Article

Closed-Form Expressions for the Matrix Exponential

F. De Zela

Departamento de Ciencias, Sección Física, Pontificia Universidad Católica del Perú, Ap.1761, Lima L32, Peru; E-Mail: fdezela@pucp.edu.pe; Tel.: +51-1-6262000; Fax: +51-1-6262085

Abstract: We discuss a method to obtain closed-form expressions of $f(A)$, where f is an analytic function and A a square, diagonalizable matrix. The method exploits the Cayley–Hamilton theorem and has been previously reported using tools that are perhaps not sufficiently appealing to physicists. Here, we derive the results on which the method is based by using tools most commonly employed by physicists. We show the advantages of the method in comparison with standard approaches, especially when dealing with the exponential of low-dimensional matrices. In contrast to other approaches that require, e.g., solving differential equations, the present method only requires the construction of the inverse of the Vandermonde matrix. We show the advantages of the method by applying it to different cases, mostly restricting the calculational effort to the handling of two-by-two matrices.

Keywords: matrix exponential; Cayley–Hamilton theorem; two-by-two representations; Vandermonde matrices

Classification: PACS 02.30.Tb, 42.25.Ja, 03.65.Fd

1. Introduction

Physicists are quite often faced with the task of calculating $f(A)$, where A is an $n \times n$ matrix and f an analytic function whose series expansion generally contains infinitely many terms. The most prominent example corresponds to $\exp A$. Usual approaches to calculate $f(A)$ consist in either truncating its series expansion, or else finding a way to "re-summate" terms so as to get a closed-form expression. There is yet another option that can be advantageously applied when dealing with an $n \times n$ matrix, and which derives from the Cayley–Hamilton theorem [1]. This theorem states that every square matrix satisfies

its characteristic equation. As a consequence of this property, any series expansion can be written in terms of the first n powers of A. While this result is surely very well known among mathematicians, it appears to be not so widespread within the physicists' community [2]. Indeed, most textbooks on quantum mechanics still resort to the Baker–Hausdorff lemma or to special properties of the involved matrices, in order to obtain closed-form expressions of series expansions [3–5]. This happens even when dealing with low-dimensional matrices, *i.e.*, in cases in which exploiting the Cayley–Hamilton theorem would straightforwardly lead to the desired result. Such state of affairs probably reflects a lack of literature on the subject that is more palatable to physicists than to mathematicians. The present paper aims at dealing with the subject matter by using language and tools that are most familiar to physicists. No claim of priority is made; our purpose is to show how well the derived results fit into the repertoire of tools that physicists routinely employ. To this end, we start addressing the simple, yet rich enough case of 2×2 matrices.

An archetypical example is the Hamiltonian $H = k\boldsymbol{\sigma} \cdot \boldsymbol{B}$ that rules the dynamics of a spin-1/2 particle subjected to a magnetic field $\boldsymbol{B}$. Here, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ denotes the Pauli spin operator and k is a parameter that provides the above expression with appropriate units. The upsurge of research in several areas of physics—most notably in quantum optics—involving two-level systems, has made a Hamiltonian of the above type quite ubiquitous. Indeed, the dynamics of any two-level system is ruled by a Hamiltonian that can be written in such a form. Hence, one often requires an explicit, closed-form expression for quantities such as $\exp(i\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma})$, where $\boldsymbol{n}$ is a unit vector. This closed-form expression can be obtained as a generalization of Euler's formula $\exp i\alpha = \cos\alpha + i\sin\alpha$. It reads

$$\exp(i\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma}) = \cos\alpha I + i\sin\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma} \tag{1}$$

with I denoting the identity operator.

Let us recall how most textbooks of quantum mechanics proceed to demonstrate Equation (1) (see, e.g., [3–5]). The demonstration starts by writing the series expansion $\exp A = \sum_k A^k/k!$ for the case $A = i\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma}$. Next, one invokes the following relationship:

$$(\boldsymbol{a} \cdot \boldsymbol{\sigma})(\boldsymbol{b} \cdot \boldsymbol{\sigma}) = (\boldsymbol{a} \cdot \boldsymbol{b})I + i(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{\sigma} \tag{2}$$

whose proof rests on $\sigma_i \sigma_j = \delta_{ij}I + i\epsilon_{ijk}\sigma_k$ (summation over repeated indices being understood). Equation (2) implies that $(\boldsymbol{n} \cdot \boldsymbol{\sigma})^{2n} = I$, and hence $(\boldsymbol{n} \cdot \boldsymbol{\sigma})^{2n+1} = \boldsymbol{n} \cdot \boldsymbol{\sigma}$. This allows one to split the power series of $\exp(i\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma})$ in two parts, one constituted by even and the other by odd powers of $i\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma}$:

$$\exp(i\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma}) = \sum_{n=0}^{\infty} \frac{(i\alpha)^{2n}}{2n!}I + \sum_{n=0}^{\infty} \frac{(i\alpha)^{2n+1}}{(2n+1)!}\boldsymbol{n} \cdot \boldsymbol{\sigma} \tag{3}$$

By similarly splitting Euler's exponential, *i.e.*,

$$\exp i\alpha = \cos\alpha + i\sin\alpha = \sum_{n=0}^{\infty} \frac{(i\alpha)^{2n}}{2n!} + \sum_{n=0}^{\infty} \frac{(i\alpha)^{2n+1}}{(2n+1)!} \tag{4}$$

one sees that Equation (3) is the same as Equation (1).

Although this standard demonstration is a relatively simple one, it seems to be tightly related to the particular properties of the operator $\boldsymbol{n} \cdot \boldsymbol{\sigma}$, as well as to our ability to "re-summate" the series expansion

so as to obtain a closed-form expression. There are several other cases [6] in which a relation similar to Equation (1) follows as a consequence of generalizing some properties of the group $SU(2)$ and its algebra to the case $SU(N)$, with $N > 2$. Central to these generalizations and to their associated techniques are both the Cayley–Hamilton theorem and the closure of the Lie algebra $su(N)$ under commutation and anti-commutation of its elements [6]. As already recalled, the Cayley–Hamilton theorem states that any $n \times n$ matrix A satisfies its own characteristic equation $p(A) = 0$, where

$$p(\lambda) = \text{Det}(\lambda I - A) = \lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0 \qquad (5)$$

is A's characteristic polynomial. From $p(A) = 0$ it follows that any power A^k, with $k \geq n$, can be written in terms of the matrices $I = A^0, A, \ldots, A^{n-1}$. Thus, any infinite series, such as the one corresponding to $\exp A$, may be rewritten in terms of the n powers $A^0, A, \ldots, A^{n-1}$. By exploiting this fact one can recover Equation (1). Reciprocally, given A, one can construct a matrix B that satisfies $\exp B = A$, as shown by Dattoli, Mari and Torre [2]. These authors used essentially the same tools as we do here and presented some of the results that we will show below, but leaving them in an implicit form. The aforementioned authors belong to a group that has extensively dealt with our subject matter and beyond it [7], applying the present techniques to cases of current interest [8]. A somewhat different approach was followed by Leonard [9], who related the Cayley–Hamilton theorem to the solution of ordinary differential equations, in order to get closed expressions for the matrix exponential. This technique can be applied to all $n \times n$ matrices, including those that are not diagonalizable. Untidt and Nielsen [10] used this technique when addressing the groups $SU(2)$, $SU(3)$ and $SU(4)$. Now, especially when addressing $SU(2)$, Leonard's approach seems to be unnecessarily involved. This is because there is a trade-off between the wide applicability of the method and its tailoring to a special case. When dealing with diagonalizable matrices, the present approach may prove more useful. Thus, one exploits not only the Cayley–Hamilton theorem, but the diagonalizability of the involved matrices as well. As a result, we are provided with a straightforward way to obtain closed-form expressions for the matrix exponential. There are certainly many other ways that are either more general [9,11] or else better suited to specific cases [12–16], but the present method is especially useful for physical applications.

