THE EYRING-KRAMERS FORMULA

FOR

POINCARÉ

AND

LOGARITHMIC SOBOLEV INEQUALITIES

Von der Fakultät für Mathematik und Informatik
der Universität Leipzig
angenommene

DISSERTATION

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM
(DR. RER. NAT.)

im Fachgebiet

Mathematik

vorgelegt

von DIPL. MATH. ANDRÉ SCHLICHTING

geboren am 19. August 1985 in Räckelwitz

Die Annahme der Dissertation wurde empfohlen von:

1. PROF. DR. STEPHAN LUCKHAUS
   Universität Leipzig
2. PROF. DR. ANTON BOVIER
   Universität Bonn

Die Verleihung des akademischen Grades erfolgt mit Bestehen der Verteidigung am 25.10.2012 mit dem Gesamtprädikat *summa cum laude*.

Acknowledgment

> The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them.
>
> *(Sir William Henry Bragg (1862-1942))*

First and foremost, I want to thank my supervisor Prof. Dr. Stephan Luckhaus. On the one hand, he gave me the chance to pursue my own ideas. On the other hand, he provided me with important mathematical impulses to "push me into the right direction" at the right time. I am extremely grateful to him for his trust and support.

A special thanks also goes also to Prof. Dr. Felix Otto who has established a stimulating atmosphere in the broad field of analysis at the Max Planck Institute, where I had the opportunity to discuss mathematical questions with him. It was he who drew the low temperature regime to my attention.

I cannot find appropriate words to express my gratitude to Dr. Georg Menz who has been available for me at any time. I enjoyed our fruitful, long-lasting discussions and his introduction to the area of functional inequalities. He also improved the presentation of this work and carefully proof-read the manuscript.

This dissertation would not have been possible unless the financial support of the IMPRS of the Max Planck Institute, the Research Academy Leipzig and the DFG project FOR 718 (Analysis and stochastics in complex physical systems). I want to especially give thanks to Prof. Dr. Wolfgang König who gave me the opportunity to participate in an intense weekend retreat of the FOR 718 in April 2012. There, discussions with Prof. Dr. Max von Renesse and Dr. Martin Slowik helped me to clarify the connection of my work to large deviations as well as to the approach to metastability via the potential theoretic approach of Prof. Dr. Anton Bovier.

There are many more people who I met at several workshops, conferences, summer schools and during guest visits at our institute and who gave helpful comments on my work. It is a great pleasure to thank everyone who helped me write my dissertation successfully. Among them, I share credit to Prof. Dr. Nicola Gigli helping to clarify the connection of the weighted transport distance to the Wasserstein distance; Prof. Dr. Martin Hairer giving the stochastic interpretation of the local coordinates at the saddle points; Dr. Arnaud Guillin bringing the Lyapunov functions to my attention; and Prof. Dr. Luigi Ambrosio for his introduction to the field of optimal transportation.

I enjoyed working in an extremely stimulating atmosphere of the Max Planck Institute very much. Numerous discussions with Jens Wohlgemuth helped me to look at mathematical problems from a rather physical perspective. Moreover, cordial thanks go to my former office colleagues Roland Seydel, Ilona Kosiuk and Björn Zocher. The "coffee/tea group" at the M-section as well as the "Mensa group" broadened my field of view among many more areas in mathematics.

Besides the atmosphere at the institute, I found also perfect working conditions and a great library. I also appreciated the good relation toward the administrative staff being always on hand with help and advice especially for proposing applications for my business trips. Here, I want to thank Miss Rietz, Miss Heid, Miss Hünninger and Miss Vandenberg.

I am not exaggerating, if I say that surviving the last three years would be quite a challenge for me without my longstanding friends accompanying me since the very beginning of my studies. Here, I especially want to mention Svend, David, Frank, Matti and Kay – the "Freiberg-clique." During my time in Leipzig I have also met my new running and cycling friends: Sebastian, Nico, Martin and Daniel who always helped me to find the balance in sports to the mental effort of mathematics. I also enjoyed an uncountable number of runs with the "Unilaufgruppe" and the members of the "Lauf Prag-Leipzig" on the occasion of the 600^{th} anniversary of the University of Leipzig in 2009.

Moreover, I owe sincere and earnest thankfulness to my family, especially to my parents for their support and letting me go, not just this time, but so many times in my life. The last, and surely the most, I want to thank my fiancée Laura with whom I shared the best and worst moments of my dissertation journey and who always boosted me morally. Thank you for your never-ending love, patience and just for being with me.

Contents

Chapter 1

Introduction

Let us consider a diffusion on a potential landscape which is given by a sufficiently smooth *Hamiltonian function* $H : \mathbb{R}^n \to \mathbb{R}$. We are interested in the regime of small noise ε. The *generator* of the diffusion has the following form

$$L = \varepsilon\Delta - \nabla H \cdot \nabla. \tag{1.1}$$

The associated *Dirichlet form* is given by

$$\mathcal{E}(f) := \int (-Lf) f \mathrm{d}\mu = \varepsilon \int |\nabla f|^2 \,\mathrm{d}\mu. \tag{1.2}$$

The corresponding diffusion ξ_t satisfies the *stochastic differential equation*

$$\mathrm{d}\xi_t = -\nabla H(\xi_t)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}B_t, \tag{1.3}$$

where B_t is the *Brownian motion* on $\mathbb{R}^n$. The last equation is also called *over-damped Langevin equation* (cf. e.g. [LL10]). Under some growth assumption on H there exists an equilibrium measure of the according stochastic process, which is called *Gibbs measure* and is given by

$$\mu(\mathrm{d}x) = \frac{1}{Z_\mu} \exp\left(-\frac{H(x)}{\varepsilon}\right)\mathrm{d}x \qquad \text{with} \qquad Z_\mu = \int \exp\left(-\frac{H(x)}{\varepsilon}\right)\mathrm{d}x. \tag{1.4}$$

With the help of the Gibbs measure and the generator we can give an evolution equation for the density of the process given in (1.3). Therefore, let us assume that $\mathrm{law}(\xi_0)$ is absolutely continuous w.r.t. to μ. Then, we can find a non-negative $f_0 \in L^1(\mu)$ with $\mathrm{law}(\xi_0) = f_0\mu$ being the initial density of the process. We call f_0 the *relative initial density* of ξ_0. After time t the density of ξ_t is given by $\mathrm{law}(\xi_t) = f_t\mu$. The relative density f_t solves the evolution equation, also called *Fokker-Planck equation* (cf. e.g. [Øks03] or [Sch10])

$$\partial_t f_t = L f_t = \varepsilon\Delta f_t - \nabla H \cdot \nabla f_t. \tag{1.5}$$

The solution f_t can be expressed in terms of the *semigroup* $(P_t)_{t\geq 0}$ generated by the operator L. The semigroup is formally given by $P_t = e^{tL}$ and satisfies the relations $P_0 = \mathrm{Id}$, $f_t = P_t f_0$ and $P_\infty f_t = \int f_0 \, \mathrm{d}\mu$. The generator L and the semigroup are both invariant and symmetric with respect to the Gibbs measure μ.

We are particularly interested in cases where H has several local minima. Then, the process shows metastable behavior for small ε in the sense that there exists a separation of scales. On the fast scale, the process converges quickly to a neighborhood of a local minimum. On the slow scale, the process stays most of the time nearby a local minimum for an exponentially long waiting time after which it eventually jumps to another local minimum.

This behavior is well known in the context of chemical reactions. The exponential waiting time follows *Arrhenius' law* [Arr89] meaning that the mean exit time from one local minimum of H to another one is exponentially large in the energy barrier between them. By now, the Arrhenius law is well-understood even for non-gradient systems by the *Freidlin-Wentzell theory [FW98]*, which is based on *large deviations*.

A refinement of the Arrhenius law is the *Eyring-Kramers formula* which additionally considers *pre-exponential* factors. The Eyring-Kramers formula for the Poincaré inequality goes back to Eyring [Eyr35] and Kramers [Kra40]. Both argue that also in high-dimensional problems of chemical reactions most reactions are nearby a single trajectory called *reaction pathway*. Evaluating the Hamiltonian along this *reaction coordinate* gives the classical picture of a double well potential (cf. Figure 1.1) in one dimension with an *energy barrier* separating the two local minima for which explicit calculations are feasible.

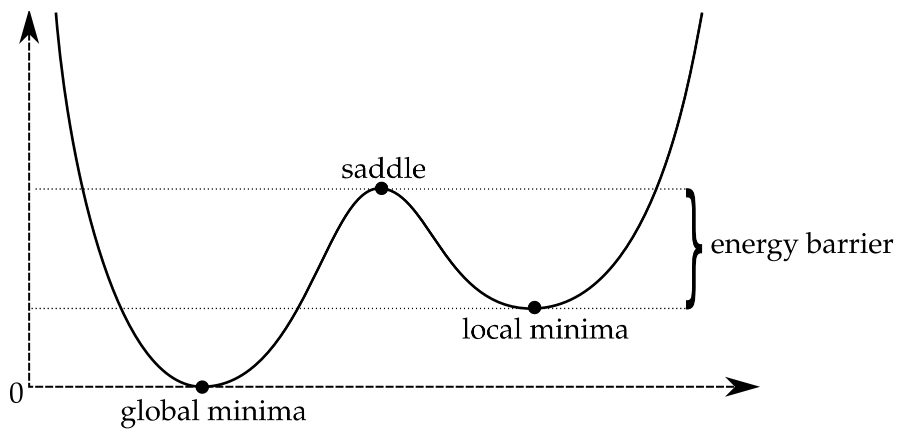

Figure 1.1.: General double-well potential H on $\mathbb{R}$.

However, a rigorous proof of the Eyring-Kramers formula for the multidimensional case was open for a long time. For a special case, where all the minima of the potential as well as all the lowest saddle points in-between have the same energy Sugiura [Sug95] defined an exponentially rescaled Markov chain on the set of minima in such a way that the pre-exponential factors become the transitions rates between the basins of the rescaled process. For the generic case Bovier, Eckhoff, Gayrard and Klein [BEGK04, BGK05] obtained first order asymptotics that are sharp in the parameter ε. They also clarified the close connection between *mean exit times*, *capacities* and the exponentially small eigenvalues of the operator L given by (1.1). The main tool of [BEGK04, BGK05] is potential theory. The small eigenvalues are related to the mean exit times of appropriate subsets of the state space. Further, the mean exit times are given by *Newtonian capacities* which can explicitly be calculated in the regime of small noise ε.

Shortly after, Helffer, Klein and Nier [HKN04, HN06, HN05] also deduced the Eyring-Kramers formula using the connection of the spectral gap estimate of the Fokker-Planck operator L given by (1.1) to the one of the *Witten Laplacian*. This approach makes it possible to get quantitative results with the help of semiclassical analysis. They deduced sharp asymptotics of the exponentially small eigenvalues of L and gave an explicit expansion in ε to theoretically any order. An overview on the Eyring-Kramers formula can be found in the review article of Berglund [Ber11].

On the one hand, this works aims to provide a new proof of the Eyring-Kramers formula for the first eigenvalue of the operator L, i.e. its *spectral gap*, and on the other hand, to extend the approach to the logarithmic Sobolev inequality, which was unknown before. We will refer to this as the Eyring-Kramers formula for Poincaré and logarithmic Sobolev inequality. Therefore, let us introduce, in the next section the Poincaré and logarithmic Sobolev inequality and explain why we are interested in the asymptotically optimal constants in these inequalities.

1.1. Poincaré and logarithmic Sobolev inequality

Definition 1.1 (PI(ϱ) and LSI(ϱ))**.** Let X be an Euclidean space. A Borel probability measure μ on X satisfies the *Poincaré inequality* with constant $\varrho > 0$, if for all test functions $f : X \to \mathbb{R}^+$

$$\operatorname{var}_\mu(f) := \int \left(f - \int f \, \mathrm{d}\mu \right)^2 \mathrm{d}\mu \le \frac{1}{\varrho} \int |\nabla f|^2 \, \mathrm{d}\mu. \qquad \mathrm{PI}(\varrho)$$

In a similar way, the probability measure μ satisfies the *logarithmic Sobolev inequality* with constant $\alpha > 0$, if for all test function $f : X \to \mathbb{R}^+$ holds

$$\operatorname{Ent}_\mu(f) := \int f \; \log f \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu \; \log \left(\int f \, \mathrm{d}\mu \right) \le \frac{1}{\alpha} \int \frac{|\nabla f|^2}{2f} \, \mathrm{d}\mu. \qquad \mathrm{LSI}(\alpha)$$

The gradient ∇ is determined by the Euclidean structure of X. Test functions are those functions for which the gradient exists and the right-hand-side in PI(ϱ) and LSI(α) is finite.

Remark 1.2 (Continuous and discrete spaces). The most common case in this work will be the continuous case $X = \mathbb{R}^n$, then the gradient ∇ is the canonical gradient identified as row vector of the partial derivatives

$$X = \mathbb{R}^n : \qquad \nabla f(x) = (\partial_1 f(x), \ldots, \partial_n f(x)).$$

For the discrete case, i.e. $X = (V, E)$ is an undirected graph or as special case $X = \mathbb{Z}^n$, holds

$$X = (V, E) : \qquad \nabla f(x) = \sum_{y \sim x} (f(y) - f(x)),$$

where $x \sim y$ means that x is adjacent to y, i.e. for $X = \mathbb{Z}^n$, that $|x - y| = 1$. A special case occurs if $X = \{0, 1\}$ is a two point space, then

$$X = \{0, 1\} : \qquad \nabla f(0) = -\nabla f(1) = f(1) - f(0)$$

Remark 1.3 (Equivalent formulations of LSI(α)). In the continuous case is LSI(α) equivalent by a change of variable $f \mapsto f^2$ to

$$\mathrm{Ent}_\mu(f^2) \leq \frac{2}{\alpha} \int |\nabla f|^2 \, \mathrm{d}\mu. \tag{1.6}$$

Note, that for a discrete state space with a discrete gradient this equivalence is no longer true. Besides the question, whether to discretize LSI(α) or (1.6), arises also the question how to discretize the *Fisher information* on the right-hand side of LSI(α), which has in the continuous setting the possible representations among others

$$I(f\mu|\mu) := \int \frac{|\nabla f|^2}{2f} \, \mathrm{d}\mu = \int \frac{f \, |\nabla f \cdot \nabla \log f|^2}{2} \, \mathrm{d}\mu = \int 2|\nabla \sqrt{f}\,|^2 \mathrm{d}\mu.$$

A more or less full investigation of the possible resulting discrete logarithmic Sobolev inequalities was done by Bobkov and Tetali [BT06].

Remark 1.4 (Relation between PI(ϱ) and LSI(α)). LSI(α) is stronger in the sense that it implies PI(ϱ). Therefore, we set $f = 1 + \eta g$ for η small and find

$$\mathrm{Ent}_\mu(f^2) = 2\eta^2 \, \mathrm{var}_\mu(g) + O(\eta^3) \quad \text{as well as} \quad \int |\nabla f|^2 \, \mathrm{d}\mu = \eta^2 \int |\nabla g|^2 \, \mathrm{d}\mu.$$

Hence, if μ satisfies LSI(α) then μ also satisfies PI(α), which always implies $\alpha \leq \varrho$.

The connection of the *Poincaré inequality* to the *spectral gap* of the operator L in (1.1) is the variational characterization of the latter one.

Lemma 1.5 (Variational characterization of the spectral gap of L)**.** *The spectral gap ϱ_{SG} of the operator L has the variational characterization*

$$\varrho_{\mathrm{SG}} := \inf_f \frac{\mathcal{E}(f)}{\mathrm{var}_\mu(f)} = \varepsilon \inf_f \frac{\int |\nabla f|^2 \, \mathrm{d}\mu}{\mathrm{var}_\mu(f)}, \tag{SG(ϱ_{SG})}$$

where the infimum runs over all non-constant test functions $f : \mathbb{R}^n \to \mathbb{R}$.
From the defintion of PI(ϱ) *and* SG(ϱ_{SG}) *follows that the operator L has a spectral gap of size $\varrho_{\mathrm{SG}} = \varrho\varepsilon$ if and only if the Gibbs measure μ satisfies* PI(ϱ) *with optimal constant ϱ.*

Remark 1.6. The right-hand side of $\mathrm{PI}(\varrho)$ given by $\int |\nabla f|^2 \,\mathrm{d}\mu$ is also a Dirichlet form, namely of the rescaled generator $\tilde{L} := \frac{1}{\varepsilon} L = \Delta - \frac{1}{\varepsilon}\nabla H \cdot \nabla$. In this case, the definition of $\widetilde{\mathrm{PI}}(\tilde{\varrho})$ and $\widetilde{\mathrm{SG}}(\tilde{\varrho}_{\mathrm{SG}})$ coincide and hence $\tilde{\varrho} = \tilde{\varrho}_{\mathrm{SG}}$. The operators L and $\tilde{L}$ consider the process ξ_t given in (1.3) on different time-scales. The operator L corresponds to the slow time-scale, where the diffusion constant vanishes and the drift stays order one. Otherwise, the operator $\tilde{L}$ corresponds to a diffusion of order one and drift increasing to infinity.

The above Lemma 1.5 is one of the main motivations to study the sharp constant in $\mathrm{PI}(\varrho)$. Moreover, the inequalities $\mathrm{PI}(\varrho)$ and $\mathrm{LSI}(\alpha)$ are already enough to show concentration properties of the associated semigroup $\{P_t\}_{t\geq 0}$, which we formulate in the following theorem.

Theorem 1.7 (Concentration induced by $\mathrm{PI}(\varrho)$ and $\mathrm{LSI}(\alpha)$)**.** *Let f_t be the relative density of the process ξ_t* (1.3) *with a relative initial density f_0, i.e. the solution to* (1.5). *Further, assume that the Gibbs measure μ satisfies* $\mathrm{PI}(\varrho)$ *with $\varrho > 0$, then it holds the exponential* concentration in variance

$$\mathrm{var}_\mu(f_t) \leq e^{-2\varepsilon\varrho t}\,\mathrm{var}_\mu(f_0). \tag{1.7}$$

Likewise, assume that the Gibbs measure μ satisfies $\mathrm{LSI}(\alpha)$ *with $\alpha > 0$, then it holds the exponential* concentration in entropy

$$\mathrm{Ent}_\mu(f_t) \leq e^{-2\varepsilon\alpha t}\,\mathrm{Ent}_\mu(f_0).$$

Proof. The result is derived from a Gronwall argument. Therefore, let us consider the time derivative of $\mathrm{var}_\mu(f_t)$

$$\partial_t \,\mathrm{var}_\mu(f_t) = 2\int f_t \partial f_t \,\mathrm{d}\mu \overset{(1.5)}{=} 2\int f_t L f_t \,\mathrm{d}\mu \overset{(1.2)}{=} -2\varepsilon \int |\nabla f|^2 \,\mathrm{d}\mu \overset{\mathrm{PI}(\varrho)}{\leq} -2\varepsilon\varrho\,\mathrm{var}_\mu(f_t).$$

An application of the Gronwall Lemma shows (1.7). Similarly, we consider the time derivative of $\mathrm{Ent}_\mu(f_t)$

$$\begin{aligned}\partial_t \,\mathrm{Ent}_\mu(f_t) &= \int \partial_t f_t \,(\log f_t + 1)\,\mathrm{d}\mu \overset{(1.5)}{=} \int L f_t\,(\log f_t + 1)\,\mathrm{d}\mu \\ &\overset{(1.1)}{=} -\varepsilon \int \nabla f_t \cdot \nabla(\log f_t)\,\mathrm{d}\mu = -\varepsilon \int \frac{|\nabla f_t|^2}{f_t}\,\mathrm{d}\mu \overset{\mathrm{LSI}(\alpha)}{\leq} -2\varepsilon\alpha\,\mathrm{Ent}_\mu(f_t).\end{aligned}$$

□

Theorem 1.7 shows that detailed knowledge about the constants ϱ and α in $\mathrm{PI}(\varrho)$ and $\mathrm{LSI}(\alpha)$, especially there ε-dependence, leads to sharp asymptotics of the long time behavior of the process ξ_t (1.3).

Besides these immediate consequences from Lemma 1.5 and Theorem 1.7, the Poincaré and logarithmic Sobolev inequality give also bounds in the *concentration of measure phenomenon* (cf. [Led99a]).

1.2. Setting and assumptions

Before starting the precise assumptions on the Hamiltonian H, we introduce the notion of a *Morse function*.

Definition 1.8 (Morse function)**.** A smooth function $H : \mathbb{R}^n \to \mathbb{R}$ is a *Morse function*, if the Hessian $\nabla^2 H$ of H is non-degenerated on the set of critical points. More precisely, for some $1 \leq C_H < \infty$ holds

$$\forall x \in \mathcal{S} := \{x \in \mathbb{R}^n : \nabla H = 0\} : \quad \frac{1}{C_H} \leq \|\nabla^2 H(x)\| \leq C_H. \tag{1.8}$$

We make the following assumption on the Hamiltonian H which despite the non-degeneracy only matter if the domain of H is unbounded. Hereby, we have to assume stronger properties for H if we want to proof a logarithmic Sobolev inequality.

Assumption 1.9 (Poincaré inequality)**.** *We assume that $H \in C^3(\mathbb{R}^n, \mathbb{R})$ is a Morse function. Further, for some constants $C_H > 0$ and $K_H \geq 0$ holds*

$$\liminf_{|x|\to\infty} |\nabla H| \geq C_H. \tag{$\mathbf{A1}_{\mathtt{PI}}$}$$

$$\liminf_{|x|\to\infty} |\nabla H|^2 - \Delta H \geq -K_H. \tag{$\mathbf{A2}_{\mathtt{PI}}$}$$

Assumption 1.10 (Logarithmic Sobolev inequality)**.** *We assume that $H \in C^3(\mathbb{R}^n, \mathbb{R})$ is a Morse function. Further, for some constants $C_H > 0$ and $K_H \geq 0$ holds*

$$\liminf_{|x|\to\infty} \frac{|\nabla H(x)|^2 - \Delta H(x)}{|x|^2} \geq C_H. \tag{$\mathbf{A1}_{\mathtt{LSI}}$}$$

$$\inf_x \nabla^2 H(x) \geq -K_H. \tag{$\mathbf{A2}_{\mathtt{LSI}}$}$$

Remark 1.11 (Discussion of assumptions). The Assumption 1.9 yields the following consequences for the Hamiltonian H:

- The condition ($\mathbf{A1}_{\mathtt{PI}}$) ensures that e^{-H} is integrable and can be normalized to a probability measure on $\mathbb{R}^n$. Hence, the Gibbs measure μ given by (1.4) is well defined.

- A combination of the condition ($\mathbf{A1}_{\mathtt{PI}}$) and ($\mathbf{A2}_{\mathtt{PI}}$) ensures that there exists a spectral gap for the operator L given by (1.1), the argument is presented in the Appendix C. Equivalently, this means by the variational characterization of the spectral gap of L (cf. Lemma 1.5) that the Gibbs measure μ given by (1.4) satisfies a Poincaré inequality for sufficiently small ε.

- Basically, the *Lyapunov-type* condition ($\mathbf{A2}_{\mathtt{PI}}$) allows to recover the spectral gap of the full Gibbs measure μ from the spectral gap of the Gibbs measure μ_R restricted to some ball B_R with radius $R > 0$ not necessarily large (cf. Chapter 3).

- The Morse Assumption (1.8) together with the growth condition ($\mathbf{A1}_{\mathrm{PI}}$) ensures that the set $\mathcal{S}$ of critical points is discrete and finite. In particular, it follows that the set of local minima $\mathcal{M} = \{m_1, \dots, m_M\}$ is also finite i.e. $M := \#\mathcal{M} < \infty$.

Similarly the Assumption 1.10 has the following consequences for the Hamiltonian H:

- To illustrate the differences between the assumptions on H for the Poincaré and the logarithmic Sobolev inequality, it shall be pointed out that ($\mathbf{A1}_{\mathrm{PI}}$) means at least linear growth at infinity for H, whereas a combination of condition ($\mathbf{A1}_{\mathrm{LSI}}$) and ($\mathbf{A2}_{\mathrm{LSI}}$) yields that H has at least quadratic growth at infinity; that is

$$\liminf_{|x|\to\infty} \frac{|\nabla H(x)|}{|x|} \geq C_H. \tag{$\mathbf{A0}_{\mathrm{LSI}}$}$$

- In addition, ($\mathbf{A1}_{\mathrm{LSI}}$) and ($\mathbf{A2}_{\mathrm{LSI}}$) imply ($\mathbf{A1}_{\mathrm{PI}}$) and ($\mathbf{A2}_{\mathrm{PI}}$), which is only an indication that $\mathrm{LSI}(\alpha)$ is stronger than $\mathrm{PI}(\varrho)$ in the sense of Remark 1.4. Especially, the Assumption 1.10 is more restrictive than the Assumption 1.9 and whenever we refer to Assumption 1.9 the properties in question hold also under the Assumption 1.10.
- Since, the condition ($\mathbf{A0}_{\mathrm{LSI}}$) ensures quadratic growth at infinity of H, $e^{-\frac{H}{\varepsilon}}$ can be normalized to a probability measure. Moreover, quadratic growth at infinity is also a necessary condition to have a logarithmic Sobolev inequalities (cf. [Roy07, Theorem 3.1.21.]).
- Again, the condition ($\mathbf{A1}_{\mathrm{LSI}}$) is a *Lyapunov type* condition, but only implying a defective WI-inequality (cf. Appendix D). To deduce a logarithmic Sobolev inequality additionally ($\mathbf{A2}_{\mathrm{LSI}}$) has to be enforced (cf. Chapter 3).

The next additional non-degeneracy assumption is not directly needed for the proof of the Eyring-Kramers formula, but more to keep the presentation feasible and clear. Let us introduce the saddle height $\widehat{H}(m_i, m_j)$ between two local minima m_i, m_j by

$$\widehat{H}(m_i, m_j) = \inf\left\{ \max_{s\in[0,1]} H(\gamma(s)) : \gamma \in C([0,1], \mathbb{R}^n), \gamma(0) = m_i, \gamma(1) = m_j \right\}.$$

Assumption 1.12 (Non-degeneracy). *There exists $\delta > 0$ such that:*

(i) The saddle height between two local minima m_i, m_j is attained at a unique critical point $s_{i,j} \in \mathcal{S}$, i.e. it holds $H(s_{i,j}) = \widehat{H}(m_i, m_j)$. The point $s_{i,j}$ is called optimal *or* communicating saddle *between the minima m_i and m_j. It follows from Assumption 1.9 that $s_{i,j}$ is a saddle point of index one, i.e. $\{x \in \mathbb{R}^n : \langle \nabla^2 H(s_{i,j})x, x\rangle \leq 0\}$ is one-dimensional.*

(ii) The set of local minima $\mathcal{M} = \{m_1, \dots, m_M\}$ is ordered such that m_1 is the global minimum and for all $i \in \{3, \dots, M\}$ yields

$$H(s_{1,2}) - H(m_2) \geq H(s_{1,i}) - H(m_i) + \delta.$$

Remark 1.13 (Weaker degenerate assumptions). In Section 4.5 we give an argument, how the Eyring-Kramers formulas will look like without the Assumption 1.12 (i), i.e. if there exist more than one *communicating saddle* between certain minima. Likewise, the role of Assumption 1.12 (ii) becomes after the deduction of the Eyring-Kramers formulas in Section 2.4.3 apparent. In Remark 2.30 we argue that explicit formulas will strongly depend on the specific model and cannot give meaningful insights for the present general case.

Remark 1.14 (Comparison with the assumptions of [BGK05]). Between the Assumption 1.9 and Assumption 1.12 and ones used by [BGK05] are two minor differences:

- For convenience, we assume that the domain of H is $\mathbb{R}^n$. As in [BGK05], our argument would also work for any open and connected subset domain $D \subset \mathbb{R}^n$ satisfying $H(x_i) \to \infty$ whenever $x_i \to x \in \partial D$.
- Note that the non-degeneracy assumption (1.8) holds for all critical points x of H. In [BGK05], the assumption (1.8) is only needed for the local minimum m_i and the saddles $s_{i,j}$. We need this slightly stronger assumption only in the alternative proof of the local Poincaré inequality (cf. Theorem 2.19 and Chapter 5). The construction of the Lyapunov function in the proof of Theorem 2.19 and Theorem 2.21 of Chapter 3 does not rely on this non-degeneracy assumptions. In the argument of [BGK05] the need of a local Poincaré inequality is circumvented by the use of regularity theory for elliptic operators.

Chapter 2

Outline and main results

This chapter contains the main results of this work and gives an outline of the proofs of the Eyring-Kramers formulas. The proof is mainly motivated by the heuristics of the splitting into two time-scales: one describing the *fast relaxation* to a local minima of H and the other one describing *exponentially long transitions* between local equilibrium states.

The heuristics motivate a *divide-and-conquer strategy*. Therefore, the first Section 2.1 consists of the observation that the splitting of a measure into a *conditional* and a *marginal* measures can be lifted to a splitting for the variance and entropy. In Section 2.2 we specify the splitting and introduce *local measures* living on the *basin of attraction* of the local minima of H. Since, we can lift this splitting to the variance and entropy, we obtain in this way *local variances and entropies* as well as *coarse-grained variances and entropies*. It turns out, that the coarse-grained variances have the form of mean-differences. However, we have to do some work for the coarse-grained entropies to obtain a discrete logarithmic Sobolev type inequality in Section 2.3, which allows for an estimate in terms of local variances and mean-differences.

At this point, we will have prepared all the ingredients to state the main results in Section 2.4, which consist of good estimates for the local variances and entropies (cf. Section 2.4.1) and very sharp estimates for the mean-differences (cf. Section 2.4.2). With the help of the main ingredients, the Eyring-Kramers formulas are simple corollaries and are deduced in Section 2.4.3. We close this chapter with a discussion of the optimality of the Eyring-Kramers formula for the logarithmic Sobolev inequality in Section 2.5.

2.1. Splitting of measures and functionals

The variance $\mathrm{var}_\mu(f)$ and relative entropy $\mathrm{Ent}_\mu(f)$ on the left-hand side of $\mathrm{PI}(\varrho)$ and $\mathrm{LSI}(\alpha)$ fall in a general class of relative distances on absolutely continuous probability measures, which can be characterized by Definition 2.1. We will see that all distances of this kind can be split into *conditional* and *marginal* measures.

Definition 2.1 (Distance functional)**.** For a convex function $\xi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ and an absolutely continuous probability measure μ on $\mathbb{R}^n$, *distance functional* $\Xi_\mu : \mathcal{D}_\mu(\xi) \to \mathbb{R}_0^+$ is defined by

$$\Xi_\mu(f) := \int \xi(f)\ \mathrm{d}\mu - \xi\left(\int f\ \mathrm{d}\mu\right), \quad \text{where} \quad \mathcal{D}_\mu(\xi) = \left\{f : \int \xi(f)\ \mathrm{d}\mu < \infty\right\}. \quad (2.1)$$

Remark 2.2. The positivity of Ξ_μ follows from the convexity of ξ since by Jensen's inequality holds $\xi\left(\int f \mathrm{d}\mu\right) \leq \int \xi(f)\ \mathrm{d}\mu$.
Furthermore, if ξ is strictly convex, then $\Xi_\mu(f) = 0$ if and only if f is constant.

Example 2.3. (i) We can set $\xi(x) := x^2$ and obtain that the distance functional Ξ_μ of Definition 2.1 becomes the *variance*

$$\Xi_\mu(f) = \int f^2\ \mathrm{d}\mu - \left(\int f\ \mathrm{d}\mu\right)^2 = \int \left(f - \int f\ \mathrm{d}\mu\right)^2 \mathrm{d}\mu = \mathrm{var}_\mu(f).$$

(ii) Likewise, we can set $\xi(x) := x \log x$ for $x \geq 0$ and $\xi(x) = +\infty$ for $x < 0$ and recover the *relative entropy* between $f\mu$ and μ, where now $\mathcal{D}_\mu(\xi) \subset \left\{f : \mathbb{R}^n \to \mathbb{R}_0^+\right\}$

$$\Xi_\mu(f) = \int f \log f\ \mathrm{d}\mu - \int f\ \mathrm{d}\mu\ \log\left(\int f\ \mathrm{d}\mu\right) = \int f\ \log \frac{f}{\int f \mathrm{d}\mu}\ \mathrm{d}\mu = \mathrm{Ent}_\mu(f).$$

In the following, X will be a subset of $\mathbb{R}^n$ or $\mathbb{R}^n$ itself.

Definition 2.4 (Conditional and marginal measure for general coordinates)**.** A family $\{\psi_z : Y_z \to X\}_{z\in Z}$ are called *general coordinates* for the absolutely continuous probability measure μ on X, if for every $x \in X$ exists exactly one pair (y, z) such that $x = \psi_z(y)$ and there exists a family of probability measures $\mu(\cdot|z)$ on Y_z for $z \in Z$ and a probability measure $\bar{\mu}(\cdot)$ on Z satisfying

$$\mu(\mathrm{d}x) = \mu(\mathrm{d}y|z)\ \bar{\mu}(\mathrm{d}z), \qquad \text{where} \quad x = \psi_z(y), \qquad \text{in the weak sense,} \qquad (2.2)$$

i.e. for any test function $f : X \to \mathbb{R}$ holds $\int_X f(x)\,\mu(\mathrm{d}x) = \int_Z \int_{Y_z} f\circ\psi_z(y)\,\mu(\mathrm{d}y|z)\,\bar{\mu}(\mathrm{d}z)$. Then we call the measures $\mu(\cdot|z)$ *conditional measures* and the measure $\bar{\mu}(\cdot)$ *marginal measure*.

The Definition 2.4 is quite general and the precise definition of $\mu(\cdot|z)$ and $\bar{\mu}(\cdot)$ heavily depends on the structure of Z and regularity properties of $\{\psi_z : Y_z \to X\}$. In general, it will rely on change of variables formulas and the coarea formula (cf. [EG92]). Therefore, let us illustrate the Definition 2.4 with two prototypical examples:

Example 2.5. (i) Let X be a measurable subset of $\mathbb{R}^n$ and let us consider some absolutely continuous probability measure μ on X. We set $Z = \{1, \ldots, m\}$ and let $\{Y_i\}_{i=1}^m$ be a measurable partition of X, i.e. $Y_i \cap Y_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^m \bar{Y}_i = \bar{X}$.

Then, we can choose $\psi_i(y) = y$ for $y \in Y_i$, i.e. the natural embedding. The conditional and marginal measures in (2.2) have the form

$$\mu(\mathrm{d}y|i) = \frac{\mathbb{1}_{Y_i}(y)}{Z_i}\,\mu(y)\,\mathrm{d}y \quad \text{and} \quad \bar{\mu} = Z_1\delta_1 + \cdots + Z_m\delta_m, \quad \text{where } Z_i = \mu(Y_i). \tag{2.3}$$

In this case, we say that μ has the representation of a *mixture*, i.e.

$$\mu = Z_1\mu(\cdot|1) + \cdots + Z_m\mu(\cdot|m).$$

(ii) For $X = \mathbb{R}^n$ with $n \geq 2$, we set $Z = S^{n-1}$, $Y_\eta = \mathbb{R}^+$ for $\eta \in S^{n-1}$ and choose $\psi_\eta(r) = r\eta$ for $r \in \mathbb{R}^+$. The conditional and marginal measures have the form

$$\mu(\mathrm{d}r|\eta) = \frac{1}{\bar{\mu}(\eta)}\,\mu(r\eta)\,\mathrm{d}r \qquad \text{and} \qquad \bar{\mu}(\mathrm{d}\eta) = \int_{\mathbb{R}^+} r^{n-1}\,\mu(r\eta)\,\mathrm{d}r\,\mathrm{d}\eta,$$

where $\mathrm{d}r$ is the one-dimensional Lebesgue measure and $\mathrm{d}\eta$ the surface element on S^{n-1}. This is just μ represented in *polar coordinates*.

Lemma 2.6 (Splitting lemma). *If $\mu(\mathrm{d}x) = \mu(\mathrm{d}y|z)\,\bar{\mu}(\mathrm{d}z)$ is a splitting of μ into conditional measures $\mu(\cdot|z)$ and the marginal measure $\bar{\mu}(\cdot)$ in general coordinates $\{\psi_z : Y_z \to X\}_{z\in Z}$ in the sense of Definition 2.4, then this splitting carries over to the distance functional of Definition 2.1 via*

$$\Xi_\mu(f) = \mathbb{E}_{\bar{\mu}(\mathrm{d}z)}\left(\Xi_{\mu(\mathrm{d}y|z)}(f\circ\psi_z(y))\right) + \Xi_{\bar{\mu}(\mathrm{d}z)}\left(\mathbb{E}_{\mu(\mathrm{d}y|z)}(f\circ\psi_z(y))\right).$$

Proof. We start with the definition of the distance functional $\Xi_\mu(f)$ from (2.1) and use the splitting of μ into conditional and marginal measures w.r.t. general coordinates given in (2.2)

$$\begin{aligned}
\Xi_\mu(f) &= \int_Z\int_{Y_z} \xi(f\circ\psi_z(y))\,\mu(\mathrm{d}y|z)\,\bar{\mu}(\mathrm{d}z) - \xi\left(\int_X f\,\mathrm{d}\mu\right)\\
&= \int_Z \left(\int_{Y_z} \xi(f\circ\psi_z(y))\,\mu(\mathrm{d}y|z) - \xi\left(\int_{Y_z} f\circ\psi_z(y)\,\mu(\mathrm{d}y|z)\right)\right)\bar{\mu}(\mathrm{d}z)\\
&\quad + \int_Z \xi\left(\int_{Y_z} f\circ\psi_z(y)\,\mu(\mathrm{d}y|z)\right)\bar{\mu}(\mathrm{d}z) - \xi\left(\int_Z\int_{Y_z} f\circ\psi_y(y)\,\mu(\mathrm{d}y|z)\,\bar{\mu}(\mathrm{d}z)\right).
\end{aligned}$$

□

Corollary 2.7. *The variance and entropy allow the splitting in general coordinates in the sense of Definition 2.4*

$$\operatorname{var}_\mu(f) = \mathbb{E}_{\bar{\mu}(\mathrm{d}z)}\left(\operatorname{var}_{\mu(\mathrm{d}y|z)}(f\circ\psi_z(y))\right) + \operatorname{var}_{\bar{\mu}(\mathrm{d}z)}\left(\bar{f}(z)\right) \tag{2.4}$$

$$\operatorname{Ent}_\mu(f) = \mathbb{E}_{\bar{\mu}(\mathrm{d}z)}\left(\operatorname{Ent}_{\mu(\mathrm{d}y|z)}(f\circ\psi_z(y))\right) + \operatorname{Ent}_{\bar{\mu}(\mathrm{d}z)}\left(\bar{f}(z)\right), \tag{2.5}$$

where $\bar{f}(z) := \mathbb{E}_{\mu(\mathrm{d}y|z)}(f\circ\psi_z(y))$.

Remark 2.8. Depending on the specific choice of the general coordinates, we will call $\operatorname{var}_{\mu(\mathrm{d}y|z)}(f\circ\psi_z(y))$ in (2.4) *microscopic, local* or *radial* variances. Similarly, the term $\operatorname{var}_{\bar{\mu}(\mathrm{d}z)}(\bar{f}(z))$ is called *macroscopic, coarse-grained* or *polar* variance. However, we will often treat the latter term not as a variance and will take advantage of its special structure induced by the according choice of the general coordinates. The according descriptions are used for the entropies appearing in (2.5).

2.2. Partition of the state space

Motivated by the fast convergence of the diffusion ξ_t given by (1.3) to metastable states, we decompose the Gibbs measure μ into a mixture of *local* Gibbs measures μ_i in the following way: To every local minimum $m_i \in \mathcal{M}$ for $i = 1, \dots, M$ we associate its *basin of attraction* Ω_i defined by

$$\Omega_i := \left\{ y \in \mathbb{R}^n : \lim_{t\to\infty} y_t = m_i, \dot{y}_t = -\nabla H(y_t), y_0 = y \right\}. \tag{2.6}$$

Up to sets of Lebesgue measure zero, the set $\mathcal{P}_M = \{\Omega_i\}_{i=1}^M$ is a partition of $\mathbb{R}^n$. We are in the setting of Example 2.5 (i) and can associate the local Gibbs measure μ_i to each element of the partition Ω_i as the restriction of μ

$$\mu_i(\mathrm{d}x) := \frac{1}{Z_i Z_\mu} \, \mathbb{1}_{\Omega_i}(x) \, \exp\left(-\frac{H(x)}{\varepsilon}\right) \mathrm{d}x, \qquad \text{where} \quad Z_i = \mu(\Omega_i). \tag{2.7}$$

The marginal measure $\bar{\mu}$ is given, according to (2.3), as a sum of Dirac measures $\bar{\mu} = Z_1\delta_1 + \cdots + Z_M\delta_M$. We note that $\sum_i Z_i = 1$, since $\{\Omega_i\}_{i=1}^M$ is a partition of $\mathbb{R}^n$ and μ a probability measure. The starting point of the argument is a representation of the Gibbs measure μ as a *mixture* of the mutual singular measures μ_i, namely

$$\mu = Z_1\mu_1 + \cdots + Z_M\mu_M. \tag{2.8}$$

We can apply Corollary 2.7 which lifts the decomposition of μ to a decomposition of the variance $\operatorname{var}_\mu(f)$ and entropy $\operatorname{Ent}_\mu(f)$. The representation below was also used in [CM10, Section 4.1].

Lemma 2.9 (Splitting of variance and entropy for partition)**.** *For all $f : \mathbb{R}^n \to \mathbb{R}$ holds*

$$\operatorname{var}_\mu(f) = \sum_{i=1}^M Z_i \operatorname{var}_{\mu_i}(f) + \sum_{i=1}^M \sum_{j>i} Z_i Z_j \left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \tag{2.9}$$

$$\operatorname{Ent}_\mu(f) = \sum_{i=1}^M Z_i \operatorname{Ent}_{\mu_i}(f) + \operatorname{Ent}_{\bar{\mu}}\left(\bar{f}\right). \tag{2.10}$$

We call the terms $\operatorname{var}_{\mu_i}(f)$ and $\operatorname{Ent}_{\mu_i}(f)$ local variance and local entropy. Furthermore, the term $\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2$ is called mean-difference and $\operatorname{Ent}_{\bar{\mu}}(\bar{f})$ denoted by coarse-grained entropy *is given by*

$$\operatorname{Ent}_{\bar{\mu}}\left(\bar{f}\right) = \sum_{i=1}^M Z_i \bar{f}_i \, \log \frac{\bar{f}_i}{\sum_{j=1}^M Z_j \bar{f}_j}, \tag{2.11}$$

where $\bar{f}(i) = \bar{f}_i = \mathbb{E}_{\mu_i}(f)$.

Proof. The equations (2.9) and (2.10) are immediate consequences of (2.4) and (2.5). The only non-obvious step is the representation of the coarse-grained variance $\operatorname{var}_{\bar{\mu}}(\bar{f})$ in terms of the second term of the right-hand side in (2.9). We will use the convention $\sum_i = \sum_{i=1}^M$ and have to check the following identity

$$\mathrm{var}_{\bar\mu}(\bar f) = \sum_i Z_i \bar f_i^{\,2} - \Big(\sum_j Z_j \bar f_j\Big)^2 \stackrel{!}{=} \sum_i \sum_{j>i} Z_i Z_j \left(\bar f_i - \bar f_j\right)^2 . \tag{2.12}$$

By expanding the square, we can write the left-hand side of (2.12) as

$$\sum_i Z_i \bar f_i^{\,2} - \sum_{i,j} Z_i Z_j \bar f_i \bar f_j = \sum_i Z_i (1 - Z_i) \bar f_i^{\,2} - 2 \sum_i \sum_{j>i} Z_i Z_j \bar f_i \bar f_j . \tag{2.13}$$

The first term on the right-hand side of (2.13) can be rewritten by using the identity $1 - Z_i = \sum_{j \neq i} Z_j$ as

$$\sum_i Z_i (1 - Z_i) \bar f_i^{\,2} = \sum_i \sum_{j<i} Z_i Z_j \bar f_i^{\,2} + \sum_i \sum_{j>i} Z_i Z_j \bar f_i^{\,2} = \sum_i \sum_{j>i} Z_i Z_j \left(\bar f_i^{\,2} + \bar f_j^{\,2}\right) . \tag{2.14}$$

Hence, we finally get the desired identity from combining (2.13) and (2.14)

$$\sum_i Z_i \bar f_i^{\,2} - \Big(\sum_j Z_j \bar f_j\Big)^2 = \sum_i \sum_{j>i} Z_i Z_j (\bar f_i^{\,2} - 2 \bar f_i \bar f_j + \bar f_j^{\,2}) = \sum_i \sum_{j>i} Z_i Z_j \left(\bar f_i - \bar f_j\right)^2 .$$

□

2.3. Discrete logarithmic Sobolev type inequalities

From (2.10) we have to estimate the *coarse-grained entropy* $\mathrm{Ent}_{\bar\mu}(\bar f)$. From the heuristic explanation, we expect that the main contribution comes from this term, which we want to treat further. If H has only two minima, we can use the following discrete logarithmic Sobolev inequality for a Bernoulli random variable, which was found by Higuchi and Yoshida [HY95] and Diaconis and Saloff-Coste [DSC96, Theorem A.2.] at the same time.

Lemma 2.10 (Optimal logarithmic Sobolev inequality for Bernoulli measures)**.** *A Bernoulli measure μ_p on $X = \{0, 1\}$, i.e. a mixture of two Dirac measures $\mu_p = p\delta_0 + q\delta_1$ with $p + q = 1$ satisfies the discrete logarithmic Sobolev inequality*

$$\mathrm{Ent}_{\mu_p}(f^2) \leq \frac{pq}{\Lambda(p,q)} (f(0) - f(1))^2 \tag{2.15}$$

with optimal constant given by the logarithmic mean

$$\Lambda(p,q) := \frac{p - q}{\log p - \log q}, \quad \textit{for } p \neq q \qquad \textit{and} \quad \Lambda(p,p) := \lim_{q \to p} \Lambda(p,q) = p.$$

Some properties of the logarithmic-mean are outlined in Appendix A.

We want to handle the general case with more than two minima. Therefore, we need to answer the question of how to generalize Lemma 2.10 to discrete measures with a state space with more than two elements. A possible answer was given by Diaconis and Saloff-Coste [DSC96, Theorem A.1.], which turns out to be not optimal for our application.

Lemma 2.11 (Logarithmic Sobolev inequality for finite discrete measure). *For $m \in \mathbb{N}$ let $\mu_m = \sum_{i=1}^m Z_i \delta_i$ be a discrete probability measure and assume that $Z_* := \min_i Z_i > 0$. Then for a function $f : \{1, \dots, m\} \to \mathbb{R}$ holds the logarithmic Sobolev inequality*

$$\operatorname{Ent}_{\mu_m}(f^2) \leq \frac{1}{\Lambda(Z_*, 1 - Z_*)} \sum_{i=1}^{m} \sum_{j>i} Z_i Z_j \left(f(i) - f(j) \right)^2 . \tag{2.16}$$

Remark 2.12. Note that the right-hand side of (2.16) is by the proof of Lemma 2.9 just the variance, hence we have

$$\operatorname{Ent}_{\mu_m}(f^2) \leq \frac{1}{\Lambda(Z_*, 1 - Z_*)} \operatorname{var}_{\mu_m}(f). \tag{2.17}$$

It turns out using the estimate (2.17) is not optimal for our application. We have to use a refined version of (2.16), which can be seen as an immediate generalization of (2.15) to the m-point case.

Lemma 2.13 (Logarithmic difference inequality for finite discrete measures). *For $m \in \mathbb{N}$ let $\mu_m = \sum_{i=1}^m Z_i \delta_i$ be a discrete probability measure and assume that $\min_i Z_i > 0$. Then for a function $f : \{1, \dots, m\} \to \mathbb{R}_0^+$ holds the logarithmic difference inequality*

$$\operatorname{Ent}_{\mu_m}(f^2) \leq \sum_{i=1}^{m} \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \left(f_i - f_j \right)^2 . \tag{2.18}$$

Proof. We conclude by induction and find that for $m = 2$ the estimate (2.18) just becomes (2.15), which shows the base case. For the inductive step, let us assume that (2.18) holds for $m \geq 2$. Then let us rewrite the entropy $\operatorname{Ent}_{\mu_{m+1}}(f^2)$ as follows

$$\begin{aligned}
\operatorname{Ent}_{\mu_{m+1}}(f^2) &= \sum_{i=1}^{m} Z_i f_i^2 \log(f_i^2) + Z_{m+1} f_{m+1}^2 \log(f_{m+1}^2) \\
&\quad - \left(\sum_{i=1}^{m} Z_i f_i^2 + Z_{m+1} f_{m+1}^2 \right) \log \left(\sum_{j=1}^{m} Z_j f_j^2 + Z_{m+1} f_{m+1}^2 \right) \\
&= (1 - Z_{m+1}) \sum_{i=1}^{m} \frac{Z_i f_i^2}{1 - Z_{m+1}} \left(\log(f_i^2) - \log \left(\sum_{j=1}^{m} \frac{Z_j f_j^2}{1 - Z_{m+1}} \right) \right) \\
&\quad + (1 - Z_{m+1}) \sum_{i=1}^{m} \frac{Z_i f_i^2}{1 - Z_{m+1}} \left(\log \left(\sum_{j=1}^{m} \frac{Z_j f_j^2}{1 - Z_{m+1}} \right) - \log \left(\sum_{i=j}^{m} Z_j f_j^2 + Z_{m+1} f_{m+1}^2 \right) \right) \\
&\quad + Z_{m+1} f_{m+1}^2 \left(\log(f_{m+1}^2) - \log \left(\sum_{i=j}^{m} Z_j f_j^2 + Z_{m+1} f_{m+1}^2 \right) \right) \\
&= (1 - Z_{m+1}) \operatorname{Ent}_{\tilde{\mu}_m}(f^2) + \operatorname{Ent}_{\nu}(\tilde{f}),
\end{aligned}$$

where the probability measure $\tilde{\mu}_m$ lives on $\{1,\ldots,m\}$ and is given by

$$\tilde{\mu}_m := \sum_{i=1}^{m} \frac{Z_i}{1-Z_{m+1}} \delta_i.$$

Further, ν is the Bernoulli measure given by $\nu := (1-Z_{m+1})\delta_0 + Z_{m+1}\delta_1$ and the function $\tilde{f} : \{0,1\} \to \mathbb{R}$ is given with values

$$\tilde{f}_0 := \sum_{i=1}^{m} \frac{Z_i f_i^2}{1-Z_{m+1}} \qquad \text{and} \qquad \tilde{f}_1 := f_{m+1}^2.$$

Now, we apply the inductive hypothesis to $\operatorname{Ent}_{\tilde{\mu}_m}(f^2)$ and arrive at

$$\begin{aligned}(1-Z_{m+1}) \operatorname{Ent}_{\tilde{\mu}_m}(f^2) &\leq (1-Z_{m+1}) \sum_{i=1}^{m} \sum_{j>i} \frac{Z_i Z_j}{(1-Z_{m+1})^2} \frac{1-Z_{m+1}}{\Lambda(Z_i, Z_j)} (f_i - f_j)^2 \\ &= \sum_{i=1}^{m} \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} (f_i - f_j)^2 ,\end{aligned}$$

where we used $\Lambda(\cdot,\cdot)$ being homogeneous of degree one in both arguments (cf. Appendix A), i.e.

$$\Lambda(\lambda a, \lambda b) = \lambda \Lambda(a,b) \qquad \text{for} \qquad \lambda, a, b > 0.$$

We can apply the inductive base to the second entropy $\operatorname{Ent}_\nu(\tilde{f})$, which in this case is the discrete logarithmic Sobolev inequality for the two-point case (2.15)

$$\operatorname{Ent}_\nu(\tilde{f}) \leq \frac{Z_{m+1}(1-Z_{m+1})}{\Lambda(Z_{m+1}, 1-Z_{m+1})} \left(\sqrt{\tilde{f}_0} - \sqrt{\tilde{f}_1}\right)^2 . \tag{2.19}$$

The last step is to recover the square differences $(f_i - f_{m+1})^2$ from $\left(\sqrt{\tilde{f}_0} - \sqrt{\tilde{f}_1}\right)^2$, which follows from an application of the Jensen inequality

$$\begin{aligned}\left(\sqrt{\tilde{f}_0} - \sqrt{\tilde{f}_1}\right)^2 &= \sum_{i=1}^{m} \frac{Z_i f_i^2}{1-Z_{m+1}} - 2 \underbrace{\sqrt{\sum_{i=1}^{m} \frac{Z_i f_i^2}{1-Z_{m+1}}}}_{\geq \sum_{i=1}^{m} \frac{Z_i f_i}{1-Z_{m+1}}} f_{m+1} + f_{m+1}^2 . \\ &\leq \sum_{i=1}^{m} \frac{Z_i}{1-Z_{m+1}} (f_i - f_{m+1})^2 .\end{aligned}$$

We obtain in combination with (2.19) the following estimate

$$\operatorname{Ent}_\nu(\tilde{f}) \leq \frac{Z_{m+1}}{\Lambda(Z_{m+1}, 1-Z_{m+1})} \sum_{i=1}^{m} Z_i (f_i - f_{m+1})^2 .$$

To conclude the assertion, we first note that $1 - Z_{m+1} = \sum_{j=1}^{m} Z_j \geq Z_j$ for $j = 1, \ldots, m$. Further, $\Lambda(a,\cdot)$ is monotone increasing for $a > 0$, i.e. $\partial_b \Lambda(a,b) > 0$ (cf. Appendix A). Both properties imply that $\Lambda(Z_{m+1}, 1-Z_{m+1}) \geq \Lambda(Z_{m+1}, Z_j)$ for $j = 1, \ldots, m$, which finally shows (2.18). □

Remark 2.14 (Relation between Lemma 2.11 and Lemma 2.13). Let us give a example, which shows that neither the estimate (2.16) is a consequence of the estimate (2.18) nor vice versa. Therefore, let us consider a tree point measure $\mu_3 = Z_1\delta_1 + Z_2\delta_2 + Z_3\delta_3$ with distribution

$$Z_1 = 1 - e^{-z_2} - e^{-z_3}, \qquad Z_2 = e^{-z_2} \qquad \text{and} \qquad Z_3 = e^{-z_3}$$

for some $0 = z_1 \ll z_2 \ll z_3$ to be fixed later. Then, it follows $Z_* = Z_3$ and therefore

$$\frac{1}{\Lambda(Z_*, 1 - Z_*)} = \frac{\log(1 - e^{-z_3}) + z_3}{1 - 2e^{-z_3}} \leq \frac{z_3 - e^{-z_3}}{1 - 2e^{-z_3}} = z_3 + O(z_3 e^{-z_3}).$$

By similar reasons, we deduce

$$\text{for } i < j\text{:} \quad \frac{1}{\Lambda(Z_i, Z_j)} = (z_j - z_i)e^{z_i} + O((z_j - z_i)e^{-(z_j - z_i)}).$$

Now, let us compare the estimate (2.16) and (2.18) by first considering a test function $f : \{1, 2, 3\} \to \mathbb{R}$ only supported on $\{1, 2\}$

$$\begin{aligned} &(2.16) \qquad \mathrm{Ent}_{\mu_3}(f^2) \leq (z_3 + O(e^{-z_3}))Z_1 Z_2 \left(f_1 - f_2\right)^2 \\ &(2.18) \qquad \mathrm{Ent}_{\mu_3}(f^2) \leq (z_2 + O(e^{-z_2}))Z_1 Z_2 \left(f_1 - f_2\right)^2, \end{aligned}$$

whereas, if f is only supported on $\{2, 3\}$, then we arrive at

$$\begin{aligned} &(2.16) \qquad \mathrm{Ent}_{\mu_3}(f^2) \leq (z_3 + O(e^{-z_3}))Z_2 Z_3 \left(f_2 - f_3\right)^2 \\ &(2.18) \qquad \mathrm{Ent}_{\mu_3}(f^2) \leq ((z_3 - z_2)e^{z_2} + O(e^{-(z_3 - z_2)}))Z_2 Z_3 \left(f_2 - f_3\right)^2. \end{aligned}$$

In the first case, we see that (2.18) gives a better bound. In the second case, if we choose z_2 large enough, we find that (2.16) yields the better estimate.

An immediate consequence of Lemma 2.11 and Lemma 2.13 is the following combined bound.

Corollary 2.15. *Let for $m \in \mathbb{N}$ be $\mu_m = \sum_{i=1}^m Z_i\delta_i$ a discrete probability measure and assume that $Z_* = \min_i Z_i > 0$. Then for a function $f : \{1, \ldots, m\} \to \mathbb{R}$ holds the estimate*

$$\mathrm{Ent}_{\mu_m}(f^2) \leq \sum_i \sum_{j>i} \frac{Z_i Z_j}{\max\left\{\Lambda(Z_*, 1 - Z_*), \Lambda(Z_i, Z_j)\right\}} \left(f(i) - f(j)\right)^2.$$

Open question: What is the optimal family of constants $\{C_{i,j}\}_{i,j=1}^m$ not depending on f in the inequality

$$\mathrm{Ent}_{\mu_m}(f^2) \leq \sum_{i=1}^m \sum_{j>i} C_{i,j} \left(f(i) - f(j)\right)^2 \qquad ? \tag{2.20}$$

Remark 2.16. The question (2.20) is related to the recent works of Maas [Maa11], Erbar and Mass [EM11] and Mielke [Mie11]. [Mie11] proved that every finite Markov chain is geodesic λ-convex w.r.t. to an entropic gradient structure. Independently, at the same time [Maa11] also found the gradient structure for finite Markov chains. Based on this work, [EM11] defined a Ricci curvature, which allows to prove modified logarithmic Sobolev inequalities similar to (2.16) in the positive curved case. In all three works the logarithmic mean plays a crucial role for defining the gradient structure as well as for obtaining bounds for the involved quantities (λ-convexity or curvature).

We are now able to estimate the coarse-grained entropy $\mathrm{Ent}_{\bar\mu}(\overline{f^2})$ occuring in the splitting of the entropy (2.10) with the help of Lemma 2.13. This generalizes the approach of [CM10, Section 4.1.] to the case of finite mixtures with more than two components.

Lemma 2.17 (Estimate of the coarse-grained entropy)**.** *The coarse-grained entropy in* (2.11) *can be estimated by*

$$\mathrm{Ent}_{\bar\mu}\left(\overline{f^2}\right) \le \sum_{i=1}^{M}\left(\sum_{j\neq i}\frac{Z_iZ_j}{\Lambda(Z_i,Z_j)}\,\mathrm{var}_{\mu_i}(f) \;+\; \sum_{j>i}\frac{Z_iZ_j}{\Lambda(Z_i,Z_j)}\left(\mathbb{E}_{\mu_i}(f)-\mathbb{E}_{\mu_j}(f)\right)^2\right)$$

where $\overline{f^2} : \{1,\dots,M\} \to \mathbb{R}$ *is given by* $\overline{f_i^2} := \mathbb{E}_{\mu_i}(f^2)$ *and* $\Lambda(Z_i,Z_j) = \frac{Z_i - Z_j}{\log Z_i - \log Z_j}$ *is the logarithmic mean between* Z_i *and* Z_j*.*

Proof. Since $\bar\mu = Z_1\delta_1 + \dots + Z_M\delta_M$ is finite discrete probability measure, we can apply Lemma 2.13 to $\mathrm{Ent}_{\bar\mu}(\overline{f^2})$

$$\mathrm{Ent}_{\bar\mu}(\overline{f^2})) \le \sum_{i=1}^{m}\sum_{j>i}\frac{Z_iZ_j}{\Lambda(Z_i,Z_j)}\left(\sqrt{\overline{f_i^2}} - \sqrt{\overline{f_j^2}}\right)^2 . \tag{2.21}$$

The square-root-mean-difference on the right-hand side of (2.21) can be estimated in several ways. [CM10] multiplies the square out and use the Jensen inequality

$$\mathbb{E}_{\mu_i}(f)\mathbb{E}_{\mu_j}(f) \le \sqrt{\mathbb{E}_{\mu_i}(f^2)\mathbb{E}_{\mu_j}(f^2)}.$$

This strategy leads to the same result as [JSTV04] obtains by using the fact that the function $(x,y) \mapsto (\sqrt{x}-\sqrt{y})^2$ is convex. For this trick they reference [KOV89]. Indeed, for two random variables X, Y such that $X \sim \mu_i$ and $Y \sim \mu_j$ by an application of Jensen's inequality holds the estimate

$$\begin{aligned}\left(\sqrt{\mathbb{E}_{\mu_i}(f^2)} - \sqrt{\mathbb{E}_{\mu_j}(f^2)}\right)^2 &= \left(\sqrt{\mathbb{E}_{\mu_i\times\mu_j}(f(X)^2)} - \sqrt{\mathbb{E}_{\mu_i\times\mu_j}(f(Y)^2)}\right)^2 \\ &\le \mathbb{E}_{\mu_i\times\mu_j}\left((f(X)-f(Y))^2\right) \\ &\le \mathbb{E}_{\mu_i}(f^2) - 2\mathbb{E}_{\mu_i}(f)\mathbb{E}_{\mu_j}(f) + \mathbb{E}_{\mu_j}(f^2) \\ &= \mathrm{var}_{\mu_i}(f) + \mathrm{var}_{\mu_j}(f) + \left(\mathbb{E}_{\mu_i}(f)-\mathbb{E}_{\mu_j}(f)\right)^2 .\end{aligned} \tag{2.22}$$

Now, we can combine (2.21) and (2.22) to arrive at

$$\begin{aligned}\operatorname{Ent}_{\bar\mu}(\overline{f^2}) &\le \sum_{i=1}^{m}\sum_{j>i}\frac{Z_iZ_j}{\Lambda(Z_i,Z_j)}\left(\operatorname{var}_{\mu_i}(f)+\operatorname{var}_{\mu_j}(f)+\left(\mathbb{E}_{\mu_i}(f)-\mathbb{E}_{\mu_j}(f)\right)^2\right)\\ &= \sum_{i=1}^{m}\sum_{j\neq i}\frac{Z_iZ_j}{\Lambda(Z_i,Z_j)}\operatorname{var}_{\mu_i}(f)+\sum_{i=1}^{m}\sum_{j>i}\frac{Z_iZ_j}{\Lambda(Z_i,Z_j)}\left(\mathbb{E}_{\mu_i}(f)-\mathbb{E}_{\mu_j}(f)\right)^2.\end{aligned}$$

□

The combination of the splitting Lemma 2.9 with the above Lemma 2.17 results in an estimate of the entropy in terms of local variances, local entropies and mean-differences.

