QUADRATIC FUNCTIONALS OF BROWNIAN MOTION, OPTIMAL CONTROL, AND THE "COLDITZ" EXAMPLE

by

L.C.G. Rogers & Z. Shi*

Queen Mary & Westfield College, University of London

Abstract: In this paper, we present a general method for computing (the Laplace transform of) the distribution of a very general quadratic functional of Brownian motion. The method is based on solving an equivalent problem of linear deterministic control (the "linear regulator"), and reduces to the well-known Riccati equation. We illustrate the method by solving completely a complicated example which appears to be intractable by other means.

Keywords and phrases: quadratic functional, Brownian motion, Colditz example, linear regulator, polymer moment-of-inertia.

1. Introduction

Recently, there has been renewed interest in the calculation of the law of quadratic functionals of Brownian motion (and Gaussian processes), as is exemplified by the problem of computing

$$E\exp\left[-\frac{\theta^2}{2}\int_0^1 (B_s-\overline{B})^2 ds\right], \tag{1}$$

where $\overline{B}\equiv\int_0^1 B_s ds$ is the centre-of-mass of the Brownian path $(B_t)_{0\le t\le 1}$. This is a problem of interest in polymer physics, where it arises from the conformation of a polymer in a pure straining flow. Although the polymer is a three-dimensional object, its moment of inertia splits into the sum of three independent contributions, thus reducing the problem to a one-dimensional setting. The one-dimensional case has received much attention (see, for example, Donati-Martin & Yor [4], Chan, Dean, Jansons & Rogers [3]), but our interest here was to extend if possible to genuinely multidimensional examples. The paper of Chan, Dean, Jansons & Rogers [3] makes clear the importance of excursion ideas and the Ray-Knight theorem in solving one-dimensional examples, and we wanted to find out whether these methods might work in higher dimensions, or, if not, what would be viable substitutes. To guide us in this,

* Supported by SERC grant number GR/H 00444

we took a test example and analysed it completely with bare-hands techniques; the end results bore no interpretation in terms of excursion theory or any other probabilistic object as far as we could see. The test problem we set ourselves was to compute

$$E \exp\left[-\frac{\theta^2}{2}\int_0^1 X_u^T K_u X_u du\right], \tag{2}$$

where X is a Brownian motion in R^2, and K_t is the matrix

$$K_t \equiv \begin{pmatrix} \cos\omega t \\ \sin\omega t \end{pmatrix} \begin{pmatrix} \cos\omega t \\ \sin\omega t \end{pmatrix}^T, \tag{3}$$

with $\omega > 0$ a fixed parameter. We shall present a very general methodology for computing things like (2), which certainly includes all examples where X is $\mathrm{BM}(R^n)$ and K is $n \times n$, symmetric, continuous. The methodology is easy to implement numerically (closed-form solutions appear to be the exception); in fact, the problem we solve is *equivalent* to one which has been much studied in optimal control for years!
To explain how the method works, define

$$p(t,x) \equiv E\left[\exp\left(-\frac{1}{2}\int_t^1 X_u^T K_u X_u du\right) \mid X_t = x\right], \tag{4}$$

where X is $\mathrm{BM}(R^n)$, K is $n \times n$ symmetric, non-negative definite (though this last is not essential). Then

$$\exp\left(-\frac{1}{2}\int_0^t X_u^T K_u X_u du\right) p(t, X_t) \ \textit{is a martingale},$$

from which Itô's formula gives

$$\frac{1}{2}\Delta p + \dot p - \frac{1}{2}x^T K_t x p = 0, \quad p(1,x) \equiv 1. \tag{5}$$

It is easy to guess the form of the solution to (5); if we try

$$p(t,x) = \exp\left[-\frac{1}{2}x^T Q_t x - \gamma_t\right], \tag{6}$$

then we find that if Q and γ solve

$$Q_t^2 - \dot Q_t - K_t = 0, \tag{7.i}$$

$$\dot\gamma_t = -\frac{1}{2}\operatorname{tr} Q_t, \tag{7.ii}$$

with initial conditions $Q_1 = 0, \gamma_1 = 0$, then (6) is indeed the solution. Thus computing p is equivalent to finding Q to solve the matrix Riccati equation (7.i). In §2, we prove the relationship between the law of a quadratic functional of a Gaussian process and the solution of an optimisation problem. This result is then specialised to the Brownian motion setting. A major advance here is that one does not need the Ray-Knight theorem or excursion theory to compute the laws of quadratic functionals.

The appearance of the matrix Riccati equation in the computation of the laws of quadratic functionals of Brownian motion turns out to be quite ancient; the one-dimensional case first appears, to our knowledge, in a paper of Cameron & Martin [2] in 1945, and the method has appeared sporadically in the literature since (see, for example, Kac [5], Liptser & Shiryaev [8] p.280). The point of view adopted in these references is to take a change-of-measure martingale

$$dZ_t = -Z_t X_t^T Q_t dX_t$$

which is solved by

$$Z_t = \exp\left[-\int_0^t X_u^T Q_u dX_u - \frac{1}{2}\int_0^t \mid Q_u X_u \mid^2 du \right];$$

this, by Itô's formula, is equal to

$$= \exp\left[-\frac{1}{2} X_t^T Q_t X_t + \frac{1}{2} X_0^T Q_0 X_0 - \frac{1}{2}\int_0^t X_u^T (Q_u^2 - \dot{Q}_u) X_u du + \frac{1}{2}\int_0^t \text{tr } Q_u du \right].$$

Now, it can be seen that if Q solves (7.i), then the law of the quadratic functional is indeed given by

$$E \exp\left(-\frac{1}{2}\int_0^1 X_u^T K_u X_u du \right) = \exp\left(-\frac{1}{2}\int_0^1 \text{tr } Q_u du \right).$$

This approach is, of course, mathematically equivalent to the one we outlined above, but neither is satisfying in the sense that they *solve* the problem without *explaining* the connection with the classical mathix Riccati equation of optimal control. We provide such an explanation in §2.

