
DOI: https://doi.org/10.4213/tvp22
THE MOMENTS OF WISHART PROCESSES
VIA ITÔ CALCULUS

Prиводятся формулы для моментов и математических ожидалей некоторых функционалов от процесса Уишарта \( X = (X_t)_{t \geq 0} \). Эти формулы выводятся с помощью исчисления Ито для матричных процессов. Полагая \( t = 1 \), можно немедленно получить аналогичные результаты для центрированных и нецентрированных матриц Уишарта.

Ключевые слова и фразы: закон Уишарта, матрица Уишарта, процесс Уишарта, моменты.

1. Introduction. The Wishart law appears in multivariate statistics as the law of the empirical variance-covariance matrix in Gaussian models. It is one of the reasons why it plays a very important role in multivariate statistics (cf. [13]). Consequently, the moments of the Wishart laws are involved in some approximation procedures related to the computation of tolerance factors of a population distributed according to a Gaussian multivariate law (cf. [10]). The knowledge of the moments is necessary in order to evaluate the power of correlation between the variables. Some moments of Wishart laws appear also in the recent theory of graphical Markov models. Wishart laws are useful in finances, namely in modelling of stochastic volatility [6], [5].

The computation of moments of centered Wishart matrices and of multilinear functionals of centered Wishart matrices (involving products of traces) has been studied by many authors for a long time (cf., for example, [11], [7], [12], and the references therein) using methods based on the Laplace transform and the representation theory of the symmetric group. The case of noncentered real Wishart matrices is more difficult and has not been studied sufficiently up to now. In this article we consider the Wishart processes \( (X_t)_{t \geq 0} \) and we pass from the «statical» point of view to the «dynamical» one. The link between the two cases is simple: just fix \( t \), for example \( t = 1 \), in Wishart process and we obtain Wishart matrix. But the difference

* LAREMA, Département de Mathématiques, Université d'Angers, France; e-mail: graczyk@univ-angers.fr; vostrik@univ-angers.fr
between the two cases is fundamental: while introducing stochastic processes we can benefit from the powerful tools of Itô calculus.

Suppose that \( B = (B_s)_{s \geq 0} \) is a Brownian \( n \times m \) matrix, that is, a random process with values in the matrix space \( \mathcal{M}_{n,m}(\mathbb{R}) \) whose components are independent standard real-valued Brownian motions.

For \( s \geq 0 \) define

\[
X_s = B_s^T B_s,
\]

where \( ^T \) denotes a transposition. Then the process \( X = (X_s)_{s \geq 0} \) is a diffusion with values in the space of nonnegative definite symmetric matrices \( S^+_m(\mathbb{R}) \) and is called the Wishart process with parameters \( (n, m, X_0) \). It can be shown that for \( n \geq m \) and \( s > 0 \), \( \det(X_s) > 0 \) \( \mathbb{P} \)-a.s. and that \( X_s \) is invertible matrix. Then we can define

\[
dV_s = X_s^{-1/2}B_s^T dB_s
\]

and by the Lévy characterization of the Brownian motion, \( V = (V_s)_{s \geq 0} \) is a Brownian \( m \times m \) matrix. Finally, we can easily see that the process \( X = (X_s)_{s \geq 0} \) satisfies the stochastic differential equation

\[
dX_s = \sqrt{X_s} dV_s + dV_s^T \sqrt{X_s} + nI_m dt,
\]

where \( I_m \) is the identity matrix of size \( m \times m \).

In the particular case \( m = 1 \), the Brownian matrix \( B_s = (B_s^{(1)}, \ldots, B_s^{(n)})^T \) is a Brownian vector with values in \( \mathbb{R}^n \) and \( X_s = (B_s^{(1)})^2 + \cdots + (B_s^{(n)})^2 \). Thus the process \( X \) is a Bessel Squared process with parameter \( n \), and the equation (1) becomes

\[
dX_s = 2\sqrt{X_s} d\beta_s + n ds,
\]

where \( \beta = (\beta_s)_{s \geq 0} \) is a one-dimensional Brownian motion.

More general Wishart processes are introduced in [1]. Let \( \alpha > 0 \) and let \( V \) be a Brownian \( m \times m \) matrix. Wishart process \( X = (X_t)_{t \geq 0} \) with parameters \( (\alpha, m, x_0) \) is defined as a solution, when it exists, of the equation

\[
\begin{cases}
    dX_t = \sqrt{X_t} dV_t + dV_t^T \sqrt{X_t} + \alpha I_m dt,
    \\
    X_0 = x_0.
\end{cases}
\]

It was shown that if \( x_0 \in S^+_m(\mathbb{R}) \) and \( \alpha \in \{1, 2, \ldots, m-1\} \cup (m-1, m+1) \), then there exists a unique weak solution of (2) and if \( \alpha \geq m+1 \), then there exists a unique strong solution of (2).

One of important properties of Wishart process is the following scaling property: for all \( c > 0 \) we have

\[
\mathcal{L}(c^{-1} X_{ct}^{x_0})_{t \geq 0} = \mathcal{L}(X_t^{x_0})_{t \geq 0}.
\]
This property follows by time change from the unicity of the solution of the equation (2) (cf. [15]).

Many other important properties of Wishart process are given in [1]. Namely, the eigenvalues of the process \( X = (X_t)_{t \geq 0} \) starting from \( x_0 \in S^+_m(\mathbb{R}) \) never collide, i.e., almost surely \( \lambda_1(t) > \cdots > \lambda_m(t) \geq 0 \) for all \( t > 0 \). If \( \alpha > m + 1 \), then \( \lambda_m(t) > 0 \) for all \( t > 0 \) almost surely and they are solutions of the following stochastic differential system:

\[
d\lambda_i(t) = 2\sqrt{\lambda_i(t)} \, d\beta_i(t) + \left( \alpha + \sum_{k \neq i} \frac{\lambda_i(t) + \lambda_k(t)}{\lambda_i(t) - \lambda_k(t)} \right) dt, \quad i = 1, \ldots, m,
\]

where \( \beta_1(\cdot), \ldots, \beta_m(\cdot) \) are independent standard Brownian motions. If \( \alpha \geq m + 1 \) we have also stochastic differential equations for the determinant \( D_t = \det(X_t) \) of \( X_t \) and its logarithm:

\[
dD_t = 2D_t \sqrt{\text{tr} \left( X_t^{-1} \right)} \, d\beta(t) + (\alpha - m + 1) D_t \, \text{tr} \left( X_t^{-1} \right) dt,
\]

\[
d(\ln D_t) = 2\sqrt{\text{tr} \left( X_t^{-1} \right)} \, d\beta(t) + (\alpha - m - 1) \, \text{tr} \left( X_t^{-1} \right) dt,
\]

where \( \beta(\cdot) \) is a standard Brownian motion. The Laplace transform of \( X_t \) coming from a Wishart process \( X = (X_t)_{t \geq 0} \) with parameters \( (\alpha, m, x_0) \) and \( x_0 \in S^+_m(\mathbb{R}) \) is given by

\[
E \exp \left( - \text{tr} \left( \theta X_t \right) \right) = (\det(I + 2t\theta))^{-\alpha/2} \exp \left( - \text{tr} \left( x_0(I + 2t\theta)^{-1} \theta \right) \right),
\]

where \( \theta \in S^+_m(\mathbb{R}) \). This expression implies that \( X_t \) has a central Wishart distribution when \( x_0 = 0 \), and a noncentral one when \( x_0 \neq 0 \).