Corollary 2.18. *The entropy of f with respect to μ on a partition $\{\Omega_i\}_{i=1}^M$ with restricted probability measures μ_i on Ω_i can be estimated by*

$$\begin{aligned}\operatorname{Ent}_\mu(f^2) \le \sum_{i=1}^{M} Z_i\operatorname{Ent}_{\mu_i}(f^2)&+\sum_{i=1}^{M}\sum_{j\neq i}\frac{Z_iZ_j}{\Lambda(Z_i,Z_j)}\operatorname{var}_{\mu_i}(f)\\ &+\sum_{i=1}^{M}\sum_{j>i}\frac{Z_iZ_j}{\Lambda(Z_i,Z_j)}\left(\mathbb{E}_{\mu_i}(f)-\mathbb{E}_{\mu_j}(f)\right)^2\end{aligned} \tag{2.23}$$

where $\Lambda(Z_i,Z_j)=\frac{Z_i-Z_j}{\log Z_i-\log Z_j}$ is the logarithmic mean between Z_i and Z_j.

2.4. Main results

The main results of this work are good estimates of the single terms on the right-hand side of (2.9) and (2.23). In detail, we need *local Poincaré* and *local logarithmic Sobolev inequalities* provided by Theorem 2.19 and Theorem 2.21. Furthermore, we need good control of the mean-differences, which will be the content of Theorem 2.23. Finally, the Eyring-Kramers formulas in Corollary 2.26 and Corollary 2.28 are simple consequences of all these representations and estimates.

2.4.1. Local Poincaré and logarithmic Sobolev inequalities

Let us now turn to the estimation of the local variances and entropies. From the heuristic understanding of the process ξ_t given by (1.3), we expect a good behavior of the local Poincaré and logarithmic Sobolev constant for the local Gibbs measures μ_i as it resembles the fast convergence of ξ_t to a neighborhood of the next local minimum. Therefore, the local variances and entropies should not contribute to the leading order expansion of the total Poincaré and logarithmic Sobolev constant of μ. This idea is quantified in the next both theorems.

Theorem 2.19 (Local Poincaré inequality). *The local measures $\{\mu_i\}_{i=1}^M$, obtained by restricting μ to the basin of attraction Ω_i of the local minimum m_i (cf.* (2.7)*), satisfy* $\mathrm{PI}(\varrho_i)$ *with*

$$\varrho_i^{-1} = O(\varepsilon).$$

Remark 2.20. Using the variational characterization of the spectral gap (cf. Lemma 1.5) one can easily see the following consequence of Theorem 2.19: The spectral gap of the diffusion ξ_t given by (1.3) reflected at the boundary of a basin of attraction Ω_i is at least of order 1. This reflects the heuristic idea of a scale separation of the diffusion ξ_t into a fast and a slow scale.

Theorem 2.21 (Local logarithmic Sobolev inequality). *The local measures $\{\mu_i\}_{i=1}^M$, obtained by restricting μ to the basin of attraction Ω_i of the local minimum m_i (cf.* (2.7)*), satisfy* $\mathrm{LSI}(\alpha_i)$ *with*

$$\alpha_i^{-1} = O(1).$$

Even if there are simple heuristics for the validity of Theorem 2.19 and Theorem 2.21, we need the elaborated machinery of *Lyapunov functions* for the proof. The reason is that our situation goes beyond the scope of the standard tools for Poincaré and logarithmic Sobolev inequalities. We outline the argument for Theorem 2.19 and Theorem 2.21 in Chapter 3. Moreover, in Chapter 5 we give an alternative proof of Theorem 2.19 completely avoiding the use of Lyapunov conditions.

Remark 2.22 (Optimality of Theorem 2.19 and 2.21). We will indicate in Section 3.3, that the results of Theorem 2.19 and Theorem 2.21 for a Gibbs measure obtained from restricting its Hamiltonian H to the basin of attraction Ω of a local minimum is the best behavior, which one can expect in general. Especially, H is not convex and can even have further critical points on the boundary of the basins of attraction. The indication is given through the one-dimensional case, where the *Muckenhoupt functional* [Muc72] and *Bobkov-Götze functional* [BG99] are available for showing upper and lower bounds for ϱ and α.

2.4.2. Mean-difference estimate

Let us now turn to the estimation of the mean-difference $\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2$. From the heuristics and the splittings (2.9) and (2.23), we expect to see in the estimation of $(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f))^2$ the exponential long waiting times of the jumps of the diffusion ξ_t given by (1.3) between the basins of attraction Ω_i. We have to estimate the mean-differences on the right-hand side of (2.9) in terms of the Dirichlet energy $\int |\nabla f|^2 \,\mathrm{d}\mu$. Namely, we have to find a good upper bound for the constant C in the inequality

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \leq C \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Therefore, we introduce in Section 4.1 a *weighted transport distance* between probability measures which yields a variational bound on the constant C. By an approximation argument (cf. Section 4.2), we give an explicit construction of a transport interpolation (cf. Section 4.3), which allows for asymptotically sharp estimates of the constant C.

Theorem 2.23 (Mean-difference estimate)**.** *The mean-differences between the measures μ_i and μ_j for $i = 1, \dots M-1$ and $j = i+1, \dots, M$ satisfy*

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|}\, e^{\frac{H(s_{i,j})}{\varepsilon}} \int |\nabla f|^2 \,\mathrm{d}\mu, \tag{2.24}$$

where $\lambda^-(s_{i,j})$ denotes the negative eigenvalue of the Hessian $\nabla^2 H(s_{i,j})$ at the 1-saddle $s_{i,j}$. The symbol $\lesssim$ means $\leq$ up to a multiplicative error term of the form

$$1 + O(\sqrt{\varepsilon}\,|\log \varepsilon|^{\frac{3}{2}}).$$

The proof of Theorem 2.23 is carried out in full detail in Chapter 4.

Remark 2.24 (Multiple minimal saddles). In Assumption 1.12, we demand that there is exactly one minimal saddle between the local minima m_i and m_j. The technique we will develop in Chapter 4 is flexible enough to handle also cases, in which there exists more than one minimal saddle between local minima. We outline the according adaptions and the resulting Theorem 4.18 in Section 4.5.

Remark 2.25 (Relation to capacity). The quantity on the right-hand side of (2.24) is the inverse of the capacity of a small neighborhood around m_i w.r.t. to a small neighborhood around m_j. The capacity is the crucial ingredient of the works [BEGK04] and [BGK05]. Therefore, we will exploit in Section 4.6 how the *weighted transport distance* (cf. Section 4.1), which is the crucial ingredient to proof (2.24), could lead to a general variational principle for obtaining lower bounds for capacities.

2.4.3. Eyring-Kramers formulas

Now, let us turn to the Eyring-Kramers formulas. Starting from the splitting obtained in Lemma 2.9 and Corollary 2.18, we will see how a combination of Theorem 2.19, Theorem 2.21 and Theorem 2.23 immediately leads to the multidimensional Eyring-Kramers formulas for the Poincaré inequality (cf. [BGK05, Theorem 1.2]) and logarithmic Sobolev inequality.

Corollary 2.26 (Eyring-Kramers formula for Poincaré inequality)**.** *The measure μ satisfies* $\mathrm{PI}(\varrho)$ *with*

$$\frac{1}{\varrho} \lesssim Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 (H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|}\, e^{\frac{H(s_{1,2})}{\varepsilon}}, \tag{2.25}$$

where $\lambda^-(s_{1,2})$ denotes the negative eigenvalue of the Hessian $\nabla^2 H(s_{1,2})$ at the 1-saddle $s_{1,2}$. Further, the order is given such that $H(m_1) \leq H(m_i)$ and $H(s_{1,2}) - H(m_2)$ is the energy barrier of the system in the sense of Assumption 1.12.

Proof. We start from the decomposition of the variance into local variances and mean-differences given by Lemma 2.9. Then, an application of Theorem 2.19 and Theorem 2.23 yields the estimate

$$
\begin{aligned}
\operatorname{var}_\mu(f) &\leq \sum_i Z_i \operatorname{var}_{\mu_i}(f) + \sum_i \sum_{j<i} Z_i Z_j \left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \\
&\lesssim O(\varepsilon) \int |\nabla f|^2 \,\mathrm{d}\mu + \sum_i \sum_{j>i} \frac{Z_i Z_j Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} \, e^{\frac{H(s_{i,j})}{\varepsilon}} \int |\nabla f|^2 \,\mathrm{d}\mu.
\end{aligned} \tag{2.26}
$$

The final step is to observe, that by Assumption 1.12, the exponential dominating term in (2.26) is given for $i = 1$ and $j = 2$. □

In [BGK05, Theorem 1.2]) it is also shown that the upper bound of (2.25) is optimal by an approximation of the harmonic function. Therefore, in the following we can assume that (2.25) holds with $\approx$ instead of $\lesssim$.

Remark 2.27 (Higher exponential small eigenvalues). The statement of [BGK05, Theorem 1.2] does not only characterize the second eigenvalue of L (i.e. the spectral gap) but also the higher exponentially small eigenvalues. In principle, these characterizations can be also obtained in the present approach: The dominating exponential modes in (2.26), i.e. those obtained by setting $i = 1$, correspond to the inverse eigenvalues of L for $j = 2, \ldots, M$. By using the variational characterization of the eigenvalues of the operator L (cf. Lemma 1.5), the other exponentially small eigenvalues may be obtained by restricting the class of test functions f to the orthogonal complement of the eigenspaces of smaller eigenvalues.

Corollary 2.28 (Eyring-Kramers formula for logarithmic Sobolev inequalities)**.** *The measure μ satisfies* LSI(α) *with*

$$
\frac{2}{\alpha} \lesssim \frac{Z_1 Z_2}{\Lambda(Z_1, Z_2)} \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 (H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|} \, e^{\frac{H(s_{1,2})}{\varepsilon}} \approx \frac{1}{\Lambda(Z_1, Z_2)} \frac{1}{\varrho}, \tag{2.27}
$$

where $\lambda^-(s_{1,2})$ denotes the negative eigenvalue of the Hessian $\nabla^2 H(s_{1,2})$ at the 1-saddle $s_{1,2}$. Further, where we assume that the ordering is given such that $H(m_1) \leq H(m_i)$ and $H(s_{1,2}) - H(m_2)$ is the energy barrier of the system in the sense of Assumption 1.12.

Proof. The starting point is the estimate in Corollary 2.18 from which we are left to bound the local entropies and variances as well as the mean-differences. The according bounds are the statements of Theorem 2.19, Theorem 2.21 and Theorem 2.23 and lead to

$$
\begin{aligned}
\operatorname{Ent}_\mu(f^2) \leq{}& O(1) \sum_{i=1}^M Z_i \int |\nabla f|^2 \,\mathrm{d}\mu_i + O(\varepsilon) \sum_{i=1}^M \sum_{j\neq i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \int |\nabla f|^2 \,\mathrm{d}\mu_i \\
&+ \sum_{i=1}^M \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} \, e^{\frac{H(s_{i,j})}{\varepsilon}} \int |\nabla f|^2 \,\mathrm{d}\mu.
\end{aligned} \tag{2.28}
$$

The first term on the right-hand side of (2.28) is just $O(1)\int |\nabla f|^2 \,\mathrm{d}\mu$. For estimating the second term, we have to take care of the expression and use the one-homogeneity of $\Lambda(\cdot,\cdot)$ (cf. Appendix A)

$$\frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} = Z_i \frac{\log \frac{Z_i}{Z_j}}{\frac{Z_i}{Z_j} - 1} = Z_i P\left(\tfrac{Z_i}{Z_j}\right), \qquad \text{where} \quad P(x) := \frac{\log x}{x-1}. \tag{2.29}$$

The function $P(x)$ is decreasing and has a logarithmic singularity in 0. Therefore, we have to check when $\frac{Z_i}{Z_j}$ becomes small. Let us therefore do an asymptotic evaluation of $\frac{Z_i}{Z_j}$, which can be deduced from

$$Z_i Z_\mu = \int_{\Omega_i} e^{-\frac{H}{\varepsilon}} \,\mathrm{d}x = \left(\frac{(2\pi\varepsilon)^{\frac{n}{2}}}{\sqrt{\det \nabla^2 H(m_i)}} + O(\varepsilon^{\frac{n+1}{2}}) \right) e^{-\frac{H(m_i)}{\varepsilon}}. \tag{2.30}$$

This formula immediately leads to the identity

$$\frac{Z_i}{Z_j} \approx \frac{\sqrt{\nabla^2 H(m_j)}}{\sqrt{\nabla^2 H(m_i)}}\, e^{-\frac{H(m_i)-H(m_j)}{\varepsilon}}, \tag{2.31}$$

which becomes exponentially small provided that $H(m_i) > H(m_j)$. In particular, using the expression (2.30) in (2.29) results in

$$\frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} = Z_i O(\varepsilon^{-1}). \tag{2.32}$$

Hence, also the second term in (2.28) can be estimated by $O(1)\int |\nabla f|^2 \,\mathrm{d}\mu$. This shows that the third term dominates the first two on an exponential scale. This leads to the estimate

$$\mathrm{Ent}_\mu(f^2) \lesssim \sum_{i=1}^{M} \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \, \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|}\, e^{\frac{H(s_{i,j})}{\varepsilon}} \int |\nabla f|^2 \,\mathrm{d}\mu.$$

From Assumption 1.12 together with (2.30) and (2.32) follows that the exponential leading order term is given for $i = 1$ and $j = 2$. □

Corollary 2.29 (Comparison of ϱ and α in special cases)**.** *Let us state two specific cases of* (2.25) *and* (2.27). *Firstly, if $H(m_1) < H(m_2)$, it holds*

$$\frac{1}{\varrho} \approx \frac{1}{\sqrt{\det \nabla^2 H(m_2)}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 (H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|}\, e^{\frac{H(s_{1,2})-H(m_2)}{\varepsilon}}, \tag{2.33}$$

$$\frac{2}{\alpha} \lesssim \left(\frac{H(m_2) - H(m_1)}{\varepsilon} + \frac{1}{2}\log\left(\frac{\det \nabla^2 H(m_1)}{\det \nabla^2 H(m_2)}\right)\right) \frac{1}{\varrho}. \tag{2.34}$$

A special case occurs when $H(m_1) = H(m_2)$ and the constants take the form

$$\frac{1}{\varrho} \approx \frac{1}{\sqrt{\det \nabla^2 H(m_1)} + \sqrt{\det \nabla^2 H(m_2)}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 (H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|}\, e^{\frac{H(s_{1,2})-H(m_2)}{\varepsilon}}, \tag{2.35}$$

$$\frac{2}{\alpha} \lesssim \frac{1}{\Lambda\left(\sqrt{\det \nabla^2 H(m_1)}, \sqrt{\det \nabla^2 H(m_2)}\right)} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 (H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|}\, e^{\frac{H(s_{1,2})-H(m_2)}{\varepsilon}}. \tag{2.36}$$

Proof. If $H(m_1) < H(m_2)$, then holds by (2.30) $Z_1 = 1 + O\left(e^{-\frac{H(m_2)-H(m_1)}{\varepsilon}}\right)$. Therefore, the factor $Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}$ evaluates with (2.30) to

$$Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \approx \frac{1}{\sqrt{\det \nabla^2 H(m_2)}} e^{-\frac{H(m_2)}{\varepsilon}},$$

which leads to (2.33). For the logarithmic Sobolev inequality, we additionally have to evaluate the factor $\frac{1}{\Lambda(Z_i,Z_j)}$ which can be done with the help of (2.31)

$$\frac{1}{\Lambda(Z_i, Z_j)} = \log \frac{Z_i}{Z_j} \left(1 + O\left(e^{-\frac{H(m_2)-H(m_1)}{\varepsilon}}\right)\right) \overset{(2.31)}{\approx} \log \left(\frac{\sqrt{\nabla^2 H(m_j)}}{\sqrt{\nabla^2 H(m_i)}} \, e^{-\frac{H(m_i)-H(m_j)}{\varepsilon}} \right).$$

That is already the estimate (2.34).
Let us turn now to the case $H(m_1) = H(m_2)$. Let us introduce the shorthand notation $\kappa_i = \sqrt{\det \nabla^2 H(m_i)}$. Then, we can evaluate Z_μ like in (2.30) and obtain by assuming $H(m_1) = H(m_2) = 0$

$$Z_\mu = \left(\frac{(2\pi\varepsilon)^{\frac{n}{2}}}{\kappa_1} + \frac{(2\pi\varepsilon)^{\frac{n}{2}}}{\kappa_2} + O(\varepsilon^{\frac{n+1}{2}}) \right).$$

Therewith, the factor $Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}$ results with (2.30) in

$$Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} = \frac{(2\pi\varepsilon)^{\frac{n}{2}}}{Z_\mu} \, \frac{Z_1 Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, \frac{Z_2 Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} = \frac{1}{\frac{1}{\kappa_1} + \frac{1}{\kappa_2}} \, \frac{1}{\kappa_1} \, \frac{1}{\kappa_2} = \frac{1}{\kappa_1 + \kappa_2},$$

which precisely leads to the expression (2.35). By using the homogeneity of $\Lambda(\cdot,\cdot)$ (cf. Appendix A) and (2.30) follows for the logarithmic Sobolev inequality

$$\frac{Z_1 Z_2}{\Lambda(Z_1, Z_2)} \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} = \frac{1}{\Lambda\left(\frac{(2\pi\varepsilon)^{\frac{n}{2}}}{Z_2 Z_\mu}, \frac{(2\pi\varepsilon)^{\frac{n}{2}}}{Z_1 Z_\mu} \right)} \approx \frac{1}{\Lambda(\kappa_2, \kappa_1)}.$$

Finally, the result (2.36) is a consequence of the symmetry of $\Lambda(\cdot,\cdot)$. $\square$

Remark 2.30 (Higher degree of symmetry). In Assumption 1.12 (ii), we only allow for two global non-degenerate minima. Equations (2.35) and (2.36) are consequences of this assumption. If there is even more symmetry in the system, then it is also possible to obtain formulas for ϱ and α. In the present generality of the assumptions, the resulting expressions will become cumbersome. Here one needs to rely on specific models and investigate the according symmetries of the model.

Remark 2.31 (Identification of α and ϱ). Remark 1.3 shows that always $\alpha \leq \varrho$. Let us introduce the shorthand notation $\kappa_i = \sqrt{\det \nabla^2 H(m_i)}$. We want to compare the case when $H(m_1) = H(m_2)$ where we observe by comparing (2.35) and (2.36)

$$1 \leq \frac{\varrho}{\alpha} \lesssim \frac{\frac{\kappa_1 + \kappa_2}{2}}{\Lambda(\kappa_1, \kappa_2)}. \tag{2.37}$$

The quotient in (2.37) consists of the arithmetic and logarithmic mean. The lower bound of 1 can also be observed by applying the logarithmic-arithmetic mean inequality from Lemma A.1. Moreover equality only holds for $\kappa_1 = \kappa_2$. Hence, only in the symmetric case with $\kappa_1 = \kappa_2$ holds $\varrho \approx \alpha$.

Remark 2.32 (Relation to mixtures). If $H(m_1) < H(m_2)$, then (2.34) gives

$$\frac{\varrho}{\alpha} \lesssim \frac{1}{2} \log\left(\frac{\kappa_2}{\kappa_1} e^{\frac{H(m_2)-H(m_1)}{\varepsilon}}\right) \approx \frac{1}{2} \left|\log Z_2\right|, \quad \text{where} \quad Z_2 = \mu(\Omega_2). \tag{2.38}$$

which shows a inverse scaling in ε. Different scaling behavior between Poincaré and logarithmic Sobolev constants was also observed by Chafaï and Malrieu [CM10] in a different context. They consider mixtures of probability measures ν_0 and ν_1 satisfying $\mathrm{PI}(\varrho_i)$ and $\mathrm{LSI}(\alpha_i)$, i.e. for $p \in [0,1]$ the measure ν_p given by

$$\nu_p = p\nu_0 + (1-p)\nu_1.$$

Then, they deduce conditions under which also ν_p satisfies $\mathrm{PI}(\varrho_p)$ and $\mathrm{LSI}(\alpha_p)$ and give bounds on the constants. They show in the one-dimensional case examples where the Poincaré constant stays bounded, whereas the logarithmic Sobolev constant blows up logarithmically, when the mixture parameter p goes to 0 or 1. The common feature of the examples they deal with is $\nu_1 \ll \nu_2$ or $\nu_2 \ll \nu_1$. This case can be generalized to the multidimensional case, where also a different scaling of the Poincaré and logarithmic Sobolev constants is observed. The details can be found in Chapter 6.

In the present case the Gibbs measure μ has also a mixture representation given in (2.8). In the two-component case it looks like

$$\mu = Z_1\mu_1 + Z_2\mu_2.$$

Let us emphasize, that $\mu_1 \perp \mu_2$. (2.38) shows also a logarithmic blow-up in the mixture parameter Z_2 for the ratio of the Poincaré and the logarithmic Sobolev constant.

2.5. Optimality of the logarithmic Sobolev constant in one dimension

In this section, we want to give a strong indication, that the result of Corollary 2.28 is optimal. Therefore, we will explicitly construct a function attaining equality in (2.27) for the one dimensional case. This is the same strategy, which was used by [BGK05] to proof the optimality of Corollary 2.26. The proceeding of this section is similar to the one of Felix Otto given in a lecture[1] proving the Eyring-Kramers formula for the Poincaré inequality in one dimension.

Let μ be a probability measure on $\mathbb{R}$ having as Hamiltonian H a generic double-well (cp. Figure 2.1). Namely, H has two minima m_1 and m_2 with $H(m_1) \leq H(m_2)$ and a saddle s in-between. Then, Theorem 2.28 shows

$$\inf_{g:\int g^2 \mathrm{d}\mu = 1} \frac{\int (g')^2 \mathrm{d}\mu}{\int g^2 \log g^2 \mathrm{d}\mu} \gtrsim \frac{\Lambda(Z_1,Z_2)}{Z_1 Z_2} \frac{\sqrt{2\pi\varepsilon}}{Z_\mu} \frac{\sqrt{|H''(s)|}}{2\pi\varepsilon}\, e^{\frac{H(s)-H(m_2)}{\varepsilon}}. \tag{2.39}$$

[1] one of the Ringvorlesungen at the MPI in November 2011

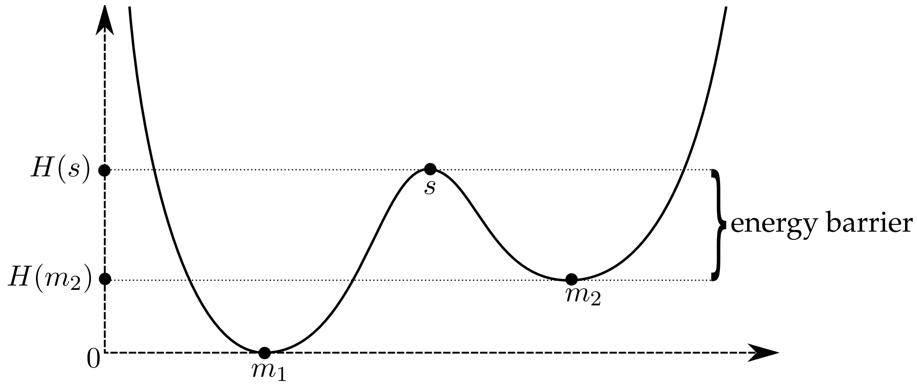

Figure 2.1.: Double-well potential H on $\mathbb{R}$ (labeled).

We have to construct a function g attaining the lower bound given in (2.39) satisfying $H(m_1) \leq H(m_2)$. We make the following ansatz for the function g and firstly define it on some small δ-neighborhoods around the minima m_1, m_2 and the saddle s:

$$g(x) := \begin{cases} g(m_1) & , x \in B_\delta(m_1) \\ g(m_1) + \frac{g(m_2)-g(m_1)}{\sqrt{2\pi\varepsilon\sigma}} \int_{m_1}^{x} e^{-\frac{(y-s)^2}{2\sigma\varepsilon}} \mathrm{d}y & , x \in B_\delta(s) \\ g(m_2) & , x \in B_\delta(m_2). \end{cases}$$

The ansatz contains the parameters $g(m_1)$, $g(m_2)$ and σ. Furthermore, we assume that in-between the δ-neighborhoods g is extended to a smooth function in a monotone fashion.

The measure μ is given by

$$\mu(\mathrm{d}x) = \frac{1}{Z_\mu} e^{-\frac{H(x)}{\varepsilon}}, \qquad \text{where} \qquad Z_\mu = \int e^{-\frac{H(x)}{\varepsilon}} \mathrm{d}x.$$

We fix Z_μ by assuming that $H(m_1) = 0$. We can represent μ as the mixture

$$\mu = Z_1\mu_1 + Z_2\mu_2, \qquad \text{where} \qquad \mu_1 = \mu \llcorner \Omega_1 \quad \text{and} \quad \mu_2 = \mu \llcorner \Omega_2,$$

hereby, $\Omega_1 = (-\infty, s)$ and $\Omega_2 = (s, \infty)$ and $Z_i = \mu(\Omega_i)$ for $i = 1, 2$, which implies $Z_1 + Z_2 = 1$. Then, for the ansatz g, we find via an asymptotic evaluation of $\int g^2 \mathrm{d}\mu$

$$\int g^2 \mathrm{d}\mu \approx Z_1 g^2(m_1) + Z_2 g^2(m_2) \stackrel{!}{=} 1. \tag{2.40}$$

This motivates the choice

$$g^2(m_1) = \frac{\tau}{Z_1} \qquad \text{and} \qquad g^2(m_2) = \frac{1-\tau}{Z_2} = \frac{1-\tau}{1-Z_1}, \quad \text{for some } \tau \in [0,1] \tag{2.41}$$

The choice (2.41) is picked to fulfill the constraint (2.40). Let us now calculate the denominator of (2.39)

$$\int g^2 \log g^2 \mathrm{d}\mu = \tau \log \frac{\tau}{Z_1} + (1-\tau) \log \frac{1-\tau}{Z_2}. \tag{2.42}$$

The final step is to evaluate the Dirichlet energy $\int (g')^2 \mathrm{d}\mu$. Therefore, we do a Taylor expansion of H around s. Furthermore, since s is a saddle, it holds $H''(s) < 0$

$$\begin{aligned}
\int (g')^2 \mathrm{d}\mu &\approx \frac{(g(m_2)-g(m_1))^2}{Z_\mu\, 2\pi\varepsilon\sigma} \int_{B_\delta(s)} e^{-\frac{(x-s)^2}{\sigma\varepsilon} - \frac{H(x)}{\varepsilon}} \mathrm{d}x \\
&\approx \frac{(g(m_2)-g(m_1))^2}{Z_\mu\, 2\pi\varepsilon\sigma} \int_{B_\delta(s)} e^{-\frac{1}{\varepsilon}\left(\frac{(x-s)^2}{\sigma} + H(s) + H''(s)\frac{(x-s)^2}{2}\right)} \mathrm{d}x \\
&\approx \frac{(g(m_2)-g(m_1))^2}{Z_\mu\, 2\pi\varepsilon\sigma}\, e^{-\frac{H(s)}{\varepsilon}} \int_{B_\delta(s)} e^{-\frac{(x-s)^2}{2\varepsilon}\left(\frac{2}{\sigma} + H''(s)\right)} \mathrm{d}x \\
&\approx \left(\sqrt{\frac{\tau}{Z_1}} - \sqrt{\frac{1-\tau}{Z_2}}\right)^2 \frac{\sqrt{2\pi\varepsilon}}{Z_\mu}\, e^{-\frac{H(s)}{\varepsilon}}\, \frac{1}{2\pi\varepsilon} \frac{1}{\sigma\sqrt{\frac{2}{\sigma} + H''(s)}},
\end{aligned} \tag{2.43}$$

where we assume that σ is small enough such that $\frac{2}{\sigma} + H''(s) > 0$. The last step is to minimize the right-hand side of (2.43) in σ, which means to maximize the expression $2\sigma + \sigma^2 H''(s)$ in σ. Elementary calculus results in $\sigma = -\frac{1}{H''(s)} = \frac{1}{|H''(s)|} > 0$ and therefore

$$\int (g')^2 \mathrm{d}\mu \approx \left(\sqrt{\frac{\tau}{Z_1}} - \sqrt{\frac{1-\tau}{Z_2}}\right)^2 \frac{\sqrt{2\pi\varepsilon}}{Z_\mu} \frac{\sqrt{|H''(s)|}}{2\pi\varepsilon}\, e^{-\frac{H(s)}{\varepsilon}}. \tag{2.44}$$

Hence, we have constructed by combining (2.42) and (2.44) an upper bound for the optimization problem (2.39) given by a one-dimensional optimization in the still free parameter $\tau \in (0,1)$

$$\inf_{g : \int g^2 \mathrm{d}\mu = 1} \frac{\int (g')^2 \mathrm{d}\mu}{\int g^2 \log g^2 \mathrm{d}\mu} \lesssim \min_{\tau \in (0,1)} \frac{\left(\sqrt{\frac{\tau}{Z_1}} - \sqrt{\frac{1-\tau}{Z_2}}\right)^2}{\tau \log \frac{\tau}{Z_1} + (1-\tau)\log\frac{1-\tau}{Z_2}} \frac{\sqrt{2\pi\varepsilon}}{Z_\mu} \frac{\sqrt{|H''(s)|}}{2\pi\varepsilon}\, e^{-\frac{H(s)}{\varepsilon}}.$$

The minimum in τ is attained according to Lemma A.3 for $\tau = Z_2$ resulting in

$$\min_{\tau \in (0,1)} \frac{\left(\sqrt{\frac{\tau}{Z_1}} - \sqrt{\frac{1-\tau}{Z_2}}\right)^2}{\tau \log \frac{\tau}{Z_1} + (1-\tau)\log\frac{1-\tau}{Z_2}} = \frac{\Lambda(Z_1, Z_2)}{Z_1 Z_2}.$$

Chapter 3

Local Poincaré and logarithmic Sobolev inequalities

In this chapter, we want to proof the local Poincaré inequality of Theorem 2.19 and the local logarithmic Sobolev inequality of Theorem 2.21. Therefore, we consider only one of the basins of attraction Ω_i for fixed i and we can omit the index i. We will write Ω and μ instead of Ω_i and μ_i respectively. Further, we assume w.l.o.g. that $0 \in \Omega$ is the unique minimum of H in Ω.

Let us begin with stating the two classical conditions for Poincaré and logarithmic Sobolev inequalities. The first one is the *Bakry-Émery criterion* which states the convexity of the Hamiltonian and positive curvature of the underlying space exhibits good mixing for the associate Gibbs measure. Since we are working in the flat space, we formulate it only in terms of the Hessian of the Hamiltonian.

Theorem 3.1 (Bakry-Émery criterion [BE85, Proposition 3, Corollaire 2])**.** *Let H be a Hamiltonian with Gibbs measure $\mu(\mathrm{d}x) = Z_\mu^{-1} e^{-\varepsilon^{-1}H(x)}\mathrm{d}x$ and assume that $\nabla^2 H(x) \geq \lambda > 0$ for all $x \in \mathbb{R}^n$. Then μ satisfies* $\mathrm{PI}(\varrho)$ *and* $\mathrm{LSI}(\alpha)$ *with*

$$\varrho \geq \frac{\lambda}{\varepsilon} \qquad \textit{and} \qquad \alpha \geq \frac{\lambda}{\varepsilon}.$$

The second condition is the *Holley-Stroock perturbation principle*, which allows to show Poincaré and logarithmic Sobolev inequalities for a very large class of measures. However, in general the constant obtained from this principle will be not optimal in terms of scaling with the temperature ε.

Theorem 3.2 (Holley-Stroock perturbation principle [HS87, p. 1184])**.** *Let H be a Hamiltonian with Gibbs measure $\mu(\mathrm{d}x) = Z_\mu^{-1} e^{-\varepsilon^{-1}H(x)}\mathrm{d}x$. Further, let ψ be a bounded function and denote by $\tilde{\mu}$ the Gibbs measure with Hamiltonian $H + \psi$, i.e.*

$$\mu(\mathrm{d}x) = \frac{1}{Z_\mu} e^{-\frac{H(x)}{\varepsilon}}\mathrm{d}x \qquad \textit{and} \qquad \tilde{\mu}(\mathrm{d}x) = \frac{1}{Z_{\tilde{\mu}}} e^{-\frac{H(x)+\psi(x)}{\varepsilon}}\mathrm{d}x.$$

Then, if μ *satisfies* $\mathrm{PI}(\varrho)$ *or* $\mathrm{LSI}(\alpha)$ *then also* $\tilde{\mu}$ *satisfy* $\mathrm{PI}(\tilde{\varrho})$ *or respectively* $\mathrm{LSI}(\tilde{\alpha})$. *Hereby the constants satisfy the relations*

$$\tilde{\varrho} \geq e^{-\frac{\operatorname{osc}\psi}{\varepsilon}}\varrho \qquad \text{and} \qquad \tilde{\alpha} \geq e^{-\frac{\operatorname{osc}\psi}{\varepsilon}}\alpha, \tag{3.1}$$

where $\operatorname{osc}\psi := \sup\psi - \inf\psi$.

For the proofs relying on semigroup theory of Theorem 3.1 and Theorem 3.2 we refer to the exposition by Ledoux [Led99b, Corollary 1.4, Corollary 1.6 and Lemma 1.2]. The only difference is, that we always explicitly express the temperature ε and consider H being ε-independent.

Let us summarize the reasons, why we cannot directly apply the above standard criteria for the Poincaré and logarithmic Sobolev inequalities to a Hamiltonian restricted to the basin of attraction of a local minimum.

- The criterion of Bakry-Émery [BE85] does not cover the present situation, because in general H is not convex on the basin of attraction Ω.
- The perturbation principle of Holley-Stroock [HS87] cannot be applied naively because it would yield an exponentially bad dependence of the Poincaré constant ϱ on ε.

Nevertheless, we will use both of them in the proof. The perturbation principle of Holley-Stroock will be used very carefully. In particular, we will compare the measure μ with a measure $\tilde{\mu}$, which is obtained from the construction of a perturbed Hamiltonian $\tilde{H}_\varepsilon$ such that $\|H - \tilde{H}_\varepsilon\|_\infty = O(\varepsilon)$ in Ω. The condition of slight perturbation allows to compare the Poincaré and logarithmic Sobolev constants of μ and $\tilde{\mu}$ upto an ε-independent factor. The second step consists of a Lyapunov argument developed by Bakry, Barthe, Cattiaux, Guillin, Wang and Wu (cf. [BBCG08], [BCG08], [CG10], [CGW09] and [CGWW09]). The Lyapunov conditions shows similarities to the characterization of the spectral gap by Donsker and Varadhan [DV76]. We will state a *Lyapunov function* for $\tilde{\mu}$, which will allow to compare the scaling behavior for the Poincaré and logarithmic Sobolev constants with the truncated Gibbs measure $\hat{\mu}_a$ (cf. Definition 3.5 and Lemma 3.6).

The following definition is motivated by the Holley-Stroock perturbation principle and becomes eminent from the subsequent Lemma 3.4.

Definition 3.3 (ε-modification $\tilde{H}_\varepsilon$ of H)**.** We say that $\tilde{H}_\varepsilon$ is a ε-modification of H, if for all ε small enough $\tilde{H}_\varepsilon$ is of class $C^2(\Omega) \cap C^0(\overline{\Omega})$ and satisfies the condition: $\tilde{H}_\varepsilon$ is ε-close to H, i.e. there exists an ε-independent constant $C_{\tilde{H}} > 0$ such that

$$\forall x \in \Omega: \ |\tilde{H}_\varepsilon(x) - H(x)| \leq C_{\tilde{H}}\varepsilon. \tag{$\tilde{\mathbf{H}}_\varepsilon$}$$

The associated modified Gibbs measure $\tilde{\mu}$ obtained from the ε-modified Hamiltonian $\tilde{H}_\varepsilon$ of H is given by

$$\tilde{\mu}(\mathrm{d}x) = \frac{1}{Z_{\tilde{\mu}}} e^{-\frac{\tilde{H}_\varepsilon}{\varepsilon}} \mathrm{d}x.$$

Lemma 3.4 (Perturbation by an ε-modification)**.** *If the ε-modified Gibbs measure $\tilde{\mu}$ satisfy* $\mathrm{PI}(\tilde{\varrho})$ *and* $\mathrm{LSI}(\tilde{\alpha})$ *then the associated measure μ also satisfies* $\mathrm{PI}(\varrho)$ *and* $\mathrm{LSI}(\alpha)$*, where the constants fulfill the estimate*

$$\varrho \geq e^{-2C_{\tilde{H}}}\tilde{\varrho} \qquad \textit{and} \qquad \alpha \geq e^{-2C_{\tilde{H}}}\tilde{\alpha}. \tag{3.2}$$

Proof. We just can apply Theorem 3.2 with H replaced by $\tilde{H}$ and $\psi = H - \tilde{H}$. Finally, observe that by ($\tilde{\mathbf{H}}_\varepsilon$) holds

$$\operatorname{osc}\psi = \sup(H - \tilde{H}) - \inf(H - \tilde{H}) \leq 2|H - \tilde{H}| \leq 2C_{\tilde{H}}\varepsilon.$$

Therewith, the bound (3.1) becomes (3.2). □

Definition 3.5 (Truncated Gibbs measure)**.** To the Gibbs measure μ we associate by $\hat{\mu}_a$ the truncated measure obtained from μ by restricting it to a ball of radius $a\sqrt{\varepsilon}$ around 0 for some $a > 0$

$$\hat{\mu}_a(\mathrm{d}x) = \frac{1}{Z_{\hat{\mu}_a}}\, \mathbb{1}_{B_{a\sqrt{\varepsilon}}}(x)\, e^{-\frac{H(x)}{\varepsilon}}\mathrm{d}x.$$

Lemma 3.6 (PI and LSI for truncated Gibbs measure)**.** *The measure $\hat{\mu}_a$ satisfies* $\mathrm{PI}(\hat{\varrho})$ *and* $\mathrm{LSI}(\hat{\alpha})$ *for ε small enough, where*

$$\frac{1}{\hat{\varrho}} = O(\varepsilon) \qquad \textit{and} \qquad \frac{1}{\hat{\alpha}} = O(\varepsilon). \tag{3.3}$$

Proof. In the local minimum 0 of Ω the Hessian of H is non-degenerated by Assumption 1.9 or 1.10. Therefore, for ε small enough, H is strictly convex in $B_{a\sqrt{\varepsilon}}$ and satisfies by the Bakry-Émery criterion (cf. Theorem 3.1) $\mathrm{PI}(\hat{\varrho})$ and $\mathrm{LSI}(\hat{\alpha})$ with $\hat{\varrho}$ and $\hat{\alpha}$ obeying the relation (3.3). □

3.1. Lyapunov conditions . . .

3.1.1. . . . for Poincaré inequality

In this subsection, we will show that there exists an ε-modified Hamiltonian $\tilde{H}_\varepsilon$ which ensures that the Poincaré constant of $\tilde{\mu}$ is of the same order as the Poincaré constant of the truncated measure $\hat{\mu}_a$ from Definition 3.5. Therefore, we will state a Lyapunov function for the measure $\tilde{\mu}$. Firstly, let us introduce the notion of a Lyapunov condition.

Definition 3.7 (Lyapunov condition for Poincaré inequality)**.** Let $H : \Omega \to \mathbb{R}$ be a Hamiltonian and let

$$\mu(\mathrm{d}x) = \frac{\mathbb{1}_\Omega(x)}{Z_\mu}\, e^{-\frac{H(x)}{\varepsilon}}\mathrm{d}x$$

denote the associated Gibbs measure μ at temperature ε. Then, $W : \Omega \to [1, \infty)$ is a Lyapunov function for H provided there exist constants $R > 0$, $b > 0$ and $\lambda > 0$ such that

$$\frac{1}{\varepsilon} LW \le -\lambda W + b\, \mathbb{1}_{\Omega_R}. \tag{3.4}$$

[BBCG08] is recommended for further information on the use of Lyapunov conditions for deducing Poincaré inequalities. The main ingredient of this technique is the following statement:

Theorem 3.8 (Lyapunov condition for PI [BBCG08, Theorem 1.4.]). *Suppose that H fulfills the Lyapunov condition* (3.4) *and that the restricted measure μ_R given by*

$$\mu_R(\mathrm{d}x) = \mu(\mathrm{d}x)\llcorner_{\Omega_R} = \frac{\mathbb{1}_{\Omega_R}(x)}{\mu(\Omega_R)}\, \mu(\mathrm{d}x), \qquad \textit{where} \qquad \Omega_R = \Omega \cap B_R$$

satisfies $\mathrm{PI}(\varrho_R)$*. Then, μ also satisfies* $\mathrm{PI}(\varrho)$ *with constant*

$$\varrho \ge \frac{\lambda}{b + \varrho_R}\, \varrho_R.$$

Proof. Having the Lyapunov condition (3.4) the proof of Theorem 3.8 becomes simple and only relies on the symmetry of L in $L^2(\mu)$. Let us just outline the argument for which more the details can be found in [BBCG08]. Let us rewrite the Lyapunov condition (3.4) and observe

$$1 \le -\frac{LW}{\varepsilon\lambda W} + \frac{b}{\lambda}\mathbb{1}_{\Omega_R}. \tag{3.5}$$

The next observation to use the symmetry of $\frac{1}{\varepsilon}L$ in $L^2(\mathrm{d}\mu)$, i.e.

$$\int f(-\frac{1}{\varepsilon}Lg)\mathrm{d}\mu = \int \langle \nabla f, \nabla g\rangle\, \mathrm{d}\mu$$

for deducing the estimate

$$\begin{aligned}\int f^2 \frac{(-LW)}{\varepsilon W}\mathrm{d}\mu &= \int \left\langle \nabla\left(\frac{f^2}{W}\right), \nabla W \right\rangle \mathrm{d}\mu \\ &= 2\int \frac{f}{W}\langle \nabla f, \nabla W\rangle\, \mathrm{d}\mu - \int \frac{f^2 \left|\nabla W\right|^2}{W^2}\mathrm{d}\mu \\ &= \int |\nabla f|^2\, \mathrm{d}\mu - \int \left|\nabla f - \frac{f}{W}\nabla W\right|^2 \mathrm{d}\mu \\ &\le \int |\nabla f|^2\, \mathrm{d}\mu.\end{aligned} \tag{3.6}$$

Therefore, we can estimate the variance by first using $\mathrm{var}_\mu(f) \le \int (f-m)^2\, \mathrm{d}\mu$ for any $m \in \mathbb{R}$ with equality if and only if $m = \int f \mathrm{d}\mu$ and then estimating with the help of (3.5) and (3.6) as well as the local Poincaré inequality for μ_R

$$\begin{aligned}\mathrm{var}_\mu(f) &\le \int (f-m)^2\, \mathrm{d}\mu \le \int (f-m)^2 \frac{-LW}{\varepsilon\lambda W} + \frac{b}{\lambda}\int_{\Omega_R} (f-m)^2\, \mathrm{d}\mu \\ &\le \frac{1}{\lambda}\int |\nabla f|^2\, \mathrm{d}\mu + \frac{b}{\lambda}\int_{\Omega_R} (f-m)^2\, \mathrm{d}\mu.\end{aligned} \tag{3.7}$$

The last step consists of setting $m = \int f \mathrm{d}\mu_R$ and observing that the last integral in the right-hand side of (3.7) becomes $\mathrm{var}_{\mu_R}(f)$, which can be estimated since μ_R satisfies $\mathrm{PI}(\varrho_R)$ by assumption. □

We want to apply Theorem 3.8 to our situation. Hence, we do not only have to verify the Lyapunov condition (3.4) but also have to investigate the dependence of the constants R, b and λ on the parameter ε.

We will use the Lyapunov approach given in Definition 3.7 and will explicitly construct an $\varepsilon-$modification $\tilde{H}_\varepsilon$ of H in the sense of Definition 3.3. More precisely, we deduce the following statement:

Lemma 3.9 (Lyapunov function for PI)**.** *Without loss of generality we may assume that $0 \in \Omega$ is the unique minimum of H in Ω. Then, there exits an $\varepsilon-$modification $\tilde{H}_\varepsilon$ of H in the sense of Definition 3.3 such that for some constant $a > 0$ large enough holds with $\lambda_0 > 0$*

$$\frac{1}{2\varepsilon}\Delta \tilde{H}_\varepsilon(x) - \frac{1}{4\varepsilon^2}|\nabla \tilde{H}_\varepsilon(x)|^2 \le -\frac{\lambda_0}{\varepsilon} \qquad \textit{for all} \quad |x| \ge a\sqrt{\varepsilon}. \tag{3.8}$$

In particular, $\tilde{H}_\varepsilon$ satisfies the Lyapunov condition (3.4) *with Lyapunov function*

$$W(x) = \exp\left(\frac{1}{2\varepsilon}\tilde{H}_\varepsilon(x)\right) \quad \textit{and constants} \quad R = a\sqrt{\varepsilon},\ b \le \frac{b_0}{\varepsilon},\ \textit{and}\ \lambda \ge \frac{\lambda_0}{\varepsilon}. \tag{3.9}$$

If the above lemma holds true the content of the local Poincaré inequality of Theorem 2.19 is just a simple consequence of a combination of Theorem 3.8 and Lemma 3.4. We will outline the proof in Section 3.2

Likewise, the statement of Lemma 3.9 directly follows from the following two observations.

Lemma 3.10. *Assume that the Hamiltonian H satisfies the Assumption* ($\mathbf{A2}_{\mathtt{PI}}$)*. Then, there is a constant $0 \le C_H < \infty$ and $0 \le \tilde{R} < \infty$ such that*

$$\frac{\Delta H(x)}{2\varepsilon} - \frac{|\nabla H(x)|^2}{4\varepsilon^2} \le -\frac{C_H}{\varepsilon} \qquad \textit{for all } |x| \ge \tilde{R}. \tag{3.10}$$

Moreover, let us assume that H is a Morse function in the sense of Definition 1.8. Additionally, let $\mathcal{S}$ denote the set of all critical points of H in Ω; that is

$$\mathcal{S} = \{y \in \Omega \mid \nabla H(y) = 0\}.$$

Then, there exists a constant $0 < c_H$ depending only on H such for $a > 0$ and ε small enough holds

$$|\nabla H(x)| \ge c_H a\sqrt{\varepsilon} \qquad \textit{for all } x \notin \bigcup_{y\in\mathcal{S}} B_{a\sqrt{\varepsilon}}(y). \tag{3.11}$$

In particular, this implies that there is a constant $C_H > 0$ such that

$$\frac{\Delta H(x)}{2\varepsilon} - \frac{|\nabla H(x)|^2}{4\varepsilon^2} \le -\frac{C_H}{\varepsilon} \qquad \textit{for all } x \in B_{\tilde{R}}(0) \setminus \bigcup_{y\in\mathcal{S}} B_{a\sqrt{\varepsilon}}(y). \tag{3.12}$$

Proof. The proof basically consists only of elementary calculations based on the non-degeneracy assumption on H. For showing (3.10) we use the assumptions ($\mathbf{A1}_{\mathtt{PI}}$) and ($\mathbf{A2}_{\mathtt{PI}}$). Therefore, we define $\tilde{R}$ such that

$$\forall\,|x| \geq \tilde{R}: \qquad |\nabla H| \geq \frac{C_H}{2} \quad \text{and} \quad |\nabla H| - \Delta H(x) \geq -2K_H.$$

Therewith, it is easy to show, that for $|x| \geq \tilde{R}$ holds

$$\frac{\Delta H(x)}{2\varepsilon} - \frac{|\nabla H(x)|^2}{4\varepsilon^2} \leq \frac{1}{\varepsilon}\left(K_H - \frac{1}{2}\left(\frac{1}{2\varepsilon} - 1\right)\frac{C_H^2}{4}\right) \leq -\frac{C_H^2}{32\varepsilon} \quad \text{for} \quad \varepsilon \leq \frac{1}{4}\frac{C_H^2}{C_H^2 + 8K_H},$$

which proves the statement (3.10). The condition (3.11) is first checked for a δ-neighborhood around the critical points $y \in \mathcal{S}$. There, by the Morse assumption on H (cp. Assumption 1.9 and Definition 1.8), we can do a Taylor expansion of H around the critical point y and find for $x \in B_\delta(y)\backslash B_{a\sqrt{\varepsilon}}(y)$

$$|\nabla H(x)| \geq \left|\lambda_{\min}(\nabla^2 H(y))\right| a\sqrt{\varepsilon} + O(\delta^2). \tag{3.13}$$

This shows, that (3.11) holds for $x \in B_\delta(y)\backslash B_{a\sqrt{\varepsilon}}(y)$. To conclude, we assume that (3.11) does not hold for some critical point y, i.e. for every $\varepsilon > 0$ and $c_H > 0$ and $a > 0$ we find $x \notin B_{a\sqrt{\varepsilon}}(y)$ such that $|\nabla H(x)| \leq c_H a\sqrt{\varepsilon}$, which by (3.13) contradicts the Morse assumption (1.8) for ε small enough. Finally, (3.12) is a conclusion of a combination of (3.10) and (3.11). □

The second observation needed for the verification of Lemma 3.9 is given by the following statement, which is the main ingredient for the proof of the local Poincaré inequalities.

Lemma 3.11. *On a basin of attraction Ω there exists an ε−modification $\tilde{H}_\varepsilon$ of H in the sense of Definition 3.3 satisfying*

(i) The modification $\tilde{H}_\varepsilon$ equals H except for small neighborhoods around the critical points except the local minimum of H, i.e.

$$\tilde{H}_\varepsilon(x) = H(x), \qquad \textit{for all } x \notin \bigcup_{y\in\mathcal{S}\backslash\{0\}} B_{a\sqrt{\varepsilon}}(y).$$

(ii) There are constants $0 < C_H$ and $a > 1$ such that for all small ε it holds

$$\frac{\Delta \tilde{H}_\varepsilon(x)}{2\varepsilon} - \frac{|\nabla \tilde{H}_\varepsilon(x)|^2}{4\varepsilon^2} \leq -\frac{C_H}{\varepsilon} \qquad \textit{for all} \quad x \in \bigcup_{y\in\mathcal{S}\backslash\{0\}} B_{a\sqrt{\varepsilon}}(y). \tag{3.14}$$

Proof of Lemma 3.11. By the property (i) of Lemma 3.11, it is sufficient to construct the ε-modification $\tilde{H}_\varepsilon$ on a small neighborhood of any critical point y, which is not the global minimum of H in Ω. By translation, we may assume w.l.o.g. that $y = 0$.
Because the Hamiltonian H is a Morse function in the sense of Definition 1.8, we may assume that u_i, $i \in \{1, \ldots, n\}$ are orthonormal eigenvectors w.r.t. the Hessian $\nabla^2 H(0)$.

The corresponding eigenvalues are denoted by λ_i, $i \in \{1, \ldots, n\}$. Additionally, we may assume w.l.o.g. that $\lambda_1, \ldots, \lambda_\ell < 0$ and $\lambda_{\ell+1}, \ldots, \lambda_n > 0$ for some $\ell \in \{1, \ldots, n\}$. If all $\lambda_i < 0$, we set $\tilde{H}_\varepsilon(x) = H(x)$ on $B_{a\sqrt{\varepsilon}}(0)$ and directly observe the desired statement (ii).
For the construction of $\tilde{H}_\varepsilon$, we need a smooth auxiliary function $\xi : [0, \infty) \to \mathbb{R}$ that satisfies

$$\xi'(z) = -1 \quad \text{for } |z| \le \frac{a}{2}\sqrt{\varepsilon} \qquad \text{and} \qquad -1 \le \xi'(z) \le 0 \quad \text{for } z \in [0, \infty) \tag{3.15}$$

as well as for some $C_\xi > 0$ and any $|z| \le a\sqrt{\varepsilon}$

$$\left|\xi''(z)\right| \le \frac{C_\xi}{\sqrt{\varepsilon}} \qquad \text{and} \qquad \xi(z) = \xi'(z) = \xi''(z) = 0 \quad \text{for } |z| \ge a\sqrt{\varepsilon}. \tag{3.16}$$

Let us choose a constant $\delta > 0$ small enough such that

$$-\tilde{\delta} := (n - 2\ell)\delta + \sum_{i=1}^{\ell} \lambda_i < 0 \qquad \text{and} \qquad \delta \le \frac{1}{2} \min\{\lambda_i : i = \ell + 1, \ldots, n\}. \tag{3.17}$$

Because $u_1, \ldots, u_n$ is an orthonormal basis of $\mathbb{R}^n$, we introduce a norm $|\cdot|_\delta$ on $\mathbb{R}^n$ by

$$|x|_\delta^2 := \sum_{i=1}^{\ell} \frac{1}{2}\delta \, |\langle u_i, x\rangle|^2 + \sum_{i=\ell+1}^{n} \frac{1}{2}(\lambda_i - \delta) \, |\langle u_i, x\rangle|^2$$

The norm $|\cdot|_\delta$ is equivalent to the standard euclidean norm $|\cdot|$ and satisfies the estimate

$$\frac{\delta}{2} |x|^2 \ \le \ |x|_\delta^2 \ \le \ \frac{\lambda_{\max}^+ - \delta}{2} |x|^2 \le, \tag{3.18}$$

where $\lambda_{\max}^+ = \max\{\lambda_i : i = \ell + 1, \ldots, n\}$. With the help of the function

$$H_b(x) := \xi\left(|x|_\delta^2\right), \tag{3.19}$$

we define the ε-modification $\tilde{H}_\varepsilon$ of H on a small neighborhood of the critical point 0 as

$$\tilde{H}_\varepsilon(x) = H(x) + H_b(x).$$

Note that by definition of H_b holds $\tilde{H}_\varepsilon(x) = H(x)$ for all $|x|_\delta \ge a\sqrt{\varepsilon}$. Therefore, the property (i) of Lemma 3.10 is satisfied by the equivalence of norms on finite dimensional vectorspaces (3.18).
For the verification of the statement (ii) of Lemma 3.10, it is sufficient to deduce the following two facts: The first one is the estimate

$$\Delta\tilde{H}(x) \le -\frac{\tilde{\delta}}{2} \qquad \text{for all } |x|_\delta \le \frac{a}{2}\sqrt{\varepsilon}. \tag{3.20}$$

The second one is that there is a constant $0 < C_H$ such that for a large $a > 1$ and small enough ε holds

$$\frac{\Delta\tilde{H}_\varepsilon(x)}{2} - \frac{|\nabla\tilde{H}_\varepsilon(x)|^2}{4\varepsilon} \le -C_H \qquad \text{for all } \ \frac{a}{2}\sqrt{\varepsilon} \le |x|_\delta \le a\sqrt{\varepsilon}. \tag{3.21}$$

Let us have a look at (3.20). Because the function ξ has derivative -1 for $|x|_\delta \leq \frac{a}{2}\sqrt{\varepsilon}$, straightforward calculation yields

$$\nabla^2 \tilde{H}(x) = \nabla^2 H(x) - \sum_{i=1}^{\ell} \delta\, u_i \otimes u_i - \sum_{i=\ell+1}^{n} (\lambda_i - \delta)\, u_i \otimes u_i.$$

Taking the trace in the above identity results in

$$\Delta \tilde{H}(x) = \Delta H(x) - \sum_{i=\ell+1}^{n} \lambda_i + (n - 2\ell)\delta.$$

By the Taylor formula there is a constant $0 \leq C < \infty$ such that

$$|\Delta H(x) - \Delta H(0)| \leq C|x|.$$

Therefore, we get for $|x|_\delta \leq \frac{a}{2}\sqrt{\varepsilon}$

$$\begin{aligned}\Delta \tilde{H}(x) &= \Delta H(0) - \sum_{i=\ell+1}^{n} \lambda_i + (n-2\ell)\delta + \Delta H(x) - \Delta H(0) \\ &\leq \sum_{i=1}^{\ell} \lambda_i + (n-2\ell)\delta + C\frac{a}{2}\sqrt{\varepsilon} \overset{(3.17)}{\leq} -\tilde{\delta} + Ca\sqrt{\varepsilon} \leq -\frac{\tilde{\delta}}{2}, \quad \text{for} \quad \sqrt{\varepsilon} \leq \frac{2\tilde{\delta}}{Ca},\end{aligned}$$

which yields the desired statement (3.20).
Let us turn to the verification of (3.21). On the one hand, straightforward calculation reveals that there exists a constant $0 < C_\Delta < \infty$ such that

$$\Delta \tilde{H}(x) \leq C_\Delta \qquad \text{for all } \frac{a}{2}\sqrt{\varepsilon} < |x|_\delta < a\sqrt{\varepsilon}. \tag{3.22}$$

Indeed, we observe

$$\begin{aligned}\Delta \tilde{H}_\varepsilon(x) &= \Delta H(x) + \xi''(|x|_\delta^2)\left|\nabla |x|_\delta^2\right|^2 + \underbrace{\xi'(|x|_\delta^2)}_{\leq 0}\ \underbrace{\Delta |x|_\delta^2}_{\geq 0} \\ &\overset{(3.16)}{\leq} \Delta H(x) + \frac{C_\xi}{\sqrt{\varepsilon}} \left| \sum_{i=1}^{\ell} \delta\, \langle u_i, x\rangle\, u_i + \sum_{i=\ell+1}^{n} (\lambda_i - \delta)\, \langle u_i, x\rangle\, u_i \right|^2 \\ &\leq \Delta H(x) + \frac{C_\xi}{\sqrt{\varepsilon}}\, \lambda_{\max}^{+}\, |x|^2 \leq C_H + \frac{C_\xi}{\sqrt{\varepsilon}}\, a^2 \varepsilon \leq C_H + C_\xi a^2 \sqrt{\varepsilon} \leq C_\Delta,\end{aligned}$$

for some C_Δ and ε small enough, which yields (3.22).

On the other hand, we will deduce that there is a constant $0 < c_\nabla < \infty$ such that

$$|\nabla \tilde{H}_\varepsilon(x)|^2 \geq c_\nabla\ a^2\ \varepsilon \qquad \text{for all } \frac{a}{2}\sqrt{\varepsilon} < |x|_\delta < a\sqrt{\varepsilon}. \tag{3.23}$$

We want to note that the observations (3.22) and (3.23) already yield the desired statement (3.21). Indeed, we get for $a^2 \geq 4\frac{C_\Delta}{c_\nabla}$

$$\frac{\Delta \tilde{H}_\varepsilon(x)}{2\varepsilon} - \frac{|\nabla \tilde{H}_\varepsilon(x)|^2}{4\varepsilon^2} \leq \frac{C_\Delta}{2\varepsilon} - \frac{c_\nabla a^2}{4\varepsilon} \leq -\frac{C_\Delta}{2\varepsilon} \qquad \text{for all } \frac{a}{2}\sqrt{\varepsilon} < |x|_\delta < a\sqrt{\varepsilon},$$

which is the desired statement (3.21). Therefore, it is only left to deduce the estimate (3.23). By the definition of $\tilde{H}_\varepsilon$ from above, we can write

$$|\nabla \tilde{H}_\varepsilon(x)|^2 = |\nabla H(x)|^2 + |\nabla H_b(x)|^2 + 2\langle \nabla H(x), \nabla H_b(x)\rangle . \tag{3.24}$$

Let us have a closer look at each term on the right-hand side of the last identity and let us start with the first term. By Taylor's formula we obtain

$$|\nabla H(x) - \nabla^2 H(0)x| \le C_\nabla |x|_\delta^2 \tag{3.25}$$

where $0 < C_\nabla < \infty$ denotes a generic constant. Therefore, we can estimate

$$|\nabla H(x)|^2 \ge |\nabla^2 H(0)x|^2 - C_\nabla a^4 \varepsilon^2 \qquad \text{for } |x|_\delta \le a\sqrt{\varepsilon}. \tag{3.26}$$

By the definition of $\lambda_1, \ldots \lambda_n$, we also know

$$|\nabla^2 H(0)x|^2 = \sum_{i=1}^n \lambda_i^2 \, |\langle u_i, x\rangle|^2 . \tag{3.27}$$

Let us have a closer look at the second term in (3.24), namely $|\nabla H_b(x)|^2$. From the definition (3.19) of $|\nabla H_b(x)|^2$ follows

$$\begin{aligned} |\nabla H_b(x)|^2 &= |\xi'(|x|_\delta^2)|^2 \left(\sum_{i=1}^{\ell} \delta^2 \, |\langle u_i, x\rangle|^2 + \sum_{i=\ell+1}^{n} (\lambda_i - \delta)^2 \, |\langle u_i, x\rangle|^2 \right) \\ &\le 2\lambda_{\max}^+ \, |x|_\delta^2 . \end{aligned} \tag{3.28}$$

Now, we turn the the analysis of the last term, namely $2\langle \nabla H(x), \nabla H_b(x)\rangle$. By using the estimates (3.25) and (3.28), we get for $|x|_\delta \le a\sqrt{\varepsilon}$.

$$\begin{aligned} &\langle \nabla H(x), \nabla H_b(x)\rangle = \left\langle \nabla^2 H(0)x, \nabla H_b(x)\right\rangle + \left\langle \nabla H(x) - \nabla^2 H(0)x, \nabla H_b(x)\right\rangle \\ &\overset{(3.25)}{\ge} \left\langle \nabla^2 H(0)x, \nabla H_b(x)\right\rangle - 2C_\nabla \lambda_{\max} |x|_\delta^3 \\ &\ge -\sum_{i=1}^{\ell} \lambda_i \delta \left|\xi'\left(|x|_\delta^2\right)\right| |\langle u_i, x\rangle|^2 - \sum_{i=\ell+1}^{n} \lambda_i(\lambda_i - \delta) \left|\xi'\left(|x|_\delta^2\right)\right| |\langle u_i, x\rangle|^2 - 2C_\nabla \lambda_{\max} a^3 \varepsilon^{\frac{3}{2}} . \end{aligned} \tag{3.29}$$

Combining now the estimates and identities (3.24), (3.26), (3.27), (3.28) and (3.29), we arrive for $|x|_\delta \le a\sqrt{\varepsilon}$ at

$$\begin{aligned} |\nabla \tilde{H}_\varepsilon(x)|^2 \ge &\sum_{i=1}^{\ell} \Big(\lambda_i - \delta \left|\xi'\left(|x|_\delta^2\right)\right|\Big)^2 |\langle u_i, x\rangle|^2 \\ &+ \sum_{i=\ell+1}^{n} \Big(\lambda_i - (\lambda_i - \delta) \left|\xi'\left(|x|_\delta^2\right)\right|\Big)^2 |\langle u_i, x\rangle|^2 - 4C_\nabla \lambda_{\max}^+ a^3 \varepsilon^{\frac{3}{2}} . \end{aligned}$$

By (3.15) holds $|\xi'(|x|_\delta^2)| \le 1$, which applied to the last inequality yields

$$|\nabla \tilde{H}_\varepsilon(x)|^2 \ge \delta^2 \sum_{i=1}^n |\langle u_i, x\rangle|^2 - 4C_\nabla \lambda_{\max}^+ a^3 \varepsilon^{\frac{3}{2}} .$$

Because $u_1, \dots u_n$ is an orthonormal basis of $\mathbb{R}^n$, the desired statement (3.23) follows for $\frac{a\sqrt{\varepsilon}}{2} \leq |x|_\delta \leq a\sqrt{\varepsilon}$ from

$$|\nabla \tilde{H}_\varepsilon(x)|^2 \geq = \delta^2|x|^2 - 4C_\nabla \lambda^+_{\max} a^3 \varepsilon^{\frac{3}{2}} \overset{(3.18)}{\geq} \frac{2\delta^2}{\lambda^+_{\max}}|x|^2_\delta - 4C_\nabla \lambda^+_{\max} a^3 \varepsilon^{\frac{3}{2}}$$
$$\geq \frac{\delta^2}{2\lambda^+_{\max}}\, a^2\varepsilon - 4C_\nabla \lambda^+_{\max} a^3 \varepsilon^{\frac{3}{2}} \geq c_\nabla a^2 \varepsilon$$

for some $c_\nabla < \frac{\delta^2}{2\lambda^+_{\max}}$ and ε small enough. □

We have collected all auxiliary results needed in the proof of Lemma 3.9.

