

# Agmon-Type Estimates for a Class of Difference Operators

Markus Klein and Elke Rosenberger

**Abstract.** We analyze a general class of self-adjoint difference operators  $H_\varepsilon = T_\varepsilon + V_\varepsilon$  on  $\ell^2((\varepsilon\mathbb{Z})^d)$ , where  $V_\varepsilon$  is a one-well potential and  $\varepsilon$  is a small parameter. We construct a Finslerian distance  $d$  induced by  $H_\varepsilon$  and show that short integral curves are geodesics. Then we show that Dirichlet eigenfunctions decay exponentially with a rate controlled by the Finsler distance to the well. This is analog to semiclassical Agmon estimates for Schrödinger operators.

## 1. Introduction

The central topic of this paper is the investigation of a rather general class of families of self-adjoint difference operators  $H_\varepsilon$  on the Hilbert space  $\ell^2((\varepsilon\mathbb{Z})^d)$ , as the small parameter  $\varepsilon > 0$  tends to zero.

The operator  $H_\varepsilon$  is given by

$$H_\varepsilon = (T_\varepsilon + V_\varepsilon), \quad \text{where} \quad T_\varepsilon = \sum_{\gamma \in (\varepsilon\mathbb{Z})^d} a_\gamma \tau_\gamma, \quad (1.1)$$

$$(\tau_\gamma u)(x) = u(x + \gamma) \quad \text{and} \quad (a_\gamma u)(x) := a_\gamma(x, \varepsilon)u(x) \quad \text{for} \quad x, \gamma \in (\varepsilon\mathbb{Z})^d \quad (1.2)$$

where  $V_\varepsilon$  is a multiplication operator, which in leading order is given by  $V_0 \in \mathcal{C}^\infty(\mathbb{R}^d)$ .

We remark that the limit  $\varepsilon \rightarrow 0$  is analog to the semiclassical limit  $\hbar \rightarrow 0$  for the Schrödinger operator  $-\hbar^2\Delta + V$ . This paper is the first in a series of papers; the aim is to develop an analytic approach to the semiclassical eigenvalue problem and tunneling for  $H_\varepsilon$  which is comparable in detail and precision to the well known analysis for the Schrödinger operator (see Simon [24, 25] and Helffer–Sjöstrand [17]). Our motivation comes from stochastic problems (see Bovier–Eckhoff–Gayrard–Klein [8, 9]). A large class of discrete Markov chains analyzed in [9] with probabilistic techniques falls into the framework of difference operators treated in this article.

We recall that sharp semiclassical Agmon estimates describing the exponential decay of eigenfunctions of appropriate Dirichlet realizations of the Schrödinger operator are crucial to analyze tunneling for the Schrödinger operator. We further recall that the original work of Agmon on the decay of eigenfunctions for second order differential operators is not in the semiclassical limit. It treats the limit  $|x| \rightarrow \infty$  (in a non bounded domain of  $\mathbb{R}^n$ ). Agmon realized in [3] that for a large class of such operators the exponential rate at which eigenfunctions decay is given by the geodesic distance in the Agmon metric. This is the Riemannian metric from Jacobi’s theorem in classical mechanics: For a Hamilton function whose kinetic energy is a positive definite quadratic form in the momenta, the projection to configuration space of an integral curve of the Hamiltonian vector field is a geodesic in the Agmon (Jacobi) metric.

This paper contains analog results for the class of operators  $H_\varepsilon$ , including a generalization of Jacobi’s theorem. It is essential that we consider these operators as semiclassical quantizations of suitable Hamilton functions and investigate the relation of these Hamilton functions to Finsler geometry. In this generality our results are new. We recall, however, that various examples extending the original framework of the semiclassical analysis in the work of Simon [24] and Helffer–Sjöstrand [17] have been analyzed: The operator  $\cos hD_x + \cos x$  in Harpers equation (see e.g. Helffer–Sjöstrand [18]), the Schrödinger operator with magnetic field (Helffer–Mohamed [15]), the Dirac and Klein–Gordon operator (see e.g. Helffer–Parisse [16], Servat [23]) and the Kac operator (Helffer [14]).

If  $\mathbb{T}^d := \mathbb{R}^d/(2\pi)\mathbb{Z}^d$  denotes the  $d$ -dimensional torus and  $b \in \mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{T}^d \times (0, 1])$ , a pseudo-differential operator  $\text{Op}_\varepsilon^{\mathbb{T}^d}(b) : \mathcal{K}((\varepsilon\mathbb{Z})^d) \rightarrow \mathcal{K}'((\varepsilon\mathbb{Z})^d)$  is defined by

$$\text{Op}_\varepsilon^{\mathbb{T}^d}(b)v(x) := (2\pi)^{-d} \sum_{y \in (\varepsilon\mathbb{Z})^d} \int_{[-\pi, \pi]^d} e^{\frac{i}{\varepsilon}(y-x)\xi} b(x, \xi; \varepsilon)v(y) d\xi, \tag{1.3}$$

where

$$\mathcal{K}((\varepsilon\mathbb{Z})^d) := \{u : (\varepsilon\mathbb{Z})^d \rightarrow \mathbb{C} \mid u \text{ has compact support}\} \tag{1.4}$$

and  $\mathcal{K}'((\varepsilon\mathbb{Z})^d) := \{f : (\varepsilon\mathbb{Z})^d \rightarrow \mathbb{C}\}$  is dual to  $\mathcal{K}((\varepsilon\mathbb{Z})^d)$  by use of the scalar product  $\langle u, v \rangle_{\ell^2} := \sum_x \bar{u}(x)v(x)$ .

We remark that under certain assumptions on the  $a_\gamma$  defining  $T_\varepsilon$  in (1.1), one has  $T_\varepsilon = \text{Op}_\varepsilon^{\mathbb{T}^d}(t(\cdot, \cdot; \varepsilon))$ , where  $t \in \mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{T}^d \times (0, 1])$  is given by

$$t(x, \xi, \varepsilon) = \sum_{\gamma \in (\varepsilon\mathbb{Z})^d} a_\gamma(x, \varepsilon) \exp\left(-\frac{i}{\varepsilon}\gamma \cdot \xi\right). \tag{1.5}$$

Here  $t(x, \xi; \varepsilon)$  is considered as a function on  $\mathbb{R}^{2d} \times (0, 1]$ , which is  $2\pi$ -periodic with respect to  $\xi$ .

Furthermore, assuming that  $a_\gamma(x, \varepsilon) = a_\gamma^{(0)}(x) + \varepsilon a_\gamma^{(1)}(x) + R_\gamma^{(2)}(x, \varepsilon)$ , where  $R_\gamma^{(2)}(x, \varepsilon) = O(\varepsilon^2)$  uniformly with respect to  $x$  and  $\gamma$ , we can write

$$\begin{aligned}
 t(x, \xi; \varepsilon) &= t_0(x, \xi) + \varepsilon t_1(x, \xi) + t_2(x, \xi; \varepsilon), \quad \text{with} \tag{1.6} \\
 t_j(x, \xi) &:= \sum_{\gamma \in (\varepsilon\mathbb{Z})^d} a_\gamma^{(j)}(x) e^{-\frac{i}{\varepsilon} \gamma \xi}, \quad j = 0, 1 \\
 t_2(x, \xi; \varepsilon) &:= \sum_{\gamma \in (\varepsilon\mathbb{Z})^d} R_\gamma^{(2)}(x, \varepsilon) e^{-\frac{i}{\varepsilon} \gamma \xi}.
 \end{aligned}$$

Thus, in leading order the symbol of  $H_\varepsilon$  is  $h_0 = t_0 + V_0$ . In its original form, neither Jacobi's theorem applies to  $h_0(x, \xi)$  nor Agmon estimates to  $H_\varepsilon$ . Our analysis is motivated by the remark in Agmon's book [3] to develop part of the theory of the Agmon metric in the more general context of Finsler geometry. It turns out that the Hamilton function  $-h_0(x, i\xi)$  (this transformation is analog to the procedure in the case of the Schrödinger operator) in a natural way induces a Finsler metric and an associated Finsler distance  $d$  on  $\mathbb{R}^d$ . This allows to formulate and prove a generalization of Jacobi's theorem (which might be some kind of lesser known folk wisdom in mathematical physics, which, however, we were unable to find in the literature) and prove an analog of the semiclassical Agmon estimates for  $H_\varepsilon$ . We remark that Finsler distances have been used for higher order elliptic differential operators in the analysis of decay of resolvent kernels and/or heat kernels, see Tintarev [26] and Barbatis [6, 7].<sup>1</sup> However, these papers do not develop a generalization of Jacobi's theorem, which turns out to be crucial in our semiclassical analysis.

We will now state our assumptions on  $H_\varepsilon$  and formulate our results more precisely.

**Hypothesis 1.1.** (a) The coefficients  $a_\gamma(x, \varepsilon)$  in (1.1) are functions

$$a : (\varepsilon\mathbb{Z})^d \times \mathbb{R}^d \times (0, 1] \rightarrow \mathbb{R}, \quad (\gamma, x, \varepsilon) \mapsto a_\gamma(x, \varepsilon), \tag{1.7}$$

satisfying the following conditions:

(i) They have an expansion

$$a_\gamma(x, \varepsilon) = a_\gamma^{(0)}(x) + \varepsilon a_\gamma^{(1)}(x) + R_\gamma^{(2)}(x, \varepsilon), \tag{1.8}$$

where  $a_\gamma^{(i)} \in \mathcal{C}^\infty(\mathbb{R}^d)$  and  $|a_\gamma^{(j)}(x) - a_\gamma^{(j)}(x + h)| = O(|h|)$  for  $j = 0, 1$  uniformly with respect to  $\gamma \in (\varepsilon\mathbb{Z})^d$  and  $x \in \mathbb{R}^d$ . Furthermore  $R_\gamma^{(2)} \in \mathcal{C}^\infty(\mathbb{R}^d \times (0, 1])$  for all  $\gamma \in (\varepsilon\mathbb{Z})^d$ .

(ii)  $\sum_\gamma a_\gamma^{(0)} = 0$  and  $a_\gamma^{(0)} \leq 0$  for  $\gamma \neq 0$

(iii)  $a_\gamma(x, \varepsilon) = a_{-\gamma}(x + \gamma, \varepsilon)$  for  $x \in \mathbb{R}^d, \gamma \in (\varepsilon\mathbb{Z})^d$

---

<sup>1</sup>M. K. thanks S. Agmon for the reference to [26], where prior to the publication of Agmon's book a Finslerian approach was used to obtain estimates on the kernel of the resolvent and the decay of the heat kernel for higher order elliptic operators, following ideas of Agmon.

- (iv) For any  $c > 0$  there exists  $C > 0$  such that for  $j = 0, 1$  uniformly with respect to  $x \in (\varepsilon\mathbb{Z})^d$  and  $\varepsilon$

$$\left\| e^{\frac{c|\cdot|}{\varepsilon}} a^{(j)}(x) \right\|_{\ell^2_\gamma((\varepsilon\mathbb{Z})^d)} \leq C \quad \text{and} \quad \left\| e^{\frac{c|\cdot|}{\varepsilon}} R^{(2)}(x) \right\|_{\ell^2_\gamma((\varepsilon\mathbb{Z})^d)} \leq C\varepsilon^2 \quad (1.9)$$

- (v)  $\text{span}\{\gamma \in (\varepsilon\mathbb{Z})^d \mid a_\gamma^{(0)}(x) < 0\} = \mathbb{R}^d$  for all  $x \in \mathbb{R}^d$ .

- (b) (i) The potential energy  $V_\varepsilon$  is the restriction to  $(\varepsilon\mathbb{Z})^d$  of a function  $\widehat{V}_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$ , which has an expansion

$$\widehat{V}_\varepsilon(x) = \sum_{l=0}^N \varepsilon^l V_l(x) + R_{N+1}(x; \varepsilon),$$

where  $V_l \in \mathcal{C}^\infty(\mathbb{R}^d)$ ,  $R_{N+1} \in \mathcal{C}^\infty(\mathbb{R}^d \times (0, \varepsilon_0])$  for some  $\varepsilon_0 > 0$  and for any compact set  $K \subset \mathbb{R}^d$  there exists a constant  $C_K$  such that  $\sup_{x \in K} |R_{N+1}(x; \varepsilon)| \leq C_K \varepsilon^{N+1}$ .

- (ii) There exist constants  $R, C > 0$  such that  $V_\varepsilon(x) > C$  for all  $|x| \geq R$  and  $\varepsilon \in (0, \varepsilon_0]$ .
- (iii)  $V_0(x)$  has exactly one non-degenerate minimum at  $x_0 = 0$  with the value  $V_0(0) = 0$ .

The following lemma couples the assumptions on the coefficients  $a_\gamma$  given in Hypothesis 1.1 with properties of the symbol  $t$  and the kinetic energy  $T_\varepsilon$ .

**Lemma 1.2.** *Assume Hypothesis 1.1 and let  $t$  and  $t_j, j = 0, 1, 2$  be defined in (1.5) and (1.6) respectively. Then:*

- (a)  $t \in \mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{T}^d \times (0, 1])$  and the estimate  $\sup_{x, \xi} |\partial_x^\alpha \partial_\xi^\beta t(x, \xi; \varepsilon)| \leq C_{\alpha, \beta}$  holds for all  $\alpha, \beta \in \mathbb{N}^d$  uniformly with respect to  $\varepsilon$ . Furthermore  $t_0$  and  $t_1$  are bounded and  $\sup_{x, \xi} |t_2(x, \xi; \varepsilon)| = O(\varepsilon^2)$ .
- (b) The  $2\pi$ -periodic function  $\mathbb{R}^d \ni \xi \mapsto t_0(x, \xi)$  is even and has an analytic continuation to  $\mathbb{C}^d$ .
- (c) At  $\xi = 0$ , for fixed  $x \in \mathbb{R}^d$  the function  $t_0$  defined in (1.6) has an expansion

$$t_0(x, \xi) = \langle \xi, B(x)\xi \rangle + O(|\xi|^4) \quad \text{as } |\xi| \rightarrow 0, \quad (1.10)$$

where  $B : \mathbb{R}^d \rightarrow \mathcal{M}(d \times d, \mathbb{R})$  is positive definite and symmetric.

- (d) The operator  $T_\varepsilon$  defined in (1.1) is symmetric, bounded (uniformly in  $\varepsilon$ ) and  $\langle u, T_\varepsilon u \rangle_{\ell^2} \geq -C\varepsilon \|u\|^2$  for some  $C > 0$ . Furthermore  $T_\varepsilon = \text{Op}_\varepsilon^{\mathbb{T}^d}(t)$  (see (1.3)).

*Remark 1.3.*  $T_\varepsilon$  being symmetric boils down to  $a_\gamma$  being real and condition (a)(iii) of Hypothesis 1.1. In the probabilistic context, which is our main motivation, the latter is a standard reversibility condition while the former is automatic for a Markov chain (see Section 4). Since  $T_\varepsilon$  is bounded,  $H_\varepsilon = T_\varepsilon + V_\varepsilon$  defined in (1.1) possesses a self adjoint realization on the maximal domain of  $V_\varepsilon$ . Abusing notation, we shall denote this realization also by  $H_\varepsilon$  and its domain by  $\mathcal{D}(H_\varepsilon) \subset \ell^2((\varepsilon\mathbb{Z})^d)$ . The associated symbol is denoted by  $h(x, \xi; \varepsilon)$ . Clearly,  $H_\varepsilon$  commutes with complex conjugation.

We will use the notation

$$\tilde{a} : \mathbb{Z}^d \times \mathbb{R}^d \ni (\eta, x) \mapsto \tilde{a}_\eta(x) := a_{\varepsilon\eta}^{(0)}(x) \in \mathbb{R} \tag{1.11}$$

and set

$$\tilde{h}_0(x, \xi) := -h_0(x, i\xi) = \tilde{t}_0(x, \xi) - V_0(x) : \mathbb{R}^{2d} \rightarrow \mathbb{R}, \tag{1.12}$$

where by Lemma 1.2 (b)

$$\tilde{t}_0(x, \xi) := -t_0(x, i\xi) = - \sum_{\eta \in \mathbb{Z}^d} \tilde{a}_\eta(x) \cosh(\eta \cdot \xi). \tag{1.13}$$

We shall now describe, how Hamilton functions such as  $\tilde{h}_0$  for fixed energy  $E$  introduce a Finsler geometry in configuration space.

**Hypothesis 1.4.** Let  $M$  be a  $d$ -dimensional smooth manifold. Let  $h \in \mathcal{C}^\infty(T^*M, \mathbb{R})$  be hyperregular and even and strictly convex in each fibre. Furthermore, let  $h(\cdot, 0)$  be bounded from above. For  $E \in \mathbb{R}$  set  $\tilde{M} := M \setminus \{h(x, 0) \geq E\}$ . Denoting the fibre derivative of  $h$  by  $\mathcal{D}_F h$ , we associate to  $h$  the energy function  $E_h(x, v) := h \circ (\mathcal{D}_F h)^{-1}(x, v)$  on  $TM$ .

The notion of fibre derivative and hyperregular are standard (see Abraham–Marsden [2]). For convenience of the reader, they are repeated in Definition 2.7.

Now Theorem 2.11 states that assuming Hypothesis 1.4

$$\ell_{h,E}(x, v) := (\mathcal{D}_F h)^{-1}(x, \tilde{v}) \cdot v, \tag{1.14}$$

where  $\tilde{v}$  is chosen such that  $E_h(x, \tilde{v}) = E$ , is a Finsler function on  $\tilde{M}$ . The most important property of  $\ell_{h,E}$  is the homogeneity

$$\ell_{h,E}(x, \lambda v) = |\lambda| \ell_{h,E}(x, v), \quad \lambda \in \mathbb{R},$$

which is analog to the homogeneity of  $|v| = \sqrt{g(v, v)}$  in the case of a Riemannian metric. This is essential to define a curve length associated to  $\ell_{h,E}$  as described in Definition 2.3 by

$$s_{\ell_{h,E}}(\gamma) := \int_a^b \ell_{h,E}(\gamma(t), \dot{\gamma}(t)) dt.$$

A Finsler geodesic is then a curve  $\gamma$  on  $M$ , for which  $s_{\ell_{h,E}}$  is extremal (see Definition 2.4).

The following theorem establishes the connection between geodesics with respect to the Finsler function  $\ell_{h,E}$  for a given hyperregular Hamilton function  $h$  and the integral curves of the associated Hamiltonian vector field  $X_h$ . It amplifies the Maupertuis principle in classical mechanics.