As we observed above, the Wishart process \( X \) is a matrix generalization of Bessel Squared processes. This has incited very recently a strong interest in studying Wishart processes. In [4] the authors prove for Wishart process some well-known properties of BESQ processes. The absolute continuity of the laws of Wishart processes with parameters \( (\alpha, m, x_0) \) when \( x_0 \) is strictly positive definite symmetric matrix and \( \alpha \geq m + 1 \) is established. Namely for any \( \nu > 0 \) and \( t > 0 \)

\[
\frac{dP_t^{m+1+2\nu}}{dP_t^{m+1}}(X) = \left( \frac{\det(X_t)}{\det(x_0)} \right)^{\nu/2} \exp \left( -\frac{\nu^2}{2} \int_0^t \text{tr} \left( X_s^{-1} \right) ds \right),
\]

where \( P_t^{(\alpha)} \) is the restriction of the law of the process \( X \) on \( \mathcal{F}_t \), where \( \mathcal{F} = (\mathcal{F}_t)_{t \geq 0} \) is natural filtration. The generalization of Hartman–Watson law was also obtained, namely the expression for conditional law \( P^{(m+1)}(\exp((-\nu^2/2) \int_0^t \text{tr} \left( X_s^{-1} \right) ds) \mid X_t = y) \) via hypergeometrical functions. Analogous problems for Wishart processes on Hermitian positive definite matrices, called Laguerre processes, were investigated in [2]. Large deviations for Wishart processes are studied in [3].
Our article is devoted to the moments and the expectations of some functionals of Wishart process. Main result, given in Section 3, shows that the moments of the Wishart matrices may be computed by recurrence, by multiplying some real matrices given explicitly. Our techniques are based on the Itô calculus and allow one to compute the moments of the form $\mathbb{E}X^k_t$ as well as the expectation of the polynomials

$$R_I(X_t) = (\text{tr}(X_t))^{i_1}(\text{tr}(X^2_t))^{i_2} \cdots (\text{tr}(X^{k+1}_t))^{i_{k+1}}X^j_t,$$

where $I$ is a multiindex $I = (j, \mathbf{i}) = (j, i_1, i_2, \ldots, i_{k+1})$. Our results are new in the case of noncentered Wishart distributions corresponding to $x_0 \neq 0$. In the centered case corresponding to $x_0 = 0$, some of our results are new. For example, the computation of $\mathbb{E}(\text{tr}(X^p_t) X^q_t)$, $p \geq 2$, $q \geq 1$, is not covered by the methods of [11] (see [11, p. 18]), but is available by our techniques. The proposed methods could be also useful in the complex case of Hermitian Wishart matrices.

From the point of view of a statistician, it would be also interesting to study the same questions for the process $\tilde{X} = (\tilde{X}_t)_{t \geq 0}$ with

$$\tilde{X}_t = \sigma X_t \sigma,$$

where $\sigma^2$ is covariance matrix. If $\sigma^2$ is known and invertible then, of course, one can multiply $\tilde{X}_t$ on the left and on the right by $\sigma^{-1}$ in order to obtain covariance matrix equal to $I_m$. Although the situation with $\sigma^2$ known may seem not very realistic, it may occur that a precise estimation of $\sigma^2$ had already been obtained in the past. These arguments make the case $\sigma^2 = I_m$ interesting from statistical point of view. The case $\sigma^2 \neq I_m$ will be studied in a forthcoming paper.

The paper is organized as follows. After a brief presentation of necessary tools of Itô calculus on matrices in Section 2, we give the main results on the moments in Section 3. The most important result is contained in Theorem 3.1. In the particular case of Squared Bessel process we obtain as a corollary the known expression for the moments of Squared Bessel process via Laguerre polynomials. In Section 4 we give some illustrative examples of computation of the moments of Wishart process.

2. Preliminaries. Let $(\Omega, \mathcal{F}, \mathbf{F}, \mathbf{P})$ be a probability space equipped with a filtration $\mathbf{F} = (\mathcal{F}_t)_{t \geq 0}$ completed with sets from $\mathcal{F}$ having $\mathbf{P}$-probability zero. To perform Itô calculus on matrix-valued processes we will interprete usual semimartingale notions (see [9] for details) in component-by-component way.

We say that a process $X$ with values in $\mathcal{M}_m(\mathbb{R})$ is a martingale (local martingale, semimartingale) if each component of $X$ is a martingale (local martingale, semimartingale) with values in $\mathbb{R}$. We say that $X$ with values
in $\mathcal{M}_m(\mathbb{R})$ is predictable and of finite variation (increasing, bounded, locally integrable) if each real-valued component of $X$ is so. Let $X = (X_t)_{t \geq 0}$ and $Y = (Y_t)_{t \geq 0}$ be two semimartingales with values in the space of $m \times m$ real matrices $\mathcal{M}_m(\mathbb{R})$. Let us denote the martingale parts of $X$ and $Y$ by $X^c$ and $Y^c$, respectively.

**Definition 2.1.** The quadratic variation $\langle X, Y \rangle$ is an $\mathcal{M}_m(\mathbb{R})$-valued locally integrable predictable stochastic process such that $X^c Y^c - \langle X, Y \rangle$ is a local martingale.

**Proposition 2.1** (cf. [14]). Let $X$ and $Y$ be two $\mathcal{M}_m(\mathbb{R})$-valued continuous semimartingales on $(\Omega, \mathcal{F}, \mathcal{F}, P)$. The following integration by parts formula holds:

$$d(X_t Y_t) = (dX_t) Y_t + X_t dY_t + d\langle X, Y \rangle_t,$$

where $\langle X, Y \rangle_t$ is given by

$$\langle X, Y \rangle_t^{ij} = \sum_{k=1}^{n} \langle X^{ik}, Y^{kj} \rangle_t.$$  

**Proof.** Apply the usual integration by parts Ito formula to each term $(X_t Y_t)^{ij} = \sum_{k=1}^{n} X_t^{ik} Y_t^{kj}$.

**Remark.** It should be noticed that $\langle X, Y \rangle$ is not, in general, a symmetric matrix. It is symmetric if $XY = YX$.

In order to perform the Ito calculus for matrix-valued processes we need also the expressions for predictable brackets of stochastic integrals with respect to Brownian matrix. To simplify the notation for integration with respect to Brownian matrix $V$ we write

$$(A \cdot V B)_t = \int_0^t A_s dV_s B_s,$$

where $A$, $B$ are predictable locally bounded matrix-valued processes. The same notation will be used for the integration with respect to the components of Brownian matrix and also for the integrals with respect to continuous martingales related to $V$.

**Proposition 2.2.** Let

$$A = (A_s)_{s \geq 0}, \quad B = (B_s)_{s \geq 0}, \quad C = (C_s)_{s \geq 0}, \quad D = (D_s)_{s \geq 0}$$

be predictable locally bounded processes with the values in $\mathcal{M}_m(\mathbb{R})$ and let $V$ be a Brownian matrix with values in $\mathcal{M}_m(\mathbb{R})$. Then

$$\langle A \cdot V B, C \cdot V D \rangle_t = \int_0^t A_s (B_s C_s)^T D_s ds,$$

$$\langle A \cdot V B, C \cdot V^T D \rangle_t = \int_0^t A_s D_s \text{tr} (B_s C_s) ds.$$
Since $V^T$ is also a Brownian $m \times m$ matrix, the same formulas are true with replacing of $V$ by $V^T$ at each position of $V$ in considered formulas.

