A method to compute the inverse of an n-block tridiagonal quasi-Hermitian matrix

1991 J. Phys.: Condens. Matter 3 7843

(http://iopscience.iop.org/0953-8984/3/40/005)

J. Phys.: Condens. Matter 3 (1991) 7843–7848. Printed in the UK

A method to compute the inverse of an n-block tridiagonal quasi-Hermitian matrix

Elena M Godfrin†
Departamento de Física de la Materia Condensada Universidad Autónoma de Madrid 28049 Cantoblanco, Madrid, Spain

Received 3 September 1990

Abstract. This paper presents a method for computing the inverse of a complex n-block tridiagonal quasi-Hermitian matrix using an adequate number of partitions of the complete matrix. This type of matrix is very usual in quantum mechanics and, more specifically, in solid state physics (e.g. interfaces and super-lattices), when the tight-binding approximation is used. The efficiency of the method is analysed by comparing the required CPU time and work-area with other techniques.

1. Introduction

The linear combination of atomic orbital methods [1, 2], especially within the tight-binding approximation, is one of the most commonly used techniques in quantum mechanics for the analysis of the properties related to elementary excitations (electrons, phonons, etc). Its application to the study of molecules and solids gives Hamiltonian operators which usually have, when written in an appropriate base, many null elements. In periodical systems such as crystal solids and super-lattices the Hamiltonian may be represented as a quasi-Hermitian‡, block-tridiagonal matrix, in many cases with a high dimension (e.g. [3]). The density of states as a function of the energy is obtained from the trace of the Green function matrix $\mathbf{G}$ [1, 2] defined as

$$\mathbf{G} = (\mathbf{E}\mathbf{I} - \mathbf{H})^{-1}$$

where $\mathbf{E}$ is a scalar and $\mathbf{I}$, the identity matrix. The numerical techniques usually used are not appropriate for solving this equation since they do not take into account the particular structure of $\mathbf{H}$ (essentially the large number of null elements), which is inefficient both in CPU time and memory requirements.

This article presents a method for computing the inverse of a complex n-block tridiagonal quasi-Hermitian matrix which considers the structure of $\mathbf{H}$ using an adequate number of partitions of the complete matrix. Its efficiency is analysed by comparing its required CPU time with other techniques. It should be mentioned that the first version of this method (which only computes the diagonal blocks of the inverse matrix) has been used in several studies of the electronic properties of semiconductor interfaces (e.g. see [4, 5]).

† Permanent address: Departamento Fuentes Renovables y Uso Racional de la Energía Comisión Nacional de Energía Atómica Av. del Libertador 8'250 1429 Buenos Aires, Argentina.

‡ $M_{ij} \in C^{p\times q}$, $M_{ij} = M_{ji}^{\dagger}$ if $i \neq j$ and $\det M_{ii} \neq 0$, where M^{+} indicates the transposed and conjugate matrix of $\mathbf{M}$.

0953-8984/91/407843+06$03.50

2. Mathematical problem

Let **M** be the n-block tridiagonal quasi-Hermitian matrix

$$\begin{pmatrix} M_{11} & M_{21}^{+} & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ M_{21} & M_{22} & M_{32}^{+} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & M_{32} & M_{33} & M_{43}^{+} & \cdots & \cdots & \cdots & \\ \vdots & & & \ddots & & & \vdots & \\ 0 & \cdots & \cdots & \cdots & 0 & M_{n-1,n-2} & M_{n-1,n-1} & M_{n,n-1}^{+} \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & M_{n,n-1} & M_{n,n} \end{pmatrix}.$$

The problem consists of finding the matrix **G** such that $\mathbf{M} \times \mathbf{G} = \mathbf{I}$, where **I** is the identity matrix (it should be mentioned that the different blocks of the matrix may not necessarily be of equal dimension).

Partitioning the matrix **G** in the same way as **M**,

$$\mathbf{G} = \begin{pmatrix} G_{11} & G_{12} & \cdots & G_{1n} \\ G_{21} & G_{22} & \cdots & G_{2n} \\ \vdots & \vdots & & \vdots \\ G_{n1} & G_{n2} & \cdots & G_{nn} \end{pmatrix}$$

and the system $\mathbf{M} \times \mathbf{G} = \mathbf{I}$ can be written as

$$\begin{cases} M_{11}G_{11} + M_{21}^{+}G_{21} & & & & & = \mathbf{I} \\ M_{21}G_{11} + M_{22}G_{21} + M_{32}^{+}G_{31} & & & & & = 0 \\ M_{32}G_{21} + M_{33}G_{31} + M_{43}^{+}G_{41} & & & & & = 0 \\ \vdots & & & & & \\ M_{n,n-1}G_{n-1,1} + M_{nn}G_{n1} & & & & & = 0 \end{cases}$$

$$\vdots$$

$$\begin{cases} M_{11}G_{1n} + M_{21}^{+}G_{2n} & & & & & = 0 \\ M_{21}G_{1n} + M_{22}G_{2n} + M_{32}^{+}G_{3n} & & & & & = 0 \\ M_{32}G_{2n} + M_{33}G_{3n} + M_{43}^{+}G_{4n} & & & & & = 0 \\ \vdots & & & & & \\ M_{n,n-1}G_{n-1,n} + M_{nn}G_{nn} & & & & & = \mathbf{I} \end{cases}$$

where **I** is the identity matrix with an adequate number of dimensions.

