Orthogonal polynomial expansion for the Lowdin transformation

1989 J. Phys. A: Math. Gen. 22 2675

(http://iopscience.iop.org/0305-4470/22/14/018)

J. Phys. A: Math. Gen. 22 (1989) 2675-2685. Printed in the UK

Orthogonal polynomial expansion for the Löwdin transformation

Yorck Leschber† and J P Draayer‡

† Physics Department, Glasgow University, Glasgow G12 8QQ, UK
‡ Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA 70803, USA

Received 14 November 1988

Abstract. A set of linearly independent functions is orthogonalised by means of the Löwdin transformation. In this paper, a simple method is presented for calculating the orthogonalisation as a polynomial of finite degree in the overlap matrix S. The method is illustrated for various cases.

1. Introduction

A familiar problem in quantum mechanics is to expand a function χ in terms of a basis spanned by a set of functions Φ_i, $i = 1, \ldots, n$, such that $\chi = \Sigma\, a_i \Phi_i$. However, in many cases the natural basis set is not orthogonal but has overlap integrals of the form

$$\Delta_{ij} = \int \Phi_i \Phi_j \, d\tau = \delta_{ij} + S_{ij}. \tag{1.1}$$

Examples of non-orthogonal functions can be found in different fields of physics. In nuclear physics, a model based on the group SU(3) provides a microscopic understanding of rotational motion [1, 2]. The basis states are labelled as

$$\psi_{SU(3)} = |(\lambda, \mu) KLM\rangle. \tag{1.2}$$

Here (λ, μ) are the quantum numbers associated with an irreducible representation (irrep) of SU(3), L is the value of the orbital angular momentum and M is its z component. There exist certain rules to deduce the values of L that are compatible with an irrep of SU(3). In general a state with a particular value of L can occur several times. The multiplicity is resolved by the label K. The simplest way is to define the states of the SU(3) model by a Hill-Wheeler projection [1]. However, it turns out that they are not orthogonal with respect to the quantum number K. Another field where non-orthogonality enters is molecular physics. Consider the case of the hydrogen molecule (H_2) which was first treated by Heitler and London [3]. They evaluated a simple two-body Hamiltonian at fixed values of the internuclear distances (the Born-Oppenheimer approximation). The wavefunction was taken as an antisymmetrised product (Slater determinant) of two 1s hydrogen wavefunctions centred about the nuclei A and B, respectively. The orbital part of these functions are in general not orthogonal but have an overlap as defined in (1).

0305-4470/89/142675+11$02.50

In principle, all calculations can be done in a non-orthogonal scheme; however, many formulae are much simpler if an orthogonal basis is used. There are different methods to generate an orthogonal function system. The best known is probably the Gram-Schmidt procedure. Denoting the original functions by Φ_i and the orthogonal functions by ϕ_i, the Gram-Schmidt method may be summarised as

$$\phi_{n+1} = \Phi_{n+1} c_{n+1,n+1} - \sum_{k=1}^{n} \phi_k c_{k,n+1} \tag{1.3}$$

$$c_{n+1,n+1} = \frac{1}{\sqrt{a_{n+1}}} \qquad c_{k,n+1} = \frac{C_{k,n+1}}{\sqrt{a_{n+1}}} \tag{1.4}$$

$$a_{n+1} = \Delta_{n+1,n+1} - \sum_{k=1}^{n} C_{n+1,k} C_{k,n+1} \tag{1.5}$$

$$C_{k,n+1} = \int \phi_k^* \Phi_{n+1} \, d\tau. \tag{1.6}$$

The first function is simply $\phi_i = \Phi_i / \sqrt{\Delta_{11}}$. The other ϕ_i are generated successively by adding Φ_i and orthogonalising against all the previous functions. Clearly the explicit form of the orthogonal basis depends on the order in which the non-orthogonal functions are used.

