LA-UR-13-24707

Approved for public release; distribution is unlimited.

<table>
<tr><td>Title:</td><td>Optimal Stochastic Control as Non-equilibrium Statistical Mechanics: Calculus of Variations over Density and Current</td></tr>
<tr><td>Author(s):</td><td>Chernyak, Vladimir Y
Chertkov, Michael
Bierkens, Joris
Kappen, Hilbert J.</td></tr>
<tr><td>Intended for:</td><td>arXiv
PRL</td></tr>
<tr><td>Issued:</td><td>2013-06-26</td></tr>
</table>

Los Alamos
NATIONAL LABORATORY
—— EST. 1943 ——

Optimal Stochastic Control as Non-equilibrium Statistical Mechanics: Calculus of Variations over Density and Current

Vladimir Y. Chernyak^{a,b}, Michael Chertkov^b, Joris Bierkens ^{c}, and Hilbert J. Kappen ^{c}

^{a}*Department of Chemistry, Wayne State University, 5101 Cass Ave,Detroit, MI 48202, USA*
^{b}*Center for Nonlinear Studies and Theoretical Division, LANL, Los Alamos, NM 87545, USA and*
^{c} *SNN Adaptive Intelligence, Radboud University, Nijmegen P.O. Box 9101, 6500 HB, Nijmegen, the Netherlands*
(Dated: June 26, 2013)

In Optimal Stochastic Control (OSC) one minimizes the average cost-to-go, that consists of the cost-of-control (amount of efforts), cost-of-space (where one wants the system to be) and the target cost (where one wants the system to arrive), for a system participating in forced and controlled Langevin dynamics. We generalize the OSC problem by introducing an additional cost-of-dynamics, characterized by a vector potential. We propose derivation of the generalized gauge-invariant Hamilton-Jacobi-Bellman equation as a variation over density and current, suggest hydrodynamic interpretation and discuss examples, e.g., ergodic control of a particle-within-a-circle, illustrating non-equilibrium space-time complexity.

In its standard setting the problem of the Optimal Stochastic Control (OSC) involves minimizing the average over stochastic trajectories of the cost-to-go, which consists of the cost of using the control field, the cost of arriving to a certain position $\boldsymbol{x}$ in the configuration space $\mathcal{M}$, described by a potential function $\varphi(\boldsymbol{x})$, and the cost accumulated along the trajectory, described by a time-dependent potential $V(\boldsymbol{x}, \tau)$. The potentials $\varphi(\boldsymbol{x})$ and $V(\boldsymbol{x}, \tau)$ can be viewed as variables dual to the particle probability distributions at the arrival time and during the time evolution, respectively, and thus interpreted as Lagrange multipliers.

In this Letter, continuing the thread of [1–6] where methods of statistical and quantum mechanics were applied to OSC, we extend the standard cost-to-go functional by adding a term, associated with a vector potential, $\boldsymbol{A}(\boldsymbol{x}, \tau)$, which leads to (i) a variational derivation of the stochastic Hamilton-Jacobi-Bellman (HJB) equation and its more general Gauge Invariant (GI) version, (ii) extend the capability of the HJB equation to treat such observables as produced work, generated entropy, and fluxes that occur in systems with non-contractible cycles in phase space.

Stochastic dynamics of a particle in a compact m-dimensional space $\mathcal{M}$, e.g., an m-dimensional torus, is described by the following Langevin equation

$$\dot{\eta}^j \doteq \frac{d}{d\tau}\eta^j = f^j(\tau, \boldsymbol{\eta}) + u^j(\tau, \boldsymbol{\eta}) + \xi^j(\tau, \boldsymbol{\eta}), \quad (1)$$

$$\langle \xi^j(\tau, \boldsymbol{\eta}) \rangle = 0, \quad \langle \xi^j(\tau, \boldsymbol{\eta}) \xi^k(\tau', \boldsymbol{\eta}) \rangle = \kappa g^{jk}\delta(\tau - \tau'), \quad (2)$$

