Trivariate support of flat-volatility forward Libor rates

FARSHID JAMSHIDIAN

Keywords: Support; Brownian motion; Libor market model; constrained functional optimization; Euler-Lagrange equation; Beltrami identity.

Abstract. This paper investigates the multivariate support of forward Libor rates in the one-factor, constant volatilities Libor market model. The comparatively simple bivariate case was solved in Jamshidian [2] in connection to the recent finding by Davis and Mataix-Pastor [1] of positive probability of negative Libor rates in the swap market model. The approach here builds on [2] but becomes really effective only in the trivariate case, and there particularly for a special "flat-volatility" case, leading to an analytic solution. The main idea is a certain recursion in the Libor market model by means of which the calculation of the support is reduced to a calculus of variation problem (with bounds on the slope).

1. Introduction and the main result

Recently, Davis and Mataix-Pastor [1] employed the Strook-Varadhan Support Theorem to show that forward Libor rates eventually become negative with positive probability in the lognormal swap market model. To pursue this phenomenon, Jamshidian [2] calculated the bivariate support of forward Libor and swap rates. For constant volatilities, the bivariate support turned out to be the region between two very simple graphs. A natural question raised by Davis and Mataix-Pastor is generalization to higher dimensions.

The topology of the multivariate support of forward Libor rates $L^1, \cdots, L^n$ is not so difficult to guess: if the covariation $[\log L^i, \log L^j]_t$ of the log-rates equals $\sigma_{ij}t$ for some constant matrix (σ_{ij}) of rank k, then the joint support $\mathcal{S}(\log L_t)$ as a closed subset of $\mathbb{R}^n$ seems to be homeomorphic to $\mathbb{R}^k \times [0,1]^{n-k}$ for $t > 0$. The precise analytic geometry of the support is a much harder problem however. This paper takes a first step in that direction.

Our approach is to reduce the calculation of the trivariate support to a calculus of variation problem and solve it for a special "flat-volatility" case. The multivariate case is studied too, but the results are less conclusive. The exposition is more explorative than formal. We use some properties from Jamshidian [2] of the Brownian motion as a "prolific process," but informally, eschewing secondary technicalities.

The calculus of variation formulation in the multivariate case $n \geq 4$ involves nonstandard equality constraints, and the first-order conditions take the form of a $2n-5$ by $2n-5$ ODE system, with little hope for an analytic solution. The formulation in the trivariate case is

Part-time Professor of Applied Mathematics, FELAB, University of Twente.

Version 13-March-2007. For possible future updates visit wwwhome.math.utwente.nl/~ jamshidianf.

standard, save for inequality constraints bounding the slope. Still, the Largrangian is too complex in the general constant-volatility trivariate case, hindering further progress.

Substantial simplification occurs in the trivariate "flat-volatility" case of (two) equal volatilities. We solve the unconstrained Euler-Lagrange equation analytically by explicitly integrating the Beltrami identity, verify the second order conditions for local maximality, and find the local maximum of the slope inequality-constrained problem. Due to certain complications, only the easier "positive energy" case $L^2_0 \leq L^3_0$ is worked out in full detail.

The analytical tractability of the problem is rather interesting. However, the paper falls short in completing the proof that the proposed solution actually works. At issue is the maximization of a functional $\mathcal{L}$ of Euler-Lagrange form $\mathcal{L}(a) = \int_0^T L(a_t, \dot{a}_t)dt$ on the space of C^1 functions $a = a_t$ on $[0, T]$, subject to given a_0 and a_T and the constraint $0 \leq \dot{a} \leq 1$. (Here $\dot{a} = \frac{da}{dt}$.) We show that $\mathcal{L}$ (restricted to the feasible set) has a unique local maximum. If we could show that $\mathcal{L}$ has a global maximum, i.e., attains its supremum, then it would follow that our solution is indeed the global maximum. But we leave this question open. As such, our analysis merely identifies a subset of the trivariate support, and the contention that this subset is the whole support hovers admittedly at conjecture level.

1.1. **The trivariate, flat-volatility, Libor market model.** For convenience, in this section (only) we denote the last three forward Libor rates L^i by

$$X := L^n, \quad Y := L^{n-1}, \quad Z := L^{n-2}.$$

The trivariate one-factor Libor market model with flat volatility $\sigma > 0$ and positive initial values $X_0, Y_0, Z_0 > 0$ is governed by the (recursive) SDE system

$$\frac{dX}{X} = \sigma dW,$$
$$\frac{dY}{Y} = -\frac{\sigma^2 X}{1+X}dt + \sigma dW,$$
$$\frac{dZ}{Z} = -\sigma^2 (\frac{Y}{1+Y} + \frac{X}{1+X})dt + \sigma dW,$$

where W is a Brownian motion in an equivalent measure, and σ is a deterministic, positive, locally square-integrable function on $[0, \infty)$. It is not difficult to show that the bivariate support $\mathcal{S}(X_t, Y_t)$ (and similarly $\mathcal{S}(Y_t, Z_t)$) is simply given by wedge between two rays:

$$\mathcal{S}(X_t, Y_t) = \{(x, y) \in [0, \infty)^2 : e^{-\int_0^t \sigma_s^2 ds} \frac{x}{X_0} \leq \frac{y}{Y_0} \leq \frac{x}{X_0}\}. \tag{1.1}$$

1.2. **The function** $A_*(T, r, r_0)$**.** First, for $0 < r \leq p < 1$, define the function $t^*(r, p)$ by

$$t^*(r, p) := r + \log(\frac{p}{r}) + \sqrt{\frac{p}{1-p}}(\arcsin(\sqrt{p}) - \arcsin(\sqrt{r}) - \sqrt{r - r^2}). \tag{1.2}$$

(One has $t^*(r,r)=0$ and $t^*(r,p)>0$ if $r<p$). Now, for $T>0$ and $0<r\le r_0\le 1$, define the function $A_*(T,r,r_0)$ (which satisfies $0\le A_*(T,r,r_0)<T$) analytically by

$$A_*(T,r,r_0) := s - r + r_0e^{-s} + \frac{(\arcsin(\sqrt{r_0e^{-s}}) + \sqrt{r_0e^{-s}-{r_0}^2e^{-2s}} - \arcsin(\sqrt{r}) - \sqrt{r-r^2})^2}{T+\log(r)-r-\log(r_0)+r_0e^{-s}}, \tag{1.3}$$

where $s=0$ if $r_0<1$ and $T\ge t^*(r,r_0)$, and otherwise s is the solution to the equation $T-s=t^*(r,r_0e^{-s})$ (which exists, is unique, and satisfies $0<s<\log(r_0/r)$).

1.3. **Statement of the main result.** Consider the trivariate flat-volatility model, and assume further that $Y_0\le X_0$. This readily implies $Y_t<X_t$ and $\frac{Z_0Y_t}{Y_0Z_t}<\frac{Y_0X_t}{X_0Y_t}$ for $t>0$.

Conditional on our conjecture on the existence of a global maximum, we contend that[1]

$$\mathcal{S}(X_t,Y_t,Z_t) = \{(x,y,z)\in(0,\infty)^3 : e^{-\int_0^t\sigma_s^2ds}\frac{x}{X_0}\le\frac{y}{Y_0}\le\frac{x}{X_0}, \quad e^{A_*(\int_0^t\sigma_s^2ds,\frac{y}{x},\frac{Y_0}{X_0})}\le\frac{Z_0y}{Y_0z}\le\frac{Y_0x}{X_0y}\}. \tag{1.4}$$

As in Eq. (1.1), the first inequality gives the bivariate support $\mathcal{S}(X_t,Y_t)$. The second inequality is explained more easily by rewriting it equivalently as

$$\frac{X_0y^2}{Y_0^2x}\le\frac{z}{Z_0}\le\frac{y}{Y_0}e^{-A_*(\int_0^t\sigma_s^2ds,\frac{y}{x},\frac{Y_0}{X_0})}.$$

This inequality describes the trivariate support as the three-dimensional region between the graphs of two functions over the bivariate support. The upper graph intersects lower one at the two line edges and covers it on both sides like a (widening) tunnel. The boundary of the trivariate support consists of the union of the two graphs. Using that $A_*\ge0$ and $A_*(T,r,r)=0$ for all r, Eq. (1.4) above is equivalent to the more succinct expression

$$\mathcal{S}(X_t,Y_t,Z_t) = \{(x,y,z)\in(0,\infty)^3 : e^{A_*(\int_0^t\sigma_s^2ds,\frac{y}{x},\frac{Y_0}{X_0})}\le\frac{Z_0y}{Y_0z}\le\frac{Y_0x}{X_0y}\le e^{\int_0^t\sigma_s^2ds}\}. \tag{1.5}$$

2. Summary of the derivation

This section outlines the main ideas of the derivation of Eq. (1.4). The first three subsections reduce the trivariate flat-volatility case to a calculus of variation problem with slope inequality constraints; they are detailed and generalized to non-flat volatility in Section 3 and partially extended to the multivariate case in Section 5. The remaining subsections derive and describe the solution, and are discussed in detail in Section 4. Some routine but tedious calculations in the sequel are delegated to an Appendix (Section 6).

[1] This support is a closed of subset $(0,\infty)^3$ viewed as the range of (X,Y,Z). The support with range viewed as $\mathbb{R}^3$ is the closure of this set in $\mathbb{R}^3$, which happens to equal the set itself union the origin.

2.1. **Libor market model recursion.** Consider, as in Section 1.1, the one-factor, flat volatility case. By rescaling time, we may assume the flat volatility σ equals 1.

We find it more convenient to determine equivalently the support of log-rates, and denote (in reverse order) $X^i := \log(L^{n+1-i})$. Our assumption of flat volatility of 1 means that

$$[X^i, X^j]_t = t \quad \text{for all } t \geq 0,\ 1 \leq i, j \leq n.$$

The no-arbitrage property of the Libor market model can be expressed by the equations

$$dX^{i+1} = dX^i - f(X^i)dt,$$

where

$$f(x) := \frac{1}{1+e^{-x}}. \qquad (f^{-1}(y) = \log\frac{y}{1-y}).$$

This specification is consistent with the SDE system of Section 1.1. Note, $f' > 0$, $f(-\infty) = 0$, and $f(\infty) = 1$. Defining the processes $A^i := X^i - X^{i+1}$ for $1 \leq i \leq n-1$, we get

$$A_t^i := X_t^i - X_t^{i+1} = A_0^i + \int_0^t f(X_s^i)ds. \tag{2.1}$$

Therefore $\dot{A}^i = f(X^i)$ (Here, $\dot{a} := \frac{da}{dt}$.) As such, the processes A^i are C^1 and strictly increasing with slope less than 1. Using that $X^i = f^{-1}(\dot{A}^i)$, we find the following recursion

$$\dot{A}^{i+1} = f(X^{i+1}) = f(X^i - A^i) = f(f^{-1}(\dot{A}^i) - A^i) = \frac{\dot{A}^i}{\dot{A}^i + e^{A^i} - \dot{A}^i e^{A^i}}.$$

2.2. **Preliminaries on the support.** In the Libor market model X^1 is a Brownian motion in an equivalent measure and hence a ***prolific process*** as in [2].[2] This easily implies that

$$\mathcal{S}(X_t^1, A_t^1, \cdots, A_t^{n-1}) = \mathbb{R} \times \mathcal{S}(A_t^1, \cdots, A_t^{n-1}).$$

The support $\mathcal{S}(X_t)$ is thus practically determined once $\mathcal{S}(A^1, \cdots, A^{n-1})$ is. We have

$$\mathcal{S}(A_t^i) = [A_0^i, A_0^i + t]$$

for each i. This follows by using X^i is prolific to show $\mathcal{S}(A_t^i)$ is connected while noting $f \approx 0$ on low paths and $f \approx 1$ on high paths of X^i to locate the interval's end points.

The support $\mathcal{S}(A_t)$ contains the diagonal line segment of length t starting from A_0, i.e., the set $\{(A_0^1 + u, \cdots, A_0^{n-1} + u) : u \in [0, t]\}$. Indeed, consider a path with X^1 hence all X^i very high on some subset of $[0, t]$ of measure nearly u and very low on nearly all of the subset's complement. Then $\int_0^t f(X_s^i)ds \approx u$ on this path for all i, implying $A_t^i \approx A_0^i + u$.

In this connection we further note that if $A_0^i \geq 0$ then $A_t^{i+1} - A_0^{i+1} < A_t^i - A_0^i$ for all $t > 0$. Indeed, then $A_t^i > 0$, which, using that f is strictly increasing, implies that

$$A_t^{i+1} - A_0^{i+1} = \int_0^t f(X_s^{i+1})ds = \int_0^t f(X_s^i - A_s^i)ds < \int_0^t f(X_s^i)ds = A_t^i - A_0^i. \quad (A_0^i \geq 0)$$

[2]We call a process X prolific if $\mathbb{P}\{\sup_{t\in[0,T]} |X_t - c(t)| < \varepsilon\} > 0$ for all $T > 0$, $\varepsilon > 0$ and continuous functions $c : [0, T] \to \mathbb{R}$ with $c(0) = X_0$. Intuitively, a prolific process can follow any continuous path.