**Theorem 1.5.** *Let  $h, E$  and  $\tilde{M}$  satisfy Hypothesis 1.4. Let  $\ell_h := \ell_{h,E}$  be as defined in (1.14) (see Theorem 2.11 for details).*

- (a) *Let  $\gamma_0 : [a, b] \rightarrow \tilde{M}$  be a base integral curve of the Hamiltonian vector field  $X_h$  with energy  $E$  (i.e.  $E_h(\gamma_0(t), \dot{\gamma}_0(t)) = E$  for all  $t \in [a, b]$ ). Then  $\gamma_0$  is a geodesic on  $\tilde{M}$  with respect to  $\ell_h$ .*

- (b) *Conversely, if  $\gamma_0$  is a geodesic on  $\widetilde{M}$  with respect to  $\ell_h$  with energy  $E_h(\gamma_0, \dot{\gamma}_0) = E$ , then  $\gamma_0$  is a base integral curve of  $X_h$ .*

The Hamilton function  $\tilde{h}_0$  introduced in (1.12) actually satisfies Hypothesis 1.4 with respect to the energy  $E = 0$  (see Corollary 2.15) and thus induces a Finsler function  $\ell := \ell_{\tilde{h}_0, 0}$  and a Finsler distance defined by

$$d_\ell(x_0, x_1) = \inf_{\gamma \in \Gamma_{0,1}(x_0, x_1)} \int_0^1 \ell(\gamma(t), \dot{\gamma}(t)) dt, \tag{1.15}$$

where  $\Gamma_{0,1}(x_0, x_1)$  denotes the set of regular curves  $\gamma$  with  $\gamma(0) = x_0$  and  $\gamma(1) = x_1$ .

**Theorem 1.6.** *There exists a neighborhood  $\Omega$  of 0 such that  $d^0(x) := d_\ell(0, x)$ , with  $d_\ell$  defined in (1.15), fulfills the generalized eikonal equation*

$$\tilde{h}_0(x, \nabla d^0(x)) = 0, \quad x \in \Omega.$$

Furthermore

$$d^0(x) - \sum_{1 \leq k \leq N} \varphi_k(x) = O(|x|^{N+1}) \quad \text{as } x \rightarrow 0, \tag{1.16}$$

where each  $\varphi_k$  is an homogeneous polynomial of degree  $k + 2$ .

In addition  $d_\ell$  is locally Lipschitz continuous, i.e.

$$|d_\ell(x, y)| \leq C |x - y|, \quad x, y \in \mathbb{R}^d, \tag{1.17}$$

where  $C$  is locally uniform in  $x$  and  $y$ .

The eikonal inequality

$$\tilde{h}_0(x, \nabla d^0(x)) \leq 0 \tag{1.18}$$

holds almost everywhere in  $\mathbb{R}^d$ .

To analyze eigenfunctions concentrated at the potential minimum  $x_0 = 0$ , we introduce a Dirichlet operator  $H_\varepsilon^\Sigma$  as follows.

**Definition 1.7.** For  $\Sigma \subset \mathbb{R}^d$  we set  $\Sigma_\varepsilon := \Sigma \cap (\varepsilon\mathbb{Z})^d$ . Any function  $u \in \ell^2(\Sigma_\varepsilon)$  can be zero extension, i.e. via  $u(x) = 0$  for  $x \notin \Sigma_\varepsilon$ , be embedded in  $\ell^2((\varepsilon\mathbb{Z})^d)$ . If we denote this embedding by  $i_{\Sigma_\varepsilon}$ , we can define the space  $\ell_{\Sigma_\varepsilon}^2 := i_{\Sigma_\varepsilon}(\ell^2(\Sigma_\varepsilon)) \subset \ell^2((\varepsilon\mathbb{Z})^d)$  and the Dirichlet operator

$$H_\varepsilon^\Sigma := \mathbf{1}_{\Sigma_\varepsilon} H_\varepsilon |_{\ell_{\Sigma_\varepsilon}^2} : \ell_{\Sigma_\varepsilon}^2 \rightarrow \ell_{\Sigma_\varepsilon}^2 \tag{1.19}$$

with domain  $\mathcal{D}(H_\varepsilon^\Sigma) = \{u \in \ell_{\Sigma_\varepsilon}^2 \mid V_\varepsilon u \in \ell_{\Sigma_\varepsilon}^2\}$ .

We now formulate our estimates of weighted  $\ell^2$ -norms of eigenfunctions of the Dirichlet operator  $H_\varepsilon^\Sigma$ . We will show that they decay exponentially at a rate controlled by the Finsler distance  $d^0(x)$ . Theorem 1.6 is crucial to prove these estimates.

**Theorem 1.8.** *Let  $\Sigma \subset \mathbb{R}^d$  be a bounded open region including the point 0 such that  $d^0 \in \mathcal{C}^2(\bar{\Sigma})$ , where  $d^0(x) := d_\ell(0, x)$  is defined by (1.15).*

*Let  $E \in [0, \varepsilon R_0]$  for  $R_0$  fixed, assume Hypothesis 1.1 and let  $H_\varepsilon^\Sigma$  denote the Dirichlet operator introduced in (1.19).*

*Then there exist constants  $\varepsilon_0, B, C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0]$  and real  $u \in \ell_{\Sigma_\varepsilon}^2$*

$$\begin{aligned} & \left\| \left(1 + \frac{d^0}{\varepsilon}\right)^{-B} e^{\frac{d^0}{\varepsilon}} u \right\|_{\ell^2} \\ & \leq C \left[ \varepsilon^{-1} \left\| \left(1 + \frac{d^0}{\varepsilon}\right)^{-B} e^{\frac{d^0}{\varepsilon}} (H_\varepsilon^\Sigma - E) u \right\|_{\ell^2} + \|u\|_{\ell^2} \right]. \end{aligned} \quad (1.20)$$

*In particular, let  $u \in \ell_{\Sigma_\varepsilon}^2$  be a normalized eigenfunction of  $H_\varepsilon^\Sigma$  with respect to the eigenvalue  $E \in [0, \varepsilon R_0]$ . Then there exist constants  $B, C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0]$*

$$\left\| \left(1 + \frac{d^0}{\varepsilon}\right)^{-B} e^{\frac{d^0}{\varepsilon}} u \right\|_{\ell^2} \leq C. \quad (1.21)$$

*Remark 1.9.* The estimate (1.20) is the analog of the sharp semiclassical Agmon estimate in Helffer–Sjöstrand [17] for the Schrödinger operator. We emphasize that our generalization in the context of general Finsler geometry is a result in the semiclassical limit, under the crucial hypothesis that both the kinetic and the potential energy have a non-degenerate minimum at  $\xi = 0, x = 0$ . In particular, we have nothing to report for an analog of the original Agmon estimate if  $\varepsilon = 1$  and  $|x| \rightarrow \infty$ .

The plan of the paper is as follows.

Section 2 is devoted to the construction and properties of a Finsler function associated to a hyperregular Hamilton function. In particular, in Subsection 2.1 we introduce the general notion of a Finsler manifold, the associated curve length and Finsler geodesics. In Subsection 2.2 we construct the absolute homogeneous Finsler function  $\ell_{h,E}$  with respect to an hyperregular Hamilton function  $h$  and a fixed energy  $E$ . In particular, we prove Theorem 2.11. The proof of Theorem 1.5 is given in Subsection 2.3. In Subsection 2.4 we prove Lemma 1.2 and we show that we can apply the results derived up to this point to the Hamilton function  $\tilde{h}_0$  defined in (1.12). Subsection 2.5 contains the proof of Theorem 1.6.

In Section 3 we show the exponential decay of the eigenfunctions of the low lying spectrum of  $H_\varepsilon$  with a rate controlled by the Finsler distance constructed in Section 2. In particular, in Subsection 3.1 we show three basic lemmata and in Subsection 3.2 we prove Theorem 1.8. In Section 4 we describe how a certain class of Markov chains fits into the framework of our hypotheses.

## 2. Finsler distance associated to $H_\varepsilon$

### 2.1. Definition and properties of Finsler manifold and Finsler metric

We introduce the general notion of a Finsler manifold and Finsler distance (for detailed description of Finsler manifolds we refer e.g. to Bao–Chern–Shen [5], Abate–Patrizio [1]).

For a manifold  $M$ ,  $\pi : TM \rightarrow M$  denotes the tangent bundle with fibre  $T_xM = \pi^{-1}(x)$ . We denote an element of  $TM$  by  $(x, v)$  where  $x \in M$  and  $v \in T_xM$ . Analogously,  $(x, \xi)$  with  $\xi \in T_x^*M$  denotes a point in the cotangent bundle  $\pi^* : T^*M \rightarrow M$ .

The canonical pairing between an element  $v \in T_xM$  and  $\xi \in T_x^*M$  is written as  $v \cdot \xi$ .

**Definition 2.1.** Let  $M$  be a  $d$ -dimensional  $\mathcal{C}^\infty$ -manifold and  $TM \setminus \{0\} := \{(x, v) \in TM \mid v \neq 0\}$  the slit tangent bundle.

- (a) A (Lagrange)-function  $F : TM \rightarrow [0, \infty)$  is called a Finsler function on  $M$ , if:
  - 1)  $F$  is of class  $\mathcal{C}^\infty(TM \setminus \{0\})$ .
  - 2)  $F(x, \lambda v) = \lambda F(x, v)$  for  $\lambda > 0$ , i.e.,  $F$  is positive homogeneous of order 1 in each fibre.
  - 3)  $g(x, v) := D_v^2 \left( \frac{1}{2} F^2 \right) |_{(x,v)}$  is positive definite as a bilinear form on  $T_xM$  for all  $(x, v) \in TM \setminus \{0\}$ .
- (b) A Finsler function  $F$  is said to be absolute homogeneous, if
  - 4)  $F(x, \lambda v) = |\lambda| F(x, v)$  for all  $\lambda \in \mathbb{R}$ ,
- (c) A manifold together with a Finsler function,  $(M, F)$ , is called a Finsler manifold.

*Remark 2.2.* (a) By Euler’s Theorem, 2) implies  $\sum_{i,j=1}^d g_{ij}(x, v) v_i v_j = F^2(x, v)$  and thus 3) implies  $F(x, v) > 0$  for  $v \neq 0$  (for details see Bao–Chern–Shen [5]).

- (b) In Agmon [3], the definition of a Finsler function is slightly more general:  $F$  is only required to be continuous and  $F(x, \cdot)$  to be positive and convex. In this paper, we use the definition of Bao–Chern–Shen [5], since the natural class of Finsler functions which we construct turn out to be Finsler in this more narrow sense. Thus we can use results of Bao–Chern–Shen [5].

A Finsler function induces a curve length on  $M$  as follows. A curve  $\gamma : [a, b] \rightarrow M$  on  $M$  is called regular, if it is  $\mathcal{C}^2$  and  $\dot{\gamma}(t) \neq 0$  for all  $t \in [a, b]$ . We introduce the Banach manifold

$$\Gamma_{a,b}(x_1, x_2) := \left\{ \gamma \in \mathcal{C}^2([a, b], M) \mid \gamma \text{ is regular and } \gamma(a) = x_1, \gamma(b) = x_2 \right\}. \tag{2.1}$$

**Definition 2.3.** For any Finsler function  $F$  on  $M$ , the curve length  $s_F : \Gamma_{a,b}(x_1, x_2) \rightarrow \mathbb{R}$  associated to  $F$  is defined as

$$s_F(\gamma) := \int_a^b F(\gamma(t), \dot{\gamma}(t)) dt.$$



For any  $\delta > 0$ , a regular variation of  $\gamma \in \Gamma_{a,b}(x_1, x_2)$  is a  $\mathcal{C}^2$ -map  $\gamma_\delta : [a, b] \times (-\delta, \delta) \rightarrow M$ , such that  $\gamma_\delta(t, 0) = \gamma(t)$  for all  $t \in [a, b]$  and  $\gamma_\delta(\cdot, u)$  is regular for each  $u \in (-\delta, \delta)$ .

Each  $\mathcal{C}^2$ -map  $\gamma_\delta : (-\delta, \delta) \rightarrow \Gamma_{a,b}(x_1, x_2)$  with  $\gamma_\delta(0) = \gamma$  can be considered as a regular variation of  $\gamma$  with fixed endpoints (i.e. with  $\gamma_\delta(a, u) = x_1$  and  $\gamma_\delta(b, u) = x_2$  for all  $u \in (-\delta, \delta)$ ). Therefore the tangent space of  $\Gamma_{a,b}(x_1, x_2)$  at a point  $\eta$  is given by

$$T_\eta \Gamma_{a,b}(x_1, x_2) = \{ \partial_u \eta_\delta|_{u=0} \mid \eta_\delta \text{ is a regular variation of } \eta \text{ with fixed endpoints} \}, \tag{2.2}$$

where  $\partial_u \eta_\delta|_{u=0}$  is considered as a vector field along  $\eta$ , i.e., as a function  $\partial_u \eta_\delta|_{u=0} : [a, b] \rightarrow TM$  such that  $\partial_u \eta_\delta|_{u=0}(t) \in T_{\eta(t)}M$ . Since the variation  $\eta_\delta$  has fixed endpoints, it follows that  $\partial_u \eta_\delta|_{u=0}(a) = \partial_u \eta_\delta|_{u=0}(b) = 0$ .

**Definition 2.4.**  $\gamma \in \Gamma_{a,b}(x_1, x_2)$  is called a geodesic with respect to the Finsler function  $F$  (or a Finsler geodesic), if  $ds_F|_\gamma = 0$ .

**Definition 2.5.** Let  $(M, F)$  denote a Finsler manifold.

- (a) The Finsler distance  $d_F(x_1, x_2) : M \times M \rightarrow [0, \infty]$  between the points  $x_1$  and  $x_2$  is defined by

$$d_F(x_1, x_2) := \inf_{\gamma \in \Gamma_{0,1}(x_1, x_2)} s_F(\gamma).$$

If  $\Gamma_{0,1}(x_1, x_2)$  is empty, the distance is defined to be infinity.

- (b) A geodesic  $\gamma$  between two points  $x_1$  and  $x_2$  is called minimal, if  $s_F(\gamma) = d(x_1, x_2)$ .

It follows easily from the definitions of a Finsler function  $F$  and the associated Finsler distance  $d_F$  that  $d_F(x_1, x_2) \geq 0$ , where equality holds if and only if  $x_1 = x_2$ . Furthermore the triangle inequality  $d_F(x_1, x_3) \leq d_F(x_1, x_2) + d_F(x_2, x_3)$  holds. If in addition the Finsler function  $F$  is absolute homogeneous, then  $d_F(x_1, x_2) = d_F(x_2, x_1)$ . Thus for an absolute homogeneous Finsler function,  $(M, d_F)$  is a metric space.

**Definition 2.6.** We denote by  $SM := TM / \sim_S$  the sphere bundle, where

$$(x, v) \sim_S (y, w), \quad \text{if } x = y \quad \text{and} \quad v = \lambda w \quad \text{for any } \lambda > 0.$$

Let  $\pi_s : TM \rightarrow SM$  denote the projection  $\pi_s(x, v) = [x, v]$ .

**2.2. The Finsler function of a hyperregular Hamilton function**

To define a Finsler distance for which an analog of Jacobi's Theorem holds, we briefly introduce the notion of fibre derivatives, hyperconvexity and hyperregularity of  $h$ . It is shown in Proposition 2.9 that hyperconvexity of  $h$  is a sufficient condition for hyperregularity.

**Definition 2.7.** (a) Let  $M$  be a manifold and  $f \in \mathcal{C}^\infty(T^*M, \mathbb{R})$ . Then for  $f_x := f|_{T_x^*M}$  the map  $\mathcal{D}_F f : T^*M \rightarrow TM$  defined by  $\mathcal{D}_F f(x, \xi) := Df_x(\xi)$  is called the fibre derivative of  $f$ .

- (b) Analogously, the fibre derivative of a function  $g \in \mathcal{C}^\infty(TM, \mathbb{R})$  is defined as  $\mathcal{D}_F g : TM \rightarrow T^*M$ ,  $\mathcal{D}_F g(x, v) := Dg_x(v)$ .
- (c) A function  $f : SM \rightarrow TM$  is called strictly fibre preserving, if  $f([x, u]) \in [x, u]$ , where  $[x, u]$  denotes the equivalence class with respect to  $\sim_S$ .
- (d) A smooth function  $h : T^*M \rightarrow \mathbb{R}$  (or  $L : TM \rightarrow \mathbb{R}$ ) is said to be hyperregular, if its fibre derivative  $\mathcal{D}_F h : T^*M \rightarrow TM$  (or  $\mathcal{D}_F L : TM \rightarrow T^*M$ ) is a diffeomorphism. For a hyperregular function  $h \in \mathcal{C}^\infty(T^*M)$ , we sometimes use the notation

$$\xi_h(x, v) := (\mathcal{D}_F h)^{-1}(x, v) \quad \text{and} \quad v_h(x, \xi) := \mathcal{D}_F h(x, \xi). \tag{2.3}$$

**Definition 2.8.** For a normed vector space  $V$  we call a function  $L \in \mathcal{C}^2(V, \mathbb{R})$  hyperconvex, if there exists a constant  $\alpha > 0$  such that

$$D^2L|_{v_0}(v, v) \geq \alpha \|v\|^2 \quad \text{for all } v_0, v \in V.$$

We recall that a strictly convex function  $L \in \mathcal{C}^2(V, \mathbb{R})$  has the properties

$$L(v_1) - L(v_2) \geq DL(v_2)(v_1 - v_2) \tag{2.4}$$

$$D^2L|_{v_0}[v, v] > 0 \tag{2.5}$$

$$(DL(v_1) - DL(v_2))(v_1 - v_2) > 0. \tag{2.6}$$

**Proposition 2.9.** *If a real valued function  $h \in \mathcal{C}^\infty(T^*M)$  is hyperconvex in each fibre  $T_x^*M$ , it is hyperregular.*

*Proof.* By definition,  $\mathcal{D}_F h$  is fibre preserving, thus in the coordinates on  $TM$  and  $T^*M$  induced from local coordinates on  $M$  at  $x_0$ , its derivative  $D\mathcal{D}_F h|_{(x_0, \xi_0)}$  is given by the  $2d \times 2d$ -matrix

$$\begin{pmatrix} \mathbf{1} & 0 \\ * & M \end{pmatrix}, \tag{2.7}$$

where  $M$  is the matrix representation of  $D_\xi^2 h|_{(x_0, \xi_0)} = D(\mathcal{D}_F h|_{T_x^*M})$ .

Since  $h$  was assumed to be hyperconvex in each fibre,  $M$  is positive definite and thus it follows from (2.7) that  $\mathcal{D}_F h$  is a local diffeomorphism.

We claim that  $Dh_x : T_x^*M \rightarrow T_x M$  is bijective for all  $x \in M$ . Since  $\mathcal{D}_F h$  is fibre preserving, this shows that  $\mathcal{D}_F h$  is a global diffeomorphism and finishes the proof. Thus we fix any  $x \in M$  and analyze  $Dh_x$ .

Since  $h_x$  is strictly convex for each  $x \in M$ , by (2.6)

$$(\xi - \eta)(Dh_x(\xi) - Dh_x(\eta)) > 0, \quad \xi, \eta \in T_x^*M, \quad \eta \neq \xi$$

holds and thus  $\mathcal{D}_F h(x, \cdot) = Dh_x$  is injective.

