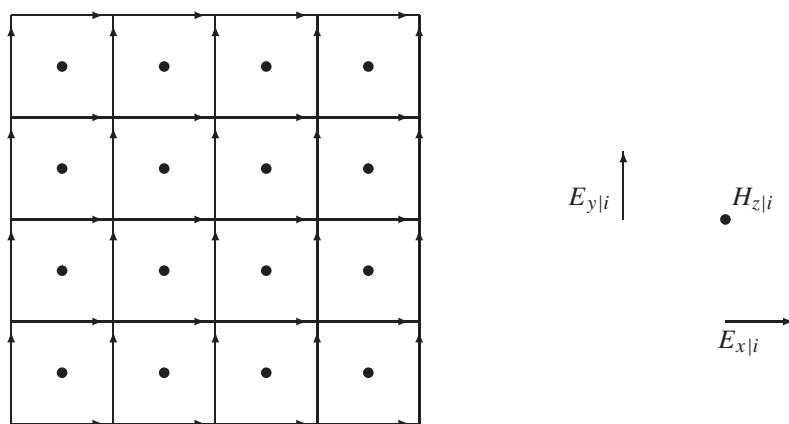


Figure 7.13. Spatial arrangement for the TE_z mode. In the right-hand picture, the fields of a single cell are depicted.

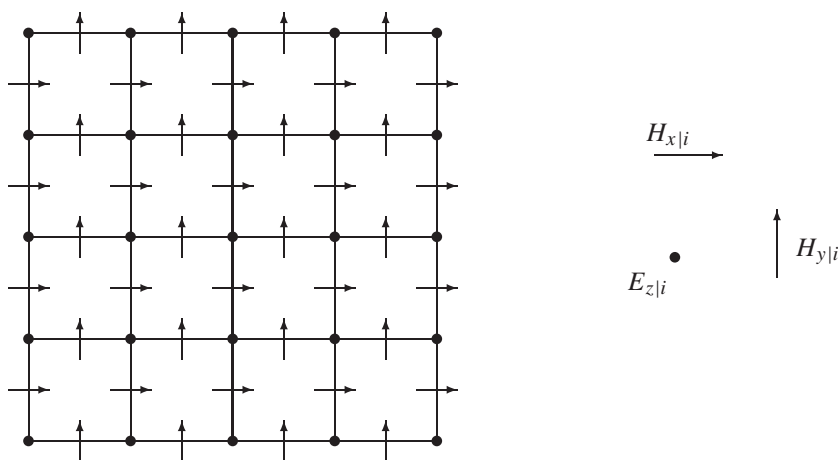


Figure 7.14. Spatial arrangement for the TM_z mode. In the right-hand picture, the fields of a single cell are depicted.

resulting from the physical dispersion. The dispersion relation of the continuous Maxwell model is given by

$$k_x^2 + k_y^2 = \left(\frac{\omega}{c}\right)^2,$$

and the numerical dispersion relation is given by

$$\left[\frac{\omega}{2c} \frac{\sin(\omega \Delta x/2)}{\omega \Delta x/2}\right]^2 = \left[\frac{k_x}{2} \frac{\sin(k_x \Delta x/2)}{k_x \Delta x/2}\right]^2 + \left[\frac{k_y}{2} \frac{\sin(k_y \Delta y/2)}{k_y \Delta y/2}\right]^2.$$

Notice that if $\Delta t \rightarrow 0$, $\Delta x \rightarrow 0$, and $\Delta y \rightarrow 0$, the numerical dispersion relation reduces to the ideal dispersion relation.

Concerning stability, stability imposes a bound on the maximum time step relative to the spatial increments as follows (Courant condition)

$$\frac{c \Delta t}{\Delta s} \leq \frac{1}{\sqrt{d}}, \quad (7.70)$$

where Δs is the length of a side of a uniform cell and d is the number of space dimensions in the problem.