Instead Löwdin defined a symmetric transformation between the two bases [4]

$$\phi_i = \sum_{k=1}^{n} \Phi_k \Delta_{ki}^{-1/2}. \tag{1.7}$$

It is therefore necessary to find an expression for $\Delta^{-1/2} = (\mathbf{1} + \boldsymbol{S})^{-1/2}$. If the matrix elements of $\boldsymbol{S}$ are small, that is if

$$\sum_q |\boldsymbol{S}_{kq}| < 1 \tag{1.8}$$

then the following Taylor series converges:

$$(\mathbf{1} + \boldsymbol{S})^{-1/2} = \mathbf{1} - \tfrac{1}{2}\boldsymbol{S} + \tfrac{3}{8}\boldsymbol{S}^2 + - \ldots. \tag{1.9}$$

If the convergence criterion is not satisfied then another matrix σ can be defined

$$\mathbf{1} + \boldsymbol{S} = (1 + s_{\max})(\mathbf{1} - \sigma) \qquad \sigma = \frac{s_{\max}}{1 + s_{\max}} (\mathbf{1} - s_{\max}^{-1} \boldsymbol{S}). \tag{1.10}$$

Here $s_{\max}$ is the largest eigenvalue of $\boldsymbol{S}$. Now a convergent power series in σ has the form

$$(\mathbf{1} + \boldsymbol{S})^{-1/2} = (1 + s_{\max})^{-1/2} (1 + \tfrac{1}{2}\sigma + \tfrac{3}{8}\sigma^2 + \ldots). \tag{1.11}$$

Therefore, in one way or another, it is possible to define the Löwdin transformation in terms of an *infinite* power series in the overlap matrix $\boldsymbol{S}$.

A second realisation of the Löwdin transformation, which avoids the Taylor series expansion, is based on the fact that the similarity transformation that diagonalises Δ also diagonalises $\Delta^{-1/2}$. If we denote the diagonal matrix containing the eigenvalues of Δ by $\boldsymbol{d}$, then

$$\boldsymbol{d} = \boldsymbol{U}^{-1} \Delta \boldsymbol{U}. \tag{1.12}$$

The matrix $\boldsymbol{U}$ contains the eigenvectors. It is trivial to calculate $\boldsymbol{d}^{-1/2}$ because it is diagonal and subsequently

$$\Delta^{-1/2} = \boldsymbol{U}\boldsymbol{d}^{-1/2}\boldsymbol{U}^{-1}. \tag{1.13}$$

If the Löwdin transformation is calculated in this way one needs to know both the eigenvalues and the eigenvectors of $\boldsymbol{S}$. In this paper we will discuss an algorithm where $\Delta^{-1/2}$ is defined as a polynomial of *finite* degree in the overlap matrix and which only requires the eigenvalues.

2. Mathematical background

The method presented here was previously developed in the context of finite group transformations. We will now briefly discuss its mathematical foundation [5]. As we have seen, the Löwdin transformation can be expressed as an infinite power series in the overlap matrix. Consider now the Cayley-Hamilton theorem which states that a matrix satisfies its own eigenvalue equation. Therefore, if $\boldsymbol{S}$ is an $n \times n$ matrix then terms like $\boldsymbol{S}^k$, $k \geqslant n$, can be expressed as a power series in $\boldsymbol{S}$ of degree $r < n$. As a consequence, the infinite power series is replaced by a finite one:

$$(\mathbf{1}+\boldsymbol{S})^{-1/2} = \sum_{k=0}^{n-1} b_k \boldsymbol{S}^k. \tag{2.1}$$

Over the years many methods for calculating the coefficients in this expansion have been developed [6] but most of these algorithms are not very attractive to the practitioner. However, a much simpler form arises if one uses a basis of orthogonal matrix polynomials rather than the monomial basis $\boldsymbol{S}^k$.

Consider the form

$$\kappa(\boldsymbol{X}, \boldsymbol{Y}) = \frac{1}{n}\,\mathrm{Trace}(\boldsymbol{X}^+\boldsymbol{Y}). \tag{2.2}$$

Orthogonal polynomials $P_k(\boldsymbol{S})$ of degree k are defined by

$$(D_k D_{k-1})^{-1/2} P_k(\boldsymbol{S}) = \begin{vmatrix} M_0 & M_1 & \cdots & M_k \\ M_1 & M_2 & \cdots & M_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ M_{k-1} & M_k & \cdots & M_{2k-1} \\ \mathbf{1} & \boldsymbol{S} & \cdots & \boldsymbol{S}^k \end{vmatrix}. \tag{2.3}$$