To show the surjectivity, we claim that, for any  $v_0 \in T_x M$ , the initial value problem

$$v_0 = \frac{d}{dt} Dh_x(\xi(t)) = D^2 h_x(\xi(t)) \cdot \dot{\xi}(t), \quad \xi(0) = 0. \tag{2.8}$$

has a solution  $\xi(t)$  for all  $t \in [0, 1]$ . Then

$$v_0 = \int_0^1 D^2h_x(\xi(t)) \cdot \dot{\xi}(t) dt = Dh_x(\xi(1)) - Dh_x(\xi(0)) = Dh_x(\xi(1))$$

and thus  $Dh_x$  is surjective.

Since  $h$  is hyperconvex, the inverse  $(D^2h_x|_{\xi(t)})^{-1}$  exists, thus (2.8) can be rewritten as

$$\dot{\xi}(t) = (D^2h_x|_{\xi(t)})^{-1} \cdot v_0, \quad \xi(0) = 0. \tag{2.9}$$

Thus (2.9) is of the form  $\dot{\xi} = F(\xi)$ , where  $F$  is locally Lipschitz. Therefore for any  $v_0 \in T_xM$ , (2.9) has a solution, which either exists for all  $t \geq 0$  or becomes infinite for a finite value of  $t$ .

In order to exclude that the curve  $\xi$  reaches infinity for some  $t < 1$ , we need the hyperconvexity of  $h$ . We choose a norm  $\|\cdot\|_{T_x^*M}$  on  $T_x^*M$  and denote by  $\|\cdot\|_{T_xM}$  the norm on  $T_xM$ , which is induced by duality. Since for fixed  $\eta \in T_x^*M = T_\xi(T_x^*M)$  the second derivative  $D^2h_x|_\xi(\eta)$  can be regarded as an element of  $T_xM$ , it follows by the hyperconvexity of  $h$  that there exists a constant  $\alpha > 0$  such that for all  $\xi \in T_x^*M$

$$\|D^2h_x|_\xi(\eta)\|_{T_xM} = \sup_{\mu \in T_x^*M} \frac{|D^2h_x|_\xi(\eta, \mu)|}{\|\mu\|_{T_x^*M}} \geq \frac{|D^2h_x|_\xi(\eta, \eta)|}{\|\eta\|_{T_x^*M}} \geq \alpha \|\eta\|_{T_x^*M}, \tag{2.10}$$

$\eta \in T_\xi(T_x^*M)$

and therefore

$$\|v\|_{T_xM} = \left\| D^2h_x|_\xi (D^2h_x|_\xi)^{-1}(v) \right\|_{T_xM} \geq \alpha \left\| (D^2h_x|_\xi)^{-1}(v) \right\|_{T_x^*M}, \tag{2.11}$$

$v \in T_xM$ .

(2.9) together with (2.11) yields

$$\|\dot{\xi}(t)\|_{T_x^*M} = \left\| (D^2h_x|_{\xi(t)})^{-1}(v_0) \right\|_{T_x^*M} \leq \frac{1}{\alpha} \|v_0\|. \tag{2.12}$$

Therefore the curve  $\xi(t)$  exists for all  $t \in [0, 1]$  and  $\mathcal{D}_F h(\xi(1)) = v_0$ . □

For any hyperregular Hamilton function  $h \in \mathcal{C}^\infty(T^*M)$ , we define the energy function  $E_h$  on  $TM$  by

$$E_h(x, v) := h \circ (\mathcal{D}_F h)^{-1}(x, v) \left( = h(x, \xi_h(x, v)) \right) \tag{2.13}$$

and the action

$$A_h : TM \rightarrow \mathbb{R}, \quad A_h(x, v) := (\mathcal{D}_F h)^{-1}(x, v) \cdot v \left( = \xi_h(x, v) \cdot v \right). \tag{2.14}$$

Then the Lagrange function

$$L_h : TM \rightarrow \mathbb{R} \quad \text{defined by} \quad L_h(x, v) = A_h(x, v) - E_h(x, v) \tag{2.15}$$

(the Legendre transform of  $h$ ) is hyperregular on  $TM$  and

$$\mathcal{D}_F L_h(x, v) = (\mathcal{D}_F h)^{-1}(x, v) \tag{2.16}$$

(by Theorem 3.6.9 in [2], the hyperregular Lagrange functions on  $TM$  and the hyperregular Hamilton functions on  $T^*M$  are in bijection). In particular, by (2.14) and (2.16),

$$A_h(x, v) = \mathcal{D}_F L_h(x, v) \cdot v. \tag{2.17}$$

**Definition 2.10.** For a smooth manifold  $M$ , a  $\mathcal{C}^\infty$ -function  $h : T^*M \rightarrow \mathbb{R}$  and  $E \in \mathbb{R}$ , we define the singular set  $S_h(E)$  by

$$S_h(E) := \{x \in M \mid h(x, 0) \geq E\}.$$

Since  $h(\cdot, 0)$  is continuous,  $S_h(E)$  is closed. Thus  $\widetilde{M} := M \setminus S_h(E)$  is again a smooth manifold.

For  $h$  hyperregular, we shall now introduce an associated Finsler function on  $\widetilde{M}$ .

**Theorem 2.11.** *Let  $M$  and  $h \in \mathcal{C}^\infty(T^*M)$  satisfy Hypothesis 1.4 and let  $E, S_h(E)$  and  $\widetilde{M}$  be as described in Definition 2.10.*

- i) *Then there exists a strictly fibre preserving  $\mathcal{C}^\infty$ -function  $\tau_E : \widetilde{S}M \rightarrow T\widetilde{M}$ , which is uniquely determined by the condition*

$$h \circ (\mathcal{D}_F h)^{-1} \circ \tau_E = E. \tag{2.18}$$

- ii) *Let  $\tilde{\tau}_E := \tau_E \circ \pi_S : T\widetilde{M} \rightarrow T\widetilde{M}$  and let  $\ell_{h,E} : T\widetilde{M} \rightarrow \mathbb{R}$  be defined by*

$$\ell_{h,E}(x, v) := (\mathcal{D}_F h)^{-1} \circ \tilde{\tau}_E(x, v) \cdot v.$$

*Then  $\ell_{h,E}$  is an absolute homogeneous Finsler function on  $\widetilde{M}$ .*

- iii) *For any regular curve  $\gamma : [a, b] \rightarrow \widetilde{M}$ , there exists a unique  $\mathcal{C}^1$ -function  $\lambda : [a, b] \rightarrow \mathbb{R}_+$  such that*

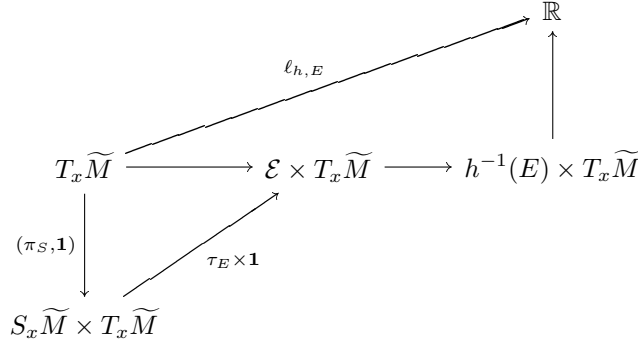
$$\tilde{\tau}_E(\gamma(t), \dot{\gamma}(t)) = (\gamma(t), \lambda(t)\dot{\gamma}(t)). \tag{2.19}$$

*Remark 2.12.* (a) Since  $\ell_{h,E}$  is defined on  $\widetilde{M} = M \setminus S(E)$ , we call  $(M, \ell_{h,E})$  a Finsler manifold with singularities.

- (b) If we continuously extend  $\ell_{h,E}$  from  $\widetilde{M}$  to  $M$  by setting  $\ell_{h,E}(x, v) = 0$  for  $x \in S(E)$ , the associated distance  $d_\ell$  is well defined on all of  $M$ . Nevertheless contrary to the case of a Finsler manifold without singularities (as described for example in Bao–Chern–Shen [5]), the geodesic curves with respect to  $\ell_{h,E}$  may have kinks at the “singular points”, which are the connected components of  $S_h(E)$ .

- (c) Geometrically, the function  $\tilde{\tau}_E$  projects an element  $(x, v)$  of the tangent bundle  $T\widetilde{M}$  to an element  $(x, \lambda v)$  in the  $(2d - 1)$ -dimensional submanifold  $\mathcal{E} = E_h^{-1}(E)$ .

(d) Schematically the functions occurring in Theorem 2.11 are illustrated in the following diagram.



$$v \xrightarrow{(\tilde{\tau}_E, \mathbf{1})} (\tilde{v}, v) \xrightarrow{(\mathcal{D}_F h)^{-1} \times \mathbf{1}} (\xi_h(\tilde{v}), v)$$

(e) With the notation (2.3),  $\ell_{h,E}(x, v)$  can be written as

$$\ell_{h,E}(x, v) = \xi_h(x, \tilde{v}) \cdot v \quad \text{where} \quad (x, \tilde{v}) = \tilde{\tau}_E(x, v) \in \mathcal{E}. \tag{2.20}$$

To prove Theorem 2.11, we need the following lemma.

**Lemma 2.13.** *In the setting of Theorem 2.11 fix  $x \in \widetilde{M}$  and  $u \in T_x \widetilde{M} \setminus \{0\}$ . Then for  $E_h : TM \rightarrow \mathbb{R}$  defined by (2.13), the function*

$$E_u : [0, \infty) \rightarrow \mathbb{R}, \quad E_u(\lambda) := E_h(x, \lambda u)$$

*is strictly increasing with  $\frac{d}{d\lambda} E_u > 0$  for  $\lambda > 0$ . Furthermore  $E_u(0) \leq E$  and  $\lim_{\lambda \rightarrow \infty} E_u(\lambda) = \infty$ .*

*Proof of Lemma 2.13.* Since  $h_x$  is even,  $Dh_x(0) = 0$ , thus  $v_h(x, 0) = 0$ ,  $\xi_h(x, 0) = 0$  and

$$E_u(0) = E_h(x, 0) = h(x, 0) < E. \tag{2.21}$$

To show that  $E_u$  is strictly increasing, we will analyze the derivative of  $E_u$  for  $\lambda > 0$ .

By definition  $\mathcal{D}_F h(x, \xi) = Dh_x(\xi)$ , thus

$$\frac{dE_u}{d\lambda} \Big|_{\lambda} = Dh_x|_{(Dh_x)^{-1}(\lambda u)} \cdot D(Dh_x)^{-1} \Big|_{\lambda u}(u). \tag{2.22}$$

We notice that

$$Dh_x|_{(Dh_x)^{-1}(\lambda u)} = \lambda u \tag{2.23}$$

and for  $L_h$  defined in (2.15) it follows from (2.16) that

$$D(Dh_x)^{-1} \Big|_{\lambda u}(u) = D(DL_{h,x}) \Big|_{\lambda u}(u). \tag{2.24}$$

Inserting (2.24) and (2.23) in (2.22) yields

$$\frac{dE_u}{d\lambda}|_\lambda = \lambda u \cdot D(DL_{h,x})|_{\lambda u}(u) = \lambda D^2L_{h,x}|_{\lambda u}(u, u), \tag{2.25}$$

where we identify linear maps from  $T_xM$  to  $T_x^*M$  with bilinear forms on  $T_xM$ . If  $h$  is strictly convex in each fibre, the same is true for  $L_h$ . Therefore by (2.25) the first derivative of  $E_u$  is strictly positive for  $\lambda \in (0, \infty)$  and thus  $E_u$  is strictly increasing.

The fact that  $\lim_{\lambda \rightarrow \infty} E_u(\lambda) = \infty$  can be seen as follows. From the strict convexity of  $h$  and since  $h(x, \xi) \geq h(x, 0)$  for all  $\xi \in T_x^*M$ , it follows that  $\lim_{\|\xi\| \rightarrow \infty} h(x, \xi) = \infty$ . Since  $h$  is hyperregular,  $\mathcal{D}_F h(x, \cdot) = Dh_x : T_xM \rightarrow T_x^*M$  is a global diffeomorphism. Thus for any norm  $\|\cdot\|_{T_xM}$  on  $T_xM$  and the induced norm  $\|\cdot\|_{T_x^*M}$  on  $T_x^*M$ , we have  $\|Dh_x(v_n)\|_{T_x^*M} \rightarrow \infty$  for any sequence  $(v_n)$  in  $T_xM$  satisfying  $\|v_n\|_{T_xM} \rightarrow \infty$  (any global diffeomorphism is proper). Thus

$$\lim_{\lambda \rightarrow \infty} E_u(\lambda) = \lim_{\lambda \rightarrow \infty} h(x, \xi_h(x, \lambda u)) = \lim_{\|\xi\| \rightarrow \infty} h(x, \xi) = \infty. \quad \square$$

*Proof of Theorem 2.11.* i) From Lemma 2.13 it follows that for fixed  $x \in \widetilde{M}$ , each ray  $[x, u] \in S\widetilde{M}$  intersects the  $\mathcal{E}_x := E_h^{-1}(E) \cap T_x\widetilde{M}$  in exactly one point  $v$ , i.e. for each ray  $[x, u]$  there is exactly one point  $(x, v) \in T_x\widetilde{M}$  such that  $E_h(x, v) = E$ . Thus  $[x, u] \mapsto (x, v)$  defines a map  $\tau_E$ , which is uniquely determined by (2.18). Clearly  $\tau_E$  is strictly fibre preserving.

To analyze the regularity of  $\tau_E$ , we will use the Implicit Function Theorem. We remark that there exists an isomorphism  $\psi : T\widetilde{M} \setminus \{0\} \rightarrow S\widetilde{M} \times \mathbb{R}_+$ , such that  $\psi^{-1}(\{[x, u]\} \times \mathbb{R}_+) = [x, u] \subset T\widetilde{M} \setminus \{0\}$ . Setting  $\widehat{E}_h := E_h \circ \psi^{-1}$  and  $\widehat{\tau}_E := \psi \circ \tau_E$ , we get  $\widehat{\tau}_E(s) = (s, \lambda(s))$ ,  $s \in S\widetilde{M}$ , for some  $\lambda : S\widetilde{M} \rightarrow \mathbb{R}_+$  and therefore  $\widehat{E}_h(s, \lambda(s)) = E$ . By Lemma 2.13,  $\frac{d\widehat{E}_h}{d\lambda}(s_0, \lambda_0) > 0$  for all  $(s_0, \lambda_0) \in S\widetilde{M} \times \mathbb{R}_+$ . Smoothness of  $\lambda$  and thus of  $\tau_E$  now follows from the Implicit Function Theorem.

iii) By i), for  $t \in [a, b]$  fixed, there is a unique  $\lambda(t) \in \mathbb{R}_+$  with (2.19). Since  $\tau_E$  is smooth (as composition of smooth maps),  $t \mapsto (\gamma(t), \lambda(t)\dot{\gamma}(t))$  is a  $\mathcal{C}^1$ -curve. Thus  $\lambda \in \mathcal{C}^1([a, b], \mathbb{R})$ .

ii) To show that  $\ell_{h,E} : T\widetilde{M} \rightarrow \mathbb{R}$  is a Finsler function on  $\widetilde{M}$ , we check the defining properties.

- 1) The regularity  $\ell_{h,E} \in \mathcal{C}^\infty(T\widetilde{M} \setminus \{0\})$  follows from the fact that  $h$  is hyperregular and the function  $\widehat{\tau}_E$  is  $\mathcal{C}^\infty$ .
- 2) To show  $\ell_{h,E}(x, \lambda v) = \lambda \ell_{h,E}(x, v)$ , ( $\lambda > 0$ ), we notice that by construction  $\widehat{\tau}_E(x, \lambda v) = \widehat{\tau}_E(x, v)$  for any  $\lambda > 0$ . Thus  $(\mathcal{D}_F h)^{-1} \circ \widehat{\tau}_E$  is homogeneous of order zero in each fibre. Since  $\xi \cdot v$  is bilinear, it follows that

$$\ell_{h,E}(x, \lambda v) = (\mathcal{D}_F h)^{-1} \circ \widehat{\tau}_E(x, \lambda v) \cdot \lambda v = \lambda \ell_{h,E}(x, v).$$

- 3) We start showing that  $\ell_{h,E}(x, v) > 0$ , ( $v \neq 0$ ). To this end we define

$$a_h = A_h \circ \mathcal{D}_F h : T^*M \rightarrow \mathbb{R}, \quad a_h(x, \xi) = \xi \cdot \mathcal{D}_F h(x, \xi).$$

Since  $h$  was assumed to be strictly convex in each fibre, one obtains from (2.6)

$$(\xi - \eta) \cdot (\mathcal{D}_F h(x, \xi) - \mathcal{D}_F h(x, \eta)) > 0, \quad \xi, \eta \in T_x^* M, \quad \eta \neq \xi.$$

Therefore choosing  $\xi = -\eta$  and using that  $h$  is even in each fibre (thus  $\mathcal{D}_F h$  is odd) yields

$$2\xi \cdot (\mathcal{D}_F h(x, \xi) - \mathcal{D}_F h(x, -\xi)) = 4\xi \cdot \mathcal{D}_F h(x, \xi) = 4a_h(x, \xi) > 0, \quad \text{for } \xi \neq 0. \quad (2.26)$$

Since  $h$  is even and strictly convex, it takes its absolute minimum at  $\xi = 0$  and thus  $\mathcal{D}_F h(x, 0) = 0$ . Since furthermore  $\mathcal{D}_F h$  is a global diffeomorphism, we get  $(\mathcal{D}_F h)^{-1}(x, v) \neq 0$  for  $v \neq 0$ . By (2.26)

$$A_h(x, v) = a_h(x, (\mathcal{D}_F h)^{-1}(x, v)) > 0, \quad v \neq 0. \quad (2.27)$$

Setting  $\tilde{\tau}_E(x, v) =: (x, \tilde{v})$ , it follows from the fact that  $\tau_E$  is strictly fibre preserving that there exists a  $\lambda > 0$  such that  $v = \lambda\tilde{v}$ . Thus by (2.27) for  $v \neq 0$

$$\ell_{h,E}(x, v) = (\mathcal{D}_F h)^{-1}(x, \tilde{v}) \cdot \lambda\tilde{v} = \lambda A_h(x, \tilde{v}) > 0. \quad (2.28)$$

To show that the matrix  $g$  is positive definite, we set  $\ell_x(v) := \ell_{h,E}(x, v) : T_x \widetilde{M} \rightarrow \mathbb{R}$ . Then we have to show that  $g_{(x,v)}(w, w) = D^2(\frac{1}{2}\ell_x^2)|_v(w, w) > 0$  for all  $v, w \in T_x \widetilde{M} \setminus \{0\}$ .