Proof of Lemma 3.9. The condition (3.8) is a consequence of (3.10) from Lemma 3.10 and of (3.14) from Lemma 3.11. Now, we verify the Lyapunov condition (3.4) and calculate with $W = \exp\left(\frac{1}{2\varepsilon}\tilde{H}_\varepsilon\right)$

$$\frac{1}{\varepsilon}\frac{LW}{W} = \frac{1}{2\varepsilon}\Delta\tilde{H}_\varepsilon + \frac{1}{4\varepsilon^2}|\nabla\tilde{H}_\varepsilon|^2 - \frac{1}{2\varepsilon^2}|\nabla\tilde{H}_\varepsilon|^2 = \frac{1}{2\varepsilon}\Delta\tilde{H}_\varepsilon - \frac{1}{4\varepsilon^2}|\nabla\tilde{H}_\varepsilon|^2.$$

Choosing $|x| \geq a\sqrt{\varepsilon} := R$ one obtains from (3.8) the estimate $\lambda \geq \frac{\lambda_0}{\varepsilon}$. If $|x| \leq a\sqrt{\varepsilon}$, we note that $\tilde{H}_\varepsilon = H$ in $B_{a\sqrt{\varepsilon}}(0)$. Furthermore, H is quadratic around 0 and therefore is bounded by $H(x) \leq C_H a^2 \varepsilon$ for $|x| \leq a\sqrt{\varepsilon}$. Using, this in the definition of W, we arrive at the bound for $|x| \leq a\sqrt{\varepsilon}$

$$W(x) = e^{\frac{1}{2\varepsilon}H(x)} \leq e^{\frac{C_H a^2}{2}}.$$

This yields the desired estimate on the constant b, namely for $|x| \leq a\sqrt{\varepsilon}$

$$\frac{1}{\varepsilon}LW(x) \leq \frac{1}{2\varepsilon}\Delta H(x)W(x) \leq \frac{C_H e^{\frac{C_H a^2}{2}}}{\varepsilon} =: \frac{b_0}{\varepsilon},$$

which finishes the proof. □

3.1.2. ... for logarithmic Sobolev inequality

The Lyapunov condition for proving a logarithmic Sobolev inequality is stronger than the one for Poincaré inequality. Nevertheless, the construction of the ε-modified Hamiltonian $\tilde{H}_\varepsilon$ from the previous section carries over and we can mainly use the same Lyapunov function as for the Poincaré inequality. The Lyapunov condition for logarithmic Sobolev inequalities goes back to the work of Cattiaux, Guillin, Wand and Wu [CGWW09]. We will apply the results in the form of the work [CGW09]. We will restate the proofs of the main results in [CGW09] for two reasons: Firstly, to adopt the notation to the low temperature regime and more importantly, to work out the explicit dependence between the constants of the Lyapunov condition, the logarithmic Sobolev constant and especially their ε-dependence.

Theorem 3.12 (Lyapunov condition for LSI [CGW09, Theorem 1.2]). *Suppose that there exists a C^2-function $W : \Omega \to [1, \infty)$ and constants $\lambda, b > 0$ such that for $L = \varepsilon\Delta - \nabla H \cdot \nabla$ holds*

$$\forall x \in \Omega : \frac{1}{\varepsilon}\frac{LW}{W} \leq -\lambda |x|^2 + b. \tag{3.30}$$

Further assume, that $\nabla^2 H \geq -K_H$ for some $K_H > 0$ and μ satisfies $\mathrm{PI}(\varrho)$, then μ satisfies $\mathrm{LSI}(\alpha)$ with

$$\frac{1}{\alpha} \leq 2\sqrt{\frac{1}{\lambda}\left(\frac{1}{2} + \frac{b + \lambda\mu(|x|^2)}{\varrho}\right)} + \frac{K_H}{2\varepsilon\lambda} + \frac{K_H(b + \lambda\mu(|x|^2)) + 2\varepsilon\lambda}{\varrho\varepsilon\lambda}. \tag{3.31}$$

where $\mu(|x|^2)$ denotes the second moment of μ.

Lemma 3.13 ([CGW09, Lemma 3.4]). *Assume that U is a non-negative locally Lipschitz function such that for some lower bounded function ϕ*

$$\frac{Le^U}{e^U} = LU + \varepsilon |\nabla U|^2 \leq -\varepsilon\phi \tag{3.32}$$

in the distributional sense. Then for any g holds

$$\int \phi g^2 \mathrm{d}\mu \leq \int |\nabla g|^2 \mathrm{d}\mu.$$

Proof. We can assume w.l.o.g. that g is smooth with bounded support and ϕ is bounded. For the verification of the desired statement, we need the symmetry of L in $L^2(\mu)$:

$$\int (-Lf) g \mathrm{d}\mu = \int f(-L) g \mathrm{d}\mu = \varepsilon \int \nabla f \cdot \nabla g \mathrm{d}\mu, \tag{3.33}$$

and the simple estimate

$$2g\nabla U \cdot \nabla g \leq |\nabla U|^2 g^2 + |\nabla g|^2. \tag{3.34}$$

An application of the assumption (3.32) yields

$$\begin{aligned} \varepsilon \int \phi g^2 \mathrm{d}\mu &\overset{(3.32)}{\leq} \int \left(-LU - \varepsilon |\nabla U|^2\right) g^2 \mathrm{d}\mu \\ &\overset{(3.33)}{=} \varepsilon \int \left(2g\nabla U \cdot \nabla g - |\nabla U|^2 g^2\right) \mathrm{d}\mu \overset{(3.34)}{\leq} \varepsilon \int |\nabla g|^2 \mathrm{d}\mu, \end{aligned}$$

which is the desired estimate. □

The proof of Theorem 3.12 relies on an interplay of some other functional inequalities, which will not occur anywhere else. Therefore, in Appendix D a condensed summary may be found.

Proof of Theorem 3.12. The argument of [CGW09] is a combination of the Lyapunov condition (3.30) leading to a defective WI inequality and the use of the HWI inequality of Otto and Villani [OV00]. In the following, we will use the measure ν given by $\nu(\mathrm{d}x) = h(x)\mu(\mathrm{d}x)$, where we can assume w.l.o.g. that ν is a probability measure, i.e. $\int h \,\mathrm{d}\mu = 1$. The first step is to estimate the Wasserstein distance in terms of the total variation (cf. Theorem D.2 and [Vil09, Theorem 6.15])

$$W_2^2(\nu,\mu) \le 2 \| |\cdot|^2 (\nu-\mu)\|_{TV}. \tag{3.35}$$

For every function g with $|g| \le \phi(x) := \lambda |x|^2$, where λ is from the Lyapunov condition (3.30) we get

$$\int g \,\mathrm{d}(\nu-\mu) \le \int \phi \,\mathrm{d}\nu + \int \phi \,\mathrm{d}\mu = \int (\lambda |x|^2 - b) h(x)\, \mu(\mathrm{d}x) + \int b \,\mathrm{d}\nu + \mu(\phi), \tag{3.36}$$

We can apply to $\int (\lambda |x|^2 - b) h \,\mathrm{d}\mu$ Lemma 3.13, where the assumption is exactly the Lyapunov condition (3.30) by choosing $U = \log W$ and arrive at

$$\int (\lambda |x|^2 - b) h \,\mathrm{d}\mu \le \int \left|\nabla \sqrt{h}\right|^2 \mathrm{d}\mu = \int \frac{|\nabla h|^2}{4h} \mathrm{d}\mu = \frac{1}{2} I(\nu|\mu), \tag{3.37}$$

by the definition of the Fisher information. Taking the supremum over g in (3.36) and combining the estimate with (3.35) and (3.37) we arrive at the defective WI inequality

$$\frac{\lambda}{2} W_2^2(\nu,\mu) \le \lambda \| |\cdot|^2 (\nu-\mu)\|_{TV} \le \frac{1}{2} I(\nu|\mu) + b + \mu(\phi). \tag{3.38}$$

The next step is to use the HWI inequality (cf. Theorem D.6 and [OV00, Theorem 3]), which holds by the assumption $\nabla^2 H \ge -K_H$

$$\mathrm{Ent}_\mu(h) \le W_2(\nu,\mu)\sqrt{2I(\nu|\mu)} + \frac{K_H}{2\varepsilon} W_2^2(\nu,\mu).$$

Substituting the defective WI inequality into the HWI inequality and using the Young inequality $ab \le \frac{\tau}{2} a^2 + \frac{1}{2\tau} b^2$ for $\tau > 0$ results in

$$\begin{aligned} \mathrm{Ent}_\mu(h) &\le \tau I(\nu|\mu) + \left(\frac{1}{2\tau} + \frac{K_H}{2\varepsilon}\right) W_2^2(\nu,\mu) \\ &\overset{(3.38)}{\le} \left(\tau + \frac{1}{2\lambda}\left(\frac{1}{\tau} + \frac{K_H}{\varepsilon}\right)\right) I(\nu|\mu) + \frac{1}{\lambda}\left(\frac{1}{\tau} + \frac{K_H}{\varepsilon}\right)(b + \mu(\phi)). \end{aligned} \tag{3.39}$$

The last inequality is of the type $\mathrm{Ent}_\mu(h) \le \frac{1}{\alpha_d} I(\nu|\mu) + B \int h \mathrm{d}\mu$ and is often called defective logarithmic Sobolev inequality $\mathrm{dLSI}(\alpha_d, B)$. It is well-known, that a defective logarithmic Sobolev inequality can be tightened by a Poincaré inequality $\mathrm{PI}(\varrho)$ to a logarithmic Sobolev inequality with constant (cf. Proposition D.9)

$$\frac{1}{\alpha} = \frac{1}{\alpha_d} + \frac{B+2}{\varrho}. \tag{3.40}$$

A combination of (3.39) and (3.40) reveals

$$\begin{aligned}\frac{1}{\alpha} &= \tau + \frac{1}{2\lambda}\left(\frac{1}{\tau} + \frac{K_H}{\varepsilon}\right) + \frac{1}{\varrho}\left(\frac{1}{\lambda}\left(\frac{1}{\tau} + \frac{K_H}{\varepsilon}\right)(b+\mu(\phi)) + 2\right) \\ &= \tau + \frac{1}{\tau}\,\frac{1}{\lambda}\left(\frac{1}{2} + \frac{b+\mu(\phi)}{\varrho}\right) + \frac{K_H}{2\varepsilon\lambda} + \frac{K_H(b+\mu(\phi)) + 2\varepsilon\lambda}{\varrho\varepsilon\lambda} = \tau + \frac{c_1}{\tau} + c_2.\end{aligned}$$

The last step is to optimize in τ, which leads to $\tau = \sqrt{c_1}$ and therefore $\frac{1}{\alpha} = 2\sqrt{c_1} + c_2$. The final result (3.31) follows by recalling the definition of $\phi(x) = \lambda\,|x|^2$. □

For proofing the Lyapunov condition (3.30) we can use the construction of an ε-modification done in Lemma 3.11.

Lemma 3.14 (Lyapunov function for LSI)**.** *There exists an ε-modification $\tilde{H}_\varepsilon$ of H satisfying the Lyapunov condition* (3.30) *with Lyapunov function*

$$W(x) = \exp\left(\frac{1}{2\varepsilon}\tilde{H}_\varepsilon(x)\right) \quad \text{and constants} \quad b = \frac{b_0}{\varepsilon},\ \text{and } \lambda \ge \frac{\lambda_0}{\varepsilon} \text{ for some } b_0, \lambda_0 > 0$$

and Hessian $\nabla^2 \tilde{H}(x) \ge -K_{\tilde{H}}$ for some $K_{\tilde{H}} \ge 0$.

The proof consists of three steps, which correspond to three regions of interests. First, we will consider a neighborhood of ∞, i.e. we will fix some $\tilde{R} > 0$ and only consider $|x| \ge \tilde{R}$. This will be the analog estimate to formula (3.10) of Lemma 3.10. Then, we will look at an intermediate regime for $a\sqrt{\varepsilon} \le |x| \le \tilde{R}$, where we will have to take special care for the neighborhoods around critical points and use the construction of Lemma 3.11. The last regime is for $|x| \le a\sqrt{\varepsilon}$, which will be the simplest case.

Therefore, besides the construction done in the proof of Lemma 3.11, we need an analogous formulation of Lemma 3.10 under the stronger assumption ($\mathbf{A1}_{\mathrm{LSI}}$).

Lemma 3.15. *Assume that the Hamiltonian H satisfies Assumption* ($\mathbf{A1}_{\mathrm{LSI}}$)*. Then, there is a constant $0 \le C_H < \infty$ and $0 \le \tilde{R} < \infty$ such that for ε small enough*

$$\frac{\Delta H(x)}{2\varepsilon} - \frac{|\nabla H(x)|^2}{4\varepsilon^2} \le -\frac{C_H}{\varepsilon}\ |x|^2 \qquad \text{for all } |x| \ge \tilde{R}. \tag{3.41}$$

We skip the proof of the Lemma 3.15, because it would work in the same way as for Lemma 3.10 and only consists of elementary calculations based on the non-degeneracy assumption on H. The only difference, is that we now demand the stronger statement (3.41), which is a consequence of the stronger assumption ($\mathbf{A1}_{\mathrm{LSI}}$) in comparison to assumption ($\mathbf{A2}_{\mathrm{PI}}$).

Now, we have collected the auxiliary statements and can proof Lemma 3.14.

Proof of Lemma 3.14. First, let us check the lower bound on the Hessian of $\tilde{H}$. We will use the same construction as of the Poincaré inequality in Lemma 3.11. Therefore, the support of $\tilde{H} - H$ is compact and $\tilde{H}$ is composed only of smooth functions, which already implies the lower bound on the Hessian for compact domains. Outside a sufficient large domain, the lower bound is just the Assumption (**A2**$_{\mathrm{LSI}}$). Now we can turn to verify the Lyapunov condition (3.30) and calculate with $W = \exp(\frac{1}{2\varepsilon}\tilde{H}_\varepsilon)$

$$\frac{1}{\varepsilon}\frac{LW}{W} = \frac{1}{2\varepsilon}\Delta\tilde{H}_\varepsilon + \frac{1}{4\varepsilon^2}|\nabla\tilde{H}_\varepsilon|^2 - \frac{1}{2\varepsilon^2}|\nabla\tilde{H}_\varepsilon|^2 = \frac{1}{2\varepsilon}\Delta\tilde{H}_\varepsilon - \frac{1}{4\varepsilon^2}|\nabla\tilde{H}_\varepsilon|^2.$$

If $|x| \geq \tilde{R}$ with $\tilde{R}$ given in Lemma 3.15, we apply (3.41) and have the Lyapunov condition fulfilled with constant $\lambda = \frac{C_H}{\varepsilon}$. This allows us to only consider $x \in B_{\tilde{R}} \cap \Omega$, which is of course bounded. In this case, Lemma 3.9 yields for $a\sqrt{\varepsilon} \leq |x| \leq \tilde{R}$ the estimate

$$\frac{1}{\varepsilon}\frac{LW}{W} \leq -\frac{\lambda_0}{\varepsilon} \leq -\frac{\lambda_0}{\tilde{R}^2\varepsilon}\;|x|^2\,. \tag{3.42}$$

For $|x| \leq a\sqrt{\varepsilon}$ holds by the representation (3.42) since H is smooth and strictly convex in $B_{a\sqrt{\varepsilon}}$ the bound

$$\frac{1}{\varepsilon}\frac{LW}{W} \leq \frac{1}{2\varepsilon}\Delta H(x) \leq \frac{b_0}{\varepsilon}. \tag{3.43}$$

A combination of (3.42) and (3.43) is the desired estimate (3.30). □

3.2. Proof of the local inequalities

In the previous Section 3.1, we were able to construct Lyapunov functions for the Hamiltonian restricted to the basin of attraction for each minimum. This is sufficient to finally prove the local Poincaré and logarithmic Sobolev inequalities of Theorem 2.19 and Theorem 2.21, which consist of mainly checking, whether the constants in the Lyapunov conditions show the right scaling behavior in ε. Let us start by restating the local Poincaré inequality.

Theorem 2.19 (Local Poincaré inequality). *The local measures $\{\mu_i\}_{i=1}^M$, obtained by restricting μ to the basin of attraction Ω_i of the local minimum m_i (cf.* (2.7)*), satisfy* $\mathrm{PI}(\varrho_i)$ *with*

$$\varrho_i^{-1} = O(\varepsilon).$$

Proof. We prove the theorem for each μ_i individually and omit the index i. The first step is the application of the Holley-Stroock perturbation principle in Lemma 3.4, which ensures that whenever $\tilde{H}_\varepsilon$ is an ε-modification of H, i.e. $\sup_{x\in\Omega}|\tilde{H}_\varepsilon(x) - H(x)| \leq C_{\tilde{H}}\varepsilon$, the Poincaré constants are of the same order in terms of scaling in ε, i.e.

$$\varrho \geq e^{-2C_{\tilde{H}}}\,\tilde{\varrho}. \tag{3.44}$$

In the next step, we construct an explicit ε-modification $\tilde{H}$ satisfying the Lyapunov condition Definition 3.7. Therefore, we can apply Theorem (3.8) with constant λ and b satisfying the bounds (3.9) from Lemma 3.9. This leads to a lower bound for $\tilde{\varrho}$ by

$$\tilde{\varrho} \geq \frac{\lambda\varrho_R}{b + \varrho_R} \geq \frac{\lambda_0\varrho_R}{b_0 + \varepsilon\varrho_R}. \tag{3.45}$$

The final step is to observe that, since $R = a\sqrt{\varepsilon}$, we can assume that the measure $\tilde{\mu}_R = \tilde{\mu}\llcorner B_{a\sqrt{\varepsilon}}$ is just the measure $\hat{\mu}_a$. Therefore, it holds $\varrho_R^{-1} = O(\varepsilon)$ by Lemma 3.6, which leads by combining the estimates (3.44) and (3.45) to the conclusion $\varrho^{-1} = O(\varepsilon)$. □

Before continuing with the proof of the local logarithmic Sobolev inequality Theorem 2.21, we want to remark, that the Lyapunov condition for the Poincaré inequality and in particular for the logarithmic Sobolev inequality imply an estimate of the second moment of μ.

Lemma 3.16 (Second moment estimate). *If H fulfills the Lyapunov condition* (3.4)*, then μ has finite second moment and it holds*

$$\int |x|^2 \, \mu(\mathrm{d}x) \leq \frac{1 + bR^2}{\lambda} \tag{3.46}$$

Proof. We use the equation (3.7) from the proof of Theorem 3.8, where have deduced that for a nice function f and any $m \in \mathbb{R}$ holds

$$\int (f - m)^2 \, \mathrm{d}\mu \leq \frac{1}{\lambda} \int |\nabla f|^2 \, \mathrm{d}\mu + \frac{b}{\lambda} \int_{\Omega_R} (f - m)^2 \, \mathrm{d}\mu.$$

We set $f(x) = |x|$ and $m = 0$ to observe the estimate (3.46). □

The construction done in Section 3.1.1 leads to the immediate results.

Corollary 3.17. *If H fulfills the assumptions* (**A1**$_{\mathrm{PI}}$) *and* (**A2**$_{\mathrm{PI}}$)*, then μ has finite second moment and it holds*

$$\int |x|^2 \, \mu(\mathrm{d}x) = O(\varepsilon).$$

Proof. We cannot apply the previous Lemma 3.16, but first have to do a change of measure to a measure $\tilde{\mu}$, where $\tilde{\mu}$ comes from an ε-modified Hamiltonian $\tilde{H}_\varepsilon$ of H

$$\int |x|^2 \, \mathrm{d}\mu \leq e^{2C_{\tilde{H}}} \int |x|^2 \, \mathrm{d}\tilde{\mu}.$$

Moreover, Lemma 3.9 ensures that $\tilde{H}_\varepsilon$ satisfies the Lyapunov condition (3.4) with constants $\lambda \geq \frac{\lambda_0}{\varepsilon}$, $b \leq \frac{b_0}{\varepsilon}$ and $R = a\sqrt{\varepsilon}$. Now, we can apply the previous Lemma 3.16 and immediately observe the result. □

Theorem 2.21 (Local logarithmic Sobolev inequality). *The local measures $\{\mu_i\}_{i=1}^M$, obtained by restricting μ to the basin of attraction Ω_i of the local minimum m_i (cf.* (2.7)*), satisfy* LSI(α_i) *with*

$$\alpha_i^{-1} = O(1).$$

Proof. For the same reason as in the proof of Theorem 2.19, we omit the index i. The first step is also the same as in the proof of Theorem 2.19. By Lemma 3.4 we obtain that, whenever $\tilde{H}_\varepsilon$ is an ε-modification of μ in the sense of Definition 3.3, the logarithmic Sobolev constants α and $\tilde{\alpha}$ of μ and $\tilde{\mu}$ satisfy $\alpha \geq \exp(-2C_{\tilde{H}})\tilde{\alpha}$.
The next step is to construct an explicit ε-modification $\tilde{H}$ satisfying the Lyapunov condition (3.30) of Theorem 3.12, which is provided by Lemma 3.14.
Additionally, the logarithmic Sobolev constant $\tilde{\alpha}$ depends on the second moment of $\tilde{\mu}$. Since $\tilde{H}_\varepsilon$ satisfies by Lemma 3.9 in particular the Lyapunov condition for the Poincaré inequality (3.4) with constants $\lambda \geq \frac{\lambda_0}{\varepsilon}$, $b \leq \frac{b_0}{\varepsilon}$ and $R = a\sqrt{\varepsilon}$, we can apply Lemma 3.16 and arrive at

$$\int |x|^2 \, \mathrm{d}\tilde{\mu} \leq \frac{1 + R^2 b}{\lambda} \leq \frac{1 + b_0 a^2}{\lambda_0} \, \varepsilon = O(\varepsilon).$$

Now, we have control on all constants occurring in (3.31) and can determine the logarithmic Sobolev constant $\tilde{\alpha}$ of $\tilde{\mu}$. Let us estimate term by term of (3.31) and use the fact from Theorem (2.19), that $\tilde{\mu}$ satisfies $\mathrm{PI}(\tilde{\varrho})$ with $\tilde{\varrho}^{-1} = O(\varepsilon)$

$$2\sqrt{\frac{1}{\lambda}\left(\frac{1}{2} + \frac{b + \lambda \tilde{\mu}(|x|^2)}{\varrho}\right)} \leq 2\sqrt{\frac{\varepsilon}{\lambda_0}\left(\frac{1}{2} + O(1)\right)} = O(\sqrt{\varepsilon}).$$

The second term evaluates to $\frac{K_H}{2\varepsilon\lambda} = O(1)$ and finally the last one

$$\frac{K_H(b + \lambda \tilde{\mu}(|x|^2)) + 2\varepsilon\lambda}{\varrho\varepsilon\lambda} = O(\varepsilon)\left(K_H\left(\frac{b_0}{\varepsilon} + O(\varepsilon)\right) + O(1)\right) = O(1).$$

A combination of all the results leads to the conclusion $\tilde{\alpha}^{-1} = O(1)$ and since $\tilde{H}_\varepsilon$ is only an ε-modification of H also $\alpha^{-1} = O(1)$. □

3.3. Optimality in one dimension

We want to close this chapter with a remark regarding the optimality of Theorem (2.19) and Theorem (2.21). We claim that the results are optimal in terms of scaling w.r.t. ε for measures obtained by restricting a Gibbs measure to the basin of attraction of a local minimum of its Hamiltonian. We will indicate the optimal behavior by estimating the Poincaré and logarithmic Sobolev constant for a characteristic energy landscape in one dimension. There exist two functionals, namely the *Muckenhoupt* and *Bobkov-Götze* functional, which allow to determine the Poincaré and logarithmic Sobolev constants for a given one dimensional measure up to a universal multiplicative factor. This is enough to show the optimal scaling behavior in ε. Let us now introduce the Muckenhoupt and Bobkov-Götze functional and after that apply them to a specific one dimensional example.

3.3.1. Muckenhoupt and Bobkov-Götze functional

Theorem 3.18 (Muckenhoupt functional). *Let μ be a probability measure on $\mathbb{R}$ with density $e^{-H(x)}$ with respect to the Lebesgue measure on its support. Then, the constants B_m^- and B_m^+ defined by*

$$\begin{aligned} B_m^- &= \sup_{x \le m} \left(\int_x^m e^{H(y)} \mathrm{d}y \int_{-\infty}^x e^{-H(y)} \mathrm{d}y \right) \\ B_m^+ &= \sup_{x \ge m} \left(\int_m^x e^{H(y)} \mathrm{d}y \int_x^{\infty} e^{H(y)} \mathrm{d}y \right) \end{aligned} \tag{3.47}$$

are finite for μ-a.e. m if and only if they are finite for one common m.

Further, μ satisfies a Poincaré inequality PI(ϱ) if and only if B_m^- and B_m^+ are finite. In this case ϱ obeys the estimate

$$\max\left\{(1 - F_\mu(m))B_m^-, F_\mu(m)B_m^+\right\} \le \varrho^{-1} \le 4 \max\left\{B_m^-, B_m^+\right\},$$

where $F_\mu(m) = \mu((-\infty, m])$. Especially, by setting $m = m^*$ with m^* the median of μ, i.e. $F_\mu(m^*) = \frac{1}{2}$, then it holds

$$\frac{1}{2} \max\left\{B_{m^*}^-, B_{m^*}^+\right\} \le \varrho^{-1} \le 4 \max\left\{B_{m^*}^-, B_{m^*}^+\right\},$$

In the original work [Muc72], the main interest was to show the existence and to find estimate on the optimal constant in weighted Hardy type inequalities. For the proof and the detailed references how Theorem 3.18 evolved from [Muc72], we refer to Section 5.3.

After bringing the Muckenhoupt functional [Muc72] in the context of Poincaré inequalities, the question arose, whether there exists an analogous functional for the logarithmic Sobolev inequality. It was discovered by Bobkov and Götze [BG99] and we will refer to it as the Bobkov-Götze functional.

Theorem 3.19 (Bobkov-Götze functional [BG99, Theorem 1.1]). *Let μ be a probability measure on $\mathbb{R}$ with density $e^{-H(x)}$ w.r.t. to the Lebesgue measure on its support. Then, the measure μ satisfies* LSI(α) *for some constant α if and only if $D^- + D^+ < \infty$. In this case, α satisfies*

$$K_0(D^- + D^+) \le \frac{1}{\alpha} \le K_1(D^- + D^+),$$

where K_0 and K_1 are certain positive constants and D^- and D^+ are given by

$$\begin{aligned} D^- &= \sup_{x<m} \left(\int_{-\infty}^x e^{-H(y)} \mathrm{d}y \log\left(\frac{\int_{-\infty}^\infty e^{-H(y)} \mathrm{d}y}{\int_{-\infty}^x e^{-H(y)} \mathrm{d}y} \right) \int_x^m e^{H(y)} \mathrm{d}y \right) \\ D^+ &= \sup_{x>m} \left(\int_x^\infty e^{-H(y)} \mathrm{d}y \log\left(\frac{\int_{-\infty}^\infty e^{-H(y)} \mathrm{d}y}{\int_x^\infty e^{-H(y)} \mathrm{d}y} \right) \int_m^x e^{H(y)} \mathrm{d}y \right) \end{aligned} \tag{3.48}$$

and m is the median of the probability measure μ.

3.3.2. Application to characteristic energy landscape

In this section, we want to estimate the Poincaré and logarithmic Sobolev constant in the regime of low temperature $\varepsilon \ll 1$ for some a Hamiltonian being itself the basin of attraction. The Hamiltonian function is given by

$$x \in [-1,1] : H(x) = 1 - (1-x^2)^2 \quad \text{and} \quad x \notin [-1,1] : H(x) = \infty. \tag{3.49}$$

The according Gibbs measure at temperature ε has the form

$$\mu(\mathrm{d}x) = \frac{1}{Z_\mu} e^{-\frac{H(x)}{\varepsilon}} \mathrm{d}x \quad \text{and} \quad Z_\mu = \int_{-1}^{1} e^{-\frac{H(x)}{\varepsilon}} \mathrm{d}x. \tag{3.50}$$

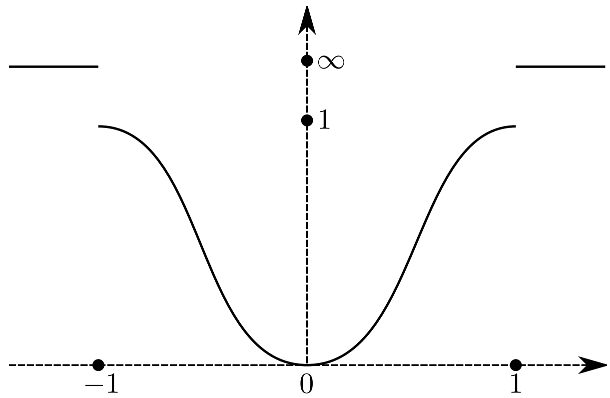

Figure 3.1.: Cartoon of a basin of attraction in one dimension

Lemma 3.20. *The Gibbs measure μ given by* (3.50) *with Hamiltonian* (3.49) *satisfies for $\varepsilon \leq 1$*

$$\frac{\varepsilon}{C_B} \leq B^+ = B^- \leq C_B \varepsilon \tag{3.51}$$

$$\frac{1}{C_D} \leq D^+ = D^- \leq C_D. \tag{3.52}$$

In particular, the measure μ satisfies $\mathrm{PI}(\varrho)$ *and* $\mathrm{LSI}(\alpha)$ *with*

$$\frac{\varepsilon}{\tilde{C}_B} \leq \frac{1}{\varrho} \leq \tilde{C}_B \varepsilon \quad \textit{and} \quad \frac{1}{\tilde{C}_D} \leq \frac{1}{\alpha} \leq \tilde{C}_D.$$

Proof. The median m of μ is equal to 0, due to μ is symmetric. Another consequence of the symmetry is that by definition (3.47) holds $B_m^+ = B_m^-$ and also definition (3.48) implies $D_m^+ = D_m^-$. Therefore, we can concentrate on the evaluation of B_m^+ and D_m^+ and consider the functions for $x \in (0,1)$

$$B^+(x) = \int_0^x e^{H(y)} \mathrm{d}y \int_x^1 e^{H(y)} \mathrm{d}y$$

$$\text{and} \quad D^+(x) = \int_m^x e^{H(y)} \mathrm{d}y \int_x^1 e^{-H(y)} \mathrm{d}y \log \left(\frac{Z_\mu}{\int_x^1 e^{-H(y)} \mathrm{d}y} \right).$$

The upper bound for $B^+(x)$ follows from the observations

$$\forall y \in [0,x] : H(y) \leq H(x) - (y-x)^2 \quad \text{and} \quad \forall y \in [x,1] : H(y) \geq H(x) + (y-x)^2. \tag{3.53}$$

Therewith, we can estimate $B^+(x)$

$$B^+(x) \overset{(3.53)}{\leq} \int_0^x e^{\frac{H(x)}{\varepsilon} - \frac{(y-x)^2}{\varepsilon}} \mathrm{d}y \int_x^1 e^{-\frac{H(x)}{\varepsilon} - \frac{(y-x)^2}{\varepsilon}} \mathrm{d}y = \int_0^x e^{-\frac{y^2}{\varepsilon}} \mathrm{d}y \int_0^{1-x} e^{-\frac{y^2}{\varepsilon}} \mathrm{d}y \leq \frac{\pi}{4}\, \varepsilon,$$

where in the last step we extended the domain of integration onto $[0,\infty)$ for both integrals and evaluated the Gaussian integrals. For the lower bound, it is enough to give a lower bound for $B^+(x)$ for one single specific x. We will use the choice $x = \frac{\sqrt{\varepsilon}}{2}$. For the estimate we use the rough bound $0 \leq H(x) \leq 2x^2$ on $[-1,1]$ and calculate

$$\begin{aligned} B^+(\tfrac{\sqrt{\varepsilon}}{2}) &\geq \int_0^{\frac{\sqrt{\varepsilon}}{2}} e^{\frac{H(y)}{\varepsilon}} \mathrm{d}y \int_{\frac{\sqrt{\varepsilon}}{2}}^1 e^{-\frac{2y^2}{\varepsilon}} \mathrm{d}y \geq \frac{\sqrt{\varepsilon}}{2} \frac{\sqrt{\varepsilon}}{2} \int_1^{\frac{2}{\sqrt{\varepsilon}}} e^{-\frac{y^2}{2}} \mathrm{d}y \\ &\geq \frac{\varepsilon}{4} \int_1^2 e^{-\frac{y^2}{2}} \mathrm{d}y =: \frac{c_H}{4}\, \varepsilon, \end{aligned}$$

where we assumed in the last estimate $\varepsilon \leq 1$. Hence, we finished the proof of (3.51) with constant $C_B = \max\left\{\frac{\pi}{4}, \frac{4}{c}\right\}$.

For the function $D^+(r)$, we first have to asymptotically evaluate Z_μ

$$Z_\mu = \sqrt{\frac{\pi}{2}}\, \sqrt{\varepsilon} + O(\varepsilon).$$

Since, we are not interested in the asymptotic sharp estimate, it is enough to use the bounds obtain by comparing with Gaussian integrals by using $x^2 \leq H(x) \leq 2x^2$ on $[-1,1]$

$$\sqrt{\frac{\pi}{2}} \sqrt{\varepsilon} \leq Z_\mu \leq \sqrt{\pi\varepsilon}. \tag{3.54}$$

Furthermore, we need an improved bound for the integral $\int_x^1 e^{-\frac{H(y)}{\varepsilon}} \mathrm{d}y$, which is again obtained by the upper bound $H(x) \leq 2x^2$ on $[-1,1]$ by comparing

$$\int_x^1 e^{-\frac{H(y)}{\varepsilon}} \mathrm{d}y \geq \int_x^1 e^{-\frac{2y^2}{\varepsilon}} \mathrm{d}y = \frac{\sqrt{\varepsilon}}{2}\, e^{-\frac{2x^2}{2}} \int_0^{\frac{2}{\sqrt{\varepsilon}}(1-x)} e^{-\frac{y^2}{2}} \mathrm{d}y. \tag{3.55}$$

For bounding the remaining Gaussian integral, we can simply use the rough estimate where the concrete constants are chosen for convenience

$$\tfrac{1}{\sqrt{2}} \min\left\{\sqrt{\tfrac{\pi}{2}}, \tfrac{1}{\sqrt{2}} r\right\} \leq \int_0^r e^{-\frac{y^2}{2}} \mathrm{d}y \leq \min\left\{\sqrt{\tfrac{\pi}{2}}, r\right\}. \tag{3.56}$$

Hence, for the estimate of $D^+(x)$ we can use the same strategy like in the estimate of $B^+(x)$, especially employing the property (3.53), but now with an improved upper bound for $\int_x^1 e^{-\frac{H(y)}{\varepsilon}}\mathrm{d}y$ to compensate the logarithmic term

$$\begin{aligned}
D^+(x) &\overset{(3.53)}{\leq} \int_0^x e^{\frac{H(x)}{\varepsilon}-\frac{(y-x)^2}{\varepsilon}}\mathrm{d}y \int_x^1 e^{-\frac{H(x)}{\varepsilon}-\frac{(y-x)^2}{\varepsilon}}\mathrm{d}y \log\left(\frac{\sqrt{\pi\varepsilon}}{\int_x^1 e^{-H(y)}\mathrm{d}y}\right) \\
&\overset{(3.55)}{\leq} \underbrace{\int_0^x e^{-\frac{y^2}{\varepsilon}}\mathrm{d}y}_{\leq\sqrt{\frac{\pi}{2}}\sqrt{\varepsilon}} \underbrace{\int_0^{1-x} e^{-\frac{y^2}{\varepsilon}}\mathrm{d}y}_{y\mapsto\sqrt{\frac{\varepsilon}{2}}\,y} \log\left(\frac{\sqrt{\pi\varepsilon}}{\frac{\sqrt{\varepsilon}}{2}\,e^{-\frac{2x^2}{2}}\int_0^{\frac{2}{\sqrt{\varepsilon}}(1-x)} e^{-\frac{y^2}{2}}\mathrm{d}y}\right) \\
&\overset{(3.56)}{\leq} \sqrt{\tfrac{\pi}{2}}\sqrt{\varepsilon}\,\sqrt{\tfrac{\varepsilon}{2}}\min\left\{\sqrt{\tfrac{\pi}{2}},\sqrt{\tfrac{2}{\varepsilon}}(1-x)\right\}\log\left(\frac{2\sqrt{\pi}e^{\frac{2x^2}{\varepsilon}}}{\frac{1}{\sqrt{2}}\min\left\{\sqrt{\frac{\pi}{2}},\sqrt{\frac{2}{\varepsilon}}(1-x)\right\}}\right) \\
&\leq \tfrac{\pi}{\sqrt{2}} + \tfrac{\pi}{2\sqrt{2}}\varepsilon\log(2\sqrt{\pi}) - \sqrt{\tfrac{\pi}{2}}\,\varepsilon\,\psi\left(\tfrac{1}{\sqrt{2}}\min\left\{\sqrt{\tfrac{\pi}{2}},\sqrt{\tfrac{2}{\varepsilon}}(1-x)\right\}\right),
\end{aligned}$$

where $\psi(x) := x\log x$, which is bounded from below by $-\frac{1}{e}$. Hence, we obtain the first part of (3.52), i.e. the bound $D^+(x) \leq \frac{\pi}{2} + O(\varepsilon)$ for $x \in [0,1]$. The lower bound can be obtained in the dual way. Observe that H also satisfies the bound $1-4(x-1)^2 \leq H(x) \leq 1-(x-1)^2 \leq 1$, which leads to the following estimate of the single factors of $D^+(x)$, where we choose $x = 1-\frac{\sqrt{\varepsilon}}{2}$ and assume again $\varepsilon \leq 1$

$$\int_0^{1-\frac{\sqrt{\varepsilon}}{2}} e^{\frac{H(y)}{\varepsilon}}\mathrm{d}y \geq e^{\frac{1}{\varepsilon}}\int_{\frac{\sqrt{\varepsilon}}{2}}^1 e^{-\frac{4y^2}{\varepsilon}}\mathrm{d}y = e^{\frac{1}{\varepsilon}}\sqrt{\varepsilon}\int_{\frac{1}{2}}^1 e^{-4y^2}\mathrm{d}y =: e^{\frac{1}{\varepsilon}}\sqrt{\varepsilon}\,c_1. \tag{3.57}$$

For the second term, we simply obtain the lower bound

$$\int_{1-\frac{\sqrt{\varepsilon}}{2}}^1 e^{-\frac{H(y)}{\varepsilon}}\mathrm{d}y \geq e^{-\frac{1}{\varepsilon}}\frac{\sqrt{\varepsilon}}{2}. \tag{3.58}$$

Additionally, we need an upper bound for the same term, where we use $H(x) \leq 1-(x-1)^2$

$$\int_{1-\frac{\sqrt{\varepsilon}}{2}}^1 e^{-\frac{H(y)}{\varepsilon}}\mathrm{d}y \leq e^{-\frac{1}{\varepsilon}}\int_0^{\frac{\sqrt{\varepsilon}}{2}} e^{-\frac{y^2}{\varepsilon}} = e^{-\frac{1}{\varepsilon}}\sqrt{\varepsilon}\int_0^{\frac{1}{2}} e^{-y^2}\mathrm{d}y =: e^{-\frac{1}{\varepsilon}}\sqrt{\varepsilon}\,c_2. \tag{3.59}$$

A combination of the estimates (3.54), (3.57), (3.58), (3.59) results in the lower bound

$$D^+(1-\tfrac{\sqrt{\varepsilon}}{2}) \geq \varepsilon\,\frac{c_1}{2}\,\log\left(\frac{\sqrt{\frac{\pi}{2}\varepsilon}}{e^{-\frac{1}{\varepsilon}}\sqrt{\varepsilon}c_2}\right) = \frac{c_1}{2} + O(\varepsilon).$$

Finally, this shows (3.52) and finishes the proof. □

Chapter 4

Mean-difference estimates – weighted transport distance

This chapter is devoted to the proof of Theorem 2.23. We want to estimate the mean-difference $\left(\mathbb{E}_{\mu_i} f - \mathbb{E}_{\mu_j} f\right)^2$ for i and j fixed. The proof consists of four steps:

In the first step, we introduce the *weighted transport distance* in Section 4.1. This distance depends on the transport speed similarly to the Wasserstein distance, but in addition weights the speed of a transported particle w.r.t. the reference measure μ. The weighted transport distance allows in general for a variational characterization of the constant C in the inequality

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \leq C \int |\nabla f|^2 \,\mathrm{d}\mu.$$

The problem of finding good estimates of the constant C is then reduced to the problem of finding a good transport between the measures μ_i and μ_j w.r.t. to the weighted transport distance.

For measures as general as μ_i and μ_j, the construction of an explicit transport interpolation is not feasible. Therefore, the second step consists of an approximation, which is done in Section 4.2. There, the restricted measures μ_i and μ_j are replaced by *simpler* measures ν_i and ν_j, namely truncated Gaussians. We show in Lemma 4.7 that this approximation only leads to higher order error terms.

The most import step, the third one, consists of the estimation of the mean-difference w.r.t. the approximations ν_i and ν_j. Because the structure of ν_i and ν_j is very simple, we can explicitly construct a transport interpolation between ν_i and ν_j (see Lemma 4.12 in Section 4.3).

The last step consists of combining the results of Lemma 4.7 and Lemma 4.12 (cf. Section 4.4). To demonstrate the flexibility of the weighted transport distance, we will deduce a mean-difference estimate, where the non-degeneracy Assumption 1.12 is not completely fulfilled in Section 4.5.

4.1. Mean-difference estimates by transport

At the moment let us consider two arbitrary measures $\nu_0 \ll \mu$ and $\nu_1 \ll \mu$. The overall aim is to estimate the mean-difference

$$(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f))^2 .$$

The starting point of the estimation is a representation of the mean-difference as a transport interpolation. This idea goes back to Chafaï and Malrieu [CM10]. However, they used a similar but non-optimal estimate for our purpose. Hence, let us consider a transport map U between ν_0 and ν_1, i.e. the push forward of ν_0 under the map U is given by $U_\sharp \nu_0 = \nu_1$. Further, let $(\Phi_s)_{s\in[0,1]}$ be a smooth interpolation between the identity and the transport map U, i.e.

$$\Phi_0 = \mathrm{Id}, \quad \Phi_1 = U, \quad \text{and} \quad (\Phi_s)_\sharp \nu_0 = \nu_s.$$

The representation of the mean-difference as a transport interpolation is attained by using the fundamental theorem of calculus, i.e.

$$(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f))^2 = \left(\int \int_0^1 \frac{\mathrm{d} f(\Phi_s)}{\mathrm{d}s} \,\mathrm{d}s \,\mathrm{d}\nu_0\right)^2 = \left(\int_0^1 \int \langle \nabla f(\Phi_s), \dot{\Phi}_s\rangle \,\mathrm{d}\nu_0 \,\mathrm{d}s\right)^2 .$$

At this point it is tempting to apply the Cauchy-Schwarz inequality in $L^2(\mathrm{d}\nu_0 \times \mathrm{d}s)$ leading to the estimate of Chafaï and Malrieu [CM10]. However, this strategy would not yield the pre-exponential factors in the Eyring-Kramers formula (2.25) (cf. Remark 4.2). On Stephan Luckhaus' advice the author realized the fact that it really matters on which integral you apply the Cauchy-Schwarz inequality. This insight lead to the following proceeding

$$\begin{aligned}
(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f))^2 &= \left(\int_0^1 \int \langle \nabla f(\Phi_s), \dot{\Phi}_s\rangle \,\mathrm{d}\nu_0 \,\mathrm{d}s\right)^2 \\
&= \left(\int_0^1 \int \langle \nabla f, \dot{\Phi}_s \circ \Phi_s^{-1}\rangle \,\mathrm{d}\nu_s \,\mathrm{d}s\right)^2 \\
&\le \left(\int |\nabla f| \int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \,\frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \,\mathrm{d}\mu\right)^2 \\
&\le \int \left(\int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \,\frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s\right)^2 \mathrm{d}\mu \int |\nabla f|^2 \,\mathrm{d}\mu. \qquad (4.1)
\end{aligned}$$

Note that in the last step we have applied the Cauchy-Schwarz inequality only in $L^2(\mathrm{d}\mu)$ and that the desired Dirichlet integral $\int |\nabla f|^2 \,\mathrm{d}\mu$ is already recovered.

The prefactor in front of the the Dirichlet energy on the right-hand side of (4.1) only depends on the transport interpolation $(\Phi_s)_{s\in[0,1]}$. Hence, we can infimize over all possible admissible transport interpolations and arrive at the following definition.

Definition 4.1 (Weighted transport distance $\mathcal{T}_\mu$)**.** Let μ be an absolutely continuous probability measure on $\mathbb{R}^n$ with connected support. Additionally, let ν_0 and ν_1 be two probability measures such that $\nu_0 \ll \mu$ and $\nu_1 \ll \mu$, then define the *weighted transport distance* by

$$\mathcal{T}_\mu^2(\nu_0, \nu_1) = \inf_{\Phi_s} \int \left(\int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \right)^2 \mathrm{d}\mu. \tag{4.2}$$

The family $(\Phi_s)_{s\in[0,1]}$ is chosen absolutely continuous in the parameter s such that $\Phi_0 = \mathrm{Id}$ on $\operatorname{supp}\nu_0$ and $(\Phi_1)_\sharp \nu_0 = \nu_1$. For a fixed family and $(\Phi_s)_{s\in[0,1]}$ and a point $x \in \operatorname{supp}\mu$ the *cost density* is defined by

$$\mathcal{A}(x) := \int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}(x)| \; \nu_s(x) \,\mathrm{d}s. \tag{4.3}$$

Remark 4.2 (Relation of $\mathcal{T}_\mu$ to [CM10]). In general, the transport cost $\mathcal{T}_\mu(\nu_0, \nu_1)$ is always smaller than the constant obtained by Chafaï and Malrieu [CM10, Section 4.6]. Indeed, applying the Cauchy-Schwarz inequality on $L^2(\mathrm{d}s)$ in (4.2) yields

$$\begin{aligned}\mathcal{T}_\mu^2(\nu_0, \nu_1) &\le \inf_{\Phi_s} \int \int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}|^2 \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \int_0^1 \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \,\mathrm{d}\mu \\ &\le \inf_{\Phi_s} \left(\sup_x \left(\int_0^1 \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu}(x) \,\mathrm{d}s \right) \int \int_0^1 |\dot{\Phi}_s|^2 \,\mathrm{d}s \,\mathrm{d}\nu_0 \right),\end{aligned}$$

where we used the assumption that $\nu_s \ll \mu$ for all $s \in [0,1]$ in the last L^1-L^∞ estimate.

Remark 4.3 (Relation of $\mathcal{T}_\mu$ to the L^2-Wasserstein distance W_2). If the support of μ is convex, we can set the transport interpolation $(\Phi_s)_{s\in[0,1]}$ to the linear interpolation map $\Phi_s(x) = (1-s)x + sU(x)$. Assuming that U is the optimal W_2-transport map between ν_0 and ν_1, the estimate in Remark 4.2 becomes

$$\mathcal{T}_\mu^2(\nu_0, \nu_1) \le \left(\sup_x \int_0^1 \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu}(x) \,\mathrm{d}s \right) W_2^2(\nu_0, \nu_1). \tag{4.4}$$

Remark 4.4 (Invariance under time rescaling). The cost density $\mathcal{A}$ given by (4.3) is independent of rescaling the transport interpolation in the parameter s. Indeed, we observe that

$$\mathcal{A}(x) = \int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}(x)| \; \nu_s(x) \,\mathrm{d}s = \int_0^T |\dot{\Phi}_t^T \circ (\Phi_t^T)^{-1}(x)| \; \nu_t^T(x) \,\mathrm{d}t,$$

where $\Phi_t^T = \Phi_{t/T}$ and $\nu_t^T = \nu_{t/T}$.

In this paragraph, we show that $\mathcal{T}_\mu(\cdot,\cdot)$ actually is a distance justifying the term *weighted transport distance*. It turns out that the distance $\mathcal{T}_\mu(\cdot,\cdot)$ is a metric on a subspace of the space of probability measures on $\mathbb{R}^n$ with finite second moment. The main restriction is that the weighted transport distance is very sensitive to support constraints, which are difficult to check for interpolations between general measures.

Proposition 4.5 ($\mathcal{T}_\mu$ as a distance). *Assume that μ is absolutely continuous w.r.t. the Lebesgue measure. Additionally, assume that μ has convex support and finite second moment. Then $\mathcal{T}_\mu(\cdot,\cdot)$ is a distance on the space*

$$\mathcal{P}_\mu := \{\nu \in \mathcal{P}(\mathbb{R}^n) : \nu \ll \mu,\ \operatorname{supp}\nu \text{ compact}\}.$$

Proof of Proposition 4.5. The symmetry follows from the observation that if $s \mapsto \Phi_s$ is an optimal interpolation between ν_0 and ν_1, then $s \mapsto \Phi_{1-s}$ is an optimal interpolation between ν_1 and ν_0.
Let us consider the definiteness. Therefore, we assume that $\mathcal{T}_\mu(\nu_0,\nu_1) = 0$, then

$$\int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s = 0, \qquad \mu\text{-a.e.}$$

Hence, by integrating w.r.t. μ and interchanging the order of integration we obtain

$$0 = \int_0^1 \int |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}\mu \, \mathrm{d}s = \int_0^1 \int |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \mathrm{d}\nu_s \, \mathrm{d}s = \int_0^1 \int |\dot{\Phi}_s| \, \mathrm{d}\nu_0 \, \mathrm{d}s.$$

This shows that $\dot{\Phi}_s = 0$ on $\operatorname{supp}\nu_0$. Therefore, it holds $\Phi_s = \mathrm{Id}|_{\operatorname{supp}\nu_0}$ resulting in $\nu_s = \nu_0$ for all $s \in [0,1]$.
Let us consider the triangle inequality. We have to show that for arbitrary measures $\nu_0, \nu_{1/2}, \nu_1 \ll \mu$ holds

$$\mathcal{T}_\mu(\nu_0,\nu_1) \le \mathcal{T}_\mu(\nu_0,\nu_{\frac{1}{2}}) + \mathcal{T}_\mu(\nu_{\frac{1}{2}},\nu_1). \tag{4.5}$$

Let $(\Phi_s)_{s\in[0,1]}$ be an interpolation between ν_0 and ν_1 such that $[0,1] \ni s \mapsto \Phi_{s/2}$ is an interpolation between ν_0 and $\nu_{1/2}$ and $[0,1] \ni s \mapsto \Phi_{s/2+1/2}$ is an interpolation between $\nu_{1/2}$ and ν_1. An application of the triangle inequality in $L^2(\mathrm{d}\mu)$ yields the estimate

$$\begin{aligned}
\mathcal{T}_\mu(\nu_0,\nu_1) &\le \left(\int \left(\int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \mathrm{d}s \right)^2 \mathrm{d}\mu \right)^{\frac{1}{2}} \\
&= \left(\int \left(\int_0^{\frac{1}{2}} |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \mathrm{d}s + \int_{\frac{1}{2}}^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \mathrm{d}s \right)^2 \mathrm{d}\mu \right)^{\frac{1}{2}} \\
&\le \left(\int \left(\int_0^{\frac{1}{2}} |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \mathrm{d}s \right)^2 \mathrm{d}\mu \right)^{\frac{1}{2}} + \left(\int \left(\int_{\frac{1}{2}}^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \mathrm{d}s \right)^2 \mathrm{d}\mu \right)^{\frac{1}{2}}.
\end{aligned}$$

Because of the invariance of the cost density under rescaling of time (cf. Remark 4.4) and the arbitrariness of the transport $(\Phi_s)_{s\in[0,1]}$, the last inequality already implies the desired triangle inequality (4.5).
In the last step we show that $\mathcal{T}_\mu(\nu_0,\nu_1) < \infty$ for $\nu_0,\nu_1 \in \mathcal{P}_\mu$. For that purpose, we apply the bound in terms of the Wasserstein distance (4.4). Then, it is sufficient to show that

$\nu \in \mathcal{P}_\mu$ has finite second moment provided that μ has finite second moment. This fact follows from the change of measure

$$\int |x|^2 \,\mathrm{d}\nu = \int |x|^2 \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu \le \left\| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right\|_{L^\infty} \int |x|^2 \,\mathrm{d}\mu < \infty.$$

In the last inequality we used the observation $\left\| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right\|_{L^\infty} < \infty$, which holds due to the compactness of the support of ν. □

Remark 4.6 (p-weighted transport distance $\mathcal{T}_{p,\mu}$). It is also possible to define a *p-weighted transport distance* for $p \in [1, \infty)$ by

$$\mathcal{T}_{p,\mu}(\nu_0, \nu_1) := \left(\inf_{(\Phi_s)_{s\in[0,1]}} \int \left(\int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \mathrm{d}s \right)^p \mathrm{d}\mu \right)^{\frac{1}{p}}.$$

4.2. Approximation of the local measures μ_i

In this subsection we show that it is sufficient to consider only the mean-difference w.r.t. some auxiliary measures ν_i approximating μ_i. More precisely, the next lemma shows that there are nice measures ν_i which are close to the measures μ_i in the sense of the mean-difference.

Lemma 4.7 (Mean-difference of approximation)**.** *Let ν_i be a truncated Gaussian centered around the local minimum m_i with covariance matrix $\Sigma_i = (\nabla^2 H(m_i))^{-1}$, more precisely*

$$\nu_i(\mathrm{d}x) = \frac{1}{Z_{\nu_i}} \, e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}} \, \mathbb{1}_{E_i}(x) \,\mathrm{d}x, \quad \textit{where} \quad Z_{\nu_i} = \int_{E_i} e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}} \,\mathrm{d}x. \tag{4.6}$$

Here and further on, we use the convention that for a matrix M and a vector x we write

$$M[x] := \langle x, Mx \rangle .$$

The restriction E_i is given by an ellipsoid

$$E_i = \{x \in \mathbb{R}^n : |\Sigma_i^{-\frac{1}{2}}(x - m_i)| \le \sqrt{2\varepsilon}\, \omega(\varepsilon)\}. \tag{4.7}$$

Additionally, assume that μ_i satisfies $\mathrm{PI}(\varrho_i)$ with $\varrho_i^{-1} = O(\varepsilon)$.
Then the following estimate holds

$$(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\mu_i}(f))^2 \le O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)) \int |\nabla f|^2 \,\mathrm{d}\mu, \tag{4.8}$$

where the function $\omega(\varepsilon) : \mathbb{R}^+ \to \mathbb{R}^+$ in (4.7) *and* (4.8) *is smooth and monotone satisfying*

$$\omega(\varepsilon) \ge |\log \varepsilon|^{\frac{1}{2}} \qquad \textit{for } \varepsilon < 1.$$

The first step towards the proof of Lemma 4.7 is the following Lemma.

Lemma 4.8. *Let ν_i be a probability measure satisfying $\nu_i \ll \mu_i$. Moreover, if μ_i satisfies $\mathrm{PI}(\varrho_i)$ for some $\varrho_i > 0$, then the following estimate holds*

$$(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\mu_i}(f))^2 \leq \frac{1}{\varrho_i} \operatorname{var}_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right) \int |\nabla f|^2 \,\mathrm{d}\mu_i. \tag{4.9}$$

Proof of Lemma 4.8. The result is a consequence from the representation of the mean-difference as a covariance. Therefore, we note that $\mathrm{d}\nu_i = \frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\,\mathrm{d}\mu_i$ since $\nu_i \ll \mu_i$ and use the Cauchy-Schwarz inequality for the covariance

$$\begin{aligned}(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\mu_i}(f))^2 &= \int f \mathrm{d}\nu_i - \int f \mathrm{d}\mu_i = \int f \,\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\,\mathrm{d}\mu_i - \int f \mathrm{d}\mu_i \underbrace{\int \frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\,\mathrm{d}\mu_i}_{=1} \\ &= \operatorname{cov}^2_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}, f\right) \leq \operatorname{var}_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right) \operatorname{var}_{\mu_i}(f).\end{aligned}$$

Using the fact that μ_i satisfies a Poincaré inequality results in (4.9). □

The above lemma tells us that we only need to construct ν_i, which approximates μ_i in variance. The following lemma provides exactly this.

Lemma 4.9 (Approximation in variance)**.** *Let the measures ν_i be given by Lemma 4.7. Then the partition sum Z_{ν_i} satisfies for ε small enough*

$$Z_{\nu_i} = (2\pi\varepsilon)^{\frac{n}{2}} \sqrt{\det \Sigma_i}\,(1 + O(\sqrt{\varepsilon})). \tag{4.10}$$

Additionally, ν_i is a good approximation in variance of μ_i, i.e.

$$\operatorname{var}_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right) = O(\sqrt{\varepsilon}\,\omega^3(\varepsilon)). \tag{4.11}$$

Proof of Lemma 4.9. The proof of (4.10) reduces to an estimate of a Gaussian integral on the complementary domain $\mathbb{R}^n \backslash E_i$. By recalling, that $(2\pi\varepsilon)^{\frac{n}{2}} \sqrt{\det \Sigma_i}$ is the normalization for a Gaussian with covariance matrix Σ_i, we arrive at

$$Z_{\nu_i} = \int_{E_i} e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}} \mathrm{d}x = (2\pi\varepsilon)^{\frac{n}{2}} \sqrt{\det \Sigma_i}\left(1 - \frac{1}{(2\pi\varepsilon)^{\frac{n}{2}} \sqrt{\det \Sigma_i}} \int_{\mathbb{R}^n\backslash E_i} e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}} \mathrm{d}x\right).$$

The integral on the complementary domain $\mathbb{R}^n \backslash E_i$ evaluates by the change of variables $x \mapsto y = (2\varepsilon\Sigma_i)^{-\frac{1}{2}}(x - m_i)$ to

$$\begin{aligned}\frac{1}{(2\pi\varepsilon)^{\frac{n}{2}} \sqrt{\det \Sigma_i}} \int_{\mathbb{R}^n\backslash E_i} e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}} \mathrm{d}x &= \frac{1}{\pi^{\frac{n}{2}}} \int_{\mathbb{R}^n\backslash B_{\omega(\varepsilon)}} e^{-y^2} \mathrm{d}y = \frac{n}{\Gamma(\frac{n}{2}+1)} \int_{\omega(\varepsilon)}^{\infty} r^{n-1} e^{-r^2} \mathrm{d}r \\ &= \frac{1}{\Gamma(\frac{n}{2})} \int_{\omega^2(\varepsilon)}^{\infty} r^{\frac{n}{2}-1} e^{-r} \mathrm{d}r = \frac{\Gamma(\frac{n}{2}, \omega^2(\varepsilon))}{\Gamma(\frac{n}{2})},\end{aligned}$$

where $\Gamma(\frac{n}{2}, \omega^2(\varepsilon))$ is the complementary incomplete Gamma function. It has the asymptotic expansion [Olv97, p. 109-112] given by

$$\Gamma\left(\frac{n}{2}, \omega^2(\varepsilon)\right) = O(e^{-\omega^2(\varepsilon)}\omega^{n-2}(\varepsilon)), \quad \text{for } \omega(\varepsilon) \geq \sqrt{n}.$$

We obtain (4.10) by the choice of $\omega(\varepsilon) \geq |\log \varepsilon|^{\frac{1}{2}}$, since the error becomes

$$O(e^{-\omega^2(\varepsilon)}\omega^{n-2}(\varepsilon)) = O(\varepsilon \,|\log \varepsilon|^{\frac{n}{2}-1}) = O(\sqrt{\varepsilon}), \qquad \text{for } \varepsilon \leq e^{-n}.$$

For the proof of (4.11), we compare the asymptotic expression for $Z_{\mu_i} = Z_i Z_\mu e^{\varepsilon^{-1} m_i}$ from (2.30) and Z_{ν_i} and obtain

$$Z_{\mu_i} = Z_{\nu_i} + O(\sqrt{\varepsilon}). \tag{4.12}$$

The relative density of ν_i w.r.t. μ_i can be estimated by Taylor expanding H around m_i. By the definition of ν_i given in (4.6), we obtain that $\Sigma_i^{-1}[y - m_i] - H_i(y) = O(|y - m_i|^3)$. This observation together with (4.12) leads to

$$\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}(y) = \frac{Z_{\mu_i}}{Z_{\nu_i}} e^{-\frac{1}{2\varepsilon}\Sigma_i^{-1}[y-m_i] + \frac{1}{2\varepsilon}H_i(y)} \mathbb{1}_{E_i}(y) = \frac{Z_{\mu_i}}{Z_{\nu_i}} e^{\frac{O(|y-m_i|^3)}{\varepsilon}} \mathbb{1}_{E_i}(y) = 1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon)).$$

Now, the conclusion directly follows from the definition of the variance

$$\begin{aligned}
\operatorname{var}_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right) &= \int_{E_i} \left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right)^2 \mathrm{d}\mu_i - \left(\int \frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\,\mathrm{d}\mu_i\right)^2 \\
&= \int_{E_i} 1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\,\mathrm{d}\mu_i - \left(\int_{E_i} \mathrm{d}\nu_i\right)^2 = 1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon)) - 1.
\end{aligned}$$

□

Proof of Lemma 4.7. A combination of Lemma 4.8 and Lemma 4.9 together with the assumption $\varrho_i^{-1} = O(\varepsilon)$ immediately reveals

$$(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\mu_i}(f))^2 \overset{(4.9),(4.11)}{\leq} O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)) \int |\nabla f|^2 \,\mathrm{d}\mu_i.$$

□

4.3. Affine transport interpolation

The aim of this section is to estimate $\left(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\nu_j}(f)\right)^2$ with the help of the weighted transport distance $\mathcal{T}_\mu(\nu_i, \nu_j)$ introduced in Section 4.1. The main result of this section estimates the weighted transport distance $\mathcal{T}_\mu(\nu_i, \nu_j)$ and is formulated in Lemma 4.12. For the proof of Lemma 4.12, we construct an explicit transport interpolation between ν_i and ν_j w.r.t. the measure μ. We start with a class of possible transport interpolations and optimize the weighted transport cost in this class.