Concerning accuracy of the FDTD scheme, we report results given in [258], where it is proved that the FDTD scheme is second-order accurate on general nonuniform rectangular grids. Consider a three-dimensional cavity given by a parallelepiped $\Omega = [0, L_x] \times [0, L_y] \times [0, L_z]$. We assume that the electric field satisfies a perfectly conducting (PEC) boundary condition on the boundary of the domain for all $t \in [0, T]$ as

$$\vec{E} \times \vec{n} = 0,$$

and we assume the following initial conditions

$$\vec{E}(\vec{r}, 0) = \vec{E}_0(\vec{r}), \quad \vec{H}(\vec{r}, 0) = \vec{H}_0(\vec{r}).$$

We also request particular regularity properties of the field solutions of the Maxwell equations (7.62) in the generic three-dimensional case as follows

$$\vec{E} \in C\left([0, T]; [C^3(\Omega)]^3\right), \quad \vec{H} \in C\left([0, T]; [C^3(\Omega)]^3\right) \cap C^1\left([0, T]; [C^2(\Omega)]^3\right).$$

Let \vec{E} and \vec{H} be the continuous fields and \vec{E}^h and \vec{H}^h be the discretized fields. With these hypotheses, it is possible to prove the following theorem [258].

Theorem 7.3. *Suppose that \vec{E} and \vec{H} are three times continuously differentiable on V , that $\partial \vec{H} / \partial t$ is twice continuously differentiable on V , and that all these derivative are continuous in time. Then, for any fixed $T > 0$, there is a constant C depending on T such that for $0 \leq t \leq T$, we have*

$$\|\vec{E}(t) - \vec{E}^h(t)\|_E + \|\vec{H}(t) - \vec{H}^h(t)\|_H \leq Ch^2, \quad (7.71)$$

where $\|\cdot\|_E$ and $\|\cdot\|_H$ are L^2 -norms on the staggered grid.

7.4.5 Perfectly Matched Layer

An FDTD scheme applies on computational domains that are finite, while we need to simulate waves that radiate outward in the unbounded free space. To model this phenomenon with a bounded computational domain, Berenger [33] defined an innovative absorbing material scheme called the perfectly matched layer (PML). The computational domain extends over this medium that generates reflection neither at the interface between the main domain and the PML nor at the end of the PML. This scheme requires matching at the interface of material properties and attenuation of the field propagating in the PML region; see Figure 7.15.

In the one-dimensional case the PML represents a lossy medium, with electric conductivity σ and magnetic conductivity σ^* and without any anisotropy; i.e., the electromagnetic characteristics do not depend on the direction. Consider a TEM mode x -directed and

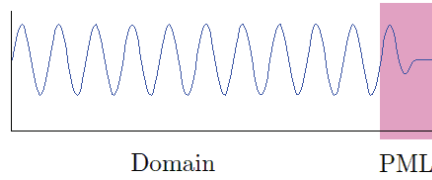


Figure 7.15. Idea for the truncation of computational domain in the one-dimensional case. At the interface between the main domain (Region 1) and the PML region (Region 2), the wave is not reflected and it is attenuated.

z -polarized in a structure similar to that depicted in Figure 7.15. Assume the wave propagating from the lossless Region 1 to Region 2 that corresponds to the PML region. We assume that the PML region ends with a PEC boundary. We assume that Region 2 is made of lossy material. The Maxwell equations that govern the wave propagation in the two regions are as follows

$$\begin{cases} \frac{\partial H_y}{\partial t} = \frac{1}{\mu_1} \frac{\partial E_z}{\partial x}, \\ \frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon_1} \frac{\partial H_y}{\partial x}, \end{cases} \quad (7.72)$$

$$\begin{cases} \frac{\partial H_y}{\partial t} = \frac{1}{\mu_2} \left(\frac{\partial E_z}{\partial x} - \sigma^* H_y \right), \\ \frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon_2} \left(\frac{\partial H_y}{\partial x} - \sigma E_z \right). \end{cases} \quad (7.73)$$

Now, consider the problem of defining different material properties at the interface of the two regions in such a way as to avoid reflection at the interface. For this purpose, consider (7.72) and (7.73) in the frequency domain. We have

$$\begin{cases} j\omega\mu_1 \tilde{H}_y = \frac{\partial \tilde{E}_z}{\partial x}, \\ j\omega\varepsilon_1 \tilde{E}_z = \frac{\partial \tilde{H}_y}{\partial x}, \end{cases} \quad (7.74)$$

$$\begin{cases} j\omega\tilde{H}_y = \frac{1}{\mu_2} \left(\frac{\partial \tilde{E}_z}{\partial x} - \sigma^* \tilde{H}_y \right), \\ j\omega\tilde{E}_z = \frac{1}{\varepsilon_2} \left(\frac{\partial \tilde{H}_y}{\partial x} - \sigma \tilde{E}_z \right). \end{cases} \quad (7.75)$$