We first remark that

$$g_{(x,v)}(w, w) = T_1(w) + T_2(w) \quad \text{where} \quad (2.29)$$

$$T_1(w) = \ell_x(v) D^2 \ell_x|_v(w, w) \quad \text{and} \quad T_2(w) = (D\ell_x|_v(w))^2.$$

By the definition of  $\ell_x$  and (2.16) it follows that

$$D\ell_x|_v(w) = D^2 L_x|_{\tilde{\tau}_E(v)}(D\tilde{\tau}_E|_v(w), v) + (Dh_x)^{-1} \circ \tilde{\tau}_E(v) \cdot w. \quad (2.30)$$

To analyze  $D\tilde{\tau}_E$ , we use that by Lemma 2.13 the function  $E_{h,x}$  is strictly increasing in each fibre. Therefore, analogue to the proof of i), there exists a smooth function  $\mu_x : T_x M \rightarrow (0, \infty)$  such that  $\tilde{\tau}_E(v) = \mu_x(v)v$  and thus

$$D\tilde{\tau}_E|_v(w) = v \cdot D\mu|_v(w) + \mu(v)w. \quad (2.31)$$

Since

$$E = E_{h,x}(\tilde{\tau}_E)(v) = E_{h,x} \circ G_x(v, \mu(v)) \quad \text{where} \quad G_x(v, \mu(v)) := \mu(v)v$$

we have  $D_v(E_{h,x} \circ G_x)(v, \mu(v)) = 0$ , leading to

$$D\mu|_v(w) = - (D_2(E_{h,x} \circ G_x)|_{(v, \mu(v))})^{-1} D_1(E_{h,x} \circ G_x)|_{(v, \mu(v))}(w). \quad (2.32)$$

Since by the definition of  $E_h$  in (2.13) and again by (2.16) we have

$$DE_{h,x}|_v(w) = Dh_x|_{(Dh_x)^{-1}(v)} \cdot D(Dh_x)^{-1}|_v(w) = D^2 L|_v(v, w), \quad (2.33)$$

it follows that

$$D_2(E_{h,x} \circ G_x)(v, \mu) = DE_{h,x}|_{\mu v}(v) = \mu D^2 L|_v(v, v). \quad (2.34)$$

By (2.33) it follows at once that

$$\begin{aligned} D_1(E_{h,x} \circ G_x)|_{(v,\mu(v))}(w) &= DE_{h,x}|_{\mu(v)v}D_1G_x|_{(\mu(v),v)}(w) \\ &= \mu^2(v)D^2L_x|_{\mu(v)v}(v, w). \end{aligned} \tag{2.35}$$

Since  $L$  is strictly convex, the symmetric bilinear form  $D^2L_x|_{\mu(v)v}(\cdot, \cdot)$  is positive definite and therefore defines a scalar product

$$\langle v, w \rangle_L := D^2L_x|_{\mu(v)v}(v, w) \quad \text{with associated norm} \quad \|v\|_L. \tag{2.36}$$

Inserting (2.35) and (2.34) in (2.32) and using (2.36) gives

$$D\mu|_v(w) = -\frac{\mu(v)}{\|v\|_L^2} \langle v, w \rangle_L. \tag{2.37}$$

Thus inserting (2.37) in (2.31) and the resulting term in (2.30) yields

$$\begin{aligned} D\ell_x|_v(w) &= \left\langle \left[ -v \frac{\mu(v)}{\|v\|_L^2} \langle v, w \rangle_L + \mu(v)w \right], v \right\rangle_L + (Dh_x)^{-1}|_{\mu(v)v} \cdot w \\ &= (Dh_x)^{-1}|_{\mu(v)v} \cdot w. \end{aligned} \tag{2.38}$$

Using (2.38) and (2.16) we get

$$D^2\ell_x|_v(w, w) = \langle (vD\mu|_v(w) + \mu(v)w), w \rangle_L$$

and inserting (2.32) gives

$$D^2\ell_x|_v(w, w) = \mu(v) \left\{ \langle w, w \rangle_L - \frac{(\langle v, w \rangle_L)^2}{\|v\|_L^2} \right\} \geq 0, \tag{2.39}$$

where the last estimate follows from the Cauchy–Schwarz Inequality. Since  $T_2$  is quadratic, (2.39) together with (2.28) gives

$$D^2\left(\frac{1}{2}\ell_x^2\right)|_v(w, w) \geq 0, \quad v, w \in T_x\widetilde{M}.$$

To prove the strict positivity, we now fix  $x \in \widetilde{M}$  and  $v \in T_x\widetilde{M} \setminus \{0\}$ . Assuming

$$g_{(x,v)}(w, w) = 0 \quad \text{for} \quad w \in T_x\widetilde{M}, \tag{2.40}$$

we have to show that  $w = 0$ . By (2.29) it follows from (2.40) that  $T_1(w) = 0$  and  $T_2(w) = 0$ . We have already seen in (2.28) that  $\ell_{h,E}(x, v) > 0$  for  $v \neq 0$ , thus  $T_1(w) = 0$  implies  $g_{(x,v)}(w, w) = 0$ , leading by (2.39) and the Cauchy–Schwarz-inequality to

$$w = \eta v \quad \text{for some} \quad \eta \in \mathbb{R}.$$

Inserting this in  $T_2$ , the homogeneity of  $\ell_x$  shows

$$0 = D\ell_x|_v(\eta v) = \eta D\ell_x|_v(v) = \eta \ell_x(v)$$

and thus by the positivity of  $\ell_x$  we get  $\eta = 0$  and thus  $w = 0$ .



4) It remains to show that  $\ell_{h,E}$  is absolute homogeneous of order one. Since  $h$  is even in each fibre,  $\mathcal{D}_F h$  and  $(\mathcal{D}_F h)^{-1}$  are odd. Thus for  $(x, v) \in \mathcal{E}_x$

$$h \circ (\mathcal{D}_F h)^{-1}(x, -v) = h \circ (\mathcal{D}_F h)^{-1}(x, v) = E \tag{2.41}$$

and (2.41) yields

$$(x, v) \in \mathcal{E} \implies (x, -v) \in \mathcal{E}. \tag{2.42}$$

Since  $\tau_E$  is strictly fibre preserving, we have for  $(x, v) \in T_x \widetilde{M}$  and some  $\lambda > 0$ , using (2.42)

$$\tilde{\tau}_E(-v) = \lambda(-v) = -\tilde{\tau}_E(v). \tag{2.43}$$

By the fact that  $(\mathcal{D}_F h)^{-1}$  is odd and (2.43) we can conclude that

$$\ell_{h,E}(-v) = (\mathcal{D}_F h)^{-1} \circ \tilde{\tau}_E(-v) \cdot (-v) = (\mathcal{D}_F h)^{-1} \circ \tilde{\tau}_E(v) \cdot v = \ell_{h,E}(v). \tag{2.44}$$

By (2.44),  $\ell_{h,E}$  is even in each fibre and thus for any  $\lambda \in \mathbb{R}$

$$\ell_{h,E}(x, \lambda v) = \ell_{h,E}(x, |\lambda|v) = |\lambda| \ell_{h,E}(x, v). \quad \square$$

**2.3. Proof of Theorem 1.4**

Step 4 of our proof is adapted from Abraham–Marsden [2] and uses the Maupertuis principle (at least implicitly).

*Step 1:*

We will show that

$$\Gamma(x_1, x_2, [a, b], E) := \left\{ (\gamma, \alpha) \mid \alpha : [a, b] \rightarrow \mathbb{R} \text{ is } \mathcal{C}^2, \frac{d\alpha}{dt} > 0, \alpha(a) = 0, \right. \\ \left. \gamma \in \Gamma_{0,\alpha(b)}(x_1, x_2) \text{ such that} \right. \\ \left. E_h(\gamma(\alpha(t)), \dot{\gamma}(\alpha(t))) = E \text{ for all } t \in [a, b] \right\}, \tag{2.45}$$

is a Banach manifold, where  $\Gamma_{a,b}(x_1, x_2)$  was introduced in (2.1).

$\Gamma(x_1, x_2, [a, b], E)$  is the set of all pairs  $(\gamma, \alpha)$ , where  $\gamma$  is a regular curve on  $\widetilde{M}$  joining the points  $x_1$  and  $x_2$  and  $\alpha$  is a change of parameter, ensuring that the curve  $(\gamma \circ \alpha, \dot{\gamma} \circ \alpha) \in TM$  (which is not equal to the lifted curve  $(\gamma \circ \alpha, \frac{d}{dt}(\gamma \circ \alpha))$ ) lies on the energy shell  $\mathcal{E} = E_h^{-1}(E)$ .

Set  $A := \{\alpha : [a, b] \rightarrow \mathbb{R} \mid \frac{d\alpha}{dt} > 0 \text{ and } \alpha(a) = 0\}$  and denote by  $\Gamma_{0,\infty}$  the space of all regular curves  $\gamma : [0, \infty) \rightarrow \widetilde{M}$ . Then  $\Gamma_{0,\infty} \times A$  is a Banach manifold. We consider the  $\mathcal{C}^1$ -mapping

$$g : \Gamma_{0,\infty} \times A \rightarrow \widetilde{M} \times \widetilde{M}, \quad (\gamma, \alpha) \mapsto (\gamma \circ \alpha(a), \gamma \circ \alpha(b)).$$

Then  $(x_1, x_2) \in \widetilde{M} \times \widetilde{M}$  is a regular value of  $g$  and

$$\Gamma([a, b], x_1, x_2) := g^{-1}(x_1, x_2) = \left\{ (\gamma, \alpha) \in \Gamma_{0,\infty} \times A \mid \gamma(\alpha(a)) = x_1, \gamma(\alpha(b)) = x_2 \right\}$$

is a submanifold of  $\Gamma_{0,\infty} \times A$ . This follows from the fact that the Inverse Function Theorem holds in Banach manifolds (see Hamilton [13]). We introduce

$$\tilde{E}_h : \mathcal{C}^1([a, b], T\tilde{M}) \rightarrow \mathcal{C}^1([a, b], \mathbb{R}), \quad \tilde{E}_h(\eta)(t) := E_h(\eta(t))$$

and for

$$\Gamma_{a,b}^{T\tilde{M}}(x_1, x_2) := \left\{ (\gamma, k\dot{\gamma}) \in \mathcal{C}^1([a, b], T\tilde{M}) \mid \gamma \in \Gamma_{a,b}(x_1, x_2), k \in \mathcal{C}^1([a, b], \mathbb{R}_+) \right\}$$

we set

$$\Phi : \Gamma([a, b], x_1, x_2) \rightarrow \Gamma_{a,b}^{T\tilde{M}}(x_1, x_2), \quad (\gamma, \alpha) \mapsto (\gamma \circ \alpha, \dot{\gamma} \circ \alpha).$$

Then  $\Phi$  is a diffeomorphism. In fact, by a straightforward calculation it is bijective, with inverse  $\Phi^{-1}(\eta, k\dot{\eta}) = (\eta \circ \alpha^{-1}, \alpha)$ , where  $\alpha(t) = \int_a^t (k(s))^{-1} ds$ .

Identifying  $E$  with the constant function  $E(t) = E$ , we obtain  $\Gamma(x_1, x_2, [a, b], E) = f^{-1}(E)$  for  $f := \tilde{E}_h \circ \Phi$ . To show that  $E$  is a regular value of  $f$  it is sufficient to show that it is a regular value of  $\tilde{E}_h$ , i.e., that for each  $v \in \mathcal{C}^1([a, b], \mathbb{R})$  (considered as a vector field along  $E \in \mathcal{C}^1([a, b], \mathbb{R})$ ), there is a vector field  $X$  along  $\eta \in \tilde{E}_h^{-1}(E)$  with  $d\tilde{E}_h|_{\eta} X = v$ . Note that  $\eta(t) = (x(t), v(t))$  with  $v(t) \neq 0$ .

Since  $DE_h(x, v) \neq 0$  for  $v \neq 0$ , there is a covering of  $[a, b]$  by open intervals  $I_j, j \in J$  and vector fields  $X_j$  along  $\eta|_{I_j}$  with  $DE_h|_{\eta(t)} X_j(t) = v(t)$  for all  $t \in I_j$ . Choosing a partition of unity  $(\chi_j)$  subordinate to  $(I_j)$ , we set  $X(t) = \sum_{j \in J} \chi_j(t) X_j(t)$ . Then  $d\tilde{E}_h|_{\eta} X = v$ . Thus  $E$  is a regular value of  $f$  and  $\Gamma([a, b], x_1, x_2, E)$  is a Banach manifold.

*Step 2:*

We construct a diffeomorphism  $b_E : \Gamma_{a,b}(x_1, x_2) \rightarrow \Gamma(x_1, x_2, [a, b], E)$ .

By Theorem 2.11, there exists for any  $\eta \in \Gamma_{a,b}(x_1, x_2)$  a unique  $\mathcal{C}^1$ -function  $\lambda : [a, b] \rightarrow \mathbb{R}_+$  such that

$$E_h(\eta(t), \lambda(t)\dot{\eta}(t)) = E.$$

Set

$$\alpha(t) := \int_a^t \frac{1}{\lambda(s)} ds \quad \text{and} \quad \gamma = \eta \circ \alpha^{-1} : [0, \alpha(b)] \rightarrow \tilde{M}, \tag{2.46}$$

then  $\alpha : [a, b] \rightarrow \mathbb{R}$  with  $\dot{\alpha} > 0$ . From  $\dot{\eta}(t) = \dot{\gamma}(\alpha(t)) \cdot \dot{\alpha}(t)$  it follows that

$$E_h(\gamma(\alpha(t)), \dot{\gamma}(\alpha(t))) = E_h(\eta(t), \lambda(t)\dot{\eta}(t)) = E, \tag{2.47}$$

i.e.  $(\gamma, \alpha) \in \Gamma(x_1, x_2, [a, b], E)$ . We can conclude that there is a bijection between Banach manifolds given by

$$\begin{aligned} b_E : \Gamma_{a,b}(x_1, x_2) &\rightarrow \Gamma(x_1, x_2, [a, b], E) \\ b_E(\eta) &= (\eta \circ \alpha^{-1}, \alpha) \quad \text{with} \\ \alpha(t) &:= \int_a^t (\lambda(s))^{-1} ds \quad \text{for} \quad \tilde{\tau}_E(\eta, \dot{\eta}) = (\eta, \lambda\dot{\eta}). \end{aligned} \tag{2.48}$$

On the other hand, if we start with  $(\gamma, \alpha) \in \Gamma(x_1, x_2, [a, b], E)$ , then  $E_h(\gamma(s), \dot{\gamma}(s)) = E$  with  $s = \alpha(t)$ . Setting  $\eta := \gamma \circ \alpha : [a, b] \rightarrow \widetilde{M}$  it follows from  $\dot{\gamma}(s) = \dot{\eta}(t)(\dot{\alpha}(t))^{-1}$  that

$$E_h\left(\eta(t), (\dot{\alpha}(t))^{-1} \dot{\eta}(t)\right) = E \quad \text{and thus} \quad \lambda(t) = (\dot{\alpha}(t))^{-1}. \tag{2.49}$$

Thus the inverse function  $b_E^{-1}$  is given by  $b_E^{-1}(\gamma, \alpha) = \gamma \circ \alpha$ .

*Step 3:*

We show that the critical points of the length functional  $s_{\ell_h}$  defined in Definition 2.3 (i.e., the geodesics of  $\ell_h$ ) are in bijection with the critical points of the action integral

$$I : \Gamma(x_1, x_2, [a, b], E) \rightarrow \mathbb{R}, \quad I(\gamma, \alpha) := \int_{\alpha(a)}^{\alpha(b)} A_h(\gamma(s), \dot{\gamma}(s)) ds, \tag{2.50}$$

where  $A_h$  denotes the action with respect to  $h$  defined in (2.14). Setting  $s = \alpha(t)$  and using (2.3) gives

$$\int_{\alpha(a)}^{\alpha(b)} A_h(\gamma(s), \dot{\gamma}(s)) ds = \int_a^b \xi_h\left(\gamma(\alpha(t)), \dot{\gamma}(\alpha(t))\right) \cdot \dot{\gamma}(\alpha(t)) \dot{\alpha}(t) dt. \tag{2.51}$$

Setting  $\eta(t) = \gamma(\alpha(t))$  and using (2.49) and the definition of  $\ell_h$  and  $s_{\ell_h}$ , we obtain from (2.50) and (2.51)

$$\begin{aligned} I(\gamma, \alpha) &= \int_a^b \xi_h\left(\eta(t), \dot{\eta}(t)(\dot{\alpha}(t))^{-1}\right) \cdot \dot{\eta}(t) dt = \int_a^b \xi_h \circ \tilde{\tau}_E(\eta(t), \dot{\eta}(t)) \cdot \dot{\eta}(t) dt \\ &= \int_a^b \ell_{h,E}(\eta(t), \dot{\eta}(t)) dt = s_{\ell_h}(\eta). \end{aligned}$$

Since  $b_E(\gamma \circ \alpha) = (\gamma, \alpha)$ , it follows that

$$s_{\ell_h} = I \circ b_E \quad \text{and thus} \quad ds_{\ell_h}|_{\eta} = dI|_{b_E(\eta)} \circ db_E|_{\eta}. \tag{2.52}$$

Since  $b_E$  is a diffeomorphism, we get

$$ds_{\ell_h}|_{\eta} = 0 \iff dI|_{b_E(\eta)} = 0 \quad \eta \in \Gamma_{a,b}(x_1, x_2). \tag{2.53}$$

*Step 4:*

We show (a).

We set  $\gamma_0(a) = x_1, \gamma_0(b) = x_2$ . If  $\gamma_0$  is a base integral curve of the Hamiltonian vector field  $X_h$  with  $E_h(\gamma_0(t), \dot{\gamma}_0(t)) = E$  for all  $t \in [a, b]$ , then  $b_E(\gamma_0) = (\gamma_0, \mathbf{1})$ , where  $\mathbf{1} : [a, b] \rightarrow [a, b]$  is defined by  $\mathbf{1}(t) = t$ .

