Opt Quant Electron (2007) 39:239–260
DOI 10.1007/s11082-007-9077-5

Computation of linear and nonlinear stationary states of photonic structures using modern iterative solvers

José Ramón Salgueiro · David Olivieri · Humberto Michinel

Received: 5 December 2006 / Accepted: 8 May 2007 / Published online: 7 June 2007

Abstract In this paper we describe efficient methods to obtain the stationary states of linear and nonlinear photonic systems, which have gained particular interest in the field of integrated and nonlinear optics. While the methods presented are directly applicable to optical physics, they are also general and should be of interest in a broad range of phenomena presently under study in other areas of physics and engineering. The strategy consists in combining the use of classical methods, such as inverse iteration or the Newton method, together with modern, nonstationary linear solvers, such as SYMMLQ or GMRES, in order to obtain efficient numerical computations to problems involving large matrices. We have selected several example problems in order to discuss the practical implementation details, not normally described in the present literature. Moreover, the problems we have selected provide a backdrop to contrast and motivate the use of different methods for systems which are symmetric and non-symmetric, single and multi-component, and also real and complex. Information relative to numerical performance of the different algorithms, including a survey for a nonsymmetric problem, which requires the adjustment of a restarting parameter for the GMRES algorithm, is also presented.

Keywords Photonic crystal fibers · Solitons · Vortices · Linear solvers · Inverse iteration · Newton methods

J. R. Salgueiro (✉) · H. Michinel
Departamento de Física Aplicada, Universidade de Vigo,
Facultade de Ciencias de Ourense, As Lagoas s/n,
32004 Ourense, Spain
e-mail: jrs@uvigo.es

D. Olivieri
Departamento de Informática,
Universidade de Vigo, E. S. de Enxenería Informática,
As Lagoas s/n, 32004 Ourense, Spain

1 Introduction

Many important problems in physics and engineering require the calculation of stationary solutions of the governing differential operators. For example, in wave optics as well as quantum mechanics the Schrödinger equation plays a fundamental role in the description of propagating beams or pulses through a variety of types of waveguides, as well as the states of quantum mechanical systems given potentials of arbitrary shape. In almost all cases of practical interest, a numerical solution for these states is required, and as such, efficient computational methods are of great interest. This is especially true for multidimensions problems or in cases where the phenomena of interest requires from a very precise description, thus demanding considerable computational requirements.

Numerical methods commonly used to solve partial differential equations are based upon the establishment of a discretized grid which covers the domain where the solution is required in order to obtain a numerical solution at these sets of points. Thus, it is well-known how the continuous differential equation is transformed, through either a finite difference or finite element method, into a set of N linear algebraic equations. For practical problems, these resulting linear systems can be extremely large, especially if they are multidimensional, yet also very sparse, due to the form of the resulting discretization equations. Direct methods, such as LU-decomposition, which need to treat all elements, thus destroying the sparse structure of the matrix, are not computationally feasible for large N due to the number of operations, $\mathcal{O}(N^3)$, and the matrix storage requirements of $\mathcal{O}(N^2)$. Consequently, suitable iterative methods, which properly exploit the sparse structure of the matrix, are necessary.

Classical iterative methods, such as Jacobi, Gauss-Seidel, Successive Over-relaxation (SOR) and others, described in many textbooks (Young 1971; Press et al. 1992; Barrett et al. 1994), normally converge slowly and are thus limited to moderately large problems. These methods are referred to as *stationary* since all iterations are performed using the same information, i.e., the iteration matrix is fixed from the beginning. If we consider that interesting problems are both multidimensional and occur at fine grid resolution, then these methods are practically useless.

A more efficient solution to these systems is found using the modern iterative solvers (Barrett et al. 1994; Greenbaum 1997; Meurant 1999). These methods, named *nonstationary* due to the changing nature of the information used to carry out each iteration, are based on the construction of a set of orthogonal vectors (one per iteration) which progressively increases the dimension of the space where the solution lies and/or restricts that corresponding to the residual. For this reason, these methods converge in a finite number of iterations, at most in N iterations, where N is the order of the matrix. However, convergence is usually achieved in much fewer iterations, thus making these methods particularly useful. Another important characteristic, which makes these *nonstationary* algorithms suitable for dealing with large matrices is the fact that they need only reference the system matrix through its product times a vector. This operation is usually carried out by means of a routine supplied by the user and the form of the matrix, as well as the way it is stored in memory, is unimportant for the algorithm, and carried out in the most efficient way depending on the problem characteristics.

The objective of this paper is to describe practical issues and subtleties of the use of modern iterative solvers for a set of related models in optics. In particular, we demonstrate how we have used iterative solvers combined with standard inverse iteration algorithms (linear problems) and the globally convergent Newton method (nonlinear problems) for obtaining the stationary solutions of several different types of photonic structures, of great interest in integrated and nonlinear optics. The methods we describe are themselves iterative and require at each iteration step the solution of a linear equation system. In particular, we have combined

standard algorithms with iterative solvers, particularly SYMMLQ and GMRES, in order to obtain significant performance gains for large N problems, while still maintaining the fine grain resolution for an accurate description of the relevant physical effects.

This paper is organized as follows. In the next section, we shall propose the general model problem, relevant to light propagation in optical dielectric structures, yet also important in other areas of physics like quantum mechanics. We shall also briefly review the inverse iteration method (Sect. 2.1) and the globally convergent Newton method (Sect. 2.2), each useful for solving linear and nonlinear stationary problems respectively, treating the main aspects of their implementation. The section ends with a summary of the most important available nonstationary methods for solving linear equation systems, discussing issues of efficiency, and types of relevant problems. The main part of the paper, in Sect. 3, is devoted to the description of problems of a variety of types, which we have solved and we use to demonstrate the power of these combined iterative methods. Finally, we consider some issues relative to the algorithm's performance and their dependence with the finess of the grid.

2 Model problems

In order to illustrate the combined iterative methods applied to linear as well as nonlinear problems, we shall initially consider a second order differential equation which describes many important problems in physics. This equation models the stationary states of an optical beam in general inhomogeneous media, and may also model a quantum system described by the time-independent Schrödinger equation with an arbitrary potential. Thus, the general model equation we shall consider takes the form,

$$\nabla^2 u(\mathbf{r}) + V(\mathbf{r}, u)u(\mathbf{r}) = \beta u(\mathbf{r}), \tag{1}$$

where $\mathbf{r}$ is the position vector, $u(\mathbf{r})$ a stationary wave function, ∇^2 the Laplace operator, $V(\mathbf{r}, u)$ a potential function which may depend only on the variables (linear problem), or also on the wave function (nonlinear problem), and β is an eigenvalue. For the quantum mechanical systems, this equation constitutes the so-called time-independent Schrödinger equation which describes the stationary states of a particle in the potential V, characterized by the energies β. Moreover, the equation is derived from the time-dependent Schrödinger equation if a time factorization is performed, considering the wave function of the form $\psi(\mathbf{r}, t) = u(\mathbf{r}) \exp(-\beta t)$.

For the case of the wave equation in optics, however, Eq. (1) describes the continuous beam propagation modes in a medium whose refractive index is given by the function V, and are characterized by the propagation constant β. The equation is derived from the Maxwell equations, after assuming the scalar approximation (Tamir 1990). If we choose z-axis as the propagation direction, the stationary modal problem arises after considering solutions for the electric field of the form $E(\mathbf{r}_\perp, z) = u(\mathbf{r}_\perp) \exp(\beta z)$, where $\mathbf{r}_\perp$ is the position vector in the transverse plane, with components x and y, and β is the z-component of the wave vector. Furthermore, we shall show in Sect. 3.5, that we can also describe the more complex case of a pulsed optical beam, with the addition of a new time variable T, and the term $\partial^2 u/\partial T^2$.

It should be noted that the stationary states may be regarded in certain cases as the modes of the system described by Eq. 1. The concept of mode implies a solution of the eigenvalue problem and the fact that those modal solutions span the space, so that every solution of the equation can be expressed as a linear combination of such modes. Consequently, the concept of modes can only be applied to the linear problem. For the special case of degenerate

solutions, the corresponding fields may not, in general, be linearly independent. However, any linear combination from the corresponding subspace of the degenerate fields may be chosen to fulfill the linear independency condition to be modes of the system.

A numerical solution of Eq. 1 requires the discretization of the variables $\mathbf{r}$, together with the functions $u(\mathbf{r})$ and $V(\mathbf{r}, u)$. This procedure results in N algebraic equations at the discrete points, whose solution is the function of interest. Indeed, in the discussion that follows, we shall be interested primarily in the nontrivial realizations of Eq. 1 in two-dimensional, although higher dimensional problems are also dealt in a similar fashion.

