A three-dimensional solver of the Schrödinger equation in momentum space for the detailed simulation of nanostructures

Marco G. Pala
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni, Università di Pisa
Giuseppe Iannaccone
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni, Università di Pisa

M.G. Pala, G. Iannaccone , *A three-dimensional solver of the Schrödinger equation in momentum space for the detailed simulation of nanostructures*, Nanotechnology, **13** , 3, pp.369-372 (2002).

INSTITUTE OF PHYSICS PUBLISHING NANOTECHNOLOGY
Nanotechnology **13** (2002) 369–372 PII: S0957-4484(02)31501-0

A three-dimensional solver of the Schrödinger equation in momentum space for the detailed simulation of nanostructures

M G Pala and G Iannaccone

Università degli Studi di Pisa, Dipartimento di Ingegneria dell'Informazione,
Via Diotisalvi 2, I-56122 Pisa, Italy

E-mail: m.pala@iet.unipi.it and g.iannaccone@iet.unipi.it

Received 4 December 2001, in final form 27 March 2002
Published 23 May 2002
Online at stacks.iop.org/Nano/13/369

Abstract
We propose a simple method for computing the single-particle eigenfunctions in nanostructures with three-dimensional confinement. The proposed procedure transfers the problem to the momentum space, solves an eigenvalue equation on a reduced wavevectors space and then transfers the solution back to the real space. We show that in such a way it is possible to obtain the eigenvectors and eigenvalues corresponding to lower energies with significant improvement in computing time and memory requirements with respect to numerical methods in the coordinate space. The method can be applied to structures with inhomogeneous effective mass and can easily include the full band structure. We have tested the code on typical confining potentials of nanostructures, in order to show the advantages and possible limitations of the proposed method.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Detailed simulation of nanoscale devices requires that quantum mechanical effects that emerge in the case of strong electron confinement be properly taken into account. The many-particle Schrödinger equation can usually be solved with density functional theory, and the local density approximation [1–3]. In such a framework, typical methods are based on the self-consistent solution of the nonlinear Poisson equation and of the single-particle Schrödinger equation.

In the case of three-dimensional confinement, the diagonalization of the Hamiltonian needed to solve the eigenvalue problem ($\mathcal{H}\Psi = \mathrm{E}\Psi$) can require huge computational resources. Indeed, if we discretize the problem domain with a three-dimensional grid of $N \times N \times N$ points we obtain a matrix of order N^3. In the case of the three-dimensional Schrödinger equation we obtain a symmetric matrix in which the non-zero elements are only in seven diagonals ($7N$). For large N the use of direct methods of diagonalization such as the Jacobi method or the Householder method [4] may become prohibitive. Since researchers are frequently interested only in eigenvalues and eigenvectors close to a particular energy level, selective methods that enable partial solutions of the eigenvalue problem have been successfully developed, such as, for example, those based on the method of powers or on the heat equation [4]. A commonly used technique consists in the so-called Lanczos methods, based on the tridiagonalization of the Hamiltonian, and in the direct solution of the tridiagonal matrix [5]. The possible convergence problems arising from the loss of orthogonality of the Lanczos vectors have been overcome by some authors with specific modifications of the standard algorithm [6]. Other authors have proposed selective relaxation methods characterized by strong stability in convergence [7], or the Hamiltonian variational approach [8].

We will show that transferring the eigenvalue problem to the momentum space provides significant advantages with respect to solving the problem in the coordinate space, in terms of computing time and memory requirements. In addition, it easily allows one to perform full band simulations, which is important in strongly confined structures, where the effective mass approximation may be inadequate.

Trellakis [9] analysed the problem of the correction effects due to the non-parabolicity in silicon-based devices. He observed that the kinetic part of the Schrödinger equation is diagonal in the momentum space representation: therefore, he proposes to write the kinetic part in the momentum space and then, by means of the Fourier transform, to transfer the kinetic energy to the coordinate space. Then he used the Arnoldi algorithm, based on the Krylov vectors, to diagonalize, for a minimum number of eigenvectors, the Hamiltonian matrix. He alternatively proposed to transfer the potential term in the momentum space and then to solve the eigenvalue problem.

