CBPF-NF-031/98

On the Coefficients of the Characteristic Polynomial

Ronaldo Rodrigues Silva*
Centro Brasileiro de Pesquisas Físicas,
Rua Dr. Xavier Sigaud 150
22290-180 Rio de Janeiro, Brazil

Abstract

It is a well-known fact that the first and last non-trivial coefficients of the characteristic polynomial of a linear operator are respectively its trace and its determinant. This work shows how to compute recursively all the coefficients as polynomial functions in the traces of successive powers of the operator. With the aid of Cayley-Hamilton's theorem the trace formulae provide a rational formula for the resolvent kernel and an operator-valued null identity for each finite dimension of the underlying vector space. The 4-dimensional resolvent formula allows an algebraic solution of the inverse-metric problem in general relativity.

PACS: 02.10, 02.30.T, 04.20.

*E-mail: ronaldor@cat.cbpf.br

1 Introduction

The eigenvalue problem arises in a variety of different branches of mathematical physics. For instance, it is well-known that quantum systems reach stationary states that are the eigenvectors of a suitable linear operator defined throughout a complex vector space representing the physical quantum states.

In the infinite dimensional case real and complex analytical methods have been developed to compute the eigenvalues of a linear operator. We mention, for instance, global variational methods, which consist of investigating stationary levels of a suitable energy functional, and local perturbative methods by means of complex analytical continuation.

On the other hand, if the underlying vector space is finite dimensional the eigenvalue problem becomes an algebraically well-posed problem by means of the characteristic polynomial. It is exactly this feature that justifies our algebraic approach in this work. Following this general device, the algebraic environment is set by the complex number field. After working out a few tools, the most important of these being Newton's identities, and setting very fundamental concepts and basic results in complex linear algebra, we succeed in achieving the mathematical core of this work: a recursive algorithm to compute the coefficients of the characteristic polynomial as algebraic functions in the traces of the successive powers of the linear operator.

With the aid of Cayley-Hamilton's theorem the trace formulae already obtained provide an operator-valued null identity for each finite dimension of the underlying vector space. As a by-product it sheds light on the algebraic structure of the associative algebra of complex linear operators. The computational skill just attained is enough to yield an operator-valued polynomial with rational coefficients for the finite-dimensional resolvent kernel, which improves a known result by revealing its rational dependence with respect to the spectral variable, as well as with respect to the linear operator.

In the context of general relativity, the 4-dimensional characteristic formula endows a polynomial expression for the volume scalar density. Furthermore, the 4-dimensional resolvent formula yields a tensor-valued third-degree polynomial for the inverse metric, thus avoiding the computacional and formal drawbacks of the Neumann series.

2 Newton's Identities

The fundamental theorem of algebra [1] together with Euclid's division algorithm imply algebraic closureness [2] of the complex number field; this basic result is encompassed without proof by

Theorem 0 *If* $p(z) = z^n + D_1 z^{n-1} + \cdots + D_{n-1} z + D_n$ *is a polynomial with complex coefficients* D_k *then there exist complex numbers* $\lambda_1, \lambda_2, \ldots, \lambda_n$*, called roots of* p*, such that* $p(z) = (z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_n)$.

As a consequence of the identity principle, the following relations between coefficients and roots hold:

$$
\begin{aligned}
-D_1 &= \lambda_1 + \lambda_2 + \cdots + \lambda_n, \\
+D_2 &= \lambda_1\lambda_2 + \cdots + \lambda_{n-1}\lambda_n, \\
&\vdots \\
(-)^k D_k &= \textstyle\sum k\text{-products of } \lambda\text{'s}, \\
&\vdots \\
(-)^n D_n &= \lambda_1\lambda_2\cdots\lambda_n.
\end{aligned} \tag{1}
$$

The right-hand side of (1) defines the elementary symmetric functions [3] in the variables $\lambda_1, \lambda_2, \ldots, \lambda_n$. Next to them, the most important symmetric functions are the sums of like powers:

$$
\begin{aligned}
T_1 &= \lambda_1 + \lambda_2 + \cdots + \lambda_n, \\
T_2 &= {\lambda_1}^2 + {\lambda_2}^2 + \cdots + {\lambda_n}^2, \\
&\vdots \\
T_k &= {\lambda_1}^k + {\lambda_2}^k + \cdots + {\lambda_n}^k, \\
&\vdots
\end{aligned} \tag{2}
$$

D_k and T_k, besides being symmetric, are homogeneous functions of degree k in the variables $\lambda_1, \lambda_2, \ldots, \lambda_n$. We shall derive a set of recursive relations connecting D's and T's by which the λ's are eliminated from (1) and (2).

To pursuit this goal we need two lemmas. The first one is an improved version of the remainder theorem.

Proposition 1 *If* $p(z) = z^n + D_1 z^{n-1} + \cdots + D_{n-1} z + D_n$ *is a polynomial with complex coefficients* D_k *and* λ *is a complex number then*

$$
\begin{aligned}
\frac{p(z) - p(\lambda)}{z - \lambda} = {} & z^{n-1} + (\lambda + D_1)\, z^{n-2} + (\lambda^2 + D_1\lambda + D_2)\, z^{n-3} + \cdots \\
& + (\lambda^{n-1} + D_1 \lambda^{n-2} + \cdots + D_{n-1}).
\end{aligned}
$$

Proof*: By induction on the degree. For n=1,* $p(z) - p(\lambda) = (z + D_1) - (\lambda + D_1) = z - \lambda$. *For generic n,* $p(z) - p(\lambda) =$

$$
\begin{aligned}
&= (z^n + D_1 z^{n-1} + \cdots + D_{n-1} z + D_n) - (\lambda^n + D_1 \lambda^{n-1} + \cdots + D_{n-1}\lambda + D_n) \\
&= z\,(z^{n-1} + D_1 z^{n-2} + \cdots + D_{n-1}) - \lambda\,(\lambda^{n-1} + D_1 \lambda^{n-2} + \cdots + D_{n-1}) \\
&= z\,((z^{n-1} + D_1 z^{n-2} + \cdots + D_{n-1}) - (\lambda^{n-1} + D_1 \lambda^{n-2} + \cdots + D_{n-1})) + \\
&\quad + (z - \lambda)(\lambda^{n-1} + D_1 \lambda^{n-2} + \cdots + D_{n-1}).
\end{aligned}
$$