In what follows, our attention shall focus upon the optical modal problem. Let the transverse spatial variables be x and y, with discretization given by x_j, where $1 \leq j \leq n$, and y_i, where $1 \leq i \leq m$, respectively. Thus, at each point of the domain (x_j, y_i) we consider samples for the solution $u_{ij} = u(x_j, y_i)$ and for the potential function $V_{i,j} = V(x_j, y_j; u_{ji})$. Substituting these defined variables into Eq. 1, with $\nabla^2 \equiv \nabla_\perp^2 \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2$, and finite-differencing these derivative terms, we obtain,

$$\begin{aligned}[-2(1+\gamma^2) + (\Delta x)^2 V_{ij}]u_{ij} + u_{ij-1} + u_{ij+1} \\ +\gamma^2(u_{i-1j} + u_{i+1j}) = (\Delta x)^2 \beta u_{ij},\end{aligned} \tag{2}$$

where $\gamma = \Delta x/\Delta y$, and the constant discretization intervals are given by Δx and Δy. More generally, a non-uniform grid spacing can also be considered when required, if we make the intervals in Eq. 2 explicitly dependent upon the grid points as $(\Delta x)_j = x_j - x_{j-1}$, $(\Delta y)_i = y_i - y_{i-1}$, and $\gamma_{ij} = (\Delta x)_j/(\Delta y)_i$.

With this substitution, the finite difference equation, Eq. 2, constitutes a system of $N = n \times m$ algebraic equations with N unknowns u_{ij}. Together with these equations, it is necessary to impose the required boundary conditions depending on the nature of the particular problem that is being solved. The most usual cases include homogeneous (Dirichlet), Neumann and periodic conditions that can be easily dealt with minor changes in the system (Meurant 1999), conveniently modifying those equations involving boundary samples to force the required behavior.

In order to make the notation more compact, the samples of the solution function as well as those corresponding to the potential function, which constitute $n \times m$ matrices, are rearranged by concatenating the columns to produce large column vectors of N rows, so that $\mathbf{u} = (u_k)$ and $\mathbf{V} = (V_k)$, $k = 1, 2, \ldots N$. Therefore, the algebraic system may be written as $\mathbf{Au} = \beta\mathbf{u}$, where $\mathbf{A}$ is the $N \times N$ matrix which may depend on the unknown vector through the dependence of the potential function V and may be consequently nonlinear. Thus, we denote the matrix as $\mathbf{A}[\mathbf{u}]$ to highlight this dependence, and the system of equations takes the form,

$$\mathbf{A}[\mathbf{u}]\mathbf{u} = \beta\mathbf{u}, \tag{3}$$

so that the system equations are written as,

$$E_k = \sum_{j=1}^{N} (\mathbf{A}[\mathbf{u}])_{kj} u_j - \beta u_k = 0, \qquad k = 1, 2, \ldots, N. \tag{4}$$

The matrix $\mathbf{A}$, as described, is in general very large and sparse, differing from zero along the main diagonal, the two adjacent diagonals, and also the diagonals at a distance n from the main diagonal. These four diagonals appear due to the coupling terms in the derivatives of the Laplace operator. Given the simple particular case we are considering at first, the matrix $\mathbf{A}$ is also symmetric. More complicated matrices arise for cases little different from the one

under consideration, and will be the subject of discussion in Sect. 3. The equations can be nonlinear, due to a potential V which depends upon the function u. For such cases, suitable iterative methods such as Newton are required. Nonetheless, if the potential does not depend on the function u, the homogeneous system (3) constitutes a linear eigenvalue problem which also requires an appropriate iterative method to be solved.

2.1 Linear stationary problems. Inverse iteration

As described, when the potential in Eq. 1 does not depend on the function $u(\mathbf{r})$, then Eq. 3 is a linear eigenvalue problem $\mathbf{Au} = \beta\mathbf{u}$ and has solutions for a set of values of β (eigenvalues), which constitute the spectrum of the operator $\mathbf{A}$. The solution of this problem will provide those eigenvalues together with the corresponding eigenvectors $\mathbf{u}$.

A number of well known methods exist to numerically obtain the eigensolutions of a matrix (Golub and Van Loan 1989; Demmel 1997). However, due to the large problem size, not all the methods are computationally suitable. A particularly successful method for large problems is the *inverse iteration* method, first proposed by Wielandt and further developed to become viable by Wilkinson (Wilkinson 1965; Ipsen 1997), in which each eigensolution is generated one at a time, starting from an initial approximation to the eigenvalue. Two characteristics of the types of problems we are considering make this method particularly useful. First, we are normally interested in the discrete part of the spectrum—the bounded states in quantum systems or guided modes in wave optics—and mostly in a restricted set of states, the fundamental or, at most, the few with smallest or largest propagation constant. Secondly, the corresponding propagation constants or eigenvalues are located close to the extreme of the potential function—the minimum for quantum systems, or maximum for wave optics—and finding a suitable initial guess for the eigenvalue is easy by means of a hand-adjustment. As an added benefit, the starting values may be available, in some cases, from approximate methods.

The inverse iteration algorithm at the step n may be summarized as:

$$(\mathbf{A} - p\mathbf{I})\mathbf{w}^{(n)} = \mathbf{u}^{(n-1)}, \tag{5}$$

$$c^{(n)} = (\langle w|w\rangle)^{1/2}, \tag{6}$$

$$\mathbf{u}^{(n)} = \mathbf{w}^{(n)}/c^{(n)}. \tag{7}$$

The key step of the algorithm is solving the linear system of Equations (Eq. 5). The vector norm $c^{(n)}$ is introduced in order to rescale the vector at each iteration, thus avoiding potential computational overflow. Furthermore, the algorithm requires the adjustment of the parameter p and converges to the eigenvector whose associated eigenvalue is closest to that parameter. Hence, the algorithm iteratively constructs one eigensolution at a time and has been considered the most successful method to calculate eigenvectors given accurate eigenvalues (Wilkinson 1965). When only an approximation to the eigenvalues is known, as is generally the case for the problems we are considering, the method is still successful and converges in only few iterations. Once the algorithm has converged, an accurate eigenvalue comes given by the so-called Rayleigh quotient:

$$\beta^{(n)} = \frac{\langle u^{(n)}|A|u^{(n)}\rangle}{\langle u^{(n)}|u^{(n)}\rangle}, \tag{8}$$

where the denominator is the unity since the field $\mathbf{u}$ was normalized. Equation 8 also gives an approximation for the eigenvalue at each iteration which can be used to update the value

of p, by setting $p^{(n)} = \beta^{(n-1)}$, having in this case a variant of the algorithm named Rayleigh quotient iteration. This often supposes an acceleration in the convergence, but if the method for solving the linear system takes advantage of computations performed only once for all iterations—for instance, matrix decompositions—the calculation of the estimate at each iteration induces a computational overload, which can be alleviated through updates only after a certain number of iterations. Using the Rayleigh iteration step the convergence is cubic compared to a quadratic convergence for no updates. However, the method requires special care, since it may fail to converge if the initial parameter p is not close enough to the eigenvalue (Szyld 1988).

The choice of the initial starting vector is not very important, except if the matrix is non-normal—i.e., if its set of eigenvectors does not constitute an orthogonal basis of the space—for which a careful choice is required. For the kind of problems we are treating, however, the matrix is normal and therefore, in practice, a randomly chosen starting vector should work sufficiently well. Nonetheless, it has been argued that a vector with ones in all the components is particularly a very good choice for almost all the cases (Ipsen 1997). An interesting case occurs for degenerate states, where the algorithm initially converges to one of them, while the remaining ones can be calculated by a subsequent iteration with a starting vector orthogonal to those previously calculated.

The solution of the linear system of equations in Eq. 5 is the key step in the algorithm, specially when considering a very large N matrix. It is here where we can exploit the power of nonstationary methods to obtain an efficient solution, which would otherwise be outside the scope of traditional methods. Thus, the complete algorithm is doubly iterative, being a combination of outer and inner iterations.

2.2 Nonlinear stationary problems

We now turn our attention to numerical methods for the nonlinear problems, which corresponds to the case in which the potential of Eq. 1 depends explicitly upon the field $u(\mathbf{r})$. This problem differs in a fundamental way from the linear eigenvalue problem described in the previous section, since β acts as a parameter describing a family of solutions. Therefore, the computational strategy changes and consists in fixing a value for β and then solving for the function $u(\mathbf{r})$.

For this case, the algebraic problem (3) consists of a set of nonlinear equations which must be solved by proper methods, in order to obtain the values of the field $\mathbf{u}$ at the grid points. One standard and successful method is the globally convergent Newton method (Dennis and Schnabel 1983; Press et al. 1992), which builds the solution iteratively from an initial guess $\mathbf{u}^{(0)}$ so that, at the iteration l, the function is updated by the expression $\mathbf{u}^{(l)} = \mathbf{u}^{(l-1)} + \delta\mathbf{u}$. The calculation of the so-called Newton step $\delta\mathbf{u}$ is carried out, at each iteration, by performing a first order Taylor expansion of the equations (4),

$$\mathbf{E}(\mathbf{u}^{(l-1)} + \delta\mathbf{u}) = \mathbf{E}(\mathbf{u}^{(l-1)}) + \sum_{j=1}^{N} \frac{\partial \mathbf{E}}{\partial u_j} \delta u_j + \cdots , \tag{9}$$

where $\partial E_k / \partial u_j = (\mathbf{J})_{kj}$ is identified as the Jacobian matrix. Thus, we may express the system of equations as $\mathbf{E}(\mathbf{u}^{(l)}) = 0$, and the expression (9) yields,

$$\mathbf{J} \cdot \delta\mathbf{u} = -\mathbf{E}, \tag{10}$$

which is a linear system of equations whose matrix is the Jacobian matrix. The solution of this system provides the correction $\delta\mathbf{u}$ to update the function at each Newton step. The Jacobian matrix, as well as the right side of (10), are evaluated replacing the last known iterate ($\mathbf{u}^{(l-1)}$) into the equations of the system (Eq. 4). The Jacobian matrix possesses the same sparse structure as the system matrix $\mathbf{A}$ and usually their elements can be calculated analytically.