Thus by (2.53) it remains to show that  $dI|_{(\gamma_0, \mathbf{1})} = 0$  for any base integral curve  $\gamma_0 \in \Gamma_{a,b}(x_1, x_2)$  of the Hamiltonian vector field  $X_h$  with energy  $E$ . The tangent space of  $\Gamma(x_1, x_2, [a, b], E)$  at a point  $(\gamma, \alpha)$  can be described by use of

variations as

$$T_{(\gamma,\alpha)}\Gamma(x_1, x_2, [a, b], E) = \left\{ \partial_u(\gamma, \alpha)_\delta|_{u=0} \mid (\gamma, \alpha)_\delta : (-\delta, \delta) \rightarrow \Gamma(x_1, x_2, [a, b], E) \right. \\ \left. \text{is } \mathcal{C}^2 \text{ with } (\gamma, \alpha)_\delta(0) = (\gamma, \alpha) \right\}. \quad (2.54)$$

We start analyzing  $dI|_{(\gamma,\alpha)}$ . We use the notation  $(\gamma(\cdot), \alpha(\cdot))_\delta(u) =: (\gamma_\delta(\cdot, u), \alpha_\delta(\cdot, u))$ . From (2.54) and (2.45) it follows that  $\alpha_\delta(a, u) = 0$ . Furthermore  $\gamma_\delta(0, u) = x_1$  and  $\gamma_\delta(\alpha_\delta(b, u), u) = x_2$  for all  $u \in (-\delta, \delta)$ . This leads to

$$\frac{d}{du}\gamma_\delta(0, u) = 0 = \frac{d}{du}\gamma_\delta(\alpha_\delta(b, u), u). \quad (2.55)$$

Since  $A_h = L_h + E_h$  by the definition (2.15) of the Lagrange function  $L_h$ , it follows from the definition (2.45) of  $\Gamma(x_1, x_2, [a, b], E)$  that

$$A_h\left(\gamma_\delta(\alpha_\delta(t, u), u), \dot{\gamma}_\delta(\alpha_\delta(t, u), u)\right) = L_h\left(\gamma_\delta(\alpha_\delta(t, u), u), \dot{\gamma}_\delta(\alpha_\delta(t, u), u)\right) + E, \quad (2.56)$$

thus the definition (2.50) of  $I$  and (2.56) yield

$$dI|_{(\gamma,\alpha)}(\partial_u(\gamma, \alpha)_\delta|_{u=0}) = \partial_u I((\gamma_\delta, \alpha_\delta))|_{u=0} \\ = \frac{d}{du} \int_{\alpha_\delta(a, u)}^{\alpha_\delta(b, u)} \left( L_h(\gamma_\delta(s, u), \dot{\gamma}_\delta(s, u)) + E \right) ds \Big|_{u=0}. \quad (2.57)$$

We get using  $\gamma_\delta(t, 0) = \gamma(t)$  and  $\alpha_\delta(t, 0) = \alpha(t)$

$$\frac{d}{du} \int_{\alpha_\delta(a, u)}^{\alpha_\delta(b, u)} \left( L_h(\gamma_\delta(s, u), \dot{\gamma}_\delta(s, u)) + E \right) ds \Big|_{u=0} \\ = \left[ \left( L_h(\gamma(\alpha(t)), \dot{\gamma}(\alpha(t))) + E \right) \cdot \partial_u \alpha_\delta|_{u=0}(t) \right]_a^b \\ + \int_0^{\alpha(b)} \frac{d}{du} L_h(\gamma_\delta(s, u), \dot{\gamma}_\delta(s, u)) \Big|_{u=0} ds. \quad (2.58)$$

For the integrand on the right hand side of (2.58) we get

$$\frac{d}{du} L_h(\gamma_\delta(s, u), \dot{\gamma}_\delta(s, u)) \Big|_{u=0} = D_\gamma L_h(\gamma(s), \dot{\gamma}(s)) \cdot \partial_u \gamma_\delta|_{u=0}(s) \\ + D_{\dot{\gamma}} L_h(\gamma(s), \dot{\gamma}(s)) \partial_u \dot{\gamma}_\delta|_{u=0}(s). \quad (2.59)$$

Since  $\partial_u \dot{\gamma}_\delta|_{u=0}(s) = \partial_s \partial_u \gamma_\delta|_{u=0}(s)$ , integration by parts for the second summand on the right hand side of (2.59) gives

$$\int_{\alpha(a)}^{\alpha(b)} \frac{d}{du} L_h(\gamma_\delta(s, u), \dot{\gamma}_\delta(s, u)) \Big|_{u=0} ds = \left[ D_\gamma L_h(\gamma(s), \dot{\gamma}(s)) \cdot \partial_u \gamma_\delta(s, u)|_{u=0} \right]_0^{\alpha(b)} \\ - \int_0^{\alpha(b)} \left( D_\gamma L_h(\gamma(s), \dot{\gamma}(s)) + \frac{d}{ds} D_{\dot{\gamma}} L_h(\gamma(s), \dot{\gamma}(s)) \right) \cdot \partial_u \gamma_\delta|_{u=0}(s) ds. \quad (2.60)$$

It follows from (2.55) that

$$\partial_u \gamma_\delta(\alpha(t), u)|_{u=0} = -\dot{\gamma}(\alpha(t)) \partial_u \alpha_\delta|_{u=0}(t), \quad t = a, b. \tag{2.61}$$

Since by (2.17) we have  $A_h(\gamma, \dot{\gamma}) = D_{\dot{\gamma}} L_h(\gamma, \dot{\gamma}) \cdot \dot{\gamma}$ , we get by (2.56)

$$\begin{aligned} & - \left[ D_{\dot{\gamma}} L_h(\gamma(\alpha(t)), \dot{\gamma}(\alpha(t))) \cdot \dot{\gamma}(\alpha(t)) \partial_u \alpha_\delta|_{u=0}(t) \right]_a^b \\ & = - \left[ \left( L_h(\gamma(\alpha(t)), \dot{\gamma}(\alpha(t))) + E \right) \cdot \partial_u \alpha_\delta|_{u=0}(t) \right]_a^b. \end{aligned} \tag{2.62}$$

Using (2.61) and (2.62), the boundary terms on the right hand side of (2.58) and (2.60) cancel. Combining (2.57), (2.58) and (2.60) yields

$$\begin{aligned} & dI|_{(\gamma, \alpha)}(\partial_u(\gamma, \alpha)_\delta|_{u=0}) \\ & = \int_0^{\alpha(b)} \left( D_\gamma L_h(\gamma(s), \dot{\gamma}(s)) - \frac{d}{ds} D_{\dot{\gamma}} L_h(\gamma(s), \dot{\gamma}(s)) \right) \cdot \partial_u \gamma_\delta|_{u=0}(s) ds. \end{aligned} \tag{2.63}$$

For  $(\gamma, \alpha) = (\gamma_0, \mathbf{1})$ , the integrand is zero, since the integral curve  $(\gamma_0, \dot{\gamma}_0)$  of  $X_h$  solves Lagranges equation and thus

$$dI|_{(\gamma_0, \mathbf{1})} = 0.$$

*Step 5:*

We show (b).

If  $\gamma_0$  is a Finslerian geodesic with energy  $E$ , then  $b_E(\gamma_0) = (\gamma_0, \mathbf{1})$  and by (2.53) the integral (2.63) is zero for each tangent vector  $\partial_u \gamma_{0, \delta}|_{u=0}$ . By standard arguments it follows that  $(\gamma_0, \dot{\gamma}_0)$  solves Lagranges equation. Thus  $\gamma_0$  is a base integral curve of  $X_h$ .  $\square$

### 2.4. Application to $H_\varepsilon$

We start with

*Proof of Lemma 1.2.* (a): These estimates and the regularity follow at once by Hypothesis 1.1, (a), (i) and (iii).

(b): By standard Fourier theory,  $t_0$  is even with respect to  $\xi$ , i.e.  $t_0(x, \xi) = t_0(x, -\xi)$ , if and only if for all  $\eta \in \mathbb{Z}^d$

$$\tilde{a}_\eta(x) = \tilde{a}_{-\eta}(x). \tag{2.64}$$

To show (2.64) we use that by Hypothesis 1.1, (a), (iv) for all  $\eta \in \mathbb{Z}^d$ ,  $x \in \mathbb{R}^d$  and  $\varepsilon \in (0, 1]$

$$\tilde{a}_\eta(x) + \varepsilon a_{\varepsilon\eta}^{(1)}(x) + R_{\varepsilon\eta}^{(2)}(x, \varepsilon) = \tilde{a}_{-\eta}(x + \varepsilon\eta) + \varepsilon a_{-\varepsilon\eta}^{(1)}(x + \varepsilon\eta) + R_{-\varepsilon\eta}^{(2)}(x + \varepsilon\eta, \varepsilon). \tag{2.65}$$

By Hypothesis 1.1, (a), (i) we have

$$\tilde{a}_{-\eta}(x + \varepsilon\eta) = \tilde{a}_{-\eta}(x) + O_\eta(\varepsilon). \tag{2.66}$$

Combining (2.65) and (2.66) leads to

$$\begin{aligned} |\tilde{a}_\eta(x) - \tilde{a}_{-\eta}(x)| &\leq \varepsilon |a_{-\varepsilon\eta}^{(1)}(x + \varepsilon\eta) - a_{\varepsilon\eta}^{(1)}(x)| + O_\eta(\varepsilon) \\ &\quad + |R_{-\varepsilon\eta}^{(2)}(x + \varepsilon\eta, \varepsilon) - R_{\varepsilon\eta}^{(2)}(x, \varepsilon)| \\ &\leq \varepsilon C_{x,\eta} + \varepsilon^2 C_{x,\eta}. \end{aligned} \tag{2.67}$$

Since the left hand side of (2.67) is independent of  $\varepsilon$ , it is equal to zero and (2.64) follows.

The analytic continuation of  $t_0$  follows at once from Hypothesis 1.1,(a),(v), since  $a_\gamma^{(0)}(x)$  are the Fourier-coefficients of  $t_0(x, \xi)$ .

(c): By (a),  $t_0(x, \xi) = \sum_{\eta \in \mathbb{Z}^d} \tilde{a}_\eta(x) \cos(\eta \cdot \xi)$ , thus its Taylor-expansion at  $\xi = 0$  yields by Hypothesis 1.1(a)(ii)

$$\sum_{\eta \in \mathbb{Z}^d} \tilde{a}_\eta(x) \left( 1 - \frac{1}{2}(\eta \cdot \xi)^2 + O(|\xi|^4) \right) = \langle \xi, B(x)\xi \rangle + O(|\xi|^4), \tag{2.68}$$

where the symmetric  $d \times d$ -matrix  $B$  is given by

$$-\frac{1}{2} \sum_{\eta} \tilde{a}_\eta(x) \eta_\nu \eta_\mu = B_{\nu\mu}(x) \quad \text{for } \mu, \nu \in \{1, \dots, d\}, \quad x \in \mathbb{R}^d. \tag{2.69}$$

Since  $\langle \xi, B(x)\xi \rangle = -\frac{1}{2\varepsilon^2} \sum_{\gamma} a_\gamma^{(0)}(x)(\xi \cdot \gamma)^2$ , by Hypothesis 1.1, (a), (iii) and (vi) the matrix  $B$  is positive definite.

(d): First we mention that by a short calculation  $\text{Op}_\varepsilon^{\mathbb{T}^d}(e^{-\frac{i}{\varepsilon}\gamma \cdot \xi}) = \tau_\gamma$ . This implies  $T_\varepsilon = \sum_{\gamma} a_\gamma \tau_\gamma = \text{Op}_\varepsilon^{\mathbb{T}^d}(t)$  as operator on  $u \in \mathcal{K}((\varepsilon\mathbb{Z})^d)$  for  $t$  given in (1.5).

*Boundedness:*

For  $u \in \ell^2((\varepsilon\mathbb{Z})^d)$ , by the Cauchy–Schwarz inequality the  $l^2$ -norm of  $T_\varepsilon u$  can be estimated as

$$\begin{aligned} \|T_\varepsilon u\|_{\ell^2}^2 &\leq \sum_{x \in (\varepsilon\mathbb{Z})^d} \left( \sum_{\gamma \in (\varepsilon\mathbb{Z})^d} |a_\gamma(x, \varepsilon)u(x + \gamma)| \right)^2 \\ &\leq \sum_{x \in (\varepsilon\mathbb{Z})^d} \left( \sum_{\gamma \in (\varepsilon\mathbb{Z})^d} |a_\gamma(x, \varepsilon)|^2 \left( \frac{|\gamma|}{\varepsilon} \right)^{d+1} \right)^{\frac{1}{2}} \\ &\quad \times \left( \sum_{\gamma \in (\varepsilon\mathbb{Z})^d} \left( \frac{|\gamma|}{\varepsilon} \right)^{-(d+1)} |u(x + \gamma)|^2 \right)^{\frac{1}{2}}. \end{aligned} \tag{2.70}$$

By (1.9), the first factor on the right hand side of (2.70) is bounded uniformly in  $x$ . Thus

$$\|T_\varepsilon u\|_{\ell^2}^2 \leq C \sum_{\eta \in \mathbb{Z}^d} |\eta|^{-(d+1)} \sum_{x \in (\varepsilon\mathbb{Z})^d} |u(x + \varepsilon\eta)|^2 \leq C \|u\|_{\ell^2}^2.$$

Thus  $T_\varepsilon$  is a bounded operator on  $\ell^2((\varepsilon\mathbb{Z})^d)$ .

*Symmetry:*

Since  $T_\varepsilon$  is bounded, it is symmetric if and only if for any  $x, \gamma \in (\varepsilon\mathbb{Z})^d$

$$\langle T_\varepsilon \delta_x, \delta_{x+\gamma} \rangle_{\ell^2} = \langle \delta_x, T_\varepsilon \delta_{x+\gamma} \rangle_{\ell^2}, \tag{2.71}$$

where  $\delta_x(y) := \delta_{xy}$ . Since the left hand side of (2.71) is equal to  $a_{-\gamma}(x + \gamma, \varepsilon)$  and the right hand side equals  $a_\gamma(x, \varepsilon)$ , the statement follows by Hypothesis 1.1(a)(iv).

*Boundedness from below:*

For  $u \in \mathcal{K}((\varepsilon\mathbb{Z})^d)$ , we write

$$\langle u, T_\varepsilon u \rangle_{\ell^2} = A[u] + B[u] \quad \text{where} \tag{2.72}$$

$$A[u] := \sum_{x \in (\varepsilon\mathbb{Z})^d} \left\{ a_0^{(0)}(x) |u(x)|^2 + \sum_{\gamma \neq 0} a_\gamma^{(0)}(x) u(x + \gamma) \bar{u}(x) \right\}$$

$$B[u] := \sum_{x, \gamma \in (\varepsilon\mathbb{Z})^d} \left( \varepsilon a_\gamma^{(1)}(x) + R_\gamma^{(2)}(x, \varepsilon) \right) \bar{u}(x) u(x + \gamma).$$

Then by the exponential decay of  $a^{(1)}$  and  $R^{(2)}$  with respect to  $\gamma$  (Hypothesis 1.1(a)(v))

$$|B[u]| \leq \sum_{x, \gamma \in (\varepsilon\mathbb{Z})^d} \left| \varepsilon a_\gamma^{(1)}(x) + R_\gamma^{(2)}(x, \varepsilon) \right| (|\bar{u}(x)|^2 + |u(x + \gamma)|^2) \leq \varepsilon C \|u\|^2. \tag{2.73}$$

By Hypothesis 1.1(iii) we have

$$A[u] = \sum_x \sum_{\gamma \neq 0} a_\gamma^{(0)}(x) (u(x + \gamma) \bar{u}(x) - |u(x)|^2) \tag{2.74}$$

$$= \frac{1}{2} \left\{ \sum_{\substack{x \\ \gamma \neq 0}} a_\gamma^{(0)}(x) (u(x + \gamma) \bar{u}(x) - |u(x)|^2) \right.$$

$$\left. + \sum_{\substack{\tilde{x} \\ \tilde{\gamma} \neq 0}} a_{-\tilde{\gamma}}^{(0)}(\tilde{x} + \tilde{\gamma}) (u(\tilde{x}) \bar{u}(\tilde{x} + \tilde{\gamma}) - |u(\tilde{x} + \tilde{\gamma})|^2) \right\}$$

$$= -\frac{1}{2} \sum_{\substack{x \\ \gamma \neq 0}} a_\gamma^{(0)}(x) |u(x) - u(x + \gamma)|^2 \geq 0,$$

where for the second step we used the symmetry of  $T_\varepsilon$  and the substitution  $\tilde{x} = x + \gamma$  and  $\tilde{\gamma} = -\gamma$  and the last estimate follows from Hypothesis 1.1(a)(iii). Inserting (2.73) and (2.74) in (2.72) gives the stated result.  $\square$

Definition 2.1 and Theorem 2.11 allow to define a metric adapted to the Hamilton operator  $H_\varepsilon$  as follows.

**Proposition 2.14.** *The Hamilton function  $\tilde{h}_0 : \mathbb{R}^{2d} \cong T^*\mathbb{R}^d \rightarrow \mathbb{R}$  defined in (1.12) is hyperconvex in each fibre.*

*Proof.* We have to show that there exists a constant  $\alpha > 0$  such that

$$\langle v, D_\xi^2 \tilde{h}_0(x, \xi)v \rangle \geq \alpha \|v\|^2 \quad \text{for all } x, \xi, v \in \mathbb{R}^d. \tag{2.75}$$

For simplicity of notation, we will skip the  $x$ -dependence of  $\tilde{h}_0$  and  $a_\gamma^{(0)}$ . We have for  $\tilde{a}$  defined in (1.11)

$$\langle v, D_\xi^2 \tilde{h}_0(\xi)v \rangle = - \sum_{\eta \in \mathbb{Z}^d} \tilde{a}_\eta (\gamma \cdot v)^2 \cosh(\gamma \cdot \xi), \quad \xi, v \in \mathbb{R}^d. \tag{2.76}$$

By Hypothesis 1.1,(vi), we can choose a basis  $\{\eta^1, \dots, \eta^d\}$ ,  $\eta^j \in \mathbb{Z}^d$  of  $\mathbb{R}^d$  with  $\tilde{a}_{\eta^i} < 0$ . Since by Hypothesis 1.1(a)(iv) each summand in (2.76) has positive sign

$$\langle v, D_\xi^2 \tilde{h}_0(\xi)v \rangle \geq - \sum_{k=1}^d \tilde{a}_{\eta^k} (\eta^k \cdot v)^2 \cosh(\eta^k \cdot \xi), \quad \xi, v \in \mathbb{R}^d. \tag{2.77}$$

We have  $C = \min_k (-\tilde{a}_{\eta^k}) > 0$ , thus (2.77) yields

$$\langle v, D_\xi^2 \tilde{h}_0(\xi)v \rangle \geq C \sum_{k=1}^d (\eta^k \cdot v)^2 = \langle v, Mv \rangle \geq 0, \quad \text{for } M = \left( C \sum_k \eta_i^k \eta_j^k \right).$$

The sum can take the value 0 only if  $v = 0$  since  $\{\eta^k\}$  is a basis of  $\mathbb{R}^d$ . Thus  $\tilde{h}_0$  is hyperconvex (the lowest eigenvalue of  $M$  gives the lower bound for its second derivative). □

Proposition 2.14 leads by Proposition 2.9 and Lemma 1.2 to the following corollary.

**Corollary 2.15.** *The Hamilton function  $\tilde{h}_0 : \mathbb{R}^{2d} \rightarrow \mathbb{R}$  defined in (1.12) is hyperregular and even and strictly convex in each fibre.*

In the setting of Theorem 2.11, we choose  $M = \mathbb{R}^d$ ,  $E = 0$  and  $h = \tilde{h}_0 = \tilde{t}_0 - V_0$ . Recall that by Hypothesis 1.1, the set  $S(0)$  of singular points with respect to the energy  $E = 0$  is given by  $S(0) = \{0\}$ .

**Definition 2.16.** For the hyperregular Hamilton function  $\tilde{h}_0$  given in (1.12), we define

$$\ell(x, v) := \begin{cases} \ell_{\tilde{h}_0, 0}(x, v), & x \in \tilde{M} := \mathbb{R}^d \setminus \{0\}, v \in T_x \tilde{M} \\ 0 & x = 0. \end{cases} \tag{2.78}$$

The associated Finsler metric  $d_\ell : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$  is given by

$$d_\ell(x_0, x_1) = \inf_{\gamma_{0,1} \in \Gamma(x_0, x_1)} \int_0^1 \ell(\gamma(t), \dot{\gamma}(t)) dt. \tag{2.79}$$

We notice that it follows from the Definition of  $\tilde{\tau}_0$  that  $\lim_{x \rightarrow 0} \tilde{\tau}_0(x, v) = (0, 0)$ . Thus  $\ell : \mathbb{R}^{2d} \rightarrow \mathbb{R}$  defined in (2.78) is continuous.