Now, we state the main idea of this optimization procedure. Recall that the measures ν_i and ν_j are truncated Gaussians by the approximation we have done in the previous Section 4.2. Hence, the measures ν_i and ν_j are characterized by their mean and covariance matrix. We will choose the transport interpolation (cf. Section 4.3.1) such that the push forward measures $\nu_s := (\Phi_s)_\sharp \nu_0$ are again truncated Gaussians. Hence, it is sufficient to optimize among all paths γ connecting the minima m_i and m_j and all covariance matrices interpolating between Σ_i and Σ_j.

4.3.1. Definition of regular affine transport interpolations

Let us state in this section the class of transport interpolation among we want to optimize the weighted transport cost.

Definition 4.10 (Affine transport interpolations)**.** Assume that the measures ν_i and ν_j are given by Lemma 4.7. In detail, $\nu_i = \mathcal{N}(m_i, \varepsilon^{-1}\Sigma_i)\llcorner E_i$ and $\nu_j = \mathcal{N}(m_j, \varepsilon^{-1}\Sigma_j)\llcorner E_j$ are truncated Gaussians centered in m_i and m_j with covariance matrices $\varepsilon^{-1}\Sigma_i$ and $\varepsilon^{-1}\Sigma_j$. The restriction E_i and E_j are given for $l = 1, 2$ by the ellipsoids

$$E_l = \{x \in \mathbb{R}^n : |\Sigma_l^{-\frac{1}{2}}(x - m_l)| \leq \sqrt{2\varepsilon}\,\omega(\varepsilon)\}, \qquad \text{where } \omega(\varepsilon) \geq |\log \varepsilon|^{\frac{1}{2}} .$$

A transport interpolation Φ_s between ν_i and ν_j is called *affine transport interpolation* if there exists

- an interpolation path $(\gamma_s)_{s\in[0,T]}$ between $m_i = \gamma_0$ and $m_j = \gamma_T$ satisfying

$$\gamma = (\gamma_s)_{s\in[0,T]} \in C^2([0,T], \mathbb{R}^n) \qquad \text{and} \qquad \forall s \in [0,T] : \dot\gamma_s \in S^{n-1}, \tag{4.13}$$

- an interpolation path $(\Sigma_s)_{s\in[0,T]}$ of covariance matrices between Σ_i and Σ_j satisfying

$$\Sigma = (\Sigma_s)_{s\in[0,T]} \in C^2([0,T], \mathbb{R}^{n\times n}_{\text{sym},+}), \quad \Sigma_0 = \Sigma_i \quad \text{and} \quad \Sigma_T = \Sigma_j,$$

such that the transport interpolation $(\Phi_s)_{s\in[0,T]}$ is given by

$$\Phi_s(x) = \Sigma_s^{\frac{1}{2}} \Sigma_0^{-\frac{1}{2}}(x - m_0) + \gamma_s. \tag{4.14}$$

Since the cost density $\mathcal{A}$ given by (4.3) is invariant under rescaling of time (cf. Remark 4.4), one can always assume that the interpolation path γ_s is parameterized by arc-length. Hence, the condition $\dot\gamma_s \in S^{n-1}$ (cf. (4.13)) is not restricting.

We want to emphasize that for an affine transport interpolation $(\Phi_s)_{s\in[0,T]}$ the push forward measure $(\Phi_s)_\sharp \nu_0 = \nu_s$ is again a truncated Gaussian $\mathcal{N}(\gamma_s, \varepsilon^{-1}\Sigma_s)\llcorner E_s$, where E_s is the support of ν_s being again an ellipsoid in $\mathbb{R}^n$ given by

$$E_s = \{x \in \mathbb{R}^n : |\Sigma_s^{-\frac{1}{2}}(x - \gamma_s)| \leq \sqrt{2\varepsilon}\,\omega(\varepsilon)\}. \tag{4.15}$$

Therewith, the partition sum of ν_s is given by (cf. (4.10))

$$Z_{\nu_s} = (2\pi\varepsilon)^{\frac{n}{2}} \sqrt{\det \Sigma_s}(1 + O(\sqrt{\varepsilon})). \tag{4.16}$$

By denoting $\sigma_s = \Sigma_s^{\frac{1}{2}}$ and using the definition (4.14) of the affine transport interpolation $(\Phi_s)_{s\in[0,T]}$, we arrive at the relations

$$\begin{aligned}\dot{\Phi}_s(x) &= \dot{\sigma}_s\sigma_0^{-1}(x - m_0) + \dot{\gamma}_s,\\ \Phi_s^{-1}(y) &= \sigma_0\sigma_s^{-1}(y - \gamma_s) + m_0,\\ \dot{\Phi}_s \circ \Phi_s^{-1}(y) &= \dot{\sigma}_s\sigma_s^{-1}(y - \gamma_s) + \dot{\gamma}_s.\end{aligned}$$

Among all possible affine transport interpolations we are considering only those satisfying the following regularity assumption.

Assumption 4.11 (Regular affine transport interpolations). *The affine transport interpolation $(\gamma_s, \Sigma_s)_{s\in[0,T]}$ belongs to the class of* regular affine transport interpolations *if the length $T < T^*$ is bounded by some uniform $T^* > 0$ large enough. Further, for a uniform constant $c_\gamma > 0$ holds*

$$\inf\{r(x, y, z) : x, y, z \in \gamma, x \neq y \neq z \neq x\} \geq c_\gamma, \tag{4.17}$$

where $r(x, y, z)$ denotes the radius of the unique circle through the three distinct points x, y and z. Furthermore, there exists a uniform constant $C_\Sigma \geq 1$ for which

$$C_\Sigma^{-1}\,\mathrm{Id} \leq \Sigma_s \leq C_\Sigma\,\mathrm{Id} \quad \textit{and} \quad \|\dot{\Sigma}_s\| \leq C_\Sigma. \tag{4.18}$$

The infimum in condition (4.17) is called *global radius of curvature* (cf. [GMSvdM02]). It ensures that a small neighborhood of size $\frac{c_\gamma}{2}$ around γ is not self-intersecting, since the infimum can only be attained for the following three cases:

(i) All three points in a minimizing sequence of (4.17) coalesce to a point at which the radius of curvature is minimal.

(ii) Two points coalesce to a single point and the third converges to another point, such that the both points are a pair of closest approach.

(iii) Two points coalesce to a single point and the third converges to the starting or ending point of γ.

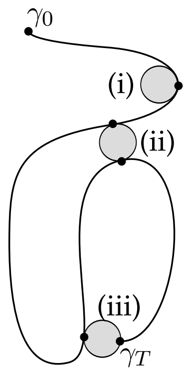

Figure 4.1.: Global radius of curvature

In the following calculations, there often occurs a multiplicative error of the form $1 + O(\sqrt{\varepsilon}\ \omega^3(\varepsilon))$. Therefore, let us introduce for convenience the notation "$\approx$" meaning "$=$" up to the multiplicative error $1 + O(\sqrt{\varepsilon}\ \omega^3(\varepsilon))$. The symbols "$\lesssim$" and "$\gtrsim$" have the analogous meaning.

Now, we can formulate the key ingredient for the proof of Theorem 2.23, namely the estimation of the weighted transport distance $\mathcal{T}_\mu(\nu_i, \nu_j)$.

Lemma 4.12. *Assume that ν_i and ν_j are given by Lemma 4.7. Then the weighted transport distance $\mathcal{T}_\mu(\nu_i, \nu_j)$ can be estimated as*

$$\begin{aligned}\mathcal{T}_\mu^2(\nu_i,\nu_j) &= \inf_{\Phi_s} \int \left(\int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right)^2 \mathrm{d}\mu \\ &\le \inf_{\Psi_s} \int \left(\int_0^1 |\dot{\Psi}_s \circ \Psi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right)^2 \mathrm{d}\mu \\ &\lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, 2\pi\varepsilon \left(\frac{\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} + \frac{T(C_\Sigma)^{\frac{n-1}{2}}}{\sqrt{2\pi\varepsilon}} e^{-\omega^2(\varepsilon)} \right) e^{\frac{H(s_{i,j})}{\varepsilon}},\end{aligned} \tag{4.19}$$

where the infimum over Ψ_s only considers regular affine transport interpolations Ψ_s in the sense of Assumption 4.11.
In particular, if we choose $\omega(\varepsilon) \ge |\log\varepsilon|^{\frac{1}{2}}$, which is enforced by Lemma 4.7, we get the estimate

$$\mathcal{T}_\mu^2(\nu_i,\nu_j) \le \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, \frac{2\pi\varepsilon\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} \, e^{\frac{H(s_{i,j})}{\varepsilon}} \, \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right). \tag{4.20}$$

The proof of Lemma 4.12 presents the core of the proof of the Eyring-Kramers formulas and consists of three steps carried out in the following sections:

- In Section 4.3.2, we carry out some preparatory work: We introduce tube coordinates on the support of the transport cost $\mathcal{A}$ given by (4.3) (cf. Lemma 4.13), we deduce a pointwise estimate on the transport cost $\mathcal{A}$ and we give a rough a priori estimate on the transport cost $\mathcal{A}$.
- In Section 4.3.3, we split the transport cost into a transport cost around the saddle and the complement. We also estimate the transport cost of the complement yielding the second summand in the desired estimate (4.19).
- In Section 4.3.4, we finally deduce a sharp estimate of the transport cost around the saddle yielding the first summand in the desired estimate (4.19).

4.3.2. Preparations and auxiliary estimates

The main reason for making the regularity Assumption 4.11 on affine transport interpolations is that we can introduce tube coordinates around the path γ. In these coordinates, the calculation of the cost density $\mathcal{A}$ given by (4.3) becomes a lot handier.

We start with defining the caps E_0^- and E_T^+ as

$$E_0^- := \{x \in E_0 : \langle x - \gamma_0, \dot{\gamma}_0 \rangle < 0\} \quad \text{and} \quad E_T^+ := \{x \in E_T : \langle x - \gamma_T, \dot{\gamma}_T \rangle > 0\},$$

The caps E_0^- and E_T^+ have no contribution to the total cost but unfortunately need some special treatment. Further, we define the slices V_s with $s \in [0, T]$

$$V_s = \{x \in \operatorname{span}\{\dot{\gamma}_s\}^\perp : |\Sigma_s^{-\frac{1}{2}} x| \le \sqrt{2\varepsilon}\,\omega(\varepsilon)\}$$

In $\operatorname{span} V_s$ we can choose a basis $e_s^2, \ldots, e_s^n$ smoothly depending on the parameter s. Especially there exists a family of rotational matrices $(Q_s)_{s\in[0,T]} \in C^2([0,T], SO(n))$ satisfying the same regularity assumption as the family $(\Sigma_\tau)_{\tau\in[0,T]}$ such that

$$Q_s e^1 = \dot\gamma_s, \qquad Q_s e^i = e_s^i, \quad \text{ for } i = 2, \ldots, n, \tag{4.21}$$

where $(e^1, \ldots, e^n)$ is the canonical basis of $\mathbb{R}^n$.

Let use now define the tube E as

$$E = \bigcup_{s\in[0,T]} (\gamma_s + V_s).$$

The support of the cost density $\mathcal{A}$ given by (4.3) is now given by

$$\operatorname{supp} \mathcal{A} = \bigcup_{s\in[0,T]} E_s = E_0^- \cup E \cup E_T^+. \tag{4.22}$$

By the definition (4.15) of E_s and the uniform bound (4.18) on Σ_s holds

$$\operatorname{diam} V_s \leq 2\sqrt{2\varepsilon C_\Sigma}\, \omega(\varepsilon). \tag{4.23}$$

Therewith, we find

$$\operatorname{supp} \mathcal{A} \subset B_{2\sqrt{2\varepsilon C_\Sigma}\,\omega(\varepsilon)}((\gamma_\tau)_{\tau\in[0,T]}) := \left\{ x \in \mathbb{R}^n : |x - \gamma_\tau| \leq 2\sqrt{2\varepsilon C_\Sigma}\, \omega(\varepsilon) \right\}.$$

The assumption (4.15) ensures that $B_{2\sqrt{2\varepsilon C_\Sigma}\,\omega(\varepsilon)}((\gamma_\tau)_{\tau\in[0,T]})$ is not self-intersecting for any ε small enough. The next lemma just states that by changing to tube coordinates in E one can asymptotically neglect the Jacobian determinant $\det J$.

Lemma 4.13 (Change of coordinates). *For the change of coordinates $(\tau, z) \mapsto x = \gamma_\tau + z_\tau$ with $z_\tau \in V_\tau$ holds for any function ξ on E that*

$$\int_E \xi(x)\, \mathrm{d}x \approx \int_0^T \int_{V_\tau} \xi(\gamma_\tau + z_\tau)\, \mathrm{d}z_\tau\, \mathrm{d}\tau.$$

Proof of Lemma 4.13. We use the representation of the tube coordinates via (4.21). Therewith, it holds that $x = \gamma_\tau + Q_\tau z$, where $z \in \{0\} \times \mathbb{R}^{n-1}$. Then, the Jacobian J of the coordinate change $x \mapsto (\tau, Q_\tau z)$ is given by

$$J = (\dot\gamma_\tau + \dot Q_\tau z, (Q_\tau)_2, \ldots, (Q_\tau)_n) \in \mathbb{R}^{n\times n},$$

where $(Q_\tau)_i$ denotes the i-th column of Q_τ. By the definition (4.21) of Q_τ follows $\dot\gamma_\tau = (Q_\tau)_1$. Hence, we have the representation $J = Q_\tau + \dot Q_\tau z \otimes e_1$. The determinant of J is then given by

$$\det J = \det\Big(Q_\tau + \dot Q_\tau z \otimes e_1\Big) = \underbrace{\det(Q_\tau)}_{=1} \det\Big(\mathrm{Id} + (Q_\tau^\top \dot Q_\tau z) \otimes e_1\Big) = 1 + \Big(Q_\tau^\top \dot Q_\tau z\Big)_1.$$

By Assumption 4.11 holds $\|\dot Q_\tau\| \leq C_\Sigma$, from which we conclude $(Q_\tau^\top \dot Q_\tau z)_{1,1} = O(z)$. Since $Q_\tau z \in V_\tau$, we get $O(z) = O(\sqrt{\varepsilon}\, \omega(\varepsilon))$ by (4.23). Hence we get

$$\det J = 1 + O(\sqrt{\varepsilon}\, \omega(\varepsilon)),$$

which concludes the proof. □

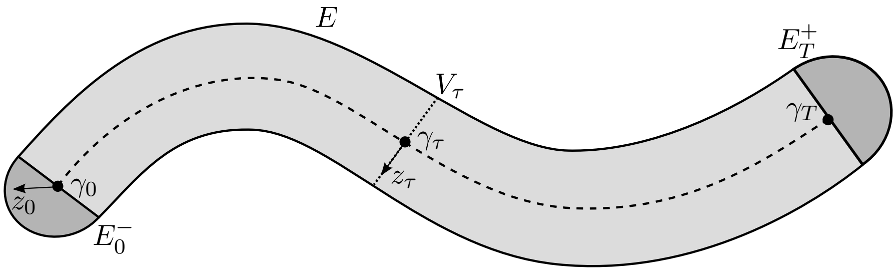

Figure 4.2.: The support of $\mathcal{A}$ in tube coordinates.

An important tool is the following auxiliary estimate.

Lemma 4.14 (Pointwise estimate of the cost-density $\mathcal{A}$)**.** *For $x \in \operatorname{supp} \mathcal{A}$ we define*

$$\tau = \underset{s\in[0,T]}{\arg\min} |x - \gamma_s| \qquad \textit{and} \qquad z_\tau := x - \gamma_\tau. \tag{4.24}$$

Then the following estimate holds

$$\mathcal{A}(x) \lesssim (2\pi\varepsilon)^{-\frac{n-1}{2}} \sqrt{\det\nolimits_{1,1}(Q_\tau^\top \tilde{\Sigma}_\tau^{-1} Q_\tau)}\; e^{-\frac{\tilde{\Sigma}_\tau^{-1}[z_\tau]}{2\varepsilon}} =: P_\tau\, e^{-\frac{\tilde{\Sigma}_\tau^{-1}[z_\tau]}{2\varepsilon}}, \tag{4.25}$$

where Q_τ is defined in (4.21) *and $\tilde{\Sigma}_\tau^{-1}$ is given by*

$$\tilde{\Sigma}_\tau^{-1} = \Sigma_\tau^{-1} - \frac{1}{\Sigma_\tau^{-1}[\dot{\gamma}_\tau]} \Sigma_\tau^{-1}\dot{\gamma}_\tau \otimes \Sigma_\tau^{-1}\dot{\gamma}_\tau. \tag{4.26}$$

Further, $\det_{1,1} A$ *is the determinant of the matrix obtained from A removing the first row and column.*

Remark 4.15. With a little bit of additionally work, one could show that (4.25) holds with "$\approx$" instead of "$\lesssim$". It follows from (4.26) that the matrix $\tilde{\Sigma}_\tau^{-1}$ is positive definite. Hence, $\mathcal{A}$ is an $\mathbb{R}^{n-1}$-dimensional Gaussian on the slice $\gamma_\tau + V_\tau$ up to approximation errors.

Proof of Lemma 4.24. We start the proof with some preliminary remarks and results. By the regularity Assumption 4.11 on the transport interpolation, we find that for all $x \in \operatorname{supp} \mathcal{A}$ holds uniformly

$$I_T(x) := \{s : E_s \ni x\} \quad \text{satisfies} \quad \mathcal{H}^1(I_T(x)) = O(\sup_{s\in[0,T]} \operatorname{diam}(E_s)) = O(\sqrt{\varepsilon}\,\omega(\varepsilon)).$$

This allows to linearize the transport interpolation around τ given in (4.24). It holds for s such that $x \in E_s$

$$\begin{aligned} \Sigma_s^{-1}[x - \gamma_s] &= \Sigma_\tau^{-1}[\gamma_\tau + z_\tau - \gamma_s] + O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)) \\ &= \Sigma_\tau^{-1}[(\tau - s)\dot{\gamma}_\tau + z_\tau] + O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)). \end{aligned} \tag{4.27}$$

For similar reasons, we can linearize the determinant $\det \Sigma_s$ and have $\det \Sigma_s = \det \Sigma_\tau + O(\sqrt{\varepsilon}\,\omega(\varepsilon))$. Finally, we have the following bound on the transport speed

$$\begin{aligned}
|\dot{\Phi}_s \circ \Phi_s^{-1}(x)| \mathbb{1}_{E_s}(x) &= \left|\dot{\sigma}_s \sigma_s^{-1}(x-\gamma_s) + \dot{\gamma}_s\right| \mathbb{1}_{E_s}(x) \\
&\le \left(\left|\dot{\sigma}_s \sigma_s^{-1}(x-\gamma_s)\right| + |\dot{\gamma}_s|\right) \mathbb{1}_{E_s}(x) \\
&\le (C_\Sigma\,|x-\gamma_s| + 1)\, \mathbb{1}_{E_s}(x) = \left(1 + O(\sqrt{\varepsilon}\,\omega(\varepsilon))\right) \mathbb{1}_{E_s}(x).
\end{aligned} \tag{4.28}$$

Let us first consider the case $x \in E$. We use (4.16), (4.27) and (4.28) to arrive with $x = \gamma_\tau + z_\tau$ where $z_\tau \in V_\tau$ at

$$\begin{aligned}
\mathcal{A}(x) &= \int_{I_T(x)} |\dot{\Phi}_s \circ \Phi_s^{-1}(x)|\, \frac{1}{Z_{\nu_s}}\, \exp\left(-\frac{1}{2\varepsilon}\Sigma_s^{-1}[x-\gamma_s]\right) \mathbb{1}_{E_s}(x)\, \mathrm{d}s \\
&\le \frac{1}{(2\pi\varepsilon)^{\frac{n}{2}}} \int_{I_T(x)} \frac{1 + O(\sqrt{\varepsilon}\,\omega(\varepsilon))}{\sqrt{\det \Sigma_s}}\, \exp\left(-\frac{1}{2\varepsilon}\Sigma_s^{-1}[x-\gamma_s]\right) \mathrm{d}s \\
&\le \frac{1}{(2\pi\varepsilon)^{\frac{n}{2}}\sqrt{\det \Sigma_\tau}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2\varepsilon}\Sigma_\tau^{-1}[(\tau - s)\dot{\gamma}_\tau + z_\tau]\right) \mathrm{d}s\, \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right) \\
&= \frac{\sqrt{\det \Sigma_\tau^{-1}}}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{\sqrt{2\pi\varepsilon}}{\sqrt{\Sigma_\tau^{-1}[\dot{\gamma}_\tau]}} \exp\left(-\frac{1}{2\varepsilon}\tilde{\Sigma}_\tau^{-1}[z_\tau]\right) \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon)\right),
\end{aligned}$$

where the last step follows by an application of a partial Gaussian integration (cf. Lemma B.1). Finally, by using the relation (B.2), we get that

$$\frac{\Sigma_\tau^{-1}}{\Sigma_\tau^{-1}[\dot{\gamma}_\tau]} = \det\nolimits_{1,1}(Q_\tau^\top \tilde{\Sigma}_\tau^{-1} Q_\tau),$$

and conclude the hypothesis for this case.

Let us now consider the case $x \in E_0^- \cup E_T^+$. For convenience, we only consider the case $x \in E_0^-$. By the definition of E_0^- holds $\tau = 0$. The integration domain $I_T(x)$ is now given by

$$I_T(x) = [0, s^*) \quad \text{with} \quad s^* = O(\sqrt{\varepsilon}\,\omega(\varepsilon)). \tag{4.29}$$

Therewith, we can estimate $\mathcal{A}(x)$ in the same way as for $x \in E$ and conclude the proof. □

We only need one more ingredient for the proof of Lemma 4.12. It is an a priori estimate on the cost density $\mathcal{A}$.

Lemma 4.16 (A priori estimates for the cost density $\mathcal{A}$)**.** *For $\mathcal{A}$ it holds:*

$$\int \mathcal{A}(x)\,\mathrm{d}x \;\lesssim\; T, \qquad \textit{and} \tag{4.30}$$

$$\mathcal{A}(x) \;\lesssim\; \left(\frac{C_\Sigma}{2\pi\varepsilon}\right)^{\frac{n-1}{2}} \qquad \textit{for } x \in \operatorname{supp}\mathcal{A}. \tag{4.31}$$

Proof of Lemma 4.16. Let us first consider the estimate (4.30). It follows from the characterization (4.22) of the support of $\mathcal{A}$ that

$$\int \mathcal{A}(x)\,\mathrm{d}x = \int_E \mathcal{A}(x)\,\mathrm{d}x + \int_{E_0^- \cup E_T^+} \mathcal{A}(x)\,\mathrm{d}x. \tag{4.32}$$

Now, we estimate the first term on the right-hand side of the last identity. Using the change to tube coordinates of Lemma 4.13 and noting that the upper bound (4.25) is a $(n-1)$-dimensional Gaussian density on V_τ for $\tau \in [0,T]$, we can easily infer that

$$\int_E \mathcal{A}(x)\,\mathrm{d}x \lesssim |\gamma| \le T$$

Let us turn to the second term on the right-hand side of (4.32). For convenience, we only consider the integral w.r.t. the cap E_0^-. It follows from (4.28) and (4.29) that

$$\begin{aligned}\int_{E_0^-} \mathcal{A}(x)\,\mathrm{d}x &\lesssim \int_{E_0^-} \int_0^1 \nu_s(x)\,\mathrm{d}s\,\mathrm{d}x = \int_0^{s^*} \int_{E_0^-} \nu_s(x)\,\mathrm{d}x\,\mathrm{d}s \\ &\lesssim \int_0^{s^*} \int \nu_s(x)\,\mathrm{d}x\,\mathrm{d}s = s^* = O(\sqrt{\varepsilon}\,\omega(\varepsilon)),\end{aligned}$$

which yields the desired statement (4.30).
Let us now consider the estimate 4.31. Note by Remark 4.15 the matrix $\tilde{\Sigma}_\tau^{-1}$ given by (4.26) is positive definite and the matrix we subtract is also positive definite. Therefore, it holds in the sense of quadratic forms

$$0 < \tilde{\Sigma}_\tau^{-1} = \Sigma_\tau^{-1} - \frac{1}{\Sigma_\tau^{-1}[\dot{\gamma}_\tau]} \Sigma_\tau^{-1} \otimes \Sigma_\tau^{-1} \le \Sigma_\tau^{-1}.$$

Now, the uniform bound (4.18) yields

$$\sqrt{\det\nolimits_{1,1}(Q_\tau^\top \tilde{\Sigma}_\tau^{-1} Q_\tau)} \le C_\Sigma^{\frac{n-1}{2}}.$$

Then, the desired statement (4.31) follows directly from the estimate (4.25). □

4.3.3. Reduction to neighborhood around the saddle

Firstly, observe that from (4.31) follows the a priori estimate

$$\frac{\mathcal{A}^2(x)}{\mu(x)} \lesssim \left(\frac{C_\Sigma}{2\pi\varepsilon}\right)^{n-1} Z_\mu\, e^{\frac{1}{\varepsilon}H(x)}. \tag{4.33}$$

Hence, on an exponential scale, the leading order contribution to the cost comes from neighborhoods of points where $H(x)$ is large. Therefore, we want to make the set, where H is comparable to its value at the optimal connecting saddle $s_{i,j}$, as small as possible. For this purpose, let us define the following set

$$\Xi_{\gamma,\Sigma} := \left\{ x \in \operatorname{supp}\mathcal{A} : H(x) \ge H(s_{i,j}) - \varepsilon\omega^2(\varepsilon) \right\}. \tag{4.34}$$

Therewith, we obtain by denoting the complement $\Xi^c_{\gamma,\Sigma} := \operatorname{supp}\mathcal{A}\backslash\Xi_{\gamma,\Sigma}$ the splitting

$$\mathcal{T}^2_\mu(\nu_i,\nu_j) \le \int_{\Xi_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)}\,\mathrm{d}x + \int_{\Xi^c_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)}\,\mathrm{d}x.$$

The integral on $\Xi^c_{\gamma,\Sigma}$ can be estimated with the a priori estimate (4.33) and Lemma 4.16 as follows

$$\begin{aligned}\int_{\Xi^c_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)}\,\mathrm{d}x &\overset{(4.34)}{\le} Z_\mu e^{\frac{H(s_{i,j})}{\varepsilon}-\omega^2(\varepsilon)} \int_{\Xi^c_{\gamma,\Sigma}} \mathcal{A}^2(x)\,\mathrm{d}x \\ &\overset{(4.31)}{\lesssim} Z_\mu e^{\frac{H(s_{i,j})}{\varepsilon}-\omega^2(\varepsilon)} \left(\frac{C_\Sigma}{2\pi\varepsilon}\right)^{\frac{n-1}{2}} \int \mathcal{A}(x)\,\mathrm{d}x \qquad (4.35)\\ &\overset{(4.30)}{\lesssim} Z_\mu e^{\frac{H(s_{i,j})}{\varepsilon}-\omega^2(\varepsilon)} \left(\frac{C_\Sigma}{2\pi\varepsilon}\right)^{\frac{n-1}{2}} T.\end{aligned}$$

We observe that estimate (4.35) is the second summand in the desired bound (4.19).

4.3.4. Cost estimate around the saddle

The aim of this subsection is to deduce the estimate

$$\int_{\Xi_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)}\,\mathrm{d}x \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, \frac{2\pi\varepsilon\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|}. \qquad (4.36)$$

Note that this estimate would yield the missing ingredient for the verification of the desired estimate (4.19).

By the non-degeneracy Assumption 1.12, we can assume that ε is small enough such that $E_0^- \cup E_T^+ \subset \Xi^c_{\gamma,\Sigma}$. It follows that $\Xi_{\gamma,\Sigma} \subset E$. We claim that the transport interpolation Φ_s can be chosen such that there exists a connected interval $I_T \subset [0,T]$ satisfying

$$\Xi_{\gamma,\Sigma} \subset \bigcup_{s\in I_T} (V_s + \gamma_s) \qquad \text{and} \qquad \mathcal{H}^1(I_T) = O(\sqrt{\varepsilon}\,\omega(\varepsilon)). \qquad (4.37)$$

Indeed, the level set $\{x \in \mathbb{R}^n : H(x) \le H(s_{i,j}) - \varepsilon\omega^2(\varepsilon)\}$ consists of at least two connected components M_i and M_j such that $m_i \in M_i$ and $m_j \in M_j$. Further, it holds

$$\operatorname{dist}(M_i, M_j) = \inf_{x\in M_i, y\in M_j} |x-y| = O(\sqrt{\varepsilon}\,\omega(\varepsilon)),$$

which follows from expanding H around $s_{i,j}$ in direction of the eigenvector corresponding to the negative eigenvalue of $\nabla^2 H(s_{i,j})$. We can choose the path γ in direction of this eigenvector in a neighborhood of size $O(\sqrt{\varepsilon}\,\omega(\varepsilon))$ around $s_{i,j}$, which shows (4.37).

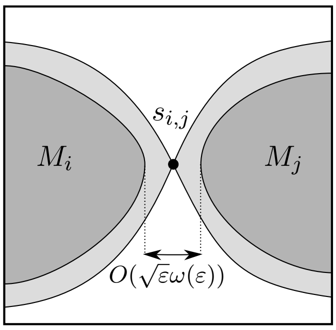

Figure 4.3.: Neighborhood around saddle

Combining the covering (4.37) and Lemma 4.13 yields the estimate

$$\int_{\Xi_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)}\,\mathrm{d}x \le \int_{I_T}\int_{V_s} \frac{\mathcal{A}^2(\gamma_s+z_s)}{\mu(\gamma_s+z_s)}\,\mathrm{d}z_s\,\mathrm{d}s. \tag{4.38}$$

Recalling the definition (4.21) of the family of rotations $(Q_\tau)_{\tau\in[0,T]}$, it holds that $z_\tau = Q_\tau z$ with $z \in \{0\}\times\mathbb{R}^{n-1}$. Hence, the following relation holds

$$\int_{I_T}\int_{V_\tau} \frac{\mathcal{A}^2(\gamma_\tau+z_\tau)}{\mu(\gamma_\tau+z_\tau)}\,\mathrm{d}z_\tau\,\mathrm{d}\tau = \int_{\{0\}\times\mathbb{R}^{n-1}}\int_{I_T} \mathbb{1}_{V_\tau}(Q_\tau z)\frac{\mathcal{A}^2(\gamma_\tau+Q_\tau z)}{\mu(\gamma_\tau+Q_\tau z)}\,\mathrm{d}\tau\,\mathrm{d}z. \tag{4.39}$$

The next step is to rewrite $H(\gamma_\tau+Q_\tau z)$. By the reason that $|z_\tau| = O(\sqrt{\varepsilon}\,\omega(\varepsilon))$ for $z_\tau\in V_\tau$ and the global non-degeneracy assumption (1.8), we can Taylor expand $H(\gamma_\tau+z_\tau)$ around $s_{i,j}=\gamma_{\tau^*}$ for $\tau\in I_T$ and $z_\tau = Q_\tau z\in V_\tau$. More precisely, we get

$$\begin{aligned} H(\gamma_\tau+Q_\tau z)-H(s_{i,j}) &= \frac{1}{2}\nabla^2 H(s_{i,j})[\gamma_\tau+Q_\tau z - s_{i,j}] + O(|\gamma_\tau+Q_\tau z-s_{i,j}|^3)\\ &= \frac{1}{2}\nabla^2 H(s_{i,j})[\gamma_\tau-\gamma_{\tau^*}] + \frac{1}{2}\nabla^2 H(s_{i,j})[Q_\tau z]\\ &\quad + \left\langle Q_\tau z, \nabla^2 H(s_{i,j})(\gamma_\tau-\gamma_{\tau^*})\right\rangle + O(|\gamma_\tau+Q_\tau z-\gamma_{\tau^*}|^3) \end{aligned}$$

Now, further expanding γ_τ and Q_τ in τ leads to

$$\gamma_\tau = \gamma_{\tau^*} + \dot\gamma_{\tau^*}(\tau-\tau^*) + O(|\tau-\tau^*|), \quad \text{and} \quad Q_\tau z = Q_{\tau^*}z + O(|\tau-\tau^*|\,|z|).$$

For the expansion of H, we arrive at the identity

$$\begin{aligned} &H(\gamma_\tau+Q_\tau z)-H(s_{i,j}) =\\ &\quad \frac{1}{2}\nabla^2 H(s_{i,j})[\dot\gamma_{\tau^*}(\tau-\tau^*)+O(|\tau-\tau^*|^2)] + \frac{1}{2}\nabla^2 H(s_{i,j})[Q_{\tau^*}z + O(|\tau-\tau^*|\,|z|)]\\ &\quad + \left\langle Q_{\tau^*}z + O(|\tau-\tau^*|\,|z|), \nabla^2 H(s_{i,j})\left(\dot\gamma_{\tau^*}(\tau-\tau^*)+O(|\tau-\tau^*|^2)\right)\right\rangle\\ &\quad + O(|\gamma_\tau+Q_\tau z-\gamma_{\tau^*}|^3)\\ &= \frac{1}{2}\nabla^2 H(s_{i,j})[\dot\gamma_{\tau^*}]\,(\tau-\tau^*)^2 + \frac{1}{2}\nabla^2 H(s_{i,j})[Q_{\tau^*}z] + \left\langle Q_{\tau^*}z, \nabla^2 H(s_{i,j})\dot\gamma_{\tau^*}\right\rangle(\tau-\tau^*)\\ &\quad + O(|\tau-\tau^*|^3, |z|\,|\tau-\tau^*|^2, |z|^2\,|\tau-\tau^*|, |z|^3). \end{aligned}$$

Using $|\tau-\tau^*| = O(\sqrt{\varepsilon}\;\omega(\varepsilon))$ and $|z| = O(\sqrt{\varepsilon}\;\omega(\varepsilon))$ we obtain for the error the estimate

$$O(|\tau-\tau^*|^3, |z|\,|\tau-\tau^*|^2, |z|^2\,|\tau-\tau^*|, |z|^3) = O(\varepsilon^{\frac{3}{2}}\;\omega^3(\varepsilon)).$$

The term $\left\langle Q_{\tau^*}z, \nabla^2 H(s_{i,j})\dot\gamma_{\tau^*}\right\rangle(\tau-\tau^*)$ in the expansion of H has no sign and has to vanish. This is only the case, if we choose $\dot\gamma_{\tau^*}$ as an eigenvector of $\nabla^2 H(s_{i,j})$ to the negative eigenvalue $\lambda^-(s_{i,j})$, because then

$$\left\langle Q_{\tau^*}z, \nabla^2 H(s_{i,j})\dot\gamma_{\tau^*}\right\rangle(\tau-\tau^*) = \lambda^-(s_{i,j})\left\langle Q_{\tau^*}z, \dot\gamma_{\tau^*}\right\rangle = 0.$$

Additionally, by this choice of $\dot\gamma_{\tau^*}$ the quadratic form $\nabla^2 H(s_{i,j})[\dot\gamma_{\tau^*}]$ evaluates to

$$\nabla^2 H(s_{i,j})[\dot\gamma_{\tau^*}] = \lambda^-(s_{i,j})|\dot\gamma_{\tau^*}|^2 = \lambda^-(s_{i,j}).$$

Therefore, we deduced the desired rewriting of $H(\gamma_\tau + Q_\tau z)$ as

$$H(\gamma_\tau + Q_\tau z) = H(s_{i,j}) - |\lambda^-(s_{i,j})|(\tau - \tau^*)^2 + \frac{1}{2}\nabla^2 H(s_{i,j})[Q_{\tau^*} z] + O(\varepsilon^{\frac{3}{2}}\ \omega^3(\varepsilon)). \tag{4.40}$$

From the regularity assumptions on the transport interpolation we can deduce that

$$\begin{aligned}\tilde\Sigma_\tau^{-1}[Q_\tau z] &= \tilde\Sigma_{\tau^*}^{-1}[Q_\tau z] + O(|\tau - \tau^*|\,|z|^2)\\ &= \tilde\Sigma_{\tau^*}^{-1}[Q_{\tau^*} z + O(|\tau-\tau^*|\,|z|)] + O(|\tau - \tau^*|\,|z|^2)\\ &= \tilde\Sigma_{\tau^*}^{-1}[Q_{\tau^*} z] + O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)).\end{aligned}$$

Then, it follows easily from the definition (4.25) of P_τ that

$$P_\tau \approx P_{\tau^*}. \tag{4.41}$$

Applying the cost estimate (4.25) of Lemma 4.14, the rewriting (4.40) of $H(\gamma_\tau + Q_\tau z)$ and the identity (4.41) yields the estimate

$$\frac{\mathcal{A}^2(\gamma_\tau + Q_\tau z)}{\mu(\gamma_\tau + Q_\tau z)} \lesssim Z_\mu e^{\frac{H(s_{i,j})}{\varepsilon}} P_{\tau^*}^2 e^{-\frac{(2\tilde\Sigma_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}))[Q_{\tau^*} z]}{2\varepsilon} - \frac{|\lambda^-(s_{i,j})|(\tau-\tau^*)^2}{2\varepsilon}}. \tag{4.42}$$

The exponentials are densities of two Gaussian, if we put an additional constraint on the transport interpolation. Namely, we postulate

$$2\tilde\Sigma_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}) > 0 \quad \text{on} \quad \operatorname{span} V_{\tau^*}$$

in the sense of quadratic forms. Note that $\operatorname{span} V_{\tau^*} = Q_{\tau^*}(\{0\} \times \mathbb{R}^{n-1}) = \operatorname{span}\{\dot\gamma_{\tau^*}\}^\perp$ is the tangent space of the stable manifold in the 1-saddle $s_{i,j}$. With this preliminary considerations we finally are able to estimate the right-hand side of (4.39) as follows

$$\begin{aligned}&\int_{\{0\}\times\mathbb{R}^{n-1}} \int_{I_T} \mathbb{1}_{V_\tau}(Q_\tau z)\, \frac{\mathcal{A}^2(\gamma_\tau + Q_\tau z)}{\mu(\gamma_\tau + Q_\tau z)}\, \mathrm{d}\tau\, \mathrm{d}z\\ &\overset{(4.42)}{\lesssim} Z_\mu\, e^{\frac{H(s_{i,j})}{\varepsilon}} \int_{\{0\}\times\mathbb{R}^{n-1}} \int_{I_T} P_{\tau^*}^2\, e^{-\frac{(2\tilde\Sigma_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}))[Q_{\tau^*} z]}{2\varepsilon} - \frac{\left|\lambda^-(s_{i,j})\right|(\tau-\tau^*)^2}{2\varepsilon}}\, \mathrm{d}\tau\, \mathrm{d}z\\ &\le Z_\mu\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, \frac{\sqrt{2\pi\varepsilon}}{\sqrt{|\lambda^-(s_{i,j})|}} \int_{\{0\}\times\mathbb{R}^{n-1}} P_{\tau^*}^2\, e^{-\frac{(2\tilde\Sigma_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}))[Q_{\tau^*} z]}{2\varepsilon}}\, \mathrm{d}z\\ &= Z_\mu\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, \frac{\sqrt{2\pi\varepsilon}}{\sqrt{|\lambda^-(s_{i,j})|}}\, P_{\tau^*}^2\, \frac{(2\pi\varepsilon)^{\frac{n-1}{2}}}{\sqrt{\det_{1,1}\left(Q_{\tau^*}^\top (2\tilde\Sigma_{\tau^*}^{-1} - \nabla^2 H(s_{i,j})) Q_{\tau^*}\right)}}\\ &= \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, \frac{2\pi\varepsilon}{\sqrt{|\lambda^-(s_{i,j})|}}\, \underbrace{\frac{\det_{1,1}(Q_{\tau^*}^\top \tilde\Sigma_{\tau^*}^{-1} Q_{\tau^*})}{\sqrt{\det_{1,1}\left(Q_{\tau^*}^\top (2\tilde\Sigma_{\tau^*}^{-1} - \nabla^2 H(s_{i,j})) Q_{\tau^*}\right)}}}_{\text{to optimize!}}.\end{aligned} \tag{4.43}$$

The final step consists of optimizing the choice of $\tilde\Sigma_{\tau^*}$. Let us use the notation $A = Q_{\tau^*}^\top \tilde\Sigma_{\tau^*}^{-1} Q_{\tau^*}$ and $B = Q_{\tau^*}^\top H(s_{i,j}) Q_{\tau^*}$. Then the minimization problem has the structure

$$\inf_{A\in\mathbb{R}^{n\times n}_{\text{sym},+}} \left\{ \frac{\det_{1,1} A}{\sqrt{\det_{1,1}(2A-B)}} : 2A - B > 0 \quad \text{on} \quad \{0\}\times\mathbb{R}^{n-1} \right\}. \tag{4.44}$$

In the appendix, we show in Lemma B.2 that the optimal value of (4.44) is attained at $\tilde{\Sigma}_{\tau^*}^{-1} = \nabla^2 H(s_{i,j})$ restricted V_{τ^*}. The optimal value is given by

$$\frac{\det_{1,1} A}{\sqrt{\det_{1,1}(2A-B)}} = \sqrt{\det_{1,1}(Q_{\tau^*}^\top \nabla^2 H(s_{i,j}) Q_{\tau^*})}.$$

Because V_{τ^*} is the tangent space of the stable manifold of the saddle $s_{i,j}$, it holds

$$\det_{1,1}(Q_{\tau^*}^\top \nabla^2 H(s_{i,j}) Q_{\tau^*}^\top) = \frac{\det(\nabla^2 H(s_{i,j}))}{\lambda^-(s_{i,j})} = \frac{|\det(\nabla^2 H(s_{i,j}))|}{|\lambda^-(s_{i,j})|}. \tag{4.45}$$

The final step is a combination of (4.38), (4.39), (4.43) and (4.45) to obtain the desired estimate (4.36).

For the verification of Lemma 4.12, it is only left to deduce the estimate (4.20). For that purpose we analyze the error terms in the estimate (4.19) i.e.

$$\mathcal{T}_\mu^2(\nu_i,\nu_j) \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, 2\pi\varepsilon \left(\underbrace{\frac{\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|}}_{=O(1)} + \underbrace{\frac{T(C_\Sigma)^{\frac{n-1}{2}}}{\sqrt{2\pi\varepsilon}} e^{-\omega^2(\varepsilon)}}_{=O(\varepsilon^{-\frac{1}{2}} e^{-\omega^2(\varepsilon)})} \right).$$

By the choice of $\omega(\varepsilon) \geq |\log\varepsilon|^{\frac{1}{2}}$, enforced by Lemma 4.7, we see that

$$O(\varepsilon^{-\frac{1}{2}} e^{-\omega^2(\varepsilon)}) = O(\sqrt{\varepsilon}).$$

Recalling, that "$\lesssim$" means "$\leq$" up to a multiplicative error of order $1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))$ we get

$$\mathcal{T}_\mu^2(\nu_i,\nu_j) \leq \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, 2\pi\varepsilon \frac{\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right) \left(1 + O(\sqrt{\varepsilon})\right).$$

The last inequality already yields the desired estimate (4.20) by using the observation

$$\left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right) \left(1 + O(\sqrt{\varepsilon})\right) = \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right).$$

Remark 4.17. Let us summarize the additional constraints on the transport interpolation besides the Assumption 4.11 of a regular affine transport interpolation to obtain the desired estimate (4.19):

- γ passes the saddle point $s_{i,j}$ at the passage time τ^* in direction of the eigenvector to the negative eigenvalue $\lambda^-(s_{i,j})$ of $\nabla^2 H(s_{i,j})$
- γ stays in the sublevel set $\left\{H(x) \leq H(s_{i,j}) - \varepsilon\omega^2(\varepsilon)\right\}$ up to a small time interval of order $\sqrt{\varepsilon}\,\omega(\varepsilon)$ around the passage time τ^*
- It holds $\Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{i,j})$ on the stable manifold of $s_{i,j}$.

4.4. Conclusion of the mean-difference estimate

With the help of Lemma 4.7 and Lemma 4.12 the proof of Theorem 2.23 is straightforward. We can estimate the mean-differences w.r.t. to the measure μ_i by introducing the means w.r.t. the approximations ν_i and ν_j.

$$\big(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\big)^2 = \big(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\nu_i}(f) + \mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\nu_j}(f) + \mathbb{E}_{\nu_j}(f) - \mathbb{E}_{\mu_j}(f)\big)^2$$

We apply the Young inequality with a weight that is motivated by the final total multiplicative error term $R(\varepsilon)$ in Theorem 2.23. More precisely,

$$\begin{aligned}\big(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\big)^2 &\le (1+\varepsilon^{\frac{1}{2}}\omega^3(\varepsilon))\big(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\nu_j}(f)\big)^2 \\ &\quad + 2(1+\varepsilon^{-\frac{1}{2}}\omega^{-3}(\varepsilon))\Big((\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\nu_i}(f))^2 + \big(\mathbb{E}_{\mu_j}(f) - \mathbb{E}_{\nu_j}(f)\big)^2\Big).\end{aligned}$$

Then, the estimate (4.8) of Lemma 4.7 yields

$$\big(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\big)^2 \le (1+\sqrt{\varepsilon}\,\omega^3(\varepsilon))\big(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\nu_j}(f)\big)^2 + O(\varepsilon)\int |\nabla f|^2 \,\mathrm{d}\mu, \tag{4.46}$$

which justifies the statement, that the approximation only leads to higher-order error terms in ε. An application of (4.1) to the estimate (4.46) transfers the mean-difference to the Dirichlet form with the help of the weighted transport distance

$$\big(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\big)^2 \le \big((1+\sqrt{\varepsilon}\,\omega^3(\varepsilon))\;\mathcal{T}_\mu^2(\nu_i,\nu_j) + O(\varepsilon)\big)\int |\nabla f|^2 \,\mathrm{d}\mu,$$

The weighted transport distance $\mathcal{T}_\mu(\nu_i,\nu_j)$ is dominating the above estimate. Finally, we arrive at the estimate

$$\big(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\big)^2 \lesssim \mathcal{T}_\mu^2(\nu_i,\nu_j)\int |\nabla f|^2 \,\mathrm{d}\mu.$$

Now, the Theorem 2.23 follows directly from an application of the estimate (4.20) of Lemma 4.12 and setting $\omega(\varepsilon) = |\log\varepsilon|^{\frac{1}{2}}$.

4.5. Transport across several saddles

In this section, we want to demonstrate that the transport technique introduces in Section 4.1 is flexible enough to handle degenerate cases, in which we are not enforcing Assumption 1.12.

4.5.1. Splitting of the transport

The transport cost $\mathcal{T}_\mu(\nu_0,\nu_1)$ from Definition 4.1 works with transport maps, or more precisely, with interpolations between transport maps. It would be desirable to translate the setting, like for the Wasserstein distance, to transport plans which admit more flexibility. However, from the structure of the definition of the weighted transport distance, it is not obvious how to achieve such a reformulation. A first step in this direction is the following idea. Let us represent the measures $\nu_0, \nu_1 \ll \mu$ as a possible infinite convex combination of measures $\{\nu_0^k\}_{k\in\mathbb{N}}$ and $\{\nu_1^i\}_{k\in\mathbb{N}}$, i.e.

$$\nu_0 = \sum_{k=1}^{\infty} p_k \nu_0^k \qquad \text{and} \qquad \nu_1 = \sum_{k=1}^{\infty} p_k \nu_1^k \qquad \text{with} \qquad \sum_{k=1}^{\infty} p_k = 1.$$

Then the mean-difference $(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f))^2$ can be decomposed into

$$(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f))^2 = \left(\sum_{k=1}^{\infty} p_k \left(\mathbb{E}_{\nu_0^k}(f) - \mathbb{E}_{\nu_1^k}(f)\right)\right)^2 .$$

Therewith, we can again do the estimate (4.1) and arrive at a slightly more general definition of the weighted transport distance

$$\mathcal{T}_\mu^2(\nu_0, \nu_1) = \inf_{\sum p_k = 1} \inf_{\Phi_s^1, \ldots, \Phi_s^l, \ldots} \int \left(\sum_{k=1}^{\infty} p_k \int_0^1 |\dot{\Phi}_s^k \circ (\Phi_s^k)^{-1}| \frac{\mathrm{d}\nu_s^k}{\mathrm{d}\mu} \,\mathrm{d}s\right)^2 \mathrm{d}\mu, \tag{4.47}$$

in which the first infimum denotes all possible convex representations of ν_0 and ν_1 such that the elements ν_0^k and ν_1^k are absolutely continuous w.r.t. to μ. The *cost density* is defined analogous to (4.3) by

$$\mathcal{A}^k(x) := \int_0^1 |\dot{\Phi}_s^k \circ (\Phi_s^k)^{-1}(x)| \, \nu_s^k(x) \,\mathrm{d}s.$$

4.5.2. Mean-difference estimate with several saddles

The preliminary consideration of Section 4.5.1 allows us to drop Assumption 1.12 (i). Therefore, let m_i and m_j be two local minima of H with domain of attraction Ω_i and Ω_j. Then let μ_i and μ_j be the measures obtained from μ by restricting to Ω_i and Ω_j. We want to recall the notion of the saddle height between m_i and m_j, which is given by

$$\widehat{H}(m_i, m_j) = \inf\left\{\max_{s\in[0,1]} H(\gamma(s)) : \gamma \in C([0,1], \mathbb{R}^n), \gamma(0) = m_i, \gamma(1) = m_j\right\}$$

Then, between m_i and m_j exists $N_{i,j} \geq 1$ distinct saddles $\{s_{i,j}^k\}_{k=1}^{N_{i,j}}$ satisfying

$$\widehat{H}(m_i, m_j) = H(s_{i,j}^1) = \cdots = H(s_{i,j}^{N_{i,j}}).$$

Therewith, we can reformulate Theorem 2.23 in the following way.

Theorem 4.18 (Mean-difference estimate in the degenerate case)**.** *The mean-differences between the measures μ_i and μ_j satisfy*

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon}{\sum_{k=1}^{N_{i,j}} \kappa_k^{-1}} e^{\frac{1}{\varepsilon}\widehat{H}(m_i,m_j)} \int |\nabla f|^2 \,\mathrm{d}\mu,$$

where $\lambda^-(s_{i,j})$ *denotes the negative eigenvalue of the Hessian* $\nabla^2 H(s_{i,j})$ *at the* 1*-saddle* $s_{i,j}$ *and* $\{\kappa_k\}_{k=1}^{N_{i,j}}$ *is given by*

$$\kappa_k = \frac{\sqrt{|\det \nabla^2 H(s_{i,j}^k)|}}{|\lambda_k^-|}.$$

Proof. The proof consists of the same steps as presented in Sections 4.2 to . Since we have all the preliminary work done and all necessary Lemmata by hand, we can present the proof in a very condensed form. Let us recall that the measures μ_i and μ_j (cf. (2.7)) are obtained from a Gibbs measure μ of a Hamiltonian H by restriction to the basin of attraction Ω_i (cf. (2.6)) of the minima m_i and m_j w.r.t. to the deterministic gradient flow. Since the proof relies on an explicit construction of a transport interpolation map, we have to use approximated measures ν_i and ν_j of μ_i and μ_j obtained by Gaussians with covariance matrices given by the Hessian $\nabla^2 H$ evaluated in the minima m_i and m_j restricted to small neighborhoods. In Lemma 4.7, where also the precise definition of ν_i and ν_j is given, we obtained the approximation in mean-difference by the estimate

$$(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\mu_i}(f))^2 \leq O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)) \int |\nabla f|^2 \,\mathrm{d}\mu. \tag{4.48}$$

Therewith, we can argue, like in the derivation of (4.46), that it is sufficient to estimate the mean-difference of the approximated measure $\left(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\nu_j}(f)\right)^2$. Therefore, we can now choose the modified transport given in (4.47). We can choose the convex combination $\nu_i = \sum_{k=1}^{N_{i,j}} p_k \nu_i$ with $\sum_{k=1}^{N_{i,j}} p_k = 1$. The strategy of the evaluation of the transport cost is the same as already presented in Section 4.3, but now we have to construct $N_{i,j}$ transport interpolations.

Every transport interpolation $(\Phi_s^k)_{s\in[0,T_k}$ for $k = 1, \ldots, N_{i,j}$ is parameterized by a path $(\gamma_s^k)_{s\in[0,T_k]}$ and a path of covariance matrices $(\Sigma_s^k)_{s\in[0,T_k]}$ satisfying for each k the Assumption 4.17. In other words, $(\Phi_s^k)_{s\in[0,T_k]}$ is for each k a *regular affine transport interpolation*. Then, we can introduce tube coordinates on every $\mathcal{A}^k$ separately as described in Section 4.3.2. Therewith, the cost densities $\mathcal{A}^k$ satisfies for $k = 1, \ldots, N_{i,j}$ the bounds obtained in Lemma 4.14 and Lemma 4.16.

Let us now use the splitting described in Section 4.3.3, where we introduced the set

$$\Xi_{\gamma,\Sigma} := \left\{ x \in \operatorname{supp} \mathcal{A} : H(x) \geq \widehat{H}(m_i, m_j) - \varepsilon\omega^2(\varepsilon) \right\}$$

and its complement $\Xi_{\gamma,\Sigma}^c := \operatorname{supp} \mathcal{A} \backslash \Xi_{\gamma,\Sigma}$. We can choose the transport paths $(\gamma_s)_{s\in[0,T_k]}$, such that γ_s^k passes nearby the saddle $s_{i,j}^k$ for some s. Then it holds that

$$\Xi_{\gamma,\Sigma} \subset \bigcup_{k=1}^{N_{i,j}} B_{C\sqrt{\varepsilon}\omega(\varepsilon)}(s_{i,j}^k), \qquad \text{for } C > 0 \text{ large enough.}$$

Hence, for ε small enough, $\Xi_{\gamma,\Sigma}$ consists of $N_{i,j}$ connected components denoted by $\Xi^k_{\gamma,\Sigma}$ for $k = 1, \dots, N_{i,j}$ and therefore the following decomposition for $\Xi_{\gamma,\Sigma}$ holds

$$\Xi_{\gamma,\Sigma} = \bigcup_{k=1}^{N_{i,j}} \Xi^k_{\gamma,\Sigma}.$$

The estimate on $\Xi^c_{\gamma,\Sigma}$ works in the same ways as already done in (4.35). With the splitting of $\Xi_{\gamma,\Sigma}$ into $\left\{\Xi^k_{\gamma,\Sigma}\right\}_{k=1}^{N_{i,j}}$ we are in the same setting like for only one optimal communicating saddle by considering every $\Xi^k_{\gamma_\Sigma}$ separately. Therefore, we can apply the estimate (4.43) and arrive after the matrix optimization procedure at

$$\mathcal{T}_\mu(\nu_i, \nu_j) \lesssim \inf_{\sum p_k = k} \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, e^{\frac{\widehat{H}(m_i,m_j)}{\varepsilon}} \, 2\pi\varepsilon \sum_{k=1}^{N_{i,j}} p_i^2 \frac{\sqrt{|\det \nabla^2 H(s^k_{i,j})|}}{|\lambda^-_k|},$$

where λ^-_k is the negative eigenvalue of $\nabla^2 H(s^k_{i,j})$. With the definition of κ_k, we end up with the optimization problem

$$\mathcal{T}_\mu(\nu_i, \nu_j) \lesssim \inf_{\sum p_k = 1} \sum_{k=1}^{N_{i,j}} p_k^2 \kappa_k. \tag{4.49}$$

The solution to (4.49) is given by

$$p_k = \frac{\kappa_k^{-1}}{\sum_{k=1}^{N_{i,j}} \kappa_k^{-1}}.$$

In this case the transport cost evaluates to

$$\mathcal{T}_\mu(\nu_i, \nu_j) \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, e^{\frac{\widehat{H}(m_i,m_j)}{\varepsilon}} \, \frac{2\pi\varepsilon}{\sum_{k=1}^{N_{i,j}} \kappa_k^{-1}}, \qquad \text{where} \qquad \kappa_k = \frac{\sqrt{|\det \nabla^2 H(s^k_{i,j})|}}{|\lambda^-_k|}.$$

The last step is to use (4.48) to obtain the estimate for the mean-difference between μ_i and μ_j as described in Section 4.4. □

4.6. Relation to capacities

The approach to obtain the Eyring-Kramers formula for the Poincaré inequality in the works [BEGK04] and [BGK05] consists of two steps, in both of them capacities play a crucial role. Sharp estimates at low temperature for the capacities are deduced in [BEGK04]. In the second part [BGK05], the capacity of certain neighborhoods of local minima of H is sharply identified with the spectrum of the generator L. Let us take a look at how these capacities are defined and how one can derive estimates for them.

Definition 4.19 (Equilibrium potential and capacity)**.** For two regular subsets $A, B \subset \mathbb{R}^n$ define the *equilibrium potential* $h_{A,B}$ as the solution of the Dirichlet problem

$$\begin{aligned} Lh_{A,B}(x) &= 0, & x &\in (A \cup B)^{\mathsf{c}} \\ h_{A,B}(x) &= 1, & x &\in A \\ h_{A,B}(x) &= 0, & x &\in B. \end{aligned}$$

Therewith, the *capacity* $\operatorname{cap}(A, B)$ is defined as

$$\operatorname{cap}(A,B) = \varepsilon \int_{(A\cup B)^{\mathsf{c}}} |\nabla h_{A,B}|^2 \, \mathrm{d}\mu = \mathcal{E}(h_{A,B}),$$

where $\mathcal{E}$ is the Dirichlet form (1.2) for the operator L (1.1).

Upper bounds for the capacity can be easily obtained by the variational characterization from the *Dirichlet principle* which states

$$\operatorname{cap}(A,B) = \inf \left\{ \mathcal{E}_{(A\cup B)^{\mathsf{c}}}(h) : h \in H^1((A\cup B)^{\mathsf{c}}), h|_A \equiv 1, h|_B \equiv 0 \right\}.$$

Therefore, we only need to have a good guess of the equilibrium potential $h_{A,B}$. However, a variational principle for obtaining lower bounds is not available at this time for the continuous case. Therefore, the work [BEGK04] rely on certain monotonicity properties of the capacity and good approximations for the equilibrium potential $h_{A,B}$.

In a recent work, concerning the discrete case, Bianchi, Bovier und Ioffe [BBI09] rediscovered a powerful dual variational representation to the *Dirichlet principle* going back to ideas of Berman and Konsowa [BK90]. It is used in [BBI09] and [BdHS10] to obtain sharp lower bounds for the capacities describing metastable behavior in Ising type models. It is therefore called discrete *Berman-Konsowa principle*. The next proposition indicates, that the weighted transport distance $\mathcal{T}_\mu$ from Definition 4.1 could be the continuous counterpart, which is not available at the moment.

Proposition 4.20 (Lower bound of the capacity in terms of $\mathcal{T}_\mu$)**.** *Let $A, B \subset \mathbb{R}^n$ be regular and disjoint. Further define by $\mu \llcorner A$ and $\mu \llcorner B$ the restriction of μ to A and B, respectively. Then, the following estimate holds*

$$\frac{\varepsilon}{\mathcal{T}_\mu^2(\mu\llcorner A, \mu\llcorner B)} \le \operatorname{cap}(A,B). \tag{4.50}$$

Proof. From the Definition 4.1 and the derivation (4.1) of the weighted transport distance follows

$$\left(\mathbb{E}_{\mu\llcorner A}(f) - \mathbb{E}_{\mu\llcorner B}(f)\right)^2 \le \mathcal{T}_\mu^2(\mu\llcorner A, \mu\llcorner B) \int |\nabla f|^2 \, \mathrm{d}\mu = \mathcal{T}_\mu^2(\mu\llcorner A, \mu\llcorner B)\varepsilon^{-1}\mathcal{E}(f),$$

where $\mathcal{E}$ is the Dirichlet form (1.2). We are free to choose the test function f, as long as the Dirichlet form is finite. Therefore, we take $f = h_{A,B}$ from Definition 4.19 and arrive at

$$1 \le \mathcal{T}_\mu^2(\mu\llcorner A, \mu\llcorner B)\varepsilon^{-1} \operatorname{cap}(A,B),$$

which finishes the proof. □

We want to close this section with some remarks on whether equality in (4.50) can be attained or not.

Remark 4.21 (Optimality of Proposition 4.20). The first observation is obvious. We can just compare the result of Theorem 2.23 with the sharp estimation of the capacity in [BEGK04, Theorem 3.1]. We observe that the bound (4.50) becomes sharp for $\varepsilon \to 0$.