In §3, we consider the optimal control problem related to our test example. Prisoners-of-war trying to escape from Colditz castle have to cross an open field before reaching the safety of woodland. In the middle of this field there is a searchlight which from time to time is turned on, and then sweeps slowly round the field with angular velocity ω. If it is turned on while an escaper is crossing the field, his comrades arrange a diversion, which takes effect in T seconds time. Until then, however, the escaper must evade detection. The faster he moves, the easier it is to avoid the searchlight beam, but the more noise he makes, thus increasing his chances of being detected. If he starts at position $x \in R^2$ and follows trajectory $(x_t)_{0\le t\le T}$, we suppose his chances of being caught are

$$\exp\left[-\frac{1}{2}\int_0^T (\theta^2 x_u^T K_u x_u + \mid \dot{x}_u \mid^2) du\right],$$

where K is given by (3). Thus the escaper has to solve the optimisation problem

$$\min \int_0^T (\theta^2 x_u^T K_u x_u + \mid \dot{x}_u \mid^2) du, \quad x_0 = x.$$

In §3 we give the explicit closed-form solution to this optimisation problem; any prisoner who could solve it in real time would deserve to escape with his life, freedom and a Fields Medal!

We remark straight away that the test example (2)-(3) was not chosen arbitrarily, nor for its relevance to real life. It is a natural example to choose because we know two limits:

$$\begin{aligned} &\lim_{\omega\to 0} E\exp\left(-\frac{\theta^2}{2}\int_0^1 X_u^T K_u X_u du\right) \\ &= E\exp\left(-\frac{\theta^2}{2}\int_0^1 (X_u^1)^2 du\right) \\ &= (\cosh\theta)^{-1/2}; \end{aligned} \tag{8}$$

and

$$\begin{aligned} &\lim_{\omega\to\infty} E\exp\left(-\frac{\theta^2}{2}\int_0^1 X_u^T K_u X_u du\right) \\ &= E\exp\left(-\frac{\theta^2}{4}\int_0^1 \{(X_u^1)^2 + (X_u^2)^2\} du\right) \\ &= (\cosh(\theta/\sqrt{2}))^{-1}. \end{aligned} \tag{9}$$

(See, for example, Karatzas & Shreve [6] p.434). These identities provide us with two checks on our answer.

2. Quadratic functionals and optimal control

The starting point for the analysis is the following elementary result (the "Fundamental Theorem of Statistics"!), which featured largely in Chan, Dean, Jansons & Rogers [3].

Lemma 1. *Let X be a Gaussian random vector in R^d with mean* 0 *and covariance V, and let S be a non-negative definite symmetric matrix. Then for any $a \in R^d$,*

$$(10)\quad E\exp\{-\frac{1}{2}(X+a)^TS(X+a)\} = \det(I+SV)^{-1/2}\exp\{-\frac{1}{2}a^T(I+SV)^{-1}Sa\}.$$

It is once again elementary to confirm that

$$(11)\qquad a^T(I+SV)^{-1}Sa = \min_x\{(x+a)^TS(x+a)+x^TV^{-1}x\},$$

so that (10) can be re-expressed as

$$(12)\qquad E\exp\{-\frac{1}{2}(X+a)^TS(X+a)\}$$

$$= E\exp(-\frac{1}{2}X^TSX)\exp[-\frac{1}{2}\min_x\{(x+a)^TS(x+a)+x^TV^{-1}x\}].$$

The examples we have in mind are formally similar to (10)-(12) if we replace the Hilbert space R^d with $L^2([0,1])$, take the quadratic functional

$$(x,Sx)\equiv\theta^2\int_0^1 x_u^TK_ux_udu,$$

and write, in place of the log-likelihood $-\frac{1}{2}x^TV^{-1}x$ of the Gaussian law on R^d, the log-likelihood of Wiener measure on $C([0,1])$, namely

$$-\frac{1}{2}\int_0^1 \mid \dot{x}_u \mid^2 du.$$

(Of course, this is only a heuristic; but, as always, it leads us to a correct result which is easy to prove rigorously by other means.) We continue to assume such differentiability as we need until the statement and proof of Theorem 1. Thus we expect that if a is some deterministic bounded measurable function, then for each $t\in[0,1]$,

$$\begin{aligned}
& E\left[\exp\left\{-\frac{1}{2}\int_t^1(X_u+a_u)^TS_u(X_u+a_u)du\right\}\mid X_t=x\right]\\
&= E\left[\exp\left\{-\frac{1}{2}\int_t^1(X_u+x+a_u)^TS_u(X_u+x+a_u)du\right\}\mid X_t=0\right]\\
&= E\left[\exp\left(-\frac{1}{2}\int_t^1 X_u^TS_uX_udu\right)\mid X_t=0\right]\\
&\quad\times\exp\left[-\frac{1}{2}\min_{y_t=0}\left\{\int_t^1[(y_u+x+a_u)^TS_u(y_u+x+a_u)+\mid\dot{y}_u\mid^2]du\right\}\right]\\
(13) &= E\left[\exp\left(-\frac{1}{2}\int_t^1 X_u^TS_uX_udu\right)\mid X_t=0\right]\ \exp\left\{-\frac{1}{2}(x+b_t)^TQ_t(x+b_t)-\psi_t\right\},
\end{aligned}$$

for some non-negative definite $n \times n$ matrix Q_t, n-vector b_t, and scalar ψ_t;

$$= \exp\left\{ -\frac{1}{2}(x+b_t)^T Q_t (x+b_t) - \psi_t - \rho_t \right\}, \tag{14}$$

