NANOSYSTEMS: PHYSICS, CHEMISTRY, MATHEMATICS, 2013, 4 (3), P. 306–312

METHOD OF SYMMETRIC POLYNOMIALS IN THE COMPUTATIONS OF SCATTERING MATRIX

Yu. N. Belyayev

Syktyvkar State University, Oktyabrskii pr. 55, Syktyvkar-167001, Russia

ybelyayev@mail.ru

PACS 02.10.Yn, 02.30.Hq, 11.55.-m

The method for calculating any analytic matrix function by means of symmetric polynomials is presented. The method of symmetric polynomials (MSP) is applied to the calculation of the fundamental matrix of a differential equations system. The scaling method is developed for computation of the scattering matrix. An analytical estimate of the scaling parameter, allowing the calculation of the matrix exponential with the required reliability and accuracy is obtained. This parameter depends on the matrix order n, the value of the matrix elements and layer thickness.

Keywords: layered media, matrix, exponential, symmetric polynomials, roundoff error, truncation error, scaling.

1. Introduction

The dynamical theory of electron diffraction in crystals is based on the equations of Howie-Whelan [1]. An analogous approach to X-ray diffraction is based on Takagi-Taupin equations [2]. Calculation of multiple-wave diffraction in the crystalline layer thickness z on these equations leads to the Cauchy problem

$$\frac{d\Psi(z)}{dz} = A\Psi(z), \quad \Psi(0) = \Psi_0, \tag{1}$$

where $\Psi(z)$ — column matrix whose components are wave functions $\psi_i(z)$, $i = 1, 2, \ldots, n$; $A \equiv \left\|a_{ij}\right\|$ — matrix of order n; column matrix Ψ_0 consists of n known ψ_{i0}.

The fundamental matrix S of the differential equations system (1) converts the wave field $\Psi(0)$ in the field $\Psi(z)$ at a depth z: $\Psi(z) = S\Psi_0$. In the theory of diffraction S is known as the scattering matrix.

If the crystal is perfect, the scattering matrix is defined as follows:

$$S = \exp(Az) \equiv I + Az + \frac{(Az)^2}{2!} + \frac{(Az)^3}{3!} + \ldots, \tag{2}$$

where I — unit matrix. If the coefficients a_{ij} of differential equations (1) are functions of z, then one way to solve the problem (1) is the partition of the interval $[0, z]$ on N subdomains: $[0 = z_0, z_1]$, $[z_1, z_2], \ldots, [z_{N-1}, z_N = z]$. The number N should be large enough so that $A \simeq A_l = const_l$ for any $z \in [z_{l-1}, z_l]$, $l = 1, 2, \ldots, N$, and the scattering matrix can be approximated by the product: $S = \prod_{l=N}^{1} \exp\bigl(A_l(z_l - z_{l-1})\bigr)$.

Calculations of $\exp(Az)$ with the aid of the Lagrange-Sylvester [3], Becker [4], Newton [5] formulas or matrix decomposition methods [6] requires the prior determination of the matrix A eigenvalues λ_j, $j = 1, 2, \ldots, n$. Another approach to the calculation of the matrix exponential is based on the use of symmetric polynomials [7]. This method is applied in the present work to calculate the scattering matrix.

2. Method of symmetric polynomials (MSP)

According to the Cayley-Hamilton theorem, any matrix $A \equiv \|a_{ij}\|$ of order n satisfies its characteristic equation $\lambda^n - \sigma_1\lambda^{n-1} + \sigma_2\lambda^{n-2} - \ldots (-1)^n\sigma_n = 0$, i.e.

$$A^n - p_1A^{n-1} - p_2A^{n-2} - \ldots - p_nI = 0\,. \tag{3}$$

Here we use the notations

$$p_j = (-1)^{j-1}\sigma_j, \quad j = 1, \ldots, n, \tag{4}$$

$I \equiv A^0$, and σ_i, $i = 1, 2, \ldots, n$, — sums of principal minors of i-th order $\det A$:

$$\sigma_1 = a_{11} + a_{22} + \ldots + a_{nn}, \quad \sigma_2 = \sum_{j>i} \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}, \quad \ldots, \quad \sigma_n = \det A\,. \tag{5}$$