The rest of the paper is organized as follows. First, we present Leonard's technique in a way that somewhat differs from the approach used in [9]. Thereafter, we show how to obtain Equation (1) by using a technique that can be generalized to diagonalizable $n \times n$ matrices, thereby introducing the method that is the main subject of the present work. As an illustration of this technique, we address some representative cases that were taken from the repertoire of classical mechanics, quantum electrodynamics, quantum optics and from the realm of Lorentz transformations. While the results obtained are known, their derivations should serve to demonstrate the versatility of the method. Let us stress once again that our aim has been to present this method by following an approach that could be appealing to most physicists, rather than to mathematically oriented readers.

2. Closed Form of the Matrix Exponential via the Solution of Differential Equations

Consider the coupled system of differential equations, given by

$$D\boldsymbol{x} \equiv \frac{d\boldsymbol{x}}{dt} = A\boldsymbol{x} \qquad (6)$$

with $\boldsymbol{x} = (x_1, \ldots, x_n)^T$ and A a constant, $n \times n$ matrix. The matrix exponential appears in the solution of Equation (6), when we write it as $\boldsymbol{x}(t) = e^{At}\boldsymbol{x}(0)$. By successive derivation of this exponential we obtain $D^k e^{At} = A^k e^{At}$. Hence, $p(D)e^{At} \equiv (D^n + c_{n-1}D^{n-1} + \ldots + c_1 D + c_0)e^{At} = p(A)e^{At} = 0$, on account of $p(A) = 0$, *i.e.*, the Cayley–Hamilton theorem. Now, as already noted, this implies that e^{At} can be expressed in terms of $A^0, A, \ldots, A^{n-1}$. Let us consider the matrix $M(t) := \sum_{k=0}^{n-1} y_k(t)A^k$, with the $y_k(t)$ being n independent solutions of the differential equation $p(D)y(t) = 0$. That is, the $y_k(t)$ solve this equation for n different initial conditions that will be conveniently chosen. We have thus that $p(D)M(t) = \sum_{k=0}^{n-1} p(D)y_k(t)A^k = 0$. Our goal is to choose the $y_k(t)$ so that $e^{At} = M(t)$. To this end, we note that $D^k e^{At}\big|_{t=0} = A^k e^{At}\big|_{t=0} = A^k$. That is, e^{At} solves $p(D)\Phi(t) = 0$ with the initial conditions $\Phi(0) = A^0, \ldots, D^{n-1}\Phi(0) = A^{n-1}$. It is then clear that we must take the following initial conditions: $D^j y_k(0) = \delta_k^j$, with $j, k \in \{0, \ldots, n-1\}$. In such a case, e^{At} and $M(t)$ satisfy both the same differential equation and the same initial conditions. Hence, $e^{At} = M(t)$.

Summarizing, the method consists in solving the n-th order differential equation $p(D)y(t) = 0$ for n different initial conditions. These conditions read $D^j y_k(0) = \delta_k^j$, with $j, k \in \{0, \ldots, n-1\}$. The matrix exponential is then given by $e^{At} = \sum_{k=0}^{n-1} y_k(t)A^k$. The standard procedure for solving $p(D)y(t) = 0$ requires finding the roots of the characteristic equation $p(\lambda) = 0$. Each root λ with multiplicity m contributes to the general solution with a term $(a_0 + a_1 t + \ldots + a_{m-1}t^{m-1})e^{\lambda t}$, the a_k being fixed by the initial conditions. As already said, this method applies even when the matrix A is not diagonalizable. However, when the eigenvalue problem for A is a solvable one, another approach can be more convenient. We present such an approach in what follows.

3. Closed Form of the Matrix Exponential via the Solution of Algebraic Equations

Let us return to Equation (1). We will derive it anew, this time using standard tools of quantum mechanics. Consider a Hermitian operator A, whose eigenvectors satisfy $A |a_k\rangle = a_k |a_k\rangle$ and span the Hilbert space on which A acts. Thus, the identity operator can be written as $I = \sum_k |a_k\rangle \langle a_k|$. One can also write $A = A \cdot I = \sum_k a_k |a_k\rangle \langle a_k|$. Moreover, $A^m = \sum_k a_k^m |a_k\rangle \langle a_k|$, from which it follows that

$$F(A) = \sum_k F(a_k) |a_k\rangle \langle a_k| \tag{7}$$

for any function $F(A)$ that can be expanded in powers of A.

Let us consider the 2×2 case $A = \boldsymbol{n} \cdot \boldsymbol{\sigma}$, with $\boldsymbol{n}$ a unit vector. This matrix has the eigenvalues ± 1 and the corresponding eigenvectors $|\boldsymbol{n}_\pm\rangle$. That is, $\boldsymbol{n} \cdot \boldsymbol{\sigma} |\boldsymbol{n}_\pm\rangle = \pm |\boldsymbol{n}_\pm\rangle$. We need no more than this to get Equation (1). Indeed, from $\boldsymbol{n} \cdot \boldsymbol{\sigma} = |\boldsymbol{n}_+\rangle \langle \boldsymbol{n}_+| - |\boldsymbol{n}_-\rangle \langle \boldsymbol{n}_-|$ and $I = |\boldsymbol{n}_+\rangle \langle \boldsymbol{n}_+| + |\boldsymbol{n}_-\rangle \langle \boldsymbol{n}_-|$, it follows that $|\boldsymbol{n}_\pm\rangle \langle \boldsymbol{n}_\pm| = (I \pm \boldsymbol{n} \cdot \boldsymbol{\sigma})/2$. Next, we consider $F(A) = \exp A = \sum_k \exp a_k |a_k\rangle \langle a_k|$, with $A = i\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma}$. The operator $i\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma}$ has eigenvectors $|\boldsymbol{n}_\pm\rangle$ and eigenvalues $\pm i\alpha$. Thus,

$$\begin{align}
\exp(i\alpha \boldsymbol{n} \cdot \boldsymbol{\sigma}) &= e^{i\alpha} |\boldsymbol{n}_+\rangle \langle \boldsymbol{n}_+| + e^{-i\alpha} |\boldsymbol{n}_-\rangle \langle \boldsymbol{n}_-| \tag{8} \\
&= \frac{1}{2}e^{i\alpha} (I + \boldsymbol{n} \cdot \boldsymbol{\sigma}) + \frac{1}{2}e^{-i\alpha} (I - \boldsymbol{n} \cdot \boldsymbol{\sigma}) \tag{9} \\
&= \left(\frac{e^{i\alpha} + e^{-i\alpha}}{2}\right) I + \left(\frac{e^{i\alpha} - e^{-i\alpha}}{2}\right) \boldsymbol{n} \cdot \boldsymbol{\sigma} \tag{10}
\end{align}$$

which is Equation (1). Note that it has not been necessary to know the eigenvectors of $A = i\alpha\boldsymbol{n} \cdot \boldsymbol{\sigma}$. It is a matter of convenience whether one chooses to express $\exp(i\alpha\boldsymbol{n} \cdot \boldsymbol{\sigma})$ in terms of the projectors $|\boldsymbol{n}_{\pm}\rangle\langle\boldsymbol{n}_{\pm}|$, or in terms of I and $\boldsymbol{n} \cdot \boldsymbol{\sigma}$.