where we have defined ρ_t by

$$\rho_t \equiv -\log E\left[\exp\left(-\frac{1}{2}\int_t^1 X_u^T S_u X_u du \right) \mid X_t = 0 \right];$$

and we also suppose b is C^1 near to 1. There would be two consequences if this were true. Firstly, if we set

$$V(t,x) \equiv \exp\left\{ -\frac{1}{2}(x+b_t)^T Q_t (x+b_t) - \psi_t - \rho_t \right\}, \tag{15}$$

then

$$\begin{aligned} M_t &\equiv E\left[\exp\left\{ -\frac{1}{2}\int_0^1 (X_u+a_u)^T S_u (X_u+a_u) du \right\} \mid \mathcal{F}_t \right] \\ &= \exp\left\{ -\frac{1}{2}\int_0^t (X_u+a_u)^T S_u (X_u+a_u) du \right\} V(t,X_t) \end{aligned} \tag{16}$$

is a martingale, and hence, by Itô's formula,

$$\frac{1}{2}\Delta V(t,x) + \dot{V}(t,x) - \frac{1}{2}(x+a_t)^T S_t (x+a_t) V(t,x) = 0,$$

or again,

$$\begin{aligned} &\frac{1}{2}(x+b_t)^T Q_t^2 (x+b_t) - \frac{1}{2}\operatorname{tr} Q_t - \frac{1}{2}(x+b_t)^T \dot{Q}_t (x+b_t) \\ &-\dot{b}_t^T Q_t (x+b_t) - \dot{\psi}_t - \dot{\rho}_t - \frac{1}{2}(x+a_t)^T S_t (x+a_t) = 0, \end{aligned} \tag{17}$$

using the form (15). Considering the terms quadratic in x, we deduce that

$$Q_t^2 - \dot{Q}_t - S_t = 0, \tag{18}$$

with the obvious boundary condition $Q_1 = 0$. Next, if we consider the terms linear in x, we learn that

$$Q_t^2 b_t - \dot{Q}_t b_t - Q_t \dot{b}_t - S_t a_t = 0,$$

or using (18) we can re-express this as

$$Q_t \dot{b}_t + S_t(a_t - b_t) = 0, \tag{19}$$

with the boundary condition $b_1 = a_1$ (since $Q_1 = 0$ and $\dot{b}_t$ remains bounded near 1 by assumption). Finally, considering the constant term in (17) yields

$$\frac{1}{2} b_t^T Q_t^2 b_t - \frac{1}{2} \operatorname{tr} Q_t - \frac{1}{2} b_t^T \dot{Q}_t b_t - \dot{b}_t^T Q_t b_t - \dot{\psi}_t - \dot{\rho}_t - \frac{1}{2} a_t^T S_t a_t = 0,$$

which can be reduced (using (18) and (19)) to

$$\frac{1}{2} \operatorname{tr} Q_t + \dot{\psi}_t + \dot{\rho}_t + \frac{1}{2}(b_t - a_t)^T S_t (b_t - a_t) = 0.$$

Thus, if we can find Q, b and $\psi + \rho$ to satisfy

$$Q_t^2 - \dot{Q}_t - S_t = 0, \quad Q_1 = 0; \tag{20.i}$$

$$Q_t \dot{b}_t + S_t(a_t - b_t) = 0, \quad b_1 = a_1; \tag{20.ii}$$

$$\frac{1}{2} \operatorname{tr} Q_t + \dot{\psi}_t + \dot{\rho}_t + \frac{1}{2}(b_t - a_t)^T S_t (b_t - a_t) = 0, \quad \psi_1 + \rho_1 = 0, \tag{20.iii}$$

then the process M defined by (16) and (15) is a martingale, and $V(t, x)$ has indeed the interpretation

$$V(t, x) = E\left[\exp\left\{-\frac{1}{2}\int_t^1 (X_u + a_u)^T S_u (X_u + a_u) du\right\} \mid X_t = x\right].$$

The second consequence, if (13) were true, would be that, since

$$\min_{y_t = x} \int_t^1 \left\{\frac{1}{2}(y_u + a_u)^T S_u (y_u + a_u) + \frac{1}{2} \mid \dot{y}_u \mid^2\right\} du = \frac{1}{2}(x + b_t)^T Q_t (x + b_t) + \psi_t,$$

we should have

$$f(t) \equiv \int_0^t \left\{\frac{1}{2}(x_u + a_u)^T S_u (x_u + a_u) + \frac{1}{2} \mid \dot{x}_u \mid^2\right\} du + \frac{1}{2}(x_t + b_t)^T Q_t (x_t + b_t) + \psi_t \tag{21}$$

is non-decreasing whatever the path $(x_t)_{0\le t\le 1}$, and is constant if x is the path minimising the cost functional. Differentiating f, we obtain (dropping the subscript 't')

$$(22)\quad \dot{f} = \frac{1}{2}(x+a)^T S(x+a) + \frac{1}{2} \mid \dot{x} \mid^2 + (\dot{x}+\dot{b})^T Q(x+b) + \frac{1}{2}(x+b)^T \dot{Q}(x+b) + \dot{\psi},$$

which is minimised over choice of $\dot{x}$ when

$$(23)\qquad \dot{x} = -Q(x+b)$$

to the value

$$\frac{1}{2}(x+a)^T S(x+a) - \frac{1}{2}(x+b)^T Q^2(x+b) + \frac{1}{2}(x+b)^T \dot{Q}(x+b) + \dot{\psi} + \dot{b}^T Q(x+b)$$

$$= \frac{1}{2}(x+b)^T(S - Q^2 + \dot{Q})(x+b) + (x+b)^T\{S(a-b) + Q\dot{b}\} + \frac{1}{2}(a-b)^T S(a-b) + \dot{\psi}.$$

Thus if (20.i)-(20.iii) hold, together with

$$(20.\text{iv})\qquad \frac{1}{2}\operatorname{tr} Q_t + \dot{\rho}_t = 0, \quad \rho_1 = 0,$$

then we do indeed have that f is non-decreasing, and constant if (23) holds (thus (23) is the differential equation for the least action path). It is clear that we need to investigate existence and uniqueness of solutions to (20). Various questions of this nature are dealt with in standard books on optimal control (see, for example, Kwakernaak & Sivan [7], or Anderson & Moore [1]) but usually only for the case $a \equiv 0$, and with assumptions on the positive-definiteness of various matrices which would not be satisfied by the rank-1 matrix S. So we here outline the proof of a result which covers what we need; the heuristics are now over.