In the equations (7.75), corresponding to Region 2, it is possible to group some terms as follows

$$\begin{cases} j\omega\mu_2 \left(1 + \frac{\sigma^*}{j\omega\mu_2} \right) \tilde{H}_y = \frac{\partial \tilde{E}_z}{\partial x}, \\ j\omega\varepsilon_2 \left(1 + \frac{\sigma}{j\omega\varepsilon_2} \right) \tilde{E}_z = \frac{\partial \tilde{H}_y}{\partial x}. \end{cases} \quad (7.76)$$

With these new equations, we can identify the reflectionless conditions. For the normal incidence case the reflection coefficient Γ is given by

$$\Gamma = \frac{\eta_1 - \eta_2}{\eta_1 + \eta_2}, \quad (7.77)$$

where η_1 and η_2 are the impedances in Region 1 and in Region 2, respectively. Note that in the one-dimensional case, there is always normal incidence to the interface.

To obtain the impedance in Region 2 it is sufficient to observe the structure of equations (7.76). We find that, analogously to Region 1, the equivalent electric permittivity ε'_2 and the magnetic permeability μ'_2 are given by

$$\varepsilon'_2 = \varepsilon_2 \left(1 + \frac{\sigma}{j\omega\varepsilon_2} \right), \quad \mu'_2 = \mu_2 \left(1 + \frac{\sigma^*}{j\omega\mu_2} \right).$$

Now, the impedances in Region 1 and in Region 2 are as follows

$$\eta_1 = \sqrt{\frac{\mu_1}{\varepsilon_1}}, \quad \eta_2 = \sqrt{\frac{\mu_2 (1 + \sigma^*/(j\omega\mu_2))}{\varepsilon_2 (1 + \sigma/(j\omega\varepsilon_2))}}.$$

Notice from (7.77) that Γ is zero if the impedances are equal. For this reason, we require that $\eta_1 = \eta_2$ and the following

$$\varepsilon_1 = \varepsilon_2, \quad \mu_1 = \mu_2.$$

Now, we have

$$\sqrt{\frac{\mu_1}{\varepsilon_1}} = \sqrt{\frac{\mu_1}{\varepsilon_1}} \sqrt{\frac{(1 + \sigma^*/(j\omega\mu_1))}{(1 + \sigma/(j\omega\varepsilon_1))}}$$

and we require that

$$\frac{\left(1 + \frac{\sigma^*}{j\omega\mu_1} \right)}{\left(1 + \frac{\sigma}{j\omega\varepsilon_1} \right)} = 1,$$

which implies $\frac{\sigma^*}{\mu_1} = \frac{\sigma}{\varepsilon_1}$, and we obtain

$$\sigma^* = \sigma \frac{\mu_1}{\varepsilon_1} = \sigma(\eta_1)^2. \quad (7.78)$$

The same analysis and properties are obtained for a PML region with a TEM mode x -directed and z -polarized.

Notice that the implementation of a one-dimensional PML region is not completed by the appropriate choice of values σ and σ^* at the interface. In fact in a continuous space, the wave impacts the PML region, crosses this layer, and is reflected by the PEC wall that ends the layer. Then, the wave comes back in the computational domain after a second crossing of the PML layer. Thus we have an effective reflection factor defined as follows

$$R = e^{-2\sigma\eta d},$$

where d is the PML thickness with a PEC wall ending.

Now, the straightforward setting of σ and σ^* involves significant numerical reflection. To reduce this reflection error, Berenger proposed to gradually increase the losses in the PML region, starting from 0 at the interface and increasing to a maximum value towards the end of the region. With this approach, the reflection factor becomes

$$R = e^{-2\eta \int_0^d \sigma(x) dx}. \quad (7.79)$$

Typically, we use a polynomial variation for the PML electric loss, like the following

$$\sigma(x) = \left(\frac{x}{d}\right)^m \sigma_{max}, \quad (7.80)$$

where d is the PML layer length, x is the position in the PML layer, varying from 0 to d , σ_{max} is the electric conductivity in $x = d$, and m is the grading exponential. The integral in (7.79) becomes

$$\int_0^d \sigma(x) dx = \int_0^d \sigma_{max} \left(\frac{x}{d}\right)^m dx = \frac{\sigma_{max}}{d^m} \left[\frac{x^{m+1}}{m+1} \right]_0^d = \frac{\sigma_{max} d}{m+1}.$$

