ABSTRACT

Title of dissertation: Optimal Control of Heat Engines in Non-equilibrium Statistical Mechanics

Yunlong Huang, 2017

Dissertation directed by: Professor P. S. Krishnaprasad
Dept. of Electrical & Computer Engineering

A heat engine is a cyclically operated statistical mechanical system which converts heat supply from a heat bath into mechanical work. The heat engine is operated by varying the system parameter. As it is operated in finite time, this non-equilibrium statistical mechanical system is a dissipative system. In this dissertation, our research focuses on two heat engines: one is a stochastic oscillator and the other is a capacitor connected to a Nyquist-Johnson resistor (a stochastically driven resistor-capacitor circuit). In the stochastic oscillator, by varying the stiffness of the potential well, the system can convert heat to mechanical work. In the resistor-capacitor circuit, the output of mechanical work is due to the change of the capacitance of the capacitor. These two heat engines are parametrically-controlled. A path in the parameter space of a heat engine is termed as a protocol.

In the first chapter of this dissertation, under the near-equilibrium assumption, with the help of linear response theory, fluctuation theorem and stochastic thermodynamics, we consider an inverse diffusion tensor in the parameter space of a heat engine. The inverse diffusion tensor of the stochastic oscillator induces a hyperbolic

space structure in the parameter space composed of the stiffness of the potential well and the inverse temperature of the heat bath. The inverse diffusion tensor of the resistor-capacitor circuit induces a Euclidean space structure in the parameter space composed of the capacitance of the capacitor and the inverse temperature of the heat bath. The average dissipation rate of a heat engine is given by a quadratic form (with a positive-definite inverse diffusion tensor) on the tangent space of the system parameter.

Along a finite-time protocol of a heat engine, besides the energy dissipation, there are two auxiliary quantities of interest: one is the extracted work of the heat engine and the other is the total heat supply from the bath to the engine. These two quantities are fundamental to the analysis of the efficiency of a heat engine. In Chapter 2, combining the energy dissipation and the extracted work of a heat engine, we introduce sub-Riemannian geometry structures underlying both heat engines.

In Chapter 3, after defining efficiency of a heat engine, we show the equivalence between an optimal control problem in the sub-Riemannian geometry of the heat engine and the problem of maximizing the efficiency of the heat engine. In this way, we bring geometric control theory to non-equilibrium statistical mechanics. In particular, we explicate the relation between conjugate point theory and the working loops of a heat engine. As a related calculation, we solve the isoperimetric problem in hyperbolic space as an optimal control problem in Chapter 4.

Based on the theoretical analysis in the first four chapters, in the final chapter of the dissertation, we adopt level set methods, mid-point approximation and shooting method to design maximum-efficiency working loops of both heat engines.

The associated efficiencies of these protocols are computed.

Optimal Control of Heat Engines in Non-equilibrium Statistical Mechanics

by

Yunlong Huang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:
Professor P. S. Krishnaprasad, Chair/Advisor
Professor S. Marcus
Professor A. Tits
Professor J. Rosenberg
Professor C. Jarzynski

Acknowledgments

First and foremost I would like to thank my advisor, Professor Krishnaprasad for giving me an invaluable opportunity to work on this research project over the past four years. It has been a pleasure to work with and learn from such an extraordinary individual. I would like to thank Prof. Steve Marcus, Prof. Andre Tits, Prof. Jonathan M. Rosenberg and Prof. Christopher Jarzynski for joining my dissertation committee, and for their feedback on this work.

I owe my deepest thanks to my family - my mother and father who have always supported me. Words cannot express the gratitude I owe them.

This research was supported by the ARL/ARO MURI Program Grant No. W911NF-13-1-0390.

Table of Contents

List of Tables v

List of Figures v

List of Notations vii

0 Introduction 1
- 0.1 Background . . . 1
- 0.2 Motivation . . . 2
  - 0.2.1 Introduction to sub-Riemannian geometry . . . 3
  - 0.2.2 Pontryagin's maximum principle . . . 4
- 0.3 Main results . . . 5

1 Inverse Diffusion Tensor of A Statistical Mechanical System 7
- 1.1 A Hamiltonian system in contact with a heat bath . . . 7
- 1.2 A port-Hamiltonian system in contact with a heat bath . . . 12
- 1.3 Linear response theory and inverse diffusion tensor . . . 16
- 1.4 Inverse diffusion tensor of the resistor-capacitor system . . . 23
  - 1.4.1 The stochastic thermodynamics . . . 23
  - 1.4.2 The fluctuation theorem . . . 27
  - 1.4.3 Dissipation from different processes . . . 35

2 Sub-Riemannian Geometry Structures of the Heat Engines 37
- 2.1 The extracted work and total heat supply . . . 37
- 2.2 Correction terms for extracted mechanical power of the heat engines . 41
- 2.3 Sub-Riemannian geometry structures of the heat engines . . . 46

3 Design of a Heat Engine: Optimal Control Approach 50
- 3.1 Optimal control of the heat engines . . . 50
- 3.2 Maximum efficiency protocols of the heat engines . . . 59
- 3.3 A working cycle for a heat engine and conjugate point theory . . . 65

4 Optimal Control in Hyperbolic Space 69
4.1 Symmetry in Poincaré upper half plane 70
4.2 Isoperimetric problem in Poincaré upper half plane 71
4.3 Isoperimetric equality in Poincaré upper half plane 74

5 Numerical Design of Working Loops 79
5.1 Reachable set of a heat engine 80
5.1.1 A stochastic oscillator example 85
5.1.2 A resistor-capacitor circuit example 90
5.2 Trajectory reconstruction 95
5.2.1 Numerical results 96
5.3 Efficiencies of the heat engines 105

6 Conclusions and Directions for Future Research 108

Bibliography 110

List of Tables

5.1 Working loop end-points of the stochastic oscillator 90
5.2 Working loop end-points of the resistor-capacitor circuit 94
5.3 Reconstructed working loops through mid-point approximation 97
5.4 Reconstructed working loops through shooting method 99
5.5 Reconstructed working loop through mid-point approximation 102
5.6 Reconstructed working loop through shooting method 103
5.7 Efficiencies of maximum efficiency working loops 107
5.8 Efficiency of maximum efficiency working loop 107

List of Figures

1.1 RC circuit 12

5.1 Reachable set of a stochastic oscillator in 3D 87
5.2 Reachable set in the space of $(\tilde{y}, \tilde{\psi}_1)$ 88
5.3 Reachable set in the space of $(\tilde{x}, \tilde{\psi}_1)$ 88
5.4 Reachable set in the space of $(\tilde{x}, \tilde{y})$ 89
5.5 Reachable set of the RC-circuit in 3D 92
5.6 Reachable set in the space of $(\tilde{\beta}, \tilde{\psi}_2)$ 93
5.7 Reachable set in the space of $(\tilde{c}, \tilde{\psi}_2)$ 93
5.8 Reachable set in the space of $(\tilde{c}, \tilde{\beta})$ 94
5.9 Reconstructed working loop through mid-point approximation 98
5.10 Reconstructed working loop through shooting method 99

5.11 Projection of reconstructed working loop of length 2.9525 100
5.12 Projection of reconstructed working loop of length 1.1126 101
5.13 Reconstructed working loop through mid-point approximation 102
5.14 Reconstructed working loop through shooting method 103
5.15 Projection of reconstructed working loop 104

List of Notations

m	Mass of the stochastic oscillator
ζ	Cartesian frictional constant in the stochastic oscillator
$B(t)$	Brownian motion at time t
δ	Dirac delta function
λ	System parameter of a heat engine
$\omega(t)$	$\mathbb{R}^d-$valued Wiener-process at time t
$\rho_{eq,\lambda}$	Probability density at equilibrium with system parameter fixed at value of λ
$\rho_{neq,\lambda}$	Probability density at non-equilibrium with system parameter at value of λ
$\langle\cdot\rangle$	Average evaluated with non-equilibrium distribution
$\langle\cdot\rangle_{eq,\lambda}$	Average evaluated with the equilibrium distribution $\rho_{eq,\lambda}$
$\mathcal{H}(\rho)$	Thermodynamical entropy of a system associated to density ρ
U	Internal energy of a system
U_{neq}	Internal energy of a system at non-equilibrium
F	Free energy of a system
F_{eq}	Free energy of a system at equilibrium
$D(\cdot\\|\cdot)$	Relative entropy between probability densities
$\\|\cdot\\|_\infty$	Supremum norm of a vector
N	A manifold
$q \in N$	A point q on the manifold N
T_qN	Tangent space at the point q on the manifold N
T_q^*N	Cotangent space at point q on the manifold N
T^*N	Cotangent bundle of the manifold N
$\langle\cdot,\cdot\rangle_d$	Duality pairing between a cotangent vector and a tangent vector

Chapter 0: Introduction

0.1 Background

The second law of thermodynamics formulated the relation between dissipation and irreversibility of a statistical mechanical system. With development of the research in non-equilibrium statistical mechanics, fluctuation theorems [1] [2] and stochastic thermodynamics [3] provide more thorough and complete understanding of dissipation and irreversibility of a statistical mechanical system. Among the fluctuation theorems, the core is to compute the irreversibility between a trajectory of the forward process of a statistical mechanical system and the time-reversed trajectory of the backward process of the system. This irreversibility is measured by the relative entropy defined in the path space of the random process associated to the system. The irreversibility is also described as the entropy production in statistical mechanics. From the view of stochastic thermodynamics, by defining the heat released by the system, work done on the system with Stratonovich integral, the first and second law of thermodynamics are proved and quantitively analyzed. These two methodologies - fluctuation theorem and stochastic thermodynamics can only be proved to be consistent with each other in some special cases [4].

0.2 Motivation

Geometric analysis appeared in the study of thermodynamics long ago (as Carathéodory principle for second law [5]). Later researchers sought to connect stochastic control and filtering to thermodynamics [6] [7] [8]. Specifically in [6], a thermodynamic system consisting of a controllable capacitor and a Nyquist-Johnson resistor [9] [10] in contact with two heat baths serves as a model of a heat engine. For a thermodynamic system, a protocol is a path in the space of its system parameter. In [6] the maximum efficiency protocol is quasi-static (Carnot cycle) and hence dissipation is not included. In this dissertation, we seek finite-time protocols of maximum efficiency, accounting for dissipation arising from the non-equilibrium setting. This requires explicit (approximate) expressions for dissipation.

Much more recently, in the work [11], the concept of an inverse diffusion tensor was first proposed. In the near-equilibrium regime, with linear response analysis [12] , this inverse diffusion tensor can describe the dissipation along a protocol in the space of system parameter. In the paper [13], the inverse diffusion tensor of a stochastic oscillator is computed as an example. The work of this thesis is motivated by these two papers [6] and [13] coming from different research fields – one from control theory and the other from statistical mechanics. In this thesis, there are two heat engine models: one is the stochastic oscillator and the other is the resistor-capacitor circuit. By computing the inverse diffusion tensor of each heat engine, we investigate optimal control of the heat engines.

In either the stochastic oscillator or the resistor-capacitor circuit, the inverse diffusion tensor in the space of system parameter induces a Riemannian manifold structure. By considering the extracted work of a heat engine, the space of the system parameter is enlarged into a new manifold. We would like to investigate the existence of a sub-Riemannian manifold structure in this enlarged manifold.

0.2.1 Introduction to sub-Riemannian geometry

The knowledge of sub-Riemmanian geometry is extensively explored in the work [14]. Here, we just introduce the elementary concepts on it.

Definition 0.2.1. *A sub-Riemannian geometry on a manifold N consists of a distribution, which is to say a vector subbundle $\Delta \subset TN$ of the tangent bundle of N, together with a fiber inner-product $\langle \cdot, \cdot \rangle$ on this subbundle.*

We call Δ the horizontal distribution. An object such as a vector field or a curve on N is called horizontal if it is tangent to Δ.

Given a collection $\{X_a\}$ of vector fields, form its *Lie hull*, the collection of all vector fields $\{X_a, [X_b, X_c], [X_a, [X_b, X_c]], \dots\}$ generated by Lie brackets of the X_a. The collection $\{X_a\}$ is *bracket generating* if this Lie hull spans the whole tangent bundle.

Definition 0.2.2. *A distribution $\Delta \subset TN$ is called bracket generating if any local horizontal frame $\{X_a\}$ for the distribution is bracket generating (over this domain).*

Theorem 0.2.3 (Chow-Rashevski theorem)**.** *If a distribution $\Delta \subset TN$ is bracket generating then the set of points that can be connected to $q \in N$ by a horizontal path is the connected component of N containing q.*

If N is a manifold with a bracket-generating distribution then any point q of N is contained in a neighborhood U, such that every $\tilde{q} \in U$ can be connected to q. With this $\Delta-$connectivity and the fiber inner product, we can say there is a *sub-Riemannian manifold structure* in the enlarged space.

0.2.2 Pontryagin's maximum principle

On the sub-Riemannian manifolds associated to our heat engines, we would like to solve associated optimal control problem by appeal to Pontryagin's maximum principle. In this subsection, we state the Pontryagin's maximum principle for fixed-endpoint control problem. For more details on Pontryagin's maximum principle and its application to different types of problems, refer to [15], and reference therein.

Control systems take the form (following [15])

$$\dot{x} = f(x, u),\ x(t_0) = x_0 \tag{0.1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control, $t \in \mathbb{R}$ is time, t_0 is the initial time and x_0 is the initial state. Both x and u are functions of time. We will consider the problem of minimizing a cost functional of the form

$$J(u) \equiv \int_{t_0}^{t_f} L(x(t), u(t))dt \tag{0.2}$$

where t_f and $x(t_f)$ are the *final* time and state, $L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is the Lagrangian. Pontryagin's maximum principle for above control system to minimize the cost functional (0.2) is:

Theorem 0.2.4 (Maximum principle for fixed-endpoint problem)**.** *Let $u^* : [t_0, t_f] \to \mathbb{R}^m$ be an optimal control and let $x^* : [t_0, t_f] \to \mathbb{R}^n$ be the corresponding optimal state* *trajectory. Then there exists a function* $p^* : [t_0, t_f] \to \mathbb{R}^n$ *and a constant* p^* *satisfying* $(p_0^*, p^*(t)) \neq (0, 0)$ *for all* $t \in [t_0, t_f]$ *and having the following properties:*

1. x^* *and* p^* *satisfy the canonical equations*

$$\dot{x}^* = H_p(x^*, u^*, p^*, p_0^*), \tag{0.3}$$

$$\dot{y}^* = -H_x(x^*, u^*, p^*, p_0^*)$$

with the boundary conditions $x^*(t_0) = x_0$ *and* $x^*(t_f) = x_1$*, where the Hamiltonian* $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ *is defined as*

$$H(x, u, p, p_0) \equiv \langle p, f(x, u) \rangle + p_0 L(x, u) \tag{0.4}$$

where $\langle \cdot, \cdot \rangle$ *in above definition is the Euclidean inner product.*

2. *For each fixed* t*, the function* $u \to H(x^*(t), u, p^*(t), p_0^*)$ *has a global maximum at* $u = u^*(t)$*, i.e., the inequality*

$$H(x^*(t), u^*(t), p^*(t), p_0^*) \geq H(x^*(t), u, p^*(t), p_0^*) \tag{0.5}$$

holds for all $t \in [t_0, t_f]$ *and all* $u \in \mathbb{R}^m$*.*

3. $\forall t \in [t_0, t_f]$, $H(x^*(t), u^*(t), p^*(t), p_0^*)$ *is a constant.*

4. *When the extremum satisfies* $p_0^* \neq 0$*, then we can take it to be equal to* -1 *and this is the case of normality.*

0.3 Main results

In Chapter 1, we first discuss the inverse diffusion tensor of the stochastic oscillator from the point of view of system theory. Moreover, in the case of the resistor-capacitor circuit, by introducing the ideas of fluctuation theorem and stochastic thermodynamics, we extend the concept of inverse diffusion tensor to have it describe the dissipation of the resistor-capacitor circuit appropriately. We also show the consistency between fluctuation theorem and stochastic thermodynamics in the circuit model.

In Chapter 2, with the inverse diffusion tensors from Chapter 1, we reveal the underlying sub-Riemannian geometry structures of the heat engines. As a consequence, in Chapter 3, in both engines, we show the equivalence between finding maximum efficiency working loops and the associated optimal control problems, which are analytically solved by appeal to the maximum principle. The same methodology is shown to recover the explicit solution to the isoperimetric problem in Poincaré upper half plane in Chapter 4. In Chapter 5, following the results of Chapter 3, we design maximum efficiency working loops numerically and the associated efficiencies are computed.

Chapter 1: Inverse Diffusion Tensor of A Statistical Mechanical System

A heat engine is a cyclically operated statistical mechanical system which converts heat supply from a heat bath into mechanical work. As it is operated in finite time, this non-equilibrium statistical mechanical system is a dissipative system. In this thesis, our research focuses on two heat engines: one is a stochastic oscillator in contact with a heat bath and the other is a capacitor connected to a Nyquist-Johnson resistor [9] [10] which is a stochastically-driven resistor-capacitor circuit.

1.1 A Hamiltonian system in contact with a heat bath

In this section, we will introduce a stochastic oscillator in contact with a heat bath, as an example of a Hamiltonian system in contact with a heat bath.

A stochastic oscillator with position ξ_1 and momentum ξ_2 is driven by Brownian motion,

$$d\xi_1 = \frac{\xi_2}{m} dt$$
$$d\xi_2 = -\zeta \frac{\xi_2}{m} dt - k\xi_1 dt + \sqrt{\frac{2\zeta}{\beta}} dB(t) \tag{1.1}$$

where $B(t)$ is standard Brownian motion, such that

$$\langle dB(t) \rangle = 0 \quad \langle dB(t) dB(t') \rangle = \delta(t - t') dt' \tag{1.2}$$

We let $\xi = (\xi_1, \xi_2)$ denote the state of the oscillator. Angled brackets indicate statistical average of Brownian motion and δ is Dirac delta function. ζ is the Cartesian friction coefficient and m is the mass of the oscillator. These are two constants. At time t, $\beta(t) = (k_B T(t))^{-1}$ is the inverse temperature of the heat bath in natural units with Boltzmann constant k_B and $k(t)$ is the stiffness of the potential well. The pair $\lambda = (\beta, k)$ is the system parameter.

The stochastic oscillator in contact with a heat bath is an example of a Hamiltonian system in contact with a heat bath,

$$d\xi = (J - G) \frac{\partial H(\xi, \lambda)}{\partial \xi} dt + \sqrt{\frac{2}{\beta}} G^{1/2} d\omega(t) \tag{1.3}$$

In (1.3), ξ is the $d-$dimensional state variable and $\omega(t)$ denotes an $\mathbb{R}^d$-valued standard Wiener-process while J is a skew-symmetric matrix ($J = -J^\intercal$) and G is a $d-$dimensional positive-semidefinite matrix, such that

$$G = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{G} \end{pmatrix} \tag{1.4}$$

where $\tilde{G}$ is a symmetric positive-definite matrix ($\tilde{G} = \tilde{G}_1 \tilde{G}_1^\intercal$) and the dimension of $\tilde{G}$ is less than or equal to d. In the case that the dimension of $\tilde{G}$ is less than d, $G^{1/2}$

is $\begin{pmatrix} 0 & 0 \\ 0 & \tilde{G}_1 \end{pmatrix}$. In the stochastic oscillator, $d = 2$ and

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \; G = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.5}$$

As a dissipative dynamical system [16], we consider the unitless term $S = \beta H$ as the storage function, where H is the Hamiltonian of the system which is a function of the state ξ and the system parameter λ. Correspondingly, the vector $\dot{\lambda}$ is the *input*, $X \equiv \frac{\partial S}{\partial \lambda}$ is the *output* and $w = \dot{\lambda}^T X$ is the *supply rate*. For a statistical mechanical system, its state probability density at equilibrium

$$\rho_{eq,\lambda}(\xi) \propto e^{-\beta H} \tag{1.6}$$

when the system is held at fixed control parameter λ. For us, dissipativeness has to do with the transition of a system which has been disturbed from equilibrium to a new equilibrium. When λ is changed to a new value, the oscillator undergoes a transition to a new equilibrium probability distribution. In general this process results in dissipation of energy.

We define the *average* dissipation rate (in the spirit of [16]) $d = \langle w \rangle - \frac{d}{dt} \langle S \rangle_{eq,\lambda}$, where $\langle \, \rangle$ is the average evaluated with non-equilibrium distribution and $\langle \, \rangle_{eq,\lambda}$ is the average evaluated with the equilibrium distribution of the statistical mechanical system with system parameter fixed at the value of λ. For the stochastic oscillator, its Hamiltonian $H(\xi, \lambda) = \frac{\xi_2^2}{2m} + \frac{k\xi_1^2}{2}$. Given that $S(t, \xi) \in \mathcal{C}^2([0, \infty) \times \mathbb{R}^2)$, by Ito's formula [17],

$$
\begin{aligned}
dS &= \frac{\partial S}{\partial t}dt + \left(\frac{\partial S}{\partial \xi_1}\right)d\xi_1 + \left(\frac{\partial S}{\partial \xi_2}\right)d\xi_2 \\
&+ \frac{1}{2}\mathrm{tr}\left(\begin{pmatrix} \frac{\partial^2 S}{\partial \xi_1^2} & \frac{\partial^2 S}{\partial \xi_1 \partial \xi_2} \\ \frac{\partial^2 S}{\partial \xi_2 \partial \xi_1} & \frac{\partial^2 S}{\partial \xi_2^2} \end{pmatrix}\left(\begin{pmatrix} 0 \\ \sqrt{\frac{2\zeta}{\beta}} \end{pmatrix}\begin{pmatrix} 0 \\ \sqrt{\frac{2\zeta}{\beta}} \end{pmatrix}^{\intercal}\right)\right)dt \qquad (1.7) \\
&= \left(H(\xi,\lambda)\dot{\beta} + \beta\dot{k}\frac{\xi_1^2}{2} + \frac{\zeta}{m} - \frac{\beta\zeta\xi_2^2}{m^2}\right)dt + \frac{\sqrt{2\zeta\beta}}{m}dB(t)
\end{aligned}
$$

By the equipartition theorem [18], such that $\left\langle \frac{k\xi_1^2}{2} \right\rangle_{eq,\lambda} = \left\langle \frac{\xi_2^2}{2m} \right\rangle_{eq,\lambda} = \frac{1}{2\beta}$.

$$\frac{d}{dt}\langle S\rangle_{eq,\lambda} = \dot{\lambda}^{\intercal}\langle X\rangle_{eq,\lambda} = \frac{\dot{\beta}}{\beta} + \frac{\dot{k}}{2k} \qquad (1.8)$$

Thus,

$$d = \dot{\lambda}^{\intercal}\left\langle X - \langle X\rangle_{eq,\lambda}\right\rangle = \dot{\lambda}^{\intercal}\langle \Delta X\rangle \qquad (1.9)$$

where the vector $\Delta X \equiv X - \langle X\rangle_{eq,\lambda}$.

The average dissipation rate defined above is closely related to the energy dissipation rate in non-equilibrium statistical mechanics. For the stochastic oscillator, the equilibrium distribution is Boltzmann distribution, $\rho_{eq,\lambda}(\xi) = \frac{e^{-\beta H(\xi,\lambda)}}{Z(\lambda)}$, where $Z(\lambda) \equiv \int e^{-\beta H(\xi,\lambda)}d\xi$ is the partition function. Moreover,

$$-\frac{\partial \ln \rho_{eq,\lambda}}{\partial \lambda} = \Delta X \qquad (1.10)$$

The thermodynamical entropy associated to ρ is

$$\mathcal{H}(\rho) = -k_B \int \rho(\xi)\ln\rho(\xi)d\xi \qquad (1.11)$$

The internal energy U of the system is the average value of the total energy $H(\xi,\lambda)$ (it is denoted below as U_{neq} when it is evaluated with non-equilibrium distribution).

Free energy

$$F = U - T\mathcal{H} \tag{1.12}$$

is the portion of internal energy which could be converted to mechanical work or dissipated into surrounding environment. At equilibrium,

$$F_{eq} = -k_B T \ln Z(\lambda) \tag{1.13}$$

For a non-equilibrium distribution $\rho_{neq,\lambda}$ with a time-dependent protocol $\lambda(t)$, its divergence from the equilibrium distribution with the same control parameter is measured by relative entropy $D(\rho_{neq,\lambda} \| \rho_{eq,\lambda}) = \int \rho_{neq,\lambda} \ln\left(\frac{\rho_{neq,\lambda}}{\rho_{eq,\lambda}}\right) d\xi$

$$\begin{aligned} D(\rho_{neq,\lambda} \| \rho_{eq,\lambda}) &= \int \rho_{neq,\lambda} \ln \rho_{neq,\lambda} d\xi - \int \rho_{neq,\lambda} \ln \rho_{eq,\lambda} d\xi \\ &= -\frac{1}{k_B}\mathcal{H}(\rho_{neq,\lambda}) - \int \rho_{neq,\lambda} \ln \frac{e^{-\beta H(\xi,\lambda)}}{Z(\lambda)} d\xi \\ &= -\frac{1}{k_B}\mathcal{H}(\rho_{neq,\lambda}) + \beta U_{neq} - \beta F_{eq} \\ &= \beta(F_{neq} - F_{eq}) \end{aligned} \tag{1.14}$$

As relative entropy is non-negative, in relaxing from non-equilibrium to equilibrium under the Fokker-Planck dynamics, the statistical mechanical system dumps (part of) its free energy to surroundings. In this case of the stochastic oscillator, the energy dissipation refers to the (unitless) loss of free energy. Thus, the time derivative of relative entropy is the energy dissipation rate of the system.

$$
\begin{aligned}
\frac{d}{dt}D(\rho_{neq,\lambda}\|\rho_{eq,\lambda}) &= \int \frac{\partial \rho_{neq,\lambda}}{\partial t}\ln\left(\frac{\rho_{neq,\lambda}}{\rho_{eq,\lambda}}\right)d\xi \\
&+ \int \rho_{neq,\lambda}\left(\frac{1}{\rho_{neq,\lambda}}\frac{\partial \rho_{neq,\lambda}}{\partial t} - \frac{\partial \ln \rho_{eq,\lambda}}{\partial t}\right)d\xi \\
&= \int \frac{\partial \rho_{neq,\lambda}}{\partial t}\ln\left(\frac{\rho_{neq,\lambda}}{\rho_{eq,\lambda}}\right)d\xi + \dot{\lambda}^T \langle \Delta X\rangle
\end{aligned}
\tag{1.15}
$$

On the right hand side of above equality, the first term in the case of stochastic oscillator [19] can be simplified as

$$
-\frac{\zeta}{\beta}\int e^{-\frac{\beta\xi_2^2}{m}}\left[\frac{1}{\rho_{neq}}\left(\frac{\partial}{\partial \xi_2}\left(e^{\frac{\beta\xi_2^2}{2m}}\rho_{neq}\right)\right)^2\right] \leq 0 \tag{1.16}
$$

and the second term is the average dissipation rate d. Therefore, d is an upper bound of the entropy dissipation rate.

1.2 A port-Hamiltonian system in contact with a heat bath

A capacitor with voltage v is connected with a Nyquist-Johnson resistor [9] [10] (as shown in Figure 1.1), based on the charge conservation principle in the circuit,

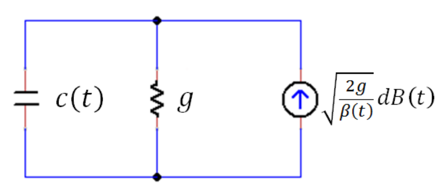

Figure 1.1: RC circuit

$$
d(cv) = -gvdt + \sqrt{\frac{2g}{\beta}}dB(t) \tag{1.17}
$$

As a consequence,

$$
dv = -\left(\frac{\dot{c}+g}{c}\right)vdt + \frac{1}{c}\sqrt{\frac{2g}{\beta}}dB(t) \tag{1.18}
$$

In this heat engine model, g is the electrical conductance of the Nyquist-Johnson resistor, which is a constant. $\beta(t)$ and $B(t)$ carry the same meaning as in the stochastic oscillator model. c is the capacitance of the capacitor and $\lambda = (c, \beta)$ is the system parameter of this heat engine.

The capacitor-resistor system can be generalized as a port-Hamiltonian system [20] in contact with a heat bath,

$$d\xi = (J - G)\frac{\partial H(\xi, \lambda)}{\partial \xi}dt + h(\xi, \lambda)udt + \sqrt{\frac{2}{\beta}}G^{1/2}d\omega(t) \tag{1.19}$$

In (1.19), ξ, $\omega(t)$, and matrices J, G, $G^{1/2}$ have the same meaning as in (1.3). For the resistor-capacitor circuit with $\lambda = (c, \beta)$,

$$\xi = v,\ J = 0,\ G = \frac{g}{c^2},\ H(v, \lambda) = \frac{cv^2}{2},\ h(v, \lambda) = -\frac{v}{c} \tag{1.20}$$

Given the unitless function of the system $\beta H(\xi, \lambda) \in \mathcal{C}^2(\mathbb{R}^d \times [0, \infty))$, where H is the Hamiltonian of the system and usually it is the total energy of the system, by Ito's rule [17]

$$\begin{aligned} d\beta H &= \frac{\partial \beta H}{\partial t}dt + \left(\frac{\partial \beta H}{\partial \xi}\right)^{\intercal} d\xi + \frac{1}{2}\mathrm{tr}\left(\mathrm{Hessian}(\beta H)\left(\frac{2}{\beta}G\right)\right)dt \\ &= \left(\frac{\partial \beta H}{\partial \lambda}\right)^{\intercal} \dot{\lambda}dt + \left(\frac{\partial \beta H}{\partial \xi}\right)^{\intercal}\left((J - G)\frac{\partial H}{\partial \xi}dt + hudt + \sqrt{\frac{2}{\beta}}G^{1/2}d\omega_t\right) \\ &+ \frac{1}{2}\mathrm{tr}\left(\mathrm{Hessian}(\beta H)\left(\frac{2}{\beta}G\right)\right)dt \end{aligned} \tag{1.21}$$

As $\langle \beta H \rangle$ is the average value of βH and $J = -J^{\intercal}$,

$$\begin{aligned} \frac{d}{dt}\langle \beta H \rangle = &\left\langle \left(\frac{\partial \beta H}{\partial \lambda}\right)^{\intercal} \right\rangle \dot{\lambda} - \left\langle \left(\frac{\partial \beta H}{\partial \xi}\right)^{\intercal} G \frac{\partial H}{\partial \xi} \right\rangle + \left\langle \left(\frac{\partial \beta H}{\partial \xi}\right)^{\intercal} h \right\rangle u \\ &+ \frac{1}{2}\mathrm{tr}\left(\mathrm{Hessian}(\beta H)\left(\frac{2}{\beta}G\right)\right) \end{aligned} \tag{1.22}$$

u is referred to as the *linear input* and $y = \left(\frac{\partial \beta H}{\partial \xi}\right)^{\intercal} h$ ($h \in \mathbb{R}^{d \times m}$) is the *linear output*. The term -yu is the power extracted by an external agent who applies the linear input. Part of this extracted power will convert to the internal energy of the system and the other part will be the *extracted mechanical power*. Move the extracted power term to the right-hand side of the above equality,

$$\begin{aligned}\frac{d}{dt}\langle\beta H\rangle-\left\langle\left(\frac{\partial\beta H}{\partial\xi}\right)^{\intercal}h\right\rangle u=\left\langle\left(\frac{\partial\beta H}{\partial\lambda}\right)^{\intercal}\right\rangle\dot{\lambda}-\left\langle\left(\frac{\partial\beta H}{\partial\xi}\right)^{\intercal}G\frac{\partial H}{\partial\xi}\right\rangle\\+\frac{1}{2}\mathrm{tr}\left(\mathrm{Hessian}(\beta H)\left(\frac{2}{\beta}G\right)\right)\end{aligned}\tag{1.23}$$

Assuming that S is the storage function of the system, $\dot{\lambda}$ is the *nonlinear input* and $\frac{\partial S}{\partial\lambda}$ is the *nonlinear output*. $w=\dot{\lambda}^{\intercal}\frac{\partial S}{\partial\lambda}$ is *nonlinear supply rate*. The difference $-\left\langle\left(\frac{\partial\beta H}{\partial\xi}\right)^{\intercal}G\frac{\partial H}{\partial\xi}\right\rangle+\frac{1}{2}\mathrm{tr}\left(\mathrm{Hessian}(\beta H)\left(\frac{2}{\beta}G\right)\right)$ is the average heat transfer between the system and the heat bath.

In the case of the resistor-capacitor circuit, at time $t>0$, we define the storage function of the system $S\equiv\beta H(v,\lambda)-\int_0^t\frac{\partial\beta H(v,\lambda)}{\partial v}h(v,\lambda)u\,ds$, which is a sum of the unitless internal energy and the accumulation of the unitless power extracted by an external agent from time 0 to time t, with the assumption $\int_0^t\left|\left\langle\frac{\partial\beta H(v,\lambda)}{\partial v}h(v,\lambda)u\right\rangle\right|<\infty$. The *nonlinear supply rate* is $w=\dot{\lambda}^{\intercal}\left(\frac{\partial\beta H(v,\lambda)}{\partial\lambda}\right)$. The *average* dissipation rate $d\equiv\langle w\rangle-\frac{d}{dt}\langle S\rangle_{eq,\lambda}$. Given that $S(t,v)\in\mathcal{C}^2([0,\infty)\times\mathbb{R})$, by Ito's formula [17],

$$dS=\frac{\partial S}{\partial t}dt+\frac{\partial S}{\partial v}dv+\frac{1}{2}\frac{\partial^2S}{\partial v^2}\frac{2g}{\beta c^2}dt\tag{1.24}$$

$$=\dot{\lambda}^{\intercal}\left(\frac{\partial\beta H}{\partial\lambda}\right)dt-\beta gv^2dt+\sqrt{2g\beta}v\,dB(t)+\frac{g}{c}dt\tag{1.25}$$

By equipartition theorem [18], $\left\langle\frac{cv^2}{2}\right\rangle_{eq,\lambda}=\frac{1}{2\beta}$,

$$\frac{d}{dt}\langle S\rangle_{eq,\lambda}=\dot{\lambda}^{\intercal}\left\langle\frac{\partial\beta H}{\partial\lambda}\right\rangle_{eq,\lambda}\tag{1.26}$$

Thus, given $X\equiv\frac{\partial S}{\partial\lambda}=\frac{\partial\beta H}{\partial\lambda}$,

$$d=\dot{\lambda}^{\intercal}\left\langle X-\langle X\rangle_{eq,\lambda}\right\rangle=\dot{\lambda}^{\intercal}\langle\Delta X\rangle\tag{1.27}$$

where $\Delta X \equiv X - \langle X \rangle_{eq,\lambda}$.