Proof. Let us prove (5). We have for $1 \leq i, j \leq m$

$$\langle A \cdot VB, C \cdot VD \rangle_t^{ij} = \sum_{k=1}^{m} \langle (A \cdot VB)^{ik}, (C \cdot VD)^{kj} \rangle_t$$

$$= \sum_{1 \leq k, l, r, l', r' \leq m} \langle A_{il} \cdot V_{lr} B_{rk}, C_{kl} \cdot V_{l'r'} D_{r'} \rangle_t$$

$$= \int_0^t (A_s(B_sC_s)^T D_s)^{ij} ds.$$

Similarly we get relation (6).

Corollary 2.1. Let $A = (A_s)_{s \geq 0}$, $C = (C_s)_{s \geq 0}$, $D = (D_s)_{s \geq 0}$ be predictable locally bounded processes with values in the space of symmetric matrices $S_m(R)$ and let $V$ be a Brownian matrix with values in $\mathcal{M}_m(R)$. Suppose that the matrices $A_s, C_s, D_s$ commute. Let the martingale $M = (M_s)_{s \geq 0}$ be defined by

$$dM_s = A_s dV_s + dV_s^T A_s.$$

Then

$$\langle M, C \cdot MD \rangle_t = \int_0^t (2A_s^2C_s + tr (C_s) A_s^2 + tr (A_s^2C_s) I_m) D_s ds,$$

where $I_m$ is identical $m \times m$ matrix.

Proof. This is a direct application of relations (5) and (6).

Proposition 2.3. Let $A = (A_s)_{s \geq 0}$, $B = (B_s)_{s \geq 0}$, $C = (C_s)_{s \geq 0}$, $D = (D_s)_{s \geq 0}$ be predictable locally bounded processes with the values in $\mathcal{M}_m(R)$ and let $V$ be a Brownian matrix with values in $\mathcal{M}_m(R)$. Then

$$\langle A \cdot VB, tr (C \cdot VD) I_m \rangle_t = \int_0^t A_s(D_sC_s)^T B_s ds,$$  (7)

$$\langle A \cdot VB, tr (C \cdot V^T D) I_m \rangle_t = \int_0^t A_s D_s C_s B_s ds,$$  (8)

where $I_m$ is identical $m \times m$ matrix. The same formulas are true with replacing $V$ by $V^T$ at each position of $V$ in considered formulas.

Proof. The proof of (7) is going in coordinate-by-coordinate way exchanging the integral and trace. Similarly we get relation (8).

Corollary 2.2. Let $A = (A_s)_{s \geq 0}$, $B = (B_s)_{s \geq 0}$, $C = (C_s)_{s \geq 0}$, $D = (D_s)_{s \geq 0}$ be predictable locally bounded processes with the values in $\mathcal{M}_m(R)$ and let $V$ be a Brownian matrix with values in $\mathcal{M}_m(R)$. Then

$$\langle tr (A \cdot VB) I_m, C \cdot VD \rangle_t = \int_0^t C_s(B_sA_s)^T D_s ds,$$  (9)

$$\langle tr (A \cdot VB) I_m, C \cdot V^T D \rangle_t = \int_0^t C_s B_s A_s D_s ds.$$  (10)
The same relations are true with replacing $V$ by $V^\top$ at each position of $V$ in considered formulas.

Proof. The matrices $A \cdot VB$ and $\text{tr} \left( C \cdot VD \right) I_m$ commute, so it is sufficient just replace $A$ by $C$ (and vice versa) and $B$ by $D$ (and vice versa) in (7) and (8).

**Corollary 2.3.** Let

$$A = (A_s)_{s \geq 0}, \quad B = (B_s)_{s \geq 0}, \quad C = (C_s)_{s \geq 0}, \quad D = (D_s)_{s \geq 0}, \quad F = (F_s)_{s \geq 0}$$

be predictable locally bounded processes with values in $S_m(\mathbb{R})$ and let $V$ be a Brownian matrix with values in $M_m(\mathbb{R})$. Suppose that the matrices $A_s, B_s, C_s, D_s, F_s$ commute. Let the martingale $M = (M_s)_{s \geq 0}$ be defined by

$$dM_s = F_s dV_s + dV_s^\top F_s.$$ 

Then

$$\langle \text{tr} (A \cdot MB), \text{tr} (C \cdot MD) \rangle_t = 4 \int_0^t A_s B_s C_s D_s F_s^2 ds.$$ 

Proof. The corollary follows directly from (9) and (10) of Corollary 2.2.

**Proposition 2.4.** Let $A = (A_s)_{s \geq 0}, B = (B_s)_{s \geq 0}, C = (C_s)_{s \geq 0}, D = (D_s)_{s \geq 0}$ be predictable locally bounded processes with the values in $M_m(\mathbb{R})$ and let $V$ be a Brownian matrix with values in $M_m(\mathbb{R})$. Then

$$\langle \text{tr} (A \cdot VB), \text{tr} (C \cdot VD) \rangle_t = \int_0^t \text{tr} \left( A_s D_s C_s B_s \right) ds,$$

(11)

$$\langle \text{tr} (A \cdot VB), \text{tr} (C \cdot V^\top D) \rangle_t = \int_0^t \text{tr} \left( A_s D_s C_s B_s \right) ds.$$ 

(12)

Proof. Relations (11) and (12) follows from (7) and (8).

**Corollary 2.4.** Let $A = (A_s)_{s \geq 0}, B = (B_s)_{s \geq 0}, C = (C_s)_{s \geq 0}, D = (D_s)_{s \geq 0}, F = (F_s)_{s \geq 0}$ be predictable locally bounded processes with values in $S_m(\mathbb{R})$ and let $V$ be a Brownian matrix with values in $M_m(\mathbb{R})$. Suppose that the matrices $A_s, B_s, C_s, D_s, F_s$ commute. Let the martingale $M = (M_s)_{s \geq 0}$ be defined by

$$dM_s = F_s dV_s + dV_s^\top F_s.$$ 

Then

$$\langle \text{tr} (A \cdot MB), \text{tr} (C \cdot MD) \rangle_t = 4 \int_0^t \text{tr} \left( A_s B_s C_s D_s F_s^2 \right) ds.$$ 

Proof. This follows from Corollary 2.3 similarly as in the proof of Proposition 2.4.
3. Moments of Wishart process. In this section we will compute the moments $\mathbb{E}X^k_t$ for the Wishart process $X = (X_t)_{t \geq 0}$ defined by (2) as well as the expectation of the polynomials

$$R_t(X_t) = (\text{tr}(X_t))^j_3 (\text{tr}(X_t^2))^i_2 \cdots (\text{tr}(X_t^{k+1}))^{i_{k+1}} X^j_t,$$

where $I$ is a multiindex $I = (j, i) = (j, i_1, i_2, \ldots, i_{k+1})$.