Conversely, as is known, σ_j are the elementary symmetric polynomials in the eigenvalues λ_j of matrix A: $\sigma_1 = \lambda_1+\lambda_2+\ldots+\lambda_n$, $\sigma_2 = \sum_{l\neq j}\lambda_l\lambda_j$, $\ldots$, $\sigma_n = \lambda_1\lambda_2\ldots\lambda_n$. Therefore, any function of the elementary symmetric polynomials σ_i is also symmetric in the eigenvalues λ_j of the matrix A.

D e f i n i t i o n 1. Functions $\mathscr{B}_g(n) \equiv \mathscr{B}_g(p_1, \ldots, p_n)$ that satisfy the recurrence relations

$$\mathscr{B}_g(n) = \begin{cases} 0, \text{ if } g = 0, 1, \ldots, n-2, \\ 1, \text{ if } g = n-1, \\ p_1\mathscr{B}_{g-1}(n) + p_2\mathscr{B}_{g-1}(n) + \ldots + p_n\mathscr{B}_{g-n}(n), \end{cases} \tag{6}$$

are called *symmetric polynomials of n-th order.*

2.1. Representation of matrix functions by means of symmetric polynomials

As a result of equality (3) any integer power $j > 0$ matrix A can be expressed as a linear combination of the first n powers A^0, $A, \ldots, A^{n-1}$:

$$A^j = \sum_{l=0}^{n-1} C_{jl}A^l, \tag{7}$$

where C_{jl} are functions of p_i.

T h e o r e m 1. *If the matrix A is nonsingular, then the coordinates C_{jl} of the matrix A^j in the basis A^0, $A, \ldots, A^{n-1}$ are represented by the formulas*:

$$\left.\begin{aligned} C_{j0} &= \mathscr{B}_{j+n-1}(n) - p_1\mathscr{B}_{j+n-2}(n) - \ldots - p_{n-1}\mathscr{B}_j(n), \\ C_{j1} &= \mathscr{B}_{j+n-2}(n) - p_1\mathscr{B}_{j+n-3}(n) - \ldots - p_{n-2}\mathscr{B}_j(n), \\ &\vdots \\ C_{j(n-2)} &= \mathscr{B}_{j+1}(n) - p_1\mathscr{B}_j(n), \\ C_{j(n-1)} &= \mathscr{B}_j(n), \end{aligned}\right\} \tag{8}$$

where j is any integer, and $\mathscr{B}_j(n)$ — symmetric polynomials of n-th order matrix A .

If $\det A = 0$, then (8) determine the coefficients C_{jl} for $j \geqslant 0$.

Relations (8) is easily verified for $j = 0, 1, \ldots, n-1$ and for other values of j formulas (8) can be proved by induction.

C o r o l l a r y 1.1.

$$C_{jl} = \sum_{g=0}^{l} p_{n-l+g}\mathscr{B}_{j-1-g}(n), \quad l = 0, \ldots, n-1, \quad \textit{where} \begin{cases} j \textit{ is any integer}, & \textit{if } \det A \neq 0, \\ j \geqslant n, & \textit{if } \det A = 0. \end{cases} \tag{9}$$

Formulas (8) and (9) are equivalent in the sense that their right-hand sides are equal to each other due to the recurrence relations (6).

Corollary 1.2. *Let $f(\zeta)$ be an entire function of complex variable ζ, then the function $f(A)$ of n-th order matrix A has the following representation*

$$f(A) = \sum_{l=0}^{n-1} A^l \left[\alpha_l + \sum_{g=0}^{l} p_{n-l+g} \sum_{j=n}^{\infty} \alpha_j \mathscr{B}_{j-1-g}(n) \right], \tag{10}$$

where symmetric polynomials $\mathscr{B}_{j-1-g}(n)$ of n-th order matrix A are defined by (6) *and α_l — coefficients of the power series*: $f(\zeta) = \sum_{j=0}^{\infty} \alpha_j \zeta^j$.