Let us now see how the above method generalizes when dealing with higher-dimensional spaces. To this end, we keep dealing with rotations. The operator $\exp(i\alpha\boldsymbol{n} \cdot \boldsymbol{\sigma})$ is a rotation operator acting on spinor space. It is also an element of the group $SU(2)$, whose generators can be taken as $X_i = i\sigma_i/2$, $i = 1, 2, 3$. They satisfy the commutation relations $[X_i, X_j] = \epsilon_{ijk}X_k$ that characterize the rotation algebra. The rotation operator can also act on three-dimensional vectors $\boldsymbol{r}$. In this case, one often uses the following formula, which gives the rotated vector $\boldsymbol{r}'$ in terms of the rotation angle θ and the unit vector $\boldsymbol{n}$ that defines the rotation axis:

$$\boldsymbol{r}' = \boldsymbol{r}\cos\theta + \boldsymbol{n}\left(\boldsymbol{n} \cdot \boldsymbol{r}\right)\left[1 - \cos\theta\right] + \left(\boldsymbol{n} \times \boldsymbol{r}\right)\sin\theta \tag{11}$$

Equation (11) is usually derived from vector algebra plus some geometrical considerations [17]. We can derive it, alternatively, by the method used above. To this end, we consider the rotation generators X_i for three-dimensional space, which can be read off from the next formula, Equation (12). The rotation matrix is then obtained as $\exp(\theta\boldsymbol{n} \cdot \boldsymbol{X})$, with

$$\boldsymbol{n} \cdot \boldsymbol{X} = \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix} \equiv M \tag{12}$$

It is straightforward to find the eigenvalues of the non-Hermitian, antisymmetric matrix M. They are 0 and $\pm i$. Let us denote the corresponding eigenvectors as $|\boldsymbol{n}_0\rangle$ and $|\boldsymbol{n}_{\pm}\rangle$, respectively. Similarly to the spin case, we have now

$$I = |\boldsymbol{n}_+\rangle\langle\boldsymbol{n}_+| + |\boldsymbol{n}_-\rangle\langle\boldsymbol{n}_-| + |\boldsymbol{n}_0\rangle\langle\boldsymbol{n}_0| \tag{13}$$
$$M = i|\boldsymbol{n}_+\rangle\langle\boldsymbol{n}_+| - i|\boldsymbol{n}_-\rangle\langle\boldsymbol{n}_-| \tag{14}$$

We need a third equation, if we want to express the three projectors $|\boldsymbol{n}_k\rangle\langle\boldsymbol{n}_k|$, $k = \pm, 0$, in terms of I and M. This equation is obtained by squaring M:

$$M^2 = -|\boldsymbol{n}_+\rangle\langle\boldsymbol{n}_+| - |\boldsymbol{n}_-\rangle\langle\boldsymbol{n}_-| \tag{15}$$

From Equations (13)–(15) we immediately obtain $|\boldsymbol{n}_{\pm}\rangle\langle\boldsymbol{n}_{\pm}| = \left(\mp iM - M^2\right)/2$, and $|\boldsymbol{n}_0\rangle\langle\boldsymbol{n}_0| = I + M^2$. Thus, we have

$$\exp(\theta M) = e^{i\theta}|\boldsymbol{n}_+\rangle\langle\boldsymbol{n}_+| + e^{-i\theta}|\boldsymbol{n}_-\rangle\langle\boldsymbol{n}_-| + e^0|\boldsymbol{n}_0\rangle\langle\boldsymbol{n}_0| \tag{16}$$
$$= I + M\sin\theta + M^2\left[1 - \cos\theta\right] \tag{17}$$

By letting M, as given in Equation (12), act on $\boldsymbol{r} = (x, y, z)^T$, we easily see that $M\boldsymbol{r} = \boldsymbol{n} \times \boldsymbol{r}$ and $M^2\boldsymbol{r} = \boldsymbol{n} \times (\boldsymbol{n} \times \boldsymbol{r}) = \boldsymbol{n}(\boldsymbol{n} \cdot \boldsymbol{r}) - \boldsymbol{r}$. Thus, on account of Equation (17), $\boldsymbol{r}' = \exp(\theta M)\boldsymbol{r}$ reads the same as Equation (11).

The general case is now clear. Consider an operator A whose matrix representation is an $N \times N$ matrix. Once the eigenvalues a_k of A (which we assume nondegenerate) have been

determined, we can write the N equations: $A^0 = I = \sum_k |a_k\rangle \langle a_k|$, $A = \sum_k a_k |a_k\rangle \langle a_k|$, $A^2 = \sum_{k=1}^{N} a_k^2 |a_k\rangle \langle a_k|, \ldots, A^{N-1} = \sum_{k=1}^{N} a_k^{N-1} |a_k\rangle \langle a_k|$, from which it is possible to obtain the N projectors $|a_k\rangle \langle a_k|$ in terms of $I, A, A^2, \ldots, A^{N-1}$. To this end, we must solve the system

$$
\begin{pmatrix}
1 & 1 & \cdots & 1 \\
a_1 & a_2 & \cdots & a_N \\
a_1^2 & a_2^2 & & a_N^2 \\
\vdots & \vdots & & \vdots \\
a_1^{N-1} & a_2^{N-1} & & a_N^{N-1}
\end{pmatrix}
\begin{pmatrix}
|a_1\rangle \langle a_1| \\
|a_2\rangle \langle a_2| \\
|a_3\rangle \langle a_3| \\
\vdots \\
|a_N\rangle \langle a_N|
\end{pmatrix}
=
\begin{pmatrix}
I \\
A \\
A^2 \\
\vdots \\
A^{N-1}
\end{pmatrix}
\tag{18}
$$

The matrix in Equation (18), with components $V_{k,i} = a_i^{k-1}$ ($k, i \in \{1, \ldots, N\}$), is a Vandermonde matrix, whose inverse can be explicitly given [18]. Once we have written the $|a_k\rangle \langle a_k|$ in terms of $I, A, \ldots A^{N-1}$, we can express any analytic function of A in terms of these N powers of A, in particular $\exp A = \sum_{k=1}^{N} \exp(a_k) |a_k\rangle \langle a_k|$. For the case $N = 4$, for instance, we have the following result:

$$
|a_1\rangle \langle a_1| = \frac{A^3 - A^2(a_2 + a_3 + a_4) + A(a_2 a_3 + a_2 a_4 + a_3 a_4) - a_2 a_3 a_4}{(a_1 - a_2)(a_1 - a_3)(a_1 - a_4)}
\tag{19}
$$

$$
|a_2\rangle \langle a_2| = \frac{A^3 - A^2(a_1 + a_3 + a_4) + A(a_1 a_3 + a_1 a_4 + a_3 a_4) - a_1 a_3 a_4}{(a_2 - a_1)(a_2 - a_3)(a_2 - a_4)}
\tag{20}
$$

$$
|a_3\rangle \langle a_3| = \frac{A^3 - A^2(a_1 + a_2 + a_4) + A(a_1 a_2 + a_1 a_4 + a_2 a_4) - a_1 a_2 a_4}{(a_3 - a_1)(a_3 - a_2)(a_3 - a_4)}
\tag{21}
$$

$$
|a_4\rangle \langle a_4| = \frac{A^3 - A^2(a_1 + a_2 + a_3) + A(a_1 a_2 + a_1 a_3 + a_2 a_3) - a_1 a_3 a_4}{(a_4 - a_1)(a_4 - a_2)(a_4 - a_3)}
\tag{22}
$$

The general solution can be written in terms of the inverse of the Vandermonde matrix V. To this end, consider a system of equations that reads like (18), but with the operators entering the column vectors being replaced by numbers, *i.e.*, $|a_j\rangle \langle a_j| \to w_j$, with $j = 1, \ldots, N$, and $A^k \to q_{k+1}$, with $k = 0, \ldots, N - 1$. The solution of this system is given by $w_j = \sum_{k=0}^{N-1} U_{j,k} q_k$, with $U = V^{-1}$, the inverse of the Vandermonde matrix. This matrix inverse can be calculated as follows [18]. Let us define a polynomial $P_j(x)$ of degree $N - 1$ as

$$
P_j(x) = \prod_{\substack{n=1 \\ n \neq j}}^{N} \frac{x - a_n}{a_j - a_n} = \sum_{k=1}^{N} U_{j,k} x^{k-1}
\tag{23}
$$

The coefficients $U_{j,k}$ of the last equality follow from expanding the preceding expression and collecting equal powers of x. These $U_{j,k}$ are the components of V^{-1}. Indeed, setting $x = a_i$ and observing that $P_j(a_i) = \delta_{ji} = \sum_{k=1}^{N} U_{j,k} a_i^{k-1} = (UV)_{j,i}$, we see that U is the inverse of the Vandermonde matrix. The projectors $|a_j\rangle \langle a_j|$ in Equation (18) can thus be obtained by replacing $x \to A$ in Equation (23). We get in this way the explicit solution

$$
|a_j\rangle \langle a_j| = \sum_{k=1}^{N} U_{j,k} A^{k-1} = \prod_{\substack{n=1 \\ n \neq j}}^{N} \frac{A - a_n}{a_j - a_n}
\tag{24}
$$

The above expression can be inserted into Equation (7), if one wants to write $F(A)$ in terms of the first N powers of A.