From the inductive hypothesis the last expression reads

$$
\begin{aligned}
& z\,(z - \lambda)\,(z^{n-2} + (\lambda + D_1)\, z^{n-3} + \cdots + (\lambda^{n-2} + D_1 \lambda^{n-3} + \cdots + D_{n-2})) + \\
& + (z - \lambda)(\lambda^{n-1} + D_1 \lambda^{n-2} + \cdots + D_{n-1}) \\
= {} & (z - \lambda)(z^{n-1} + (\lambda + D_1) z^{n-2} + \cdots + (\lambda^{n-2} + D_1 \lambda^{n-3} + \cdots + D_{n-2})\, z + \\
& + (\lambda^{n-1} + D_1 \lambda^{n-2} + \cdots + D_{n-1})).
\end{aligned}
$$

The second lemma is a contribution of differential calculus to algebra.

Proposition 2 *If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the complex roots of the nth-degree polynomial p and p' is its derivative polynomial, then*

$$p'(z) = \sum_{k=1}^{n} \frac{p(z)}{z - \lambda_k}.$$

Proof: *p can be written as $p(z) = d_0\,(z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_n)$, from which one obtains its derivative as*

$$\begin{aligned} p'(z) &= d_0\,(z-\lambda_2)(z-\lambda_3)\cdots(z-\lambda_n) + d_0\,(z-\lambda_1)(z-\lambda_3)\cdots(z-\lambda_n) + \cdots \\ &\quad + d_0\,(z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_{n-1}) \\ &= \frac{p(z)}{z-\lambda_1} + \frac{p(z)}{z-\lambda_2} + \cdots + \frac{p(z)}{z-\lambda_n}. \end{aligned}$$

Thus, we are able to get

Theorem 1 (Newton's formulae [4]) *If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the complex roots of the polynomial $p(z) = z^n + D_1\,z^{n-1} + D_2\,z^{n-2} + \cdots + D_{n-1}\,z + D_n$ with complex coefficients D_k and $T_k = {\lambda_1}^k + {\lambda_2}^k + \cdots + {\lambda_n}^k$ for $k = 1, 2, \ldots, n$ then the following relations hold:*

$$\begin{aligned} &T_1 + D_1 = 0, \\ &T_2 + D_1\,T_1 + 2\,D_2 = 0, \\ &\qquad\qquad\vdots \\ &T_k + D_1\,T_{k-1} + \cdots + D_{k-1}\,T_1 + k\,D_k = 0, \\ &\qquad\qquad\vdots \\ &T_n + D_1\,T_{n-1} + \cdots + D_{n-1}\,T_1 + n\,D_n = 0. \end{aligned}$$

Proof: *Propositions 1 and 2 yield*

$$\begin{aligned} & nz^{n-1} + (n-1)D_1 z^{n-2} + (n-2)D_2 z^{n-3} + \cdots + D_{n-1} \\ &= \sum_{k=1}^{n} \Big\{ z^{n-1} + (\lambda_k + D_1)z^{n-2} + (\lambda_k^2 + D_1\lambda_k + D_2)z^{n-3} + \cdots \\ &\qquad + (\lambda_k^{n-1} + D_1\lambda_k^{n-2} + \cdots + D_{n-2}\lambda_k + D_{n-1}) \Big\} \\ &= nz^{n-1} + (T_1 + nD_1)z^{n-2} + (T_2 + D_1T_1 + nD_2)z^{n-3} + \cdots \\ &\quad + (T_{n-1} + D_1T_{n-2} + \cdots + D_{n-2}T_1 + nD_{n-1}). \end{aligned}$$

Equating coefficients we obtain

$$\begin{aligned} (n-1)D_1 &= T_1 + nD_1, \\ (n-2)D_2 &= T_2 + D_1T_1 + nD_2, \\ &\vdots \\ (n-k)D_k &= T_k + D_1T_{k-1} + \cdots + D_{k-1}T_1 + nD_k, \\ &\vdots \\ D_{n-1} &= T_{n-1} + D_1T_{n-2} + \cdots + D_{n-2}T_1 + nD_{n-1}, \end{aligned}$$

from which follow the $(n-1)$ *first relations to be shown. Since each* λ_k *is a root of* p *it verifies* $0=\lambda_k^n+D_1\lambda_k^{n-1}+\cdots+D_{n-1}\lambda_k+D_n$. *The addition of these relations yields*

$$0=\sum_{k=1}^{n}\left\{\lambda_k^n+D_1\lambda_k^{n-1}+\cdots+D_{n-1}\lambda_k+D_n\right\}=T_n+D_1T_{n-1}+\cdots+D_{n-1}T_1+nD_n,$$

which is the remaining formula to be proved.

3 The Trace Formulae

The characteristic polynomial [5] of a complex linear operator $\mathbf{T}$ in a finite dimensional complex vector space $\mathbb{V}$ is defined by $p(z)=\det(z\mathbf{I}-\mathbf{T})$, where $\mathbf{I}$ is the identity operator in $\mathbb{V}$. The degree of p as a polynomial in the complex variable z equals the dimension of $\mathbb{V}$ as a complex vector space; its roots are called the characteristic values of $\mathbf{T}$ and the set of characteristic values is called the spectrum of $\mathbf{T}$.