For the non-equilibrium distribution $\rho_{neq,\lambda}$ of the resistor-capacitor system associated with a time-dependent protocol $\lambda(t)$, as discussed in the case of the stochastic oscillator, the relative entropy $D(\rho_{neq,\lambda} \| \rho_{eq,\lambda})$ describes the dissipation of the system into its surrounding environment, through the loss of free energy. In the case of the resistor-capacitor circuit, as $\rho_{eq,\lambda}(v) \propto e^{-\frac{\beta c v^2}{2}}$,

$$\begin{aligned}
\frac{d}{dt} D\left(\rho_{neq,\lambda} \| \rho_{eq,\lambda}\right) &= \int \frac{\partial \rho_{neq,\lambda}}{\partial t} \ln\left(\frac{\rho_{neq,\lambda}}{\rho_{eq,\lambda}}\right) dv \\
&+ \int \rho_{eq,\lambda} \left(\frac{1}{\rho_{neq,\lambda}} \frac{\partial \rho_{neq,\lambda}}{\partial t} - \frac{\partial \ln \rho_{eq,\lambda}}{\partial t}\right) dv \\
&= \int \frac{\partial \rho_{neq,\lambda}}{\partial t} \ln\left(\frac{\rho_{neq,\lambda}}{\rho_{eq,\lambda}}\right) dv + \dot{\lambda}^{\intercal} \langle \Delta X \rangle \\
&= \int \frac{\partial}{\partial v}\left(\frac{\dot{c}}{c} v \rho_{neq,\lambda}\right) \ln\left(\frac{\rho_{neq,\lambda}}{\rho_{eq,\lambda}}\right) dv \\
&+ \int \left(\frac{\partial}{\partial v}\left(\frac{g}{c} v \rho_{neq,\lambda}\right) + \frac{1}{2}\frac{\partial^2}{\partial v^2}\left(\frac{2g}{\beta c^2}\rho_{neq,\lambda}\right)\right) \ln\left(\frac{\rho_{neq,\lambda}}{\rho_{eq,\lambda}}\right) dv \\
&+ \dot{\lambda}^{\intercal} \langle \Delta X \rangle
\end{aligned} \tag{1.28}$$

The first term on the right of the last equality

$$\begin{aligned}
\int \frac{\partial}{\partial v}\left(\frac{\dot{c}}{c} v \rho_{neq,\lambda}\right) \ln\left(\frac{\rho_{neq,\lambda}}{\rho_{eq,\lambda}}\right) dv &= \frac{\dot{c}}{c} - \frac{\dot{c}}{c}\left\langle -\beta c v^2 \right\rangle \\
&= \left\langle h(v,\lambda) \frac{\partial \beta E}{\partial v} u \right\rangle - \left\langle h(v,\lambda) \frac{\partial \beta E}{\partial v} u \right\rangle_{eq,\lambda}
\end{aligned} \tag{1.29}$$

Based on the calculation in [19] (Chapter 2, page 7), the second term on the right of the last equality

$$\begin{aligned}
&\int \left(\frac{\partial}{\partial v}\left(\frac{g}{c} v \rho_{neq,\lambda}\right) + \frac{1}{2}\frac{\partial^2}{\partial v^2}\left(\frac{2g}{\beta c^2}\rho_{neq,\lambda}\right)\right) \ln\left(\frac{\rho_{neq,\lambda}}{\rho_{eq,\lambda}}\right) dv = \\
&-\frac{c^2}{\beta g} \int \left(\beta c v \rho_{neq,\lambda} + \frac{\partial_{neq,\lambda}}{\partial v}\right)^2 \frac{1}{\rho_{neq}} dv \leq 0
\end{aligned} \tag{1.30}$$

Thus, (1.28) shows

$$\dot{\lambda}^{\intercal} \langle \Delta X \rangle \geq \frac{d}{dt} D(\rho_{neq,\lambda} \| \rho_{eq,\lambda}) + \left(- \left\langle h(v,\lambda) \frac{\partial \beta H}{\partial v} u \right\rangle + \left\langle h(v,\lambda) \frac{\partial \beta H}{\partial v} u \right\rangle_{eq,\lambda} \right) \tag{1.31}$$

By viewing $-\left\langle h(v,\lambda)\frac{\partial \beta H}{\partial v} u\right\rangle$ as the power extracted by an external agent, the difference between its non-equilibrium average and the quasi-static one is the extra extracted power due to the finite-time protocols. The average dissipation rate is an upper-bound approximation of the entropy production of a finite-time protocol which comprises of two sources: one is the dissipation due to the free energy difference at the end of a protocol and the other is the extra power extracted by the external agent along the protocol.

1.3 Linear response theory and inverse diffusion tensor

In both heat engine models, the average dissipation rate is $d = \dot{\lambda} \langle \Delta X \rangle$. To analyze the average dissipation rate d, it is crucial to compute an expression for $\langle \Delta X \rangle$. Assuming that λ is varied smoothly and the statistical mechanical system is always near its equilibrium, linear response theory is the standard framework for understanding $\langle \Delta X \rangle$ [12]. The analysis in this section is mainly stated in [19] and [11] and it is presented here with greater detail. The key idea of this analysis is to use a discrete-time protocol to approximate the continuously time-varying one.

Assuming our statistical mechanical system is at equilibrium at time $t = 0$, with control parameter λ_0, it is operated with a protocol $\lambda(t)$, where $\lambda(0) = \lambda_0$. At time 0, the system is perturbed by $\dot{\lambda}(0)$. This process is approximated as follow where linear response theory is applicable.

The statistical mechanical system reaches equilibrium from time $-\infty$ to time $t = 0$ with control parameter $\lambda_0 + (\lambda_0 - \lambda(t_1))$ (later we denote $\lambda_0 - \lambda(t_1)$ as $\Delta\lambda$). At time $t = 0$, $\Delta\lambda$ vanishes and at time t_1 ($t_1 > 0$), the system is under investigation with the control parameter λ_0 as the reference value. To have this approximation process meaningful for our real process, let $\lim_{t_1 \to 0} \frac{\lambda_0 - (\lambda_0 + (\lambda_0 - \lambda(t_1)))}{t_1} = \dot{\lambda}(0)$.

In the approximation process, from time -∞ to time $t = 0$, the system arrives at an equilibrium distribution with control parameter $\lambda_0 + \Delta\lambda$. So, the equilibrium probability distribution at time $t = 0$ is $\rho_{eq,\lambda_0+\Delta\lambda}(\xi) = \frac{e^{-\beta H(\xi,\lambda_0+\Delta\lambda)}}{Z(\lambda_0+\Delta\lambda)}$ and the equilibrium distribution with control parameter λ_0 is $\rho_{eq,\lambda_0}(\xi) = \frac{e^{-\beta H(\xi,\lambda_0)}}{Z(\lambda_0)}$. By the linearization of βH,

$$\begin{aligned} \frac{\rho_{eq,\lambda_0+\Delta\lambda}(\xi)}{\rho_{eq,\lambda_0}(\xi)} &= \frac{e^{-\beta H(\lambda_0+\Delta\lambda,\xi)}}{Z(\lambda_0+\Delta\lambda)} \frac{Z(\lambda_0)}{e^{-\beta H(\lambda_0,\xi)}} \\ &\approx (1 - X^{\intercal}\Delta\lambda + O(\Delta\lambda^2))(1 + \langle X^{\intercal} \rangle_{eq,\lambda} \Delta\lambda + O(\Delta\lambda^2)) \end{aligned} \tag{1.32}$$

Following static linear response theory,

$$\rho_{eq,\lambda_0+\Delta\lambda} \approx \rho_{eq,\lambda_0} - \rho_{eq,\lambda_0}\Delta X^{\intercal}\Delta\lambda. \tag{1.33}$$

At time t_1, $\langle \Delta X \rangle = \int \rho(\xi, t_1)\Delta X d\xi = \int \rho_{eq,\lambda_0+\Delta\lambda}(\xi_0) \int \Delta X \rho(\xi, t_1|\xi_0, \lambda_0 + \Delta\lambda) d\xi d\xi_0$. Denote $\Delta X(t_1) = \int \Delta X \rho(\xi, t_1|\xi_0, \lambda_0 + \Delta\lambda) d\xi$ and use the result in (1.33),

$$\langle \Delta X \rangle = \int \rho_{eq,\lambda_0+\Delta\lambda}(\xi_0)\Delta X(t_1) d\xi_0 \approx -\langle \Delta X(t_1)\Delta X^{\intercal} \rangle_{eq,\lambda_0} \cdot \Delta\lambda \tag{1.34}$$

On the other hand, based on dynamic linear (step) response theory, at $t_1 > 0$, $\langle \Delta X \rangle \approx \int_{-\infty}^{t_1} \chi(t_1 - t')(\lambda(t') - \lambda(t_1))dt'$. From time $-\infty$ to time 0, $\lambda(t') - \lambda(t_1) = \Delta\lambda$. Do change of variable $s = t_1 - t'$ and $ds = -dt'$.

$$\langle \Delta X \rangle \approx \int_{t_1}^{\infty} \chi(s) ds \cdot \Delta\lambda \tag{1.35}$$

where χ is the linear response kernel. Compare the results from both static linear response theory (1.34) and dynamic linear response theory (1.35), $\int_{t_1}^{\infty} \chi(s) ds = -\langle \Delta X(t_1) \Delta X^{\intercal} \rangle_{eq,\lambda_0}$. Assuming that $\langle \Delta X(s) \Delta X^{\intercal} \rangle$ is differentiable against s and $\lim_{s \to \infty} \langle \Delta X(s) \Delta X^{\intercal} \rangle = 0$,

$$\chi(s) = \frac{d}{ds} \langle \Delta X(s) \Delta X^{\intercal} \rangle_{eq,\lambda_0} \tag{1.36}$$

Look into the integral $\langle \Delta X(s) \Delta X^{\intercal} \rangle_{eq,\lambda_0}$ carefully.

$$\langle \Delta X(s) \Delta X^{\intercal} \rangle_{eq,\lambda_0} = \int \left(\int \Delta X(\xi) \rho(\xi, s | \xi_0, 0) d\xi \right) \rho_{eq,\lambda_0}(\xi_0) \Delta X^{\intercal}(\xi_0) d\xi_0 \tag{1.37}$$

As $\Delta\lambda \to 0$,

$$\langle \Delta X(s) \Delta X^{\intercal} \rangle_{eq,\lambda_0} = \int \int \Delta X(\xi) \Delta X^{\intercal}(\xi_0) \rho(\xi, s; \xi_0, 0) d\xi d\xi_0 \tag{1.38}$$

where $\rho(\xi, s; \xi_0, 0)$ is the joint probability density of ξ and ξ_0.

Let us apply dynamic linear response theory to the real process with the approximated linear response kernel (1.36). Assuming that $\lim_{s \to \infty} \langle \Delta X(s) \Delta X^{T} \rangle = 0$, use change of variable $s = t_1 - t'$ and integrate-by-parts in the integral of dynamic linear response.

$$
\begin{aligned}
\langle \Delta X \rangle &\approx \int_{-\infty}^{t_1} \chi(t_1 - t') \cdot (\lambda(t') - \lambda(t_1)) dt^{'} \\
&= \int_0^{\infty} \chi(s) \cdot (\lambda(t_1 - s) - \lambda(t_1)) ds \\
&= \int_0^{\infty} \frac{d}{ds} \left\langle \Delta X(s) \Delta X^T \right\rangle_{eq,\lambda_0} \cdot (\lambda(t_1 - s) - \lambda(t_1)) \\
&= \int_0^{\infty} \left\langle \Delta X(s) \Delta X^T \right\rangle_{eq,\lambda_0} \cdot \dot{\lambda}(t_1 - s) ds
\end{aligned} \tag{1.39}
$$

In our heat engines, there are two time scales: one is fast time scale of fluctuations of the statistical mechanical systems and the other is slow time scale of the system parameter. The system parameter varies slowly to have the system operating near its corresponding equilibrium.

Theorem 1.3.1. *For the heat engines, we have following assumption:*

1. *At time s, every element in the covariance matrix $\langle \Delta X(s) \Delta X^{\intercal} \rangle$ is upper-bounded with an exponential decaying term $Be^{-s/\tau}$, where B and τ are positive constants.*

2. *The protocol is a function which is slowly varying with time as $\lambda(\epsilon t)$, where $\epsilon \ll 1$. Thus, $\dot{\lambda} = \epsilon \lambda'|_{\epsilon t}$ and $\ddot{\lambda} = \epsilon^2 \lambda''|_{\epsilon t}$. We assume that there exists $r > 0$ and $K_1 > 0$, such that for $\forall s \in [0, r\tau]$, $\|\dot{\lambda}(\epsilon(t_1 - s)) - \dot{\lambda}(\epsilon t_1)\|_{\infty} < \epsilon^2 K_1$ and $e^{-r} < \epsilon$. Also, there exists $K_2 > 0$, such that $\forall s \in [0, \infty)$, $\|\dot{\lambda}(\epsilon(t_1 - s)) - \dot{\lambda}(\epsilon t_1)\|_{\infty} < \epsilon K_2$, where $\| \, \|_{\infty}$ stands for supremum norm, applied to vector λ.*

Then, up to order $O(\epsilon^2)$, we have an inverse diffusion tensor to approximate the average dissipation rate d.

Proof.

$$\begin{aligned}
&\langle \Delta X \rangle - \int_0^\infty \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} ds \cdot \dot\lambda(t_1) = \\
&\int_0^\infty \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} \cdot \dot\lambda(t_1 - s) ds - \int_0^\infty \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} ds \cdot \dot\lambda(t_1) = \\
&\int_0^\infty \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} \cdot \left(\dot\lambda(t_1 - s) - \dot\lambda(t_1) \right) ds
\end{aligned} \tag{1.40}$$

Thus,

$$\begin{aligned}
&\| \langle \Delta X \rangle - \int_0^\infty \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} ds \cdot \dot\lambda(t_1) \|_\infty \leq \\
&\int_0^\infty \| \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} \cdot \left(\dot\lambda(t_1 - s) - \dot\lambda(t_1) \right) \|_\infty ds \\
&= \int_0^{r\tau} \| \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} \cdot \left(\dot\lambda(t_1 - s) - \dot\lambda(t_1) \right) \|_\infty ds \\
&+ \int_{r\tau}^\infty \| \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} \cdot \left(\dot\lambda(t_1 - s) - \dot\lambda(t_1) \right) \|_\infty ds
\end{aligned} \tag{1.41}$$

By calculation, if n is the dimension of $\dot\lambda$,

$$\int_0^\infty \| \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} \cdot \left(\dot\lambda(t_1 - s) - \dot\lambda(t_1) \right) \|_\infty ds \leq \frac{nB\epsilon^2(K_1 + K_2)}{\tau} + O(\epsilon^3) \tag{1.42}$$

Thus, up to order $O(\epsilon^2)$, at time t_1, $\langle \Delta X \rangle$ is approximated by $\int_0^\infty \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} ds \cdot \dot\lambda(t_1)$. Moreover, because of smoothness of $\lambda(t)$, $\lim_{t_1 \to 0} \dot\lambda(t_1) = \dot\lambda(0)$

$$\langle \Delta X \rangle \approx \int_0^\infty \langle \Delta X(s) \Delta X^\intercal \rangle_{eq,\lambda_0} ds \cdot \dot\lambda(0) \tag{1.43}$$

$$d = \dot\lambda^\intercal(0) \langle \Delta X \rangle \approx \left[\frac{d\lambda^\intercal}{dt} \right]_{t=0} g[\lambda_0] \left[\frac{d\lambda}{dt} \right]_{t=0} \tag{1.44}$$

with the inverse diffusion tensor

$$g_{ij} \equiv \int_0^\infty \langle \Delta X_j(s) \Delta X_i \rangle_{eq,\lambda_0} ds \tag{1.45}$$

$\square$

Let us return to the case of the stochastic oscillator. Based on the calculation in [13], in the space of control parameter $\lambda = (\beta, k)$, its inverse diffusion tensor has constant negative sectional curvature $-\zeta/m$. The ranges of k and β are both positive. By a theorem from Riemannian Geometry [21], there is an isometric mapping between this constant negative-curvature submanifold and a submanifold of Poincaré upper half plane.

$$
\begin{aligned}
x &\equiv \frac{1}{4\beta k}, \quad y \equiv \frac{1}{2\beta\zeta}\sqrt{\frac{m}{k}} \\
d &= \frac{m}{\zeta}\frac{\dot{x}^2 + \dot{y}^2}{y^2}
\end{aligned}
\tag{1.46}
$$

The dissipation rate in the unit $(k_B T/\text{sec})$ in space (x, y) is

$$
\tilde{d} = \frac{d}{\beta} = \zeta\left(\frac{\dot{x}^2 + \dot{y}^2}{x}\right) = \begin{pmatrix} \dot{x} & \dot{y} \end{pmatrix} \begin{pmatrix} \frac{\zeta}{x} & 0 \\ 0 & \frac{\zeta}{x} \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}
\tag{1.47}
$$

This positive-definite tensor $\begin{pmatrix} \frac{\zeta}{x} & 0 \\ 0 & \frac{\zeta}{x} \end{pmatrix}$ can be viewed as a metric tensor in the space of (x, y). The geometric energy of a curve in (x, y) space has the statistical mechanical meaning of the energy dissipation following a protocol.

In the case of the resistor-capacitor circuit, $\dot{\lambda} = (\dot{c}, \dot{\beta})^\intercal$, the associated conjugate force is $X = \frac{\partial}{\partial\lambda}\frac{\beta c v^2}{2} = (\frac{\beta v^2}{2}, \frac{c v^2}{2})^\intercal$. As a slow-perturbed process, the circuit has an approximate dynamics

$$
dv = -\frac{g}{c} v dt + \frac{1}{c}\sqrt{\frac{2g}{\beta}} dB(t)
\tag{1.48}
$$

The associated solution v is in a form of $v_h + v_p$, where v_h is the homogeneous part, which depends only on the initial condition $v(0)$ and is independent of the Brownian motion. The non-homogeneous part v_p has vanishing initial conditions but depends linearly on the stochastic driven force. For the stochastic driven force $dB(t)$, it is easy to show that the non-homogeneous part v_p does not contribute to the equilibrium time correlation function $\langle \delta X_j(0)\delta X_i(t')\rangle$ [19]. As $v_h(t') = e^{\frac{g}{c}t'}v(0)$

$$\begin{aligned}\left\langle \delta v^2(0)\delta v^2(t')\right\rangle_{eq,\lambda} &\equiv \left\langle v^2(t')v^2(0)\right\rangle_{eq,\lambda} - \left\langle v^2(t')\right\rangle_{eq,\lambda}\left\langle v^2(0)\right\rangle_{eq,\lambda} \\ &= e^{-\frac{2g}{c}t'}\left(\left\langle v^4(0)\right\rangle_{eq,\lambda} - \left\langle v^2(0)\right\rangle_{eq,\lambda}^2\right)\end{aligned} \tag{1.49}$$

Because of equipartition theorem [18], $\langle v^2(0)\rangle_{eq,\lambda} = \frac{1}{\beta c}$ and $\langle v^4(0)\rangle_{eq,\lambda} = \frac{3}{\beta^2 c^2}$.

$$\begin{aligned}\left\langle \delta v^2(0)\delta v^2(t')\right\rangle_{eq,\lambda} &= 2e^{-\frac{2g}{c}t'}\frac{1}{\beta^2c^2} \\ \int_0^\infty 2e^{-\frac{2g}{c}t'}\frac{1}{\beta^2c^2}dt' &= \frac{1}{\beta^2 cg}\end{aligned} \tag{1.50}$$

Thus, the inverse diffusion tensor of the resistor-capacitor circuit is

$$d = \begin{pmatrix}\dot{c} & \dot{\beta}\end{pmatrix}\begin{pmatrix}\frac{1}{4cg} & \frac{1}{4\beta g} \\ \frac{1}{4\beta g} & \frac{c}{4\beta^2 g}\end{pmatrix}\begin{pmatrix}\dot{c} \\ \dot{\beta}\end{pmatrix} \tag{1.51}$$

The associated tensor is only positive-semidefinite. It will not describe the entropy production of the finite-time protocols properly and other dissipation source needs to be considered for this system.

Remark 1.3.2. *In our example of the stochastic oscillator, note that the average dissipation rate d has a positive-definite inverse diffusion tensor. Indeed, it is a dissipative dynamic system in the spirit of [16].*

1.4 Inverse diffusion tensor of the resistor-capacitor system

Based on the discussion about the inverse diffusion tensor of the resistor-capacitor system, it is seen that besides the dissipation from the free energy difference at the end of a protocol and the extra power done on an external agent, another dissipation source along the protocol needs to be considered. In non-equilibrium statistical mechanics, there are two methodologies to investigate this dissipation source: one is *stochastic thermodynamics* [3] and the other is *fluctuation theorem* [1] [2]. We will apply these two methodologies to the resistor-capacitor system in parallel. It will be shown that these two methodologies are consistent with each other in this example.

1.4.1 The stochastic thermodynamics

Given the resistor-capacitor system with

$$dv = -\left(\frac{\dot{c}+g}{c}\right)vdt + \frac{1}{c}\sqrt{\frac{2g}{\beta}}dB(t) \tag{1.52}$$

By change of variable $\tilde{v} = c\sqrt{\beta}v$,

$$d\tilde{v} = -\left(-\frac{\dot{\beta}}{2\beta} + \frac{g}{c}\right)\tilde{v}dt + \sqrt{2g}dB(t) \tag{1.53}$$

We have a new system dynamics with a constant diffusion coefficient. Based on the stochastic thermodynamics theory, we use the notation

$$d\tilde{v} = -gFdt + \sqrt{2g}dB(t) \tag{1.54}$$

where the *systematic force* $F = \left(-\frac{\dot{\beta}}{2\beta g} + \frac{1}{c}\right)\tilde{v}$. As $\frac{\beta c v^2}{2} = \frac{\tilde{v}^2}{2c}$ is the unitless (stored) potential energy of the capacitor, $\frac{\partial}{\partial \tilde{v}}\frac{\tilde{v}^2}{2c} = \frac{\tilde{v}}{c}$ is the *conservative force* and $f = -\frac{\dot{\beta}}{2\beta g}\tilde{v}$ is the *non-conservative force*.

We define the *heat released* by the system from time 0 to time t_f is

$$\delta Q \equiv -\int_0^{t_f} \dot{\tilde{v}} \circ F dt = -\int_0^{t_f} \dot{\tilde{v}} \circ \left(-\frac{\dot{\beta}}{2\beta g}\tilde{v} + \frac{\tilde{v}}{c}\right) dt \tag{1.55}$$

where $\circ$ stands for Stratonovich integral. The work done on the system is defined as

$$\delta W \equiv \int_0^{t_f} \left(\frac{\partial}{\partial t}\frac{\tilde{v}^2}{2c} - f \circ \dot{\tilde{v}}\right) dt \tag{1.56}$$

Based on these two definition, the *first law of thermodynamics* is satisfied with the unitless potential energy of the capacitor $\frac{\tilde{v}^2}{2c}$ by

$$\delta W - \delta Q = \int_0^{t_f} d\left(\frac{\tilde{v}^2}{c}\right) = \frac{\tilde{v}^2}{2c}(t_f) - \frac{\tilde{v}^2}{2c}(0) \tag{1.57}$$

Based on the equation (1.53), we introduce several vector fields:

$$\begin{aligned} b &= -\left(-\frac{\dot{\beta}}{2\beta} + \frac{g}{c}\right)\tilde{v} \\ \bar{b} &= -\left(-\frac{\dot{\beta}}{2\beta} + \frac{g}{c}\right)\tilde{v} - 2g\frac{\partial}{\partial \tilde{v}}\ln\rho(\tilde{v},t) \\ w_1 &= \frac{b+\bar{b}}{2} = -\left(-\frac{\dot{\beta}}{2\beta} + \frac{g}{c}\right)\tilde{v} - g\frac{\partial}{\partial \tilde{v}}\ln\rho(\tilde{v},t) \\ w_2 &= \frac{b-\bar{b}}{2} = g\frac{\partial}{\partial \tilde{v}}\ln\rho(\tilde{v},t) \end{aligned} \tag{1.58}$$

where $\rho(\tilde{v},t)$ is the probability density of $\tilde{v}$ at time t. (With an abuse of the notation, let $\tilde{v}$ also denote the value of the random variable $\tilde{v}$.) The expectation of the released heat is

$$Q \equiv \langle \delta Q \rangle = \left\langle \int_0^{t_f} -\dot{\tilde{v}} \circ \left(-\frac{\dot{\beta}}{2\beta g} + \frac{1}{c}\right)\tilde{v} dt \right\rangle = \frac{1}{g}\left\langle \int_0^{t_f} \dot{\tilde{v}} \circ b dt \right\rangle \tag{1.59}$$

By switching from Stratonovich integral to Ito's integral,

$$
\begin{aligned}
Q &= \frac{1}{g}\left\langle \int_0^{t_f} bd\tilde{v} + g\frac{\partial}{\partial\tilde{v}}bdt \right\rangle \\
&= \frac{1}{g}\left\langle \int_0^{t_f} \left(b^2 + g\frac{\partial}{\partial\tilde{v}}b\right) dt \right\rangle
\end{aligned}
\tag{1.60}
$$

As $b = w_1 + w_2$, $\left\langle \int_0^{t_f} b^2 dt \right\rangle = \left\langle \int_0^{t_f} (w_1 + w_2)^2 dt \right\rangle$. Also,

$$
\begin{aligned}
\left\langle \int_0^{t_f} g\frac{\partial}{\partial\tilde{v}}bdt \right\rangle &= \int_0^{t_f}\int_{-\infty}^{+\infty} g\left(\frac{\partial}{\partial\tilde{v}}b\right)\rho(\tilde{v},t)d\tilde{v}dt \\
&= \int_0^{t_f}\int_{-\infty}^{+\infty} g\rho(\tilde{v},t)db(\tilde{v})dt \\
&= \int_0^{t_f}\left(g\rho(\tilde{v},t)b(\tilde{v})|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} gb(\tilde{v})\left(\frac{\partial}{\partial\tilde{v}}\rho(\tilde{v},t)\right)\rho(\tilde{v},t)\right)dt
\end{aligned}
\tag{1.61}
$$

Assuming that when $|\tilde{v}| \to 0$, $\rho(\tilde{v},t)$ converges to zero fast enough, such that $g\rho(\tilde{v},t)b(\tilde{v})|_{-\infty}^{+\infty} = 0$ and

$$
\left\langle \int_0^{t_f} g\frac{\partial}{\partial\tilde{v}}bdt \right\rangle = -\left\langle \int_0^{t_f} gb\frac{\partial}{\partial\tilde{v}}\rho(\tilde{v},t)dt \right\rangle = -\left\langle \int_0^{t_f} (w_1+w_2)w_2dt \right\rangle
\tag{1.62}
$$

Thus, by combining (1.60) and (1.61), the average released heat

$$
\begin{aligned}
Q &= \frac{1}{g}\left\langle \int_0^{t_f} \left((w_1+w_2)^2 - (w_1+w_2)\,w_2\right) dt \right\rangle \\
&= \frac{1}{g}\left\langle \int_0^{t_f} (w_1+w_2)\,w_1dt \right\rangle = \frac{1}{g}\left\langle \int_0^{t_f} \left(w_1^2 + w_1w_2\right) dt \right\rangle
\end{aligned}
\tag{1.63}
$$

Next, we would like to discuss the physical meaning of the term $\frac{1}{g}\left\langle \int_0^{t_f} w_1w_2dt \right\rangle$. Given the stochastic process,

$$
d\tilde{v} = bdt + \sqrt{2g}dB(t)
\tag{1.64}
$$

its Fokker-Planck equation is

$$\begin{aligned}
\frac{\partial}{\partial t}\rho(\tilde{v},t) &= -\frac{\partial}{\partial \tilde{v}}\left(\rho(\tilde{v},t)b\right) + g\frac{\partial^2}{\partial \tilde{v}^2}\rho(\tilde{v},t) \\
&= -\frac{\partial}{\partial \tilde{v}}\left(\rho(\tilde{v},t)b - g\frac{\partial}{\partial \tilde{v}}\rho(\tilde{v},t)\right) \\
&= -\frac{\partial}{\partial \tilde{v}}\left(\rho(\tilde{v},t)\left(b - g\frac{\partial}{\partial \tilde{v}}\ln\rho(\tilde{v},t)\right)\right) = -\frac{\partial}{\partial \tilde{v}}\left(\rho(\tilde{v},t)w_1\right)
\end{aligned} \tag{1.65}$$

On the other hand, as the information-theoretical entropy (with an abuse of the notation $\mathcal{H}$) $\mathcal{H}(\rho(\tilde{v},t)) = -\int_{-\infty}^{+\infty}\rho(\tilde{v},t)\ln\rho(\tilde{v},t)d\tilde{v}$ is a function of time t,

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{H} &= -\frac{\partial}{\partial t}\int_{-\infty}^{+\infty}\rho(\tilde{v},t)\ln\rho(\tilde{v},t)d\tilde{v} \\
&= -\int_{-\infty}^{+\infty}\left(\left(\frac{\partial}{\partial t}\rho(\tilde{v},t)\right)\ln\rho(\tilde{v},t) + \frac{\partial}{\partial t}\rho(\tilde{v},t)\right)d\tilde{v}
\end{aligned} \tag{1.66}$$

Due to the conservation of the integral of probability density, $\int_{-\infty}^{+\infty}\frac{\partial}{\partial t}\rho(\tilde{v},t)d\tilde{v} = 0$. Based on the Fokker-Planck equation,

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{H} &= -\int_{-\infty}^{+\infty}\left(\frac{\partial}{\partial t}\rho(\tilde{v},t)\right)\ln\rho(\tilde{v},t)d\tilde{v} \\
&= \int_{-\infty}^{+\infty}\frac{\partial}{\partial \tilde{v}}\left(\rho(\tilde{v},t)w_1\right)\ln\rho(\tilde{v},t)d\tilde{v} \\
&= \rho(\tilde{v},t)w_1\ln\rho(\tilde{v},t)|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty}\rho(\tilde{v},t)w_1\frac{\partial}{\partial \tilde{v}}\ln\rho(\tilde{v},t)d\tilde{v}
\end{aligned} \tag{1.67}$$

Assuming that when $|\tilde{v}| \to \infty$, $\rho(\tilde{v},t)w_1\ln\rho(\tilde{v},t)|_{-\infty}^{+\infty} = 0$,

$$\frac{\partial}{\partial t}\mathcal{H} = -\int_{-\infty}^{+\infty}\frac{1}{g}\rho(\tilde{v},t)w_1w_2d\tilde{v} = -\frac{1}{g}\left\langle(w_1w_2)\right\rangle \tag{1.68}$$

$$\begin{aligned}
Q &= \frac{1}{g}\int_0^{t_f}\left\langle w_1^2\right\rangle dt + \frac{1}{g}\int_0^{t_f}\left\langle(w_1w_2)\right\rangle dt \\
&= \frac{1}{g}\int_0^{t_f}\left\langle w_1^2\right\rangle dt - \mathcal{H}(t_f) + \mathcal{H}(0)
\end{aligned} \tag{1.69}$$

Therefore,

$$Q + \mathcal{H}(t_f) - \mathcal{H}(0) = \frac{1}{g}\int_0^{t_f} \langle w_1^2 \rangle \, dt \tag{1.70}$$

The total entropy produced from time 0 to time t_f is $\int_0^{t_f} \frac{\langle w_1^2 \rangle}{g} dt$ and the entropy production rate is $\frac{\langle w_1^2 \rangle}{g}$. Taking expectation on both sides of (1.57),

$$\langle \delta W \rangle - Q = \left\langle \frac{\tilde{v}^2}{2c}(t_f) \right\rangle - \left\langle \frac{\tilde{v}^2}{2c}(0) \right\rangle \tag{1.71}$$

Replacing Q with (1.70),

$$\begin{aligned} \langle \delta W \rangle &= \left\langle \frac{\tilde{v}^2}{2c}(t_f) \right\rangle - \left\langle \frac{\tilde{v}^2}{2c}(0) \right\rangle - \mathcal{H}(t_f) + \mathcal{H}(0) + \frac{1}{g}\int_0^{t_f} \langle w_1^2 \rangle \, dt \\ &= \left(\left\langle \left(\frac{\tilde{v}^2}{2c} \right)(t_f) \right\rangle - \mathcal{H}(t_f) \right) - \left(\left\langle \left(\frac{\tilde{v}^2}{2c} \right)(0) \right\rangle - \mathcal{H}(0) \right) + \frac{1}{g}\int_0^{t_f} \langle w_1^2 \rangle \, dt \end{aligned} \tag{1.72}$$

At time t, the unitless free energy $F(t) = \left\langle \left(\frac{\tilde{v}^2}{2c} \right)(t) \right\rangle - \mathcal{H}(t)$. We have the unitless version of the *second law of thermodynamics*,

$$\langle \delta W \rangle \geq F(t_f) - F(0) \tag{1.73}$$

To estimate the entropy production rate in the near-equilibrium scenario, where we can estimate the non-equilibrium distribution with the equilibrium distribution, such that $p(\tilde{v}, t) \propto \exp\left(-\frac{\tilde{v}^2}{2\sigma_{eq}^2} \right)$ and $\sigma_{eq}^2 = \langle \tilde{v}^2 \rangle_{eq,\lambda} = c$,

$$\langle w_1^2 \rangle_{eq,\lambda} = \left\langle \left(-\left(-\frac{\dot{\beta}}{2\beta} + \frac{g}{c} \right) \tilde{v} - g\frac{\partial}{\partial \tilde{v}} \ln p(\tilde{v}, t) \right)^2 \right\rangle_{eq,\lambda} = \frac{c\dot{\beta}^2}{4\beta^2} \tag{1.74}$$

The entropy production rate near the equilibrium is $\frac{c\dot{\beta}^2}{4g\beta^2}$.