So far, we have assumed that the eigenvalues of A are all nondegenerate. Let us now consider a matrix M with degenerate eigenvalues. As before, we deal with a special case, from which the general formalism can be easily inferred. Let M be of dimension four and with eigenvalues λ_1 and λ_2, which are two-fold degenerate. We can group the projectors as follows:

$$I = (|e_1\rangle \langle e_1| + |e_2\rangle \langle e_2|) + (|e_3\rangle \langle a_3| + |e_4\rangle \langle e_4|) \tag{25}$$

$$M = \lambda_1 (|e_1\rangle \langle e_1| + |e_2\rangle \langle e_2|) + \lambda_2 (|e_3\rangle \langle a_3| + |e_4\rangle \langle e_4|) \tag{26}$$

It is then easy to solve the above equations for the two projectors associated with the two eigenvalues. We obtain

$$|e_1\rangle \langle e_1| + |e_2\rangle \langle e_2| = \frac{\lambda_2 I - M}{\lambda_2 - \lambda_1} \tag{27}$$

$$|e_3\rangle \langle a_3| + |e_4\rangle \langle e_4| = \frac{\lambda_1 I - M}{\lambda_1 - \lambda_2} \tag{28}$$

We can then write

$$e^M = \frac{1}{\lambda_1 - \lambda_2} \left[\left(\lambda_1 e^{\lambda_2} - \lambda_2 e^{\lambda_1} \right) I + \left(e^{\lambda_1} - e^{\lambda_2} \right) M \right] \tag{29}$$

We will need this result for the calculation of the unitary operator that defines the Foldy–Wouthuysen transformation, our next example. It is now clear that in the general case of degenerate eigenvalues, we can proceed similarly to the nondegenerate case, but solving $n < N$ equations.

4. Examples

Let us now see how the method works when applied to some well-known cases. Henceforth, we refer to the method as the Cayley–Hamilton (CH)-method, for short. Our aim is to show the simplicity of the required calculations, as compared with standard techniques.

4.1. The Foldy–Wouthuysen Transformation

The Foldy-Wouthuysen transformation is introduced [19] with the aim of decoupling the upper (φ) and lower (χ) components of a bispinor $\psi = (\varphi, \chi)^T$ that solves the Dirac equation $i\hbar \partial \psi / \partial t = H\psi$, where $H = -i\hbar c \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} + \beta m c^2$. Here, β and $\boldsymbol{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$ are the 4×4 Dirac matrices:

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix} \tag{30}$$

The Foldy–Wouthuysen transformation is given by $\psi' = U\psi$, with [19]

$$U = \exp\left(\frac{\theta}{2} \beta \boldsymbol{\alpha} \cdot \boldsymbol{p} \right) \tag{31}$$

We can calculate U by applying Equation (29) for $M = \theta \beta \boldsymbol{\alpha} \cdot \boldsymbol{p}/2 = (\theta |\boldsymbol{p}|/2)\beta \boldsymbol{\alpha} \cdot \boldsymbol{n}$, where $\boldsymbol{n} = \boldsymbol{p}/|\boldsymbol{p}|$. The eigenvalues of the 4×4 matrix $\beta \boldsymbol{\alpha} \cdot \boldsymbol{n}$ are $\pm i$, each being two-fold degenerate. This follows from noting that the matrices

$$\beta \boldsymbol{\alpha} \cdot \boldsymbol{n} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \cdot \boldsymbol{n} \\ -\boldsymbol{\sigma} \cdot \boldsymbol{n} & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{32}$$

have the same eigenvalues. Indeed, because $(\boldsymbol{\sigma} \cdot \boldsymbol{n})^2 = \mathbf{1}$, the above matrices share the characteristic equation $\lambda^2 + 1 = 0$. Their eigenvalues are thus $\pm i$. The eigenvalues of $M = \theta \beta \boldsymbol{\alpha} \cdot \boldsymbol{p}/2$ are then $\lambda_{1,2} = \pm i\theta \left| \boldsymbol{p} \right|/2$. Replacing these values in Equation (29) we obtain

$$\exp\left(\frac{\theta}{2}\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}\right) = \frac{1}{i\theta\left|\boldsymbol{p}\right|}\left[\frac{i\theta\left|\boldsymbol{p}\right|}{2}\left(e^{-i\theta\left|\boldsymbol{p}\right|/2}+e^{i\theta\left|\boldsymbol{p}\right|/2}\right)I + \left(e^{i\theta\left|\boldsymbol{p}\right|/2}-e^{-i\theta\left|\boldsymbol{p}\right|/2}\right)\frac{\theta\left|\boldsymbol{p}\right|}{2}\beta\boldsymbol{\alpha}\cdot\boldsymbol{n}\right] \quad (33)$$

$$= \cos\left(\left|\boldsymbol{p}\right|\theta/2\right) + \sin\left(\left|\boldsymbol{p}\right|\theta/2\right)\beta\boldsymbol{\alpha}\cdot\frac{\boldsymbol{p}}{\left|\boldsymbol{p}\right|} \quad (34)$$

The standard way to get this result requires developing the exponential in a power series. Thereafter, one must exploit the commutation properties of $\boldsymbol{\alpha}$ and β in order to group together odd and even powers of θ. This finally leads to the same closed-form expression that we have arrived at after some few steps.

4.2. Lorentz-Type Equations of Motion

The dynamics of several classical and quantum systems is ruled by equations that can be cast as differential equations for a three-vector $\boldsymbol{S}$. These equations often contain terms of the form $\boldsymbol{\Omega}\times$. An example of this is the ubiquitous equation

$$\frac{d\boldsymbol{S}}{dt} = \boldsymbol{\Omega} \times \boldsymbol{S} \quad (35)$$

Equation (35) and its variants have been recently addressed by Babusci, Dattoli and Sabia [20], who applied operational methods to deal with them. Instead of writing Equation (35) in matrix form, these authors chose to exploit the properties of the vector product by defining the operator $\hat{\boldsymbol{\Omega}} := \boldsymbol{\Omega}\times$. The solution for the case $\partial\boldsymbol{\Omega}/\partial t = \boldsymbol{0}$, for instance, was obtained by expanding $\exp(t\hat{\boldsymbol{\Omega}})$ as an infinite series and using the cyclical properties of the vector product in order to get $\boldsymbol{S}(t)$ in closed form. This form is nothing but Equation (11) with the replacements $\boldsymbol{r}' \to \boldsymbol{S}(t)$, $\boldsymbol{r} \to \boldsymbol{S}(0)$ and $\theta \to \Omega t$, where $\Omega := |\boldsymbol{\Omega}|$. We obtained Equation (11) without expanding the exponential and without using any cyclic properties. Our solution follows from writing Equation (35) in matrix form, *i.e.*,

$$\frac{d\boldsymbol{S}}{dt} = \Omega M \boldsymbol{S} \quad (36)$$

where M is given by Equation (12) with $\boldsymbol{n} = \boldsymbol{\Omega}/\Omega$. The solution $\boldsymbol{S}(t) = \exp(M\Omega t)\boldsymbol{S}(0)$ is then easily written in closed form by applying the CH-method, as in Equation (11). The advantages of this method show up even more sharply when dealing with some extensions of Equation (36). Consider, e.g., the non-homogeneous version of Equation (35):

$$\frac{d\boldsymbol{S}}{dt} = \boldsymbol{\Omega} \times \boldsymbol{S} + \boldsymbol{N} = \Omega M \boldsymbol{S} + \boldsymbol{N} \quad (37)$$