If on the other hand $A^i_0 < 0$, then for small t the reverse equality $A^{i+1}_t - A^{i+1}_0 > A^i_t - A^i_0$ holds. But, as t grows, A^i_t becomes positive, and for large t we again get the less-than inequality. This contrast illustrates why the case $A^i_0 \geq 0$ is simpler than the case $A^i_0 < 0$.

2.3. **Reduction of trivariate support to a calculus of variation problem.** We now specialize to the case $n = 3$. Using that X^1 is prolific, it is easy to show that for each $a^1 \in \mathcal{S}(A^1_t)$ the set $\{(a^1, a^2) \in \mathbb{R}^2 : a^2 \in \mathcal{S}(A^2_t)\}$ is connected. This clearly implies that

$$\mathcal{S}(A^1_t, A^2_t) = \{(a^1, a^2) \in \mathcal{S}(A^1_t) \times \mathbb{R} : \; A^2_*(t, a^1) \leq a^2 \leq A^{2^*}(t, a^1)\}, \tag{2.2}$$

where

$$A^2_*(t, a^1) := \operatorname{ess} \inf_{\{\omega \in \Omega : A^1_t(\omega) = a^1\}} A^2_t(\omega),$$

$$A^{2^*}(t, a^1) := \operatorname{ess} \sup_{\{\omega \in \Omega : A^1_t(\omega) = a^1\}} A^2_t(\omega).$$

In general $A^{2^*}(t, a^1) - A^2_0 \geq a^1 - A^1_0$ since as mentioned above the diagonal line segment is contained in the support. Equality holds if $A^1_0 \geq 0$ since then $A^2_t - A^2_0 < A^1_t - A^1_0$ by Section 2.2, implying $A^{2^*}(t, a^1) - A^2_0 \leq a^1 - A^1_0$. Therefore, we conclude that

$$A^{2^*}(t, a^1) - A^2_0 = a^1 - A^1_0 \quad \text{if} \quad A^1_0 \geq 0. \tag{2.3}$$

Assuming henceforth $A^1_0 \geq 0$ (i.e., $L^n_0 \geq L^{n-1}_0$), our task thus reduces to calculation of the probabilistically defined infimum function $A^2_*(t, a^1)$. To this end we note that the set of paths of A^1 is dense in the subspace of C^1 functions a_t satisfying $a_0 = A^1_0$ and $0 < \dot{a} < 1$, because X^1 is prolific and $\dot{A}^2 = f(X^1)$. The recursion of Section 2.1 thus implies that

$$A^2_*(T, a^1) - A^2_0 = \inf_{\{a \in \mathcal{C}^1[0,T] : \, a_0 = A^1_0, \, a_T = a^1, \, 0 < \dot{a} < 1\}} \int_0^T \frac{\dot{a}_t}{\dot{a}_t + e^{a_t} - \dot{a}_t e^{a_t}} dt,$$

where $\mathcal{C}^1[s, T]$ denotes the Banach space of C^1 functions $a = a_t$ on $[s, T]$, $0 \leq s < T$, with norm $|a_s| + \sup_{t \in [s,T]} |\dot{a}_t|$. Since the closure in $\mathcal{C}^1[0, T]$ of the set $\{0 < \dot{a} < 1\}$ is the set $\{0 \leq \dot{a} \leq 1\}$, and since the nonlinear functional that we are minimizing is continuous (in fact smooth), the infimum is the same with binding inequality $0 \leq \dot{a} \leq 1$:

$$A^2_*(T, a^1) - A^2_0 = \inf_{\{a \in \mathcal{C}^1[0,T] : \, a_0 = A^1_0, \, a_T = a^1, \, 0 \leq \dot{a} \leq 1\}} \int_0^T \frac{\dot{a}_t}{\dot{a}_t + e^{a_t} - \dot{a}_t e^{a_t}} dt. \tag{2.4}$$

*We **conjecture** that this infimum is attained at some C^1 function a_t in the feasible set.*

Actually, we find it more natural that instead of minimizing $A^2_T(\omega)$ subject to $A^1_T(\omega) = a^1$ to equivalently maximize $A^1_T(\omega) - A^2_T(\omega)$ subject to $A^1_T(\omega) = a^1$. In other words, instead of the function $A^2_*(T, a^1)$, we will equivalently determine the function

$$B^*(T, a^1) := a^1 - A^1_0 - (A^2_*(T, a^1) - A^2_0) = \sup_{\{a \in \mathcal{C}^1[0,T] : \, a_0 = A^1_0, \, a_T = a^1, \, 0 \leq \dot{a} \leq 1\}} \int_0^T L(a_t, \dot{a}_t) dt,$$

where the Lagrangian L is defined as

$$L(a, \dot{a}) := \dot{a} - \frac{\dot{a}}{\dot{a} + e^a - \dot{a}e^a} = \frac{\dot{a}(1-\dot{a})(e^a - 1)}{\dot{a} + e^a - \dot{a}e^a}. \tag{2.5}$$

2.4. **Solution of the unconstrained problem.** Let $0 \le s < T$. We first need to solve the optimization problem *without* the constraint $0 \le \dot{a} \le 1$ of locally maximizing $\int_s^T L(a_t, \dot{a}_t)dt$ over $a \in \mathcal{C}^1[s, T]$ subject to given a_s and a_T. As C^1 functions can be approximated by C^2 functions, if we find a local maximum among C^2 functions, it will be a local maximum among C^1 functions as well. The first order condition (vanishing of functional differential) for a C^2 local optimum $a = a_t$ is the **Euler-Lagrange equation**

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{a}} = \frac{\partial L}{\partial a}.$$

Since the Lagrangian L is time independent, the second order Euler-Lagrange ODE can be reduced to a first order ODE by the **Beltrami identity** $\frac{d}{dt}(L - \dot{a}\frac{\partial L}{\partial \dot{a}}) = 0$. So, along any local optimum a, the ***energy*** E is conserved, where

$$E := L - \dot{a}\frac{\partial L}{\partial \dot{a}} = \frac{\dot{a}^2(e^a - 1)}{(\dot{a} + e^a - \dot{a}e^a)^2} = \text{constant of motion.} \tag{2.6}$$

Assuming $E > 0$, or equivalently $a_s > 0$ and $\dot{a}_s \neq 0$, we rewrite the Beltrami identity as

$$\dot{a}\left(\sqrt{\frac{e^{-a} - e^{-2a}}{E}} + 1 - e^{-a}\right) = 1.$$

Multiplying both side by dt and integrating by a change of variable to da, we get

$$\int_{a_s}^{a_t} \left(\sqrt{\frac{e^{-a} - e^{-2a}}{E}} + 1 - e^{-a}\right)da = t - s.$$

But, the left hand side can be integrated explicitly, and we find that

$$t = s + a_t + e^{-a_t} - a_s - e^{-a_s} + \frac{1}{\sqrt{E}}\left(\arcsin(e^{-\frac{a_s}{2}}) + \sqrt{e^{-a_s} - e^{-2a_s}} - \arcsin(e^{-\frac{a_t}{2}}) - \sqrt{e^{-a_t} - e^{-2a_t}}\right).$$

This gives the inverse of the solution a given a_s and E (equivalently, given a_s and $\dot{a}_s$): the value a_t at time t is obtained by solving this equation. Replacing t by T, we can also view it as determining the energy E of the solution that takes a given a_s to a given a_T.

Much can be inferred. All solutions extend to time ∞ (they don't explode) and are convex for large t with slope approaching 1. If $0 < \dot{a}_s \le 1$ (equivalently $0 < E \le e^{a_s} - 1$), then the solution remains constrained thereafter, in fact $0 < \dot{a} < 1$ on (s, ∞). As such, the *highest* solution that also satisfies $\dot{a} \le 1$ on $[s, \infty)$ is obtained by choosing $\dot{a}_s = 1$, or equivalently, $E = e^{a_s} - 1$. Then a is strictly concave at s and has a unique inflection point $u > s$. Such a solution can be extended to below s when $s > 0$, but there it will satisfy $\dot{a} > 1$ due to the concavity at s. We can also infer that there exists a (unique) solution a with given a_s and a_T that satisfies $\dot{a} \le 1$ on $[s, T]$ if and only if $T - s \ge \tau^*(a_s, a_T)$, where

$$\tau^*(\alpha, a) := a - \alpha + e^{-a} + \frac{\arcsin(e^{-\frac{\alpha}{2}}) - \arcsin(e^{-\frac{a}{2}}) - \sqrt{e^{-a} - e^{-2a}}}{\sqrt{e^{\alpha} - 1}}. \tag{2.7}$$

2.5. **The optimal functional value.** A similar integration yields that at an extremum

$$\int_s^T L(a_t, \dot{a}_t)dt = a_T - a_s + e^{-a_T} - e^{-a_s} \tag{2.8}$$
$$- \frac{(\arcsin(e^{-\frac{a_s}{2}}) + \sqrt{e^{-a_s} - e^{-2a_s}} - \arcsin(e^{-\frac{a_T}{2}}) - \sqrt{e^{-a_T} - e^{-2a_T}})^2}{T - s - a_T - e^{-a_T} + a_s + e^{-a_s}}.$$

2.6. **The second-order conditions.** A solution of the Euler-Lagrange equation is a strict local maximum if for all non-identically zero C^1 functions b with $b_s = b_T = 0$ one has,

$$\int_s^T ((\frac{\partial^2 L}{\partial a^2} - \frac{d}{dt}\frac{\partial^2 L}{\partial \dot{a} \partial a})b_t^2 + \frac{\partial^2 L}{\partial \dot{a}^2}\dot{b}_t^2)dt < 0.$$

A sufficient condition for this is that $\frac{\partial^2 L}{\partial a^2} - \frac{d}{dt}\frac{\partial^2 L}{\partial \dot{a}\partial a} < 0$ and $\frac{\partial^2 L}{\partial \dot{a}^2} < 0$. Both conditions hold in our case. Indeed, the second is straightforward, while multiplying by $\dot{a} > 0$ and applying the chain and product rules the first inequality is equivalent to

$$0 > \dot{a}(\frac{\partial^2 L}{\partial a^2} - \frac{d}{dt}\frac{\partial^2 L}{\partial \dot{a}\partial a}) = \frac{d}{dt}(\frac{\partial L}{\partial a} - \dot{a}\frac{\partial^2 L}{\partial \dot{a}\partial a}),$$

In our case this holds implying the unconstrained solution is a local maximum. Indeed,

$$\frac{d}{dt}(\frac{\partial L}{\partial a} - \dot{a}\frac{\partial^2 L}{\partial \dot{a}\partial a}) = -\frac{E\dot{a}e^a}{(e^a - 1)^2}(1 + \sqrt{E}\sqrt{e^a - 1}) < 0. \tag{2.9}$$

2.7. **Existence and uniqueness of the constrained local maximum.** Assuming $A_0^1 \geq 0$, we can now show that there exists a unique local maximum to the optimization problem $\int_0^T L(a_t, \dot{a}_t)dt$, subject to $a_0 = A_0^1$, $a_T = a^1$, and $0 \leq \dot{a} \leq 1$ on $[0, T]$, and characterize it analytically. The uniqueness follows easily using that $\ddot{a}_s < 0$ for a solution $a = a_t$ to the Euler-Lagrange equation that satisfies $\dot{a}_s = 1$ at some s. If our conjecture that a global maximum exists is true, then it follows this local maximum is the global maximum.

2.8. **Analytic characterization of the constrained local maximum.** If $A_0^1 > 0$ and $T \geq \tau^*(A_0^1, a^1)$ then, by the discussion preceding Eq. (2.7), there exists a unique solution $a = a_t$ of the Euler-Lagrange equation with $a_0 = A_0^1$, $a_T = a^1$, and moreover this solution satisfies $0 < \dot{a}_t < 1$ for all $t > 0$ and is a local maximum. The local maximum value $\int_0^T L(a_t, \dot{a}_t)dt$ is then given by equation (2.8) with $s = 0$ and $a_s = A_0^1$.

Otherwise, if $A_0^1 = 0$ or $T < \tau^*(A_0^1, a^1)$, we define $s > 0$ to be the unique solution to

$$T - s = \tau^*(A_0^1 + s, a^1),$$

and define the asserted unique local optimum a by $a_t = A_0^1 + t$ on $[0, s]$ and we define on $[s, T]$ by the unique solution of the Euler-Lagrange equation with $a_s = A_0^1 + s$ and $a_T = a^1$. The optimal value $\int_0^T L(a_t, \dot{a}_t)dt$ is again given by equation (2.8), now with $a_s = A_0^1 + s$.

2.9. **Constrained local maximality.** The second-order conditions for local maximality were shown to hold on $(s, T]$ where the constraints are inactive, i.e., $\dot{a}_t < 1$. It remains to show the (first-order) **Karush-Kuhn-Tucker conditions**, i.e., that the **Lagrange multipliers** λ_t of the active constraints $\dot{a}_t = 1$ are *nonnegative* for all $t \in [0, s]$ (as we are maximizing with inequalities "$\dot{a} \leq 1$".) (The objective functional differential is a linear combination (here an integral) with coefficients λ_t of the constraint functional differentials.)