Proof. If $f(\zeta)$ is an analytic function on the whole complex plane, then the expansion $f(\zeta) = \sum_{j=0}^{\infty} \alpha_j \zeta^j$ remains valid when replacing the complex variable ζ on a square matrix A: $f(A) = \sum_{j=0}^{\infty} \alpha_j A^j$. We substitute in the last equality the expression of the matrix A^j according to formulas (7) and (9). Equality (10) - is the result of this substitution.

□

Corollary 1.3. *For any n-th order matrix A and scalar z*

$$\exp(Az) = \sum_{l=0}^{n-1} (Az)^l \frac{1}{l!} \left[1 + l! \sum_{g=0}^{l} p_{n-l+g} \sum_{j=n}^{\infty} \frac{1}{j!} \mathscr{B}_{j-1-g}(n) \right], \tag{11}$$

where p_g, $g = 1, \ldots, n$, and $\mathscr{B}_j(n)$ are symmetric polynomials of matrix Az.

2.2. Estimation of symmetric polynomials modulus

The elementary symmetric polynomial σ_j is the sum of the principal minors of j-th order determinant of the matrix A. The number of such minors is C_n^j. One minor of j-th order contains $j!$ terms, each of which is a product of the j matrix elements a_{ij}. We denote by $\max|a_{gl}|$ the greatest value of the modulus of the matrix elements a_{gl}. Therefore, using the definition (4), we obtain the following relations

$$|p_j| \leqslant p_{j_M} = \frac{n!}{(n-j)!} \left(\max|a_{gl}|\right)^j \ , \quad j = 1, 2, \ldots, n. \tag{12}$$

Theorem 2.

$$|\mathscr{B}_j(n)| < \frac{n}{2n-1} \, x^{j-n+1}, \tag{13}$$

where

$$x = (2n-1) \max|a_{gl}| \, . \tag{14}$$

Proof. Using (12) and definition (6), we obtain

$$\begin{aligned} |\mathscr{B}_j(n)| = \left| \sum_{l=1}^{n} p_l \mathscr{B}_{j-l}(n) \right| &\leqslant p_{1_M} |\mathscr{B}_{j-1}(n)| + p_{2_M} |\mathscr{B}_{j-2}(n)| + \ldots + p_{n_M} |\mathscr{B}_{j-n}(n)| \leqslant \\ &\leqslant p_{1_M} \Big(p_{1_M} |\mathscr{B}_{j-2}(n)| + p_{2_M} |\mathscr{B}_{j-3}(n)| + \ldots + p_{n_M} |\mathscr{B}_{j-n-1}(n)| \Big) + \\ &\quad + p_{2_M} |\mathscr{B}_{j-2}(n)| + \ldots + p_{(n-1)_M} |\mathscr{B}_{j-n+1}(n)| + p_{n_M} |\mathscr{B}_{j-n}(n)| = \\ &= \frac{p_{1_M}^2 + p_{2_M}}{p_{1_M}} \sum_{l=1}^{n} c_l |\mathscr{B}_{j-1-l}(n)| \, . \end{aligned} \tag{15}$$

Here

$$c_l = \frac{p_{lM}}{p_{1M}^2 + p_{2M}} \left(p_{1M}^2 + \frac{p_{1M} p_{(l+1)M}}{p_{lM}} \right), \; l = 0, 1, \ldots, n-1; \quad c_n = p_{nM} \frac{p_{1M}^2}{p_{1M}^2 + p_{2M}}.$$

Applying (12) we find: $\dfrac{p_{1M}^2 + p_{2M}}{p_{1M}} = \dfrac{2n-1}{n} p_{1M}$ and $\dfrac{p_{1M} p_{(l+1)M}}{p_{lM}} < p_{2M}$. Consequently, $c_l < p_{lM}$, $l = 0, \ldots, n$. These relations allow us to rewrite the formula (15) as follows

$$|\mathscr{B}_j(n)| \leqslant \sum_{l=1}^{n} p_{lM} |\mathscr{B}_{j-l}(n)| < x \sum_{l=1}^{n} p_{lM} |\mathscr{B}_{j-l-1}(n)|. \tag{16}$$

where x is defined by (14). The last inequality in (16) is recursive. Continuation of this inequality gives: $|\mathscr{B}_j(n)| < x^{j-n} \sum_{l=1}^{n} p_{lM} |\mathscr{B}_{n-l}(n)| = p_{1M} x^{j-n}$, which is equivalent to (13).