These definitions and properties lead to

Proposition 3 *The trace of a complex linear operator* $\mathbf{T}$ *equals the sum of its characteristic values counted with multiplicities.*
Proof: $\det(z\mathbf{I}-\mathbf{T})=(z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_n)$. *By the matrix representation* ${T^i}_j$ *of* $\mathbf{T}$ *we cast the left-hand side as* $z^n-({T^1}_1+{T^2}_2+\cdots+{T^n}_n)\,z^{n-1}+$ *terms of order* $\leq n-2$. *In the right-hand side we have* $z^n-(\lambda_1+\lambda_2+\cdots+\lambda_n)\,z^{n-1}+$ *terms of order* $\leq n-2$. *Equating coefficients we obtain*

$${T^1}_1+{T^2}_2+\cdots+{T^n}_n=\lambda_1+\lambda_2+\cdots+\lambda_n.$$

The roots of unity in the complex field yield the following factorization lemma in the associative algebra of complex linear operators.

Proposition 4 *If* $\mathbf{T}$ *and* $\mathbf{W}$ *are commuting complex linear operators and* $\theta=\mathrm{e}^{i2\pi/k}=\cos\frac{2\pi}{k}+i\,\sin\frac{2\pi}{k}$ *with* k *a positive integer, then*

$$\mathbf{T}^k-\mathbf{W}^k=(\mathbf{T}-\mathbf{W})(\mathbf{T}-\theta\,\mathbf{W})(\mathbf{T}-\theta^2\mathbf{W})\cdots(\mathbf{T}-\theta^{k-1}\mathbf{W}).$$

Proof: *As* $\mathbf{T}$ *and* $\mathbf{W}$ *commute we collect terms as*

$$\begin{aligned}&(\mathbf{T}-\mathbf{W})(\mathbf{T}-\theta\,\mathbf{W})(\mathbf{T}-\theta^2\mathbf{W})\cdots(\mathbf{T}-\theta^{k-1}\mathbf{W})=\\&\qquad=\mathbf{T}^k+d_1\mathbf{T}^{k-1}\mathbf{W}+d_2\mathbf{T}^{k-2}\mathbf{W}^2+\cdots+d_{k-1}\mathbf{T}\mathbf{W}^{k-1}+d_k\mathbf{W}^k,\end{aligned}$$

where the complex numbers $d_1,\,d_2,\,\ldots,\,d_k$ *are the coefficients of the polynomial* $(z-1)(z-\theta)(z-\theta^2)\cdots(z-\theta^{k-1})$. *But* $\{1,\,\theta,\,\theta^2,\,\ldots,\,\theta^{k-1}\}$ *is the set of roots of the polynomial* z^k-1, *as* $(\theta^j)^k=\left((\mathrm{e}^{i2\pi/k})^j\right)^k=1$ *for each* $j=0,1,\ldots,k-1$. *Therefore we conclude that* $z^k-1=(z-1)(z-\theta)\cdots(z-\theta^{k-1})$, *which implies* $d_k=-1$ *and* $d_j=0$ *for* $j=1,\,\ldots,\,k-1$.

From the multiplicative and homogeneity properties of determinants we get

Proposition 5 *If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the characteristic values of the complex linear operator $\mathbf{T}$ then ${\lambda_1}^k, {\lambda_2}^k, \ldots, {\lambda_n}^k$ are the characteristic values of the linear operator $\mathbf{T}^k$.*
Proof: *For the particular case where $\mathbf{W} = z\mathbf{I}$, proposition 4 yields*

$$\mathbf{T}^k - z^k\mathbf{I} = (\mathbf{T} - z\mathbf{I})\,(\mathbf{T} - \theta z\mathbf{I}) \cdots \left(\mathbf{T} - \theta^{k-1}z\mathbf{I}\right),$$
$$(-)^{k-1}\left(z^k\mathbf{I} - \mathbf{T}^k\right) = (z\mathbf{I} - \mathbf{T})\,(\theta z\mathbf{I} - \mathbf{T}) \cdots \left(\theta^{k-1}z\mathbf{I} - \mathbf{T}\right).$$

Taking determinants of both sides and defining $p_k(w) = \det\left(w\mathbf{I} - \mathbf{T}^k\right)$ the last equality reads $(-)^{(k-1)n}p_k(z^k) = p_1(z)p_1(\theta z)\cdots p_1(\theta^{k-1}z) =$

$$\begin{aligned}
&= (z - \lambda_1)\cdots(z - \lambda_n)\,(\theta z - \lambda_1)\cdots(\theta z - \lambda_n)\cdots(\theta^{k-1}z - \lambda_1)\cdots(\theta^{k-1}z - \lambda_n)\\
&= (-)^{kn}(\lambda_1 - z)(\lambda_1 - \theta z)\cdots(\lambda_1 - \theta^{k-1}z)\cdots(\lambda_n - z)(\lambda_n - \theta z)\cdots(\lambda_n - \theta^{k-1}z)\\
&= (-)^{kn}({\lambda_1}^k - z^k)\cdots({\lambda_n}^k - z^k)\\
&= (-)^{(k-1)n}(z^k - {\lambda_1}^k)\cdots(z^k - {\lambda_n}^k).
\end{aligned}$$

Thus $p_k(w) = (w - {\lambda_1}^k)(w - {\lambda_2}^k)\cdots(w - {\lambda_n}^k)$.

