

TESI MAGISTRALE

Optimal stochastic transport

Stefano Bo

October 7, 2011

- Relatore interno
Prof. Antonio **Celani**
- Relatore esterno
Prof. Erik **Aurell**

Collaboratori
Prof. R. Eichhorn
Prof. P. Muratore - Ginan-
neschi

Contents

1	Stochastic thermodynamics and optimal control	5
1.1	Introduction to stochastic energetics	5
1.1.1	Physical perspective	7
1.1.2	Mathematical perspective	10
1.2	Introduction to Bellman equation	16
2	Stochastic control and optimal transport	19
2.1	Optimization procedure	19
2.2	Solution of Burgers equation	27
2.3	Computation of the optimal work with given final control	29
3	Exactly solvable optimization problems	33
3.1	Controlling the center of a quadratic potential	33
3.1.1	Solution by Schmiedl and Seifert	33
3.1.2	Solution via AMM	35
3.2	Controlling the stiffness of a quadratic potential	35
3.2.1	Solution by Schmiedl and Seifert	35
3.2.2	Solution via AMM	37
4	Brownian motors and mesoscopic engines	39
4.1	Rectifying motors	39
4.2	Molecular Motors	41
4.3	Classic thermodynamic cycles in stochastic environments	42
4.4	Optimal work for the computation of efficiency at maximum power of a mesoscopic Carnot engine	44
4.4.1	Optimal work for a non-equilibrium initial state	44
4.4.2	Explicit solution of a Carnot mesoscopic engine via AMM	46
5	Optimization in a temperature changing setting	50
5.1	Definition of heat	50
5.2	Temperature depending on time only	51
5.2.1	Heat Optimization for time dependent temperature	51
5.2.2	Entropy production optimization for time dependent temperature	52

5.3	Temperature dependent on time and space	54
5.3.1	Heat optimization for time and space dependent temperature	55
5.3.2	Entropy production optimization for time and space dependent tem- perature	56
6	Minimum entropy cycles	60
6.1	The role of the cost function	60
6.2	Constant probability distribution	61
6.3	A pin-point approach to minimum entropy cycles	65
A	Gaussianity of the distribution for linear drifts	69
	References	71

Synopsis

The following thesis deals with the thermodynamics of small, out of equilibrium systems. The starting point is a recent Letter to Physics Reviews by Aurell, Mejia-Monasterio, Muratore-Ginanneschi [1] (we will denote it as AMM in the following). The paper develops a powerful technique to solve isothermal optimal control problems for thermodynamic stochastic functionals. For the optimization it makes use of Bellman optimality principle. The striking result is the mapping of the involved stochastic transport problem into the deterministic mass transport and Burgers equations of fluid dynamics. The optimization can be performed for any probability distribution and shape of the controlling potential. In this sense it represents a generalization of the solutions derived by Seifert and Schmiedl for the harmonic case [2] which had revealed peculiar properties of stochastic transport such as discontinuous optimal protocols. Mesoscopic thermodynamics can be used to study the energetic properties of Brownian motors. It is also possible to provide a formal thermodynamical description for molecular motors which actually exist in nature (see for example the review by Jülicher, Ajdari and Prost [9]). It is interesting to apply the optimization method to such motors and investigate their optimal properties and their structural limitations. In fact, comparing the characteristics of the theoretically optimal motors with the ones existing in nature can give some further hints about their biological functions or the constraints under which they have to operate. One can, for instance, compute the efficiency at maximum power and see if the molecular motors operate close to the theoretical limit or not. If not, one can speculate that the function of the motor does not require a high efficiency at maximum power or that it is biologically difficult to approach the limit. To take a step into this direction we consider the optimization of the mesoscopic Carnot cycle described by Schmiedl and Seifert in [3] and compute its efficiency at maximum power making use of the AMM optimization method. As for a classical Carnot engine there are two adiabatic and two isothermal transformations. The non-isothermal steps call for a generalization of the AMM optimization method. The contribution of this thesis consists of this generalization of [1] to the non-isothermal case which is suitable for the study of generic thermodynamic cycles. Particular attention is paid to which are the conditions and the functionals allowing the solution to be mapped into Burgers equation. To get a closer insight we consider separately the case of temperature depending on time only and the one of temperature depending both on space and time. The results show how the optimization of heat for temperature changing settings does not lead to Burgers equation and its simplifying consequences. If one considers instead the optimization of entropy production in the

environment, the solution can again be mapped into Burgers equation for some modified space and time variables. The change of variables is different if temperature depends on time only or also on space. It is worth remarking that, operating the change of variables directly on the initial SDE for temperature depending on time only and on space only, the definition of the heat functional is mapped into the definition of entropy production in the environment. These results suggest that the arising of Burgers equation is connected to the optimization of entropy production and solves heat only for the isothermal case in which it differs from entropy production in the environment just by a constant. To conclude the study we suggest some applications of this entropy production optimization. Since it holds for non-constant temperatures it is possible to compute the minimum entropy production for any thermodynamic cycle. We initially show a trivial cycle which displays some non trivial differences from its macroscopic equivalent. Finally we compute the work associated with this minimum entropy cycles and compare it to the optimal work for the isothermal case.

The thesis is structured as follows. The first chapter provides the readers with an outline of stochastic energetics and the optimization method by Bellman. Chapter 2 contains a detailed presentation of the AMM optimization method. Explicit use of its results is made in chapter 3 to solve in an alternative way the optimal transport problems proposed by Schmiedl and Seifert in [2]. In chapter 4 some of the basic features of stochastic motors are presented and the example of a mesoscopic Carnot engine is studied in details. The generalization to non-isothermal settings of the optimization method is developed and presented in chapter 5. The last chapter contains some applications of the generalized method.

Chapter 1

Stochastic thermodynamics and optimal control

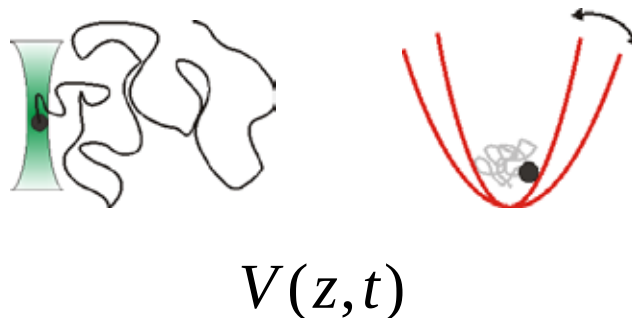
Before starting the actual description of the work it is useful to introduce and frame the basic concepts of stochastic thermodynamics and the optimization techniques that will later be extensively used.

1.1 Introduction to stochastic energetics

The following introduction to stochastic energetics is based on the works by Sekimoto [4],[5] and Ch  trite, Gaw  dzki [6]. The aim of such field is to study the energy exchanges and statistical properties of small systems in analogy with macroscopic equilibrium thermodynamics. Quantities such as mesoscopic internal energy, heat, work and entropy production are therefore introduced. It is important to sketch what is meant by small systems and we refer here to [14]. The distinctive trait of a small system is to involve small amounts of energy (of the order of $10 \div 1000 K_B T$). This feature can be met by systems made of few components (e.g. molecules) and therefore low energy or by considering short time scales for more extensive ones¹. For such settings, fluctuations play a more relevant role than for macroscopic systems. In fact, from statistical mechanics we know that the relative amplitude of fluctuations scales with one over the square root of the number of components ($\mathcal{O}\sqrt{\frac{1}{N}}$). For classical thermodynamic systems, which typically consist of 10^{23} particles, this quantity vanishes. For small systems instead, fluctuations are clearly not negligible. As a consequence, single realizations can differ strongly from mean values (which are the quantities more closely related to the macroscopic ones) and it is important to specify which of them we are considering. The properties and fluctuations of large systems are mostly determined by environmental conditions (pressure, temperature, etc.) whereas the stronger fluctuations of small ones carry some information about the actual structure of

¹Conversely a molecular sized system considered over a very long time can be considered as a macroscopic one.

the system. Studying these extreme fluctuations sheds therefore some light on the arrangement of the system and its non-equilibrium properties. Experiments on small systems can therefore be used to investigate non-equilibrium thermodynamics. A typical experimental set up that has been widely studied displays a single colloidal particle suspended in a solution and controlled by an optical trap (see also chapter 3). It is possible to record the



$$V(z, t)$$

Figure 1.1: Optical tweezers inducing a potential on a mesoscopic particle

motion of the micron sized particle and to measure the amount of energy discharged in the solution (heat). Since the intensity and the focus of the optical trap are determined by the experimenter we also have information on the work applied on the particle.

Biological fluctuating settings can be described as small systems. Molecular motors (see

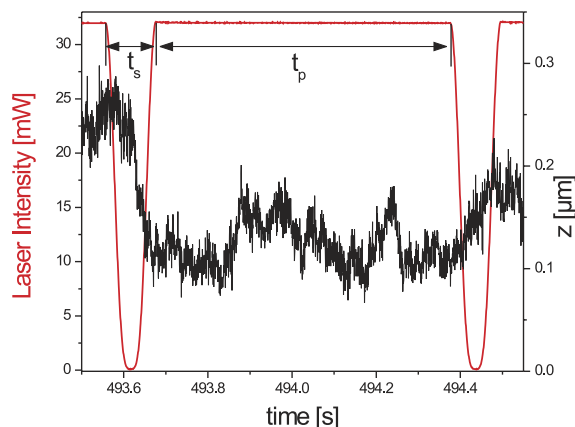


Figure 1.2: Measured tweezers intensity and particle trajectory as in [15]

chapter 4) are in fact small in size (few nanometers) and they operate with small energies (tens of $k_B T$) and extremely short time scales. An experimental study of molecular motors can be performed by controlling the two ends of a DNA chain with optical tweezers and measure the sliding motion of RNA Polymerase (an enzyme responsible for copying the DNA sequence into RNA). Such motion is in fact determined by the action of a molecular motor.

Making use of optical tweezers it is also possible to stretch and unzip the double helix

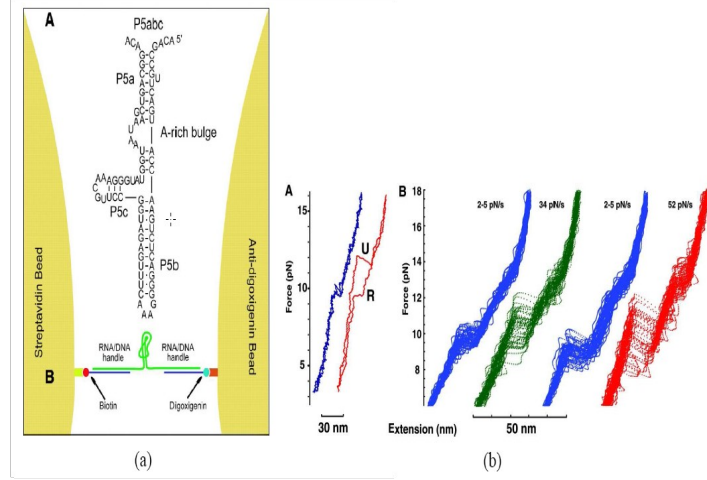


Figure 1.3: (a) Sequence and secondary structure of RNA. (b) Force-extension unfolding curves as depicted in [16]

of DNA molecules or RNA hairpins. Measuring the forces applied through the tweezers it is possible to evaluate the coupling forces between base pairs and the conformational and statistical properties of different sequences.

In order to give a wider overview we choose to present both Sekimoto and Ch  trite-Gaw  dzki approaches. The former provides an intuitive definition of stochastic thermodynamics starting from the physics involved in the one dimension Langevin equation. The latter develops an extremely general mathematical formalism for multidimensional diffusion processes which is more suitable for the generalizations we present in chapter 5 and 6.

1.1.1 Physical perspective

Our starting point is a one dimensional Langevin Equation in the overdamped (i.e. negligible inertia) case.²

$$\frac{dx}{dt} = -\frac{1}{\gamma} \frac{dU}{dx} + \sqrt{\frac{2}{\gamma\beta}} \dot{w}_t \quad (1.1)$$

Where \dot{w}_t is a Gaussian white correlated noise with the following properties $\langle \dot{w}_t \rangle = 0$ and $\langle \dot{w}_t \dot{w}_{t'} \rangle = \delta(t - t')$. β is the inverse temperature and γ the viscous friction. The relation between the factor in front of the noise term (squared root of the diffusion coefficient) and the viscous friction is known as Einstein relation. We will consider it to hold true throughout this work even in non equilibrium settings. However, several studies about fluctuation relations outside of equilibrium have been carried out and we point to [8] for

²The underdamped version can be of interested when considering heat leaking fluxes in temperature changing settings

a review. To proceed in our study it is useful to tell the various components involved in Langevin equation and clearly define their meaning. One way to do it is to consider three of them: the system, the thermal environment and the external system.

- The system is a properly defined part of the world described in terms of the coordinate x . It contributes to the equation with the term $\frac{dx}{dt}$ and its characteristic timescale is the mesoscopic one. In a single colloidal particle experiment would be the particle itself.
- The thermal environment is a sort of heat bath and is fully described by its temperature. It returns instantaneously to equilibrium if perturbed and takes into no account the influences from the systems and the conservation laws for total energy, mass or momentum. It couples with the system through the coefficient γ and the temperature which determine the magnitude of the fluctuating force. Its timescale is the microscopic one. In an experiment would be the solution in which the particle moves.
- The external system is at the macroscopic level and acts on the system by externally controlling the potential energy through some control parameter. In the equation it is responsible for the drift term $-\frac{1}{\gamma} \frac{dU}{dx}$ and can be experimentally realized as an optical trap of variable intensity.

If we consider the energy balance between the three different components we can derive a relation analogous to the first law of thermodynamics. Let us first focus on the interaction between the system and the thermal environment. Through noise, the thermal environment exerts a (random) force on the particle which is then coupled to the environment via the friction term $\gamma \frac{dx}{dt}$. The balance of forces exerted by the thermal environment onto the system is then given by $-\gamma \frac{dx}{dt} + \gamma \sqrt{\frac{2}{\gamma\beta}} \dot{w}_t$. If the law of action reaction holds, the particle reacts with an opposite force to the environment i.e. $-(-\gamma \frac{dx}{dt} + \gamma \sqrt{\frac{2}{\gamma\beta}} \dot{w}_t)$. If we now consider $dx(t)$ as the evolution of the position of the particle during a time dt we can define the infinitesimal work done by the reaction force of the particle on the system. Since we have considered only the “dissipative” terms, i.e. the ones coupling the heat bath to the system we can regard this work as the heat exchanged between the thermal environment and the system in the infinitesimal step.

$$\delta Q = -(-\gamma \frac{dx}{dt} + \gamma \sqrt{\frac{2}{\gamma\beta}} \dot{w}_t) \circ dx \quad (1.2)$$

Where \circ indicates the Stratonovich product. Some attention must be paid to the signs. Here we define heat as the energy damped by the system into the environment. This convention is opposite to the usual thermodynamic one where the δQ is the heat absorbed by the system. Other works adopt different sign conventions (e.g. [5]). It is also interesting to note that rearranging the above formula one can compute the released heat during a

process taking place from time t_0 to t_f

$$Q = - \int_{t_0}^{t_f} \left(-\gamma \frac{dx}{dt} + \gamma \sqrt{\frac{2}{\gamma\beta}} \dot{w}_t \right) \circ \frac{dx}{dt} \circ dt \quad (1.3)$$

Let us now consider the interaction of all three components and their energy balance. From Langevin equation we have

$$-\left(-\gamma \frac{dx}{dt} + \gamma \sqrt{\frac{2}{\gamma\beta}} \dot{w}_t - \frac{dU}{dx} \right) \circ dx = 0 \quad (1.4)$$

In case the potential energy does not depend on any external control parameter the last term in brackets simplifies to the total differential of the potential U . Hence the energy balance now reads

$$\delta Q + dU = 0 \quad (1.5)$$

If we consider also the action of the external system on the potential through a control parameter λ equation (1.4) takes the form

$$-\left(-\gamma \frac{dx}{dt} + \gamma \sqrt{\frac{2}{\gamma\beta}} \dot{w}_t - \frac{\partial U(x, \lambda)}{\partial x} \right) \circ dx = 0 \quad (1.6)$$

To recover the total differential of U , dU , in the formula one needs to add on both sides the quantity $\frac{\partial U(x, \lambda)}{\partial \lambda} d\lambda$:

$$\delta Q + dU = \frac{\partial U}{\partial \lambda} d\lambda \quad (1.7)$$

The quantity $\frac{\partial U(x, \lambda)}{\partial \lambda} d\lambda$ can be defined as the infinitesimal work done by the external system through changes in the potential U by changes of the control parameter λ . Again the sign is opposite to the usual thermodynamic description.

$$\delta W \equiv \frac{\partial U}{\partial \lambda} d\lambda \quad (1.8)$$

Some care is required for the correct understanding of the definition. The work defined above does not represent the work done on the particle by the potential U but the one done by the external agent on the system. In fact if we consider a constant potential applying a constant force f_0 on the particle, the work $f_0 dx$ is done on the particle but according to our definition (1.8) we have no mesoscopic work done by the external system on the system. We count $-f_0 dx$ as part of the change in internal energy $\frac{\partial U(x, \lambda)}{\partial x} \circ dx = dU$. In fact, being in the overdamped case, we consider it as the dissipation of potential energy by the friction due to the motion of the particle and it is therefore taken into account in the heat term. In contrast, if the external system changes the potential, work is done on the system.

With such definitions of heat (1.2) and work (1.8) we are able to express the energy balance in terms of the first law of thermodynamics:

$$dU = \delta W - \delta Q \quad (1.9)$$

1.1.2 Mathematical perspective

The work by Ch  trite and Gaw  dzki [6] deals with general multidimensional diffusion processes and adopts a notation that differs from the one used in the other sections of this thesis. To maintain the generality and make explicit contact with the original paper we keep their notation and present a table of equivalence (tab 1.1) to the one dimension Langevin equation and the formalism used in the other sections.

They define stochastic thermodynamic quantities as specific cases of fluctuation relations for Langevin dynamics. Rather than considering the energy balance to obtain the first law of thermodynamics, they focus on irreversibility and derive the second law. The key quantity defined is therefore entropy production in relation to irreversibility. Such a definition is derived by comparing the expectation values over a stochastic process forward in time with a backward one. Their differences and relations arise because of irreversibility and are therefore a way of quantifying it. For isothermal transformations, heat can then be derived in relation to entropy.

This section aim is to illustrate the line of reasoning behind [6] while listing the main results needed to define entropy production. We start by reporting their notation in the definition of diffusive processes and expectation values and follow their derivation of backward processes. For more rigorous and extensive explanations and proofs refer directly to the paper.

Let us consider the forward process and its related formalism and operators. The starting point is the definition of a generic Stratonovich stochastic differential equation as

$$\dot{x} = u_t(x) + v_t(x) \quad (1.10)$$

where $u_t(x)$ represents the deterministic drift and $v_t(x)$ the time dependent noise term being a Gaussian random vector field with zero mean and $\langle v_t^i(x) v_s^j(y) \rangle = \delta(t - s) D_t^{ij}(x, y)$.

The equivalent of (1.10) in the It   formalism is given by

$$\dot{x} = u_t(x) + \tilde{u}_t(x) + v_t(x) \quad (1.11)$$

Where

$$\tilde{u}_t^i(x) = \frac{1}{2} \partial_{x^j} D_t^{ij}(x, y)|_{x=y} \quad (1.12)$$

is the Stratonovich drift.

As usual, we can define the generator of the process described by the SDE (1.10) and it reads

$$L_t = \hat{u}_t^i \partial_i + \frac{1}{2} \partial_j d_t^{ij} \partial_i \quad (1.13)$$

where

$$d_t^{ij}(x) = D_t^{ij}(x, x) \quad (1.14)$$

is the diffusion matrix and

$$\hat{u}_t^i(x) = u_t^i(x) - \frac{1}{2} \partial_{y^j} D_t^{ij}(x, y)|_{x=y} \quad (1.15)$$

is the modified drift taking care of the position of the diffusion matrix with respect to the two space derivatives. It is worth remarking that

$$\partial_j d_t^{ij} = \partial_{xj} D_t^{ij}(x, x) = \partial_{xj} D_t^{ij}(x, y)|_{x=y} + \partial_{yj} D_t^{ij}(x, y)|_{x=y} \quad (1.16)$$

Let us now consider constrained expectation values of functionals of x_t obeying (1.10) for a process starting in position x at time t_0 . They can be expressed in terms of the transition probabilities of the process. In fact, they write

$$E_{x,t_0} g(x_t) = \int P_{t_0,t}(x, dy) g(y) \equiv (P_{t_0,t} g)(x) \quad (1.17)$$

This relation defines the transition probabilities $P_{t_0,t}(x, dy)$ and the operator $P_{t_0,t}$

$$P_{t_0,t}(x, dy) = E_{x,t_0} \delta(x_t - y) dy \quad (1.18)$$

which satisfy the Chapman-Kolmogorov rule

$$\int P_{t_0,t}(x, dy) P_{t,t'}(y, dz) = P_{t_0,t'}(x, dz) \quad (1.19)$$

The time evolution of the expectation value can be proven to be determined by the generator of the diffusion process introduced in equation (1.13)

$$\frac{d}{dt} E_{x,t_0} g(x_t) = E_{x,t_0} (L_t g)(x_t) \quad (1.20)$$

As a consequence, the Fokker-Planck equation describing the evolution of the probability density for this diffusion process is determined by the adjoint of the generator

$$\partial_t m_t = L_t^\dagger m_t \quad (1.21)$$

Where m is the probability density.