Remark 4.22 (Modified weighted transport distance). It may be important for the non-low temperature regime, to use a slightly modified definition of the weighted transport distance. Instead of applying the Cauchy-Schwarz inequality to $\langle \nabla f, \dot{\Phi}_s \circ \Phi_s^{-1} \rangle$ in (4.1) it might be better to use this estimate

$$\left| \left\langle \nabla f, \dot{\Phi}_s \circ \Phi_s^{-1} \right\rangle \right| \leq |\nabla f| \left| \left\langle \frac{\nabla f}{|\nabla f|}, \dot{\Phi}_s \circ \Phi_s^{-1} \right\rangle \right| \leq |\nabla f| \sup_{\xi \in S^{n-1}} \left| \left\langle \xi, \dot{\Phi}_s \circ \Phi_s^{-1} \right\rangle \right|. \tag{4.51}$$

Combining (4.51) with (4.1) leads to the improved estimate

$$\begin{aligned} (\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f))^2 &= \left(\int_0^1 \int \langle \nabla f, \dot{\Phi}_s \circ \Phi_s^{-1} \rangle \, \mathrm{d}\nu_s \, \mathrm{d}s \right)^2 \\ &\leq \left(\int |\nabla f| \sup_{\xi \in S^{n-1}} \int_0^1 \left\langle \xi, \dot{\Phi}_s \circ \Phi_s^{-1} \right\rangle \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \, \mathrm{d}\mu \right)^2 \\ &\leq \int \left(\sup_{\xi \in S^{n-1}} \int_0^1 \left\langle \xi, \dot{\Phi}_s \circ \Phi_s^{-1} \right\rangle \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right)^2 \mathrm{d}\mu \int |\nabla f|^2 \, \mathrm{d}\mu. \end{aligned} \tag{4.52}$$

Furthermore, the factor in front of the Dirichlet form in (4.52) can be simplified to

$$\begin{aligned} \sup_{\xi \in S^{n-1}} \int_0^1 \left\langle \xi, \dot{\Phi}_s \circ \Phi_s^{-1} \right\rangle \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s &= \sup_{\xi \in S^{n-1}} \left\langle \xi, \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1} \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right\rangle \\ &= \left| \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1} \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right|. \end{aligned}$$

This leads to the definition of the modified weighted transport distance

$$\widetilde{\mathcal{T}}_\mu(\nu_0, \nu_1) := \inf_{(\Phi_s)_{s \in [0,1]}} \int \left| \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1} \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right|^2 \mathrm{d}\mu, \tag{4.53}$$

where $(\Phi_s)_{s \in [0,1]}$ is a transport interpolation from ν_0 to ν_1 absolutely continuous in s. In the low temperature regime, i.e. ε small, the integral in $\mathrm{d}s$ of (4.53) concentrates around some s^* and only the value $\Phi_{s^*} \circ \Phi_{s^*}^{-1}$ contributes to the integral. Therefore, in this case the modified weighted transport distance coincides with Definition 4.1 of the weighted transport distance. However, in general it always holds the relation $\widetilde{\mathcal{T}}_\mu(\nu_0, \nu_1) \leq \mathcal{T}_\mu(\nu_0, \nu_1)$, which is obvious from the estimate (4.51). This relation improves the bound (4.50) to

$$\frac{\varepsilon}{\mathcal{T}_\mu^2(\mu \llcorner A, \mu \llcorner B)} \leq \frac{\varepsilon}{\widetilde{\mathcal{T}}_\mu^2(\mu \llcorner A, \mu \llcorner B)} \leq \operatorname{cap}(A, B).$$

Remark 4.23 (Benamou-Brenier formula). In this dissertation, the weighted transport distance is a tool to get quantitative mean-difference estimates. However, more theoretical insights, like the question which spaces it metrizises and convexity properties would be worthwhile. Therefore, under suitable regularity assumptions on the class of transport interpolations $\{\Phi_s\}_{s\in[0,1]}$, for details it shall be referred to Chapter 8 of the book from Ambrosio, Gigli and Savaré [AGS05], a *Benamou-Brenier type formula* for the modified weighted transport distance $\widetilde{\mathcal{T}}_\mu(\cdot,\cdot)$ from Remark 4.22 could be obtained as follows

$$\widetilde{\mathcal{T}}^2_\mu(\nu_0,\nu_1) := \inf_{v\in\mathcal{V}(\nu_0,\nu_1)} \int \left| \int_0^1 v_s \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right|^2 \mathrm{d}\mu,$$

where $\mathcal{V}(\nu_0,\nu_1)$ is the set of admissible vector fields

$$\mathcal{V}(\nu_0,\nu_1) := \left\{ v_s \in L^1(\nu_s) : \text{ such that } \partial_s\nu_s + \nabla\cdot(v_s\nu_s) = 0 \right\}. \tag{4.54}$$

The continuity equation in (4.54) is understood in the sense of distributions, i.e.

$$\int_0^1 \int \left(\partial_s\varphi(s,x) + \langle v_s(x), \nabla\varphi(s,x)\rangle \right) \nu_s(\mathrm{d}x) \, \mathrm{d}s = 0, \quad \forall \varphi \in C_0^\infty([0,1]\times\mathbb{R}^n).$$

Chapter 5

The local Poincaré inequality revisited

In this chapter, we are considering only one of the basins of attraction Ω_i and therefore we omit the index i. We will write Ω and μ instead of Ω_i and μ_i, respectively. Further, we assume w.l.o.g. that $0 \in \Omega$ is the unique minimum in Ω.

We start with some heuristics for the validity of Theorem 2.19, the local Poincaré inequality. We consider the Gibbs measure μ restricted to a basin of attraction Ω given by

$$\mu(\mathrm{d}x) = \frac{\mathbb{1}_\Omega(x)}{Z_\mu} \exp\left(-\frac{H(x)}{\varepsilon}\right) \mathrm{d}x.$$

We have to show that this restricted Gibbs measure μ satisfies PI(ϱ) with constant $\varrho^{-1} = O(\varepsilon)$. We argue as follows: On a basin of attraction Ω can only be one local minimum of H, which is located w.l.o.g. at 0. Therefore, on Ω cannot exist another metastable state of the diffusion ξ_t given by (1.3). This means that for small noise ε the diffusion ξ_t slides down on the energy landscape without any obstacle until it reaches a small region around the local minimum of H at 0. This heuristically implies that only the small region around 0 is important for the Poincaré constant ofμ. Therefore, the Gibbs measure μ restricted to the basin of attraction Ω and the Gibbs measure μ restricted to a small neighborhood of 0 should have the same Poincaré constant in terms of scaling in ε. However, around the local minimum 0 the Hamiltonian H is strictly convex due to non-degeneracy assumption (1.8). Now, an application of the criterion of Bakry-Émery [BE85] – it connects convexity of $\frac{H}{\varepsilon}$ to the Poincaré constant of μ – yields that the Gibbs measure μ restricted to a small region around 0 should satisfy the PI(ϱ) with constant $\varrho^{-1} = O(\varepsilon)$.

Let us turn to the main idea of the proof of Theorem 2.19. In chapter 3 we succeeded to prove Theorem 2.19 by using the Lyapunov condition. There, we constructed an explicit Lyapunov function (cf. Section 3.1.1) with subtle properties (cf. Lemma 3.11), especially around critical points x, i.e. $\nabla H(x) = 0$.

Here, we want to pursue a different idea, which completely avoids the use of a Lyapunov argument and is self-containing. Moreover, the following method will allow a more precise quantitative control on the Poincaré constant. This could be particular interesting of the energy landscape consists just of one degenerate minimum, then the method shall be able to show that the Poincaré constant is dominated by the diffusion on the minimum manifold for low temperature.

The proof is motivated by the one-dimensional case, i.e. $H : \mathbb{R} \to \mathbb{R}$. In this situation, Theorem 2.19 can be deduced by using the Muckenhoupt functional (cf. Section 5.3 and [Muc72]), which determines the Poincaré constant in one dimension up to a universal factor. Because of the non-existence of a multidimensional version of the Muckenhoupt functional, we have to reduce the multidimensional case to the one-dimensional case.

For this purpose, we introduce on Ω polar like coordinates $\psi_\eta(r) \in \Omega$ for $\eta \in S^{n-1}$ and $r \geq 0$ (cf. Section 5.1). The coordinates $\psi_\eta(r)$ are going to be a small perturbation of the coordinates resulting from the deterministic gradient-flow

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi_t = -\nabla H(\xi_t),$$

parameterized by arc-length i.e. $|\dot{\psi}_\eta(r)| = 1$. In these coordinates, the restricted Gibbs measures becomes

$$\mu(\mathrm{d}r, \mathrm{d}\eta) = \frac{\mathbb{1}_\Omega(\psi_\eta(r))}{Z_\mu}\, j_\eta(t)\, \exp\left(-\frac{1}{\varepsilon}H(\psi_\eta(r))\right)\, \mathrm{d}r\, \mathrm{d}\eta, \tag{5.1}$$

where $j_\eta(r)$ is the Jacobian determinant of the coordinate transformation.

Now, we carry out a two-scale argument similar to the one used for the proof of the Eyring-Kramers formula. So, the restricted Gibbs measure $\mu(\mathrm{d}r, \mathrm{d}\eta)$ is decomposed into

$$\mu(\mathrm{d}r, \mathrm{d}\eta) = \mu(\mathrm{d}r|\eta)\, \hat{\mu}(\mathrm{d}\eta), \tag{5.2}$$

where the conditional measures $\mu(\mathrm{d}r|\eta)$ and the marginal $\hat{\mu}(\mathrm{d}\eta)$ are given according to Definition 2.4 by

$$\mu(\mathrm{d}r|\eta) = \frac{1}{Z_\mu \hat{\mu}(\eta)} e^{-\frac{1}{\varepsilon}H(\psi_\eta(r))} j_\eta(r)\, \mathrm{d}r \tag{5.3}$$

$$\hat{\mu}(\mathrm{d}\eta) = \frac{1}{Z_\mu} \int e^{-\frac{1}{\varepsilon}H(\psi_\eta(r))} j_\eta(r)\, \mathrm{d}r\, \mathrm{d}\eta. \tag{5.4}$$

Therewith, we get by an application of Corollary 2.7

$$\begin{aligned} \operatorname{var}_{\mu(\mathrm{d}x)}(f(x)) &= \operatorname{var}_{\mu(\mathrm{d}r,\mathrm{d}\eta)}(f(\psi_\eta(r))) \\ &= \mathbb{E}_{\hat{\mu}(\mathrm{d}\eta)}\left(\operatorname{var}_{\mu(\mathrm{d}r|\eta)}(f(\psi_\eta(r)))\right) + \operatorname{var}_{\hat{\mu}(\mathrm{d}\eta)}\left(\mathbb{E}_{\mu(\mathrm{d}r|\eta)}(f(\psi_\eta(r)))\right). \end{aligned} \tag{5.5}$$

Let us consider the first term on the right-hand side of the last equation. Note that the conditional measures $\mu(\mathrm{d}r|\eta)$ are one-dimensional. Hence, we are able to deduce

Poincaré inequalities for the conditional measures $\mu(\mathrm{d}r|\eta)$ for η fixed using the Muckenhoupt functional. Of course, this step is very sensitive to the choice of the coordinates $\{\psi_\eta\}_{\eta \in S^{n-1}}$. However, a careful and non-trivial construction of the coordinates yields the desired scaling of the Poincaré constant. The next statement contains the existence of *good* coordinates $\{\psi_\eta\}_{\eta \in S^{n-1}}$. At this point we have to enforce slightly changed properties of H at infinity in comparison to Assumption 1.9 to obtain certain monotonicity properties of the Hamiltonian evaluated along $\psi_\eta(\cdot)$ for fixed η. Namely, for this chapter we assume instead of (**A1**$_{\mathrm{PI}}$) and (**A2**$_{\mathrm{PI}}$) the following assumptions:

Assumption 5.1. *We assume that $H \in C^3(\mathbb{R}^n, \mathbb{R})$ is a Morse function. Further for some constants $C_H > 0$ and $K_H \geq 0$ holds*

$$\liminf_{|x|\to\infty} |\nabla H| \geq C_H. \tag{$\mathbf{A1}_{\mathrm{PI}}$}$$

$$\liminf_{|x|\to\infty} |\nabla H| - \nabla \cdot \left(\frac{\nabla H}{|\nabla H|}\right) \geq -K_H \tag{$\mathbf{A2'}_{\mathrm{PI}}$}$$

Proposition 5.2 (Existence of mixing coordinates)**.** *Assume that H satisfies the Assumption 5.1. Then, there exist coordinates $\{\psi_\eta\}_{\eta\in S^{n-1}}$ on Ω such that the conditional measures $\mu(\mathrm{d}r|\eta)$ are radial-mixing measures in the sense of Definition 5.14.*

The definition of radial-mixing measures is technically and for the moment means that the one-dimensional measures $\mu(\mathrm{d}r|\eta)$ have a good Poincaré constant (cf. Proposition 5.3). This is achieved, because the Hamiltonian evaluated along coordinates lines $H(\psi_\eta(r))$ fulfills certain monotonicity assumptions and the Jacobian determinant $j_\eta(r)$ of the coordinate transform shows a controlled blowup behavior, not causing any metastabilites.

Proposition 5.3 (PI for radial-mixing measures)**.** *Let $\mu(\cdot|\eta)$ be a radial-mixing measure in the sense of Definition 5.14. Then the measure $\mu(\cdot|\eta)$ satisfies* $\mathrm{PI}(\varrho(\eta))$ *with constant*

$$\frac{1}{\varrho(\eta)} = O(\varepsilon) \qquad \textit{for} \quad \varepsilon \to 0, \qquad \textit{uniformly in} \quad \eta \in S^{n-1}.$$

The proof of Proposition 5.3 is carried out in the Section 5.3. It is based on some properties of radial-mixing measures that are outlined in Section 5.1.

Remark 5.4 (Special case one dimension). The splitting (5.2) is not necessary in one dimension, since there we can directly apply the radial Poincaré inequality of Proposition 5.3 to the restricted measure μ itself. In particular, note that the measure $\hat{\mu}$ would live on $S^0 = \{-1, 1\}$ being disconnected and not satisfy any mixing properties.

Now, let us consider the second term on the right-hand side of (5.5), which is given by

$$\mathrm{var}_{\hat{\mu}(\mathrm{d}\eta)}(\hat{f}(\eta)) = \int \left(\hat{f}(\eta) - \int \hat{f}(\theta)\, \hat{\mu}(\mathrm{d}\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta) \tag{5.6}$$

by using the notation

$$\hat{f}(\eta) = \mathbb{E}_{\mu(\mathrm{d}r|\eta)}(f(\psi_\eta(r))).$$

Unfortunately, an application of the same strategy as in [GOVW09] would not yield the desired estimate. The reason is that, following [GOVW09], one would apply a Poincaré inequality for the marginal measure $\hat{\mu}(\mathrm{d}\eta)$ and then estimate the coarse-grained gradient by the full gradient via a covariance estimate. However, the Poincaré constant for the marginal measure $\hat{\mu}$ already is at least of order 1, since $\hat{\mu}$ lives on the sphere S^{n-1}. Therefore, this strategy could not yield the desired Poincaré inequality $\mathrm{PI}(\hat{\varrho})$ with constant of order $\hat{\varrho}^{-1} = O(\varepsilon)$.

For that reason we use a different strategy. Following the proof of Corollary 2.26, we represent the polar variance given by (5.6) as mean-difference. More precisely, the polar variance takes the form

$$\operatorname{var}_{\hat{\mu}(\mathrm{d}\eta)}(\hat{f}(\eta)) = \frac{1}{2} \int\int \left(\hat{f}(\eta) - \hat{f}(\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta)\, \hat{\mu}(\mathrm{d}\theta). \tag{5.7}$$

From this starting point, the procedure is somehow similar to the proof of the mean-difference estimate of Theorem 2.23. However, the procedure needs some more evolved ingredients due to the fact that the marginal measure $\hat{\mu}$ lives on the continuous state space S^{n-1} and the conditional measures $\mu(\mathrm{d}r|\eta)$ have one-dimensional support. The argument is outlined in detail in Section 5.4, in which the following statement is deduced.

Proposition 5.5 (Polar mean-difference estimate). *It holds the estimate*

$$\int\int \left(\hat{f}(\eta) - \hat{f}(\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta)\, \hat{\mu}(\mathrm{d}\theta) \leq O(\varepsilon) \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Now, we have provided all the ingredients that are needed for the proof of Theorem 2.19.

Proof of Theorem 2.19. The starting point of the proof is formula (5.5). Proposition 5.2 allows us to apply the local Poincaré inequality of Proposition 5.3, which leads to the estimate

$$\begin{aligned}\operatorname{var}_{\mu(\mathrm{d}x)}(f(x)) &\leq O(\varepsilon) \int\int |\partial_r f(\psi_\eta(r))|^2 \,\mu(\mathrm{d}r|\eta)\, \hat{\mu}(\mathrm{d}\eta) + \operatorname{var}_{\hat{\mu}(\mathrm{d}\eta)}(\hat{f}(\eta)) \\ &= O(\varepsilon) \int\int |\nabla f(\psi_\eta(r))|^2 \underbrace{|\dot{\psi}_\eta(r)|^2}_{=1} \mu(\mathrm{d}r, \mathrm{d}\eta) + \operatorname{var}_{\hat{\mu}(\mathrm{d}\eta)}(\hat{f}(\eta)).\end{aligned}$$

We now need to consider the second term on the right-hand side of the last inequality. The polar variance $\operatorname{var}_{\hat{\mu}(\mathrm{d}\eta)}(\hat{f}(\eta))$ is expressed by (5.7) as a mean-difference. Hence, an application of the estimate of Proposition 5.5 yields the desired statement. □

It is only left to verify the ingredient used in the proof of Theorem 2.19. In Section 5.1, we introduce the notion of radial-mixing measures and state some important properties of them. In Section 5.2, we carry out the proof of Proposition 5.2. In Section 5.3, we will prove Proposition 5.3. Finally, in Section 5.4, we state the proof of Proposition 5.5.

5.1. Local coordinates and radial-mixing measures

In this subsection we discuss the right choice of local coordinates $(\psi_\eta)_{\eta\in S^{n-1}}$ in order to apply the two-scale proof of Theorem 2.19. Moreover, we introduce the notion of radial-mixing measures (cf. Definition 5.14) and state some properties, which are needed in the proof of Proposition 5.3 in Section 5.3.

An obvious question is: Why are we not using the coordinates $\psi_\eta(r)$ resulting from the deterministic gradient-flow

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi_t = -\nabla H(\xi_t),$$

reparameterized by arc-length i.e. $|\dot{\psi}_\eta(r)| = 1$? The Poincaré inequality of μ is strongly connected to ergodic properties of the diffusion ξ_t given by (1.3). Intuitively, these coordinates should describe the diffusion ξ_t very well as we are only interested in the regime of vanishing noise ε. For this reason, deterministic gradient-flow coordinates $\psi_\eta(r)$ should be a natural choice of coordinates.

However, we are not using these deterministic gradient-flow coordinates because the conditional measures

$$\mu(\mathrm{d}r|\eta) = \frac{1}{Z_\mu \hat{\mu}(\eta)} e^{-\frac{1}{\varepsilon}H(\psi_\eta(r))} j_\eta(r)\,\mathrm{d}r$$

given by (5.3) will not show the right scaling of the Poincaré constant. A detailed analysis reveals that problems only arise in neighborhoods of critical points x, i.e. $\nabla H(x) = 0$. Therefore, we will only change the coordinates in neighborhoods around critical points (cf. (**$\tilde{\mathbf{H}}$1**) of Assumption 5.6). Heuristically, this observation seems to be plausible as, at critical points, the gradient of H becomes comparable to the noise. Hence, at least there, stochastic effects have to be taken into account. The main stochastic effect to be considered is that the noise does not have a preferred direction. Thus, it is better to smoothly transform the deterministic gradient-flow coordinates around local extremes to polar coordinates (cf. (**$\tilde{\mathbf{H}}$2**) of Assumption 5.6).

However, a more serious problem arises if one considers directions $\eta \in S^{n-1}$ such that the coordinate lines $\psi_\eta(r)$ almost hit a saddle point of H. For such directions η, the Hamiltonian

$$H_\eta(r) := \frac{1}{\varepsilon}H(\psi_\eta(r)) - \log j_\eta(r) \tag{5.8}$$

of the measure $\mu(\mathrm{d}r|\eta)$ has a metastability for the deterministic gradient-flow coordinates $\psi_\eta(r)$, since $j_\eta(r)$ becomes unbounded to $+\infty$. Here, the following stochastic effect has to be taken into account: If one follows a trajectory of the diffusion ξ_t given by (1.3) starting anywhere above the saddle, one will almost never get too close to the saddle point. The reason is that the process ξ_t will be pushed away from the saddle as soon as the noise becomes comparable to $|\nabla H(\xi_t)|$. This stands in contrast to deterministic gradient-flow coordinates, where you can get as close to the saddle as you want. Heuristically, this shows that deterministic gradient-flow coordinates are not appropriate at the saddle for the stochastic dynamics.

As you can see by Figure 5.1, the deterministic gradient-flow coordinates are adjusted around the saddles in such a way that a particle will never get close to the saddle by following a flow-line starting from above the saddle. Technically, these ideas are manifested in the properties (**H̃3**), (**H̃4**) and (**H̃5**). Later, these properties guarantee in the proof of Proposition 5.3 that in the Hamiltonian $H_\eta(r)$ there are no metastabilites for any direction η (cf. Section 5.3).

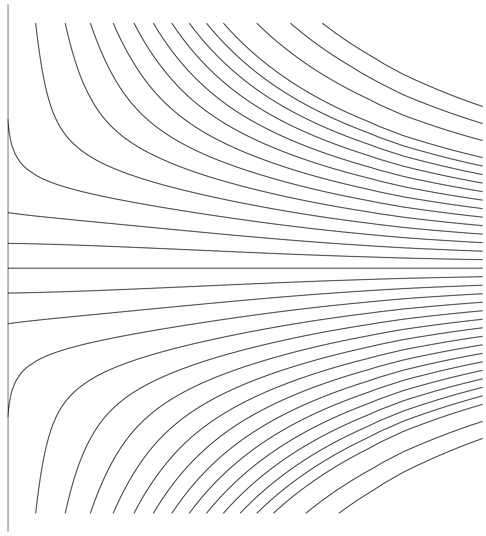

Figure 5.1.: Deterministic gradient-flow coordinates obtained from $\tilde{H}$

As we have outlined above, the desired coordinates $\psi_\eta(r)$ will be slight modifications of the coordinates resulting from the deterministic gradient-flow w.r.t. the Hamiltonian H. An easy way to construct such coordinates is to consider a small perturbation $\tilde{H}$ of the Hamiltonian H. Then, we define the modified coordinates $\psi_\eta(r)$ by the deterministic gradient-flow w.r.t. the perturbation $\tilde{H}$. In the next assumption we state some properties of the perturbation $\tilde{H}$ that allow to derive coordinates from it.

Assumption 5.6 (Perturbation $\tilde{H}$ of H)**.** *We assume that the perturbation $\tilde{H}$ is of class $C^2(\Omega\setminus\{0\}) \cap C^0(\overline{\Omega})$ and satisfies the conditions:*

(i) $\tilde{H}$ equals H up to small δ-neighborhoods around the critical points of H, i.e. by recalling that $\mathcal{S}$ is the set of critical points of H we have

$$\forall x \notin \bigcup_{y\in\mathcal{S}} B_\delta(y) : \ \tilde{H}(x) = H(x). \tag{$\tilde{\mathbf{H}}$1}$$

(ii) $\tilde{H}$ restricted to a ball around the local minimum located in 0 of radius $\frac{\delta}{2} > 0$ is spherically symmetric and linear, i.e.

$$\forall x \in B_{\frac{\delta}{2}}(0) : \ \tilde{H}(x) = \tilde{H}(0) + |x| \,. \tag{$\tilde{\mathbf{H}}$2}$$

(iii) $|\nabla\tilde{H}|$ is uniformly bounded from below in $\Omega\setminus\{0\}$, i.e.

$$\exists c_{\tilde{H}} > 0 \ \forall x \in \Omega\setminus\{0\} : \ |\nabla\tilde{H}(x)| \geq c_{\tilde{H}}. \tag{$\tilde{\mathbf{H}}$3}$$

By using the Assumption 5.6, we are able to define the coordinates $\psi_\eta(r)$ as follows.

Definition 5.7 (Local coordinates)**.** Assume that $\tilde{H}$ satisfies the Assumption 5.6. For $r \in \mathbb{R}$, we consider the associated flow $\Psi_r : \Omega\setminus\{0\} \to \Omega\setminus\{0\}$ defined as the solution of

$$\Psi_0(x) = x \qquad \text{and} \qquad \dot{\Psi}_r(x) = \tilde{F}(\Psi_r(x)) := \frac{\nabla\tilde{H}(\Psi_r(x))}{|\nabla\tilde{H}(\Psi_r(x))|}. \tag{5.9}$$

For $\eta \in S^{n-1}$, the coordinates $\psi_\eta : (0, T_\eta) \to \mathbb{R}^n$ are defined by

$$\psi_\eta(r) = \Psi_{r-\frac{\delta}{2}}(\tfrac{\delta}{2}\eta), \tag{5.10}$$

where the domain $(0, T_\eta)$ is chosen to be maximal and possibly $T_\eta = \infty$. Therefore, we define Θ_0 as the set of finite coordinate lines and Θ_∞ the set of infinite coordinate lines

$$\Theta_0 = \{\eta : T_\eta < \infty\} \qquad \text{and} \qquad \Theta_\infty = \{\eta : T_\eta = \infty\}.$$

Remark 5.8. Note that the coordinates ψ_η are well-defined by the property (**$\tilde{\mathbf{H}}$2**). Indeed, the coordinates $\psi_\eta(r)$ satisfy for $r \in (0, \frac{\delta}{2})$ the equation

$$\dot{\psi}_\eta(r) = \frac{\nabla \tilde{H}(\psi_\eta(r))}{|\nabla \tilde{H}(\psi_\eta(r))|} = \eta. \tag{5.11}$$

Therefore, we find $\Psi_{r-\frac{\delta}{2}}(\frac{\delta}{2}\eta) = r\eta$ for $r \in (0, \frac{\delta}{2})$. This shows that $\psi_\eta(r)$ are polar coordinates around the local minimum.

Now, we have to show some basic properties following for the coordinates of Definition 5.7.

Proposition 5.9 (Last visit times)**.** *For an open bounded set $A \subset \Omega$ define the last visit time by*

$$\tau(A) := \sup_{x \in A} \sup \{t : \Psi_t(x) \in A\}. \tag{5.12}$$

Then it holds

$$\tau(A) \le c_{\tilde{H}}^{-1} \operatorname*{osc}_{x \in A} \tilde{H}(x).$$

Proof. Let us consider the solution of the gradient flow $\dot{\psi}(s) = \tilde{F}(\psi(s))$ with $\psi(0) = x \in A$. We get for the difference

$$\tilde{H}(\psi(T)) - \tilde{H}(\psi(0)) = \int_0^T \partial_s \tilde{H}(\psi(s)) \mathrm{d}s \overset{(5.9)}{=} \int_0^T |\nabla \tilde{H}(\psi(s))| \mathrm{d}s \ge c_{\tilde{H}} T.$$

Hence we have for every $T < \tau(A)$ the bound

$$T \le c_{\tilde{H}}^{-1} \left(\tilde{H}(\psi(T)) - \tilde{H}(\psi(0)) \right) \le c_{\tilde{H}}^{-1} \operatorname*{osc}_{x \in A} \tilde{H}.$$

□

Proposition 5.10. *The gradient flow lines in the family Θ_0 are uniform bounded in space and time, i.e.*

$$\exists R^* : \sup_{\eta \in \Theta_0} \sup_{t \in (0, T_\eta)} |\psi_\eta(t)| \le R^* < \infty \qquad \text{and} \qquad \exists T^* : \sup_{\eta \in \Theta_0} T_\eta \le T^* < \infty. \tag{5.13}$$

Proof. Let us fix some $\eta \in \Theta_0$. Then for some $s > 0$ holds $\{\psi_\eta(s)\}_{s\in(T_\eta - s, T_\eta)} \subset B_\delta(y)$ for some critical point $y \in \mathcal{S}$ of H, because else it would be possible to prolong ψ_η for $t \geq T_\eta$, which contradicts the maximality of T_η. Hence in particular $\{\psi_\eta(s)\}_{s\in(0,T_\eta)} \subset \{H \leq l^*\}$ with $l^* = \sup\{H(x) : x \in \bigcup_{y\in\mathcal{S}} B_\delta(y)\}$. By the growth assumption (**A1**$_{\mathrm{PI}}$) holds that $\{H \leq l^*\}$ is bounded and R^* can be chosen such that $B_{R^*} \supset \{H \leq l^*\}$. In particular $\sup\{T_\eta : \eta \in \Theta_0\} \leq \tau(\{H \leq l^*\}) =: T^* < \infty$ by (5.12). □

Now, we choose $R \geq R^*$ with R^* given by (5.13), yielding in particular that $\tilde{H}(x) = H(x)$, such that by assumption (**A2′**$_{\mathrm{PI}}$) holds

$$\forall |x| \geq R : \qquad |\nabla H| - \nabla \cdot \frac{H(x)}{|\nabla H(x)|} \geq -2K_H \quad \text{and} \quad |\nabla H| \geq \tfrac{C_H}{2}. \tag{5.14}$$

Therewith, we define $\Omega_R := \Omega \cap B_R(0)$ and with $\tau_\eta(\Omega_R) := \sup\{t : \psi_\eta(t) \in \Omega_R\}$ the last visit time of the gradient flow line ψ_η in Ω_R holds

$$\tilde{T}_\eta := \min\{T_\eta, \tau_\eta(\Omega_R)\} = \begin{cases} T_\eta & , \eta \in \Theta_0 \\ \tau_\eta(\Omega_R) & , \eta \in \Theta_\infty. \end{cases} \tag{5.15}$$

By definition holds that

$$\tilde{T}_\eta \leq \tau(\Omega_R) =: \tilde{T}^*.$$

Further note, that if Ω is bounded, then $\tilde{T}_\eta = T_\eta$ for all $\eta \in S^{n-1}$.

As we have explained above, it is very important to have a good understanding of the Hamiltonian

$$H_\eta(r) = \frac{1}{\varepsilon} H(\psi_\eta(r)) - \log j_\eta(r)$$

of the conditional measure $\mu(\mathrm{d}r|\eta)$. In the next statement, we consider the evolution of the second term of the last identity, namely the Jacobian determinant $j_\eta(r)$.

Lemma 5.11 (Evolution of the Jacobian determinant)**.** *For $x \in \Omega\setminus\{0\}$ let $D\Psi_r(x)$ be the Jacobian of the flow Ψ_r given by* (5.9)*. Then the Jacobian determinant*

$$j_\eta(r) := \det D\Psi_{r-\frac{\delta}{2}}(\tfrac{\delta}{2}\eta)$$

evaluated along the gradient-flow lines $\psi_\eta(r)$ (cf. (5.10)*) satisfies the evolution*

$$\forall r \in (0, T_\eta) : \qquad \frac{\mathrm{d}}{\mathrm{d}r} \log j_\eta(r) = \nabla \cdot \tilde{F}(\psi_\eta(r)). \tag{5.16}$$

Proof. Differentiating the identity (5.9) leads to

$$D\dot{\Psi}_r(x) = D\tilde{F}(\Psi_r(x))\, D\Psi_r(x).$$

As $D\Psi_r$ is regular, we can multiply by $(D\Psi_r)^{-1}(x)$ and take the trace

$$\operatorname{tr}\left(D\dot{\Psi}_r(x)(D\Psi_r)^{-1}(x)\right) = \operatorname{tr} D\tilde{F}(\Psi_r(x)) = \nabla \cdot \tilde{F}(\Psi_r(x)).$$

By Jacobi's formula (B.3) the left-hand side of the last identity can be written as

$$\frac{\mathrm{d}}{\mathrm{d}r} \log \det D\Psi_r(x) = \operatorname{tr}\left(D\dot{\Psi}_r(x) D\Psi_r^{-1}(x)\right),$$

which yields the desired formula (5.16). □

With the help of the last statement, we are able to formulate a condition (cf. (**H̃4**) below) on the perturbed Hamiltonian $\tilde{H}$, which ensures that the Jacobian determinant $j_\eta(r)$ behaves in the right way when applying the Muckenhoupt functional later (cf. Proposition 5.3).

Lemma 5.12 (Bounds on the Jacobian determinant)**.** *Assume that the perturbed Hamiltonian $\tilde{H}$ satisfies the Assumption 5.6 and that the local coordinates $\psi_\eta(r)$ are given by Definition 5.7. Additionally, assume that $\tilde{H}$ satisfies the condition*

$$\forall r \in (\tfrac{\delta}{2}, \tilde{T}_\eta): \ \nabla \cdot \tilde{F}(\psi_\eta(r)) \ \leq C_{\tilde{H}} \qquad (\tilde{\mathbf{H}}\mathbf{4})$$

for some constant $C_{\tilde{H}} > 0$. Then the Jacobian determinant $j_\eta(r)$ satisfies

$$\forall r \in (0, \tfrac{\delta}{2}] : j_\eta(r) = r^{n-1}, \qquad (\mathbf{j1})$$

$$\forall r \in (\tfrac{\delta}{2}, \tilde{T}_\eta) : \log j_\eta(r) = \log j_\eta^+(r) - \log j_\eta^-(r), \ \textit{with} \ \log \frac{j_\eta^+(r)}{j_\eta(\delta/2)} \leq T^* C_{\tilde{H}} \qquad (\mathbf{j2})$$

$$\textit{and } j_\eta^+, j_\eta^- \textit{ monotone increasing functions in } (0, \tilde{T}_\eta)$$

Proof of Lemma 5.12. The property (**j1**) is a direct consequence of the fact that the coordinates $\psi_\eta(r)$ are spherical coordinates for $r < \frac{\delta}{2}$ by Remark 5.8.
Let us consider the property (**j2**). By the fundamental theorem of calculus and the identity (5.16) we obtain for $r \in (\frac{\delta}{2}, \tilde{T}_\eta)$

$$\begin{aligned}
\log j_\eta(r) &= \log j_\eta(\tfrac{\delta}{2}) + \int_{\frac{\delta}{2}}^{r} \nabla \cdot \tilde{F}(\psi_\eta(s)) \, \mathrm{d}s \\
&= \log \left(j_\eta(\tfrac{\delta}{2}) e^{\int_{\delta/2}^{r} |\nabla \cdot \tilde{F}(\psi_\eta(s))|_+ \mathrm{d}s} \right) - \log \left(e^{\int_{\delta/2}^{r} |\nabla \cdot \tilde{F}(\psi_\eta(s))|_- \mathrm{d}s} \right) \\
&=: \log j_\eta^+(r) - \log j_\eta^-(r),
\end{aligned}$$

where $|x|_+ = \max\{0, x\}$ is the positive part and $|x|_- = \max\{0, -x\}$ the negative part of x. For $r \in (0, \frac{\delta}{2})$ we can set $j_\eta^+(r) := j_\eta(r)$ and $j_\eta^-(r) := 1$. Then, the monotonicity properties of $j_\eta^\pm(r)$ are obvious from their definitions. Furthermore, with the help of the bound (**H̃4**) and the uniform upper bound $T_\eta \leq T^*$ from (5.13) holds

$$\log \frac{j_\eta^+(r)}{j_\eta(\frac{\delta}{2})} \leq T^* C_{\tilde{H}}.$$

□

We still need to understand the first term in the Hamiltonian H_η given by (5.8), namely $H(\psi_\eta(r))$. The growth estimates (**h1**), (**h2**) and (**h3**) from below are one of the main ingredients to apply of the Muckenhoupt functional in the proof of Proposition 5.3 (see Section 5.3).

Lemma 5.13 (Properties of the radial Hamiltonian)**.** *Assume that the perturbed Hamiltonian $\tilde{H}$ satisfies the Assumption 5.6 and that the local coordinates $\psi_\eta(r)$ are given by Definition 5.7. Additionally, assume that $\tilde{H}$ satisfies the condition*

$$\langle \nabla H(\psi_\eta(r)), \nabla \tilde{H}(\psi_\eta(r))\rangle \geq c_{\tilde{H}} \min\{1, r, T_\eta - r\} |\nabla \tilde{H}(\psi_\eta(r))| \qquad (\tilde{\mathbf{H}}\mathbf{5})$$

for some constant $0 < c_{\tilde{H}}$. If $T_\eta = \infty$, we mean $\min\{1, r, T_\eta - r\} = \min\{1, r\}$.
For $\eta \in S^{n-1}$ we define the radial Hamiltonian function $h_\eta : [0,\infty) \to \mathbb{R} \cup \{\infty\}$ by

$$h_\eta(r) := \begin{cases} H(\psi_\eta(r)), & \text{for } 0 \leq r \leq T_\eta \\ \infty, & \text{for } T_\eta \leq r. \end{cases}$$

Then, the following estimates hold uniformly in η:

$$\exists \lambda_1 > 0\, \forall s \in [0, \tilde{T}_\eta) : H(\psi_\eta(r)) \geq H(\psi_\eta(s)) + \lambda_1 (r-s)^2, \quad \forall r \in [s, \tilde{T}_\eta] \qquad (\mathbf{h1})$$

$$\exists \lambda_2 > 0\, \forall s \in (0, \tilde{T}_\eta] : H(\psi_\eta(r)) \leq H(\psi_\eta(sr)) - \lambda_2 (r-s)^2, \quad \forall r \in [0, s]. \qquad (\mathbf{h2})$$

Further, if $\eta \in \Theta_\infty$, then we have for $\varepsilon \leq \frac{1}{2}\frac{C_H}{2C_H + 4K_H}$

$$\forall t \in (\tilde{T}_\eta, \infty) : \frac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{\varepsilon} h_\eta(t) - \ln j_\eta(t)\right) \geq \frac{C_H}{4}\frac{1}{\varepsilon} \qquad \text{for } \varepsilon \leq \frac{1}{2}\frac{C_H}{C_H + 4K_H}, \qquad (\mathbf{h3})$$

where C_H and K_H are from assumptions ($\mathbf{A1}_{\mathtt{PI}}$) and ($\mathbf{A2'}_{\mathtt{PI}}$).

Proof of Lemma 5.13. For $r \leq \tilde{T}_\eta$ it follows from Definition 5.7 and ($\tilde{\mathbf{H}}\mathbf{5}$) that

$$\frac{\mathrm{d}}{\mathrm{d}r} H(\psi_\eta(r)) = \left\langle \nabla H(\psi_\eta(r)), \frac{\nabla \tilde{H}(\psi_\eta(r))}{|\nabla \tilde{H}(\psi_\eta(r))|}\right\rangle \geq c_{\tilde{H}} \min\{1, r, T_\eta - r\}.$$

Hence on $(0, \tilde{T}_\eta)$, the radial Hamiltonian $h_\eta(r)$ is strictly increasing. Additionally, we have with (5.13) that $T_\eta \leq T^* < \infty$ holds uniformly in η. Therefore, it is easy to deduce (**h1**) and (**h2**) with constants λ_1 and λ_2 uniformly in η.

For the last property (**h3**) let $\eta \in \Theta_\infty$. Then for $t \geq \tilde{T}_h$ it follows by (5.15) that $\psi_\eta(t) \notin \Omega_R \supset L(l^*)$ and therefore $\tilde{H}(\psi_\eta(t)) = H(\psi_\eta(t))$. Hence, we can use the estimates (5.14) from which we conclude

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{1}{\varepsilon} h_\eta(r) - \ln j_\eta(r)\right) &= \frac{1}{\varepsilon}\nabla H(\psi_\eta(r)) \cdot \frac{\nabla \tilde{H}(\psi_\eta(r))}{|\nabla \tilde{H}(\psi_\eta(r))|} - \nabla \cdot \frac{\nabla \tilde{H}(\psi_\eta(r))}{|\nabla \tilde{H}(\psi_\eta(r))|} \\ &= \left(\frac{1}{\varepsilon} - 1\right) |\nabla H(\psi_\eta(r))| + \left(|\nabla H(\psi_\eta(r))| - \nabla \cdot \frac{\nabla H(\psi_\eta(r))}{|\nabla H(\psi_\eta(r))|}\right) \\ &\overset{(5.14)}{\geq} \left(\frac{1}{\varepsilon} - 1\right)\frac{C_H}{2} - 2K_H = \frac{1}{\varepsilon}\frac{C_H}{4}. \end{aligned}$$

We choose ε small enough such that

$$\frac{1}{\varepsilon}\frac{C_H}{4} - \frac{C_H}{2} - 2K_H \geq 0 \quad \text{hence} \quad \varepsilon \leq \frac{1}{2}\frac{C_H}{C_H + 4K_H}.$$

□

Now, we have provided all the ingredients needed for the proof of Proposition 5.3 in Section 5.3. For convenience, we make the following definition.

Definition 5.14 (Radial-mixing measures)**.** Assume that the perturbed Hamiltonian $\tilde{H}$ satisfies (**H̃1**), (**H̃2**) and (**H̃3**) of Assumption 5.6 and that the local coordinates $\psi_\eta(r)$ are given by Definition 5.7. Additionally, assume that $\tilde{H}$ satisfies the conditions (**H̃4**) and (**H̃5**). Then, the conditional measures

$$\mu(\mathrm{d}r|\eta) = \frac{1}{\hat{\mu}(\eta)} j_\eta(r) \exp\left(-\frac{1}{\varepsilon} H(\psi_\eta(r))\right) \mathrm{d}r$$

are called *radial-mixing measures*.

The existence of radial-mixing measures is stated in Proposition 5.2, which is verified in the next section. The term *mixing* in the definition of radial-mixing measures is justified by Proposition 5.3, which proofs that the Poincaré constant $\varrho(\eta)$ of $\mu(\mathrm{d}r|\eta)$ satisfies $\varrho^{-1}(\eta) = O(\varepsilon)$.

5.2. Construction of the perturbed Hamiltonian $\tilde{H}$

Let us recall, that we only consider one basin of attraction Ω_i (cf. (2.6)) and therefore omit the index i. We assume w.l.o.g. that $0 \in \Omega$ is the local minimum of H in Ω. In this section we state the proof of Proposition 5.2, which is restated at this point for the convenience of the reader.

Proposition 5.2 (Existence of mixing coordinates)**.** *Assume that H satisfies the Assumption 5.1. Then, there exist coordinates $\{\psi_\eta\}_{\eta \in S^{n-1}}$ on Ω such that the conditional measures $\mu(\mathrm{d}r|\eta)$ are radial-mixing measures in the sense of Definition 5.14.*

By Definition 5.14, we have to show that on a basin of attraction Ω (cf. (2.6)) there exists a perturbation $\tilde{H}$ of the Hamiltonian H satisfying the conditions (**H̃1**)-(**H̃5**). Because of (**H̃1**), the perturbation $\tilde{H}$ is only allowed to differ from the Hamiltonian H on small neighborhoods around the critical points $z \in \Omega$ of H. Therefore, we obtain the perturbation $\tilde{H}$ by a local construction on a neighborhood of every critical point. Hence the properties (**H̃1**) and (**H̃2**) will be automatically satisfied by construction. If one stays away from critical points $z \in \Omega$ of H, the properties (**H̃3**) and (**H̃5**) are automatically satisfied if we take there $\tilde{H} = H$. Furthermore, the property (**H̃4**) needs only to be checked on the bounded domain Ω_R, which is again clear as long as one stays away from critical points of H. Therefore, we only have to verify the properties (**H̃3**)-(**H̃5**) near the critical points. We distinguish two cases: Section 5.2.1 is devoted to the construction of $\tilde{H}$ at a local extremum of H and Section 5.2.2 considers the construction of $\tilde{H}$ at the saddle points of H.

5.2.1. Perturbed Hamiltonian $\tilde{H}$ at local extremes

In this section, we will construct the perturbed Hamiltonian $\tilde{H}$ satisfying the conditions (**H̃1**)-(**H̃5**) in a small neighborhood of the local minimum $0 \in \Omega$ of the Hamiltonian H. We omit the case where the critical point z is a local maxima. The reason is that the construction of $\tilde{H}$ would be very similar to the construction of $\tilde{H}$ around a local minimum by considering $-H$ instead of H.

We denote by $\lambda_{\min}$ the smallest eigenvalue of $\nabla^2 H(0)$, which is strictly positive by the non-degeneracy Assumption 1.9. Hence, we can choose $\delta > 0$ such that H is strictly convex on $B_\delta(0)$. Let $\xi \in C^\infty(\mathbb{R}^+, [0,1])$ be a decreasing smooth step function satisfying the conditions

$$\forall r \in [0, 1/2] : \xi(r) = 1, \quad \forall r \in (1/2, 1) : \xi'(r) < 0, \quad \text{and} \quad \forall r \in [1, \infty) : \xi(r) = 0. \tag{5.17}$$

For convenience, we set $\xi_\delta(r) = \xi(\frac{r}{\delta})$. Motivated by property (**H̃2**), we define the correction $\tilde{H}$ of H around the minimum as

$$\tilde{H}(x) := \xi_\delta(|x|)\,(|x| - \delta) + (1 - \xi_\delta(|x|))\,H(x), \qquad \text{for } x \in B_\delta(0). \tag{5.18}$$

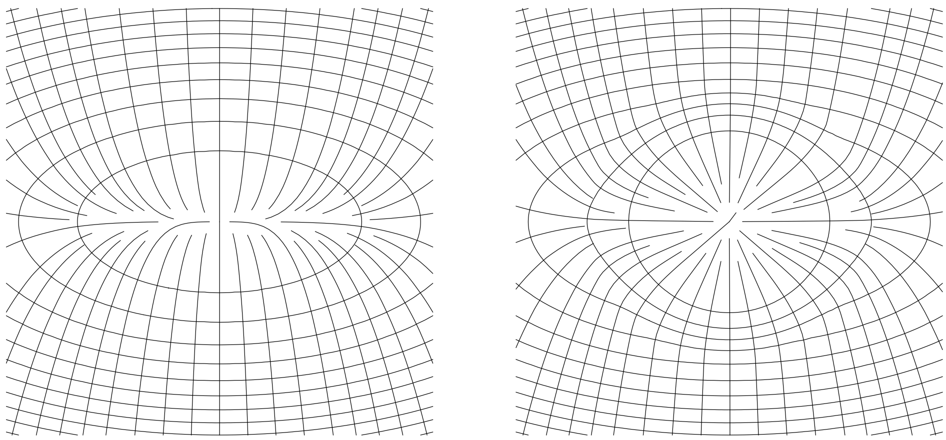

Figure 5.2.: Contours and flow lines derived from H (left) and $\tilde{H}$ (right).

Now, we state the main result of this section.

Lemma 5.15 (Perturbation $\tilde{H}$ around local extrema)**.** *For δ small enough, the function $\tilde{H}$ satisfies the properties* (**H̃1**)-(**H̃5**) *in the neighborhood $B_\delta(0)$ around the minimum $0 \in \Omega$.*

Proof. The property (**H̃1**) is fulfilled by construction.
The property (**H̃2**) is satisfied by the definition (5.18) of $\tilde{H}$ and the definition (5.17) of ξ, which ensures that the gradient-flow coordinates of $\tilde{H}$ coincide with polar coordinates on the ball $B_{\delta/2}(0)$.

The property (**H̃3**) is verified by a straightforward argument, where direct calculation leads to the identity

$$\nabla \tilde{H}(x) = a(x)\, \frac{x}{|x|} + b(x)\, \nabla H(x), \quad \text{with} \tag{5.19}$$
$$a(x) = \delta^{-1} \xi_\delta' \,(|x|)\; (|x| - \delta - H(x)) + \xi_\delta(|x|) \quad \text{and} \quad b(x) = 1 - \xi_\delta(|x|).$$

For the verification of (**H̃3**), we will need two observations.
The first one is $a(x) > \xi_\delta(|x|)$, which follows for small enough δ from the assumptions on H, namely $H(0) = 0$ and $H(x) \geq 0$ as well as the monotonicity (5.17) of ξ, i.e. $\xi' \leq 0$.
The second observation is that for $x \in B_\delta(0)$ holds

$$\left\langle \frac{x}{|x|}, \nabla H(x) \right\rangle = |x| \, \left\langle \frac{x}{|x|}, \nabla^2 H(0)\, \frac{x}{|x|} \right\rangle + O(x^2) \geq \lambda_{\min} |x| + O(\delta^2) \geq \frac{\lambda_{\min}}{2}\, |x| \,, \tag{5.20}$$

which follows directly from Taylor expansion and choosing δ sufficiently small. In the last inequality, $\lambda_{\min} > 0$ denotes the smallest eigenvalue of $\nabla^2 H(0)$.
Let us now verify (**H̃3**). The identity (5.19) directly yields for $x \in B_\delta(0)$

$$\begin{aligned} |\nabla \tilde{H}(x)| &= a^2(x) + 2\, a(x)\, b(x)\, \left\langle \frac{x}{|x|}, \nabla^2 H(x) \right\rangle + b^2(x)\, |\nabla H(x)|^2 \\ &\overset{(5.20)}{\geq} a^2(x) + 2\, a(x)\, b(x)\, \frac{\lambda_{\min}}{2} |x| + b^2(x)\, |\nabla H(x)|^2. \end{aligned}$$

We consider two cases. In the first case, we assume that $|x| < \frac{\delta}{4}$. Then we get

$$|\nabla \tilde{H}(x)| \geq a^2(x) + O(\delta^2) \overset{(5.17)}{\geq} \xi\,(1/4) + O(\delta^2) \geq c_{\tilde{H}} > 0.$$

In the second case, namely $\frac{\delta}{4} \leq |x| \leq \delta$, we get

$$\begin{aligned} |\nabla \tilde{H}(x)| &\geq b^2(x)\, |\nabla H(x)|^2 + O(\delta^2) \\ &\overset{(5.17)}{\geq} \left(1 - \xi\left(\tfrac{1}{4}\right)\right)^2 \min_{\frac{\delta}{4} \leq |x| \leq \delta} |\nabla H(x)|^2 + O(\delta^2) \geq c_{\tilde{H}} > 0, \end{aligned}$$

which yields the desired property (**H̃3**).

Now, let us turn to the property (**H̃4**). It suffices to show that for $x \in B_\delta(0) \backslash B_{\delta/2}(0)$ holds

$$\left| \nabla \cdot \frac{\nabla \tilde{H}}{|\nabla \tilde{H}|}(x) \right| \leq C.$$

However, this easily follows from the observation that

$$\begin{aligned} \left| \nabla \cdot \frac{\nabla \tilde{H}}{|\nabla \tilde{H}|}(x) \right| &= \frac{1}{|\nabla \tilde{H}(x)|}\, \left| \Delta \tilde{H}(x) - \left\langle \frac{\nabla \tilde{H}}{|\nabla \tilde{H}|}(x), \nabla^2 \tilde{H}(x)\, \frac{\nabla \tilde{H}}{|\nabla \tilde{H}|}(x) \right\rangle \right| \\ &\overset{(\tilde{\mathbf{H}}3)}{\leq} \frac{n-1}{c}\, \|\nabla^2 \tilde{H}(x)\| \end{aligned}$$

and the fact that $\tilde{H}(x)$ is by construction a smooth function on $B_\delta(0)\backslash B_{\delta/2}(0)$.

Finally, let us verify the property (**H̃5**). It is sufficient to show that

$$\forall t \text{ such that } \psi_\eta(t) \in B_\delta(0): \quad \left\langle \nabla H(\psi_\eta(t)), \nabla \tilde{H}(\psi_\eta(t)) \right\rangle \geq c_{\tilde{H}}\, t\, |\nabla \tilde{H}(\psi_\eta(t))|,$$

where the local coordinates $\psi_\eta(t)$ are given by (5.11). Using (5.19) we get

$$\begin{aligned} \left\langle \nabla H(x), \nabla \tilde{H}(x) \right\rangle &= a(x) \left\langle \nabla H(x), \frac{x}{|x|} \right\rangle + b(x)\, |\nabla H(x)|^2 \\ &\geq a(x) \left\langle \nabla H(x), \frac{x}{|x|} \right\rangle + b(x)\, |\nabla H(x)| \left\langle \nabla H(x), \frac{x}{|x|} \right\rangle \\ &\overset{(5.20)}{\geq} (a(x) + b(x)\, |\nabla H(x)|)\, \frac{\lambda_{\min}}{2} |x| \\ &\overset{(5.19)}{\geq} \frac{\lambda_{\min}}{2}\, |x|\, |\nabla \tilde{H}(x)|. \end{aligned}$$

By the last inequality, it is only left to show that for some constant $c > 0$

$$|\psi_\eta(t)| \geq c\, t, \qquad \text{for all } \psi_\eta(t) \in B_\delta(0). \tag{5.21}$$

It follows from the definition (5.11) of $\psi_\eta(t)$ that

$$|\psi_\eta(t)| = \int_0^t \partial_s |\psi_\eta(s)|\, ds = \int_0^t \left\langle \frac{\psi_\eta(s)}{|\psi_\eta(s)|}, \frac{\nabla \tilde{H}(\psi_\eta(s))}{|\nabla \tilde{H}(\psi_\eta(s))|} \right\rangle ds. \tag{5.22}$$

For $\psi_\eta(s) \in B_\delta(0)$ it follows from (5.19) that $|\nabla \tilde{H}(\psi_\eta(s))| \leq C$. Therefore, we can estimate

$$\begin{aligned} &\left\langle \frac{\psi_\eta(s)}{|\psi_\eta(s)|}, \frac{\nabla \tilde{H}(\psi_\eta(s))}{|\nabla \tilde{H}(\psi_\eta(s))|} \right\rangle \\ &\quad \overset{(5.19)}{=} \frac{1}{|\nabla \tilde{H}(\psi_\eta(s))|} \left(a(\psi_\eta(s)) + b(\psi_\eta(s)) \left\langle \frac{\psi_\eta(s)}{|\psi_\eta(s)|}, \nabla H(\psi_\eta(s)) \right\rangle \right) \\ &\quad \overset{(5.20)}{\geq} \frac{1}{C} \left(a(\psi_\eta(s)) + b(\psi_\eta(s))\, \frac{\lambda_{\min}}{2}\, |\psi_\eta(s)| \right) \\ &\quad \geq c_{\tilde{H}} > 0, \end{aligned} \tag{5.23}$$

where the last line is deduced with the same argument as used for (**H̃3**). A combination of (5.22) and (5.23) yields the desired estimate (5.21) and concludes the proof. □

5.2.2. Perturbed Hamiltonian $\tilde{H}$ at the saddles

In this section, we construct a local a perturbation $\tilde{H}$ of H satisfying the conditions (**H̃1**)-(**H̃5**) nearby saddle points. Note that every saddle s lies on the boundary of Ω. We will locally deform H in such a way that either trajectories will leave Ω nearby the saddle

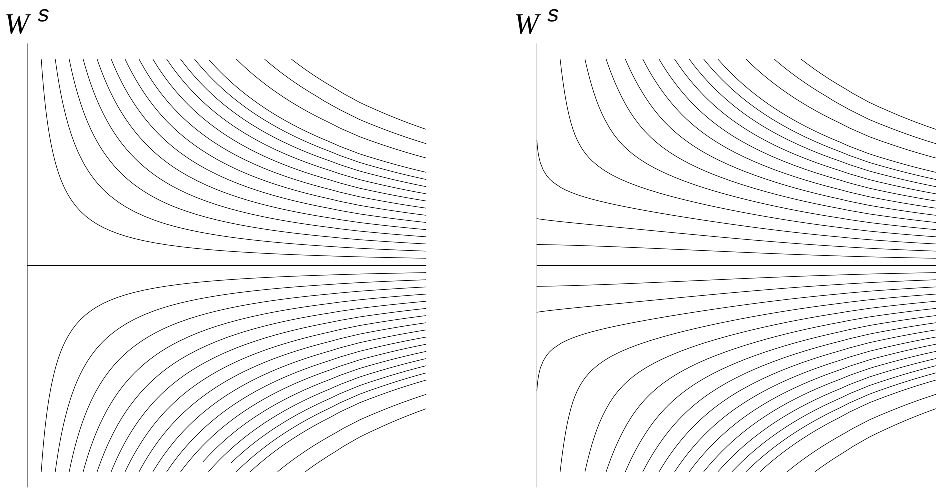

Figure 5.3.: Flow lines derived from H (left) and $\tilde{H}$ (right).

or they will avoid coming close to the saddle (cf. Figure 5.3). We achieve this behavior by introducing a kink on the stable manifold (cf. Figure 5.5).

We assume w.l.o.g. that the saddle point s lies at the origin 0. The stable and unstable manifold of the saddle at 0 are given by

$$W^{\mathrm{s}} := \{y_0 \in \mathbb{R}^n : \dot{y}_t = -\nabla H(y_t), y_t \to 0\}$$
$$W^{\mathrm{u}} := \{y_0 \in \mathbb{R}^n : \dot{y}_t = \nabla H(y_t), y_t \to 0\}.$$

In order to deform the trajectories as indicated in Figure 5.3, we need a better description of the stable and unstable manifold of the saddle at 0. This description is provided by the stable-manifold theorem (cf. Theorem 5.17 below).

Definition 5.16 (Stable $\mathcal{E}^s$ and unstable $\mathcal{E}^u$ subspaces of $\nabla^2 H(0)$)**.** By the non-degeneracy assumption (1.8) there is a number k and an orthogonal matrix Q such that

$$\nabla^2 H(0) = Q\,\mathrm{diag}(\lambda_1^-, \dots, \lambda_k^-, \lambda_1^+, \dots, \lambda_{n-k}^+)Q^\top,$$

where $\lambda_1^-, \dots, \lambda_k^- < 0$ and $\lambda_1^+, \dots, \lambda_{n-k}^+ > 0$. Therefore, the stable $\mathcal{E}^s$ and unstable $\mathcal{E}^u$ subspaces of $\nabla^2 H(0)$ are given by

$$\mathcal{E}^{\mathrm{u}} := \left\{Q(y,0)^\top \in \mathbb{R}^n : y \in \mathbb{R}^k\right\}$$
$$\mathcal{E}^{\mathrm{s}} := \left\{Q(0,y)^\top \in \mathbb{R}^n : y \in \mathbb{R}^{n-k}\right\}.$$

Theorem 5.17 (Stable-manifold theorem [Tes12, Theorem 9.4, p. 259])**.** *For a small neighborhood $\mathcal{U}$ of the saddle 0 the local stable-manifold $W^{\mathrm{s}}_{\mathrm{loc}} := W^{\mathrm{s}} \cap \mathcal{U}$ is a smooth manifold tangent to $\mathcal{E}^{\mathrm{s}}$ at 0. Moreover, there exist neighborhoods $\mathcal{U}^{\mathrm{s}} \subseteq \mathcal{E}^{\mathrm{s}}$, $\mathcal{U}^{\mathrm{u}} \subseteq \mathcal{E}^{\mathrm{u}}$ and a smooth map $\mathtt{h}_{\mathtt{s}} : \mathcal{U}^{\mathrm{s}} \to \mathcal{U}^{\mathrm{u}}$ satisfying*

$$\mathtt{h}_{\mathtt{s}}(0) = 0 \quad \textit{and} \quad D\mathtt{h}_{\mathtt{s}}(0) = 0 \tag{5.24}$$

such that for

$$\mathcal{U} := \{x_{\mathtt{s}} + \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}}) + x_{\mathtt{u}} :\ x_{\mathtt{s}} \in \mathcal{U}^{\mathrm{s}},\ x_{\mathtt{u}} \in \mathcal{U}^{\mathrm{u}}\}$$

the local stable-manifold $W^{\mathrm{s}}_{\mathrm{loc}} = W^{\mathrm{s}} \cap \mathcal{U}$ *allows the representation*

$$W^{\mathrm{s}}_{\mathrm{loc}} = \{x_{\mathtt{s}} + \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}}) : x_{\mathtt{s}} \in \mathcal{U}^{\mathrm{s}}\}\,.$$

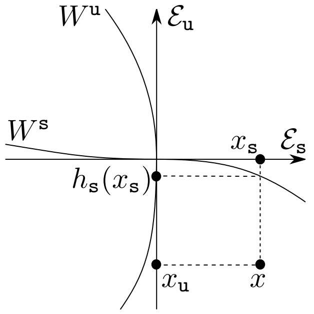

Figure 5.4.: Illustration of the stable-manifold theorem

Using the stable-manifold theorem we are able to introduce the following *local coordinates* on the saddle by the map (cf. Figure 5.4)

$$\xi(x_{\mathtt{s}}, x_{\mathtt{u}}) : \mathcal{U}^{\mathrm{s}} \times \mathcal{U}^{\mathrm{u}} \to \mathcal{U} \subset \mathbb{R}^n \qquad \xi(x_{\mathtt{s}}, x_{\mathtt{u}}) \mapsto x_{\mathtt{s}} + \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}}) + x_{\mathtt{u}}.$$

With the orthogonal projections $\Pi_{\mathtt{s}}$ and $\Pi_{\mathtt{u}}$ onto the stable linear $\mathcal{E}^{\mathrm{s}}$ and unstable linear $\mathcal{E}^{\mathrm{u}}$ manifolds

$$\Pi_{\mathtt{s}} : \mathcal{U} \to \mathcal{U}^{\mathrm{s}} \qquad \Pi_{\mathtt{s}} = Q\,\mathrm{diag}(\underbrace{0,\ldots,0}_{k\text{-times}},\ \underbrace{1\ldots,1}_{(n-k)\text{-times}}\ )Q^{\top}$$

$$\Pi_{\mathtt{u}} : \mathcal{U} \to \mathcal{U}^{\mathrm{u}} \qquad \Pi_{\mathtt{u}} = Q\,\mathrm{diag}(\underbrace{1,\ldots,1}_{k\text{-times}},\ \underbrace{0\ldots,0}_{(n-k)\text{-times}}\ )Q^{\top}$$

we obtain the reverse coordinate transformation by

$$\mathcal{U} \ni x \mapsto (\Pi_{\mathtt{s}} x, \Pi_{\mathtt{u}} x - \mathtt{h}_{\mathtt{s}}(\Pi_{\mathtt{s}} x)) \in \mathcal{U}^{\mathrm{s}} \times \mathcal{U}^{\mathrm{u}}.$$

We will construct the Hamiltonian $\tilde{H}$ by adding a perturbation to the Hamiltonian H. This additive perturbation is obtained with the help of the local coordinates from above and two auxiliary functions $p, q \in C^2(\mathbb{R}^+, [0,1])$ satisfying

$$p(0) = 1, \quad p'(0) = 0, \quad \forall\, t \in (0,1):\ p'(t) < 0, \tag{A$_p$}$$

$$\forall\, t \in [0,1]:\ \left|p'(t)\right| \le 8t, \tag{B$_p$}$$

$$\forall\, t \in [0,1]:\ p(t) \ge (1-2t)_+, \tag{C$_p$}$$

and

$$q(0) = 1, \quad \forall t \in (0,1):\ q'(t) < 0, \tag{A$_q$}$$

$$\forall\, t \in [0,1]:\ (1-2t)_+ \le \left|q'(t)\right| \le 2. \tag{B$_q$}$$

Example 5.18. By straightforward calculation one can show that the functions

$$p(t) = (1-t^2)^3\, \mathbb{1}_{[0,1)}(t) \qquad \text{and} \qquad q(t) = (1-t)(1-t^2)^2\, \mathbb{1}_{[0,1)}(t)$$

satisfy the conditions (A_p)-(C_p) and (A_q)-(B_q), respectively

The next statment contains the main result of this section.