**2.5. Proof of Theorem 1.5**

In order to prove Theorem 1.6, we notice that if  $d_\ell$  is locally Lipschitz continuous, it is differentiable almost everywhere in both arguments (Rademacher Theorem).

*Step 1:* We prove (1.17).

By the triangle inequality and the definition of  $d_\ell(x, y)$ , we have for any  $v \in \mathbb{R}^d$  with  $|v| = 1$  and  $\delta > 0$

$$d^0(x + \delta v) - d^0(x) \leq d_\ell(x, x + \delta v) \leq \int_0^1 \ell(\gamma_0(t), \dot{\gamma}_0(t)) dt, \tag{2.80}$$

where  $\gamma_0(t) = x + t\delta v$ . For this special curve we get by the homogeneity of  $\ell$

$$\int_0^1 \ell(\gamma_0(t), \dot{\gamma}_0(t)) dt \leq \sup_{t \in [0,1]} \ell(x + t\delta v, \delta v) = \delta \sup_{t \in [0,1]} \ell(x + t\delta v, v), \tag{2.81}$$

where by a slight abuse of notation  $v$  is considered as an element of  $T_{x+t\delta v}\mathbb{R}^d$ . Thus (2.80) together with (2.81) proves (1.17).

*Step 2:* We prove (1.18).

By (2.80) and (2.81) we have for any  $v \in \mathbb{R}^d$  with  $|v| = 1$  almost everywhere in  $x \in \mathbb{R}^d$

$$\begin{aligned} \nabla d^0(x) \cdot v = \partial_v d^0(x) &= \lim_{\delta \rightarrow 0} \frac{d^0(x + \delta v) - d^0(x)}{\delta} \\ &\leq \lim_{\delta \rightarrow 0} \sup_{t \in [0,1]} \ell(x + t\delta v, v) = \ell(x, v). \end{aligned} \tag{2.82}$$

Note that  $\nabla d^0(x)$  can be considered as an element of  $T_x^*\mathbb{R}^d$ . Since both sides in (2.82) are positive homogeneous of order one with respect to  $v$ , we can extend the inequality to all  $v \in \mathbb{R}^d$ . Using (2.20), the Finsler function  $\ell$  can be written as  $\ell(x, v) = \xi_{\tilde{h}_0}(x, \tilde{v}) \cdot v$ , where  $v$  is considered as an element of  $T_x M$  and will be written as  $(x, v)$ . It follows from (2.82) that

$$(\xi_{\tilde{h}_0}(x, \tilde{v}) - \nabla d^0(x)) \cdot v \geq 0, \quad (x, v) \in TM \text{ a.e. on } M. \tag{2.83}$$

Since  $\tilde{h}_0(x, \xi)$  is differentiable, real valued and convex in each fibre, by (2.4) the inequality

$$\tilde{h}_0(x, \xi) \geq \tilde{h}_0(x, \eta) + D_\eta \tilde{h}_0(x, \eta) \cdot (\xi - \eta)$$

holds for all  $x, \xi, \eta \in \mathbb{R}^d$ . Thus by setting  $\xi = \xi_{\tilde{h}_0}(x, \tilde{v})$  and  $\eta = \nabla d^0(x)$ , we get for all  $(x, v) \in TM$  the estimate

$$\tilde{h}_0(x, \xi_{\tilde{h}_0}(x, \tilde{v})) \geq \tilde{h}_0(x, \nabla d^0(x)) + D_\xi \tilde{h}_0(x, \nabla d^0(x)) \cdot (\xi_{\tilde{h}_0}(x, \tilde{v}) - \nabla d^0(x)), \tag{2.84}$$

where  $(x, \tilde{v}) \in \mathcal{E}$  is associated to  $(x, v)$ . The left hand side of (2.84) is by definition of  $\tilde{v}$  equal to zero. Choosing  $(x, v) = D_\xi \tilde{h}_0(x, \nabla d^0(x))$  in equation (2.84) yields

$$0 \geq \tilde{h}_0(x, \nabla d^0(x)) + v \cdot (\xi_{\tilde{h}_0}(x, \tilde{v}) - \nabla d^0(x)).$$

Using (2.83), this proves (1.18).

*Step 3:* We prove (1.16) (the eikonal equality):

We consider the generalized eikonal equation

$$\tilde{h}_0(x, \nabla\varphi(x)) = \tilde{t}_0(x, \nabla\varphi(x)) - V_0(x) = 0. \tag{2.85}$$

Choose coordinates such that  $t_0(x_0, \xi) = |\xi|^2 + O(|\xi|^3)$  and  $V_0(x) = \sum_{\nu=1}^d \lambda_\nu^2 x_\nu^2 + O(|x|^3)$ . It is proven in [22] that there exists a unique positive  $\mathcal{C}^\infty$ -function  $\varphi$  defined in a neighborhood  $\Omega$  of 0, solving (2.85) such that  $\varphi$  has an expansion as asymptotic series

$$\varphi(x) \sim \sum_{\nu=1}^d \frac{\lambda_\nu}{2} x_\nu^2 + \sum_{k \geq 1} \varphi_k(x), \quad x \in \Omega, \tag{2.86}$$

where each  $\varphi_k$  is an homogeneous polynomial of order  $k + 2$ .

In particular, denote by  $F_t$  the flow of the Hamiltonian vector field  $X_{\tilde{h}_0}$ . Then the Local Stable Manifold Theorem [2] tells us that there is an open neighborhood  $\mathcal{N}$  of  $(0, 0)$  such that the two submanifolds

$$\Lambda_\pm (X_{\tilde{h}_0}, (0, 0)) := \{ (x, \xi) \in T^*\mathbb{R}^d \mid F_t(x, \xi) \rightarrow (0, 0) \text{ for } t \rightarrow \mp\infty \} \tag{2.87}$$

exist and are unique in  $\mathcal{N}$ . They are called stable ( $\Lambda_-$ ) and unstable ( $\Lambda_+$ ) manifold of  $X_{\tilde{h}_0}$  of the critical point  $(0, 0)$ . Moreover they are of dimension  $d$  and contained in  $\tilde{h}_0^{-1}(0)$ . It is shown in [22] that  $\Lambda_\pm$  are Lagrangian manifolds in  $T^*\mathbb{R}^d$  and that the outgoing manifold can be parametrized as  $\Lambda_+ = \{ (x, \nabla\varphi(x)) \mid x \in \Omega \}$ . Thus for a given  $x \in \Omega$  there exists an integral curve  $\hat{\gamma}_0 := (\gamma_0, \nabla\varphi(\gamma_0)) \subset \Lambda_+$  of the Hamiltonian vector field  $X_{\tilde{h}_0}$ , parametrized by  $[-\infty, 0]$  such that  $\hat{\gamma}_0(0) = (x, \nabla\varphi(x))$  and  $\lim_{t \rightarrow -\infty} \hat{\gamma}_0(t) = (0, 0)$ . Since  $\hat{\gamma}_0$  is an integral curve of  $X_{\tilde{h}_0}$ , it follows from Hamilton's equations that

$$(\gamma_0, \dot{\gamma}_0) = \mathcal{D}_F \tilde{h}_0(\gamma_0, \nabla\varphi(\gamma_0))$$

and therefore

$$\left( \mathcal{D}_F \tilde{h}_0 \right)^{-1} (\gamma_0, \dot{\gamma}_0) = (\gamma_0, \nabla\varphi(\gamma_0)).$$

Thus

$$\frac{d}{dt} \varphi \circ \gamma_0 = \nabla\varphi|_{\gamma_0} \cdot \dot{\gamma}_0 = \left( \mathcal{D}_F \tilde{h}_0 \right)^{-1} (\gamma_0, \dot{\gamma}_0) \cdot \dot{\gamma}_0. \tag{2.88}$$

Since  $\hat{\gamma}_0$  is an integral curve,  $(\gamma_0(t), \dot{\gamma}_0(t)) \in \mathcal{E}$  for all  $t$ . Therefore  $\tilde{\tau}_0(\gamma_0, \dot{\gamma}_0) = (\gamma_0, \dot{\gamma}_0)$  and it follows at once from (2.88) and the definition of  $\ell$  that

$$\frac{d}{dt} \varphi \circ \gamma_0 = \ell(\gamma_0, \dot{\gamma}_0). \tag{2.89}$$

The point  $x = 0$  is a singular point of the Finsler manifold  $(\mathbb{R}^d, \ell)$ , thus the base integral curve  $\gamma_0 : [-\infty, 0] \rightarrow \Omega \ni 0$  of  $X_{\tilde{h}_0}$  is not a regular curve on a Finsler manifold in the sense of Definition 2.3. To avoid this difficulty, we restrict the curve  $\gamma_0$  to  $[-T, 0]$  and set  $y_T := \gamma_0(-T)$ . Then by (2.89)

$$\varphi(x) - \varphi(y_T) = \int_{-T}^0 \ell(\gamma_0(t), \dot{\gamma}_0(t)) dt. \tag{2.90}$$

By Proposition 1.5 the base integral curve  $\gamma_0$  of  $X_{\tilde{h}_0}$  is a geodesic with respect to the associated Finsler function  $\ell$ . It is a basic theorem in Finsler Geometry (see Abate–Patrizio [1], Theorem 1.6.6), that geodesics, which are short enough, actually minimize the curve length among all  $\mathcal{C}^\infty$ -curves (or  $\mathcal{C}^2$ -curves) with the same endpoints. Thus the length of any short geodesic joining  $x$  and  $y$  is for  $|x - y|$  sufficiently small equal to the Finsler distance  $d_\ell(x, y)$  and

$$\varphi(x) - \varphi(y_T) = d_\ell(y_T, x). \tag{2.91}$$

Since  $y_T \rightarrow 0, T \rightarrow \infty$  and  $d_\ell$  and  $\varphi$  are continuous in  $x = 0$ , we get

$$\varphi(x) = d^0(x) \quad x \in \Omega$$

for  $|x|$  sufficiently small. □

### 3. Weighted estimates for Dirichlet eigenfunctions

#### 3.1. Preliminary results

**Lemma 3.1.** *Assume Hypothesis 1.1 and, for  $\Sigma \subset \mathbb{R}^d$ , let  $H_\varepsilon^\Sigma$  denote the Dirichlet operator introduced in Definition 1.7. Let  $\varphi : \Sigma \rightarrow \mathbb{R}$  be Lipschitz and constant outside some bounded set. Then for any real valued  $v \in \mathcal{D}(H_\varepsilon^\Sigma)$*

$$\begin{aligned} \langle (e^{\frac{\varphi}{\varepsilon}} H_\varepsilon^\Sigma e^{-\frac{\varphi}{\varepsilon}}) v, v \rangle_{\ell^2} &= \langle (V_\varepsilon + V_{\varepsilon, \Sigma}^\varphi) v, v \rangle_{\ell^2} \\ &\quad - \frac{1}{2} \sum_{x \in \Sigma} \sum_{\gamma \in \Sigma'_\varepsilon(x)} a_\gamma(x, \varepsilon) \cosh\left(\frac{1}{\varepsilon}(\varphi(x) - \varphi(x + \gamma))\right) (v(x) - v(x + \gamma))^2, \end{aligned}$$

where  $\Sigma'_\varepsilon(x) := \{\gamma \in (\varepsilon\mathbb{Z})^d \mid x + \gamma \in \Sigma\}$  and

$$V_{\varepsilon, \Sigma}^\varphi(x) := \sum_{\gamma \in \Sigma'_\varepsilon(x)} a_\gamma(x, \varepsilon) \cosh\left(\frac{1}{\varepsilon}(\varphi(x) - \varphi(x + \gamma))\right), \tag{3.1}$$

where the sum on the right hand side converges.

*Proof.* By use of the symmetry of  $T_\varepsilon$  (Lemma 1.2) and since  $v$  and  $\varphi$  are assumed to be real valued and  $e^{\pm \frac{\varphi}{\varepsilon}} v \in \mathcal{D}(H_\varepsilon^\Sigma)$ , we have

$$\begin{aligned} &\langle (e^{\frac{\varphi}{\varepsilon}} \mathbf{1}_{\Sigma_\varepsilon} T_\varepsilon \mathbf{1}_{\Sigma_\varepsilon} e^{-\frac{\varphi}{\varepsilon}}) v, v \rangle_{\ell^2} \\ &= \frac{1}{2} [\langle \mathbf{1}_{\Sigma_\varepsilon} T_\varepsilon \mathbf{1}_{\Sigma_\varepsilon} e^{-\frac{\varphi}{\varepsilon}} v, e^{\frac{\varphi}{\varepsilon}} v \rangle_{\ell^2} + \langle e^{-\frac{\varphi}{\varepsilon}} v, \mathbf{1}_{\Sigma_\varepsilon} T_\varepsilon \mathbf{1}_{\Sigma_\varepsilon} e^{\frac{\varphi}{\varepsilon}} v \rangle_{\ell^2}] \\ &= \frac{1}{2} \sum_{x, x+\gamma \in \Sigma} a_\gamma(x, \varepsilon) \left( e^{\frac{1}{\varepsilon}(\varphi(x) - \varphi(x+\gamma))} + e^{-\frac{1}{\varepsilon}(\varphi(x) - \varphi(x+\gamma))} \right) v(x + \gamma)v(x). \end{aligned}$$

Writing  $v(x+\gamma)v(x) = v^2(x) - \frac{1}{2} (2v^2(x) - 2v(x+\gamma)v(x))$  yields by the definition of  $V_{\varepsilon, \Sigma}^\varphi$

$$\begin{aligned} & \langle (e^{\frac{\varphi}{\varepsilon}} T_\varepsilon e^{-\frac{\varphi}{\varepsilon}}) v, v \rangle_{\ell^2} \\ &= \langle V_{\varepsilon, \Sigma}^\varphi v, v \rangle_{\ell^2} - \frac{1}{2} \sum_{x, x+\gamma \in \Sigma} a_\gamma(x, \varepsilon) \cosh\left(\frac{1}{\varepsilon}(\varphi(x) - \varphi(x+\gamma))\right) \\ & \quad \times (2v^2(x) - 2v(x)v(x+\gamma)). \end{aligned} \tag{3.2}$$

By Hypothesis 1.1 we have  $a_\gamma(x, \varepsilon) = a_{-\gamma}(x+\gamma, \varepsilon)$ . Thus by use of the substitutions  $x' = x + \gamma$  and  $\gamma' = -\gamma$  together with the fact that  $\cosh$  is even

$$\begin{aligned} & \sum_{x, x+\gamma \in \Sigma} a_\gamma(x, \varepsilon) \cosh\left(\frac{1}{\varepsilon}(\varphi(x) - \varphi(x+\gamma))\right) v^2(x) \\ &= \sum_{x', x'+\gamma' \in \Sigma} a_{-\gamma'}(x' + \gamma', \varepsilon) \cosh\left(\frac{1}{\varepsilon}(\varphi(x' + \gamma') - \varphi(x'))\right) v^2(x' + \gamma') \\ &= \sum_{x', x'+\gamma' \in \Sigma} a_{\gamma'}(x', \varepsilon) \cosh\left(\frac{1}{\varepsilon}(\varphi(x') - \varphi(x' + \gamma'))\right) v^2(x' + \gamma'). \end{aligned} \tag{3.3}$$

Inserting (3.3) into (3.2) gives

$$\begin{aligned} & \langle (e^{\frac{\varphi}{\varepsilon}} \mathbf{1}_{\Sigma_\varepsilon} T_\varepsilon \mathbf{1}_{\Sigma_\varepsilon} e^{-\frac{\varphi}{\varepsilon}}) v, v \rangle_{\ell^2} \\ &= \langle V_{\varepsilon, \Sigma}^\varphi v, v \rangle_{\ell^2} - \frac{1}{2} \sum_{x, x+\gamma \in \Sigma} a_\gamma(x, \varepsilon) \cosh\left(\frac{1}{\varepsilon}(\varphi(x) - \varphi(x+\gamma))\right) \\ & \quad \times (v^2(x) - 2v(x)v(x+\gamma) + v^2(x+\gamma)). \end{aligned}$$

Since  $V_\varepsilon$  commutes with  $e^{-\frac{\varphi}{\varepsilon}}$ , the stated equality follows. The convergence of the series in (3.1) follows from the decay of  $a_\gamma(x, \varepsilon)$  with respect to  $\gamma$  (Hypothesis 1.1(v)) together with the assumptions on  $\varphi$  and the mean value theorem.  $\square$

Lemma 3.1 leads to the following norm estimate, which will be used later on to prove Theorem 1.8.