Theorem 1. (i) *Let $t \longmapsto S_t$ be a bounded measurable map from R^+ into the cone of $n \times n$ non-negative-definite symmetric matrices. Then the equation*

$$(20.\text{i})\qquad Q_t^2 - \dot{Q}_t - S_t = 0,\ (t \le 1);\quad Q_1 = 0,$$

has a unique solution Q, which is symmetric and non-negative-definite. The solution Q has the two interpretations

$$(24) \qquad E\left[\exp\left\{-\frac{1}{2}\int_t^1 X_u^T S_u X_u du\right\} \mid X_t = x\right] = \exp\{-\frac{1}{2}x^T Q_t x - \rho_t\},$$

and

$$(25) \qquad \min_{y_t=x} \frac{1}{2}\int_t^1 \left\{y_u^T S_u y_u + \mid \dot{y}_u \mid^2\right\} du = \frac{1}{2}x^T Q_t x,$$

where we have defined

$$(26) \qquad \rho_t \equiv \int_t^1 \frac{1}{2} \operatorname{tr} Q_u du.$$

(ii) *If we have also that for every* $\epsilon > 0$,

$$(27) \qquad \int_{1-\epsilon}^1 S_u du \;\; \textit{is of full rank},$$

then Q_t *is positive-definite for all* $t < 1$.
(iii) *If we assume* (27) *and take bounded measurable* $a : [0,1] \longmapsto R^n$, *then there exists* $b : [0,1] \longmapsto R^n$ *and* $\psi : [0,1] \longmapsto R$ *such that*

$$E\left[\exp\left\{-\frac{1}{2}\int_t^1 (X_u + a_u)^T S_u (X_u + a_u) du\right\} \mid X_t = x\right]$$

$$(28) \qquad = \exp\{-\frac{1}{2}(x + b_t)^T Q_t (x + b_t) - \psi_t - \rho_t\},$$

and

$$(29) \quad \min_{y_t=x} \frac{1}{2}\int_t^1 \left\{(y_u + a_u)^T S_u (y_u + a_u) + \mid \dot{y}_u \mid^2\right\} du = \frac{1}{2}(x + b_t)^T Q_t (x + b_t) + \psi_t.$$

The functions b, ψ *solve*

$$(20.\text{ii}) \qquad Q_t \dot{b}_t + S_t(a_t - b_t) = 0, \quad (t < 1);$$

$$(20.\text{iii}) \qquad \dot{\psi}_t + (b_t - a_t)^T S_t (b_t - a_t) = 0.$$

The equation (20.ii) *has at most one bounded solution.*

Proof. (i) To begin with, let us assume that S is continuous, $S_t \geq \epsilon I$ for all t (here, $\epsilon > 0$ is small but fixed). We seek to solve

$$Q_t = \int_t^1 (S_u - Q_u^2) du, \tag{30}$$

which we can always do for t in some neighbourhood of 1, by the classical method of succesive approximations. The solution Q is clearly symmetric. It is *a priori* possible that Q might be explode, but Q will be unique until explosion, as the coefficients of the ODE are locally Lipschitz. We show that $Q_t > 0$ for all $t < 1$ by taking

$$\tau = \sup\{t < 1 : Q_t \text{ is not positive-definite } \},$$

and taking $x \neq 0$ such that $x^T Q_\tau x = 0$. We then see that

$$\frac{d}{dt} x^T Q_t x \Big|_{t=\tau} = -x^T S_\tau x < 0,$$

which contradicts the hypothesis that $x^T Q_t x > 0$ for $\tau < t < 1$.
Since Q is positive-definite, we see that

$$x^T Q_t x \leq \int_t^1 x^T S_u x du \leq c(1-t)\|x\|^2, \tag{31}$$

for some constant c, and thus the solution Q does *not* explode, and uniqueness for all time follows. The interpretations (24) and (25) are justified by the arguments (15)-(16) and (21)-(23) respectively.
To extend now to bounded measurable S, we may, by Egorov's theorem, approximate such S by continuous uniformly elliptic $S^{(n)}$ for which (24), (25), (30) and (31) hold. The bound (31) ensures that by passing to a subsequence we may assume the solutions $Q^{(n)}$ converge, and the limiting forms of (24) and (25) are seen to hold.
(ii) To see that Q_t must be positive-definite for $t < 1$, take the interpretation (24) and consider what happens to either side as $|\, x \,| \to \infty$.
(iii) Again, we shall prove this assuming initially that S is continuous and uniformly elliptic, and a is continuous. We can begin by assuming also that S and a are piecewise constant functions, constant on each interval $((j-1)2^{-n}, j2^{-n}]$. Next, we take the values $X(j2^{-n})$, form the continuous piecewise-linear interpretation $X^{(n)}$ and consider the quadratic functional

$$\frac{1}{2}\int_t^1 (X_u^{(n)} + a_u)^T S_u (X_u^{(n)} + a_u) du$$

and the analogous deterministic problem. These are now both discrete-time problems, so we may directly apply (10), (11) and (12) to deduce the forms (28) and (29). These expressions keep the same form in the limit. We may now approximate general continuous S and a by piecewise constant S, a and the form of the expressions (28), (29) in the limit is unchanged. The differential equations (20.ii) and (20.iii) can now be derived by considering a change from t to $t+h$ in the quadratic functional (28) or in the minimisation problem (29); this analysis establishes that b and ψ are C^1, which is not immediately apparent from their derivation. We omit the routine but tedious details of this analysis. Passing from continuous to bounded measurable S and a present no further problems, (20.i)-(20.iii) must now be understood in integrated form.
The final assertion that (20.ii) has at most one bounded solution is clear when we consider the martingale (16) - which *is* a martingale, because (20) is satisfied. This shows that the equality (28) is valid, so b is uniquely determined (recall that $Q_t > 0$).