The Newton method is quadratically convergent if the initial guess is *sufficiently* close to the solution. In practice, the condition of *sufficiently* close is not met often and therefore additional techniques are required to improve the method and make it globally convergent. A classical technique consists in defining a function $f = (1/2)\mathbf{E} \cdot \mathbf{E}$ and rescaling the Newton step if necessary, so as to make the value of the function f decrease, i. e. the solution, which fulfils $f = 0$, is reached following the so-called *descent directions*, yet avoiding a big step that may take the iterate too far away (for further details see (Dennis and Schnabel 1983; Press et al. 1992). The principal problem with this method, however, is the possibility for the algorithm to get into a local minimum of the function f. In practice, however, this happens rarely if the qualitative shape of the solution is known and the initial guess is built resembling such a shape. We shall see in the problems that we describe in Sect. 3, that convergence may be reached with this global strategy by simply a judicious choice of the initial functional shape. It should be mentioned that improvements of this technique for obtaining global convergence are available, including the use of information obtained from the algorithm which solves the linear system to obtain the Newton step (Bellavia and Morini 2001).

Again, due to a large size of the matrix $\mathbf{J}$, the system (10) can only be solved iteratively. Therefore, the power of the nonstationary iterative methods are determinant to efficiently get the solution, and the complete algorithm is again composed of a set of outer and inner iterations.

2.3 Modern iterative linear solvers

For both linear and nonlinear kernels, the problems are reduced to efficiently solving linear equation systems with very large sparse matrices. There exist a number of nonstationary iterative methods, which may be efficient for obtaining the solution for each step of the inverse iteration or Newton methods. The choice of a particular nonstationary method is highly dependent upon the characteristics of a particular system matrix. Given an adequate nonstationary method for a given system matrix, these methods offer sufficient computational efficiency so as to support the requirements of a double (inner-outer) iterative process for reasonably sized problems.

The general idea for most all such methods is the construction of a set of orthogonal vectors (one per iteration) which progressively increases the dimension of the space where the solution lies. These vectors are initially generated by means of a Krylov sequence and then orthogonalised by a Gram-Schmidt, or a related technique (Greenbaum 1997; Meurant 1999). These nonstationary methods give rise to a practical convergence in a number of iterations much smaller than N, and the form of the matrix or the way it is stored or accessed in memory is unimportant for the algorithm, since it only needs to refer the matrix through its product times a vector. Each of these characteristics make these methods a natural choice for solution of large N problems.

The first decision to chose a nonstationary method is based upon properties of the system matrix, since not all the methods work with a general matrix. The most important properties to this concern are symmetry and the definite or indefinite character. Not only is a problem

possessing a symmetric matrix much easier to solve, it is also often easier when the matrix is definite, i. e. when all the eigenvalues have the same sign. Thus, for systems with symmetric and definite matrix, the Conjugate Gradient (CG) is an efficient method and is often a very good choice. For cases where the matrix is symmetric but indefinite, methods such as the Symmetric LQ (SYMMLQ) or Minimum Residual (MINRES) (Paige and Saunders 1975) are both suitable choices, although for some applications, especially when the matrix is close to singularity, as often happens for the inverse iteration, the MINRES method has been considered to be slightly better (Dul 1998). Nonetheless, from our experience, we have also determined that SYMMLQ works sufficiently well even for these problems.

On the other hand, a host of methods exists for general indefinite and non symmetric matrices. Amongst them are methods such as: Chebyshev Iteration, Conjugate Gradients on Normal Equations (CGNE), Generalized Minimal Residuals (GMRES), BiConjugate Gradient (BiCG), Biconjugate Gradient Stabilized (Bi-CGSTAB), Quasi-Minimal Residual (QMR) and Conjugate Gradient Squared (CGS), and there are many others. With each of the methods mentioned, there are advantages and drawbacks that must be considered when deciding whether to use a particular method over another. A proper selection of a method is very often only possible after experimentation. In fact, for most of the methods it is possible to find examples where a particular method is the best and other examples where it is the worst (Nachtigal et al. 1992). The BiCG method may have convergence problems, requires the transpose matrix A^T along with A, and it is not useful when the matrix is close to the singularity. QMR (Freund and Nachtigal 1991), CGS (Sonneveld 1989), and Bi-CGSTAB (van der Vorst 1992) are refined versions of BiCG method, each one improving some of the cited drawbacks, but none improve all of them at a time or in a full satisfactory way (Barrett et al. 1994). Generally, GMRES, developed by Saad and Schultz (Saad 1986), is a method which works fine in most situations and is considered one of the most robust algorithms. We have selected this algorithm for solving the non-symmetric example given in Sect. 3.5.

The most important drawback with GMRES is that it needs to store the sequence of orthogonal vectors which are progressively generated iteration after iteration. This means that after certain number of iterations the storage requirements are considerable and computational efficiency is severely affected, so that a restarting operation must be undertaken. This consists in re-initiating the iterative process, but with the initial guess chosen to be the last approximate solution obtained from the previous set of iterations. The number of iterations before restarting, turns out to be a critical value both for convergence and computational efficiency. If it is too low, the algorithm may need a high number of iterations to converge or may fail to converge at all. If it is too large, iterations are slower and memory resources are unnecessary spoiled. Preliminary experimentation usually allows to easily fix this parameter for a particular kind of problems.

A further improvement which can be made is use of a *preconditioner*, which is a matrix $\mathbf{M}$ that approximates the inverse of the original matrix $\mathbf{A}$ and for which the system $\mathbf{M} \cdot \mathbf{x} = \mathbf{b}$ is easy to solve. Thus, the preconditioned system $\mathbf{M}^{-1} \cdot \mathbf{A} = \mathbf{M}^{-1}\mathbf{b}$ has a matrix $\mathbf{M}^{-1} \cdot \mathbf{A}$ with more favorable spectral properties (lower condition number) and is solved with less iterations. The use of a well-suited preconditioner may lead to a great increase in the performance of the iterative methods. Several different preconditioners have been described in the literature (see Refs. Barrett et al. 1994; Greenbaum 1997; Meurant 1999). Some are more suitable for general classes of matrices, such as those based upon stationary methods—Jacobi, block Jacobi, Gauss-Seidel, SOR, Symmetric SOR (SSOR) among others—or those based upon incomplete decompositions of the system matrix (in order to maintain a reasonable sparsity), or also upon multigrid methods. Moreover, preconditioners have been developed for specific problems or

problem-types, for example the multigrid techniques to use in a combined Newton-GMRES algorithm (Knoll and Rider 1999). For the numerical problems we present in Sect. 3, the use of a preconditioner was not essential, except for the case of the non-symmetric problem, and they can be managed with a reasonable efficiency without preconditioning. Nevertheless, the use of even the simpler versions of a preconditioner leads to a substantial improvement in the convergence rate.

3 Stationary state calculation in photonic structures

In this section, we present several examples of physical problems taken from wave optics, that we successfully solved by means of the techniques described in the previous sections. As previously remarked, although the problems we solve are from wave optics, these methods can readily be applied to a great number of problems in physics and engineering. Our selection allows us to illustrate several issues related to different forms of the matrices and the methods required to solve. In particular, through the following subsections, we treat the modes of a linear rectangular waveguide, the nonlinear stationary states of a dual-core directional coupler made of a photonic crystal fiber (PCF), a nonlinear single-core PCF but considering a two-component vector system, the vortex states in such a PCF, and a non-symmetric problem consisting in a spatio-temporal soliton system.

3.1 Linear symmetric systems: channel waveguides

The first problem we are going to consider corresponds to a simple linear two-dimensional rectangular waveguide, which is a particular example of a channel waveguide (Tamir 1990). For this case, the potential function in Eq. 1 is defined as,

$$V(x, y) = \begin{cases} n; & |x| \leq a, \quad |y| \leq b, \\ n_s; & \text{otherwise}, \end{cases} \tag{11}$$

where n and n_s are the refractive indices inside the core and in the substrate, while a and b are the dimensions of the core. The condition $n > n_s$ is required in order for the waveguide to be able to confine light and support a set of bounded states. Formally, this means that the spectrum of the operator which constitutes the kernel of Eq. 1 possesses a discrete set of eigenvalues.