□

Remark. For matrix Az, where z some scalar, the inequality (13) remains valid, but the parameter x is defined in this case by the expression

$$x = (2n-1) \max |a_{gl} z| . \tag{17}$$

Corollary 2.1.

$$\left| \sum_{j=N}^{\infty} \frac{1}{j!} C_{jl} \right| < \frac{x^{-l} n}{(2n-1)^{n-l+1}} \sum_{g=0}^{l} \frac{n!}{(l-g)!(2n-1)^g} \sum_{j=N}^{\infty} \frac{x^j}{j!} <$$

$$< \frac{x^{N-l} n (N+1)}{(2n-1)^{n-l+1} N! (N+1-x)} \sum_{g=0}^{l} \frac{n!}{(l-g)!(2n-1)^g}, \quad \text{if} \quad x < N+1. \tag{18}$$

3. Computation of the matrix exponential

According to relations (11) and (9)

$$\exp(Az) \simeq \mathscr{E}(J) \equiv \left\| e_{ik}(J) \right\| = \sum_{l=0}^{n-1} (Az)^l \left[\frac{1}{l!} + \mathscr{C}_l(J) \right], \tag{19}$$

where $J \geqslant n$ and

$$\mathscr{C}_l(J) = \sum_{j=n}^{J} \frac{1}{j!} C_{jl} = \sum_{g=0}^{l} p_{n-l+g} \sum_{j=n}^{J} \frac{1}{j!} \mathscr{B}_{j-1-g}(n) . \tag{20}$$

The approximation in (19) is replaced by the exact equality if J goes to infinity.

The calculation of the matrix exponential by formulas (19)-(20) can be performed using the following algorithm.

1. First, the consecutive computations $1/(2!), \ldots, 1/(J!)$ are done.
2. Powers of the matrix Az are computed from the second to n-th inclusive, and traces of these matrices
$$s_g = \operatorname{tr} (Az)^g, \quad g = 1, \ldots, n$$
are calculated.
3. Sequential calculation of the coefficient p_g, defined by (4) can be performed by Newton's formula [3]:
$$g p_g = s_g - p_1 s_{g-1} - \ldots - p_{g-1} s_g, \; g = 1, \ldots, n.$$
In particular, $p_1 = s_1$, $p_2 = (s_2 - p_1 s_1)/2$, $p_3 = (s_3 - p_1 s_2 - p_2 s_1)/3, \ldots$.

4. After that, the symmetric polynomials $\mathscr{B}_l(n)$ for $l = n, n+1, \ldots, J-1$ are calculated by recurrence formulas (6).
5. The calculation of the sums $\Sigma_g \equiv \sum_{j=n}^{J} \mathscr{B}_{j-1-g}(n)/(j!)$ for $g = 0, 1, \ldots, n-1$.
6. Substitution of the values which were found in the previous steps in the formula (19) and calculation

$$\exp(Az) = I[1 + p_n\Sigma_0] + (Az)[1 + (p_{n-1}\Sigma_0 + p_n\Sigma_1)] + \ldots + \\ + (Az)^{n-1}\left[\frac{1}{(n-1)!} + (p_1\Sigma_0 + p_2\Sigma_1 + \ldots + p_n\Sigma_{n-1})\right].$$

3.1. Estimation of roundoff errors

Multiplication of numbers generates roundoff errors. Therefore, the accuracy of the scattering matrix computations can be estimated by the number of multiplications which are used in the calculations. The six steps listed above require for their implementation the following number of multiplications, respectively: $N_{1_1} = J-1$, $N_{1_2} = n^3(n-1)$, $N_{1_3} = (n-1)(n+2)/2$, $N_{1_4} = n(J-n)$, $N_{1_5} = nJ - n(3n-1)/2$, $N_{1_6} = n^3 + n(n+1)/2$. On the whole $N_1 = (2n+1)J + n^4 - n^2 - 2 - 3n(n-1)/2$.