Remarks. (i) Theorem 1 provides a powerful method for calculating the laws of additive functionals. We show it at work on (1), the moment of inertia of a Brownian polymer.
If we can compute

$$\begin{aligned} & E\exp\left(-\frac{1}{2}\theta^2\int_0^1 (B_s+x)^2 ds\right) \\ (32) \qquad & = E\exp\left(-\frac{1}{2}\theta^2\int_0^1 B_s^2 ds - x\theta^2\int_0^1 B_s ds - \frac{1}{2}\theta^2 x^2\right), \end{aligned}$$

then if we multiply by $\exp(\theta^2 x^2/2)$ and mix over x with $N(0,\theta^{-2})$ distribution, we obtain (1); this trick was used extensively in [3]. But it is simple to compute (32); just take $S \equiv \theta^2$ in Theorem 1, solve (20.i) to obtain

$$Q_t = \theta\tanh\theta(1-t)$$

and then

$$\rho_t \equiv \int_t^1 \frac{1}{2}\,\mathrm{tr}\,Q_u du = \frac{1}{2}\log\cosh\theta(1-t).$$

Thus

$$E\exp\left(-\frac{1}{2}\theta^2\int_0^1 (B_s+x)^2 ds\right) = (\cosh\theta)^{-1/2}\exp\left\{-\frac{1}{2}x^2\theta\tanh\theta\right\}.$$

Mixing over x with an $N(0,\theta^{-2})$ distribution yields the result

$$E\exp\left(-\frac{1}{2}\theta^2\int_0^1 (B_s-\overline{B})^2 ds\right) = (\theta \text{ cosech } \theta)^{1/2}.$$

See also Donati-Martin & Yor [4], Chan, Dean, Jansons & Rogers [3] for other proofs of this remarkable fact. Notice that we have not needed the Ray-Knight theorem.
(ii) The situation for $a \neq 0$ is much more untidy than one would wish. One would like to say that (20.ii) has a unique solution, but this is made difficult by the fact that $Q_t \to 0$ as $t \uparrow 1$. Indeed, in the one-dimensional example just considered, with $a \equiv 0$, we find that

$$b(t) \equiv 0 \text{ and } b(t) = \text{ cosech } \theta(1-t)$$

are both solutions to (20.ii). The final uniqueness assertion is some comfort, but it is not clear that the b appearing in (28) and (29) necessarily should be bounded! Since our application in §3 does not need these complications, we now lay them to one side.

3. The Colditz example

In view of Theorem 1, to compute (2) we have to solve the matrix Riccati equation

$$Q_t^2 - \dot{Q}_t - \theta^2 K_t = 0, \quad Q_1 = 0, \tag{33}$$

where K is given by (3). Before attacking this, we rework the problem into a more manageable form. If we set

$$R_t \equiv \begin{pmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{pmatrix}, \quad U_t \equiv R_t^T Q_t R_t,$$

then the matrix Ricatti equation becomes

$$\dot{U}_t = U_t^2 + \omega U_t C - \omega C U_t - \theta^2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

where

$$C \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Now rephrasing this in time-to-go by setting $V_t \equiv U_{1-t}$, and expressing

$$V_t \equiv S_t^T \Lambda_t S_t, \quad S_t \equiv \begin{pmatrix} \cos\phi_t & -\sin\phi_t \\ \sin\phi_t & \cos\phi_t \end{pmatrix}, \quad \Lambda_t = \begin{pmatrix} \lambda_1(t) & 0 \\ 0 & \lambda_2(t) \end{pmatrix},$$

we obtain a differential equation for the eigenvalues λ_1, λ_2 of Q, and for the phase ϕ:

$$
\begin{aligned}
\dot\Lambda &\equiv \frac{d}{dt}(SVS^T)\\
&= -\Lambda^2 + \theta^2 \begin{pmatrix}\cos\phi\\ \sin\phi\end{pmatrix}\begin{pmatrix}\cos\phi & \sin\phi\end{pmatrix} + (\dot\phi C + \omega SCS^T)\Lambda - \Lambda(\dot\phi C + \omega SCS^T),
\end{aligned}
$$

which yields

$$
\begin{aligned}
&(34.\mathrm{i}) && \dot\lambda_1 = -\lambda_1^2 + \theta^2\cos^2\phi;\\
&(34.\mathrm{ii}) && \dot\lambda_2 = -\lambda_2^2 + \theta^2\sin^2\phi;\\
&(34.\mathrm{iii}) && (\lambda_1-\lambda_2)(\dot\phi+\omega) + \theta^2\sin\phi\cos\phi = 0.
\end{aligned}
$$

with boundary conditions $\lambda_1(0) = \lambda_2(0) = 0$. No boundary condition needs to be imposed on ϕ, since $\lambda_1(0) = \lambda_2(0) = 0$ together with (34) implies that Q solves (33), and we already know that there exists a unique solution.
For notational simplicity, we shall write throughout this section:

$$
\begin{aligned}
&\beta = \sqrt{\omega^2+\theta^2};\\
&\sigma = \sqrt{(\beta-\omega)(\beta+3\omega)};\\
&\tau = \sqrt{(\beta+\omega)(\beta-3\omega)};\\
&r(t) = \tau^2(\beta+\omega)\cosh\sigma t + \sigma^2(\beta-\omega)\cosh\tau t - 8\omega^2\beta;\\
&A = \sqrt{\dot r^2 + 2\theta^2 r^2 - 2r\ddot r}\\
&C = (\beta-3\omega)\cosh\sigma t + (\beta+3\omega)\cosh\tau t - 2\beta.
\end{aligned}
$$

