Theor Chem Acc (2012) 131:1109
DOI 10.1007/s00214-012-1109-y

REGULAR ARTICLE

Mayer's orthogonalization: relation to the Gram-Schmidt and Löwdin's symmetrical scheme

Péter R. Nagy · Péter R. Surján · Ágnes Szabados

Received: 6 August 2011 / Accepted: 13 October 2011 / Published online: 8 February 2012

Abstract A method introduced by Mayer (Theor Chem Acc 104:163, 2000) for generating an orthogonal set of basis vectors, perpendicular to an arbitrary start vector, is examined. The procedure provides the complementary vectors in closed form, expressed with the components of the start vector. Mayer's method belongs to the family of orthogonalization schemes, which keep an arbitrary vector intact without introducing any non-physical sequence-dependence. It is shown that Mayer's orthogonalization is recovered by performing a two-step combination of the Gram-Schmidt and Löwdin's symmetrical orthogonalization. Processor time requirement of constructing Mayer's orthonormal set is proportional to $\sim N^2$, in contrast to the rough $\sim N^3$ CPU requirement of performing either a full Gram-Schmidt or Löwdin's symmetrical orthogonalization. Utility of Mayer's orthogonalization is demonstrated on an electronic structure application using perturbation theory to improve multiconfigurational wavefunctions.

Keywords Orthogonalization · Mayer vectors

1 Introduction

Orthonormality of basis vectors is a practical need when solving linear algebraic problems inspired by physics or chemistry, since it yields the expressions in their most simple form. Given an overlapping but non-redundant vector set in an $N+1$-dimensional Euclidean space

$$\{\mathbf{v}_i\}_{i=0}^{N}$$

with

$$S_{ij} = \mathbf{v}_i^T \mathbf{v}_j \neq \delta_{ij},$$
$$\det \mathbf{S} \neq 0,$$

we wish to obtain an orthonormal (ON) set

$$\{\mathbf{u}_i\}_{i=0}^{N}$$

satisfying

$$\mathbf{u}_i^T \mathbf{u}_j = \delta_{ij}.$$

Column vectors are denoted by small, boldface letters and matrices by capital boldface letters. Italic letters are used to represent scalars.

A straightforward procedure leading to a set of $\mathbf{u}_i$'s starting from $\mathbf{v}_i$'s is to perform successive projections, that is, the Gram-Schmidt scheme. An alternative method, frequently applied in problems of chemical physics, is the symmetric orthogonalization introduced by Löwdin [1] (also attributed to Landshoff and Wannier [2, 3]), which transforms the overlapping set to an orthonormal one by matrix $\mathbf{S}^{-1/2}$. There exist numerous further orthogonalization procedures, like Löwdin's canonical [4] (also known as Scheinler-Wigner method [5]), those arising from the Householder or Givens rotation schemes [6] or the approach recently suggested by Chaturvedi [7]. Actual choice of the technique used in a given problem is governed by the specific properties of orthogonalization methods. If one wishes to keep an element of the initial set, say $\mathbf{v}_0$, intact, Gram-Schmidt orthogonalization or Householder transformation may be the method of choice. The resemblance theorem applying to Löwdin's symmetric orthogonalization [8, 9] may be of advantage when conservation of the original vectors to the maximum possible extent is of interest.

P. R. Nagy · P. R. Surján · Á. Szabados (✉)
Laboratory of Theoretical Chemistry, Eötvös University, POB 32, Budapest 1518, Hungary
e-mail: szabados@chem.elte.hu

Recently, an orthogonalization procedure was obtained by Mayer [10, 11] for the special case where the overlapping set is provided by N orthonormal vectors and an additional single vector, non-orthogonal to the previous ones. In practice, such a situation arises when a start vector is available, expanded on an $N+1$ dimensional orthonormal set. A typical task then is to generate N orthonormal vectors, orthogonal to the start vector. A completely analytical solution to this problem was described by Mayer, a technique termed Mayer-orthogonalization hereafter. Relation of Mayer's procedure to the generally applicable orthogonalization schemes has not been explored. Closer inspection reveals that Mayer-orthogonalization shows combined properties, characteristic to the Gram-Schmidt and to Löwdin's orthogonalization.[1] The former is manifested by the fact that $\mathbf{v}_0$ is conserved. An impression of the latter is conveyed by observing the expression of Mayer-orthogonalized vectors (vide infra), which are clearly treated on an equal footing. In addition, numerical experience shows that Mayer-orthogonalized vectors exhibit great similarity to the basis vectors, used to expand $\mathbf{v}_0$, among certain circumstances.

