Reforming the Wishart characteristic function

Eberhard Mayerhofer,
University of Limerick, Department of Mathematics and Statistics,
Castletroy, County Limerick, T94 T9PX Ireland.
e-mail: eberhard.mayerhofer@ul.ie

Abstract: The literature presents the characteristic function of the Wishart distribution on $m \times m$ matrices as an inverse power of the determinant of the Fourier variable, the exponent being the positive, real shape parameter.

I demonstrate that only for $2 \times 2$ matrices, this expression is unambiguous – in this case the complex range of the determinant excludes the negative real line. When $m \geq 3$, the range of the determinant contains closed lines around the origin, hence a single branch of the complex logarithm does not suffice to define the determinant’s power. To resolve this issue, I give the correct analytic extension of the Laplace transform, by exploiting the Fourier-Laplace transform of a Wishart process.

MSC 2010 subject classifications: Primary 62H10; secondary 60E10, 60B20.

Keywords and phrases: Wishart distribution, characteristic function, analytic continuation.

1. Introduction and Statement of Theorem

One of the best known multivariate distributions is the Wishart distribution $p(d\xi; \alpha)$, explored first by Wishart in 1928 as the distribution of the covariance matrix of a normally distributed sample [21]:

1. For $\alpha$ being a positive half-integer $n/2$, it is the distribution of the covariance of a sample of size $n$ from the $m$-variate standard normal distribution. More explicitly, let $x_1, \ldots, x_n$ be an i.i.d. sequence of standard normally distributed $m$ vectors. Then the sum of (almost surely) rank one matrices

$$x_1x_1^\top + \ldots + x_nx_n^\top,$$

is Wishart distributed with shape parameter $n/2$.

2. For shape parameters $\alpha > (m - 1)/2$, it can be defined in terms of its density:

$$p(\xi; \alpha) = \frac{\det(\xi)^{\alpha - \frac{m+1}{2}} e^{-\text{tr}(\xi/2)}}{2^{m\alpha} \Gamma_m(\alpha)}.$$

1 Using standard normal distributions (zero mean, zero correlations) amounts to the standardization with a scale (matrix) parameter $V = I_m$ and avoids a series expansions of the density in terms of zonal polynomials.

2 All notation is summarized in the end of the section.
The above range of shape parameters $\alpha$ is indeed maximal (see, e.g., [20, 17, 10, 14]); the Wishart distribution exists, if and only if $\alpha$ lies in

$$\Lambda_m := \left\{ \frac{1}{2}, \ldots, \frac{m-2}{2} \right\} \cup \left[ \frac{m-1}{2} \right].$$

This set is typically referred to as Gindikin ensemble (in honor of Gindikin who discovered this parameterization in the more general setting of homogeneous cones ([9])).

The characteristic function $\int_{S_m} e^{i \text{tr}(\xi v)} p(d\xi; \alpha)$ of the Wishart distribution is typically quoted as the “map”

$$\Phi_\alpha : S_m \rightarrow \mathbb{C} : v \mapsto \det(I_m - 2iv)^{-\alpha}. \tag{1.2}$$

Many (standard) references in multivariate statistics state and prove this formula, at least for positive half-integers $\alpha$, e.g. Anderson (18989 citations)$^4$, who claims that it can be shown that formula (1.2) holds, in general (that is, for appropriate Fourier-Laplace variables, see [1, 7.3.1, eq. (10) and the subsequent paragraph]), Muirhead (5127 citations) [18, Theorem 10.3.3], Gupta (1306 citations) [11, Theorem 3.3.7], Eaton (839 citations) [6, Proposition 8.3 (iii)] and also the already mentioned work [20].

To my knowledge, the literature does not offer a rigorous definition of the determinant’s power in (1.2) (e.g. through the complex power). Most of the references use explicitly the product formula of the characteristic function to reduce the computation of the characteristic function to a sample of size one: To obtain the formula for general half-integer $\alpha = n/2$, they thus compute

$$\left( \det(I_m - 2iv)^{-\frac{1}{2}} \right)^n = \det(I_m - 2iv)^{-\alpha}$$

thereby overlooking that the exponential has a complex period, and therefore the general complex power does in general not satisfy the functional equation $z^{\alpha} z^{\beta} = z^{\alpha+\beta}$. This brief note exemplifies the issues arising with the formula (1.2), and it provides a clean interpretation for $\Phi_\alpha$, without using the complex power:

**Theorem 1.1.** The characteristic function of the Wishart distribution with shape parameter $\alpha \in \Lambda_m$ is given by

$$\int_{S_m} e^{i \text{tr}(\xi v)} p(d\xi; \alpha) = e^{\alpha \int_0^1 \text{tr}(I_m - 2iv)^{-1}(2iv)dt}, \quad v \in S_m.$$

In the following, I summarize some notation used in this paper, and compute explicitly some characteristic functions, to demonstrate the issue. These are followed by a proof of the Theorem using the Fourier-Laplace transform of Wishart processes, and some final remarks.