The moments M_k are related to the traces of $\boldsymbol{S}^k$

$$M_k = \frac{1}{n}\,\mathrm{Trace}(\boldsymbol{S}^k). \tag{2.4}$$

The normalisation factors D_k are calculated by replacing $\boldsymbol{S}^m$ by M_{m+k} in the determinant. The orthonormality of the polynomials is expressed as

$$\kappa[P_k(\boldsymbol{S}), P_l(\boldsymbol{S})] = \frac{1}{n}\,\mathrm{Trace}[P_k^+(\boldsymbol{S})P_l(\boldsymbol{S})] = \delta_{kl}. \tag{2.5}$$

It can also be shown that

$$P_k(\boldsymbol{S}) = 0 \qquad k \geqslant n. \tag{2.6}$$

This result follows from the Cayley-Hamilton theorem in combination with the determinant properties. The (q, k) element in the determinant is essentially $\boldsymbol{S}^{q+k}$ which, for $k \geqslant n$, can be rewritten as $\Sigma\, a_r \boldsymbol{S}^{q+k-n+r}$. This result holds also for the moments, because the trace of the sum is equal to the sum of the traces. The coefficients are independent of the row q and therefore every element in the kth column can be expressed as a linear combination of the previous n columns and so the determinant vanishes.

Using these results it follows that the Löwdin transformation can be written as

$$(\mathbf{1}+\boldsymbol{S})^{-1/2} = \sum_{k=0}^{n-1} c_k P_k(\boldsymbol{S}). \tag{2.7}$$

The coefficients are determined using the orthogonality of the polynomials

$$c_k = \kappa[P_k^+(\boldsymbol{S}), (\mathbf{1}+\boldsymbol{S})^{-1/2}] = \frac{1}{n}\,\mathrm{Trace}[P_k^+(\boldsymbol{S})(\mathbf{1}+\boldsymbol{S})^{-1/2}]. \tag{2.8}$$

The c_k depend only on the conjugacy class of $\boldsymbol{S}$, that is they are invariant under similarity transformations. If we denote $\boldsymbol{R} = \boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^{-1}$ as the conjugate of $\boldsymbol{S}$ then

$$\begin{aligned}(\mathbf{1}+\boldsymbol{R})^{-1/2} &= \boldsymbol{C}(\mathbf{1}+\boldsymbol{S})^{-1/2}\boldsymbol{C}^{-1} \\ &= \boldsymbol{C}\sum_{k=0}^{n-1} c_k P_k(\boldsymbol{S})\boldsymbol{C}^{-1} \\ &= \sum_{k=0}^{n-1} c_k P_k(\boldsymbol{R}).\end{aligned} \tag{2.9}$$

The coefficients can be calculated for the simplest possible form of the matrix $\boldsymbol{R} = \boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^{-1}$ which will be in the cases considered here the diagonal form.

3. Applications

As the first example, consider the two-centre problem which is relevant in molecular physics as well as in quark-nuclear physics [7]. For the moment we do not consider any explicit form for the wavefunctions, other than that they are centred about two points A and B which are a distance R apart. The overlap integral and matrix are

$$\varepsilon = \int \Phi_A \Phi_B \,\mathrm{d}\tau \qquad \boldsymbol{S} = \begin{bmatrix} 0 & \varepsilon \\ \varepsilon & 0 \end{bmatrix}. \tag{3.1}$$

The orthogonal polynomials are derived from this 2×2 matrix

$$P_0(\boldsymbol{S}) = \mathbf{1} \qquad P_1(\boldsymbol{S}) = \frac{1}{\varepsilon}\boldsymbol{S}. \tag{3.2}$$

The coefficients are determined by using the diagonal matrix $\boldsymbol{R} = \mathrm{diag}(+\varepsilon, -\varepsilon)$

$$c_0 = \frac{1}{2\sqrt{1-\varepsilon^2}}(\sqrt{1+\varepsilon}+\sqrt{1-\varepsilon}) \qquad c_1 = \frac{1}{2\sqrt{1-\varepsilon^2}}(\sqrt{1-\varepsilon}-\sqrt{1+\varepsilon}). \tag{3.3}$$

With these results the Löwdin transformation $\Delta^{-1/2} = c_0 P_0(\boldsymbol{S}) + c_1 P_1(\boldsymbol{S})$ is

$$(\mathbf{1}+\boldsymbol{S})^{-1/2} = \frac{1}{2\sqrt{1-\varepsilon^2}}\begin{bmatrix} \sqrt{1+\varepsilon}+\sqrt{1-\varepsilon} & \sqrt{1-\varepsilon}-\sqrt{1+\varepsilon} \\ \sqrt{1-\varepsilon}-\sqrt{1+\varepsilon} & \sqrt{1+\varepsilon}+\sqrt{1-\varepsilon} \end{bmatrix}. \tag{3.4}$$

This is of course a well known result [8].