As described in Sect. 2, the application of a finite different scheme generates a set of linear equations of the form of Eq. 2. After a re-ordering operation, which transforms variables and functions on the 2-dimensional grid into large column vectors $u_{ij} \to \mathbf{u}_k$ and $V_{ij} \to V_k$, with $k = 1, \ldots, N$, with $N = n \times m$, we obtain the difference equation system,

$$\begin{aligned} &u_{k+1} + u_{k-1} - 2u_k + \gamma^2(u_{k+n} + u_{k-n} - 2u_k) + \\ &+(\Delta x)^2 V_k u_k = (\Delta x)^2 \beta u_k, \qquad k = 1, \ldots, N, \end{aligned} \tag{12}$$

which can be written in the form of Eq. 3, where the matrix $\mathbf{A}$ does not depend on the field $\mathbf{u}$ and takes the form,

$$\mathbf{A} = \left(\begin{array}{cccccc} a_k & b_k & c_k & & & \\ b_k & a_k & b_k & c_k & & \\ c_k & b_k & a_k & b_k & c_k & \\ & c_k & b_k & a_k & b_k & c_k \\ & & c_k & b_k & a_k & b_k \\ & & & c_k & b_k & a_k \end{array} \right) \tag{13}$$

where,

$$a_k = -2(1+\gamma^2) + (\Delta x)^2 V_k, \qquad b_k = 1, \quad c_k = \gamma^2. \tag{14}$$

This matrix is symmetric and represents the well-known tridiagonal form with two fringes separated n positions from the main diagonal. The storage requirements of the matrix are reduced to those of the main diagonal, due to the fixed constant value for the off-diagonal elements.

Due to the symmetric nature of the matrix $\mathbf{A}$, and the fact that it is indefinite, the SYMMLQ or MINRES are excellent candidate methods, in combination with the inverse iteration (Eq. 5), to solve the linear algebraic system at each iteration. Note, that the matrix $\mathbf{A} - p\mathbf{I}$ would only be definite if parameter p were chosen larger than the highest eigenvalue. With such a choice it would be possible to use the CG method, but the method would be restricted to the computation of the fundamental mode.

Examples of stationary states calculated with this technique, using a grid size of 300×300 points (N=90000), are shown in Fig. 1. In all cases, we started with an initial vector formed with ones in all the components. We have found that, as expected, the starting vector is not very important in obtaining the solution. This is in contrast with other classical methods used for obtaining these solutions, such as the famous Iteration Eigenfunction Expansion Method (Jablonski and Sowinski 1989), which requires an initial guess that approximately resembles the mode shape.

For obtaining the fundamental state, corresponding to the largest eigenvalue, the obvious choice for the parameter p, given that no other approximation is available, is the maximum of the refractive index. For obtaining initial guesses for all other states, p may be obtained by testing by hand a small set of parameters, if no other approximation is available. This methodology is usually efficient in practice, since the number of guided modes (discrete spectrum) is small and convergence is very fast.

3.2 Nonlinear photonic crystal fibers

The second problem we consider is to calculate the optical modes of a *dual-core* Photonic Crystal Fiber (PCF) coupler (Russell 2003; Knight 2003), which consists of a Kerr-type

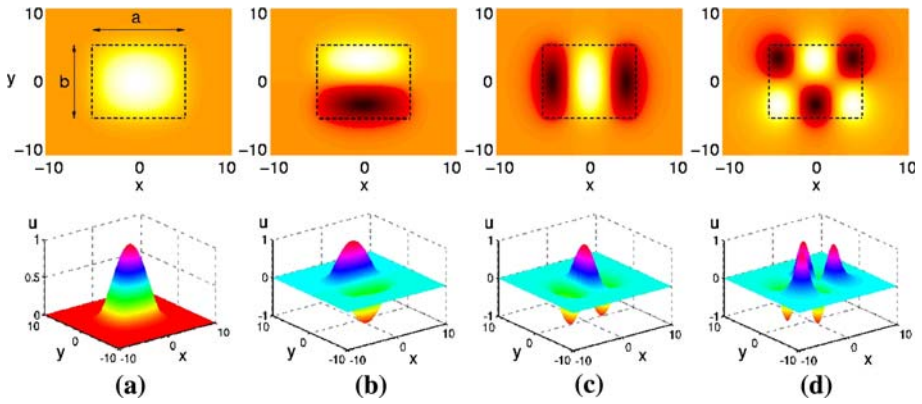

Fig. 1 Some of the different stationary states (modes) of a linear rectangular waveguide with a core of refractive index $n = 3$ in a substrate of $n_s = 1$. Core dimensions are $a = 8$, $b = 6$. Propagation constants for the states $\beta_{\nu\mu}$, where μ and ν indicate the number of nodes in each direction, are (**a**) $\beta_{00} = 2.7126$ (fundamental mode), (**b**) $\beta_{01} = 2.1961$, (**c**) $\beta_{20} = 1.8623$ and (**d**) $\beta_{21} = 1.3702$

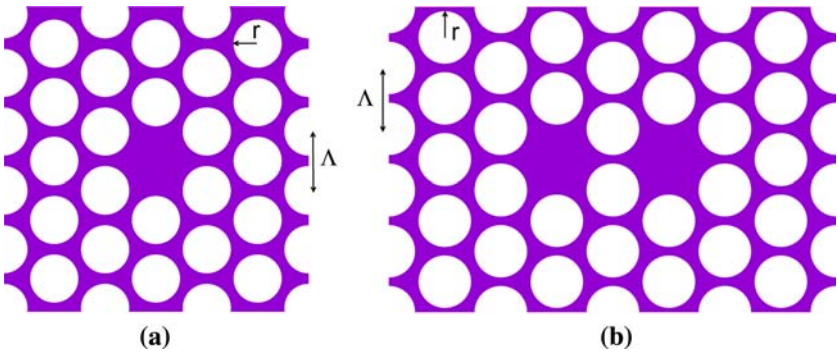

Fig. 2 Sketch of a photonic crystal fiber (PCF) structure (transverse plane) with a triangular lattice geometry. (**a**) Single-core PCF; (**b**) dual-core PCF coupler. Λ and r are respectively the pitch (distance between holes) and hole radius, which characterize the structure

nonlinear material giving rise to a potential that is proportional to the intensity (squared modulus) of the field. The PCF is a cylindrical structure, similar to a conventional optical fiber, consisting of an array of air holes running parallel to the fiber axis and with a central defect (lack of the central hole), which acts as the core. Figure 2 shows the cross-sectional geometry of such devices for the particular case of a triangular lattice, illustrating both a single (Fig. 2a) and a double (Fig. 2b) core structures.

Several techniques (Ferrando et al. 2003) have been reported in the literature for calculating the different types of stationary states that can exist in PCF structures, which normally take advantage of the symmetries of the single core device. A particularly simple and efficient solution can be obtained using the globally convergent Newton method described in Sect. 2.2, which is also suitable for less symmetric problems, as is the case with the dual-core PCF (Salgueiro and Kivshar 2005).

In order to obtain the stationary solutions for the nonlinear dual-core PCF, we shall assume that the nonlinear effect is present in the substrate material but not in the holes (made usually of air). Therefore, the nonlinear potential term may be described by the expression,

$$V(x, y) = n_a + f(x, y)(\delta + u^2), \tag{15}$$

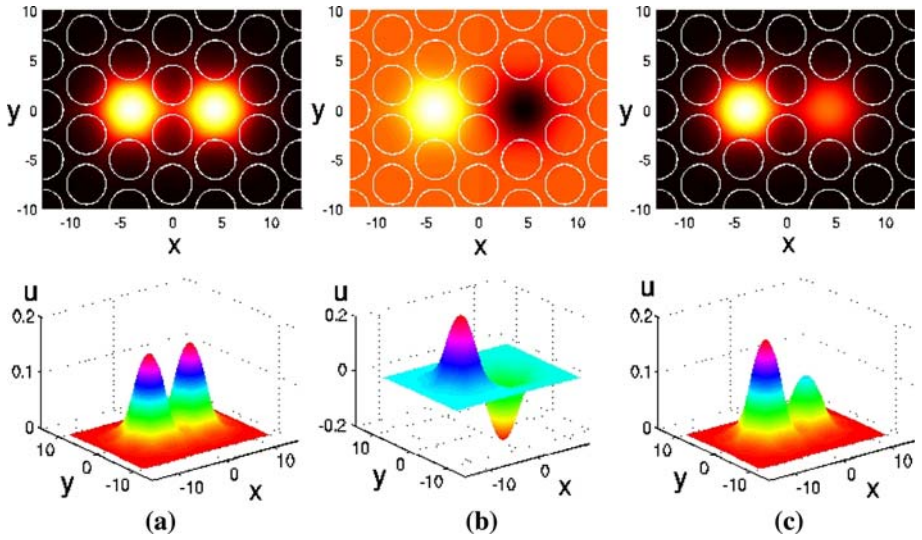

Fig. 3 Two dimensional (top row) and three-dimensional (bottom) plots of different stationary states of a nonlinear dual-core photonic crystal fiber coupler. (**a**) Symmetric, (**b**) antisymmetric, (**c**) asymmetric. Parameters were chosen as $\Lambda = 5$ (pitch), $r = 2$ (hole radius), $n_s = 1$, $n_a = 0$ and $\beta = 0.74$

where $\delta = n_s - n_a$, being n_s and n_a respectively the refractive index in the substrate and in the holes. The function $f(x, y)$ describes the hole lattice with the two defects, and assigns $f = 1$ in the substrate and $f = 0$ in the holes.