where $\tau \in [t; T]$; $\boldsymbol{f}$ is the "force" field, deterministic and assumed known; $\boldsymbol{u}$ is the "control" field which as we will see below is subject to our optimization/choice; and $\boldsymbol{\xi}$ is a δ-correlated in time, zero-mean Gaussian random field, whose correlations are fully expressed via a strictly positive symmetric matrix, $\kappa g(\boldsymbol{\eta})$, where κ measures the noise strength, and g can be viewed as a (generally space-time dependent) metric in the configuration space $\mathcal{M}$, with $g_{ij}g^{jk} = g^{kj}g_{ji} = \delta_i^k$, where we use standard in theoretical physics covariant notations, i.e., assuming summation over repeating pairs of sub/superscripts, and applying the metric to relate vectors to co-vectors, e.g. $f_i = g_{ij}f^j$.

We consider a problem describing the optimal choice of the control vector field, $\boldsymbol{u}$, in Eq. (1)

$$\mathcal{C}(t, \boldsymbol{x}; T) \doteq \min_{\{\boldsymbol{u}\}} C(\{\boldsymbol{u}\}; t, \boldsymbol{x}; T), \quad C(\{\boldsymbol{u}\}; t, \boldsymbol{x}; T) \equiv \left\langle \varphi(\eta(T)) + \int_t^T d\tau \left(\frac{1}{2} h_{ij}u^i u^j + V + A_j \dot{\eta}^j \right) \right\rangle, \quad (3)$$

$$\langle B(t', \eta(t')) \rangle \doteq \frac{\int_{\boldsymbol{\eta}(t)=\boldsymbol{x}} \mathcal{D}\boldsymbol{\eta} \exp\left(-\kappa \int_t^{t'} d\tau (\dot{\eta}^i - f^i - u^i)g_{ij}(\dot{\eta}^j - f^j - u^j) \right) B(t', \eta(t'))}{\int_{\boldsymbol{\eta}(t)=\boldsymbol{x}} \mathcal{D}\boldsymbol{\eta} \exp\left(-\kappa \int_t^{t'} d\tau (\dot{\eta}^i - f^i - u^i)g_{ij}(\dot{\eta}^j - f^j - u^j) \right)}, \quad (4)$$

where Eq. (4) defines averaging over stochastic trajectories in terms of a path integral; all co-vector and tensor fields in the integrand of Eq. (3) may depend explicitly on τ and $\boldsymbol{x}$; minimization/variation over $\{\boldsymbol{u}\}$ is functional, i.e. we minimize over all $u(\tau, \eta(\tau))$; and Eq. (4), stated as a path integral over

$\eta(\tau)$ defines averaging over the stochastic trajectories evolving according to Eq. (1).

The meaning of the four terms under the average in the cost-to-go $C(\{\boldsymbol{u}\}; t, \boldsymbol{x}; T)$ is as follows: the first local term describes the target cost, i.e. the cost for the system to arrive

at the final moment of time T at $\boldsymbol{\eta}(T)$; the second term (which is also the first integral term) defines the cost-of-control; the third term stands for the cost-of-space, as it measures the cost depending on where the system stays in the phase space during the entire interval; finally, the last term in $C(\{\boldsymbol{u}\}; t, \boldsymbol{x}; T)$, as shown in Eq. (3), represents the cost-of-dynamics, i.e. it is sensitive to how the system is moving in phase space during the period of interest (the cost is zero if the system does not move). The first three terms in Eq. (3) are standard in control theory, while the fourth term is new. It is also natural (exploiting theoretical physics jargon and intuition) to refer to V and $\boldsymbol{A}$ in Eq. (3) as the scalar and vector potentials, respectively. Obviously, the average cost-to-go $\mathcal{C}$ depends functionally on V and $\boldsymbol{A}$.