1.4.2 The fluctuation theorem

In parallel with the stochastic thermodynamics analysis, we can analyze the non-equilibrium behavior of the resistor-capacitor system with the fluctuation theorem [1] [2]. In fluctuation theorem analysis, we would like to discuss the irreversibility of the statistical mechanical process by comparing the trajectories of the forward process with their time-reversed trajectories of a backward process.

Given the forward process from time 0 to time t_f, $dv = -\left(\frac{\dot{c}+g}{c}\right) v dt + \frac{1}{c}\sqrt{\frac{2g}{\beta}} dB(t)$, we have the backward process: $dv^* = -\check{b}(v^*)dt + \frac{1}{c^*}\sqrt{\frac{2g}{\beta^*}} dB(t)$, where $c^*(t) = c(t_f - t)$, $\beta^*(t) = \beta(t_f - t)$ and

$$\begin{aligned}\check{b}(v^*, t) &\equiv -\left(\frac{g - \dot{c}^*}{c^*}\right) v^*(t) - \frac{2g}{c^{*2}\beta^*}(t)\frac{\partial}{\partial v^*} \ln \rho(v^*, t) \\ &= -\left(\frac{g + \dot{c}}{c}\right)(t_f - t)v^*(t) - \frac{2g}{\beta c^2}(t_f - t)\frac{\partial}{\partial v^*} \ln \rho(v^*, t)\end{aligned} \tag{1.75}$$

The Fokker-Planck equation associated to the forward process is

$$\frac{\partial}{\partial t}\rho(v, t) = -\frac{\partial}{\partial v}\left(-\left(\frac{\dot{c} + g}{c}\right) v(t)\rho(v, t)\right) + \frac{g}{c^2\beta}(t)\frac{\partial^2}{\partial v^2}\rho(v, t) \tag{1.76}$$

From time 0 to time t_f, with the initial distribution $\rho(v, 0)$, we have he solution of the Fokker-Planck equation $\rho(v, t)$. By replacing t with $t_f - t$ in $\rho(v, t)$,

$$\begin{aligned}\frac{\partial}{\partial t}\rho(v, t_f - t) &= -\frac{\partial}{\partial(t_f - t)}\rho(v, t_f - t) \\ &= -\left(\frac{\dot{c} + g}{c}\right)(t_f - t)\frac{\partial}{\partial v}\left(v\rho(v, t_f - t)\right) - \frac{g}{c^2\beta}(t_f - t)\frac{\partial^2}{\partial v^2}\rho(v, t_f - t)\end{aligned} \tag{1.77}$$

By replacing v and $\rho(v, t_f - t)$ by v^* and $\rho(v^*, t)$, above is exactly the Fokker-Planck equation of backward process:

$$\begin{aligned}\frac{\partial}{\partial t}\rho(v^*, t) &= -\frac{\partial}{\partial v^*}\left(-\check{b}(v^*, t)\rho(v^*, t)\right) + \frac{g}{c^{*2}\beta^*}(t)\frac{\partial^2}{\partial v^{*2}}p(v^*, t) \\ &= -\left(\frac{g + \dot{c}}{c}\right)(t_f - t)\frac{\partial}{\partial v^*}\left(v^*\rho(v^*, t)\right) - \frac{g}{c^2\beta}(t_f - t)\frac{\partial^2}{\partial v^{*2}}\rho(v^*, t)\end{aligned} \tag{1.78}$$

Thus, following the backward process with initial distribution $\rho(v^*, 0) = \rho(v, t_f)$, $\forall t \in [0, t_f]$ and $\forall V \in \mathbb{R}$,

$$\rho(v^* = V, t) = \rho(v = V, t_f - t) \tag{1.79}$$

With the initial condition $\rho(v^* = V, t' = 0) = \rho(v = V, t'' = t_f)$, it is known that for any t', $t'' \in [0, t_f]$ satisfying $t_f - t' = t''$, $p(v^* = V, t') = p(v = V, t'')$. By observing the probabilities of v^* and v, it is impossible to distinguish between a result of the forward process from time 0 to time t'' and a result of the backward process from time 0 to time t', such that $t_f - t' = t''$.

Theorem 1.4.1. *With the change of variable $\tilde{v} = c\sqrt{\beta}v$, the backward process v^* will produce the backward process $\tilde{v}^*$.*

Proof. Given $\tilde{v}^* = c^*\sqrt{\beta^*}v^*$,

$$\begin{aligned}
d\tilde{v}^* &= \dot{c^*}\sqrt{\beta^*}v^*dt + c^*\left(\frac{\dot{\beta^*}}{2\sqrt{\beta^*}}\right)v^*dt + c^*\sqrt{\beta^*}dv^* \\
&= \frac{\dot{c^*}}{c^*}\tilde{v}^*dt + \frac{\dot{\beta^*}}{2\beta^*}\tilde{v}^*dt \\
&+ c^*\sqrt{\beta^*}\left(\left(\frac{g - \dot{c^*}}{c^*}\right)v^*dt + \frac{2g}{\beta^* c^{*2}}\frac{\partial}{\partial v^*}\ln p(v^*, t)dt + \frac{1}{c^*}\sqrt{\frac{2g}{\beta^*}}dB(t)\right) \\
&= \left(-\frac{\dot{\beta}}{2\beta} + \frac{g}{c}\right)(t_f - t)\tilde{v}^*dt + \frac{2g}{\sqrt{\beta}c}(t_f - t)\frac{\partial}{\partial v^*}\ln p(v^*, t) + \sqrt{2g}dB(t)
\end{aligned} \tag{1.80}$$

As $\frac{\partial}{\partial v^*} = \frac{\partial \tilde{v}^*}{\partial v^*}\frac{\partial}{\partial \tilde{v}^*} = c^*\sqrt{\beta^*}\frac{\partial}{\partial \tilde{v}^*}$ and $p(\tilde{v}^*, t) = p(v^*, t)|\frac{\partial v^*}{\partial \tilde{v}^*}| = \frac{1}{c^*\sqrt{\beta^*}}p(v^*, t)$,

$$d\tilde{v}^* = \left(-\frac{\dot{\beta}}{2\beta} + \frac{g}{c}\right)(t_f - t)\tilde{v}^*dt + 2g\frac{\partial}{\partial \tilde{v}^*}\ln p(\tilde{v}^*, t) + \sqrt{2g}dB(t) \tag{1.81}$$

It is the backward process of $\tilde{v}$. $\square$

The core of fluctuation theorem is to measure the irreversibility between the trajectories of the forward process and their time-reversed trajectories of the backward process by relative entropy:

Definition 1.4.2. *If $dP(v(t), t \in [0, t_f])$, $dP(v^*(t), t \in [0, t_f])$, $dP(\tilde{v}(t), t \in [0, t_f])$ and $dP(\tilde{v}^*, t \in [0, t_f])$ are the probability measures of trajectories v, v^*, $\tilde{v}$ and $\tilde{v}^*$ from time 0 to time t_f and v^* and $\tilde{v}^*$ are the time-reversed trajectories of v and $\tilde{v}$, the relative entropy in path space between the trajectories (v) of the forward process and their time-reversed trajectories (v^*) of the backward process*

$$\begin{aligned} &D(dP(v(t), t \in [0, t_f]) \| dP(v^*(t), t \in [0, t_f])) \equiv \\ &\int \ln \frac{dP(v(t), t \in [0, t_f])}{dP(v^*(t), t \in [0, t_f])} dP(v(t), t \in [0, t_f]) \end{aligned} \tag{1.82}$$

and the relative entropy between the trajectories ($\tilde{v}$) of the forward process and their time-reversed trajectories ($\tilde{v}^*$) of the backward process

$$\begin{aligned} &D(dP(\tilde{v}(t), t \in [0, t_f]) \| dP(\tilde{v}^*(t), t \in [0, t_f])) \\ &\equiv \int \ln \frac{dP(\tilde{v}(t), t \in [0, t_f])}{dP(\tilde{v}^*(t), t \in [0, t_f])} dP(\tilde{v}(t), t \in [0, t_f]) \end{aligned} \tag{1.83}$$

In the following theorem, we would like to prove that under the change of variables $\tilde{v} = c\sqrt{\beta}v$ and $\tilde{v}^* = c^*\sqrt{\beta^*}v^*$, the relative entropy is invariant.

Theorem 1.4.3. *As stated in the above definition, v^* and $\tilde{v}^*$ are time-reversed trajectories of v and $\tilde{v}$. If $dP(v(t), t \in [0, t_f])$ and $dP(v^*(t))$ are mutually absolutely continuous [22] and so are $dP(\tilde{v}(t), t \in [0, t_f])$ and $dP(\tilde{v}^*(t), t \in [0, t_f])$,*

$$\ln \frac{dP(v(t), t \in [0, t_f])}{dP(v^*(t), t \in [0, t_f])} = \ln \frac{dP(\tilde{v}(t), t \in [0, t_f])}{dP(\tilde{v}^*(t), t \in [0, t_f])} \tag{1.84}$$

Moreover, if above ratio is an Ito integral of v or $\tilde{v}$ from time 0 to time t_f,

$$\begin{aligned} &D\left(dP(v(t), t \in [0, t_f]) \| dP(v^*(t), t \in [0, t_f])\right) = \\ &D(dP(\tilde{v}(t), t \in [0, t_f]) \| dP(\tilde{v}^*, t \in [0, t_f])) \end{aligned} \tag{1.85}$$

In other words, the relative entropy is invariant under the change of variables from (v, v^*) to $(\tilde{v}, \tilde{v}^*)$.

Proof. Given the time duration $\tilde{T} = [0, t_f]$, n is an positive integer and $\Delta t = \frac{t_f}{2^n}$, let $\tilde{T}_n = \{i\Delta t | i = \{0, \cdots, 2^n\}\}$ be an increasing sequence of the finite subset of $\tilde{T}$, such that $\tilde{T}_n$ becomes dense in $\tilde{T}$ as $n \to \infty$. Let $\mathcal{P}_n$ denote the σ-algebra generated by $\{v(t) : t \in \tilde{T}_n\}$ (as $v^*(t) = v(t_f - t)$, it is also the σ-algebra generated by $\{v^*(t) : t \in \tilde{T}_n\}$). Based on [22],

$$\begin{aligned} L_n &= \left\langle \frac{dP(v(t), t \in [0, t_f])}{dP(v^*(t), t \in [0, t_f])} \middle| \mathcal{P}_n \right\rangle \text{ w.r.t. } dP(v^*(t), t \in [0, t_f]) \\ &= \frac{P(\{v(t) : t \in \tilde{T}_n\})}{P(\{v^*(t) : t \in \tilde{T}_n\})} \end{aligned} \tag{1.86}$$

Similarly,

$$\tilde{L}_n = \frac{P(\{\tilde{v}(t) : t \in \tilde{T}_n\})}{P(\{\tilde{v}^*(t) : t \in \tilde{T}_n\})} \tag{1.87}$$

As $\tilde{v} = c\sqrt{\beta}v$ and $\tilde{v}^* = c^*\sqrt{\beta^*}v^*$,

$$\frac{P(\{\tilde{v}(t) : t \in \tilde{T}_n\})}{P(\{\tilde{v}^*(t) : t \in \tilde{T}_n\})} = \frac{\left(\prod_{i=0}^{2^n} \frac{1}{\sqrt{\beta(i\Delta t)}c(i\Delta t)}\right) P(\{v(t) : t \in \tilde{T}_n\})}{\left(\prod_{i=0}^{2^n} \frac{1}{\sqrt{\beta^*(i\Delta t)}c^*(i\Delta t)}\right) P(\{v^*(t) : t \in \tilde{T}_n\})} \tag{1.88}$$

Because $c^*(i\Delta t) = c(2^n\Delta t - i\Delta t)$ and $\beta^*(i\Delta t) = \beta(2^n\Delta t - i\Delta t)$,

$$\begin{aligned} \prod_{i=0}^{2^n} \frac{1}{\sqrt{\beta(i\Delta t)}c(i\Delta t)} &= \prod_{i=0}^{2^n} \frac{1}{\sqrt{\beta^*(i\Delta t)}c^*(i\Delta t)} \\ \tilde{L}_n = \frac{P(\{\tilde{v}(t) : t \in \tilde{T}_n\})}{P(\{\tilde{v}^*(t) : t \in \tilde{T}_n\})} &= \frac{P(\{v(t) : t \in \tilde{T}_n\})}{P(\{v^*(t) : t \in \tilde{T}_n\})} = L_n \end{aligned} \tag{1.89}$$

Moreover, based on [22],

$$\begin{aligned} \lim_{n\to\infty} L_n &= \frac{dP(v(t), t \in [0, t_f])}{dP(v^*(t), t \in [0, t_f])} \qquad \text{a.s. and in 1-mean} \\ \lim_{n\to\infty} \tilde{L}_n &= \frac{dP(\tilde{v}(t), t \in [0, t_f])}{dP(\tilde{v}^*(t), t \in [0, t_f])} \qquad \text{a.s. and in 1-mean} \end{aligned} \tag{1.90}$$

Thus,

$$\begin{aligned}\frac{dP(v(t), t\in[0,t_f])}{dP(v^*(t), t\in[0,t_f])} &= \frac{dP(\tilde{v}(t), t\in[0,t_f])}{dP(\tilde{v}^*(t), t\in[0,t_f])} \\ \ln\frac{dP(v(t), t\in[0,t_f])}{dP(v^*(t), t\in[0,t_f])} &= \ln\frac{dP(\tilde{v}(t), t\in[0,t_f])}{dP(\tilde{v}^*(t), t\in[0,t_f])}\end{aligned} \tag{1.91}$$

If the ratio is an Ito integral of v or $\tilde{v}$ from time 0 to time t_f, i.e.

$$\begin{aligned}\ln\frac{dP(v(t), t\in[0,t_f])}{dP(v^*(t), t\in[0,t_f])} &= \int_0^{t_f} f_1(v,t)dt + \int_0^{t_f} f_2(v,t)dB(t) \\ \ln\frac{dP(\tilde{v}(t), t\in[0,t_f])}{dP(\tilde{v}^*(t), t\in[0,t_f])} &= \int_0^{t_f} f_1(\tilde{v},t)dt + \int_0^{t_f} f_2(\tilde{v},t)dB(t)\end{aligned} \tag{1.92}$$

As $\tilde{v}(t) = c(t)\sqrt{\beta(t)}v(\mathrm{t}),\ \rho(v,t)dv = \rho(\tilde{v},t)d\tilde{v}$

$$\begin{aligned}&D\left(dP(v(t), t\in[0,t_f])\|dP(v^*(t), t\in[0,t_f])\right) = \left\langle \ln\frac{dP(v(t), t\in[0,t_f])}{dP(v^*(t), t\in[0,t_f])}\right\rangle \\ &= \int_0^{t_f}\langle f_1(v,t)\rangle\, dt + \int_0^{t_f}\langle f_2(v,t)dB(t)\rangle \\ &= \int_0^{t_f}\langle f_1(v,t)\rangle\, dt = \int_0^{t_f}\langle f_1(\tilde{v},t)\rangle\, dt \\ &= \left\langle \ln\frac{dP(\tilde{v}(t), t\in[0,t_f])}{dP(\tilde{v}^*(t), t\in[0,t_f])}\right\rangle \\ &= D(dP(\tilde{v}(t), t\in[0,t_f])\|dP(\tilde{v}^*(t), t\in[0,t_f]))\end{aligned} \tag{1.93}$$

□

To compute the relative entropy between the trajectories of the forward process and the time-reversed trajectories of the backward process, we would like to compare each of them with a diffusion driven by a Wiener process alone. First, let us compare

$$d\tilde{v} = bdt + \sqrt{2g}dB(t),\ b = -\left(-\frac{\dot{\beta}}{2\beta} + \frac{g}{c}\right) \tag{1.94}$$

with

$$dy = \sqrt{2g}dB(t) \tag{1.95}$$

Given the initial value $\tilde{v}(0) = y(0) = \tilde{v}_0$, from time 0 to time t_f, by Girsanov's theorem [17], assuming that $\left\langle \exp\left(\frac{1}{2}\int_0^{t_f} \frac{b^2}{2g} dt\right)\right\rangle < \infty$,

$$\frac{dP(\tilde{v}(t), t \in [0, t_f]|\tilde{v}_0)}{dP(y(t) = \tilde{v}(t), t \in [0, t_f]|\tilde{v}_0)} = \exp\left(\int_0^{t_f} \frac{b}{\sqrt{2g}} dB(t) - \frac{b^2}{4g} dt\right) \tag{1.96}$$

$$= \exp\left(\int_0^{t_f} \frac{b}{\sqrt{2g}} \circ dB(t) - \frac{\nabla b}{2} dt - \frac{b^2}{4g} dt\right)$$

Secondly, we would like to compare the backward process

$$d\tilde{v}^* = -\bar{b}dt + \sqrt{2g}dB(t),\ \bar{b} = -\left(-\frac{\dot{\beta}}{2\beta} + \frac{g}{c}\right)(t_f - t)\tilde{v}^* dt - 2g\frac{\partial}{\partial \tilde{v}^*} \ln \rho(\tilde{v}^*, t) \tag{1.97}$$

with a diffusion process

$$dy^* = \sqrt{2g}dB(t) \tag{1.98}$$

Given the initial condition $\tilde{v}^*(0) = y^*(0) = \tilde{v}_{t_f}$, by Girsanov's theorem again [17], assuming that $\left\langle \exp\left(\frac{1}{2}\int_0^{t_f} \frac{\bar{b}^2}{2g} dt\right)\right\rangle < \infty$,

$$\frac{dP(\tilde{v}^*(t), t \in [0, t_f]|\tilde{v}_{t_f})}{dP(y^*(t) = \tilde{v}^*(t), t \in [0, t_f]|\tilde{v}_{t_f})} = \exp\left(\int_0^{t_f} -\frac{\bar{b}}{\sqrt{2g}} dB(t) - \frac{\bar{b}^2}{4g} dt\right) \tag{1.99}$$

$$= \exp\left(\int_0^{t_f} -\frac{\bar{b}}{\sqrt{2g}} \circ dB(t) + \frac{\nabla \bar{b}}{2} dt - \frac{\bar{b}^2}{4g} dt\right)$$

As $\tilde{v}^*(t) = \tilde{v}(t_f - t)$ and due to the reversibility of Wiener process, $dP(y^*(t) = \tilde{v}^*(t), t \in [0, t_f]|\tilde{v}_{t_f}) = dP(y(t) = \tilde{v}(t), t \in [0, t_f]|\tilde{v}_0)$,

$$\left\langle \ln \frac{dP(\tilde{v}(t), t \in [0, t_f]|\tilde{v}_0)}{dP(\tilde{v}^*(t), t \in [0, t_f]|\tilde{v}_{t_f})}\right\rangle = \left\langle \int_0^{t_f} \left(\frac{b + \bar{b}}{2g}\right) \circ d\tilde{v} - \int_0^{t_f} \frac{\nabla_{\tilde{v}}(b + \bar{b}) + (2g)^{-1}\left(b^2 - \bar{b}^2\right)}{2} dt\right\rangle \tag{1.100}$$

Let us first consider the integral $-\int_0^{t_f} \frac{\nabla_{\tilde{v}}\left(b+\bar{b}\right)+(2g)^{-1}(b^2-\bar{b}^2)}{2} dt$. By Fokker-Planck equation (1.65), it is known that

$$\frac{\partial}{\partial t}\rho(\tilde{v}, t) = -\frac{\partial}{\partial \tilde{v}}\left(\rho(\tilde{v}, t)\left(\frac{b + \bar{b}}{2}\right)\right) \tag{1.101}$$

Thus,

$$\begin{aligned}
\frac{\partial}{\partial t}\ln\rho(\tilde{v},t) &= \frac{1}{\rho(\tilde{v},t)}\frac{\partial}{\partial t}\rho(\tilde{v},t) \\
&= \frac{1}{\rho(\tilde{v},t)}\left(-\frac{\partial}{\partial\tilde{v}}\left(\frac{b+\bar{b}}{2}\right)\right) \\
&= -\nabla_{\tilde{v}}\left(\frac{b+\bar{b}}{2}\right) - \frac{(2g)^{-1}(b^2-\bar{b}^2)}{2}
\end{aligned} \tag{1.102}$$

The integral is

$$-\int_0^{t_f}\frac{\nabla_{\tilde{v}}\left(b+\bar{b}\right)+(2g)^{-1}\left(b^2-\bar{b}^2\right)}{2}dt = \int_0^{t_f}\frac{\partial}{\partial t}\ln\rho(\tilde{v},t)dt = \ln\rho(\tilde{v},t_f) - \ln\rho(\tilde{v},0) \tag{1.103}$$

The expectation

$$\left\langle \ln\frac{dP(\tilde{v}(t),t\in[0,t_f]|\tilde{v}_0)}{dP(\tilde{v}^*(t),t\in[0,t_f]|\tilde{v}_{t_f})}\right\rangle = \left\langle\int_0^{t_f}\left(\frac{b+\bar{b}}{2g}\right)\circ d\tilde{v}\right\rangle + \langle\ln\rho(\tilde{v},t_f)\rangle - \langle\ln\rho(\tilde{v},0)\rangle \tag{1.104}$$

Therefore,

$$\begin{aligned}
&D\left(dP(\tilde{v}(t),t\in[0,t_f])\|dP(\tilde{v}^*(t),t\in[0,t_f])\right) \\
=&\left\langle\ln\frac{dP(\tilde{v}(t),t\in[0,t_f]|\tilde{v}_0)}{dP(\tilde{v}^*(t),t\in[0,t_f]|\tilde{v}_{t_f})}\right\rangle + \left\langle\ln\frac{\rho(\tilde{v},0)}{\rho(\tilde{v},t_f)}\right\rangle \\
=&\left\langle\int_0^{t_f}\left(\frac{b+\bar{b}}{2g}\right)\circ d\tilde{v}\right\rangle = \int_0^{t_f}\frac{\left\langle\left(b+\bar{b}\right)^2\right\rangle}{4g}dt
\end{aligned} \tag{1.105}$$

The entropy production rate $\frac{\langle(b+\bar{b})^2\rangle}{4g} = \frac{\langle w_1^2\rangle}{g}$ which is the consistent with the result from the stochastic thermodynamics analysis. In summary, the entropy production rate in the regime where the system is operated near its equilibrium is $\frac{c\dot{\beta}^2}{4g\beta^2}$. Combining this result with the positive-semidefinite inverse diffusion tensor from last section, we have a new inverse diffusion tensor to express the entropy production rate

$$d = \begin{pmatrix} \dot{c} & \dot{\beta} \end{pmatrix} \begin{pmatrix} \frac{1}{4cg} & \frac{1}{4\beta g} \\ \frac{1}{4\beta g} & \frac{c}{2\beta^2 g} \end{pmatrix} \begin{pmatrix} \dot{c} \\ \dot{\beta} \end{pmatrix} \tag{1.106}$$

and this new inverse diffusion tensor is positive-definite. Taking it as a metric tensor in the space of (c, β), it is seen that the sectional curvature is zero. In other words, this is a flat parameter surface. Furthermore, the average dissipation rate in the unit of $k_B T/\text{sec}$ is

$$\tilde{d} = \begin{pmatrix} \dot{c} & \dot{\beta} \end{pmatrix} \begin{pmatrix} \frac{1}{4cg\beta} & \frac{1}{4\beta^2 g} \\ \frac{1}{4\beta^2 g} & \frac{c}{2\beta^3 g} \end{pmatrix} \begin{pmatrix} \dot{c} \\ \dot{\beta} \end{pmatrix} \tag{1.107}$$

Remark 1.4.4. *Given that the new average dissipation rate d is positive, it tells us that the resistor-capacitor system produces entropy into its surroundings once it is perturbed from equilibrium by changing the system parameter $\lambda = (c, \beta)$. In the sense of statistical mechanics, the system is a dissipative system.*

1.4.3 Dissipation from different processes

When a system is perturbed from equilibrium with probability density $\rho_{eq,\lambda}$ by a change of system parameter of $\delta\lambda$, it will first arrive at a non-equilibrium probability density $\rho_{neq,\lambda+\delta\lambda}$ and then it will relax to the new equilibrium associated to the equilibrium probability density $\rho_{eq,\lambda+\delta\lambda}$. As $D(\rho_{neq,\lambda+\delta\lambda} \| \rho_{eq,\lambda+\delta\lambda}) = \beta(\lambda + \delta\lambda)\,(F_{neq,\lambda+\delta\lambda} - F_{eq,\lambda+\delta\lambda})$, where $\beta(\lambda+\delta\lambda)$ is the inverse temperature of the heat bath with system parameter $\lambda + \delta\lambda$, the first dissipation is due to the relaxation of the system from $\rho_{neq,\lambda+\delta\lambda}$ to $\rho_{eq,\lambda+\delta\lambda}$, which is described by the inverse diffusion tensor introduced in the work [11]. The dissipation due to the perturbation from $\rho_{eq,\lambda}$

to $\rho_{neq,\lambda+\delta\lambda}$ is accounted for by stochastic thermodynamics and fluctuation theorem in parallel. Thus, dissipation comes from two different processes. The entropy production rate computed by stochastic thermodynamics and fluctuation theorem is distinct from the entropy production rate computed by the original inverse diffusion tensor from [11].

Secondly, the entropy production rate due to the change of β from fluctuation theorem (also from stochastic thermodynamics) is $\frac{c\dot{\beta}^2}{4g\beta^2}$. Analogously for the colloidal particle dynamics which is the limit of the stochastic oscillator when $m \to 0$, the entropy production rate term associated with $\dot{\beta}$ is $\frac{\zeta k\dot{\beta}^2}{4\beta^2}$. Comparing it with the entropy production rate from the inverse diffusion tensor of the stochastic oscillator associated with $\dot{\beta}^2$, which is $\frac{\dot{\beta}^2}{\beta^2}\left(\frac{m}{\zeta} + \frac{\zeta}{4k}\right)$. When k is a very small number, as we shall see in an example in Chapter 5, the term $\frac{\zeta\dot{\beta}^2}{4k\beta^2}$ is much larger than $\frac{\zeta k\dot{\beta}^2}{4\beta^2}$. It is an appropriate approximation to exclude the dissipation from the fluctuation theorem for the stochastic oscillator case.

Also, it should be pointed out that stochastic thermodynamics is not a methodology to compute the dissipation of the system with kinetic energy, and fluctuation theorem itself is not a methodology to compute the dissipation of the system whose diffusion coefficient matrix is singular, as Girsanov's theorem cannot be directly applied to it.

Chapter 2: Sub-Riemannian Geometry Structures of the Heat Engines

2.1 The extracted work and total heat supply

As a heat engine is a cyclically operated machine which can extract heat from its heat bath to output mechanical work, it is essential to define the *extracted work* of the engine and the *total heat supply* from heat bath to it for efficiency analysis of the heat engine. For our stochastic oscillator and the resistor-capacitor circuit, these quantities will be defined as the basis for later investigation of efficiencies of these heat engines. For the stochastic oscillator,

$$\begin{aligned} d\xi_1 &= \frac{\xi_2}{m}dt \\ d\xi_2 &= -k\xi_1 dt - \zeta\frac{\xi_2}{m}dt + \sqrt{\frac{2\zeta}{\beta}}dB(t) \end{aligned} \tag{2.1}$$

Given that its internal energy U_{neq} is the average value of $H = \frac{k\xi_1^2}{2} + \frac{\xi_2^2}{2m}$, by Ito's rule [17],

$$\begin{aligned} dH &= \frac{\partial H}{\partial k}\dot{k}dt + \frac{\partial H}{\partial \xi_1}d\xi_1 + \frac{\partial H}{\partial \xi_2}d\xi_2 + \frac{1}{2}\text{tr}\left(\begin{pmatrix} \frac{\partial^2 H}{\partial \xi_1^2} & \frac{\partial^2 H}{\partial \xi_1 \partial \xi_2} \\ \frac{\partial^2 H}{\partial \xi_1 \partial \xi_2} & \frac{\partial^2 H}{\partial \xi_2^2} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & \frac{2\zeta}{\beta} \end{pmatrix}\right)dt \\ &= \frac{\dot{k}\xi_1^2}{2}dt - \zeta\frac{\xi_2^2}{m^2}dt + \frac{\xi_2}{m}\sqrt{\frac{2\zeta}{\beta}}dB(t) + \frac{\zeta}{m\beta}dt \end{aligned} \tag{2.2}$$

Taking average on both sides of above equation,

$$\dot{U}_{neq} = \left\langle \frac{\dot{k}\xi_1^2}{2} \right\rangle - \left\langle \zeta \frac{\xi_2^2}{m^2} \right\rangle + \frac{\zeta}{m\beta} \tag{2.3}$$

where $\dot{U}_{neq}$ is the rate of change of the internal energy, $-\left\langle \frac{\dot{k}\xi_1^2}{2} \right\rangle$ is the *extracted mechanical power* of the stochastic oscillator. In consequence, from time 0 to time t_f, the integral $\int_0^{t_f} -\left\langle \frac{\dot{k}\xi_1^2}{2} \right\rangle dt$ is the *extracted work* of the stochastic oscillator. The difference between the fluctuation and dissipation $\frac{\zeta}{m\beta} - \left\langle \zeta \frac{\xi_2^2}{m^2} \right\rangle = \dot{U}_{neq} - \left\langle \frac{\dot{k}\xi_1^2}{2} \right\rangle$ is the heat flux associated with the heat bath. If it is positive, the heat flux is injected from the heat bath to the engine. If it is negative, the heat flux is flowing from the heat engine to the heat bath. Thus, the integral of the positive heat flux from time 0 to time t_f

$$\int_0^{t_f} \left(\dot{U}_{neq} - \left\langle \frac{\dot{k}\xi_1^2}{2} \right\rangle \right) \mathbb{1} \left\{ \dot{U}_{neq} - \left\langle \frac{\dot{k}\xi_1^2}{2} \right\rangle > 0 \right\} dt \tag{2.4}$$

is the *heat supply* from the heat bath to the heat engine. If we approximate the integrands inside these two integrals with equilibrium averages ($U_{eq,\lambda} = \frac{1}{\beta}$ and $\left\langle \frac{k\xi_1^2}{2} \right\rangle_{eq,\lambda} = \frac{1}{2\beta}$) in the near-equilibrium regime,

$$\begin{aligned} &\int_0^{t_f} -\left\langle \frac{\dot{k}\xi_1^2}{2} \right\rangle_{eq,\lambda} dt = \int_0^{t_f} -\frac{\dot{k}}{2\beta k} dt \\ &\int_0^{t_f} \left(\dot{U}_{eq,\lambda} - \left\langle \frac{\dot{k}\xi_1^2}{2} \right\rangle_{eq,\lambda} \right) \mathbb{1} \left\{ \dot{U}_{eq,\lambda} - \left\langle \frac{\dot{k}\xi_1^2}{2} \right\rangle_{eq,\lambda} > 0 \right\} dt \\ &\quad = \int_0^{t_f} \left(-\frac{\dot{\beta}}{\beta^2} - \frac{\dot{k}}{2\beta k} \right) \mathbb{1} \left\{ -\frac{\dot{\beta}}{\beta^2} - \frac{\dot{k}}{2\beta k} > 0 \right\} dt \end{aligned} \tag{2.5}$$

Moreover, besides the heat supply from the heat bath to the heat engine, due to the fact that the engine is operating in finite-time, the heat bath will also provide the extra energy dissipated into the surroundings by the engine which accompanies heat transfer between the engine and the heat bath. The *total heat supply* from the heat bath is the sum of the heat supply from the heat bath to the heat engine and the dissipated energy of the heat engine, which in the near-equilibrium regime, approximately equals to

$$\int_0^{t_f} \left(\left(-\frac{\dot{\beta}}{\beta^2} - \frac{\dot{k}}{2\beta k} \right) \mathbb{1} \left\{ -\frac{\dot{\beta}}{\beta^2} - \frac{\dot{k}}{2\beta k} > 0 \right\} + \begin{pmatrix} \dot{\beta} & \dot{k} \end{pmatrix} g[\lambda] \begin{pmatrix} \dot{\beta} \\ \dot{k} \end{pmatrix} \right) dt \tag{2.6}$$

where $g[\lambda]$ is the inverse diffusion tensor of the stochastic oscillator at $\lambda = (\beta, k)$.