To illustrate the effects of the orthogonalisation we choose two simple Gaussians as basis functions

$$\Phi_A = \frac{1}{\sqrt{a^3\pi^{3/2}}} \exp\left(-\frac{(\boldsymbol{r}-\boldsymbol{A})^2}{2a^2}\right) \qquad \Phi_B = \frac{1}{\sqrt{a^3\pi^{3/2}}} \exp\left(-\frac{(\boldsymbol{r}-\boldsymbol{B})^2}{2a^2}\right). \tag{3.5}$$

If the distance between the centres is R then the overlap integral has the value

$$\varepsilon = \exp\left(-\frac{R^2}{4a^2}\right). \tag{3.6}$$

In figure 1 the non-orthogonal functions are compared with the orthogonal ones. For completeness we have also included the results of the Gram–Schmidt procedure. For large separations, the orthogonalisation has little effect. As the centres approach one another the form of the orthogonal functions depends strongly on the method used.

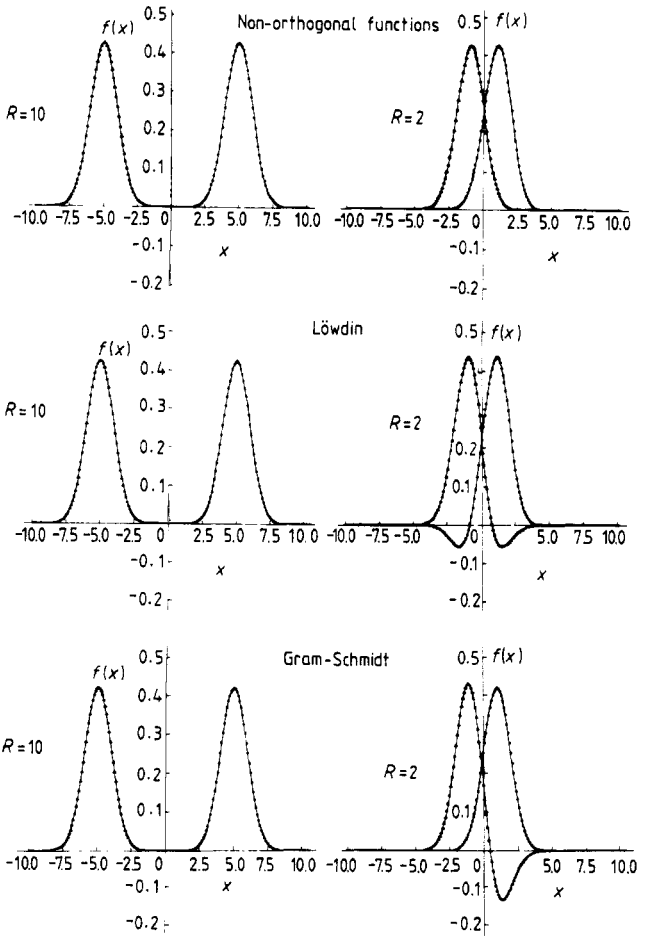

Figure 1. The effects of the Löwdin and Gram–Schmidt orthogonalisation are shown for two Gaussian functions. The original functions which are centred on the x axis at $x = \pm R/2$ are shown in the top part as a function of x. The orthogonalised functions are shown beneath. In each case the calculations were done for two different centre separations, $R = 2$ and $R = 10$.

In these graphs it appears that the Löwdin transformation changes the original functions much less than the Gram-Schmidt method. This is more than a casual impression. It has been proved that the Löwdin procedure is the best scheme inasmuch as it changes the original functions as little as possible [9].