Following [7] we define the so-called average density and average current-density (hereafter referred to as just density and current) for the Langevin dynamics given by Eq. (1):

$$\rho(\{\boldsymbol{u}\}; t, \boldsymbol{x}; T) \doteq \int_t^T \frac{d\tau}{T-t} \langle \delta(\boldsymbol{\eta}(\tau) - \boldsymbol{x}) \rangle, \qquad (5)$$

$$\boldsymbol{J}(\{\boldsymbol{u}\}; t, \boldsymbol{x}; T) \doteq \int_t^T \frac{d\tau}{T-t} \langle \dot{\boldsymbol{\eta}} \delta(\boldsymbol{\eta}(\tau) - \boldsymbol{x}) \rangle. \qquad (6)$$

In what follows we will simplify the notations, dropping the dependence of the density and current on T. Utilizing Eq. (4), and defining averaging over stochastic trajectories, one finds that the density satisfies the Fokker-Planck equation, which can also be interpreted as the continuity equation relating the current and density,

$$\partial_t \rho + \sqrt{g} \partial_i \frac{g^{ij}}{\sqrt{g}} J_j = 0, \quad J_i = -\kappa \partial_i \rho + (f_i + u_i)\rho, \quad (7)$$

where $g \doteq \det(g^{ij})$, and $\partial_i \doteq \partial_{x^i}$. Naturally, ρ is nonnegative and, according to Eq. (7), properly normalized at any time, $\int_{\mathcal{M}} d\boldsymbol{x} \rho/\sqrt{g} = 1$. Note that the second equation in (7) can be inverted, thus expressing the control field $\boldsymbol{u}$ explicitly in terms of the current and density. Therefore, applying Eqs. (5,6,7) to Eq. (3), one arrives at the following expression in terms of the density and current for the average (still not yet optimized with respect to $\boldsymbol{u}$) cost-to-go

$$C = C_0 + \int_{\mathcal{M}} \frac{d\boldsymbol{x}}{\sqrt{g}} \left(\varphi(\boldsymbol{x})\rho(T, \boldsymbol{x}) + \int_t^T \left(V\rho + g^{ij} A_i J_j \right) \right), \quad (8)$$

$$C_0 \doteq \int_{\mathcal{M}} \frac{d\boldsymbol{x}}{\sqrt{g}} \int_t^T d\tau \frac{h^{ij}(\kappa \partial_i \rho + J_i - f_i \rho)(\kappa \partial_i \rho + J_i - f_i \rho)}{2\rho}, \quad (9)$$

where the average cost-of-control C_0 does not depend on the scalar and vector potentials, and is defined as an explicit functional of ρ, $\boldsymbol{J}$ and function of $t, \boldsymbol{x}$ and T, i.e. $C_0(\{\rho\}, \{\boldsymbol{J}\}; t, \boldsymbol{x}; T)$. Thus, the optimization task in Eq. (3) translates into minimizing the r.h.s. of Eq. (8) under the conditions of Eqs. (7). Introducing a functional Lagrangian multiplier $\Phi(\tau, \mathbf{x})$ for the first equation in (7), adding the corresponding term to Eq. (8), we compute the variations over

$\boldsymbol{J}(\tau, \boldsymbol{x})$ and $\rho(\tau, \boldsymbol{x})$, under condition $\delta\rho(t; \boldsymbol{x}) = 0$ (the initial density is fixed). Combining the two variation equations results in the following closed form equation

$$\begin{aligned} \partial_\tau \Phi &= \kappa\sqrt{g}\partial_i \frac{g^{ij}}{\sqrt{g}}(\partial_j \Phi + A_j) + g^{ij} f_i(\partial_j \Phi + A_j) \\ &+ V - \frac{1}{2} g^{ik} h_{kl} g^{lj}(\partial_i \Phi + A_i)(\partial_j \Phi + A_j)), \quad (10) \end{aligned}$$

which should be solved backwards in time with the "initial" condition $\Phi(T, \boldsymbol{x}) = \varphi(\boldsymbol{x})$ that originates from the contact term in the variation over $\rho(T)$.