We show that the Lagrange multipliers λ_t of a constrained C^1 local optimum (for any Lagrangian $L(a, \dot{a})$) with constraints $\dot{a}_t = 1$ active for t in the interval $[0, s]$ is given by

$$\lambda_t = \int_t^s \left(\frac{\partial L}{\partial a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{a}}\right) dt. \qquad (0 \leq t \leq s)$$

In our case the Karush-Kuhn-Tucker conditions $\lambda_t \geq 0$ hold, since we can easily calculate

$$\lambda_t = e^{A_0^1}(e^s - e^t) \geq 0. \qquad (0 \leq t \leq s) \tag{2.10}$$

This concludes the sketch of the derivation, for it is now straightforward to assemble the findings of this section and arrive at Equation (1.4) (see Appendix 6.1 for details).

3. Reduction of trivariate support to calculus of variation

After a review of some well-known facts about the Libor market model, this section derives a recursion in the one-factor, constant volatilities case. Using that each log-forward rate is a Brownian motion under equivalent measure and hence "prolific", we then calculate the bivariate support and derive some properties of the multivariate support. Finally, we reduce the calculation of the trivariate support to a calculus and variation problem, generalizing the formulation (2.4) above to constant volatilities. We will return to the multivariate support in Sec. 5 after solving the trivariate flat-volatility case in Sec. 4.

3.1. **One-factor Libor Market Model.** We fix a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with trivial $\mathcal{F}_0$, and take as given an $(n+1)$-tuple $(B^1, \cdots, B^{n+1})$ of positive functions $B^i = B^i_t(\omega) > 0$ on $\Omega \times [0, \infty)$ which is ***arbitrage free*** in the sense that for every $1 \leq j \leq n+1$ and $T > 0$ there exists an equivalent measure $\mathbb{P}^j$ (depending certainly on j, possibly on T, and not necessarily unique) under which $\frac{B^i}{B^j}$ is a (right-continuous) martingale on $[0, T]$ for every $1 \leq i \leq n+1$. (As is well-known, a simple change of measure argument shows that if such a measure $\mathbb{P}^j$ exists for some j then it exists for all j). It follows (e.g., as in [2]) that $\frac{B^i}{B^j}$ is automatically a positive semimartingale with positive left limits for all i, j.

Thinking of B^i as the price process of the zero-coupon bond maturing in year i, the annual forward Libor rate L^i process - necessarily a semimartingale - is defined by

$$L^i := \frac{B^i}{B^{i+1}} - 1. \qquad (1 \leq i \leq n) \tag{3.1}$$

Evidently, L^i is a $\mathbb{P}^{i+1}$-martingale on $[0, T]$ for all i, as are the processes $\frac{B^{i-1}}{B^{i+1}} - \frac{B^i}{B^{i+1}} = L^{i-1}(1 + L^i) =: M$. By Itô's product rule, $dM = L^{i-1}_{-} dL^i + (1 + L^i_{-}) dL^{i-1} + d[L^{i-1}, L^i]$.

Dividing by $1+L^i_-$ (which is positive) and integrating, it follows that for $i \geq 2$ the process

$$L^{i-1} + \int_0^\cdot \frac{d[L^{i-1}, L^i]}{1+L^i_-} \text{ is a } \mathbb{P}^{i+1}\text{-local martingale on } [0,T]. \qquad (2 \leq i \leq n) \tag{3.2}$$

Henceforth, we assume a positive, one-factor model, meaning here positive initial values $L^i_0 > 0$ and that there exists *positive* optional processes $\sigma^1, \cdots, \sigma^n$ such that for all i, j,

$$d[L^i, L^j] = \sigma^i \sigma^j L^i L^j dt. \qquad (1 \leq i, j \leq n) \tag{3.3}$$

This implies that all L^i are continuous and positive, $\int_0^T {\sigma_t^i}^2 dt = [\int_0^\cdot \frac{dL^i}{L^i}]_T < \infty$ a.s. for all $T > 0$, and that the process $W^i := \int_0^\cdot \frac{dL^i}{\sigma^i L^i}$ is a $\mathbb{P}^{i+1}$-Brownian motion on $[0,T]$ for each i, because clearly $[W^i]_t = t$ and W^i is a $\mathbb{P}^{i+1}$-local martingale on $[0,T]$ as L^i is.

We have, $W^i = W^{i-1} + \int_0^\cdot \frac{\sigma^i L^i}{1+L^i_-} dt$. Indeed, (3.2) and (3.3) easily imply that the right hand side is a $\mathbb{P}^{i+1}$ local martingale on $[0,T]$. Hence, so is $N := W^{i-1} + \int_0^\cdot \frac{\sigma^i L^i}{1+L^i_-} dt - W^i$. But (3.3) gives $[N] = 0$. Thus $N = 0$ (everywhere as T is arbitrary), as desired. It follows

$$\frac{dL^{i-1}}{L^{i-1}} = -\frac{\sigma^{i-1}\sigma^i L^i}{1+L^i} dt + \frac{\sigma^{i-1}}{\sigma_i} \frac{dL^i}{L^i}. \qquad (2 \leq i \leq n) \tag{3.4}$$

Starting at $i = n$ and telescoping, we can also write this in the well-known SDE form

$$\frac{dL^j}{L^j} = -\sum_{i=j+1}^n \frac{\sigma^i \sigma^j L^i}{1+L^i} dt + \sigma^j dW^n. \qquad (W^n := \int_0^\cdot \frac{dL^n}{\sigma^n L^n}, \ \ 1 \leq j \leq n)$$

We find it more convenient to work with logarithm of Eq. (3.4) but in reverse order. Set

$$X^i := \log(L^{n+1-i}), \qquad \sigma_i := \sigma^{n+1-i}. \tag{3.5}$$

(So, $[X^i] = \int_0^\cdot \sigma_i^2 dt$). Since by Itô's formula $dX^{n+1-j} = \frac{dL^j}{L^j} - \frac{1}{2}{\sigma^j}^2 dt$, from (3.4) we conclude

$$\frac{dX^{i+1}}{\sigma_{i+1}} = (\frac{\sigma_i - \sigma_{i+1}}{2} - \sigma_i f(X^i))dt + \frac{dX^i}{\sigma_i}, \tag{3.6}$$

where, as in the previous section,

$$f(x) = \frac{1}{1+e^{-x}}. \qquad (x \in \mathbb{R})$$

(What follows, except (3.11), holds for any C^1 function f with $f' > 0$, $\inf f = 0$, $\sup f = 1$.)

3.2. **The Recursion.** *Henceforth, we assume that σ_i are (positive) constants.*[3]

Eq. (3.6) provides a recursion for X^i, but X^i have non-smooth paths, and the differential is stochastic. There is a similar (albeit more complex) recursion for the C^1 processes

$$A^i := \frac{X^i}{\sigma_i} - \frac{X^{i+1}}{\sigma_{i+1}}. \qquad (1 \leq i \leq n-1) \tag{3.7}$$

[3]The apparently more general case of constant volatility ratios $\frac{\sigma_i}{\sigma_j}$ is reduced to the constant volatility case (further satisfying $\sigma_1 = 1$) by the deterministic change of time $t \mapsto \int_0^t \sigma_1^2(s)ds$.

We can rewrite equation (3.6) in terms of A^i as

$$\dot{A}^i := \frac{dA^i}{dt} = \frac{\sigma_{i+1} - \sigma_i}{2} + \sigma_i f(X^i). \tag{3.8}$$

Solving for X^i gives,

$$X^i = f^{-1}(\frac{\dot{A}^i}{\sigma_i} + \frac{\sigma_i - \sigma_{i+1}}{2\sigma_i}).$$

Substituting this in (3.7) we get,

$$X^{i+1} = \sigma_{i+1}(\frac{X^i}{\sigma_i} - A^i) = \frac{\sigma_{i+1}}{\sigma_i} f^{-1}(\frac{\dot{A}^i}{\sigma_i} + \frac{\sigma_i - \sigma_{i+1}}{2\sigma_i}) - \sigma_{i+1}A^i.$$

Substituting back into (3.8) with i changed to $i+1$, we arrive at an interesting recursion:

$$\dot{A}^{i+1} = L_i(A^i, \dot{A}^i), \qquad (1 \le i \le n-2) \tag{3.9}$$

where

$$L_i(a, \dot{a}) := \frac{\sigma_{i+2} - \sigma_{i+1}}{2} + \sigma_{i+1} f(\frac{\sigma_{i+1}}{\sigma_i} f^{-1}(\frac{\dot{a}}{\sigma_i} + \frac{\sigma_i - \sigma_{i+1}}{2\sigma_i})) - \sigma_{i+1} a). \tag{3.10}$$

In our case with $f(x) = \frac{1}{1+e^{-x}}$, we easily calculate explicitly (see Appendix 6.2 for details)

$$L_i(a, \dot{a}) = \frac{\sigma_{i+2} - \sigma_{i+1}}{2} + \frac{\sigma_{i+1}(\sigma_i - \sigma_{i+1} + 2\dot{a})^{\frac{\sigma_{i+1}}{\sigma_i}}}{(\sigma_i - \sigma_{i+1} + 2\dot{a})^{\frac{\sigma_{i+1}}{\sigma_i}} + (\sigma_i + \sigma_{i+1} - 2\dot{a})^{\frac{\sigma_{i+1}}{\sigma_i}} e^{\sigma_{i+1} a}}. \tag{3.11}$$

3.3. **Prolific processes and the support.** As the volatilities are assumed positive constants, each L^i is a $\mathbb{P}^{i+1}$-geometric Brownian motion and so X^i is a $\mathbb{P}^{n-i+2}$-Brownian motion with drift, at least on bounded intervals $[0, T]$. This implies that X^i is ***prolific***, meaning

$$\mathbb{P}\{\sup_{t \in [0,T]} |X^i_t - c(t)| < \varepsilon\} > 0$$

for all $T > 0$, $\varepsilon > 0$, and continuous functions $c : [0, T] \to \mathbb{R}$ with $c(0) = X^i_0$. Indeed, a Girsanov's change of measure reduces this to the case $c = 0$ (and without drift), where the probability above admits a well-known closed-form expression as an infinite sum of positive integrals. (See, e.g., [2], Proposition 2.2.) Intuitively, a prolific process can follow any conceivable continuous path, at least approximately. More precisely, its set of paths is dense in the space of continuous functions (anchored at time 0) with the sup norm.

The support $\mathcal{S}(L^1_t, \cdots, L^n_t)$ is clearly determined once $\mathcal{S}(X^1_t, A^1, \cdots, A^{n-1})$ is.[4] But,

$$\mathcal{S}(X^1_t, A^1_t, \cdots, A^i_t) = \mathbb{R} \times \mathcal{S}(A^1_t, \cdots, A^i_t). \quad (t > 0) \tag{3.12}$$

[4]Recall, the (topological) ***support*** $\mathcal{S}(R)$ of an n-dimensional random variable R is defined by

$$\mathcal{S}(R) := \{r \in \mathbb{R}^n : \mathbb{P}\{R \in U\} > 0 \text{ for all open subsets } U \text{ of } \mathbb{R}^n \text{ containing } r\}.$$

Equivalently, $\mathcal{S}(R)$ is the smallest *closed* subset of $\mathbb{R}^n$ such that $R \in \mathcal{S}(R)$ a.s. The reduction of $\mathcal{S}(L)$ to $\mathcal{S}(X^1, A)$ uses the property that if $g : \mathbb{R}^n \to \mathbb{R}^m$ is continuous then $\mathcal{S}(g(R))$ is the closure of $g(\mathcal{S}(R))$.

To see this, let $x^1 \in \mathbb{R}$. By (3.8), A_t^i is an affine function of $\int_0^t f(X_s^i)ds$. As X^1 is prolific, we can change X_s^1 along each path for s very near t so that $X_t^1 = x^1$, yet the integrals $\int_0^t f(X_s^i)ds$ and hence A_t^i change very little. More precisely, given $\omega \in \Omega$ and $\varepsilon > 0$, we can find $\omega' \in \Omega$ and a small $\delta > 0$ such that $X_s^1(\omega) = X_s^1(\omega')$ for $s \leq t - \delta$, $X_t^1(\omega') = x^1$, and $|A^j(\omega') - A^j(\omega)| < \varepsilon$ for all j. This readily implies the "$\supset$" (the "$\subset$" is obvious).

It follows from (3.12) that the multivariate support is determined once $\mathcal{S}(A_t^1, \cdots, A_t^i)$ is. In particular, the bivariate support $\mathcal{S}(X_t^1, X_t^2)$ is determined from $\mathcal{S}(A_t^1)$. We claim

$$\mathcal{S}(A_t^i) = [A_0^i + \frac{\sigma_{i+1} - \sigma_i}{2}t\,,\; A_0^i + \frac{\sigma_{i+1} + \sigma_i}{2}t]. \tag{3.13}$$

Indeed, integrating Eq. (3.8) gives

$$A_t^i = A_0^i + \frac{\sigma_{i+1} - \sigma_i}{2}t + \sigma_i \int_0^t f(X_s^i)ds. \tag{3.14}$$

Since $0 < f < 1$, we infer that $\frac{\sigma_{i+1}-\sigma_i}{2}t \leq A_t^i - A_0^i \leq \frac{\sigma_{i+1}+\sigma_i}{2}t$ and hence "$\subset$" holds. For the converse, we note that the integral $\int_0^t f(X_s^i)ds$ can be made very near 0 or very near t by choosing very low (near $-\infty$) or very high (near ∞) paths of X^i. By (3.14), this shows that "$\supset$" follows once we show $\mathcal{S}(A_t^i)$ is connected. And this is so because for any two states $\omega_0, \omega_1 \in \Omega$, we can clearly join any $A_t^i(\omega_0)$ to any $A_t^i(\omega_1)$ by the continuous curve $p \mapsto A_t^i(\omega_p)$, where, using that X^i prolific, for $0 < p < 1$ the state ω_p is so chosen and uniquely characterized as to satisfy $X_s^i(\omega_p) = f^{-1}((1-p)f(X_s^i(\omega_0)) + pf(X_s^i(\omega_1)))$ for $s \leq t$.