Since $X = (X_t)_{t \geq 0}$ is a continuous semimartingale, its $k$th power $X^k = (X^k_t)_{t \geq 0}$ is also a continuous semimartingale for any $k \in \mathbb{N}^*$. According to semimartingale decomposition,

$$X^k_t = M_t^{(k)} + A_t^{(k)},$$

where $M_t^{(k)} = (M_t^{(k)})_{t \geq 0}$ is a continuous local martingale and $A_t^{(k)} = (A_t^{(k)})_{t \geq 0}$ is a predictable process of finite variation. When $k = 1$ we write simply $X_t = M_t + A_t$ with

$$dM_t = \sqrt{X_t} dV_t + dV_t^T \sqrt{X_t},$$

$$dA_t = \alpha I_m \, dt.$$

**Proposition 3.1.** For any $k \in \mathbb{N}^*$ and $t \geq 0$

$$dM_t^{(k+1)} = dM_t X^k_t + X_t dM_t^{(k)},$$

$$dA_t^{(k+1)} = dA_t X^k_t + X_t dA_t^{(k)} + d\langle M, M^{(k)} \rangle_t$$

and consequently

$$dM_t^{(k+1)} = \sum_{r=0}^{k} X^r_t dM_t X^{k-r}_t,$$

$$dA_t^{(k+1)} = \sum_{r=0}^{k} X^r_t dA_t X^{k-r}_t + \sum_{r=1}^{k} X^{k-r} d\langle M, M^{(r)} \rangle_t.$$  

**Proof.** Relations (13) and (14) are a direct consequence of the integration by part formula from Proposition 2.1

$$dX^{k+1} = d(X \cdot X^k) = dX_t X^k_t + X_t dX^k_t + d\langle M, M^{(k)} \rangle_t$$

$$= dM_t X^k_t + X_t dM_t^{(k)} + dA_t X^k_t + X_t dA_t^{(k)} + d\langle M, M^{(k)} \rangle_t,$$

where we identify the martingale and the predictable parts. Relations (15) and (16) follow then by induction. Proposition 3.1 is proved.

From semimartingale decomposition of $(X_t^{(k+1)})_{t \geq 0}$ we have

$$\mathbb{E}X_t^{(k+1)} = \mathbb{E}A_t^{(k+1)} + X_0^{k+1},$$
where \((\tau_n)\) is a localizing sequence to make \(M^{(k+1)}\) martingale. Using the fact that the trace of Wishart process is BESQ process with parameter \(\alpha m\) and the fact that the supremum of BESQ process on finite interval is integrable in any positive power, we can show the uniform integrability of \((\|X_{t,\tau_n}^{(k+1)}\|)_{n \geq 1}\) and \((\|A_{t,\tau_n}^{(k+1)}\|)_{n \geq 1}\). So, passing to the limit, we obtain that

\[
EX^{(k+1)}_t = EA^{(k+1)}_t + X^{k+1}_0.
\]

It means that for our task it is sufficient to explicite the process \(A^{(k+1)}\).

To do this, we need to compute the predictable brackets \((M, M^{(r)})_t\). We compute, at the same occasion, some other useful brackets, related to \(M^{(r)}\) and involving traces.

**Proposition 3.2.** Let \(r, l \in \mathbb{N}^*\) and \(t \geq 0\). We have

\[
(M, M^{(r)})(t) = (2r + m) \int_0^t X_r^r \, ds + 2 \sum_{q=1}^{r-1} \int_0^t \text{tr} (X^q_s) X^{r-q}_s \, ds
\]

\[
+ \int_0^t \text{tr} (X^r_s) I_m \, ds,
\]

\[
(M, M^{(r)})(t) = (2r + m) \int_0^t X_r^r \, ds + 2 \sum_{q=1}^{r-1} \int_0^t \text{tr} (X^q_s) X^{r-q}_s \, ds
\]

\[
+ \int_0^t \text{tr} (X^r_s) I_m \, ds.
\]

**Proof.** Using (15) with \(k = r - 1\) and Corollary 2.1 with \(A_s = \sqrt{X_s}\) we get

\[
(M, M^{(r)})(t) = \left( M, \sum_{q=0}^{r-1} X^q \cdot MX^{r-q-1} \right)_t
\]

\[
= \sum_{q=0}^{r-1} \int_0^t (2X^q_s + \text{tr} (X^q_s) X_s + \text{tr} (X^{q+1}_s) I_m) X^{r-q-1}_s \, ds
\]

\[
= \sum_{q=0}^{r-1} \int_0^t (2X^r_s + \text{tr} (X^q_s) X^{r-q}_s + \text{tr} (X^{q+1}_s) X^{r-q-1}_s) \, ds
\]

\[
= 2r \int_0^t X^r_s \, ds + 2 \sum_{q=1}^{r-1} \int_0^t \text{tr} (X^q_s) X^{r-q}_s + \int_0^t \text{tr} (X^0_s) X^r_s \, ds
\]

\[
+ \int_0^t \text{tr} (X^r_s) I_m \, ds
\]

\[
= (2r + m) \int_0^t X^r_s \, ds + 2 \sum_{q=1}^{r-1} \int_0^t \text{tr} (X^q_s) X^{r-q}_s \, ds
\]

\[
+ \int_0^t \text{tr} (X^r_s) I_m \, ds
\]
and (17) follows. Corollary 2.3 with \( F_s = \sqrt{X_s} \) and (15) imply that

\[
\langle \text{tr} (M^{(r)}) I_m, M^{(l)} \rangle = \sum_{i=0}^{r-1} \sum_{j=0}^{l-1} \langle \text{tr} (X^i \cdot M X^{r-i-1}) I_m, X^j \cdot M X^{l-j-1} \rangle_t
\]

\[
= 4 \sum_{i=0}^{r-1} \sum_{j=0}^{l-1} \int_0^t X_s^{r+i-l} \, ds = 4rl \int_0^t X_s^{r+i-l} \, ds
\]

that gives (18). Relation (19) follows in the same way from Corollary 2.4.

**Proposition 3.3.** Let \( k \in \mathbb{N}^* \) and \( t \geq 0 \). We have

\[
A_t^{(k)} = (\alpha k + (k-1)(k+m)) \int_0^t X_s^{k-1} \, ds + \sum_{r=1}^{k-1} (2r - 1) \int_0^t \text{tr} (X_s^{k-r}) X_s^{r-1} \, ds.
\]

**Proof.** Putting together (16) and (17) and using \( dA_t = \alpha I_m \, dt \) one obtains

\[
dA_t^{(k)} = \sum_{r=0}^{k-1} X_t^{r} \, dA_t X_t^{k-r-1} + \sum_{r=1}^{k-1} X_t^{k-r-1} \langle \text{tr} (M^{(r)}) \rangle_t = \alpha k X_t^{k-1} \, dt
\]

\[
+ \sum_{r=1}^{k-1} X_t^{k-r-1} \left[ (2r + m) X_t^{r} + 2 \sum_{q=1}^{r-1} \text{tr} (X_t^{q}) X_t^{r-q} + \text{tr} (X_t^{r}) I_m \right] \, dt
\]

\[
= \left( \alpha k + (k-1) m + (k-1) k \right) X_t^{k-1} \, dt
\]

\[
+ \sum_{r=1}^{k-1} (2k - 2r - 1) \text{tr} (X_t^{r}) X_t^{k-r-1} \, dt
\]

\[
= \left( \alpha k + (k-1)(k+m) \right) X_t^{k-1} \, dt + \sum_{r=1}^{k-1} (2r - 1) \text{tr} (X_t^{k-r}) X_t^{r-1} \, dt.
\]

Proposition 3.3 is proved.