As proved below, Mayer's orthogonalization is in fact equivalent to a two-step combination of the Gram-Schmidt and Löwdin's method. Since Löwdin orthogonalization involves the construction of an inverse square root matrix, it is remarkable that Mayer-orthogonalized vectors can be expressed in a closed form. We show here that due to the simple form of the overlap matrix occurring in the Löwdin-orthogonalization step, its inverse square root can be constructed analytically.

2 Theory

2.1 Mayer's orthogonalization

In a study related to Jacobi-rotations, Mayer was lead to the matrix [10]

$$\mathbf{U}=\begin{pmatrix} v_0 & -v_1 & \dots & \dots & -v_i & \dots & -v_N \\ v_1 & 1-\frac{v_1^2}{1+v_0} & -\frac{v_1v_2}{1+v_0} & \dots & -\frac{v_1v_i}{1+v_0} & \dots & -\frac{v_1v_N}{1+v_0} \\ \vdots & -\frac{v_2v_1}{1+v_0} & \ddots & & -\frac{v_2v_i}{1+v_0} & & \vdots \\ \vdots & \vdots & & \ddots & \vdots & & \vdots \\ v_j & -\frac{v_jv_1}{1+v_0} & -\frac{v_jv_2}{1+v_0} & \dots & -\frac{v_jv_i}{1+v_0} & \dots & -\frac{v_jv_N}{1+v_0} \\ \vdots & \vdots & & & \vdots & \ddots & \vdots \\ v_N & -\frac{v_Nv_1}{1+v_0} & \dots & \dots & -\frac{v_Nv_i}{1+v_0} & \dots & 1-\frac{v_N^2}{1+v_0} \end{pmatrix}, \quad (1)$$

with $i \neq j$. In Mayer's original formulation, matrix $\mathbf{U}$ is expressed with rotation angles $\{\varphi_i\}_{i=0}^N$. These quantities are in one to one mapping with $\{v_i\}_{i=0}^N$ via the relations $v_0 = \cos\varphi, v_i = \varphi_i \sin(\varphi)/\varphi$ and $\varphi^2 = \sum_{i=1}^N \varphi_i^2$.

With scalars $v_0, \dots, v_N$ satisfying

$$\sum_{i=0}^{N} v_i^2 = 1,$$

matrix $\mathbf{U}$ is unitary

$$\mathbf{U}^T\mathbf{U} = \mathbf{I},$$

where $\mathbf{I}$ stands for the identity matrix. As a consequence, columns of $\mathbf{U}$ can be considered as vectors forming an ON set. Basis vectors used for representing matrix $\mathbf{U}$ denoted by $\{\mathbf{e}_i\}_{i=0}^N$, the first element of the ON set can be written as

$$\mathbf{u}_0 = \sum_{i=0}^{N} \mathbf{e}_i\, v_i. \quad (2)$$

Further elements $\{\mathbf{u}_i\}_{i=1}^N$ of the ON set are given by the expression

$$\mathbf{u}_i = -\mathbf{e}_0\, v_i + \sum_{j=1}^{N} \mathbf{e}_j \left(\delta_{ij} - \frac{v_i v_j}{1+v_0}\right), \quad i = 1, \dots, N. \quad (3)$$

Orthonormality $\mathbf{u}_i^T\ \mathbf{u}_j = \delta_{ij}$ is straightforward to check, utilizing $\mathbf{e}_i^T\ \mathbf{e}_j = \delta_{ij}$.

Since all elements of the set $\{\mathbf{u}_i\}_{i=1}^N$ are expressed with the components of vector $\mathbf{u}_0$, the above construction can be viewed as a technique for generating an ON and complementary vector set via Eq. (3) to a normalized start vector, Eq. (2). Note that the start vector $\mathbf{u}_0$ may be selected as any vector of an initial overlapping set $\{\mathbf{v}_i\}_{i=0}^N$. Apart from this single vector, other $\mathbf{v}_i$'s play no role when constructing matrix $\mathbf{U}$.