Making use of the definition of the operator $P_{t_0,t}$ (1.17) in relation (1.20) we can obtain a differential equation

$$\partial_t P_{t_0,t} = P_{t_0,t} L_t \quad (1.22)$$

which, with the initial condition $P_{t_0,t_0} = 1$, allows us to write the operator in terms of the time ordered exponential

$$P_{t_0,t} = \overrightarrow{\mathcal{T}} \exp\left[\int_{t_0}^t L_s ds\right] \quad (1.23)$$

As mentioned in the opening lines, we want to quantify irreversibility by relating the characteristics of processes forward and backward in time. To do so we need to define a diffusion process which can be considered as the backward version of the original one. The time inversion of a process taking a time T is the transformation:

$$(t, x) \rightarrow (T - t, x^*) \equiv (t^*, x^*) \quad (1.24)$$

Before proceeding to the inversion it has to be noted that, in dissipative dynamics, the two time directions are not equivalent and special care must be taken. In fact, if one was to naively revert the diffusion processes in time one would face anti-dissipative dynamics which make no physical sense. In other words, dissipation specifies a preferential time direction. To take this into account we split the deterministic drift into two parts: a dissipative one $u_{t,+}$ and a conservative one $u_{t,-}$

$$u_t = u_{t,+} + u_{t,-} \quad (1.25)$$

that behave differently upon time inversion avoiding non-physical processes. If we denote the time reversed process as x'_t we can write its evolution in terms of the reversed SDE

$$\dot{x}' = u'_t(x') + v'_t(x') \quad (1.26)$$

and the different rule of inversion for the two components of the drift

$$\begin{aligned} u'_{t,+}(x) &= \partial_k x^{*i}(x^*) u_{t^*,+}^k(x^*) \\ u'_{t,-}(x) &= -\partial_k x^{*i}(x^*) u_{t^*,-}^k(x^*) \end{aligned} \quad (1.27)$$

avoid the possibility of anti-dissipative dynamics. The sign of the inversion of the random force is not relevant as its covariance is not affected by the choice and

$$D_t^{ij}(x, y) = \partial_k x^{*i}(x^*) D_{t^*}^{kl}(x^*, y^*) \partial_l x^{*j}(x^*) \quad (1.28)$$

The process generator can be defined also for the reversed case and split in a conservative and a dissipative part too

$$L'_t = L'_{t,+} + L'_{t,-} \quad (1.29)$$

where

$$\begin{aligned} L'_{t,+} &= \hat{u}_{t,+}^i \partial_i + \frac{1}{2} \partial_j d_t^{ij} \partial_i \\ L'_{t,-} &= \hat{u}_{t,-}^i \partial_i \end{aligned} \quad (1.30)$$

It is interesting to notice that upon definition of the involution operator R acting on functions

$$Rf(x) = f(x^*) \quad (1.31)$$

we can express the reversed process generator in a clearer fashion

$$L'_t = RL_{t^*,+}R - RL_{t^*,-}R \quad (1.32)$$

The averages taken over the backward process with initial condition $x'_{t_0} = x$ are described by

$$E'_{x,t_0} g(x'_t) = (P'_{x,t_0} g)(x) \quad (1.33)$$

$$P'_{x,t_0} = \overrightarrow{\mathcal{T}} \exp \left[\int_{t_0}^t L'_s ds \right] \quad (1.34)$$

In the abstract mathematical formulation of diffusion processes there are several ways of choosing the splitting between conservative and dissipative drift in (1.25) according to the specific problem addressed. The different possibilities, their meaning and their consequences are precisely described in [6]. For the sake of this thesis we restrict to the study of the overdamped Langevin equation which has a straightforward physical interpretation. Considering the physical meaning of overdamped dynamics, we can take the conservative drift as vanishing and the total drift coincides with the dissipative ones (see table 1.1).

After having defined the backward process, it is necessary for the definition of thermodynamic functionals, to find a way of relating its expectation values with the ones taken over the corresponding forward process. According to the Cameron-Martin-Girsanov theorem it is possible to express averages taken over a perturbed process y_t in terms of the original one x_t weighted with some exponential functional $-\int_{t_0}^t \mathcal{J}_s ds$

$$E_{x,t_0} g(y) = E_{x,t_0} g(x) e^{-\int_{t_0}^t \mathcal{J}_s ds} \quad (1.35)$$

What is then needed is an expression for the generator of the backward process as a perturbed version of the forward one. If this is found, the backward and forward process are related and the functional weight indicates the difference between the two time directions and can be therefore thought as a measure of irreversibility.

If we consider the perturbed version of the generator of the forward process

$$L_t^1 = L_t - 2\hat{u}_{t,+}^i \partial_i - (\partial_i \hat{u}_{t,+}^i) + (\partial_i u_{t,-}^i) \quad (1.36)$$

we notice that is possible to express it in terms of the backward generator as

$$R(L_t^1)^\dagger R = RL_{t,+}R - RL_{t,-}R = L_{t*}' \quad (1.37)$$

so that

$$L_t^1 = (RL_{t*}'R)^\dagger \quad (1.38)$$

In analogy with (1.23), we can write the perturbed evolution operator as

$$P_{t_0,t}^1 = \overrightarrow{\mathcal{T}} \exp\left[\int_{t_0}^t L_s^1 ds\right] = (RP_{t*,t_0}'R)^\dagger \quad (1.39)$$

which in terms of operator kernels yields

$$dx P_{t_0,t}^1(x, dy) = dy P_{t*,t_0}'(y^*, dx^*) \quad (1.40)$$

As already mentioned, through a joint use of the Cameron-Martin-Girsanov theorem and the Feynman-Kac formula cf [6, Appendix D] we can express the kernel $P_{t_0,t}^1(x, dy)$ as an expectation for the forward process exponentially weighted by a functional. Recalling (1.35) and (1.18) we get

$$P_{t_0,t}^1(x, dy) = E_{x,t_0} e^{-\int_{t_0}^t \mathcal{J}_s ds} \delta(x_t - y) dy \quad (1.41)$$

where if $d_t^{ij}(x)$ is invertible for all x and t

$$\mathcal{J}_t = 2\hat{u}_{t,+}(x_t)d_t^{-1}(x_t) \circ \dot{x}_t - 2\hat{u}_{t,+}(x_t)d_t^{-1}(x_t)u_{t,-}(x_t) - (\nabla u_{t,-})(x_t) \quad (1.42)$$

and the first product is meant in the Stratonovich way. Notice that for overdamped Langevin dynamics only $\hat{u}_{t,+}$ is non zero and the expression greatly simplifies (see table 1.1).

Given relation (1.40) between the perturbed process P^1 and the backward one P' we have

$$dx E_{x,t_0} e^{-\int_{t_0}^t \mathcal{J}_s ds} \delta(x_t - y) dy = P'_{t^*,t_0^*}(y^*, dx^*) \quad (1.43)$$

Which is the result we were looking for: an explicit relation between the transition probability of the backward process and the expectation in the forward process weighted with an exponential factor.

In order to investigate the thermodynamic meaning of this functional weight it is useful to introduced its modified version \mathcal{W} . If we express the probability density in terms of exponential as

$$\begin{aligned} m(x, 0) &= e^{-\varphi_0(x)} \\ m(x, T) &= e^{-\varphi_T(x)} \end{aligned} \quad (1.44)$$

associated to the time inversed ones

$$\begin{aligned} m'(x, 0) &= e^{-\varphi'_0(x)} = e^{-\varphi_T(x^*)} \\ m'(x, T) &= e^{-\varphi'_T(x)} = e^{-\varphi_0(x^*)} \end{aligned} \quad (1.45)$$

we can introduce a new functional

$$\mathcal{W} = \Delta\varphi + \int_0^T \mathcal{J}_t dt = \log m(x_0) - \log m(x_T) + \int_0^T \mathcal{J}_t dt \quad (1.46)$$

Such a functional leads to generalizations of equation (1.43) which can be used to derive Jarzynski equality. For the derivation we refer to [6] and list here the main results. It can be proven that for a generic functional $\mathcal{F}(x_t)$

$$\langle \mathcal{F}(x_t) e^{-\mathcal{W}} \rangle = \langle \mathcal{F}(\tilde{x}_t) \rangle' \quad (1.47)$$

where $\tilde{x}_t = x_{t^*}^*$ is the reversed trajectory for the reversed time of the backward process and $\langle \rangle'$ denotes an average over the reversed probability measure.

In case $\mathcal{F}(x_t) = 1$ we obtain the equality

$$\langle e^{-\mathcal{W}} \rangle = 1 \quad (1.48)$$

which, for some specifications of \mathcal{W} , such as the one in Langevin dynamics for which $\mathcal{W} = \beta(W - \Delta F)$, yields Jarzynski equality

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F} \quad (1.49)$$

From equation (1.48) and the Jensen inequality follows also

$$\langle \mathcal{W} \rangle \geq 0 \quad (1.50)$$

The averaged functional $\langle \mathcal{W} \rangle$ can be shown to be the relative entropy of the trajectory measure of the backward process with respect to the forward one. In fact the relative entropy of a probability measure $\nu(dx)$ with respect to an other one $\mu(dx)$ is defined as

$$S(\mu|\nu) = \int \log \frac{\nu(dx)}{\mu(dx)} \mu(dx) \quad (1.51)$$

and, since equation (1.47) can be seen as a relation between the probability measures M of the forward and backward process

$$e^{-\mathcal{W}} M(dx) = M'(d\tilde{x}) \quad (1.52)$$

we can now see that upon definition of $M'(d\tilde{x}) = \tilde{M}'(dx)$

$$S(\tilde{M}'|M) = \int \mathcal{W} \mu(dx) = \langle \mathcal{W} \rangle \quad (1.53)$$

If we now consider the probability density $m(x, T)$ as the evolution corresponding to $m(x, 0)$ at time T , the relative entropy takes the meaning of global entropy production during the forward process. The positivity of $\langle \mathcal{W} \rangle$ states now the positivity of the global entropy production and therefore inequality (1.50) is a version of the second law of thermodynamics. In order to get a deeper thermodynamic understanding of the terms involved let us write explicitly the average of \mathcal{W} from its definition of equation (1.46)

$$\langle \mathcal{W} \rangle = \langle \log m(x_0) \rangle - \langle \log m(x_T) \rangle + \left\langle \int_0^T \mathcal{J}_t dt \right\rangle \quad (1.54)$$

By definition $-\langle \log m(x_T) \rangle$ is the entropy associated with a probability distribution $m(x_T)$ and therefore the first two terms represent the change of entropy of the fixed time distribution along the process. Since the total sum yields the global entropy production the term

$$\Delta S_{env} = \int_0^T \langle J_t \rangle dt \quad (1.55)$$

must account for the mean entropy production in the environment related to the irreversibility of the process. We can then write

$$\langle \mathcal{W} \rangle = \sigma(\mu_T) - \sigma(\mu_0) + \Delta S_{env} \quad (1.56)$$

In case of an isothermal process we can express heat in terms of the entropy production from standard thermodynamic reasoning

$$\frac{1}{\beta} \int_{t_0}^t \mathcal{J}_s ds = \int_{t_0}^t \frac{1}{\beta} \mathcal{J}_s ds = Q \quad (1.57)$$

For the overdamped Langevin equation of (1.1) the functional reduces to

$$\mathcal{J}_t = (-\beta \partial_x U_t)(x_t) \circ \dot{x}_t \quad (1.58)$$

and consequently heat becomes

$$Q = \int_{t_0}^{t_f} -(\partial_x U_t) \circ \dot{x}_t dt \quad (1.59)$$

which is equal to (1.3) if we consider that from Langevin equation (1.1) $\frac{1}{\gamma} \frac{dU}{dx} = -\frac{dx}{dt} + \sqrt{\frac{2}{\gamma\beta}} \dot{w}_t$. It is remarkable as, through a completely different reasoning from Sekimoto the same result is recovered.

1.2 Introduction to Bellman equation

We briefly sketch the concepts underlying dynamic programming and Bellman equation in the field of optimal control. To illustrate the basic ideas we focus on the deterministic case and refer to [7] for a thorough exposition and the study of the stochastic case. The problem consists of finding the control to be operated on a system in order to minimize a suitably defined cost function. Typically the evolution of the system is described by a differential equation involving the external control b_t

$$\frac{dx}{dt} = f_t(x, b) \quad (1.60)$$

In the absence of a terminal cost, the cost function is defined as

$$J_t(x, b) = \int_t^{t_f} L_s(x, b) ds \quad (1.61)$$

where $L_s(x, b)$ is the running cost. The intuition of dynamic programming is to decompose the control interval in sub-intervals and to consider the last one as already optimal. The optimization then reduces to considering only the first step and is therefore simplified. An iteration of the procedure allows to optimize the whole control interval. To formalize the idea it is fundamental to introduce the concept of value function.

$$V_t(x) = \min_b \{J_t(x, b)\} \quad (1.62)$$

The value function at time t and position x is the minimum cost achievable by any possible control needed to run the process from time t to the end of the control interval. It therefore corresponds to the cost function of the process having already been optimized. Given the value function we want to compute the optimal cost of the whole process. Since the global cost consists of the sum of the costs of the sub intervals we can write for $t' < t$

$$J_{t'}(y, b) = \int_{t'}^t L_s(y, b) ds + \int_t^{t_f} L_s(x, b) ds \quad (1.63)$$

The corresponding value function can be proven to be

$$V_{t'}(y) = \min_b \left\{ \int_{t'}^{t_f} L_s(y, b) ds \right\} + V_t(x) \quad (1.64)$$

and its expression is called Bellman equation. We are therefore able to compute the optimal cost of the process starting at time t' . It is evident that the optimization concerns now only the first term and is clearly simpler. If we iterate the process we are then able to compute the minimum cost of the whole process. Moreover, assuming the value function to be differentiable we can consider an infinitesimal time step and obtain the dynamic programming equation as

$$\partial_t V_{t'}(x) + \min_b \{ L_t(x, b) + f_t(x, b) \partial_x V_t(x) \} = 0 \quad (1.65)$$

Object	CG formalism	Thesis formalism and 1D Langevin equivalent
Deterministic drift	$u_t^i(x)$	$\frac{b_t(x)}{\gamma} = -\frac{1}{\gamma} \frac{\partial U}{\partial x}$
Noise term	$v_t^i(x)$	$\sqrt{\frac{2}{\gamma\beta_t(x)}} \dot{w}_t$
Diffusion matrix	$D_t^{ij}(x, y)$	$\frac{2}{\gamma} \sqrt{\frac{1}{\beta_t(x)\beta_s(y)}}$
Diffusion derivative	$\partial_j d^{ij} = \partial_{x^j} D_t^{ij}(x, x)$	$\frac{2}{\gamma} \partial_x \left(\frac{1}{\beta_t(x)} \right)$
Noise induced drift	$\tilde{u}_t^i = \frac{1}{2} \partial_{x^j} D_t^{ij}(x, y) _{x=y}$	$\frac{1}{2\gamma} \partial_x \left(\frac{1}{\beta_t(x)} \right)$
Modified drift	$\hat{u}_t^i(x) = u_t^i(x) - \frac{1}{2} \partial_{y^j} D_t^{ij}(x, y) _{x=y}$	$\frac{b_t(x)}{\gamma} - \frac{1}{2\gamma} \partial_x \left(\frac{1}{\beta_t(x)} \right)$
Diffusion generator	$L_t = u_t^i(x) \partial_i + \frac{1}{2} \partial_j d_t^{ij} \partial_i$	$\frac{1}{\gamma} [b_t(x) + \frac{1}{2} \partial_x \left(\frac{1}{\beta_t(x)} \right)] \partial_x + \frac{1}{2\gamma\beta_t(x)} \partial_x^2$
Transition probability	$P_{t_0,t}(x, dy)$	$P(y, t x, t_0) dy$
Expectation value	$Eg(x_t)$	$\langle g(x_t) \rangle$
Conservative part of drift	$u_{t,-}(x)$	0
Dissipative part of drift	$u_{t,+}(x)$	$\frac{b_t(x)}{\gamma}$
Instantaneous entropy production	\mathcal{J}_t	$\beta b \circ \dot{x}_t$

Table 1.1: Table of equivalence between the general formalism of [6] and the specific case considered in other sections

Chapter 2

Stochastic control and optimal transport

In this chapter we are going to carefully analyze the results presented in [1] and describe them from a slightly different point of view. Let us consider a stochastic process transporting a mesoscopic particle. The transport is induced by varying the potential trapping the particle. From the definitions of section 1.1.2 we can see that the heat released and the work needed along the process obviously depends on the potential which is driving it. The general aim of the paper is to individuate the external control on the potential which optimizes these thermodynamic quantities. This would provide information on which is the way to transport a particle spending the least work possible. Instead of minimizing the functionals via Euler Langrange they make use of the Bellman equation.

2.1 Optimization procedure

Let us start by expressing the drift as $b_t = -\partial_x U_t$ so that (1.1) becomes

$$\dot{x}_t = \frac{1}{\gamma} b_t + \sqrt{\frac{2}{\gamma\beta}} \dot{w}_t \quad (2.1)$$

where w_t is a Wiener process with zero mean and variance equal to dt . In this chapter we restrict to constant values of the diffusion coefficient and postpone the study of space and time dependent ones to chapter 5.

From definition (1.59) we can write an expression for the heat released from the system into the thermal environment during a process starting at time t_0 and ending at time t_f as

$$\delta Q = \int_{t_0}^{t_f} b_t \circ \dot{x}_t dt \quad (2.2)$$

Similarly, we can get the expression for work needed for a process taking place between time t_0 and t_f from equation (1.8)

$$\delta W = \int_{t_0}^{t_f} \frac{\partial U(x, \lambda_t)}{\partial \lambda} \cdot \frac{d\lambda}{dt} dt \quad (2.3)$$

Given the stochastic nature of the process, single trajectory thermodynamic values fluctuate widely and we therefore focus on average quantities. We can define the average of a generic quantity G , over the process x_t as

$$\langle G_t(x_t) \rangle = \int dx m_t(x) G_t(x) = E G_t(x_t) \quad (2.4)$$

where $m_t(x)$ is a probability density evolving according to the Fokker Planck equation corresponding to equation (1.1). For the study of mean values it is more suitable to express the stochastic quantities in terms of Itô integrals. In fact their non anticipating properties simplify greatly the computation (see for example [11]). The thermodynamics quantities derived in section 1.1 are defined in terms of Stratonovich integrals and we therefore need to translate them into the Itô formalism.

Let us now consider the definition of heat (2.2). The fact that the integral is defined in the sense of Stratonovich means

$$\int \dot{x}_t \cdot b_t dt = \int [x_{t+dt} - x_t] \left[\frac{b_{t+dt}(x) + b_t(x)}{2} \right] \quad (2.5)$$

To convert it to an Itô formulation let us evaluate all terms at time t and get rid of the anticipativity of the process. We can start by Taylor expanding b about x

$$\frac{b_{t+dt}(x) + b_t(x)}{2} \simeq b_t(x) + \frac{1}{2} \partial_x b_t(x) \cdot (x_{t+dt} - x_t) \quad (2.6)$$

Now let us consider the full product in (2.5) making use of (1.1).

$$\begin{aligned} & \left[\frac{b}{\gamma} dt + \sqrt{\frac{2}{\gamma\beta}} dw_t \right] \left[b_t(x) + \frac{1}{2} \partial_x b_t(x) \cdot (x_{t+dt} - x_t) \right] = \\ &= \left[\frac{b}{\gamma} dt + \sqrt{\frac{2}{\gamma\beta}} dw_t \right] \left[b_t(x) + \frac{1}{2} \partial_x b_t(x) \cdot \left(\frac{b}{\gamma} dt + \sqrt{\frac{2}{\gamma\beta}} dw_t \right) \right] = \\ &= \frac{b^2}{\gamma} dt + \frac{b}{\gamma} \partial_x b \frac{b}{2\gamma} (dt)^2 + \frac{b}{2\gamma} \partial_x b \sqrt{\frac{2}{\gamma\beta}} dw_t dt + \\ &+ b \sqrt{\frac{2}{\gamma\beta}} dw_t + \frac{b}{2\gamma} \partial_x b \sqrt{\frac{2}{\gamma\beta}} dw_t dt + \frac{2}{2\gamma\beta} \partial_x b (dw_t)^2 \end{aligned} \quad (2.7)$$

We can now take the average and should note that for non anticipative processes the functions are independent of the time increments and we can factorize them in separate averages. For example, for the products involving dw_t we have

$$\langle b \sqrt{\frac{2}{\gamma\beta}} dw_t \rangle = \langle b \sqrt{\frac{2}{\gamma\beta}} \rangle \langle dw_t \rangle \quad (2.8)$$

and we can exploit the fact that by definition

$$\begin{aligned} \langle dw_t \rangle &= 0 \\ \langle (dw_t)^2 \rangle &= dt \end{aligned} \quad (2.9)$$

Dropping the terms of order higher than dt and keeping in mind that $dw_t \sim O(\sqrt{dt})$ we obtain the expected amount of released heat in the Itô convention.

$$\langle \delta Q \rangle = \int_{t_0}^{t_f} \frac{dt}{\gamma} \langle b_t^2 + \frac{1}{\beta} \partial_x b_t \rangle \quad (2.10)$$

In order to cast the optimization of released heat in terms of Bellman equation it is useful to introduce the control functional $S_t(x) = \langle \delta Q \rangle_{x,t} = E_{x,t} Q$. This is the expected heat released, not during the whole process, but starting from an intermediate time t in position x and evolving to t_f .