A number of remarks with regards to Eq. (10) is in order. First, combining the equation emerging in the result of variation over $\boldsymbol{J}$ with the second relation in Eq. (7) one arrives at the following explicit expression of the optimal control field $\boldsymbol{u}$ via the optimal Φ

$$u_i = -\bar{h}_i^k(\partial_k \Phi + A_k), \quad \bar{h}_{ij}h^{jk} = \delta_i^k, \quad \bar{h}_i^k = \bar{h}_{ij}g^{jk}. (11)$$

Second, Eq. (10) is gauge invariant under simultaneous transformation of the scalar and vector potentials: $A_i \to A_i + \partial_i\phi, V \to V + \partial_\tau\phi$, where $\phi(\tau, \boldsymbol{x})$ is an arbitrary scalar function. Third, the Lagrangian multiplier solving Eq. (10) actually coincides with the optimal average cost-to-go function, $\mathcal{C}(t, \boldsymbol{x}; T) = \Phi(t, \boldsymbol{x})$. The relation follows from multiplying Eq. (10) by ρ, integrating the result over the $d\boldsymbol{x}/\sqrt{g}$ and also over τ in the $[t; T]$ interval, and then comparing the final expression with Eq. (8). Forth, the terms in Eq. (10) that contain the gauge field can be obviously absorbed into the force field $\boldsymbol{f}$ and scalar potential V, resulting in the celebrated stochastic Bellman-Hamilton-Jacobi (HJB) equation of the control theory, which means that Eq. (10) is a particular case of the standard HJB equation, and access to an efficient solver of the latter provides a way to solve Eq. (10). On the other hand, the equivalence between $\mathcal{C}$ and Φ means that one can also replace Φ in Eq. (10), thus discovering that Eq. (10) can be viewed as a Gauge Invariant (GI) generalization of the HJB, rather than a particular case, since it allows to consider control over a broader set of phenomena (e.g., work/entropy generation, fluxes, etc.). Finally, combining Eqs. (10,11) we arrive at the following version of the GI-HJB equation (10) stated in terms of the control field

$$\begin{aligned} &- \partial_\tau u_i - \underbrace{\frac{1}{2}\bar{h}_i^k \partial_k h_{kl} h^{lm} h^{kn} u_m u_n}_{\text{"self-advection"}} - \underbrace{\bar{h}_i^k \partial_k f_j h^{jm} u_m}_{\text{"advection by force"}} \\ &= \underbrace{\kappa\bar{h}_i^k \partial_k \sqrt{g}\partial_l h^{ln} u_n/\sqrt{g}}_{\text{"dissipation"}} + \underbrace{\bar{h}_i^k (\partial_\tau A_k - \partial_k V)}_{\text{"pumping/constraints"}}, \quad (12) \end{aligned}$$

where, to avoid tedious expressions, we assumed that both h and g metrics are time-independent.

All terms in Eq. (12) allow for a very natural hydrodynamic interpretation, where the optimal control field, $\boldsymbol{u}$, is interpreted up to a metric-dependent re-normalization as the "velocity" field that evolves backwards in time in a compact space with curvature g. The second term on the l.h.s. of

Eq. (12), also the only nonlinear term in the equation, describes the "self-advection of velocity by itself". Continuing the hydrodynamic analogy, one interprets the second term on the l.h.s as "advection by an external field $\boldsymbol{f}$". Then, the first term on the r.h.s. stands for the dissipation/viscosity, induced by the noise in the original Langevin equation (1). Finally, the last term on the r.h.s. represents pumping/injection, it may also represent constraints, e.g. expressing relations between pressure and density, the phase space hydrodynamics. Details and consequences of the ultimate relation between the control and hydrodynamics will be discussed elsewhere [8].

Scalar and vector potentials, as well as $\varphi(\boldsymbol{x})$, can also be viewed as functional Lagrangian multipliers used to fix specific forms of the density and current functions. (Note, however, that in this formulation ρ and J are not fully arbitrary but consistent with each other through the continuity equation, i.e. the first equation of Eqs. (7).) Therefore, under fixed and consistent scalar and vector potentials no additional optimization in Eq. (3) is needed, the control field is completely defined by the second equation of (7), and then the average cost is just the cost-of-control, C_0, given by Eq. (9).