After performing the discretisation and our standard reordering of variables and functions into column vectors, a nonlinear system of equations of the form Eq. 3 is obtained. The system matrix $\mathbf{A}[\mathbf{u}]$ takes on the same form as obtained previously by Eqs. (13) and (14), except that now terms along the main diagonal which correspond to the nonlinear potential given by Eq. 15 must be considered.

Since the potential, and consequently the system of equations given by (4), are known in a simple analytical form, the partial derivatives of those equations with respect to the unknown variables (u_k) are easy to calculate. Hence, the Jacobian matrix, given by $J_{jk} = \partial E_j / \partial u_k$, needed for the Newton steps, has the same structure as the system matrix. The only difference is the main diagonal which may be written as,

$$J_{kk} = \frac{\partial E_k}{\partial u_k} = -2(1 + \gamma^2) + (\Delta x)^2 \left[n_a + f_k(\delta + 3u_k^2) - \beta \right], \tag{16}$$

together with the remaining off-diagonal elements, $J_{k\,k-1} = J_{k\,k+1} = 1$ and $J_{k\,k-n} = J_{k\,k+n} = \gamma^2$, which are unchanged. On the other hand, since homogeneous boundary conditions are considered on all four edges of the domain, zeros appear on the two diagonals adjacent to the main one at points located at positions $n, 2n, 3n, \ldots, N$.

Again, due to the symmetry of the problem and the fact that the definiteness of the Jacobian matrix cannot be guaranteed, a solution using the SYMMLQ or MINRES is possible. Figure 3 shows the three lowest order states of such a PCF device, which correspond to the symmetric, antisymmetric and asymmetric states. We remark that the asymmetric field only exists in nonlinear devices if the value of parameter β is high enough. The fields of Fig. 3 were calculated using the globally convergent Newton method combined with the SYMMLQ, at each Newton step, on a grid of 300×300 points ($N = 90000$). The initial guess consisted of

two superposed Gaussian functions, each one centered at one of the cores of the coupler. Both Gaussians had the same amplitude for the symmetric state and inverted-sign amplitudes for the antisymmetric state. For the asymmetric state, one of the amplitudes was chosen to be a fraction of the other. Owing to the globally convergence strategy, a simple hand-adjustment of the peak amplitudes and widths of the starting Gaussians was enough to easily achieve convergence in few steps.

3.3 Multicomponent systems: vector solitons in photonic crystal fibers

The next problem we consider is a variation on the previous, yet contributes an additional numerical complexity in order to further illustrating the methods. Thus, instead of a single propagating beam, we now consider the coupling interaction of two co-propagating beams within a nonlinear single-core PCF (Fig. 2a). Of particular interest is the incoherent interaction of two fields, achieved with either beams of different frequency or orthogonal polarization. Given the two field components u_1 and u_2, the two-beam system or *vector soliton* can be described by the model,

$$\begin{aligned} \nabla^2 u_1 + \left[n_a + f(x,y)(\delta + u_1^2 + \mu u_2^2) - \beta_1\right] u_1 = 0, \\ \nabla^2 u_2 + \left[n_a + f(x,y)(\delta + u_2^2 + \mu u_1^2) - \beta_2\right] u_2 = 0, \end{aligned} \tag{17}$$

where β_1 and β_2 are the propagation constants for u_1 and u_2, respectively. As is easily verified, this model is composed of two equations of the form of Eq. 1, one for each component, but with an additional nonlinear cross-term, which couples both field components (Salgueiro et al. 2005), with coupling constant μ. As before, $\delta = n_s - n_a$ and $f(x, y)$ describes the PCF.

As before, we express the finite difference equations through the reordering of variables and functions in large column vectors, so that the two field components are now written $u_{1,k}$ and $u_{2,k}$. Additionally, both of the field vectors are further concatenated into another column vector with $2N$ components: $\mathbf{q} = (\mathbf{u_1}^T | \mathbf{u_2}^T)^T$, so that we may write the algebraic nonlinear system in an analogous form as Eq. 3, $\mathbf{A[q]q} = 0$. Now $\mathbf{A}$ is a $2N \times 2N$ matrix with two diagonal blocks, each one with the same structure as (13), and with two additional diagonals, separated N elements from the main one, due to the coupling between the components. The Jacobian matrix, necessary to carry out the Newton iteration process, can be calculated analytically in a similar fashion as in the previous case and also has the same symmetric structure of the matrix $\mathbf{A}$,

$$\mathbf{J} = \frac{\partial \mathbf{E}}{\partial q_k} = \begin{pmatrix} & & & \\ a_k & c_k & & d_k \\ & b_k & & \\ b_k & & & \\ & c_k & & \\ & & & \\ & & & c_k \\ d_k & & a_k & b_k \\ & & b_k & \\ & & c_k & \end{pmatrix} \tag{18}$$

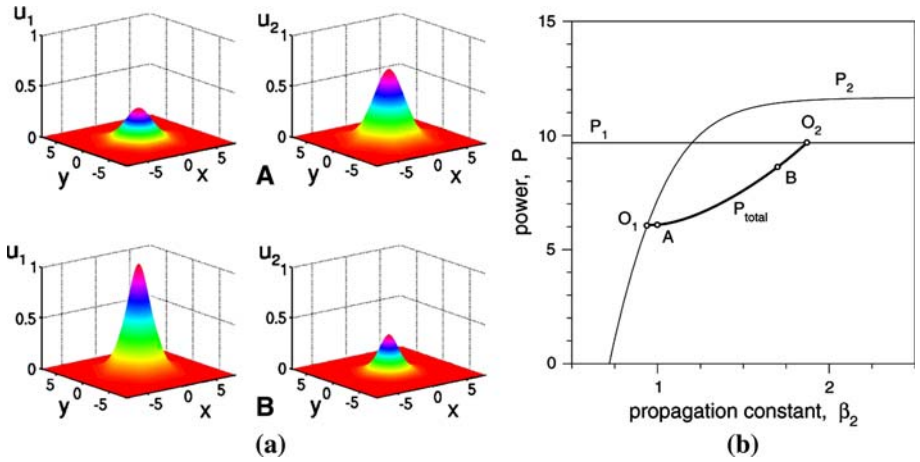

Fig. 4 Nonlinear vector solitons for the single-core PCF (same physical parameters of the previous example and $\beta_1 = 3$, $\mu = 2$). (**a**) The two components of two different stationary states corresponding to $\beta_2 = 2.85$ (label A) and $\beta_2 = 3.35$ (label B). (**b**) Power curve (thick line) of the family of vector solitons obtained varying the parameter β_2. This line starts and ends respectively at the bifurcation points O_1 and O_2. Thin lines are the power curves for the scalar solitons (one of the components zero). Points A and B correspond to the examples showed in subfigure (**a**)

where,

$$a_k = -2(1+\gamma^2) + (\Delta x)^2 n_a + \\ + (\Delta x)^2 \begin{cases} f_k(\delta + 3q_k^2 + \mu q_{k+N}^2) - \beta_1; & 1 \le k \le N \\ f_{k-N}(\delta + 3q_k^2 + \mu q_{k-N}^2) - \beta_2; & N+1 \le k \le 2N, \end{cases} \tag{19}$$

$$b_k = 1 \tag{20}$$

$$c_k = \gamma^2 \tag{21}$$

$$d_k = 2\mu(\Delta x)^2 f_k q_k q_{k+N}. \tag{22}$$

Due to the imposed homogeneous conditions, there exist zeros at points separated by n positions along the diagonals adjacent to the main one.

We now turn our attention to methods for solving this case. Once again, the symmetry of the problem allows us to use a solver such as SYMMLQ to get the solution at each Newton step. In Fig. 4(a) we show two examples obtained for a single core PCF with the same physical characteristics and using the same grid size as treated previously, yet now for two co-propagating beams. Each example represents a solution for different values of parameter β_2, where $\beta_1 = 3$ and $\mu = 2$ are held constant. In this case it is necessary to storage the main diagonal elements along with those belonging to the diagonal coming from the coupling term.

An interesting point, which we wish to highlight with this example, is the possibility of obtaining a family of solutions depending upon only small changes in the values of the parameters. In fact, in this case a solution, just obtained, may act as an excellent initial guess for obtaining the next, and leading to a rapidly convergent Newton method. Thus, families of solutions are easily calculated, such as those illustrated in Fig. 4(b), obtained by varying the parameter β_2. In the Fig. 4, the total power of the field, defined as $P = \int (u_1^2 + u_2^2)\, dxdy$, is plotted and evaluated in each case by numerical integration. The examples illustrated in

Fig. 4(a) correspond to the marked points A and B on the thick line of the curve that describes the family of vector solitons. This family starts and ends at points O_1 and O_2, where one of the components vanishes, thus converting into a scalar soliton. The families of scalar solitons are described by the thin lines. That is why those points are named *bifurcation points*.