Let us write explicitly the expected released heat over a random process representing our global cost

$$\langle \delta Q \rangle = E[\delta Q] = \int E_{x_0, t_0} \left[\int_{t_0}^{t_f} \frac{dt}{\gamma} (b_t^2 + \frac{1}{\beta} \partial_x b_t) \right] m_{t_0}(x_0) dx_0 \quad (2.11)$$

Where it is more evident that it is an average over the possible positions at initial time weighed with the appropriate probability density. The control functional (the cost of the operation from an intermediate time t to the end) is the conditional expectation

$$S_t(x) = E_{x,t} \left[\int_t^{t_f} \frac{dt}{\gamma} (b_t^2 + \frac{1}{\beta} \partial_x b_t) \right] \quad (2.12)$$

And we can introduce the state averaged control functional

$$\bar{S}_t = \int dx m_t(x) S_t(x) \quad (2.13)$$

Where obviously $\bar{S}_{t_0} = \langle \delta Q \rangle$. The control functional $S_t(x)$ is suitable for a study along the line of Bellman principle of optimality (cf section 1.2). In fact, we are going to consider the process starting at time t as optimized, so that $S_t(x)$ plays the role of the value function (1.62), and propagate it backwards in time by a step dt . We then optimize over the step we have taken and iterate the process. To carry out the optimization of the step we need to consider a process starting from position y at time $t - dt$ and passing through x at time t .

$$\begin{aligned} S_{t-dt}(y) &= E_{y,t-dt} \left[\int_{t-dt}^{t_f} \frac{dt'}{\gamma} (b_{t'}^2(x) + \frac{1}{\beta} \partial_x b_{t'}(x)) \right] = \\ &E_{y,t-dt} \left[\int_{t-dt}^t \frac{dt'}{\gamma} (b_{t'}^2(x) + \frac{1}{\beta} \partial_x b_{t'}(x)) + \int_t^{t_f} \frac{dt'}{\gamma} (b_{t'}^2(x) + \frac{1}{\beta} \partial_x b_{t'}(x)) \right] \end{aligned} \quad (2.14)$$

The first term between braces features the expectation value of an integral over an infinitesimal interval constrained to the initial position. We can approximate it with the value of the integrand function at the initial position y multiplied by dt since the other contributions will be of higher order.

$$E_{y,t-dt} \left[\int_{t-dt}^t \frac{dt'}{\gamma} (b_{t'}^2(x) + \frac{1}{\beta} \partial_x b_{t'}(x)) \right] \simeq \frac{dt}{\gamma} (b^2(y) + \frac{1}{\beta} \partial_x b(y)) \quad (2.15)$$

For the second term in (2.14) we can make use of the relations between propagators and expectation values (1.17) and Chapman-Kolmogorov property (1.19)

$$\begin{aligned}
E_{y,t-dt}[\int_t^{t_f} \frac{dt'}{\gamma} (b_{t'}^2(x) + \frac{1}{\beta} \partial_x b_{t'}(x))] &= \\
= \int dx' dx P(x, t|y, t-dt) P(x', t'|x, t) \int_t^{t_f} \frac{dt'}{\gamma} (b_{t'}^2(x) + \frac{1}{\beta} \partial_x b_{t'}(x)) &= \\
= \int dx P(x, t|y, t-dt) E_{x,t}[\int_t^{t_f} \frac{dt'}{\gamma} (b_{t'}^2(x) + \frac{1}{\beta} \partial_x b_{t'}(x))] &= \\
= \int dx P(x, t|y, t-dt) S_t(x) & \quad (2.16)
\end{aligned}$$

This shows how (2.14) is an iterative equation and can therefore be later used to derive Bellman equation. To compute the value of (2.16) we need an approximate expansion of the propagator for small times. Since the dynamics follow equation (2.1) and we know that dw is a white random Gaussian term with zero mean and variance dt we can obtain the approximate expansion

$$P(x, t|y, t-dt) \propto \exp[-\frac{1}{2dt} \frac{\beta\gamma}{2} (x_t - y_{t-dt} - \frac{b_t dt}{\gamma} + \frac{\partial_x b_t dx dt}{2\gamma})^2] \quad (2.17)$$

For small time steps we can stop at a first order approximation, neglecting the term $\partial_x b dx dt$. This corresponds to considering the drift as time-independent over small time intervals and therefore a Gaussian propagator. Since the propagator represents the transition probability from state y we need to express $S_t(x)$ in terms of $S_t(y)$ and this can be done by Taylor expanding around y .

$$\begin{aligned}
\int dx P(x, t|y, t-dt) S_t(x) &\simeq \\
\simeq \int dx [S_t(y) + \partial_x S(y) \cdot (x-y) + \frac{1}{2} (x-y) \partial_x^2 S(y) \cdot (x-y)] e^{-\frac{\beta\gamma}{4dt} (x-y-\frac{bdt}{\gamma})^2} &= \\
= S_t(y) + \frac{dt}{\gamma} b \partial_x S(y) + \frac{dt}{\beta\gamma} \partial_x^2 S(y) & \quad (2.18)
\end{aligned}$$

Where the last step exploited the fact that the propagator is a normalized Gaussian with defined mean and variance. Adding this expression to (2.15) and dropping the explicit space dependence gives

$$S_{t-dt} - S_t = dt(\frac{b}{\gamma} \partial_x S + \frac{1}{\beta\gamma} \partial_x^2 S + \frac{b^2}{\gamma} + \frac{\partial b}{\beta\gamma}) \quad (2.19)$$

Leading to

$$-\partial_t S - \frac{b}{\gamma} \partial_x S - \frac{1}{\beta\gamma} \partial_x^2 S = \frac{b^2}{\gamma} + \frac{\partial b}{\beta\gamma} \quad (2.20)$$

Which is the Hamilton-Jacobi-Bellman (HJB) equation for the problem. The left hand side displays the adjoint of the Fokker Planck operator describing the backward evolution of the process. This backward time direction accounts for the time derivative and the diffusion term ∂_x^2 having the same sign. The right hand side is instead the running cost of the considered problem. In the search for an optimal functional, one can always derive a HJB equation but, of course, different evolution operators and different running costs result in different HJB equations. When considering a diffusion process the specific form of the Langevin equation determines its related Fokker Planck operator, therefore changes in its form result in changes of the left hand side of equation (2.20).

We are now interested in finding the optimal drift which minimizes the released heat. Carrying on along the line of Bellman optimality principle, we split the control interval. Let us consider the state averaged control \bar{S}_{t-dt} . We assume to know the control b' optimizing the process from time t to the final time t_f so that we can consider it as the value function (1.62). To optimize \bar{S}_{t-dt} we then have to find the control b optimizing the step from $t-dt$ to t . The process in the interval between t_0 and $t-dt$ is governed by a yet unspecified control b'' .

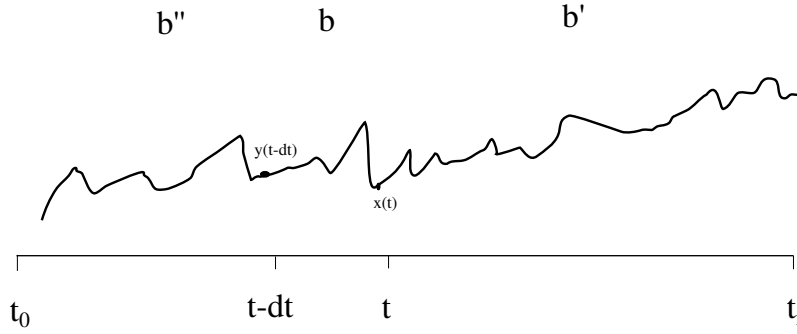


Figure 2.1: Splitting of the control interval for a generic stochastic process

We therefore obtain Bellman equation for the value function \bar{S}

$$\begin{aligned} \bar{S}_{t-dt} &= \int dy m_{t-dt}(y) S_{t-dt}(y) = \\ &dt \int dy \left(\frac{b^2}{\gamma} + \frac{\partial b}{\beta \gamma} \right) m_{t-dt}(y) + \int dy m_{t-dt}(y) \int dx P(x, t|y, t-dt) S_t(x) \end{aligned} \quad (2.21)$$

Since in the second term we are integrating over all possible starting states of the propagator we can directly change the probability density to that of the arrival state

$$\int dy m_{t-dt}(y) \int dx P(x, t|y, t-dt) S_t(x) = \int dx m_t(x) S_t(x) \quad (2.22)$$

and Taylor expanding m in time around $t-dt$ we get to the expression

$$m_t(x) \simeq m_{t-dt}(x) + dt \partial_t m_{t-dt}(x) \quad (2.23)$$

so that

$$\begin{aligned} \bar{S}_{t-dt} = & \frac{dt}{\gamma} \int dx (b^2 + \frac{\partial b}{\beta}) m_{t-dt}(x) + \\ & \int dx m_{t-dt}(x) S_t(x) + dt \int \partial_t m_{t-dt}(x) S_t(x) dx \end{aligned} \quad (2.24)$$

We note that in the first integral in the second line there is no dependence on b since the probability density up to time $t - dt$: $m_{t-dt}(x)$ depends on the parameter b'' and that $S_t(x)$ depends on b' . We denote it as A and remark that it is not relevant to the search for the optimal b . We now express the time derivative of m through the Fokker Planck equation corresponding to equation (1.1),

$$\partial_t m = -\frac{\partial_x(bm)}{\gamma} + \frac{1}{\beta\gamma} \partial_x^2 m \quad (2.25)$$

obtaining

$$\begin{aligned} \bar{S}_{t-dt} = & A + dt \int dx (+\frac{b^2}{\gamma} + \frac{\partial_x b}{\beta\gamma}) m_{t-dt}(x) + \\ & + dt \int S_t(x) (-\frac{\partial_x[bm_{t-dt}(x)]}{\gamma} + \frac{1}{\beta\gamma} \partial_x^2 m_{t-dt}(x)) dx \end{aligned} \quad (2.26)$$

Integrating the last term by parts we move the derivatives to $S_t(x)$, changing the sign of the first order one. We then get to the following expression

$$\bar{S}_{t-dt} = A + \frac{dt}{\gamma} \int dx (b^2 + \frac{\partial_x b}{\beta} + \frac{1}{\beta} \partial_x^2 S + b \partial_x S) m_{t-dt}(x) \quad (2.27)$$

To find the optimal drift it is now fundamental to use the explicit form of the functional we want to optimize. In fact we now differentiate with respect to b and we would get a different optimal drift for each different functional. Let us notice that since we have restricted the optimization to this infinitesimal interval the derivative over b does not indicate a functional derivative but the differentiation over the value of b in the interval.

$$\frac{\partial \bar{S}_{t-dt}}{\partial b} = \frac{dt}{\gamma} \int dx (2b + \frac{\partial}{\partial b} (\frac{\partial_x b}{\beta}) + \partial_x S) m_{t-dt}(x) \quad (2.28)$$

We do not get any contribution from the differentiation of the probability density $m_{t-dt}(x)$ as the process, up to $t - dt$, is controlled by b'' and therefore independent of b . The term in ∂b needs to be integrated by parts yielding:

$$\frac{\partial \bar{S}_{t-dt}}{\partial b} = \frac{dt}{\gamma} \int dx (2b - \frac{1}{\beta} \frac{\partial_x m}{m} + \partial_x S) m_{t-dt}(x) \quad (2.29)$$

And introducing the potential corresponding to equilibrium part of m as

$$R = \frac{1}{\beta} \log m \quad (2.30)$$

we finally get

$$\frac{\partial \bar{S}_{t-dt}}{\partial b} = \frac{dt}{\gamma} \int dx (2b - \partial_x R + \partial_x S) m_{t-dt}(x) \quad (2.31)$$

Hence the optimal value of b resulting in a minimum released heat is

$$b^* = \frac{1}{2} \partial_x (R - S) \quad (2.32)$$

If we now plug this optimal drift into the Fokker Planck equation (2.25) we get

$$\partial_t m = -\partial_x \left[\frac{\partial_x (R - S)}{2\gamma} m \right] + \frac{1}{\beta\gamma} \partial_x^2 m \quad (2.33)$$

For a generic Fokker Planck equation it is always possible to split the drift into an equilibrium part and remainder one and recover the deterministic transport equation with the remainder part as a velocity. Explicitly, splitting the drift into

$$b = \partial_x R + \partial_x \psi \quad (2.34)$$

The Fokker Planck now reads

$$\begin{aligned} \partial_t m + \partial_x \left[\frac{\partial_x (R + \psi)}{\gamma} m \right] - \frac{1}{\beta\gamma} \partial_x^2 m = \\ \partial_t m + \partial_x^2 \frac{(R + \psi)}{\gamma} m + \frac{\partial_x (R + \psi)}{\gamma} \partial_x m - \frac{1}{\beta\gamma} \partial_x^2 m \end{aligned} \quad (2.35)$$

and considering that

$$\partial R = \frac{1}{\beta} \frac{\partial m}{m} \quad (2.36)$$

$$\partial_x^2 R = \frac{1}{\beta} \left(\frac{\partial_x^2 m}{m} - \left(\frac{\partial m}{m} \right)^2 \right) \quad (2.37)$$

We see that the derivative of the drift cancels the diffusion term and we get to

$$\partial_t m + \frac{1}{\beta\gamma} \partial_x^2 m - \frac{1}{\gamma\beta} \frac{\partial_x m}{m} + \frac{1}{\gamma} \partial_x [(\partial_x \psi) m] + \frac{1}{\gamma\beta} \frac{\partial_x m}{m} = \frac{1}{\beta\gamma} \partial_x^2 m \quad (2.38)$$

Which after straightforward simplifications reduces to the deterministic transport equation in the gradient of the remainder term of the drift ψ :

$$\partial_t m + \frac{1}{\gamma} \partial_x [(\partial_x \psi) m] = 0 \quad (2.39)$$

What happens to the Hamilton-Jacobi-Bellman equation (2.20) for the optimal drift (2.32) upon substitution of the splitting (2.34) is instead characteristic of this optimization.

$$\partial_t S + \partial_x S \frac{\partial_x(R + \psi)}{2\gamma} + \partial_x^2 S \frac{1}{\beta\gamma} = -\frac{[\partial_x(R + \psi)]^2}{4\gamma} - \frac{\partial_x^2(R + \psi)}{2\beta\gamma} \quad (2.40)$$

Equation (2.32) enforces a relation between the functional S and the optimal drift. By splitting the drift, we can express the functional S in terms of the equilibrium potential R and ψ

$$S = -(2\psi + 2R - R) = -2\psi - R \quad (2.41)$$

where the relation holds for the optimal drift only. Plugging it in the HJB (2.40) we get

$$\begin{aligned} -2\partial_t \psi - \partial_t R - \partial_x \frac{(\psi + R)}{\gamma} \partial_x(2\psi + R) - \frac{\partial_x^2(2\psi + R)}{\beta\gamma} = \\ -\frac{[\partial_x(R + \psi)]^2}{\gamma} - \frac{\partial_x^2(\psi + R)}{\beta\gamma} \end{aligned} \quad (2.42)$$

Which expanding and simplifying yields

$$\begin{aligned} -2\partial_t \psi - \partial_t R - 2\frac{\partial_x \psi \partial_x \psi}{\gamma} - 2\frac{\partial_x \psi \partial_x R}{\gamma} - \frac{\partial_x \psi \partial_x R}{\gamma} - \\ \frac{\partial_x R \partial_x R}{\gamma} - \frac{\partial_x^2 \psi}{\beta\gamma} = -\frac{(\partial_x R)^2}{\gamma} - \frac{(\partial_x \psi)^2}{\gamma} - 2\frac{\partial_x \psi \partial_x R}{\gamma} \end{aligned} \quad (2.43)$$

Which after immediate simplification becomes

$$-2\partial_t \psi - \partial_t R - \frac{\partial_x \psi \partial_x \psi}{\gamma} - \frac{\partial_x \psi \partial_x R}{\gamma} = \frac{\partial_x^2 \psi}{\beta\gamma} \quad (2.44)$$

If we now express explicitly R in terms of m

$$\begin{aligned} -2\partial_t \psi - \frac{\partial_t m}{\beta m} - \frac{[\partial_x \psi]^2}{\gamma} - \frac{\partial_x \psi \partial_x m}{\beta\gamma m} - \frac{\partial_x^2 \psi}{\beta\gamma} = \\ -2\partial_t \psi - \frac{[\partial_x \psi]^2}{\gamma} - \frac{1}{\beta m} (\partial_t m + \frac{\partial_x \psi \partial_x m}{\gamma} + \frac{\partial_x^2 \psi}{\beta\gamma} m) = 0 \end{aligned} \quad (2.45)$$

Which includes the Fokker Planck equation (2.39) and therefore reduces to

$$\partial_t \psi + \frac{[\partial_x \psi]^2}{2\gamma} = 0 \quad (2.46)$$

That, surprisingly, is the inviscid Burgers equation for the velocity potential $\frac{\psi}{\gamma}$.

$$\partial_t v + (v \partial_x) v = 0 \quad (2.47)$$

On the other hand equation (2.39) is the equation of mass transport by the corresponding velocity field. The complicated original optimization problem has been reduced to the solution of Burgers equation for which several numerical methods already exist. The forward Fokker-Planck evolution determines the initial conditions while the backward evolution derived from Bellman equations enforces final ones. It is worth remarking that optimization via a Bellman-like procedure is possible for a variety of functionals but only in specific cases does one recover an auxiliary problem reducing to Burgers equation.

2.2 Solution of Burgers equation

It is noteworthy that, since the Burgers equation in (2.46) is the result of mixed forwards-backwards dynamics, it is not reasonable to regularize the problem by adding $\nu \partial_x^2 \psi$ or $-\nu \partial_x^2 \psi$ since choosing one of the two would mean choosing a direction in time. Without shocks the solutions of Burgers equation describe a free-streaming motion. This can be easily seen by considering the velocity as the derivative of the potential velocity field $v_i = \frac{1}{\gamma} \partial_i \psi$. Differentiating Burgers equation we get

$$\frac{1}{\gamma} \partial_i (\partial_t \psi - \frac{[\partial_{x_j} \psi]^2}{2\gamma}) = \partial_t v_i + \frac{1}{\gamma^2} \frac{\partial \psi}{\partial x_j} \frac{\partial^2 \psi}{\partial x_i \partial x_j} \quad (2.48)$$

And since ψ is a potential field we can commute the order of the derivatives getting to

$$\partial_t v_i + \frac{1}{\gamma} \frac{\partial^2 \psi}{\partial x_i \partial x_j} v_j = \partial_t v_i + \frac{\partial v_i}{\partial x_j} v_j \quad (2.49)$$

From equation (2.49) one can see that the convective derivative of the velocity is zero. This means that the solution will be a straight line and that the evolution is linear in time. Once on a streamline a particle will not leave it. It is therefore possible to define a Lagrangian map connecting the initial position of each particle with its final one. Denoting as $a = x_{t_i}$ the initial position.

$$x_{t_f}(a) = a + (t_f - t_i) v_{t_i}(a) \quad (2.50)$$

The fact that every particle is associated with one single velocity line reduces the problem of reconstructing the dynamics associated with the transport equation (2.39) to the one of assigning the initial positions to the final ones [13]. Such an interpretation points towards a connection with the known optimal transport problem proposed by Monge. The problem dates back to 1781 and originated from civil engineering. It consists of finding the optimal way of carrying earth from cuts to fills.