This is the form taken by the Lorentz equation of motion when the electromagnetic field is given by scalar and vector potentials reading $\Phi = -\boldsymbol{E} \cdot \boldsymbol{r}$ and $\boldsymbol{A} = \boldsymbol{B} \times \boldsymbol{r}/2$, respectively [20]. The solution of Equation (37) is easily obtained by acting on both sides with the "integrating (operator-valued) factor" $\exp(-\Omega M t)$. One then readily obtains, for the initial condition $\boldsymbol{S}(0) = \boldsymbol{S}_0$,

$$\boldsymbol{S}(t) = e^{\Omega M t}\boldsymbol{S}_0 + \int_0^t e^{\Omega M(t-s)}\boldsymbol{N}ds \quad (38)$$

The matrix exponentials in Equation (38) can be expressed in their eigenbasis, as in Equation (16). For a time-independent $\boldsymbol{N}$, the integral in Equation (38) is then trivial. An equivalent solution is given in [20], but written in terms of the evolution operator $\hat{U}(t) = \exp(t\hat{\Omega})$ and its inverse. Inverse operators repeatedly appear within such a framework [20] and are often calculated with the help of the Laplace transform identity: $\hat{A}^{-1} = \int_0^\infty \exp(-s\hat{A})ds$. Depending on $\hat{A}$, this could be not such a straightforward task as it might appear at first sight. Now, while vector notation gives us additional physical insight, vector calculus can rapidly turn into a messy business. Our strategy is therefore to avoid vector calculus and instead rely on the CH-method as much as possible. Only at the end we write down our results, if we wish, in terms of vector products and the like. That is, we use Equations (13)–(17) systematically, in particular Equation (16) when we need to handle $\exp(\theta M)$, e.g., within integrals. The simplification comes about from our working with the eigenbasis of $\exp(\theta M)$, *i.e.*, with the eigenbasis of M. Writing down the final results in three-vector notation amounts to expressing these results in the basis in which M was originally defined, *cf.* Equation (12). Let us denote this basis by $\{|\boldsymbol{x}\rangle, |\boldsymbol{y}\rangle, |\boldsymbol{z}\rangle\}$. The eigenvectors $|\boldsymbol{n}_\pm\rangle$ and $|\boldsymbol{n}_0\rangle$ of M are easily obtained from those of X_3, *cf.* Equation (12). The eigenvectors of X_3 are, in turn, analogous to those of Pauli's σ_y, namely $|\pm\rangle = (|\boldsymbol{x}\rangle \mp i|\boldsymbol{y}\rangle)/\sqrt{2}$, plus a third eigenvector that is orthogonal to the former ones, that is, $|0\rangle = |\boldsymbol{z}\rangle$. In order to obtain the eigenvectors of $\boldsymbol{n} \cdot \boldsymbol{X}$, with $\boldsymbol{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, we apply the rotation $\exp(\phi X_3)\exp(\theta X_2)$ to the eigenvectors $|\pm\rangle$ and $|0\rangle$, thereby getting $|\boldsymbol{n}_\pm\rangle$ and $|\boldsymbol{n}_0\rangle$, respectively. All these calculations are easily performed using the CH-method.

Once we have $|\boldsymbol{n}_\pm\rangle$ and $|\boldsymbol{n}_0\rangle$, we also have the transformation matrix T that brings M into diagonal form: $T^{-1}MT = M_D = \mathrm{diag}(-i, 0, i)$. Indeed, T's columns are just $|\boldsymbol{n}_-\rangle$, $|\boldsymbol{n}_0\rangle$ and $|\boldsymbol{n}_+\rangle$. After we have carried out all calculations in the eigenbasis of M, by applying T we can express the final result in the basis $\{|\boldsymbol{x}\rangle, |\boldsymbol{y}\rangle, |\boldsymbol{z}\rangle\}$, thereby obtaining the desired expressions in three-vector notation. Let us illustrate this procedure by addressing the evolution equation

$$\frac{d\boldsymbol{S}}{dt} = \boldsymbol{\Omega} \times \boldsymbol{S} + \lambda\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{S}) \tag{39}$$

In matrix form, such an equation reads

$$\frac{d\boldsymbol{S}}{dt} = \Omega M \boldsymbol{S} + \lambda(\Omega M)^2 \boldsymbol{S} = [\Omega M + \lambda(\Omega M)^2]\boldsymbol{S} \equiv A\boldsymbol{S} \tag{40}$$

The solution is given by $\boldsymbol{S}(t) = \exp(At)\boldsymbol{S}_0$. The eigenbasis of A is the same as that of M. We have thus

$$\exp(At) = e^{(i\Omega - \lambda\Omega^2)t}|\boldsymbol{n}_+\rangle\langle\boldsymbol{n}_+| + e^{(-i\Omega - \lambda\Omega^2)t}|\boldsymbol{n}_-\rangle\langle\boldsymbol{n}_-| + |\boldsymbol{n}_0\rangle\langle\boldsymbol{n}_0| \tag{41}$$

The projectors $|\boldsymbol{n}_k\rangle\langle\boldsymbol{n}_k|$ can be written in terms of the powers of A by solving the system

$$I = |\boldsymbol{n}_+\rangle\langle\boldsymbol{n}_+| + |\boldsymbol{n}_-\rangle\langle\boldsymbol{n}_-| + |\boldsymbol{n}_0\rangle\langle\boldsymbol{n}_0| \tag{42}$$

$$A = (i\Omega - \lambda\Omega^2)|\boldsymbol{n}_+\rangle\langle\boldsymbol{n}_+| - (i\Omega + \lambda\Omega^2)|\boldsymbol{n}_-\rangle\langle\boldsymbol{n}_-| \tag{43}$$

$$A^2 = (i\Omega - \lambda\Omega^2)^2|\boldsymbol{n}_+\rangle\langle\boldsymbol{n}_+| + (i\Omega + \lambda\Omega^2)^2|\boldsymbol{n}_-\rangle\langle\boldsymbol{n}_-| \tag{44}$$

Using $A = \Omega M + \lambda(\Omega M)^2$ and $A^2 = -2\lambda\Omega^3 M + (1 - \lambda^2\Omega^2)(\Omega M)^2$, and replacing the solution of the system (42)–(44) in Equation (41) we get

$$\exp(At) = I + e^{-\lambda\Omega^2 t}\sin(\Omega t)M + [1 - e^{-\lambda\Omega^2 t}\cos(\Omega t)]M^2 \tag{45}$$

Finally, we can write the solution $\boldsymbol{S}(t) = \exp(At)\boldsymbol{S}_0$ in the original basis $\{|\boldsymbol{x}\rangle, |\boldsymbol{y}\rangle, |\boldsymbol{z}\rangle\}$, something that in this case amounts to writing $M\boldsymbol{S}_0 = \boldsymbol{n} \times \boldsymbol{S}_0$ and $M^2\boldsymbol{S}_0 = \boldsymbol{n}\,(\boldsymbol{n} \cdot \boldsymbol{S}_0) - \boldsymbol{S}_0$. Equation (39) was also addressed in [20], but making use of the operator method. The solution was given in terms of a series expansion for the evolution operator. In order to write this solution in closed form, it is necessary to introduce sin- and cos-like functions [20]. These functions are defined as infinite series involving two-variable Hermite polynomials. The final expression reads like Equation (11), but with sin and cos replaced by the aforementioned functions containing two-variable Hermite polynomials. Now, one can hardly unravel from such an expression the physical features that characterize the system's dynamics. On the other hand, a solution given as in Equation (45) clearly shows such dynamics, in particular the damping effect stemming from the λ-term in Equation (39), for $\lambda > 0$. Indeed, Equation (45) clearly shows that the state vector $\boldsymbol{S}(t) = \exp(At)\boldsymbol{S}_0$ asymptotically aligns with $\boldsymbol{\Omega}$ while performing a damped Larmor precession about the latter.