---

$^3$In his introductory paper [21], Wishart himself uses densities rather than characteristic functions.

$^4$Google Scholar Data.
1.1. Notation

In this paper, $\mathbb{R}$ denotes the real line, $i = \sqrt{-1}$ the imaginary unit. For integers $m \geq 1$ I use the following matrix spaces: $S_m$, the real valued symmetric $m \times m$ matrices, $S^+_m$ the positive semidefinite ones, and $S^{++}_m$ the positive definite matrices. $I_m$ is the unit $m \times m$ matrix, and for a given matrix $A$, $\det(A)$ denotes its determinant and $\text{tr}(A)$ its trace.

The main branch of the logarithm of a complex number $z$ is defined implicitly by $z = re^{i\varphi}$, where $\varphi \in (-\pi, \pi]$, in other words, $\log(z) = \log(r) + i\varphi$.

2. The problem emerges with size

No issues arise for $m = 1$, where the Wishart distribution coincides with the Gamma distribution with characteristic function

$$(1 - iv\theta)^{-k},$$

where $\theta = 2$ and $k = \alpha$. In this case it is obvious that $1 - 2iv$ has only strictly positive real part, whence for general real shape parameter $k$, the characteristic function should be understood by using the main branch of the logarithm,

$$(1 - 2iv)^{-k} = e^{-k\log(1 - 2iv)}.$$

Similarly, for dimension $m = 2$, an explicit computation shows that the range of $\det(I_2 - 2iv)$ does not contain the ray $(-\infty, 1)^5$. Therefore, again, with the main branch of the complex logarithm,

$$\det(I_2 - 2iv)^{-\alpha} := e^{-\alpha \log(\det(I_2 - 2iv))}.$$  

But for any dimension $m \geq 3$, the range of the map $\det(I_m - 2iv)$ contains the curve

$$c : [-\sqrt{3}, \sqrt{3}], \quad c(t) := (1 - it)^3$$

(set $2v = \text{diag}(t, t, t, 0, \ldots, 0)$). This curve starts in the complex plane at $(-8, 0)$, goes through third and fourth quadrant, intersecting the real axis at $(1, 0)$, and continues as its mirror image through first and second quadrant, terminating at its starting point $(-8, 0)$.

Therefore, the definition

$$\det(I_m - iv)^{-\alpha} := \exp(-\alpha \log(\det(I_m - 2iv)), \quad v \in S_m$$

of the characteristic function is ambiguous and a-priori it is not clear that it is complex analytic, because the power is only analytic on a simply connected domain.

---

5 In fact, it does not contain the open parabola $y^2 < 4(1 - x)$, which can be shown by computing the determinant using the eigenvalues of $v$. 

---
Making the above example more concrete, let me use the main branch of the complex logarithm to define 

\[ z^\alpha = e^{\alpha \log(z)} . \]

Then I get for \( t = \pm \sqrt{3}, (1 \pm \sqrt{3}i)^3 = -8 \). Therefore, for \( v = \mp \sqrt{3}I \) and \( \alpha = 1/2 \),

\[
\det(I_3 - 2iv)^{1/2} = ((1 \pm i\sqrt{3})^3)^{1/2} = -8 = 2\sqrt{2}(e^{i\pi})^{1/2} = 2\sqrt{2}i. \tag{2.2}
\]

But this should be the reciprocal value of the characteristic function of the sum of three independent gamma distributed random variables (the diagonal elements of a Wishart distributed random matrix), hence be of the form

\[
((1 \pm i\sqrt{3})^{1/2})^3 = (\sqrt{2}e^{\pm i\pi/3})^3 = \pm 2\sqrt{2}i
\]

which obviously does only agree with (2.2) for the Fourier variable \( v_- \). So for different Fourier variable one needs different branches to get the correct value of the characteristic function. Similar examples can be produced for arbitrary dimension \( m \geq 4 \) and for other values of \( \alpha \).