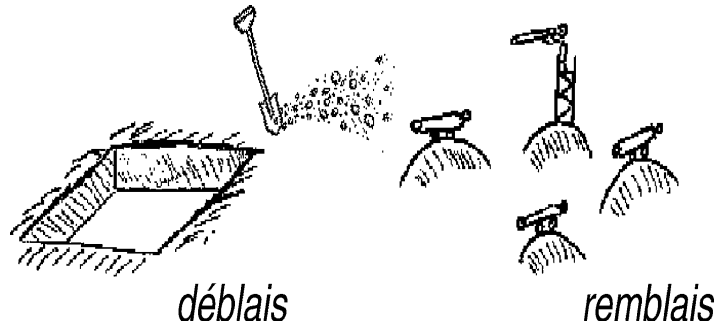


Figure 2.2: Cartoon of Monge problem as depicted in [13]

It can be shown [12] that the solutions of Burgers equation correspond to the solutions of Monge problem in case its cost function is quadratic in the distance between initial and

final positions

$$K = \gamma \int \frac{(a - x_f)^2}{(t_f - t_i)} m_{t_i}(a) da \quad (2.51)$$

constrained with the evolution of probability (2.39) and mass conservation. A variational analysis of the cost shows that it is minimum if the velocity is a gradient. So that again

$$v = \frac{1}{\gamma} \nabla \psi \quad (2.52)$$

and therefore

$$x_{t_f}(a) - a = \frac{(t_f - t_i)}{\gamma} \nabla \psi_{t_i}(a) \quad (2.53)$$

Which, means

$$\psi_{t_f}(x_f) = \psi_{t_i}(a) + \gamma \frac{(a - x_f)^2}{2(t_f - t_i)} \quad (2.54)$$

In more than one dimension this is described by

$$\Psi_{t_i}(a) = \frac{1}{2} a^2(x) - \frac{(t_f - t_i)}{\gamma} \psi_{t_f}(x_f) \quad (2.55)$$

$$a = \nabla \Psi_{t_i}(a) \quad (2.56)$$

Let us remark that mass conservation

$$\begin{aligned} m_{t_i}(a) da &= m_{t_f}(x_f) dx_f \\ M_i(a) &= M_f(x) \end{aligned} \quad (2.57)$$

implies that the cumulative mass functions are equal. As a consequence, solutions of the assignment problem solve Monge-Ampere equation

$$\det \frac{\partial a}{\partial x_f} = \det \frac{\partial^2 \Psi_{t_i}(a)}{\partial x^j \partial x^i} = \frac{m_{t_f}(x_f)}{m_{t_i}(a)} \quad (2.58)$$

The positivity of the probability densities results in the Hessian of $\Psi_{t_i}(a)$ being everywhere positive and therefore the function being convex. If we now interpolate it between initial and final state and define its values at intermediate times

$$\psi_t(x) = \psi_{t_f}(x_f) + \gamma \frac{(x - x_f)^2}{2(t_f - t)} \quad (2.59)$$

we can check that it is a solution of Burgers equation as

$$\partial_x \psi = \gamma \frac{x - x_f}{t_f - t} \quad (2.60)$$

$$\partial_t \psi = -\gamma \left[\frac{x - x_f}{t_f - t} \right]^2 \quad (2.61)$$

2.3 Computation of the optimal work with given final control

It is possible to make use of the optimization method for heat to compute the optimal work for a given initial state and given final potential. In fact, through the first law of thermodynamics (1.9) we can express work in terms of heat as

$$\langle \delta W \rangle = \langle dU \rangle + \langle \delta Q \rangle = \int m_{t_f}(x) U_{t_f}(x) dx - \int m_{t_i}(a) U_{t_i}(a) da + (\bar{S}_{t_i} - \bar{S}_{t_f}) \quad (2.62)$$

The optimization has now to include also the changes in internal energy and will therefore be different. Recalling equation (2.41) for the optimal drift, we can expand the expression for optimal heat as

$$\begin{aligned} \bar{S}_{t_i} - \bar{S}_{t_f} &= \int (-2\psi_{t_i}(a) - R_{t_i}) m_{t_i}(a) da - \int (-2\psi_{t_f}(x) - R_{t_f}) m_{t_f}(x) dx = \\ &- \int \frac{1}{\beta} \log [m_{t_i}(a)] m_{t_i}(a) da + \int \frac{1}{\beta} \log [m_{t_f}(x)] m_{t_f}(x) dx + \\ &- 2 \int \psi_{t_i}(a) m_{t_i}(a) da + 2 \int \psi_{t_f}(x) m_{t_f}(x) dx = \\ &T \sigma_{t_i}(a) - T \sigma_{t_f}(x) - 2 \int \psi_{t_i}(a) m_{t_i}(a) da + 2 \int \psi_{t_i}(a) m_{t_f}(x) dx + 2 \int \frac{(a - x_f)^2}{2(t_f - t_i)} m_{t_f}(x) dx \end{aligned} \quad (2.63)$$

The condition in equation (2.57) allows the two integrals containing ψ to cancel off so that, for an isothermal transformation, we can write work as

$$\langle \delta W \rangle = \Delta U - T \Delta \sigma + \int \frac{(a - x_f)^2}{t_f - t_i} m_{t_f}(x) dx \quad (2.64)$$

Which makes very much thermodynamic sense, giving

$$\langle \delta W \rangle = \Delta F + \int \frac{(a - x_f)^2}{t_f - t_i} m_{t_f}(x) dx \quad (2.65)$$

This is an expression of work as the difference of free energy between two states plus some dissipative, irreversible work. The interpretation of the cost function as dissipated work is one of the findings of [1]. It is now evident that the optimal solution for work is reached by solving the assignment problem for the inverse Lagrangian map minimizing the cost function and the boundary energy terms. This kind of problem has been widely studied and efficient solving techniques have been found for it.

Let us now explicitly compute the value of equation (2.62) for the case of a quadratic potential in one dimension in an isothermal process. This can be done analytically if the final probability distribution is Gaussian too. Let us consider a process starting from an

equilibrium state with a quadratic potential centered in the origin namely $U_{t_i}(a) = \frac{1}{2}a^2$. The initial distribution is then Gaussian with zero mean and variance $\frac{1}{\beta}$:

$$\begin{aligned} m_{t_i}(a) &= \frac{e^{-\beta U(a)}}{N} = \sqrt{\frac{\beta}{2\pi}} e^{-\beta \frac{a^2}{2}} \\ R_{t_i}(a) &= -\frac{a^2}{2} + \frac{1}{2\beta} \log \frac{\beta}{2\pi} \end{aligned} \quad (2.66)$$

The assumption of a final Gaussian distribution may seem too simplifying as, considering the control moving the center of the non equilibrium distribution in one direction, one would expect a deformation in its shape and the arising of skewness. However, for a quadratic potential (i.e. linear drift), a Gaussian distribution with moving mean can satisfy the Fokker-Planck equation associated to a Langevin . The distribution can therefore remain Gaussian along its whole evolution. This useful property is explicitly proven in Appendix A and ensures that the final distribution is Gaussian. The distributions and potentials needed for the to the computation of the optimal work are only the initial and final time ones (see equation 2.62). We can consider a generic quadratic final potential centered in μ and with a stiffness c

$$U_{t_f}(x) = \frac{c}{2}(x - \mu)^2 \quad (2.67)$$

The final distribution might not have equilibrated yet with the final potential but it is Gaussian and we denote its (yet unknown) mean by q and variance by $\frac{1}{\beta r}$.

$$\begin{aligned} m_{t_f}(x) &= \sqrt{\frac{\beta r}{2\pi}} e^{-\beta r \frac{(x-q)^2}{2}} \\ R_{t_f}(x) &= -\frac{r(x-q)^2}{2} - \frac{1}{2\beta} \log \frac{2\pi}{\beta r} \end{aligned} \quad (2.68)$$

It is important to note that if we let the distribution equilibrate no additional work is done since the external potential is not changed and this makes the second term in equation (1.8) zero. Relaxation to equilibrium contributes only to the released heat.

It is now possible to compute all the terms in equation (2.62) starting from the average final internal energy:

$$\begin{aligned} \langle U_{t_f}(x) \rangle &= \int \frac{c}{2}(x - \mu)^2 \frac{e^{-\beta r \frac{(x-q)^2}{2}}}{N} dx = \\ &= \frac{c}{2N} \int (x - q + q - \mu)^2 e^{-\beta r \frac{(x-q)^2}{2}} dx = \\ &= \frac{c}{2N} \int [(x - q)^2 + 2(x - q)(q - \mu) + (q - \mu)^2] e^{-\beta r \frac{(x-q)^2}{2}} dx = \\ &= \frac{c}{2} \left[\frac{1}{\beta r} + (q - \mu)^2 \right] \end{aligned} \quad (2.69)$$

Where in the last step we have made use of properties of Gaussian integrals and that $(q - \mu)$ is just constant. Then for the average initial internal energy

$$-\langle U_{t_i}(a) \rangle = - \int \frac{1}{2} a^2 m_{t_i}(a) da = - \int \frac{1}{2} a^2 \frac{e^{-\beta \frac{a^2}{2}}}{N} da = - \frac{1}{2\beta} \quad (2.70)$$

And for the entropy difference

$$\begin{aligned} & - \int \frac{1}{\beta} \log [m_{t_i}(a)] m_{t_i}(a) da + \int \frac{1}{\beta} \log [m_{t_f}(x)] m_{t_f}(x) dx = \\ & - \frac{1}{\beta N_i} \int \left[-\beta \frac{a^2}{2} - \frac{1}{2} \log \frac{2\pi}{\beta} \right] e^{-\beta \frac{a^2}{2}} da + \\ & \frac{1}{\beta N_f} \int \left[-\frac{\beta r(x-q)^2}{2} - \frac{1}{2} \left(\log \frac{2\pi}{\beta} - \log r \right) \right] e^{-\beta r \frac{(x-q)^2}{2}} dx = \\ & - \frac{1}{\beta} \left(-\frac{1}{2} - \frac{1}{2} \log \frac{2\pi}{\beta} \right) + \frac{1}{\beta} \left(-\frac{1}{2} - \frac{1}{2} \log \frac{2\pi}{\beta} + \frac{1}{2} \log r \right) = \\ & = \frac{1}{2\beta} \log r \end{aligned} \quad (2.71)$$

For this quadratic example we can analytically compute the cost term. The initial position a is related to the final one q by the inverse Lagrangian map describing the process moving the probability distribution from the origin to the final position: we have that $a(q) = 0$. To compute the integral, we need to know what initial position corresponds to a final position slightly moved from the center of the final distribution: $a(q + \Delta q)$. In one dimension we know that, from mass conservation in the transport equation (2.57), the cumulative mass functions of the initial and final states are equal. Moreover, for a Gaussian distribution with mean μ and variance σ^2 the following holds: $M(x) = \frac{1}{2} [\text{erf}(\frac{x-\mu}{\sigma\sqrt{2}})]$. We then have

$$a(x)\sqrt{\beta} = (x - q)\sqrt{\beta r} \quad (2.72)$$

which for $x = q + \Delta q$ yields

$$a(q + \Delta q) = (\Delta q)\sqrt{r} \quad (2.73)$$

This relation is rigorous in one dimension but is an ansatz for higher dimensions. By substituting $x = q + \Delta q$ with q fixed and Δq as integrating variable in the cost function from equation (2.64) we get to

$$\begin{aligned} & \int \frac{(a(x) - x_f)^2}{t_f - t_i} m_{t_f}(x) dx = \\ & = \int \frac{(a(q + \Delta q) - q - \Delta q)^2}{t_f - t_i} \frac{1}{N} e^{-\frac{r\beta}{2}(\Delta q)^2} d(\Delta q) = \\ & = \int \frac{(-q - \Delta q(1 - \sqrt{r}))^2}{t_f - t_i} \frac{1}{N} e^{-\frac{r\beta}{2}(\Delta q)^2} d(\Delta q) = \\ & = \frac{1}{t_f - t_i} \left(q^2 + \frac{(1 - \sqrt{r})^2}{r\beta} \right) \end{aligned} \quad (2.74)$$

Adding all the terms together we can finally write the expression for the optimal work

$$\langle \delta W \rangle = \left(\frac{c}{r} - 1\right) \frac{1}{2\beta} + \frac{c}{2}(q - \mu)^2 + \frac{\log r}{2\beta} + \frac{1}{t_f - t_i} \left(q^2 + \frac{(1 - \sqrt{r})^2}{r\beta}\right) \quad (2.75)$$

This expression depends on the final distribution and, since the final state is not given, we can look for the optimal parameters r, q that minimize it. The unidimensional quadratic potential setting is the same used by Schmiedl and Seifert in [2] and described in the next chapter. The procedure followed so far to find the optimal work corresponds to solving the Euler-Lagrange equation as done by Schmiedl and Seifert. The two approaches yield exactly the same results. Optimizing also the final distribution is analogous to choosing optimal constants in the solution of the Euler-Lagrange equation. It is as if in equation (2.75) we had specified the form and time dependence of the optimal protocol and want now to choose the coefficients that would optimize the function. If we now differentiate with respect to q we can find the optimal value for the final distribution mean.

$$\frac{d\langle \delta W \rangle}{dq} = c(q - \mu) + \frac{2q}{t_f - t_i} \quad (2.76)$$

Which gives

$$q^* = \frac{c(t_f - t_i)}{c(t_f - t_i) + 2} \mu = \frac{1}{1 + \frac{2}{c(t_f - t_i)}} \mu \quad (2.77)$$

This shows how the mean position moves towards the center of the final potential and actually reaches it for a quasi-static process (i.e. $t_f - t_i \rightarrow \infty$). One can also ask for the optimal value of r which is obtained by differentiating over r :

$$\frac{d\langle \delta W \rangle}{dr} = \frac{1}{2\beta r} \left(1 - \frac{c}{r}\right) + \frac{1}{\beta(t_f - t_i)} \frac{\sqrt{r} - 1}{-r^2} = \frac{1}{2\beta r^2} \left[r - c + 2\frac{\sqrt{r}}{t_f - t_i} - \frac{2}{t_f - t_i}\right] \quad (2.78)$$

Equating it to zero we have a second order equation in \sqrt{r} resulting in

$$\sqrt{r^*} = \frac{1}{(t_f - t_i)} [-1 + \sqrt{1 + 2(t_f - t_i) + c(t_f - t_i)^2}] \quad (2.79)$$

It is worth remarking that the distribution is determined by the control and that there is a one to one relation between optimal distribution parameters and optimal control as shown in equations (2.77), (2.79), (A.10), (A.11). Hence, an optimization over the distribution parameters results in finding the optimal control as well.

In case we are interested in minimizing the heat released between fixed final and initial states we would have to consider only

$$\langle \delta Q \rangle = -T\Delta\sigma + \int \frac{(a - x_f)^2}{t_f - t_i} m_{t_f}(x) dx \quad (2.80)$$

and since the given initial and final state determine fully the entropy term $-T\Delta\sigma$, the optimization then reduces to the one of the cost function $\int \frac{(a - x_f)^2}{t_f - t_i} m_{t_f}(x) dx$ which is found by solving the assignment problem as described in section 2.2. The differentiation over the parameters of the final distribution is meaningless as they are given by the problem and can therefore not be optimized.

Chapter 3

Exactly solvable optimization problems

In this chapter we briefly describe the examples introduced by Seifert and Schmiedl in [2]. The solutions for optimal protocols presented there display unexpected discontinuities rising interesting questions on optimal stochastic control. The AMM method [1] described in the previous chapter (2) can be seen as a generalization of these examples and we use it to solve them. The approach in [2] is to cast the definition of the work functional (2.3) in terms of the parameters of its Gaussian distribution and to minimize it via Euler-Lagrange equation. Before starting, let us remark that they use a slightly different notation: they denote the time variable as τ , the total control interval $t_f - t_i \equiv t$ and the mean position of the particle as u . Moreover, for simplicity reasons in their computations β , the diffusion coefficient and the viscous friction are taken to be 1.

3.1 Controlling the center of a quadratic potential

3.1.1 Solution by Schmiedl and Seifert

Their first example deals with a moving laser trap described by a quadratic potential with time dependent center and constant stiffness.

$$U_t(x) = \frac{1}{2}(x - \lambda_t)^2 \quad (3.1)$$

Here λ_t is the time dependent protocol responsible for the particle transport. By comparing it to equation (2.67) we see that the final control λ_{t_f} corresponds to the center of the potential at final time which we have previously denoted as μ . The stiffness of the trap c is here constant and equal to be 1. Seifert and Schmiedl start by averaging the Langevin equation with this sort of potential obtaining an important relationship between the mean value of the distribution at intermediate time q_t and the protocol

$$\dot{q}_t = \lambda_t - q_t \quad (3.2)$$

Which is equivalent to what found in (A.10) and enforces boundary conditions on $\dot{q}(t)$

$$\dot{q}(t_f) = \lambda_{t_f} - q(t_f) \quad (3.3)$$

$$\dot{q}_0 = \lambda_0 - q_0 = 0 \quad (3.4)$$

They are now able to express the mean value for work (2.3) in terms of the mean position of the particle

$$\langle W(\lambda_t, x) \rangle = \int_0^{t_f} \dot{q}(\dot{q} + \ddot{q})dt = \int_0^{t_f} \dot{q}^2 dt + \frac{1}{2}[\dot{q}^2]_0^{t_f} \quad (3.5)$$

Its associated Euler Lagrange equation reads

$$\ddot{q} = 0 \quad (3.6)$$

And it is solved by a linear evolution of the mean of the distribution

$$q_t = lt \quad (3.7)$$

Where l is a multiplying coefficient and the constant term is zero because of the initial condition $q(0) = 0$. Substituting this into the expression for mean work (3.5), recalling the boundary conditions (3.3) they get

$$\langle \delta W \rangle = l^2 t_f + \frac{(\lambda_f - l t_f)^2}{2} \quad (3.8)$$

This is equivalent to the expression found in equation (2.75). There we had further optimized by looking for the optimal final distribution, differentiating over q and r . Similarly, here, after having found the optimal time dependence of the mean position q , it is possible to compute the coefficient l minimizing the functional. Hence they differentiate (3.8) over l finding

$$l^* = \frac{\lambda_{t_f}}{t_f + 2} \quad (3.9)$$

which gives an optimal work:

$$\langle \delta W^* \rangle = \frac{\lambda_{t_f}^2}{t_f + 2} \quad (3.10)$$

By means of equation (3.2) it is possible to compute the expression for the optimal protocol

$$\lambda_t^* = \frac{\lambda_{t_f}(t + 1)}{t_f + 2} \quad (3.11)$$

which surprisingly displays two jumps at the beginning and at the end of the control interval. This discontinuity results by simply imposing continuity for the mean position q and plugging into equation (3.2) the optimal l . The boundary conditions on \dot{u} require in fact discontinuities in its evolution.

3.1.2 Solution via AMM

Let us now tackle the problem making use of the results of [1] described in the previous chapter. We start by noting that, if during the process the potential stiffness does not change, also the variance of the distribution remains constant. In fact, from equation (2.79) we have that for $c = 1$ we get $r^* = 1$. From equation (2.77) we can compute the optimal final mean position

$$q^* = \frac{t_f}{t_f + 2} \lambda_{t_f} \quad (3.12)$$

and consequently the optimal mean work

$$\langle \delta W^* \rangle = \frac{\lambda_{t_f}^2}{2} \left(\frac{t_f}{t_f + 2} - 1 \right)^2 + \frac{1}{t_f} \frac{t_f^2 \lambda_{t_f}^2}{(t_f + 2)^2} = \frac{\lambda_{t_f}^2}{t_f + 2} \quad (3.13)$$

Which is exactly what found by Seifert and Schmiedl.

So far, the result concerns the final state and carries no information about the time dependent protocol. Nonetheless, we know that the optimal transport problem can be described by Burgers equation which has a linear time evolution. Consequently the transport is linear in time and so will be the evolution of the mean and of the standard deviation of the distribution. Hence

$$q_t^* = \frac{t}{t_f} q^* \quad (3.14)$$

and, through (3.2), we can compute the optimal protocol.

$$\lambda_t^* = q_t^* + \frac{dq_t^*}{dt} = \frac{(t+1)}{t_f} q_t^* = \lambda_{t_f} \frac{t+1}{t_f + 2} \quad (3.15)$$

3.2 Controlling the stiffness of a quadratic potential

3.2.1 Solution by Schmiedl and Seifert

The second example in [2] concerns a potential with a fixed center and a variable stiffness

$$U_t(x) = \lambda_t \frac{x^2}{2} \quad (3.16)$$

Schmiedl and Seifert start by obtaining the equation of motion for the variance by multiplying the Langevin equation with x^2 and averaging:

$$\dot{w} = -2\lambda w + 2 \quad (3.17)$$

Which is equivalent to what found in (A.12) and which enables them to express the mean work in terms of the variance of the distribution

$$\begin{aligned}
\langle W \rangle &= \int_0^{t_f} \dot{\lambda} \frac{w}{2} dt = \\
&= [\lambda \frac{w}{2}]_0^{t_f} - \int_0^{t_f} \lambda \frac{\dot{w}}{2} dt = \\
&= [\lambda \frac{w}{2}]_0^{t_f} - \int_0^{t_f} \frac{\dot{w}}{2w} dt + \frac{1}{4} \int_0^{t_f} \frac{\dot{w}^2}{w} dt = \\
&= \frac{1}{2} [\lambda w - \ln w]_0^{t_f} + \frac{1}{4} \int_0^{t_f} \frac{\dot{w}^2}{w} dt
\end{aligned} \tag{3.18}$$

To minimize it they again make use of Euler-Lagrange equation which reads

$$\dot{w}^2 - 2w\ddot{w} = 0 \tag{3.19}$$

The equation is solved by a quadratically time dependent variance of the following type

$$w_t = c_1(1 + c_2 t)^2 \tag{3.20}$$

The first constant is determined by the initial distribution and we have $c_1 = 1/\lambda_1$. The second one is chosen in order to minimize work. They proceed by plugging the time dependent variance (3.17) into the expression of work (3.18)

$$\langle W \rangle = \frac{(c_2 t_f)^2}{\lambda_{t_i} t_f} - \log(1 + c_2 t_f) + \frac{\lambda_{t_f}}{2\lambda_{t_i}} (1 + c_2 t_f)^2 - \frac{1}{2} \tag{3.21}$$

and differentiating over c_2 they find

$$c_2^* = \frac{-1 - \lambda_{t_f} t_f + \sqrt{1 + 2\lambda_{t_i} t_f + \lambda_{t_f} \lambda_{t_i} t_f^2}}{t_f(2 + \lambda_{t_f} t_f)} \tag{3.22}$$

and therefore an optimal work of

$$\langle W^* \rangle = \frac{1 + \lambda_{t_f} t_f - \sqrt{1 + 2\lambda_{t_i} t_f + \lambda_{t_f} \lambda_{t_i} t_f^2}}{2\lambda_{t_i} t_f + \lambda_{t_f} \lambda_{t_i} t_f^2} - \log \frac{1 + \sqrt{1 + 2\lambda_{t_i} t_f + \lambda_{t_f} \lambda_{t_i} t_f^2}}{(2 + \lambda_{t_f} t_f)} \tag{3.23}$$

Through the relation derived in (3.20) they are able to express the optimal protocol as

$$\lambda_t^* = \frac{\lambda_0 - c_2^*(1 + c_2^* t)}{(1 + c_2^* t)^2} \tag{3.24}$$

3.2.2 Solution via AMM

If we now try to solve the problem by the general optimization method shown in equation (2.75) we recover the same result. The procedure is analogous to the one followed for the first example but we are now looking for the optimal r which is connected to the standard deviation σ by $\sigma_{t_f} = \frac{1}{\sqrt{r}}$. In [2], for simplicity, β and γ are taken to be 1.