Proposition 3.3 implies a recurrence relation between the expectations \( \mathbf{E} X_t^k = \mathbf{E} A_t^{(k)} + X_t^k \) and \( \mathbf{E} \text{tr} (X_t^{k-r}) X_t^{r-1} \) for all \( r = 1, \ldots, k \). Only one of these terms, when \( k = r \), has the same form as \( \mathbf{E} X_t^k \), namely \( \mathbf{E} X_t^{k-1} \). An application of the Itô formula to \( \text{tr} (X_t^{k-r}) X_t^{r-1} \) will in turn lead to more complicated expressions, for example of the form \( \text{tr} (X_t^{k-r-j}) \text{tr} (X_t^{j}) X_t^{r-2} \), in its predictable part.

In order to give an effective recurrence formula for \( \mathbf{E} X_t^{k+1} \) we are thus obliged to work with a vector of matrix-valued stochastic processes, which is stable by the application of the Itô formula.

For a multiindex \( I = (j, i) = (j, i_1, i_2, \ldots, i_{k+1}) \in \mathbb{N}^{k+2} \) such that

\[
j + \sum_{r=1}^{k+1} ri_r = k + 1
\]
and a matrix $B \in M_m(\mathbb{R})$ we introduce the matrix-valued polynomials

$$R_j(B) = (\text{tr}(B))^{i_1} (\text{tr}(B^2))^{i_2} \cdots (\text{tr}(B^{k+1}))^{i_{k+1}} B^j.$$  

We can arrange the multiindices $I$ decreasing in the lexicographical order and define the vector $Q^{(k+1)}(B)$:

$$Q^{(k+1)}(B) = \begin{pmatrix} B^{k+1} \\ \text{tr}(B) B^k \\ (\text{tr}(B))^2 B^{k-1} \\ \text{tr}(B^2) B^{k-1} \\ \vdots \\ \text{tr}(B^{k+1}) I_m \end{pmatrix}.$$  

The dimension of the vector $Q^{(k)}(B)$ is equal to

$$d_k = \sum_{n=0}^k p(n),$$

where $p(n)$ is the number of partitions of the integer $n$ into a sum of positive integers (we put $p(0) = 1$). We have $p(1) = 1$, $p(2) = 2$, $p(3) = 3$, $p(4) = 5$, so $d_0 = 1$, $d_1 = 2$, $d_2 = 4$, $d_3 = 7$, $d_4 = 12$ (cf. [16]).

First of all we remark that for each fixed $t \geq 0$ all components of $Q^{(k+1)}(X_t)$ are integrable. In fact, since $(\text{tr}(X_t))_{t \geq 0}$ is a Bessel process with parameter $a m$, $\text{tr}(X_t)$ has the moments of all positive order and for all $k \in \mathbb{N}$ and $0 \leq r \leq k + 1$ we obtain

$$E\|X_t^k\| = E\left(\sqrt{(X_t^k, X_t^k)}\right) = E\sqrt{\text{tr}(X_t^{2k})} \leq E(\text{tr}(X_t))^k < \infty,$$

$$E\|X_t^{r-1} \text{tr}(X_t^{k+1-r})\| = E\left(\sqrt{\text{tr}(X_t^{2r-2}) \text{tr}(X_t^{k+1-r})}\right) \leq E(\text{tr}(X_t))^k < \infty.$$  

**Theorem 3.1.** For all $k \in \mathbb{N}$ there exists a $d_{k+1} \times d_k$ matrix $D_k$ such that for all $t > 0$

$$E Q^{(k+1)}(X_t) = Q^{(k+1)}(X_0) + D_k \int_0^t E Q^{(k)}(X_s) \, ds. \quad (21)$$

Consequently,

$$E Q^{(k+1)}(X_t) = Q^{(k+1)}(X_0) + \sum_{r=1}^{k+1} \frac{t^r}{r!} D_k D_{k-1} \cdots D_{r+1} Q^{(k-r+1)}(X_0) \quad (22)$$

with the convention $Q^{(0)}(X_0) = I_m$ and with the matrices $D_l$, $l = 0, 1, \ldots, k$, explicitly given below in Table 1.

In particular, when $X_0 = 0$, we have

$$E Q^{(k+1)}(X_t) = \frac{t^{k+1}}{(k+1)!} D_k D_{k-1} \cdots D_0 I_m. \quad (23)$$
Table 1. Coefficients $\tilde{D}_{I'I'}$ for the matrix $D_k$

<table>
<thead>
<tr>
<th>Line</th>
<th>$I$</th>
<th>$I'$</th>
<th>$\tilde{D}_{I'I'}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$j \geq 1$</td>
<td>$(j - 1, i)$</td>
<td>$\alpha_j + (j - 1)(j + m)$</td>
</tr>
<tr>
<td>2</td>
<td>$j \geq 2$</td>
<td>$(l, 1 + e_j - l, -1)$</td>
<td>$2l + 1$ for some $l$, $l = 0, \ldots, j - 2$</td>
</tr>
<tr>
<td>3</td>
<td>$j \geq 1, i_r \geq 1$</td>
<td>$(j + r - 1, i - e_r)$</td>
<td>$4jr_i_r$ for some $r$, $2 \leq r \leq k + 1$</td>
</tr>
<tr>
<td>4</td>
<td>$i_1 \geq 1$</td>
<td>$(j, i - e_1)$</td>
<td>$(4j + \alpha m) i_1$</td>
</tr>
<tr>
<td>5</td>
<td>$i_r \geq 1$</td>
<td>$(j, i + e_{r-1} - e_r)$</td>
<td>$(ar + (r - 1)(r + m) + m) i_r$ for some $r$, $1 &lt; r \leq k + 1$</td>
</tr>
<tr>
<td>6</td>
<td>$i_r \geq 1$</td>
<td>$(j, i - e_r + e_{l-1} + e_r - l)$</td>
<td>$(2l - 1) i_r$ for some $l$, $2 \leq l &lt; r$</td>
</tr>
<tr>
<td>7</td>
<td>$i_r \geq 2$</td>
<td>$(j, i - 2e_r + e_{2r-1})$</td>
<td>$2r^2(i_r - 1) i_r$ for some $r$, $2 &lt; r \leq k + 1$</td>
</tr>
<tr>
<td>8</td>
<td>$i_r, i_l \geq 1$</td>
<td>$(j, i - e_r - e_l + e_{r+l-1})$</td>
<td>$4ri_r i_l$ for some $r, l$, $1 \leq l &lt; r \leq k + 1$</td>
</tr>
</tbody>
</table>

Before proving Theorem 3.1 let us give some information about the coefficients of the matrix $D_k$.