3.4 Complex fields: vortex states in PCFs

An analogous case with a two-component field is obtained when we intend to calculate vortex states of a non-homogeneous structure, i. e. fields which present a phase dislocation and whose wave envelope must be described by a complex field (Ferrando et al. 2004). Formally, a vortex is a field defined by a function of the form $\psi(r, \phi) = v(r, \phi) \exp(i\ell\phi)$, being $r = (x^2+y^2)^{1/2}$, $\phi = \arctan(y/x)$ the cilyndrical coordinates and ℓ is an integer describing the number of the phase windings, referred to as the *vorticity*, or in other contexts as the *topological charge*. The function $v(r, \phi)$ may be real and ϕ-independent if the system has cylindrical symmetry, however, in such a case the problem is reduced to one-dimension and consequently easier to solve. Nevertheless, if the problem does not present full cylindrical symmetry, as in the case of the PCF lattice considered in the previous sections, then the function $v(r, \phi)$ is complex, contributing to the phase in a way dependent on the medium geometry.

Using Cartesian coordinates to describe the field $u(x, y) = u_1(x, y)+i\, u_2(x, y)$, we carry out a substitution into our model equation (1) and, after considering the real and imaginary terms separately, we obtain a two equation system describing both components u_1 and u_2. If the problem is linear, both equations are decoupled and actually the same equation,

$$\nabla^2 u + [n_a + \delta f(x, y)]\, u = \beta u, \tag{23}$$

where u is each one of the components u_1 or u_2. In this case, the problem is reduced to obtaining the degenerate states of a one component system, and then to combine them into the real and imaginary parts in order to form different types of vortex-like solutions. If nonlinear terms exist, both equations will be coupled and of the form (17), with $\mu = 1$ and $\beta_1 = \beta_2 = \beta$ (if the nonlinear term is of the Kerr type). After discretisation, a nonlinear algebraic problem is obtained, whose Jacobian matrix has a form of the type illustrated in Eqs. (18–22).

In Fig. 5, two examples of vortices in PCF structures are shown. The top row corresponds to the linear case, for $\ell = 1$, and was obtained by inverse iteration combined with the SYMMLQ solver. In this case there are two degenerate dipole fields (as shown in Fig. 5a–b), which describe the system in a real azimuthal basis which gives rise to the vortex (Fig. 5c–d) after forming the real and imaginary part of the field from those dipoles. For producing the degenerate states, the algorithm converged initially to one of the dipoles and then, starting with an orthogonal vector obtained by the Gram-Schmidt technique, a further inverse iteration process produced the other dipole. In this case, due to the degeneracy, starting with an initial vector that vaguely resembles one of the dipoles, for instance a function with an azimuthal node such as $u^{(0)}(r, \phi) = \exp(-r^2/w^2) \sin(\ell\phi)$, where w is an uncritical constant, helps to achieve a much faster convergence. Besides, in this case, it is easy to obtain a proper parameter p just calculating the Rayleigh quotient (Eq. 8) from this initial vector.

Shown in bottom row of Fig. 5, a vortex with $\ell = 2$ is given for the PFC structure. The quadrupole fields of Fig. 5(e)-(f) constitute the real and imaginary parts of the vortex field (Fig. 5g–h), where a Kerr-type nonlinearity has been considered with the form given by the potential in Eq. 15. The nonlinear problem was solved for a grid size of 300×300 using the Newton method starting with an initial guess chosen as a Gaussian multiplied times a first order monomial and with an azimuthal phase factor, i.e., $u_0 = Ar^\ell \exp(-r^2/w^2) \exp(i\ell\phi)$,

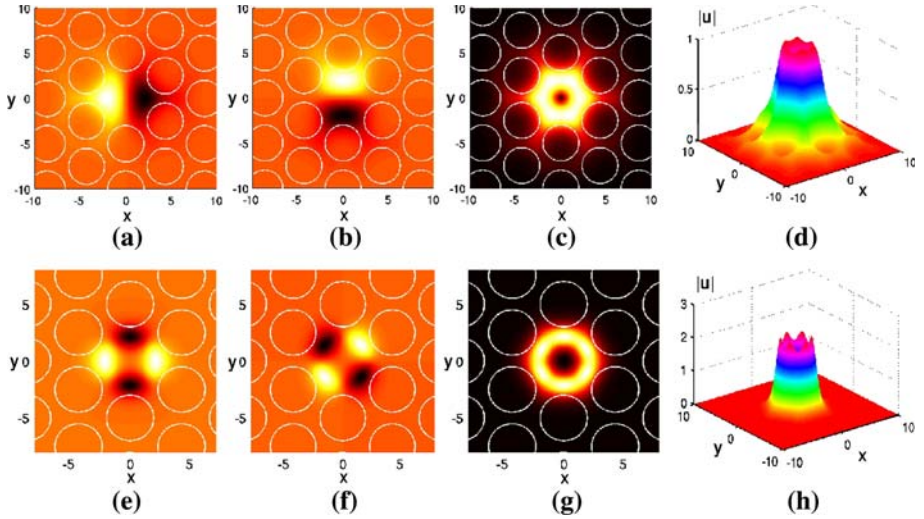

Fig. 5 Vortex states in a photonic crystal fiber. Top row: linear vortex for $\ell = 1$, (**a**–**b**) dipole field, (**c**–**d**) vortex field. The calculated propagation constant is $\beta = 1.18$. Bottom row: nonlinear vortex for $\ell = 2$ and $\beta = 3$, (**e**–**f**) quadrupole fields, (**g**–**h**) vortex field

where A and w are constants to adjust. This function is chosen to resemble the form of the amplitude and phase of the solution. A simple hand-adjustment of the parameters A and w leads to an easy convergence of the Newton method.

Instead of solving for both separate components, as we have done in this nonlinear problem, an alternative and often preferred method is to directly implement the algorithm using complex arithmetic. Therefore, the full complex field is obtained without the need of splitting it in its real and imaginary parts, and the complex algebra is handled directly by using a complex data type. The SYMMLQ algorithm can be implemented directly in complex arithmetic and is an eligible technique when the system matrix is hermitian. In our case the matrix is not hermitian, since it is still symmetric, and this fact would require the use of a more general—and often less efficient—method such as GMRES. However, the SYMMLQ or MINRES algorithms are suitable for symmetric complex matrices if they are implemented defining the inner vector product as $\langle v|w\rangle = \sum_k v_k w_k$, instead of the usual definition, with the first vector conjugated $\langle v|w\rangle = \sum_k v_k^* w_k$ (Barrett et al. 1994).

3.5 Asymmetric problem: optical bullets

All the examples we have presented have lead to symmetric matrices and it was possible to solve them very efficiently by employing the SYMMLQ solver. For non-symmetric matrices, however, this is not possible, and we require solvers appropriate to dealing with non-symmetric problems (see Sect. 2.3), which in general are not so efficient. This situation can arise physically if we consider the propagation of a temporally finite-width light pulse through a bulk homogeneous medium. This finite duration pulse is what we refer to as a wave packet and is constructed from a linear combination of waves at different frequency. Thus, it gives rise to dispersion phenomena as it propagates. This dispersion is formally described introducing an additional second order time-derivative term in the differential Eq. 1. Therefore, if we consider an homogeneous (bulk) nonlinear Kerr-type medium, we

may write our model as:

$$\sigma \frac{\partial^2 u}{\partial T^2} + \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + u^3 = \beta u. \tag{24}$$

where σ is a dispersion coefficient, which will be negative in case of a normal dispersion and positive in case of an anomalous dispersion. The solutions are localized in space and time and they are named *spatiotemporal solitons* or, due to their shape, *optical bullets* (Kivshar and Agrawal 2003).

Computationally, the problem shares several aspects with those described previously. Indeed, since this problem differs from those treated previously by only the addition of a second-order derivative, it is still symmetric. Now, however, there appear two new fringe diagonals in the system matrix in order to account for the coupling between contiguous terms, which are due to the new second order derivative. Nevertheless, due to the introduction of the third dimension, the problem is inherently three-dimensional, making the system matrix grow as a multiple of the number of discretization points in this third dimension. In practice, the impact of the third dimension is considerable and the grid step must be larger, so that the problem is kept at a reasonable size. Another possibility is to appeal to symmetries in order to reduce the dimension of the problem. If the medium is homogeneous as in the case we are considering, the problem possesses cylindrical symmetry and we may use a cylindrical coordinate system to obtain a solution where the azimuthal coordinate has been factored,

$$u(r, \phi, T) = w(r, T) \exp(i\ell\phi), \tag{25}$$

where the integer ℓ is the vorticity, and thus, if $\ell \neq 0$, we have vortex-like solutions. Substituting this expression into Eq. 24 we obtain,

$$\sigma \frac{\partial^2 w}{\partial T^2} + \frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} - \frac{\ell^2}{r^2} w + w^3 = \beta w. \tag{26}$$

which is two-dimensional, but gives rise to a non-symmetric matrix due to the presence of the first derivative. Therefore, the important reduction of one of the dimensions in the problem is done at a cost of producing a non symmetric problem. A solution to the resulting non-symmetric problem is particularly interesting, since these types of problems also commonly arise in many other areas of physics such as fluid dynamics or diffusive-convective systems.