Since σ_{t_f} describes the distribution at final time, to recover the time dependent protocol, we need to introduce a time dependence as we have done in the previous example. The solution of Burgers equation imposes a linear time dependence for the standard deviation

$$\sigma_t = 1 + \frac{t}{t_f}(\sigma_{t_f} - 1) \quad (3.25)$$

Which again satisfies Euler Lagrange (3.19) since

$$\begin{aligned} w_t &= \sigma_t^2 = \left(1 + \frac{t}{t_f}(\sigma_{t_f} - 1)\right)^2 \\ \dot{w} &= 2 \frac{(\sigma_{t_f} - 1)}{t_f} \left(1 + \frac{t}{t_f}(\sigma_{t_f} - 1)\right) \\ \ddot{w} &= 2 \frac{(\sigma_{t_f} - 1)^2}{t_f^2} \end{aligned} \quad (3.26)$$

result in

$$\begin{aligned} \dot{w}^2 - 2w\ddot{w} &= \\ &= 4 \frac{(\sigma_{t_f} - 1)^2}{t_f^2} \left(1 + \frac{t}{t_f}(\sigma_{t_f} - 1)\right)^2 - 4 \left(1 + \frac{t}{t_f}(\sigma_{t_f} - 1)\right)^2 \frac{(\sigma_{t_f} - 1)^2}{t_f^2} = 0 \end{aligned} \quad (3.27)$$

We are now going to use the optimal r^* derived in equation (2.79) considering that c corresponds to λ_{t_f} and that the initial variance is one $w(0) = 1$ as the initial stiffness $\lambda_{t_i} = 1$. Since the variance is the square of the standard deviation we have

$$w_t = \sigma_t^2 = \left[1 + \frac{t}{t_f}(\sigma_{t_f} - 1)\right]^2 = c_1(1 + c_2 t)^2 \quad (3.28)$$

we can then express c_2 in terms of σ_{t_f} and explicit its form

$$\begin{aligned}
c_2^* &= \frac{\sigma_{t_f} - 1}{t_f} = \frac{1 - \sqrt{r^*}}{t_f \sqrt{r^*}} = \frac{t_f - \sqrt{\lambda_{t_f} t_f^2 + 2t_f + 1} + 1}{t_f (\sqrt{\lambda_{t_f} t_f^2 + 2t_f + 1} - 1)} \\
&= \frac{(t_f - \sqrt{\lambda_{t_f} t_f^2 + 2t_f + 1} + 1)(\sqrt{\lambda_{t_f} t_f^2 + 2t_f + 1} + 1)}{t_f (\lambda_{t_f} t_f^2 + 2t_f)} \\
&= \frac{t_f + t_f \sqrt{\lambda_{t_f} t_f^2 + 2t_f + 1} + 1 + \sqrt{\lambda_{t_f} t_f^2 + 2t_f + 1} - \sqrt{\lambda_{t_f} t_f^2 + 2t_f + 1} - (\lambda_{t_f} t_f^2 + 2t_f + 1)}{t_f^2 (\lambda_{t_f} t_f + 2)} \\
&= \frac{-1 - \lambda_{t_f} t_f + \sqrt{\lambda_{t_f} t_f^2 + 2t_f + 1}}{t_f (\lambda_{t_f} t_f + 2)} \tag{3.29}
\end{aligned}$$

Which is what found by Schmiedl and Seifert expressed in equation (3.22) in case of an initial variance and potential stiffness equal to one. If we consider equation (2.75) and plug in the optimal r (2.79) with the conditions of this example i.e. $\mu = 0$ and $c = \lambda_{t_f}$ we get

$$\langle W^* \rangle = \frac{1 + \lambda_{t_f} t_f - \sqrt{1 + 2t_f + \lambda_{t_f} t_f^2}}{2t_f + \lambda_{t_f} t_f^2} + \log \left[\frac{1}{t} (-1 + \sqrt{1 + 2t_f + \lambda_{t_f} t_f^2}) \right] \tag{3.30}$$

which again corresponds to the case of an initial variance and potential stiffness equal to one of equation (3.23).

Chapter 4

Brownian motors and mesoscopic engines

Stochastic mesoscopic settings allow for the design of a variety of engines and motors. As the ones in classical thermodynamics, such devices, transform thermal (or chemical) energy into work or vice versa. Some of these models make an active use of the noisy, fluctuating environment whereas others are just classical thermodynamic cycles realized in the mesoscopic realm.

4.1 Rectifying motors

Let us begin by briefly sketching some of the fundamental features and some examples of Brownian motors that structurally differ from macroscopic ones. Their distinguishing as done by Schmiedl and Seifert

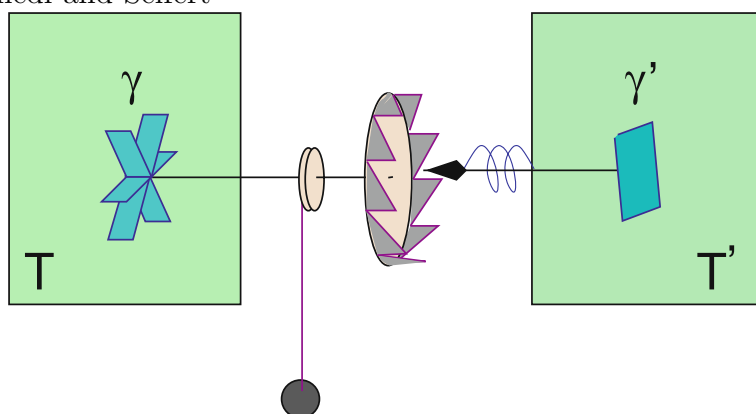


Figure 4.1: Feynman Ratchet as depicted in [5]

characteristic is the use of the random fluctuating forces to induce directed motion. This aim is reached through the introduction of some asymmetry that rectifies the random

fluctuations into a chosen direction. The asymmetry can be brought about by means of asymmetric potentials or of asymmetric temperature profiles or by exploiting some more involved cross-coupling between the motion and the thermodynamic forces.

Historically relevant works described in [5] are the Feynman ratchet and the Büttiker Landauer ratchet. We hint the underlying basic concepts here and point at [5] for actual and detailed description.

The Feynman ratchet (fig. 4.1) operates between two different heat reservoirs:

Fluctuations of the hot reservoir (T) randomly move the vane that rotates the ratchet wheel while fluctuations in the cold one (T') shake the pawl. The asymmetric design of the ratchet wheel makes rotation in one direction more likely and this results in a net average rotation that can be used to lift a load.

The Büttiker Landauer ratchet (fig.4.2) involves a single mesoscopic particle in a periodic potential. The thermal environment switches between two different temperatures with the same periodicity of the potential

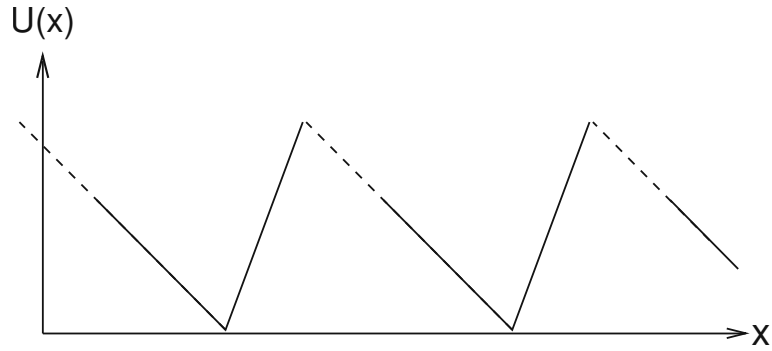


Figure 4.2: Büttiker Landauer Ratchet as depicted in [5]

The dashed line indicates a temperature T' lower than T denoted by the full line. We see that in the region immediately before (left of) the potential barrier the temperature is higher than immediately after (right of) the barrier. A higher temperature means stronger intensity of the random forces and this results in the particle having a higher probability of jumping over the barrier. Since it is now more likely for the particle to jump over the barrier when the particle is at its left than when it is at its right we have an average flux of particles rightwards. This shows how asymmetric periodical setting of temperatures (and therefore fluctuations) in an asymmetric periodic potential is capable of rectifying the noise and inducing directed transport.

Among isothermal Brownian motors the flashing ratchet (fig.4.3) is a very instructive example.

It consists of a single mesoscopic particle subject to an asymmetric potential (e.g. a saw-tooth function) and make explicit use of thermal noise. The potential flashes i.e. it is periodically turned on and off.

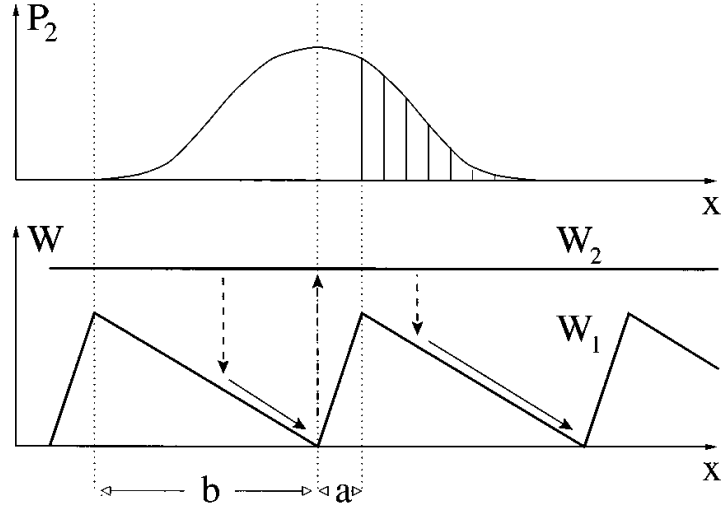


Figure 4.3: Flashing Ratchet as depicted in [9] The top shows the probability of overcoming the position of the barrier during the diffusive step. The bottom shows the potential profile.

The particle therefore switches from motion determined by the potential to free diffusion. The distance between the bottom of the potential and the barrier top to its right is labeled as a whereas the distance to its left one is b . The ratio $\frac{b}{a}$ indicates the degree of asymmetry towards the right. During the potential controlled step the particle drifts downhill and reaches on average the bottom of the potential which is located next to the barrier (given the asymmetry) of the potential. The potential is now switched off and the particle freely diffuses. When the potential is switched on again the particle has a probability equivalent to the hatched area of the Gaussian in (fig.4.3) of having diffused to the next potential well. Once there, it will drift again towards its minimum. On average this induces a directed motion. Crucial for this model is the asymmetry of the potential (with a centered minimum of the potential there would be no net transport), the noisy environment and the choice of the flashing time. In fact, in the potential step, the particle must have enough time to drift to the bottom of the potential well but not enough to spontaneously jump out of the well by effect of the fluctuations (shorter time than the Kramer one). Also, the diffusive step shall last long enough to allow diffusion for a distance a that is needed for particle to get right of the barrier but shall last less than the time needed for diffusing a distance b which might lead the particle to be in the well to the left of the starting one. The energy needed for the transport is given by the energy needed to flash the potential and can be provided by chemical reactions.

4.2 Molecular Motors

Many biological settings are properly described by typical mesoscopic stochastic dynamics [9]. Of particular interest are those systems which are responsible for the transport of

substances inside the cell. Such systems can be defined as molecular motors. They operate along filaments which serve as railroads. The most studied ones are myosins which move on actin filaments and kinesins and dyneins which use microtubules as support. For purposes of a highly idealized model the filaments can be defined as rigid, periodical ($\sim 10nm$) and polar structures. The motors attach to them by two “heads” of size comparable to the one of the period. Inside a cell the temperature is nearly constant and this rules out the motor modeled by the Feynman Ratchet from the ones actually realizable in nature. Local temperature inhomogeneities are possible in a cell, but since the typical periodicity of a filament is of a few tens of nanometers, the temperature gradients needed to implement a Büttiker Landauer Ratchet would decay in times of the orders of microseconds making this motor unfit for transport. The flashing ratchet represents instead a concrete possibility for providing motion in isothermal, stochastic settings. The somewhat abstract flashing feature can be realized by considering a two state system. In the first state the particle is subject to the potential whereas in the second one freely diffuses¹. The transition between the two states is made possible by the chemical energy supplied by ATP or ADP hydrolisation and it is at its expenses that the motion is possible. The transition rates are determined by changes in the chemical potential.

Extracellular motions such as muscle contraction can be modeled by the simultaneous action of many motors. We refer to [9] for a thorough and detailed exposition and restrict to listing some the results here. Such models exhibit a richer variety of possible behaviour and new features can arise. For example, if we consider several particles in the two state potential described before, all rigidly linked to a backbone, collective effect such as spontaneous symmetry breaking can come into play and net motion is possible even for symmetrical potentials. The results remind strongly of the Ising model for a ferro-magnet where the velocity of the resulting motion plays the role of magnetization, the viscous load to be transported the one of the external magnetic field and a measure of the chemical energy used plays the role of the inverse temperature β .

4.3 Classic thermodynamic cycles in stochastic environments

In stochastic environments it is also possible to define cyclic transformations in analogy to classic thermodynamics. To carry out the analogy we need to define the mesoscopic objects corresponding to the classical ones. The role of the working gas that can be compressed or expanded by a piston is now played by a colloidal particle in an optical trap (analogous to the piston) of variable intensity. The source of heating that, classically is fire is now laser heating. Work, heat and internal energy are defined as in section 1.1. These engines differ substantially from the rectifying ones since they do not make any explicit use of the fluctuations but operate only taking them into account (if not despite them). The

¹The states do not necessarily need to be two and their potentials can be of more general form than the one here described

role of noise is important in the computation of their characteristic and introduce possible cycles in which no work is extracted but does not change their structure. Moreover, these mesoscopic thermal engines do not provide directed motion but extract work intended as change on the controlling potential as in eq. (1.8).

A mesoscopic Carnot cycle is presented in [3]. It displays two isothermal and two adiabatic

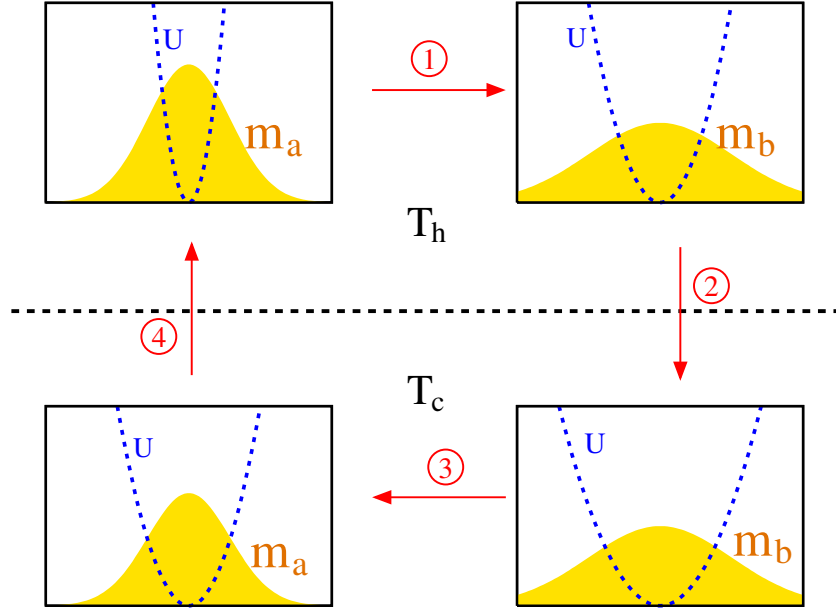


Figure 4.4: Mesoscopic Carnot Cycle as depicted in [3]

steps. In the first isothermal expansion, work is extracted since the potential is widened. An instantaneous adiabatic cooling then occurs and, while the distribution does not have time to change, the potential is allowed to be modified as the time scale for the optical trap are much smaller than the ones needed for the distribution to equilibrate. An isothermal compression takes now place and work is needed to accomplish it. For the process to be a cycle, its final distribution must equal the one of the beginning of the cycle. Finally an adiabatic transition to the initial temperature is performed. As for classic thermodynamics, the fact that the two isothermal steps are performed at different temperatures allows the engine to extract work from the heat bath.

A mesoscopic Stirling cycle can also be defined and it has been experimentally realized in [15]. For such cycle the instantaneous adiabatic steps are substituted by instantaneous isochoric ones in which the potential (playing the role of the volume) is kept fixed.

4.4 Optimal work for the computation of efficiency at maximum power of a mesoscopic Carnot engine

Since these mesoscopic engines are formalized in terms of stochastic work and heat with the same definition of [1] it is interesting to apply this optimization procedure to the transformations they involve. Such study will indicate the optimal protocols allowing the engine to produce the most work possible and this will give information on its maximum efficiency. Moreover, since the solution of [1] deals with non-equilibrium transformations, we are able to compute non-equilibrium and finite-time thermodynamic quantities such as the efficiency at maximum power.

4.4.1 Optimal work for a non-equilibrium initial state

To proceed to such an extension of the method some care must be taken as, in a cycle, the initial distribution of a step of the engine is the same as the final one of the previous step. This means that if the final distribution of a step has not had time to equilibrate with the potential (feature that we had previously considered) also the initial distribution of the following step will be a non-equilibrium one (whereas we had assumed an equilibrium starting point). We therefore need to generalize the formalism to settings displaying non-equilibrium initial states. This leads to a slight modification of the average work of equation (2.75).

Furthermore, all thermodynamic engines operate between different temperatures and extract work from this difference. For AMM solution to describe these features we need to extend it to processes with non constant temperature and therefore with varying β and consequently diffusion coefficient. This is not straightforward and affects the optimization procedure solution at various points. We focus here on the isothermal step and postpone such discussion to chapter 5.

Let us now extend AMM to the case of a non equilibrium initial state. We will here consider an initial distribution with a generic variance w_{t_i} instead of the $\frac{1}{\beta}$ as used for simplicity before. With a slight change of notation we will describe the final distribution in terms of its variance $w_{t_f} = \frac{1}{\beta r}$ instead of the parameter r . Let us consider which terms of equation (2.75) are affected by a non-equilibrium initial state. The averages taken at initial time, equation (2.70) and the entropy term (2.71) are now clearly different since the initial distribution is changed. A different initial variance affects also the inverse Lagrangian map connecting x_f to x_i because of the relation involving the error functions at initial and final state.

To avoid confusion let us recompute the optimal work step by step for a non-equilibrium initial state with initial potential $U_{t_i} = \frac{\lambda_{t_i}}{2}x^2$ and final one $U_{t_f} = \frac{\lambda_{t_f}}{2}(x - \mu)^2$. Note that this is not the most general setting as we are considering an initial potential and distribution centered in the origin whereas this is not always the case. Anyway, mesoscopic thermodynamic cycles of interest and specifically the example in [3] deal with origin cen-

tered potentials and distributions and we therefore restrict our generalization to such case. The average final internal energy (2.69) is not affected as only the initial settings are now different. In terms of w_{t_f} it takes the form

$$\langle U_{t_f} \rangle = \frac{\lambda_{t_f}}{2} [w_{t_f} + (q - \mu)^2] \quad (4.1)$$

The initial internal energy gives

$$-\langle U_{t_i} \rangle = -\frac{\lambda_{t_i}}{2} w_{t_i} \quad (4.2)$$

in case of non origin centered distribution and potential it takes a form analogous to (4.1). For the entropy change we have

$$\begin{aligned} & - \int \frac{1}{\beta} \log [m_{t_i}(a)] m_{t_i}(a) da + \int \frac{1}{\beta} \log [m_{t_f}(x)] m_{t_f}(x) dx = \\ & = -\frac{1}{\beta N_i} \int \left[-\frac{a^2}{2w_{t_i}} - \frac{1}{2} \log 2\pi w_{t_i} \right] e^{-\frac{a^2}{2w_{t_i}}} da + \\ & + \frac{1}{\beta N_f} \int \left[-\frac{(x-q)^2}{2w_{t_f}} - \frac{1}{2} (\log 2\pi w_{t_f}) \right] e^{-\frac{(x-q)^2}{2w_{t_f}}} dx = \\ & = -\frac{1}{\beta} \left(-\frac{1}{2} - \frac{1}{2} \log 2\pi w_{t_i} + \frac{1}{2} + \frac{1}{2} \log 2\pi w_{t_f} \right) = \\ & = \frac{1}{2\beta} \log \frac{w_{t_i}}{w_{t_f}} \end{aligned} \quad (4.3)$$

Where N is the normalization constant of the distribution. The entropy change has the same form if the initial distribution and potential are not centered in the origin.

The cost function is affected only by changes in the Lagrangian map

$$\frac{a(x)}{\sqrt{w_{t_i}}} = \frac{(x-q)}{\sqrt{w_{t_f}}} \quad (4.4)$$

which for $x = q + \Delta q$ yields

$$a(q + \Delta q) = (\Delta q) \frac{\sqrt{w_{t_i}}}{\sqrt{w_{t_f}}} \quad (4.5)$$

Leading to

$$\begin{aligned} & \int \frac{(a(x) - x_f)^2}{t_f - t_i} m_{t_f}(x) dx = \\ & \frac{1}{t_f - t_i} (q^2 + (1 - \frac{\sqrt{w_{t_i}}}{\sqrt{w_{t_f}}})^2 w_{t_f}) = \\ & \frac{1}{t_f - t_i} (q^2 + (\sqrt{w_{t_f}} - \sqrt{w_{t_i}})^2) \end{aligned} \quad (4.6)$$

In case of a non centered distribution instead of having only q^2 we would have the difference between the initial and final mean position. It is important to remember that in general the variance of the distribution depends on the temperature. Adding all the terms together we get the expression for the optimal work between to non-equilibrium states.

$$\langle \delta W \rangle = \frac{\lambda_{t_f}}{2} [w_{t_f} + (q - \mu)^2] - \frac{\lambda_{t_i}}{2} w_{t_i} + \frac{1}{2\beta} \log \frac{(w_{t_i})}{(w_{t_f})} + \frac{1}{t_f - t_i} (q^2 + (\sqrt{w_{t_f}} - \sqrt{w_{t_i}})^2) \quad (4.7)$$

4.4.2 Explicit solution of a Carnot mesoscopic engine via AMM

With the tools derived above we can now consider the Carnot engine described in [3] and solve it in the case of a quadratic potential. Let us see in detail the steps of the model:

1. Isothermal transition at high temperature T_H taking a time t_1 . The initial potential is $U_{t_i} = \lambda_{t_i} \frac{x^2}{2}$ and the final one is $U_{t_1} = \lambda_{t_1} \frac{x^2}{2}$. The process changes a distribution centered in the origin with variance w_a to another, still centered in the origin but now with variance w_b .
2. Adiabatic instantaneous step from high temperature T_H to a cold one T_C . Being the transformation instantaneous the distribution does not have time to change so the variance is still w_b . The potential is allowed to change and takes the value $U_{t_1+} = \lambda_{t_1+} \frac{x^2}{2}$.
3. Isothermal transition at low temperature T_C taking a time t_3 . For the process to be cyclic we need the final distribution of this step to be equal to the initial one of step 1. Therefore the variance will be w_a again. This constrains the possible values of the potential and the temporal length of the step. The final potential is $U_{t_1+t_3} = \lambda_{t_1+t_3} \frac{x^2}{2}$.
4. Adiabatic instantaneous step from the low temperature T_C to the original hot one T_H . Again the distribution does not change but the potential does. For the process to be a cycle it needs to match the initial one $\lambda_{t_1+t_3+} = \lambda_{t_i}$.