Another way to see the "$\supset$" is that given any element $a^i = A_0^i + \frac{\sigma_{i+1}-\sigma_i}{2}t + \sigma_i s$ in the right-hand side of (3.13), Eq. (3.14) shows that we can always construct a state $\omega \in \Omega$ such that $A_t^i(\omega) \approx a^i$, e.g., by choosing ω such that $X^i(\omega)$ is very high on (say) $[0, s]$ and very low on most of $(s, t]$, which easily implies the "$\supset$". Moreover, if $X^i(\omega)$ is very high (low) for some i it is very high (low) for all i. This argument actually shows more strongly that $\mathcal{S}(A_t^1, \cdots, A_t^i)$ contains the line segment $\{(A_0^j + \frac{\sigma_{j+1}-\sigma_j}{2}t + \sigma_j u)_{j=1}^i \in \mathbb{R}^i :\ u \in [0, t]\}$.

3.4. **The trivariate support.** By Eq. (3.12), the trivariate support is determined once $\mathcal{S}(A_t^1, A_t^2)$ is. The latter is clearly given by Eq. (2.2) once we show that for any $a^1 \in \mathcal{S}(A_t^1)$ the set $\{a^2 \in \mathbb{R} : (a^1, a^2) \in \mathcal{S}(A_t^1, A_t^2)\}$ is connected.

To this end, let $a^1 \in \mathcal{S}(A_t^1)$ and ω_0 and ω_1 be two states such that $A_t^1(\omega_0) = A_t^1(\omega_1) = a^1$. To show pathwise connectedness, we must find a continuous curve contained in $\mathcal{S}(A_t^1, A_t^2)$ that joins $(a^1, A_t^2(\omega_0))$ to $(a^1, A_t^2(\omega_1))$. For $0 \leq p \leq 1$ define the continuous function $c_p(s) = f^{-1}((1-p)f(X_s^1(\omega_0)) + pf(X_s^1(\omega_1))$ for $s \leq t$. Note $c_0(s) = X_s^1(\omega_0)$ and $c_1(s) = X_s^1(\omega_1)$ for all s. Since X^1 is prolific, for each $0 < p < 1$ there is (loosely speaking) a state $\omega_p \in \Omega$ such that $X_s^1(\omega_p) = c_p(s)$ for all $s \leq t$. It remains to show that $A_t^1(\omega_p) = a^1$, for then the continuous curve, $p \mapsto (a^1, A_t^2(\omega_p))$ is clearly contained in $\mathcal{S}(A_t^1, A_t^2)$ and connects $(a^1, A_t^2(\omega_0))$ to $(a^1, A_t^2(\omega_1))$. But this follows from the linearity of the integral: using (3.14),

$$A_t^1(\omega_p) - A_0^1 - \frac{\sigma_{i+1} - \sigma_i}{2}t = \sigma_1 \int_0^t f(X_s^1(\omega_p))ds$$

$$= \sigma_1 \int_0^t f(c_p(s))ds = (1-p)\sigma_1 \int_0^t f(X_s^1(\omega_0))ds + p\sigma_1 \int_0^t f(X_s^1(\omega_1))ds$$
$$= (1-p)(A_t^1(\omega_0) - A_0^1 - \frac{\sigma_{i+1} - \sigma_i}{2} t) + p(A_t^1(\omega_1) - A_0^1 \frac{\sigma_{i+1} - \sigma_i}{2} t)$$
$$= a^1 - A_0^1 - \frac{\sigma_{i+1} - \sigma_i}{2} t.$$

Hence, $A_t^1(\omega_p) = a^1$, as desired. This establishes Eq. (2.2). The problem is thus reduced to calculating the infimum function $A_*^2(T, a^1) := \operatorname{ess\,inf}_{\{\omega \in \Omega: A_t^1(\omega) = a^1\}} A_t^2(\omega)$, and the similarly supremum function $A^{2^*}(T, a^1)$ entering Eq. (2.2). But, Eq. (3.9) with $i = 1$ yields

$$A_*^2(T, a^1) - A_0^2 = \inf_{\{a \in \mathcal{C}^1[0,T]:\ a_0 = A_0^1,\, a_T = a^1,\ \frac{\sigma_3 - \sigma_2}{2} \le \dot{a} \le \frac{\sigma_3 + \sigma_2}{2}\}} \int_0^T L_1(a_t, \dot{a}_t)dt, \tag{3.15}$$

where the Lagrangian L_1 is defined by (3.10) or more explicitly (3.11) with $i = 1$. The problem thus reduces to this calculus of variation minimization plus a similar maximization in regards to the supremum function $A^{2^*}(T, a^1)$. Due to the complexity of L_1, we cannot solve this problem in general. (It is straightforward to write down the Beltrami identity explicitly, but the resulting ODE appears too complicated to admit an analytic solution.)

But, when $\sigma_1 = \sigma_2$ the Lagrangian (3.11) simplifies substantially to

$$L_1(a, \dot{a}) := \frac{\sigma_3 - \sigma_2}{2} + \frac{\sigma_2 \dot{a}}{\dot{a} + e^{\sigma_2 a} - \dot{a} e^{\sigma_2 a}}. \qquad (\sigma_1 = \sigma_2)$$

The constant term $\frac{\sigma_3 - \sigma_2}{2}$ in this Largrangian is innocuous, as it does not effect the minimization. Moreover, by a change of time unit we may assume $\sigma_2 = 1$. So, the minimization is equivalent to that with the simpler Lagrangian $\frac{\dot{a}}{\dot{a} + e^a - \dot{a}e^a}$. As mentioned in Section 2.3, this is clearly equivalent to maximization with respect to the Lagrangian

$$L(a, \dot{a}) := \dot{a} - \frac{\dot{a}}{\dot{a} + e^a - \dot{a}e^a},$$

which is the subject of the next section. We finally recall from Section 2.3 that the problem is further simplified by assuming $A_0^1 \ge 0$, for then $A^{2^*}(T, a^1) - A_0^2 = a^1 - A_0^1$ by Eq. (2.3).

4. Solution of the calculus of variation problem

As discussed above, only the trivariate constant-volatility case outlined is Section 2 is analytically tractable. This section is devoted to this case. Specifically, given $T > 0$, $A_0^1 \ge 0$, and $A_0^1 \le a^1 \le A_0^1 + T$, the problem is to maximize

$$\sup_{\{a \in \mathcal{C}^1[0,T]:\ a_0 = A_0^1,\, a_T = a^1,\, 0 \le \dot{a} \le 1\}} \int_0^T L(a_t, \dot{a}_t)dt, \tag{4.1}$$

where

$$L(a, \dot{a}) := \frac{\dot{a}(1 - \dot{a})(e^a - 1)}{\dot{a} + e^a - \dot{a}e^a}. \tag{4.2}$$

(The Lagrangian $L(a, \dot{a})$ is singular where $\dot{a}(1 - e^{-a}) = 1$, but there $\dot{a} > 1$ if $a > 0$ while $\dot{a} < 0$ if $a < 0$ (and a never becomes zero); so, L is well-defined on the strip $0 \le \dot{a} \le 1$.)

Note, $L = 0$ if either $a = 0$ or $\dot{a} = 0$ or $\dot{a} = 1$. We are primarily interested in the first quadrant $a, \dot{a} > 0$. There, $0 < L < \dot{a}$ if $0 < \dot{a} < 1$ (but $L < 0$ if $1 < \dot{a} < \frac{1}{1-e^{-a}}$).

4.1. **Partial derivatives of the Lagrangian.** (See Appendix (6.3) for the details.)

$$\frac{\partial L}{\partial \dot{a}} = 1 - \frac{e^a}{(\dot{a} + e^a - \dot{a}e^a)^2};$$

$$\frac{\partial L}{\partial a} = \frac{e^a \dot{a}(1-\dot{a})}{(\dot{a} + e^a - \dot{a}e^a)^2};$$

$$\frac{\partial^2 L}{\partial \dot{a}^2} = -\frac{2e^a(e^a - 1)}{(\dot{a} + e^a - \dot{a}e^a)^3};$$

$$\frac{\partial^2 L}{\partial a \partial \dot{a}} = \frac{e^a(e^a - \dot{a} - \dot{a}e^a)}{(\dot{a} + e^a - \dot{a}e^a)^3};$$

$$\frac{\partial^2 L}{\partial a^2} = \frac{\dot{a}(1-\dot{a})e^a(\dot{a} - e^a + \dot{a}e^a)}{(\dot{a} + e^a - \dot{a}e^a)^3}.$$

4.2. **The Euler-Lagrange equation.** Any local optimum $a = a_t$ of the inequality $0 \leq \dot{a} \leq 1$ constrained problem must satisfy the first-order conditions on the open subset of $[0, T]$ where neither inequality is active. As such, it is necessary to study the unconstrained problem more generally on intervals $[s, T]$, for $0 \leq s < T$, namely, to solve the problem of maximizing $\int_s^T L(a_t, \dot{a}_t)dt$ over $a \in \mathcal{C}^1[s, T]$ subject to given a_s and a_T.

The first order conditions are that the differential of the objective functional vanishes at any local optimum a, that is for all C^1 curves (variations) $b = b_t$ with $b_s = b_T = 0$ one has

$$\frac{d}{d\varepsilon}|_{\varepsilon=0} \int_s^T L(a_t + \varepsilon b_t, \dot{a}_t + \varepsilon \dot{b}_t)dt = 0.$$

Following the long-established procedure, we differentiate under the integral sign and then integrate by parts, while noting that boundary terms vanish as $b_0 = b_T = 0$, to get

$$0 = \int_0^T (\frac{\partial L}{\partial a} b_t + \frac{\partial L}{\partial \dot{a}} \dot{b}_t)dt = \int_0^T (\frac{\partial L}{\partial a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{a}})b_t dt.$$

As b_t was arbitrary, we arrive at the Euler-Lagrange equation of calculus of variation:

$$\frac{\partial L}{\partial a} = \frac{d}{dt}\frac{\partial L}{\partial \dot{a}} = \frac{\partial^2 L}{\partial a \partial \dot{a}}\dot{a} + \frac{\partial^2 L}{\partial \dot{a}^2}\ddot{a}. \tag{4.3}$$

If we do not constraint the solution value a_T at time T, then b_T is also free to vary, and accordingly the boundary term $\frac{\partial L}{\partial \dot{a}}(a_T, \dot{a}_T)b_T$ contributes as well. Since b_T is arbitrary, this implies that the transversality condition $\frac{\partial L}{\partial \dot{a}}(a_T, \dot{a}_T) = 0$ must hold too. By the formula for $\frac{\partial L}{\partial \dot{a}}$, this is equivalent to $\dot{a}_T + e^{a_T} - \dot{a}_T e^{a_T} = e^{\frac{a_T}{2}}$, which simplifies to $\dot{a}_T(1 + e^{-\frac{a_T}{2}}) = 1$.

In our case, the Euler-Lagrange equation is (by Section (4.1))

$$\frac{\dot{a}(1-\dot{a})e^a}{(\dot{a} + e^a - \dot{a}e^a)^2} + \frac{d}{dt}\frac{e^a}{(\dot{a} + e^a - \dot{a}e^a)^2} = 0.$$

In particular, if a is constant, i.e., $\dot{a} = 0$, then a is clearly a solution. We will soon see that any nonconstant solution is either strictly increasing or strictly decreasing.