**Lemma 3.2.** *Assume Hypothesis 1.1 and, for  $\Sigma \subset \mathbb{R}^d$ , let  $H_\varepsilon^\Sigma$  denote the Dirichlet operator introduced in Definition 1.7. For  $E \geq 0$  fixed, let  $F_\pm : \Sigma \rightarrow [0, \infty)$  be a pair of functions such that  $F(x) := F_+(x) + F_-(x) > 0$  and*

$$F_+^2(x) - F_-^2(x) = \widehat{V}_\varepsilon(x) + V_{\varepsilon, \Sigma}^\varphi(x) - E, \quad x \in \Sigma, \tag{3.4}$$

where  $V_{\varepsilon, \Sigma}^\varphi(x)$  is given in (3.1). Then for  $v \in \mathcal{D}(H_\varepsilon^\Sigma)$  real-valued with  $Fv \in \ell_{\Sigma_\varepsilon}^2$  and  $\varphi : \Sigma \rightarrow \mathbb{R}$  Lipschitz and constant outside some bounded set, we have for some  $C > 0$

$$\|Fv\|_{\ell^2}^2 \leq 4 \left\| \frac{1}{F} \left( e^{\frac{\varphi}{\varepsilon}} (H_\varepsilon^\Sigma - E) e^{-\frac{\varphi}{\varepsilon}} \right) v \right\|_{\ell^2}^2 + 8\|F_-v\|_{\ell^2}^2 + C\varepsilon\|v\|^2. \tag{3.5}$$

*Proof.* First observe that

$$\|Fv\|_{\ell^2}^2 \leq 2(\|F_+v\|_{\ell^2}^2 + \|F_-v\|_{\ell^2}^2) = 2(\|F_+v\|_{\ell^2}^2 - \|F_-v\|_{\ell^2}^2) + 4\|F_-v\|_{\ell^2}^2. \quad (3.6)$$

By (3.4) one has

$$\|F_+v\|_{\ell^2}^2 - \|F_-v\|_{\ell^2}^2 = \left\langle (\widehat{V}_\varepsilon + V_{\varepsilon,\Sigma}^\varphi - E)v, v \right\rangle_{\ell^2}. \quad (3.7)$$

Hypothesis 1.1(a)(iii) and (v) yields by straightforward calculation

$$-\frac{1}{2} \sum_{x,x+\gamma \in \Sigma} a_\gamma(x, \varepsilon) \cosh\left(\frac{1}{\varepsilon}(\varphi(x) - \varphi(x + \gamma))\right) (v(x) - v(x + \gamma))^2 \geq -C\varepsilon\|v\|^2, \quad (3.8)$$

since  $|\varphi(x + \gamma) - \varphi(x)| \leq |\gamma| \sup_{y \in K} |D\varphi(y)|$  for some compact set  $K \subset \Sigma$ . Thus it follows from Lemma 3.1 and (3.8) that

$$\left\langle (\widehat{V}_\varepsilon + V_{\varepsilon,\Sigma}^\varphi - E)v, v \right\rangle_{\ell^2} - C\varepsilon\|v\|^2 \leq \left\langle (e^{\frac{\varphi}{\varepsilon}}(H_\varepsilon^\Sigma - E)e^{-\frac{\varphi}{\varepsilon}})v, v \right\rangle_{\ell^2}. \quad (3.9)$$

(3.7) and (3.9) yield by use of the Cauchy–Schwarz inequality

$$\begin{aligned} & 2(\|F_+v\|_{\ell^2}^2 - \|F_-v\|_{\ell^2}^2) - C\varepsilon\|v\|^2 \\ & \leq 2\left\langle (e^{\frac{\varphi}{\varepsilon}}(H_\varepsilon^\Sigma - E)e^{-\frac{\varphi}{\varepsilon}})v, v \right\rangle_{\ell^2} \\ & \leq 2\sqrt{2} \left\| \frac{1}{F} \left( e^{\frac{\varphi}{\varepsilon}}(H_\varepsilon^\Sigma - E)e^{-\frac{\varphi}{\varepsilon}} \right) v \right\|_{\ell^2} \frac{1}{\sqrt{2}} \|Fv\|_{\ell^2} \\ & \leq v2 \left\| \frac{1}{F} \left( e^{\frac{\varphi}{\varepsilon}}(H_\varepsilon^\Sigma - E)e^{-\frac{\varphi}{\varepsilon}} \right) v \right\|_{\ell^2}^2 + \frac{1}{2} \|Fv\|_{\ell^2}^2. \end{aligned} \quad (3.10)$$

Inserting (3.10) into (3.6) we get

$$\|Fv\|_{\ell^2}^2 \leq 2 \left\| \frac{1}{F} \left( e^{\frac{\varphi}{\varepsilon}}(H_\varepsilon^\Sigma - E)e^{-\frac{\varphi}{\varepsilon}} \right) v \right\|_{\ell^2}^2 + \frac{1}{2} \|Fv\|_{\ell^2}^2 + 4\|F_-v\|_{\ell^2}^2 + C\varepsilon\|v\|^2.$$

This proves (3.5). □

**Lemma 3.3.** *Let  $\Sigma \subset \mathbb{R}^d$  be an open bounded region including the point 0 and such that  $d^0 \in \mathcal{C}^2(\overline{\Sigma})$ , where  $d^0(x) := d_\ell(0, x)$  is defined in (2.79). Let  $\chi \in \mathcal{C}^\infty(\mathbb{R}_+, [0, 1])$  such that  $\chi(r) = 0$  for  $r \leq \frac{1}{2}$  and  $\chi(r) = 1$  for  $r \geq 1$ . In addition we assume that  $0 \leq \chi'(r) \leq \frac{2}{\log 2}$ . For  $B > 0$  we define  $g : \Sigma \rightarrow [0, 1]$  by*

$$g(x) := \chi\left(\frac{d^0(x)}{B\varepsilon}\right), \quad x \in \Sigma \quad (3.11)$$

and set

$$\Phi(x) := d^0(x) - \frac{B\varepsilon}{2} \log\left(\frac{B}{2}\right) - g(x) \frac{B\varepsilon}{2} \log\left(\frac{2d^0(x)}{B\varepsilon}\right), \quad x \in \Sigma. \quad (3.12)$$

Then there exists a constant  $C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0]$

$$|\partial_\nu \partial_\mu \Phi(x)| \leq C. \quad (3.13)$$

Furthermore for any  $B > 0$  there is  $C' > 0$  such that

$$e^{\frac{d^0(x)}{\varepsilon}} \frac{1}{C'} \left(1 + \frac{d^0(x)}{\varepsilon}\right)^{-\frac{B}{2}} \leq e^{\frac{\Phi(x)}{\varepsilon}} \leq e^{\frac{d^0(x)}{\varepsilon}} C' \left(1 + \frac{d^0(x)}{\varepsilon}\right)^{-\frac{B}{2}}. \tag{3.14}$$

*Proof.* We write for simplicity  $d(x) := d^0(x)$ . First we notice that there exists a  $C > 0$  such that for  $\alpha \in \mathbb{N}^d, |\alpha| \leq 2$

$$|\partial^\alpha g(x)| \leq C\varepsilon^{-\frac{\alpha}{2}}, \quad x \in \Sigma. \tag{3.15}$$

Here one uses that by (1.16)  $d(x) = O(|x|^2)$  and  $\nabla d(x) = O(|x|)$  as  $|x| \rightarrow 0$ , thus  $x = O(\sqrt{\varepsilon})$  and  $\nabla d(x) = O(\sqrt{\varepsilon})$  on  $\text{supp } \nabla g \subset \{x \in \Sigma \mid \frac{B\varepsilon}{2} \leq d(x) \leq B\varepsilon\}$ . By the definition (3.12) of  $\Phi$  we have

$$\partial_\nu \partial_\mu \Phi(x) = \partial_\nu \partial_\mu d(x) - \partial_\nu \partial_\mu \left( g(x) \frac{B\varepsilon}{2} \log \left( \frac{2d(x)}{B\varepsilon} \right) \right) = A_1 + A_2 + A_3 \tag{3.16}$$

with

$$\begin{aligned} A_1 &:= \partial_\nu \partial_\mu d(x) \\ A_2 &:= - \left\{ (\partial_\nu \partial_\mu g)(x) \frac{B\varepsilon}{2} \log \left( \frac{2d(x)}{B\varepsilon} \right) + (\partial_\nu g)(x) \frac{B\varepsilon}{2d(x)} (\partial_\mu d)(x) \right. \\ &\quad \left. + (\partial_\mu g)(x) \frac{B\varepsilon}{2d(x)} (\partial_\nu d)(x) \right\} \\ A_3 &:= g(x) \frac{B\varepsilon}{2d(x)} \left( \frac{(\partial_\nu d)(x)(\partial_\mu d)(x)}{d(x)} + (\partial_\nu \partial_\mu d)(x) \right). \end{aligned}$$

Since  $\Sigma$  is bounded, all derivatives of  $d$  are at least bounded by a constant independent of  $\varepsilon$ , thus  $A_1$  is bounded.

Each summand in  $A_2$  includes a derivative of  $g$  and is therefore supported in the region  $\frac{B\varepsilon}{2} < d(x) < B\varepsilon$ . Thus  $1 < \frac{2d(x)}{B\varepsilon} < 2$  and from (1.16), it follows as above that  $\partial_\nu d(x) = O(\sqrt{\varepsilon})$ . By (3.15)  $A_2$  is bounded.

To estimate  $A_3$ , we introduce a constant  $\delta > 0$  such that  $\{x \in \Sigma \mid d(x) < \delta\} \subset \Omega$  and  $\delta \geq \varepsilon_0 B$  and analyze the regions  $d(x) < \delta$  and  $d(x) \geq \delta$  separately.

*Case 1:  $d(x) < \delta$ :*

By Theorem 1.6, we have  $\partial_\nu d(x) = O(|x|)$  and  $\partial_\nu \partial_\mu d(x) = O(1)$ . Thus there exists a constant  $M > 0$  such that

$$\sum_{\nu, \mu} \left| \frac{(\partial_\nu d)(x)(\partial_\mu d)(x)}{d(x)} \right| + |(\partial_\nu \partial_\mu d)(x)| < M \quad \text{for } \delta \text{ small enough.}$$

Since in addition for  $d(x) > \frac{B\varepsilon}{2}$  (on the support of  $g$ ), the term  $\frac{B\varepsilon}{2d}$  is bounded by 1,  $A_3$  is bounded by a constant independent of  $\varepsilon$ .

*Case 2:  $d(x) \geq \delta$ :*

We use that the derivatives of  $d$  are bounded on  $\Sigma$  and that  $\frac{1}{d(x)} \leq \frac{1}{\delta}$ .

Combining Case 1 and 2 we get the boundedness of  $A_3$  and thus (3.13).

To see (3.14), we first note that by definition

$$e^{\frac{\Phi(x)}{\varepsilon}} = e^{\frac{d(x)}{\varepsilon}} \left(\frac{B}{2}\right)^{-\frac{B}{2}(g(x)-1)} \left(\frac{d(x)}{\varepsilon}\right)^{-\frac{B}{2}g(x)}. \tag{3.17}$$

We notice that for any  $y \geq 0$  and for any  $B > 0$  there exists  $\tilde{C} > 0$  such that

$$\frac{1}{\tilde{C}} \leq \frac{y\chi(\frac{y}{B})}{1+y} \leq \tilde{C}.$$

Setting  $y = \frac{d(x)}{\varepsilon}$ , this leads to (3.14). □

**3.2. Proof of Theorem 1.7**

We partly follow the ideas in the proof of Proposition 5.5 in Helffer–Sjöstrand [17].

Let

$$t_0^\Sigma(x, \xi) := \sum_{\gamma \in \Sigma'_\varepsilon(x)} a_\gamma^{(0)}(x) \cos\left(\frac{1}{\varepsilon}\gamma \cdot \xi\right), \quad (x, \xi) \in \Sigma \times \mathbb{T}^d, \tag{3.18}$$

where  $\Sigma'_\varepsilon(x) := \{\gamma \in (\varepsilon\mathbb{Z})^d \mid x + \gamma \in \Sigma\}$ . We notice that

$$t_0(x, i\xi) \leq t_0^\Sigma(x, i\xi), \tag{3.19}$$

since  $a_\gamma^{(0)} \leq 0$  for  $\gamma \neq 0$ . In the following we write for simplicity  $d(x) := d^0(x)$ . By Theorem 1.6, for any  $B > 0$  we may choose  $\varepsilon_B > 0$  such that for all  $\varepsilon < \varepsilon_B$

$$V_0(x) + t_0(x, i\nabla d(x)) = 0, \quad x \in \Sigma \cap d^{-1}([0, B\varepsilon]), \tag{3.20}$$

By (3.11) and (3.12)

$$\nabla\Phi(x) = \nabla d(x) \left\{ 1 - \frac{B\varepsilon}{2d(x)}\chi\left(\frac{d(x)}{B\varepsilon}\right) - \frac{1}{2}\chi'\left(\frac{d(x)}{B\varepsilon}\right) \log\left(\frac{2d(x)}{B\varepsilon}\right) \right\}. \tag{3.21}$$

*Step 1:* We shall show that there is  $C_0 > 0$  independent of  $B$  such that

$$V_0(x) + t_0^\Sigma(x, i\nabla\Phi) \geq \begin{cases} 0, & x \in \Sigma \cap d^{-1}([0, B\varepsilon]) \\ \frac{B}{C_0}\varepsilon, & x \in \Sigma \cap d^{-1}([B\varepsilon, \infty)) \end{cases} \tag{3.22}$$

*Case 1:*  $d(x) \leq \frac{B\varepsilon}{2}$

Since  $\chi(x) = \chi'(x) = 0$  and the eikonal equation (2.85) holds, we get

$$V_0(x) + t_0(x, i\nabla\Phi(x)) = V_0(x) + t_0(x, i\nabla d(x)) = 0, \tag{3.23}$$

$$x \in \Sigma \cap d^{-1}\left(\left[0, \frac{B\varepsilon}{2}\right]\right).$$

which by (3.19) leads at once to the first estimate in (3.22) in Case 1.

*Case 2:*  $d(x) \geq B\varepsilon$

Since  $\chi'(x) = 0$  in this region, we have by (3.21)

$$\nabla\Phi(x) = \nabla d(x) \left(1 - \frac{B\varepsilon}{2d(x)}\right). \tag{3.24}$$

By Lemma 2.14,  $t_0(x, i\xi) = -\tilde{t}_0(x, \xi)$  is concave with respect to  $\xi$ , therefore

$$t_0(x, \lambda i\xi + (1 - \lambda)i\eta) \geq \lambda t_0(x, i\xi) + (1 - \lambda)t_0(x, i\eta) \quad \text{for } 0 \leq \lambda \leq 1, \quad \xi, \eta \in \mathbb{R}^d. \quad (3.25)$$

We have  $0 \leq (1 - \frac{B\varepsilon}{2d(x)}) \leq 1$ . Thus choosing  $\lambda = (1 - \frac{B\varepsilon}{2d(x)})$  and  $\eta = 0$  in (3.25) and using  $t_0(x, 0) = 0$  for all  $x \in (\varepsilon\mathbb{Z})^d$ , by (3.24) we get the estimate

$$\begin{aligned} V_0(x) + t_0(x, i\nabla\Phi(x)) &\geq V_0(x) + \left(1 - \frac{B\varepsilon}{2d(x)}\right) t_0(x, i\nabla d(x)) \\ &\geq V_0(x) \left(1 - \left(1 - \frac{B\varepsilon}{2d(x)}\right)\right) \\ &= V_0(x) \frac{B\varepsilon}{2d(x)}, \end{aligned} \quad (3.26)$$

where for the second estimate we used that by Theorem 1.6 the eikonal inequality  $t_0(x, i\nabla d(x)) \geq -V_0(x)$  holds. It follows from Theorem 1.6 and Hypothesis (1.1)(b) respectively that  $d(x) = O(|x|^2)$  and  $V_0(x) = O(|x|^2)$  for  $|x| \rightarrow 0$ . Since the region  $\Sigma$  was assumed to be bounded, it thus follows that there exists a constant  $C_0 > 0$  such that

$$C_0^{-1} \leq \frac{V_0(x)}{2d(x)} \leq C_0, \quad x \in \Sigma \cap d^{-1}([B\varepsilon, \infty)). \quad (3.27)$$

Combining (3.19), (3.26) and (3.27), we finally get the second estimate in (3.22).

*Case 3:*  $\frac{B\varepsilon}{2} < d(x) < B\varepsilon$

We define

$$f_1(x) := \frac{B\varepsilon}{2d(x)} \chi\left(\frac{d(x)}{B\varepsilon}\right) \quad \text{and} \quad f_2(x) := \frac{1}{2} \chi'\left(\frac{d(x)}{B\varepsilon}\right) \log\left(\frac{2d(x)}{B\varepsilon}\right),$$

such that by (3.21)

$$\nabla\Phi(x) = \nabla d(x)(1 - f_1(x) - f_2(x)). \quad (3.28)$$

Since  $1 < \frac{2d(x)}{B\varepsilon} < 2$ ,  $f_1$  and  $f_2$  are non-negative and therefore  $1 - f_1(x) - f_2(x) \leq 1$ . In addition it follows that  $0 \leq f_1(x) \leq 1$  and by the assumption  $\chi'(r) \leq \frac{2}{\log 2}$  we get  $0 \leq f_2(x) \leq 1$ . Therefore  $0 \leq f_1(x) + f_2(x) \leq 2$  and thus the estimate

$$|1 - f_1(x) - f_2(x)| \leq 1 \quad (3.29)$$

holds. Setting  $\lambda(x) := 1 - f_1(x) - f_2(x)$  it follows from (3.28) and (3.29) that

$$\nabla\Phi(x) = \lambda(x)\nabla d(x) \quad \text{with} \quad |\lambda(x)| \leq 1 \quad x \in \mathbb{R}^d. \quad (3.30)$$

Thus again from (3.25) (with  $\eta = 0$  and  $\xi = \nabla d(x)$ ) together with (3.30), (3.19) and the fact that  $t_0$  is even with respect to  $i\xi$  it follows that

$$V_0(x) + t_0^\Sigma(x, i\nabla\Phi(x)) \geq V_0(x) + |\lambda(x)|t_0(x, i\nabla d(x)) \geq V_0(1 - |\lambda(x)|), \quad (3.31)$$

where for the second step we used (3.20). Since  $|\lambda(x)| \leq 1$  and  $V_0 \geq 0$ , (3.31) gives the first estimate in (3.22) in Case 3.



Step 2: We shall show

$$\widehat{V}_\varepsilon(x) + V^\Phi(x) \geq \begin{cases} -C_5 \varepsilon & \text{for } x \in \Sigma \cap d^{-1}([0, B\varepsilon]) \\ \left(\frac{B}{C_0} - C_5\right) \varepsilon & \text{for } x \in \Sigma \cap d^{-1}([B\varepsilon, \infty)). \end{cases} \quad (3.32)$$

for some  $C_5 > 0$  independent of  $B$ , where  $V^\Phi := V_{\varepsilon, \Sigma}^\Phi$  is defined in (3.1).

We write

$$\begin{aligned} \widehat{V}_\varepsilon(x) + V^\Phi(x) &= \left(\widehat{V}_\varepsilon(x) - V_0(x)\right) + \left(V^\Phi(x) - t_0^\Sigma(x, i\nabla\Phi(x))\right) \\ &\quad + \left(V_0(x) + t_0^\Sigma(x, i\nabla\Phi(x))\right) \end{aligned} \quad (3.33)$$

and give estimates for the differences in the first two brackets on the right hand side.

By Hypothesis 1.1 and since  $\Sigma$  is bounded, there exists a constant  $C_1 > 0$  such that

$$\widehat{V}_\varepsilon(x) - V_0(x) \geq -C_1\varepsilon, \quad x \in \Sigma. \quad (3.34)$$

We shall show that

$$\left|V^\Phi(x) - t_0^\Sigma(x, i\nabla\Phi(x))\right| \leq \varepsilon C_4. \quad (3.35)$$

Then inserting (3.35), (3.34) and (3.22) in (3.33) proves (3.32).

Setting (see (3.1))

$$V_0^\Phi(x) := \sum_{\gamma \in \Sigma'(x)} a_\gamma^{(0)}(x) \cosh\left(\frac{1}{\varepsilon}(\Phi(x) - \Phi(x + \gamma))\right),$$

we write

$$\begin{aligned} V^\Phi(x) - t_0^\Sigma(x, i\nabla\Phi(x)) &= (V^\Phi(x) - V_0^\Phi(x)) + \left(V_0^\Phi(x) - t_0^\Sigma(x, i\nabla\Phi(x))\right) \\ &=: D_1(x) + D_2(x) \end{aligned}$$

and analyze the two summands on the right hand side separately. Since  $\Phi$  is Lipschitz and constant outside of some bounded set, it follows from Hypothesis 1.1(a) (as in the proof of (3.8)) that for some  $\tilde{C} > 0$

$$\left|D_1(x)\right| = \left| \sum_{\gamma \in \Sigma'(x)} \left(\varepsilon a_\gamma^{(1)}(x) + R_{\gamma}^{(2)}(x, \varepsilon)\right) \cosh\left(\frac{1}{\varepsilon}(\Phi(x) - \Phi(x + \gamma))\right) \right| \leq \tilde{C}\varepsilon. \quad (3.36)$$

uniformly with respect to  $x$ .