Collecting the former results we are ready to achieve the **recursive formulae** enclosed by

Theorem 2 *If $p(z) = z^n + D_1 z^{n-1} + \cdots + D_{n-1}z + D_n$ is the characteristic polynomial of the complex linear operator $\mathbf{T}$ and $T_k = \mathrm{trace}\,(\mathbf{T}^k)$ then*

$$T_k + D_1\,T_{k-1} + \cdots + D_{k-1}\,T_1 + k\,D_k = 0, \quad k = 1, 2, \ldots, n.$$

Proof: *It is a straightforward consequence of theorem 1 together with propositions 3 and 5.*

The key content of theorem 2 shall be stood out by the **trace formulae statement: the coefficients of the characteristic polynomial of a linear operator can be recursively computed as polynomial functions in the traces of its successive powers.**

The **trace formulae** $D_k = D_k(T_1, T_2, \ldots, T_{k-1}, T_k)$ for the coefficients of the characteristic polynomial are listed below for k up to 10.

$$\begin{aligned}
D_1 &= -T_1,\\
D_2 &= \tfrac{1}{2}{T_1}^2 - \tfrac{1}{2}T_2,\\
D_3 &= -\tfrac{1}{6}{T_1}^3 + \tfrac{1}{2}T_1T_2 - \tfrac{1}{3}T_3,\\
D_4 &= \tfrac{1}{24}{T_1}^4 - \tfrac{1}{4}{T_1}^2T_2 + \tfrac{1}{3}T_1T_3 + \tfrac{1}{8}{T_2}^2 - \tfrac{1}{4}T_4,\\
D_5 &= -\tfrac{1}{120}{T_1}^5 + \tfrac{1}{12}{T_1}^3T_2 - \tfrac{1}{6}{T_1}^2T_3 - \tfrac{1}{8}T_1{T_2}^2 + \tfrac{1}{4}T_1T_4 + \tfrac{1}{6}T_2T_3 - \tfrac{1}{5}T_5,\\
D_6 &= \tfrac{1}{720}{T_1}^6 - \tfrac{1}{48}{T_1}^4T_2 + \tfrac{1}{18}{T_1}^3T_3 + \tfrac{1}{16}{T_1}^2{T_2}^2 - \tfrac{1}{8}{T_1}^2T_4 - \tfrac{1}{6}T_1T_2T_3\\
&\quad + \tfrac{1}{5}T_1T_5 - \tfrac{1}{48}{T_2}^3 + \tfrac{1}{8}T_2T_4 + \tfrac{1}{18}{T_3}^2 - \tfrac{1}{6}T_6,\\
D_7 &= -\tfrac{1}{5040}{T_1}^7 + \tfrac{1}{240}{T_1}^5T_2 - \tfrac{1}{72}{T_1}^4T_3 - \tfrac{1}{48}{T_1}^3{T_2}^2 + \tfrac{1}{24}{T_1}^3T_4 + \tfrac{1}{12}{T_1}^2T_2T_3\\
&\quad - \tfrac{1}{10}{T_1}^2T_5 + \tfrac{1}{48}T_1{T_2}^3 - \tfrac{1}{8}T_1T_2T_4 - \tfrac{1}{18}T_1{T_3}^2 + \tfrac{1}{6}T_1T_6 - \tfrac{1}{24}{T_2}^2T_3\\
&\quad + \tfrac{1}{10}T_2T_5 + \tfrac{1}{12}T_3T_4 - \tfrac{1}{7}T_7,
\end{aligned} \tag{3}$$

$$
\begin{aligned}
D_8 = \; & \tfrac{1}{40320}T_1{}^8 - \tfrac{1}{1440}T_1{}^6T_2 + \tfrac{1}{360}T_1{}^5T_3 + \tfrac{1}{192}T_1{}^4T_2{}^2 - \tfrac{1}{96}T_1{}^4T_4 - \tfrac{1}{36}T_1{}^3T_2T_3 \\
& + \tfrac{1}{30}T_1{}^3T_5 - \tfrac{1}{96}T_1{}^2T_2{}^3 + \tfrac{1}{16}T_1{}^2T_2T_4 + \tfrac{1}{36}T_1{}^2T_3{}^2 - \tfrac{1}{12}T_1{}^2T_6 + \tfrac{1}{24}T_1T_2{}^2T_3 \\
& - \tfrac{1}{10}T_1T_2T_5 - \tfrac{1}{12}T_1T_3T_4 + \tfrac{1}{7}T_1T_7 + + \tfrac{1}{384}T_2{}^4 - \tfrac{1}{32}T_2{}^2T_4 - \tfrac{1}{36}T_2T_3{}^2 \\
& + \tfrac{1}{12}T_2T_6 + \tfrac{1}{15}T_3T_5 + \tfrac{1}{32}T_4{}^2 - \tfrac{1}{8}T_8, \\
D_9 = \; & -\tfrac{1}{362880}T_1{}^9 + \tfrac{1}{10080}T_1{}^7T_2 - \tfrac{1}{2160}T_1{}^6T_3 - \tfrac{1}{960}T_1{}^5T_2{}^2 + \tfrac{1}{480}T_1{}^5T_4 + \tfrac{1}{144}T_1{}^4T_2T_3 \\
& - \tfrac{1}{120}T_1{}^4T_5 + \tfrac{1}{288}T_1{}^3T_2{}^3 - \tfrac{1}{48}T_1{}^3T_2T_4 - \tfrac{1}{108}T_1{}^3T_3{}^2 + \tfrac{1}{36}T_1{}^3T_6 - \tfrac{1}{48}T_1{}^2T_2{}^2T_3 \\
& + \tfrac{1}{20}T_1{}^2T_2T_5 + \tfrac{1}{24}T_1{}^2T_3T_4 - \tfrac{1}{14}T_1{}^2T_7 - \tfrac{1}{384}T_1T_2{}^4 + \tfrac{1}{32}T_1T_2{}^2T_4 + \tfrac{1}{36}T_1T_2T_3{}^2 \\
& - \tfrac{1}{12}T_1T_2T_6 - \tfrac{1}{15}T_1T_3T_5 - \tfrac{1}{32}T_1T_4{}^2 + \tfrac{1}{8}T_1T_8 + \tfrac{1}{144}T_2{}^3T_3 - \tfrac{1}{40}T_2{}^2T_5 \\
& - \tfrac{1}{24}T_2T_3T_4 + \tfrac{1}{14}T_2T_7 - \tfrac{1}{162}T_3{}^3 + \tfrac{1}{18}T_3T_6 + \tfrac{1}{20}T_4T_5 - \tfrac{1}{9}T_9, \\
D_{10} = \; & \tfrac{1}{3628800}T_1{}^{10} - \tfrac{1}{80640}T_1{}^8T_2 + \tfrac{1}{15120}T_1{}^7T_3 + \tfrac{1}{5760}T_1{}^6T_2{}^2 - \tfrac{1}{2880}T_1{}^6T_4 - \tfrac{1}{720}T_1{}^5T_2T_3 \\
& + \tfrac{1}{600}T_1{}^5T_5 - \tfrac{1}{1152}T_1{}^4T_2{}^3 + \tfrac{1}{192}T_1{}^4T_2T_4 + \tfrac{1}{432}T_1{}^4T_3{}^2 - \tfrac{1}{144}T_1{}^4T_6 + \tfrac{1}{144}T_1{}^3T_2{}^2T_3 \\
& - \tfrac{1}{60}T_1{}^3T_2T_5 - \tfrac{1}{72}T_1{}^3T_3T_4 + \tfrac{1}{42}T_1{}^3T_7 + \tfrac{1}{768}T_1{}^2T_2{}^4 - \tfrac{1}{64}T_1{}^2T_2{}^2T_4 - \tfrac{1}{72}T_1{}^2T_2T_3{}^2 \\
& + \tfrac{1}{24}T_1{}^2T_2T_6 + \tfrac{1}{30}T_1{}^2T_3T_5 + \tfrac{1}{64}T_1{}^2T_4{}^2 - \tfrac{1}{16}T_1{}^2T_8 - \tfrac{1}{144}T_1T_2{}^3T_3 + \tfrac{1}{40}T_1T_2{}^2T_5 \\
& + \tfrac{1}{24}T_1T_2T_3T_4 - \tfrac{1}{14}T_1T_2T_7 + \tfrac{1}{162}T_1T_3{}^3 - \tfrac{1}{18}T_1T_3T_6 - \tfrac{1}{20}T_1T_4T_5 + \tfrac{1}{9}T_1T_9 \\
& - \tfrac{1}{3840}T_2{}^5 + \tfrac{1}{192}T_2{}^3T_4 + \tfrac{1}{144}T_2{}^2T_3{}^2 - \tfrac{1}{48}T_2{}^2T_6 - \tfrac{1}{30}T_2T_3T_5 - \tfrac{1}{64}T_2T_4{}^2 \\
& + \tfrac{1}{16}T_2T_8 - \tfrac{1}{72}T_3{}^2T_4 + \tfrac{1}{21}T_3T_7 + \tfrac{1}{24}T_4T_6 + \tfrac{1}{50}T_5{}^2 - \tfrac{1}{10}T_{10}.
\end{aligned}
$$