2.2 Relation to the Gram-Schmidt and Löwdin's symmetrical scheme

A peculiar feature of Mayer's orthogonalization is that a single vector

$$\mathbf{v}_0 = \mathbf{u}_0 = \sum_{i=0}^{N} \mathbf{e}_i\, v_i \quad (4)$$

is given at start. In order to have $N+1$ initial vectors, one may consider basis vectors $\{\mathbf{e}_i\}_{i=1}^N$, as further elements of the initial set

$$\mathbf{v}_i = \mathbf{e}_i, \quad i = 1, \dots, N.$$

Provided that $v_0 \neq 0$, the set $\{\mathbf{v}_i\}_{i=0}^N$ is linearly independent and complete in the $N+1$ dimensional Euclidean space.

[1] For brevity, we use the term 'Löwdin-orthogonalization' to refer to Löwdin's symmetrical scheme.

We now devise a two-step orthogonalization of the set $\{\mathbf{v}_i\}_{i=0}^N$. At first, let us take vectors $\{\mathbf{e}_i\}_{i=1}^N$ and set them orthogonal to $\mathbf{v}_0$. This can be achieved by applying the projection

$$\mathbf{P} = \mathbf{I} - \mathbf{v}_0\mathbf{v}_0^T,$$

as if taking one step by the Gram-Schmidt algorithm:

$$\mathbf{e}_i' = \mathbf{e}_i - \mathbf{v}_0\left(\mathbf{v}_0^T\mathbf{e}_i\right) = \mathbf{e}_i - \mathbf{v}_0 v_i, \quad i = 1,\ldots,N.$$

Vector $\mathbf{v}_0$ is assumed to be normalized. Next, let us subject vectors $\{\mathbf{e}_i'\}_{i=1}^N$ to Löwdin orthogonalization. For this end, the overlap matrix of the projected vectors is constructed according to

$$S_{ik}' = \mathbf{e}_i'^T\mathbf{e}_k' = \left(\mathbf{e}_i^T - v_i\mathbf{v}_0^T\right)\left(\mathbf{e}_k - \mathbf{v}_0 v_k\right) = \delta_{ik} - v_i v_k. \tag{5}$$

Once the components v_i are given, it is possible to compute the inverse square root of $\mathbf{S}'$ to yield the elements of the ON set as

$$\mathbf{u}_i = \sum_{k=1}^N (S')_{ki}^{-1/2}\mathbf{e}_k', \quad i = 1,\ldots,N. \tag{6}$$

In the present case, the simple structure of the overlap matrix in Eq. (5) opens a way to construct $(\mathbf{S}')^{-1/2}$ in closed form, yielding

$$(\mathbf{S}')^{-1/2} = \mathbf{I} - (\mathbf{S}' - \mathbf{I})\frac{1}{v_0(v_0+1)}. \tag{7}$$

Derivation of Eq. (7) is presented in the "Appendix".

Making use of the above result, Löwdin-orthogonalized vectors of Eq. (6) can be expressed as

$$\mathbf{u}_i = \sum_{k=1}^N \left(\delta_{ki} + \frac{v_i v_k}{v_0(v_0+1)}\right)\left(\mathbf{e}_k - \mathbf{v}_0 v_k\right).$$

By utilizing $\sum_{k=1}^N v_k^2 = 1 - v_0^2$ and expansion (Eq. 4) for $\mathbf{v}_0$, the above formula can be simplified to get

$$\mathbf{u}_i = \sum_{k=1}^N \left(\delta_{ik} - \frac{v_i v_k}{1+v_0}\right)\mathbf{e}_k - v_i\mathbf{e}_0,$$

an expression that agrees with Mayer's vectors of Eq. (3). Keeping in mind that matching of $\mathbf{u}_0$ of Eq. (4) with $\mathbf{v}_0$ of Eq. (2) was supposed at start, we see that Mayer's orthogonalization is recovered by the following two-step procedure:

(i) projection of N orthonormal basis vectors to become orthogonal to a selected unit vector (a Gram-Schmidt step),
(ii) Löwdin orthogonalization of the resulting N-dimensional, overlapping set.