The case $\partial\boldsymbol{\Omega}/\partial t \neq \boldsymbol{0}$ is more involved and generally requires resorting to Dyson-like series expansions, e.g., time-ordered exponential integrations. While this subject lies beyond the scope of the present work, it should be mentioned that the CH-method can be advantageously applied also in this context. For instance, time-ordered exponential integrations involving operators of the form $A + B(t)$ do require the evaluation of $\exp A$. Likewise, disentangling techniques make repeated use of matrix exponentials of single operators [21]. In all these cases, the CH-method offers a possible shortcut.

4.3. The Jaynes–Cummings Hamiltonian

We address now a system composed by a two-level atom and a quantized (monochromatic) electromagnetic field. Under the dipole and the rotating-wave approximations, the Hamiltonian of this system reads (in standard notation)

$$H = \frac{\hbar}{2}\omega_0\sigma_z + \hbar\omega a^\dagger a + \hbar g\left(a^\dagger\sigma_- + a\sigma_+\right) \tag{46}$$

Let us denote the upper and lower states of the two-level atom by $|a\rangle$ and $|b\rangle$, respectively, and the Fock states of the photon-field by $|n\rangle$. The Hilbert space of the atom-field system is spanned by the basis $\mathcal{B} = \{|a, n\rangle, |b, n\rangle, n = 0, 1, \ldots\}$. The states $|a, n\rangle$ and $|b, n\rangle$ are eigenstates of the unperturbed Hamiltonian $H_0 = \hbar\omega_0\sigma_z/2 + \hbar\omega a^\dagger a$. The interaction Hamiltonian $V = \hbar g\left(a^\dagger\sigma_- + a\sigma_+\right)$ couples the states $|a, n\rangle$ and $|b, n + 1\rangle$ alone. Hence, H can be split into a sum: $H = \sum_n H_n$, with each H_n acting on the subspace $\mathrm{Span}\{|a, n\rangle, |b, n + 1\rangle\}$. Within such a subspace, H_n is represented by the 2×2 matrix

$$H_n = \hbar\omega\left(n + \frac{1}{2}\right)I + \hbar\begin{pmatrix} \delta/2 & g\sqrt{n+1} \\ g\sqrt{n+1} & -\delta/2 \end{pmatrix} \tag{47}$$

where $\delta = \omega_0 - \omega$.

A standard way [22] to calculate the evolution operator $U = \exp(-iHt/\hbar)$ goes as follows. One first writes the Hamiltonian in the form $H = H_1 + H_2$, with $H_1 = \hbar\omega\left(a^\dagger a + \sigma_z/2\right)$ and $H_2 = \hbar\delta\sigma_z/2 + \hbar g\left(a^\dagger\sigma_- + a\sigma_+\right)$. Because $[H_1, H_2] = 0$, the evolution operator can be factored as $U = U_1 U_2 = \exp(-iH_1t/\hbar)\exp(-iH_2t/\hbar)$. The first factor is diagonal in $\mathrm{Span}\,\mathcal{B}$. The second factor can be expanded in a Taylor series. As it turns out, one can obtain closed-form expressions for the even and the odd

powers of the expansion. Thus, a closed-form for U_2 can be obtained as well. As can be seen, this method depends on the realization that Equation (46) can be written in a special form, which renders it possible to factorize U.

Let us now calculate U by the CH-method. We can exploit the fact that H splits as $H = \sum_n H_n$, with $[H_n, H_m] = 0$, and write $U = \prod_n U_n = \prod_n \exp(-iH_n t/\hbar)$. Generally, a 2×2 Hamiltonian H has eigenvalues of the form $E_\pm = \hbar(\lambda_0 \pm \lambda)$. We have thus

$$I = |+\rangle \langle +| + |-\rangle \langle -| \tag{48}$$

$$H/\hbar = (\lambda_0 + \lambda) |+\rangle \langle +| + (\lambda_0 - \lambda) |-\rangle \langle -| \tag{49}$$

so that

$$\exp(-iHt/\hbar) = \exp(-i\lambda_+ t) |+\rangle \langle +| + \exp(-i\lambda_- t) |-\rangle \langle -| \tag{50}$$

$$= \frac{e^{-i\lambda_0 t}}{\lambda} \left[(i\lambda_0 \sin \lambda t + \lambda \cos \lambda t) I - i(\sin \lambda t) \frac{H}{\hbar} \right] \tag{51}$$

In our case, H_n has eigenvalues $E_n^\pm = \hbar\omega(n + 1/2) \pm \hbar\sqrt{\delta^2/4 + g^2(n+1)} \equiv \hbar\omega(n+1/2) \pm \hbar R_n$. Whence,

$$\exp(-iH_n t/\hbar) = \frac{e^{-i\omega(n+1/2)t}}{R_n} \left[\left(i\omega(n + \frac{1}{2}) \sin (R_n t) + R_n \cos (R_n t) \right) I - i \sin(R_n t) \frac{H_n}{\hbar} \right] \tag{52}$$

Replacing H_n from Equation (47) in the above expression we get

$$\exp(-iH_n t/\hbar) = e^{-i\omega(n+1/2)t} \left[\cos (R_n t) I - \frac{i \sin (R_n t)}{2R_n} \begin{pmatrix} \delta & 2g\sqrt{n+1} \\ 2g\sqrt{n+1} & -\delta \end{pmatrix} \right] \tag{53}$$

This result enables a straightforward calculation of the evolved state $|\psi(t)\rangle$ out of a general initial state

$$|\psi(0)\rangle = \sum_n C_{a,n} |a, n\rangle + C_{b,n+1} |b, n+1\rangle \tag{54}$$

Equation (53) refers to a matrix representation in the two-dimensional subspace $\text{Span}\{|a, n\rangle, |b, n+1\rangle\}$. Let us focus on

$$\cos (R_n t) I = \begin{pmatrix} \cos (R_n t) & 0 \\ 0 & \cos (R_n t) \end{pmatrix} \tag{55}$$

This matrix is a representation in subspace $\text{Span}\{|a, n\rangle, |b, n+1\rangle\}$ of the operator

$$\cos \left(t\sqrt{\widehat{\varphi} + g^2} \right) |a\rangle\langle a| + \cos \left(t\sqrt{\widehat{\varphi}} \right) |b\rangle\langle b| \tag{56}$$

where $\widehat{\varphi} := g^2 a^\dagger a + \delta^2/4$. Proceeding similarly with the other operators that enter Equation (53) and observing that $\sin (R_n t) R_n^{-1} \sqrt{n+1} = \langle n| i \sin(t\sqrt{\widehat{\varphi} + g^2})(\sqrt{\widehat{\varphi} + g^2})^{-1} a |n+1\rangle$, *etc.*, we readily obtain

$$\exp(-iHt/\hbar) = e^{-i\omega(a^\dagger a + \frac{1}{2})t} \begin{pmatrix} \cos \left(t\sqrt{\widehat{\varphi} + g^2} \right) - \frac{i\delta \sin\left(t\sqrt{\widehat{\varphi}+g^2}\right)}{2\sqrt{\widehat{\varphi}+g^2}} & -\frac{ig \sin(t\sqrt{\widehat{\varphi}+g^2})}{\sqrt{\widehat{\varphi}+g^2}} a \\ -\frac{ig \sin\left(t\sqrt{\widehat{\varphi}}\right)}{\sqrt{\widehat{\varphi}}} a^\dagger & \cos \left(t\sqrt{\widehat{\varphi}} \right) + \frac{i\delta \sin\left(t\sqrt{\widehat{\varphi}}\right)}{2\sqrt{\widehat{\varphi}}} \end{pmatrix} \tag{57}$$

where the 2×2 matrix refers now to the atomic subspace $\text{Span}\{|a\rangle, |b\rangle\}$. One can see that the CH-method reduces the amount of calculational effort invested to get Equation (53), as compared with other approaches [22].