Lemma 5.19 (Perturbation $\tilde{H}$ around saddles)**.** *For constants $a, \delta > 0$ we define the additive perturbation as*

$$H^+_{\mathrm{loc}} : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R} \qquad H^+_{\mathrm{loc}}(s,u) = a\delta^2 p\left(\frac{s}{\delta}\right) q\left(\frac{u}{\delta}\right),$$

where the functions p and q satisfy (A_p)-(C_p) *and* (A_q)-(B_q)*, respectively. Let us define the modified Hamiltonian $\tilde{H}$ on $\mathcal{U} := \{x \in \mathbb{R}^n : |\Pi_{\mathtt{s}} x| < \delta, |\Pi_{\mathtt{u}} x - h_{\mathtt{s}}(\Pi_{\mathtt{s}} x)| < \delta\}$ by*

$$\tilde{H}(x) := H(x) + H^+_{\mathrm{loc}}(|\Pi_{\mathtt{s}} x|, |\Pi_{\mathtt{u}} x - h_{\mathtt{s}}(\Pi_{\mathtt{s}} x)|). \tag{5.25}$$

Then, there exists $a > 0$ and $\delta > 0$ small enough such that $\tilde{H}$ on $\mathcal{U}$ satisfies the properties (**H̃1**)-(**H̃5**)*.*

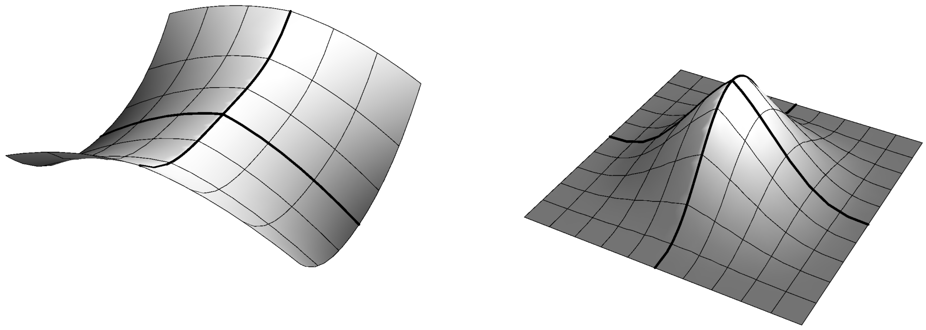

Figure 5.5.: $\tilde{H}$ (left) with a kink on the stable manifold and H^+_{loc} (right).

Proof. The properties (**H̃1**) and (**H̃2**) are fulfilled by construction.

Argument for (**H̃3**): We have to show that for $x \in \mathcal{U}$ holds

$$|\nabla \tilde{H}(x)| \geq c_{\tilde{H}} > 0. \tag{5.26}$$

Let $x = x_{\mathtt{s}} + x_{\mathtt{u}} \in \mathcal{U} \backslash W^{\mathtt{s}}_{\mathrm{loc}}$ with $x_{\mathtt{s}} = \Pi_{\mathtt{s}} x$ and $x_{\mathtt{u}} = \Pi_{\mathtt{u}} x$, hence $x_{\mathtt{u}} \neq \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})$. Further, set $s = |x_{\mathtt{s}}|$ and $u = |x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})|$. Then, differentiation of (5.25) using the equation (B.4) from the Appendix B.5 yields

$$\nabla \tilde{H}(x) = \nabla H(x) + \partial_s H^+_{\mathrm{loc}}(s,u) \frac{x_{\mathtt{s}}}{s} + \partial_u H^+_{\mathrm{loc}}(s,u) \left(\Pi_{\mathtt{u}} - D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\Pi_{\mathtt{s}}\right)^\top \frac{x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})}{u}. \tag{5.27}$$

Let us note that since $\Pi_{\mathtt{u}}$ is an orthogonal projection and since $D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}}) : \mathcal{E}^{\mathtt{s}} \to \mathcal{E}^{\mathtt{u}}$, it holds $D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\Pi_{\mathtt{s}} = D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})$ and $(\Pi_{\mathtt{u}} - D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\Pi_{\mathtt{s}})^{\top} = \Pi_{\mathtt{u}} - D\mathtt{h}_{\mathtt{s}}^{\top}(x_{\mathtt{s}})$. Let us expand $\nabla H(x)$ around the saddle point

$$\nabla H(x) = \nabla H(x_s + x_u) = \nabla^2 H(0) x_s + \nabla^2 H(0)(x_u - \mathtt{h}_{\mathtt{s}}(x_s)) + R_2(x), \tag{5.28}$$

where the remainder term $R_2(x)$ satisfies $|R_2(x)| \leq (C_H + C_h)\,|x|^2 \leq (C_H + C_h)\delta^2$ as $|\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})| \leq C_h\,|x_{\mathtt{s}}|^2$ by (5.24). We note that the partial derivatives of the additive perturbation are given by

$$\partial_s H_{\mathrm{loc}}^{+}(s,u) = a\delta\, p'\left(\frac{s}{\delta}\right) q\left(\frac{u}{\delta}\right) \quad \text{and} \quad \partial_u H(s,u) = a\delta\, p\left(\frac{s}{\delta}\right) q'\left(\frac{u}{\delta}\right).$$

Combining all expressions yields that the the gradient of the perturbed Hamiltonian can be written as

$$\nabla \tilde{H}(x) = N_{\mathtt{s}}(x) + N_{\mathtt{u}}(x) + R_2(x), \tag{5.29}$$

where the terms $N_{\mathtt{s}}(x)$ and $N_{\mathtt{u}}(x)$ are given by

$$N_{\mathtt{s}}(x) = \left(\nabla^2 H(0) s + a\delta\, p'\left(\frac{s}{\delta}\right) q\left(\frac{u}{\delta}\right) \mathrm{Id}\right) \frac{x_{\mathtt{s}}}{s} - a\delta\, p\left(\frac{s}{\delta}\right) q'\left(\frac{u}{\delta}\right) D\mathtt{h}_{\mathtt{s}}^{\top}(x_{\mathtt{s}}) \frac{x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})}{u}, \tag{5.30}$$

$$N_{\mathtt{u}}(x) = \left(\nabla^2 H(0) u + a\delta p\left(\frac{s}{\delta}\right) q'\left(\frac{u}{\delta}\right) \mathrm{Id}\right) \frac{x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})}{u}. \tag{5.31}$$

We want to point out that $\langle N_{\mathtt{s}}(x), N_{\mathtt{u}}(x)\rangle = 0$ because $N_{\mathtt{s}}(x) \in \mathcal{E}^{\mathtt{s}}$ and $N_{\mathtt{u}}(x) \in \mathcal{E}^{\mathtt{u}}$ ($\nabla^2 H(0)$ leaves $\mathcal{E}^s$ and $\mathcal{E}^u$ invariant). Therefore, we have

$$|\nabla \tilde{H}(x)|^2 = |N_{\mathtt{s}}(x)|^2 + |N_{\mathtt{u}}(x)|^2 + O(\delta^3) \geq \max\{|N_{\mathtt{s}}(x)|^2, |N_{\mathtt{u}}(x)|^2\} + O(\delta^3). \tag{5.32}$$

We proceed by determining lower bounds for $|N_{\mathtt{s}}(x)|$ and $|N_{\mathtt{u}}(x)|$. For $|N_{\mathtt{s}}(x)|$ we get

$$|N_{\mathtt{s}}(x)| \geq \lambda_{\min}^{+} s - a\delta\, \left|p'\left(\frac{s}{\delta}\right)\right| q\left(\frac{u}{\delta}\right) - a\delta\, p\left(\frac{s}{\delta}\right) \left|q'\left(\frac{u}{\delta}\right)\right|\, \|D\mathtt{h}_{\mathtt{s}}^{\top}(x_{\mathtt{s}})\|, \tag{5.33}$$

where $\lambda_{\min}^{+}$ denotes the smallest positive eigenvalue of the Hessian $\nabla^2 H(0)$. We employ now the assumptions (B_p) and (A_q) to bound the second term in (5.33) by

$$a\delta \left|p'\left(\frac{s}{\delta}\right)\right| q\left(\frac{u}{\delta}\right) \leq 8as. \tag{5.34}$$

Likewise, we use the assumptions (A_p) and (B_q) and the additional observation that $\left|D\mathtt{h}_{\mathtt{s}}^{\top}(x_{\mathtt{s}})\right| \leq C_h\,|x_{\mathtt{s}}|$ by (5.24) to bound the third term in (5.33) by

$$a\delta\, p\left(\frac{s}{\delta}\right) \left|q'\left(\frac{u}{\delta}\right)\right|\, |D\mathtt{h}_{\mathtt{s}}^{\top}(x_{\mathtt{s}})| \leq 2a\delta C_h s \leq 2as, \qquad \text{for } \delta \leq C_h^{-1}. \tag{5.35}$$

Hence, we can deduce a lower bound for $|N_{\mathtt{s}}(x)|$ by choosing a small enough. Indeed, we observe

$$|N_{\mathtt{s}}(x)| \geq \left(\lambda_{\min}^{+} - 8a - 2a\right) s \geq \delta\, \frac{\lambda_{\min}^{+}}{2}\, \frac{s}{\delta} \quad \text{for} \quad a \leq \frac{\lambda_{\min}^{+}}{20}. \tag{5.36}$$

Hence for $a \leq \frac{\lambda^+_{\min}}{20}$, a combination of (5.32) and (5.36) yields the desired estimate (5.26) if $s \geq \frac{\delta}{4}$.
However, if $s \leq \frac{\delta}{4}$, we have to take a closer look at the term $|N_{\mathtt{u}}(x)|$. Note that by definition of $\mathcal{U}^{\mathrm{u}}$, the Hessian $\nabla^2 H(0)$ is strictly negative definite on $\mathcal{U}^{\mathrm{u}}$. Let $\lambda^-_{\min} < 0$ denote by the smallest negative eigenvalue of $\nabla^2 H(0)$ on $\mathcal{U}^{\mathrm{u}}$ in modulus. Using the definition (5.31) of $N_{\mathtt{u}}(x)$< we can deduce the lower bound

$$|N_{\mathtt{u}}(x)| \geq |\lambda^-_{\min}|u + a\delta p\left(\frac{s}{\delta}\right)\left|q'\left(\frac{s}{\delta}\right)\right| \overset{(\mathrm{C}_p),(\mathrm{B}_q)}{\geq} |\lambda^-_{\min}|u + a\delta\left(1-2\frac{s}{\delta}\right)_+\left(1-2\frac{u}{\delta}\right)_+$$
$$\geq \min\left\{\frac{|\lambda^-_{\min}|\delta}{2}, a\delta\left(1-2\frac{s}{\delta}\right)_+\right\} \tag{5.37}$$

where the last step follows from either setting $u = 0$ or $u = \frac{\delta}{2}$. Hence, if $s \leq \frac{\delta}{4}$, the desired estimate (5.26) follows from a combination of (5.32) and (5.37).

Argument for ($\tilde{\mathbf{H}}\mathbf{4}$): Although property ($\tilde{\mathbf{H}}\mathbf{4}$) is formulated in a pathwise version, it is enough to show the pointwise estimate

$$\forall x \in \Omega_R \backslash B_{\delta/2}(0) : \nabla \cdot \frac{\nabla \tilde{H}(x)}{|\nabla \tilde{H}(x)|} \leq C_{\tilde{H}}, \tag{5.38}$$

where $\Omega_R = \Omega \cap B_R(0)$ and R is given in (5.14). Direct calculation shows that

$$\nabla \cdot \frac{\nabla \tilde{H}}{|\nabla \tilde{H}|} = \frac{1}{|\nabla \tilde{H}|}\left(\Delta \tilde{H} - \left\langle \frac{\nabla \tilde{H}}{|\nabla \tilde{H}|}, \nabla^2 \tilde{H}\frac{\nabla \tilde{H}}{|\nabla \tilde{H}|}\right\rangle\right).$$

We observe that if $\tilde{\lambda}_1(x) \leq \cdots \leq \tilde{\lambda}_n(x)$ denote the ordered eigenvalues of $\nabla^2 \tilde{H}(x)$, then the following bound holds

$$\Delta \tilde{H} - \left\langle \frac{\nabla \tilde{H}}{|\nabla \tilde{H}|}, \nabla^2 \tilde{H}\frac{\nabla \tilde{H}}{|\nabla \tilde{H}|}\right\rangle \leq \sum_{i=2}^n \tilde{\lambda}_i(x) \leq (n-1)\tilde{\lambda}_n(x). \tag{5.39}$$

For the estimation of $\tilde{\lambda}_n(x)$, we use the variational formulation of the eigenvalues

$$\tilde{\lambda}_n(x) = \sup_{\eta \in S^{n-1}} \left\langle \eta, \nabla^2 \tilde{H}(x)\eta \right\rangle. \tag{5.40}$$

Below, we will show that for some constant $C_{\tilde{H}} > 0$

$$\left\langle \eta, \nabla^2 \tilde{H}(x)\eta \right\rangle \leq C_{\tilde{H}}. \tag{5.41}$$

In combination with (5.39) and (5.40), this estimate already yields the desired statement (5.38). Now, we verify the estimate (5.41). For that purpose, let us have a closer look at $\nabla \tilde{H}$, which can be written as (cf. (5.27))

$$\nabla \tilde{H}(x) = \underbrace{\nabla H(x) + \partial_s H^+_{\mathrm{loc}}(s,u)\frac{x_{\mathtt{s}}}{s}}_{=:T_1} + \underbrace{\partial_u H^+_{\mathrm{loc}}(s,u)\left(\Pi_{\mathtt{u}} - D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\right)^\top \frac{x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})}{u}}_{=:T_2}. \tag{5.42}$$

Hence, we can write $\nabla^2 \tilde{H}$ as

$$\nabla^2 \tilde{H} = DT_1 + DT_2. \tag{5.43}$$

For the moment, we skip the details for the estimation of DT_1 and first consider the term DT_2. Straightforward calculation shows that

$$DT_2 = D\left(\partial_u H^+_{\mathrm{loc}}(s,u)\left(\Pi_{\mathtt{u}} - D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\right)^\top \frac{x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})}{u}\right) = T_3 + T_4,$$

where the terms T_3 and T_4 are given by

$$(T_3)_{i,j} = \sum_{k=1}^{n} \partial_{x_i}\left(\partial_u H^+_{\mathrm{loc}}(s,u)\left(\Pi_{\mathtt{u}} - D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\right)^\top\right)_{jk}\left(\frac{x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})}{u}\right)_k$$

and

$$T_4 = \partial_u H^+_{\mathrm{loc}}(s,u)\left(\Pi_{\mathtt{u}} - D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\right)^\top D\left(\frac{x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})}{u}\right).$$

Using the assumptions (A$_p$), (B$_p$), (A$_q$) and (B$_q$) as well as the properties of $\mathtt{h}_{\mathtt{s}}$ given by (5.24), one can deduce that the term T_3 satisfies for $\eta \in S^{n-1}$ the estimate

$$|\langle \eta, T_3 \eta\rangle| \leq C. \tag{5.44}$$

The same procedure applies also to the term DT_1, which can be similarly estimated like T_3

$$|\langle \eta, DT_1 \eta\rangle| \leq C. \tag{5.45}$$

Let us now consider the term T_4. With the help of the matrix $A = \Pi_{\mathtt{u}} - D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})$ and the vector $v = \frac{x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})}{u} \in \mathcal{E}^{\mathtt{u}} \cap S^{n-1}$ we find using the equation (B.5) from Appendix B.5 the relation

$$Dv = \frac{1}{u}\left(\mathrm{Id} - v \otimes v\right) A = \frac{1}{u}\left(\Pi_{\mathtt{u}} - v \otimes v\right) A,$$

since $A : \mathbb{R}^n \to \mathcal{E}^{\mathtt{u}}$. Therewith, we can rewrite the term T_4 as

$$T_4 = \frac{\partial_u H^+_{\mathrm{loc}}(s,u)}{u} A^\top \left(\Pi_{\mathtt{u}} - v \otimes v\right) A.$$

Because $\|D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\| = O(s) = O(\delta)$ and $D\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}}) : \mathcal{E}^{\mathtt{s}} \to \mathcal{E}^{\mathtt{u}}$ (cf. (5.24)), we have that $A > 0$ on $\mathcal{E}^{\mathtt{u}}$ and $A = 0$ on $\mathcal{E}^{\mathtt{s}}$, which yields

$$A^\top (\Pi_u - v \otimes v) A \geq 0.$$

Finally, because $\frac{\partial_u H^+_{\mathrm{loc}}(s,u)}{u} \leq 0$, this means that T_4 is negative definite, i.e.

$$\langle \eta, T_4 \eta \rangle \leq 0. \tag{5.46}$$

A combination of (5.43), (5.45), (5.44) and (5.46) yields the desired statement (5.41).

Argument for (**$\tilde{\mathbf{H}}$5**): We conclude in several steps. Firstly, we deduce a pointwise estimate

$$\langle \nabla H(x), \nabla \tilde{H}(x)\rangle \geq \frac{(\lambda^+_{\min})^2}{2} s^2 + |\lambda^-_{\min}| a\delta \left(1 - \frac{2u}{\delta}\right)_+ p_\delta(s) u + \left(\lambda^-_{\min}\right)^2 u^2 + O(s^2\delta) + O(u^2\delta). \tag{5.47}$$

Then, in the second step, we get rid of the error terms, as they can be compensated by the first and second term of (5.47), respectively. This leads to the estimate

$$\langle \nabla H(x), \nabla \tilde{H}(x)\rangle \geq \frac{(\lambda^+_{\mathrm{min}})^2}{4} s^2 + |\lambda^-_{\mathrm{min}}| a\delta \left(1 - \frac{2u}{\delta}\right)_+ p_\delta(s) u + \frac{\left(\lambda^-_{\mathrm{min}}\right)^2}{2} u^2. \tag{5.48}$$

If we for the moment assume that (5.47) respectively (5.48) holds, we still need to recombine this estimate (5.48) with the trajectories. Below, we will deduce two differential inequalities for $u(t)$ and $s(t)$ given by

$$s(t) := |\Pi_{\mathtt{s}} \psi_\eta(t)| \qquad \text{and} \quad u(t) := |\Pi_{\mathtt{u}} \psi_\eta(t) - \mathtt{h}_{\mathtt{s}} \left(\Pi_{\mathtt{s}} \psi_\eta(t)\right)| .$$

Namely, we will show that for δ sufficiently small and a satisfying (5.36) holds

$$1 \geq \dot{s}(t) \geq \frac{\lambda^+_{\mathrm{min}}}{2} s(t) + O(\delta^2), \tag{5.49}$$

$$1 + O(\delta^2) \geq -\dot{u}(t) \geq a(\delta - 2u(t))_+ p_\delta(s(t)) + O(\delta^2). \tag{5.50}$$

The deduction of (5.49) and (5.50) is similar to the estimate of $N_{\mathtt{s}}$ and $N_{\mathtt{u}}$. By direct calculation and term-wise estimation of $|\nabla \tilde{H}|$ in (5.42), we find that $|\nabla \tilde{H}| = O(\delta)$ in $\mathcal{U}$. Hence we can assume from now that $|\nabla \tilde{H}| \leq 1$. The starting point for all estimates is the representation of $\nabla \tilde{H}$ given in (5.27) and its decomposition into $N_{\mathtt{s}}, N_{\mathtt{u}}$ and the error R_2 from (5.29)

$$\begin{aligned}
\dot{s}(t) &= \frac{1}{s(t)} \left\langle \Pi_{\mathtt{s}} \psi_\eta(t), \dot{\psi}_\eta(t) \right\rangle = \frac{1}{s(t) |\nabla \tilde{H}(\psi_\eta(t))|} \left\langle \Pi_{\mathtt{s}} \psi_\eta(t), \nabla \tilde{H}(\psi_\eta(t)) \right\rangle \\
&= \frac{1}{s(t) |\nabla \tilde{H}(\psi_\eta(t))|} \left(\langle \Pi_{\mathtt{s}} \psi_\eta(t), N_{\mathtt{s}}(\psi_\eta(t)) \rangle + \langle \Pi_{\mathtt{s}} \psi_\eta(t), R_2(x) \rangle \right) \\
&\geq \frac{1}{s(t)} \langle \Pi_{\mathtt{s}} \psi_\eta(t), N_{\mathtt{s}}(\psi_\eta(t)) \rangle + O(\delta^2).
\end{aligned}$$

We use the representation of $N_{\mathtt{s}}$ from (5.30) and follow along the lines of (5.33) and (5.36) to estimate $\langle \psi_\eta(t), N_s(\psi_\eta(t))\rangle$

$$\begin{aligned}
\left\langle \frac{\Pi_{\mathtt{s}} \psi_\eta(t)}{s(t)}, N_s(\psi_\eta(t)) \right\rangle &= \left\langle \frac{\Pi_{\mathtt{s}} \psi_\eta(t)}{s(t)}, \nabla^2 H(0) \psi_\eta(t) \right\rangle + \partial_s H^+_{\mathrm{loc}}(s(t), u(t)) \frac{|\Pi_{\mathtt{s}} \psi_\eta(t)|^2}{s(t)^2} \\
&\quad - \partial_u H^+_{\mathrm{loc}}(s, u) \left\langle \frac{\Pi_{\mathtt{s}} \psi_\eta(t)}{s(t)}, D\mathtt{h}_{\mathtt{s}}^\top (\Pi_{\mathtt{s}} \psi_\eta(t)) \frac{\Pi_{\mathtt{u}} \psi_\eta(t) - \mathtt{h}_{\mathtt{s}}(\Pi_{\mathtt{s}} \psi_\eta(t))}{u(t)} \right\rangle \\
&\overset{(5.33)}{\geq} \lambda^+_{\mathrm{min}} s(t) - a\delta \left| p'\left(\frac{s(t)}{\delta}\right) \right| q\left(\frac{u(t)}{\delta}\right) - a\delta p\left(\frac{s(t)}{\delta}\right) q'\left(\frac{u(t)}{\delta}\right) \left| D\mathtt{h}_{\mathtt{s}}^\top (\Pi_{\mathtt{s}} \psi_\eta(t)) \right| \\
&\overset{(5.34),(5.35)}{\geq} \lambda^+_{\mathrm{min}} s(t) - 8as(t) - 2aC_h s(t).
\end{aligned}$$

Using the same choice of a as in (5.36) shows (5.49). For the differential inequality (5.50) we have to take care of the sign of $\dot{u}(t)$

$$\begin{aligned}
-\dot{u}(t) &= -\frac{\mathrm{d}}{\mathrm{d}t}\left|\Pi_{\mathtt{u}}\psi_\eta(t) - \mathtt{h}_{\mathtt{s}}(\Pi_{\mathtt{s}}\psi_\eta(t))\right| \\
&= -\left\langle \frac{\Pi_{\mathtt{u}}\psi_\eta(t) - \mathtt{h}_{\mathtt{s}}(\Pi_{\mathtt{s}}\psi_\eta(t))}{u(t)}, \left(\Pi_{\mathtt{u}} - D\mathtt{h}_{\mathtt{s}}(\Pi_{\mathtt{s}}\psi_\eta(t))\,\Pi_{\mathtt{s}}\right)\dot{\psi}_\eta(t)\right\rangle \\
&\geq -\left\langle \frac{\Pi_{\mathtt{u}}\psi_\eta(t) - \mathtt{h}_{\mathtt{s}}(\Pi_{\mathtt{s}}\psi_\eta(t))}{u(t)}, \frac{\Pi_{\mathtt{u}}\nabla\tilde{H}(\psi_\eta(t))}{|\nabla\tilde{H}(\psi_\eta(t))|}\right\rangle - \|D\mathtt{h}_{\mathtt{s}}(\Pi_{\mathtt{s}}\psi_\eta(t))\| \left|\frac{\Pi_{\mathtt{s}}\nabla\tilde{H}(\psi_\eta(t))}{|\nabla\tilde{H}(\psi_\eta(t))|}\right|.
\end{aligned}$$

Let us first note, that $\|D\mathtt{h}_{\mathtt{s}}(\Pi_{\mathtt{s}}\psi_\eta(t))\| = O(|\Pi_{\mathtt{s}}\psi_\eta(t)|) = O(\delta)$ by (5.24). Furthermore, we have by (5.29) that $\Pi_{\mathtt{s}}\nabla\tilde{H}(\psi_\eta(t)) = N_s(\psi_\eta(t)) + \Pi_{\mathtt{s}}R_2(\psi_\eta(t))$ and by (5.30) we can easily deduce $|N_{\mathtt{s}}(x)| = O(\delta)$ and $|\Pi_{\mathtt{s}}R_2(x)| = O(\delta^2)$, which in comparison with $|\nabla\tilde{H}| \geq c_{\tilde{H}}$ yields

$$\|D\mathtt{h}_{\mathtt{s}}(\Pi_{\mathtt{s}}\psi_\eta(t))\| \left|\frac{\Pi_{\mathtt{s}}\nabla\tilde{H}(\psi_\eta(t))}{|\nabla\tilde{H}(\psi_\eta(t))|}\right| = O(\delta^2). \tag{5.51}$$

Therefore, we are left to consider the first term. Since $\Pi_{\mathtt{u}}\nabla\tilde{H}(\psi_\eta(t)) = N_{\mathtt{u}}(\psi_\eta(t)) + \Pi_{\mathtt{u}}R_2(\psi_\eta(t))$ with $|\Pi_{\mathtt{u}}R_2(x)| = O(\delta^2)$, it is enough to use a similar estimate like (5.37)

$$\begin{aligned}
-\left\langle \frac{\Pi_{\mathtt{u}}\psi_\eta(t) - h_s(\Pi_{\mathtt{s}}\psi_\eta(t))}{u(t)}, \frac{N_{\mathtt{u}}(\psi_\eta(t))}{|\nabla\tilde{H}(\psi_\eta(t))|}\right\rangle &\geq |\lambda^-_{\min}|u(t) + a\delta p\left(\frac{s(t)}{\delta}\right)\left|q'\left(\frac{s(t)}{\delta}\right)\right| \\
&\overset{(\mathrm{B}_q)}{\geq} |\lambda^-_{\min}|u(t) + a\delta\left(1 - 2\frac{u(t)}{\delta}\right)_+ p\left(\frac{s(t)}{\delta}\right),
\end{aligned}$$

which concludes the proof of the lower bound in (5.50) and the upper bound is a simple consequence of the error estimate (5.51).

Now we have proven the auxiliary estimates for the trajectories and can continue with the proof of (**H̃5**) under the assumption, that (5.48) holds. Let us introduce the entrance time E_η of $\psi_\eta(t)$ into $\operatorname{supp} H^+_{\mathrm{loc}}$ where we can assume that $\operatorname{supp} H^+_{\mathrm{loc}} \subset B_{2\delta}$. Let us assume that δ is chosen small enough such that $\dot{s}(t) \geq \frac{\lambda^+_{\min}\delta}{8} > 0$ whenever $s(t) \geq \frac{\delta}{2}$. Then, we can consider two cases: First, there exists $t_0 \in (E_\eta, T_\eta)$ with $s(t_0) \geq \frac{\delta}{2}$ and secondly, there exists no such t_0. In the first case, there is nothing to show, since we find that $s(t)$ is monotone and with (5.48) holds

$$\forall t \in (t_0, T_\eta) : \langle\nabla H(\psi_\eta(t)), \nabla\tilde{H}(\psi_\eta(t))\rangle \geq \frac{(\lambda^+_{\min})^2}{4}s(t_0)^2 \geq \frac{(\lambda^+_{\min})^2}{4}\frac{\delta^2}{4}. \tag{5.52}$$

Now, the second case and we assume that $s(t) < \frac{\delta}{2}$ for all $t \in (E_\eta, T_\eta)$. Then it follows from (5.50) for $u(t) \leq \frac{\delta}{4}$

$$-\dot{u}(t) \geq \frac{a\delta}{2}p\left(\frac{1}{2}\right) + O(\delta^2) \geq c_u\delta > 0 \quad \text{for some } c_u > 0 \text{ and } \delta \text{ small enough.} \tag{5.53}$$

In particular, the estimate (5.53) implies

$$u(t) = u(T_\eta) - \int_t^{T_\eta} \dot{u}(s)\mathrm{d}s \overset{(5.53)}{\geq} c_u\delta(T_\eta - t), \tag{5.54}$$

since $u(T_\eta) = 0$. We can use this observation together with the estimate (5.48) and arrive for $u \le \frac{\delta}{4}$ at

$$\langle \nabla H(\psi_\eta(t)), \nabla \tilde{H}(\psi_\eta(t)) \rangle \ge |\lambda^-_{\min}|\; a\, \frac{\delta}{2}\, p\left(\frac{1}{2}\right) u(t) \overset{(5.54)}{\ge} |\lambda^-_{\min}|\; a\, \frac{\delta}{2}\, p\left(\frac{1}{2}\right)\, c_u \delta\, (T_\eta - t). \tag{5.55}$$

Now, the conclusion (**$\tilde{\mathbf{H}}$5**) follows with (5.52) and (5.55) from the observation $|\nabla \tilde{H}| = O(\delta)$ in $\mathcal{U}$ and especially $|\nabla \tilde{H}| \le 1$ for δ chosen appropriate small enough.

To finish the proof, we still have to show (5.47). We start with (5.28) and (5.29) to deduce

$$\begin{aligned}\left\langle \nabla H(x), \nabla \tilde{H}(x) \right\rangle &= \left\langle \nabla^2 H(0)(x_{\mathtt{s}} + x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})) + R_2(x), N_{\mathtt{s}}(x) + N_{\mathtt{u}}(x) + R_2(x) \right\rangle \\ &= \left\langle \nabla^2 H(0)(x_{\mathtt{s}} + x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})), N_{\mathtt{s}}(x) + N_{\mathtt{u}}(x) \right\rangle \\ &\quad + \left\langle R_2(x), N_s(x) + N_u(x) + \nabla^2 H(0)(x_{\mathtt{s}} + x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})) + R_2(x) \right\rangle.\end{aligned}$$

Let us first estimate the order of the error terms containing $R_2(x)$. With (5.28) we find that $|R_2(x)| = O(|x|^2) = O(s^2) + O(u^2)$. Further, we have $|N_s(x)| = O(\delta)$ as well as $|N_u(x)| = O(\delta)$. Furthermore, it holds $|\mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})| = O(s^2)$ and of course $|x| = O(\delta)$. Therefore, the total approximation error can be estimated by

$$\left\langle R_2(x), N_s(x) + N_u(x) + \nabla^2 H(0)(x_{\mathtt{s}} + x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}}) \right\rangle = O(\delta s^2) + O(\delta u^2).$$

We find an estimate for $\left\langle \nabla^2 H(0)(x_{\mathtt{s}} + x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})), N_{\mathtt{s}}(x) + N_{\mathtt{u}}(x) \right\rangle$ by using the representation of $N_{\mathtt{s}}$ in (5.30) and $N_{\mathtt{u}}$ in (5.31) and going along the lines of (5.33) and (5.37)

$$\begin{aligned}&\left\langle \nabla^2 H(0)(x_{\mathtt{s}} + x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})), N_{\mathtt{s}}(x) + N_{\mathtt{u}}(x) \right\rangle \\ &= \left\langle \nabla^2 H(0) x_{\mathtt{s}}, N_{\mathtt{s}} \right\rangle + \left\langle \nabla^2 H(0)\left(x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\right), N_{\mathtt{u}}(x) \right\rangle \\ &\ge \left\langle \nabla^2 H(0) x_{\mathtt{s}}, \nabla^2 H(0) x_{\mathtt{s}} \right\rangle \\ &\qquad - a\delta \left| p'\left(\frac{s}{\delta}\right)\right| q\left(\frac{u}{\delta}\right) \left\langle \nabla^2 H(0) x_{\mathtt{s}}, \frac{x_{\mathtt{s}}}{s} \right\rangle \\ &\qquad - a\delta p\left(\frac{s}{\delta}\right) \left| q'\left(\frac{u}{\delta}\right)\right| \left|\nabla^2 H(0) x_{\mathtt{s}}\right| \left| D\mathtt{h}_{\mathtt{s}}^\top(x_{\mathtt{s}})\right| \\ &\quad + \left\langle \nabla^2 H(0)\left(x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\right), \nabla^2 H(0)\left(x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\right) \right\rangle \\ &\qquad + a\delta p\left(\frac{s}{\delta}\right) \left| q'\left(\frac{u}{\delta}\right)\right| \left\langle -\nabla^2 H(0)\left(x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})\right), \frac{x_{\mathtt{u}} - \mathtt{h}_{\mathtt{s}}(x_{\mathtt{s}})}{u} \right\rangle \\ &\ge (\lambda^+_{\min})^2 s^2 - 8a\lambda^+_{\max} s^2 - 2a\delta\lambda^+_{\max} C_h s^2 + (\lambda^-_{\min})^2 u^2 + a\delta p\left(\frac{s}{\delta}\right)\left(1 - \frac{2u}{\delta}\right)_+ \left|\lambda^-_{\min}\right| u,\end{aligned}$$

where $\lambda^+_{\max}$ is the largest eigenvalue of $\nabla^2 H(0)$. We finish by bounding the first three terms from below. Therefore, we assume, like in (5.35), that $\delta \le C_h^{-1}$ and obtain

$$(\lambda^+_{\min})^2 s^2 - 8a\lambda^+_{\max} s^2 - 2a\lambda^+_{\max}\delta C_h s^2 \ge \left((\lambda^+_{\min})^2 - 10a\lambda^+_{\max}\right) s^2 \ge \frac{(\lambda^+_{\min})^2}{2} s^2,$$

where we set $a \le \frac{(\lambda^+_{\min})^2}{20\lambda^+_{\max}}$ in the last estimate, which enforces the assumption on a in (5.36) by a factor $\frac{\lambda^+_{\min}}{\lambda^+_{\max}}$. Altogether, this shows (5.47) and finishes the proof. □

5.3. Radial Poincaré inequality via Muckenhoupt functional

The aim of this section is the proof of Proposition 5.3. In the argument we apply the Muckenhoupt functional [Muc72]. Firstly, let us restate the main result of [Muc72] in Theorem 5.20 below and clarify the connection of the Muckenhoupt functional to the Poincaré inequality in Proposition 5.21. For the sake of completeness and notation we also state the proofs of Theorem 5.20 and Proposition 5.21 in full detail.

The following theorem is a weighted Hardy inequality discovered by [Muc72] in a more general context with a sharp characterization of the optimal constant up to a factor 4.

Theorem 5.20 (Muckenhoupt functional [Muc72, Theorem 4]). *Let μ be an absolutely continuous measure on $\mathbb{R}^+$, then there exists a constant $C < \infty$ for which*

$$\forall f : \mathbb{R}^+ \to \mathbb{R} \qquad \int_0^\infty \left(\int_0^x f(t)\, \mathrm{d}t \right)^2 \mu(\mathrm{d}x) \leq C \int_0^\infty f^2 \mu(\mathrm{d}x), \tag{5.56}$$

if and only if

$$B = \sup_{r>0} \int_r^\infty \mathrm{d}\mu \int_0^r \frac{1}{\mu(x)}\, \mathrm{d}x < \infty. \tag{5.57}$$

Furthermore, the optimal C obeys the estimate

$$B \leq C \leq 4B. \tag{5.58}$$

Proof of Theorem 5.20. The proof consists of two steps. Firstly, we show that (5.57) implies (5.56) with constant $C \leq 4B$. In the second step, we will show that (5.56) also implies (5.57) by the the estimate $B \leq C$, which then establishes (5.58).

Define the function $G(x) = \int_0^x \frac{1}{\mu(y)} \mathrm{d}y$. The first step is to apply Cauchy-Schwarz inequality to the inner integral

$$\left(\int_0^x f(t) \mathrm{d}t \right)^2 \leq \int_0^x |f(t)|^2 \sqrt{G(t)}\, \mu(\mathrm{d}t) \int_0^x \frac{1}{G(t)\mu(t)}\, \mathrm{d}t \tag{5.59}$$

Now, note that with $2 \frac{\mathrm{d}}{\mathrm{d}t} \sqrt{G(t)} = \frac{1}{\sqrt{G(t)}\mu(t)}$ the last integral evaluates to

$$\int_0^x \frac{1}{G(t)\mu(t)}\, \mathrm{d}t = 2\sqrt{G(x)} \leq 2\sqrt{B} \left(\int_x^\infty \mathrm{d}\mu \right)^{-\frac{1}{2}}, \tag{5.60}$$

where the last estimate follows from the definition of the constant B. Combining (5.60) and (5.59), integrating the resulting inequality w.r.t. to μ and using Fubini's theorem to interchange the integration, we arrive at

$$\int_0^\infty \left(\int_0^x f(t)\, \mathrm{d}t \right)^2 \mu(\mathrm{d}x) \leq 2\sqrt{B} \int_0^\infty |f(t)|^2 \sqrt{G(t)} \left(\int_t^\infty \left(\int_x^\infty \mathrm{d}\mu \right)^{-\frac{1}{2}} \mu(\mathrm{d}x) \right) \mu(\mathrm{d}t). \tag{5.61}$$

Now, we observe that

$$2\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_x^\infty \mathrm{d}\mu\right)^{\frac{1}{2}} = -\left(\int_x^\infty \mathrm{d}\mu\right)^{-\frac{1}{2}} \mu(x).$$

Therewith, the inner integrals in (5.61) evaluate to

$$\int_t^\infty \left(\int_x^\infty \mathrm{d}\mu\right)^{-\frac{1}{2}} \mu(\mathrm{d}x) = -2\int_t^\infty \frac{\mathrm{d}}{\mathrm{d}x}\left(\int_x^\infty \mathrm{d}\mu\right)^{\frac{1}{2}} \mathrm{d}x = 2\left(\int_t^\infty \mathrm{d}\mu\right)^{\frac{1}{2}} \leq 2\,\frac{\sqrt{B}}{\sqrt{G(t)}}, \tag{5.62}$$

by the definition of B. Combining (5.62) and (5.61) results in the desired upper bound.

For the lower bound assume that f is non-negative. We can bound the left-hand side in (5.56) for fixed $r > 0$ from below by

$$\int_r^\infty \mathrm{d}\mu \left(\int_0^r f(t)\,\mathrm{d}t\right)^2 \leq \int_0^\infty \left(\int_0^x f(t)\,\mathrm{d}t\right)^2 \mu(\mathrm{d}x) \leq C\int_0^\infty |f(x)|^2\,\mu(\mathrm{d}x). \tag{5.63}$$

Now, we want to show that for any r holds

$$\int_r^\infty \mathrm{d}\mu \int_0^r \frac{1}{\mu(x)}\,\mathrm{d}x \leq C, \tag{5.64}$$

which implies $B \leq C$. If $\int_0^r \frac{1}{\mu(x)}\,\mathrm{d}x = 0$, then the estimate (5.64) is immediate. If $\int_0^r \frac{1}{\mu(x)}\,\mathrm{d}x = \infty$, then there exists a function $f(x)$ with $\int_0^r |f(x)|^2\,\mathrm{d}\mu(x) < \infty$ and $\int_0^r f(x)\mathrm{d}x = \infty$. Further, (5.63) implies that $\int_r^\infty \mathrm{d}\mu = 0$, which also yields (5.64). Finally, if $0 < \int_0^r \frac{1}{\mu(x)}\,\mathrm{d}x$, set $f(x) = \frac{1}{\mu(x)}$, we observe that (5.63) becomes

$$\int_r^\infty \mathrm{d}\mu \left(\int_0^r \frac{1}{\mu(x)}\,\mathrm{d}x\right)^2 \leq C\int_0^r \frac{1}{\mu(x)}\,\mathrm{d}x,$$

which results, after dividing by $\int_0^r \frac{1}{\mu(x)}\mathrm{d}x$, in the desired estimate (5.64). □

From the previous theorem it is easy to obtain the following proposition, which determines the Poncaré inequality in one dimension up to a factor of 8 for a general measure (cf. Remark 5.22). The argument is similar to the work of Bobkov and Götze [BG99], which considered the logarithmic Sobolev inequality in the continuous case and also to the work of Miclo [Mic99], which considered the Poincaré inequality in the discrete case.

Proposition 5.21 (From Muckenhoupt functional to Poincaré inequality)**.** *Let μ be a probability measure on $\mathbb{R}$ absolutely continuous w.r.t. to the Lebesgue measure. Then the constants B_m^- and B_m^+ defined by*

$$\begin{aligned} B_m^- &= \sup_{x\leq m}\left(\int_x^m \frac{1}{\mu(y)}\mathrm{d}y \int_{-\infty}^x \mathrm{d}\mu\right) \\ B_m^+ &= \sup_{x\geq m}\left(\int_m^x \frac{1}{\mu(y)}\mathrm{d}y \int_x^\infty \mathrm{d}\mu\right) \end{aligned} \tag{5.65}$$

are finite for μ-a.e. m if and only if they are finite for one common m.

Further, μ satisfies a Poincaré inequality PI(ϱ) if and only if B_m^- and B_m^+ are finite. In this case ϱ obeys the estimate

$$\max\left\{(1-F_\mu(m))B_m^-, F_\mu(m)B_m^+\right\} \le \varrho^{-1} \le 4\max\left\{B_m^-, B_m^+\right\},$$

where $F_\mu(m) = \mu((-\infty, m])$.

Remark 5.22. A common cited version (e.g. Fougères [Fou04]) of the above proposition is obtained by setting $m = m^*$ with m^* given as median of the probability measure μ, i.e. $\mu((-\infty, m^*]) = \frac{1}{2}$. In this case, the estimate of Proposition 5.21 has the form

$$\tfrac{1}{2}\max\left\{B_{m^*}^-, B_{m^*}^+\right\} \le \varrho^{-1} \le 4\max\left\{B_{m^*}^-, B_{m^*}^+\right\}.$$

This gives a general characterization of the spectral gap or Poincaré constant up to a factor of 8. However, the median is hard to compute in general and the result stated in Proposition 5.21 is easier to obtain upper bounds.

Proof of Proposition 5.21. In the first step we will show that if m^* exists such that $B_{m^*}^-$ and $B_{m^*}^+$ are finite, then B_m^- and B_m^+ are finite for μ-a.e. m. Therefore, we write $B_m^-(x) = \int_x^m \frac{1}{\mu(y)}\mathrm{d}y \int_{-\infty}^x \mathrm{d}\mu$ such that with this notation $B_m^- = \sup_{x\le m} B_m^-(x)$ and similarly for $B_m^+(x)$. By symmetry it is enough to consider $m \ge m^*$. Then, an immediate consequence is $B_m^+ \le B_{m^*}^+$. To obtain a bound for $B_m^-(x)$ with $x \le m$, we split up the integral in the definition of $B_m^-(x)$ and obtain

$$\begin{aligned} B_m^-(x) = \int_x^m \frac{1}{\mu(y)}\mathrm{d}y \int_{-\infty}^x \mathrm{d}\mu &\le \mathbb{1}_{(-\infty,m^*]}(x)\int_x^{m^*} \frac{1}{\mu(y)}\mathrm{d}y\int_{-\infty}^x \mathrm{d}\mu + \int_{m^*}^m \frac{1}{\mu(y)}\mathrm{d}y\int_{-\infty}^x \mathrm{d}\mu \\ &\le \mathbb{1}_{(-\infty,m^*]}(x) B_{m^*}^-(x) + B_{m^*}^+(m)\frac{F_\mu(x)}{1-F_\mu(m)} \le B_{m^*}^- + B_{m^*}^+\frac{F_\mu(x)}{1-F_\mu(m)}, \end{aligned}$$

which is finite for m with $0 < F_\mu(m) < 1$, hence for μ-a.e. m.

Now, we want to deduce the upper bound for the inverse spectral gap, which follows by noting that for an arbitrary $m \in \mathbb{R}$ holds

$$\operatorname{var}_\mu(f) \le \int_{-\infty}^{\infty} (f - f(m))^2\,\mathrm{d}\mu = \int_{-\infty}^{m} (f-f(m))^2\,\mathrm{d}\mu_m^- + \int_m^\infty (f-f(m))^2\,\mathrm{d}\mu_m^+ \tag{5.66}$$

where we introduced the measures

$$\mu_m^-(\mathrm{d}x) = \mathbb{1}_{(-\infty,m)}(x)\mu(\mathrm{d}x) \qquad \text{and} \qquad \mu_m^+ = \mathbb{1}_{(m,+\infty)}(x)\mu(\mathrm{d}x).$$

We can apply Theorem 5.20 to both of the last integrals in (5.66) with the measure μ substituted by $\mu_m^\pm$ and obtain the estimate (5.56) with a constant $C_m^\pm$, which is bounded by $4B_m^\pm$. For μ_m^+ this leads to

$$\int_m^\infty (f-f(m))^2\,\mathrm{d}\mu_m^+ = \int_m^\infty \left(\int_m^y f'(x)\mathrm{d}x\right)^2 \mathrm{d}\mu_m^+(y) \le 4B_m^+\int_m^\infty \left|f'\right|^2 \mathrm{d}\mu_m^+,$$

and similarly for μ_m^-. This proves out right after again combining both integrals

$$\operatorname{var}_\mu(f) \leq 4B_m^- \int_{-\infty}^{m} |f'|^2 \, \mathrm{d}\mu_m^- + 4B_m^+ \int_m^\infty |f'|^2 \, \mathrm{d}\mu_m^+ \leq 4 \max\left\{B_m^-, B_m^+\right\} \int_{-\infty}^{\infty} |f'|^2 \, \mathrm{d}\mu$$

the upper bound.

Once again we start with equation (5.66) and consider the integral over (m, ∞). Now we choose $0 < \tau < C_m^+$ and a function g_τ such that for the optimal constant C_m^+ in

$$\int_m^\infty (f - f(m))^2 \, \mathrm{d}\mu \leq C_m^+ \int_m^\infty |f'|^2 \, \mathrm{d}\mu$$

holds

$$\int_m^\infty (g_\tau - g_\tau(m))^2 \, \mathrm{d}\mu \geq \left(C_m^+ - \tau\right) \int_m^\infty |g_\tau'|^2 \, \mathrm{d}\mu. \tag{5.67}$$

These g_τ exists thanks to the optimality. By Theorem 5.20 we know that $B_m^+ \leq C_m^+$. We can assume w.l.o.g. $g_\tau(m) = 0$, else we could consider $\tilde{g}_\tau(x) = g_\tau(x) - g_\tau(m)$. Now, we set $g_\tau(x) = 0$ for $x < m$ and note $\mu(\{g_\tau = 0\}) \geq \mu((-\infty, m])) = F_\mu(m)$. Therewith, it follows

$$\left(\int g_\tau \mathrm{d}\mu\right)^2 \leq \mu(\{g_\tau \neq 0\}) \int g_\tau^2 \mathrm{d}\mu \leq (1 - F_\mu(m)) \int g_\tau^2 \mathrm{d}\mu.$$

This results in the estimate

$$\begin{aligned}\operatorname{var}_\mu(g_\tau) &= \int g_\tau^2 \mathrm{d}\mu - \left(\int g_\tau \mathrm{d}\mu\right)^2 \geq F_\mu(m) \int g_\tau^2 \mathrm{d}\mu \\ &\overset{(5.67)}{\geq} F_\mu(m)(C_m^+ - \tau) \int |g_\tau'|^2 \, \mathrm{d}\mu \geq F_\mu(m)(B_m^+ - \tau) \int |g_\tau'|^2 \, \mathrm{d}\mu.\end{aligned}$$

Hence, $\varrho^{-1} \geq F_\mu(m)(B_m^+ - \tau)$ and sending $\tau \to 0$ the desired result follows. The case for the integral on $(-\infty, m)$ follows similarly by symmetry. □

We introduce an additional tool, which allows to compare the constants B_m^- and B_m^+ in (5.65) of the Muckenhoupt functional for different measures. This will allow us to prove the scaling behavior of different Hamiltonian at low temperature, only knowing certain monotonicity properties.

Lemma 5.23 (Comparison principle for Muckenhoupt functional)**.** *For an one-dimensional Hamiltonian $H : \mathbb{R} \to \mathbb{R}$ and a function $\psi : \mathbb{R} \to \mathbb{R}$ increasing on (m, ∞) with $m \in \mathbb{R}$ let us consider the family $\{\nu_\lambda\}_{\lambda \in [0,\infty)}$ of Gibbs measures*

$$\nu_\lambda(\mathrm{d}x) := \frac{1}{\int e^{-\lambda\psi(x) - H(x)} \mathrm{d}x} e^{-\lambda\psi(x) - H(x)} \mathrm{d}x.$$

Then the constants

$$B_{m,\lambda}^+ = \sup_{x \geq m} \left(\int_x^\infty e^{-\lambda\psi(t) - H(t)} \mathrm{d}t \int_m^x e^{\lambda\psi(t) + H(t)} \mathrm{d}t \right)$$

satisfy for $\lambda_0 \le \lambda_1$ the relation

$$B^+_{m,\lambda_1} \le B^+_{m,\lambda_0} \tag{5.68}$$

If ψ is monotone decreasing, the constant $B^-_{m,\lambda}$ satisfies the relation (5.68).

Proof. Firstly, let us consider the Poincaré inequality. Let us recall the characterization of the PI constant by the Muckenhoupt functional. It is sufficient to show that for any $\lambda > 0$ and $x \ge m$

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} B^+_{m,\lambda}(x) \le 0.$$

Direct calculation yields the formula

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} B^+_{m,\lambda}(x) = \int_x^\infty e^{-\lambda\psi - H}\mathrm{d}t \int_m^x \psi e^{\lambda\psi + H}\mathrm{d}t - \int_x^\infty \psi e^{-\lambda\psi - H}\mathrm{d}t \int_m^x e^{\lambda\psi + H}\mathrm{d}t.$$

The negativity of the right-hand side follows directly from the observation that there exist $m \le t_1 \le x$ and $x \le t_2 \le \infty$ such that

$$\frac{\int_m^x \psi e^{\lambda\psi+H}\mathrm{d}t}{\int_m^x e^{\lambda\psi+H}\mathrm{d}t} - \frac{\int_x^\infty \psi e^{-\lambda\psi-H}\mathrm{d}t}{\int_x^\infty e^{-\lambda\psi-H}\mathrm{d}t} = \psi(t_1) - \psi(t_2) \le 0,$$

where we applied the monotonicity of ψ in the last step. □

Now, we have completed all the preparatory work and can directly proceed to the proof Proposition 5.3, which is restated at this point.

Proposition 5.3 (PI for radial-mixing measures). *Let $\mu(\cdot|\eta)$ be a radial-mixing measure in the sense of Definition 5.14. Then the measure $\mu(\cdot|\eta)$ satisfies $\mathrm{PI}(\varrho(\eta))$ with constant*

$$\frac{1}{\varrho(\eta)} = O(\varepsilon) \qquad \text{for} \quad \varepsilon \to 0, \qquad \text{uniformly in} \quad \eta \in S^{n-1}.$$

Proof of Proposition 5.3. We deduce the desired statement by an application of Proposition 5.21. Therefore, we have to show that $B^\pm_m = O(\varepsilon)$ for a particular choice of m. For this purpose, let us set $m = \sqrt{\frac{\varepsilon}{h''_\eta(0)}}$. Additionally, we assume that ε is small enough such that $m \le \frac{\delta}{2}$ with $\frac{\delta}{2}$ given by the conditions (**H̃1**) and (**H̃2**).

The measure $\mu(\cdot|\eta)$ is absolutely continuous and has by denoting $h_\eta(r) := H(\psi_\eta(r)$ the form

$$\mu(\mathrm{d}r) = \frac{1}{Z_\mu \hat\mu(\eta)} j_\eta(r) e^{-\frac{h_\eta(r)}{\varepsilon}}\mathrm{d}r, \text{ where } \hat\mu(\eta) = \frac{1}{Z_\mu}\int j_\eta(r) e^{-\frac{h_\eta(r)}{\varepsilon}}\mathrm{d}r.$$

Since H is quadratic in 0, it holds $h_\eta(0) = h'_\eta(0) = 0$ and by (**H̃2**) and the non-degenerate Assumption 1.12 holds $h''_\eta(0) = \langle \eta, \nabla^2 H(0)\eta\rangle \ge \lambda_{\min} > 0$. We start with the estimation of $B^-_m(r)$ for $r \le m = O(\sqrt{\varepsilon})$. Then, according to (**j1**) of Lemma 5.12 holds $j_\eta(r) = r^{n-1}$.

Now, we use the expansion $h_\eta(r) = h''_\eta(0)\frac{r^2}{2} + O(r^3) = h''_\eta(0)\frac{r^2}{2} + O(\varepsilon^{\frac{3}{2}})$ to find the estimate

$$\begin{aligned} B^-_m(r) &= \int_0^r j_\eta(t) \exp\left(-\frac{1}{\varepsilon} h_\eta(t)\right) \mathrm{d}t \int_r^m \frac{1}{j_\eta(t)} \exp\left(\frac{1}{\varepsilon} h_\eta(t)\right) \mathrm{d}t \\ &= \int_0^r t^{n-1} \exp\left(-\frac{h''_\eta(0)t^2}{2\varepsilon} + O(\sqrt{\varepsilon})\right) \mathrm{d}t \int_r^m t^{-n+1} \exp\left(\frac{h''_\eta(0)t^2}{2\varepsilon} + O(\sqrt{\varepsilon})\right) \mathrm{d}t \end{aligned}$$

We continue with the variable substitution $r = sm$ and $t = zm$ and arrive at

$$\begin{aligned} B^-_m(r) &\le m^2 \int_0^s z^{n-1} \exp\left(-\frac{z^2}{2}\right) \mathrm{d}z \int_s^1 z^{-n+1} \exp\left(\frac{z^2}{2}\right) \mathrm{d}z \left(1 + O(\sqrt{\varepsilon})\right) \\ &\le \varepsilon \frac{E_n}{h''_\eta(0)} (1 + O(\sqrt{\varepsilon})). \end{aligned}$$

The constant E_n is given as upper bound in the estimate

$$\int_0^s z^{n-1} \underbrace{e^{-\frac{z^2}{2}}}_{\le 1} \mathrm{d}z \int_s^1 z^{-n+1} \underbrace{e^{\frac{z^2}{2}}}_{\le \sqrt{e}} \mathrm{d}z \le \sqrt{e}\, \frac{s^n}{n} \begin{cases} \log(\frac{1}{s}) & , n = 2 \\ \frac{1-s^{n-2}}{(n-2)s^{n-2}} & , n > 2 \end{cases}.$$

Now, we bound the right-hand side of the above estimate. If $n = 2$, the function $s \mapsto s^2 \log \frac{1}{s}$ attains its maximum for $s = \frac{1}{\sqrt{e}}$ with value $\frac{1}{2e}$. Hence, we have $E_2 = \frac{1}{4\sqrt{e}}$. For $n > 2$, the function $s \mapsto s^2(1 - s^{n-2})$ attains its maximum for $s = \left(\frac{2}{n}\right)^{\frac{1}{n-2}}$ with value $\left(\frac{2}{n}\right)^{\frac{2}{n-2}} \left(1 - \frac{2}{n}\right) \le \frac{n-2}{n}$. Hence, we can set $E_n = \frac{\sqrt{e}}{n^2}$.

Now, we continue with the estimation of $B^+_m(r)$ for $r \ge m$. We extend the decomposition of the Jacobian determinant $\log j_\eta(r) = \log j^+_\eta(r) - \log j^-_\eta(r)$ given in **(j2)** of Lemma 5.12 from $r \in (0, \tilde{T}_\eta)$ to $r \in (\delta/2, \infty)$ by setting

$$\forall r \in (\tilde{T}_\eta, \infty): \quad j^-_\eta(r) = j^-_\eta(\tilde{T}_\eta) \quad \text{and} \quad \log j^+_\eta(r) = \log j_\eta(r) + \log j^-_\eta(\tilde{T}_\eta).$$

Then, with the monotonicity from **(j2)** holds for $j^+_\eta(r)$ the bound from below and above

$$\forall r \in (\delta/2, \tilde{T}_\eta): \quad c_j := \left(\tfrac{\delta}{2}\right)^{n-1} = j_\eta(\tfrac{\delta}{2})) \le j^+_\eta(r) \le j_\eta(\tfrac{\delta}{2}) e^{T^* C_{\tilde{H}}} = c_j e^{T^* C_{\tilde{H}}} := C_j. \tag{5.69}$$

Furthermore, j^-_η is monotone increasing on $(0, T_\eta)$ by construction. We define now for $\lambda \in [0, 1]$ the modified Muckenhoupt functional

$$\tilde{B}^+_{m,\lambda}(r) := \int_m^r e^{\frac{1}{\varepsilon} h_\eta(t) - \log j^+_\eta(r) + \lambda \log j^-_\eta(r)} \mathrm{d}t \int_r^{T_\eta} e^{-\frac{1}{\varepsilon} h_\eta(t) + \log j^+_\eta(r) - \lambda \log j^-_\eta(r)} \mathrm{d}t.$$

Then, by the comparison principle of Lemma 5.23 and the monotonicity of j^-_η holds

$$B^+_m(r) = \tilde{B}^+_{m,1}(r) \le \tilde{B}^+_{m,0}(r) = \int_m^r e^{\frac{1}{\varepsilon} h_\eta(t) - \log j^+_\eta(r)} \mathrm{d}t \int_r^{T_\eta} e^{-\frac{1}{\varepsilon} h_\eta(t) + \log j^+_\eta(r)} \mathrm{d}t. \tag{5.70}$$

Now, we have to consider the two cases $m \le r < \tilde{T}_\eta$ and $r \in (\tilde{T}_\eta, T_\eta)$. Note, that by the definition of $\tilde{T}_\eta$ given in (5.15), $\neq T_\eta$ only for $T_\eta = \infty$. Let us start and assume $m \le r < \tilde{T}_\eta$. Then, we decompose $\tilde{B}^+_{m,0}(r)$ from (5.70) as follows

$$\tilde{B}^+_{m,0}(r) \le \int_m^r \frac{1}{j^+_\eta(t)} e^{\frac{1}{\varepsilon}h_\eta(t)} \mathrm{d}t \left(\int_r^{\tilde{T}_\eta} j^+_\eta(t) e^{-\frac{1}{\varepsilon}h_\eta(t)} \mathrm{d}t + \int_{\tilde{T}_\eta}^{T_\eta} e^{-\frac{1}{\varepsilon}h_\eta(t) + \log j^+_\eta(t)} \mathrm{d}t \right), \quad (5.71)$$

The last integral in (5.71) can be estimated in general for some $T \in (\tilde{T}_\eta, T_\eta)$ by using (**h3**) of Proposition 5.13, because from the definition of $j^+_\eta(r)$ follows, that $\partial_t \log j^+_\eta(t) = \partial_t \log j_\eta(t)$ for $t \in (\tilde{T}_\eta, T_\eta)$

$$\int_T^{T_\eta} e^{-\frac{1}{\varepsilon}h_\eta(t) + \log j^+_\eta(t)} \mathrm{d}t \le j^+_\eta(T) e^{-\frac{1}{\varepsilon}h_\eta(T)} \int_T^\infty e^{-\frac{C_H(t-T)}{4\varepsilon}} \mathrm{d}t \le \frac{4}{C_H}\, \varepsilon\, j^+_\eta(T) e^{-\frac{1}{\varepsilon}h_\eta(T)}. \quad (5.72)$$

Especially, by setting $T = \tilde{T}_\eta$, we can employ the bound (5.69) and observe

$$\int_{\tilde{T}_\eta}^{T_\eta} e^{-\frac{1}{\varepsilon}h_\eta(t) + \log j^+_\eta(t)} \mathrm{d}t \le \frac{4C_j}{C_H}\, \varepsilon\, e^{-\frac{1}{\varepsilon}h_\eta(\tilde{T}_\eta)}. \quad (5.73)$$

To the first two integrals in (5.71), we can directly apply the bound (5.69). Furthermore, using the properties (**h1**) and (**h2**) of Proposition 5.13 and the just observed estimate (5.73) leads to

$$\begin{aligned} \tilde{B}^+_{m,0}(r) &\le \frac{1}{c_j} \int_m^r e^{\frac{1}{\varepsilon}h_\eta(r) - \lambda_2 \frac{(r-t)^2}{\varepsilon}} \mathrm{d}t \left(C_j \int_r^{\tilde{T}_\eta} e^{-\frac{1}{\varepsilon}h_\eta(r) - \lambda_1 \frac{(r-t)^2}{\varepsilon}} \mathrm{d}t + \frac{4C_j}{C_H}\, \varepsilon\, e^{-\frac{1}{\varepsilon}h_\eta(T)} \right) \\ &\le \frac{\sqrt{\pi\varepsilon}}{2c_j\sqrt{\lambda_2}} e^{\frac{1}{\varepsilon}h_\eta(r)} \left(\frac{\sqrt{\pi\varepsilon}C_j}{2\sqrt{\lambda_1}} e^{\frac{1}{\varepsilon}h_\eta(r)} + \frac{4C_j}{C_H}\, \varepsilon\, e^{-\frac{1}{\varepsilon}h_\eta(T)} \right) = \varepsilon \frac{\pi C_j}{4c_j\sqrt{\lambda_1\lambda_2}} + O(\varepsilon^{\frac{3}{2}}), \end{aligned}$$

where we just estimated the partial Gaussians integrals on the whole half-line.

For the last case $r \in (\tilde{T}_\eta, \infty)$, we split the first integral in $\tilde{B}^+_{m,0}(r)$ into two parts

$$\tilde{B}^+_{m,0}(r) = \left(\int_m^{\tilde{T}_\eta} \frac{1}{j^+_\eta(t)} e^{\frac{1}{\varepsilon}h_\eta(t)} \mathrm{d}t + \int_{\tilde{T}_\eta}^r e^{\frac{1}{\varepsilon}h_\eta(t) - \log j^+_\eta(t)} \mathrm{d}t \right) \int_r^\infty e^{-\frac{1}{\varepsilon}h_\eta(t) + \log j^+_\eta(t)} \mathrm{d}t. \quad (5.74)$$

We have already estimated the second integral factor in (5.72). Let us continue with estimating the first factor of (5.74). By using the lower bound $j^+_\eta(r) \ge c_j$ from (5.69) we get with the monotonicity of $h_\eta(t)$ that

$$\int_m^{\tilde{T}_\eta} \frac{1}{j^+_\eta(t)} e^{\frac{1}{\varepsilon}h_\eta(t)} \mathrm{d}t \le e^{\frac{1}{\varepsilon}h_\eta(\tilde{T}_\eta)} \frac{\tilde{T}_\eta - m}{c_j}. \quad (5.75)$$

For the second term in the first factor we employ again property (**h3**) of Proposition 5.13 to estimate

$$\int_{\tilde{T}_\eta}^r \frac{1}{j^+_\eta(t)} e^{\frac{1}{\varepsilon}h_\eta(t)} \mathrm{d}t \le \frac{1}{j^+_\eta(r)} e^{\frac{1}{\varepsilon}h_\eta(r)} \int_{\tilde{T}_\eta}^r e^{-\frac{C_H(r-t)}{4\varepsilon}} \mathrm{d}r \le \frac{4}{C_H j^+_\eta(r)} e^{\frac{1}{\varepsilon}h_\eta(r)} \varepsilon. \quad (5.76)$$

The final step consists of starting from the representation (5.70) to use a combination of the estimates (5.73) and (5.75) as well as the estimates (5.72) and (5.76) to arrive at

$$B_m^+(r) \leq \frac{\tilde{T}_\eta - m}{c_j} \frac{4C_J}{C_H} \varepsilon + \left(\frac{4}{C_H}\varepsilon\right)^2 = O(\varepsilon).$$

□

5.4. A polar mean-difference estimate

This section is devoted to the proof of Proposition 5.5. We have to show the estimate

$$\iint \left(\hat{f}(\eta) - \hat{f}(\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta)\hat{\mu}(\mathrm{d}\theta) \leq O(\varepsilon) \int |\nabla f|^2 \,\mathrm{d}\mu.$$

The procedure is similar to the proof of the mean-difference estimate of Theorem 2.23. Analogous to section 4.2, we show in the first step that it is sufficient to consider the mean-difference estimate w.r.t. the simpler measures $\nu(\cdot|\eta)$ that are Gaussian approximations of the measures $\mu(\cdot|\eta)$. This step is the content of Lemma 5.24. Following the ideas of Section 4.1, we estimate in the second step the mean-difference w.r.t. the approximations $\nu(\cdot|\eta)$ by using a transport argument. This step is the content of Lemma 5.25. Compared to Section 4.1, we have to argue more carefully since the support of the measures $\nu(\cdot|\eta)$ is only one-dimensional. Therefore, we need an additionally ingredient in the proof of Lemma 5.25: It is an identity for the spherical mean and is provided by Lemma 5.29.