Of course, the right side of (2.1) is well-defined, if one is willing to restrict the values of the Fourier variable to be of sufficiently small size, relative to \( \alpha \), because in a complex neighborhood of the identity matrix, the power of the determinant will be complex analytic.

Beyond that, as suggested by the above examples, the expression \( \det(I_m - iv)^{-\alpha} \) could be interpreted by using the eigenvalues of the Fourier variable \( u \), due to the fact that the determinant is invariant under actions of the orthogonal group. But this approach will lead to a fairly messy definition. The literature has done computation along these lines (see [19] and the references therein). The above examples challenge the results of [19], however. Besides, the complex analyticity of the characteristic function - as is proved below - cannot easily be read off from a formula that uses a multitude of branches of the complex logarithm.

The next section resolves the problem, by proving the semi-explicit formula for the characteristic function stated in Theorem 1.1.

### 3. Proof of Theorem 1.1

There is an elegant way of interpreting the determinant formula which is inspired by the theory of Wishart processes (introduced by [2], extended to a more general context in [3, 4]). The Wishart process \( X = (X_t)_{t \geq 0} \) (here without linear drift term) is the solution to the stochastic differential equation (SDE)

\[
dX_t = \sqrt{X_t} dB_t + dB_t^T \sqrt{X_t} + 2\alpha I_m dt, \quad X_0 = x \in S_m^+,
\]

where all products are in terms of the matrix multiplication, and \( B \) is an \( m \times m \) matrix of standard Brownian motions. This SDE indeed has a unique global weak solution if \( \alpha \geq (m-1)/2 \), or \( \text{rank}(x) \leq 2\alpha \), and \( \alpha \in \{1/2, \ldots, (m-2)/2\} \), see [10, Theorem 1.3]. Furthermore, by [16, Proposition 2.8 and Lemma 2.9] for any \( u \in S_m^+ \)

\[
\mathbb{E}[e^{-\text{tr}(uX_t)}] = e^{-\phi(t,u)-\text{tr}(\psi(t,u)x)}, \quad t \geq 0,
\]
where the matrix valued function $\psi(t, u)$ satisfies the initial value problem

$$\partial_t \psi(t, u) = -2\psi(t, u)^2, \quad \psi(t = 0, u) = u,$$

and thus is of the explicit form $\psi(t, u) = (I_m + 2tu)^{-1}u$. The function $\phi(t, u)$ satisfies the trivial equation

$$\partial_t \phi(t, u) = 2\alpha \text{tr}(\psi(t, u)), \quad \phi(t = 0, u) = 0,$$

and thus equals

$$\phi = \alpha \log(\det(I_m + 2tu)). \quad (3.1)$$

In particular, for $x = 0$, the Wishart SDE has unique global weak solution (3.1), hence it satisfies at $t = 1$

$$e^{-\phi(1, u)} = \det(I_m + 2u)^{-\alpha}.$$

This is exactly the Laplace transform of the Wishart distribution.

Consider now the (trivial) initial value problem for the function $\phi(t, u - iv)$:

$$[0, 1] \times D \rightarrow \mathbb{C}, \quad D = \{u - iv \in S_m + iS_m \mid u \in -I_m/2 + S^+\}.$$

$$\partial_t \phi(t, u - iv) = \alpha \text{tr}((I_m + 2t(u - iv))^{-1}(2u - 2iv)), \quad \phi(t = 0, u - iv) = 0.$$

The IVP defined in (3.2) has also unique solutions for any $u \in D$ with blow-up beyond $t = 1$, because $(I_m + 2t(u - iv))$ is non-singular for each $t \in [0, 1].$ Therefore, the solution of (3.2) can be written semi-explicitly as

$$\phi(1, u + iv) := \alpha \int_0^1 \text{tr}((I_m + 2t(u - iv))^{-1}(2u - 2iv))dt.$$

This solution is complex analytic in $d(d + 1)/2$ variables (the entries of $u - iv$), because the right-side of (3.2) is (cf. [5, Theorem 10.8.2]). Hence

$$e^{-\phi(1,u-iv)}$$

constitutes an analytic extension of $\det(I_m + 2u)^{-\alpha}$ to the complex strip $D$. Now $e^{-\phi(1,u-iv)}$ and $\int_{S^+_m} e^{-\text{tr}((u - iv)\xi)} p(\xi; \alpha) d\xi$ are both complex analytic in several variables on $D$, and the two functions agree for $v = 0$ and $u \in S^+_m$. An open domain in $S_m$ is a set of uniqueness, hence the two functions agree (cf. [5, (9.4.4)]).