We propose a code that takes the information on the confining potential in the coordinate space (x-space), but solves the equation in the momentum space (k-space). The steps of this approach consists of transforming the input data through a Fourier transform (in particular, we will use the sine Fourier transform in order to enforce Dirichlet boundary conditions), solving a reduced eigenvalue problem in the k-space and then transferring back the eigenfunctions to the x-space. If we are interested only in the low-energy eigenfunctions, it is possible to strongly reduce the dimension of the Hamiltonian matrix in the k-space with no loss of accuracy in the eigenfunctions and eigenvalues.

2. Theory

Let us explain our approach for a one-dimensional problem. In the x-space we can describe a generic function $f(x)$ with a series $f(x) = \sum_n a_n v_n(x)$, that consists of a linear combination of the vectors of the orthogonal basis $\{v_n(x)\}$. A natural choice for the basis functions would be the complex exponentials $\{e^{ikx}\}$ that imply the use of the fast Fourier transform (FFT) algorithm. Nevertheless, in order to enforce Dirichlet boundary conditions for the eigenfunctions, we use the sine Fourier transform.

However, it is clear that, if the given function $f(x)$ is quite similar to the first vectors of the basis, it can be reasonably approximated by a series truncated to lower order. In our code we calculate the eigenfunctions by solving the Hamiltonian in the k-space and suppose that for a good approximation of the first eigenfunctions it is sufficient to diagonalize a reduced matrix, in which we retain only terms associated with smaller k. The advantage, from the point of view of memory requirements (that scale as N, where N is the matrix dimension) and of computing time (that scales between N^2 and N^3) is evident.

As an example, we first consider the eigenfunctions of a triangular confining potential, for which we have the exact analytical eigenfunctions. In figure 1 we plot the norm of the error associated with the first low-energy eigenfunctions obtained with different discretizations. With dotted curves we indicate the errors due to a solution in the x-space, versus the number N of discretization points in the x-space. Solutions in the k-space are obtained by first discretizing the equation in the x-space on 512 grid points, by transferring the problem into the k-space with the sine Fourier transform, by solving the eigenvalue problem on a minor of dimension N of the complete matrix, and finally by transferring back the solution to the x-space. The norm of the error associated with the solution in the k-space is plotted with solid curves as a function of N. It is worth noting that the computing time basically depends only

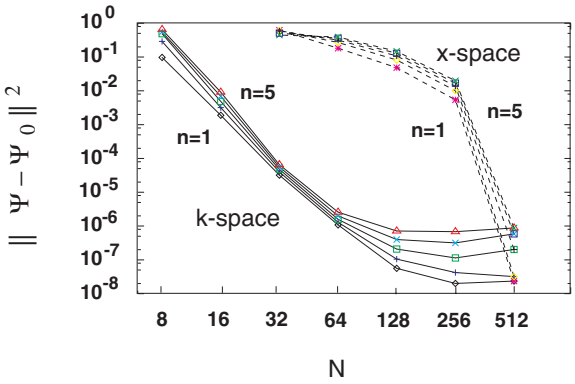

Figure 1. Two-norm of the difference between the analytic eigenfunctions Ψ_0 (Airy functions) and those calculated from the diagonalization of the reduced Hamiltonian of order N in the k-space (solid curves) and those obtained from the diagonalization of the Hamiltonian in the x-space on N grid points (dashed curves). The confining potential is a triangular well. Index n indicates the nth eigenvector ($n = 1$ is the lowest).

on N. It can been seen, for example, that if an error of 10^{-4} is required, we need to solve an eigenvalue problem of order $N = 512$ in the coordinate space, and of only order $N = 32$ in the k-space. Since the computing timescales between N^2 and N^3, the time saving is a factor between 64 and 512.

The single-particle three-dimensional Schrödinger equation in the x-space

$$-\frac{\hbar^2}{2}\left[\frac{\partial}{\partial x}m_x^{-1}(x,y,z)\frac{\partial}{\partial x}+\frac{\partial}{\partial y}m_y^{-1}(x,y,z)\frac{\partial}{\partial y}+\frac{\partial}{\partial z}m_z^{-1}(x,y,z)\frac{\partial}{\partial z}\right]\Psi(x,y,z)+\mathrm{V}(x,y,z)\,\Psi(x,y,z)=\mathrm{E}\,\Psi(x,y,z), \tag{1}$$

in the k-space becomes a very complicated expression that is the generalization of the following one-dimensional formula:

$$\int \mathrm{d}k'_x\,\mathcal{F}_s\{\Psi\}(k'_x)\times\left\{\frac{\hbar^2}{4}k_x\,k'_x[\mathcal{F}_c\{m_x^{-1}\}(k_x-k'_x)+\mathcal{F}_c\{m_x^{-1}\}(k_x+k'_x)]+\tfrac{1}{2}[\mathcal{F}_c\{V\}(k_x-k'_x)-\mathcal{F}_c\{V\}(k_x+k'_x)]\right\}=\mathrm{E}\,\mathcal{F}_s\{\Psi\}(k_x) \tag{2}$$

with obvious notation for the sine transform $\mathcal{F}_s$ and cosine transform $\mathcal{F}_c$. For example, the term relative to the sine Fourier transform of $\frac{\partial}{\partial x}m_x^{-1}(x,y,z)\frac{\partial}{\partial x}\Psi(x,y,z)$ is

$$\begin{aligned}\iiint \mathrm{d}k'_x\,\mathrm{d}k'_y\,\mathrm{d}k'_z\,\mathcal{F}_s\{\Psi\}(k'_x,k'_y,k'_z)\\ \times\tfrac{1}{8}k_x\,k'_x\,[&\mathcal{F}_c\{m_x^{-1}\}(k_x-k'_x,k_y-k'_y,k_z-k'_z)\\ -&\mathcal{F}_c\{m_x^{-1}\}(k_x-k'_x,k_y-k'_y,k_z+k'_z)\\ +&\mathcal{F}_c\{m_x^{-1}\}(k_x+k'_x,k_y-k'_y,k_z-k'_z)\\ -&\mathcal{F}_c\{m_x^{-1}\}(k_x+k'_x,k_y-k'_y,k_z+k'_z)\\ -&\mathcal{F}_c\{m_x^{-1}\}(k_x-k'_x,k_y+k'_y,k_z-k'_z)\\ +&\mathcal{F}_c\{m_x^{-1}\}(k_x-k'_x,k_y+k'_y,k_z+k'_z)\\ -&\mathcal{F}_c\{m_x^{-1}\}(k_x+k'_x,k_y+k'_y,k_z-k'_z)\\ +&\mathcal{F}_c\{m_x^{-1}\}(k_x+k'_x,k_y+k'_y,k_z+k'_z)]\end{aligned}$$

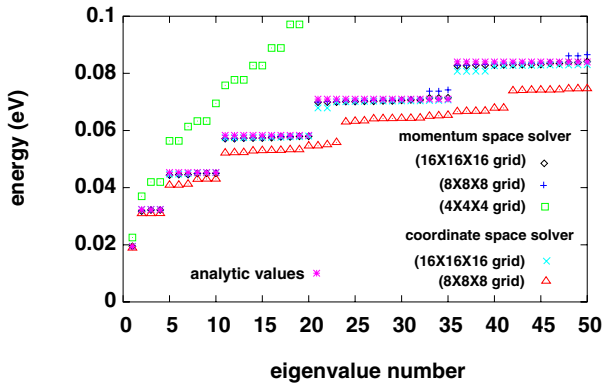

Figure 2. The lowest 50 eigenvalues of a three-dimensional harmonic oscillator corresponding to different truncated bases in the x-space or in the k-space, compared with the analytic eigenvalues. A truncated basis of 8^3 elements in the k-space provides the same accuracy as a basis of 16^3 elements in the x-space.

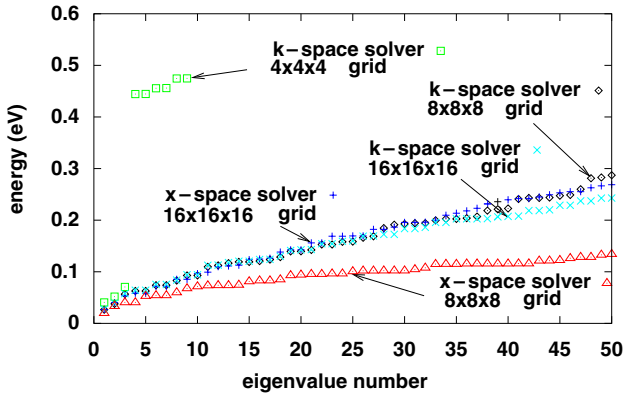

Figure 3. The lowest 50 eigenvalues of a cylindrical confining corresponding to different truncated bases in the x-space or the k-space. The cylinder has a circular basis with radius 10 nm and a height of 10 nm, and potential barriers of 1 eV.