In the case of the resistor-capacitor system,

$$dv = -\left(\frac{\dot{c} + g}{c} \right) v dt + \frac{1}{c} \sqrt{\frac{2g}{\beta}} dB(t) \tag{2.7}$$

The total energy of the capacitor is the potential energy stored in the capacitor $H = \frac{cv^2}{2}$. By Ito's rule [17] again,

$$\begin{aligned} dH &= \frac{\partial}{\partial c} \frac{cv^2}{2} \dot{c} dt + \frac{\partial H}{\partial v} dv + \frac{1}{2} \frac{\partial^2}{\partial v^2} \frac{cv^2}{2} \frac{2g}{c^2 \beta} dt \\ &= -\frac{\dot{c}v^2}{2} dt - gv^2 dt + v \sqrt{\frac{2g}{\beta}} dB(t) + \frac{g}{c\beta} dt \end{aligned}$$

As the internal energy of the capacitor $U_{neq} = \left\langle \frac{cv^2}{2} \right\rangle$, we have the *First law of thermodynamics*,

$$\dot{U}_{neq} = -\left\langle \frac{\dot{c}v^2}{2} \right\rangle + \frac{g}{c\beta} - \langle gv^2 \rangle \tag{2.8}$$

where $\dot{U}_{neq}$ is the rate of change of the internal energy, $\left\langle \frac{\dot{c}v^2}{2} \right\rangle$ is the *extracted mechanical power* of the capacitor. The integral from time 0 to time t_f

$$\int_0^{t_f} \left\langle \frac{\dot{c}v^2}{2} \right\rangle dt \tag{2.9}$$

is the *extracted work* of the capacitor. Similar to the stochastic oscillator case, $\frac{g}{c\beta} - \langle gv^2 \rangle = \dot{U}_{neq} + \left\langle \frac{\dot{c}v^2}{2} \right\rangle$ is the heat flux from the heat bath and the engine.

Similar to the case of the stochastic oscillator, the integral

$$\int_0^{t_f} \left(\dot{U}_{neq} + \left\langle \frac{\dot{c}v^2}{2} \right\rangle \right) \mathbb{1} \left\{ \dot{U}_{neq} + \left\langle \frac{\dot{c}v^2}{2} \right\rangle > 0 \right\} dt \tag{2.10}$$

is the *heat supply* from the heat bath to the heat engine from time 0 to time t_f. Assuming that the resistor-capacitor circuit is operated near the equilibrium, we can approximate the integrands in these two integrals with the equilibrium ones ($\left\langle \frac{cv^2}{2} \right\rangle_{eq,\lambda} = \frac{1}{2\beta}$), such as,

$$\begin{aligned} &\int_0^{t_f} \left\langle \frac{\dot{c}v^2}{2} \right\rangle_{eq,\lambda} dt = \int_0^{t_f} \frac{\dot{c}}{2c\beta} dt \\ &\int_0^{t_f} \left(\dot{U}_{eq,\lambda} + \left\langle \frac{\dot{c}v^2}{2} \right\rangle_{eq,\lambda} \right) \mathbb{1} \left\{ \dot{U}_{eq,\lambda} + \left\langle \frac{\dot{c}v^2}{2} \right\rangle_{eq,\lambda} > 0 \right\} dt = \\ &\int_0^{t_f} \left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} \right) \mathbb{1} \left\{ -\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} > 0 \right\} dt \end{aligned} \tag{2.11}$$

Taking the dissipation from the heat engine to its surroundings as a portion of the *total heat supply* from the heat bath, the total heat supply from the heat bath is, in the near-equilibrium regime,

$$\int_0^{t_f} \left(\left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} \right) \mathbb{1} \left\{ -\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} > 0 \right\} + \begin{pmatrix} \dot{c} & \dot{\beta} \end{pmatrix} \begin{pmatrix} \frac{1}{4cg\beta} & \frac{1}{4g\beta^2} \\ \frac{1}{4g\beta^2} & \frac{c}{2g\beta^3} \end{pmatrix} \begin{pmatrix} \dot{c} \\ \dot{\beta} \end{pmatrix} \right) dt \tag{2.12}$$

In this section, based on the *First law of thermodynamics* and Ito's rule, we have given definition of the extracted work of the heat engines and the total heat supply from the heat bathes in both systems. These are basic quantities in the analysis of the efficiencies of our heat engines.

2.2 Correction terms for extracted mechanical power of the heat engines

In the previous section, we approximated non-equilibrium averages with equilibrium averages with the near-equilibrium assumption of both heat engines. However, there are still differences between the equilibrium averages and the near-equilibrium averages. The discussion over these differences is inspired by the work [23]. In this section, we will give the correction term for the extracted mechanical power of each heat engine.

In the case of the stochastic oscillator, from time $-\infty$ to time 0, k and β are fixed. Then the oscillator is at the corresponding thermodynamic equilibrium. The dynamics of the average of (ξ_1, ξ_2)

$$\frac{d}{dt}\begin{pmatrix} \langle \xi_1 \rangle \\ \langle \xi_2 \rangle \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{m} \\ -k & -\frac{\zeta}{m} \end{pmatrix}\begin{pmatrix} \langle \xi_1 \rangle \\ \langle \xi_2 \rangle \end{pmatrix} \tag{2.13}$$

where $\langle\,\rangle$ is the average evaluated with non-equilibrium distribution and $\langle\,\rangle_{eq,\lambda}$ is the average evaluated with the equilibrium distribution of the thermodynamic system with the system parameter λ. As matrix $\begin{pmatrix} 0 & \frac{1}{m} \\ -k & -\frac{\zeta}{m} \end{pmatrix}$ has negative eigenvalues, $(\langle \xi_1 \rangle, \langle \xi_2 \rangle)$ converges to the zero vector exponentially. At time 0, we can assume $(\langle \xi_1 \rangle, \langle \xi_2 \rangle) = (0, 0)$. Thus,

$$\Sigma(t) = \begin{pmatrix} \langle \xi_1^2(t) \rangle & \langle \xi_1 \xi_2(t) \rangle \\ \langle \xi_1 \xi_2(t) \rangle & \langle \xi_2^2(t) \rangle \end{pmatrix} \tag{2.14}$$

is the covariance matrix of the stochastic oscillator. By Ito's formula [17],

$$\dot{\Sigma}(t) = A(t)\Sigma(t) + \Sigma(t)A^{\intercal}(t) + BB^{\intercal}(t) \tag{2.15}$$

where

$$A(t) = \begin{pmatrix} 0 & \frac{1}{m} \\ -k(t) & -\frac{\zeta}{m} \end{pmatrix}$$

$$B(t) = \begin{pmatrix} 0 \\ \sqrt{\frac{2\zeta}{\beta(t)}} \end{pmatrix} \quad BB^{\intercal}(t) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{2\zeta}{\beta(t)} \end{pmatrix}$$

If at time t, the system is at thermodynamic equilibrium, $\langle \xi_1^2(t) \rangle_{eq,\lambda(t)} = \frac{1}{k(t)\beta(t)}$, $\langle \xi_1(t)\xi_2(t) \rangle_{eq,\lambda(t)} = 0$ and $\langle \xi_2^2(t) \rangle_{eq,\lambda(t)} = \frac{m}{\beta(t)}$. The covariance matrix $\Sigma_{eq,\lambda(t)}$ comprises these terms and

$$A(t)\Sigma_{eq,\lambda(t)} + \Sigma_{eq,\lambda(t)}A^{\intercal}(t) + BB^{\intercal}(t) = 0 \tag{2.16}$$

At time t ($t > 0$), the system is perturbed by $(\dot{k}, \dot{\beta})$. At time $t+\epsilon$ ($0 < \epsilon \ll 1$), $\Sigma(t+\epsilon) = \Sigma_{eq,\lambda(t)} + \Delta\Sigma(t,\epsilon)$. Here $\Delta\Sigma(t,\epsilon)$ is the non-equilibrium correction term and $k(t+\epsilon) = k(t) + \epsilon\dot{k}$, $\beta(t+\epsilon) = \beta(t) + \epsilon\dot{\beta}$. The matrix $\Delta\Sigma(t,\epsilon)$ is symmetric and

$$\Delta\Sigma(t,\epsilon) = \begin{pmatrix} \Delta\sigma_{11}(t,\epsilon) & \Delta\sigma_{12}(t,\epsilon) \\ \Delta\sigma_{21}(t,\epsilon) & \Delta\sigma_{22}(t,\epsilon) \end{pmatrix} \tag{2.17}$$

Therefore,

$$\dot{\Sigma}(t+\epsilon) = A(t+\epsilon)\Sigma(t+\epsilon) + \Sigma(t+\epsilon)A^{\intercal}(t+\epsilon) + BB^{\intercal}(t+\epsilon) \tag{2.18}$$

As

$$\begin{pmatrix} 0 & \frac{1}{m} \\ -(k(t)+\epsilon\dot{k}) & -\frac{\zeta}{m} \end{pmatrix} = \epsilon \begin{pmatrix} 0 & 0 \\ -\dot{k} & 0 \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{m} \\ -k(t) & -\frac{\zeta}{m} \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 \\ 0 & \frac{2\zeta}{\beta(t)+\epsilon\dot{\beta}} \end{pmatrix} = \epsilon \begin{pmatrix} 0 & 0 \\ 0 & -\frac{2\zeta\dot{\beta}}{\beta(t)^2} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \frac{2\zeta}{\beta(t)} \end{pmatrix} + O(\epsilon^2) \tag{2.19}$$

$$\begin{aligned}\dot{\Sigma}(t+\epsilon) = \epsilon \begin{pmatrix} 0 & 0 \\ -\dot{k} & 0 \end{pmatrix} \Sigma_{eq,\lambda(t)} + A(t)\Delta\Sigma(t,\epsilon) + \epsilon\Sigma_{eq,\lambda(t)} \begin{pmatrix} 0 & -\dot{k} \\ 0 & 0 \end{pmatrix} \\ + \Delta\Sigma(t,\epsilon)A^{\intercal}(t) + \epsilon \begin{pmatrix} 0 & 0 \\ 0 & -\frac{2\zeta\dot{\beta}}{\beta^2(t)} \end{pmatrix} + O(\epsilon^2)\end{aligned} \tag{2.20}$$

If $\epsilon \to 0$, equation (2.20) becomes

$$\dot{\Sigma}(t) = A(t)\Delta\Sigma(t,0) + \Delta\Sigma(t,0)A^{\intercal}(t) \tag{2.21}$$

On the other hand,

$$\dot{\Sigma}(t) = \dot{\Sigma}_{eq,\lambda(t)} + \frac{d}{dt}\Delta\Sigma(t,0) \tag{2.22}$$

where,

$$\dot{\Sigma}_{eq,\lambda} = \begin{pmatrix} \frac{d}{dt}\frac{1}{k(t)\beta(t)} & 0 \\ 0 & \frac{d}{dt}\frac{m}{\beta(t)} \end{pmatrix}$$

$$\frac{d}{dt}\Delta\Sigma(t,0) = \begin{pmatrix} \frac{d}{dt}\Delta\sigma_{11}(t,0) & \frac{d}{dt}\Delta\sigma_{12}(t,0) \\ \frac{d}{dt}\Delta\sigma_{12}(t,0) & \frac{d}{dt}\Delta\sigma_{22}(t,0) \end{pmatrix}$$

Based on (2.21) and (2.22)

$$A(t)\Delta\Sigma(t,0) + \Delta\Sigma(t,0)A^{T}(t) - \frac{d}{dt}\Delta\Sigma(t,0) = \dot{\Sigma}_{eq,\lambda(t)} \tag{2.23}$$

In the form of a linear system,

$$\frac{d}{dt}\begin{pmatrix}\frac{1}{k(t)\beta(t)}\\ 0\\ \frac{m}{\beta(t)}\end{pmatrix} = -\frac{d}{dt}\begin{pmatrix}\Delta\sigma_{11}(t,0)\\ \Delta\sigma_{12}(t,0)\\ \Delta\sigma_{22}(t,0)\end{pmatrix} + \begin{pmatrix}0 & \frac{2}{m} & 0\\ -k(t) & -\frac{\zeta}{m} & \frac{1}{m}\\ 0 & -2k(t) & -\frac{2\zeta}{m}\end{pmatrix}\begin{pmatrix}\Delta\sigma_{11}(t,0)\\ \Delta\sigma_{12}(t,0)\\ \Delta\sigma_{22}(t,0)\end{pmatrix} \tag{2.24}$$

$$R(t) = \begin{pmatrix}0 & \frac{2}{m} & 0\\ -k(t) & -\frac{\zeta}{m} & \frac{1}{m}\\ 0 & 2k(t) & -\frac{2\zeta}{m}\end{pmatrix} \tag{2.25}$$

Its inverse is

$$R^{-1}(t) = -\frac{m^2}{4k\zeta}\begin{pmatrix}\frac{2\zeta^2}{m^2}+\frac{2k}{m} & \frac{4\zeta}{m^2} & \frac{2}{m^2}\\ -\frac{2\zeta k}{m} & 0 & 0\\ 2k^2 & 0 & \frac{2k}{m}\end{pmatrix} \tag{2.26}$$

Thus,

$$R^{-1}(t)\frac{d}{dt}\begin{pmatrix}\frac{1}{k(t)\beta(t)}\\ 0\\ \frac{m}{\beta(t)}\end{pmatrix} = \left(\mathbb{1} - R^{-1}(t)\frac{d}{dt}\right)\begin{pmatrix}\Delta\sigma_{11}(t,0)\\ \Delta\sigma_{12}(t,0)\\ \Delta\sigma_{22}(t,0)\end{pmatrix} \tag{2.27}$$

As the velocities $\dot{k}$ and $\dot{\beta}$ are very small, up to the first order time-derivative,

$$\begin{aligned}\begin{pmatrix}\Delta\sigma_{11}(t,0)\\ \Delta\sigma_{12}(t,0)\\ \Delta\sigma_{22}(t,0)\end{pmatrix} &\approx R^{-1}(t)\frac{d}{dt}\begin{pmatrix}\frac{1}{k(t)\beta(t)}\\ 0\\ \frac{m}{\beta(t)}\end{pmatrix}\\ &= -\frac{m^2}{4k(t)\zeta}\begin{pmatrix}\frac{2\zeta^2}{m^2}+\frac{2k(t)}{m} & \frac{4\zeta}{m^2} & \frac{2}{m^2}\\ -\frac{2\zeta k(t)}{m} & 0 & 0\\ 2k^2(t) & 0 & \frac{2k(t)}{m}\end{pmatrix}\begin{pmatrix}\frac{d}{dt}\frac{1}{k(t)\beta(t)}\\ 0\\ \frac{d}{dt}\frac{m}{\beta(t)}\end{pmatrix}\end{aligned} \tag{2.28}$$

Based on this approximation, we have the non-equilibrium correction term

$$
\begin{aligned}
\Delta\sigma_{11}(t,0) &\approx -\frac{m^2}{4k(t)\zeta}\left(\left(\frac{2\zeta^2}{m^2}+\frac{2k(t)}{m}\right)\left(\frac{d}{dt}\frac{1}{k(t)\beta(t)}\right)+\frac{2}{m^2}\left(\frac{d}{dt}\frac{m}{\beta(t)}\right)\right) \\
&= \frac{m^2}{4k(t)\zeta}\left(\frac{2\zeta^2+2mk}{m^2k^2(t)\beta(t)}\dot{k}+\frac{2\zeta^2+4mk}{m^2k(t)\beta^2(t)}\dot{\beta}\right) \\
&= \frac{\zeta^2+mk(t)}{2\zeta k^3(t)\beta(t)}\dot{k}+\frac{\zeta^2+2mk(t)}{2\zeta k^2(t)\beta^2(t)}\dot{\beta}
\end{aligned}
\tag{2.29}
$$

At time t, the extracted power with above non-equilibrium correction term is

$$
\begin{aligned}
\dot{\psi} &= -\left\langle\frac{\partial}{\partial k}\frac{k\xi_1^2(t)}{2}\right\rangle\dot{k} = -\frac{\sigma_{11}(t)}{2}\dot{k} \\
&\approx -\frac{\dot{k}}{2}\left(\frac{1}{k(t)\beta(t)}+\frac{\zeta^2+mk(t)}{2\zeta k^3(t)\beta(t)}\dot{k}+\frac{\zeta^2+2mk(t)}{2\zeta k^2(t)\beta^2(t)}\dot{\beta}\right) \\
&= -\frac{\dot{k}}{2k(t)\beta(t)}-\frac{\zeta^2+mk(t)}{4\zeta k^3(t)\beta(t)}\dot{k}^2-\frac{\zeta^2+2mk(t)}{4\zeta k^2(t)\beta^2(t)}\dot{k}\dot{\beta}
\end{aligned}
\tag{2.30}
$$

Remark 2.2.1. *If $\dot{\beta}=0$, (i.e. the temperature is a constant), the extracted mechanical power with non-equilibrium correction term is less than the estimation in the near-equilibrium regime.*

In the case of the resistor-capacitor system,

$$
dv = -\left(\frac{\dot{c}+g}{c}\right)vdt+\frac{1}{c}\sqrt{\frac{2g}{\beta}}dB(t)
\tag{2.31}
$$

Considering the quadratic term v^2,

$$
\begin{aligned}
dv^2 &= 2vdv+\frac{1}{2}\frac{\partial^2 v^2}{\partial v^2}\frac{1}{c^2}\frac{2g}{\beta}dt \\
&= -2\left(\frac{\dot{c}+g}{c}\right)v^2dt+\frac{2v}{c}\sqrt{\frac{2g}{\beta}}dB(t)+\frac{2g}{c^2\beta}dt
\end{aligned}
$$

The second order moment $\sigma=\langle v^2\rangle$ follows the dynamics

$$
\frac{d}{dt}\sigma = -\frac{2\dot{c}}{c}\sigma-\frac{2g}{c}\sigma+\frac{2g}{c^2\beta}
\tag{2.32}
$$

When $\dot{c}\equiv 0$, the equilibrium value of $\sigma_{eq,\lambda(t)}$ is $\frac{1}{c(t)\beta(t)}$. As a near-equilibrium approximation, $\sigma(t+\epsilon)=\sigma_{eq,\lambda(t)}+\Delta\sigma(t,\epsilon)$, $c(t+\epsilon)=c(t)+\epsilon\dot{c}$, $\beta(t+\epsilon)=\beta(t)+\epsilon\dot{\beta}$

and

$$\dot{\sigma}_{eq,\lambda(t)}(t) + \frac{d}{dt}\Delta\sigma(t,\epsilon) = -\frac{2(\dot{c}+g)}{c(t)+\epsilon\dot{c}}(\sigma(t)_{eq,\lambda(t)} + \Delta\sigma(t,\epsilon)) + \frac{2g}{(c(t)+\epsilon\dot{c})^2(\beta(t)+\epsilon\dot{\beta})} \tag{2.33}$$

As $0 < \epsilon \ll 1$

$$\frac{1}{c(t)+\epsilon\dot{c}} \approx \frac{1}{c(t)} - \epsilon\frac{\dot{c}}{c^2(t)} + O(\epsilon^2) \tag{2.34}$$
$$\frac{1}{\beta(t)+\epsilon\dot{\beta}} \approx \frac{1}{\beta(t)} - \epsilon\frac{\dot{\beta}}{\beta^2(t)} + O(\epsilon^2)$$

Given that $\epsilon \to 0$ and $\sigma_{eq,\lambda(t)} = \frac{1}{c(t)\beta(t)}$,

$$\left(\frac{d}{dt} + \frac{2(\dot{c}+g)}{c(t)}\right)\Delta\sigma(t,0) = -\frac{\dot{c}}{c^2(t)\beta(t)} + \frac{\dot{\beta}}{c(t)\beta^2(t)} \tag{2.35}$$

With the assumption that velocities $\dot{c}$ and $\dot{\beta}$ are very small, $\frac{2(\dot{c}+g)}{c(t)} \neq 0$ and

$$\Delta\sigma(t,0) = \left(1 + \frac{c(t)}{2(\dot{c}+g)}\frac{d}{dt}\right)^{-1}\frac{c\dot{\beta} - \dot{c}\beta}{2(\dot{c}+g)c\beta^2}(t) \tag{2.36}$$

Thus, up to the 0^{th} order of the differential order $\frac{d}{dt}$, $\Delta\sigma(t,0) \approx \frac{c(t)\dot{\beta} - \dot{c}\beta(t)}{2(\dot{c}+g)c(t)\beta^2(t)}$. The extracted mechanical power by the capacitor with the correction term is

$$\frac{1}{2}\dot{c}\left\langle v^2\right\rangle \approx \frac{1}{2}\dot{c}\left(\sigma_{eq,\lambda(t)} + \Delta\sigma(t,0)\right) = \frac{\dot{c}}{2c(t)\beta(t)} + \frac{c(t)\dot{c}\dot{\beta} - \beta(t)\dot{c}^2}{4gc(t)\beta^2(t)} \tag{2.37}$$

Remark 2.2.2. *Similar to the stochastic oscillator case, when β is a constant, the extracted mechanical power with the correction term will be less than the extracted mechanical power which is only evaluated with equilibrium average.*

2.3 Sub-Riemannian geometry structures of the heat engines

In the previous chapter, it was shown that with proper assumptions, the inverse diffusion tensor induces the structure of a Riemannian manifold in the space of the system parameter of a heat engine. A path in the space of the system parameter λ, is denoted as the protocol Λ. The geometric energy of the path carries the meaning of energy dissipation of that protocol.

Besides the dissipation of energy in the analysis of non-equilibrium thermodynamic system, there are auxiliary quantities ψ of interest, such as the average extracted work of a heat engine, which are functionals of a protocol and system parameter's velocity along the protocol. The system parameter space M is enlarged to a higher dimensional manifold N of pair (λ, ψ). Let π denote the projection from N to M. Assuming π is a surjective submersion, π_* is the corresponding push-forward action $\pi_* : T_{(\lambda,\psi)}N \to T_\lambda M$. π_* and metric tensor g on M induce a positive semidefinite inner product $\langle\, ,\, \rangle_N$ on the manifold N. $\forall v_1,\, v_2 \in T_{(\lambda,\psi)}N$, its inner product is

$$\langle v_1, v_2 \rangle_N = g(\pi_* v_1, \pi_* v_2) \tag{2.38}$$

As the number of control variables is smaller than the dimension of $T_{(\lambda,\psi)}N$, we ask if a sub-Riemannian manifold structure might exist in the space of (λ, ψ). This requires controllability. Once we write down the dynamics explicitly, as a control system, controllability depends on the Chow-Rashevski theorem [14]. We will analyze both the stochastic oscillator and the resistor-capacitor circuit.

At time t, the mechanical power $\dot{\psi}$ is extracted from the oscillator by varying the stiffness of potential well during the transition between equilibrium states,

$$\dot{\psi}_1 = \left\langle -\frac{\partial\left(\frac{k\xi_1^2}{2}\right)}{\partial k}\dot{k}\right\rangle_{eq,\lambda} = \frac{\zeta^2}{m}\left(\frac{y^2}{x^2}\dot{x} - \frac{y}{x}\dot{y}\right) \tag{2.39}$$

In the case of the resistor-capacitor circuit, its extracted mechanical power is

$$\dot{\psi}_2 = \left\langle \frac{\dot{c}v^2}{2} \right\rangle_{eq,\lambda} = \frac{\dot{c}}{2c\beta} \tag{2.40}$$

In linear response regime, the average dissipation rate $\tilde{d}$ of each heat engine is positive while the system parameter of the engine is varied. The average dissipation rate with $k_B T/\text{sec}$ unit of the stochastic oscillator is

$$\tilde{d}_1 = \frac{\zeta(\dot{x}^2 + \dot{y}^2)}{x} \tag{2.41}$$

which defines the metric tensor on the system parameter manifold of (x, y). $(\sqrt{\frac{x}{\zeta}}, 0)^T$ and $(0, \sqrt{\frac{x}{\zeta}})^T$ are orthonormal vectors on the tangent space to the control parameter manifold. The dynamics of the heat engine can be formulated as a control system,

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\psi}_1 \end{pmatrix} = u \begin{pmatrix} \sqrt{\frac{x}{\zeta}} \\ 0 \\ \frac{\zeta^{3/2}}{m}\frac{y^2}{x^{3/2}} \end{pmatrix} + v \begin{pmatrix} 0 \\ \sqrt{\frac{x}{\zeta}} \\ -\frac{\zeta^{3/2}}{m}\frac{y}{\sqrt{x}} \end{pmatrix} \tag{2.42}$$

where u and v now take the role of control signals.

Taking m and ζ as units in the controllability analysis, for the vector fields $f_1 = (\sqrt{x}, 0, \frac{y^2}{x^{3/2}})^T$ and $f_2 = (0, \sqrt{x}, -\frac{y}{\sqrt{x}})^T$, the Lie bracket $[\,,\,]$ of f_1 and f_2 gives $f_3 = [f_1, f_2] = (0, \frac{1}{2}, -\frac{3y}{2x})^T$. The rank $(f_1, f_2, f_3) = 3$ indicates that the system (3.1) is completely non-holonomic (controllable). From above definitions, it follows that $\tilde{d} = u^2 + v^2$ and hence, the problem finding a protocol minimizing dissipation is an optimal control problem for (3.1) with quadratic cost functional.

In the case of the resistor-capacitor circuit, the average dissipation rate in

k_BT/sec unit is

$$\tilde{d}_2 = \begin{pmatrix} \dot{c} & \dot{\beta} \end{pmatrix} \begin{pmatrix} \frac{1}{4cg\beta} & \frac{1}{4g\beta^2} \\ \frac{1}{4g\beta^2} & \frac{c}{2g\beta^3} \end{pmatrix} \begin{pmatrix} \dot{c} \\ \dot{\beta} \end{pmatrix} \tag{2.43}$$

which defines the metric tensor of the manifold of (c, β). The control system of the circuit is

$$\begin{pmatrix} \dot{c} \\ \dot{\beta} \\ \dot{\psi}_2 \end{pmatrix} = u \begin{pmatrix} 2\sqrt{gc\beta} \\ 0 \\ \sqrt{\frac{g}{c\beta}} \end{pmatrix} + v \begin{pmatrix} -2\sqrt{gc\beta} \\ 2\beta^{3/2}\sqrt{\frac{g}{c}} \\ -\sqrt{\frac{g}{c\beta}} \end{pmatrix} \tag{2.44}$$

Taking g as the unit, there are two vector fields $f_1 = (2\sqrt{c\beta},\, 0,\, \frac{1}{\sqrt{c\beta}})^\intercal$ and $f_2 = (-2\sqrt{c\beta},\, 2\beta^{3/2}\frac{1}{\sqrt{c}},\, -\frac{1}{\sqrt{c\beta}})^\intercal$. The Lie bracket of f_1 and f_2 is $f_3 = (-2\beta,\, -\frac{2\beta^2}{c},\, \frac{1}{c})^\intercal$ and the rank$(f_1, f_2, f_3) = 3$ indicates that the control system of the resistor-capacitor circuit is completely non-holonomic (controllable). Again, $\tilde{d}_2 = u^2 + v^2$ and the problem of finding a protocol minimizing dissipation is an optimal control problem with quadratic cost functional.

Chapter 3: Design of a Heat Engine: Optimal Control Approach

3.1 Optimal control of the heat engines

As shown in Chapter 1, a parametrically-controlled heat engine can convert heat flow into mechanical work. The engine is modeled by focusing on the control of the system parameter. In the case of the stochastic oscillator,

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\psi}_1 \end{pmatrix} = u_1 \begin{pmatrix} \sqrt{\frac{x}{\zeta}} \\ 0 \\ \frac{\zeta^{3/2}}{m}\frac{y^2}{x^{3/2}} \end{pmatrix} + v_1 \begin{pmatrix} 0 \\ \sqrt{\frac{x}{\zeta}} \\ -\frac{\zeta^{3/2}}{m}\frac{y}{\sqrt{x}} \end{pmatrix} \tag{3.1}$$

where (x, y) is the system parameter and $\psi_1(t)$ is the mechanical work extracted over the interval $[0, t]$. Correspondingly, extracted power $\dot{\psi}_1 = u_1\frac{\zeta^{3/2}}{m}\frac{y^2}{x^{3/2}} - v_1\frac{\zeta^{3/2}}{m}\frac{y}{\sqrt{x}} = \frac{\zeta^2}{m}\left(\frac{y^2}{x^2}\dot{x} - \frac{y}{x}\dot{y}\right)$.

In the case of the resistor-capacitor circuit,

$$\begin{pmatrix} \dot{c} \\ \dot{\beta} \\ \dot{\psi}_2 \end{pmatrix} = u_2 \begin{pmatrix} 2\sqrt{gc\beta} \\ 0 \\ \sqrt{\frac{g}{c\beta}} \end{pmatrix} + v_2 \begin{pmatrix} -2\sqrt{gc\beta} \\ 2\beta^{3/2}\sqrt{\frac{g}{c}} \\ -\sqrt{\frac{g}{c\beta}} \end{pmatrix} \tag{3.2}$$

where $\psi_2(t)$ is the mechanical work extracted over the interval $[0, t]$. The extracted mechanical power is $\dot{\psi}_2 = u_2\sqrt{\frac{g}{c\beta}} + v_2\sqrt{\frac{g}{c\beta}} = \frac{\dot{c}}{2c\beta}$.

The total heat supply from the heat bath to the engine is consumed in two ways. Some of the heat flows to the engine as useful supply (some of it is converted to extracted work and some of it is converted to the internal energy of the system) and the remaining part is dissipated to the engine's surroundings. Based on the discussion in Chapter 1, $\tilde{d}_i = u_i^2 + v_i^2$ ($i = 1$ for the stochastic oscillator and $i = 2$ for the resistor-capacitor circuit) is the average dissipation rate in $k_B T/\text{sec}$. Given a protocol Λ (i.e. a trajectory on the space of system parameter), from time 0 to time t_f, along a trajectory in the space of (x, y), the total heat supply in the stochastic oscillator system is

$$Q_1 = \int_0^{t_f} \left(\left(-\frac{\dot{\beta}}{\beta^2} - \frac{\dot{k}}{2\beta k} \right) \mathbb{1}\{-\frac{\dot{\beta}}{\beta^2} - \frac{\dot{k}}{2\beta k} > 0\} + \tilde{d}_1 \right) dt \tag{3.3}$$

where $\mathbb{1}$ is an indicator function. Expressing Q_1 in the coordinate of (x, y), it is

$$Q_1 = \int_0^{t_f} \left(\left(\frac{\zeta^2}{m} \frac{y}{x} \dot{y} \right) \mathbb{1}\{\frac{y}{x}\dot{y} > 0\} + \zeta \left(\frac{\dot{x}^2 + \dot{y}^2}{x} \right) \right) dt \tag{3.4}$$

In the resistor-capacitor system, the total heat supply

$$Q_2 = \int_0^{t_f} \left(\left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} \right) \mathbb{1}\left\{ -\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} > 0 \right\} + \tilde{d}_2 \right) dt =$$
$$\int_0^{t_f} \left(\left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} \right) \mathbb{1}\left\{ -\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} > 0 \right\} + \begin{pmatrix} \dot{c} & \dot{\beta} \end{pmatrix} \begin{pmatrix} \frac{1}{4gc\beta} & \frac{1}{4g\beta^2} \\ \frac{1}{4g\beta^2} & \frac{c}{2g\beta^3} \end{pmatrix} \begin{pmatrix} \dot{c} \\ \dot{\beta} \end{pmatrix} \right) dt \tag{3.5}$$

Efficiency η of a heat engine is the ratio of the extracted work of the heat engine to the total heat supply from the heat bath. The efficiency of the stochastic oscillator is,

$$\begin{aligned}\eta_1 &= \frac{\int_0^{t_f} \dot{\psi}_1 dt}{\int_0^{t_f} \left(\left(\frac{\zeta^2}{m}\frac{y}{x}\dot{y}\right) \mathbb{1}\{\frac{y}{x}\dot{y} > 0\} + \tilde{d}_1\right) dt} \\ &= \frac{\int_0^{t_f} \frac{\zeta^2}{m}\left(\frac{y^2}{x^2}\dot{x} - \frac{y}{x}\dot{y}\right) dt}{\int_0^{t_f} \left(\left(\frac{\zeta^2}{m}\frac{y}{x}\dot{y}\right) \mathbb{1}\{\frac{y}{x}\dot{y} > 0\} + \tilde{d}_1\right) dt}\end{aligned} \tag{3.6}$$

The efficiency of the resistor-capacitor circuit is

$$\begin{aligned}\eta_2 &= \frac{\int_0^{t_f} \dot{\psi}_2 dt}{\int_0^{t_f} \left(\left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}\right) \mathbb{1}\left\{-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} > 0\right\} + \tilde{d}_2\right) dt} \\ &= \frac{\int_0^{t_f} \frac{\dot{c}}{2c\beta} dt}{\int_0^{t_f} \left(\left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}\right) \mathbb{1}\left\{-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} > 0\right\} + \tilde{d}_2\right) dt}\end{aligned} \tag{3.7}$$

If the extracted work is prescribed, to maximize the efficiency of a heat engine, the total heat supply (i.e. the denominator in η_1 or η_2) must be minimized. This can be first formulated as a problem in the calculus of variations on the sub-Riemannian manifold N_i (where N_1 is the manifold for the stochastic oscillator and N_2 is the manifold for the resistor-capacitor circuit):

Find $t \to \gamma_1(t) = (x(t), y(t), \psi_1(t))$, a curve in N_1 such that,

$$J_1 = \int_0^{t_f} \left(\left(\frac{\zeta^2}{m}\frac{y}{x}\dot{y}\right) \mathbb{1}\left\{\frac{y}{x}\dot{y} > 0\right\} + \tilde{d}_1\right) dt \tag{3.8}$$

is **minimized** subject to

$$\int_0^{t_f} \dot{\psi}_1 dt = A \text{ and end-points are specified.}$$

Or

Find $t \to \gamma_2(t) = (c(t), \beta(t), \psi_2(t))$, a curve in N_2 such that,

$$J_2 = \int_0^{t_f} \left(\left(-\frac{\dot{\beta}}{\beta^2} + \frac{\dot{c}}{2c\beta} \right) \mathbb{1} \left\{ -\frac{\dot{\beta}}{\beta^2} + \frac{\dot{c}}{2c\beta} > 0 \right\} + \tilde{d}_2 \right) dt \tag{3.9}$$

is **minimized** subject to

$$\int_0^{t_f} \dot{\psi}_2 dt = A \text{ and end-points are specified.}$$

As a heat engine is a cyclically operated machine, *working loops* are a type of protocols of special interest. In the stochastic oscillator, the necessary condition for the working loops is: $x(0) = x(t_f)$ and $y(0) = y(t_f)$ and in the resistor-capacitor circuit, the necessary condition for the working loops is: $c(0) = c(t_f)$ and $\beta(0) = \beta(t_f)$.