We can thus write down the following **characteristic formulae** for dimensions up to 5:

$$
\begin{aligned}
\det\,(z\mathbf{I} - \mathbf{T}) &= z - T_1 \\
\det\,(z\mathbf{I} - \mathbf{T}) &= z^2 - T_1 z + \frac{1}{2}(T_1{}^2 - T_2), \\
\det\,(z\mathbf{I} - \mathbf{T}) &= z^3 - T_1 z^2 + \frac{1}{2}(T_1{}^2 - T_2)z - \frac{1}{6}(T_1{}^3 - 3T_1T_2 + 2T_3), \\
\det\,(z\mathbf{I} - \mathbf{T}) &= z^4 - T_1 z^3 + \frac{1}{2}(T_1{}^2 - T_2)z^2 - \frac{1}{6}(T_1{}^3 - 3T_1T_2 + 2T_3)z + \\
&\quad + \frac{1}{24}(T_1{}^4 - 6T_1{}^2T_2 + 8T_1T_3 + 3T_2{}^2 - 6T_4) \qquad (4) \\
\det\,(z\mathbf{I} - \mathbf{T}) &= z^5 - T_1 z^4 + \frac{1}{2}(T_1{}^2 - T_2)z^3 - \frac{1}{6}(T_1{}^3 - 3T_1T_2 + 2T_3)z^2 + \\
&\quad + \frac{1}{24}(T_1{}^4 - 6T_1{}^2T_2 + 8T_1T_3 + 3T_2{}^2 - 6T_4)z + \\
&\quad - \frac{1}{120}(T_1{}^5 - 10T_1{}^3T_2 + 20T_1{}^2T_3 + 15T_1T_2{}^2 - 30T_1T_4 - 20T_2T_3 + 24T_5)
\end{aligned}
$$

4 Null Identities

The polynomial $p(z) = d_0 z^n + d_1 z^{n-1} + \cdots + d_{n-1} z + d_n$ with complex coefficients d_k is said to annihilate the complex linear operator $\mathbf{T}$ if $p(\mathbf{T}) = d_0 \mathbf{T}^n + d_1 \mathbf{T}^{n-1} + \cdots + d_{n-1}\mathbf{T} + d_n \mathbf{I} = \mathbf{O}$, the identically null operator.

The knowledge of the principal ideal [6] of polynomials that annihilate a linear operator $\mathbf{T}$ is essential to attain computational skill in the associative algebra generated by $\mathbf{T}$. To ensure this goal we state without proof one of the fundamental results in linear algebra.