where

$$\mathcal{F}_s\{f\}(k_x,k_y,k_z) \equiv \iiint \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ \sin(k_x x)\ \sin(k_y y) \times \sin(k_z z)\ f(x,y,z) \tag{3}$$

and

$$\mathcal{F}_c\{f\}(k_x,k_y,k_z) \equiv \iiint \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ \cos(k_x x)\ \cos(k_y y) \times \cos(k_z z)\ f(x,y,z) \tag{4}$$

are, respectively, the sine transform and cosine transform of $f(x,y,z)$ and we have another similar expression for the other part of the Hamiltonian.

From the discretization of the three-dimensional generalization of equation (2) we find the matrix of the complete eigenvalue problem. Then we diagonalize only the sub-matrix corresponding to smaller k vectors.

3. Simulations

In this section we show examples for evaluating the efficiency of the three-dimensional Schrödinger solver in the momentum space for typical confining potentials in nanostructures.

As a first test structure we consider a three-dimensional harmonic potential $V(x,y,z) = \frac{1}{2}m\omega^2(x^2+y^2+z^2)$. The parameter values are $m = 0.067 \times 9.1 \times 10^{-31}$ kg (GaAs effective mass), $\omega = 2 \times 10^{13}$ Hz and the domain size a cube of dimensions $200 \times 200 \times 200$ nm^3. We compare the performance of our code and with that of codes based on the solution in the coordinate space, in terms of accuracy and computing time. The Schrödinger equation has been discretized in the coordinate space on a grid of $16 \times 16 \times 16$ points. Then, the whole matrix of order 16^3 has been transferred into the k space by means of the FFT. In the k-space, we reduced the eigenvalue problem by keeping only the matrix coefficients corresponding to low-k terms. In figure 2 we plot the first 50 energy eigenvalues obtained from a basis in the k-space of $16 \times 16 \times 16$ vectors, on a basis of $8 \times 8 \times 8$ and on a basis of $4 \times 4 \times 4$. In the same figure we also plot the exact analytic solution and the values obtained from an x-space solver on a grid of $16 \times 16 \times 16$ points and $8 \times 8 \times 8$ points. From figure 2 it can be noticed that the solution on a reduced set of 8^3 vectors in the k-space provides better accuracy than the solution on a grid of 16^3 points in the x-space.

Table 1. The CPU running time of an 800 MHz Pentium III required for the solution of the eigenvalue problem for a parabolic potential defined on a grid of $16 \times 16 \times 16$ points in the real space. The computing time for solving the problem in the k-space includes the FFT of the equation, the complete diagonalization and the anti-FFT of the eigenvectors.

Sub matrix dimension	Computing time
$16 \times 16 \times 16$ in the k-space	89 min (complete diag.)
$16 \times 16 \times 8$ in the k-space	10 min (complete diag.)
$16 \times 8 \times 8$ in the k-space	64 s (complete diag.)
$8 \times 8 \times 8$ in the k-space	7 s (complete diag.)
$16 \times 16 \times 16$ in the x-space	95 min (complete diag.)
$16 \times 16 \times 16$ in the x-space	20 min (first ten eigenvalues)

Table 1 shows the CPU running times required by a 800 MHz Pentium III PC, with the solver based on wavevector grids of different size. Well known routines based on the Householder methods[1] are used for the complete diagonalization, while the routine DNLASO [10], based on a Lanczos method, is used for extraction of the lowest ten eigenvalues. The solution on the reduced set of 8^3 vectors in the k space requires 7 s, while the solution on 16^3 points in the coordinate space requires 95 min for the complete diagonalization, or 20 min for extracting only the lowest ten eigenvalues.

Let us underline the fact that the main advantage comes from the fact that we can use a fine discretization in the x-space, in order not to lose the details of the confining potential, because the computational cost of performing an FFT of a matrix of order 16^3 is negligible (smaller than 50 ms).

With the use of the proposed routine, the computing time for the solution of the single-particle Schrödinger equation becomes negligible with respect to the total time for the self-consistent solution of the Poisson–Schrödinger equation. Therefore, we can perform a complete diagonalization of the Hamiltonian with a simple well known code[2].