Theorem 3.1.1. *If a non-contant working loop* $\gamma_1^* \in C^\infty$ *is an optimal solution to problem* (3.8) *and along* $\gamma_1^* : [0, t_f] \to N_1$*, there are finite number of* $t_i \in [0, t_f](i = 1, 2, \ldots, m)$ *when* $\frac{y}{x}\dot{y}$ *is zero,* γ_1^* *is also a solution to the following optimal control problem with dynamics* (3.1).

$$\begin{aligned} \underset{(u_1,v_1)\in L^2_{[0,t_f]}\times L^2_{[0,t_f]}}{Min} \quad & J = \int_0^{t_f} \left(\frac{u_1^2 + v_1^2}{2} \right) dt \\ \textit{subject to} \quad & \int_0^{t_f} \dot{\psi}_1 dt = A, \textit{ and end-points are specified.} \end{aligned} \tag{3.10}$$

Proof. If there are m points along $\gamma_1^* : [0, t_f] \to N_1$, such that $\frac{y}{x}\dot{y}(t_i) = 0$ $(i = 1, 2, \ldots, m)$, in general (assuming that $\frac{y}{x}\dot{y}(0)$ and $\frac{y}{x}\dot{y}(t_f)$ are non-zero), we can decompose γ_1^* to $(m+1)$ pieces, such as $\gamma_{1,0}^* : [t_0, t_1] \to N_1, \ldots, \gamma_{1,m}^* : [t_m, t_{m+1}] \to N_1$, where we set $t_0 = 0$ and $t_{m+1} = t_f$.

Moreover, as from time 0 to time t_f, we have a loop in the parameter space (x, y),

$$\int_0^{t_f} \dot{\psi}_1 dt = \int_0^{t_f} \left(\frac{d}{dt} \frac{1}{\beta(x,y)} + \dot{\psi}_1 \right) dt = \int_0^{t_f} \frac{\zeta^2}{m} \frac{y}{x} \dot{y} dt \tag{3.11}$$

which indicates that the extracted work over $[0, t_f]$ being a prescribed value is equivalent to the integral $\frac{\zeta^2}{m}\frac{y}{x}\dot{y}$ being the same prescribed value. As γ_1^* is the optimal solution to problem (3.8) from time 0 to time t_f, each $\gamma_{1,i}^* : [t_i, t_{i+1}] \to N_1$ $(i = 0, 1, \ldots, m)$ should be the optimal solution to following problem.

$$\begin{aligned} \underset{\gamma_{1,i}}{\text{Min}} \quad & J = \int_{t_i}^{t_{i+1}} \left(\frac{\zeta^2}{m} \frac{y}{x} \dot{y} \mathbb{1} \left\{ \frac{y}{x} \dot{y} > 0 \right\} + \tilde{d}_1 \right) dt \\ \text{subject to:} \quad & \gamma_{1,i}(t_i) = \gamma_1^*(t_i) \text{ and } \gamma_{1,i}(t_{i+1}) = \gamma_1^*(t_{i+1}) \end{aligned} \tag{3.12}$$

If during $[t_i, t_{i+1}]$, $\frac{y}{x}\dot{y}$ is non-negative. The augmented Lagrangian along γ_i^* is

$$L_i = (1 + \rho_i) \frac{\zeta^2}{m} \frac{y}{x} \dot{y} + \zeta \frac{\dot{x}^2 + \dot{y}^2}{x} \tag{3.13}$$

where ρ_i is the Lagrange multiplier and it is a constant. As an optimal solution, $\gamma_{1,i}^*$ must follow Euler-Lagrange equation, such that

$$\begin{aligned} (1 + \rho_i) \frac{\zeta}{m} \frac{y\dot{y}}{x^2} - \frac{\dot{x}^2}{x^2} + \frac{\dot{y}^2}{x^2} + \frac{2\ddot{x}}{x} &= 0 \\ (1 + \rho_i) \frac{\zeta}{m} \frac{y\dot{x}}{x} - \frac{2\ddot{y}}{x} + \frac{2\dot{x}\dot{y}}{x^2} &= 0 \end{aligned} \tag{3.14}$$

In time interval $[t_{i+1}, t_{i+2}]$, $\frac{y}{x}\dot{y}$ is either non-negative or non-positive. If $\frac{y}{x}\dot{y}$ is non-negative, it is seen with ρ_{i+1} as the Lagrange multiplier and L_{i+1} as the augmented Lagrangian, $\gamma^*_{1,i+1}$ should be a solution of following Euler-Lagrange equation.

$$\begin{aligned}(1+\rho_{i+1})\frac{\zeta}{m}\frac{y\dot{y}}{x^2}-\frac{\dot{x}^2}{x^2}+\frac{\dot{y}^2}{x^2}+\frac{2\ddot{x}}{x}&=0\\(1+\rho_{i+1})\frac{\zeta}{m}\frac{y\dot{x}}{x}-\frac{2\ddot{y}}{x}+\frac{2\dot{x}\dot{y}}{x^2}&=0\end{aligned}\tag{3.15}$$

Comparing equations in (3.14) and (3.15) at time t_{i+1}, because of γ^*_1's smoothness,

$$\begin{aligned}(\rho_i-\rho_{i+1})\frac{y\dot{y}}{x^2}&=0\\(\rho_i-\rho_{i+1})\frac{y\dot{x}}{x}&=0\end{aligned}\tag{3.16}$$

either $\rho_i=\rho_{i+1}$ or $\dot{x}=0$ and $\dot{y}=0$. It is seen that L_i is time-independent. Along $\gamma^*_{1,i}$, $E_i\equiv\frac{\partial L_i}{\partial\dot{x}}\dot{x}+\frac{\partial L_i}{\partial\dot{y}}\dot{y}-L_i=\zeta\frac{\dot{x}^2+\dot{y}^2}{x}$ is a conserved quantity. Similarly, along $\gamma^*_{1,i+1}$, $E_{i+1}=\zeta\frac{\dot{x}^2+\dot{y}^2}{x}$ is also conserved. If $\dot{x}=0$ and $\dot{y}=0$ at time t_{i+1}, E_i and E_{i+1} are both zero along $\gamma^*_{1,i}$ and $\gamma^*_{1,i+1}$. Thus, in an interval $[t_i,t_{i+2}]$ we have a static point. If $t_i=0$ and $t_{i+2}=t_f$, the whole trajectory γ^*_1 is constant, which does not satisfy γ^*_1's non-constancy assumption. The only plausible solution to (3.25) is $\rho_i=\rho_{i+1}$.

If $\frac{y}{x}\dot{y}$ is non-positive from time t_{i+1} to time t_{i+2}, the augmented Lagrangian is $\rho_{i+1}\frac{\zeta^2}{m}\frac{y}{x}\dot{y}+\zeta\frac{\dot{x}^2+\dot{y}^2}{x}$ and its Euler-Lagrange equation is

$$\begin{aligned}\rho_{i+1}\frac{\zeta}{m}\frac{y\dot{y}}{x^2}-\frac{\dot{x}^2}{x^2}+\frac{\dot{y}^2}{x^2}+\frac{2\ddot{x}}{x}&=0\\\rho_{i+1}\frac{\zeta}{m}\frac{y\dot{x}}{x}-\frac{2\ddot{y}}{x}+\frac{2\dot{x}\dot{y}}{x^2}&=0\end{aligned}\tag{3.17}$$

Again, following a discussion similar to the one for solving equations in (3.25), $1+\rho_i=\rho_{i+1}$. Thus, along each piece of γ^*_1, the augmented Lagrangian $L=\rho\frac{\zeta^2}{m}\frac{y}{x}\dot{y}+\zeta\frac{\dot{x}^2+\dot{y}^2}{x}$ and ρ is a constant. It indicates that γ^*_1 is an optimal solution of following calculus of variations problem.

$$\underset{\gamma_1}{\text{Min}} \quad J = \int_0^{t_f} \zeta \frac{\dot{x}^2 + \dot{y}^2}{x} dt$$
$$\text{subject to:} \quad \int_0^{t_f} \frac{\zeta^2}{m} \left(\frac{y^2}{x^2} \dot{x} - \frac{y}{x} \dot{y} \right) dt \text{ and end-points are prescribed.} \tag{3.18}$$

Based on (3.1), $\zeta \frac{\dot{x}^2+\dot{y}^2}{x} = u_1^2 + v_1^2$. Apparently, problem (3.8) is equivalent to problem (3.10). □

In the case of the resistor-capacitor circuit, we have a similar result:

Theorem 3.1.2. *If a non-contant working loop $\gamma_2^* \in C^\infty$ is an optimal solution to problem* (3.9) *and along $\gamma_2^* : [0, t_f] \to N_2$, there are finite number of $t_i \in [0, t_f] (i = 1, 2, \ldots, m)$ when $-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}$ is zero, γ_2^* is also a solution to the following geometric optimal control problem with dynamics* (3.2).

$$\underset{(u_2,v_2) \in L^2_{[0,t_f]} \times L^2_{[0,t_f]}}{\mathit{Min}} \quad J = \int_0^{t_f} \left(\frac{u_2^2 + v_2^2}{2} \right) dt$$
$$\textit{subject to} \quad \int_0^{t_f} \dot{\psi}_2 dt = A, \ \textit{and end-points are specified.} \tag{3.19}$$

Proof. If there are m points along $\gamma_2^* : [0, t_f] \to N_2$, such that $-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}(t_i) = 0$ $(i = 1, 2, \ldots, m)$, in general (assuming that $-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}(0)$ and $-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}(t_f)$ are non-zero), we can decompose γ_2^* to $(m+1)$ pieces, such as $\gamma_{2,0}^* : [t_0, t_1] \to N_2$, $\ldots$, $\gamma_{2,m}^* : [t_m, t_{m+1}] \to N_2$, where we set $t_0 = 0$ and $t_{m+1} = t_f$.

Moreover, as from time 0 to time t_f, we have a loop in the parameter space (c, β),

$$\int_0^{t_f} \dot{\psi}_2 dt = \int_0^{t_f} \left(\frac{d}{dt} \frac{1}{2\beta} + \dot{\psi}_2 \right) dt = \int_0^{t_f} \left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} \right) dt \tag{3.20}$$

which indicates that the extracted work over $[0, t_f]$ being a prescribed value is equivalent to the integral $-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}$ being the same prescribed value. As γ_2^* is the optimal solution to problem (3.9) from time 0 to time t_f, each $\gamma_{2,i}^* : [t_i, t_{i+1}] \to N_2$ $(i = 0, 1, \ldots, m)$ should be the optimal solution to following problem.

$$\begin{aligned} \underset{\gamma_{i,2}}{\text{Min}} \quad J &= \int_{t_i}^{t_{i+1}} \left(\left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} \right) \mathbb{1} \left\{ -\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} > 0 \right\} + \tilde{d}_2 \right) dt \\ \text{subject to:} \quad & \gamma_{i,2}(t_i) = \gamma_2^*(t_i) \text{ and } \gamma_{2,i+1}(t_{i+1}) = \gamma_2^*(t_{i+1}) \end{aligned} \tag{3.21}$$

If during $[t_i, t_{i+i}]$, $-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}$ is non-negative. The augmented Lagrangian along γ_i^* is

$$L_i = (1 + \rho_i) \left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} \right) + \frac{\dot{c}^2}{4cg\beta} + \frac{\dot{c}\dot{\beta}}{2g\beta^2} + \frac{c\dot{\beta}^2}{2g\beta^3} \tag{3.22}$$

where ρ_i is the Lagrange multiplier and it is a constant. As an optimal solution, $\gamma_{2,i}^*$ must follow Euler-Lagrange equation, such that

$$\begin{aligned} -(1 + \rho_i) \frac{\dot{\beta}}{2c\beta^2} + \frac{\ddot{c}}{2cg\beta} + \frac{\ddot{\beta}}{2g\beta^2} - \frac{3\dot{\beta}^2}{2g\beta^3} - \frac{\dot{c}^2}{4c^2 g\beta} - \frac{\dot{c}\dot{\beta}}{2cg\beta^2} &= 0 \\ (1 + \rho_i) \frac{\dot{c}}{2c\beta^2} + \frac{\ddot{c}}{2g\beta^2} + \frac{2c\ddot{\beta}}{g\beta^3} + \frac{\dot{c}\dot{\beta}}{g\beta^3} + \frac{\dot{c}^2}{4cg\beta^2} - \frac{3c\dot{\beta}^2}{2g\beta^4} &= 0 \end{aligned} \tag{3.23}$$

In time interval $[t_{i+1}, t_{i+2}]$, $-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}$ is either non-negative or non-positive. If $-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}$ is non-negative, it is seen with ρ_{i+1} as the Lagrange multiplier and L_{i+1} as the augmented Lagrangian, $\gamma_{2,i+1}^*$ should be a solution of following Euler-Lagrange equation.

$$\begin{aligned} -(1 + \rho_{i+1}) \frac{\dot{\beta}}{2c\beta^2} + \frac{\ddot{c}}{2cg\beta} + \frac{\ddot{\beta}}{2g\beta^2} - \frac{3\dot{\beta}^2}{2g\beta^3} - \frac{\dot{c}^2}{4c^2 g\beta} - \frac{\dot{c}\dot{\beta}}{2cg\beta^2} &= 0 \\ (1 + \rho_{i+1}) \frac{\dot{c}}{2c\beta^2} + \frac{\ddot{c}}{2g\beta^2} + \frac{2c\ddot{\beta}}{g\beta^3} + \frac{\dot{c}\dot{\beta}}{g\beta^3} + \frac{\dot{c}^2}{4cg\beta^2} - \frac{3c\dot{\beta}^2}{2g\beta^4} &= 0 \end{aligned} \tag{3.24}$$

Comparing equations in (3.23) and (3.24) at time t_{i+1}, because of γ_2^*'s smoothness,

$$
\begin{aligned}
(\rho_i - \rho_{i+1})\frac{\dot{\beta}}{2c\beta^2} &= 0 \\
(\rho_i - \rho_{i+1})\frac{\dot{c}}{2c\beta^2} &= 0
\end{aligned}
\tag{3.25}
$$

either $\rho_i = \rho_{i+1}$ or $\dot{c} = 0$ and $\dot{\beta} = 0$. It is seen that L_i is time-independent. Along $\gamma_{2,i}^*$, $E_i \equiv \frac{\partial L_i}{\partial \dot{c}}\dot{c} + \frac{\partial L_i}{\partial \dot{\beta}}\dot{\beta} - L_i = \frac{\dot{c}^2}{4cg\beta} + \frac{\dot{c}\dot{\beta}}{2g\beta^2} + \frac{c\dot{\beta}^2}{2g\beta^3}$ is a conserved quantity. Similarly, along $\gamma_{2,i+1}^*$, $E_{i+1} = \frac{\dot{c}^2}{4cg\beta} + \frac{\dot{c}\dot{\beta}}{2g\beta^2} + \frac{c\dot{\beta}^2}{2g\beta^3}$ is also conserved. If $\dot{c} = 0$ and $\dot{\beta} = 0$ at time t_{i+1}, E_i and E_{i+1} are both zero along $\gamma_{2,i}^*$ and $\gamma_{2,i+1}^*$. Thus, in an interval $[t_i, t_{i+2}]$ we have a static point. If $t_i = 0$ and $t_{i+2} = t_f$, the whole trajectory γ_2^* is constant, which does not satisfy γ_2^*'s non-constancy assumption. The only plausible solution to (3.25) is $\rho_i = \rho_{i+1}$.

If $-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}$ is non-positive from time t_{i+1} to time t_{i+2}, the augmented Lagrangian is $\rho_{i+1}\left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}\right) + \frac{\dot{c}^2}{4cg\beta} + \frac{\dot{c}\dot{\beta}}{2g\beta^2} + \frac{c\dot{\beta}^2}{2g\beta^3}$ and its Euler-Lagrange equation is

$$
\begin{aligned}
-\rho_{i+1}\frac{\dot{\beta}}{2c\beta^2} + \frac{\ddot{c}}{2cg\beta} + \frac{\ddot{\beta}}{2g\beta^2} - \frac{3\dot{\beta}^2}{2g\beta^3} - \frac{\dot{c}^2}{4c^2g\beta} - \frac{\dot{c}\dot{\beta}}{2cg\beta^2} &= 0 \\
\rho_{i+1}\frac{\dot{c}}{2c\beta^2} + \frac{\ddot{c}}{2g\beta^2} + \frac{2c\ddot{\beta}}{g\beta^3} + \frac{\dot{c}\dot{\beta}}{g\beta^3} + \frac{\dot{c}^2}{4cg\beta^2} - \frac{3c\dot{\beta}^2}{2g\beta^4} &= 0
\end{aligned}
\tag{3.26}
$$

Again, following a discussion similar to the one for solving equations in (3.25), $1 + \rho_i = \rho_{i+1}$. Thus, along each piece of γ_2^*, the augmented Lagrangian $L = \rho\left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta}\right) + \frac{\dot{c}^2}{4cg\beta} + \frac{\dot{c}\dot{\beta}}{2g\beta^2} + \frac{c\dot{\beta}^2}{2g\beta^3}$ and ρ is a constant. It indicates that γ_2^* is an optimal solution of following calculus of variations problem.

$$
\begin{aligned}
\underset{\gamma_2}{\text{Min}} \quad J &= \int_0^{t_f} \frac{\dot{c}^2}{4cg\beta} + \frac{\dot{c}\dot{\beta}}{2g\beta^2} + \frac{c\dot{\beta}^2}{2g\beta^3} dt \\
\text{subject to:} \quad & \int_0^{t_f} \frac{\dot{c}}{2c\beta} dt \text{ and end-points are prescribed.}
\end{aligned}
\tag{3.27}
$$

Based on (3.2), $\frac{\dot{c}^2}{4cg\beta} + \frac{\dot{c}\dot{\beta}}{2g\beta^2} + \frac{c\dot{\beta}^2}{2g\beta^3} = u_2^2 + v_2^2$. Apparently, problem (3.9) is equivalent to problem (3.19). $\square$

Remark 3.1.3. *Sivak and Crooks [11] suggested that obtaining efficient performance of molecular machines requires minimization of dissipation. Theorems 3.1.1 and 3.1.2 give a mathematical basis for why this must be so.*

3.2 Maximum efficiency protocols of the heat engines

On the sub-Riemannian manifold N_i ($i = 1$ for the stochastic oscillator and $i = 2$ for the resistor-capacitor circuit) of a heat engine, for a trajectory $\gamma_i : t \in [0, t_f] \to N_i$, $\dot{\gamma}_i \in T_{\gamma_i(t)}N_i$ is its tangent vector and $p_i \in T^*_{\gamma_i(t)}N_i$ is a cotangent vector, such that $p_i : \dot{\gamma}_i \in T_{\gamma_i(t)}N_i \to \mathbb{R}$ is the duality pairing $\langle p_i, \dot{\gamma}_i \rangle_d$. In canonical coordinates, $p_i = (p_{i,1}, p_{i,2}, p_{i,3})^\intercal$ and we let $\langle p_i, \dot{\gamma}_i \rangle_d = p_i^\intercal \dot{\gamma}_i$. The pair $\tilde{\lambda}_i = (p_i, \gamma_i) \in T^*N_i$, where $T^*N_i = \{T_q^*N_i | \ \forall q \in N_i\}$ is the cotangent bundle of N_i. In this Chapter, we would like to consider optimal solutions to problem (3.10) and problem (3.19) with non-zero controls. In other words, these optimal solutions are regular extremals. As shown in Chapter 2, in the case of the stochastic oscillator $f_{1,1} = (\sqrt{\frac{x}{\zeta}}, 0, \frac{\zeta^{3/2}}{m}\frac{y^2}{x^{3/2}})^\intercal$ and $f_{1,2} = (0, \sqrt{\frac{x}{\zeta}}, -\frac{\zeta^{3/2}}{m}\frac{y}{\sqrt{x}})^\intercal$ denote vector fields of the control system (3.1) and in the case of the resistor-capacitor circuit $f_{2,1} = (2\sqrt{gc\beta}, 0, \sqrt{\frac{g}{c\beta}})$ and $f_{2,2} = (-2\sqrt{gc\beta}, 2\beta^{3/2}\sqrt{\frac{g}{c}}, -\sqrt{\frac{g}{c\beta}})$. In either case, following maximum principle [25] [26], at time $t \in [0, t_f]$, the sub-Riemannian Hamiltonian is the smooth function of T^*N_i defined as follows

$$H_i : T^*N_i \to \mathbb{R}, \quad H_i(\tilde{\lambda}_i) = \max_{u_i, v_i} \left(\langle p_i, u_i f_{i,1} + v_i f_{i,2} \rangle_d - \frac{u_i^2 + v_i^2}{2} \right). \tag{3.28}$$

So, in the case of the stochastic oscillator, the optimal control in canonical coordinates is

$$
\begin{aligned}
u_1^* &= \langle p_1, f_{1,1}\rangle_d = \sqrt{\frac{x}{\zeta}} p_{1,1} + \frac{\zeta^{3/2}}{m}\frac{y^2}{x^{3/2}} p_{1,3} \\
v_1^* &= \langle p_1, f_{1,2}\rangle_d = \sqrt{\frac{x}{\zeta}} p_{1,2} - \frac{\zeta^{3/2}}{m}\frac{y}{\sqrt{x}} p_{1,3}
\end{aligned}
\tag{3.29}
$$

and in the case of the resistor-capacitor circuit, it is

$$
\begin{aligned}
u_2^* &= \langle p_2, f_{2,1}\rangle_d = 2\sqrt{gc\beta} p_{2,1} + \sqrt{\frac{g}{c\beta}} p_{2,3} \\
v_2^* &= \langle p_2, f_{2,2}\rangle_d = -2\sqrt{gc\beta} p_{2,1} + 2\beta^{3/2}\sqrt{\frac{g}{c}} p_{2,2} - \sqrt{\frac{g}{c\beta}} p_{2,3}
\end{aligned}
\tag{3.30}
$$

Correspondingly,

$$
H_i(p_i, \gamma_i(t)) = \frac{1}{2}\left((p_i^\intercal f_{i,1})^2 + (p_i^\intercal f_{i,2})^2\right) = \frac{1}{2}\left((u_i^*)^2 + (v_i^*)^2\right) \tag{3.31}
$$

In particular, every regular extremal trajectory is smooth and it is a solution of the Hamiltonian system $\dot{\tilde{\lambda}}_i = \vec{H}_i(\tilde{\lambda}_i)$, where $\vec{H}_i$ is the Hamiltonian vector field. Here $\{\,,\}$ is Poisson bracket. In canonical coordinates, for a real-valued function f of $\tilde{\lambda}_i$,

$$
\vec{H}_i(f) = \{f, H_i\} = \left(\frac{\partial H_i}{\partial p_i}\right)^\intercal \frac{\partial f}{\partial \gamma_i} - \left(\frac{\partial H_i}{\partial \gamma_i}\right)^\intercal \frac{\partial f}{\partial p_i} \tag{3.32}
$$

If $\tilde{\lambda}_i(0)$ is the initial condition, the solution of the Hamiltonian system is $\tilde{\lambda}_i(t) = e^{t\vec{H}_i}(\tilde{\lambda}_i(0))$ and

$$
\begin{aligned}
\dot{p}_i &= -\frac{\partial H_i}{\partial \gamma_i} = -u_i^* D_{\gamma_i}(p^\intercal f_{i,1}) - v_i^* D_{\gamma_i}(p^\intercal f_{i,2}) \\
\dot{\gamma}_i &= \frac{\partial H_i}{\partial p_i} = (p^\intercal f_{i,1}) f_{i,1} + (p^\intercal f_{i,2}) f_{i,2} = u_i^* f_{i,1} + v_i^* f_{i,2}
\end{aligned}
\tag{3.33}
$$

where $p_{i,3}$ is observed to be a constant. As $\dot{H}_i = \{H_i, H_i\} = 0$, H_i is also a constant along the trajectory of $\vec{H}_i$. u_i^* and v_i^* can be parameterized as $\sqrt{2H_i}\cos\phi_i$ and $\sqrt{2H_i}\sin\phi_i$.

$$\dot{u^*}_i = \{u_i^*, H_i\} = v_i^*\{u_i^*, v_i^*\}$$

$$\dot{v^*}_i = \{v_i^*, H_i\} = u_i^*\{v_i^*, u_i^*\} \tag{3.34}$$

Thus, in the case of the stochastic oscillator,

$$\dot{\phi}_1 = -\{u_1^*, v_1^*\} = \frac{p_{1,2}}{2\sqrt{\zeta}} - \frac{3\zeta^{3/2}}{2m}\frac{y}{x}p_{1,3} \tag{3.35}$$

and in the case of the resistor-capacitor circuit,

$$\dot{\phi}_2 = -\{u_2^*, v_2^*\} = -2g\beta p_{2,1} - 2\frac{g\beta^2}{c}p_{2,2} + \frac{g}{c}p_{2,3} \tag{3.36}$$

To replace $p_{1,2}$ in the equation of $\dot{\phi}_1$ in the stochastic oscillator, from $v_1^* = \sqrt{2H_1}\sin\phi_1 = \sqrt{\frac{x}{\zeta}}p_{1,2} - \frac{\zeta^{3/2}}{m}\frac{y}{\sqrt{x}}p_{1,3}$, we get $p_{1,2} = \frac{\zeta^2}{m}\frac{y}{x}p_{1,3} + \sqrt{\frac{2H_1\zeta}{x}}\sin\phi_1$. The necessary condition for the heat engine (3.1) to work with maximum efficiency gives a trajectory on N_1 with dynamics:

$$\dot{x} = \sqrt{\frac{2H_1x}{\zeta}}\cos\phi_1;\ \dot{y} = \sqrt{\frac{2H_1x}{\zeta}}\sin\phi_1$$

$$\dot{\phi}_1 = \frac{1}{2}\sqrt{\frac{2H_1}{x}}\sin\phi_1 - \frac{\zeta^{3/2}}{m}\frac{y}{x}p_{1,3} \tag{3.37}$$

In the case of the resistor-capacitor circuit, from $u_2^* = \sqrt{2H_2}\cos\phi_2 = 2\sqrt{gc\beta}p_{2,1} + \sqrt{\frac{g}{c\beta}}p_{2,3}$ and $v_2^* = \sqrt{2H_2}\sin\phi_2 = -2\sqrt{gc\beta}p_{2,1} + 2\beta^{3/2}\sqrt{\frac{g}{c}}p_{2,2} - \sqrt{\frac{g}{c\beta}}p_{2,3}$, we have

$$-2g\beta p_{2,1} = -\sqrt{\frac{2H_2g\beta}{c}}\cos\phi_2 + \frac{g}{c}p_{2,3} \tag{3.38}$$

$$-2g\frac{\beta^2}{c}p_{2,2} = -\sqrt{\frac{2H_2g\beta}{c}}\cos\phi_2 - \sqrt{\frac{2H_2g\beta}{c}}\sin\phi_2$$

The necessary condition for the heat engine (3.2) to work with maximum efficiency gives a trajectory on N_2 with dynamics:

$$\begin{aligned}
\dot{c} &= 2\sqrt{2H_2gc\beta}\cos\phi_2 - 2\sqrt{2H_2gc\beta}\sin\phi_2; \\
\dot{\beta} &= 2\sqrt{\frac{2H_2g}{c}}\beta^{3/2}\sin\phi_2 \\
\dot{\phi}_2 &= -2\sqrt{\frac{2H_2g\beta}{c}}\cos\phi_2 - \sqrt{\frac{2H_2g\beta}{c}}\sin\phi_2 + \frac{2g}{c}p_{2,3}
\end{aligned} \tag{3.39}$$

Definition 3.2.1. *In general, from $\gamma(0) \in N$ and along a regular extremal trajectory of $\vec{H}$ determined by Pontryagin maximum principle, at time t, we define the t-exponential map of the optimal solution*

$$\mathcal{E}_{\gamma(0)} : \mathbb{R}^+ \times T^*_{\gamma(0)}N \to N, \;\; \mathcal{E}_{\gamma(0)}(t, \tilde{\lambda}(0)) = \pi\left(e^{t\vec{H}}(\tilde{\lambda}(0))\right) \tag{3.40}$$

where π is a projection from T^*N to N.

In the case of the heat engines, as $\gamma_i(0)$ is fixed, the information of $\tilde{\lambda}_i(0)$ is determined by $\phi_i(0)$, H_i and $p_{i,3}$. We can write $\tilde{\lambda}_i(0) = \tilde{\lambda}_i(\phi_i(0), H_i, p_{i,3})$.

Also, along a regular extremal in the time interval $[0, t_f]$, $\forall t \in [0, t_f]$, if $\gamma_i^*(t)$ belongs to a bounded region and $\dot{\gamma}_i^*(t)$ is piece-wise continuous and satisfies a Lipschitz condition in that region, there exists a unique solution $\gamma_i^*(t)$ to the Cauchy problem

$$\dot{\gamma}_i^* = u_i^* f_{i,1} + v_i^* f_{i,2}, \gamma_i^*(0) \text{ is fixed} \tag{3.41}$$

The t-exponential map of a heat engine can also be defined as mapping

$$F : \mathbb{R}^+ \times L^2_{[0,t_f]} \times L^2_{[0,t_f]} \to N_i, \; F(t, u_i^*, v_i^*) = \gamma_i^*(0) + \int_0^t (u_i^* f_{i,1} + v_i^* f_{i,2})\, dt' \tag{3.42}$$

At time t, the partial differential of the t-exponential map at (u_i^*, v_i^*) is the map

$$D_{(u_i^*,v_i^*)}F : L^2_{[0,t_f]} \times L^2_{[0,t_f]} \to T_{\gamma_i(t)}N_i,\ D_{(u_i^*,v_i^*)}F(u_{i,p}, v_{i,p}) = \int_0^t P^t_{t'*}(u_{i,p}f_{i,1} + v_{i,p}f_{i,2})dt' \tag{3.43}$$

where $P^t_\tau : \gamma_i^*(\tau) \to \gamma_i^*(t)$ is the flow generated by (u_i^*, v_i^*), $P^t_{\tau *}$ is its push-forward operator [26] and $(u_{i,p}, v_{i,p})$ is the perturbation in control (further elaborated in the proof of proposition 3.2.2). After introducing the concept of t-exponential map, we state a local property of geodesic sphere on N_i, which can be viewed as Gauss' lemma [21] in sub-Riemannian manifold N_i. This theorem and the proof is a special case of the work in [26].

Proposition 3.2.2. *Along a regular extremal starting from $\gamma_i(0)$, $\tilde{\lambda}_i(t) = e^{t\vec{H}_i}(\tilde{\lambda}_i(0))$. The corresponding cotangent vector $p_i(t)$ of $\tilde{\lambda}_i(t)$ annihilates the tangent space to the sub-Riemannian front at $\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i(0))$.*

Proof. $\tilde{\lambda}_i^s(0)$ is a smooth variation of initial co-vector $\tilde{\lambda}_i(0) = \tilde{\lambda}_i(\phi_i(0), H_i, p_{i,3})$, such that $\lambda_i^0(0) = \lambda(0)_i$. As a consequence, $\forall t' \in [0, t]$, $\tilde{\lambda}_i^s(t') = e^{t'\vec{H}_i}(\tilde{\lambda}_i^s(0))$ and $\tilde{\lambda}_i^s(t') = (p_i^s, \gamma_i^s)(t')$. The pair (u_i^s, v_i^s) is the optimal control associated to $\tilde{\lambda}_i^s$, where $u_i^s(t') = \langle p_i^s, f_{i,1}\rangle_d (t')$ and $v_i^s(t') = \langle p_i^s, f_{i,2}\rangle_d (t')$.