4.3. **Beltrami identity.** Taking the total differential of L, and then applying the Euler-Lagrange equation followed by the product and chain rules,

$$\frac{dL}{dt} = \frac{\partial L}{\partial a}\dot{a} + \frac{\partial L}{\partial \dot{a}}\ddot{a} = \dot{a}\frac{d}{dt}\frac{\partial L}{\partial \dot{a}} + \frac{\partial L}{\partial \dot{a}}\ddot{a} = \frac{d}{dt}(\dot{a}\frac{\partial L}{\partial \dot{a}}).$$

It follows that $\frac{d}{dt}(L - \dot{a}\frac{\partial L}{\partial \dot{a}}) = 0$, or equivalently, the energy E is a constant, where

$$E := L - \dot{a}\frac{\partial L}{\partial \dot{a}} = \text{constant of motion} \tag{4.4}$$

Conversely, the Beltrami identity (4.4) implies the Euler-Lagrange equation, at least on the open set $\dot{a}_t \neq 0$. In our case,[5]

$$E = \frac{\dot{a}^2(e^a - 1)}{(\dot{a} + e^a - \dot{a}e^a)^2}. \tag{4.5}$$

Note that $E > 0$ if and only if $a > 0$ and $\dot{a} \neq 0$. If a is a constant solution, then $\dot{a} = 0$, so $E = 0$. Conversely, any solution a with zero energy E is constant. Indeed, if $E = 0$ then $\dot{a}(e^a - 1) = 0$ everywhere. So on the open set $\{t : \dot{a}_t \neq 0\}$ we must have $e^a - 1 = 0$, i.e., $a = 0$, implying that on this set too $\dot{a} = 0$, so it must be empty. Therefore if $a_u = 0$ or $\dot{a}_u = 0$ for any $u \in [s, T]$, then $E = 0$ so $\dot{a} = 0$. (Interestingly, there is a continuous function a on $[s, \infty)$ with $a_s = 0$, which is a solution on (s, ∞) but satisfies $\dot{a}_s = \infty$.)

Clearly $E \geq 0$ if and only if $a \geq 0$, and $E < 0$ if and only if $a < 0$ and $\dot{a} \neq 0$. We do not pursue negative energy solutions here, but it seems they can be treated similarly.

We saw above that a nonconstant solution is either strictly increasing or strictly decreasing. We assume the former; specifically we assume $\dot{a} > 0$ and $a > 0$ (implying $E > 0$).

Equation (4.5) obviously implies that $\dot{a}_u = 1$ if and only if $E = e^{a_u} - 1$, i.e., $u = \log(1 + E)$. Likewise, $\dot{a} < 1$ if and only if $E < e^a - 1$. In particular, if $\dot{a}_s \leq 1$ then $\dot{a}_t < 1$ for $t > s$ (since $E \leq e^{a_s} - 1 < e^{a_t} - 1$). Significantly, this means that *a solution that is initially constrained will remain constrained for all time.* It also follows that if $t > \log(1 + E)$ then $\dot{a}_t < 1$ (since $E < e^{a_t} - 1$), and if $t < \log(1 + E)$ then $\dot{a}_t > 1$. So, at the point $u = \log(1 + E)$ where $\dot{a}_u = 1$, the solution a is strictly concave, i.e., $\ddot{a}_u < 0$.

Since E is a constant, the sign of $\dot{a} + e^a - \dot{a}e^a$ is also a constant, for if it changes, it must become zero at some time, implying infinite energy E. We are interested in and assume the positive-sign case, i.e., $\dot{a} + e^a - \dot{a}e^a > 0$, or equivalently, $\dot{a}(1 - e^{-a}) < 1$, for this condition necessarily holds when $\dot{a} \leq 1$. Note that it then follows from Section (4.1) that $\frac{\partial^2 L}{\partial \dot{a}^2} < 0$.

Solving for $\dot{a} > 0$ in terms of E, we can rewrite the Beltrami identity as

$$\dot{a}(\sqrt{\frac{e^{-a} - e^{-2a}}{E}} + 1 - e^{-a}) = 1. \tag{4.6}$$

[5]In more detail, $E = \frac{\dot{a}e^a}{(\dot{a}+e^a-\dot{a}e^a)^2} - \frac{\dot{a}}{\dot{a}+e^a-\dot{a}e^a} = \frac{\dot{a}^2(e^a-1)}{(\dot{a}+e^a-\dot{a}e^a)^2}$.

We saw the transversality condition for (the widest) a_T is $\frac{1}{\dot a_T} = 1 + e^{-\frac{a_T}{2}}$. In view of (4.6), this is easily equivalent to $e^{a_T} - 1 = E(e^{\frac{a_T}{2}} + 1)^2$, a quadratic equation for $e^{\frac{a_T}{2}}$.

We can calculate $\ddot a$ directly from (4.6). One finds that (see Appendix 6.4 for details)

$$\ddot a_t = 0 \text{ if and only if } \dot a_t = 1 - \frac{1}{e^{a_t}-1} \text{ if and only if } E = \frac{(\frac{1}{2}e^{a_t}-1)^2}{e^{a_t}-1};$$

$$\ddot a_t < 0 \text{ if and only if } \dot a_t > 1 - \frac{1}{e^{a_t}-1} \text{ if and only if } \sqrt{E} > \frac{\frac{1}{2}e^{a_t}-1}{\sqrt{e^{a_t}-1}}.$$

In particular it follows that a solution a is convex for large t and has a unique inflection point, at which point the slope is less that 1. This is consistent with before: the slope of a solution a can be greater than or equal to 1 only at points where a is strictly concave.

4.4. **The unconstrained solution.** Multiplying both sides of Eq. (4.6) by dt and integrating from s to t while changing variables by substituting $\dot a dt = da$, gives

$$\int_{a_s}^{a_t} (\sqrt{\frac{e^{-a}-e^{-2a}}{E}} + 1 - e^{-a})da = t - s.$$

But, it is easy to verify that (see Appendix 6.5 for details)

$$\sqrt{e^{-a}-e^{-2a}}da = -d(\arcsin(e^{-\frac{a}{2}}) + \sqrt{e^{-a}-e^{-2a}}).$$

Substituting into the integral, integrating the exact differential, and rearranging yields

$$t = s + a_t - a_s + e^{-a_t} - e^{-a_s} \tag{4.7}$$
$$+\frac{1}{\sqrt{E}}(\arcsin(e^{-\frac{a_s}{2}}) + \sqrt{e^{-a_s}-e^{-2a_s}} - \arcsin(e^{-\frac{a_t}{2}}) - \sqrt{e^{-a_t}-e^{-2a_t}}).$$

For fixed s, a_s, and E, consider the right-hand-side as a function $h = h(a_t)$ of a_t. Then the equation states $t = h(a_t)$, that is $a_t = h^{-1}(t)$. Both a and h are strictly increasing functions. For large a we have $h(a) \approx a+$ constant. Indeed, all (nonconstant) terms in the expression above for $h(a_t)$ contain the term e^{-a_t} and go to zero for large a_t, except for the first term, which is a_t itself. In particular, solutions do not explode and $a_\infty = \infty$.

With t replaced by T, the equation (4.7) yields explicitly the unique energy E, namely

$$\sqrt{E} = \frac{\arcsin(e^{-\frac{a_s}{2}}) + \sqrt{e^{-a_s}-e^{-2a_s}} - \arcsin(e^{-\frac{a_T}{2}}) - \sqrt{e^{-a_T}-e^{-2a_T}}}{T - s - a_T - e^{-a_T} + a_s + e^{-a_s}}, \tag{4.8}$$

of the solution $a = a_t$ that joins a given a_s at time s to a given a_T at time T, establishing both existence and uniqueness. But, the solution may violate the constraint $\dot a \leq 1$. Starting from a_s at time s, the highest a_T that can be reached at time T with a constrained solution is attained by the solution a^* with the highest allowable slope at time s, namely with $\dot a^*_s = 1$, or equivalently with energy $E = e^{a_s} - 1$. (As noted previously, then $0 < \dot a_t < 1$ for $t > s$.) This highest solution a^* is obtained by substituting $E = e^{a_s} - 1$ in equation (4.7). After a cancellation due to the equality $e^{-a_s}\sqrt{E} = \sqrt{e^{-a_s}-e^{-2a_s}}$, we get $t = s + \tau^*(a_s, a^*_t)$, where

$$\tau^*(\alpha, a) := a - \alpha + e^{-a} + \frac{\arcsin(e^{-\frac{\alpha}{2}}) - \arcsin(e^{-\frac{a}{2}}) - \sqrt{e^{-a} - e^{-2a}}}{\sqrt{e^{\alpha} - 1}}. \tag{4.9}$$

A given a_T is reached by the highest solution a^* at the time $s + \tau^*(a_s, a_T)$, that is, $a_T = a^*_{s+\tau^*(a_s,a_T)}$. If $T < s + \tau^*(a_s, a_T)$ then, since a^* is increasing, $a_T > a^*_T$. So in this case a_T is too high, implying that the solution a that joins a_s to a_T fails the constraint, namely $\dot{a}_s > 1$. Contrariwise, if $T \geq s + \tau^*(a_s, a_T)$, then $a_T \leq a^*_T$, implying that the solution that joins a_s to a_T lies below a^*; hence $\dot{a}_s \leq a^*_s = 1$, implying $\dot{a} < 1$ on $(s, T]$.

In short, a constrained solution a (i.e., satisfying $\dot{a} \leq 1$) on $[s, T]$ with given (positive) values a_s and a_T exists (and is unique as given above) if and only if $T \geq s + \tau^*(a_s, a_T)$.

4.5. **The optimal value.** Using $L(a, \dot{a}) = \dot{a} - \frac{\dot{a}}{\dot{a}+e^a-\dot{a}e^a}$ and the definition (4.5) of energy,

$$L(a, \dot{a}) = \dot{a} - \frac{\sqrt{E}}{\sqrt{e^a - 1}}.$$

Integrating with respect to time t, then changing variable to a by using (4.6) to substitute $dt = (\frac{\sqrt{e^{-a}-e^{-2a}}}{\sqrt{E}} + 1 - e^{-a})da$, followed by simplification and then explicit integration yields,

$$\int_s^T L(a_t, \dot{a}_t)dt = a_T - a_s - \int_s^T \frac{\sqrt{E}dt}{\sqrt{e^{a_t} - 1}}$$

$$= a_T - a_s - \int_{a_s}^{a_T} \frac{\sqrt{e^{-a} - e^{-2a}} + \sqrt{E}(1 - e^{-a})}{\sqrt{e^a - 1}} da$$

$$= a_T - a_s - \int_{a_s}^{a_T} (e^{-a} + \sqrt{E}\sqrt{e^{-a} - e^{-2a}})da$$

$$= a_T - a_s + e^{-a_T} - e^{-a_s} - \sqrt{E}(\arcsin(e^{-\frac{a_s}{2}}) + \sqrt{e^{-a_s} - e^{-2a_s}} - \arcsin(e^{-\frac{a_T}{2}}) - \sqrt{e^{-a_T} - e^{-2a_T}}).$$

(Above, we used the same formula as the previous section for the integral $\int \sqrt{e^{-a} - e^{-2a}}da$.)

Substituting for $\sqrt{E}$ from (4.8) yields the critical value at an extremum in closed form:

$$\int_s^T L(a_t, \dot{a}_t)dt = a_T - a_s + e^{-a_T} - e^{-a_s} \tag{4.10}$$

$$- \frac{(\arcsin(e^{-\frac{a_s}{2}}) + \sqrt{e^{-a_s} - e^{-2a_s}} - \arcsin(e^{-\frac{a_T}{2}}) - \sqrt{e^{-a_T} - e^{-2a_T}})^2}{T - s - a_T - e^{-a_T} + a_s + e^{-a_s}}.$$

4.6. **Second order conditions.** A sufficient condition for a solution $a = a_t$ to be a strict local maximum is that for all non-identically zero C^1 functions b with $b_s = b_T = 0$ one has

$$0 > \frac{d^2}{d\varepsilon^2}|_{\varepsilon=0} \int_s^T L(a_t + \varepsilon b_t, \dot{a}_t + \varepsilon \dot{b}_t) dt$$

$$= \frac{d}{d\varepsilon}|_{\varepsilon=0} \int_s^T (\frac{\partial L}{\partial a}(a_t + \varepsilon b_t, \dot{a}_t + \varepsilon \dot{b}_t) b_t + \frac{\partial L}{\partial \dot{a}}(a_t + \varepsilon b_t, \dot{a}_t + \varepsilon \dot{b}_t) \dot{b}_t) dt$$

$$= \int_s^T (\frac{\partial^2 L}{\partial a^2} b_t^2 + 2\frac{\partial^2 L}{\partial \dot{a} \partial a} b_t \dot{b}_t + \frac{\partial^2 L}{\partial \dot{a}^2} \dot{b}_t^2) dt.$$

This would hold if L had a negative definite Hessian, but our case fails this test.[6] Instead, substitute $2b\dot{b} = \dot{b^2}$ in the middle term and integrate by parts to rewrite the condition as

$$\int_s^T ((\frac{\partial^2 L}{\partial a^2} - \frac{d}{dt}\frac{\partial^2 L}{\partial \dot{a} \partial a}) b_t^2 + \frac{\partial^2 L}{\partial \dot{a}^2} \dot{b}_t^2) dt < 0. \tag{4.11}$$

A sufficient condition for this is that $\frac{\partial^2 L}{\partial a^2} - \frac{d}{dt}\frac{\partial^2 L}{\partial \dot{a} \partial a} < 0$ and $\frac{\partial^2 L}{\partial \dot{a}^2} < 0$. Both conditions hold in our case. Indeed, the second was pointed out in Section 4.3. As for the first condition, multiplying both of its sides by $\dot{a} > 0$, it is equivalent to

$$0 > \dot{a}(\frac{\partial^2 L}{\partial a^2} - \frac{d}{dt}\frac{\partial^2 L}{\partial \dot{a} \partial a}) = \frac{d}{dt}(\frac{\partial L}{\partial a} - \dot{a}\frac{\partial^2 L}{\partial \dot{a} \partial a}), \tag{4.12}$$

where the equality follows by applying the chain and product rules similarly as in the derivation of Beltrami identity. In our case, a simple calculation shows (see Appendix 6.6),

$$\frac{\partial L}{\partial a} - \dot{a}\frac{\partial^2 L}{\partial \dot{a} \partial a} = E(\frac{1}{e^a - 1} + \frac{2\sqrt{E}}{\sqrt{e^a - 1}} - 1),$$

which is clearly decreasing as a is increasing. Specifically,

$$\frac{d}{dt}(\frac{\partial L}{\partial a} - \dot{a}\frac{\partial^2 L}{\partial \dot{a} \partial a}) = -\frac{E\dot{a}e^a}{(e^a - 1)^2}(1 + \sqrt{E}\sqrt{e^a - 1}) < 0, \tag{4.13}$$

with which we conclude that the unique extremum found in Sec. 4.4 is a local maximum.