As the second example we choose the three-centre problem which as been studied with regard to the non-additivity of molecular forces [10]. If we denote the overlaps S_{12}, S_{13} and S_{23} by α, β and γ, respectively, then the 3×3 overlap matrix is

$$\boldsymbol{S}=\begin{bmatrix} 0 & \alpha & \beta \\ \alpha & 0 & \gamma \\ \beta & \gamma & 0 \end{bmatrix}. \tag{3.7}$$

The orthogonal polynomials are

$$P_0(\boldsymbol{S})=\mathbf{1} \qquad P_1(\boldsymbol{S})=\frac{1}{\sqrt{M_2}}\boldsymbol{S} \qquad P_2(\boldsymbol{S})=\frac{M_2\boldsymbol{S}^2-M_3\boldsymbol{S}-M_2^2\mathbf{1}}{\sqrt{M_2(\frac{1}{2}M_2^3-M_3^2)}}. \tag{3.8}$$

Since $\boldsymbol{S}$ is symmetric and real its eigenvalues are real. Because the matrix is traceless it has the secular equation

$$\lambda_\alpha^3-q\lambda_\alpha-r=0. \tag{3.9}$$

The quantities q and r are given by

$$q=\tfrac{1}{2}\,\mathrm{Trace}(\boldsymbol{S}^2)=\alpha^2+\beta^2+\gamma^2 \tag{3.10}$$

$$r=\tfrac{1}{3}\,\mathrm{Trace}(\boldsymbol{S}^3)=2\alpha\beta\gamma. \tag{3.11}$$

The roots of a third-order polynomial are determined using the standard formulae [11]

$$\lambda_\rho=2\sqrt{\frac{q}{3}}\cos\left(\frac{\theta-2\pi\rho}{3}\right) \tag{3.12}$$

with $\rho=1, 2, 3$ and

$$\theta=\tan^{-1}\left(\frac{4q^3}{27r^2}-1\right)^{1/2}. \tag{3.13}$$

For the explicit calculations of the coefficients we again choose Gaussian functions centred about three points A, B and C. All possible configurations can be described by a triangle but we consider only isosceles triangles in the xy plane. The centres A and B are fixed at $\pm x/2$ and the centre C is at the coordinates $(0, y, 0)$. The coefficients c_k are calculated for different values of y. The results are shown in figure 2. The curves are symmetric about $y=0$, which describes a linear arrangement of the three centres. If the centres form an equilateral triangle then the coefficient c_2 is zero. In that case the overlaps are all equal and $\boldsymbol{S}$ has a double root at $-\alpha$, and as a consequence $\boldsymbol{S}^2$ is not independent of $\mathbf{1}$ and $\boldsymbol{S}$. For small deviations from $y=x$ we define $\varepsilon=\alpha-\beta=\alpha-\gamma$ and expand the c_2 in ε. The result is

$$\lim_{\varepsilon\to 0} c_2=\frac{1}{3}\sqrt{\frac{2}{27}}\left(\frac{2}{\sqrt{1+2\alpha}}+\frac{5\alpha-2}{\sqrt{(1-\alpha)^3}}\right)\frac{\varepsilon}{\alpha}. \tag{3.14}$$

The Löwdin transformation is calculated for the different triangular arrangements and compared with the results of a Taylor expansion. In order to learn something about the convergence of the latter we consider a second- and a tenth-order expansion. As