This GI approach also allows to consider less restrictive cases with constraints which are linear in the density and current. For example, interesting problem is – find the least expansive (in terms of the average cost) control obeying the detailed balance, i.e. with the zero current $\boldsymbol{J} = 0$ for all $\tau \in [t; T]$: $\min_{\{\rho\}} C_0(\{\rho\}, \{\boldsymbol{0}\}; t, \boldsymbol{x}; T)$. Note that, due to the continuity equation, ρ should be time-independent. The most interesting setting corresponds to the stationary force $\boldsymbol{f}$ and control $\boldsymbol{u}$ fields, when we let the system equilibrate and minimize the average cost-to-go in the long-time limit. This corresponds to minimizing with respect to ρ the functional C_0, given by Eq. (9), with a proper Lagrange term added to ensure normalization, in the case $\boldsymbol{J} = 0$ and time-independent ρ and $\boldsymbol{f}$. The above minimization leads to the following equation

$$-\frac{\kappa^2}{\rho}\sqrt{g}\partial_i\frac{h^{ij}}{\sqrt{g}}\partial_j\rho + \frac{\kappa^2}{2\rho^2}h^{ij}(\partial_i\rho)(\partial_j\rho) + U = E \quad (13)$$

with E being the Lagrange multiplier. Here in Eq. (13) we introduced the potential $U = \frac{1}{2}h^{ij}f_if_j + \kappa\sqrt{g}\partial_i\left(h^{ij}f_j/\sqrt{g}\right)$. Under the substitution, $\psi = \sqrt{\rho}$, Eq. (13), adopts a form of the Schrödinger equation, $-(\hbar^2/2)\sqrt{g}\partial_i(h^{ij}/\sqrt{g})\partial_j\psi + U\psi = E\psi$, where the newly introduced notation, $\hbar = 2\kappa$, aims at making a relation of the OSC problem to the Quantum Mechanics obvious. A straightforward calculation shows that the cost-of-control rate, i.e., thee average cost-to-go per unit time in the long time limit is given by E. Therefore, the problem of minimizing the cost of being at equilibrium is reduced to finding the ground state of a Schrödinger equation, with the low-noise case corresponding to the quasi-classical limit.

Returning to the GI-HJB setting, Eq. (12) or Eq. (10) are nonlinear, however in the case when the two metrics, g and h, characterizing the statistics of noise and the cost-of-control respectively, are proportional, i.e. $g^{kj} = 2\kappa q h^{kj}$, where q is the scalar proportionality coefficient, the gauge-invariant HJB

Eq. (10) turns into the linear equation upon substituting, $\Psi \doteq \exp(-q\Phi)$:

$$-\partial_\tau\Psi = \kappa\sqrt{g}(\partial_k - qA_k)\frac{g^{kj}}{\sqrt{g}}(\partial_j - qA_j)\Psi$$
$$+ (g^{kj}f_k(\partial_j - qA_j) - qV)\Psi. \quad (14)$$

The linear Eq. (14) should be solved backwards in time starting with $\Psi(T, \boldsymbol{x}) = \exp(-q\varphi(\boldsymbol{x}))$. This reduction from the nonlinear Eq. (10) to the linear Eq. (14) generalizes what has been explored in the recent papers [3–6, 9–11], devoted to the so-called path-integral and Kullback-Leibler controls. Note also that this transformation is akin the so-called Cole-Hopf transformation of the Burgers equation into an auxiliary linear equation of the Shrödinger type [12]. (See also [13] and references therein for a discussion of the Cole-Hopf transformation applied to a HJB type of equation in the context of the dynamic programming approach in the portfolio optimization, which is a branch of the operation research.)