4.7. **The unique constrained optimum.** We now return to the original problem of maximizing $\int_0^T L(a_t, \dot{a}_t) dt$, subject to $a_0 = A_0^1$, $a_T = a^1$, and the constraint $0 \le \dot{a} \le 1$ on $[0, T]$. We require $A_0^1 \le a^1 \le A_0^1 + T$, i.e., $a^1 \in \mathcal{S}(A_T^1)$, for otherwise there is no constrained function joining A_0^1 and a^1. The boundary cases $a^1 = A_0^1$ and $a^1 = A_0^1 + T$ are trivial: there is a single constrained function $a = a_t$ such that $a_0 = A_0^1$ and $a_T = A_0^1$ (resp. $a_T = A_0^1 + T$), namely $a_t = A_0^1$ (resp. $a_t = A_0^1 + t$). So, we assume henceforth that $A_0^1 < a^1 < A_0^1 + T$.

If $A_0^1 > 0$ and $T \ge \tau^*(A_0^1, a^1)$, then it follows from Section 4.4 with $s = 0$ that the unique solution a found there that joins A_0^1 at $t = 0$ to a^1 at $t = T$ is already constrained,

[6] When the Hessian of $L(a, \dot{a})$ is negative definite, or equivalently L is strictly concave, then the functional $\mathcal{L}(a) = \int_s^T L(a_t, \dot{a}_t) dt$ is strictly concave and any extremum is the unique global maximum. The condition is equivalent to $\frac{\partial^2 L}{\partial a^2} + \frac{\partial^2 L}{\partial \dot{a}^2} < 0$ and $\frac{\partial^2 L}{\partial a^2}\frac{\partial^2 L}{\partial \dot{a}^2} > (\frac{\partial^2 L}{\partial \dot{a} \partial a})^2$ everywhere, both of which fail (slightly) in our case.

namely satisfies $0 < \dot{a} \leq 1$ on $[0, T]$. There is nothing more to do in this case: the optimal functional value is given by equation (4.10) with $s = 0$, and the problem is solved.

But if $A_0^1 = 0$, then any solution a to the Euler-Lagrange equation on $[0, T]$ with $a_0 = A_0^1$ is identically zero. Likewise, as saw, if $A_0^1 > 0$ but $T < \tau^*(A_0^1, a^1)$, then the solution on $[0, T]$ that joints A_0^1 and a^1 satisfies $\dot{a}_0 > 1$, i.e., violates the constraint. In these cases, a^1 is too large to be reached by a constrained solution on $[0, T]$. The highest a_*^1 that can be reached at time T by a constrained solution on $[0, T]$ is that attained by the solution a with highest allowable slope at time 0, namely with $\dot{a}_0 = 1$; that is, a_*^1 is the solution to the equation $T = \tau^*(A_0^1, a_*^1)$. It is intuitive that if $a^1 > a_*^1$ then the best that can be done is to commence at A_0^1 with the slope of 1 and stay along the line $A_0^1 + t$, for as long time s as it takes so that the solution on $[s, T]$ with $a_s = A_0^1 + s$ and $\dot{a}_s = 1$ finally matches $a_T = a^1$. By section 4.4, this s is precisely the solution to the equation $T - s = \tau^*(A_0^1 + s, a^1)$.

Assuming $A_0^1 = 0$ or else $A_0^1 > 0$ but $T < \tau^*(A_0^1, a^1)$, let us first prove the uniqueness by showing that any constrained optimum a must be of the above form, that is, $a_t = A_0^1 + t$ on $[0, s]$, and a_t satisfies the Euler-Lagrange equation on $[s, T]$ with $\dot{a}_s = 1$. The reason is that, as observed in Section 4.3, a solution is strictly concave at any time with slope 1.

Indeed, assume a is a constrained optimum. As before $\dot{a} > 0$ everywhere, for otherwise a would be constant, contradicting $a^1 > A_0^1$. Since a is an optimum, it follows that a satisfies the first-order conditions on the open set $U := \{t \in [0, T] : \dot{a}_t < 1\}$, i.e., is a solution of the Euler-Lagrange equation there. Set $s := \sup\{t \in [0, T] : \dot{a}_t = 1\}$. We claim that $a_t = A_0^1 + t$ on $[0, s]$, from which it clearly follows that $U = (s, T]$. Assume otherwise, i.e., that $\dot{a}_u < 1$ for some $u < s$. Let $v = \inf\{t \geq u : \dot{a}_t = 1\}$. Then $\dot{a}_t < 1$ on (u, v). Hence a satisfies the Euler-Lagrange equation on (u, v). But $\dot{a}_v = 1$ by continuity of $\dot{a}$. Therefore (by extending a as a solution beyond v if necessary), a is strictly concave at v by Section 4.3, implying $\dot{a}_t > 1$ for t near but less than v. This contradiction proves the claim.

We have thus shown that a constrained maximum, if it exists as we have conjectured, is necessarily unique and given as above, namely it is initially affine linear with slope 1 until some $s \geq 0$, and afterwards it equals the analytic solution to the Euler-Lagrange equation. Moreover $s = 0$ if $A_0^1 > 0$ and $T \geq \tau^*(A_0^1, a^1)$, and otherwise s is solution to the equation

$$T - s = \tau^*(A_0^1 + s, a^1). \tag{4.14}$$

Since $L(a, 1) = 0$, we further conclude that, conditioned on our conjecture being true, the maximum value itself is given in all cases by equation (4.10) with $a_s = A_0^1 + s$.

4.8. **Constrained local maximality.** Conditioned on our conjecture that a global maximum exists, we showed it is unique and identified it analytically in the previous section. Whether or not the conjecture is true, we argue in this section that the proposed solution a is at least a local maximum. We already showed in Section 4.6 that the second-order conditions for local maximality hold on $(s, T]$ where the constraints are inactive, i.e., $\dot{a} < 1$. What remains to be shown is the Karush-Kuhn-Tucker maximality conditions, i.e., that the Lagrange multiplier λ_t of each active inequality constraint $\dot{a}_t \leq 1$, $t \in [0, s]$, is nonnegative.

As in finite dimensions, the Karush-Kuhn-Tucker conditions are established by verifying that the differential of the objective functional is a linear combination of the differentials of the constraint functionals with nonnegative coefficients $\lambda_t \geq 0$. (The nonnegativity is because we are maximizing and the inequalities are "$\leq$".) A difference here is that there are a continuum of (linear) constraints $\dot{a}_t \leq 1$, one for each $t \in [0, s]$; as such, the Lagrange multipliers λ_t form a function on $[0, s]$ and the linear combination is really an integral.

For each $t \in [0, s]$, the differential of the constraint functional $a \mapsto \dot{a}_t$ is the continuous linear functional $b \mapsto \dot{b}_t$ on the subspace of $\mathcal{C}^1[0, T]$ consisting of variations b with $b_0 = b_T = 0$. As such, a "linear combination" of these differential with coefficients λ_t is represented by the continuous linear functional $b \mapsto \int_0^s \lambda_t \dot{b}_t dt$. Integrating by parts, since $b_0 = 0$, this is the same as $b \mapsto \lambda_s b_s - \int_0^s \dot{\lambda}_t b_t dt$. On the other hand, the differential of the objective functional (at any point a) is identified as in Section 4.2 by the linear functional $b \mapsto \int_0^T (\frac{\partial L}{\partial a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{a}}) b_t dt$. Setting them equal, since b was arbitrary, we conclude that our analytic solution a is a local maximum if there exists a nonnegative C^1 function $\lambda_t \geq 0$ on $[0, s]$ such that

$$\frac{\partial L}{\partial a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{a}} = \lambda_s \delta_s - 1_{[0,s]} \dot{\lambda}$$

on $[0, T]$ in the sense of distribution, where δ_s is the Dirac delta function at s: $\langle \delta_s, b \rangle = b_s$.

The equation implies that, as expected, the Euler-Lagrange equation holds on $(s, T]$, for the right hand side vanishes there. When $\lambda_s \neq 0$, the δ-function singularity at s indicates that a is not C^1 at s. But in our case a is C^1 by construction (with $\dot{a}_s = 1$), which forces λ_s to be zero. Indeed, the left hand side above has only a step-function singularity at s (as $\ddot{a}_{s-} = 0$, $\ddot{a}_{s+} < 0$), so the right hand side cannot have a δ singularity, requiring $\lambda_s = 0$. The local maximality conditions thus become that for some C^1 function $\lambda \geq 0$ with $\lambda_s = 0$,

$$\frac{\partial L}{\partial a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{a}} = -1_{[0,s]} \dot{\lambda}. \qquad (\lambda_s = 0)$$

Integrating and using $\lambda_t = -\int_t^s \dot{\lambda}_u du$ (since $\lambda_s = 0$) yields that for $t \leq s$,

$$\lambda_t = \int_t^s (\frac{\partial L}{\partial a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{a}}) dt. \qquad (0 \leq t \leq s) \tag{4.15}$$

Therefore, the Karush-Kuhn-Tucker condition $\lambda_t \geq 0$ for local maximality is equivalent to

$$\int_t^s (\frac{\partial L}{\partial a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{a}}) dt \geq 0. \qquad (0 \leq t \leq s) \tag{4.16}$$

Our solution a satisfies this condition. Specifically, the Lagrange multiplier function (at a) is given by $\lambda_t = e^{A_0^1}(e^s - e^t)$, $t < s$, which is positive. Indeed, since $\dot{a} = 1$ on $[0, s]$, by the derivatives formulae in Section 4.1, we have $\frac{\partial L}{\partial a} = 0$ while $\frac{\partial L}{\partial \dot{a}} = 1 - e^a$ on $[0, s]$, yielding

$$\lambda_t = \int_t^s (\frac{\partial L}{\partial a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{a}}) dt = -(1 - e^a)|_t^s = e^{a_s} - e^{a_t} = e^{A_0^1}(e^s - e^t) > 0. \qquad (0 \leq t < s)$$

Due to the strict inequality, it follows that our solution is in fact a *strict* local maximum.

5. The multivariate support

Here, we continue the discussion of Section 3 beyond the trivariate case. The geometry of the feasible set is no longer so simple as there are now nontrivial equality constraints.

5.1. **The calculus of variation formulation.** The multivariate support $\mathcal{S}(A_t^1, \cdots, A_t^{i+1})$ is in principle governed by an induction. The inductive step requires that for each $(a^1, \cdots, a^i) \in \mathcal{S}(A_t^1, \cdots, A_t^i)$ the set $\{a^{i+1} \in \mathbb{R} : (a^1, \cdots, a^{i+1}) \in \mathcal{S}(A_t^1, \cdots, A_t^{i+1})\}$ be connected. This appears to be more difficult for $i \geq 2$ than the trivariate case $i = 1$ proved previously, and we just assume it here. Then, as in the trivariate case, the support is clearly given by

$$\begin{aligned}\mathcal{S}(A_t^1, \cdots, A_t^{i+1}) = \{(a^1, \cdots, a^{i+1}) \in \mathcal{S}(A_t^1, \cdots, A_t^i) \times \mathbb{R} :& \\ A_*^{i+1}(t, a^1, \cdots, a^i) \leq a^{i+1} \leq A^{i+1^*}(t, a^1, \cdots, a^i)\},&\end{aligned} \tag{5.1}$$

where

$$A_*^{i+1}(t, a^1, \cdots, a^i) := \operatorname{ess} \inf_{\{\omega \in \Omega:\, A_t^1(\omega) = a^1, \cdots, A_t^i(\omega) = a^i\}} A_t^{i+1}(\omega), \tag{5.2}$$

and supremum function A^{i+1^*} is defined similarly. Recall from Eq. (3.9) that for all i,

$$\dot{A}^{i+1} = L_i(A^i, \dot{A}^i).$$

Let us first look at the case $i = 2$. Clearly, $A_*^3(T, a^1, a^2)$ is determined by minimizing a functional defined on a subset of the space of $\mathbb{R}^2$-valued curves $(a_1, a_2) \in \mathcal{C}^1[0, T] \times \mathcal{C}^1[0, T]$. The functional, given by $(a_1, a_2) \mapsto \int_0^T L_2(a_2(t), \dot{a}_2(t))dt$, actually depends only on the second component a_2, but the feasible set is defined by the equality constraint $\dot{a}_2 = L_1(a_1, \dot{a}_1)$, in addition to the previous inequality constraint $\frac{\sigma_3 - \sigma_2}{2} \leq \dot{a}_1 \leq \frac{\sigma_3 - \sigma_2}{2}$, as well as by given initial and terminal values $a_1(0) = A_0^1$, $a_2(0) = A_0^2$, $a_1(T) = a^1$, $a_2(T) = a^2$.