It is worth remarking that, in order to make the process cycle, some conditions constrain the control. Let us highlight them once more. First, the final distribution in step 3 must be equal to the one at the beginning of the cycle. The joint effect of distribution at the beginning of the step, time duration of the step, final potential and temperature must account for that. Furthermore the final potential of the cycle must equal the starting one but this is a more evident and easy to fulfill requirement.

The final aim is to compute the efficiency at maximum power of the engine. This quantity is interesting because the (unconstrained) maximum efficiency of an engine is achieved for quasi-static transformations which, taking an infinite time, yield vanishing power. One is usually interested in engines capable of supplying power and therefore it is more relevant to consider the efficiency at maximum power which is a finite time quantity. In analogy with the solution proposed in [3] we split the problem in two steps: first we look for the protocol which enables the greatest power production and then we compute the value of

efficiency associated to it. To derive the protocol for maximum power we first find the functional time dependence of the variance which optimizes work production (as done in [1]) and then optimize over time duration and initial and final distributions the functional corresponding to power production which is different from the one of work and will therefore lead to different protocols.

Let us now consider each step and write down its value of the optimal work. The first step reads

$$\langle \delta W \rangle_1 = \frac{\lambda_{t_1}}{2} w_b - \frac{\lambda_{t_i}}{2} w_a + \frac{1}{2} \frac{1}{\beta_H} \log \frac{w_a}{w_b} + \frac{1}{t_1} (\sqrt{w_b} - \sqrt{w_a})^2 \quad (4.8)$$

As discussed before, the method developed in [1] is not suitable for the optimization of processes taking place in temperature changing settings. Moreover, steps 2 and 4 are instantaneous and therefore the distributions do not have time to change. This means that there is no time nor room for an optimal protocol, in fact, even if we allow the potential to change it would just display a discontinuous jump. The specific choice of the potential jump is also irrelevant to the final computation of work as, for the process to be cyclic all the contributions given by potential changes must be null.

$$\sum_{i=1}^4 \Delta U_{t_i} = \sum_{i=1}^4 \Delta_i(\lambda w) = 0 \quad (4.9)$$

The steps are adiabatic which means that there is no heat exchange and therefore the work reduces to the change in potential energy.

$$\langle \delta W \rangle_2 = \Delta U_2 = \frac{\lambda(t_{1+})}{2} w_b - \frac{\lambda_{t_1}}{2} w_b \quad (4.10)$$

Then step 3 is similar the first one

$$\langle \delta W \rangle_3 = \frac{\lambda_{t_3+t_1}}{2} w_a - \frac{\lambda(t_{1+})}{2} w_b + \frac{1}{2} \frac{1}{\beta_C} \log \frac{w_b}{w_a} + \frac{1}{t_3} (\sqrt{w_a} - \sqrt{w_b})^2 \quad (4.11)$$

and finally step 4 resembles step 2

$$\langle \delta W \rangle_4 = \Delta U_4 = \frac{\lambda_{t_3+t_1+}}{2} w_a - \frac{\lambda_{t_3+t_1}}{2} w_a \quad (4.12)$$

Adding them together we get the total work done on the system for a cycle. Since it is a cyclic process, the energy contribution must cancel out and this indeed happens when we impose the cycling condition i.e. that the final potential energy equals the initial one $\frac{\lambda(t_3+t_1+)}{2} = \lambda_{t_i}$.

$$\begin{aligned} \langle \delta W \rangle_{1,2,3,4} &= \frac{1}{2} \frac{1}{\beta_H} \log \frac{w_a}{w_b} + \frac{1}{t_1} (\sqrt{w_b} - \sqrt{w_a})^2 + \\ &+ \frac{1}{2} \frac{1}{\beta_C} \log \frac{w_b}{w_a} + \frac{1}{t_3} (\sqrt{w_a} - \sqrt{w_b})^2 = \\ &= \left(\frac{1}{t_1} + \frac{1}{t_3} \right) (\sqrt{w_a} - \sqrt{w_b})^2 + \left(\frac{1}{\beta_H} - \frac{1}{\beta_C} \right) \frac{1}{2} \log \frac{w_a}{w_b} \end{aligned} \quad (4.13)$$

Which is in perfect agreement with what found in [3, equation (15)]. If we were to proceed in the optimization as in section 3.2 we would further optimize this process by differentiating over the intermediate variance w_b and, (whether not specified by the problem) the initial variance w_a . Since we now want to maximize power, we must differentiate its functional and not the one of work. The power supplied by the considered engine is

$$P = \frac{-W}{t_1 + t_3} = \frac{\frac{1}{2}(T_H - T_C) \log \frac{w_b}{w_a} - \frac{(t_3 + t_1)}{t_1 t_3} (\sqrt{w_a} - \sqrt{w_b})^2}{t_1 + t_3} =$$

$$= \frac{1}{2(t_1 + t_3)} (T_H - T_C) \log \frac{w_b}{w_a} - \frac{(\sqrt{w_a} - \sqrt{w_b})^2}{t_1 t_3} \quad (4.14)$$

It is now interesting to look for the optimal duration for the cycle as it has a more involved role than in the case of heat or work. In fact, whereas on one hand a slow cycle allows for greater denominator (work production), a fast one minimizes the denominator in (4.14) increasing the power. Let us then differentiate over the duration of steps t_1, t_3 . Given their symmetrical roles the two times would have the same optimal value, we can therefore impose $t_1 = t_3$ and differentiating (4.14) over t_1 we get

$$\frac{dP}{dt_1} = \frac{d[\frac{1}{4t_1}(T_H - T_C) \log \frac{w_b}{w_a} - \frac{(\sqrt{w_a} - \sqrt{w_b})^2}{t_1^2}]}{dt_1} =$$

$$-(\frac{1}{t_1^2})[\frac{1}{4}(T_H - T_C) \log \frac{w_b}{w_a} - 2\frac{(\sqrt{w_a} - \sqrt{w_b})^2}{t_1}] \quad (4.15)$$

giving an optimal time time interval of

$$t_1^* = t_3^* = 8 \frac{(\sqrt{w_a} - \sqrt{w_b})^2}{(T_H - T_C) \log \frac{w_b}{w_a}} \quad (4.16)$$

which means that the work done on the system at maximum power is

$$\langle \delta W^* \rangle_{1,2,3,4} = -\frac{1}{4}(T_H - T_C) \log \frac{w_b}{w_a} \quad (4.17)$$

To evaluate the efficiency it is not relevant to compute the optimal variances w_b and w_a as their contribution to the work is canceled by their contribution to the heat uptake which is

$$\langle \delta Q^* \rangle_1 = \bar{S}_{t_i} - \bar{S}_{t_1} = \frac{T_H}{2} \log \frac{w_a}{w_b} + \frac{1}{t_1^*} (\sqrt{w_b} - \sqrt{w_a})^2 =$$

$$\frac{T_H}{2} \log \frac{w_a}{w_b} + (T_H - T_C) \frac{1}{8} \log \frac{w_b}{w_a} = (3T_H + T_C) \frac{1}{8} \log \frac{w_a}{w_b} \quad (4.18)$$

Again in agreement with what found in [3].

We can finally explicitly write the efficiency at maximum power

$$\eta^* = \frac{\frac{1}{4}(T_H - T_C) \log \frac{w_a}{w_b}}{(3T_H + T_C) \frac{1}{8} \log \frac{w_a}{w_b}} = \frac{2(T_H - T_C)}{3T_H + T_C} \quad (4.19)$$

It is interesting to express it in terms of Carnot efficiency $\eta_C = 1 - \frac{T_C}{T_H}$

$$\eta^* = \frac{2\eta_C}{4 - \eta_C} \quad (4.20)$$

It is also noteworthy that, as stressed in [3], it differs from the expression of efficiency at maximum power derived by Curzon-Ahlborn.

Chapter 5

Optimization in a temperature changing setting

Several problems of physical and biological interest display processes taking place in temperature varying environments (see chapter 4). In the following chapter we discuss the differences that arise with respect to the isothermal case and suggests a generalization of the AMM optimization procedure [1]. We focus on which features persist despite the non-constant temperature and especially on the arising of Burgers equation. It is important to note that changing the temperature affects also the viscous friction γ but for simplicity reasons we now neglect this effect.

5.1 Definition of heat

When temperature is spatially inhomogeneous it can act as a driving force contributing to the drift experienced by the mesoscopic particle. It becomes crucial to specify if the noise term in Langevin equation is meant in the Itô or Stratonovich way. In fact the two formalisms differ by a term $\frac{1}{2}\partial_x(\frac{1}{\beta})$ which alters the stochastic functionals of section 1.1 and the optimization technique of chapter 2. For instance, if Langevin equation is intended in the Itô convention

$$\gamma \frac{dx}{dt} = -\partial_x U(x, \lambda) + \sqrt{\frac{2}{\beta(x)}} \cdot \dot{w}_t \quad (5.1)$$

its Stratonovich equivalent reads

$$\gamma \frac{dx}{dt} = -\partial_x \left(\frac{1}{2\beta(x)} + U(x, \lambda) \right) + \sqrt{\frac{2}{\beta(x)}} \circ \dot{w}_t \quad (5.2)$$

the first law of thermodynamics derived in (1.9) becomes

$$dE = \delta W - \delta Q \quad (5.3)$$

where E is defined as

$$E = \frac{1}{2\beta(x)} + U(x, \lambda) \quad (5.4)$$

and heat is defined as

$$\delta Q = - \int \partial_x \left(\frac{1}{2\beta(x)} + U(x, \lambda) \right) dx \quad (5.5)$$

The definition of work is unchanged unless we consider the temperature to be controllable through a parameter λ' . In this case it reads

$$\delta W = \int \partial_\lambda (U(x, \lambda)) d\lambda + \int \partial_{\lambda'} \left(\frac{1}{2\beta(x, \lambda')} \right) d\lambda' \quad (5.6)$$

Beside these definition issues, a non constant temperature (and consequently diffusion coefficient) brings about more fundamental differences in the optimization described in [1]. As mentioned in the closing of section 1.1 a constant temperature allows a close relation between entropy production in the environment and heat released (1.57) which makes it possible to define heat from the functional \mathcal{J} . One could in fact say that, in the isothermal case, heat optimization corresponds to the optimization of the functional \mathcal{J} and therefore to the minimization of the irreversibility of the process. This correspondence is lost as soon as β is not constant and can therefore not filter across the integral. In the following sections we analyze in details what happens for the optimization of both heat and entropy production in the environment and highlight their differences.

5.2 Temperature depending on time only

We start by considering the case of temperature depending on time but constant in space. The time dependence of temperature spoils the validity of relation (1.57) but does not introduce any spurious drift nor ambiguities in the definition of Langevin equation (see eq 5.1 and 5.2) as $\partial_x \left(\frac{1}{\beta_t} \right) = 0$

5.2.1 Heat Optimization for time dependent temperature

We are now interested in the minimization of average heat released during a transformation taking place with a time dependent temperature. The general idea is to perform the optimization including temperature as a parameter. As a result we will be able to compute the optimal control to be exerted on the potential as a function of temperature for any given temperature protocol. We follow the same optimization procedure described in section 2.1 for the constant temperature case. Since the derivation is performed over infinitesimal time steps, the temperature time dependence introduces only higher order corrections and does not modify the first results and we recover the same HJB equation (2.20). Also the expression for the optimal drift (2.21) and the splitting of the drift (2.34) mapping Fokker-Planck equation into deterministic transport (2.39) remain unchanged. What does change

is that in the definition of the equilibrium potential (2.30). The previously constant factor $\frac{1}{\beta}$ now depends on time

$$R_t(x) = \frac{1}{\beta_t} \log m_t(x) \quad (5.7)$$

This plays a role in the mapping of HJB into Burgers equation. In fact it involves a time derivative of S which, according to equation 2.41, includes R . Making explicit the dependence we have

$$S_t(x) = -2\psi_t(x) - R_t(x) = -2\psi_t(x) - \frac{1}{\beta_t} \log m_t(x) \quad (5.8)$$

so that now

$$\partial_t S_t(x) = -2\partial_t \psi_t(x) - \frac{1}{\beta_t} \partial_t \log m_t(x) - \partial_t \left(\frac{1}{\beta_t} \right) \log m_t(x) \quad (5.9)$$

Since all the other terms involved in (2.20) have not changed, the extra term $\partial_t \left(\frac{1}{\beta_t} \right) \log m_t(x)$ given by $\partial_t S$ is not compensated. As a result, for a time dependent temperature there is a source term in Burgers equation

$$\partial_t \psi + \frac{[\partial_x \psi]^2}{2\tau} = -\frac{1}{2} \partial_t \left(\frac{1}{\beta_t} \right) \quad (5.10)$$

5.2.2 Entropy production optimization for time dependent temperature

Let us now consider entropy production in the environment as defined in equation (1.55). For our Langevin equation we can write

$$\langle \Delta S_{ENV} \rangle = \int_{t_i}^{t_f} \langle \beta_t b(x_t) \circ dx_t \rangle \quad (5.11)$$

To perform an optimization analogous to AMM we have to start by defining the value function, corresponding to what S (2.12) was for heat. It now takes the meaning of entropy produced in a process taking place from an intermediate time t to the the final one t_f

$$\Sigma_t(x) = \int_t^{t_f} (\beta_t b_t^2 + \partial_x b_t) dt \quad (5.12)$$

expressed in the Itô convention where again

$$\bar{\Sigma}_{t_i} = \langle \Delta S_{ENV} \rangle \quad (5.13)$$

As mentioned in the previous section, the HJB equation can be obtained also for a time dependent temperature. Since we are now optimizing a different functional we have a different RHS with respect to equation (2.20). In fact it has a multiplicative factor of β

$$-\partial_t \Sigma - \frac{b}{\gamma} \partial_x \Sigma - \frac{1}{\beta \gamma} \partial_x^2 \Sigma = \beta \frac{b^2}{\gamma} + \frac{\partial b}{\gamma} \quad (5.14)$$

Following the same steps as AMM we can derive the expression for the optimal drift

$$b^* = \frac{1}{2}(\partial_x R - \frac{1}{\beta}\partial_x \Sigma) = \frac{1}{2}\partial_x(R - \frac{1}{\beta}\Sigma) \quad (5.15)$$

It is of course similar to what found for S (2.21) but in terms of $\frac{1}{\beta}\Sigma$. This is dimensionally correct since Σ is an entropy and not an energy. Upon the same splitting of the drift (2.34) into an equilibrium part and a remainder we can again express

$$\Sigma = -\beta R - 2\beta\psi \quad (5.16)$$

As for the isothermal case, the splitting maps Fokker Planck equation into deterministic transport induced by the velocity field ψ (2.39).

The problem we had encountered when taking the time derivative of S in the case of heat is now avoided by the fact that Σ contains a term in βR and not only R . Substituting in the HJB (5.14) we therefore get

$$\partial_t(\beta\psi) + \frac{1}{2\gamma\beta}(\partial_x(\beta\psi))^2 = 0 \quad (5.17)$$

which upon definition of

$$\beta\psi = \varphi \quad (5.18)$$

and time re-parametrization as

$$\beta\partial_t = \partial_\tau \quad (5.19)$$

becomes

$$\partial_\tau(\varphi) + \frac{1}{2\gamma}\partial_x(\varphi)^2 = 0 \quad (5.20)$$

Which is Burgers equations for the modified time. The deterministic transport equation in the new variables reads again

$$\partial_\tau m + \frac{1}{\gamma}\partial_x(\partial_x \varphi m) = 0 \quad (5.21)$$

These two equations fully describe the transport process and can be solved via MAK reducing to an assignment problem of initial and final conditions minimizing the cost function in terms of τ

$$\int \frac{(x_i - x_f)^2}{2(\tau_f - \tau_i)} m_{t_i} dx_i \quad (5.22)$$

From this first generalization it seems that the arising of Burgers equation is connected to the optimization of entropy production and it is met by heat optimization only in the isothermal case.

It is interesting to apply the change of variables (5.19) to the SDE describing the process (i.e. Langevin equation). Considering that $dt = \beta d\tau$ we get

$$dx_\tau = \frac{\beta b}{\gamma} d\tau + \sqrt{\frac{2}{\gamma}} dw_\tau \quad (5.23)$$

and that its expression for heat yields

$$\langle Q(\tau) \rangle = \int_{\tau_i}^{\tau_f} \frac{1}{\gamma} \langle (\beta^2 b^2 + \beta \partial_x b) d\tau \rangle = \int_{t_i}^{t_f} \frac{1}{\gamma} \langle \beta b^2 + \partial_x b \rangle dt = \langle \Delta S_{ENV} \rangle \quad (5.24)$$

which is the expression for entropy production (5.11) for the usual SDE .

5.3 Temperature dependent on time and space

Let us now take the generalization of AMM a step further and consider the case of temperature depending both on time and space. If temperature depends not only on time but also on space more involved modifications are necessary and it becomes fundamental to specify if we are considering a Stratonovich or an Itô Langevin equation (see section 5.1). In the following we consider a Stratonovich Langevin process

$$dx_t = \frac{dt}{\gamma} b + \sqrt{\frac{2}{\gamma \beta(x, t)}} \circ dw_t \quad (5.25)$$

which in the Itô formulation reads

$$dx_t = \frac{dt}{\gamma} \left(b + \frac{1}{2} \partial_x \left(\frac{1}{\beta(x, t)} \right) \right) + \sqrt{\frac{2}{\gamma \beta(x, t)}} \cdot dw_t \quad (5.26)$$

The Fokker Planck equation corresponding to (5.25) displays to the spurious drift involved in the Itô SDE

$$\partial_t m = -\frac{1}{\gamma} \partial_x \left[\left(b + \frac{1}{2} \partial_x \left(\frac{1}{\beta} \right) \right) m \right] + \partial_x^2 \left(\frac{m}{\gamma \beta} \right) \quad (5.27)$$

It is interesting to note that inhomogeneous temperature scenarios also affect the generality of the Gaussianity of the particle distribution in a quadratic potential shown before equation (2.68). In fact, the thermally induced drift alters the symmetries of the potential and lead to non-Gaussian space dependent mean and variance. In the formalism, the differences arise from the new form and space dependence of Fokker-Planck equation (5.27) and results in the following conditions (analogous to A.10, A.12) coming respectively from the constants, the first order and the second order balance in x

$$\dot{w} = 2 \left[D - \left(\frac{c}{\gamma} + \partial_x^2 D \right) w \right] \quad (5.28)$$

$$\dot{q} = \frac{c}{\gamma} (\mu - q) - \frac{3}{2} \partial_x D \quad (5.29)$$

$$\dot{w} = 2 \left[D - \frac{c}{\gamma} w \right] \quad (5.30)$$

5.3.1 Heat optimization for time and space dependent temperature

We focus first on the optimization of average released heat. Since the expression for heat is written in the Itô notation it is affected by the presence of the spurious drift and, compared to (2.2), has an extra term

$$\langle Q_{SD} \rangle = \int_{t_i}^{t_f} \frac{dt}{\gamma} \langle b^2 + \frac{1}{2} b \partial_x \left(\frac{1}{\beta} \right) + \frac{\partial_x b}{\beta} \rangle \quad (5.31)$$

If we carry out the usual AMM optimization procedure for this new functional and Fokker-Planck equation, the HJB equation (2.20) describing the evolution of the functional S changes on both sides. The left hand side displays the additional spurious drift of the Fokker Planck equation and the right hand one bears the consequences of the different definition of heat.

$$-\partial_t S_{SD} - \frac{b}{\gamma} \partial_x S_{SD} - \frac{1}{2} \partial_x \left(\frac{1}{\beta} \right) \partial_x S_{SD} - \frac{1}{\beta \gamma} \partial_x^2 S_{SD} = \frac{b^2}{\gamma} + \frac{1}{2\gamma} b \partial_x \left(\frac{1}{\beta} \right) + \frac{\partial b}{\beta \gamma} \quad (5.32)$$

Since the derivation of the evolution of \bar{S} involves the term $\partial_t m$ (2.24), it is also affected by the change in Fokker-Planck equation.

$$\begin{aligned} \bar{S}_{SD}(t - dt) &= \frac{dt}{\gamma} \int dx \left(b^2 + \frac{1}{2} b \partial_x \left(\frac{1}{\beta} \right) + \frac{\partial b}{\beta} \right) m_{t-dt}(x) + \\ &+ \int dx m_{t-dt}(x) S_{SD}(x, t) + dt \int \partial_t m_{t-dt}(x) S_{SD}(x, t) dx \end{aligned} \quad (5.33)$$

from which, recalling that now β depends also on space when integrating by parts, we get

$$\begin{aligned} \bar{S}_{SD}(t - dt) &= A + \frac{dt}{\gamma} \int dx \left[b^2 + \frac{1}{2} b \partial_x \left(\frac{1}{\beta} \right) - \frac{b}{m} \partial_x \left(\frac{m}{\beta} \right) + \right. \\ &\left. + \frac{1}{\beta} \partial_x^2 S_{SD} + b \partial_x S_{SD} + \frac{1}{2} \partial_x \left(\frac{1}{\beta} \right) \partial_x S_{SD} \right] m_{t-dt}(x) \end{aligned} \quad (5.34)$$

And finding the extremal condition by differentiating with respect to b we get

$$2b + \frac{1}{2} \partial_x \left(\frac{1}{\beta} \right) - \frac{1}{m} \partial_x \left(\frac{m}{\beta} \right) + \partial_x S_{SD} = 0 \quad (5.35)$$

yielding an optimal drift

$$b_{SD}^* = \frac{1}{2} \left(\frac{1}{\beta} \frac{\partial_x m}{m} + \frac{1}{2} \partial_x \left(\frac{1}{\beta} \right) - \partial_x S_{SD} \right) \quad (5.36)$$

Another difference introduced by an inhomogeneous temperature is that the equilibrium potential R_{SD} takes a more involved form than (2.30). In fact, if we consider Fokker Planck equation at equilibrium ($\partial_t m = 0$) and denote the drift as $-\partial_x U = b$ we have

$$-(-\partial_x U + \frac{1}{2} \partial_x \left(\frac{1}{\beta} \right)) m_{eq} + \partial_x \left(\frac{m_{eq}}{\beta} \right) = 0 \quad (5.37)$$

which integrated yields

$$\log [m_{eq}\sqrt{\frac{1}{\beta}}] = - \int \partial_x U \beta dx \quad (5.38)$$

The potential gradient associated with the equilibrium is therefore given by

$$\partial_x R_{SD} = \frac{1}{\beta} \partial_x \log [m\sqrt{\frac{1}{\beta}}] = \frac{\partial_x m}{m\beta} + \frac{1}{2} \partial_x \left(\frac{1}{\beta}\right) \quad (5.39)$$

With this definition of $\partial_x R_{SD}$ the splitting of the drift (2.34) maps the Fokker Planck equation (5.27) into the deterministic transport (2.39). It is interesting to note that the expression for the optimal drift (5.36) can now be expressed as

$$b_{SD}^* = \frac{1}{2} \partial_x (R_{SD} - S_{SD}) \quad (5.40)$$

which is the same as what found in the isothermal case (2.21) but with a different expression for the equilibrium drift $\partial_x R$ and the running cost S .