[1] Routines `tred2` and `tqli` from [4].
[2] See footnote 1.

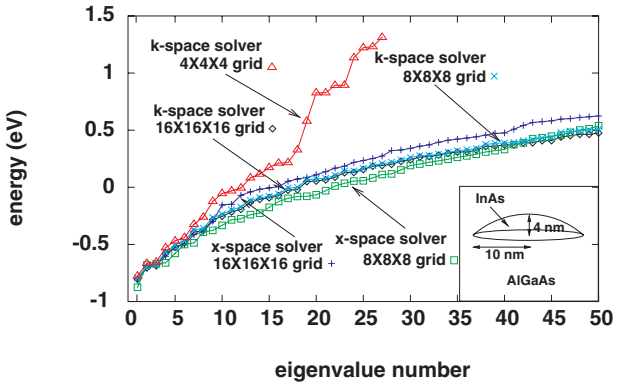

Figure 4. The lowest 50 confining eigenvalues corresponding to a lens-shaped InAs quantum dot with a circular basis of radius 10 nm and a height of 4 nm (shown in the inset). The best trade-off between computing time and accuracy is obtained with a basis of 8^3 elements.

It is clear that the method takes advantage of the similarity between the sines of lower order and the eigenfunctions of the Hamiltonian related to smaller energies. Therefore, it provides the best performance in the case of strongly confined systems, where the confining region has a shape similar to the domain in which the Schrödinger equation is solved.

The second example we consider is a quantum dot defined by a cylindrical potential, with a circular base of radius 10 nm, a height of 10 nm and a confining energy of 1 eV. The data presented in figure 3 give the same results as figure 2, and shows that it is possible decrease the dimension of the Hamiltonian in the k-space to $8 \times 8 \times 8$ points with the same accuracy of a solution on a $16 \times 16 \times 16$ grid, while this would not be possible in the x-space. In this case, the potential in the x-space is discretized on a grid of 32^3 points.

As a third example we have simulated a lens-shaped quantum dot of InAs embedded in $Al_{0.2}Ga_{0.8}As$, with a radius of 10 nm and a height of 4 nm, shown in the inset of figure 4. The potential well has a height of 1.05 eV. The data presented in figure 4 gives the same results as the data presented in figures 2 and 3. Again, a solution on a basis of 8^3 elements in the k space provides the same accuracy of a solution on 16^3 points in the coordinate space.

4. Conclusions

We have presented a method to solve the three-dimensional Schrödinger equation for systems with variable effective mass based on the expansion of the sine Fourier series. We have shown that this method allows us to solve the eigenvalue problem on a reduced matrix in the k-space, with negligible loss of accuracy, and significant saving in terms of computing time. This method also has the advantage of allowing full band simulations, which may be important for nanoscale structures when the parabolic approximation of the dispersion relation does not hold. We have shown that if we are interested only in the lower eigenvalues, the order of the matrix can be typically reduced by a factor 8, which typically means a corresponding reduction of the computing time by about two orders of magnitude. The same method is presently being applied to the solution of the Schrödinger equation with open boundary conditions.

Acknowledgment

Support from the IST NANOTCAD project (EC contract IST-1999-10828) is gratefully acknowledged.

References

[1] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864–73
[2] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133–8
[3] von Barth U and Hedin L 1972 *J. Phys. C: Solid State Phys.* **5** 1629–42
[4] Press W H, Vettering W T, Teulolsky S A and Flannery B P 1992 *Numerical Recipes in FORTRAN 77, The Art of Scientific Computing* (Cambridge: Cambridge University Press)
[5] Lanczos C 1956 *Applied Analysis* (Englewood Cliffs, NJ: Prentice-Hall)
[6] Grosso G, Martinelli L and Parravicini G P 1995 *Phys. Rev.* B **51** 13 033–8
[7] Presilla C and Tambini U 1995 *Phys. Rev.* E **52** 4495–8
[8] Dagotto E and Moreo A 1985 *Phys. Rev.* D **31** 865–70
[9] Trellakis A 2000 *PhD Dissertation* University of Illinois, Urbana-Champaign
[10] Scott D S and Parlett B N *Routine DNLASO of the LASO Package* Computer Sciences Department, University of Texas at Austin (available at www.netlib.org)