So, with a polynomial grading the reflection factor is given by

$$R = e^{-2\eta \sigma_{max} d/(m+1)}. \quad (7.81)$$

Using this equation with chosen m , d and the desired reflection factor, we can solve the equation for the maximum value of σ at the end of the PML layer. Then, solving (7.81) for σ_{max} , we obtain

$$\sigma_{max} = -\frac{(m+1) \ln(R)}{2\eta d}.$$

With σ_{max} , we can determine σ^* with (7.78). The PML is completely definite.

Typical values of m are chosen to be between 3 and 4, whereas the values of d and R depend on the applications. For example, R may be equal to e^{-8} or e^{-16} ; d may correspond to 5 to 10 cells or more.

Exponential Time Differencing in the PML Region

Consider a computational domain surrounded by a PML layer, with conductivities σ and σ^* , related by (7.78). We can see this relation from another point of view: σ^*/μ and σ/ε are in [Hz], so μ/σ^* and ε/σ are in [sec]. Thus, we can write (7.78) in this manner:

$$\tau = \frac{\varepsilon}{\sigma}, \quad \tau^* = \frac{\mu}{\sigma^*} \quad \Rightarrow \quad \tau = \tau^*.$$

Typically, to have a small size of the PML layer, σ is chosen large, so τ is of order $O(10^{-12})$ s; i.e., the wave attenuation in the PML region can be very rapid. But in the computational domain, the time scale of typical FDTD simulation is $T_c > 10^{-9}$ s. Then, $\tau \ll T_c$. If we set the FDTD parameters to resolve τ , it follows that T_c is overresolved. This implies that, because of the Courant criterion, the spatial cell size Δx overresolves the fields. On the other hand, if we consider only T_c and not τ , the standard FDTD could fail.

Then, we need to introduce an approximation scheme that gradually changes in a way so as to accommodate the variation of the lossy material in the PML region. The resulting

scheme is known as the exponential time differencing (ETD) method; see [275]. The ETD scheme is used in the implementation of the FDTD scheme instead of the standard FDTD such that the setting in the computational grid is that of the classic FDTD scheme.

Consider a TEM mode x -directed and z -polarized in a lossy medium. We have

$$\begin{cases} \frac{\partial E_z}{\partial t} + \frac{\sigma}{\varepsilon} E_z = \frac{1}{\varepsilon} \frac{\partial H_y}{\partial x}, \\ \frac{\partial H_y}{\partial t} + \frac{\sigma^*}{\mu} H_y = \frac{1}{\mu} \frac{\partial E_z}{\partial x}. \end{cases}$$

We focus on the first equation, but the reasoning is the same for the second equation and for the TEM mode x -directed and y -polarized. We can observe that

$$\frac{\partial}{\partial t} \left(e^{(\sigma/\varepsilon)t} E_z \right) = \left(\frac{\partial E_z}{\partial t} + \frac{\sigma}{\varepsilon} E_z \right) e^{(\sigma/\varepsilon)t}.$$

This means that we can rewrite the considered equation, after multiplying both sides by $e^{(\sigma/\varepsilon)t}$ as follows

$$\frac{\partial}{\partial t} \left(e^{(\sigma/\varepsilon)t} E_z \right) = e^{(\sigma/\varepsilon)t} \left(\frac{1}{\varepsilon} \frac{\partial H_y}{\partial x} \right).$$

Using the central difference formula in the point $(i - 1/2, n)$, we have

$$\begin{aligned} & \frac{e^{(\sigma_{i-1/2}/\varepsilon)(n+1/2)\Delta t} E_{z|i-1/2}^{n+1/2} - e^{(\sigma_{i-1/2}/\varepsilon)(n-1/2)\Delta t} E_{z|i-1/2}^{n-1/2}}{\Delta t} \\ &= \frac{1}{\varepsilon} \frac{H_{y|i}^n - H_{y|i-1}^n}{\Delta x} e^{(\sigma_{i-1/2}/\varepsilon)(n)\Delta t}, \end{aligned}$$

that is,

$$\begin{aligned} & e^{(\sigma_{i-1/2}/\varepsilon)(n+1/2)\Delta t} E_{z|i-1/2}^{n+1/2} \\ &= e^{(\sigma/\varepsilon)(n-1/2)\Delta t} E_{z|i-1/2}^{n-1/2} + \frac{\Delta t}{\varepsilon} \frac{H_{y|i}^n - H_{y|i-1}^n}{\Delta x} e^{(\sigma_{i-1/2}/\varepsilon)(n)\Delta t}. \end{aligned} \quad (7.82)$$