Another interesting feature of the proportional case is that in the long-time limit the rate $C_0/(T - t)$ of the optimal average cost becomes equal, up to a simple re-scaling, to the so-called Cramér (or large-deviation) functional describing the negative log-probability of density and current fluctuations, $\tilde{\rho}, \tilde{J}$ (defined as the expressions under the average signs on the r.h.s. of Eqs. (5,6), see [7] for details): $\log(-\mathcal{P}(\{\tilde{\rho}\}, \{\tilde{\boldsymbol{J}}\}; 0, \boldsymbol{x}; T)) = qC_0(\{\rho\}, \{\boldsymbol{J}\}; 0, \boldsymbol{x}; T)$.

An important feature of GI-HJB equation in the context of ergodic control, i.e., the long time limit with stationary fields, is the principal capability to optimize over the fluxes. Flux over a non-contractible cycle is defined as the number of times the system goes over the cycle divided by the (long) time T, or equivalently as the integral over the current density $\boldsymbol{J}$ over the corresponding non-contractible $(m - 1)$-dimensional surface (see [7] for details). This can be done by solving the stationary version of Eq. (10) with $V = 0$ and curvature free, i.e., $\partial_iA_j - \partial_jA_i = 0$, vector potential, which is globally still not a gradient $A_i \neq \partial_i\varphi$.

Next we discuss our enabling ergodic case example of a particle moving along a simple circle of length $L = 2\pi$ with constant $g = h = 1$. Note, however, that in this special (and not fully representative) case the flux density and the current density coincide ($J(x) = \text{const}$). This case is analyzed by combining the stationary version of Eq. (12) with the second expression in Eq. (7), resulting in

$$u^2/2 + fu + \kappa\partial_x u = -E, \quad -\kappa\partial_x\rho + (f + u)\rho = J, \quad (15)$$

where E and J should be treated as constants, with the periodic boundary conditions $u(x + 2\pi) = u(x)$ and $\rho(x + 2\pi) = \rho(x)$ imposed. It is convenient to perform analysis implicitly by fixing the value of E, solving Eqs. (15), and thus determining the value of the flux J. This analysis, illustrated in Figs. 1a-c, suggests few observations. First, the cases with and without flux-fixing control are significantly different. Since the bare (without control) flux was smaller (simply zero in

(a)$f(x) = -2\cos(x)$, $\kappa = 0.5$, $J = 0.2$ and $\lim_{T \to \infty} \mathcal{C}/T \approx 2.32$

(b)$f(x) = -2\cos(x)$, $\kappa = 2.5$, $J = 0.2$ and resulting in $\lim_{T \to \infty} \mathcal{C}/T \approx 1.81$

(c)$f(x) = 1 - 2\cos(x)$, $\kappa = 0.5$, $J = 0.09$ and resulting in $\lim_{T \to \infty} \mathcal{C}/T \approx 0.36$
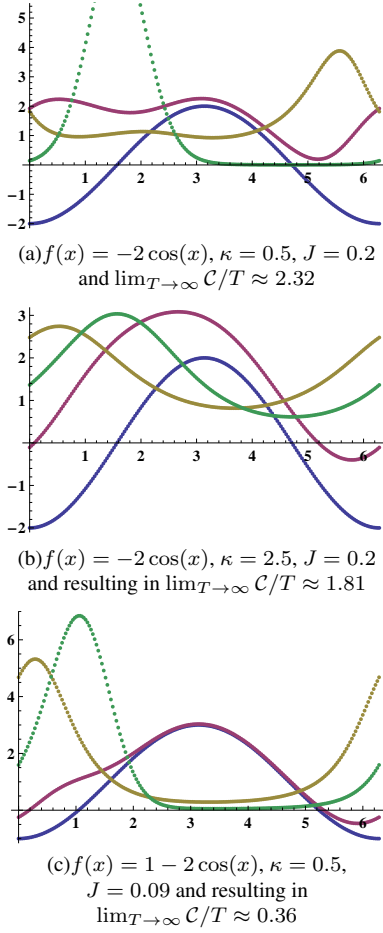
FIG. 1: Three illustrative examples of the 1d (particle on the circle) ergodic control with fixed flux (zero in the cases (a,b) and nonzero in the case (c). The color coding of the curves is as follows: bright green and dark green curves show $10 * \rho(x)$ in the bare (without control) and optimal control cases respectively; blue and purple curves show $f(x)$ and $f(x) + u(x)$ respectively.