**Coefficients of the matrix $D_k$.** Let $k \geq 0$ be fixed and let $D_k = (D_{II'})$, where $I = (j, i)$ and $I' = (j', i')$ are multiindices verifying $i = (i_1, i_2, \ldots, i_k)$, $j + \sum_{r=1}^{k+1} r_i_r = k + 1$ and $i' = (i'_1, i'_2, \ldots, i'_k)$, $j' + \sum_{r=1}^{k} r'i'_r = k$. As we will see in the proof of Theorem 3.1, the coefficients $D_{I'I'}$ are equal to zero except of the cases mentioned in Table 1. To give different from zero coefficients $D_{I'I'}$, we denote by $e_1, e_2, \ldots, e_{k+1}$ the vectors of canonical basis of $\mathbb{R}^{k+1}$ and we extend $I'$ without changing the notation to $I' = (j', i', 0)$, to make the dimensions of $I$ and $I'$ equal. Table 1 contains 8 lines and, hence, 8 types of coefficients corresponding to the particular conditions verified by multiindex $I$. For example, the condition $j \geq 2$ on the second line of Table 1 means that only $I = (j, i)$ with $j \geq 2$ are concerned. Similarly, the condition $i_r \geq 2$ in the 7th line of Table 1 means that this line applies only to multiindices $I = (j, i)$ with at least one coefficient $i_r \geq 2$ for some $r \geq 1$.

The third column of Table 1 gives the multiindices $I'$ corresponding to nonzero coefficients. The first tree lines of the table contain the coefficients...
mainly related to the changing of \( j \) in multiindex \( I = (j, i) \). The lines 4–8 give the coefficients coming from the changing of the second part \( i \) in multiindex \( I = (j, i) \).

To obtain the coefficient \( D_{II'} \) corresponding to multiindices \( I, I' \) one has to select all lines concerning \( I, I' \) and add the corresponding values of \( \tilde{D}_{II'} \).

Let us illustrate the use of Table 1 by some examples. Take \( k = 10 \) and \( I = (11, 0) \). Since \( i = 0 \), only the two first lines of Table 1 apply to this multiindex. For \( I' = (10, 0) \) using the first line we obtain \( D_{II'} = 11a + 10m + 110 \), where \( a, m \) are the parameters of Wishart process. For \( I' = (0, e_{10}), (1, e_9), \ldots, (9, e_1) \) we obtain using the second line the coefficients \( D_{II'} = 1, 3, \ldots, 19 \), respectively. Still for \( k = 10 \), if \( I = (4, 2e_1 + e_2 + e_3) \), then all the lines apply, some of them several times. Let \( I' = (4, e_1 + e_2 + e_3) \) then we see that the 4th line applies once, 7th line applies once (because \( i_1 \geq 2 \) and \( i_2 < 2, i_3 < 2 \)) and that the 8th line applies twice (for \( l = 1 \) and \( r = 2, 3 \)). So, the coefficient \( D_{II'} = 2(16 + am) + 4 + 16 + 24 = 2am + 76 \).

To obtain the coefficients of the matrix \( D_k \) one should write the multiindices \( I \) and \( I' \) in lexicographical order and then use Table 1 to find \( D_{I,I'} \).

Let \( k = 3 \). Then we have

\[
Q^{(3)}(X_t) = \begin{pmatrix}
X_t^3 \\
X_t^2 \text{tr} X_t \\
X_t (\text{tr} X_t)^2 \\
X_t \text{tr} (X_t^2) \\
(\text{tr} X_t)^3 I_m \\
\text{tr} X_t (\text{tr} X_t^2) I_m \\
\text{tr} (X_t^3) I_m \\
\end{pmatrix}
\]

that corresponds to multiindices \( I \) in the following order:

\((3; 0, 0, 0), (2; 1, 0, 0), (1; 2, 0, 0), (1; 0, 1, 0), (0; 3, 0, 0), (0; 1, 1, 0), (0; 0, 0, 1)\).

Doing the same for \( k = 2 \) we can find that the multiindices \( I' \) are

\((2; 0, 0), (1; 1, 0), (0; 2, 0), (0; 0, 1)\).

The reader will easily check that the matrix \( D_2 \) needed in the computation of the moments \( \mathbf{E} Q^{(3)}(X_t) \) of the third order is equal to

\[
D_2 = \begin{pmatrix}
3a + 2m + 6 & 3 & 0 & 1 \\
am + 8 & 2a + m + 2 & 1 & 0 \\
0 & 2am + 12 & a & 0 \\
8 & 2a + 2m + 2 & 0 & a \\
0 & 0 & 3am + 12 & 0 \\
0 & 0 & 2a + 2m + 2 & am + 8 \\
0 & 0 & 3 & 3a + 3m + 6 \\
\end{pmatrix}
\]
Proof of Theorem 3.1. Let us first present the idea of the proof of Theorem 3.1. We have

$$
\mathbf{E}Q^{(k+1)}(X_t) = \begin{pmatrix}
\mathbf{E}X_{t}^{k+1} \\
\mathbf{E} \text{tr} (X_t) X_{t}^{k} \\
\vdots \\
\mathbf{E} \text{tr} (X_{t}^{k+1}) I_m
\end{pmatrix} = \begin{pmatrix}
\mathbf{E}A(X_{t}^{k+1}) \\
\mathbf{E}A(\text{tr} (X_t) X_{t}^{k}) \\
\vdots \\
\mathbf{E}A(\text{tr} (X_{t}^{k+1})) I_m
\end{pmatrix} + \mathbf{Q}^{(k+1)}(X_0),
$$

where $A(Y)$ denotes the predictable part in semimartingale decomposition of $Y$. Now, using the Itô calculus, along the same lines as in Proposition 3.3, we will show that

$$
\begin{pmatrix}
A(X_{t}^{k+1}) \\
A(\text{tr} (X_t) X_{t}^{k}) \\
\vdots \\
A(\text{tr} (X_{t}^{k+1})) I_m
\end{pmatrix} = \int_0^t D_k \mathbf{Q}^{(k)}(X_s) \, ds
$$

and then relation (21) follows by taking the expected value. We deduce from it that

$$
\frac{d}{dt} \mathbf{E}Q^{(k+1)}(X_t) = D_k \mathbf{E}Q^{(k)}(X_t)
$$

and, by induction,

$$
\frac{d^r}{dt^r} \mathbf{E}Q^{(k+1)}(X_t) = D_k \cdots D_{k+1-r} \mathbf{E}Q^{(k+1-r)}(X_t).
$$

In particular, for $r = k + 1$, we have that the derivative of order $k + 1$

$$
\frac{d^{k+1}}{dt^{k+1}} \mathbf{E}Q^{(k+1)}(X_t) = D_k \cdots D_0 I_m
$$

is a constant independent of $t$. We see in this way that the matrix-valued function $U(t) = \mathbf{E}Q^{(k+1)}(X_t)$ is a polynomial of order $k + 1$ in $t$, so it may be written in the Taylor form

$$
U(t) = \sum_{r=0}^{k+1} \frac{U^{(r)}(0) t^r}{r!}
$$

and (22) follows.

We must therefore justify (24). Let us write

$$
r_i(X_t) = (\text{tr} X_t)^{i_1}(\text{tr} X_{t}^2)^{i_2} \cdots (\text{tr} X_{t}^{k+1})^{i_{k+1}}
$$

so that

$$
R_i(X_t) = R_{(i,i)}(X_t) = r_i(X_t) X^i_t.
$$

We have, by Proposition 2.1,

$$
dR_i(X_t) = dr_i(X_t) X^i_t + r_i(X_t) dX^i_t + d\langle r_i(X) I_m, X^i \rangle_t.
$$
Consequently, writing the semimartingale decomposition

\[ r_t(X_t) = M_t(r_t(X)) + A_t(r_t(X)), \]

we have

\[
dA_t(R_t(X_t)) = dA_t(r_t(X_t)) X_t^j + r_t(X_t) dA_t^{(j)} + d\langle M(r_t(X)) I_m, M^{(j)} \rangle_t
\]
\[
= (T_t^{(1)} + T_t^{(2)} + T_t^{(3)}) dt. \tag{25}
\]

We already know \( A_t^{(j)} \) and \( M_t^{(j)} \). We need only to compute explicitly \( M_t(r_t(X)) \) and \( A_t(r_t(X)) \).