Introducing matrix $\mathbf{V}$ that performs the linear transformation of the Gram-Schmidt step (i):

$$\mathbf{v}_0 = \sum_{j=0}^N \mathbf{e}_j V_{j0},$$

$$\mathbf{v}_i = \sum_{j=0}^N \mathbf{e}_j V_{ji}, \; i = 1,\ldots,N,$$

the above, two-step orthogonalization procedure is characterized by the transformation

$$\mathbf{V}\mathbf{S}^{-1/2}, \tag{8}$$

where the overlap matrix

$$\mathbf{S} = \mathbf{V}^T\mathbf{V}$$

is the direct sum of the one-dimensional identity matrix and $\mathbf{S}'$:

$$\mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{S}' \end{pmatrix}. \tag{9}$$

Matrix $\mathbf{V}\ \mathbf{S}^{-1/2}$ of Eq. (8) is unitary (it connects the orthonormal sets $\{\mathbf{e}_i\}_{0=1}^N$ and $\{\mathbf{u}_i\}_{i=0}^N$) and parametrized by N components of vector $\mathbf{v}_0$. (One component can be taken fixed, due to normalization). Regarding Mayer's matrix $\mathbf{U}$, it is also a unitary matrix, parametrized by N rotation angles φ_i in the form [10, 11]

$$\mathbf{U} = e^{\mathbf{A}},$$

with

$$\mathbf{A} = \begin{pmatrix} 0 & -\varphi_1 & \ldots & -\varphi_N \\ \varphi_1 & 0 & \ldots & 0 \\ \ldots & 0 & \ldots & 0 \\ \varphi_N & 0 & \ldots & 0 \end{pmatrix},$$

mapping between angles φ_i and vector components v_i is as given in Sect. 2.1. Since both $\mathbf{V}\ \mathbf{S}^{-1/2}$ and $\mathbf{U}$ are unitary, containing N parameters, it is obvious that the two are related by a unitary transformation. The non-trivial finding of the present study is that they are equal:

$$\mathbf{U} = \mathbf{V}\mathbf{S}^{-1/2}.$$

2.3 Some properties of Mayer-orthogonalization

Closer examination of Eq. (1) tells that constructing matrix $\mathbf{U}$, that is, constructing the complementary orthonormal set to $\mathbf{v}_0$ requires multiplications proportional in number to N^2. This is in contrast to the rough $\sim N^3$ processor time requirement of a full Gram-Schmidt or a Löwdin orthogonalization in general. Reduction of the exponent of scaling is a benefit of the special problem given at the outset. Analytical solution of the Löwdin-orthogonalization step, for example, relies on the diadic product structure of the offdiagonal elements of $\mathbf{S}'$. It is to be noted that Mayer-orthogonalization is not the only way to achieve a better

than cubic orthogonalization for this problem. The extreme sparsity of the initial overlap matrix (differing from **I** in just one row and column) may, for example, be exploited by the application of sparse matrix techniques in numerical orthogonalization procedures to reach this aim.

Treatment of near linear dependence in the original set is a delicate question of orthogonalization techniques. In the present context, there is just one possibility for redundancy: the start vector $\mathbf{v}_0$ may become close to one vector of the N-dimensional ON set, say $\mathbf{e}_2$. The behaviour of Mayer-orthogonalization in this case can easily be checked by considering a three-component start vector $v_0 = v_1 = \epsilon, v_2 = \sqrt{1-2\epsilon^2}$ and taking the limit $\epsilon \to 0$. Carrying out the Gram-Schmidt and the Löwdin step successively, one runs into trouble for $\epsilon = 0$, since $\mathbf{e}_2'$ becomes zero and induces divergence of $(\mathbf{S}')^{-1/2}$. For this reason, $v_0 = 0$ has to be excluded if taking the Gram-Schmidt and the Löwdin step successively. It is however interesting to observe that the analytically available combination of the two, matrix **U** itself, stays stable for $\epsilon = 0$. The effect of choosing an unfortunate pivot (v_0 instead of v_2) is simply $\mathbf{u}_2 \to -\mathbf{e}_0$ and $\mathbf{u}_0 \to \mathbf{e}_2$ as $\epsilon \to 0$. For $\epsilon = 0$ the final ON set is therefore composed of $\mathbf{e}_2$, $\mathbf{e}_1$ and $-\mathbf{e}_0$ what is desirable. It is worth to compare this result with Cholesky decomposition (CD), since in case of Gram-Schmidt orthogonalization, linear dependence is very well-handled by CD [12, 13]. For this end, the overlap matrix of the four vectors $\mathbf{v}_0$, $\mathbf{e}_0$, $\mathbf{e}_1$ and $\mathbf{e}_2$ can be constructed and subjected to Cholesky decomposition. This yields the same three vectors for $\epsilon = 0$ as Mayer-orthogonalization, apart from the sign. The fourth Cholesky-vector becomes exactly zero.