After applying the usual finite different scheme, taking n and m samples for the variables r and T respectively, a nonlinear algebraic system is obtained whose matrix, after rearranging the functions in the usual column-vector form, has the well-known tridiagonal with two fringes structure, but now with non-symmetric structure. Moreover, the Jacobian matrix has the same structure. We will take the finite difference interval size as Δr and ΔT, respectively for each of the variables, and set $\gamma = \Delta r/\Delta T$. Additionally, the axis r starts at the point $r = 0$, since it only takes positive values, so that the samples for variable r are of the form $r_k = \tilde{k}\Delta r$ where $\tilde{k} = (k \mod n)$. Thus, the diagonals of the Jacobian matrix are given by,

$$J_{kk} = -2(1 + \sigma\gamma^2) + (\Delta r)^2(3w_{\tilde{k}}^2 - \beta) + \frac{1}{\tilde{k}}\left(1 - \frac{\ell^2}{\tilde{k}}\right), \tag{27}$$

$$J_{kk-1} = 1, \qquad J_{kk+1} = 1 - \frac{1}{\tilde{k}}, \tag{28}$$

$$J_{kk-n} = J_{kk-n} = \sigma\gamma^2. \tag{29}$$

A special remark concerning the boundary condition at $r = 0$ is now considered. When $\ell \neq 0$, the condition is also homogeneous and is treated like the other ones. However, for

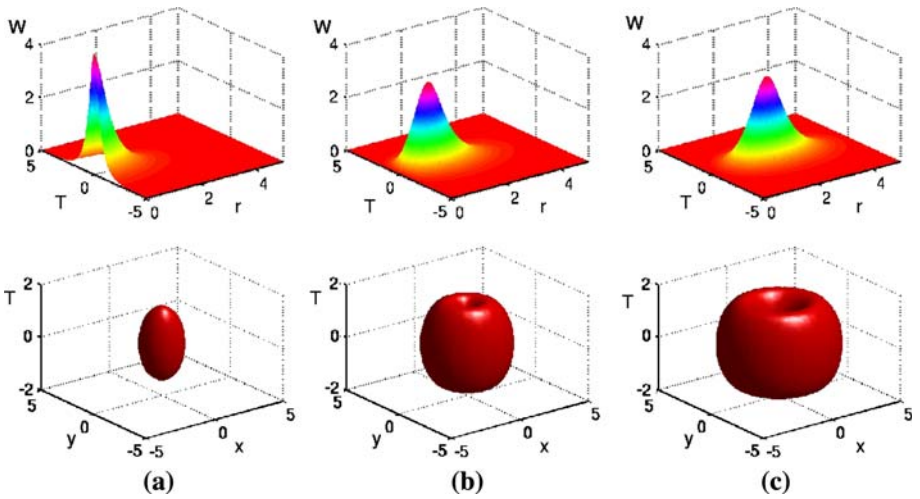

Fig. 6 Three dimensional stationary states for a nonlinear homogeneous Kerr medium. Top row: field amplitude versus radial (r) and temporal (T) coordinates; bottom row: 3-D (x,y,T)-isosurfaces. (**a**) $\ell = 0$, $\beta = 7.83$. (**b**) $\ell = 1$, $\beta = 10.66$. (**c**) $\ell = 2$, $\beta = 11.25$

the particular case where $\ell = 0$, it is necessary to impose a Neumann boundary condition, $(\partial w/\partial r)(r \to 0) \to 0$, which requires the addition of a unity along the main diagonal and the suppression of the terms obtained from the first derivative. Hence, for those boundary points the main diagonal is,

$$J_{kk} = -1 + 2\sigma\gamma^2 + (\Delta r)^2(3w_k^2 - \beta); \qquad \text{if } \tilde{k} = 1. \tag{30}$$

In Fig. 6, three examples of the calculations of spatiotemporal solitons are presented, for the case where $\sigma = 1$ (anomalous dispersion) and for $\ell = 0$ (fundamental), $\ell = 1$ and $\ell = 2$ (vortices), respectively. The function is shown, in the plane (r, T) as a 3-D plot, together with an isosurface (showing the bullet shape) for a particular value of the function to give a spatial view which depends upon the three variables (x, y, T).

For these computations, we have combined the globally convergent Newton Method together with the GMRES solver. In this case, to reasonably improve the efficiency we have used a preconditioner formed by the tridiagonal part of the Jacobian matrix, which can be solved in $\mathcal{O}(N)$ operations by LU-decomposition (Press et al. 1992). This constitutes a block-Jacobi preconditioner. As a initial guess we took a two-dimensional function with Gaussian shape in the T-direction and Gaussian times a monomial shape in the r-direction, in order to resemble the solution. A hand-adjustment of the parameters (peak amplitude and widths) was enough to make the algorithm converge easily.

4 Performance issues

One important topic when one deals with computational methods is the algorithm performance in terms of iterations and running times to solve the different problems. The running times are, of course, a function of the machine used to carry out the computations and a better measurement should be given in terms of machine flops. Less precise but quite informative is the number of iterations that an algorithm needs to undertake to reach the solution of a

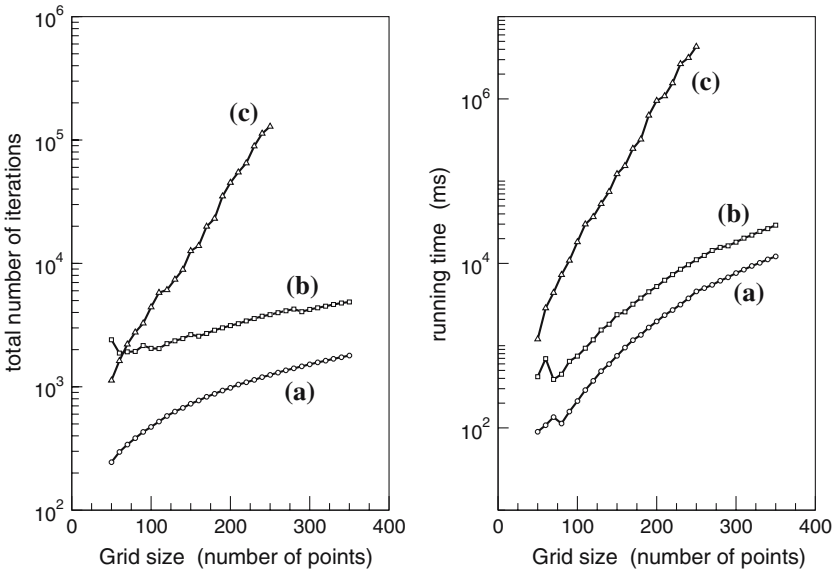

Fig. 7 Performance of different example problems described as a function of the grid size. Shown is the total accumulated number of inner iterations for a particular solver (left plot) and the total running time on our computer (right plot). The different curves correspond to the following cases: (**a**) linear problem for a dual-core PCF coupler solved by inverse iteration, (**b**) the nonlinear problem for the same structure solved by Newton method using SYMMLQ, and (**c**) the case of the optical bullet solved by Newton using GMRES

particular problem with a given precision. This is problem-dependent and also very often a function of the number of samples or finess of the calculation grid. In this section we will give representative results of the number of iterations and running times on our machine, for some of the former representative problems, as well as a comparison to demonstrate that modern solvers are definitively better than the classical methods. Finally, we also address the proper adjustment of the restarting parameter for the problem using the GMRES solver.

The total number of iterations and running times for our computer (INTEL Pentium IV D, CPU 2.8 GHz, 1 Gb RAM) are plotted in Fig. 7 versus the number of mesh points for different problems. We have used a square ($N \times N$) grid with the same number of points (N) in each dimension, so that the grid size given in the horizontal axis is N. According to these simulation results, one observes that only few seconds are necessary in order to obtain the different solutions, even for a relatively large number of grid points, thereby affirming the utility of these methods.