For comparison, calculations according formula (2)

$$\exp(Az) = I + (Az) + (Az)\frac{(Az)}{2} + (Az)\frac{(Az)}{2}\frac{(Az)}{3} + \ldots + (Az)\frac{(Az)}{2}\frac{(Az)}{3}\cdots\frac{(Az)}{J}$$

require $N_2 = (n^3 + n^2)(J-1)$ multiplications.

It is easy to verify that $N_1 < N_2$ if $J > n$. For these values of J, the computation $\exp(Az)$ by the MSP generates a smaller roundoff error than calculations by formula (2). Asymptotics of the ratio of N_2 to N_1 is characterized by

$$\lim_{J\to\infty}\frac{N_2}{N_1} = \frac{n^3+n^2}{2n+1} > \frac{n^2}{2}\,.$$

3.2. Truncation error

Let us consider error caused by truncation of series

$$\sum_{j=n}^{\infty}\frac{1}{j!}\mathscr{B}_{j-1-g}(n) \approx \sum_{j=n}^{n+N}\frac{1}{j!}\mathscr{B}_{j-1-g}(n)\,,$$

and substitution of exact formula (11) for (19).

D e f i n i t i o n 2. *The relative truncation error* of the matrix exponential is

$$\epsilon(J) = \max\left|\frac{e_{ik}(\infty) - e_{ik}(J)}{e_{ik}(\infty)}\right|. \tag{21}$$

Obviously

$$|e_{ik}(\infty) - e_{ik}(J)| < \left|\sum_{l=0}^{n-1}\frac{(a_{ik}z)^l}{l!}\right| \max\left|l!\sum_{j=J+1}^{\infty}\frac{1}{j!}C_{jl}\right|, \tag{22}$$

and

$$|e_{ik}(\infty)| = \left|\sum_{l=0}^{n-1}\frac{(a_{ik}z)^l}{l!}\right| |1+T| > \left|\sum_{l=0}^{n-1}\frac{(a_{ik}z)^l}{l!}\right| (1-|T|), \tag{23}$$

where

$$|T| = \left| \left(\sum_{l=0}^{n-1} \frac{(a_{ik}z)^l}{l!}\, l!\, \mathscr{C}_l(\infty) \right) \Big/ \left(\sum_{l=0}^{n-1} \frac{(a_{ik}z)^l}{l!} \right) \right| \leqslant \max \left| l! \sum_{j=n}^{\infty} \frac{1}{j!} C_{jl} \right| . \tag{24}$$

For $x < 1$ from (18) and (24) it follows

$$|T| < \frac{x^{n-l} n(n+1)(n-1)!}{(2n-1)^{n-l+1} n!(n+1-x)} \sum_{g=0}^{n-1} \frac{n!}{(n-1-g)!(2n-1)^g} < \frac{1}{2} .$$

From this relation and (23) we find

$$|e_{ik}(\infty)| > \left| \sum_{l=0}^{n-1} \frac{(a_{ik}z)^l}{l!} \right| \frac{1}{2} > \frac{xn(n+1)(n-1)!}{(2n-1)^2\, n!(n+1-x)}\, \Sigma , \tag{25}$$

and from (22) and (18) —

$$|e_{ik}(\infty) - e_{ik}(N+n)| < \frac{n!(N+n+2)x^{N+2}}{(2n-1)^2 (N+n+1)!(N+n+2-x)}\, \Sigma . \tag{26}$$

Here $\Sigma = \left| \sum_{l=0}^{n-1} \frac{(a_{ik}z)^l}{l!} \right| \sum_{g=0}^{n-1} \frac{n!}{(n-1-g)!(2n-1)^g}$.