Let \( F \) be the function \( \mathbb{R}^{k+1} \to \mathbb{R} \) defined by

\[ F(y_1, \ldots, y_{k+1}) = \prod_{r=1}^{k+1} y_r^r. \]

Denote

\[ Y^{(i)} := tr(X_t^i) = tr(M_t^{(i)}) + tr(A_t^{(i)}). \]

Then \( r_t(X_t) = F(Y_t^{(1)}, \ldots, Y_t^{(k+1)}) \). By the Itô formula,

\[
dr_t(X_t) = \sum_{r=1}^{k+1} \frac{\partial F}{\partial y_r}(Y_t^{(1)}, \ldots, Y_t^{(k+1)}) dY_t^{(r)}
\]
\[
+ \frac{1}{2} \sum_{i,r=1}^{k+1} \frac{\partial^2 F}{\partial y_i \partial y_r}(Y_t^{(1)}, \ldots, Y_t^{(k+1)}) d\langle Y^{(i)}, Y^{(r)} \rangle_t
\]
\[
= \sum_{r=1}^{k+1} i_r r_{i-r} (X_t) d tr(M_t^{(r)})
+ \sum_{r=1}^{k+1} i_r r_{i-r} A_t^{(r)} d tr(A_t^{(r)})
\]
\[
+ \frac{1}{2} \sum_{r=1}^{k+1} r_{i-r} (X_t) d\langle tr M^{(i)}, tr M^{(r)} \rangle_t
\]
\[
+ \frac{1}{2} \sum_{r=1}^{k+1} r_{i-r} (X_t) d\langle tr M^{(r)}, tr M^{(r)} \rangle_t.
\]

We see that

\[
dM_t(r_t(X)) = \sum_{r=1}^{k+1} i_r r_{i-r} (X_t) d tr(M_t^{(r)}) \tag{26}
\]

and

\[
dA_t(r_t(X)) = \sum_{r=1}^{k+1} i_r r_{i-r} (X_t) d tr(A_t^{(r)})
\]
\[
+ \sum_{r=2}^{k+1} \sum_{l=1}^{r-1} i_l r_{i-r} d\langle tr M^{(l)}, tr M^{(r)} \rangle_t.
\]
The moments of Wishart processes via Itô calculus

\[
\begin{align*}
&\quad+ \frac{1}{2} \sum_{r=1}^{k+1} i_r (i_r - 1) r r_{1-2e_r} (X_t) \, d\langle \text{tr } M^{(r)}, \text{tr } M^{(r)} \rangle_t \\
&= (S_t^{(1)} + S_t^{(2)} + S_t^{(3)}) \, dt.
\end{align*}
\]

(27)

Using the previously obtained formulas for \(dM_t^{(r)}, dA_t^{(r)}, \) and \(d\langle \text{tr } M^{(l)}, \text{tr } M^{(r)} \rangle_t \), we can now write explicitly the three terms of (25).

**Term \( T_t^{(1)} \).** We use Propositions 3.2 and 3.3. According to (27), \( T_t^{(1)} \) has three terms, \( T_t^{(1)} = S_t^{(1)} X_t^j + S_t^{(2)} X_t^j + S_t^{(3)} X_t^j \). We have

\[
S_t^{(1)} X_t^j = \left\{ i_1 \alpha \mu r_{1-e_1} (X_t) \\
+ \sum_{r=2}^{k+1} i_r r_{1-e_r} (X_t) \left[ (\alpha r + (r - 1)(r + m) + m) Y_t^{(r-1)} \\
\quad+ \sum_{r=2}^{k+1} \sum_{l=2}^{r-1} (2l - 1) Y_t^{(r-l)} Y_t^{(l-1)} \right] \right\} X_t^j
\]

\[
\quad+ \sum_{r=2}^{k+1} i_r (\alpha r + (r - 1)(r + m) + m) R_{(j,i-e_r+e_r-1)} (X_t)
\]

\[
\quad+ \sum_{r=3}^{k+1} \sum_{l=2}^{r-1} i_r (2l - 1) R_{(j,i-e_r+e_r-1+e_r-1)} (X_t).
\]

This gives the \( \alpha \mu i_1 \) term of the 4th line, and the 5th, and the 6th lines of the table of the coefficients of \( D_{II} \). Furthermore,

\[
S_t^{(2)} X_t^j = \sum_{1 \leq i \leq r} i i_i r_{1-e_i-e_r} (X_t) 4 r l Y_t^{(r+l-1)} X_t^j
\]

\[
= 4 \sum_{1 \leq i \leq r} l i_i i_i R_{(j,i-e_i-e_r+e_r+1)} (X_t).
\]

The condition \( i_l, i_r \geq 1 \) ensures that \( r + l \leq k + 1 \) so that \( r + l - 1 \leq k \) and the multiindex in the last expression is well defined. We obtain the 8th line of the table.

Finally,

\[
S_t^{(3)} X_t^j = \frac{1}{2} \sum_{r=1}^{k+1} i_r (i_r - 1) r r_{1-2e_r} (X_t) 4 r^2 Y_t^{(2r-1)} X_t^j
\]

\[
= \sum_{r=1}^{k+1} 2 r^2 i_r (i_r - 1) R_{(j,i-2e_r+e_2-1)} (X_t).
\]

This implies the 7th line of the table.
Term $T_t^{(2)}$. Using Proposition 3.3, we get

$$T_t^{(2)} = (\alpha j + (j - 1)(j + m)) X_t^{j-1} r_1(X_t) + \sum_{l=0}^{j-2} (2l + 1) Y_t^{(j-l-1)} X_t^l r_1(X_t)$$

$$= (\alpha j + (j - 1)(j + m)) R_{(j-1,1)}(X_t) + \sum_{l=0}^{j-2} (2l + 1) R_{(l,1+e_{j-1-1})}(X_t).$$

We obtain the first and second lines of the table.

Term $T_t^{(3)}$. We use relation (26) and Proposition 3.2. We have

$$T_t^{(3)} dt = d\langle M(r_1(X_t)) I_m, M^{(j)} \rangle_t = d\left\langle \sum_{r=1}^{k+1} i_r^r r_{i-r}(X) \tr M^{(r)}, M^{(j)} \right\rangle_t$$

$$= \sum_{r=1}^{k+1} i_r^r r_{i-r}(X_t) 4rj X_t^{r+j-1} dt = \sum_{r=1}^{k+1} 4irj r_{(j+r-1,i-e_r)}(X_t) dt.$$ 

We get the third line of the table (for $r \geq 2$) and the $4j_1$ term of the fourth line for $r = 1$. Theorem 3.1 is proved.

The following corollary concerns the case $m = 1$. We obtain by Itô calculus the moments of the non centered gamma distributions, which were calculated before using the Laplace transform (cf. [8]).