As for the case of time dependent temperature, the substitution of the optimal and the split drift in HJB equation (5.32) does not result in a complete cancellation. Upon derivation of

$$R = \int^x \frac{1}{\beta} \partial_x \log [m\sqrt{\frac{1}{\beta}}] dx = \int^x \frac{1}{\beta} \partial_x \log [m] + \frac{1}{2} \frac{1}{\beta} \quad (5.41)$$

we get to Burgers equation plus an involved combination of five other terms.

$$\begin{aligned} 2\partial_t \psi + \frac{1}{\gamma} (\partial_x \psi)^2 + \partial_t \int^x \frac{1}{\beta} \partial_x \log [m] dx + \frac{1}{2} \partial_t \left(\frac{1}{\beta}\right) + \\ + \frac{1}{\gamma} \frac{\partial_x \psi \partial_x m}{m\beta} + \frac{1}{\gamma} \partial_x \psi \partial_x \left(\frac{1}{\beta}\right) + \frac{1}{\beta\gamma} \partial_x^2 \psi = 0 \end{aligned} \quad (5.42)$$

which can be re-written as

$$-2\partial_t \psi - \frac{1}{\gamma} (\partial_x \psi)^2 = \partial_t \int^x \frac{1}{\beta} \partial_x \log [m] dx + \frac{1}{2} \partial_t \left(\frac{1}{\beta}\right) + \frac{1}{\gamma m} \partial_x [(\partial_x \psi) \frac{m}{\beta}] \quad (5.43)$$

5.3.2 Entropy production optimization for time and space dependent temperature

Let us now consider the case of minimization of entropy production into the environment. The functional is defined in expression (1.55) and can be put in an Itô fashion by explicitly carrying out the Stratonovich products. In showing it we adopt the notation of section 1.1.2

$$\begin{aligned} S_{env} &= \int \langle \mathcal{J}_t \rangle dt = \int 2 \langle [\hat{u} d^{-1} + \frac{1}{2} \partial_x (\hat{u} d^{-1}) dx] dx \rangle = \\ &= \int 2 \langle [\hat{u} d^{-1} + \frac{1}{2} \partial_x (\hat{u} d^{-1}) \sqrt{d} dw_t] [(u + \tilde{u}) dt + \sqrt{d} dw_t] \rangle \end{aligned} \quad (5.44)$$

neglecting the terms of higher order, taking the average and exploiting the relation

$$u + \tilde{u} - \frac{1}{2d} \partial_x d = \hat{u} \quad (5.45)$$

deriving from 1.16 we get to

$$\int \langle \mathcal{J}_t \rangle dt = \int \langle 2\hat{u}d^{-1}\hat{u} + \partial_x \hat{u} \rangle dt \quad (5.46)$$

and making use of the explicit expressions of \hat{u} and d (see Table 1.1) we obtain

$$S_{env} = \int \langle \mathcal{J}_t \rangle dt = \int \langle \beta b^2 + \partial_x b - \beta b \partial_x \left(\frac{1}{\beta}\right) + \frac{\beta}{4} [\partial_x \left(\frac{1}{\beta}\right)]^2 - \frac{1}{2} \partial_x^2 \left(\frac{1}{\beta}\right) \rangle \frac{dt}{\gamma} \quad (5.47)$$

We can now apply the usual AMM optimization to entropy production in the environment, carefully considering the space dependence of temperature and the new functional.

The HJB equation now reads

$$\begin{aligned} -\partial_t \Sigma_{SD} - \frac{b}{\gamma} \partial_x \Sigma_{SD} - \frac{1}{2} \partial_x \left(\frac{1}{\beta}\right) \partial_x \Sigma_{SD} - \frac{1}{\beta \gamma} \partial_x^2 \Sigma_{SD} = \\ = \frac{1}{\gamma} (\beta b^2 + \partial_x b - \beta b \partial_x \left(\frac{1}{\beta}\right) + \frac{\beta}{4} [\partial_x \left(\frac{1}{\beta}\right)]^2 - \frac{1}{2} \partial_x^2 \left(\frac{1}{\beta}\right)) \end{aligned} \quad (5.48)$$

and the expression for $\bar{\Sigma}_{SD}$ is

$$\begin{aligned} \bar{\Sigma}_{SD}(t - dt) = A' + \frac{dt}{\gamma} \int [\frac{1}{\beta} \partial_x^2 \Sigma_{SD} + b \partial_x \Sigma_{SD} + \frac{1}{2} \partial_x \left(\frac{1}{\beta}\right) \partial_x \Sigma_{SD} + \\ + \beta b^2 + \partial_x b - \beta b \partial_x \left(\frac{1}{\beta}\right) + \frac{\beta}{4} [\partial_x \left(\frac{1}{\beta}\right)]^2 - \frac{1}{2} \partial_x^2 \left(\frac{1}{\beta}\right)] m_{t-dt}(x) dx \end{aligned} \quad (5.49)$$

where A' denotes a term independent of b and we have expressed the time dependence in brackets to avoid confusion in the subscript.

Looking for the extremal value of b we differentiate obtaining

$$\partial_x \Sigma_{SD} + 2\beta b - \frac{\partial_x m}{m} - \beta \partial_x \left(\frac{1}{\beta}\right) = 0 \quad (5.50)$$

corresponding to an optimal

$$b_{SDent}^* = \frac{1}{2} \left(\frac{\partial_x m}{\beta m} + \partial_x \left(\frac{1}{\beta}\right) - \frac{1}{\beta} \partial_x \Sigma_{SD} \right) \quad (5.51)$$

Making use of relation (5.39) we can express it as

$$b_{SDent}^* = \frac{1}{2} \left(\partial_x R_{SD} + \frac{1}{2} \partial_x \left(\frac{1}{\beta}\right) - \frac{1}{\beta} \partial_x \Sigma_{SD} \right) \quad (5.52)$$

where it is evident that the space dependence of temperature contributes with an extra term to the expression of the optimal drift. This did not happen for the optimization of heat (5.40).

We want now to express Σ_{SD} in terms of R_{SD} and ψ making use of the relation for the optimal drift.

$$\partial_x \Sigma_{SD} = -\beta \partial_x R_{SD} + \frac{\beta}{2} \partial_x \left(\frac{1}{\beta} \right) - 2\beta \partial_x \psi = -\partial_x \log m - 2\beta \partial_x \psi \quad (5.53)$$

which means that

$$\Sigma_{SD} = -\log m - 2 \int \beta \partial_x \psi dx \quad (5.54)$$

Substituting this and the usual split for the drift (2.34) in the HJB (5.48) and making use of equation (2.39) we get to

$$2\partial_t \int \beta \partial_x \psi dx + \frac{\beta}{\gamma} (\partial_x \psi)^2 = 0 \quad (5.55)$$

If we consider that $\partial_x \psi = \frac{1}{\beta} \partial_x \int \beta \partial_x \psi dx$ and define

$$\Phi = \int \beta \partial_x \psi dx \quad (5.56)$$

we get to

$$\partial_t \Phi + \frac{1}{2\gamma\beta} (\partial_x \Phi)^2 = 0 \quad (5.57)$$

which is analogous to the equation found for the time depending temperature (5.20).

The equation for transport for m (2.39) in terms of the new field Φ reads

$$\partial_t m + \frac{1}{\gamma} \partial_x \left[\left(\frac{1}{\beta} \partial_x \Phi \right) m \right] = 0 \quad (5.58)$$

Multiplying both sides by $\frac{1}{\beta}$ and operating the following change of variables

$$\frac{1}{\beta} \partial_t = \partial_\tau \quad (5.59)$$

$$\frac{1}{\beta} \partial_x = \partial_\chi \quad (5.60)$$

we can express it as

$$\partial_\tau m + \frac{1}{\gamma} \partial_\chi \left[(\partial_\chi \Phi) m \right] = 0 \quad (5.61)$$

This change of variables is different from the one used for the time dependent temperature (5.19) but is still suitable for mapping equation (5.57) into Burgers equation

$$\partial_\tau \Phi + \frac{1}{2\gamma} (\partial_\chi \Phi)^2 = 0 \quad (5.62)$$

This result confirms that the AMM method leads to Burgers equation when applied to the minimization of entropy production. The possibility of solving the optimization problem for heat through Burgers equation is then limited to the isothermal case.

In analogy to what done in the closing of section 5.2 we try to perform the change of variables (5.59) directly in the SDE (5.26) and consider its related expression for heat. Through Itô lemma we obtain

$$d\mathcal{X}_\tau = \frac{\partial \mathcal{X}}{\partial \mathcal{T}} d\mathcal{T} + \frac{d\mathcal{T}}{\gamma} [b - \frac{1}{2} \partial_x (\frac{1}{\beta})] + \sqrt{\frac{2}{\gamma}} dw_\tau \quad (5.63)$$

$$d\mathcal{T} = \beta dt \quad (5.64)$$

The presence of the time derivative of \mathcal{X} does not allow the change of variables to map heat in what is entropy production. If we instead restrict to the case of temperature depending on space only, the troublesome term disappear. Upon interpretation of $b - \frac{1}{2} \partial_x (\frac{1}{\beta})$ as the modified potential drift¹ we can define the modified heat as

$$\begin{aligned} \langle \delta Q(\mathcal{X}, \mathcal{T}) \rangle &= \langle \int [b - \frac{1}{2} \partial_x (\frac{1}{\beta})] \circ d\mathcal{X}_\tau \rangle = \\ &= \int \langle \beta b^2 + \partial_x b - \beta b \partial_x (\frac{1}{\beta}) + \frac{\beta}{4} [\partial_x (\frac{1}{\beta})]^2 - \frac{1}{2} \partial_x^2 (\frac{1}{\beta}) \rangle \frac{dt}{\gamma} \end{aligned} \quad (5.65)$$

which is exactly the expression for entropy production in the environment of (5.47). Considering also the result of (5.24) we can remark that when temperature does not depend on both space and time the change of variables needed to map the auxiliary problem into Burgers equation transforms the Langevin SDE in in another one of which the expression for heat corresponds to entropy production.

¹The interpretation is not straightforward as we are including also the noise induced drift. See section 5.1

Chapter 6

Minimum entropy cycles

In this chapter we apply the optimization techniques for entropy production developed in chapter 5 to cyclic processes. The possibility of carrying out the optimization for any temperature protocol makes it a suitable tool to study the mesoscopic realizations of any thermodynamic engine. In view of considerations regarding efficiency at maximum power we compute the work yielded by the minimum entropy producing cycles. A first study reveals that zero entropy producing cycles do not provide mean work and therefore a method for minimizing non vanishing entropy is suggested. For simplicity reasons we restrict to the case of temperature depending only on time.

6.1 The role of the cost function

From the findings of section 5.2.2 we know we can make use of Burgers equation for the auxiliary function φ in the modified time τ (5.20) to solve the optimization problem of entropy production in the environment. In analogy to (2.62) we can also express entropy production as

$$\Delta S_{env} = \bar{\Sigma}_{t_0} - \bar{\Sigma}_{t_f} \quad (6.1)$$

which, recalling (5.16) can be expressed as

$$\begin{aligned} \Delta S_{env} &= \\ &= - \int \log [m_{t_i}(a)] m_{t_i}(a) da - 2 \int \varphi_i m_{t_i} dx_i + \int \log [m_{t_f}(x)] m_{t_f}(x) dx - 2 \int \varphi_f m_{t_f} dx_f = \\ &= \sigma(t_0) - \sigma(t_f) + \int \frac{(x_i - x_f)^2}{(\tau_f - \tau_i)} m_{t_f}(x) dx \end{aligned} \quad (6.2)$$

The cost function takes therefore the meaning of overall entropy production during the process. It is the functional $\langle \mathcal{W} \rangle$ of equation (1.56) defined in section 1.1.2. Let us consider it in some more detail

$$\langle \mathcal{W} \rangle = \int \frac{(x_i - x_f)^2}{(\tau_f - \tau_i)} m_{t_i} dx_i \quad (6.3)$$

Intuitively it approaches zero in two cases:

- Infinite denominator which corresponds to a quasi-static transformation
- Vanishing numerator which corresponds to not changing the position

For transport problems between different positions, the vanishing numerator trivially says that if there is no transport there is no cost. For a cycle, the initial and final state always coincides therefore we can have minimum entropy production also in finite time. For such cycle the probability distribution will remain unchanged during the process. To keep the position fixed despite the change in temperature we have to control the potential in a way that would compensate the temperature effects on the distribution. Changing the potential means that we are exerting work on the system. Even though the whole problem might seem trivial since we are considering a cycle that is actually still, it is interesting to compute in details the values of entropy production and work needed or produced.

6.2 Constant probability distribution

Let us consider a harmonic potential with controllable stiffness i.e. an Ornstein Uhlenbeck process with time dependent drift and diffusion coefficient.

$$U = \frac{\lambda_t}{2} x^2 \quad (6.4)$$

In equation (A.12) we have derived the equation of motion for a Gaussian distribution in terms of the control parameters of a quadratic potential. The result holds also in case of a time dependent temperature and, for the example we are considering now, reads

$$\dot{w}_t = 2\left(\frac{1}{\gamma\beta_t} - \frac{\lambda_t}{\gamma} w_t\right) \quad (6.5)$$

It can be easily seen that to balance the effect of the temperature and keep the distribution still, the spring constant must take the form

$$\lambda_t = \frac{1}{\beta_t \sigma^2} \quad (6.6)$$

Where σ^2 is the variance of the initial distribution. The SDE describing the process is now

$$dx_t = -\frac{1}{\gamma\sigma^2\beta_t} x_t dt + \sqrt{\frac{2}{\gamma\beta_t}} dw_t \quad (6.7)$$

With these tools we can proceed to the computation of entropy production and work. From relation (1.58) we have

$$S_{env} = - \int \beta_t \frac{\partial U}{\partial x} \circ dx_t = - \int \beta_t \lambda_t x_t \circ dx_t = -\frac{1}{\sigma^2} \int x_t \circ dx_t \quad (6.8)$$

Which is the integral of a total differential and therefore is zero if initial and final state coincides. The first result is that entropy production is null along every trajectory. Let us now consider the work needed to keep the distribution still. Plugging (6.4) in the definition (2.3) we get

$$W = \int_{t_i}^{t_f} \frac{\partial U}{\partial \lambda_t} \dot{\lambda}_t dt = \frac{1}{2\sigma^2} \int_{t_i}^{t_f} x_t^2 \frac{d}{dt}(\beta_t^{-1}) dt \quad (6.9)$$

The integrand is no longer a total differential and therefore not necessarily zero. If we instead compute the mean work

$$\langle W \rangle = \frac{1}{2\sigma^2} \int_{t_i}^{t_f} \langle x_t^2 \rangle \frac{d}{dt}(\beta_t^{-1}) dt \quad (6.10)$$

since $\langle x_t^2 \rangle = \sigma^2$ and it is constant, we have again the integral of a total differential which, for a cycle, is zero.

We are facing a process where the entropy production is zero for every trajectory and the work needed is zero only in average. To sketch the idea of how certain trajectories can require or produce work without entropy generation let us consider the following hand-waving argument.

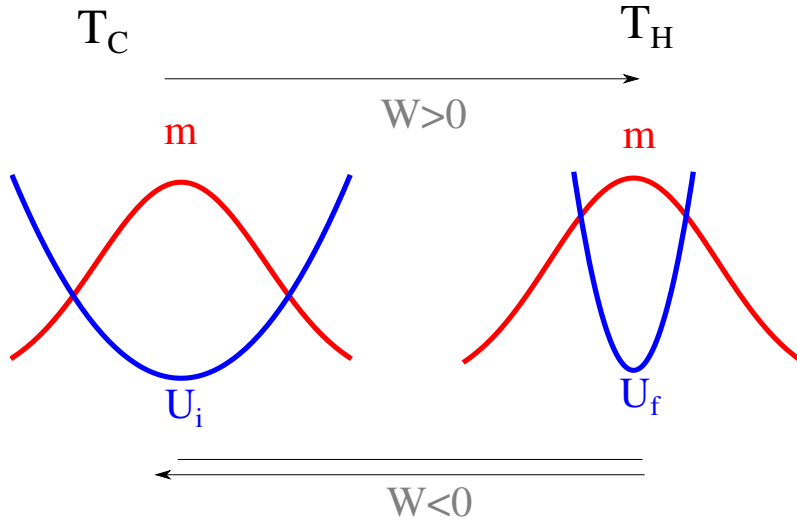


Figure 6.1: Schematic view of a simple temperature protocol

Let us imagine the simplest temperature protocol starting from a low value and monotonically increasing to a high value and then monotonically decreasing back to the starting point. During the increasing temperature phase ($\frac{d}{dt}(\beta_t^{-1}) > 0$) we have to exert work in order to keep still the distribution (shrinking the potential) and during the cooling phase ($\frac{d}{dt}(\beta_t^{-1}) < 0$) we are extracting work (broadening the potential). If we perform the heating step very quickly (with respect to the characteristic time of x_t) we can consider the value

x_t^2 as constant along the integration and easily compute the integral. On the other hand, if the cooling step is carried out slowly, the value of x_t^2 will tend to the variance σ^2 which is also constant. Therefore for this simple protocol we have approximately

$$\begin{aligned} W_{SP} &\approx \frac{1}{2\sigma^2} [x_{t_i}^2 \int_{t_i(\beta_C)}^{t_{i+}(\beta_H)} \frac{d}{dt}(\beta_t^{-1}) dt + \sigma^2 \int_{t_{i+}(\beta_H)}^{t_f(\beta_C)} \frac{d}{dt}(\beta_t^{-1}) dt] = \\ &= \frac{1}{2\sigma^2} (x_{t_i}^2 - \sigma^2)(\beta_H^{-1} - \beta_C^{-1}) \end{aligned} \quad (6.11)$$

Where t_i is the time at which the fast heating step takes place and the integration extremes $t_k(\beta_J)$ explicitly show the temperature of the system at the specified time.

It is again evident that the average is zero. It is also showing that if the initial position is lesser than the standard deviation (which for a Gaussian distribution happens for roughly 68% of the cases) the work is negative and this corresponds to the cycle performing work on the environment. Hence, though in average we are not extracting any work, most trajectories produce work. It is conceivable that certain biological mechanisms are not concerned with mean work but with single trajectory work and might therefore adopt cycles of this sort.

In order to get a clearer picture of the statistics of work we can compute the values of its central moments.

$$\langle W^2 \rangle = \left(\frac{1}{2\sigma^2}\right)^2 \int \int dt_1 dt_2 \frac{d}{dt_1}(\beta_{t_1}^{-1}) \frac{d}{dt_2}(\beta_{t_2}^{-1}) \langle x_{t_1}^2 x_{t_2}^2 \rangle \quad (6.12)$$

$$\langle W^3 \rangle = \left(\frac{1}{2\sigma^2}\right)^3 \int \int \int dt_1 dt_2 dt_3 \frac{d}{dt_1}(\beta_{t_1}^{-1}) \frac{d}{dt_2}(\beta_{t_2}^{-1}) \frac{d}{dt_3}(\beta_{t_3}^{-1}) \langle x_{t_1}^2 x_{t_2}^2 x_{t_3}^2 \rangle \quad (6.13)$$

The values inside the averages can be expressed in terms of products of correlation functions $\langle x_{t_i} x_{t_j} \rangle$ through Wick's theorem. As a preliminary study let us start by computing the values for the specific temperature protocol described above.