3 Illustration

Mayer's orthogonalization may be of use in various areas of quantum chemistry. Whenever an ON set of vectors has to be generated, based on a single start vector, this may be a method of choice. The application presented below gives one example for such a situation. With this illustration, we merely wish to demonstrate the usefulness of Mayer-orthogonalization. It is out of the scope of the present study to compare Mayer's method with other orthogonalization techniques in terms of scaling or numerical behaviour.

A typical case when the present orthogonalization problem arises is perturbation theory (PT) based on a multireference function. The common starting point of the diverse techniques available in this field of electronic structure theory [14–17] is a zero-order wavefunction ($\mathbf{v}_0$), often arising as the eigenvector of some model Hamiltonian. Construction of PT corrections to this start vector necessitates a complete set of zero-order vectors. In multiconfiguration PT (MCPT) [18–20], this situation was previously handled by considering excited determinants as zero-order excited states. The non-zero-overlap between the zero-order functions emerges by projecting the determinants out of the multiconfigurational reference function. This overlap was treated previously by constructing the reciprocal set and adopting the biorthogonal formulation of Rayleigh-Schrödinger PT. Mayer's vectors give a completely new approach to this problem, via producing a complete, ON set of zero-order functions. In contrast to the previous non-Hermitean formulation, Mayer's vectors facilitate a Hermitean zero-order Hamiltonian and the use of the standard second-order formula

$$E^{(2)} = -\sum_{i=1}^{N} \frac{|\mathbf{v}_0^T \mathbf{H} \mathbf{u}_i|^2}{E_i - E_0}, \tag{10}$$

with $\mathbf{u}_i$ being Mayer's vectors. In what follows, an example is given for this PT approach, starting from the antisymmetrized product of strongly orthogonal geminals [21] (APSG) wavefunction, as zero-order solution. Energy levels E_i appearing in Eq. (10) are eigenfunctions of the zero-order Hamiltonian. In the present case, these quantities are computed in a Møller-Plesset-like manner, that is, as sums of suitably chosen one-particle energies. For more details on the possible definitions of E_i, we refer to previous studies on multiconfigurational PT [20]. The dissociation curve plotted in Fig. 1 for the N_2 molecule reflects that the use of Mayer's vectors give a reliable MCPT method that captures the essential physics of the problem.

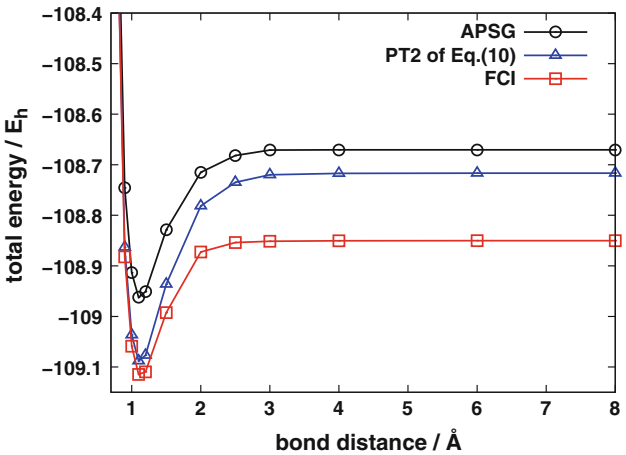

Fig. 1 Dissociation energy profile of the N_2 molecule obtained by APSG-based second-order PT in Dunning's valence double-zeta basis [22]. The full configuration interaction (FCI) is plotted as benchmark. Core electrons are frozen. The APSG wavefunction involves 3 geminals, with two orbitals assigned to each. This essentially agrees with a generalized valence bond (GVB) wavefunction, producing eight terms in the determinantal expansion. Mayer's orthogonalization is performed in this 8-dimensional space. Vectors $\mathbf{u}_i$ falling out of this space are simple determinants

Acknowledgments The authors are indebted to Professor I. Mayer (Budapest) for a detailed inspection and instructive critical remarks on the present study. This work has been supported by the Hungarian National Research Fund (OTKA), grant numbers K-81588 and K-81590. The European Union and the European Social Fund have also provided financial support to the project under the grant agreements TÁMOP 4.2.1./B-09/1/KMR-2010-0003 and 4.2.2./B-10/1-2010-0030.