As H_i is a constant along a regular extremal, let us consider a family of initial cotangent vectors $\tilde{\lambda}_i^s(0) \in H_i^{-1}\left(\frac{1}{2}\right)$, where $H_i^{-1}\left(\frac{1}{2}\right) = \left\{\tilde{\lambda}_i^s | (u_i^s)^2 + (v_i^s)^2 = 1\right\}$. With this assumption, $\tilde{\lambda}_i^s(0) = \tilde{\lambda}_i^s(\phi_i^s(0), p_{i,3}^s)$. To prove the theorem, it is enough to show that at time t

$$\left\langle p_i^0(t), \frac{d}{ds}|_{s=0} \mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i^s(0)) \right\rangle_d = 0 \tag{3.44}$$

As shown in (3.42), the t-exponential map $\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i^s(0))$ is a functional of

(u_i^s, v_i^s) (i.e. $\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i^s(0)) = F(t, u_i^s, v_i^s)$, with a fixed $\gamma_i(0)$).

$$\frac{d}{ds}|_{s=0}\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i^s(0)) = \frac{d}{ds}|_{s=0}F(t, u_i^s, v_i^s) = D_{(u_i^*, v_i^*)}F(u_{i,p}, v_{i,p}), \tag{3.45}$$

where $u_{i,p} = \frac{du_i^s}{ds}|_{s=0}$ and $v_{i,p} = \frac{dv_i^s}{ds}|_{s=0}$. Notice that since $(u_i^s)^2 + (v_i^s)^2 = 1$, $(u_{i,p}, v_{i,p})$ is orthogonal to (u_i^*, v_i^*):

$$\int_0^t (u_{i,p}u_i^* + v_{i,p}v_i^*)\, dt' = 0 \tag{3.46}$$

Thus,

$$\begin{aligned}\left\langle p_i^0(t), \frac{d}{ds}|_{s=0}\mathcal{E}_{\gamma_i^s(0)}(t, \tilde{\lambda}_i^s(0))\right\rangle_d &= \left\langle p_i^0(t), D_{(u_i^*, v_i^*)}F(u_{i,p}, v_{i,p})\right\rangle_d \\ &= \left\langle p_i^0(t), \int_0^t P_{t'\,*}^{t}\,(u_{i,p}f_{i,1} + v_{i,p}f_{i,2})\, dt'\right\rangle_d \\ &= \int_0^t \left\langle P_{t'}^{t*}p_i^0(t), u_{i,p}f_{i,1} + v_{i,p}f_{i,2}\right\rangle_d dt' \\ &= \int_0^t (u_{i,p}u_i^* + v_{i,p}v_i^*)\, dt' = 0\end{aligned} \tag{3.47}$$

where $P_{t'}^{t*}$ is the dual of $P_{t'\,*}^{t}$ and it is a pull-back operator. $\square$

Here, we expand the result in the proof a little bit further. As $\tilde{\lambda}_i(0) = \tilde{\lambda}_i(\phi_i(0), p_{i,3})$, the variation $\tilde{\lambda}_i^s(0)$ is due to the variation of $(\phi_i^s(0), p_{i,3}^s)$.

$$\begin{aligned}\frac{d}{ds}|_{s=0}\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i^s(0)) &= \frac{d}{ds}|_{s=0}\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i(\phi_i^s(0), p_{i,3}^s)) \\ &= \left(\frac{d\phi_i^s(0)}{ds}\right)|_{s=0}\frac{\partial}{\partial\phi_i^s(0)}|_{s=0}\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i(\phi_i^s(0), p_{i,3}^s)) \\ &+ \left(\frac{dp_{i,3}^s}{ds}\right)|_{s=0}\frac{\partial}{\partial p_{i,3}^s}|_{s=0}\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i(\phi_i^s(0), p_{i,3}^s))\end{aligned} \tag{3.48}$$

As $\phi_i^s(0)$ and $p_{i,3}^s$ are independent variables, we get separately

$$\left\langle p_i^0(t), \frac{\partial}{\partial \phi_i^s(0)}|_{s=0}\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i(\phi_i^s(0), p_{i,3}^s))\right\rangle_d = 0$$
$$\left\langle p_i^0(t), \frac{\partial}{\partial p_{i,3}^s}|_{s=0}\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i(\phi_i^s(0), p_{i,3}^s))\right\rangle_d = 0 \tag{3.49}$$

3.3 A working cycle for a heat engine and conjugate point theory

In the last section, we saw that protocol (working loop) for a heat engine of maximum efficiency, under the condition of being a regular extremal, satisfies the Hamilton equations associated to the Pontryagin maximum principle. Among these optimal working loops, there is a special group called *working cycles*. By analyzing the working cycles with conjugate point theory, the optimality of working loops is investigated.

Definition 3.3.1. *For a heat engine operating along optimal working loops, from time 0 to time t_f, we say the protocol is a working cycle if it satisfies the following conditions.*

(1) For the stochastic oscillator: $x(0) = x(t_f)$, $y(0) = y(t_f)$ and for the resistor-capacitor circuit: $c(0) = c(t_f)$, $\beta(0) = \beta(t_f)$;

(2) For both the stochastic oscillator and the resistor-capacitor circuit: $\phi_i(t_f) = \phi_i(0) - 2\pi$ and ϕ_i is monotonically decreasing.

For the stochastic oscillator, along a loop in the space of (x, y),

$$\psi_1(t_f) = \int \frac{\zeta^2}{m}\frac{y^2}{x^2}dx - \frac{\zeta^2}{m}\frac{y}{x}dy = \iint -\frac{\zeta^2}{m}\frac{y}{x^2}dxdy \tag{3.50}$$

For the resistor-capacitor circuit, along a loop in the space of (c, β),

$$\psi_2(t_f) = \int \frac{dc}{2c\beta} = -\iint \frac{1}{2c\beta^2} dc d\beta \tag{3.51}$$

Condition 1 *and* 2 *are sufficient conditions to have the working cycle to be a periodic and clock-wise trajectory in the space of* (x, y) *or* (c, β)*, by which we can extract positive mechanical work from the heat bath by operating the stochastic oscillator system or the resistor-capacitor circuit. Moreover, the time* t_f *is the period of the working cycle.*

AS for the geodesics in Riemannian geometry, along a geodesic on a sub-Riemannian manifold, at time t, the image of t-exponential map may turn out to be a conjugate point, beyond where the geodesics will lose its optimality. The conjugate point theory of sub-Riemannian geometry is well-developed in [26], we apply this knowledge to our heat engine problem.

Definition 3.3.2. *In general, fix* $\gamma(0) \in N$ *and consider the t-exponential map* $\mathcal{E}_{\gamma(0)}(t, \lambda(0))$*. A point* $\gamma(t) \neq \gamma(0)$ *is said conjugate to* $\gamma(0)$ *if* $\gamma(t)$ *is a critical value for* $\mathcal{E}_{\gamma(0)}(t, \lambda(0))$ *at time* t*.*

Following the definition of a working cycle for our heat engine models, we will prove that at time t_f, $\gamma_i(t_f)$ is a conjugate point to $\gamma_i(0)$ ($i = 1$ for the stochastic oscillator and $i = 2$ for the resistor-capacitor circuit).

Theorem 3.3.3. *Along a regular extremal of one of our heat engine models and satisfying conditions 1 and 2 in the definition of a working cycle,* $\gamma_i(t_f) \neq \gamma_i(0)$ *and it is conjugate to* $\gamma_i(0)$*.*

Proof. Along a working cycle of the stochastic oscillator, as x and y are positive, $\psi_1(t_f) = \int \frac{\zeta^2}{m}\frac{y^2}{x^2}dx - \frac{\zeta^2}{m}\frac{y}{x}dy = \iint -\frac{\zeta^2}{m}\frac{y}{x^2}dxdy > 0$. $\gamma_1(t_f) \neq \gamma_1(0)$.

Similarly, along a working cycle of the resistor-capacitor circuit, as c and β are positive, $\psi_2(t_f) = \int \frac{dc}{2c\beta} = -\iint \frac{1}{2c\beta^2}dcd\beta > 0$. Hence, $\gamma_2(t_f) \neq \gamma_2(0)$.

Once we choose the initial condition $\gamma_i(0)$ for a working cycle, its t-exponential map $\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i(0))$ is a function of time t and $\tilde{\lambda}_i(0) = \tilde{\lambda}_i(\phi_i(0), p_{i,3})$. (Again, we take $H_i = \frac{1}{2}$.) If $\gamma_i(t_f) = \mathcal{E}_{\gamma_i(0)}(t_f, \tilde{\lambda}_i(0))$ will be conjugate to $\gamma_i(0)$, its differential is not surjective. To prove the differential is not surjective, it is enough to prove that

$$\frac{\partial}{\partial\phi(0)}|_{t=t_f}\mathcal{E}_{\gamma_i(0)}(t, \tilde{\lambda}_i(0)) = 0. \tag{3.52}$$

At time t_f, as a working cycle is an optimal working loop, for the stochastic oscillator: $x(0) = x(t_f)$, $y(0) = y(t_f)$ and for the resistor-capacitor circuit: $c(0) = c(t_f)$, $\beta(0) = \beta_{t_f}$. Thus, along a regular extremal of a heat engine model,

$$\begin{aligned} \frac{\partial x(t_f)}{\partial\phi_1(0)} &= \frac{\partial y(t_f)}{\partial\phi_1(0)} = 0 \\ \text{or } \frac{\partial c(t_f)}{\partial\phi_2(0)} &= \frac{\partial\beta(t_f)}{\partial\phi_2(0)} = 0 \end{aligned} \tag{3.53}$$

As ϕ_1 is monotonically decreasing ($\dot{\phi}_1 < 0$), x and y can be re-parameterized with ϕ_1. Thus,

$$\begin{aligned} \psi_1(t_f) &= \int_0^{t_f} \frac{\zeta^2}{m}\left(\frac{y^2}{x^2}\dot{x} - \frac{y}{x}\dot{y}\right)dt \\ &= \int_{\phi_1(0)}^{\phi_1(0)-2\pi} \frac{\zeta^2}{m}\left(\frac{y^2}{x^2}\frac{dx}{d\phi_1} - \frac{y}{x}\frac{dy}{d\phi_1}\right)d\phi_1 \end{aligned} \tag{3.54}$$

By calculation, as a working cycle is a periodic trajectory in the space of (x, y), it is observed $\frac{\partial\psi_1(t_f)}{\partial\phi_1(0)} = 0$ and $\frac{\partial}{\partial\phi_1(0)}\mathcal{E}(t_f, \tilde{\lambda}_1(0))$ is a zero-vector. Similarly, it can be proved that $\frac{\partial}{\partial\phi_2(0)}\mathcal{E}(t_f, \tilde{\lambda}_2(0))$ is also a zero-vector. Therefore, in each model, rank of the differential of the t-exponential map is less than 3 and a working cycle will gives us a conjugate point to $\gamma_i(0)$ at time t_f. $\square$

Moreover, the appearance of conjugate point will affect the optimality of a regular extremal. On that, without proof we state a theorem below. For more details, please refer to [26]

Theorem 3.3.4. *Let γ_i^* be a working cycle. $\gamma_i^*(t)$ is not conjugate to $\gamma^*(0)$ for every $0 \leq t < t_f$. The extremal γ_i is a strong minimum and the period time t_f of the working cycle is called conjugate time of this extremal.*

Remark 3.3.5. *For a optimal working loop of a heat engine, whose monotonically decreasing phase angle ϕ_i does not reach $\phi_i(0) - 2\pi$, it is a strong minimum over the time interval $[0, t_f]$.*

Remark 3.3.6. *ψ_1 is on an area moment for a metric conformal to Poincaré metric, so the constraint is on area-moment. The related sub-Riemannian problem of Poincaré length minimization subject to Poincaré area constraint is classical - the hyperbolic isoperimetric problem. The well-known solution is arrived in next chapter using the Pontryagin maximum principle, exploiting symmetry that is clear when the problem is mapped to $SL(2, \mathbb{R})$.*

Chapter 4: Optimal Control in Hyperbolic Space

In Chapter 3, we solved the optimal control problems associated to the maximum efficiency working loops of both heat engines by appeal to Pontryagin maximum principle. The solutions are in the form of ordinary differential equations, whose integration is of interest to design the working loops in practice. The analytic integrability of a solution is desired. With this property, the design of a working loop is simpler than using numerical methods, and accuracy of the design is also well guaranteed. Even though analytic integrability of the solution to the maximum efficiency working loop problem of an engine does not hold, integrability of the solution to a related problem can be a stepping stone for further analysis and design. In this chapter, we recover the isoperimetric equality in Poincaré upper half plane as the result of solving an optimal control problem related to the maximum efficiency working loop problem of the stochastic oscillator, because the extracted mechanical work along a working loop of the stochastic oscillator is on an area moment for a metric conformal to Poincaré metric.

4.1 Symmetry in Poincaré upper half plane

In the discussion in Chapter 1 on the stochastic oscillator, the system parameter manifold with (β, k) is isometrically mapped to a submanifold of Poincaré upper half plane $\mathbb{H}$ with coordinate of (x, y). Along a curve $\gamma : t \in [0, t_f] \to \mathbb{H}$ with its tangent vector $\dot{\gamma} = (\dot{x}, \dot{y})^\intercal$, the length of γ, $l(\gamma) = \int_0^{t_f} \sqrt{\frac{\dot{x}^2+\dot{y}^2}{y^2}}dt$. If γ is a loop, its enclosed area is $\oint_0^{t_f} \frac{\dot{x}}{y}dt = \iint \frac{dxdy}{y^2}$. For the ease in later investigation, we scale the enclosed area by $-\frac{1}{2}$, such that $\theta = \int_0^{t_f} -\frac{\dot{x}}{2y}dt$. With a pair of control (u, v), dynamics of (x, y, θ) is

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{pmatrix} = u \begin{pmatrix} y \\ 0 \\ -\frac{1}{2} \end{pmatrix} + v \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \tag{4.1}$$

As $f_1 = (y, 0, -1/2)^\intercal$ and $f_2 = (0, y, 0)^\intercal$ are vector fields in (4.1), (4.1) can be verified to be completely non-holonomic (i.e. controllable). Thus there is a sub-Riemannian geometry structure M in the coordinates of (x, y, θ). The control (u, v) will steer a curve in M and its length is measured by $\int \sqrt{u^2+v^2}dt = \int \sqrt{\frac{\dot{x}^2+\dot{y}^2}{y^2}}dt$, which is equal to the length of γ in $\mathbb{H}$.

On the other hand, the matrix group $G = SL(2, \mathbb{R})$ consists of all elements g of the form

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{where } a, b, c, d \in \mathbb{R} \text{ and } ad - bc = 1. \tag{4.2}$$

By QR factorization

$$g = \begin{pmatrix} \sqrt{y} & \frac{x}{\sqrt{y}} \\ 0 & \frac{1}{\sqrt{y}} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{4.3}$$

In other words, there is a local diffeomorphism between M and $SL(2,\mathbb{R})$. A basis for the Lie algebra $\mathfrak{g} = sl(2,\mathbb{R})$ is

$$A_1 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \; A_2 = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \; A_3 = \begin{pmatrix} 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \tag{4.4}$$

With $[A_1, A_2] = -A_3$ and inner product $\langle A_i, A_j \rangle \equiv 2\, trace\,(A_i A_j^T)$,

$$\dot{g} = g(u_1 A_1 + u_2 A_2) \text{ and its control is } \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -\sin 2\theta & \cos 2\theta \\ \cos 2\theta & \sin 2\theta \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{4.5}$$

It is clear that (4.5) is left-invariant and completely non-holonomic. Moreover, $\int \sqrt{u_1^2 + u_2^2}\, dt = \int \sqrt{u^2 + v^2}\, dt$. Thus the length of a curve in $SL(2,\mathbb{R})$ with the control (u_1, u_2) is equal to the length of the corresponding curve in $\mathbb{H}$. The symmetry of M is exposed in (4.5) and it will be helpful for further analysis.

4.2 Isoperimetric problem in Poincaré upper half plane

To minimize the length of a curve (loop) $\gamma : t \in [0, T] \to \mathbb{H}$ whose enclosed area is prescribed, is the isoperimetric problem in Poincaré upper half plane. In the language of geometric optimal control and based on Holder's inequality, for a curve in M, this is the same as

$$\underset{(u,v)\in L^2\times L^2}{\text{Min}} \quad J = \int_0^T \left(\frac{u^2+v^2}{2}\right) dt \tag{4.6}$$

$$\text{subject to:} \quad x(0) = x(T),\ y(0) = y(T), \theta(T) \text{ are specified}$$

For each trajectory in M, there is a corresponding trajectory in $SL(2,\mathbb{R})$ with the same length. The problem (4.6) is equivalent to

$$\underset{(u_1,u_2)\in L^2\times L^2}{\text{Min}} \quad J = \int_0^T \left(\frac{u_1^2+u_2^2}{2}\right) dt \tag{4.7}$$

$$\text{subject to:} \quad \dot{g} = g(u_1A_1 + u_2A_2), g(0) \text{ and } g(T) \text{ are specified}$$

In (4.7), the Lagrangian $L = \frac{1}{2}(u_1^2 + u_2^2)$ is G-invariant. $\dot{g} = g\xi \in T_gG$, where $\xi(= u_1A_1 + u_2A_2) \in \mathfrak{g}$. Correspondingly, $p \in T_g^*G$ and $\langle p, \dot{g}\rangle_d \in \mathbb{R}$. Following the maximum principle for optimal control on a differentiable manifold [24], we define the Hamiltonian as $H(p, g) = \max\limits_{u_1,u_2} (\langle p, g\xi\rangle_d - L)$.

Given $\mu(= \mu_1A_1^\flat + \mu_2A_2^\flat + \mu_3A_3^\flat) \in \mathfrak{g}^*$ where $\{A_1^\flat, A_2^\flat, A_3^\flat\}$ is a basis of $\mathfrak{g}^*$ such that $\left\langle A_i^\flat, A_j\right\rangle_d = \delta_{ij}$ $(i, j = 1, 2, 3)$.

$$\langle p, g\xi\rangle_d = \langle p, T_eL_g \cdot \xi\rangle_d = \left\langle T_eL_g^* \cdot p, \xi\right\rangle_d = \langle \mu, \xi\rangle_d\,, \tag{4.8}$$

optimal controls satisfying $u_1^* = \mu_1$ and $u_2^* = \mu_2$. $H(p, g) = h(\mu_1, \mu_2) = \frac{\mu_1^2+\mu_2^2}{2}$ is G-invariant. Thus the Hamiltonian vector $\vec{H}$ on T^*G induced by H can be reduced to a Hamiltonian vector field on $\mathfrak{g}^*$. This reduction is Lie-Poisson reduction [27].

Let $\tilde{\mu} = (\mu_1, \mu_2, \mu_3)$ denote coordinate representation of $\tilde{\mu}$ with basis $\{A_1^\flat, A_2^\flat, A_3^\flat\}$,

$$\dot{\tilde{\mu}} = \Lambda(\tilde{\mu})\nabla h, \text{ where } \nabla h = \left(\frac{\partial h}{\partial \mu_1}, \frac{\partial h}{\partial \mu_2}, \frac{\partial h}{\partial \mu_3}\right)^\intercal \tag{4.9}$$

where $\Lambda(\tilde{\mu})$ is the Poisson tensor. As $[A_1\, A_2] = -A_3$; $[A_1\, A_3] = -A_2$; $[A_2\, A_3] = A_1$, the non-vanishing structure constants are $\Gamma_{12}^3 = -1$, $\Gamma_{13}^2 = -1$, $\Gamma_{23}^1 = 1$.

$$\Lambda(\tilde{\mu}) = \begin{pmatrix} 0 & \mu_3 & \mu_2 \\ -\mu_3 & 0 & -\mu_1 \\ -\mu_2 & \mu_1 & 0 \end{pmatrix} \tag{4.10}$$

There is a Casimir invariant function $f(\tilde{\mu}) = \mu_1^2 + \mu_2^2 - \mu_3^2$, as ∇f is in the kernel of $\Lambda(\tilde{\mu})$. Express (4.9) explicitly,

$$\begin{aligned} \dot{\mu}_1 &= \mu_2\mu_3 \\ \dot{\mu}_2 &= -\mu_1\mu_3 \\ \dot{\mu}_3 &= 0 \end{aligned} \tag{4.11}$$

Apparently, μ_3 is a constant and the solution of (μ_1, μ_2) is

$$\begin{aligned} \mu_1 &= A\cos(\mu_3 t) + B\sin(\mu_3 t) \\ \mu_2 &= -A\sin(\mu_3 t) + B\cos(\mu_3 t) \end{aligned} \tag{4.12}$$

where A and B are constants depending on the initial condition. As optimal controls $u_1^* = \mu_1, u_2^* = \mu_2$ and based on (4.5), the optimal control (u^*, v^*) in (4.6) is

$$\begin{aligned} \begin{pmatrix} u^* \\ v^* \end{pmatrix} &= \begin{pmatrix} -\sin(2\theta) & \cos(2\theta) \\ \cos(2\theta) & \sin(2\theta) \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \\ \dot{\theta} &= -\frac{u^*}{2} \end{aligned} \tag{4.13}$$

Let $\sqrt{A^2+B^2} = K$ and $\phi = \mu_3 t + 2\theta + \alpha$, where $\sin\alpha = \frac{A}{K}$ and $\cos\alpha = \frac{B}{K}$. A solution of the isoperimetric problem in $\mathbb{H}$ is given by

$$\dot{x} = K\cos(\phi)y;\ \dot{y} = K\sin(\phi)y;\ \dot{\phi} = -K\cos\phi + \mu_3 \tag{4.14}$$

4.3 Isoperimetric equality in Poincaré upper half plane

A curve $t \to (x(t), y(t))$ satisfying (4.14) is the solution of isoperimetric problem in $\mathbb{H}$ and the integration of (4.14) will satisfy the isoperimetric equality. In this section, it will be seen that Euclidean curvature κ of this curve which is viewed as a curve in the Euclidean plane is the key element in deciding the integrability of (4.14).

Following the maximum principle to solve (4.7), Let P denote the pair $(T^*G, \{\,,\,\})$, which is a Poisson manifold, where $\{\,,\,\}$ is the canonical Poisson bracket. As $H = \frac{\mu_1^2+\mu_2^2}{2}$ is G-invariant on P, by Noether's theorem [28] there is a momentum map which is conserved on the trajectories of the Hamiltonian vector field $\vec{H}$.

Theorem 4.3.1. *The Euclidean curvature of a solution to* (4.14) *is a linear combination of terms of momentum map on P and hence it is also conserved.*

Proof. Following the construction of momentum maps in [28], the momentum map on P is $J : T^*G \to \mathfrak{g}^*$. If $(p, g) \in T^*G$ and $\tilde{\xi} \in \mathfrak{g}$,

$$\left\langle J, \tilde{\xi} \right\rangle_d (p, g) = \left\langle p, \tilde{\xi}_p(g) \right\rangle_d \tag{4.15}$$

where $\tilde{\xi}_G(g)$ is the infinitesimal action of $\tilde{\xi}$ on G and explicitly $\tilde{\xi}_G(g) = \frac{d}{d\epsilon}[\exp(\epsilon\tilde{\xi}) \cdot$

$g]|_{\epsilon=0} = T_eR_g \cdot \tilde{\xi}$.

$$\begin{aligned}\left\langle J, \tilde{\xi} \right\rangle_d (p, g) &= \left\langle T_gL_{g^{-1}}{}^* \cdot \mu, T_eR_g \cdot \tilde{\xi} \right\rangle_d \\ &= \left\langle \mu, T_gL_{g^{-1}} \cdot T_eR_g \cdot \tilde{\xi} \right\rangle_d = \left\langle \mu, g^{-1}\xi g \right\rangle_d\end{aligned} \tag{4.16}$$

As $\{A_1, A_2, A_3\}$ is a basis of $\mathfrak{g}$, there are three linearly independent terms J_1, J_2 and J_3 in J. Moreover, for $i = 1, 2, 3$, the basis element $A_i^\flat \in \mathfrak{g}^*$ is represented by A_i and $\left\langle A_i^\flat, A_j \right\rangle_d = 2\text{trace}\left(A_i A_j^\intercal\right) = \delta_{ij}$. Therefore, J_1, J_2 and J_3 are expressed in the coordinates of $G \times \mathfrak{g}^*$ as

$$\begin{aligned}J_1 &= \mu_1(ad + bc) + \mu_2(bd - ac) - \mu_3(ac + bd) \\ J_2 &= -\mu_1(ab - cd) + \frac{1}{2}\mu_2(a^2 - c^2 - b^2 + d^2) + \frac{1}{2}\mu_3(a^2 - c^2 + b^2 - d^2) \\ J_3 &= -\mu_1(ab + cd) + \frac{\mu_2}{2}(a^2 + c^2 - b^2 - d^2) + \frac{\mu_3}{2}(a^2 + b^2 + c^2 + d^2)\end{aligned} \tag{4.17}$$

Because of QR factorization in $SL(2, \mathbb{R})$ (4.3), they can also be expressed in the coordinates of (x, y, θ) in $\mathbb{H}$ as

$$\begin{aligned}J_1 &= \frac{\dot{y}}{y} + \frac{x\dot{x}}{y^2} - \mu_3\frac{x}{y} \\ J_2 &= \frac{(y^2 - x^2 + 1)\dot{x}}{2y^2} - \frac{\dot{x}y}{y} + \frac{\mu_3}{2}\left(\frac{y^2 + x^2 - 1}{y}\right) \\ J_3 &= -\frac{x\dot{y}}{y} + \frac{(y^2 - x^2 - 1)\dot{x}}{2y^2} + \frac{\mu_3}{2}\left(\frac{y^2 + x^2 + 1}{y}\right)\end{aligned} \tag{4.18}$$

On the other hand, the Euclidean curvature κ from (4.14) is

$$\kappa = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}} = \frac{\dot{\phi}}{Ky} = -\frac{\dot{x}}{Ky^2} + \frac{\mu_3}{Ky} \tag{4.19}$$

It is clear that from (4.18) $\frac{1}{K}(J_3 - J_2) = -\frac{\dot{x}}{Ky^2} + \frac{\mu_3}{Ky} = \kappa$. Hence κ is conserved along the trajectories of the Hamiltonian vector field $\vec{H}$. $\square$

Remark 4.3.2. *The conservation of Euclidean curvature could be understood from the view of Euler-Lagrange principle. As an optimal control problem on* $\mathbb{H}$, $\frac{1}{2}\left(\frac{\dot{x}^2+\dot{y}^2}{y^2}\right)+\lambda\frac{\dot{x}}{2y}$ *is the augmented Lagrangian* L_λ *where* λ *is the Lagrange multiplier which is a constant. By observation,* x *is a cyclic coordinate in the augmented Lagrangian. Following Euler-Lagrange principle,* $\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial x} = 0$. $\frac{\dot{x}}{y^2} + \frac{\lambda}{2}\frac{1}{y}$ *is a conserved quantity. Taking* $\lambda = -2\mu_3$*, it indicates that the Euclidean Curvature is conserved.*

By reconstruction of (4.14), it is desired to have a cycle on $\mathbb{H}$ with perimeter L and enclosed hyperbolic area A. The relation between L and A will be the isoperimetric equality on Poincaré upper half plane.

First and foremost, to have a cycle on $\mathbb{H}$ satisfying (4.14) it is necessary to have

$$\begin{aligned} x(0) = x(T),\ y(0) = y(T) \\ \dot{x}(0) = \dot{x}(T),\ \dot{y}(0) = \dot{y}(T) \end{aligned} \tag{4.20}$$

Thus $\phi(0) = \phi(T) + 2k\pi,\ k = \pm1, \pm2, \cdots$, which indicates that along the cycle, $\dot{\phi} = -\sqrt{2H}\cos\phi + \mu_3$ does not have any equilibrium point (i.e. $|\mu_3| > \sqrt{2H}$). After this discussion on the relation between constants H and μ_3, we will present the control-theoretical proof of isoperimetric equality on Poincaré upper half plane, which has appeared in [29] from mathematical community.

Theorem 4.3.3. *The isoperimetric equality of Poincaré upper half plane is* $L = \sqrt{A^2 + 4\pi A}$*, where* L *is the perimeter of a loop and* A *is the enclosed hyperbolic area of the loop.*

Proof. Without losing generality, assume the curve in (4.14) is unit-speed in $\mathbb{H}$, i.e. K = 1 and $\mu_3 > 1$, ϕ is monotonically increasing from its initial value $\phi(0)$ to $\phi(0) + 2\pi$.

$$\dot{x} = \cos(\phi)y;\ \dot{y} = \sin(\phi)y;\ \dot{\phi} = -\cos\phi + \mu_3 \tag{4.21}$$

It is proved in 4.3.1 that the Euclidean curvature $\kappa = \frac{\dot{\phi}}{y}$ is a nonzero constant. $\frac{d\phi}{dt} = \kappa y$ and (4.21) is re-parameterized by ϕ.

$$\frac{dx}{d\phi} = \frac{1}{\kappa}\cos\phi,\ \frac{dy}{d\phi} = \frac{1}{\kappa}\sin\phi \tag{4.22}$$

It is easy to verify that (4.21) is a closed and positive-oriented loop in (x, y) space. The enclosed area of the loop is

$$\begin{aligned} A &= \int \frac{\dot{x}}{y} dt = \int_{\phi(0)}^{\phi(0)+2\pi} \left(\frac{\cos\phi}{\kappa y(0) + \cos\phi(0) - \cos\phi} \right) d\phi \\ &= \int_{\phi(0)}^{\pi-} \left(\frac{\cos\phi}{\kappa y(0) + \cos\phi(0) - \cos\phi} \right) d\phi \\ &+ \int_{\pi+}^{\phi(0)+2\pi} \left(\frac{\cos\phi}{\kappa y(0) + \cos\phi(0) - \cos\phi} \right) d\phi \\ &= \left(\sqrt{\frac{(\kappa y(0) + \cos\phi(0))^2}{|(\kappa y(0) + \cos\phi(0))^2 - 1|}} - 1 \right) 2\pi \end{aligned} \tag{4.23}$$

On the other hand, as a unit speed curve and $dt = \frac{d\phi}{\dot{\phi}}$, its perimeter is

$$\begin{aligned} L &= \int_{\phi(0)}^{\phi(0)+2\pi} \frac{d\phi}{\dot{\phi}} = \int_{\phi(0)}^{\phi(0)+2\pi} \frac{d\phi}{\kappa y(0) + \cos\phi(0) - \cos\phi} \\ &= \int_{\phi(0)}^{\pi-} \frac{d\phi}{\kappa y(0) + \cos\phi(0) - \cos\phi} \\ &+ \int_{\pi+}^{\phi(0)+2\pi} \frac{d\phi}{\kappa y(0) + \cos\phi(0) - \cos\phi} \\ &= 2\pi \sqrt{\left(\frac{\kappa y(0) + \cos\phi(0)}{\sqrt{|(\kappa y(0) + \cos\phi(0))^2 - 1|}} \right)^2 - 1} \end{aligned} \tag{4.24}$$

Comparing both (4.23) and (4.24), isoperimetric equality in Poincaré upper half plane is $L = \sqrt{A^2 + 4\pi A}$. $\square$

Chapter 5: Numerical Design of Working Loops

In Chapter 3, for each heat engine model, we showed the necessary condition of the maximum efficiency working loops. Also, based on the conjugate point theory analysis, a sufficient condition for the optimality of the maximum efficiency working loops is proposed. As a cyclically operated machine, the working loops of a heat engine are of special interest to design. The *time span* of a working loop is a key quantity in the design. To have this important information, in section 1, we apply *level set methods* to solve the reachability problem of each engine. By fixing the initial condition of a protocol, we can have the time spans and ending points of working loops which start from the same point in the parameter space with initially zero extracted mechanical work.

With the end-points and time span of the working loops, in section 2, we can reconstruct these working loops by *mid-point approximation*. To reduce the differences between the starting points and ending points in the space of the system parameter along each reconstructed trajectory, we refine the trajectories with *shooting method*. After above three steps, we have the maximum-efficiency working loops of both heat engines. The efficiencies of each engine working along different working loops are calculated and compared in section 3.

5.1 Reachable set of a heat engine

Over a time interval $[0, t]$, the working loops of the stochastic oscillator satisfy the condition $x(0) = x(t)$ and $y(0) = y(t)$ and the working loops of the resistor-capacitor circuit satisfy the condition $c(0) = c(t)$ and $\beta(0) = \beta(t)$.

As both heat engines are controllable, starting from a point $q_1^0 = (x^0, y^0, \psi_1 = 0)$ or $q_2^0 = (c^0, \beta^0, \psi_2 = 0)$, over the time interval $[0, t]$, the ending points of all minimum path dissipation protocols in the space of (x, y, ψ_1) (for the stochastic oscillator) or (c, β, ψ_2) (for the resistor-capacitor circuit) of a heat engine form a set

$$\{\gamma_i^*(t) | \gamma_i^* : \mathbb{R} \to N_i \text{ is a regular extremal and } \gamma_i^*(0) = q_i^0\} \tag{5.1}$$

This set is the reachable set of the heat engine at time t. $\forall t \in [0, t_{\max}]$, with fixed q_i^0, the collection of reachable sets

$$R(q_i^0)_{\tau \leq t_{\max}} = \cup_{\tau \leq t_{\max}} \{\gamma_i^*(\tau) | \gamma_i^* : \mathbb{R} \to N_i \text{ is a regular extremal and } \gamma_i^*(0) = q_i^0\} \tag{5.2}$$

is the set of all reachable points till time $t_{\max}$.

Among this set $R(q_i^0)$, we can select the points which satisfy the working loop condition. Thus, given a heat engine model, the first step to design the working loops is to compute the collection of reachable sets till $t_{\max}$. For design purpose, $t_{\max}$ is sufficiently large. The reachable sets of some control problems can be computed by level set methods [30] [31]. In the following discussion, we will show that our control problems are numerically solvable by level set methods.

Given a heat engine ($i = 1$ for the stochastic oscillator and $i = 2$ for the resistor-capacitor circuit),

$$\dot{\gamma}_i = u_i f_{i,1} + v_i f_{i,2} \tag{5.3}$$

where curve $\gamma_i : \mathbb{R} \to N_i$ and N_i is the sub-Riemannian manifold of a heat engine. (u_i, v_i) is the control and $(f_{i,1}, f_{i,2})$ are the vector fields. Over the time interval $[0, t_f]$, the dissipation of a protocol Λ_i (scaled by a constant $\frac{1}{2}$)

$$J(\gamma_i) = \frac{1}{2} \int_0^{t_f} \left(u_i^2 + v_i^2\right) dt \tag{5.4}$$

Compare with the associated arc-length of the curve on the sub-Riemannian manifold N_i

$$l(\gamma_i) = \int_0^{t_f} \sqrt{u_i^2 + v_i^2} dt \tag{5.5}$$

Geometrically, it is known that the dissipation J is the *geometric energy* of the curve γ_i. As in Riemannian geometry, in [26], *Lemma 3.47*, it is proved that minimizing the arc-length of the curve is equivalent to minimizing the geometric energy of the curve.