The proof of Theorem 1.1 is complete.

---

6By Ostrowski and Tausski (cf. [12, (4.4)]), a complex matrix $X + iY$, where $X, Y$ are real and $X$ is positive definite, satisfies $|\det(X + iY)| \geq \det(X)$. 
4. Final remarks

The results of this paper can easily be extended to the non-central Wishart distribution, the difference being a multiplicative exponential term in the characteristic function (instead of $x = 0$ use in Section 3 a general starting point $X_0 = x$ for the Wishart process, respecting the rank condition, see [10]).

Analytic continuation arguments similar to the above are extensively used in the recent literature on affine Markov processes, e.g. in the papers [3, 4, 8, 13].

References

[1] Anderson, T.W. An introduction to multivariate statistical analysis. New York: Wiley, 1958.
[2] Marie-France Bru. Wishart processes. Journal of Theoretical Probability 4, no. 4 (1991): 725-751.
[3] Christa Cuchiero, Damir Filipović, Eberhard Mayerhofer and Josef Teichmann. Affine processes on positive semidefinite matrices. The Annals of Applied Probability, 21(2), 397–463, 2011.
[4] Christa Cuchiero, Martin Keller-Ressel, Eberhard Mayerhofer, and Josef Teichmann. Affine Processes on Symmetric Cones. Journal of Theoretical Probability, 29(2), 359–422, 2016.
[5] Jean Dieudonné. Foundations of modern analysis. Pure and Applied Mathematics, P. Smith and S. Eilenberg, Eds. New York: Academic Press 10, 1969.
[6] Morris L Eaton. Multivariate statistics: a vector space approach. John Wiley & Sons, Inc., 605 Third Ave., New York 10158, 1983.
[7] Jacques Faraut and Adam Korányi. Analysis on symmetric cones. Oxford: Clarendon Press, 1994.
[8] Damir Filipović and Eberhard Mayerhofer. Affine diffusion processes: Theory and applications. Advanced Financial Modelling 8 (2009): 1-40.
[9] Simon G Gindikin. Invariant generalized functions in homogeneous domains. Functional Analysis and its Applications, 9(1), 50–52, 1975.
[10] Piotr Graczyk, Jacek Malecki, and Eberhard Mayerhofer. A Characterization of Wishart Processes and Wishart Distributions. Stochastic Processes and their Applications, 128(4), 1386–1404, 2018.
[11] Gupta, Arjun K., and Daya K. Nagar. Matrix variate distributions. Chapman and Hall/CRC, 2018.
[12] Charles R Johnson. Matrices whose Hermitian part is positive definite. Diss. California Institute of Technology, 1972.
[13] Martin Keller-Ressel and Eberhard Mayerhofer. Exponential moments of affine processes. The Annals of Applied Probability 25, no. 2 (2015): 714-752.
[14] Gérard Letac and Hélène Massam. The Laplace transform $(\det s)^{-p}\exp tr(s^{-1}w)$ and the existence of non-central Wishart distributions, Journal of Multivariate Analysis 163, 96–110, 2018.
[15] Eberhard Mayerhofer. On the existence of non-central Wishart distributions. Journal of Multivariate Analysis 114 (2013): 448-456.

[16] Eberhard Mayerhofer. Wishart Processes and Wishart Distributions: An Affine Processes Point of View. Modern methods in multivariate statistics, Travaux en Cours, Hermann, Paris, 2014.

[17] Eberhard Mayerhofer. On Wishart and noncentral Wishart distributions on symmetric cones. Transactions of the American Mathematical Society, in print (2019).

[18] Robb J. Muirhead. Aspects of Multivariate Statistical Theory, volume 197. John Wiley & Sons, New York, 2005.

[19] Albert H Nuttall. Closed Form Characteristic Function for general Complex Second-Order Form in Correlated Complex Gaussian Random Variables. Naval Undersea Warfare Center Detachment, New London, Connecticut, Technical Report 10, 403 (1993).

[20] Shyamal Das Peddada and Donald St. P. Richards. Proof of a Conjecture of ML Eaton on the Characteristic Function of the Wishart distribution. The Annals of Probability, 19(2), 868–874, 1991.

[21] John Wishart. The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population. Biometrika, 20A, 32–52, 1928.