4.4. Bispinors and Lorentz Transformations

As a further application, let us consider the representation of Lorentz transformations in the space of bispinors. In coordinate space, Lorentz transformations are given by $\widetilde{x}^\mu = \Lambda^\mu_\nu x^\nu$ (Greek indices run from 0 to 3), with the Λ^μ_ν satisfying $\Lambda^\mu_\nu \Lambda^\tau_\sigma \eta^{\nu\sigma} = \eta^{\mu\tau}$. Here, $\eta^{\mu\nu}$ represents the metric tensor of Minkowsky space ($\eta^{00} = -\eta^{11} = -\eta^{22} = \eta^{33} = 1$, $\eta^{\mu\nu} = 0$ for $\mu \neq \nu$). A bispinor $\psi(x)$ transforms according to [19]

$$\widetilde{\psi}(\widetilde{x}) = \widetilde{\psi}(\Lambda x) = S(\Lambda)\psi(x) \tag{58}$$

with

$$S(\Lambda) = \exp B \tag{59}$$
$$B = -\frac{1}{4}V^{\mu\nu}\gamma_\mu\gamma_\nu \tag{60}$$

The $V^{\mu\nu} = -V^{\nu\mu}$ are the components of an antisymmetric tensor, which has thus six independent components, corresponding to the six parameters defining a Lorentz transformation. The quantities $\gamma_\mu = \eta_{\mu\nu}\gamma^\nu$ satisfy $\gamma^\mu\gamma^\nu + \gamma^\mu\gamma^\nu = 2\eta^{\mu\nu}$. The quantities $\gamma_\mu\gamma_\nu$ are the generators of the Lorentz group. $S(\Lambda)$ is not a unitary transformation, but satisfies

$$S^{-1} = \gamma_0 S^\dagger \gamma_0 \tag{61}$$

For the following, it will be advantageous to define

$$p_i = \gamma_0\gamma_i, \ i = 1, 2, 3 \tag{62}$$
$$q_1 = \gamma_2\gamma_3, \ q_2 = \gamma_3\gamma_1, \ q_3 = \gamma_1\gamma_2 \tag{63}$$

We call the p_i Pauli generators and the q_i quaternion generators. The pseudoscalar $\gamma_5 := \gamma_0\gamma_1\gamma_2\gamma_3$ satisfies $\gamma_5^2 = -1$, $\gamma_5\gamma_\mu = -\gamma_\mu\gamma_5$, so that it commutes with each generator of the Lorentz group:

$$\gamma_5(\gamma_\mu\gamma_\nu) = (\gamma_\mu\gamma_\nu)\gamma_5 \tag{64}$$

This means that quantities of the form $\alpha + \beta\gamma_5$ ($\alpha, \beta \in \mathbb{R}$) behave like complex numbers upon multiplication with p_i and q_i. We denote the subspace spanned by such quantities as the complex-like subspace $C_{\boldsymbol{i}}$ and set $\boldsymbol{i} \equiv \gamma_5$. Noting that $\boldsymbol{i}\,p_i = q_i$ and $\boldsymbol{i}\,q_i = -p_i$, the following multiplication rules are easily derived:

$$q_iq_j = \epsilon_{ijk}q_k - \delta_{ij} \tag{65}$$
$$p_ip_j = -\epsilon_{ijk}q_k + \delta_{ij} = -q_iq_j = -\boldsymbol{i}\epsilon_{ijk}p_k + \delta_{ij} \tag{66}$$
$$p_iq_j = \epsilon_{ijk}p_k + \boldsymbol{i}\delta_{ij} = \boldsymbol{i}(-\epsilon_{ijk}q_k + \delta_{ij}) \tag{67}$$

The following commutators can then be straightforwardly obtained:

$$[q_i, q_j] = 2\epsilon_{ijk}q_k \tag{68}$$
$$[p_i, p_j] = -2\epsilon_{ijk}q_k = -2\boldsymbol{i}\epsilon_{ijk}p_k \tag{69}$$
$$[p_i, q_j] = 2\epsilon_{ijk}p_k \tag{70}$$

They make clear why we dubbed the p_i as Pauli generators. Noting that they furthermore satisfy

$$p_i p_j + p_j p_i = 2\delta_{ij} \tag{71}$$

we see the correspondence $\boldsymbol{i} \rightarrow i$, $p_k \rightarrow -\sigma_k$, with i being the imaginary unit and σ_k the Pauli matrices. These matrices, as is well-known, satisfy $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$ and the anticommutation relations $\sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij}$, which follow from $\sigma_i \sigma_j = i\epsilon_{ijk}\sigma_k + \delta_{ij}$.

We can write now $S(\Lambda) = \exp(-\frac{1}{4}V^{\mu\nu}\gamma_\mu\gamma_\nu)$ in terms of p_i and q_i:

$$B = \sum_{i=1}^{3}(\alpha^i p_i + \beta^i q_i) \tag{72}$$

Here, we have set $\alpha^i = -V^{0i}/4$ and $\beta^k \epsilon_{ijk} = -V^{ij}/4$. We can write B in terms of the Pauli-generators alone:

$$B = \sum_{i=1}^{3}(\alpha^i + \boldsymbol{i}\beta^i)p_i \equiv \sum_{i=1}^{3} z^i p_i \tag{73}$$

Considering the isomorphism $p_k \leftrightarrow -\sigma_k$, we could derive the expression for $S(\Lambda) = \exp B$ by splitting the series expansion into even and odd powers of B, and noting that

$$B^2 = (\boldsymbol{\alpha}^2 - \boldsymbol{\beta}^2) + (2\boldsymbol{\alpha} \cdot \boldsymbol{\beta})\,\boldsymbol{i} \equiv z^2 \in C_{\boldsymbol{i}} \tag{74}$$

where $\boldsymbol{\alpha}^2 \equiv \boldsymbol{\alpha} \cdot \boldsymbol{\alpha}$, $\boldsymbol{\beta}^2 \equiv \boldsymbol{\beta} \cdot \boldsymbol{\beta}$, and $\boldsymbol{\alpha} \cdot \boldsymbol{\beta} \equiv \sum_{i=1}^{3} \alpha^i \beta^i$. We have then that $B^3 = z^2 B$, $B^4 = z^4$, $B^5 = z^4 B$, ... This allows us to write

$$
\begin{aligned}
\exp(B) &= 1 + B + \frac{z^2}{2!} + \frac{z^2}{3!}B + \frac{z^4}{4!} + \frac{z^4}{5!}B + ... = \\
&\left(1 + \sum_{n=1}^{\infty}\frac{z^{2n}}{(2n)!}\right) + B\left(1 + \frac{z^2}{3!} + \frac{z^4}{5!} + ...\right) = \cosh z + \frac{\sinh z}{z}B
\end{aligned} \tag{75}
$$

As in the previous examples, also in this case the above result can be obtained more directly by noting that $B = \sum_{i=1}^{3}(\alpha^i + \boldsymbol{i}\beta^i)p_i \leftrightarrow -\sum_{i=1}^{3}(\alpha^i + i\beta^i)\sigma_i$. This suggests that we consider $\exp(-\boldsymbol{f} \cdot \boldsymbol{\sigma})$, with $\boldsymbol{f} = \boldsymbol{\alpha} + i\boldsymbol{\beta} \in \mathbb{C}$. The matrix $\boldsymbol{f} \cdot \boldsymbol{\sigma}$ has the (complex) eigenvalues

$$\lambda_\pm = \pm\sqrt{\boldsymbol{\alpha}^2 - \boldsymbol{\beta}^2 + 2i\boldsymbol{\alpha} \cdot \boldsymbol{\beta}} \equiv \pm z \tag{76}$$

Writing $|\boldsymbol{f}_\pm\rangle$ for the corresponding eigenvectors, *i.e.*, $\boldsymbol{f} \cdot \boldsymbol{\sigma}\,|\boldsymbol{f}_\pm\rangle = \lambda_\pm\,|\boldsymbol{f}_\pm\rangle$, we have that

$$I = |\boldsymbol{f}_+\rangle\langle\boldsymbol{f}_+| + |\boldsymbol{f}_-\rangle\langle\boldsymbol{f}_-| \tag{77}$$

$$\boldsymbol{f} \cdot \boldsymbol{\sigma} = \lambda_+\,|\boldsymbol{f}_+\rangle\langle\boldsymbol{f}_+| + \lambda_-\,|\boldsymbol{f}_-\rangle\langle\boldsymbol{f}_-| \tag{78}$$