Finally, substitution of expressions (25) and (26) in the definition (21) gives the proof of the following theorem.

T h e o r e m 3. *If* $\max |a_{gl}z| < 1/(2n-1)$, *the relative truncation error* $\epsilon(N+n)$ *of the matrix exponential* $\exp(Az)$ *satisfies the inequality*

$$\epsilon(N+n) < \frac{n!(N+n+2)x^{N+1}}{(N+n+1)!(N+n+1)} ,$$

where x *is defined by* (17).

3.3. Scaling exponent

Using the fundamental property of the exponential function $\exp A = [\exp(A/m)]^m$, we represent the scattering matrix as follows: $S = X^m$, where

$$X \simeq \sum_{l=0}^{n-1} \left(\frac{Az}{m} \right)^l \frac{1}{l!} \left[1 + l! \sum_{g=0}^{l} p_{n-l+g} \sum_{j=n}^{n+N} \frac{1}{j!} \mathscr{B}_{j-l-g}(n) \right] , \tag{27}$$

m - integer; $p_i = (-1)^{i-1}\sigma_i$; σ_i, $i = 1, 2, \ldots, n$, and $\mathscr{B}_l(n)$ are elementary symmetric polynomials and symmetric polynomials of n-th order of matrix (Az/m), respectively.

C o r o l l a r y 3.1. *The relative truncation error of the matrix* X *calculation according to the formula* (27) $\epsilon < \dfrac{n!(N+n+2)\xi^{N+1}}{(N+n+1)!(N+n+1)}$, *provided that* $\xi = (2n-1)\dfrac{\max |a_{jl}z|}{m} < 1$.

E x a m p l e. Let the scaling factor m for matrix Az of order $n = 4$ is the smallest integer satisfying the condition $m \geqslant 10(2n-1)\max |a_{jl}z|$. In this case $\xi < 0.1$, and the calculation by formula (27) with the value $N = 2$ gives the matrix X with a relative truncation error less than 10^{-5}.

When $m > n + 1$, the most efficient way of calculating the matrix X^m, as shown in [7], is to use the theorem (8):

$$X^m = \sum_{l=0}^{n-1} X^l \sum_{g=0}^{l} (-1)^{n-l+g-1} \sigma_{n-l+g} \mathscr{B}_{m-1-g}(n),$$

where σ_j and $\mathscr{B}_l(n)$ are symmetric polynomials of matrix X.

4. Conclusion

Among the dozens of techniques that were used to calculate the matrix exponential, the method of symmetric polynomials (MSP) has significant advantages. One of the most important of them - is carrying out calculations using analytical estimates. Another is that these estimates do not depend on the eigenvalues of the matrix. Accuracy of the scattering matrix $S = \exp(Az)$ calculation by the method of scaling is determined by the truncation errors in the calculation of matrix X and roundoff errors in the computation of the matrix X^m. MSP allows one to control the first error and minimize the latter.

References

[1] Cowley J. M. *Diffraction Physics*. North-Holland Pub. Co., Amsterdam, 410 pp. (1975).
[2] Pinsker Z. G. *Dynamical scattering of X-rays in perfect crystals*. Springer-Verlag, Heidelberg, 511 pp. (1978).
[3] Gantmacher F. R. *The Theory of Matrices*. Nauka, Moscow, 552 pp. (1988).
[4] Angot A. C*ompléments de mathématiques a l'usage des ingénieurs de l'élektrotechnique et des télécommunications*. Masson, Paris, 868 pp. (1982).
[5] MacDuffee C. C. *The Theory of Matrices*. Chelsea, New York, 128 pp. (1956).
[6] Faddeev D. K., Faddeeva V. N. *Computational methods of linear algebra*. Nauka, Moscow, 656 pp. (1963).
[7] Belyayev Yu. N. Calculations of transfer matrix by means of symmetric polynomials. Proceedings of the International Conference "Days on Diffraction 2012", St.Petersburg, Russia May 28 – June 1, 2012. P. 36–41.