Theorem 3 (Cayley-Hamilton [7]) *The characteristic polynomial of a linear operator annihilates it.*

Joining Cayley-Hamilton's theorem with trace formulae statement we conclude that **for each finite dimension of the underlying vector space there is a fundamental null identity in the associative algebra of linear operators.** For instance, in dimensions up to 4, the characteristic formulae (4) yield the following **null identities:**

$$
\begin{aligned}
&\mathbf{T} - T_1 \mathbf{I} = \mathbf{O},\\
&\mathbf{T}^2 - T_1 \mathbf{T} + \frac{1}{2}({T_1}^2 - T_2)\mathbf{I} = \mathbf{O},\\
&\mathbf{T}^3 - T_1 \mathbf{T}^2 + \frac{1}{2}({T_1}^2 - T_2)\mathbf{T} - \frac{1}{6}({T_1}^3 - 3T_1T_2 + 2T_3)\mathbf{I} = \mathbf{O},\\
&\mathbf{T}^4 - T_1 \mathbf{T}^3 + \frac{1}{2}({T_1}^2 - T_2)\mathbf{T}^2 - \frac{1}{6}({T_1}^3 - 3T_1T_2 + 2T_3)\mathbf{T} +\\
&\qquad + \frac{1}{24}({T_1}^4 - 6{T_1}^2 T_2 + 8T_1T_3 + 3{T_2}^2 - 6T_4)\mathbf{I} = \mathbf{O}.
\end{aligned}
$$

5 The Finite-Dimensional Resolvent Kernel

The resolvent of a complex linear operator $\mathbf{T}$ is the operator-valued function $\mathbf{R}$ defined by $\mathbf{R}(z) = (z\mathbf{I} - \mathbf{T})^{-1}$. It is a well-known fact that $\mathbf{R}$ is an operator-valued analytic function outside the spectrum of $\mathbf{T}$. We shall refine such result in the finite dimensional case showing that, in this case, $\mathbf{R}$ is an operator-valued rational function completely tied by a finite set of complex-valued rational functions, whose coefficients are exactly the trace formulae.

In this vein we need an improved version of the remainder theorem in the associative algebra of complex linear operators, whose content is the extension of proposition 1 to operator-valued polynomials.

Proposition 6 *If $p(w) = w^n + D_1 w^{n-1} + \cdots + D_{n-1} w + D_n$ is a polynomial with complex coefficients D_k and $\mathbf{T}$ and $\mathbf{W}$ are commuting complex linear operators then*

$$
\begin{aligned}
p(\mathbf{W}) - p(\mathbf{T}) = (\mathbf{W} - \mathbf{T}) \big[&\mathbf{T}^{n-1} + (\mathbf{W} + D_1\mathbf{I})\mathbf{T}^{n-2} + (\mathbf{W}^2 + D_1\mathbf{W} + D_2\mathbf{I})\mathbf{T}^{n-3} +\\
&+ \cdots + (\mathbf{W}^{n-1} + D_1\mathbf{W}^{n-2} + \cdots + D_{n-1}\mathbf{I})\mathbf{I}\big] .
\end{aligned}
$$

Proof: *by induction on the degree. For $n = 1$ we have*

$$p(\mathbf{W}) - p(\mathbf{T}) = (\mathbf{W} + D_1\mathbf{I}) - (\mathbf{T} + D_1\mathbf{I}) = \mathbf{W} - \mathbf{T}.$$

For generic n we obtain

$$\begin{aligned} p(\mathbf{W}) - p(\mathbf{T}) &= \mathbf{W}(\mathbf{W}^{n-1} + D_1\mathbf{W}^{n-2} + \cdots + D_{n-1}\mathbf{I}) + D_n\mathbf{I} + \\ &\quad -\mathbf{T}(\mathbf{T}^{n-1} + D_1\mathbf{T}^{n-2} + \cdots + D_{n-1}\mathbf{I}) - D_n\mathbf{I} \\ &= (\mathbf{W} - \mathbf{T})(\mathbf{W}^{n-1} + D_1\mathbf{W}^{n-2} + \cdots + D_{n-1}\mathbf{I}) + \\ &\quad + \mathbf{T}\Big[(\mathbf{W}^{n-1} + D_1\mathbf{W}^{n-2} + \cdots + D_{n-1}\mathbf{I}) + \\ &\qquad -(\mathbf{T}^{n-1} + D_1\mathbf{T}^{n-2} + \cdots + D_{n-1}\mathbf{I})\Big]. \end{aligned}$$

From the inductive hypothesis the right-hand side of the last equality reads

$$\begin{aligned} &(\mathbf{W} - \mathbf{T})\,(\mathbf{W}^{n-1} + D_1\mathbf{W}^{n-2} + \cdots + D_{n-1}\mathbf{I}) \\ &\quad + \mathbf{T}(\mathbf{W} - \mathbf{T})\left[\mathbf{T}^{n-2} + (\mathbf{W} + D_1\mathbf{I})\mathbf{T}^{n-3} + \cdots \right. \\ &\qquad \left. + (\mathbf{W}^{n-2} + D_1\mathbf{W}^{n-3} + \cdots + D_{n-2}\mathbf{I})\mathbf{I}\right] \\ &= (\mathbf{W} - \mathbf{T})\left[\mathbf{T}^{n-1} + (\mathbf{W} + D_1\mathbf{I})\mathbf{T}^{n-2} + \cdots + (\mathbf{W}^{n-2} + D_1\mathbf{W}^{n-3} + \right. \\ &\qquad \left. + \cdots + D_{n-2}\mathbf{I})\mathbf{T} + (\mathbf{W}^{n-1} + D_1\mathbf{W}^{n-2} \cdots + D_{n-1}\mathbf{I})\mathbf{I}\right]. \end{aligned}$$

We are ready to achieve a **rational formula** for the resolvent kernel.