**Corollary 3.1.** Let $(X_s)_{s \geq 0}$ be a Bessel Squared process with parameter $\alpha \in \mathbb R^+$ starting at $x_0 > 0$, i.e., a solution of the stochastic differential equation

$$\begin{cases}
  dX_s = 2\sqrt{X_s} dB_s + \alpha ds, \\
  X_0 = x_0,
\end{cases}$$

where $\beta = (B_s)_{s \geq 0}$ is a one-dimensional Brownian motion. Then for $k \in \mathbb N$

$$EX_t^k = k! t^k L_k^{(\alpha-1)}\left(-\frac{x_0}{t}\right),$$

where $\{L_k^{(\alpha-1)}(x)\}_{k \geq 0}$ is the orthogonal family of Laguerre polynomials with parameter $\alpha - 1$ (cf. [17]). In particular, when $x_0 = 0$,

$$EX_t^k = (\alpha)_k t^k,$$

where $(x)_k = x(x+1) \cdots (x+k-1)$ is the Pochhammer symbol.

**Proof.** By Proposition 3.3, for $m = 1$,

$$A_t^{(k+1)} = c_k \int_0^t X_s^k ds, \quad \text{where } c_k = \alpha + (\alpha + 1) k + k^2. \quad (29)$$

Arguing as in the beginning of the proof of Theorem 3.1 we see that

$$EX_t^{k+1} = x_0^{k+1} + \sum_{r=1}^{k+1} \frac{t^r}{r!} c_k c_{k-1} \cdots c_{k-r+1} x_0^{k-r+1}$$

$$= x_0^{k+1} \left(1 + \sum_{r=1}^{k+1} \frac{1}{r!} \left(\frac{t}{x_0}\right)^r c_k c_{k-1} \cdots c_{k-r+1}\right).$$
Since \( c_l = \alpha + (\alpha + 1) l + l^2 = (l + 1)(l + \alpha) \), we have
\[
c_k c_{k-1} \cdots c_{k-r+1} = \frac{(k+1)!}{(k+1-r)!} (k+\alpha-r+1)_r.
\]
Consequently, changing the summation index \( l = k+1-r \) and using the usual notation \( \binom{s}{n} = (s-n+1)_n/n! \) for \( s \in \mathbb{R}, n \in \mathbb{N} \), we get according to (5.1.6) in [17] that
\[
EX^{k+1}_t = (k+1)! x_0^{k+1} \sum_{r=0}^{k+1} \frac{1}{(k+1-r)!} \left( \frac{t}{x_0} \right)^r (k+\alpha-r+1)_r.
\]
Using
\[
x_0^{k+1} \left( \frac{t}{x_0} \right)^{k+1-l} \binom{k+\alpha}{k+1-l} = t^{k+1} \left( \frac{x_0}{t} \right)^l \binom{k+1+(\alpha-1)}{k+1-l}
\]
and
\[
I_{k+1}^{(\alpha-1)} \left( - \frac{x_0}{t} \right) = \sum_{l=0}^{k+1} \frac{1}{l!} \left( \frac{x_0}{t} \right)^l \binom{k+1+(\alpha-1)}{k+1-l},
\]
we obtain the result.

4. Some examples. To illustrate the application of our results let us calculate the moments of the Wishart law using Theorem 3.1 and the description of the coefficients \( D_{ll'} \) in the table. For \( k = 1 \),
\[
EQ^{(1)}(X_t) = \begin{pmatrix} \mathbf{EX}_t \cr \mathbf{E} \mathbf{tr} (X_t) \mathbf{I}_m \end{pmatrix} = \begin{pmatrix} X_0 \cr \mathbf{tr} (X_0) \mathbf{I}_m \end{pmatrix} + t \begin{pmatrix} \alpha \cr \alpha m \end{pmatrix} \mathbf{I}_m.
\]
This follows directly from (2) and also from Theorem 3.1 for \( k = 1 \). We have
\[
D_0 = \begin{pmatrix} \alpha \\ \alpha m \end{pmatrix}.
\]
Note that we understand \( D_0 \mathbf{I}_m \) as
\[
D_0 \mathbf{I}_m = \begin{pmatrix} \alpha \mathbf{I}_m \\ \alpha m \mathbf{I}_m \end{pmatrix}.
\]
When \( k = 2 \), we obtain from Theorem 3.1
\[
D_1 = \begin{pmatrix} 2\alpha + m + 2 & 1 \\ \alpha m + 4 & \alpha \\ 0 & 2\alpha m + 4 \\ 0 & 2\alpha + 2m + 2 \end{pmatrix}.
\]
We have, for any $X_0 \in S_m^+$,

$$\text{EQ}^{(2)}(X_t) = \begin{pmatrix} \mathbb{E}X_t^2 \\
\mathbb{E}\text{tr}(X_t)X_t \\
\mathbb{E}(\text{tr}X_t)^2I_m \\
\mathbb{E}\text{tr}(X_t^2)I_m
\end{pmatrix} = \begin{pmatrix} X_0^2 \\
\text{tr}(X_0)X_0 \\
(\text{tr}X_0)^2I_m \\
\text{tr}(X_0^2)I_m
\end{pmatrix} + tD_1(t) \begin{pmatrix} X_0 \\
\text{tr}(X_0)I_m \\
(\text{tr}X_0)I_m \\
\text{tr}(X_0^2)I_m
\end{pmatrix} + \frac{t^2}{2} D_1D_0I_m. $$

In particular, in the centered case $X_0 = 0$,

$$\text{EQ}^{(2)}(X_t) = \begin{pmatrix} \mathbb{E}X_t^2 \\
\mathbb{E}\text{tr}(X_t)X_t \\
\mathbb{E}(\text{tr}X_t)^2I_m \\
\mathbb{E}\text{tr}(X_t^2)I_m
\end{pmatrix} = \frac{t^2}{2} D_1D_0I_m$$

$$= t^2 \begin{pmatrix} (\alpha^2 + \alpha m + \alpha)I_m \\
(\alpha^2m + 2\alpha)I_m \\
(\alpha^2m^2 + 2\alpha m)I_m \\
(\alpha^2 + \alpha m + \alpha) mI_m
\end{pmatrix}. $$

For the values of $k \geq 3$ the computation of the coefficients $D_{l,l'}$ of matrices $D_k$ may easily be programmed in the simplest computational computer software, e.g., XMaple. For instance, for $k = 4$ and in the centered case,

$$\mathbb{E}X_t^4 = (6am^2 + 6\alpha^2m + 6\alpha^2m^2 + 21\alpha^2 + 20\alpha + 17\alpha^2m$$
$$+ 6\alpha^3 + \alpha m + \alpha^4 + 21\alpha m)t^4I_m.$$ 

This formula was obtained by different methods in [11]. In the particular noncentered case $X_0 = I_m$ and $t = 1$

$$\mathbb{E}X_t^4 = (161 + 138\alpha + 119\alpha m + \alpha^4 + 152m + 4m^3 + m^3\alpha + 41\alpha^2m$$
$$+ 30\alpha m^2 + 6\alpha^2m + 6\alpha^2m^2 + 51\alpha^2 + 44m^2 + 10\alpha^3)I_m.$$ 

5. Acknowledgments. We are very grateful to A. Shiryaev and M. Yor for helpful remarks and comments.

REFERENCES

9. Жакод Ж., Ширяев А. Н. Предельные теоремы для случайных процессов. Т. 1, 2. М.: Физматлит, 1994, 544 с., 368 с.