For general i, one arrives in a similar way to the more succinct formulation,

$$A_*^{i+1}(T, a^1, \cdots, a^i) - A_0^{i+1} = \inf_{(a_1, \cdots, a_i) \in \mathcal{C}_i^1[0,T]} \int_0^T L_i(a_i(t), \dot{a}_i(t))dt, \tag{5.3}$$

where

$$\begin{aligned}\mathcal{C}_i^1[0, T] := \{(a_1, \cdots, a_i) \in \mathcal{C}^1[0, T] \times \cdots \times \mathcal{C}^1[0, T] :\ a_k(0) = A_0^k,\ a_k(T) = a^k,& \\ k = 1, \cdots, i\,;\ \frac{\sigma_3 - \sigma_2}{2} \leq \dot{a}_1 \leq \frac{\sigma_3 + \sigma_2}{2};\ \dot{a}_{j+1} = L_j(a_j, \dot{a}_j),\ j = 1, \cdots, i-1\}.&\end{aligned}$$

The complication over the trivariate case $i = 1$ is that the feasible set now consists of $\mathbb{R}^i$-valued functions $(a_1(t), \cdots, a_i(t))$ subject to the $i-1$ equality constraints $\dot{a}_{j+1} = L_j(a_j, \dot{a}_j)$ (which incidently imply $\frac{\sigma_{j+2} - \sigma_{j+1}}{2} \leq \dot{a}_{j+1} \leq \frac{\sigma_{j+2} + \sigma_{j+1}}{2}$ since L_j satisfies this inequality).

Let $a = (a_1, \cdots, a_i)$ be a point in the feasible set $\mathcal{C}_i^1[0, T]$ where the inequality is not binding, i.e., $\frac{\sigma_3 - \sigma_2}{2} < \dot{a}_1 < \frac{\sigma_3 + \sigma_2}{2}$. Then, the tangent space to the submanifold $\mathcal{C}_i^1[0, T]$ at a is the set of variations $b = (b_1, \cdots, b_i) \in \mathcal{C}^1[0, T] \times \cdots \times \mathcal{C}^1[0, T]$ such that $b_k(0) = b_k(T) = 0$ for all k and $\frac{d}{d\varepsilon}|_{\varepsilon=0}(\dot{a}_{j+1} + \varepsilon \dot{b}_{j+1}) = \frac{d}{d\varepsilon}|_{\varepsilon=0} L_j(a_j + \varepsilon b_j, \dot{a}_j + \varepsilon \dot{b}_j)$ for $j \leq i - 1$, or equivalently,

$$\dot{b}_{j+1} = \frac{\partial L_j}{\partial a_j} b_j + \frac{\partial L_j}{\partial \dot{a}_j} \dot{b}_j. \qquad (1 \leq j \leq i-1) \tag{5.4}$$

5.2. **The first-order conditions.** Since the objective function is $\int_0^T L_i(a_i(t), \dot{a}_i(t))dt$, by the usual integration by parts, the first-order conditions for local optimality require that

$$\int_0^T (\frac{\partial L_i}{\partial a_i} - \frac{d}{dt}\frac{\partial L_i}{\partial \dot{a}_i}) b_i dt = 0 \tag{5.5}$$

for all variations $b = (b_1, \cdots, b_i)$ tangent to the feasible set at the solution a. Unless $i = 1$, the function b_i is not free, but rather is subject to the constraint (5.4) with $j = i - 1$. So, unlike the trivariate case $i = 1$, we cannot conclude from (5.5) the standard Euler-Lagrange equation for L_i when $i \geq 2$. Instead, we integrate (5.5) by parts and then substitute from (5.4) for $\dot{b}_i$. We get an integrand involving b_{i-1}, and repeat integration by parts and substitution from (5.4) until we reach b_1. Then we end up with a factor times b_1 whose integral is zero, implying the factor equals zero since there is no constraint on b_1. Denoting

$$\dot{k}_i = \frac{\partial L_i}{\partial a_i} - \frac{d}{dt}\frac{\partial L_i}{\partial \dot{a}_i}, \tag{5.6}$$

Eq. (5.5) becomes $\int_0^T \dot{k}_i b_i dt = 0$. Integration by parts yields $\int_0^T k_i \dot{b}_i dt = 0$, as the boundary terms drop out. Substituting for $\dot{b}_i$ from (5.4) and integrating by parts again gives,

$$\begin{aligned} 0 = \int_0^T k_i \dot{b}_i dt &= \int_0^T k_i (\frac{\partial L_{i-1}}{\partial a_{i-1}} b_{i-1} + \frac{\partial L_{i-1}}{\partial \dot{a}_{i-1}} \dot{b}_{i-1}) dt \\ = \int_0^T (k_i \frac{\partial L_{i-1}}{\partial a_{i-1}} &- \frac{d}{dt}(k_i \frac{\partial L_{i-1}}{\partial \dot{a}_{i-1}})) b_{i-1} dt = \int_0^T \dot{k}_{i-1} b_{i-1} dt, \end{aligned}$$

where

$$\dot{k}_{i-1} := k_i \frac{\partial L_{i-1}}{\partial a_{i-1}} - \frac{d}{dt}(k_i \frac{\partial L_{i-1}}{\partial \dot{a}_{i-1}}).$$

The next iteration consists similarly of integration by parts of the above equation $0 = \int_0^T \dot{k}_{i-1} b_{i-1} dt$ to get $\int_0^T k_{i-1} \dot{b}_{i-1} dt = 0$, followed by substitution for $\dot{b}_{i-1}$ from (5.4), followed by another integration by parts to yield $\int_0^T \dot{k}_{i-2} b_{i-2} dt = 0$, with $\dot{k}_{i-2}$ defined similarly.

We continue in this way, setting along the way at the j-th stage

$$\dot{k}_j = k_{j+1} \frac{\partial L_j}{\partial a_j} - \frac{d}{dt}(k_{j+1} \frac{\partial L_j}{\partial \dot{a}_j}), \qquad (1 \leq j \leq i-1) \tag{5.7}$$

while noting that first order conditions become $\int_0^T \dot{k}_j b_j dt = 0$, until we reach $j = 1$. At the final stage $j = 1$, the first-order conditions have become $\int_0^T \dot{k}_1 b_1 dt = 0$. Since b_1 was arbitrary (subject only to $b_1(0) = b_1(T) = 0$), we conclude that $\dot{k}_1 = 0$, that is,

$$k_2 \frac{\partial L_1}{\partial a_1} = \frac{d}{dt}(k_2 \frac{\partial L_1}{\partial \dot{a}_1}) \qquad (i \geq 2). \tag{5.8}$$

For fixed i, Equations (5.6), (5.7), (5.8), together with the constraint equations

$$\dot{a}_{j+1} = L_j(a_j, \dot{a}_j), \qquad (1 \leq j \leq i-1) \tag{5.9}$$

furnish $2i - 1$ differential equations for the $2i - 1$ unknown functions $a_1, \cdots, a_i, k_2, \cdots, k_i$. As such, the first-order conditions consist precisely of this $2i - 1$ by $2i - 1$ ODE system.

In the trivariate case $i = 1$, we have as before a one-dimensional (second-order) ODE for a_1, namely, Eq. (5.6) with $\dot{k}_1 = 0$. For $i = 2$, we get a 3×3 ODE system for a_1, a_2, k_2:

$$\dot{k}_2 = \frac{\partial L_2}{\partial a_2} - \frac{d}{dt}\frac{\partial L_2}{\partial \dot{a}_2}, \qquad (i = 2)$$
$$k_2 \frac{\partial L_1}{\partial a_1} = \frac{d}{dt}(k_2 \frac{\partial L_1}{\partial \dot{a}_1}),$$
$$\dot{a}_2 = L_1(a_1, \dot{a}_1).$$

We can if desired eliminate k_2 by combining the first two ODEs into a single ODE[7]

$$\frac{d}{dt}(\frac{(\frac{\partial L_2}{\partial a_2} - \frac{d}{dt}\frac{\partial L_2}{\partial \dot{a}_2})\frac{\partial L_1}{\partial \dot{a}_1}}{\frac{\partial L_1}{\partial a_1} - \frac{d}{dt}\frac{\partial L_1}{\partial \dot{a}_1}}) = \frac{\partial L_2}{\partial a_2} - \frac{d}{dt}\frac{\partial L_2}{\partial \dot{a}_2}. \qquad (i = 2)$$

This third-order ODE, together with $\dot{a}_2 = L_1(a_1, \dot{a}_1)$, provides a 2×2 ODE for (a_1, a_2).

We note that only for $i \geq 3$ is the ODE (5.7) manifest in full general form, e.g.,

$$\dot{k}_2 = k_3 \frac{\partial L_2}{\partial a_2} - \frac{d}{dt}(k_3 \frac{\partial L_2}{\partial \dot{a}_2}). \qquad (i \geq 3)$$

5.3. **An integral of motion.** We obtain a conserved quantity E by multiplying both sides of Eq. (5.7) by $\dot{a}_j$ and using the product and the chain rules to simplify as in the Beltrami identity, and then telescoping the resulting equations. For $1 \leq j \leq i-1$ we have,

$$\begin{aligned}\dot{k}_j \dot{a}_j &= k_{j+1}\frac{\partial L_j}{\partial a_j}\dot{a}_j - \dot{a}_j \frac{d}{dt}(k_{j+1}\frac{\partial L_j}{\partial \dot{a}_j}) \\ &= k_{j+1}\frac{dL_j}{dt} - \frac{d}{dt}(k_{j+1}\dot{a}_j \frac{\partial L_j}{\partial \dot{a}_j}) \\ &= \frac{d}{dt}(k_{j+1}(L_j - \dot{a}_j \frac{\partial L_j}{\partial \dot{a}_j})) - \dot{k}_{j+1}\dot{a}_{j+1}.\end{aligned}$$

The above is also true for $j = i$ provided we define $k_{i+1} := 1$. Starting with $j = 1$, where we know $\dot{k}_1 = 0$, substituting from the above, and continuing to telescope over j, we get

$$\begin{aligned}0 = \dot{k}_1\dot{a}_1 &= \frac{d}{dt}(k_2(L_1 - \dot{a}_1\frac{\partial L_1}{\partial \dot{a}_1})) - \dot{k}_2\dot{a}_2 \\ &= \frac{d}{dt}(k_2(L_1 - \dot{a}_1\frac{\partial L_1}{\partial \dot{a}_1}) - k_3(L_2 - \dot{a}_2\frac{\partial L_2}{\partial \dot{a}_2})) - \dot{k}_3\dot{a}_3 = \cdots \\ &= \frac{d}{dt}(k_2(L_1 - \dot{a}_1\frac{\partial L_1}{\partial \dot{a}_1}) - k_3(L_2 - \dot{a}_2\frac{\partial L_2}{\partial \dot{a}_2}) + \cdots + (-1)^i k_i(L_{i-1} - \dot{a}_{i-1}\frac{\partial L_{i-1}}{\partial \dot{a}_{i-1}}) + (-1)^{i+1}(L_i - \dot{a}_i\frac{\partial L_i}{\partial \dot{a}_i})).\end{aligned}$$

[7] Use that the second equation is equivalent to the equation below, and then use the first equation twice.

$$k_2(\frac{\partial L_1}{\partial a_1} - \frac{d}{dt}\frac{\partial L_1}{\partial \dot{a}_1}) = \dot{k}_2 \frac{\partial L_1}{\partial \dot{a}_1}.$$

Thus the total differential is zero along a solution, and we conclude

$$E := \sum_{j=1}^{i-1} (-1)^j k_{j+1}(L_j - \dot{a}_j \frac{\partial L_j}{\partial \dot{a}_j}) + (-1)^i (L_i - \dot{a}_i \frac{\partial L_i}{\partial \dot{a}_i}) = \text{constant of motion.} \quad (5.10)$$

For example, for $i = 2$, we have

$$E := -k_2(L_1 - \dot{a}_1 \frac{\partial L_1}{\partial \dot{a}_1}) + (L_2 - \dot{a}_2 \frac{\partial L_2}{\partial \dot{a}_2}). \qquad (i = 2)$$

Solving for k_2, differentiating, and then comparing with $\dot{k}_2 = \frac{\partial L_2}{\partial a_2} - \frac{d}{dt}\frac{\partial L_2}{\partial \dot{a}_2}$, we get

$$\frac{d}{dt}(\frac{L_2 - \dot{a}_2 \frac{\partial L_2}{\partial \dot{a}_2} - E}{L_1 - \dot{a}_1 \frac{\partial L_1}{\partial \dot{a}_1}}) = \frac{\partial L_2}{\partial a_2} - \frac{d}{dt}\frac{\partial L_2}{\partial \dot{a}_2}. \qquad (i = 2)$$

As in the previous subsection, this combined with $\dot{a}_2 = L_1(a_1, \dot{a}_1)$ gives a 2×2 ODE system for (a_1, a_2), but it represents an improvement, as the ODE is second-order now. Despite this, an analytic solution is evidently formidable, even in the flat volatility case.