Provided the ingredients are valid, the proof of Proposition 5.5 consists of a straightforward application of Lemma 5.24 and Lemma 5.25.

Now, we turn to the first step. Let us introduce the approximation for the measures $\mu(\mathrm{d}r|\eta)$ and $\hat{\mu}(\mathrm{d}\eta)$. Recall that in Section 4.2 we introduced the truncated Gaussian measures ν_i on any domain of attraction Ω_i (cf. (4.6)). In this section we had the convention to omit the index i and considered Ω to be a domain of attraction. We also assumed w.l.o.g. that the unique local minimum of H on Ω is located at $0 \in \Omega$. Using these conventions we may write the approximation ν of μ on Ω as

$$\nu(\mathrm{d}x) = \frac{1}{Z_\nu} \mathbb{1}_E(x) e^{-\frac{\nabla^2 H(0)[x]}{2\varepsilon}} \mathrm{d}x, \quad \text{with} \quad Z_\nu = \int_E e^{-\frac{\nabla^2 H(0)[x]}{2\varepsilon}} \mathrm{d}x,$$

where the set E is given by

$$E = \{x \in \mathbb{R}^n : |\sqrt{\nabla^2 H(0)}\, x| \leq \sqrt{2\varepsilon}\, \omega(\varepsilon)\}.$$

In the new coordinates given by $\psi_\eta(r)$, the measure $\nu(\mathrm{d}x)$ becomes (cf. (5.1))

$$\nu(\mathrm{d}r, \mathrm{d}\eta) = \frac{1}{Z_\nu} \mathbb{1}_E(\psi_\eta(r))\, r^{n-1}\, e^{-\frac{\nabla^2 H(0)[\eta] r^2}{2\varepsilon}}\, \mathrm{d}r\, \mathrm{d}\eta.$$

In the last identity we already applied the fact that for ε small enough the Jacobian determinant is given by $j_\eta(r) = r^{n-1}$ (cf. (**j1**) of Lemma 5.12). Following (5.2), (5.3) and (5.4), we decompose the measure $\nu(\mathrm{d}r, \mathrm{d}\eta)$ according to

$$\nu(\mathrm{d}r, \mathrm{d}\eta) = \nu(\mathrm{d}r|\eta)\, \hat{\nu}(\mathrm{d}\eta),$$

where the conditional measures $\nu(\mathrm{d}r|\eta)$ and the marginal $\hat{\nu}(\mathrm{d}\eta)$ are given by

$$\begin{aligned}\nu(\mathrm{d}r|\eta) &= \frac{1}{Z_\nu \hat{\nu}(\eta)}\, \mathbb{1}_E(\psi_\eta(r))\, r^{n-1}\, e^{-\frac{\nabla^2 H(0)[\eta] r^2}{2\varepsilon}}\, \mathrm{d}r \quad \text{and} \\ \hat{\nu}(\mathrm{d}\eta) &= \frac{1}{Z_\nu} \int r^{n-1}\, e^{-\frac{\nabla^2 H(0)[\eta] r^2}{2\varepsilon}}\, \mathrm{d}r\, \mathrm{d}\eta.\end{aligned}$$

Now, we are able to formulate the first ingredient of the proof of Proposition 5.5.

Lemma 5.24 (Approximation by truncated Gaussians.)**.** *Using the notation*

$$\hat{f}(\eta) := \int f(\psi_\eta(r))\mu(\mathrm{d}r|\eta) \qquad \textit{and} \qquad \tilde{f}(\eta) := \int f(\psi_\eta(r))\nu(\mathrm{d}r|\eta),$$

it holds for any function f

$$\iint \left(\hat{f}(\eta) - \hat{f}(\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta)\hat{\mu}(\mathrm{d}\theta) \leq 3 \iint \left(\tilde{f}(\eta) - \tilde{f}(\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta)\hat{\mu}(\mathrm{d}\theta) + O(\varepsilon) \int |\nabla f|^2\, \mathrm{d}\mu.$$

Proof of Lemma 5.24. Using the triangle inequality we can estimate

$$\left(\hat{f}(\eta) - \hat{f}(\theta)\right)^2 \leq 3 \left(\left(\hat{f}(\eta) - \tilde{f}(\eta)\right)^2 + \left(\tilde{f}(\eta) - \tilde{f}(\theta)\right)^2 + \left(\tilde{f}(\theta) - \hat{f}(\theta)\right)^2 \right).$$

By applying the same strategy as in Lemma 4.7, we can estimate the first and third term of the last inequality by

$$\left(\hat{f}(\eta) - \tilde{f}(\eta)\right)^2 \leq \frac{\mathrm{var}_{\mu(\cdot|\eta)}\left(\frac{\mathrm{d}\nu(\cdot|\eta)}{\mathrm{d}\mu(\cdot|\eta)}\right)}{\varrho_\eta} \int |\partial_r f(\psi_\eta(r))|^2\, \mathrm{d}\mu(\mathrm{d}r|\eta),$$

where ϱ_η is the Poincaré constant of $\mu(\cdot|\eta)$. By Proposition 5.3, we know that

$$\varrho_\eta^{-1} = O(\varepsilon).$$

Additionally, one can show as in Lemma 4.9 that

$$\mathrm{var}_{\mu(\cdot|\eta)}\left(\frac{\mathrm{d}\nu(\cdot|\eta)}{\mathrm{d}\mu(\cdot|\eta)}\right) = O(\sqrt{\varepsilon}\, \omega^3(\varepsilon)).$$

Finally, we only have to express $\partial_r f$ in terms of the full gradient ∇f. Because of the parametrization of the coordinates $\psi_\eta(r)$ by arc-length, we have $|\dot{\psi}_\eta(r)| = 1$. This yields the identity

$$|\partial_r f(\psi_\eta(r))| \leq |\nabla f(\psi_\eta(r))|\, |\dot{\psi}_\eta(r)| = |\nabla f(\psi_\eta(r))|\,,$$

which completes the argument. □

Now, let us turn to the second ingredient of the proof of Proposition 5.5, namely:

Lemma 5.25 (Polar mean-difference estimate)**.** *Using the notation*

$$\tilde{f}(\eta) := \int f(\psi_\eta(r))\nu(\mathrm{d}r|\eta),$$

it holds for any function f

$$\iint \left(\tilde{f}(\eta) - \tilde{f}(\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta)\, \hat{\mu}(\mathrm{d}\theta) \le O(\varepsilon) \int |\nabla f|^2 \,\mathrm{d}\mu. \tag{5.77}$$

The last statement is verified by using a *polar transport interpolation* argument. Before turning to the proof of Lemma 5.25, some preparatory work has to be done.

For the transport argument showing (5.77), we embed $\nu(\cdot|\eta)$ and $\mu(\cdot|\eta)$ along ηr as one-dimensional measures in $\mathbb{R}^n$ and note that $\nu(\cdot|\eta) \ll \mu(\cdot|\eta)$ for all $\eta \in S^{n-1}$. Then, we can still define for $\eta, \theta \in S^{n-1}$ a transport distance like $\mathcal{T}_\mu$ (cf. Definition 4.1) by considering the Radon-Nikodym derivative in the cost density w.r.t. $\mu(\cdot|\eta)$.

We denote by $\sphericalangle(\eta,\theta) \subset S^{n-1}$ the geodesic on S^{n-1} connecting η and θ parameterized on $[0, |\sphericalangle(\eta,\theta)|]$, where $|\sphericalangle(\eta,\theta)|$ is the arc-length of the geodesic, bounded by π for any $\eta, \theta \in S^{n-1}$. Note that $\sphericalangle(\eta,\theta)$ is unique for almost all $\eta, \theta \in S^{n-1}$. Let $(\eta_s)_{s\in[0,|\sphericalangle(\eta,\theta)|]}$ be its constant speed parametrization. Then consider the transport interpolation defined on $\operatorname{supp}\nu(\cdot|\eta_0)$ by

$$\Phi_s(\eta_0 r) := \sigma_s^{\frac{1}{2}} \sigma_0^{-\frac{1}{2}} \eta_s r \qquad \text{where} \qquad \sigma_s^{-1} = \nabla^2 H(0)[\eta_s]. \tag{5.78}$$

The transport interpolation is chosen such that

$$(\Phi_s)_\sharp \nu(\cdot|\eta_0) = \nu(\cdot|\eta_s).$$

In the proof of Lemma 5.25 we need an estimate on the transport speed, namely:

Lemma 5.26. *With the definitions from above, it holds*

$$|\dot{\Phi}_s \circ \Phi_s^{-1}(r\eta_s)| \le r\sqrt{1 + \frac{1}{4}\left(\frac{\lambda_{\max}}{\lambda_{\min}} - 1\right)^2}, \tag{5.79}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ is the largest and smallest eigenvalue of $\nabla^2 H(0)$, respectively.

Proof of Lemma 5.26. First, for the derivative and the inverse of Φ_s, we have that

$$\dot{\Phi}_s(\eta_0 r) = \dot{\eta}_s \sigma_s^{\frac{1}{2}} \sigma_0^{-\frac{1}{2}} r - \eta_s \frac{1}{2} \frac{\dot{\sigma}_s}{\sqrt{\sigma_s \sigma_0}} r \quad \text{and} \quad \Phi_s^{-1}(\eta_s r) = \eta_0 \sigma_s^{-\frac{1}{2}} \sigma_0^{\frac{1}{2}} r.$$

Hence, the composition is given by

$$\dot{\Phi}_s \circ \Phi_s^{-1}(\eta_s r) = \dot{\eta}_s r - \eta_s \frac{1}{2}\frac{\dot{\sigma}_s}{\sigma_s} r.$$

By using $\langle \eta_s, \dot\eta_s \rangle = 0$ and $|\eta_s| = |\dot\eta_s| = 1$ we obtain the estimate

$$\left|\dot\Phi_s \circ \Phi_s^{-1}(\eta_s r)\right| = r\sqrt{|\dot\eta_s|^2 + \frac{1}{4}|\eta_s|^2 \left|\frac{\dot\sigma_s}{\sigma_s}\right|^2} = r\sqrt{1 + \frac{1}{4}\left|\frac{\dot\sigma_s}{\sigma_s}\right|^2}.$$

From the definition of σ_s (5.78), we arrive at

$$\left|\frac{\dot\sigma_s}{\sigma_s}\right| = \frac{2\left|\langle \eta_s, \nabla^2 H(0)\dot\eta_s\rangle\right|}{\nabla^2 H(0)[\eta_s]} \leq \frac{2\left|\langle \eta_s, \nabla^2 H(0)\dot\eta_s\rangle\right|}{\lambda_{\min}}.$$

Now, we use the spectral decomposition of $\nabla^2 H(0) = \sum_{i=1}^n \lambda_i v_i \otimes v_i$ and write $\eta_s = \sum_{i=1}^n a_i v_i$ as well as $\dot\eta_s = \sum_{i=1}^n b_i v_i$. Therewith, we obtain

$$\langle \eta_s, \nabla^2 H(0)\dot\eta_s\rangle = \sum_{i=1}^n \lambda_i a_i b_i \leq \lambda_{\max}\sum_{i=1}^n a_i b_i[a_i b_i > 0] - \lambda_{\min}\sum_{i=1}^n |a_i|\,|b_i|\,[a_i b_i < 0], \quad (5.80)$$

where $[a_i b_i < 0] = 1$ if $a_i b_i < 0$ and 0 else. From $\langle \eta_s, \dot\eta_s\rangle = 0$ follows $\sum_{i=1}^n a_i b_i = 0$ and therefore

$$\sum_{i=1}^n a_i b_i [a_i b_i > 0] = -\sum_{i=1}^n a_i b_i [a_i b_i < 0].$$

In addition, the observation

$$\sum_{i=1}^n a_i b_i[a_i b_i > 0] + \sum_{i=1}^n |a_i|\,|b_i|\,[a_i b_i < 0] \leq \frac{1}{2}\sum_{i=1}^n a_i^2 + \frac{1}{2}\sum_{i=1}^n b_i^2 = \frac{1}{2}\left(|\eta_s| + |\dot\eta_s|\right) = 1,$$

implies $\sum_{i=1}^n |a_i|\,|b_i|\,[a_i b_i < 0] \leq \frac{1}{2}$. A combination with (5.80) results in

$$\langle \eta_s, \nabla^2 H(0)\dot\eta_s\rangle \leq (\lambda_{\max} - \lambda_{\min})\sum_{i=1}^n a_i b_i[a_i b_i > 0] \leq \frac{1}{2}(\lambda_{\max} - \lambda_{\min}),$$

which shows the result. □

The next lemma contains the main contribution to the proof of Lemma 5.25. It is the estimation of the mean-difference $(\tilde f(\eta) - \tilde f(\theta))$ by applying a transport argument similar to the argument outlined in Section 4.1.

Lemma 5.27. *For any $\eta, \theta \in S^{n-1}$ holds*

$$\left(\tilde f(\eta) - \tilde f(\theta)\right)^2 \leq O(\varepsilon)\int_{\sphericalangle(\eta,\theta)}\int |\nabla f|^2 \;\mu(\mathrm{d}r|\xi)\;\mathcal{H}^1(\mathrm{d}\xi),$$

uniformly in η, θ. We want to recall that $\sphericalangle(\eta,\theta)$ denotes the geodesic on S^{n-1} connecting η and θ.

Proof. We start estimating the mean-difference in the same spirit as in (4.1), namely

$$\begin{aligned}\left(\tilde{f}(\eta)-\tilde{f}(\theta)\right)^2 &= \left(\int_0^{|\sphericalangle(\eta,\theta)|}\int \frac{\mathrm{d}f(\Phi_s(r\eta_0))}{\mathrm{d}s}\,\nu(\mathrm{d}r|\eta_0)\,\mathrm{d}s\right)^2 \\ &\leq \left(\int_0^{|\sphericalangle(\eta,\theta)|}\int |\nabla f(r\eta_s)|\,|\dot{\Phi}_s\circ\Phi_s^{-1}(r\eta_s)|\,\nu(\mathrm{d}r|\eta_s)\,\mathrm{d}s\right)^2 \\ &\overset{(5.79)}{\leq} \kappa_H\left(\int_0^{|\sphericalangle(\eta,\theta)|}\int |\nabla f(r\eta_s)|\; r\,\frac{\nu(\mathrm{d}r|\eta_s)}{\mu(\mathrm{d}r|\eta_s)}\,\mu(\mathrm{d}r|\eta_s)\,\mathrm{d}s\right)^2,\end{aligned}$$

where $\kappa_H := 1+\frac{1}{4}\left(\frac{\lambda_{\max}}{\lambda_{\min}}-1\right)^2$ from (5.79). From this point, we estimate more along the lines of Remark 4.2 by using Cauchy-Schwarz w.r.t. $L^2(\mu(\mathrm{d}r|\eta_s))$ and Jensen's inequality for the outer integral

$$\begin{aligned}\left(\tilde{f}(\eta)-\tilde{f}(\theta)\right)^2 &\leq \kappa_H\,|\sphericalangle(\eta,\theta)|\int_0^{|\sphericalangle(\eta,\theta)|}\int |\nabla f(r\eta_s)|^2\,\frac{\nu(\mathrm{d}r|\eta_s)}{\mu(\mathrm{d}r|\eta_s)}\,\mu(\mathrm{d}r|\eta_s)\int r^2\,\nu(\mathrm{d}r|\eta_s)\,\mathrm{d}s \\ &\approx \kappa_H\pi\int_0^{|\sphericalangle(\eta,\theta)|}\int |\nabla f|^2\,\mu(\mathrm{d}r|\eta_s)\int r^2\,\nu(\mathrm{d}r|\eta_s)\,\mathrm{d}s\end{aligned}$$

by using again the error estimate on the approximation given in Lemma 4.9 and therefore $\frac{\nu(\mathrm{d}r|\eta_s)}{\mu(\mathrm{d}r|\eta_s)} = 1+O(\sqrt{\varepsilon}|\log^{\frac{3}{2}}(\varepsilon)|)$. Finally, note that the second moment of $\nu(r|\eta_s)$ is $O(\varepsilon)$ uniformly in η_s, due to the non-degeneracy assumption on the Hessian $\nabla^2 H(0)$ in the minimum (cf. Assumption 1.9). □

Before turning to the proof of Lemma 5.25 we need two more ingredients. The next lemma contains an asymptotic characterization of the relative density of the marginal measure $\hat{\mu}$ w.r.t. the uniform probability measure on S^{n-1} denoted by ς.

Lemma 5.28. *The relative density of $\hat{\mu}$ w.r.t. the uniform probability measure on S^{n-1} denoted by ς can be estimated for any $\xi\in S^{n-1}$ as*

$$\frac{\mathrm{d}\hat{\mu}}{\mathrm{d}\varsigma}(\xi)\approx\frac{\sqrt{\det\nabla^2 H(0)}}{(\nabla^2 H(0)[\xi])^{\frac{n}{2}}}.$$

In particular, the last identity yields for all $\xi\in S^{n-1}$ the uniform bound

$$\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{-\frac{n}{2}} \lesssim \frac{\mathrm{d}\hat{\mu}}{\mathrm{d}\varsigma}(\xi) \lesssim \left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{\frac{n}{2}}. \tag{5.81}$$

Proof of Lemma 5.28. By the assumption (**$\tilde{\mathbf{H}}$2**) the coordinates are spherically around the local minimum. Hence, expanding H in 0 in the direction ξ yields

$$H(r\xi) = H(0)+\nabla^2 H(0)[\xi]\frac{r^2}{2}+O(r^3).$$

Now, we choose Z_μ such that $H(0) = 0$. Then it holds

$$\begin{aligned}
\hat{\mu}(\xi) &= \frac{1}{Z_\mu} \int_0^{T_\xi} e^{-\frac{H(\psi_\xi(r))}{\varepsilon}} j_\xi(r)\, \mathrm{d}r \\
&= \frac{1}{Z_\mu} \int_0^{\sqrt{\varepsilon}\,\omega(\varepsilon)} r^{n-1}\, e^{-\frac{\nabla^2 H(0)[\xi] r^2}{2\varepsilon}}\, \mathrm{d}r \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon)\right) \;+\; \frac{1}{Z_\mu} \int_{\sqrt{\varepsilon}\,\omega(\varepsilon)}^{T_\xi} e^{-\frac{H(\psi_\xi(r))}{\varepsilon}}\, \mathrm{d}r \\
&= \frac{1}{Z_\mu} \frac{1}{2} \left(\frac{2\varepsilon}{\nabla^2 H(0)[\xi]} \right)^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right) \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon)) + O(e^{-\omega^2(\varepsilon)}\right).
\end{aligned}$$

By noting that $Z_\mu \approx (2\pi\varepsilon)^{\frac{n}{2}} \left(\det \nabla^2 H(0)\right)^{-\frac{1}{2}}$, we arrive at

$$\hat{\mu}(\xi) \approx \frac{\sqrt{\det \nabla^2 H(0)}}{(\nabla^2 H(0)[\xi])^{\frac{n}{2}}} \; \frac{\Gamma\left(\frac{n}{2}\right)}{2\pi^{\frac{n}{2}}} \approx \frac{\sqrt{\det \nabla^2 H(0)}}{(\nabla^2 H(0)[\xi])^{\frac{n}{2}}}\, \varsigma(\xi),$$

where we used the fact that the density $\varsigma(\xi)$ is given by the constant

$$\varsigma(\xi) = \left(\mathcal{H}^{n-1}(S^{n-1})\right)^{-1} = \frac{\Gamma\left(\frac{n}{2}\right)}{2\pi^{\frac{n}{2}}}.$$

□

The last remaining ingredient for the proof of Lemma 5.25 is the following statement on multiple integrals on the sphere.

Lemma 5.29 (Spherical mean)**.** *Let ς be the normalized uniform measure on S^{n-1}. Then it holds for all $F : S^{n-1} \to L^1(S^{n-1}, \varsigma)$ that*

$$\int_{S^{n-1}} \int_{S^{n-1}} \int_{\sphericalangle(\theta,\eta)} F(\xi)\, \mathcal{H}^1(\mathrm{d}\xi)\, \varsigma(\mathrm{d}\eta)\, \varsigma(\mathrm{d}\theta) = \frac{\pi}{2} \int_{S^{n-1}} F(\theta)\, \varsigma(\mathrm{d}\theta).$$

Proof. Choose $e \in S^{n-1}$ fixed. Then by the transitivity of SO^n on S^{n-1}, exists for all $\theta \in S^{n-1}$ a rotation $g_\theta \in SO^n$ such that $g_\theta \theta = e$. Then, the result follows from the variable substitutions first $\xi \mapsto g_\theta \xi$ and second $\eta \mapsto g_\theta \eta$

$$\begin{aligned}
\int_{S^{n-1}} \int_{S^{n-1}} \int_{\sphericalangle(\theta,\eta)} F(\xi)\, \mathcal{H}^1(\mathrm{d}\xi)\, \varsigma(\mathrm{d}\eta)\, \varsigma(\mathrm{d}\theta) &= \int_{S^{n-1}} \int_{S^{n-1}} \int_{\sphericalangle(e, g_\theta^{-1}\eta)} F(g_\theta \xi)\, \mathcal{H}^1(\mathrm{d}\xi)\, \varsigma(\mathrm{d}\eta)\, \varsigma(\mathrm{d}\theta) \\
&= \int_{S^{n-1}} \int_{S^{n-1}} \int_{\sphericalangle(e,\eta)} F(g_\theta \xi)\, \mathcal{H}^1(\mathrm{d}\xi)\, \varsigma(\mathrm{d}\eta)\, \varsigma(\mathrm{d}\theta)
\end{aligned}$$

A change of the integration order from $\mathcal{H}^1(\mathrm{d}\xi)\ \varsigma(\mathrm{d}\eta)\ \varsigma(\mathrm{d}\theta)$ to $\varsigma(\mathrm{d}\theta)\ \mathcal{H}^1(\mathrm{d}\xi)\ \varsigma(\mathrm{d}\eta)$ together with the transitivity of SO^n on S^{n-1} results in

$$\int_{S^{n-1}} F(g_\theta \xi)\, \varsigma(\mathrm{d}\theta) = \int_{S^{n-1}} F(\theta)\, \varsigma(\mathrm{d}\theta), \quad \text{independently of } \xi.$$

The last step is to observe, that the remaining double integral is the mean length of a geodesic on S^{n-1}

$$\int_{S^{n-1}} \int_{\sphericalangle(e,\eta)} \mathcal{H}^1(\mathrm{d}\xi)\,\varsigma(\mathrm{d}\eta) = \int_{S^{n-1}} |\sphericalangle(e,\eta)|\,\varsigma(\mathrm{d}\eta) = \frac{\pi}{2}.$$

□

Finally, we can turn to the proof of Lemma 5.25

Proof of Lemma 5.25. Integrating the inequality of Lemma 5.27 w.r.t. $\hat{\mu}(\eta)$ and $\hat{\mu}(\theta)$ results in the estimate

$$\iint \left(\tilde{f}(\eta) - \tilde{f}(\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta)\hat{\mu}(\mathrm{d}\theta) \leq O(\varepsilon) \iint \int_{\sphericalangle(\eta,\theta)} \int |\nabla f(r\xi)|^2 \,\mu(\mathrm{d}r|\xi)\; \mathcal{H}^1(\mathrm{d}\xi)\; \hat{\mu}(\mathrm{d}\eta)\hat{\mu}(\mathrm{d}\theta).$$

Hence, we arrive by setting $F(\xi) = \int |\nabla f(r\xi)|^2 \,\mu(\mathrm{d}r|\xi)$ at

$$\iint \left(\tilde{f}(\eta) - \tilde{f}(\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta)\hat{\mu}(\mathrm{d}\theta) \leq O(\varepsilon) \iint \int_{\sphericalangle(\eta,\theta)} F(\xi)\; \mathcal{H}^1(\mathrm{d}\xi)\; \varsigma(\mathrm{d}\eta)\varsigma(\mathrm{d}\theta). \tag{5.82}$$

Applying Lemma 5.29 to (5.82) and changing the measure back with the help of (5.81) results into

$$\begin{aligned}\iint \left(\tilde{f}(\eta) - \tilde{f}(\theta)\right)^2 \hat{\mu}(\mathrm{d}\eta)\hat{\mu}(\mathrm{d}\theta) &\leq O(\varepsilon) \int\int |\nabla f(\psi_\eta(r))|^2\; \mu(\mathrm{d}r|\eta)\; \hat{\mu}(\mathrm{d}\eta) \\ &= O(\varepsilon) \int |\nabla f|^2 \,\mathrm{d}\mu,\end{aligned}$$

which yields the desired statement (5.77). □

Chapter 6

Poincaré and logarithmic Sobolev inequalities for mixtures

Despite being the last chapter of this dissertation, this chapter is the prelude of the proof of Eyring-Kramers formula and the first contact of the author to functional inequalities.

We consider a mixture of two probability measures μ_0 and μ_1 on $\mathbb{R}^n$, i.e. a measure μ_p of the form

$$\mu_p = p\mu_0 + (1-p)\mu_1, \quad \text{where } p \in [0,1]. \tag{6.1}$$

Hereby, we assume that both of the measures μ_0 and μ_1 are absolutely continuous w.r.t. to the Lebesgue measure and their supports are nested, i.e. $\operatorname{supp}\mu_0 \subseteq \operatorname{supp}\mu_1$ or $\operatorname{supp}\mu_1 \subseteq \operatorname{supp}\mu_0$. Under these assumptions at least one measure is absolutely continuous to the other one

$$\mu_0 \ll \mu_1 \quad \text{or} \quad \mu_1 \ll \mu_0.$$

In addition, it holds the change of measure formula

$$\mathrm{d}\mu_0 = \frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_1}\,\mathrm{d}\mu_1 \qquad \text{or} \qquad \mathrm{d}\mu_1 = \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\,\mathrm{d}\mu_0.$$

Mixtures of probability measures are also studied in the work of Chafaï and Malrieu [CM10]. The aim is to deduce simple criterions under which the measure μ_p (6.1) satisfies $\mathrm{PI}(\varrho_p)$ and $\mathrm{LSI}(\alpha_p)$ knowing that $\mu_{\{0,1\}}$ satisfy $\mathrm{PI}(\varrho_{\{0,1\}})$ and $\mathrm{LSI}(\alpha_{\{0,1\}})$. In one dimension, they constructed a functional criterion depending on the distribution function of the measures μ_0 and μ_1 expressing the Poincaré and logarithmic Sobolev constant of the mixture. There, they observe for certain mixtures a logarithmically blow-up of the logarithmic Sobolev constant in p, whereas the Poincaré constant stays bounded. This behavior corresponds to the one we already observed in the Eyring-Kramers formulas in Corollary 2.29 (cf. Remark 2.32).

In this chapter, a part of the results from the work of Chafaï and Malrieu [CM10] are extended to the multidimensional case. We will deduce a simple estimate for the Poincaré and logarithmic Sobolev constant for the case, where at least one of the measures μ_0 and μ_1 is absolutely continuous to the other one. The estimate will be optimal in the scaling behavior of the mixture parameter p, i.e. we will observe a logarithmic blow-up behavior in p for the logarithmic Sobolev constant, whereas the Poincaré constant stays bounded. However, sometimes the principle asserts a blow-up of the logarithmic Sobolev constant, when the constant actually stays bounded. In general, this phantom blow-ups can be ruled out by a combination of the Bakry-Émery criterion (cf. Theorem 3.1 with the Holley-Stroock perturbation principle (cf. Theorem 3.2).

Let us first introduce the principle in the multidimensional case for the Poincaré (cf. Section 6.1) and logarithmic Sobolev (cf. Section 6.2) inequality and then illustrate the proceeding on certain examples of mixtures (cf. Section 6.3). We will close this chapter with an outlook how the weighted transport distance could be used in the framework of mixtures, especially in the case, where the mixture components μ_0 and μ_1 are singular.

6.1. Poincaré inequality

We start the argument with an easy but powerful observation, if the supports of μ_0 and μ_1 coincide.

Lemma 6.1 (Mean-difference as covariance). *If $\mu_0 \ll \mu_1$ and $\mu_1 \ll \mu_0$, i.e. if $\operatorname{supp}\mu_0 = \operatorname{supp}\mu_1$, then for any $\vartheta \in [0,1]$ and any test function $f : \mathbb{R}^n \to \mathbb{R}$ holds*

$$\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f) = -\vartheta \operatorname{cov}_{\mu_0}\left(f, \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right) + (1-\vartheta) \operatorname{cov}_{\mu_1}\left(f, \frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_1}\right). \tag{6.2}$$

Proof. By the change of measure formula we observe that the covariances above are just the difference of the expectation on the right-hand side

$$\operatorname{cov}_{\mu_0}\left(f, \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right) = \mathbb{E}_{\mu_0}\left(f\, \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right) - \mathbb{E}_{\mu_0}(f)\mathbb{E}_{\mu_0}\left(\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right) = \mathbb{E}_{\mu_1}(f) - \mathbb{E}_{\mu_0}(f)$$

and likewise for $\operatorname{cov}_{\mu_1}(f, \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0})$. □

Remark 6.2. The above Lemma was the first observation in generalizing [CM10, Lemma 4.3] to the multidimensional case, but demanding that both measures are absolutely continuous to each other. In [CM10] an optimal control of the mean-difference in one-dimension in terms of the distribution functions F_i of the measures μ_i, not necessarily absolutely continuous to each other, was deduced:

$$(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f))^2 \le I(p) \int |f'|^2 \,\mathrm{d}\mu_p, \quad \text{where } I(p) := \int \frac{(F_1(x) - F(0))^2}{p\mu_1(x) + q\mu_0(x)} \mathrm{d}x. \tag{6.3}$$

The ultimate answer of the mean-difference estimate is given by the weighted transport distance introduced in Chapter 4 (cf. Section 6.4). However, let us investigate under which conditions a strategy using the representation (6.2) leads to good results.

The subsequent strategy is: We will use in (6.2) a Cauchy-Schwarz inequality to arrive at the product of two variances. Then, we can use $\mathrm{PI}(\varrho_0)$ or $\mathrm{PI}(\varrho_1)$. The parameter ϑ leaves some freedom to optimize the resulting expression. This allows us to prove the following theorem, which is the generalization of [CM10, Theorem 4.4] to the multidimensional case for the Poincaré inequality provided μ_0 and μ_1 are absolutely continuous to each other.

Theorem 6.3 (PI for absolutely continuous mixtures)**.** *Assume that μ_0 and μ_1 satisfy $\mathrm{PI}(\varrho_0)$ resp. $\mathrm{PI}(\varrho_1)$ and are absolutely continuous to each other, then for all $p,q \geq 0$ with $p+q=1$ the mixture measure $\mu_p = p\mu_0 + q\mu_1$ satisfies $\mathrm{PI}(\varrho_p)$ with*

$$\frac{1}{\varrho_p} \leq \begin{cases} \frac{1}{\varrho_0} & , \frac{\varrho_1}{\varrho_0} \geq 1 + pc_{10} \\ \frac{1}{\varrho_1} & , \frac{\varrho_0}{\varrho_1} \geq 1 + qc_{01} \\ \frac{pc_{10}+pqc_{01}c_{10}+qc_{01}}{p\varrho_0 c_{10}+q\varrho_1 c_{01}} & , \textit{else}. \end{cases} \tag{6.4}$$

where

$$c_{01} = \mathrm{var}_{\mu_0}\left(\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right) \quad \textit{and} \quad c_{10} = \mathrm{var}_{\mu_1}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_1}\right).$$

Proof. We decompose the variance of f w.r.t. μ_p

$$\mathrm{var}_{\mu_p}(f) = p\,\mathrm{var}_{\mu_0}(f) + q\,\mathrm{var}_{\mu_1}(f) + pq\left(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f)\right)^2.$$

The first two terms are just the expectation of the conditioned variances. The second term is the variance of a Bernoulli random variable. By applying Lemma 6.1 to the mean-difference and estimating the square by using some $\eta > 0$

$$(a+b)^2 \leq (1+\eta)a^2 + (1+\tfrac{1}{\eta})b^2$$

we obtain

$$\begin{aligned}
\mathrm{var}_\mu(f) &\leq p\,\mathrm{var}_{\mu_0}(f) + q\,\mathrm{var}_{\mu_1}(f) + \\
&\quad + pq\left((1+\eta)\vartheta^2\,\mathrm{cov}^2_{\mu_0}\left(f,\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right) + \left(1+\tfrac{1}{\eta}\right)(1-\vartheta)^2\,\mathrm{cov}_{\mu_1}\left(f,\frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_1}\right)\right) \\
&\leq \left(1+(1+\eta)\vartheta^2 qc_{01}\right)p\,\mathrm{var}_{\mu_0}(f) + \left(1+\left(1+\tfrac{1}{\eta}\right)(1-\vartheta)^2 pc_{10}\right)q\,\mathrm{var}_{\mu_1}(f) \\
&\leq \frac{1+(1+\eta)\vartheta^2 qc_{01}}{\varrho_0}\int |\nabla f|^2\, p\mathrm{d}\mu_0 + \frac{1+\left(1+\frac{1}{\eta}\right)(1-\vartheta)^2 pc_{10}}{\varrho_1}\int |\nabla f|^2\, q\mathrm{d}\mu_1 \\
&\leq \max\left\{\frac{1+(1+\eta)\vartheta^2 qc_{01}}{\varrho_0}, \frac{1+\left(1+\frac{1}{\eta}\right)(1-\vartheta)^2 pc_{10}}{\varrho_1}\right\}\int |\nabla f|^2\,\mathrm{d}\mu.
\end{aligned} \tag{6.5}$$

Option 1: Optimizing in η and ϑ. W.l.o.g. we assume $\varrho_0 \geq \varrho_1$. The other case can be always obtained by interchanging the roles of μ_0 and μ_1. If $\varrho_0 > \varrho_1$, then we can set $\vartheta = 1$ and send $\eta \to 0$ and are optimal as long as

$$\frac{1+qc_{01}}{\varrho_0} \leq \frac{1}{\varrho_1},$$

This corresponds to the second case in (6.4). By symmetry the first case follows if $\varrho_1 \geq \varrho_0$.

Now, we assume $\varrho_0 \geq \varrho_1$ and $\varrho_0 \leq (1+qc_{01})\varrho_1$. For every $\vartheta \in (0,1)$ we can choose a unique $\eta_* > 0$ such that both terms in the $\max$ of the right-hand side in (6.5) are equal. Because $qc_{01} > 0$ and $pc_{10} > 0$ we can minimize the sum of the coefficients in front $\alpha(\vartheta) = (1+\eta)\vartheta^2 + (1+\frac{1}{\eta})(1-\vartheta)^2$ in ϑ as a function of η. This leads to $\vartheta_* = \frac{1}{1+\eta}$ and we find

$$\alpha(\vartheta^*) = \frac{1}{1+\eta} + \frac{\eta}{1+\eta} = 1.$$

Hence, we observed that $s = (1+\eta_*)\vartheta_*^2 = \frac{1}{1+\eta_*} \in (0,1)$ and $(1+\frac{1}{\eta_*})(1-\vartheta_*)^2 = \frac{\eta_*}{1+\eta_*} = 1-s$. Thus, the problem can be rephrased: Find $s_* \in (0,1)$ which solves

$$\frac{1+sqc_{01}}{\varrho_0} = \frac{1+(1-s)pc_{10}}{\varrho_1}.$$

This s_* is given by

$$s_* = \frac{(1+pc_{10})\varrho_0 - \varrho_1}{p\varrho_0 c_{10} + q\varrho_1 c_{01}}$$

The according value of the $\max$ in (6.5) is given by

$$\frac{1+s_*qc_{01}}{\varrho_0} = \frac{pc_{10} + \frac{\varrho_1}{\varrho_0}qc_{01} + (1+pc_{10})qc_{01} - \frac{\varrho_1}{\varrho_0}qc_{01}}{p\varrho_0 c_{10} + q\varrho_1 c_{01}} = \frac{pc_{10} + pqc_{01}c_{10} + qc_{01}}{p\varrho_0 c_{10} + q\varrho_1 c_{01}}.$$

Option 2: Calculus. First let us solve the quadratic equation in η, which occurs from setting both terms of the $\max$ of the right-hand side in (6.5) equal to each other. To keep the constants feasible, we set $a_0 = \vartheta^2 qc_{01}$ and $a_1 = (1-\vartheta)^2 pc_{10}$. The solution is given by

$$\eta = \frac{1}{2a_0\varrho_1}\left((1+a_1)\varrho_0 - (1+a_0)\varrho_1 + \sqrt{\left((1+a_0)\varrho_1 - (1+a_1)\varrho\right)^2 + 4a_0a_1\varrho_0\varrho_1}\right),$$

Note that the other solution of the quadratic equation is given by $\frac{1}{\eta}$. After substituting η back into the max, we obtain an expression in ϑ

$$\begin{aligned} m(\vartheta) = \frac{1}{2\varrho_0\varrho_1}\Bigg(&(1+(1-\vartheta)^2c_1)\varrho_0 + (1+\vartheta^2c_0)\varrho_1 + \\ &+ \sqrt{\left((1+(1-\vartheta)^2c_1)\varrho_0 - (1+\vartheta^2c_0)\varrho_1\right)^2 + 4c_0c_1\varrho_0\varrho_1\vartheta^2(1-\vartheta)^2}\Bigg), \end{aligned}$$

where we write now $c_0 = qc_{01}$ and $c_1 = pc_{10}$. Now, we can minimize $m(\vartheta)$ in ϑ. For the derivative we find

$$m'(\vartheta) = 0 \quad \Leftrightarrow \quad (\varrho_0 - \varrho_1)(1-\vartheta)\vartheta\left((1+c_0\vartheta)\varrho_1 - (1+c_1(1-\vartheta))\varrho_0\right) = 0.$$

Thus we find three zeros if $\varrho_0 \neq \varrho_1$

$$\vartheta_0 = 0, \qquad \vartheta_1 = 1, \qquad \vartheta_* = \frac{\varrho_0 - \varrho_1 + c_1\varrho_0}{c_1\varrho_0 + c_0\varrho_1}.$$

The first two zeros evaluate to

$$m(0) = \begin{cases} \frac{1+c_1}{\varrho_1} & , \varrho_1 \leq (1+c_1)\varrho_0 \\ \frac{1}{\varrho_0} & , \varrho_1 \geq (1+c_1)\varrho_0 \end{cases} \quad \text{and} \quad m(1) = \begin{cases} \frac{1+c_0}{\varrho_0} & , \varrho_0 \leq (1+c_0)\varrho_1 \\ \frac{1}{\varrho_1} & , \varrho_0 \geq (1+c_0)\varrho_1 \end{cases}.$$

Further, m evaluates at the third zero to

$$m(\vartheta_*) = \frac{1}{\varrho_0 c_1 + \varrho_1 c_0} \begin{cases} c_1 + c_0 c_1 + c_0 & , \varrho_1 \leq (1+c_1)\varrho_0 \vee \varrho_0 \leq (1+c_0)\varrho_1 \\ c_1 \frac{\varrho_0}{\varrho_1} + \frac{(\varrho_0-\varrho_1)^2}{\varrho_0\varrho_1} + c_0 \frac{\varrho_1}{\varrho_0} & , \varrho_1 \geq (1+c_1)\varrho_0 \wedge \varrho_0 \geq (1+c_0)\varrho_1. \end{cases}$$

The last step is to find the minimum for every case. If we consider the case $\varrho_1 \leq (1+c_1)\varrho_0$ we find that

$$m(\vartheta_*) = \frac{c_1 + c_0 c_1 + c_0}{\varrho_0 c_1 + \varrho_1 c_0} \leq \frac{1}{\varrho_1} \frac{c_1 + c_0 c_1 + c_0}{\frac{c_1}{1+c_1} + c_0} \overset{c_0=0}{\leq} \frac{1+c_1}{\varrho_1} = m(0).$$

In the same manner we can show if $\varrho_0 \leq (1+c_0)\varrho_1$ that $m(\vartheta_*) \leq m(1)$. These estimates show the first two cases in the definition of $\frac{1}{\varrho}$. For the third case, let $\varrho_1 \geq (1+c_1)\varrho_0$ and $\varrho_0 \geq (1+c_0)\varrho_1$ then

$$m(\vartheta_*) = \frac{c_1 \frac{\varrho_0}{\varrho_1} + \frac{(\varrho_0-\varrho_1)^2}{\varrho_0\varrho_1} + c_0 \frac{\varrho_1}{\varrho_0}}{\varrho_0 c_1 + \varrho_1 c_0} \geq \frac{c_1(1+c_0) + c_0(1+c_1)}{\varrho_0(c_1 + \frac{c_0}{1+c_0})} \overset{c_1=0}{\geq} \frac{1+c_0}{\varrho_0} \overset{c_0=0}{\geq} \frac{1}{\varrho_0} = m(0).$$

Finally the other case is obtained in in a similar manner. □

Remark 6.4. Note that the constants c_{01} and c_{10} can be rewritten among others as

$$c_{01} = \int (\mu_1\mu_0 - 1)^2 \,\mathrm{d}\mu_0 = \int \frac{(\mu_1 - \mu_0)^2}{\mu_0} \mathrm{d}x = \int \frac{\mu_1^2}{\mu_0^2} \mathrm{d}\mu_0 - 1 = \int \frac{\mu_1^2}{\mu_0} \mathrm{d}x - 1$$

and similarly for c_{10}. This distance is known as χ^2*-distance* on the space of probability measures (cf. [GS02]). The χ^2-distance is a rather weak distance. It bounds many other probability distances. Among them is also the relative entropy. Therefore, we note that $\log$ is a concave function and apply Jensen's inequality to the definition of the relative entropy

$$\mathrm{Ent}_{\mu_0}\left(\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right) = \int \mu_1 \log \frac{\mu_1}{\mu_0} \mathrm{d}x \leq \log\left(\int \frac{\mu_1^2}{\mu_0} \mathrm{d}x\right) = \log(1 + c_{01}) \leq c_{01},$$

where the last inequality is the estimate $\log(x) \leq x - 1$. We will also observe, that the χ^2-distance already becomes infinite for two centered Gaussian with covariance matrices differing by a factor bigger than two (cf. (6.13)).

Remark 6.5. From the proof of Theorem 6.3 we find that the expression for $\frac{1}{\varrho}$ in the last case of (6.4) can be bounded by

$$\max\left\{\frac{1}{\varrho_0}, \frac{1}{\varrho_1}\right\} \leq \frac{pc_{10} + pqc_{01}c_{10} + qc_{01}}{p\varrho_0 c_{10} + q\varrho_1 c_{01}} \leq \max\left\{\frac{1+qc_{01}}{\varrho_0}, \frac{1+pc_{10}}{\varrho_1}\right\}. \tag{6.6}$$

Remark 6.6. If we are in the symmetric case, where $\varrho_0 = \varrho_1 = \tilde{\varrho}$ and $c_{01} = c_{10} = c$, then we obtain

$$\frac{1}{\varrho_p} = \frac{1 + cpq}{\tilde{\varrho}}.$$

Corollary 6.7. *If $\mu_0 \ll \mu_1$, where μ_0 and μ_1 satisfy $\mathrm{PI}(\varrho_0)$ and $\mathrm{PI}(\varrho_1)$, then for all $p, q \geq 0$ with $p + q = 1$ the mixture measure $\mu = p\mu_0 + q\mu_1$ satisfies $\mathrm{PI}(\varrho_p)$ with*

$$\frac{1}{\varrho_p} = \max\left\{\frac{1}{\varrho_0}, \frac{1 + pc_{10}}{\varrho_1}\right\}.$$

Likewise, if $\mu_1 \ll \mu_0$, then we obtain the bound

$$\frac{1}{\varrho_p} = \max\left\{\frac{1}{\varrho_1}, \frac{1 + qc_{01}}{\varrho_0}\right\}.$$

Proof. The result is obtained by considering the $\max$ of the right-hand side in (6.5). There we let $\eta \to \infty$ as well as set $\vartheta = 0$ to find the conclusion. □

6.2. Logarithmic Sobolev inequality

In Section 2.3 we derived in Theorem 2.18 an estimate for the entropy of a general mixture with finite many components. If we only consider the mixture μ_p consisting of two components μ_0 and μ_1, then (2.23) simplifies to

$$\begin{aligned}\mathrm{Ent}_{\mu_p}(f^2) &\leq p\,\mathrm{Ent}_{\mu_0}(f^2) + q\,\mathrm{Ent}_{\mu_1}(f^2)\\ &\quad + \frac{pq}{\Lambda(p,q)}\left(\mathrm{var}_{\mu_0}(f) + \mathrm{var}_{\mu_1}(f) + (\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f))^2\right).\end{aligned} \tag{6.7}$$

Now, the right-hand side of (6.7) consists of quantities we can estimate under the assumption that μ_0 and μ_1 additionally satisfy $\mathrm{LSI}(\alpha_0)$ and $\mathrm{LSI}(\alpha_1)$. The following theorem provides an extension of the result [CM10, Theorem 4.4] to the multidimensional case for the logarithmic Sobolev inequality in the case that μ_0 and μ_1 are absolutely continuous to each other.

Theorem 6.8 (LSI for absolutely continuous mixtures)**.** *Assume that μ_0 and μ_1 satisfy $\mathrm{LSI}(\alpha_0)$ resp. $\mathrm{LSI}(\alpha_1)$ and are absolutely continuous to each other, then for all $p, q \geq 0$ with $p + q = 1$ the mixture measure $\mu_p = p\mu_0 + q\mu_1$ satisfies $\mathrm{LSI}(\alpha_p)$ with*

$$\frac{1}{\alpha_p} = \begin{cases} \frac{1+qL_p}{\alpha_0} & , \frac{\alpha_1}{\alpha_0} \geq 1 + pL_p(1 + c_{10}(1 + qL_p)) \\ \frac{1+pL_p}{\alpha_1} & , \frac{\alpha_0}{\alpha_1} \geq 1 + qL_p(1 + c_{01}(1 + pL_p)) \\ \frac{p(1+qL_p)c_{10} + pqL_pc_{01}c_{10} + q(1+pL_p)c_{01}}{p\alpha_0c_{10} + q\alpha_1c_{01}} & , \textit{else}, \end{cases}$$

where

$$c_{01} = \mathrm{var}_{\mu_0}\left(\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right) \quad \textit{and} \quad c_{10} = \mathrm{var}_{\mu_1}\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_1}\right)$$

The abbreviation L_p is used for the inverse logarithmic mean

$$L_p := \frac{1}{\Lambda(p,q)} = \frac{\log p - \log q}{p-q}.$$

Proof. The starting point is the splitting obtained from (6.7). We can estimate the variances and mean-difference in (6.7) in the same way as in the proof (6.5) of Theorem 6.3 by noting that $\mathrm{LSI}(\alpha)$ implies $\mathrm{PI}(\alpha)$

$$\begin{aligned}
&\mathrm{Ent}_{\mu_p}(f^2) \\
&\leq \frac{1}{\alpha_0}\left(1 + qL_p(1+(1+\eta)\vartheta^2 c_{01}\right)\int |\nabla f|^2 \, p\mathrm{d}\mu_0 \\
&\quad + \frac{1}{\alpha_1}\left(1 + pL_p(1+(1+\tfrac{1}{\eta})(1-\vartheta)^2 c_{10}\right)\int |\nabla f|^2 \, q\mathrm{d}\mu_1 \\
&\leq \max\left\{\frac{1+qL_p(1+(1+\eta)\vartheta^2 c_{01})}{\alpha_0}, \frac{1+pL_p(1+(1+\frac{1}{\eta})(1-\vartheta)^2 c_{10})}{\alpha_1}\right\}\int |\nabla f|^2 \,\mathrm{d}\mu_p.
\end{aligned}$$

Let us introduce the reduced logarithmic Sobolev constants

$$\tilde{\alpha}_0 = \frac{\alpha_0}{1+qL_p} \qquad \text{and} \qquad \tilde{\alpha}_1 = \frac{\alpha_1}{1+pL_p}.$$

Furthermore, we define the constants $\tilde{c}_0$ and $\tilde{c}_1$ as

$$\tilde{c}_{01} = \frac{c_{01}L_p}{1+qL_p} \qquad \text{and} \qquad \tilde{c}_{10} = \frac{c_{10}L_p}{1+pL_p}.$$

Therewith, the bound takes the form

$$\mathrm{Ent}_{\mu_p}(f^2) \leq \max\left\{\frac{1+(1+\eta)\vartheta^2\tilde{c}_0}{\tilde{\alpha}_0}, \frac{1+(1+\frac{1}{\eta})(1-\vartheta)^2\tilde{c}_1}{\tilde{\alpha}_1}\right\}\int |\nabla f|^2 \,\mathrm{d}\mu_p. \tag{6.8}$$

The estimate (6.8) has the same structure as the estimate (6.5), where $\tilde{\alpha}_i$ now plays the role of ϱ_i and $\tilde{c}_{01}$, $\tilde{c}_{10}$ the role of c_{01}, c_{10}. Hence, we can use the optimization procedure from the proof of Theorem 6.3. The last step consists of translating the constants $\tilde{\alpha}_i$ and $\tilde{c}_{01}$, $\tilde{c}_{10}$ back to the original ones. □

Remark 6.9. The inverse logarithmic mean $L_p = \frac{1}{\Lambda(p,q)}$ logarithmically blows up for $p \to \{0,1\}$.

Remark 6.10. If $c_{01} = c_{10} = \tilde{c}$ and $\alpha_0 = \alpha_1 = \tilde{\alpha}$ holds the bound

$$\frac{1}{\alpha_p} \leq \frac{1+pqL_p(\tilde{c}+2)}{\tilde{\alpha}}.$$

Corollary 6.11. *If $\mu_0 \ll \mu_1$, where μ_0 and μ_1 satisfy $\mathrm{LSI}(\alpha_0)$ resp. $\mathrm{LSI}(\alpha_1)$, then for all $p, q \geq 0$ with $p+q=1$ the mixture measure $\mu = p\mu_0 + q\mu_1$ satisfies $\mathrm{LSI}(\alpha_p)$ with*

$$\frac{1}{\alpha_p} \leq \max\left\{\frac{1+qL_p}{\alpha_0}, \frac{1+pL_p(1+c_{10})}{\alpha_1}\right\} \tag{6.9}$$

Likewise, if $\mu_1 \ll \mu_0$, then we obtain the bound

$$\frac{1}{\alpha_p} \leq \max\left\{\frac{1+pL_p}{\alpha_1}, \frac{1+qL_p(1+c_{01})}{\alpha_0}\right\} \tag{6.10}$$

Remark 6.12. Note that in this case the best one can get is a logarithmically blow-up for $p \to \{0,1\}$. The following examples will show, that one of the blow-ups in the $\max$ will be artificial and with the help of the Bakry-Émery criterion and the perturbation lemma of Holley-Stroock one can rule out one case and show that the logarithmic Sobolev constant for the mixture will be bounded for $p \to 0$ in (6.9) resp. for $p \to 1$ in (6.10).

6.3. Examples

We want to compare the results from Theorem 6.3 and Theorem 6.8 with the ones obtained in [CM10, Section 4.5] for some specific examples. In [CM10] was observed for specific examples, that the Poincaré constant can stay bound, whereas the logarithmic Sobolev constant logarithmically blows up in the mixture ratio p going to zero or one. The proof of the upper bound relies on the functional (6.3) from Remark 6.2 and the lower bound is obtained via the Bobkov-Götze functional (cf. Section 3.3). The simple criterions of Theorem 6.3 and Theorem 6.8 can only give upper bounds for the multi-dimensional case, when the at least one of the mixture component is absolutely continuous to the other. However, it is still possible to obtain the optimal results in terms of scaling in the mixture parameter.

6.3.1. Mixture of two Gaussian measures with equal covariance matrix

Let us consider the mixtures of two Gaussians $\mu_0 := \mathcal{N}(0,\Sigma)$ and $\mu_1 := \mathcal{N}(y,\Sigma)$, for some $y \in \mathbb{R}^n$ and $\Sigma \geq \sigma\,\mathrm{Id}$ with $\sigma > 0$ a strictly positive definite covariance matrix. Then, μ_0 and μ_1 satisfy $\mathrm{PI}(\sigma)$ and $\mathrm{LSI}(\sigma)$ by the Bakry-Émery criterion (cf. Theorem 3.1), i.e. $\varrho_0 = \alpha_0 = \varrho_1 = \alpha_1 = \frac{1}{\sigma}$. Further, we can explicitly calculate the χ^2-distance between μ_0 and μ_1

$$\begin{aligned} c_{01} = c_{10} &= \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det\Sigma}} \int e^{-x\Sigma^{-1}x} e^{\frac{1}{2}(x-y)\Sigma^{-1}(x-y)} \mathrm{d}x - 1 \\ &= e^{y\Sigma^{-1}y} \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det\Sigma}} \int e^{-\frac{1}{2}(x+y)\Sigma^{-1}(x+y)} \mathrm{d}x - 1 = e^{|\Sigma^{-\frac{1}{2}}y|^2} - 1 \leq e^{\frac{|y|^2}{\sigma}} - 1. \end{aligned}$$

By using Remark 6.6, we obtain

$$\frac{1}{\varrho_p} \leq \left(1 + pq\left(e^{\frac{|y|^2}{\sigma}} - 1\right)\right)\sigma. \tag{6.11}$$

Likewise, the logarithmic Sobolev inequality follows from Remark 6.10

$$\frac{1}{\alpha_p} \leq \left(1 + pqL_p\left(e^{\frac{|y|^2}{\sigma}} + 1\right)\right)\sigma.$$

By noting that $pq \le pqL_p \le \frac{1}{4}$, both constants stay uniformly bounded in p.

In [CM10, Corollary 4.7] is deduced the following bound for $\frac{1}{\varrho_p}$ for the mixture of two one-dimensional standard Gaussians, i.e. $\sigma = 1$ in (6.11) and we set $a = |y|$

$$\frac{1}{\varrho_{p,\mathtt{CM}}} \le 1 + pqa^2 \left(\Phi(a) e^{a^2} + \frac{a}{\sqrt{2\pi}} e^{\frac{a^2}{2}} + \frac{1}{2} \right), \tag{6.12}$$

where $\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{y^2}{2}} \mathrm{d}y$. The bounds (6.11) and (6.12) are of the same exponential order e^{a^2} in a and only differ on the pre-exponential scale, where (6.11) gives a better bound than (6.12). This example show, that in the case where μ_0 and μ_1 are absolutely continuous to each others as well as the tail behavior of the both measures is the same, then Theorem 6.3 and Theorem 6.8 give good results and generalize the bound of [CM10] to the multidimensional case.

6.3.2. Mixture of a Gaussian and sub-Gaussian measure

Let us consider $\mu_1 = \mathcal{N}(0, \Sigma)$ where $\Sigma \ge \sigma \,\mathrm{Id}$ is strictly positive definite and let μ_0 be such that for some $\kappa \ge 1$ the densities satisfy $\mu_0 \le \kappa \mu_1$ in the pointwise sense. Thus, by the Bakry-Émery criterion (cf. Theorem 3.1), we have $\varrho_1 = \alpha_1 = \frac{1}{\sigma}$. Further, as upper bound for c_{10} we find

$$c_{10} = \mathrm{var}_{\mu_1}\left(\frac{\mu_0}{\mu_1}\right) = \int \left(\frac{\mu_0}{\mu_1}\right)^2 \mathrm{d}\mu_1 - 1 \le \kappa^2 - 1.$$

Hence, an application of Corollary 6.7 for the Poincaré constant leads to the estimate

$$\frac{1}{\varrho_p} \le \max\left\{ \frac{1}{\varrho_0}, (1 + p(\kappa^2 - 1))\sigma \right\}.$$

Similarly, Corollary 6.11 gives the following bound for the logarithmic Sobolev constant of the mixture measure μ_p

$$\frac{1}{\alpha_p} \le \max\left\{ \frac{1 + qL_p}{\alpha_0}, (1 + pL_p\kappa^2)\sigma \right\}.$$

Note that the logarithmic Sobolev constant blows up logarithmically for $p \to \{0, 1\}$, which is a consequence of the the missing information on μ_0.

6.3.3. Mixture of two Gaussians with equal mean

We consider $\mu_0 = \mathcal{N}(0, \mathrm{Id})$ and $\mu_1 = \mathcal{N}(0, \sigma\,\mathrm{Id})$. Then, we have by the Bakry-Émery criterion (cf. Theorem 3.1) $\varrho_0 = \alpha_0 = 1$ and $\varrho_1 = \alpha_1 = \frac{1}{\sigma}$. For calculating the χ^2-distance, we can use the spherical symmetry and arrive at the one dimensional integral

$$c_{01} = \int \frac{\mu_1^2}{\mu_0} \mathrm{d}x - 1 = \frac{\mathcal{H}^{n-1}(S^{n-1})}{(2\pi)^{\frac{n}{2}}\sigma^n} \int_{\mathbb{R}^+} r^{n-1} e^{-\left(\frac{1}{\sigma} - \frac{1}{2}\right) r^2} \mathrm{d}r - 1.$$

The integral exists only for $\sigma < 2$. In this case, we can evaluate the above integral and simplify the resulting expression. Furthermore, the constant c_{10} is given by the duality $\sigma \mapsto \sigma^{-1}$. Hence, we arrive at

$$c_{01} = \begin{cases} \frac{1}{(\sigma(2-\sigma))^{\frac{n}{2}}} - 1 & ,\sigma < 2 \\ +\infty & ,\sigma \geq 2 \end{cases} \qquad \text{and} \qquad c_{10} = \begin{cases} \frac{1}{(\sigma^{-1}(2-\sigma^{-1}))^{\frac{n}{2}}} - 1 & ,\sigma > \frac{1}{2} \\ +\infty & ,\sigma \leq \frac{1}{2} \end{cases}. \tag{6.13}$$

If $\sigma \leq \frac{1}{2}$, i.e. when $c_{10} = \infty$, we can only employ the bound given in Corollary 6.7

$$\frac{1}{\varrho_p} \leq \max\{\sigma, 1 + qc_{01}\} = \max\left\{\sigma, (1-q) + \frac{q}{(\sigma(2-\sigma))^{\frac{n}{2}}}\right\} = p + \frac{q}{(\sigma(2-\sigma))^{\frac{n}{2}}}$$

Similarly, if $\sigma^2 \geq 2$, i.e. when $c_{01} = \infty$, we obtain

$$\frac{1}{\varrho_p} \leq \max\{1, (1 + pc_{10})\sigma\} \leq \sigma\left(q + \frac{p}{(\sigma^{-1}(2-\sigma^{-1}))^{\frac{n}{2}}}\right).$$

If $\frac{1}{2} < \sigma < 2$, we could apply the interpolation bound of Theorem 6.3. However, to see the scaling behavior better, we use the estimate (6.6) of Remark 6.5, where we use the symmetry under $\sigma \mapsto \frac{1}{\sigma}$ to deduce the bound

$$\frac{1}{\varrho_p} \leq \begin{cases} p + \frac{q}{(\sigma(2-\sigma))^{\frac{n}{2}}} & ,\sigma \leq 1 \\ \sigma\left(q + \frac{p}{(\sigma^{-1}(2-\sigma^{-1}))^{\frac{n}{2}}}\right) & ,\sigma \geq 1 \end{cases} \tag{6.14}$$

Likewise, by applying Corollary 6.11, we obtain the bound for the logarithmic Sobolev constant

$$\frac{1}{\alpha_p} \leq \begin{cases} 1 + \frac{qL_p}{(\sigma(2-\sigma))^{\frac{n}{2}}} & ,\sigma \leq 1 \\ \sigma\left(1 + \frac{pL_p}{(\sigma^{-1}(2-\sigma^{-1}))^{\frac{n}{2}}}\right) & .\sigma \geq 1 \end{cases} \tag{6.15}$$

The bound (6.15) blows up logarithmically for $p \to \{0, 1\}$. However, the case $\sigma = 1$ allows the combined bound $\frac{1}{\alpha_p} \leq 1 + \min\{p, q\} L_p$, which stays bonded, this behavior could be extended to the range $\sigma \in (\frac{1}{2}, 2)$ thanks to (6.13) and the interpolation bound of Theorem 6.8.

Again we want to compare the result (6.14) with the ones of [CM10, Section 4.5.2.], which states that for some $C > 0$ and $\sigma > 1$

$$\frac{1}{\varrho_{p,\mathtt{CM}}} \leq \sigma + Cp^{\frac{1}{\sigma-1}}.$$

In general, depending on the constant C the bound (6.11) is better for σ small, whereas the scaling in σ is better for (6.12), namely linear instead of $\sigma^{\frac{3}{2}}$ as in (6.11).