Let us consider the following functions:

$$
\begin{aligned}
&(35.\mathrm{i}) && \phi = -\frac{1}{2}\arcsin\frac{2\theta^2\omega C}{A};\\
&(35.\mathrm{ii}) && \lambda_1 = \frac{\dot r + A}{2r};\\
&(35.\mathrm{iii}) && \lambda_2 = \frac{\dot r - A}{2r}.
\end{aligned}
$$

It is the purpose of this section to prove the following explicit form for the solution to (34).

Theorem 2. *The functions λ_1, λ_2 and ϕ defined in* (35) *solve* (34).

By analytic continuation, we can suppose without loss of generality that $\beta > 3\omega$ throughout this section. First of all, we have to justify our definitions of A and ϕ, for we always work on R in this section. We observe that $\dot r^2(0)+2\theta^2 r^2(0)-2r(0)\ddot r(0)=0$, and that

$$\begin{aligned}\frac{d}{dt}(\dot r^2+2\theta^2 r^2-2r\ddot r) &= 2r(2\theta^2\dot r - r^{(3)})\\ &= 2r\theta^4\left[\sigma(\beta-3\omega)\sinh(\sigma t)+\tau(\beta+3\omega)\sinh(\tau t)\right]\\ &\geq 0,\end{aligned}$$

which implies that $\dot r^2+2\theta^2 r^2-2r\ddot r$ is non-negative. Thus A is well-defined. Let us write down

$$\begin{aligned} r &= \tau^2(\beta+\omega)\cosh\sigma t+\sigma^2(\beta-\omega)\cosh\tau t-8\omega^2\beta;\\ \dot r &= \sigma\tau^2(\beta+\omega)\sinh\sigma t+\sigma^2\tau(\beta-\omega)\sinh\tau t;\\ \ddot r &= \theta^2\tau^2(\beta+3\omega)\cosh\sigma t+\theta^2\sigma^2(\beta-3\omega)\cosh\tau t;\\ r^{(3)} &= \theta^2\sigma\tau^2(\beta+3\omega)\sinh\sigma t+\theta^2\sigma^2\tau(\beta-3\omega)\sinh\tau t;\\ r^{(4)} &= \theta^4(\beta-3\omega)(\beta+3\omega)^2\cosh\sigma t+\theta^4(\beta-3\omega)^2(\beta+3\omega)\cosh\tau t.\end{aligned}$$

We next prove the following lemma:

Lemma 2. *We have,*

$$4\theta^8\omega^2C^2 = 2\theta^6 r^2-2\theta^4 r\ddot r-3\theta^4\dot r^2-(r^{(3)})^2+4\theta^2\dot r r^{(3)}. \tag{36}$$

Proof of Lemma 2. It follows that

$$\begin{aligned} &2\theta^6 r^2-2\theta^4 r\ddot r = 2\theta^4 r(\theta^2 r-\ddot r)\\ &= 2\theta^6 r[\tau^2(\beta+\omega)\cosh\sigma t+\sigma^2(\beta-\omega)\cosh\tau t-8\omega^2\beta\\ &\qquad -\tau^2(\beta+3\omega)\cosh\sigma t-\sigma^2(\beta-3\omega)\cosh\tau t]\\ &= 2\theta^6 r[-2\omega\tau^2\cosh\sigma t+2\omega\sigma^2\cosh\tau t-8\omega^2\beta]\\ &= 4\theta^6\omega[\tau^2(\beta+\omega)\cosh\sigma t+\sigma^2(\beta-\omega)\cosh\tau t-8\omega^2\beta]\\ &\qquad\times[-\tau^2\cosh\sigma t+\sigma^2\cosh\tau t-4\omega\beta]\\ &= 4\theta^6\omega[-\tau^4(\beta+\omega)\cosh^2\sigma t+\theta^2\tau^2(\beta+3\omega)\cosh\sigma t\cosh\tau t\\ &\qquad -4\omega\beta\tau^2(\beta+\omega)\cosh\sigma t-\theta^2\sigma^2(\beta-3\omega)\cosh\sigma t\cosh\tau t\\ &\qquad +\sigma^4(\beta-\omega)\cosh^2\tau t-4\omega\beta\sigma^2(\beta-\omega)\cosh\tau t\\ &\qquad +8\omega^2\beta\tau^2\cosh\sigma t-8\omega^2\beta\sigma^2\cosh\tau t+32\omega^3\beta^2]\\ &= 4\theta^6\omega[-\tau^4(\beta+\omega)\cosh^2\sigma t+\sigma^4(\beta-\omega)\cosh^2\tau t\\ &\qquad +2\theta^2\omega(\beta-3\omega)(\beta+3\omega)\cosh\sigma t\cosh\tau t\\ &\qquad -4\theta^2\omega\beta(\beta-3\omega)\cosh\sigma t-4\theta^2\omega\beta(\beta+3\omega)\cosh\tau t+32\omega^3\beta^2];\end{aligned}$$

$$\begin{aligned}
3\theta^2\dot{r} - r^{(3)} &= \theta^2[\sigma\tau^2(3\beta + 3\omega - \beta - 3\omega)\sinh\sigma t + \sigma^2\tau(3\beta - 3\omega - \beta + 3\omega)\sinh\tau t] \\
&= 2\theta^2\beta[\sigma\tau^2\sinh\sigma t + \sigma^2\tau\sinh\tau t];
\end{aligned}$$

$$\begin{aligned}
\theta^2\dot{r} - r^{(3)} &= \theta^2[\sigma\tau^2(\beta + \omega - \beta - 3\omega)\sinh\sigma t + \sigma^2\tau(\beta - \omega - \beta + 3\omega)\sinh\tau t] \\
&= 2\theta^2\omega[-\sigma\tau^2\sinh\sigma t + \sigma^2\tau\sinh\tau t].
\end{aligned}$$