Appendix

On the expression of $(\mathbf{S}')^{-1/2}$

Here, we construct $(\mathbf{S}')^{-1/2}$ of Eq. (7) in a closed form, by expanding it in matrix Taylor series around the N dimensional identity matrix, $\mathbf{I}$:

$$(\mathbf{S}')^{-1/2} = \mathbf{I} + \sum_{i=1}^{\infty} (\mathbf{S}' - \mathbf{I})^i (-1)^i \frac{1}{i!} \prod_{j=1}^{i} \left(-\frac{1}{2} + j \right). \tag{11}$$

To evaluate this expression, we first show by complete induction that

$$(\mathbf{S}' - \mathbf{I})^{i+1} = \left(v_0^2 - 1\right)^i (\mathbf{S}' - \mathbf{I}), \quad i \geq 1. \tag{12}$$

The first power of $\mathbf{S}' - \mathbf{I}$ is simply:

$$(\mathbf{S}' - \mathbf{I})_{lk} = -v_l v_k.$$

The second power is calculated as:

$$\begin{aligned} (\mathbf{S}' - \mathbf{I})^2_{lk} &= v_l v_k \sum_{j=1}^{N} v_j^2 = v_l v_k \left(1 - v_0^2\right) \\ &= \left(v_0^2 - 1\right)(\mathbf{S}' - \mathbf{I})_{lk}. \end{aligned} \tag{13}$$

Based on Eq. (13), we formulate the hypothesis

$$(\mathbf{S}' - \mathbf{I})^i = \left(v_0^2 - 1\right)^{i-1} (\mathbf{S}' - \mathbf{I}),$$

and take the induction step:

$$\begin{aligned} (\mathbf{S}' - \mathbf{I})^{i+1} &= (\mathbf{S}' - \mathbf{I})^i (\mathbf{S}' - \mathbf{I}) = \left(v_0^2 - 1\right)^{i-1} (\mathbf{S}' - \mathbf{I})(\mathbf{S}' - \mathbf{I}) \\ &= \left(v_0^2 - 1\right)^i (\mathbf{S}' - \mathbf{I}), \end{aligned}$$

that was to be demonstrated.

Let us now substitute Eq. (12) into the series expansion of Eq. (11):

$$\begin{aligned} (\mathbf{S}')^{-1/2} &= \mathbf{I} + (\mathbf{S}' - \mathbf{I}) \sum_{i=1}^{\infty} \left(v_0^2 - 1\right)^{(i-1)} (-1)^i \frac{1}{i!} \prod_{j=1}^{i} \left(-\frac{1}{2} + j \right) \\ &= \mathbf{I} - (\mathbf{S}' - \mathbf{I}) c \end{aligned}$$

where the constant

$$c = \sum_{i=1}^{\infty} \left(1 - v_0^2\right)^{(i-1)} \frac{1}{i!} \prod_{j=1}^{i} \left(-\frac{1}{2} + j \right) \tag{14}$$

depends only on v_0. To evaluate c, we first introduce the Gamma-function into our formula, making use of the expressions

$$\frac{1}{i!} = \frac{1}{\Gamma(i+1)}$$

$$\prod_{j=1}^{i} \left(-\frac{1}{2} + j \right) = \frac{1}{\sqrt{\pi}} \Gamma\left(i + \frac{1}{2} \right).$$

Substitution into Eq. (14) results

$$\begin{aligned} c &= \sum_{i=1}^{\infty} \left(1 - v_0^2\right)^{(i-1)} \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(i + \frac{1}{2}\right)}{\Gamma(i+1)} \\ &= \sum_{i=1}^{\infty} \frac{\left(1 - v_0^2\right)^{(i-1)}}{\sqrt{\pi}} \frac{\Gamma\left(i + \frac{1}{2}\right)}{\Gamma(i+1)} \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}. \end{aligned} \tag{15}$$

Next, we take advantage of the Gamma-function-based definition of the binomial coefficient to write

$$\binom{i - \frac{1}{2}}{i} = \frac{\Gamma\left(i + \frac{1}{2}\right)}{\Gamma(i+1)\Gamma\left(\frac{1}{2}\right)}. \tag{16}$$

Making use of Eq. (16) and substituting $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ into Eq. (15), we get

$$\begin{aligned} c &= \sum_{i=1}^{\infty} \left(1 - v_0^2\right)^{(i-1)} \binom{i - \frac{1}{2}}{i} \\ &= \frac{1}{1 - v_0^2} \sum_{i=1}^{\infty} \left(1 - v_0^2\right)^i \binom{i - \frac{1}{2}}{i} \end{aligned} \tag{17}$$