the cases of Figs. 1a,b) than the resulting flux under control, the density distribution is significantly more spread out in the control case, also showing appearance of some additional structure (two local maxima in density). Second, comparing Fig. 1a and Fig. 1b, different in diffusion only, we observe that increase in diffusion spreads up the density distribution, resulting in the average decrease of the cost-to-go. We observe that the extra diffusion helps advection to boost the particle transport (flux) with less cost. Third, in the case of Fig. 1c the increasing flux leads to the control field splitting into two components, one modifying the potential (divergence free) component of the force, $f(x)$, and the other enhancing the constant/flux contribution to the force. We conclude that the optimal flux control cannot be explained simply as adjusting the gradient (potential) part of the force, or vice versa as adjusting only the constant contribution leaving the potential

intact.

In this letter we focused on the physics analysis and interpretations of the density/current variational formulation of the OSC. A reader interested in related mathematically rigorous results is advised to consult with [14]. The research of VYC has received support from the NSF under grant agreement no. CHE-1111350. The work at LANL was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy under Contract No. DE-AC52-06NA25396. The research of JB and HJK was funded by the FP7/2007-2013 program under the grant no. 231495.

———

[1] W. H. Fleming. Exit probabilities and optimal stochastic control. *Applied Mathematics and Optimization*, 4(1):329–346, 1977.
[2] Sanjoy K. Mitter. Non-linear filtering and stochastic mechanics. In *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, volume 78, pages 479–503. NATO Advanced Study Institutes Series, 1981.
[3] H.J. Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of Statistical Mechanics: Theory and Experiment*, page P11011, 2005.
[4] E. Todorov. Linearly-solvable markov decision problems. In *In Advances in neural information processing systems*, volume 19, page 13691376. Cambridge: MIT Press, 2007.
[5] E. Todorov. General duality between optimal control and estimation. In *In 47th IEEE conference on decision and control*, page 42864292, 2008.
[6] H.J. Kappen. Optimal control theory and the linear bellman equation. *Inference and Learning in Dynamic Models*, pages 363–387, 2011.
[7] V. Y. Chernyak, M. Chertkov, S.V. Malinin, and R. Teodorescu. Non-equilibrium thermodynamics and topology of currents. *Journal of Statistical Physics*, 137(1):109–147, 2009.
[8] M. Chertkov and V. Chernyak. Fluid mechanics as a single-particle control. *Work in Progress*, 2012.
[9] K. Dvijotham and E. Todorov. A unifying framework for linearly solvable control. In *Proceedings of the 27th Annual Conference on Uncertainty in Artificial Intelligence (UAI-11)*, pages 179–186, Corvallis, Oregon, 2011. AUAI Press.
[10] J. Bierkens and H.J. Kappen. Explicit solution of relative entropy weighted control. *submitted; arXiv:1205.6946*, 2012.
[11] H.J. Kappen, V. Gomez, and M. Opper. Optimal control as a graphical model inference problem. *Machine Learning Journal*, pages 1–11, 2012.
[12] E. Hopf. The partial differential equation $u_t + uu_x = \mu u_{xx}$. *Communications of Pure and Applied Mathematics*, 3:201–230, 1950.
[13] T. Zariphopoulou. Optimal investment and consumption models with non-linear stock dynamics. *Math. Methods Oper. Res.*, 50(2):271–296, 1999.
[14] J. Bierkens, V. Y. Chernyak, M. Chertkov, and H. J. Kappen. Linear PDEs and eigenvalue problems corresponding to ergodic stochastic optimization problems on compact manifolds. *arxiv:1303.0126*, 2013.