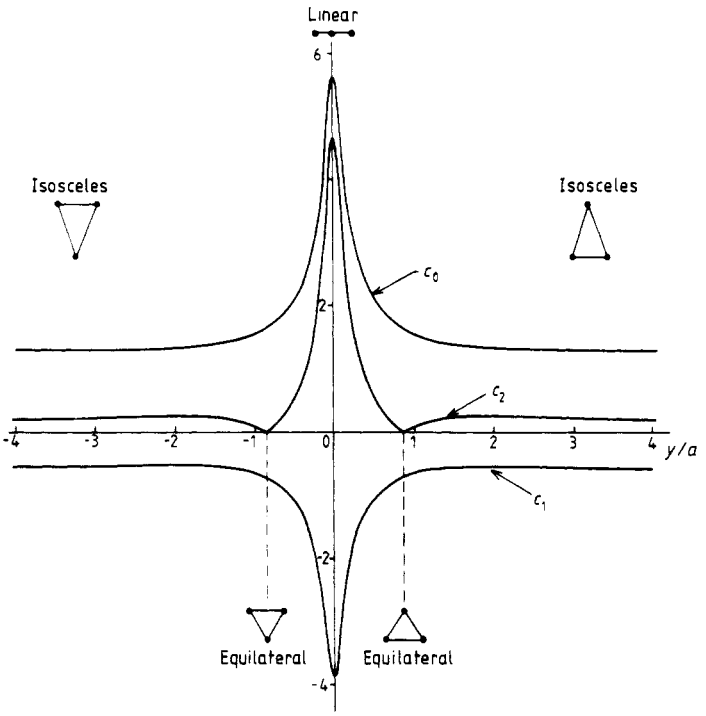

Figure 2. The coefficients c_1, c_2, and c_3 for the three-centre problem are calculated for triangular arrangements where centres A and B are fixed on the x axis at $\pm\frac{1}{2}$ (size parameter $a = 1$). The third centre C is positioned on the y axis. In general the centres form an isosceles triangle. If centre C is at $y = \pm(3/4)^{0.5}$ then the centres form an equilateral triangle and c_2 vanishes, as described in the text.

measures for comparisons we choose the centroid and the variance of the resulting matrices. They are defined by [12]

$$\varepsilon = \frac{1}{N}\,\mathrm{Trace}(\Delta^{-1/2}) \tag{3.15}$$

$$\sigma^2 = \frac{1}{N}\,\mathrm{Trace}[(\Delta^{-1/2} - \varepsilon)^2]. \tag{3.16}$$

The results of these comparisons are shown in figures 3 and 4 respectively. Both graphs indicate that the convergence of the Taylor series is rather poor as soon as the overlap between the wavefunctions of the three centres becomes comparable. This shows how much redundancy can be involved in using the Taylor series. Although the entire matrix function depends only on $\mathbf{1}$, $\boldsymbol{S}$ and $\boldsymbol{S}^2$, the Taylor series may need many more terms to give a good result.

4. A recursive algorithm for generating the matrix polynomials and evaluating the matrix function

As long as the matrices involved are small the easiest way to proceed is to construct the matrix polynomials explicitly using Szegö's determinant. We present in this paragraph an alternative method for a recursive generation of the polynomials. In that

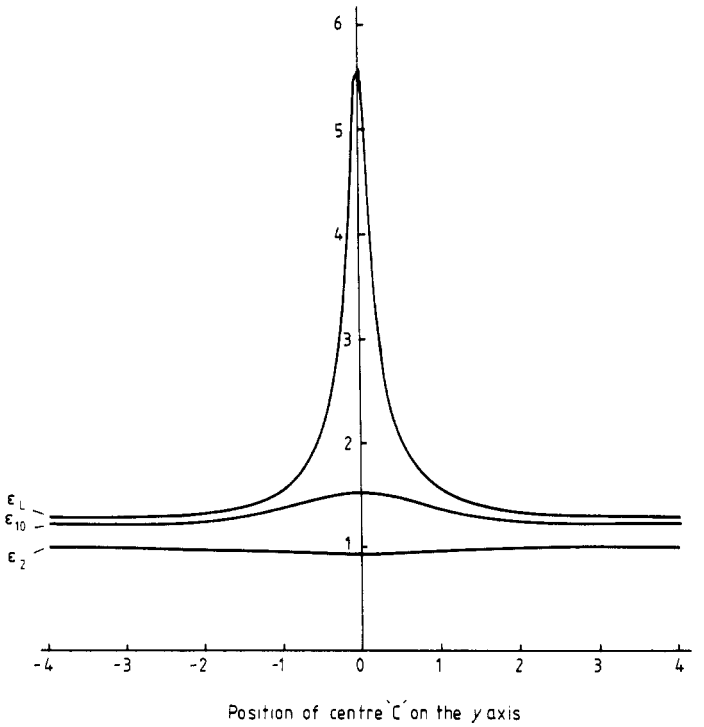

Figure 3. The centroids of the exact Löwdin transformation are compared with those from a second- and tenth-order Taylor expansion. The symbol ε_L refers to the centroid of the exact Löwdin transformation while ε_{10} and ε_2 are the results for the appropriate Taylor expansion.

way one avoids calculating a large number of determinants. As shown for example in [13] the polynomials from Szegö's determinant satisfy a three-term recursion relation:

$$\boldsymbol{S}P_k(\boldsymbol{S}) = \beta_{k-1}P_{k-1}(\boldsymbol{S}) + \alpha_k P_k(\boldsymbol{S}) + \beta_k P_{k+1}(\boldsymbol{S}). \tag{4.1}$$