Solving with respect to the next field value, i.e., dividing both sides by $e^{(\sigma_{i-1/2}/\varepsilon)(n+1/2)\Delta t}$, we obtain

$$E_{z|i-1/2}^{n+1/2} = E_{z|i-1/2}^{n-1/2} e^{-\sigma_{i-1/2}\Delta t/\varepsilon} + \frac{\Delta t}{\varepsilon \Delta x} (H_{y|i}^n - H_{y|i-1}^n) e^{-\sigma_{i-1/2}\Delta t/(2\varepsilon)}. \quad (7.83)$$

We can now use a Taylor approximation. Developing e^{-x} and approximating that for a sufficiently small x , we obtain

$$e^{-x} \cong e^{-x} \Big|_0 - e^{-x} \Big|_0 x = 1 - x,$$

that is,

$$x = 1 - e^{-x}.$$

Thus, we have

$$\frac{1}{\sigma} \frac{\Delta t \sigma}{\varepsilon} = \frac{1}{\sigma} (1 - e^{-\Delta t \sigma/\varepsilon}).$$

Applying these considerations in (7.83), we obtain

$$E_{z|i-1/2}^{n+1/2} = E_{z|i-1/2}^{n-1/2} e^{-\sigma_{i-1/2} \Delta t / \varepsilon} + \frac{(1 - e^{-\Delta t \sigma_{i-1/2} / \varepsilon})}{\sigma_{i-1/2} \Delta x} e^{-\sigma_{i-1/2} \Delta t / (2\varepsilon)} (H_{y|i}^n - H_{y|i-1}^n).$$

In the same manner we derive the other equations. Thus, the set of equations for a TEM mode x -directed and y -polarized is given by

$$E_{z|i-1/2}^{n+1/2} = E_{z|i-1/2}^{n-1/2} e^{-\sigma_{i-1/2} \Delta t / \varepsilon} + \frac{(1 - e^{-\Delta t \sigma_{i-1/2} / \varepsilon})}{\sigma_{i-1/2} \Delta x} e^{-\sigma_{i-1/2} \Delta t / (2\varepsilon)} (H_{y|i}^n - H_{y|i-1}^n),$$

$$H_{y|i}^{n+1} = H_{y|i}^n e^{-\sigma_i^* \Delta t / \mu} + \frac{(1 - e^{-\Delta t \sigma_i^* / \mu})}{\sigma_i^* \Delta x} e^{-\sigma_i^* \Delta t / (2\mu)} (E_{z|i+1/2}^{n+1/2} - E_{z|i-1/2}^{n+1/2}),$$

whereas the TEM mode x -directed and y -polarized set of equations is as follows

$$E_{y|i-1/2}^{n+1/2} = E_{y|i-1/2}^{n-1/2} e^{-\sigma_{i-1/2} \Delta t / \varepsilon} + \frac{(1 - e^{-\Delta t \sigma_{i-1/2} / \varepsilon})}{\sigma_{i-1/2} \Delta x} e^{-\sigma_{i-1/2} \Delta t / (2\varepsilon)} (H_{z|i}^n - H_{z|i-1}^n),$$

$$H_{z|i}^{n+1} = H_{z|i}^n e^{-\sigma_i^* \Delta t / \mu} + \frac{(1 - e^{-\Delta t \sigma_i^* / \mu})}{\sigma_i^* \Delta x} e^{-\sigma_i^* \Delta t / (2\mu)} (E_{y|i+1/2}^{n+1/2} - E_{y|i-1/2}^{n+1/2}).$$

For a detailed discussion on the PML scheme and ETD scheme in two spatial dimensions we refer the reader to [246].