Along a unit-speed minimizer γ_i^* $(u_i^2 + v_i^2 = 1)$ from Chapter 3, starting from point A to point B on the manifold N_i, over the time interval $[0, t_f]$, the length $l(\gamma_i^*) = t_f$. Traveling along the extremal trajectory at a different speed ν_i from A to B, the minimum length (i.e. the *distance* between two points on the sub-Riemannian manifold) is invariant but the dissipation along the extremal (i.e. the geometric energy)

$$J = \frac{l^2}{2t_f/\nu_i} = \frac{t_f \nu_i}{2} \tag{5.6}$$

Thus, instead of directly computing the reachable sets of a heat engine with minimum dissipation, we can compute the reachable sets of the heat engine with the minimum arc-length. Based on the information from the reachable sets, we can compute the minimum dissipation by (5.6).

To minimize the arc-length for a heat engine, by Pontryagin's maximum principle,

$$H_i : T^*N_i \to \mathbb{R}, \quad H_i(\tilde{\lambda}_i) = \max_{u_i, v_i} \left(\langle p_i, u_i f_{i,1} + v_i f_{i,2} \rangle_d - \sqrt{u_i^2 + v_i^2} \right) \tag{5.7}$$

It follows that the optimal control is

$$\begin{aligned} u_i^* &= \langle p_i, f_{i,1} \rangle_d \sqrt{u_i^{*2} + v_i^{*2}} \\ v_i^* &= \langle p_i, f_{i,2} \rangle_d \sqrt{u_i^{*2} + v_i^{*2}} \end{aligned} \tag{5.8}$$

Bring the (u_i^*, v_i^*) into H_i, the Hamiltonian is zero. Moreover, as it is shown in [26], the minimizers are of constant speed. In the case of unit-speed extremals ($\sqrt{u_i^{*2} + v_i^{*2}} = 1$),

$$H_i = \sqrt{\langle p_i, f_{i,1} \rangle_d^2 + \langle p_i, f_{i,2} \rangle_d^2} - 1 = 0 \tag{5.9}$$

which gives us

$$\sqrt{\langle p_i, f_{i,1} \rangle_d^2 + \langle p_i, f_{i,2} \rangle_d^2} = 1 \tag{5.10}$$

On the other hand, by the approach of dynamic programming, to minimize the cost function,

$$l(t, \gamma_i(t), u_i, v_i) = \int_t^{t_f} \sqrt{u_i^2 + v_i^2} dt \tag{5.11}$$

where $L(t, \gamma_i(t), u_i(t), v_i(t)) = \sqrt{u_i^2(t) + v_i^2(t)}$ is the Lagrangian and γ_i is a curve starting from a point in N_i and $\gamma_i(t_f) = q_i^0$. We have the value function over the time interval $[0, t_f]$

$$\hat{V}(t, \gamma_i(t)) \equiv \inf_{(u_i, v_i)} l(t, \gamma_i(t), u_i(t), v_i(t)) \tag{5.12}$$

A sufficient condition for optimality of the geodesics on the sub-Riemannian manifold is: there is a $\mathcal{C}^1$ function $\hat{V} : [0, t_f] \times \mathbb{R}^n \to \mathbb{R}$ satisfying Hamilton-Jacobi-Bellman equation [15]

$$-\frac{\partial \hat{V}}{\partial t}(t, \gamma_i(t)) = \inf_{(u_i, v_i)} \left\{ L(t, \gamma_i(t), u_i(t), v_i(t)) + \left\langle \frac{\partial \hat{V}}{\partial \gamma_i}(t, \gamma_i(t)), \dot{\gamma}_i \right\rangle_d \right\} \tag{5.13}$$

($\forall t \in [0, t_f)$ and $\forall \gamma_i(t) \in N_i$) and the boundary condition

$$\hat{V}(t_f, \gamma_i(t_f)) = 0 \tag{5.14}$$

Supposing that there exists a control (u_i^*, v_i^*) and the corresponding trajectory $\gamma_i^* : [0, t_f] \to N_i$, with a given initial condition, satisfies the equation everywhere

$$\begin{aligned} L(t, \gamma_i^*(t), u_i^*(t), v_i^*(t)) + \left\langle \frac{\partial \hat{V}}{\partial \gamma_i}|_{\gamma_i^*}, \dot{\gamma}_i^* \right\rangle_d \\ = \inf_{(u_i, v_i)} \left\{ L(t, \gamma_i(t), u_i(t), v_i(t)) + \left\langle \frac{\partial \hat{V}}{\partial \gamma_i}, \dot{\gamma}_i \right\rangle \right\} \end{aligned} \tag{5.15}$$

compare (5.15) with Pontryagin maximum principle,

$$H_i \left(t, \gamma_i^*, u_i^*, v_i^*, -\frac{\partial \hat{V}}{\partial \gamma_i}|_{\gamma_i^*} \right) = \max_{u_i, v_i} \left(\langle p_i, u_i f_{i,1} + v_i f_{i,2} \rangle_d - \sqrt{u_i^2 + v_i^2} \right) \tag{5.16}$$

It is seen that along a regular extremal

$$p_i = -\frac{\partial \hat{V}}{\partial \gamma_i}|_{\gamma_i^*} \tag{5.17}$$

In the space of (x, y, ψ_1) or (c, β, ψ_2), based on [32], there exists a function $\tilde{V}_i : (t, q_i) \to \mathbb{R}$, which is monotonically decreasing with time, where $q_i \in N_i$ and $V_i(t(q_i), q_i) = 0$ is equivalent to the fact that there $\exists t'$ and $\hat{V}(t', \gamma_i^*(t') = q_i) = t(q_i)$ (i.e. the distance between the q_i and q_i^0 is $t(q_i)$). Thus, on the level surface $\tilde{V}_i(t(q_i), q_i) = 0$

$$\frac{d\tilde{V}_i}{dq_i} = \frac{\partial \tilde{V}_i}{\partial q_i} + \frac{\partial \tilde{V}_i}{\partial t} \frac{\partial t}{\partial q_i} = 0 \tag{5.18}$$

Given that $\frac{\partial \tilde{V}_i}{\partial t} < 0$,

$$\frac{\partial \hat{V}}{\partial \gamma_i}|_{\gamma_i^*(t')=q_i} = \frac{\partial t}{\partial q_i} = -\frac{\partial \tilde{V}_i}{\partial q_i}\frac{\partial t}{\partial \tilde{V}_i} \tag{5.19}$$

$$\text{along the extremal } \gamma_i^*,\ p_i(t') = -\frac{\partial \hat{V}}{\partial \gamma_i}|_{\gamma_i^*(t')=q_i} = \frac{\partial \tilde{V}_i}{\partial q_i}\frac{\partial t}{\partial \tilde{V}_i}$$

Bringing this expression of p_i into (5.10)

$$\sqrt{\left\langle \frac{\partial \tilde{V}_i}{\partial q_i}\frac{\partial t}{\partial \tilde{V}_i}, f_{i,1} \right\rangle_d^2 + \left\langle \frac{\partial \tilde{V}_i}{\partial q_i}\frac{\partial t}{\partial \tilde{V}_i}, f_{i,2} \right\rangle_d^2} = 1 \tag{5.20}$$

we have a Hamilton-Jacobi equation,

$$\frac{\partial \tilde{V}_i}{\partial t} + \sqrt{\left\langle \frac{\partial \tilde{V}_i}{\partial q_i}, f_{i,1} \right\rangle_d^2 + \left\langle \frac{\partial \tilde{V}_i}{\partial q_i}, f_{i,2} \right\rangle_d^2} = 0 \tag{5.21}$$

In the case of the stochastic oscillator, the Hamilton-Jacobi equation

$$\frac{\partial \tilde{V}_1}{\partial t} + \sqrt{\left\langle \frac{\partial \tilde{V}_1}{\partial q_1}, \begin{pmatrix} \sqrt{\frac{x}{\zeta}} \\ 0 \\ \frac{\zeta^{3/2}}{m}\frac{y^2}{x^{3/2}} \end{pmatrix} \right\rangle_d^2 + \left\langle \frac{\partial \tilde{V}_1}{\partial q_1}, \begin{pmatrix} 0 \\ \sqrt{\frac{x}{\zeta}} \\ -\frac{\zeta^{3/2}}{m}\frac{y}{\sqrt{x}} \end{pmatrix} \right\rangle_d^2} = 0 \tag{5.22}$$

In the case of the resistor-capacitor circuit, the Hamilton-Jacobi equation

$$\frac{\partial \tilde{V}_2}{\partial t} + \sqrt{\left\langle \frac{\partial \tilde{V}_2}{\partial q_2}, \begin{pmatrix} 2\sqrt{gc\beta} \\ 0 \\ \sqrt{\frac{g}{c\beta}} \end{pmatrix} \right\rangle_d^2 + \left\langle \frac{\partial \tilde{V}_2}{\partial q_2}, \begin{pmatrix} -2\sqrt{gc\beta} \\ 2\beta^{3/2}\sqrt{\frac{g}{c}} \\ -\sqrt{\frac{g}{c\beta}} \end{pmatrix} \right\rangle_d^2} = 0 \tag{5.23}$$

Both of them are numerically solvable by level set methods and the reachable sets are realizable.

5.1.1 A stochastic oscillator example

Given that

$$
\begin{aligned}
d\xi_1 &= \frac{\xi_2}{m}dt \\
d\xi_2 &= -\zeta\frac{\xi_2}{m}dt - k\xi_1 dt + \sqrt{\frac{2\zeta}{\beta}}dB(t)
\end{aligned}
\tag{5.24}
$$

and choosing k_0 as the initial value of k, then as a deterministic oscillator, its natural frequency $\omega_0 = \sqrt{\frac{k_0}{m}}$. To have this deterministic oscillator to be critically damped, $\zeta = 2\sqrt{mk_0}$ and β_0 is the initial value of the inverse temperature. In the space (x, y), the initial value is

$$
\begin{aligned}
x_0 &= \frac{1}{4\beta_0 k_0} \\
y_0 &= \frac{1}{\beta_0\zeta}\sqrt{\frac{m}{k_0}} = \frac{1}{2\beta_0 2\sqrt{mk_0}}\sqrt{\frac{m}{k_0}} = \frac{1}{4\beta_0 k_0} = x_0
\end{aligned}
\tag{5.25}
$$

It indicates that x and y are of the same scale and of the same unit meter^2. Choosing $k_0 = 0.25 \times 10^{-7} N/\text{meter} = 10^{-7} kg/\text{sec}^2$, $m = 1ng = 10^{-12}kg$ and $\beta_0 = \frac{1}{k_B T_0} = 2.473 \times 10^{20}(\text{sec})^2/(\text{meter}^2 kg)$ ($T_0 = 293K$ and $k_B = 1.38 \times 10^{-23}(\text{meter}(\text{sec})^{-1})^2 kg/K)$),

$$
x_0 = y_0 = 4.0436 \times 10^{-14}\text{meter}^2 \tag{5.26}
$$

Rescaling x and y by 10^{-14}, we have

$$
\begin{aligned}
x &= \tilde{x} \times 10^{-14}, \quad y = \tilde{y} \times 10^{-14} \\
\dot{x} &= \dot{\tilde{x}} \times 10^{-14}, \quad \dot{y} = \dot{\tilde{y}} \times 10^{-14}
\end{aligned}
\tag{5.27}
$$

The average dissipation rate in the unit of $k_B T/sec$ is

$$
\tilde{d} = \zeta\left(\frac{\dot{x}^2 + \dot{y}^2}{x}\right) = \zeta\left(\frac{\dot{\tilde{x}}^2 \times 10^{-28} + \dot{\tilde{y}}^2 \times 10^{-28}}{\tilde{x} \times 10^{-14}}\right) = \zeta\left(\frac{\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2}{\tilde{x}}\right) \times 10^{-14} \tag{5.28}
$$

and the extracted power is

$$\dot{\psi}_1 = \frac{\zeta^2}{m}\left(\frac{y^2}{x^2}\dot{x} - \frac{y}{x}\dot{y}\right) = \frac{\zeta^2}{m}\left(\frac{\tilde{y}^2}{\tilde{x}^2}\dot{\tilde{x}} - \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\right) \times 10^{-14} \tag{5.29}$$

In consequence, the efficiency η_1 over a time interval $[0, t_f]$ is

$$\begin{aligned}\eta_1 &= \frac{\int_0^{t_f} \dot{\psi}_1 dt}{\int_0^{t_f}\left(\left(\frac{\zeta^2}{m}\frac{y}{x}\dot{y}\right)\mathbb{1}\left\{\frac{y}{x}\dot{y} > 0\right\} + \tilde{d}_1\right) dt} \\ &= \frac{\int_0^{t_f} \frac{\zeta^2}{m}\left(\frac{\tilde{y}^2}{\tilde{x}^2}\dot{\tilde{x}} - \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\right) \times 10^{-14} dt}{\int_0^{t_f}\left(\frac{\zeta^2}{m}\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}} \times 10^{-14}\mathbb{1}\left\{\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}} > 0\right\} + \zeta\left(\frac{\dot{\tilde{x}}^2+\dot{\tilde{y}}^2}{\tilde{x}}\right) \times 10^{-14}\right) dt} \\ &= \frac{\int_0^{t_f} \frac{\zeta}{m}\left(\frac{\tilde{y}^2}{\tilde{x}^2}\dot{\tilde{x}} - \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\right) dt}{\int_0^{t_f}\left(\frac{\zeta}{m}\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\mathbb{1}\left\{\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}} > 0\right\} + \left(\frac{\dot{\tilde{x}}^2+\dot{\tilde{y}}^2}{\tilde{x}}\right)\right) dt}\end{aligned} \tag{5.30}$$

Based on the theorems in Chapter 3, having $\frac{\zeta}{m}$ as a constant weighting factor between the extracted mechanical work and the dissipation, the scaled extracted mechanical power $\dot{\tilde{\psi}}_1 = \frac{\tilde{y}^2}{\tilde{x}^2}\dot{\tilde{x}} - \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}$ and $\tilde{\psi}_1(0) = 0$, it is straightforward to show that to find a maximum efficiency working loop of the stochastic oscillator is equivalent to the problem of finding an optimal protocol which minimizes

$$J = \int_0^{t_f} \frac{\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2}{\tilde{x}} dt \tag{5.31}$$

while $\int_0^{t_f}\left(\frac{\tilde{y}^2}{\tilde{x}^2}\dot{\tilde{x}} - \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\right) dt$ is a prescribed value and $(\tilde{x}(0), \tilde{y}(0)) = (\tilde{x}(t_f), \tilde{y}(t_f))$. The associated sub-Riemannian structure is given by

$$\begin{pmatrix}\dot{\tilde{x}} \\ \dot{\tilde{y}} \\ \dot{\tilde{\psi}}_1\end{pmatrix} = \tilde{u}_1\begin{pmatrix}\sqrt{\tilde{x}} \\ 0 \\ \frac{\tilde{y}^2}{\tilde{x}^{3/2}}\end{pmatrix} + \tilde{v}_1\begin{pmatrix}0 \\ \sqrt{\tilde{x}} \\ -\frac{\tilde{y}}{\sqrt{\tilde{x}}}\end{pmatrix} \tag{5.32}$$

where $(\tilde{u}_1, \tilde{v}_1)$ is the new control and the new average dissipation rate is $\tilde{u}_1^2 + \tilde{v}_1^2$. The weighting factor in our example $\frac{\zeta}{m} = \frac{2\sqrt{mk_0}}{m} = 10$. Following the discussion on the

Hamilton-Jacobi equation (5.22), with the new $\tilde{V}_1$, the Hamilton-Jacobi equation is given by

$$\frac{\partial \tilde{V}_1}{\partial t} + \sqrt{\left\langle \frac{\partial \tilde{V}_1}{\partial q_1}, \begin{pmatrix} \sqrt{\tilde{x}} \\ 0 \\ \frac{\tilde{y}^2}{\tilde{x}^{3/2}} \end{pmatrix} \right\rangle_d^2 + \left\langle \frac{\partial \tilde{V}_1}{\partial q_1}, \begin{pmatrix} 0 \\ \sqrt{\tilde{x}} \\ -\frac{\tilde{y}}{\sqrt{\tilde{x}}} \end{pmatrix} \right\rangle_d^2} = 0 \tag{5.33}$$

Rescaling (x, y, ψ_1) to $(\tilde{x}, \tilde{y}, \tilde{\psi}_1)$ is for the ease of implementing the level set methods by having the range of the $\tilde{x}$, $\tilde{y}$ and $\tilde{\psi}_1$ in the same scale. Having the $(\tilde{x}(0), \tilde{y}(0), \tilde{\psi}(0)) = (4, 4, 0)$ with associated units, up to arc-length of 1, the reachable set of the stochastic oscillator in 3D is given by figure 5.1.

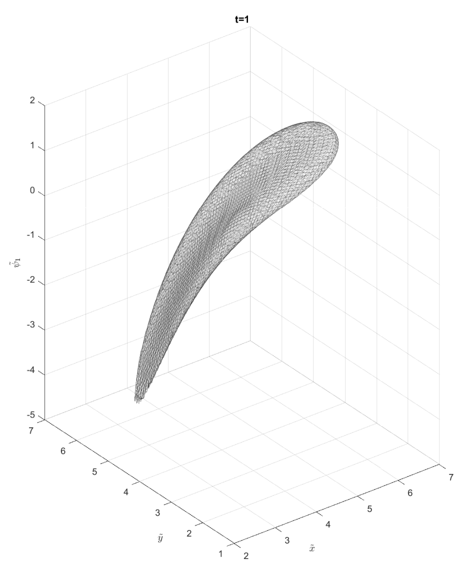

Figure 5.1: Reachable set of a stochastic oscillator in 3D

The view of the reachable set through $\tilde{x}$-axis is given by figure 5.2.

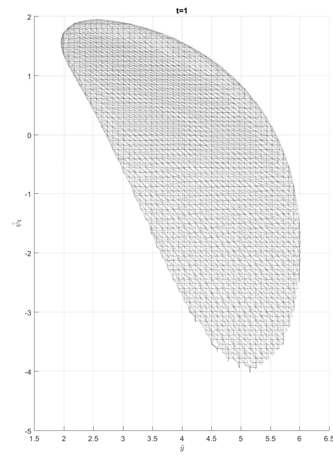

Figure 5.2: Reachable set in the space of $(\tilde{y}, \tilde{\psi}_1)$

The view of the reachable set through $\tilde{y}$-axis is given by figure 5.3.

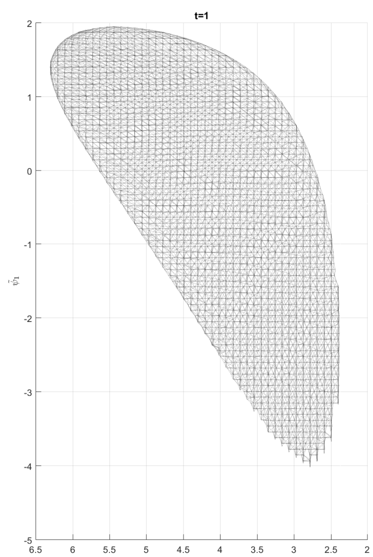

Figure 5.3: Reachable set in the space of $(\tilde{x}, \tilde{\psi}_1)$

The view of the reachable set through $\tilde{\psi}_1$-axis is given by figure 5.4.

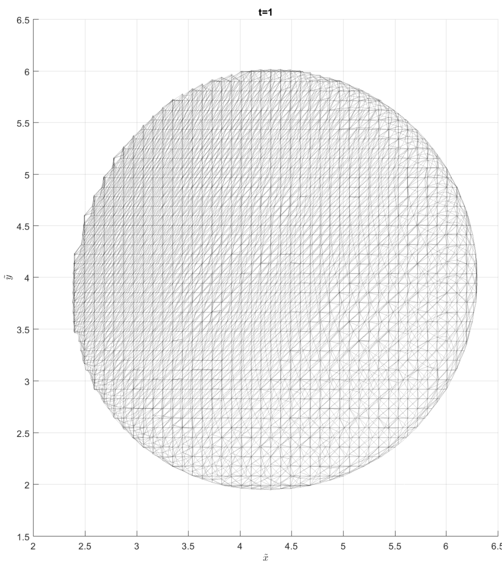

Figure 5.4: Reachable set in the space of $(\tilde{x}, \tilde{y})$

The reachable set provides the information about the distance between reachable points with $(\tilde{x}, \tilde{y}) = (4, 4)$ and positive $\tilde{\psi}_1$ coordinates and the center of the reachable set $(\tilde{x}, \tilde{y}, \tilde{\psi}_1) = (4, 4, 0)$. We display the distance information in table 5.1. Due to the discretization of $\tilde{\psi}_1$, the coordinate of $\tilde{\psi}_1$ is of step size 0.0784.

Table 5.1: Working loop end-points of the stochastic oscillator

Point number	distance	$\tilde{\psi}_1 - coordinate$
1	1.1126	0.1207
2	1.5302	0.1991
3	1.8331	0.2775
4	2.0515	0.356
5	2.2134	0.4344
6	2.3471	0.5128
7	2.4982	0.5913
8	2.6546	0.6697
9	2.8082	0.7481
10	2.9525	0.8266

5.1.2 A resistor-capacitor circuit example

Choose the initial values of the capacitance c_0 and the inverse temperature of the heat bath β_0 as

$$\begin{aligned} c_0 &= 60\text{pF} = 6 \times 10^{-11}\text{F} \\ \beta_0 &= \frac{1}{k_B T_0} = \frac{1}{1.38 \times 10^{-23} \times 293}((\text{meter})^{-1}\text{sec})^2/kg \\ &= 0.34 \times 10^{21}((\text{meter})^{-1}\text{sec})^2 kg^{-1} \end{aligned} \tag{5.34}$$

Thus, rescale the parameter as

$$c = \tilde{c} \times 10^{-11}, \quad \beta = \tilde{\beta} \times 10^{21} \tag{5.35}$$

and the velocities are rescaled as

$$\dot{c} = \dot{\tilde{c}} \times 10^{-11}, \quad \dot{\beta} = \dot{\tilde{\beta}} \times 10^{21} \tag{5.36}$$

Assuming that $g = 10^{-10}\text{ohm}^{-1}$, efficiency of the resistor-capacitor circuit over the time interval $[0, t_f]$

$$\begin{aligned}\eta_2 &= \frac{\int_0^{t_f} \frac{\dot{c}}{2c\beta} dt}{\int_0^{t_f} \left(\left(-\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} \right) \mathbb{1}\left\{ -\frac{\dot{\beta}}{2\beta^2} + \frac{\dot{c}}{2c\beta} > 0 \right\} + \left(\frac{\dot{c}^2}{4gc\beta} + \frac{\dot{c}\dot{\beta}}{2g\beta^2} + \frac{c\dot{\beta}^2}{2g\beta^3} \right) \right) dt} \\ &= \frac{\int_0^{t_f} \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}} dt}{\int_0^{t_f} \left(\left(-\frac{\dot{\tilde{\beta}}}{2\tilde{\beta}^2} + \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}} \right) \mathbb{1}\left\{ -\frac{\dot{\tilde{\beta}}}{2\tilde{\beta}^2} + \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}} > 0 \right\} + \left(\frac{\dot{\tilde{c}}^2}{4\tilde{c}\tilde{\beta}} + \frac{\dot{\tilde{c}}\dot{\tilde{\beta}}}{2\tilde{\beta}^2} + \frac{\tilde{c}\dot{\tilde{\beta}}^2}{2\tilde{\beta}^3} \right) \times 10^{-1} \right) dt}\end{aligned} \tag{5.37}$$

Based on the theorems in Chapter 3, having 10^{-1} as a constant weighting factor between the extracted work and the dissipation, to maximize the efficiency η_2 is equivalent to the problem of finding an optimal protocol to minimize

$$J = \int_0^{t_f} \left(\frac{\dot{\tilde{c}}^2}{4g\tilde{c}\tilde{\beta}} + \frac{\dot{\tilde{c}}\dot{\tilde{\beta}}}{2g\tilde{\beta}^2} + \frac{\tilde{c}\dot{\tilde{\beta}}^2}{2g\tilde{\beta}^3} \right) dt \tag{5.38}$$

while the extracted mechanical work $\int_0^{t_f} \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}} dt$ is a prescribed value and $(\tilde{c}(0), \tilde{\beta}(0)) = (\tilde{c}(t_f), \tilde{\beta}(t_f))$. The corresponding sub-Riemannian structure is given by

$$\begin{pmatrix} \dot{\tilde{c}} \\ \dot{\tilde{\beta}} \\ \dot{\tilde{\psi}}_2 \end{pmatrix} = \tilde{u}_2 \begin{pmatrix} 2\sqrt{\tilde{c}\tilde{\beta}} \\ 0 \\ \sqrt{\frac{1}{\tilde{c}\tilde{\beta}}} \end{pmatrix} + \tilde{v}_2 \begin{pmatrix} -2\sqrt{\tilde{c}\tilde{\beta}} \\ 2\frac{\tilde{\beta}^{3/2}}{\tilde{c}^{1/2}} \\ -\sqrt{\frac{1}{\tilde{c}\tilde{\beta}}} \end{pmatrix} \tag{5.39}$$

where $(\tilde{u}_2, \tilde{v}_2)$ is the new control and scaled average dissipation rate is $\tilde{u}_2^2 + \tilde{v}_2^2$. Also $\tilde{\psi}_2$ is the scaled extracted mechanical power. Based on the Hamilton-Jacobi equation (5.23) of the resistor-capacitor circuit, the new Hamilton-Jacobi equation associated to above optimal control problem is given by

$$\frac{\partial \tilde{V}_2}{\partial t} + \sqrt{\left\langle \frac{\partial \tilde{V}_2}{\partial q_2}, \begin{pmatrix} 2\sqrt{\tilde{c}\tilde{\beta}} \\ 0 \\ \sqrt{\frac{1}{\tilde{c}\tilde{\beta}}} \end{pmatrix} \right\rangle_d^2 + \left\langle \frac{\partial \tilde{V}_2}{\partial q_2}, \begin{pmatrix} -2\sqrt{\tilde{c}\tilde{\beta}} \\ 2\tilde{\beta}^{3/2}\sqrt{\frac{1}{\tilde{c}}} \\ -\sqrt{\frac{1}{\tilde{c}\tilde{\beta}}} \end{pmatrix} \right\rangle_d^2} = 0 \tag{5.40}$$

Having the $(\tilde{c}(0), \tilde{\beta}(0), \tilde{\psi}_2(0)) = (6, 0.34, 0)$, up to the arc-length of 0.5, the reachable set of the resistor-capacitor circuit in 3-D is given by figure 5.5,

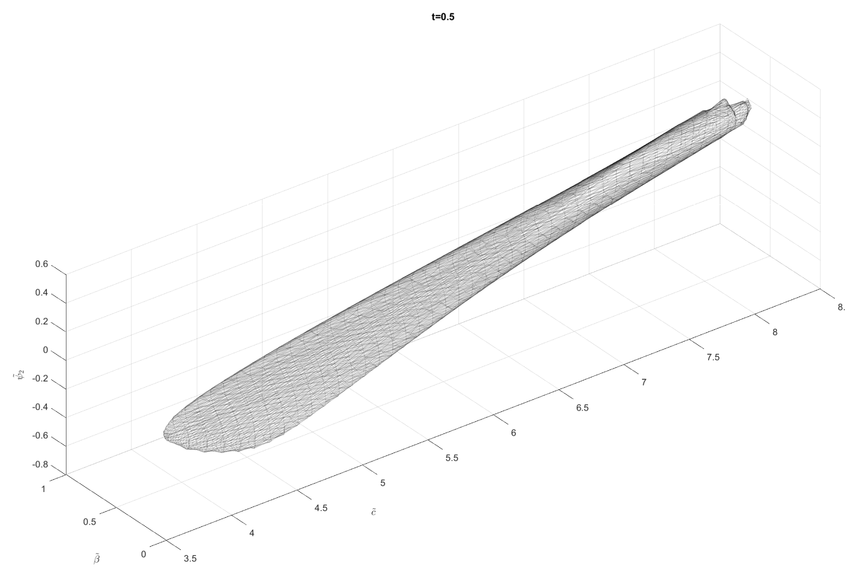

Figure 5.5: Reachable set of the RC-circuit in 3D

The view of the reachable set through $\tilde{c}$-axis is given by figure 5.6

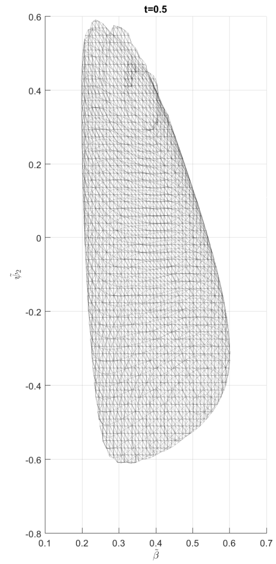

Figure 5.6: Reachable set in the space of $(\tilde{\beta}, \tilde{\psi}_2)$

The view of the reachable set through $\tilde{\beta}$-axis is given by figure 5.7

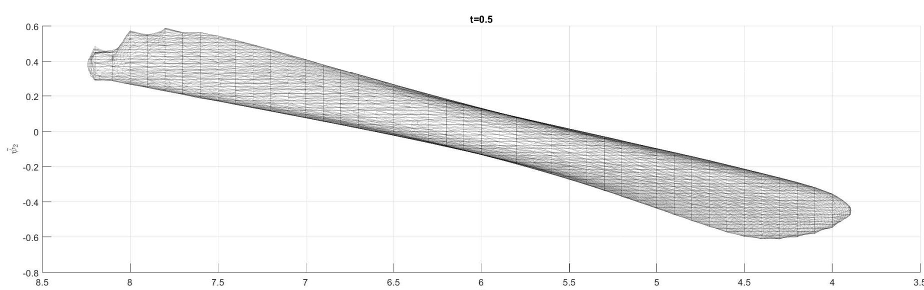

Figure 5.7: Reachable set in the space of $(\tilde{c}, \tilde{\psi}_2)$

The view of the reachable set through $\tilde{\psi}_2$-axis is given by figure 5.8

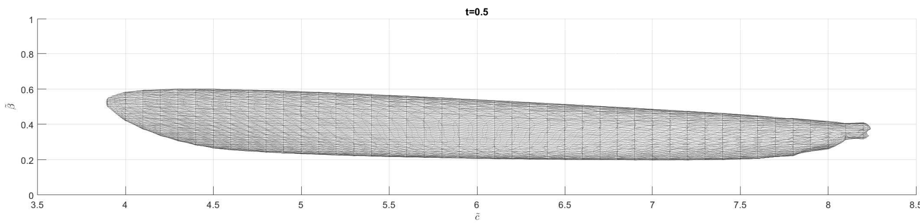

Figure 5.8: Reachable set in the space of $(\tilde{c}, \tilde{\beta})$

For the resistor-capacitor circuit, we only acquire one data point whose $\tilde{\psi}_2$ coordinate is positive and the information on the distance of that data point and its $\tilde{\psi}_2$ coordinate is listed in tabel 5.2.

Table 5.2: Working loop end-points of the resistor-capacitor circuit

Point number	distance	$\tilde{\psi}_2 - coordinate$
1	1.7182	0.1353

Comparing with the information collected for the stochastic oscillator which has 10 data points, the scarcity of data point of the circuit is due to the sub-Riemannian structure. Comparing $\dot{\psi}_1$ and $\dot{\psi}_2$, with $\tilde{x}$, $\tilde{y}$, $\tilde{c}$, $\tilde{\beta}$ in the scale $O(1)$, ψ_2 grows much slower than ψ_1 with the increment of the distance along the geodesics. Inside the given ranges of $\tilde{x}$, $\tilde{y}$, $\tilde{c}$ and $\tilde{\beta}$, we will collect less desired data points for the resistor-capacitor circuit.