So

$$\begin{aligned}
&-3\theta^4\dot{r}^2 - (r^{(3)})^2 + 4\theta^2\dot{r}r^{(3)} = -(3\theta^2\dot{r} - r^{(3)})(\theta^2\dot{r} - r^{(3)}) \\
&\quad = 4\theta^4\omega\beta[\sigma^2\tau^4\sinh^2\sigma t - \sigma^4\tau^2\sinh^2\tau t] \\
&\quad = 4\theta^6\omega[\beta\tau^2(\beta - 3\omega)(\beta + 3\omega)\sinh^2\sigma t - \beta\sigma^2(\beta - 3\omega)(\beta + 3\omega)\sinh^2\tau t] \\
&\quad = 4\theta^6\omega[\beta\tau^2(\beta - 3\omega)(\beta + 3\omega)\cosh^2\sigma t \\
&\qquad\quad - \beta\sigma^2(\beta - 3\omega)(\beta + 3\omega)\cosh^2\tau t + 4\omega\beta^2(\beta - 3\omega)(\beta + 3\omega)].
\end{aligned}$$

Therefore

$$\begin{aligned}
&\quad 2\theta^6 r^2 - 2\theta^4 r\ddot{r} - 3\theta^4\dot{r}^2 - (r^{(3)})^2 + 4\theta^2\dot{r}r^{(3)} \\
&= 4\theta^6\omega[\tau^2(\beta - 3\omega)(\beta(\beta + 3\omega) - (\beta + \omega)^2)\cosh^2\sigma t \\
&\qquad + \sigma^2(\beta + 3\omega)((\beta - \omega)^2 - \beta(\beta - 3\omega))\cosh^2\tau t \\
&\qquad + 2\theta^2\omega(\beta - 3\omega)(\beta + 3\omega)\cosh\sigma t\cosh\tau t - 4\theta^2\omega\beta(\beta - 3\omega)\cosh\sigma t \\
&\qquad - 4\theta^2\omega\beta(\beta + 3\omega)\cosh\tau t + 4\omega\beta^2(8\omega^2 + (\beta - 3\omega)(\beta + 3\omega))] \\
&= 4\theta^6\omega[\theta^2\omega(\beta - 3\omega)^2\cosh^2\sigma t + \theta^2\omega(\beta + 3\omega)^2\cosh^2\tau t \\
&\qquad + 2\theta^2\omega(\beta - 3\omega)(\beta + 3\omega)\cosh\sigma t\cosh\tau t \\
&\qquad - 4\theta^2\omega\beta(\beta - 3\omega)\cosh\sigma t - 4\theta^2\omega\beta(\beta + 3\omega)\cosh\tau t + 4\theta^2\omega\beta^2] \\
&= 4\theta^8\omega^2[(\beta - 3\omega)^2\cosh^2\sigma t + (\beta + 3\omega)^2\cosh^2\tau t \\
&\qquad + 2(\beta - 3\omega)(\beta + 3\omega)\cosh\sigma t\cosh\tau t \\
&\qquad - 4\beta(\beta - 3\omega)\cosh\sigma t - 4\beta(\beta + 3\omega)\cosh\tau t + 4\beta^2] \\
&= 4\theta^8\omega^2 C^2,
\end{aligned}$$

completing the proof of Lemma 2.

Proof of Theorem 2. It follows from (36) that

$$\begin{aligned}
\left(\frac{2\theta^2\omega C}{A}\right)^2 &= \frac{4\theta^8\omega^2C^2}{A^2\theta^4} \\
&= \frac{2\theta^6r^2 - 2\theta^4r\ddot{r} - 3\theta^4\dot{r}^2 - (r^{(3)})^2 + 4\theta^2\dot{r}r^{(3)}}{A^2\theta^4} \\
&= 1 - \frac{4\theta^4\dot{r}^2 - 4\theta^2\dot{r}r^{(3)} + (r^{(3)})^2}{A^2\theta^4} \\
&= 1 - \left(\frac{2\theta^2\dot{r} - r^{(3)}}{A\theta^2}\right)^2,
\end{aligned}$$

which is strictly smaller than 1 for every $t > 0$. Thus our definition of ϕ is also justified. When t is positive, ϕ is strictly between $-\pi/4$ and 0. We also observe from the previous calculation that

$$\cos(2\phi) = \frac{2\theta^2\dot{r} - r^{(3)}}{A\theta^2}\ . \tag{37}$$

Let us now prove (34). By definition of λ_1, we have

$$\begin{aligned}
\dot{\lambda}_1 + \lambda_1^2 &= \frac{(\ddot{r} + \dot{A})r - (\dot{r} + A)\dot{r}}{2r^2} + \frac{\dot{r}^2 + A^2 + 2\dot{r}A}{4r^2} \\
&= \frac{r\ddot{r} + A^{-1}r^2(2\theta^2\dot{r} - r^{(3)}) - \dot{r}^2 - \dot{r}A + \dot{r}^2 + \theta^2r^2 - r\ddot{r} + \dot{r}A}{2r^2} \\
&= \frac{2\theta^2\dot{r} - r^{(3)}}{2A} + \frac{\theta^2}{2},
\end{aligned}$$