By shifting the summation index

$$c = \frac{1}{1 - v_0^2} \left(-1 + \sum_{i=0}^{\infty} \left(1 - v_0^2\right)^i \binom{i - \frac{1}{2}}{i} \right), \tag{18}$$

we recognize the power series

$$\frac{1}{(1-x)^{\alpha+1}} = \sum_{i=0}^{\infty} \binom{i + \alpha}{i} x^i,$$

with $x = 1 - v_0^2$ and $\alpha = -\frac{1}{2}$. This simplifies Eq. (18) to become

$$c = \frac{1}{1 - v_0^2} \left(-1 + \frac{1}{\left[1 - \left(1 - v_0^2\right)\right]^{1/2}} \right) = \frac{1}{\left(1 - v_0^2\right)} \frac{1 - \sqrt{v_0^2}}{\sqrt{v_0^2}}$$

At this point, we have to distinguish two cases depending on the sign of v_0.

If $v_0 > 0$, then $\sqrt{v_0^2} = v_0$,

$$c = \frac{1}{v_0(1 + v_0)},$$

leading to the desired formula, Eq. (7) for $(\mathbf{S}')^{-1/2}$:

$$(\mathbf{S}')^{-1/2} = \mathbf{I} - (\mathbf{S}' - \mathbf{I})\frac{1}{v_0(v_0+1)}. \quad (19)$$

If $v_0 < 0$, then $\sqrt{v_0^2} = -v_0$,

$$c = \frac{1}{v_0(v_0-1)}$$

and $(\mathbf{S}')^{-1/2}$ becomes

$$(\mathbf{S}')^{-1/2} = \mathbf{I} - (\mathbf{S}' - \mathbf{I})\frac{1}{v_0(v_0-1)}. \quad (20)$$

Both expressions of $(\mathbf{S}')^{-1/2}$ are completely in accord with a previous result [18] on the inverse of $\mathbf{S}'$, which reads:

$$(\mathbf{S}')^{-1} = \mathbf{I} - (\mathbf{S}' - \mathbf{I})\frac{1}{v_0^2}.$$

By taking the square of either Eq. (19) or Eq. (20), $(\mathbf{S}')^{-1}$ as given above can be easily recovered.

References

1. Löwdin PO (1950) J Chem Phys 18:365
2. Löwdin PO (1992) Adv Quantum Chem 23:83
3. Wannier GH (1937) Phys Rev 52(3):0191
4. Löwdin PO (1956) Adv Phys 5(17):1
5. Schweinl HC, Wigner EP (1970) J Math Phys 11(5):1693
6. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1996) Numerical recipes in Fortran 90 (2nd ed.): the art of parallel scientific computing. Cambridge University Press, New York, NY
7. Chaturvedi S, Kapoor AK, Srinivasan V (1998) J Phys A-Math Gen 31(19):L367
8. Löwdin PO (1970) Adv Quantum Chem 5:185
9. Mayer I (2002) Int J Quantum Chem 90(1):63. doi:10.1002/qua.981
10. Mayer I (2000) Theor Chem Acc 104:163
11. Mayer I (2003) Simple theorems, proofs, and derivations in quantum chemistry. Kluwer, New York
12. Beebe NH, Linderberg J (1977) Int J Quantum Chem 12:683
13. Pedersen TB, Aquilante F, Lindh R (2009) Theor Chem Acc 124:1
14. Durand P, Malrieu JP (1987) Adv Chem Phys 67:1
15. Roos B, Andersson K, Fülscher M, Malmqvist PÅ, Serrano-Andrés L, Pierloot K, Merchán M (1996) Adv Chem Phys 93:219
16. Hoffmann MR, Datta D, Das S, Mukherjee D, Szabados Á, Rolik Z, Surján PR (2009) J Chem Phys 131:204104
17. Pulay P (2011) Int J Quantum Chem 111:3273
18. Rolik Z, Szabados Á, Surján PR (2003) J Chem Phys 119:1922
19. Kobayashi M, Szabados Á, Nakai H, Surján PR (2010) J Chem Theory Comput 6:2024
20. Surján P, Rolik Z, Szabados Á, Kőhalmi D (2004) Ann Phys (Leipzig) 13:223
21. Surján PR (1999) Topics Curr Chem 203:63
22. Dunning TH Jr. (1989) J Chem Phys 90:1007