The case is much simpler than the general one as the term $x_{t_i}^2$ is outside the integral and, since it always refers to the initial time of the cycle, the correlation reduces to the variance

$$\begin{aligned} \langle W_{SP}^2 \rangle &= \left(\frac{1}{2\sigma^2}\right)^2 \langle (x_{t_i}^2 - \sigma^2)^2 \rangle (\beta_H^{-1} - \beta_C^{-1})^2 = \\ &= \frac{1}{2} \left(\frac{1}{\sigma^2}\right)^2 \sigma^4 (\beta_H^{-1} - \beta_C^{-1})^2 = \frac{1}{2} (\beta_H^{-1} - \beta_C^{-1})^2 \end{aligned} \quad (6.14)$$

Similarly

$$\langle W_{SP}^3 \rangle = (\beta_H^{-1} - \beta_C^{-1})^3 \quad (6.15)$$

The distribution is skewed to the right and therefore asymmetric. Together with the considerations about the probability of being less than one standard deviation away from the mean, this shows that negative values of work during single trajectories are more common but smaller in absolute value than the positive ones.

For the opposite protocol (it sufficient to swap the subscripts of the temperatures in the formula) we have that most of the trajectories need external work to be done. The asymmetry

of the work distribution should not be surprising as it is compensating for the asymmetry of the temperature protocol.

To compute variance and skewness for a generic temperature protocol we need to derive the expression for the time correlation of the process. This is possible for this Ornstein-Uhlenbeck process and yields in case $t_2 > t_1$

$$\langle x_{t_1} x_{t_2} \rangle = e^{-\frac{1}{\gamma\sigma^2} \int_{t_1}^{t_2} \frac{dr}{\beta_r}} (x_{t_0}^2 e^{-\frac{2}{\gamma\sigma^2} \int_{t_0}^{t_1} \frac{dr}{\beta_r}} + \frac{2}{\gamma\sigma^2} \int_{t_0}^{t_1} \frac{ds}{\beta_s} e^{-\frac{2}{\gamma\sigma^2} \int_s^{t_1} \frac{dr}{\beta_r}}) \quad (6.16)$$

If we now consider the time re-parametrization of (5.19) we have that

$$d\tau = \frac{dt}{\beta_t} \quad (6.17)$$

and denoting

$$\begin{aligned} \tau(t_1) &= \tau_1 \\ \tau(t_2) &= \tau_2 \end{aligned} \quad (6.18)$$

$$\langle x_{t_1} x_{t_2} \rangle = (x_{\tau_0}^2 - 1) e^{-\frac{1}{\gamma\sigma^2}(\tau_2 + \tau_1 - 2\tau_0)} + e^{-\frac{|\tau_2 - \tau_1|}{\gamma\sigma^2}} \quad (6.19)$$

Making use of Wick's theorem we can express

$$\langle x_{t_1}^2 x_{t_2}^2 \rangle = \langle x_{t_1}^2 \rangle \langle x_{t_2}^2 \rangle + 2 \langle x_{t_1} x_{t_2} \rangle^2 \quad (6.20)$$

The terms involving the variance give no contribution, as shown before, and expressing it in terms of τ using relation (5.19), the work variance reads

$$\langle W^2 \rangle = \frac{1}{2\sigma^4} \int \int d\tau_1 d\tau_2 \frac{d}{d\tau_1} (\beta_{\tau_1}^{-1}) \frac{d}{d\tau_2} (\beta_{\tau_2}^{-1}) [(x_{\tau_0}^2 - 1) e^{-\frac{1}{\gamma\sigma^2}(\tau_2 + \tau_1 - 2\tau_0)} + e^{-\frac{|\tau_2 - \tau_1|}{\gamma\sigma^2}}]^2 \quad (6.21)$$

For the third moment we have to deal with a term $\langle x_{t_1}^2 x_{t_2}^2 x_{t_3}^2 \rangle$ that we can again expand and, neglecting the terms involving variances, we have

$$\begin{aligned} \langle W^3 \rangle &= \frac{1}{\sigma^6} \int \int \int d\tau_1 d\tau_2 d\tau_3 \frac{d}{d\tau_1} (\beta_{\tau_1}^{-1}) \frac{d}{d\tau_2} (\beta_{\tau_2}^{-1}) \frac{d}{d\tau_3} (\beta_{\tau_3}^{-1}) \cdot \\ &\cdot [(x_{\tau_0}^2 - 1) e^{-\frac{1}{\gamma\sigma^2}(\tau_2 + \tau_1 - 2\tau_0)} + e^{-\frac{|\tau_2 - \tau_1|}{\gamma\sigma^2}}] \cdot \\ &\cdot [(x_{\tau_0}^2 - 1) e^{-\frac{1}{\gamma\sigma^2}(\tau_3 + \tau_1 - 2\tau_0)} + e^{-\frac{|\tau_3 - \tau_1|}{\gamma\sigma^2}}] \cdot \\ &\cdot [(x_{\tau_0}^2 - 1) e^{-\frac{1}{\gamma\sigma^2}(\tau_3 + \tau_2 - 2\tau_0)} + e^{-\frac{|\tau_3 - \tau_2|}{\gamma\sigma^2}}] \end{aligned} \quad (6.22)$$

Its positivity varies according to the chosen temperature protocol.

6.3 A pin-point approach to minimum entropy cycles

The previous section showed that minimizing the cost function over an unconstrained cycle yields a null average work. This can be understood if we consider the thermodynamic role played by the cost function. From equation (6.3) we see that it represents the overall entropy production during a process. The minimization of such functional differs from the one of work. In fact, in order to minimize the work needed to carry out a transformation, one is interested in minimizing the irreversible work (which is always positive and is associated with the entropy production in the environment) and having the most negative reversible work (associated with the difference in entropies at initial and final state $\Delta\sigma$).

$$\langle W \rangle = \langle \Delta U \rangle - T\Delta\sigma + W_{irr} \quad (6.23)$$

The difference is more evident if we consider a cycle (for simplicity we consider the Carnot one)

$$\langle W \rangle_{cycle} = -(T_H - T_C)\Delta\sigma + W_{irr}(T_H) + W_{irr}(T_C) \quad (6.24)$$

To have a negative value for work (corresponding to work done by the cycle) one needs a positive (and large) value of $\Delta\sigma$ which is possible only if the distribution changes during the isothermal steps. One then needs to minimize the entropy production in the environment but to maximize the difference in entropy between intermediate state. Global entropy production is the sum of the two quantities and therefore its unconstrained minimization does not take this feature into consideration. In fact, for a cycle, it reduces to the minimization of entropy production in the environment

$$\Delta S = \sum S_{env} \quad (6.25)$$

which is a positive function. For a non quasi-static process it can vanish, and therefore be minimum, only if the distribution is constant at all times so that the numerator of the cost function (6.3) equals zero. A constant distribution results in zero reversible work as shown in (6.24).

In order to maximize work production using minimum entropy it is necessary to enforce a change of distribution along the cycle and then proceed to minimize entropy under this constraint. This results again in the minimization of the irreversibility of the process and can be solved by the assignment problem minimizing the non vanishing cost function. One of the possible ways to enforce the constraint is to split the cycle into two transformations imposing that the distribution at the end of the first one (i.e. at the beginning of the second one) is different from the initial one.

To shed light on this sort of procedure let us apply this pin-point solution to the example of a quadratic potential with controllable stiffness (6.4). We are going to compute the work associated with the constrained minimum entropy for the two steps and then add them together. We follow a procedure similar to the one described in [3] and in section 4.4 and express work in terms of the probability variance and derive its optimal functional form.

From the optimization of entropy production into the environment we know (cf equation 5.20) that the optimal transport problem is mapped into Burgers equation in the transformed time τ . As a result, the evolution of the standard deviation minimizing entropy production is linear in τ and we can write

$$\sigma_\tau = \sigma_{t_i} \left(1 + \frac{\tau - \tau_i}{\tau_f - \tau_i} \left(\frac{\sigma_{t_f}}{\sigma_{t_i}} - 1 \right) \right) \quad (6.26)$$

$$w_\tau = w_{t_i} \left(1 + \frac{\tau - \tau_i}{\tau_f - \tau_i} \left(\sqrt{\frac{w_{t_f}}{w_{t_i}}} - 1 \right) \right)^2 \quad (6.27)$$

we can also compute

$$\frac{d}{dt} w_\tau = 2w_{t_i} \left(1 + \frac{\tau - \tau_i}{\tau_f - \tau_i} \left(\sqrt{\frac{w_{t_f}}{w_{t_i}}} - 1 \right) \right) \left(\sqrt{\frac{w_{t_f}}{w_{t_i}}} - 1 \right) \frac{1}{\beta_t(\tau_f - \tau_i)} \quad (6.28)$$

Where the last factor comes from the relation in eq (5.19) and equation (6.5) enforces boundary conditions.

Inverting equation (6.5) we can write down the expression for the optimal control parameter.

$$\lambda_t = \gamma \frac{\frac{2}{\gamma\beta_t} - \frac{d}{dt} w_\tau}{2w_\tau} = \frac{1}{\beta w_\tau} - \frac{\gamma}{2} \frac{\frac{d}{dt} w_\tau}{w_\tau} \quad (6.29)$$

Its time derivative reads

$$\frac{d\lambda}{dt} = \frac{d}{dt} \left(\frac{1}{\beta w_\tau} \right) - \frac{\gamma}{2} \frac{\frac{d^2 w_\tau}{dt^2}}{w_\tau} + \frac{\gamma}{2} \left(\frac{\frac{d}{dt} w_\tau}{w_\tau} \right)^2 \quad (6.30)$$

From equation (6.29) we can express the second order derivative of the variance in terms of the control parameter as

$$\frac{d^2 w_\tau}{dt^2} = 2 \frac{d}{dt} \left(\frac{1}{\beta \gamma} \right) - 2 \frac{d}{dt} (\lambda_t w_\tau) \quad (6.31)$$

To compute the mean work needed to run the first transformation of the cycle we make use of the definition 2.3 noting that for this potential

$$\frac{\partial U}{\partial \lambda} = \frac{1}{2} x_t^2 \quad (6.32)$$

The expression for the average work for the protocol minimizing entropy production is then

$$\begin{aligned} \langle W \rangle &= \frac{1}{2} \int_{t_i}^{t_f} \frac{d\lambda}{dt} w_t dt = \\ &= \frac{1}{2} \int_{t_i}^{t_f} w_t \frac{d}{dt} \left(\frac{1}{w_t \beta_t} \right) dt - \frac{1}{2} \left[\frac{1}{\beta_t} \right]_{t_i}^{t_f} + \frac{1}{2} [\lambda_t w_t]_{t_i}^{t_f} + \frac{\gamma (\sqrt{w_{t_f}} - \sqrt{w_{t_i}})^2}{(\tau_f - \tau_i)^2} \int_{t_i}^{t_f} \frac{1}{\beta^2} dt \end{aligned} \quad (6.33)$$

and integrating by part the first term we get to

$$\langle W \rangle = -\frac{1}{2} \int_{t_i}^{t_f} \frac{1}{w_t \beta_t} \frac{d}{dt}(w_t) dt + \frac{1}{2} [\lambda_t w_t]_{t_i}^{t_f} + \frac{\gamma(\sqrt{w_{t_f}} - \sqrt{w_{t_i}})^2}{(\tau_f - \tau_i)^2} \int_{t_i}^{t_f} \frac{1}{\beta^2} dt \quad (6.34)$$

In order to understand the meaning of each term and to make an explicit connection with our previous results let us consider the isothermal case for which $\tau = \frac{t}{\beta}$. We then get to

$$\langle W_{iso} \rangle = \frac{1}{2\beta} \log \frac{w_{t_i}}{w_{t_f}} + \frac{1}{2} [\lambda_t w_t]_{t_i}^{t_f} + \frac{\gamma}{(t_f - t_i)} (\sqrt{w_{t_f}} - \sqrt{w_{t_i}})^2 \quad (6.35)$$

which coincides with (4.7) for an origin centered potential and can be rewritten as

$$\langle W_{iso} \rangle = -T \Delta \sigma + \langle \Delta U \rangle + W_{irr} \quad (6.36)$$

For an isothermal step, entropy minimization yields the same result of work optimization before the differentiation over the final distribution coefficients¹. This equivalence shows that, upon similar interpretation of the adiabatic steps, the formalism can be used to solve the Carnot cycle described in [3].

In order to conclude the computation of the work associated with the cycle we have to sum the two steps. We denote the initial variance and time as w_{t_0} , t_0 the intermediate ones as w_{t_1} , t_1 and the final ones as $w_{t_2} = w_{t_0}$, t_2 . Upon piecewise definition of the variance

$$w_\tau = \begin{cases} w_0(1 + \frac{\tau - \tau_0}{\tau_1 - \tau_0}(\sqrt{\frac{w_1}{w_0}} - 1))^2 & \text{for } \tau_0 < \tau < \tau_1 \\ w_1(1 + \frac{\tau - \tau_1}{\tau_2 - \tau_1}(\sqrt{\frac{w_0}{w_1}} - 1))^2 & \text{for } \tau_1 < \tau < \tau_2 \end{cases}$$

We obtain

$$\begin{aligned} \langle W_{cycle} \rangle = & - \int_{t_0}^{t_2} \frac{1}{\beta} \frac{d}{dt} (\log w_\tau) dt + \\ & + \gamma (\sqrt{w_1} - \sqrt{w_0})^2 \left(\frac{1}{(\tau_1 - \tau_0)^2} \int_{t_0}^{t_1} \frac{1}{\beta_t^2} dt + \frac{1}{(\tau_2 - \tau_1)^2} \int_{t_1}^{t_2} \frac{1}{\beta_t^2} dt \right) \end{aligned} \quad (6.37)$$

which can be evaluated for any temperature protocol and consequently used to compute the work production of any thermodynamic cycle. Some further care is required in the definition and interpretation of instantaneous transformations with finite changes in temperatures such as the adiabatic steps in [3].

¹Entropy optimization is analogous to heat optimization and can be performed between given final and initial states. A further optimization on the final distribution is therefore meaningless.

Concluding Remarks

The main finding of the present work is that for entropy production in the environment the AMM optimization method leads to Burgers equation also for non constant temperatures whereas for released heat only in the isothermal case. When temperature is constant, heat differs from entropy production just by a constant factor and the two optimizations are fundamentally identical. These considerations suggest that it is entropy production in the environment which is intrinsically connected with Burgers equation.

The application of the generalized method for entropy production optimization to thermodynamic cycles with asymmetric temperature protocols shows the existence of isentropic cycles for which the average work is zero but most of the trajectories produce work. Moreover, explicit use of Burgers equation is made to solve pin-pointed formalizations of thermodynamic cycles. This represents a tool for the computation of work production associated with minimum entropy producing for any temperature profile. With this expression at hand one can look for which is the temperature protocol yielding the maximum amount of work.

Appendix A

Gaussianity of the distribution for linear drifts

Let us consider a distribution and a quadratic potential at time t .

$$U_t(x) = \frac{c_t}{2}(x - \mu_t)^2 \quad (\text{A.1})$$

$$m_t(x) = \sqrt{\frac{\beta r_t}{2\pi}} e^{-\beta r_t \frac{(x - q_t)^2}{2}} \quad (\text{A.2})$$

$$b_t(x) = -c_t(x - \mu_t) \quad (\text{A.3})$$

The Fokker-Planck equation for the considered process reads:

$$\partial_t m = \partial_x \left[\frac{c_t}{\gamma} (x - \mu_t) m \right] + \partial_x^2 (Dm) \quad (\text{A.4})$$

Considering that for the Gaussian distribution (A.1)

$$\partial_x m = -\beta r_t (x - q_t) m \quad (\text{A.5})$$

$$\partial_x^2 m = (\beta r_t)^2 (x - q_t)^2 m - \beta r_t m \quad (\text{A.6})$$

$$\begin{aligned} \partial_t m &= \left\{ \frac{\dot{r}_t}{2\sqrt{r_t}} \sqrt{\frac{\beta}{2\pi}} - \sqrt{\frac{r_t \beta}{2\pi}} \left[\frac{\beta \dot{r}_t}{2} (x - q_t)^2 + (x - q_t) \beta r_t \dot{q}_t \right] \right\} e^{-\beta r_t \frac{(x - q_t)^2}{2}} = \\ &= \left[\frac{\dot{r}_t}{2r_t} - \frac{\beta \dot{r}_t}{2} (x - q_t)^2 - (x - q_t) \beta r_t \dot{q}_t \right] m \end{aligned} \quad (\text{A.7})$$

We get to

$$\begin{aligned} &\left[\frac{\dot{r}_t}{2r_t} - \frac{\beta \dot{r}_t}{2} (x - q_t)^2 + (x - q_t) \beta r_t \dot{q}_t - \frac{c_t}{\gamma} + \right. \\ &\left. \frac{c_t}{\gamma} \beta r_t (x - q_t)(x - \mu_t) - D(\beta r_t)^2 (x - q_t)^2 + D\beta r_t \right] m = 0 \end{aligned} \quad (\text{A.8})$$

The equality must be satisfied by the terms in every order of x . To explicitly write the requirement let us express $(x - \mu_t) = (x - q_t) + (q_t - \mu_t)$ and factorize in $(x - q_t)$:

$$(x - q_t)^2 \left(-\frac{\beta \dot{r}}{2} - D\beta^2 r^2 + \frac{c_t}{\gamma} \beta r \right) + \beta r_t (x - q_t) \left(-\dot{q}_t + \frac{c_t}{\gamma} (q_t - \mu_t) \right) + \left(\frac{\dot{r}}{2r} + D\beta r - \frac{c_t}{\gamma} \right) = 0 \quad (\text{A.9})$$

The conditions on the second order term and on the constants are the same. The solution therefore reduces to two first order ODEs.

$$\dot{q}_t = \frac{c_t}{\gamma} (\mu_t - q_t) \quad (\text{A.10})$$

describing the evolution of the mean of the distribution which tracks the center of the potential with a velocity proportional to the stiffness of the potential divided by the viscous friction. And

$$\dot{r}_t = -2D\beta r_t^2 + 2\frac{c_t}{\gamma} r_t \quad (\text{A.11})$$

Which makes more immediate sense if expressed in terms of the variance of the distribution $w_t = \frac{1}{\beta r_t}$

$$\dot{w}_t = 2 \left(D - \frac{c_t}{\gamma} w \right) \quad (\text{A.12})$$

Which in the stationary case gives $w = \frac{\gamma D}{c}$ and in the absence of the potential (i.e. $c = 0$) reduces to ordinary diffusive case $w_t = 2\dot{D}t$.

Bibliography

- [1] Aurell E., Mejia-Monasterio C., Muratore-Ginanneschi P. *Optimal protocols and optimal transport in stochastic thermodynamics*, Phys. Rev. Lett. 106, 250601 (2011)
- [2] T. Schmiedl and U. Seifert, *Optimal finite-time processes in stochastic thermodynamics*, Phys. Rev. Lett. 98, 108301 (2007)
- [3] T. Schmiedl and U. Seifert *Efficiency at maximum power: An analytically solvable model for stochastic heat engines* EPL 81 20003 (2008)
- [4] Sekimoto, K. *Langevin equation and thermodynamics*, Prog. Theor. Phys. Suppl. 180, 17 (1998)
- [5] Sekimoto K., *Stochastic Energetics* , Lect. Notes Phys. 799 (Springer, Berlin Heidelberg 2010)
- [6] Ch  trite R., Gaw  dzki K., *Fluctuation relations for diffusion processes* Communications in Mathematical Physics, 2007, 282, 469-518
- [7] Fleming, W. H. & Mete Soner, H. *Controlled Markov processes and viscosity solutions* Birkhuser, 2006, 25, 428
- [8] U. Marini Bettolo Marconi; A. Puglisi; L. Rondoni; A. Vulpiani *Fluctuation-Dissipation: Response Theory in Statistical Physics* (2008). Physics Reports 461 (46): 111195.
- [9] J  licher, F and Ajdari, A and Prost, J *Modeling molecular motors* Rev. Mod. Phys 69: 1269–1282 (1997)
- [10] Blickle, V *Private communication*
- [11] C.W. Gardiner, *Handbook of Stochastic Methods*, New York: Springer, 1983
- [12] J.-D. Benamou, Y. Brenier: *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. 84 (1999), 375-393
- [13] Y. Brenier, U. Frisch, M. Henon, G. Loeper, S. Matarrese, R. Mohayaee, and A. Sobolevskii, *Early Universe as a convex oprimization problem* Mon. Not. R. Astron. Soc. 346, 501 (2003)

- [14] F. Ritort, *Nonequilibrium fluctuations in small systems: from physics to biology*, Advances in Chemical Physics, vol. 137, 31-123 (2008). Ed. Stuart. A. Rice, Wiley publications
- [15] V. Blickle, T. Speck, L. Helden, U. Seifert, and C. Bechinger, *Thermodynamics of a Colloidal Particle in a Time-Dependent Nonharmonic Potential*, Phys. Rev. Lett. 96, 070603 (2006)
- [16] J. Liphardt, S. Dumont, SB. Smith, I. Tinoco, Jr. and CJ. Bustamante, *Equilibrium Information from Nonequilibrium Measurements in an Experimental Test of Jarzynski's Equality*, Science 296, 1832-1835 (2002)