The set of matrices $G_{i,j}$ is obtained by solving the matricial equation system and, thus, the inverse of **M** is computed.

3. Problem resolution

The general resolution for a general n-block matrix **M** is very tedious and does not contribute to the clarification, thus only the 4-block problem will be solved.

Therefore, let **M** and **G** be such that

$$\begin{pmatrix} M_{11} & M_{21}^{+} & 0 & 0 \\ M_{21} & M_{22} & M_{32}^{+} & 0 \\ 0 & M_{32} & M_{33} & M_{43}^{+} \\ 0 & 0 & M_{43} & M_{44}^{+} \end{pmatrix} \begin{pmatrix} G_{11} & G_{12} & G_{13} & G_{14} \\ G_{21} & G_{22} & G_{23} & G_{24} \\ G_{31} & G_{32} & G_{33} & G_{34} \\ G_{41} & G_{42} & G_{43} & G_{44} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & \mathbf{I} \end{pmatrix}$$

or, equivalently,

$$\begin{aligned} &M_{11}G_{11} + M_{21}^{+}G_{21} = \mathbf{I} \\ &M_{21}G_{11} + M_{22}G_{21} + M_{32}^{+}G_{31} = 0 \\ &M_{32}G_{21} + M_{33}G_{31} + M_{43}^{+}G_{41} = 0 \\ &M_{43}G_{31} + M_{44}G_{41} = 0 \end{aligned} \tag{1}$$

$$\begin{aligned} &M_{11}G_{12} + M_{21}^{+}G_{22} = 0 \\ &M_{21}G_{12} + M_{22}G_{22} + M_{32}^{+}G_{32} = \mathbf{I} \\ &M_{32}G_{22} + M_{33}G_{32} + M_{43}^{+}G_{42} = 0 \\ &M_{43}G_{32} + M_{44}G_{42} = 0 \end{aligned} \tag{2}$$

$$\begin{aligned} &M_{11}G_{13} + M_{21}^{+}G_{23} = 0 \\ &M_{21}G_{13} + M_{22}G_{23} + M_{32}^{+}G_{33} = 0 \\ &M_{32}G_{23} + M_{33}G_{33} + M_{43}^{+}G_{43} = \mathbf{I} \\ &M_{43}G_{33} + M_{44}G_{43} = 0 \end{aligned} \tag{3}$$

$$\begin{aligned} &M_{11}G_{14} + M_{21}^{+}G_{24} = 0 \\ &M_{21}G_{14} + M_{22}G_{24} + M_{32}^{+}G_{34} = 0 \\ &M_{32}G_{24} + M_{33}G_{34} + M_{43}^{+}G_{44} = 0 \\ &M_{43}G_{34} + M_{44}G_{44} = \mathbf{I} \end{aligned} \tag{4}$$

Solving (1) and (2), and defining

$$\begin{aligned} X_4 &= 0 \\ X_3 &= M_{43}^{+}[M_{44} - X_4]^{-1}M_{43} \\ X_2 &= M_{32}^{+}[M_{33} - X_3]^{-1}M_{32} \\ X_1 &= M_{21}^{+}[M_{22} - X_2]^{-1}M_{21} \end{aligned}$$

and

$$\begin{aligned} Y_1 &= 0 \\ Y_2 &= M_{21}[M_{11} - Y_1]^{-1}M_{21}^{+} \\ Y_3 &= M_{32}[M_{22} - Y_2]^{-1}M_{32}^{+} \\ Y_4 &= M_{43}[M_{33} - Y_3]^{-1}M_{43}^{+} \end{aligned}$$

the following expressions are obtained:

$$G_{41} = -[M_{44} - X_4]^{-1} M_{43} G_{31}$$
$$G_{31} = -[M_{33} - X_3]^{-1} M_{32} G_{21}$$
$$G_{21} = -[M_{22} - X_2]^{-1} M_{21} G_{11}$$
$$G_{11} = [M_{11} - X_1 - Y_1]^{-1}$$

and

$$G_{42} = -[M_{44} - X_4]^{-1} M_{43} G_{32}$$
$$G_{32} = -[M_{33} - X_3]^{-1} M_{32} G_{22}$$
$$G_{22} = [M_{22} - X_2 - Y_2]^{-1}$$
$$G_{12} = -[M_{11} - Y_1]^{-1} M_{21}^{+} G_{22}.$$