Together with $P_0 = \mathbf{1}$ and $P_{-1} = 0$ this specifies the orthogonal matrix polynomials. The coefficients are determined from

$$\alpha_k = \frac{1}{N}\,\mathrm{Trace}[\boldsymbol{S}P_k^2(\boldsymbol{S})] \tag{4.2}$$

$$\beta_k^2 = \frac{1}{N}\,\mathrm{Trace}[(\boldsymbol{S} - \alpha_k\mathbf{1})P_k(\boldsymbol{S}) - \beta_{k-1}P_{k-1}(\boldsymbol{S})]^2. \tag{4.3}$$

We recognise that formula (4.1) together with (4.2) and (4.3) describe essentially the Lanczos algorithm [14]. The difference is that here a series of matrices rather than vectors is generated. In order to calculate the expansion coefficients c_i the polynomials are required in their diagonal forms. These are generated by replacing in (4.1) the matrix $\boldsymbol{S}$ by the matrix $\boldsymbol{R}$ which contains the eigenvalues of $\boldsymbol{S}$. Now the formulae (4.1)-(4.3) are equivalent to those arising in Forsythe's [15] algorithm to determine orthogonal polynomials over a discrete set of points. After the coefficients have been calculated the full matrix function can be evaluated. This exercise is simplified by

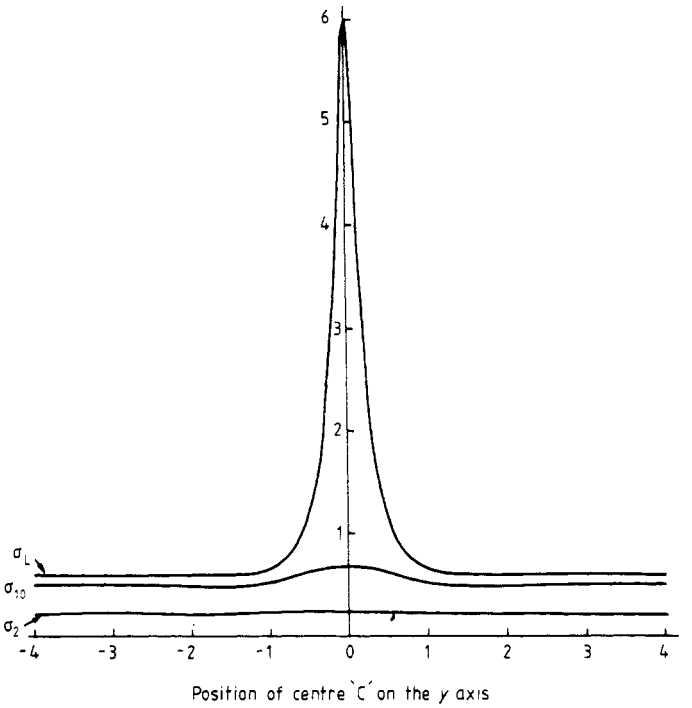

Figure 4. The standard deviations of the Löwdin transformation are compared with those from a second- and tenth-order Taylor expansion. The notation is as in figure 3.

using Clenshaw's recurrence formula [16] (see also [11] p 143). In general we are concerned with a summation of the kind

$$f(\boldsymbol{S}) = \sum_{k=0}^{M} c_k P_k(\boldsymbol{S}) \tag{4.4}$$

where $M \leqslant N-1$, with N the order of the matrix. Between the P_k there exist recurrence relations of the type

$$P_{k+1}(\boldsymbol{S}) = A_k(\boldsymbol{S}) P_k(\boldsymbol{S}) + B_k(\boldsymbol{S}) P_{k-1}(\boldsymbol{S}). \tag{4.5}$$

The coefficients A_k and B_k are matrices themselves. Their explicit form can easily be determined from (4.1). Provided that the quantities c_k are known then the matrices Y_k are defined by

$$Y_{M+1} = Y_{M+2} = 0 \tag{4.6}$$

$$Y_k = A_k(\boldsymbol{S}) Y_{k+1} + B_{k+1}(\boldsymbol{S}) Y_{k+2} + c_k \mathbf{1}. \tag{4.7}$$

Here the index k runs from $M, M-1, \ldots, 1$. Each of these equations is solved for c_k and inserted into the summation formula. Finally, the function $f(\boldsymbol{S})$ can be written as

$$f(\boldsymbol{S}) = [B_1(\boldsymbol{S}) Y_2 + c_0 \mathbf{1}] P_0(\boldsymbol{S}) + Y_1 P_1(\boldsymbol{S}). \tag{4.8}$$

Effectively only the two simplest matrix polynomials are required explicitly. The dependence of $f(\boldsymbol{S})$ on higher powers of $\boldsymbol{S}$ is contained in the Y_1 and Y_2.