which, according to (37), is equal to $\theta^2(\cos(2\phi) + 1)/2 = \theta^2\cos^2\phi$. Thus (34.i) is satisfied by our λ_1 and ϕ. A similar argument can be used to verify (34.ii). Now we turn to (34.iii). It follows from the definition of r and θ that

$$\begin{aligned}
&\theta^4r - 2\theta^2\ddot{r} + r^{(4)} \\
&= \theta^4[\tau^2(\beta+\omega)\cosh\sigma t + \sigma^2(\beta-\omega)\cosh\tau t - 8\omega^2\beta \\
&\qquad - 2\tau^2(\beta+3\omega)\cosh\sigma t - 2\sigma^2(\beta-3\omega)\cosh\tau t \\
&\qquad + (\beta-3\omega)(\beta+3\omega)^2\cosh\sigma t + (\beta-3\omega)^2(\beta+3\omega)\cosh\tau t] \\
&= \theta^4[4\omega^2(\beta-3\omega)\cosh\sigma t + 4\omega^2(\beta+3\omega)\cosh\tau t - 8\omega^2\beta] \\
&= 4\theta^4\omega^2C. \qquad (38)
\end{aligned}$$

On the other hand, we have

$$
\begin{aligned}
\dot{\phi} &= -\frac{1}{2\sin 2\phi}\frac{d}{dt}\left(\frac{2\theta^2\dot{r} - r^{(3)}}{A\theta^2}\right) \\
&= \frac{A}{4\theta^2\omega C}\frac{(2\theta^2\ddot{r} - r^{(4)})A - (2\theta^2\dot{r} - r^{(3)})\dot{A}}{A^2\theta^2} \\
&= \frac{(2\theta^2\ddot{r} - r^{(4)})A^2 - r(2\theta^2\dot{r} - r^{(3)})^2}{4\theta^4\omega C A^2}.
\end{aligned}
$$

Therefore we get that

$$
\begin{aligned}
&(\lambda_1 - \lambda_2)(\dot{\phi} + \omega) + \theta^2 \sin\phi\cos\phi \\
&= \frac{A}{r}\left(\frac{(2\theta^2\ddot{r} - r^{(4)})A^2 - r(2\theta^2\dot{r} - r^{(3)})^2}{4\theta^4\omega C A^2} + \omega\right) - \frac{\theta^4\omega C}{A} \\
&= \frac{(2\theta^2\ddot{r} - r^{(4)})A^2 - r(2\theta^2\dot{r} - r^{(3)})^2 + 4\theta^4\omega^2 C A^2 - 4r\theta^8\omega^2 C^2}{4r\theta^4\omega C A} \\
&= \frac{(2\theta^2\ddot{r} - r^{(4)})A^2 - r(2\theta^6 r^2 - 2\theta^4 r\ddot{r} + \theta^4\dot{r}^2) + 4\theta^4\omega^2 C A^2}{4r\theta^4\omega C A^2} \qquad (\text{ by } (36)) \\
&= \frac{(2\theta^2\ddot{r} - r^{(4)} - \theta^4 r + 4\theta^4\omega^2 C)A^2}{4r\theta^4\omega C A^2} \\
&= 0;
\end{aligned}
$$

here, the last equality is due to (38). Theorem 2 is thus proved .

We observe that

$$
\operatorname{tr} Q_t = \lambda_1(1-t) + \lambda_2(1-t) = \frac{\dot{r}(1-t)}{r(1-t)}.
$$

If we follow the notation of (7), we get:

$$
\gamma(t) = \frac{1}{2}\log\frac{r(1-t)}{r(0)}.
$$

Therefore by (4) we obtain the following

Theorem 3. *We have,*

$$
E\exp\left(-\frac{\theta^2}{2}\int_0^1 X_u^T K_u X_u du\right) = \left(\frac{r(1)}{2\beta(\beta - 3\omega)(\beta + 3\omega)}\right)^{-1/2};
$$

with

$$
r(1) = \tau^2(\beta + \omega)\cosh\sigma + \sigma^2(\beta - \omega)\cosh\tau - 8\omega^2\beta.
$$

Remark. Taking $\omega = 0$ and $\omega = +\infty$, we get (8) and (9) respectively.

Acknowledgement. We are very grateful to Frank Pijpers and Thomas Wolf for helpful discussions on the example in §3, and to Erwin Bolthausen, Alain Sznitman and Marc Yor for further references.

References

[1] Anderson, B.D.O. & Moore, J.B.: *Linear Optimal Control.* Prentice-Hall, New Jersey, 1971.

[2] Cameron, R.H. & Martin, W.T.: Evaluation of various Wiener integrals by use of certain Sturm-Liouville differential equations. *Bull. Amer. Math. Soc.* **51** (1945) 73-90.

[3] Chan, T., Dean, D., Jansons, K.M. & Rogers, L.C.G.: On polymer conformations in elongational flows. *In preparation.*

[4] Donati-Martin, C. & Yor, M.: Fubini's theorem for double Wiener integrals and the variance of the Brownian path. *Ann. Inst. Henri Poincaré* **27** (1991) 181-200.

[5] Kac, M.: *Integration in Function Spaces and Some of Its Applications.* Accademia Nazionale dei Lincei, Scula Normale Superiore, Pisa, 1980.

[6] Karatzas, I. & Shreve, S.E.: *Brownian Motion and Stochastic Calculus.* Springer, New York, 1988.

[7] Kwakernaak, H. & Sivan, R.: *Linear Optimal Control Systems.* Wiley, New York, 1972.

[8] Liptser, R.S. & Shiryaev, A.N.: *Statistics of Random Processes.* Vol. I: *General Theory.* Springer, New York, 1977.

L.C.G. Rogers & Z. Shi
School of Mathematical Sciences
Queen Mary & Westfield College
Mile End Road
London E1 4NS