6.3.4. Mixture of uniform and Gaussian measure

Let us consider $\mu_0 = \mathcal{N}(0,1)$ and $\mu_1 = \mathcal{U}(B_1)$, where B_1 is the unit ball around zero. We have $\varrho_0 = 1$ and $\varrho_1 \geq \frac{\pi^2}{\operatorname{diam}(B_1)^2} = \frac{\pi^2}{4}$ by the result of [PW60]. Further, it holds $\mu_1 \ll \mu_0$ and we can calculate

$$c_{01} + 1 = \int \left(\frac{\mu_1}{\mu_0}\right)^2 \mathrm{d}\mu_0 = \frac{1}{\mathcal{H}^n(B_1)^2} \int_{B_1} \frac{1}{\mu_0} \mathrm{d}x = \frac{(2\pi)^{\frac{n}{2}} \mathcal{H}^{n-1}(\partial B_1)}{\mathcal{H}^n(B_1)^2} \int_0^1 r^{n-1} e^{\frac{r^2}{2}} \mathrm{d}r. \tag{6.16}$$

The volume $\mathcal{H}^n(B_1)$ and the surface area $\mathcal{H}^{n-1}(\partial B_1)$ of the n-sphere satisfy the following relations

$$\frac{\mathcal{H}^{n-1}(\partial B_1)}{\mathcal{H}^n(B_1)} = n \qquad \text{and} \qquad \frac{(2\pi)^{\frac{n}{2}}}{\mathcal{H}^n(B_1)} = 2^{\frac{n}{2}} \Gamma\left(\frac{n}{2} + 1\right) =: g_n. \tag{6.17}$$

The integral on the right-hand side in (6.16) can be bounded below by $\frac{1}{n}$ and above by $\frac{\sqrt{e}}{n}$. Hence, we found the bound

$$g_n \leq c_{01} + 1 \leq \sqrt{e} g_n.$$

Poincaré inequality: We apply Corollary 6.7 to find

$$\frac{1}{\varrho_p} \leq \max\left\{\frac{1}{\varrho_1}, 1 + q c_{01}\right\} \leq p + q\sqrt{e} g_n, \tag{6.18}$$

where we note in the last inequality, that $\frac{4}{\pi^2} \leq p + q\sqrt{e} g_n$ for $n \geq 1$ and all $p \in [0,1]$. We only get an upper uniform in p.

Logarithmic Sobolev inequality: Let us begin by noting that $\alpha_0 = 1$ by the Bakry-Émery criterion and $\alpha_1 \geq \frac{2}{e}$, which is a consequence of a combination of the Bakry-Émery criterion (cf. Theorem 3.1) with the Hooley-Stroock perturbation principle (cf. Theorem 3.2) and an optimization[1]. We use Corollary 6.11 and obtain the bound

$$\frac{1}{\alpha_p} \leq \max\left\{\frac{1 + pL_p}{\alpha_1}, \frac{1 + qL_p(1 + c_{01})}{\alpha_0}\right\} \leq \max\left\{\frac{(1 + pL_p)e}{2}, 1 + qL_p\sqrt{e} g_n)\right\}. \tag{6.19}$$

Again we have a logarithmically blow-up of the bound for $p \to \{0,1\}$. Let us show that the blow-up for $p \to 1$ is artificial.

Comparison with Bakry-Émery and Holley-Stroock: We want to decompose the Hamiltonian of μ_p into a convex function and some error term. Therefore, we can write

$$\begin{aligned} H_p(x) &:= -\log \mu_p(x) = -\log\left(\frac{p}{(2\pi)^{\frac{n}{2}}} e^{-\frac{|x|^2}{2}} + \frac{1-p}{\omega_n} \chi_{B_1(0)}(x)\right) \\ &= -\log\left(e^{-\frac{|x|^2}{2} + \frac{1}{2}} + \frac{1-p}{p} \frac{(2\pi)^{\frac{n}{2}}}{\omega_n} \sqrt{e} \chi_{B_1(0)}(x)\right) + C_{p,n} \\ &= \frac{|x|^2 - 1}{2} - \psi_p(x) + \tilde{C}_{p,n}, \end{aligned} \tag{6.20}$$

[1]Therefore, we compare μ_1 with the measure $\nu_\lambda(x) = \frac{1}{Z_\lambda} \exp\left(-\lambda |x|^2 + \frac{\lambda}{2}\right)$ on B_1 with $\lambda > 0$. Then, one can easy check, that $\operatorname{osc}_{x \in B_1} |\mu_1(x) - \nu_\lambda(x)| = \frac{\lambda}{2}$ and that ν_λ satisfies LSI(2λ), hence μ_1 satisfies LSI($2\lambda e^{-\lambda}$) for all $\lambda > 0$. Optimizing the expression $2\lambda e^{-\lambda}$ in λ let us conclude that μ_1 satisfies LSI($\frac{2}{e}$).

where

$$\psi_p(x) := \left(\log \left(e^{-\frac{|x|^2}{2}+\frac{1}{2}} + \frac{1-p}{p} \frac{(2\pi)^{\frac{n}{2}}}{\omega_n} \sqrt{e} \right) + \frac{|x|^2-1}{2} \right) \chi_{B_1(0)}(x).$$

The function ψ_p is radially monotone towards the boundary of B_1 and we find the bound

$$0 \leq \psi_p(x) \leq \log \left(1 + \frac{1-p}{p} \frac{(2\pi)^{\frac{n}{2}}}{\omega_n} \sqrt{e} \right) = \psi_p(x)|_{\partial B_1(0)}. \tag{6.21}$$

From (6.20) we compare H_p with the convex potential $\frac{|x|^2-1}{2}$ and use the bound (6.21) on the perturbation ψ_p to find that μ_p satisfies $\mathrm{PI}(\tilde{\varrho}_p)$ and $\mathrm{LSI}(\tilde{\alpha}_p)$ with

$$\frac{1}{\tilde{\varrho}_p} \leq \frac{1}{\tilde{\alpha}_p} \leq 1 + \frac{1-p}{p} \sqrt{e}\, g_n, \tag{6.22}$$

where g_n is the same constant as in (6.17). Now, this bound only blows up for $p \to 0$. But the blow-up is like $\frac{1}{p}$. Furthermore, we have not such a detailed information on the Poincaré constant as in (6.18) and can not detect the different behavior, therefore we have to combine both approaches.

Combination: A combination of the bound obtained in (6.19) with the bound from (6.22) results in the improved bound

$$\frac{1}{\alpha} \leq C_n(1 + qL_p g_n), \quad \text{with } C_n \text{ some universal constant,} \tag{6.23}$$

which only logarithmically blows up for $p \to 0$.

Conclusion: In this example as a consequence the Poincaré constant and logarithmic Sobolev constant may have different scaling behavior for $p \to 0$. Indeed, [CM10] show for this specific mixture in the one dimensional case that the logarithmic Sobolev constant can be bounded below by

$$C\, |\log p| \leq \frac{1}{\alpha},$$

for p small enough and a constant C independent of p. In one dimension, lower bounds are accessible via the Bobkov-Götze functional (cf. Section 3.3). Hence the bound (6.23) is optimal in the one dimensional case, which strongly indicates also optimality for higher dimension in terms of scaling in the mixture ration p.

To conclude, we have the following rule of thumb: On the one hand, the Bakry-Émery criterion in combination with the Hooley-Stroock perturbation principle is effective for detecting blow-ups of the logarithmic Sobolev constant for mixtures, but has in general the wrong scaling behavior in the mixing parameter p. On the other hand, the criterion presented in Theorem 6.8 provides the right scaling of the blow-up, but also sees artificial blow-ups, if the components of the mixture become singular in the sense of the χ^2-distance.

6.4. Outlook: Modified weighted transport distance

As already pointed out in the beginning of the chapter, the presented approach to mixtures was the prelude to the Eyring-Kramers formula. The discovery of the weighted transport distance presented in Chapter 4 and especially the modified weighted transport distance from Remark 4.22 allows to get estimates of the mean-difference, which can also be applied to mixtures. We define for an absolutely continuous measure μ and $\nu_i \ll \mu$ for $i = 0, 1$ the modified weighted transport distance $\widetilde{\mathcal{T}}_\mu(\nu_0, \nu_1)$ by

$$\widetilde{\mathcal{T}}_\mu(\nu_0, \nu_1) := \inf_{(\Phi_s)_{s\in[0,1]}} \int \left| \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1} \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \right|^2 \mathrm{d}\mu, \tag{6.24}$$

where $(\Phi_s)_{s\in[0,t]}$ is a transport interpolation from ν_0 to ν_1 absolutely continuous in s.

Remark 6.13. In the one-dimensional case, we observe by using the *Brenier-Banamou* formula for ν_s of Remark 4.23, that ν_s solves the conservation law

$$\partial_s \nu_s(y) + \partial_y \left(\dot{\Phi}_s \circ \Phi_s^{-1}(y) \nu_s(y) \right) = 0 \tag{6.25}$$

An integration of (6.25) in s on $(0, 1)$ and in y on $(-\infty, x)$ results in

$$\int_{-\infty}^x \nu_1(y)\,\mathrm{d}y - \int_{-\infty}^x \nu_0(y)\,\mathrm{d}y + \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1}(x)\, \nu_s(x)\,\mathrm{d}s = 0.$$

Hence, we observe that the inner integral of (4.53) does not depend on the specific transport interpolation, but only on the difference of the distribution functions $F_i := \int_{-\infty}^x \nu_i(y)\,\mathrm{d}y$, i.e. we obtain the representation

$$\widetilde{\mathcal{T}}_\mu(\nu_0, \nu_1) = \int \frac{(F_1(x) - F_0(x))^2}{\mu(x)} \mathrm{d}x. \tag{6.26}$$

We observe that the functional (6.26) is nothing else than the function $I(p)$ of (6.3) used by [CM10, Theorem 4.4]. Hence, one can regard the weighted transport distance (6.24) as the multidimensional generalization.

Appendix A

Properties of the logarithmic mean Λ

In this part of the appendix, we collect some properties of the logarithmic mean $\Lambda(\cdot,\cdot)$. Let us start with a collection of some essential properties for this chapter. A more complete study can be found in [Car72] and the recent review [Bha08].

Let us first recall the definition of $\Lambda(\cdot,\cdot) : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$

$$\Lambda(a,b) = \frac{a-b}{\log a - \log b}, \quad a \neq b \qquad \text{and} \qquad \Lambda(a,a) = a.$$

The value of $\Lambda(a,b)$ can also be defined by the logarithmic average of a and b

$$\Lambda(a,b) = \int_0^1 a^s\, b^{1-s}\, \mathrm{d}s = \frac{1}{\log a - \log b}\ \left[a^s b^{1-s}\right]_{s=0}^1 . \tag{A.1}$$

The equation (A.1) justifies the statement, that $\Lambda(\cdot,\cdot)$ is a mean, since one immediately recovers the simple bounds $\min\{a,b\} \leq \Lambda(a,b) \leq \max\{a,b\}$. Furthermore, the following representations hold for $1/\Lambda(\cdot,\cdot)$

$$\frac{1}{\Lambda(a,b)} = \int_0^1 \frac{\mathrm{d}\tau}{\tau a + (1-\tau)b} = \int_0^\infty \frac{\mathrm{d}\tau}{(a+\tau)(b+\tau)} \tag{A.2}$$

Some immediate properties are:

- $\Lambda(\cdot,\cdot)$ is symmetric
- $\Lambda(\cdot,\cdot)$ is homogeneous of degree one, i.e. for $\lambda > 0$ holds $\Lambda(\lambda a, \lambda b) = \lambda\Lambda(a,b)$.

The derivatives of $\Lambda(\cdot,\cdot)$ are given by straight-forward calculus

$$\partial_a \Lambda(a,b) = \frac{1}{\log a - \log b}\left(1 - \frac{\Lambda(a,b)}{a}\right) > 0 \qquad \text{and}$$
$$\partial_b \Lambda(a,b) = \frac{1}{\log b - \log a}\left(1 - \frac{\Lambda(a,b)}{b}\right) > 0.$$

Hence $\Lambda(\cdot,\cdot)$ is strictly monotone increasing in both arguments.

The following result is almost classical.

Lemma A.1. *The logarithmic mean can be bounded below by the geometric mean and above by the arithmetic mean*

$$\sqrt{ab} \leq \Lambda(a,b) \leq \frac{a+b}{2}, \tag{A.3}$$

with equality if and only if $a = b$.

There exists at least four proofs of the inequality A.3

- [Car72, Theorem 1] uses the representation (A.2)
- [Mie11, Appendix A] starts with (A.1) and uses the convexity of $s \mapsto a^s b^{1-s}$
- [Bha08] gives an argument by simple calculus.
- Again [Bha08] relates the terms in question to hyperbolic trigonometric functions, which allow for a quantification of the error, in the case with no equality. We will present his proof here.

Proof. Since w.l.o.g. $a, b > 0$, we can switch to exponential variables and set $a = e^x$ and $b = e^x$. Therewith we arrive for the quotient of geometric and logarithmic mean at

$$\frac{\sqrt{ab}}{\Lambda(a,b)} = e^{\frac{x+y}{2}} \frac{x-y}{e^x - e^y} = \frac{\frac{x-y}{2}}{e^{\frac{x-y}{2}} - e^{\frac{y-x}{2}}} = \frac{\frac{x-y}{2}}{\sinh\left(\frac{x-y}{2}\right)}. \tag{A.4}$$

It is easy to verify, that the function $t \mapsto \frac{t}{\sinh t}$ is symmetric and strictly decreasing in $|t|$, hence it has a unique maximum for $t = 0$ with 1. This proves $\sqrt{ab} \leq \Lambda(a,b)$ with equality only if $a = b$.
By the same reasoning, we obtain for the quotient of arithmetic and logarithmic mean in exponential variables

$$\frac{\frac{a+b}{2}}{\Lambda(a,b)} = \frac{\frac{x-y}{2}}{\tanh\left(\frac{x-y}{2}\right)}.$$

Again, one can check that the function $t \mapsto \frac{t}{\tanh t}$ is symmetric and strictly increasing in $|t|$, hence it has a unique minimum for $t = 0$ with value 1. This proves $\frac{a+b}{2} \geq \Lambda(a,b)$ with equality only if $a = b$. □

The bounds in (A.3) are good, if a is of the same order as b, whereas the following bound is particular good if $\frac{a}{b}$ becomes very small or very large.

Lemma A.2. *It holds for $p \in (0,1)$ the following bound*

$$\frac{\Lambda(p,1-p)}{p(1-p)} < \min\left\{\frac{1}{p\log\frac{1}{p}}, \frac{1}{(1-p)\log\frac{1}{1-p}}\right\}. \tag{A.5}$$

Proof. Let us first consider the case $0 < p < \frac{1}{2}$. Then, it is enough to show, that

$$\frac{\Lambda(p,1-p)}{p(1-p)}\, p \log \frac{1}{p} = \frac{(1-2p)\log\frac{1}{p}}{(1-p)\log\frac{1-p}{p}} \overset{!}{<} 1. \tag{A.6}$$

This follows easily from the following lower bound on the denominator

$$(1-p)\log\frac{1-p}{p} = (1-2p)\log\frac{1}{p} + p\log\frac{1}{p} - (1-p)\log\frac{1}{1-p} > (1-2p)\log\frac{1}{p},$$

since $p\log\frac{1}{p} - (1-p)\log\frac{1}{1-p} > 0$ for $0 < p < \frac{1}{2}$. The case $\frac{1}{2} < p < 1$ follows by symmetry under the variable change $p \mapsto 1-p$. It remains to check the case $p = \frac{1}{2}$. The left-hand side of (A.6) evaluates for $p = \frac{1}{2}$ to

$$\lim_{p\to\frac{1}{2}} \frac{\Lambda(p,1-p)}{p(1-p)}\, p\log\frac{1}{p} = \log 2 < 1.$$

□

The logarithmic mean also occurs in the following optimization problem, which appears in the proof of the optimality of the Eyring-Kramers formula for the logarithmic Sobolev constant in one dimension (cf. Section 2.5).

Lemma A.3. *For* $p \in (0,1)$ *holds*

$$\min_{t\in(0,1)} \frac{\left(\sqrt{\frac{t}{p}} - \sqrt{\frac{1-t}{1-p}}\right)^2}{t\log\frac{t}{p} + (1-t)\log\frac{1-t}{1-p}} = \frac{\Lambda(p,1-p)}{p(1-p)}. \tag{A.7}$$

The minimum in (A.7) *is attained for* $t = 1-p$.

Proof. Let us introduce the function $f_p : (0,1) \to \mathbb{R}$ and $g_p : (0,1) \to \mathbb{R}$ given by the nominator and denominator of (A.7)

$$f_p(t) := \left(\sqrt{\tfrac{t}{p}} - \sqrt{\tfrac{1-t}{1-p}}\right)^2 \qquad \text{and} \qquad g_p(t) := t\log\tfrac{t}{p} + (1-t)\log\tfrac{1-t}{1-p}.$$

It is easy to verify, that the following relations for the derivatives hold true

$$f_p'(t) = \left(\sqrt{\tfrac{t}{p}} - \sqrt{\tfrac{1-t}{1-p}}\right)\left(\tfrac{1}{\sqrt{tp}} + \tfrac{1}{\sqrt{(1-p)(1-t)}}\right), \qquad g_p'(t) = \log\tfrac{t}{p} - \log\tfrac{1-t}{1-p}, \tag{A.8}$$

$$f_p''(t) = \sqrt{\tfrac{(1-t)t}{(1-p)p}}\ \tfrac{1}{2(1-t)^2t^2} > 0, \qquad g_p''(t) = \tfrac{1}{(1-t)t} > 0.$$

Hence, both functions f_p an g_p are strictly convex and have a unique minimum for $t = p$, where they are both zero. The derivative of the quotient of f_p and g_p has the form

$$h_p'(t) := \left(\frac{f_p(t)}{g_p(t)}\right)' = \frac{1}{g_p^2(t)}\left(f_p'(t)g_p(t) - f_p(t)g_p'(t)\right) \tag{A.9}$$

The representation (A.8) for g_p' leads to

$$\big(f_p'(t)g_p(t) - f_p(t)g_p'(t)\big) = \big(tf_p'(t) - f_p(t)\big)\log\tfrac{t}{p} + \big((1-t)f_p'(t) + f_p(t)\big)\log\tfrac{1-t}{1-p}. \quad \text{(A.10)}$$

Now, we can use (A.8) for f_p' to find

$$\begin{aligned} tf_p'(t) - f_p(t) &= \left(\sqrt{\tfrac{t}{p}} - \sqrt{\tfrac{1-t}{1-p}}\right)\left(\sqrt{\tfrac{t}{p}} + \tfrac{t}{\sqrt{(1-p)(1-t)}} - \sqrt{\tfrac{t}{p}} + \sqrt{\tfrac{1-t}{1-p}}\right) \\ &= \tfrac{1}{\sqrt{(1-p)(1-t)}}\left(\sqrt{\tfrac{t}{p}} - \sqrt{\tfrac{1-t}{1-p}}\right) \end{aligned} \quad \text{(A.11)}$$

and likewise

$$(1-t)f_p'(t) + f_p(t) = \tfrac{1}{\sqrt{tp}}\left(\sqrt{\tfrac{t}{p}} - \sqrt{\tfrac{1-t}{1-p}}\right). \quad \text{(A.12)}$$

Using (A.11) and (A.12) in (A.10) leads by (A.9) to

$$h_p'(t) = \frac{1}{g_p^2(t)}\underbrace{\left(\sqrt{\frac{t}{p}} - \sqrt{\frac{1-t}{1-p}}\right)}_{=:v_p(t)}\underbrace{\left(\frac{\log\frac{t}{p}}{\sqrt{(1-p)(1-t)}} + \frac{\log\frac{1-t}{1-p}}{\sqrt{tp}}\right)}_{=:w_p(t)}.$$

Since $g_p(p) = g_p'(p) = 0$ and $g_p''(p) > 0$, the function $\frac{1}{g_p^2(t)}$ has a pole of order 4 in $t = p$. Moreover, the function $v_p(t)$ has a simple zero in $t = p$. We have to do some more investigations for the function $w_p(t)$. First, we observe that $w_p(t)$ can be rewritten as

$$w_p(t) = \underbrace{\frac{t-p}{\sqrt{(1-t)t(1-p)p}}}_{=:\hat{w}_p(t)}\underbrace{\left(\frac{\sqrt{tp}\,\log\frac{t}{p}}{(t-p)} - \frac{\sqrt{(1-t)(1-p)}\,\log\frac{1-t}{1-p}}{(p-t)}\right)}_{:=\tilde{w}_p(t)}.$$

The function $\tilde{w}_p(t)$ can be expressed in terms of the logarithmic mean

$$\tilde{w}_p(t) = \frac{\sqrt{tp}}{\Lambda(t,p)} - \frac{\sqrt{(1-t)(1-p)}}{\Lambda(1-t,1-p)} \quad \text{(A.13)}$$

and is measuring the defect in the geometric-logarithmic mean inequality (A.3). Let us switch to exponential variables and set

$$x(t) := \log\sqrt{\frac{t}{p}} \quad \text{and} \quad y(t) := \log\sqrt{\frac{1-t}{1-p}}.$$

Note, that either $x(t) \le 0 \le y(t)$ for $t \le p$ or $y(t) \le 0 \le x(t)$ for $t \ge 0$ with equality only for $t = p$. Therewith, (A.13) becomes with the same argument as in (A.4)

$$\tilde{w}_p(t) = \frac{x(t)}{\sinh(x(t))} - \frac{y(t)}{\sinh(y(t))}.$$

By making use of the fact, that the function $x \mapsto \frac{x}{\sinh x}$ is symmetric, monotone decreasing in $|x|$ and has a unique maximum in 1, we can conclude that

$$\tilde{w}_p(t) = 0 \qquad \text{if and only if} \qquad x(t) = -y(t).$$

The solutions to the equation $x(t) = -y(t)$ are given for $t \in \{p, 1-p\}$. Let us first consider the case $t = p$, then $x(t) = y(t) = 0$ and $w_p(p)$ is a zero of order 2, since the function $x \mapsto \frac{x}{\sinh(x)}$ is strictly concave for $t = 0$. Now, we can go back to $h_p'(t)$ and argue with the representation

$$\lim_{t \to p} h_p'(t) = \lim_{t \to p} \frac{v_p(t)\hat{w}_p(t)\tilde{w}_p(t)}{g_p^2(t)} \overset{!}{\neq} 0.$$

This is a consequence of counting the zeros for $t = p$ in the nominator and denominator according to their order. For the denominator $g_p^2(p)$ is a zero of order 4. For the nominator we have $v_p(p)$ is a zero of order 1, $\hat{w}_p(p)$ is a zero of order 1 and $\tilde{w}_p(p)$ is a zero of order 2, which leads in total again to a zero of order 4 exactly compensating the zero of the denominator.
The other case is $t = 1-p$. Let us evaluate $h_p(1-p)$, which is given by

$$h_p(1-p) = \frac{\frac{1}{p(1-p)}\left(p-(1-p)\right)^2}{(1-p)\log\frac{1-p}{p} + p\log\frac{p}{1-p}} = \frac{1}{p(1-p)}\frac{(p-(1-p))^2}{(p-(1-p))\log\frac{p}{1-p}} = \frac{\Lambda(p,1-p)}{p(1-p)}.$$

Since, $t = 1-p$ is the only critical point of $h_p(t)$ inside $(0,1)$, it remains to check whether the boundary values are larger than $h_p(1-p)$. They are given by

$$\lim_{t \to 0} h_p(t) = \frac{1}{(1-p)\log\frac{1}{1-p}} \qquad \text{and} \qquad \lim_{t \to 1} h_p(t) = \frac{1}{p\log\frac{1}{p}}.$$

We observe, that the demanded inequality to be in a global minimum

$$h_p(1-p) = \frac{\Lambda(p,1-p)}{p(1-p)} \overset{!}{<} \min\left\{\frac{1}{p\log\frac{1}{p}}, \frac{1}{(1-p)\log\frac{1}{1-p}}\right\}$$

is just (A.5) of Lemma A.2. □

Appendix B

Gaussian integrals and linear algebra

B.1. Partial Gaussian integrals

This section is devoted to proof the representation for partial or incomplete Gaussian integrals. Lemma (B.1) is an ingredient to evaluate the weighted transport cost in Section 4.3.

Lemma B.1 (Partial Gaussian integral)**.** *Let $\Sigma^{-1} \in \mathbb{R}^{n\times n}_{\mathrm{sym},+}$ be a symmetric positive definite matrix and let $\eta \in S^{n-1}$ be a unit vector. Therewith, $\left\{r\eta + z^\perp\right\}_{r\in\mathbb{R}}$ is for $z^\perp \in \operatorname{span}\{\eta\}^\perp$ an affine subspace of $\mathbb{R}^n$. The integral of a centered Gaussian w.r.t. to this subspace is given by*

$$\int_{\mathbb{R}} \exp\left(-\frac{1}{2}\Sigma^{-1}[r\eta + z^\perp]\right) \mathrm{d}r = \frac{\sqrt{2\pi}}{\sqrt{\Sigma^{-1}[\eta]}} \exp\left(-\tilde{\Sigma}^{-1}[z^\perp]\right),$$
$$\textit{with} \quad \tilde{\Sigma}^{-1} = \Sigma^{-1} - \frac{\Sigma^{-1}\eta \otimes \Sigma^{-1}\eta}{\Sigma^{-1}[\eta]}.$$

Proof. To evaluate this integral on an one-dimensional subspace of $\mathbb{R}^n$, we have to expand the quadratic form $\Sigma^{-1}[r\eta + z^\perp]$ and arrive at the relation

$$\begin{aligned}
&\int_{\mathbb{R}} \exp\left(-\frac{1}{2}\Sigma^{-1}[r\eta + z^\perp]\right) \mathrm{d}r \\
&= \exp\left(-\frac{1}{2}\Sigma^{-1}[z^\perp]\right) \int_{\mathbb{R}} \exp\left(-\frac{r^2}{2}\Sigma^{-1}[\eta] + r\left\langle \eta, \Sigma^{-1}z^\perp\right\rangle\right) \mathrm{d}r \\
&= \exp\left(-\frac{1}{2}\Sigma^{-1}[z^\perp]\right) \frac{\sqrt{2\pi}}{\sqrt{\Sigma^{-1}[\eta]}} \exp\left(\frac{\left\langle \eta, \Sigma^{-1}z^\perp\right\rangle^2}{2\Sigma^{-1}[\eta]}\right) \\
&= \frac{\sqrt{2\pi}}{\sqrt{\Sigma^{-1}[\eta]}} \exp\left(-\frac{1}{2}\left(\Sigma^{-1} - \frac{\Sigma^{-1}\eta \otimes \Sigma^{-1}\eta}{\Sigma^{-1}[\eta]}\right)[z^\perp]\right),
\end{aligned}$$

which concludes the hypothesis. □

B.2. Subdeterminants, adjugates and inverses

Let $A \in \mathbb{R}^{n\times n}_{\text{sym},+}$, then define for $\eta \in S^{n-1}$ the matrix

$$\tilde{A} := A - \frac{A\eta \otimes A\eta}{A[\eta]}. \tag{B.1}$$

The matrix $\tilde{A}$ has at least rank $n-1$, since we subtracted from the positive definite matrix A a rank-1 matrix. Further, from the representation it is immediate, that $\tilde{A}$ has rank $n-1$ if and only if η is an eigenvector of A. In this case $\ker A = \operatorname{span} \eta$. It is easy to show that

$$\tilde{A} > 0 \quad \text{on } \operatorname{span}\{\eta\}^{\perp}.$$

Let $V = \operatorname{span}\{\eta\}^{\perp}$ be the $(n-1)$-dimensional subspace perpendicular to η. Then for a matrix $A \in \mathbb{R}^{n\times n}_{\text{sym},+}$ we want to calculate the determinant of A restricted to this subspace V. This determinant is obtained by first choosing $Q \in SO^n$ such that $Q(\{0\}\times\mathbb{R}^{n-1}) = V$ and then evaluating the determinant of the minor consisting of the $(n-1)\times(n-1)$ lower right submatrix of $Q^\top AQ$ denoted by $\det_{1,1}(Q^\top AQ)$. Hence, we have

$$\det{}_{1,1}(Q^\top AQ), \qquad \text{with } Q \in SO(n) : Q^\top \eta = e^1 = (1,0,\ldots,0)^\top.$$

Since $V = \operatorname{span}\{\eta\}^{\perp}$, it follows that the first column of Q is given by η and we can decompose $Q^\top AQ$ into

$$Q^\top AQ = \begin{pmatrix} A[\eta] & \widehat{Q^\top A\eta} \\ \widehat{Q^\top A\eta}^\top & \widehat{Q^\top AQ} \end{pmatrix},$$

where for a matrix M, $\widehat{M}$ is the lower right $(n-1)\times(n-1)$ submatrix of M and for a vector v, $\widehat{v}$ the $(n-1)$ lower subvector of v. Therewith, we find a similarity transformation which applied to $Q^\top AQ$ results in

$$\begin{aligned}
\det A = \det Q^\top AQ &= \det\left(\begin{pmatrix} A[\eta] & \widehat{Q^\top A\eta} \\ \widehat{Q^\top A\eta}^\top & \widehat{Q^\top AQ} \end{pmatrix}\begin{pmatrix} 1 & -\frac{\widehat{Q^\top A\eta}}{A[\eta]} \\ 0 & \mathrm{Id}_{n-1} \end{pmatrix}\right) \\
&= \det\begin{pmatrix} A[\eta] & 0 \\ \widehat{Q^\top A\eta}^\top & \widehat{Q^\top AQ} - \frac{\widehat{A\eta}\otimes\widehat{A\eta}}{A[\eta]} \end{pmatrix} = A[\eta]\det{}_{1,1}\left(Q^\top AQ - \frac{Q^\top A\eta \otimes Q^\top A\eta}{A[\eta]}\right).
\end{aligned}$$

The determinant of the minor is given by

$$\det{}_{1,1}\left(Q^\top AQ - \frac{Q^\top A\eta \otimes Q^\top A\eta}{A[\eta]}\right) = \det{}_{1,1}\left(Q^\top\left(A - \frac{A\eta\otimes A\eta}{A[\eta]}\right)Q\right).$$

Hence, by the definition (B.1) of $\tilde{A}$ and the subdeterminant we found the identity

$$\det A = A[\eta]\det{}_{1,1}(Q^\top AQ). \tag{B.2}$$

B.3. A matrix optimization

Lemma B.2. *Let $B \in \mathbb{R}^{n\times n}_{\mathrm{sym},+}$, then it holds*

$$\inf_{A\in\mathbb{R}^{n\times n}_{\mathrm{sym},+}} \left\{ \frac{\det A}{\sqrt{\det(2A-B)}} : 2A > B \right\} = \sqrt{\det B}$$

and for the optimal A holds $A = B$.

Proof. We note that

$$\frac{\det A}{\sqrt{\det(2A-B)}} = \frac{1}{\sqrt{\det(A^{-1})\det(2\,\mathrm{Id} - A^{-\frac{1}{2}}BA^{-\frac{1}{2}})}}.$$

Therewith, it is enough to maximize the radical of the root. Therefore, we substitute $A^{-\frac{1}{2}} = CB^{-\frac{1}{2}}$ with $C > 0$ not necessarily symmetric and observe that $A^{-\frac{1}{2}} = B^{-\frac{1}{2}}C^\top$. We obtain

$$\det(A^{-1})\det(2\,\mathrm{Id} - A^{-\frac{1}{2}}BA^{-\frac{1}{2}}) = \det(B^{-1})\det(CC^\top)\det(2\,\mathrm{Id} - CC^\top).$$

Note, that $CC^\top \in \mathbb{R}^{n\times n}_{\mathrm{sym},+}$ and it is enough to calculate

$$\sup_{\tilde{C}\in\mathbb{R}^{n\times n}_{\mathrm{sym},+}} \left\{ \det(\tilde{C})\det(2\,\mathrm{Id} - \tilde{C}) : \tilde{C} < 2\,\mathrm{Id} \right\}.$$

From the constraint $0 < \tilde{C} < 2\,\mathrm{Id}$ we can write $\tilde{C} = \mathrm{Id} + D$, where D is symmetric and satisfies $-\mathrm{Id} < D < \mathrm{Id}$ in the sense of quadratic forms. From here, we finally observe

$$\det(\tilde{C})\det(2\,\mathrm{Id} - \tilde{C}) = \det(\mathrm{Id} + D)\det(\mathrm{Id} - D) = \det(\mathrm{Id} - D^2).$$

Since $D^2 \geq 0$, we find the optimal $\tilde{C}$ given by Id, which yields that $A = B$. □

B.4. Jacobi's formula

Lemma B.3 (Jacobi's formula)**.** *Let $\mathbb{R} \ni t \mapsto \Phi_t \in \{A \in \mathbb{R}^{n\times n} : \det A \neq 0\}$ be a differentiable function, then*

$$\frac{\mathrm{d}}{\mathrm{d}t} \log\det\Phi_t = \mathrm{tr}\left(\Phi_t^{-1}\dot{\Phi}_t\right). \tag{B.3}$$

Proof. We first note that the determinant of $\Phi(t)$ is a multilinear function d of the columns $\phi_t^1, \ldots, \phi_t^n$, i.e. $\det\Phi_t = d(\phi_t^1, \ldots, \phi_t^n)$ Then, it follows

$$\frac{\mathrm{d}}{\mathrm{d}t}\det\Phi_t = d(\dot{\phi}_t^1, \phi_t^2, \ldots, \phi_t^n) + \cdots + d(\phi_t^1, \ldots, \phi_t^{n-1}, \dot{\phi}_t^n).$$

Now, the proof consists of two steps. We first proof the identity (B.3) for $\Phi_t = \mathrm{Id}$ and then generalize this result. If we assume w.l.o.g. that $\Phi_0 = \mathrm{Id}$. By expanding the determinant $d(\dot{\phi}_t^1, \phi_t^2, \dots, \phi_t^n)$ along its first column it immediately follows that

$$d(\dot{\phi}_t^1, \phi_t^2, \dots, \phi_t^n) = \dot{\phi}_t^{1,1}.$$

From here we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}t} \det \Phi_t = \operatorname{tr} \dot{\Phi}_t.$$

Now, let $\Phi_t = A$ be a general invertible matrix. Hence, we can apply the result from the first step to $A^{-1}\Phi_t$ and arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t} \det \left(A^{-1}\Phi_t\right) = \operatorname{tr}\left(A^{-1}\dot{\Phi}_t\right).$$

The results follows by substituting A back. □

B.5. Jacobi matrices

For a smooth function $f : \mathbb{R}^n \to \mathbb{R}^n$ denotes $Df(x)$ the *Jacobi matrix* of the partial derivatives of f in $x \in \mathbb{R}^n$ given by

$$Df(x) := \left(\frac{\mathrm{d}f_i}{\mathrm{d}x_j}(x)\right)_{i,j=1}^n.$$

Lemma B.4. *Let $A, B \in \mathbb{R}^{n\times n}$, then it holds*

$$\nabla |Ax + f(Bx)| = (A + Df(x)B)^\top \frac{Ax + f(Bx)}{|Ax + f(Bx)|}, \tag{B.4}$$

$$D\frac{f(x)}{|f(x)|} = \frac{1}{|f(x)|}\left(\mathrm{Id} - \frac{f(x)}{|f(x)|} \otimes \frac{f(x)}{|f(x)|}\right) Df(x). \tag{B.5}$$

Proof. Let us first check the relation (B.4) and calculate the partial derivative

$$\frac{\mathrm{d}\,|Ax + f(Bx)|}{\mathrm{d}x_i} = \frac{1}{2\,|Ax + f(Bx)|} \sum_j \frac{\mathrm{d}}{\mathrm{d}x_i}\left(\sum_k A_{jk}x_k + f_j(Bx)\right)^2 \tag{B.6}$$

The inner derivative of (B.6) evaluates to

$$\frac{\mathrm{d}}{\mathrm{d}x_i}\left(\sum_k A_{jk}x_k + f_j(Bx)\right)^2 = 2\left(\sum_k A_{jk}x_k + f_j(Bx)\right)\left(A_{ji} + \frac{\mathrm{d}f_j(Bx)}{\mathrm{d}x_i}\right). \tag{B.7}$$

The derivative of $f_j(Bx)$ becomes

$$\frac{\mathrm{d}f_j(Bx)}{\mathrm{d}x_i} = \frac{\mathrm{d}f_j\left(\sum_k B_{1k}x_k, \dots, \sum_k B_{nk}x_k\right)}{\mathrm{d}x_i} = \sum_{k=1}^n \partial_k f_j(Bx)B_{ki} = (Df(Bx)B)_{ji}. \tag{B.8}$$

Hence, a combination of (B.6), (B.7) and (B.8) leads to

$$\begin{aligned}\frac{\mathrm{d}\,|Ax+f(Bx)|}{\mathrm{d}x_i} &= \frac{1}{|Ax+f(Bx)|}\sum_j \left((Ax)_j + f_j(Bx)\right)\left(A_{ji}(Df(Bx)B)_{ji}\right) \\ &= \sum_j (A+Df(Bx)B)^\top_{ij}\frac{(Ax+f(Bx))_j}{|Ax+f(Bx)|},\end{aligned}$$

which shows (B.4). For the equation (B.5), let us first consider the Jacobian of the function $F(x) = \frac{x}{|x|}$, which is given by

$$DF(x) = \frac{1}{|x|}\left(\mathrm{Id} - \frac{x}{|x|}\otimes\frac{x}{|x|}\right).$$

Then, by the chain rule, we observe that

$$D\frac{f(x)}{|f(x)|} = D(F\circ f)(x) = DF(f(x))Df(x),$$

which is just (B.5). □

Appendix C

Existence of a spectral gap for L

In this short part of the appendix, we state a standard argument that ensures that the conditions ($\mathbf{A1}_{\mathtt{PI}}$) and ($\mathbf{A2}_{\mathtt{PI}}$) imply a spectral gap. More precisely, we make the following definition of the spectral gap

Definition C.1. We say that the operator $L = \varepsilon\Delta - \nabla H \cdot \nabla$ given by (1.1) satisfies a spectral gap of order $\varrho > 0$ i.e. SG(ϱ), if L has discrete non-negative spectrum and the order eigenvalues λ_i satisfy

$$\lambda_0 = 0 < \lambda_1 \leq \lambda_2 \leq \ldots \qquad \text{and} \qquad \lambda_1 \geq \varrho.$$

We take over the argument of [Kun02, Proposition 3.7] and adapt it to the case of small noise ε.

Proposition C.2 (Existence of spectral gap)**.** *Assume that H satisfies* ($\mathbf{A1}_{\mathtt{PI}}$) *and* ($\mathbf{A2}_{\mathtt{PI}}$) *with constants $C_H, K_H > 0$.*
Then for all $\varepsilon \leq \frac{1}{2}\frac{C_H^2}{2c+K_H+C_H^2}$ with $c > 0$ the operator $L = \varepsilon\Delta - \nabla H \cdot \nabla$ given by (1.1) *satisfies* $\mathrm{SG}(\varrho)$ *for some $\varrho > 0$.*

Proof. The operator $L = \varepsilon\Delta - \nabla H \cdot \nabla$ on $L^2(\mu)$ can be transformed into a *Schrödinger operator* with potential

$$L_S := -\varepsilon\Delta + \tfrac{1}{4\varepsilon}\left|\nabla H\right|^2 - \tfrac{1}{2}\Delta H$$

on $L^2(\mathrm{d}x)$. This can be seen by using the unitary transformation $U : L^2(\mathrm{d}x) \to L^2(\mu)$ given by $f \mapsto \exp(\frac{H}{2\varepsilon})f$. So, by partial integration we get

$$\begin{aligned}\mathcal{E}(Uf, Ug) &= \varepsilon \int \nabla f \cdot \nabla g + \tfrac{1}{4\varepsilon^2}|\nabla H|^2 fg + \tfrac{1}{2\varepsilon}\nabla H \cdot (f\nabla g + g\nabla f)\,\mathrm{d}x \\ &= \int \varepsilon\nabla f \cdot \nabla g + \left(\tfrac{1}{4\varepsilon}|\nabla H|^2 - \tfrac{1}{2}\Delta H\right) fg\,\mathrm{d}x = \int (L_S f) g\,\mathrm{d}x.\end{aligned}$$

For the Schrödinger operator L_S it is known (see e.g. [BS91, Theorem 3.1]) that

$$\liminf_{|x|\to\infty} \left(\tfrac{1}{4\varepsilon}|\nabla H|^2 - \tfrac{1}{2}\Delta H\right) \geq c > 0 \tag{C.1}$$

is a sufficient condition to have a discrete spectrum on $(-\infty, c)$ and in addition for every $c' < c$ and $C' < \infty$ to have a finite spectrum on $(-C', c')$. The condition (C.1) is implied by the assumption (**A2**$_{\mathtt{PI}}$) for $\varepsilon \leq \frac{1}{2}\frac{C_H^2}{2c+K_H+C_H^2}$. Since the transformation U was unitary L_S and $-L$ have the same spectrum. Hence, L has also a discrete spectrum on $(-c, \infty)$ and as it is a non-negative operator it has a discrete spectrum on $[0, \infty)$ which is finite on $[0, C')$ for every $C' < \infty$, which implies a spectral gap. □

Appendix D

Some more functional inequalities

We already introduced in Section 1.1 the both functional inequalities $\mathrm{PI}(\varrho)$ and $\mathrm{LSI}(\alpha)$. The inequalities $\mathrm{PI}(\varrho)$ and $\mathrm{LSI}(\alpha)$ can be thought as the extremes of a whole family of inequalities, from which we want introduce at least two more in this short chapter.

D.1. Horizontal and vertical distances

Let us introduce a new functional inequality incorporating the Wasserstein transportation distance, which will be in-between the Poincaré and logarithmic Sobolev inequality. The interplay of the different functional inequalities was discovered by Otto and Villani [OV00].

Definition D.1 (Wasserstein distance)**.** For any two probability measures μ, ν on an Euclidean space X, the *Wasserstein distance* of between μ and ν is defined by the formula

$$W_2^2(\nu,\mu) = \inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2 \, \pi(\mathrm{d}x,\mathrm{d}y),$$

where $\Pi(\nu,\mu)$ is the set of all *couplings*, i.e. all measures π on $\mathbb{R}^n \times \mathbb{R}^n s$ with first marginal ν and second marginal μ, i.e. $\int_{\mathbb{R}^n} \pi(\cdot,\mathrm{d}y) = \nu(\cdot)$ and $\int_{\mathbb{R}^n} \pi(\mathrm{d}x,\cdot) = \mu(\cdot)$.

Since the Wasserstein distance measures the displacement between two measure, it can be thought as a *horizontal* distance[1] on the space of probability measures. On the contrary, classical distances like the total variation, variance or relative entropy are *vertical* distances, since they measure the pointwise difference of the densities between two measures. Often, one is interested in the interplay between a horizontal and vertical distances and how a distance of the one kind can be bounded by a distance of the other kind. The following theorem provides a simple and in general rough bound of the

[1]The notion of horizontal and vertical distances is adopted from a talk of Nicola Gigli on the recent preprint [AGS12]

Wasserstein distance between two measures in terms of the second moment of the total variation of the difference of the two measures.

Theorem D.2 (Control by total variation [Vil09, Theorem 6.15])**.** *Let μ and ν be two probability measures on an Euclidean space X, then*

$$W_2^2(\nu,\mu) \le 2\int |x|^2 \; |\nu-\mu|(\mathrm{d}x) = 2\left\| |\cdot|^2 \, (\nu-\mu) \right\|_{TV}.$$

More difficult is the question, whether a horizontal distance can be bounded by an *infinitesimal* distance, like the Dirichlet energy or Fisher information, which somehow measure the local relative fluctuations between two measures. The prototype and extensively studied inequality of this type is the *transportation-information* inequality.

Definition D.3 (Transportation-information inequality WI)**.** A probability measure μ on an Euclidean space X satisfies $\mathrm{WI}(\rho)$ with constant $\rho > 0$, if for all test functions $f > 0$ with $\int f \mathrm{d}\mu = 1$ holds

$$W_2^2(f\mu,\mu) \le \frac{1}{\rho^2}\int \frac{|\nabla f|^2}{f}\,\mathrm{d}\mu. \qquad \mathrm{WI}(\rho)$$

In the abbreviation WI, W stands for the Wasserstein distance and I stands for the Fisher information.

It turns out, that the WI inequality is just in-between the Poincaré and logarithmic Sobolev inequality.

Lemma D.4 (Relation between $\mathrm{LSI}(\rho), \mathrm{WI}(\rho)$ and $\mathrm{PI}(\rho)$)**.** *Let μ be a probability measure on an Euclidean space X. Then the following implications hold*

$$\mu \textit{ satisfies } \mathrm{LSI}(\rho) \qquad \Rightarrow \qquad \mu \textit{ satisfies } \mathrm{WI}(\rho) \qquad \Rightarrow \qquad \mu \textit{ satisfies } \mathrm{PI}(\rho),$$

where all of the implications are strict.

Remark D.5. The first implication in $\mathrm{WI}(\rho)$ is on of the result in [OV00]. An example satisfying $\mathrm{WI}(\rho)$ but not $\mathrm{LSI}(\rho)$ was constructed in [CG06]. For the second implication, one uses a linearization argument, like we already presented in Remark 1.4. To proof that the implication is sharp, consider the measure $\mu(\mathrm{d}x) = Z^{-1}\exp(-|x|)\mathrm{d}x$ on the real line. Then, the condition [Goz07, Theorem 6] states that μ does not satisfy $\mathrm{WI}(\rho)$, but for instance by the Muckenhoupt functional in Theorem 5.21 one can check, that μ satisfies $\mathrm{PI}(\rho)$.

The Poincaré inequality as well as logarithmic Sobolev inequality is also in this class and bounds a vertical distance, i.e. the variance respectively the relative entropy, by an infinitesimal distance, i.e. the Dirichlet form respectively the Fisher information. An inequality showing the interplay between all three kinds of distances, i.e. vertical, horizontal and infinitesimal, was discovered by Otto and Villani [OV00]. The name HWI-inequality comes from the quantities in question, since the inequality bounds the relative entropy H in terms of the Wasserstein distance W and the Fisher information I.

Theorem D.6 (HWI inequality [OV00, Theorem 3]). *Let $\mu(\mathrm{d}x) = e^{-H(x)}\mathrm{d}x$ a probability measure on $\mathbb{R}^n$, with finite moments of order 2, such that $H \in C^2(\mathbb{R}^n)$, $\nabla^2 H \geq K_H$, $K_H \in \mathbb{R}$ (not necessarily positive). Then, for all test functions f with $\int f \mathrm{d}\mu = 1$ holds*

$$\mathrm{Ent}_\mu(f) = H(f\mu|\mu) \leq W_2(f\mu,\mu)\sqrt{2I(f\mu|\mu)} - \frac{K_H}{2} W_2^2(f\mu,\mu). \qquad \text{HWI}$$

Remark D.7 (A variance estimate in terms of $\mathcal{T}_\mu$). A special case of the "mean-difference" estimate occurs, by setting $\nu_0 = g\mu$ and $\nu_1 = \mu$, where $g \geq 0$ and $\int g \ \mathrm{d}\mu = 1$, then we arrive at the following covariance estimate

$$\mathrm{cov}_\mu^2(f,g) = \left(\mathbb{E}_{g\mu}(f) - \mathbb{E}_\mu(f)\right)^2 \leq \mathcal{T}_\mu^2(g\mu,\mu) \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Finally, setting $f = g$ results in

$$\mathrm{var}_\mu(f) \leq \mathcal{T}_\mu(f\mu,\mu)\sqrt{\int |\nabla f|^2 \,\mathrm{d}\mu}. \qquad \text{(D.1)}$$

The estimate (D.1) has the same structure as the HWI inequality in the sense, that it connects a vertical with the product of a horizontal and the square root of an infinitesimal distance. However, the estimate (D.1) does not demand a lower bound on the Hessian of the exponential density of μ. Moreover, from Remark 4.21 it is clear, the $\mathcal{T}_\mu(f\mu,\mu)$ can be replaced by the modified weighted transport distance $\widetilde{\mathcal{T}}_\mu(f\mu,\mu)$ given in (4.53).

D.2. Defective logarithmic Sobolev inequality

Let us present how a *defective* logarithmic Sobolev inequality can be tightened to a logarithmic Sobolev inequality with the help of a Poincaré inequality.

Definition D.8 (Defective logarithmic Sobolev inequality $\mathrm{dLSI}(\alpha_d, B)$). A probability measure μ on $\mathbb{R}^n$ satisfies the *defective logarithmic Sobolev inequality* $\mathrm{dLSI}(\alpha_d, B)$ with constants $\alpha_d, B > 0$, if for all test function $f : \mathbb{R}^n \to \mathbb{R}^+$ holds

$$\mathrm{Ent}_\mu(f) \leq \frac{1}{\alpha_d}\int \frac{|\nabla f|^2}{2f} \mathrm{d}\mu + B \int f \mathrm{d}\mu. \qquad \mathrm{dLSI}(\alpha_d, B)$$

Proposition D.9 ($\mathrm{dLSI}(\alpha_d, B)$ and $\mathrm{PI}(\varrho)$ imply $\mathrm{LSI}(\alpha)$). *Assume μ satisfies $\mathrm{dLSI}(\alpha_d, B)$ and $\mathrm{PI}(\varrho)$, then μ satisfies $\mathrm{LSI}(\alpha)$ with*

$$\frac{1}{\alpha} = \frac{1}{\alpha_d} + \frac{B+2}{\varrho}.$$

Proof. The argument is from [Led99a] and is a simple consequence of the estimate

$$\mathrm{Ent}_\mu(f^2) \leq \mathrm{Ent}_\mu\left((f - \mathbb{E}_\mu(f))^2\right) + 2\,\mathrm{var}_\mu(f),$$

which is due to [Rot86] and [DS89]. An application of $\mathrm{dLSI}(\alpha_d, B)$ leads to

$$\mathrm{Ent}_\mu(f^2) \leq \frac{1}{\alpha_d}\int 2\,|\nabla f|^2 \,\mathrm{d}\mu + (B+2)\,\mathrm{var}_\mu(f).$$

The result follow from applying $\mathrm{PI}(\varrho)$ to the variance in the last term. □

List of symbols and abbreviations

$\mathrm{d}x$	n-dimensional Lebesgue measure
$\mathcal{H}^k(\mathrm{d}x)$	k-dimensional Hausdorff measure
$\mu(\cdot)$	density of of the probability measure μ
$\mu(\mathrm{d}x)$	short form for $\mu(\cdot)\mathrm{d}x$
$\mu(\cdot\|y)$	conditional measure obtained from μ by conditioning on y
$\nu \ll \mu$	ν is absolutely continuous w.r.t. μ
$\operatorname{supp}\mu$	the support of μ, $\operatorname{supp}\mu = \overline{\{\mu(x) > 0\}}$
$\mathbb{1}_\Omega$	characteristic function on Ω
$\mu \llcorner \Omega$	restriction of μ onto $\Omega \subset \operatorname{supp}\mu$, $(\mu \llcorner \Omega)(\mathrm{d}x) := \frac{\mathbb{1}_\Omega(x)}{\mu(\Omega)}\mu(x)\mathrm{d}x$
$\mathbb{E}_\mu(f)$	expectation of f under μ: $\int f \mathrm{d}\mu$
$\operatorname{var}_\mu(f)$	variance of f w.r.t. μ: $\int \left(f^2 - \int f \mathrm{d}\mu\right) \mathrm{d}\mu$
$\operatorname{Ent}_\mu(f)$	relative entropy of f w.r.t. μ: $\int f \log \frac{f}{\int f \mathrm{d}\mu} \mathrm{d}\mu$
$\Phi_\sharp \mu$	push-forward measure under $\Phi : \mathbb{R}^n \to \mathbb{R}^n \supset U$, $(\Phi_\sharp \mu)(U) = \mu(\Phi^{-1}(U))$
$\mathcal{N}(m, \Sigma)$	multivariate normal distribution in $\mathbb{R}^n$ with mean $m \in \mathbb{R}^n$ and covariance matrix $\Sigma \in \mathbb{R}^{n\times n}_{\mathrm{sym},+}$
$A^\top$	matrix transpose
$\operatorname{tr} A$	trace of A
$\langle x, y\rangle := x \cdot y$	Euclidean scalar product between $x, y \in \mathbb{R}^n$
$\|\cdot\|$	Euclidean norm
$B_r(x)$	Euclidean ball of radius r around x
$B_r := B_r(0)$	Euclidean ball of radius r around 0
$A[x]$	testing a quadratic form $A \in \mathbb{R}^{n\times n}$ by $x \in R^n$: $A[x] = \langle x, Ax\rangle$
$A \leq B$	for some $A, B \in \mathbb{R}^{n\times n}$ means $A[x] \leq B[x]$ for all $x \in \mathbb{R}^n$
$\det_{i,j} A$	determinant of the matrix A with i-th row and j-th column removed
$\mathbb{R}^{n\times n}_{\mathrm{sym},+}$	set of symmetric positive definite matrices, i.e. $A = A^\top$ and $A > 0$.
S^{n-1}	unit sphere in $\mathbb{R}^n$, i.e. $\{\eta \in \mathbb{R}^n : \|x\| = 1\}$
$\sphericalangle(\eta, \theta)$	geodesic on S^{n-1} between $\eta, \theta \in S^{n-1}$
$SO(n)$	rotational matrices in $\mathbb{R}^n$, $Q \in \mathbb{R}^{n\times n}$: $Q^{-1} = Q^\top$, $\det Q = 1$.
$\nabla^2 H$	Hessian of $H : \mathbb{R}^n \to \mathbb{R}$
$\nabla \cdot F$	the divergence of $F : \mathbb{R}^n \to \mathbb{R}^n$
$\dot\psi_t = \partial_t \psi_t$	derivative in parameter $t \in \mathbb{R}$
$\omega(\varepsilon)$	a smooth monotone decreasing function satisfying $\geq \|\log\varepsilon\|^{\frac{1}{2}}$ for $\varepsilon < 1$
$\approx, \lesssim, \gtrsim$	$=, \leq, \geq$ up to multiplicative error of the form $1 + O(\sqrt{\varepsilon}\, \omega^3(\varepsilon))$.
C, c	constants only depending on the dimension n and H
C_f, c_f	constants additionally depending on the function or variable f

List of Figures

List of Statements

Bibliography

[AGS05] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zürich. Birkhäuser-Verlag, Basel, 2005.

[AGS12] L. Ambrosio, N. Gigli, and G. Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *arXiv:1109.0222*, pages 1–60, 2012.

[Arr89] S. Arrhenius. On the Reaction Velocity of the Inversion of Cane Sugar by Acids. *Zeitschrift für physikalische Chemie*, 4:226ff, 1889.

[BBCG08] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electronic Communications in Probability*, 13:60–66, 2008.

[BBI09] A. Bianchi, A. Bovier, and D. Ioffe. Sharp asymptotics for metastability in the random field Curie-Weiss model. *Electronic Journal of Probability*, 14:1541–1603, July 2009.

[BCG08] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *Journal of Functional Analysis*, 254(3):727–759, February 2008.

[BdHS10] A. Bovier, F. den Hollander, and C. Spitoni. Homogeneous nucleation for Glauber and Kawasaki dynamics in large volumes at low temperatures. *The Annals of Probability*, 38(2):661–713, March 2010.

[BE85] D. Bakry and M. Émery. *Diffusions hypercontractives*. Séminaire de probabilités, XIX, Springer, Berlin, 1985.

[BEGK04] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in Reversible Diffusion Processes I: Sharp Asymptotics for Capacities and Exit Times. *Journal of the European Mathematical Society*, 6(4):399–424, 2004.

[Ber11] N. Berglund. Kramers' law: Validity, derivations and generalisations. *arXiv:1106.5799*, 2011(x):24, June 2011.

[BG99] S. G. Bobkov and F. Götze. Exponential Integrability and Transportation Cost Related to Logarithmic Sobolev Inequalities. *Journal of Functional Analysis*, 163(1):1–28, April 1999.

[BGK05] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes II: precise asymptotics for small eigenvalues. *Journal of the European Mathematical Society*, 7(1):69–99, 2005.

[Bha08] R. Bhatia. The logarithmic mean. *Resonance*, 13(6):583–594, 2008.

[BK90] K. A. Berman and M. H. Konsowa. Random Paths and Cuts, Electrical Networks, and Reversible Markov Chains. *SIAM Journal on Discrete Mathematics*, 3(3):311, 1990.

[BS91] F. A. Berezin and M. A. Shubin. *The Schrödinger equation*. Springer, 1991.

[BT06] S. G. Bobkov and P. Tetali. Modified Logarithmic Sobolev Inequalities in Discrete Settings. *Journal of Theoretical Probability*, 19(2):289–336, 2006.

[Car72] B. C. Carlson. The Logarithmic Mean. *The American Mathematical Monthly*, 79(6):615, June 1972.

[CG06] P. Cattiaux and A. Guillin. On quadratic transportation cost inequalities. *Journal de Mathématiques Pures et Appliquées*, 86(4):342–361, October 2006.

[CG10] P. Cattiaux and A. Guillin. Functional Inequalities via Lyapunov conditions. *arXiv:1001.1822*, 2:1–11, January 2010.

[CGW09] P. Cattiaux, A. Guillin, and L.-M. Wu. A note on Talagrand's transportation inequality and logarithmic Sobolev inequality. *Probability Theory and Related Fields*, 148(1-2):285–304, June 2009.

[CGWW09] P. Cattiaux, A. Guillin, F.-y. Wang, and L.-M. Wu. Lyapunov conditions for Super Poincaré inequalities. *Journal of Functional Analysis*, 256(6):1821–1841, March 2009.

[CL04] E. Carlen and M. Loss. Logarithmic Sobolev inequalities and spectral gaps. In *Recent advances in the theory and applications of mass transport*, volume 353 of *Contemp. Math.*, pages 53–60. Amer. Math. Soc., Providence, RI, 2004.

[CM10] D. Chafaï and F. Malrieu. On fine properties of mixtures with respect to concentration of measure and Sobolev type inequalities. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 46(1):72–96, 2010.

[DS89] J.-D. Deuschel and D. W. Stroock. *Large deviations*, volume 137 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1989.

[DSC96] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability*, 6(3):695–750, August 1996.

[DV76] M. D. Donsker and S. R. S. Varadhan. On the principal eigenvalue of second-order elliptic differential operators. *Communications on Pure and Applied Mathematics*, 29(6):595–621, 1976.

[EG92] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[EM11] M. Erbar and J. Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *arXiv:1111.2687*, page 37, November 2011.

[Eyr35] H. Eyring. The Activated Complex in Chemical Reactions. *The Journal of Chemical Physics*, 3(2):107, 1935.

[Fou04] P. Fougères. Spectral Gap for log-Concave Probability Measures on the Real Line. In M. Émery, M. Ledoux, and M. Yor, editors, *Séminaire de Probabilités XXXVIII*, pages 95–123. Springer Berlin / Heidelberg, 2004.

[FW98] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems.*. Grundlehren der Mathematischen Wissenschaften. 260. New York, NY, 1998.

[GMSvdM02] O. Gonzalez, J. Maddocks, F. Schuricht, and H. von der Mosel. Global curvature and self-contact of nonlinearly elastic curves and rods. *Calculus of Variations and Partial Differential Equations*, 14(1):29–68, 2002.

[GOV87] A. Galves, E. Olivieri, and M. E. Vares. Metastability for a class of dynamical systems subject to small random perturbations. *The Annals of Probability*, 15:1288–1305, 1987.

[GOVW09] N. Grunewald, F. Otto, C. Villani, and M. G. Westdickenberg. A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 45(2):302–351, May 2009.

[Goz07] N. Gozlan. Characterization of Talagrand's like transportation-cost inequalities on the real line. *Journal of Functional Analysis*, 250(2):400–425, September 2007.

[GS02] A. L. Gibbs and F. E. Su. On Choosing and Bounding Probability Metrics. *International Statistical Review / Revue Internationale de Statistique*, 70(3):419, December 2002.

[HKN04] B. Helffer, M. Klein, and F. Nier. Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach. *Matemática Contemporânea*, 26:41–85, 2004.

[HN05] B. Helffer and F. Nier. *Hypoelliptic Estimates and Spectral Theory for Fokker-Planck Operators and Witten Laplacians*. Springer-Verlag, Berlin, 2005.

[HN06] B. Helffer and F. Nier. Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach: the case with boundary. *Mémoires de la Société Mathématique de France. Nouvelle Série*, 105:vi+89, 2006.

[HS87] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *Journal of Statistical Physics*, 46(5-6):1159–1194, 1987.

[HY95] Y. Higuchi and N. Yoshida. Analytic conditions and phase transition for Ising models. *Lecture Notes in Japanese*, 1995.

[JSTV04] M. Jerrum, J.-B. Son, P. Tetali, and E. Vigoda. Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains. *The Annals of Applied Probability*, 14(4):1741–1765, November 2004.

[KOV89] C. Kipnis, S. Olla, and S. R. S. Varadhan. Hydrodynamics and large deviation for simple exclusion processes. *Communications on Pure and Applied Mathematics*, 42(2):115–137, March 1989.

[Kra40] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[Kun02] A. Kunz. On Extremes of Multidimensional Stationary Diffusion Processes in Euclidean Norm. *Preprint:* `http://de.scientificcommons.org/43322528`, 2002.

[Led99a] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. *Séminaire de probabilités de Strasbourg*, 33:120–216, 1999.

[Led99b] M. Ledoux. Logarithmic sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités XXXV*, number X, pages 167–194. 1999.

[LL10] F. Legoll and T. Lelièvre. Effective dynamics using conditional expectations. *Nonlinearity*, 23(9):2131–2163, September 2010.

[Maa11] J. Maas. Gradient flows of the entropy for finite Markov chains. *Journal of Functional Analysis*, 261(8):2250–2292, October 2011.

[Mic95] L. Miclo. Comportement de spectres d'opérateurs de Schrödinger à basse température. *Bulletin des sciences mathématiques*, 119(6):529–553, 1995.

[Mic99] L. Miclo. An example of application of discrete Hardy's inequalities. *Markov Process. Relat. Fields*, 5(3):319–330, 1999.

[Mie11] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *WIAS Preprint*, (1650):1–31, 2011.

[Muc72] B. Muckenhoupt. Hardy's inequality with weights. *Studia Mathematica*, 44:31–38, 1972.

[Øks03] B. Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003.

[Olv97] F. W. J. Olver. *Asymptotics and Special Functions*. A. K. Peters, Wellesley, MA, 1997.

[OV00] F. Otto and C. Villani. Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality. *Journal of Functional Analysis*, 173(2):361–400, June 2000.

[PW60] L. E. Payne and H. F. Weinberger. An optimal Poincaré inequality for convex domains. *Archive for Rational Mechanics and Analysis*, 5(1):286–292, January 1960.

[Rot86] O. Rothaus. Hypercontractivity and the Bakry-Emery criterion for compact Lie groups. *Journal of Functional Analysis*, 65(3):358–367, 1986.

[Roy07] G. Royer. *An initiation to logarithmic Sobolev inequalities*, volume 14 of *SMF/AMS Texts and Monographs*. American Mathematical Society, Providence, RI, 2007.

[Sch10] Z. Schuss. *Theory and Applications of Stochastic Processes*, volume 170 of *Applied Mathematical Sciences*. Springer New York, New York, NY, 2010.

[Sug95] M. Sugiura. Metastable behaviors of diffusion processes with small parameter. *Journal of the Mathematical Society of Japan*, 47(4):755–788, 1995.

[Tes12] G. Teschl. *Ordinary differential equations and Dynamical Systems*. American Mathematical Society, XXX edition, 2012.

[Vil09] C. Villani. *Optimal Transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, Heidelberg, 2009.