We used the formalism presented here to calculate the Löwdin transformation for systems with N vectors ($N = 5, 10, \ldots, 25$). The vectors were generated at random and normalised to unity and the overlap matrix was calculated by taking the scalar products between the different vectors. In order to avoid any problems with a singularity of the Löwdin transformation the matrix elements of the overlap matrix were multiplied by 0.95. The results are shown in table 1. The accuracy of the Löwdin transform was checked by calculating the quantity

$$\boldsymbol{d} = (\mathbf{1}+\boldsymbol{S})^{-1/2}(\mathbf{1}+\boldsymbol{S})(\mathbf{1}+\boldsymbol{S})^{-1/2}. \tag{4.9}$$

The result should be the identity matrix with centroid 1 and width 0. In columns 4 and 5 of table 1 the numerical results are shown in the different cases.

Table 1. The results of the Löwdin transformation for systems of different sizes. The first column gives the dimension of the matrix, the second and third columns the centroid and width of the Löwdin transformation as calculated from the eigenvalues of $\boldsymbol{S}$. The fourth and fifth columns give the same measures for the form $\boldsymbol{d} = (\mathbf{1}+\boldsymbol{S})^{-0.5}(\mathbf{1}+\boldsymbol{S})(\mathbf{1}+\boldsymbol{S})^{-0.5}$ which should be compared with the results of the identity matrix.

N	ε_L	σ_L	ε_d	σ_d
5	1.6975	1.4151	1.0000	5.5511 E − 17
10	1.6056	1.1470	1.0000	3.0405 E − 17
15	1.7006	1.3145	1.0000	7.3084 E − 17
20	1.5712	1.1389	1.0000	8.5998 E − 17
25	1.5999	1.1274	1.0000	1.0878 E − 16

5. Conclusion

We have presented an alternative method for calculating matrix polynomials which is based on the Cayley-Hamilton theorem and the use of orthogonal polynomials. The coefficients in the expansion are determined by simple trace operations. We gave two different methods to generate the polynomial basis, a direct one based on Szegö's determinant and a second recursive method which we found more useful for a computational implementation.

References

[1] Elliott J P 1958 *Proc. R. Soc.* A **245** 128, 562
[2] Leschber Y and Draayer J P 1987 *Phys. Lett.* **190B** 1
[3] Heitler W and London F 1927 *Z. Phys.* **44** 455
[4] Löwdin P O 1956 *Adv. Phys.* **5** 1
[5] Leschber Y, Draayer J P and Rosensteel G 1986 *J. Phys. A: Math. Gen.* **19** L307
[6] Rinehart R F 1955 *Am. Math. Mon.* **62** 395
[7] Harvey M 1981 *Nucl. Phys.* A **352** 326
[8] Kotani M, Ohno K and Kayama K 1961 *Handbuch der Physik*, vol 37/2 (Berlin: Springer) pp 138-40
[9] Aiken J, Jonassen H and Aldrich H 1975 *J. Chem. Phys.* **62** 2745
[10] Rosen P 1953 *J. Chem. Phys.* **21** 1007

[11] Press W H, Flannery B P, Teukolsky S A and Vetterling W T 1986 *Numerical Recipes* (Cambridge: Cambridge University Press) p 146
[12] Wong S S M 1986 *Nuclear Statistical Spectroscopy* (Oxford: Oxford University Press)
[13] Akhiezer N I 1965 *The Classical Moment Problem* (London: Oliver and Boyd)
[14] Parlett B N 1980 *The Symmetric Eigenvalue Problem* (Englewood Cliffs, NJ: Prentice-Hall)
[15] Forsythe G E 1957 *J. Soc. Indust. Appl. Math.* **5** 74
[16] Clenshaw C W 1955 Mathematics Tables **9** 118