Theorem 4 *If $p(w) = w^n + D_1 w^{n-1} + \cdots + D_{n-1}w + D_n$ is the characteristic polynomial of the complex linear operator $\mathbf{T}$ and z is any complex number not belonging to the spectrum of $\mathbf{T}$ then*

$$\begin{aligned} (z\mathbf{I} - \mathbf{T})^{-1} &= \frac{1}{p(z)}\mathbf{T}^{n-1} + \frac{z + D_1}{p(z)}\mathbf{T}^{n-2} + \frac{z^2 + D_1 z + D_2}{p(z)}\mathbf{T}^{n-3} + \cdots \\ &\quad + \frac{z^{n-1} + D_1 z^{n-2} + \cdots + D_{n-1}}{p(z)}\mathbf{I}. \end{aligned}$$

Proof: *It is enough to consider Cayley-Hamilton's theorem for the linear operator $\mathbf{T}$ and set $\mathbf{W} = z\mathbf{I}$ in proposition 6; notice that in such case $\mathbf{W}^k + D_1\mathbf{W}^{k-1} + \cdots + D_k\mathbf{I} = (z^k + D_1 z^{k-1} + \cdots + D_k)\mathbf{I}$ for each $k = 1, 2, \ldots, n$.*

The joint content of theorem 4 and the trace formulae statement shall be pointed out as : **for each finite dimension of the underlying vector space there is a fundamental rational formula for the resolvent kernel of a linear operator.** For instance, in dimensions up to 4, the trace formulae (3) provide the following **resolvent formulae:**

$$
\begin{aligned}
(z\mathbf{I}-\mathbf{T})^{-1} &= (z-T_1)^{-1}\mathbf{I},\\
(z\mathbf{I}-\mathbf{T})^{-1} &= \left(z^2 - T_1 z + \frac{1}{2}({T_1}^2 - T_2)\right)^{-1}\left[\mathbf{T} + (z-T_1)\mathbf{I}\right],\\
(z\mathbf{I}-\mathbf{T})^{-1} &= \left(z^3 - T_1 z^2 + \frac{1}{2}({T_1}^2 - T_2)z - \frac{1}{6}({T_1}^3 - 3T_1T_2 + 2T_3)\right)^{-1}\Bigg[\mathbf{T}^2 + \\
&\qquad + (z-T_1)\mathbf{T} + \left(z^2 - T_1 z + \frac{1}{2}({T_1}^2 - T_2)\right)\mathbf{I}\Bigg],\\
(z\mathbf{I}-\mathbf{T})^{-1} &= \Bigg(z^4 - T_1 z^3 + \frac{1}{2}({T_1}^2 - T_2)z^2 - \frac{1}{6}({T_1}^3 - 3T_1T_2 + 2T_3)z + \\
&\qquad + \frac{1}{24}({T_1}^4 - 6{T_1}^2T_2 + 8T_1T_3 + 3{T_2}^2 - 6T_4)\Bigg)^{-1}\Bigg[\mathbf{T}^3 + \\
&\qquad + (z-T_1)\mathbf{T}^2 + \left(z^2 - T_1 z + \frac{1}{2}({T_1}^2 - T_2)\right)\mathbf{T} + \Big(z^3 + \\
&\qquad - T_1 z^2 + \frac{1}{2}({T_1}^2 - T_2)z - \frac{1}{6}({T_1}^3 - 3T_1T_2 + 2T_3)\Big)\mathbf{I}\Bigg].
\end{aligned} \tag{5}
$$

6 Application to General Relativity

In general relativity the fundamental physical object is an effective geometry mathematically represented by a non-degenerate covariant tensor field $\mathbf{g}$ of the second rank, defined throughout a suitable manifold. In a local coordinate system $\mathbf{x} = (x^\alpha)$ the metric tensor can be written as $\mathbf{g} = g_{\mu\nu}\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu$.

Investigations on general relativity are frequently carried out under the assumption that there exists some background geometry $\overset{o}{\mathbf{g}} = \overset{o}{g}_{\mu\nu}\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu$. The metric properties to be assumed on $\overset{o}{\mathbf{g}}$ vary depending on the gravitational scenario. In order to perform calculations it suffices to set $\overset{o}{\mathbf{g}}$ a Ricci-flat metric, $Ric\,[\overset{o}{\mathbf{g}}] = 0$. However, to interpret the results as physically meaningful it is generally agreed that one should require $\overset{o}{\mathbf{g}}$ to be a flat metric, $Riem\,[\overset{o}{\mathbf{g}}] = 0$. Even this case is sometimes thought to be too broad, as some authors claim to set harmonic coordinates [8] or even cartesian coordinates [9]. We do not take into account here any suplementary conditions on the background geometry $\overset{o}{\mathbf{g}}$.

The effective and background geometries are related by a tensor field $\mathbf{h} = h_{\mu\nu}\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu$ by means of a connecting equation, the generally accepted form of which being $\mathbf{g} = \overset{o}{\mathbf{g}} + \mathbf{h}$. The scalar density $\sqrt{-g}\mathrm{d}^4x$ associated with $\mathbf{g}$ requires us to compute $\det(\mathbf{g}) = \det(\overset{o}{\mathbf{g}} + \mathbf{h})$. This can be easily achieved by means of

Proposition 7 *If* $\mathbf{g} = \overset{o}{\mathbf{g}} + \mathbf{h}$, $\mathbf{H} = \overset{o}{\mathbf{g}}{}^{-1}\mathbf{h}$, $H_k = \mathrm{trace}\left(\mathbf{H}^k\right)$ *and the underlying manifold is 4-dimensional then*

$$
\begin{aligned}
\frac{\det(\mathbf{g})}{\det(\overset{o}{\mathbf{g}})} = {} & 1 + H_1 + \frac{1}{2}({H_1}^2 - H_2) + \frac{1}{6}({H_1}^3 - 3H_1H_2 + 2H_3) + \\
& + \frac{1}{24}({H_1}^4 - 6{H_1}^2H_2 + 8H_1H_3 + 3{H_2}^2 - 6H_4).
\end{aligned}
$$

Proof: *From* $\mathbf{g} = \overset{o}{\mathbf{g}} + \mathbf{h} = \overset{o}{\mathbf{g}} (\mathbf{I} + \mathbf{H})$ *it follows that* $\det(\mathbf{g}) = \det(\overset{o}{\mathbf{g}}) \det(\mathbf{I} + \mathbf{H})$. *Now it is enough to consider the 4-dimensional characteristic formula in (4) with* $z = 1$ *and* $\mathbf{T} = -\mathbf{H}$; *notice that* $\mathbf{T}^k = (-)^k \mathbf{H}^k$ *implies* $T_k = (-)^k H_k$.