Solving for $|\boldsymbol{f}_\pm\rangle\langle\boldsymbol{f}_\pm|$, we get

$$|\boldsymbol{f}_\pm\rangle\langle\boldsymbol{f}_\pm| = \frac{zI \pm \boldsymbol{f} \cdot \boldsymbol{\sigma}}{2z} \tag{79}$$

We apply now the general decomposition $\exp A = \sum_n \exp a_n \, |a_n\rangle \langle a_n|$ to the case $A = -\boldsymbol{f} \cdot \boldsymbol{\sigma}$. The operator $\exp(-\boldsymbol{f} \cdot \boldsymbol{\sigma})$ has eigenvectors $|\boldsymbol{f}_\pm\rangle$ and eigenvalues $\exp(\mp z)$. Thus,

$$\exp(-\boldsymbol{f} \cdot \boldsymbol{\sigma}) = e^{-z} |\boldsymbol{f}_+\rangle \langle \boldsymbol{f}_+| + e^{z} |\boldsymbol{f}_-\rangle \langle \boldsymbol{f}_-| \tag{80}$$

$$= \frac{e^{-z}}{2z}(zI + \boldsymbol{f} \cdot \boldsymbol{\sigma}) + \frac{e^{z}}{2z}(zI - \boldsymbol{f} \cdot \boldsymbol{\sigma}) \tag{81}$$

$$= \left(\frac{e^z + e^{-z}}{2}\right) I - \left(\frac{e^z - e^{-z}}{2z}\right) \boldsymbol{f} \cdot \boldsymbol{\sigma} \tag{82}$$

$$= \cosh z - \frac{\sinh z}{z} \boldsymbol{f} \cdot \boldsymbol{\sigma} \tag{83}$$

which is equivalent to Equation (75) via the correspondence $\cosh(z) + \sinh(z)B/z \leftrightarrow \cosh(z) - \sinh(z)\boldsymbol{f} \cdot \boldsymbol{\sigma}/z$. We have thus obtained closed-form expressions for $\exp(-\boldsymbol{f} \cdot \boldsymbol{\sigma})$, with $\boldsymbol{f} = \boldsymbol{\alpha} + i\boldsymbol{\beta} \in \mathbb{C}^3$, *i.e.*, for the elements of $SL(2, \mathbb{C})$, the universal covering group of the Lorentz group. It is interesting to note that the elements of $SL(2, \mathbb{C})$ are related to those of $SU(2)$ by extending the parameters $\boldsymbol{\alpha}$ entering $\exp(i\boldsymbol{\alpha} \cdot \boldsymbol{n}) \in SU(2)$ from the real to the complex domain: $i\boldsymbol{\alpha} \to \boldsymbol{\alpha} + i\boldsymbol{\beta}$. Standard calculations that are carried out with $SU(2)$ elements can be carried out similarly with $SL(2, \mathbb{C})$ elements [15]. A possible realization of $SU(2)$ transformations occurs in optics, by acting on the polarization of light with the help of birefringent elements (waveplates). If we also employ dichroic elements like polarizers, which absorb part of the light, then it is possible to implement $SL(2, \mathbb{C})$ transformations as well. In this way, one can simulate Lorentz transformations in the optical laboratory [23]. The above formalism is of great help for designing the corresponding experimental setup.

5. Conclusions

The method presented in this paper—referred to as the Cayley–Hamilton method—proves advantageous for calculating closed-form expressions of analytic functions $f(A)$ of an $n \times n$ matrix A, in particular matrix exponentials. The matrix A is assumed to be a diagonalizable one, even though only its eigenvalues are needed, not its eigenvectors. We have recovered some known results from classical and quantum mechanics, including Lorentz transformations, by performing the straightforward calculations that the method prescribes. In most cases, the problem at hand was reshaped so as to solve it by dealing with two-by-two matrices only.

Acknowledgements

Conflicts of Interest

The author declares no conflict of interest.

References

1. Gantmacher, F.R. *The Theory of Matrices*; Chelsea Publishing Company: New York, NY, USA, 1960; p. 83.

2. Dattoli, G.; Mari, C.; Torre, A. A simplified version of the Cayley-Hamilton theorem and exponential forms of the 2×2 and 3×3 matrices. *Il Nuovo Cimento* **1998**, *180*, 61–68.

3. Cohen-Tannoudji, C.; Diu, B.; Laloë, F. *Quantum Mechanics*; John Wiley & Sons: New York, NY, USA, 1977; pp. 983–989.

4. Sakurai, J.J. *Modern Quantum Mechanics*; Addison-Wesley: New York, NY, USA, 1980; pp. 163–168.

5. Greiner, W.; Müller, B. *Quantum Mechanics, Symmetries*; Springer: New York, NY, USA, 1989; p. 68.

6. Weigert, S. Baker-Campbell-Hausdorff relation for special unitary groups $SU(N)$. *J. Phys. A* **1997**, *30*, 8739–8749.

7. Dattoli, G.; Ottaviani, P.L.; Torre, A.; Vásquez, L. Evolution operator equations: Integration with algebraic and finite-difference methods. Applications to physical problems in classical and quantum mechanics and quantum field theory. *Riv. Nuovo Cimento* **1997**, *20*, 1–133.

8. Dattoli, G.; Zhukovsky, K. Quark flavour mixing and the exponential form of the Kobayashi–Maskawa matrix. *Eur. Phys. J. C* **2007**, *50*, 817–821.

9. Leonard, I. The matrix exponential. *SIAM Rev.* **1996**, *38*, 507–512.

10. Untidt, T.S.; Nielsen, N.C. Closed solution to the Baker-Campbell-Hausdorff problem: Exact effective Hamiltonian theory for analysis of nuclear-magnetic-resonance experiments. *Phys. Rev. E* **2002**, *65*, doi:10.1103/PhysRevE.65.021108.

11. Moore, G. Orthogonal polynomial expansions for the matrix exponential. *Linear Algebra Appl.* **2011**, *435*, 537–559.

12. Ding, F. Computation of matrix exponentials of special matrices. *Appl. Math. Comput.* **2013**, *223*, 311–326.

13. Koch, C.T.; Spence, J.C.H. A useful expansion of the exponential of the sum of two non-commuting matrices, one of which is diagonal. *J. Phys. A Math. Gen.* **2003**, *36*, 803–816.

14. Ramakrishna, V.; Zhou, H. On the exponential of matrices in $su(4)$. *J. Phys. A Math. Gen.* **2006**, *39*, 3021–3034.

15. Tudor, T. On the single-exponential closed form of the product of two exponential operators. *J. Phys. A Math. Theor.* **2007**, *40*, 14803–14810.

16. Siminovitch, D.; Untidt, T.S.; Nielsen, N.C. Exact effective Hamiltonian theory. II. Polynomial expansion of matrix functions and entangled unitary exponential operators. *J. Chem. Phys.* **2004**, *120*, 51–66.

17. Goldstein, H. *Classical Mechanics*, 2nd ed.; Addison-Wesley: New York, NY, USA, 1980; pp. 164–174.

18. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipees in FORTRAN, The Art of Scientific Computing*, 2nd ed.; Cambridge University Press: Cambridge, UK, 1992; pp. 83–84.

19. Bjorken, J.D.; Drell, S.D. *Relativistic Quantum Mechanics*; McGraw-Hill: New York, NY, USA, 1965.

20. Babusci, D.; Dattoli, G.; Sabia, E. Operational methods and Lorentz-type equations of motion. *J. Phys. Math.* **2011**, *3*, 1–17.
21. Puri, R.R. *Mathematical Methods of Quantum Optics*; Springer: New York, NY, USA, 2001; pp. 8–53.
22. Meystre, P.; Sargent, M. *Elements of Quantum Optics*, 2nd ed.; Springer: Berlin, Germany, 1999, pp. 372–373.
23. Kim, Y.S.; Noz, M.E. Symmetries shared by the Poincaré group and the Poincaré sphere. *Symmetry* **2013**, *5*, 233–252.