G_{i3} and G_{i4} (for $1 \leqslant i \leqslant 4$) are obtained in a similar way. The general rule for a general n-block matrix **M** is immediately derived:

$$G_{ii} = [M_{ii} - X_i - Y_i]^{-1}$$
$$G_{ij} = C_i G_{i-1,j} \qquad \text{for } i > j$$
$$G_{ij} = D_i G_{i+1,j} \qquad \text{for } i < j$$

for $1 \leqslant i, j \leqslant n$, where

$$X_n = 0$$
$$X_{n-i} = M^{+}_{n-i+1,n-i}[M_{n-i+1,n-i+1} - X_{n-i+1}]^{-1} M_{n-i+1,n-i} \qquad \text{for } 1 \leqslant i \leqslant n-1$$
$$Y_1 = 0$$
$$Y_{i+1} = M_{i+1,i}[M_{ii} - Y_i]^{-1} M^{+}_{i+1,i} \qquad \text{for } 2 \leqslant i \leqslant n$$
$$C_i = -[M_{ii} - X_i]^{-1} M_{i,i-1} \qquad \text{for } 1 \leqslant i \leqslant n-1$$
$$D_i = -[M_{ii} - Y_i]^{-1} M^{+}_{i+1,i} \qquad \text{for } 2 \leqslant i \leqslant n.$$

4. Computational solution

The computer code of this algorithm has been written in FORTRAN77, as a self-consistent routine using a band storage mode for the matrix **M** (only the diagonal and infra-diagonal blocks are stored); the required work-area† is less than or equal to $2 \times \text{MDIM}^2 \times (3 + \text{NB})$, where MDIM is the maximum block dimension and NB is the number of blocks of the matrix **M**.

The alternative methods selected for comparison with this result are the LEQT2C and MA23A routines from the IMSL and Harwell libraries respectively (the only ones available in our computer centre which can accomplish the desired task), and a routine

† The storage area for the matrix **M** is not included.

Table 1. Size of the required work-area, in units of the length of a complex number (typically, 8 byte in single precision), for three matrix dimensions. The block dimensions used are: (*a*) 5 × 5; (*b*) 10 × 10.

Matrix dimension	30	50	100
MA23A	1050	2750	10500
This method	450(*a*)	650(*a*)–1600(*b*)	1150(*a*)–2600(*b*)

Table 2. CPU time involved, in seconds. The block dimensions used are: (*a*) 5 × 5; (*b*) 10 × 10.

Matrix dimension	30	50	100
MA23A	0.54	1.88	7.70
INV	0.42	1.84	13.87
This method	0.26(*a*)	0.58(*a*)–1.25(*b*)	1.82(*a*)–3.74(*b*)

called INV, which is the simplest version of the inversion algorithm. All of them use a full storage mode for the matrix **M**.

The LEQT2C routine cannot compute the inverse matrix for the different **M** proposed (the return code obtained indicates that the matrix is algorithmically singular).

The numerical values obtained by the algorithm presented in this article and by the MA23A and INV routines coincide, at least, until the fourth significant digit. The CPU times required in a BASF-768 computer and the size of required work-area by the methods are given in tables 1 and 2.

5. Conclusions

Considering the results described in the preceding section, this method allows a more efficient use of memory and computational time for the complex n-block tridiagonal quasi-Hermitian matrices. This method, with adequate modifications, can also be applied to solve the systems $\mathbf{GM} = \mathbf{B}$, where **M** is a matrix as described, especially if **B** is a diagonal matrix.

It is important to notice that the calculations of the matrices **X** and **Y** are independent, as is the case for the infra-diagonal and supra-diagonal blocks; therefore, this method can be used in vectorial computers with the corresponding computer code.

Acknowledgments

This work has been partially supported by Scientific Cooperation Contract No. ST2J-0254-7-E (EDB) between the European Communities and the Universidad Autónoma de Madrid (Spain). The author wishes to thank Professor F Flores for his support and stimulation during the development of this work. She also acknowledges A Muñoz and J C Durán for drawing her attention to the problem discussed in this paper and for using this method in their work since 1986, and especially the second one, for reading the manuscript with fruitful criticism, which contributed towards clarification of its contents.

References

[1] Ziman J M 1963 *Principles of the Theory of Solids* (Cambridge: Cambridge University Press)
[2] Ashcroft N W and Mermin N D 1981 *Solid State Physics* (New York: Holt-Saunders)
[3] Guinea F, Sanchez-Dehesa J and Flores F 1983 *J. Phys. C: Solid State Phys.* **16** 6499
[4] Muñoz A, Durán J C and Flores F 1987 *Surf. Sci.* **181** L200
[5] Muñoz A, Pérez R, Durán J C and Flores F 1989 *Surf. Sci.* **211/212** 503