7.4.6 An Inverse Scattering Scenario

In this section, we report results of validation of the inverse scattering solution procedure based on the PDE optimization framework; see [246] for details and additional results. Consider the computational domain is a square whose edges are three times larger than the edges of a scatter square domain D . We take the edges of D corresponding to a length of 20 cells. The number of time steps is set to $n_{max} = 200$ and the time step size is $\Delta t = 10^{-12}$ s. The time domain is $(0, T)$, where $T = n_{max} \Delta t$ s. The spatial increment, Δ , is equal in the x - and y -directions and it is set equal to $\Delta = 2c\Delta t = 5.99 \cdot 10^{-4}$ m, such that it satisfies the Courant limit discussed in Section 7.4.4. Thus, the length of the computational square domain is 0.036 m. The electromagnetic properties of the free space are given by $\mu_0 = 4\pi \cdot 10^{-7}$ H/m and $\varepsilon_0 = 8.854212 \cdot 10^{-12}$ F/m.

The measurement points and the source points are placed in the region surrounding the domain D . We consider four different incidence points and eight measurement points, i.e., $N = 4$ and $K = 8$. See Figure 7.16.

To generate the incidence fields, hard sources are used. This means a variation of the magnetic field is set directly where the transmitters are placed. For every n the excitation of the magnetic field was taken equal to the following

$$H_z(t) = \begin{cases} \eta_0^{-1} \left[0.3 \sin\left(\frac{2\pi t}{T_1}\right) + \sin\left(\frac{2\pi t}{T_2}\right) \right], & 0 \leq t \leq T_1, \\ 0, & T_1 < t. \end{cases} \quad (7.84)$$

For the evaluation of the inversion algorithm, we consider two hidden scatterers that are two squares with relative electrical permittivities of $\varepsilon_r = 1.5$ and $\varepsilon_r = 2$, respectively.

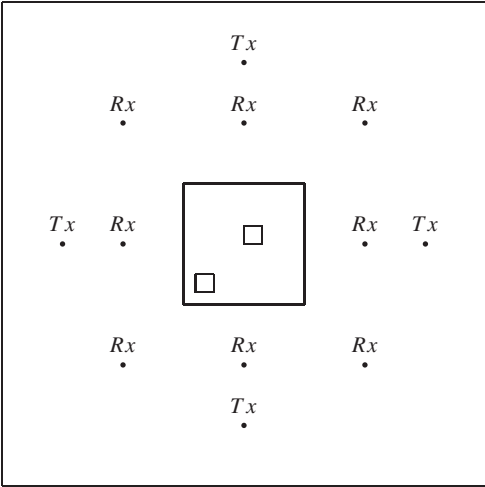


Figure 7.16. Position of transmitters (Tx) and receivers (Rx) in the domain. Two scatterers are depicted in the domain D .

With these scatterers, the associated direct scattering problem is solved and the resulting fields are recorded at the receivers for all times in $(0, T)$ and adding noise to avoid inverse crime in the construction of the synthetic data which is used as measurements. We have a white Gaussian noise such that the signal-to-noise ratio is $SNR = 10$ dB. In Figure 7.17, the two objects used to design the synthetic data are depicted (the smaller square corresponds to $\varepsilon_r = 2$) together with the reconstructed objects (after 20 steps of a steepest descent scheme).

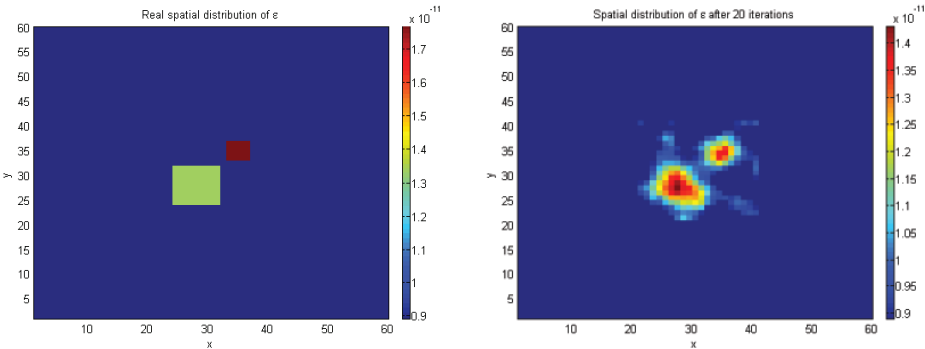


Figure 7.17. Two-dimensional visualization of the two square scatterer scenario (left) and two-dimensional visualization of the reconstructed objects.