The Levi-Civita connection associated with $\mathbf{g}$ requires us to compute $\mathbf{g}^{-1} = \left(\overset{o}{\mathbf{g}} + \mathbf{h}\right)^{-1} = (\mathbf{I} + \mathbf{H})^{-1} \overset{o}{\mathbf{g}}^{-1}$, where $\mathbf{H} = \overset{o}{\mathbf{g}}^{-1} \mathbf{h}$. The known explicit form is given by the Neumann series [10]

$$(\mathbf{I} + \mathbf{H})^{-1} = \mathbf{I} - \mathbf{H} + \mathbf{H}^2 - \mathbf{H}^3 + \cdots$$

In local coordinates it reads

$$g^{\mu\nu} = \left(\delta^\mu{}_\beta - \overset{o}{g}{}^{\mu\alpha} h_{\alpha\beta} + \overset{o}{g}{}^{\mu\alpha} h_{\alpha\lambda} \overset{o}{g}{}^{\lambda\epsilon} h_{\epsilon\beta} - \overset{o}{g}{}^{\mu\alpha} h_{\alpha\lambda} \overset{o}{g}{}^{\lambda\epsilon} h_{\epsilon\rho} \overset{o}{g}{}^{\rho\sigma} h_{\sigma\beta} + \cdots \right) \overset{o}{g}{}^{\beta\nu},$$

where $g^{\mu\nu}$ and $\overset{o}{g}{}^{\mu\nu}$ are well-defined by $g^{\mu\alpha} g_{\alpha\nu} = \delta^\mu{}_\nu = \overset{o}{g}{}^{\mu\beta} \overset{o}{g}{}_{\beta\nu}$.

Besides convergence requirements on the above series, we stress that such expression clearly leads to technical difficulties when developing a Lagrangian variational formalism. These problems were already dealt with in the literature, the proposed solution being to modify the above form of the connecting equation [11]. We show how to completely overcome such drawbacks by means of

Proposition 8 *If* $\mathbf{g} = \overset{o}{\mathbf{g}} + \mathbf{h}$, $\mathbf{H} = \overset{o}{\mathbf{g}}^{-1} \mathbf{h}$, $H_k = \operatorname{trace}\left(\mathbf{H}^k\right)$ *and the underlying manifold is 4-dimensional then*

$$\begin{aligned}\mathbf{g}^{-1} = \Big(& 1 + H_1 + \frac{1}{2}(H_1{}^2 - H_2) + \frac{1}{6}(H_1{}^3 - 3H_1H_2 + 2H_3) + \\ & + \frac{1}{24}(H_1{}^4 - 6H_1{}^2H_2 + 8H_1H_3 + 3H_2{}^2 - 6H_4)\Big)^{-1} \Big[- \mathbf{H}^3 + \\ & + (1 + H_1)\mathbf{H}^2 - \left(1 + H_1 + \frac{1}{2}(H_1{}^2 - H_2)\right) \mathbf{H} + \Big(1 + \\ & + H_1 + \frac{1}{2}(H_1{}^2 - H_2) + \frac{1}{6}(H_1{}^3 - 3H_1H_2 + 2H_3)\Big) \mathbf{I} \Big] \overset{o}{\mathbf{g}}^{-1}.\end{aligned}$$

Proof: *From* $\mathbf{g} = \overset{o}{\mathbf{g}} + \mathbf{h} = \overset{o}{\mathbf{g}} (\mathbf{I} + \mathbf{H})$ *it follows that* $\mathbf{g}^{-1} = (\mathbf{I} + \mathbf{H})^{-1} \overset{o}{\mathbf{g}}^{-1}$. *Now it is enough to consider the 4-dimensional resolvent formula in (5) with* $z = 1$ *and* $\mathbf{T} = -\mathbf{H}$; *notice that* $\mathbf{T}^k = (-)^k \mathbf{H}^k$ *implies* $T_k = (-)^k H_k$.

Acknowledgements

The author thanks Dr. Renato Klippert Barcellos for having collaborated with him in this work, for valuable discussions, and for the implementation of the computer algebra algorithm. The author owes special thanks to Dr. Nelson Pinto Neto and Dr. Vitorio Alberto De Lorenci for the encouragement, and for having drawn the attention of him to the relevance of the subject of the present work. The author is also indebted to Carla da Fonseca-Bernetti for helpful corrections of the manuscript.

This work was financially supported by CNPq.

References

[1] J. V. Uspensky, *Theory of Equations* p. 293, McGraw-Hill Book Company, NY (1948).

[2] S. Lang, *Algebra* rev. printing p. 169, Addison Wesley Publishing Company, London (1971).

[3] B. L. van der Warden, *Algebra* Vol. I 5$^{\text{th}}$ ed. p. 99, Springer-Verlag, NY (1991).

[4] See Uspensky, *op. cit.* p. 260.

[5] S. Mac Lane and G. Birkhoff, *Algebra* 4$^{\text{th}}$ printing, p. 310, The MacMillan Company, London (1970); see also Lang, *op. cit.* p. 399.

[6] See Mac Lane and Birkhoff, *op. cit.* p. 152 and Lang, *op. cit.* p. 58.

[7] See Mac Lane and Birkhoff, *op. cit.* p. 316 and Lang, *op. cit.* p. 400.

[8] V. Fock, *The Theory of Space, Time and Gravitation*, 2$^{\text{nd}}$ rev. ed. p. 369, Pergamon Press, Oxford (1966).

[9] L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics V. 2: The Classical Theory of Fields*, 4$^{\text{th}}$ rev. english ed. p. 330, Butterworth-Heinemann Ltd, Oxford (1975).

[10] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I. Functional Analysis*, p. 191, Academic Press, NY (1972).

[11] L. P. Grishchuk, A. N. Petrov and A. D. Popova, *Commun. Math. Phys.* **94**, 379 (1984).