5.2 Trajectory reconstruction

For the stochastic oscillator, from Chapter 3, the necessary condition for a unit-speed regular extremal in the space of $(\tilde{x}, \tilde{y}, \tilde{\psi}_1)$ is given by

$$\begin{aligned} \dot{\tilde{x}} &= \sqrt{\tilde{x}}\cos\phi_1,\ \dot{\tilde{y}} = \sqrt{\tilde{x}}\sin\phi_1 \\ \dot{\phi}_1 &= \frac{1}{2\sqrt{\tilde{x}}}\cos\phi_1 - \frac{\tilde{y}}{\tilde{x}}p_{1,3} \end{aligned} \tag{5.41}$$

where $p_{1,3}$ is a constant. The complete information $(\tilde{x}(t), \tilde{y}(t), \tilde{\psi}_1(t))$ of the regular extremal can be recovered from above reduced dynamics by quadrature. Given the time span t_f of the extremal and the number of steps N (we use $N = 75$ in our thesis work) used in the mid-point approximation, the step size $h = \frac{t_f}{N}$ and the reduced dynamics (5.41) is approximated as

$$\begin{aligned} \frac{\tilde{x}^{i+1} - \tilde{x}^i}{h} &= \sqrt{\frac{\tilde{x}^{i+1} + \tilde{x}^i}{2}}\cos\left(\frac{\phi_1^{i+1} + \phi_1^i}{2}\right) \\ \frac{\tilde{y}^{i+1} - \tilde{y}^i}{h} &= \sqrt{\frac{\tilde{x}^{i+1} + \tilde{x}^i}{2}}\sin\left(\frac{\phi_1^{i+1} + \phi_1^i}{2}\right) \\ \frac{\phi_1^{i+1} - \phi_1^i}{h} &= \frac{1}{\sqrt{2(\tilde{x}^{i+1} + \tilde{x}^i)}}\cos\left(\frac{\phi_1^{i+1} + \phi_1^i}{2}\right) - \left(\frac{\tilde{y}^{i+1} + \tilde{y}^i}{\tilde{x}^{i+1} + \tilde{x}^i}\right)p_{1,3} \end{aligned} \tag{5.42}$$

where $\tilde{x}^i, \tilde{y}^i, \phi^i$ are the values of x, y, ϕ at time ih $(i = 0, 1, \ldots, N)$. There are x^0, …, x^N, y^0, …, y^N, ϕ^0, …, ϕ^N and $p_{1,3}$, in total $(3N + 4)$ variables. Based on the definition of a working loop

$$\tilde{x}^0 = \tilde{x}^N = 4,\ \tilde{y}^0 = \tilde{y}^N = 4, \tag{5.43}$$

There are $3N$ unknowns, $X = \left(\tilde{x}^0, \ldots, \tilde{x}^{N-1}, \tilde{y}^0, \ldots, \tilde{y}^{N-1}, \phi^0, \ldots, \phi^N, p_{1,3}\right)^{\mathsf{T}}$ in the $3N$ equations (5.42). The problem can be numerically solved by *Newton-Raphson* *method*. In the solution, we pick up the values of $\phi_1(0)$ and $p_{1,3}$ and bring them to (5.41) with $(\tilde{x}(0), \tilde{y}(0)) = (4, 4)$. Thus, the ordinary differential equation (5.41) from time 0 to time t_f can be numerically solved. Using *shooting method*, by varying the initial value of ϕ_1 around the solution from mid-point approximation, the differences between $(\tilde{x}(0), \tilde{y}(0))$ and $(\tilde{x}(t_f), \tilde{y}(t_f))$ is minimized to have the numerical solution to be close to a working loop.

For the resistor-capacitor circuit, the necessary condition for a unit-speed maximum-efficiency working loop is given by

$$\begin{aligned}
\dot{\tilde{c}} &= 2\sqrt{\tilde{c}\tilde{\beta}}\cos\phi_2 - 2\sqrt{\tilde{c}\tilde{\beta}}\sin\phi_2 \\
\dot{\tilde{\beta}} &= 2\sqrt{\frac{1}{\tilde{c}}}\tilde{\beta}^{3/2}\sin\phi_2 \\
\dot{\phi}_2 &= -2\sqrt{\frac{\tilde{\beta}}{\tilde{c}}}\cos\phi_2 - \sqrt{\frac{\tilde{\beta}}{\tilde{c}}}\sin\phi_2 + \frac{2}{\tilde{c}}p_{2,3}
\end{aligned} \tag{5.44}$$

where $p_{2,3}$ is a constant. By applying mid-point approximation and shooting method as shown in the case of the stochastic oscillator, we can also have the numerical solution of the working loop for the resistor-capacitor circuit.

5.2.1 Numerical results

In this subsection, we display the numerical results after each step (i.e. mid-point approximation and shooting method). For the stochastic oscillator, after the mid-point approximation, we have the initial values of ϕ_1s and the values of $p_{1,3}$s of different working loops in table 5.3

As an example, for the working loop with length of 2.9525, the mid-point approxi-

Table 5.3: Reconstructed working loops through mid-point approximation

Point number	distance	$\tilde{\psi}_1 - coordinate$	$\phi_1(0)$	$p_{1,3}$
1	1.1126	0.1207	10.0348	4.6472
2	1.5302	0.1991	10.0607	3.1298
3	1.8331	0.2775	10.0534	2.4688
4	2.0515	0.3560	10.0337	2.1159
5	2.2134	0.4344	10.0095	1.9005
6	2.3471	0.5128	16.2691	1.7457
7	2.4982	0.5913	9.9729	1.5918
8	2.6546	0.6697	9.9652	1.4519
9	2.8082	0.7481	9.9595	1.3306
10	2.9525	0.8266	9.9537	1.2287

mation result is given by figure 5.9.

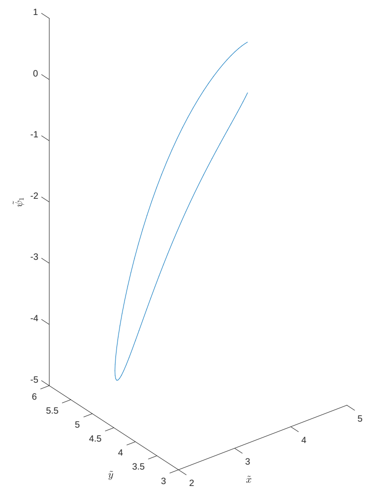

Figure 5.9: Reconstructed working loop through mid-point approximation

Based on the information from the trajectory reconstruction of the stochastic oscillator, we can implement shooting method, by which, we try to minimize the differences between the starting points and ending points of the different working loops in parameter space of $(\tilde{x}, \tilde{y})$ which are measured separately by Euclidean distances as shown in table 5.4.

As an example, the numerical result of the working loop with length of 2.9525 and extracted mechanical work of 0.8266 after the shooting method in 3D is given by figure 5.10. The projection of the working loop in the space of $(\tilde{x}, \tilde{y})$ is given by figure 5.11.

Table 5.4: Reconstructed working loops through shooting method

Point number	Difference in $\tilde{x}$	Difference in $\tilde{y}$	Extracted work	Heat supply
1	0.0058	0.0026	0.1207	0.8319
2	0.0184	0.0051	0.1991	1.2085
3	0.0441	0.0088	0.2775	1.5055
4	0.05999	0.0086	0.3560	1.7363
5	0.0637	0.0063	0.4344	1.9179
6	0.069	0.0038	0.5128	2.0771
7	0.0730	0.0002	0.5913	2.2568
8	0.0964	0.0006	0.6697	2.4464
9	0.1275	0.0052	0.7481	2.6362
10	0.1611	0.0114	0.8266	2.8185

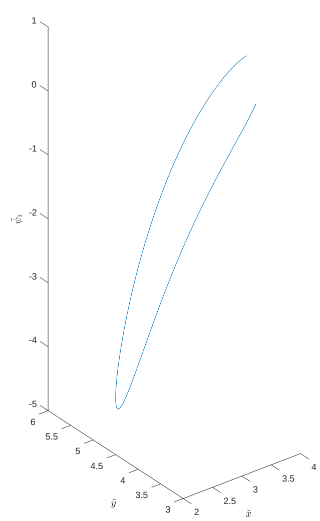

Figure 5.10: Reconstructed working loop through shooting method

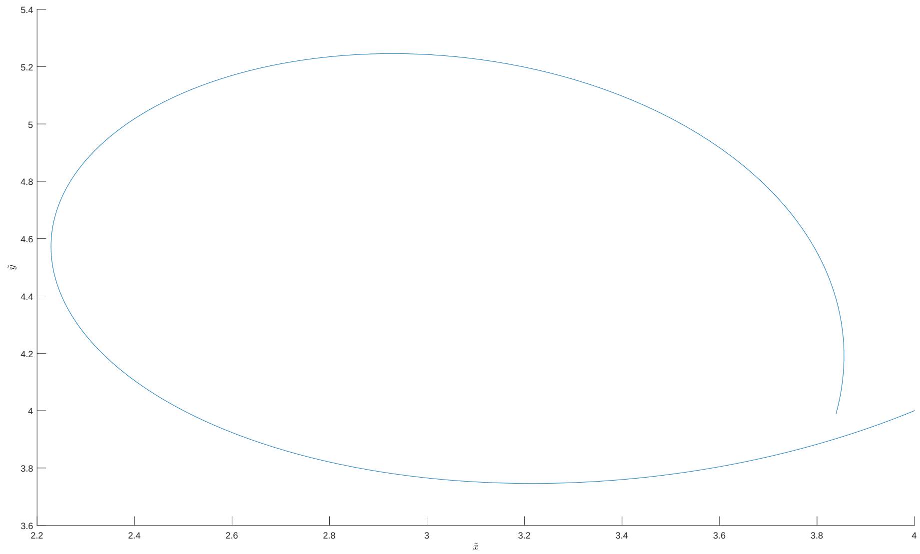

Figure 5.11: Projection of reconstructed working loop of length 2.9525

The difference between $(\tilde{x}(0), \tilde{y}(0))$ and $(\tilde{x}(t_f), \tilde{y}(t_f))$ in the figure 5.11 is noticeable. One reason is that the step-size in mid-point approximation is increasing with the length of the loop (i.e. t_f). The working loop of the shortest distance ($t_f = 1.1126$) provides the smallest error, which is given by figure 5.12. Along $\tilde{x}$ direction, the ratio of $|\tilde{x}(0) - \tilde{x}(t_f)|$ to the diameter of the working loop (i.e. the biggest difference along the working loop in $\tilde{x}$ coordinate) is smaller than 0.01 and so is the ratio between $|\tilde{y}(0) - \tilde{y}(t_f)|$ and the diameter of the loop in $\tilde{y}$ direction.

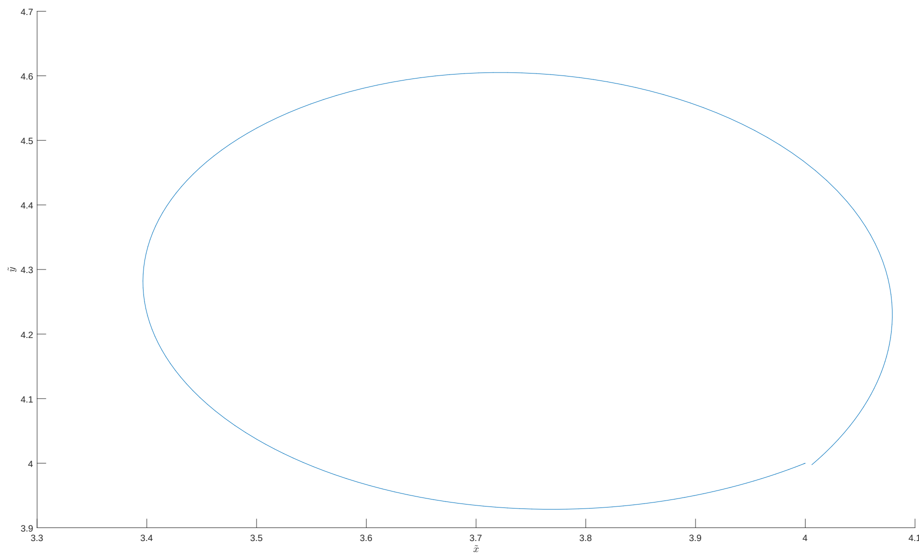

Figure 5.12: Projection of reconstructed working loop of length 1.1126

At time t, referring to the coordinate $(\tilde{x}(t), \tilde{y}(t))$ along the working loop of length 1.1126 at unit speed, the perfect working loop is approximated with an error in the scale of $(0.01\tilde{x}(t), 0.01\tilde{y}(t))$. The associated efficiency is as follows:

$$\eta_1 = \frac{\int_0^{t_f} \left(\frac{\tilde{y}^2}{\tilde{x}^2}\dot{\tilde{x}} - \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}} \right) dt}{\int_0^{t_f} \left(\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\mathbb{1}\left\{ \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}} > 0 \right\} + \frac{m}{\zeta}\left(\frac{\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2}{\tilde{x}} \right) \right) dt} \tag{5.45}$$

It turns out that the error of the extracted work is in the scale of $0.01\psi_1(t_f)$ and the error of the total heat supply is in the scale of $0.01 \int_0^{t_f} \left(\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\mathbb{1}\left\{ \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}} > 0 \right\} + \frac{m}{\zeta}\left(\frac{\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2}{\tilde{x}} \right) \right) dt$. Referring to the associated perfect working loop of length 1.126, the efficiency error of our reconstructed working loop is in the order of 0.001. Hence, the accuracy of efficiency associated to our reconstructed working loop is up to 0.01.

To reduce the errors of our reconstructed working loops, there are several procedures to improve our algorithms. On the level set methods, we may have denser mesh grid in the space of $(\tilde{x}, \tilde{y}, \tilde{\psi}_1)$ for higher accuracy. On the mid-point approximation, the error could be reduced by increasing the step number and on the shooting method, it is possible to vary $p_{i,3}$ to gain higher accuracy.

For the resistor-capacitor circuit, after the mid-point approximation, we have the information of the $\phi_2(0)$ and $p_{2,3}$ and the reconstructed working loop in table 5.5 and figure 5.13.

Table 5.5: Reconstructed working loop through mid-point approximation

Point number	distance	$\tilde{\psi}_2 - coordinate$	$\phi_2(0)$	$p_{2,3}$
1	1.7182	0.1353	9.4998	-9.1532

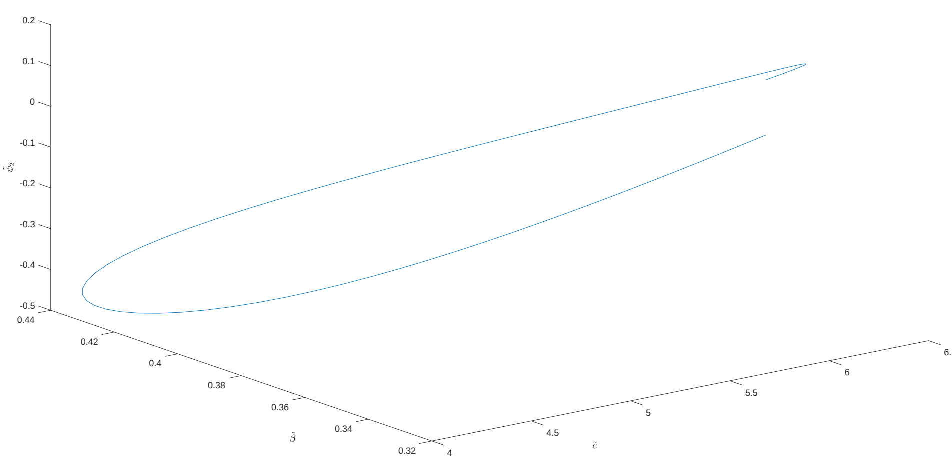

Figure 5.13: Reconstructed working loop through mid-point approximation

Similar to the case of the stochastic oscillator, based on above information, we implement shooting method to minimize the differences between the starting point and ending point of the working loop which are measured separately by Euclidean distance.

Table 5.6: Reconstructed working loop through shooting method

Point number	Difference in $\tilde{c}$	Difference in $\tilde{\beta}$	Extracted work	Heat supply
1	0.9741	0.0294	0.1353	0.9765

The plot of the working loop after shooting method is given by figure 5.14. The projection of the working loop is given by figure 5.15.

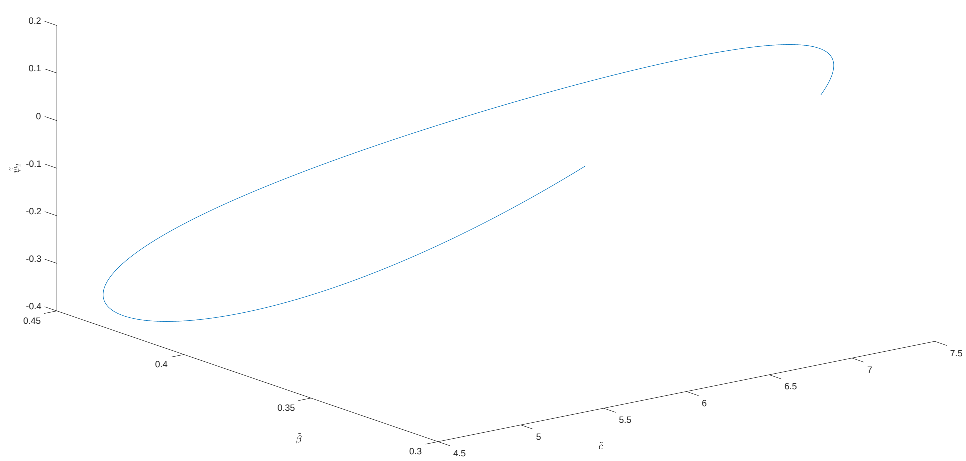

Figure 5.14: Reconstructed working loop through shooting method

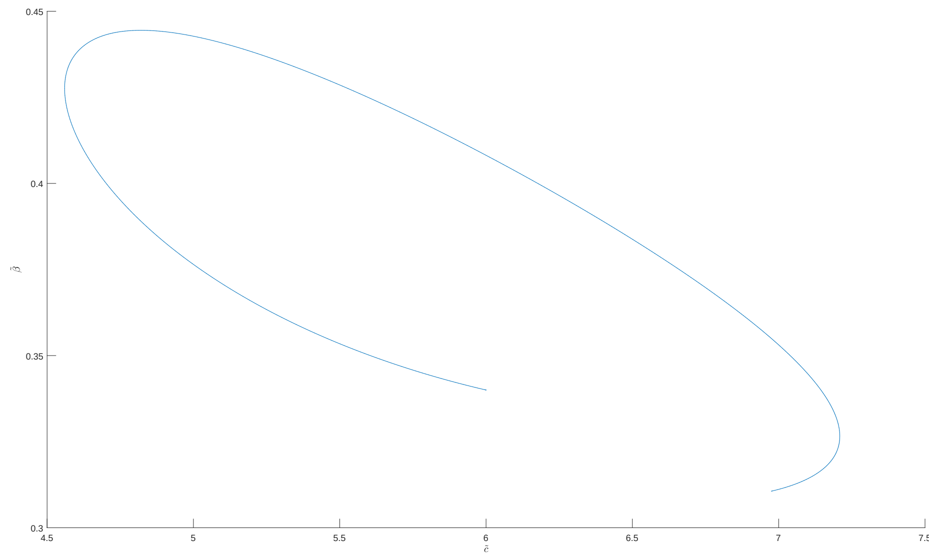

Figure 5.15: Projection of reconstructed working loop

Remark 5.2.1. *With all the sample points collected from both models, after implementing shooting method to reconstruct the associated working loops, it is shown that phase angle* ϕ_i *is monotonically decreasing with time along each loop and the phase angle difference between the end-points are smaller than* 2π*. Based on the conjugate point analysis in Chapter 3, it turns out that there is no conjugate point along each working loop and these solutions are optimal.*

5.3 Efficiencies of the heat engines

By defining η_1 and η_2 in (5.30) and (5.37),

$$
\begin{aligned}
\eta_1 &= \frac{\int_0^{t_f} \left(\frac{\tilde{y}^2}{\tilde{x}^2}\dot{\tilde{x}} - \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\right) dt}{\int_0^{t_f} \left(\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\mathbb{1}\left\{\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}} > 0\right\} + \frac{m}{\zeta}\left(\frac{\dot{\tilde{x}}^2+\dot{\tilde{y}}^2}{\tilde{x}}\right)\right) dt} \\
\eta_2 &= \frac{\int_0^{t_f} \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}} dt}{\int_0^{t_f} \left(\left(-\frac{\dot{\tilde{\beta}}}{2\tilde{\beta}^2} + \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}}\right)\mathbb{1}\left\{-\frac{\dot{\tilde{\beta}}}{2\tilde{\beta}^2} + \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}} > 0\right\} + \left(\frac{\dot{\tilde{c}}^2}{4\tilde{c}\tilde{\beta}} + \frac{\dot{\tilde{c}}\dot{\tilde{\beta}}}{2\tilde{\beta}^2} + \frac{\tilde{c}\dot{\tilde{\beta}}^2}{2\tilde{\beta}^3}\right) \times 10^{-1}\right) dt}
\end{aligned}
\tag{5.46}
$$

For a heat engine, the total heat supply is the sum of the heat supply from the heat engine and the dissipation along the working loop. Based on (5.6), along the maximum efficiency working loop, the dissipation along the loop numerically equals to product of the distance of the loop and the speed of the protocol ν_i. Based on (5.46), along a maximum efficiency working loop of a stochastic oscillator

$$
\begin{aligned}
\eta_1 &= \frac{\int_0^{t_f} \left(\frac{\tilde{y}^2}{\tilde{x}^2}\dot{\tilde{x}} - \frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\right) dt}{\int_0^{t_f} \left(\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}}\mathbb{1}\left\{\frac{\tilde{y}}{\tilde{x}}\dot{\tilde{y}} > 0\right\} + \frac{m}{\zeta}\left(\frac{\dot{\tilde{x}}^2+\dot{\tilde{y}}^2}{\tilde{x}}\right)\right) dt} \\
&= \frac{\int_0^{t_f} \left(\frac{\tilde{y}^2}{\tilde{x}^{3/2}}\tilde{u}_1^* - \frac{\tilde{y}}{\sqrt{\tilde{x}}}\tilde{v}_1^*\right) dt}{\int_0^{t_f} \frac{\tilde{y}}{\sqrt{\tilde{x}}}\tilde{v}_1^*\mathbb{1}\left\{\frac{\tilde{y}}{\sqrt{\tilde{x}}}\tilde{v}_1^* > 0\right\} dt + \frac{m}{\zeta}\int_0^{t_f} \left((\tilde{u}_1^*)^2 + (\tilde{v}_1^*)^2\right) dt}
\end{aligned}
\tag{5.47}
$$

If the unit-speed optimal control is $(\tilde{u}_{1,0}^*, \tilde{v}_{1,0}^*)$ with $t_{f,0}$, any other optimal control can be expressed as $(\nu_1\tilde{u}_{1,0}^*, \nu_1\tilde{v}_{1,0}^*)$ with ${\nu_1}^{-1}t_{f,0}$.

$$
\begin{aligned}
\eta_1 &= \frac{\int_0^{\nu_1^{-1}t_{f,0}} \left(\frac{\tilde{y}^2}{\tilde{x}^{3/2}}\nu_1\tilde{u}_{1,0}^* - \frac{\tilde{y}}{\sqrt{\tilde{x}}}\nu_1\tilde{v}_{1,0}^*\right) dt}{\int_0^{\nu_1^{-1}t_{f,0}} \left(\frac{\tilde{y}}{\sqrt{\tilde{x}}}\nu_1\tilde{v}_{1,0}^*\mathbb{1}\left\{\frac{\tilde{y}}{\sqrt{\tilde{x}}}\nu_1\tilde{v}_{1,0}^* > 0\right\} + \frac{m}{\zeta}\nu_1^2\left((\tilde{u}_1^*)^2 + (\tilde{v}_1^*)^2\right)\right) dt} \\
&= \frac{\int_0^{t_{f,0}} \left(\frac{\tilde{y}^2}{\tilde{x}^{3/2}}\tilde{u}_{1,0}^* - \frac{\tilde{y}}{\sqrt{\tilde{x}}}\tilde{v}_{1,0}^*\right) dt}{\int_0^{t_{f,0}} \frac{\tilde{y}}{\sqrt{\tilde{x}}}\tilde{v}_{1,0}^*\mathbb{1}\left\{\frac{\tilde{y}}{\sqrt{\tilde{x}}}\tilde{v}_{1,0}^* > 0\right\} dt + \frac{m}{\zeta}\nu_1 t_{f,0}}
\end{aligned}
\tag{5.48}
$$

To operate the system in unit speed, $\nu_1 = 1$, in the example of the stochastic oscillator, $\frac{m}{\zeta} = 0.1$,

$$\eta_1 = \frac{\int_0^{t_{f,0}} \left(\frac{\tilde{y}^2}{\tilde{x}^{3/2}} \tilde{u}_{1,0}^* - \frac{\tilde{y}}{\sqrt{\tilde{x}}} \tilde{v}_{1,0}^* \right) dt}{\int_0^{t_{f,0}} \frac{\tilde{y}}{\sqrt{\tilde{x}}} \tilde{v}_{1,0}^* \mathbb{1} \left\{ \frac{\tilde{y}}{\sqrt{\tilde{x}}} \tilde{v}_{1,0}^* > 0 \right\} dt + 0.1 t_{f,0}} \tag{5.49}$$

Similarly, for the resistor-capacitor circuit,

$$\begin{aligned} \eta_2 &= \frac{\int_0^{t_f} \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}} dt}{\int_0^{t_f} \left(\left(-\frac{\dot{\tilde{\beta}}}{2\tilde{\beta}^2} + \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}} \right) \mathbb{1} \left\{ -\frac{\dot{\tilde{\beta}}}{2\tilde{\beta}^2} + \frac{\dot{\tilde{c}}}{2\tilde{c}\tilde{\beta}} > 0 \right\} + \left(\frac{\dot{\tilde{c}}^2}{4\tilde{c}\tilde{\beta}} + \frac{\dot{\tilde{c}}\dot{\tilde{\beta}}}{2\tilde{\beta}^2} + \frac{\tilde{c}\dot{\tilde{\beta}}^2}{2\tilde{\beta}^3} \right) \times 10^{-1} \right) dt} \\ &= \frac{\int_0^{t_{f,0}} \frac{\dot{\tilde{c}}^*}{2\tilde{c}^* \tilde{\beta}^*} dt}{\int_0^{t_{f,0}} \left(-\frac{\dot{\tilde{\beta}}^*}{2(\tilde{\beta}^*)^2} + \frac{\dot{\tilde{c}}^*}{2\tilde{c}^*\tilde{\beta}^*} \right) \mathbb{1} \left\{ -\frac{\dot{\tilde{\beta}}^*}{2(\tilde{\beta}^*)^2} + \frac{\dot{\tilde{c}}^*}{2\tilde{c}^*\tilde{\beta}^*} > 0 \right\} + 0.1 t_{f,0}} \end{aligned} \tag{5.50}$$

where $t_{f,0}$, $\tilde{c}^*$ and $\tilde{\beta}^*$ are associated to the unit-speed maximum efficiency working loop of the resistor-capacitor circuit. Thus, the efficiencies of the stochastic oscillator along different working loops are listed in table 5.7.

The efficiency of the circuit along a working loop is listed in table 5.8.

Remark 5.3.1. *By comparing the efficiencies of the stochastic oscillator along different loops, it is seen that, the more energy the system dissipates into the environment, the more mechanical work it can extract and the efficiently of the working loop is higher.*

Remark 5.3.2. *Based on the equations* (5.48) *and* (5.50) *of heat engine efficiencies, in the near-equilibrium regime, it is shown that the extracted mechanical work and the heat supply along a working loop is independent of the speed ν_i and the dissipation is proportional to it. Hence, the faster the heat engine is operated along a maximum efficiency working loop, the higher mechanical power which is the ratio of the extracted mechanical work per loop to the time span of the working loop, the engine will produce. Meanwhile, the efficiency of the engine will be lower.*

Table 5.7: Efficiencies of maximum efficiency working loops

Point number	Extracted work	Heat supply	Dissipation	η_1
1	0.1207	0.8319	1.1126	0.1280
2	0.1991	1.2085	1.5302	0.1462
3	0.2775	1.5055	1.8331	0.1643
4	0.3560	1.7363	2.0515	0.1834
5	0.4344	1.9179	2.2134	0.2031
6	0.5128	2.0771	2.3471	0.2218
7	0.5913	2.2568	2.4982	0.2359
8	0.6697	2.4464	2.6546	0.2470
9	0.7481	2.6362	2.8082	0.2565
10	0.8266	2.8185	2.9525	0.2655

Table 5.8: Efficiency of maximum efficiency working loop

Point number	Extracted work	Heat supply	Dissipation	η_2
1	0.1353	0.9765	1.7182	0.1178

Chapter 6: Conclusions and Directions for Future Research

This thesis, using the stochastic oscillator and resistor-capacitor circuit as examples, built a connection between geometric control theory and non-equilibrium statistical mechanics. This was done by first showing that in a heat engine, extraction of work from thermal fluctuations is associated to a sub-Riemannian structure derived from the average-dissipation-rate construct in statistical physics. Extracting work efficiently is a constrained optimal control problem – minimize dissipation for prescribed work in a loop. Recognizing that in the case of the stochastic oscillator, this constraint is an area-moment over a loop suggested investigating the isoperimetric problem on Poincaré upper half plane by optimal control methods as a preparation. Eventually, following the theoretical analysis, we designed maximum efficiency working loops for both heat engines numerically and associated efficiencies were computed.

Our work only touched upon a small class of control problems in non-equilibrium statistical mechanics. For example, in both heat engines, the protocols are open-loop. There are physics experiments and theoretical analyses showing that with feedback, a system can extract work from a single heat bath of a constant temperature [33] [34]. The framework of fluctuation theorem and stochastic thermodynamics used in our thesis may be applicable to systems with feedback. Also, the framework of fluctuation theorem and stochastic thermodynamics can be adapted to investigate far-from equilibrium behaviors of a system, which turns out to be related to optimal transport problem [35]. To tackle these challenging problems in the future requires more advanced understanding and exploration in control theory, optimization and numerical methods.

Bibliography

[1] C. Jarzynski, "Nonequilibrium equality for free energy differences", *PRL 78, 2690(1997)*

[2] G. Crooks, "Entropy production fluctuation theorem and the non-equilibrium work relation for free energy differences", *Physical Review E **60**, 2721(1999).*

[3] U. Seifert, *Lecture Notes: 'Soft Matter. From Synthetic to Biological Materials'*, (39th IFF Spring School, Institute of Solid State Research, Research Center Julich, 2008)

[4] R. Edgar, *Irreversibility and dissipation in microscopic systems* (Springer International Publishing, 2014)

[5] S. Chandrasekhar, "Chapter I: the laws of thermodynamics," *An introduction to the study of stellar structure*, (University of Chicago Press, 1936)

[6] R. W. Brockett and J. C. Willems, "Stochastic control and the second law of thermodynamics", *Proceedings of the IEEE Conference on Decision and Control, San Diego, California*, pp.1007-1011, 1978

[7] R. W. Brockett, "Control of stochastic ensembles", *The Åström Symposium on Control (eds: Björn Wittenmark, Anders Rantzer)*, pp.199-216, Studentlitteratur, 1999

[8] S. K. Mitter and N. J. Newton, "Information and entropy flow in the Kalman–Bucy filter", *Journal of Statistical Physics*, vol.118, pp.145-176, 2005

[9] H. Nyquist, "Thermal Agitation of Electric Charge in Conductors", *Phys. Rev.*, vol. 32, 110 1928

[10] J. Johnson, “Thermal Agitation of Electricity in Conductors”, *Phys. Rev.*, vol. 32, 97, 1928

[11] D. A. Sivak and G. E. Crooks, “Thermodynamic metric and optimal paths”, *PRL 108, 190602(2012)*

[12] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, New York, 2001)

[13] P. R. Zulkowski, D. A. Sivak, G. E. Crooks and M. R. DeWeese “Geometry of thermodynamic control,” *Physical Review E **86**, 041148 (2012)*

[14] R. Montgomery, *A Tour of Subriemannian geometries, Their Geodesics and Applications* (American Mathematical Society, 2002)

[15] D. Liberzon, *Calculus of variations and optimal control theory*, Princeton University Press, 2012

[16] J. C. Willems,“Dissipative dynamical systems part I: General theory,” *Arch. Rational Mech. Anal.*, vol. 45, pp.321-351 1972

[17] B. Øksendal, *Stochastic differential equations: an introduction to applications* (Fifth edition, Springer-Verlag, 2000)

[18] R. K. Pathria and P. D. Beale, *Statistical mechanics* (Third edition, Elsevier, Burlington MA, 2011)

[19] P. R. Zulkowski, *The geometry of thermodynamic control* (PhD thesis, 2014)

[20] A. van der Schaft, “Port-Hamiltonian systems: an introductory survey”, *International Congress of Mathematicians*, pp. 1339-1365, 2006

[21] M. P. do Carmo, *Riemannian Geometry* (Birkhäuser, Boston, 1992)

[22] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems* (Springer-Verlag, New York, 1985)

[23] D. Mandal and C. Jarzynski, “Analysis of slow transitions between nonequilibrium steady states”, *Journal of Statistical Mechanics: Theory and Experiment*, June 2016

[24] V. Jurdjevic *Geometric control theory* (Cambridge, UK: Cambridge University Press) 1997

[25] A. Agrachev and Y. Sachkov, *Control theory from the geometric viewpoint*, (Springer, German) 2004

[26] A. Agrachev, D. Barilari, U. Boscain *Introduction to Riemannian and Sub-Riemannian geometry* Preprint 2014

[27] P. S. Krishnaprasad , "Optimal control and poisson reduction, " *ISR Technical Research Report, T.R. 93-87, University of Maryland*

[28] A. M. Bloch *Nonholonomic mechanics and Control* (New York, U.S.A: Springer-Verlag) 2003

[29] C. Adams and F. Morgan, "Isoperimetric curves on hyperbolic surfaces, " *Proceedings of the American Mathematical Society*, vol. 127, no. 5, pp.1347-1356, 1999

[30] S. Osher and R. Fedkiw, *Level set methods and dynamic implicit surfaces* Applied Mathematical Sciences Vol. 153, Springer, 2002

[31] I. Mitchell, "The flexible, extensible and efficient toolbox of level set methods," *Journal of Scientific Computing*, vol. 35, no. 2-3, pp. 300-329 (June 2008)

[32] S. Osher, "A level set formulation for the solution of the Dirichlet problem for Hamiltonian Jacobi equations, " *SIAM Journal of Mathematical Analysis*, vol 24, no. 5, pp. 1145-1152, 1993

[33] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki and M. Sano, "Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality," *Nature Physics* **6**, 988-9920 (2010)

[34] T. Sagawa and M. Ueda, "Generalized Jarzynksi Equality under Nonequilibrium Feedback Control," *Phys. Rev. Lett.* **104**, 090602 (2010)

[35] E. Aurell, C. Mejia-Monasterio and P. Muratore-Ginanneschi, "Optimal protocols and optimal transport in stochastic thermodynamics," *Phys. Rev. Lett* **106** 250601 (2011)