In Fig. 8 we show the comparison results between modern and classical solvers. First we note that among the problems we have taken into account in previous sections, we may apply classical solvers only to the first linear example for calculating the fundamental mode. In fact, for applying those algorithms it is a necessary condition that the system matrix is at least diagonal dominant. This conditions cannot be assured for the rest of the problems treated. The fact that the modern solvers are applicable to a broader variety of problems is the first reason because those algorithms are superior. Anyway, for the sake of comparison, we present results for classical (Jacobi and Gauss-Seidel) and modern (SYMMLQ) solvers for the calculation of the fundamental state of the linear dual-core PCF coupler. In the plot we see that the total number of iterations for the case of using SYMMLQ is dramatically less

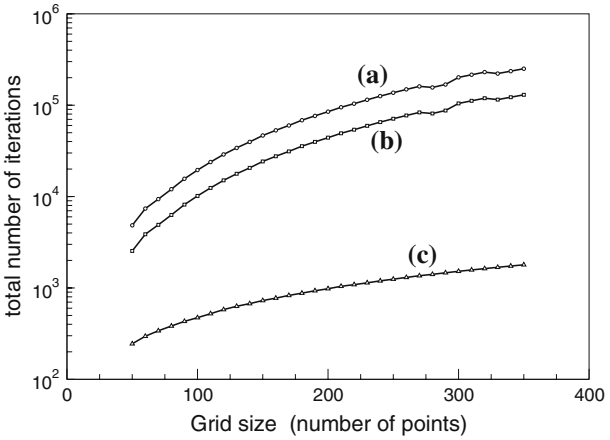

Fig. 8 Performance comparison between the modern solver (SYMMLQ) and two of the classical solvers (Jacobi and Gauss-Seidel) for the calculation of the lowest state of the linear dual-core PCF coupler, using the inverse iteration method: (**a**) Jacobi, (**b**) Gauss-Seidel, (**c**) SYMMLQ

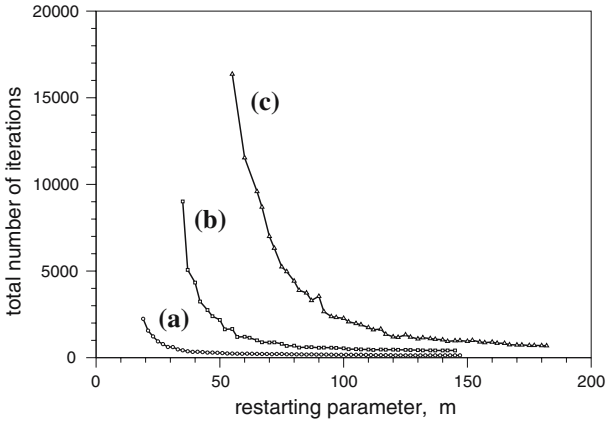

Fig. 9 Total number of iterations versus restart value m for different grid finess. (**a**) $N = 100 \times 100$, (**b**) $N = 200 \times 200$, (**c**) $N = 300 \times 300$

(by two orders of magnitude) than the other methods. Concerning the classical methods it is shown that the Gauss-Seidel method is slightly better, as expected.

Finally we consider the issue of adjustment of the number of iterations (m) before restarting the GMRES algorithm, in order to have maximum numerical efficiency. We have carried out a set of experiments in order to estimate the optimal choice of this parameter for the problem type presented in Sect. 3.5. This numerical study has also shown how the restarting parameter influences convergence and how it depends upon the coarseness of the grid spacing. Thus, for the optical bullets problem, we measured total number of iterations carried out to solve the linear systems by summing over all Newton step iterations, as a function of restart values for different problem sizes. The results are shown in Fig. 9.

Several interesting ideas can be drawn from the Fig. 9. First, there exists a minimum restart value, below which the problem will not converge. Thus, the minimum restart values for our example are 19, 35 and 55 for the cases of a grid of $N = 100 \times 100$, $N = 200 \times 200$, and $N = 300 \times 300$ respectively. A second point is that, for a given problem size, there corresponds an optimal restart value for which placing a larger value will not improve performance. In fact,

when restarting value m is sufficiently increased, the remaining residual is strongly reduced or even convergence is achieved before the algorithm is restarted for the first time, producing no further improvement, and so the computing time and number of iterations saturates. This means that there is no point in setting a value of m much larger than that for which saturation is achieved, since efficiency is not improved and memory unnecessary spoiled. In the plot it is observed a strong dependence on the grid size, being saturation obtained for larger restarting values as the grid size is increased.

5 Conclusions

In this paper we have found the solution for a number of stationary problems of present interest in guided linear and nonlinear optics. We obtained solutions to the optical waveguide stationary states by standard methods such as inverse iteration for linear problems and by Newton methods for nonlinear problems. At each iteration, where we have to obtain a costly linear system solution, we can considerably improve efficiency with the use of modern nonstationary iterative solvers. Due to the iterative nature of those linear solvers the full algorithm resulted in a double inner-outer iterative method. Examples of different characteristics were considered, discussed and solved to demonstrate the power of these algorithms, which may be extensible to many other different problems in other areas of science and engineering.

Acknowledgements J. R. Salgueiro acknowledges a Ramón y Cajal contract granted by the Secretaría de Estado de Universidades e Investigación of Spain. This work was supported by the Ministerio de Educación y Ciencia (project FIS2004-02466) and Xunta de Galicia (projects PGIDIT04TIC383001PR y PGIDIG06PXIB239155PR). D. Olivieri wishes to thank the partial support by the grant 05VI-IC02 of the Universidade de Vigo and the grant TIN2005-08863-C03 of the Spanish Ministerio de Educación y Ciencia.

References

Barrett, R., Berry, M., Chan, T.F., Demmel, J., Donato, J.M., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., van der, H. Vorst : Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods. SIAM, Philadelphia, PA (1994)

Bellavia, S., Morini, B.: A globally convergent Newton-GMRES subspace method for systems of nonlinear equations. SIAM J. Sci. Comput **23**(3), 940–960 (2001)

Demmel, J.W.: Applied Numerical Linear Algebra. SIAM, , Philadelphia, PA (1997)

Dennis, J.E. Jr., Schnabel, R.B.: Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice Hall series in computational mathematics, Prentice-Hall, Englewood Cliffs, NJ (1983)

Dul, F.A.: MINRES and MINERR are better than SYMMLQ in eigenpair computations. SIAM J. Sci. Comput. **19**(6), 1767–1782 (1998)

Ferrando, A., Zacarés, M., Fernández de Córdoba, P., Binosi, D., Monsoriu, J.: Spatial soliton formation in photonic crystal fibers. Opt. Exp. **11**(5), 452–459 (2003)

Ferrando, A., Zacarés, M., Fernández de Córdoba, P., Binosi, D., Monsoriu, J.: Vortex solitons in photonic crystal fibers. Opt. Exp. **12**(5), 817–822 (2004)

Freund, R.W., Nachtigal, N.M.: QMR: A quasi-minimal residual method for non hermitian linear systems. Num. Math **60**(3), 315–339 (1991)

Golub, G.H., Van Loan, C.F.: Matrix Computations. 2nd ed. The Johns Hopkins University Press, London (1989)

Greenbaum, A.: Iterative Methods for Solving Linear Systems. SIAM, Philadelphia, PA (1997)

Ipsen, I.C.F.: Computing an eigenvector with inverse iteration. SIAM Rev. **39**(2), 254–291 (1997)

Jablonski, T.F., Sowinski, M.J.: Analysis of dielectric guiding structures by the iterative eigenfunction expansion method. IEEE Trans. Microwave Theory Tech. **37**(1), 63–70 (1989)

Kivshar, Y.S., Agrawal, G.P.: Optical Solitons: from Fibers to Photonic Crystals. Academic Press, San Diego (2003)

Knight, J.C.: Photonic crystal fibres. Nature **424**, 847–851 (2003)
Knoll, D.A., Rider, W.J.: A multigrid preconditioned Newton-Krylov method. SIAM J. Sci. Comput. **21**(2), 691–710 (1999)
Meurant, G.: Computer Solution of Large Linear Systems. Elsevier Science B. V., Amsterdam (1999)
Nachtigal, N.M., Reddy, S.C., Trefethen, L.N.: How fast are nonsymmetric matrix iterations. SIAM J. Matrix Anal. Appl. **13**(3), 778–795 (1992)
Paige, C.C., Saunders, M.A.: Solution of sparse indefinite systems of linear equations. SIAM J. Numer. Anal. **12**(4), 617–629 (1975)
Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes in C. 2nd ed. Cambridge University Press, Cambridge (1992)
Russell, P.: Photonic crystal fibers. Science **299**, 358–362 (2003)
Saad, M.H., Schultz, Y.: GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci Statist. Comput. **7**(3), 856–869 (1986)
Salgueiro, J.R., Kivshar, Y.S.: Nonlinear dual-core photonic crystal fiber couplers. Opt. Lett. **30**(14), 1858–1860 (2005)
Salgueiro, J.R, Kivshar, Y.S, Pelinovski, D.E, Simón, V., Michinel, H.: Spatial vector solitons in nonlinear photonic crystal fibers. Stud. Appl. Math. **115**, 157–171 (2005)
Sonneveld, P.: CGS, a fast Lanczos-type solver for nonsymmetric linear systems. SIAM J. Sci. Statist. Comput. **10**(1), 36–52 (1989)
Szyld, D.B.: Ctiteria for combining inverse and Rayleigh quotient iteration. SIAM J. Numer. Anal. **25**(6), 1369–1375 (1988)
Tamir, T. (ed.): Guided-Wave Optoelectronics. 2nd ed. Springer-Verlag, Berlin (1990)
van der, H.A. Vorst : Bi-CGSTAB: A fast and smooth converging variant of Bi-CG for the solution of nonsymmetric linear systems. SIAM J. Sci. Statist. Comput **13**(2), 631–644 (1992)
Wilkinson, J.H.: The algebraic eigenvalue problem. Oxford University Press, Oxford (1965)
Young, D.M.: Iterative Solution of Large Linear Systems. Academic Press, New York (1971)