6. Appendix

6.1. **Details of Eq. (1.4) (c.f. end of Section 2).** Recall, $\mathcal{S}(A_t^1) = [A_0^1, A_0^1 + t]$. Moreover, since $A_0^1 \geq 0$, by Eq. (2.2), $A^{2*}(t, a^1) - A_0^2 = a^1 - A_0^1$. Hence, by Eq. (2.1)

$$\mathcal{S}(A_t^1, A_t^2) = \{(a^1, a^2) \in \mathbb{R}^2 : 0 \leq a^1 - A_1^0 \leq t \ ; \ A_*^2(t, a^1) - A_0^2 \leq a^2 - A_0^2 \leq a^1 - A_0^1\}.$$

Using that $A_t^i := X_t^i - X_t^{i+1}$ and $\mathcal{S}(X_t^1, A_t^1, A_t^2) = \mathbb{R} \times \mathcal{S}(A_t^1, A_t^2)$, it follows that

$$\mathcal{S}(X_t^1, X_t^2, X_t^3) = \{(x^1, x^2, x^3) : 0 \leq x^1 - x^2 - A_0^1 \leq t, \ A_*^2(t, x^1 - x^2) - A_0^2 \leq x^2 - x^3 - A_0^2 \leq x^1 - x^2 - A_0^1\}.$$

Exponentiating, while denoting $X = e^{X^1}$, $Y = e^{X^2}$, $Z = e^{X^3}$ gives

$$\mathcal{S}(X_t, Y_t, Z_t) = \{(x, y, z) \in (0, \infty)^3 : e^{-\int_0^t \sigma_s^2 ds} \frac{x}{X_0} \leq \frac{y}{Y_0} \leq \frac{x}{X_0}, \ e^{A_*^2(t, \log \frac{x}{y}) - A_0^2} \leq \frac{Z_0 y}{Y_0 z} \leq \frac{Y_0 x}{X_0 y}\}.$$

Therefore to prove Eq. (1.4) (with $\sigma = 1$), it remains to show that

$$A_*^2(t, \log \frac{x}{y}) - A_0^2 = A_*(t, \frac{y}{x}, \frac{Y_0}{X_0}).$$

This is equivalent to showing

$$A_2^*(t, -\log(r)) - A_0^2 = A_*(t, r, r_0),$$

where

$$r := \frac{y}{x} \ , \quad r_0 := \frac{Y_0}{X_0} = e^{-A_0^1}.$$

But, this is a matter of translating the analytic solution in Section 2.8 to the notation of Section 2.1. First, we must show that the solution s to $T - s = \tau^*(A_0^1 + s, a^1)$ is the same as that to $T - s = t^*(r, r_0 e^{-s})$. Indeed, the definitions $t^*(r, p)$ (Eq. (1.2)) and $\tau^*(\alpha, a)$ (Eq. (2.7)) give that $t^*(r, p) = \tau^*(\alpha, a)$, provided $r = e^{-a}$ and $p = e^{-\alpha}$. Therefore, $\tau^*(A_0^1 + s, a^1) = t^*(r, r_0 e^{-s})$ because $e^{-A_0^1 - s} = r_0 e^{-s}$ and $r = e^{-a^1}$ (as $a^1 = \log x - \log y$).

Second, we must show that if we substitute $a_s = s + A_0^1$ in equation (2.8) and then subtract (2.8) from $a_T - A_1^0$ (which results in $A_*^2(T, a_T) - A_0^2$), we obtain $A_*(t, r, r_0)$. That is, we must show that

$$A_*(t, r, r_0) = -A_1^0 + a_s - e^{-a_T} + e^{-a_s}$$
$$+\frac{(\arcsin(e^{-\frac{a_s}{2}}) + \sqrt{e^{-a_s} - e^{-2a_s}} - \arcsin(e^{-\frac{a_T}{2}}) - \sqrt{e^{-a_T} - e^{-2a_T}})^2}{T - s - a_T - e^{-a_T} + a_s + e^{-a_s}},$$

where

$$r = \frac{y}{x} = e^{-a_T}, \quad r_0 := \frac{Y_0}{X_0} = e^{-A_0^1}, \quad a_s = s + A_0^1 \quad (e^{-a_s} = r_0 e^{-s}).$$

But, this is quite immediate from the definition $A_*(T, r, r_0)$ given by Eq. (1.3).

6.2. **Details of calculation $L_i(a, \dot{a})$ (c.f. Eq. (3.11)).** Using that $f(x) := \frac{1}{1+e^{-x}}$ and $e^{-\alpha f^{-1}(y)} = (\frac{1-y}{y})^\alpha$, we calculate

$$f(\alpha f^{-1}(\gamma + \frac{1-\alpha}{2}) - \beta) = \frac{1}{1 + e^{-\alpha f^{-1}(\gamma + \frac{1-\alpha}{2}) + \beta}}$$
$$= \frac{1}{1 + (\frac{1-\gamma-\frac{1-\alpha}{2}}{\gamma+\frac{1-\alpha}{2}})^\alpha e^\beta} = \frac{1}{1 + (\frac{1+\alpha-2\gamma}{2\gamma+1-\alpha})^\alpha e^\beta}$$
$$= \frac{(2\gamma + 1 - \alpha)^\alpha}{(2\gamma + 1 - \alpha)^\alpha + (1 + \alpha - 2\gamma)^\alpha e^\beta}.$$

Equation (3.11) now follows easily by setting

$$\alpha := \frac{\sigma_{i+1}}{\sigma_i}, \quad \gamma = \frac{\dot{a}}{\sigma_i}, \quad \beta = \sigma_{i+1} a.$$

6.3. **Details of partial derivatives (cf. Sec. 4.1).** Recall,

$$L(a, \dot{a}) = \dot{a} - \frac{\dot{a}}{\dot{a} + e^a - \dot{a}e^a}.$$

Hence, differentiating with respect to $\dot{a}$,

$$\frac{\partial L}{\partial \dot{a}} = 1 - \frac{\dot{a} + e^a - \dot{a}e^a - \dot{a}(1 - e^a)}{(\dot{a} + e^a - \dot{a}e^a)^2} = 1 - \frac{e^a}{(\dot{a} + e^a - \dot{a}e^a)^2}.$$

Differentiating with respect to $\dot{a}$ again,

$$\frac{\partial^2 L}{\partial \dot{a}^2} = 2e^a \frac{1 - e^a}{(\dot{a} + e^a - \dot{a}e^a)^3} = -\frac{2e^a(e^a - 1)}{(\dot{a} + e^a - \dot{a}e^a)^3}.$$

While differentiating with respect to a,

$$\frac{\partial^2 L}{\partial a \partial \dot{a}} = -\frac{e^a(\dot{a} + e^a - \dot{a}e^a) - 2e^a(e^a - \dot{a}e^a)}{(\dot{a} + e^a - \dot{a}e^a)^3} = \frac{e^a(e^a - \dot{a} - \dot{a}e^a)}{(\dot{a} + e^a - \dot{a}e^a)^3}.$$

Similarly, differentiating L with respect to a,

$$\frac{\partial L}{\partial a} = \dot{a}\frac{e^a - \dot{a}e^a}{(\dot{a} + e^a - \dot{a}e^a)^2} = \frac{e^a \dot{a}(1 - \dot{a})}{(\dot{a} + e^a - \dot{a}e^a)^2};$$

Differentiating with respect to a again,

$$\frac{\partial^2 L}{\partial a^2} = \dot{a}(1-\dot{a})\frac{e^a(\dot{a}+e^a-\dot{a}e^a) - 2e^a(e^a-\dot{a}e^a)}{(\dot{a}+e^a-\dot{a}e^a)^3} = \frac{\dot{a}(1-\dot{a})e^a(\dot{a}-e^a+\dot{a}e^a)}{(\dot{a}+e^a-\dot{a}e^a)^3}.$$

6.4. **Details of $\ddot{a}_t \leq 0$ (c.f. Sec 4.3.)** Set $K := \sqrt{E}$. We can rewrite Eq. (4.6) as

$$\dot{a} = \frac{K}{\sqrt{e^{-a}-e^{-2a}} + K(1-e^{-a})}.$$

Differentiating $\dot{a}$ as a function of a, we get

$$\ddot{a} = -K\dot{a}\frac{\frac{2e^{-2a}-e^{-a}}{2\sqrt{e^{-a}-e^{-2a}}} + Ke^{-a}}{(\sqrt{e^{-a}-e^{-2a}} + K(1-e^{-a}))^2}.$$

Since $\dot{a} > 0$, it follows that $\ddot{a} \leq 0$ if and only if the numerator is nonnegative, i.e.,

$$\frac{2e^{-2a}-e^{-a}}{2\sqrt{e^{-a}-e^{-2a}}} + Ke^{-a} \geq 0.$$

This is in turn equivalent to

$$K \geq \frac{1-2e^{-a}}{2\sqrt{e^{-a}-e^{-2a}}} = \frac{\frac{1}{2}e^a - 1}{\sqrt{e^a-1}}.$$

This proves the second condition. Substituting from the definition of energy in Eq. (4.5) for K easily yields the first condition, namely that $\ddot{a} \leq 0$ if and only if $\dot{a} \geq 1 - \frac{1}{e^a-1}$.

6.5. **Details of the solution (c.f. Sec. 4.4).** Set $x = e^{-a}$. Then $dx = -xda$. Hence,

$$\sqrt{e^{-a}-e^{-2a}}\,da = -\sqrt{\frac{1}{x}-1}\,dx.$$

But, since

$$\frac{d}{dy}\arcsin(y) = \frac{1}{\cos(\arcsin(y))} = \frac{1}{\sqrt{1-y^2}},$$

we have

$$\frac{d}{dx}(\arcsin(\sqrt{x}) + \sqrt{x(1-x)}) = \frac{1}{\sqrt{1-x}}\frac{1}{2\sqrt{x}} + \frac{1-2x}{2\sqrt{x(1-x)}}.$$

$$= \frac{1-x}{\sqrt{x(1-x)}} = \sqrt{\frac{1}{x}-1}.$$

It follows that, as desired,

$$\sqrt{e^{-a}-e^{-2a}}\,da = -d(\arcsin(e^{-\frac{a}{2}}) + \sqrt{e^{-a}-e^{-2a}}).$$

6.6. **Details of second order conditions.** Using the partial derivatives formulae of Section 4.1, we calculate

$$\frac{\partial L}{\partial a} - \dot{a}\frac{\partial^2 L}{\partial \dot{a}\partial a} = \frac{e^a \dot{a}(1-\dot{a})}{(\dot{a}+e^a-\dot{a}e^a)^2} - \dot{a}\frac{e^a(e^a-\dot{a}-\dot{a}e^a)}{(\dot{a}+e^a-\dot{a}e^a)^3}$$
$$= \frac{\dot{a}e^a}{(\dot{a}+e^a-\dot{a}e^a)^3}((1-\dot{a})(\dot{a}+e^a-\dot{a}e^a)-(e^a-\dot{a}-\dot{a}e^a))$$
$$= \frac{\dot{a}^2 e^a}{(\dot{a}+e^a-\dot{a}e^a)^3}(2+\dot{a}e^a-\dot{a}-e^a)$$
$$= \frac{\dot{a}^2}{(\dot{a}+e^a-\dot{a}e^a)^2}(\frac{2e^a}{\dot{a}+e^a-\dot{a}e^a}-e^a)$$
$$= \frac{E}{e^a-1}(2+2\sqrt{E}\sqrt{e^a-1}-e^a)$$
$$= E(\frac{1}{e^a-1}+\frac{2\sqrt{E}}{\sqrt{e^a-1}}-1),$$

as claimed. For the penultimate equality we used, for the first factor that by Eq. (4.5),

$$\frac{\dot{a}^2}{(\dot{a}+e^a-\dot{a}e^a)^2} = \frac{E}{e^a-1},$$

and for the second factor we used

$$\frac{e^a}{\dot{a}+e^a-\dot{a}e^a} = 1+\sqrt{E}\sqrt{e^a-1}.$$

The latter is seen by noting that first by Eq. (4.5) and then by Eq. (4.6), we have

$$\frac{\sqrt{1-e^a}}{\dot{a}+e^a-\dot{a}e^a} = \frac{\sqrt{E}}{\dot{a}} = \sqrt{e^{-a}-e^{-2a}}+\sqrt{E}(1-e^{-a}).$$

Multiplying both sides by $\frac{e^a}{\sqrt{1-e^a}}$ gives the claimed formula.

References

[1] Davis, M.H.A., and Mataix-Pastor, V.: "Negative Libor rates in the Swap Market Model." *Finance and Stochastics* 11 (2), 181-193, (2007).

[2] Jamshidian, F.: "Bivariate support of forward Libor and swap rates." Univ. of Twente working paper, (2005). To appear in *Mathematical Finance*.

[3] Strook, D.W., and Varadhan S.R.S.: "On the support of diffusion processes with applications to the strong maximum principle." *Proc. 6th Berkeley Symp. Math. Statist. Prob.* 3, 333-359, (1972).