We have for  $x \in \Sigma$

$$\left|D_2(x)\right| \leq \sum_{\gamma \in \Sigma'_\varepsilon(x)} |a_\gamma^{(0)}(x)| \left| \cosh\left(\frac{1}{\varepsilon}(\Phi(x) - \Phi(x + \gamma))\right) - \cosh\left(\frac{1}{\varepsilon}\gamma \nabla\Phi(x)\right) \right|. \quad (3.37)$$

By the mean value theorem for  $\cosh z$  with  $z_0 = \frac{1}{\varepsilon}\gamma\nabla\Phi(x)$  and  $z_1 = \frac{1}{\varepsilon}(\Phi(x) - \Phi(x + \gamma))$ , we get from  $|\sinh x| \leq e^{|x|}$

$$\begin{aligned} & \left| \cosh\left(\frac{1}{\varepsilon}(\Phi(x) - \Phi(x + \gamma))\right) - \cosh\left(\frac{1}{\varepsilon}\gamma\nabla\Phi(x)\right) \right| \\ & \leq \sup_{t \in [0,1]} e^{|\frac{1}{\varepsilon}\{(\Phi(x) - \Phi(x + \gamma))t + \gamma\nabla\Phi(x)(1-t)\}|} \left| \frac{1}{\varepsilon}\{(\Phi(x) - \Phi(x + \gamma)) + \gamma\nabla\Phi(x)\} \right|. \end{aligned} \tag{3.38}$$

By (1.17) and the definition (3.12) of  $\Phi$  there exist constants  $c_1, c_2 > 0$  such that  $|\Phi(x) - \Phi(x + \gamma)| \leq c_1|\gamma|$  and  $|\gamma\nabla\Phi(x)| \leq c_2|\gamma|$ ,  $x \in \Sigma$ ,  $\gamma \in \Sigma'_\varepsilon(x)$ . (3.39)

(3.39) gives a constant  $D > 0$  such that

$$e^{|\frac{1}{\varepsilon}\{(\Phi(x) - \Phi(x + \gamma))t + \gamma\nabla\Phi(x)(1-t)\}|} \leq e^{\frac{D}{\varepsilon}|\gamma|}. \tag{3.40}$$

By second order Taylor-expansion

$$\frac{1}{\varepsilon} |(\Phi(x) - \Phi(x + \gamma)) + \gamma\nabla\Phi(x)| \leq \sup_{t \in [0,1]} \frac{1}{\varepsilon} \sum_{\nu, \mu=1}^d |\gamma_\nu \gamma_\mu \partial_\nu \partial_\mu \Phi(x + t\gamma)|. \tag{3.41}$$

Inserting (3.13) into (3.41) shows that there exists a constant  $C_3 > 0$  independent of the choice of  $B$  such that for all  $\varepsilon \in (0, \varepsilon_0]$

$$\frac{1}{\varepsilon} |(\Phi(x) - \Phi(x + \gamma)) + \gamma\nabla\Phi(x)| \leq \frac{C_3}{\varepsilon} |\gamma|^2. \tag{3.42}$$

By (1.9), inserting (3.40) and (3.42) in (3.37) we get for any  $A > 0$  with  $\eta = \frac{\gamma}{\varepsilon}$

$$\begin{aligned} |V_0^\Phi(x) - t_0^\Sigma(x, -i\nabla\Phi)| & \leq \sum_{\frac{\gamma}{\varepsilon} \in \Sigma'_\varepsilon(x)} e^{-A|\eta|} e^{D|\eta|} C_3 \varepsilon |\eta|^2 \\ & \leq \varepsilon \sum_{\eta \in \mathbb{Z}^d} e^{-|\eta|D'} C_3 |\eta|^2 \leq \varepsilon C_4, \end{aligned}$$

where  $A - D = D' > 0$ . This together with (3.36) gives (3.35).

*Step 3:* We prove (1.20) by use of Lemma 3.2.

Choosing  $B \geq C_0(1 + R_0 + C_5)$ , we have

$$\left(\frac{B}{C_0} - C_5\right) \varepsilon - E \geq \varepsilon, \quad E \in [0, \varepsilon R_0]. \tag{3.43}$$

Let

$$\Omega_- := \{x \in \Sigma \mid \widehat{V}_\varepsilon(x) + V^\Phi(x) - E < 0\} \quad \text{and} \quad \Omega_+ := \Sigma \setminus \Omega_-, \tag{3.44}$$

then from (3.43) it follows that  $\Omega_- \subset \{d(x) < \varepsilon B\}$  and by (3.32)

$$|\widehat{V}_\varepsilon(x) + V^\Phi(x)| \leq \varepsilon \max\{C_5, R_0\} \quad \text{for all } x \in \Omega_-. \tag{3.45}$$

We define the functions  $F_\pm : \Sigma \rightarrow [0, \infty)$  by

$$F_+(x) := \sqrt{\varepsilon \mathbf{1}_{\{d(x) < B\varepsilon\}}(x) + (\widehat{V}_\varepsilon(x) + V^\Phi(x) - E) \mathbf{1}_{\Omega_+}(x)} \tag{3.46}$$

and

$$F_-(x) := \sqrt{\varepsilon \mathbf{1}_{\{d(x) < B\varepsilon\}}(x) + (E - \widehat{V}_\varepsilon(x) - V^\Phi(x)) \mathbf{1}_{\Omega_-}(x)}. \tag{3.47}$$

Then  $F_\pm$  are well defined and furthermore there exists a constant  $C > 0$  such that

$$F := F_+ + F_- \geq C \sqrt{\varepsilon} > 0, \quad F_- = O(\sqrt{\varepsilon}) \quad \text{and} \quad F_+^2 - F_-^2 = \widehat{V}_\varepsilon + V^\Phi - E. \tag{3.48}$$

Lemma 3.2 yields with the choice  $v = e^{\frac{\Phi}{\varepsilon}} u$

$$\left\| F e^{\frac{\Phi}{\varepsilon}} u \right\|_{\ell^2}^2 \leq 4 \left\| \frac{1}{F} e^{\frac{\Phi}{\varepsilon}} (H_\varepsilon^\Sigma - E) u \right\|_{\ell^2}^2 + 8 \left\| F_- e^{\frac{\Phi}{\varepsilon}} u \right\|_{\ell^2}^2 + C\varepsilon \|u\|^2. \tag{3.49}$$

By (3.14) and (3.48)

$$\left\| F e^{\frac{\Phi}{\varepsilon}} u \right\|_{\ell^2}^2 \geq C\varepsilon \left\| \left(1 + \frac{d}{\varepsilon}\right)^{-\frac{B}{2}} e^{\frac{d}{\varepsilon}} u \right\|_{\ell^2}^2 \tag{3.50}$$

and

$$\left\| \frac{1}{F} e^{\frac{\Phi}{\varepsilon}} (H_\varepsilon^\Sigma - E) u \right\|_{\ell^2}^2 \leq C\varepsilon^{-1} \left\| \left(1 + \frac{d}{\varepsilon}\right)^{-\frac{B}{2}} e^{\frac{d}{\varepsilon}} (H_\varepsilon^\Sigma - E) u \right\|_{\ell^2}^2. \tag{3.51}$$

Since  $\Omega_- \subset \{d(x) < B\varepsilon\}$  it follows from the definition of  $F_-$  that  $\frac{d(x)}{\varepsilon} \leq C$  on the support of  $F_-$ . Therefore by (3.14) and (3.48) there exists a constant  $C > 0$  such that

$$\left\| F_- e^{\frac{\Phi}{\varepsilon}} u \right\|_{\ell^2}^2 \leq C\varepsilon \|u\|_{\ell^2}^2. \tag{3.52}$$

Inserting (3.50), (3.51) and (3.52) in equation (3.49) yields with  $\tilde{B} := \frac{B}{2}$

$$\tilde{C}\varepsilon \left\| \left(1 + \frac{d}{\varepsilon}\right)^{-\tilde{B}} e^{\frac{d}{\varepsilon}} u \right\|_{\ell^2}^2 \leq \varepsilon^{-1} \left\| \left(1 + \frac{d}{\varepsilon}\right)^{-\tilde{B}} e^{\frac{d}{\varepsilon}} (H_\varepsilon^\Sigma - E) u \right\|_{\ell^2}^2 + \varepsilon \|u\|_{\ell^2}^2.$$

This proves (1.20).

*Step 4:* We prove (1.21).

If  $u$  is an eigenfunction of  $H_\varepsilon^\Sigma$  with eigenvalue  $E$ , then the first summand on the right hand side of (1.20) vanishes. The normalization of  $u$  leads therefore to (1.21).  $\square$

### 4. Application to Markov chains

An example of self adjoint difference operators as analyzed above are generators of certain Markov chains and jump processes. For the sake of the reader we briefly recall some relevant facts on Markov chains.

A Markov chain on  $(\varepsilon\mathbb{Z})^d$  is described by means of a “transition matrix”  $P_\varepsilon : (\varepsilon\mathbb{Z})^d \times (\varepsilon\mathbb{Z})^d \rightarrow [0, 1]$ .  $P_\varepsilon$  is a stochastic matrix, i.e.,

$$\sum_{y \in (\varepsilon\mathbb{Z})^d} P_\varepsilon(x, y) = 1, \quad x \in (\varepsilon\mathbb{Z})^d. \tag{4.1}$$

We assume that  $P_\varepsilon$  satisfies a detailed balance condition, i.e.,

$$\mu_\varepsilon(x)P_\varepsilon(x, y) = \mu_\varepsilon(y)P_\varepsilon(y, x) \tag{4.2}$$

with respect to some family  $\{\mu_\varepsilon\}_{\varepsilon \in (0, \varepsilon_0]}$  of probability measures on  $(\varepsilon\mathbb{Z})^d$ . Then  $(1 - P_\varepsilon)$  defines a self adjoint (diffusion) operator on  $\ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)$  via

$$(1 - P_\varepsilon)u(x) = u(x) - \sum_{y \in (\varepsilon\mathbb{Z})^d} P_\varepsilon(x, y)u(y).$$

In fact  $P_\varepsilon$  is a bounded operator on  $\ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)$  with  $\|P_\varepsilon\| = 1$ . To see this, we first notice that by (4.1)

$$\begin{aligned} |P_\varepsilon u(x)|^2 &\leq \left( \sum_{y \in (\varepsilon\mathbb{Z})^d} P_\varepsilon(x, y) \right) \left( \sum_{\gamma} P_\varepsilon(x, y) |u(y)|^2 \right) \\ &= \sum_{\gamma} P_\varepsilon(x, y) |u(y)|^2. \end{aligned}$$

This yields by (4.2), the Fubini-Theorem and again (4.1)

$$\begin{aligned} \|P_\varepsilon u\|_{\ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)}^2 &= \sum_{x \in (\varepsilon\mathbb{Z})^d} \mu_\varepsilon(x) |P_\varepsilon u(x)|^2 \leq \sum_x \mu_\varepsilon(x) \sum_y P_\varepsilon(x, y) |u(y)|^2 \\ &= \sum_y \left( \sum_x P_\varepsilon(y, x) \right) \mu_\varepsilon(y) |u(y)|^2 \\ &= \|u\|_{\ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)}^2, \end{aligned}$$

thus  $\|P_\varepsilon\| \leq 1$ . Since the constant function  $u(x) = 1$  belongs to  $\ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)$  and fulfills  $\|P_\varepsilon u\|_{\ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)} = \|u\|_{\ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)}$ , this proves that  $\|P_\varepsilon\| = 1$

The symmetry of  $P_\varepsilon$  follows from the reversibility condition (4.2), since for  $u, v \in \ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)$

$$\begin{aligned} \langle u, P_\varepsilon v \rangle_{\ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)} &= \sum_{x \in (\varepsilon\mathbb{Z})^d} \mu_\varepsilon(x) u(x) \sum_{y \in (\varepsilon\mathbb{Z})^d} P_\varepsilon(x, y) v(y) \\ &= \sum_y \sum_x \mu_\varepsilon(y) P_\varepsilon(y, x) u(x) v(y) = \langle P_\varepsilon u, v \rangle_{\ell^2((\varepsilon\mathbb{Z})^d, \mu_\varepsilon)}. \end{aligned}$$

Conjugation with respect to the measure  $\mu_\varepsilon$  induces a bounded self adjoint operator  $H_\varepsilon := \mu_\varepsilon^{\frac{1}{2}}(1 - P_\varepsilon)\mu_\varepsilon^{-\frac{1}{2}}$  on  $\ell^2((\varepsilon\mathbb{Z})^d)$ , whose restriction to  $\mathcal{K}((\varepsilon\mathbb{Z})^d)$  is given by

$$H_\varepsilon u(x) = u(x) - \mu_\varepsilon^{\frac{1}{2}}(x) \sum_{\gamma} P_\varepsilon(x, x+\gamma) \mu_\varepsilon^{-\frac{1}{2}}(x+\gamma) u(x+\gamma), \quad u \in \mathcal{K}((\varepsilon\mathbb{Z})^d). \tag{4.3}$$

Note that  $\mathcal{K}((\varepsilon\mathbb{Z})^d)$  is dense in  $\ell^2((\varepsilon\mathbb{Z})^d)$  and  $H_\varepsilon$  is linear continuous and is therefore completely determined by (4.3).

**Proposition 4.1.** *The operator  $H_\varepsilon := \mu_\varepsilon^{\frac{1}{2}}(1 - P_\varepsilon)\mu_\varepsilon^{-\frac{1}{2}}$  on  $\ell^2((\varepsilon\mathbb{Z})^d)$  is of the form (1.1) and fulfills Hypothesis 1.1 (a)(iii). If the coefficients  $a_\gamma$  have an expansion (1.8), they also fulfill (ii).*

*Proof.* Setting

$$T_\varepsilon(x) := \sum_{\gamma \neq 0} \mu_\varepsilon^{\frac{1}{2}}(x) P_\varepsilon(x, x + \gamma) \mu_\varepsilon^{-\frac{1}{2}}(x + \gamma) (\mathbf{1} - \tau_\gamma) \tag{4.4}$$

$$V_\varepsilon(x) := \sum_{\gamma \neq 0} \mu_\varepsilon^{\frac{1}{2}}(x) P_\varepsilon(x, x + \gamma) \left( \mu_\varepsilon^{-\frac{1}{2}}(x) - \mu_\varepsilon^{-\frac{1}{2}}(x + \gamma) \right) \tag{4.5}$$

we have the standard form  $H_\varepsilon = T_\varepsilon + V_\varepsilon$ , where  $V_\varepsilon$  is a potential energy (a multiplication operator) and  $T_\varepsilon$  is of the form described in (1.1) with

$$a_0(x, \varepsilon) = \sum_{\gamma \neq 0} \mu_\varepsilon^{\frac{1}{2}}(x) P_\varepsilon(x, x + \gamma) \mu_\varepsilon^{-\frac{1}{2}}(x + \gamma) \geq 0 \tag{4.6}$$

$$a_\gamma(x, \varepsilon) = -\mu_\varepsilon^{\frac{1}{2}}(x) P_\varepsilon(x, x + \gamma) \mu_\varepsilon^{-\frac{1}{2}}(x + \gamma), \quad \gamma \neq 0. \tag{4.7}$$

Since  $P_\varepsilon(x, y)$  and  $\mu_\varepsilon(x)$  are non-negative numbers, it follows at once that  $a_\gamma(x, \varepsilon) \leq 0$  for all  $\gamma \neq 0$  and that  $\sum a_\gamma(x, \varepsilon) = 0$ . Thus under the assumption (1.8) it follows that Hypothesis 1.1 (a)(ii) holds. The detailed balance condition for  $P_\varepsilon$  ensures the symmetry of  $H_\varepsilon$  and thus of  $T_\varepsilon$ . By (2.71), this leads to  $a_\gamma(x, \varepsilon) = a_{-\gamma}(x + \gamma, \varepsilon)$  (Hypothesis 1.1(a)(iii)).  $\square$

*Remark 4.2.* The other conditions given in Hypothesis 1.1 lead to analog conditions on the transition matrix  $P_\varepsilon$  and the reversible measure  $\mu_\varepsilon$ . For example, condition (a)(iv) on the exponential decay of  $a_\gamma$  with respect to  $\gamma$ , must be reflected by the fact that  $P_\varepsilon(x, y)$  is assumed to be exponential small for  $|x - y|$  large. Furthermore in order to fulfill (a)(i), the measure  $\mu_\varepsilon$  should be slowly varying. Condition (a)(v) is a kind of ergodicity condition, which guarantees that jumps in each direction are possible. It follows at once from (4.4) and (4.5) that for a general probabilistic operator, the potential energy can be written in terms of the kinetic energy and the measure  $\mu_\varepsilon$  as

$$V_\varepsilon(x) = -\mu_\varepsilon^{-\frac{1}{2}}(x) \left( T_\varepsilon \mu_\varepsilon^{\frac{1}{2}} \right) (x). \tag{4.8}$$

The assumptions on  $V_0$  given in Hypothesis 1.1 are conditions on the pair  $(\mu_\varepsilon, P_\varepsilon)$ . The class of Markov chains satisfying these conditions is more general than the class of Markov chains analyzed in Bovier–Eckhoff–Gayraud–Klein [9], if the Markov chain acts on  $(\varepsilon\mathbb{Z})^d$ .

### References

[1] M. Abate, G. Patrizio, *Finsler Metrics – A Global Approach*, LNM 1591, Springer, 1994.  
 [2] R. Abraham, J. E. Marsden, *Foundations of Mechanics*, 2.ed., The Benjamin / Cummings Pub.Comp., 1978.

- [3] S. Agmon, *Lectures on Exponential Decay of Solutions of Second-order Elliptic Equations: Bounds on Eigenfunctions of N-Body Schrödinger Operators*, Mathematical Notes 29, Princeton University Press, 1982.
- [4] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2.ed., Springer-Verlag, 1989.
- [5] D. Bao, S.-S. Chern, Z. Shen, *An Introduction to Riemann–Finsler Geometry*, GTM 200, Springer, 2000.
- [6] G. Barbatis, *Sharp heat Kernel bounds and Finsler-type metrics*, Quart. J. Math. Oxford **2**, 49 (1998), 261–277.
- [7] G. Barbatis, *Explicit estimates on the fundamental solution of higher-order parabolic equations with measurable coefficients*, Journal of Diff. Equations **174** (2001), 442–463.
- [8] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, *Metastability in stochastic dynamics of disordered mean-field models*, Probab. Theory Relat. Fields **119** (2001), 99–161.
- [9] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, *Metastability and low lying spectra in reversible Markov chains*, Comm. Math. Phys. **228** (2002), 219–255.
- [10] M. Dimassi, J. Sjöstrand, *Spectral Asymptotics in the Semi- Classical Limit*, London Mathematical Society Lecture Note Series 268, Cambridge University Press, 1999.
- [11] L. C. Evans, R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, 1992.
- [12] M. Giaquinta, S. Hildebrandt, *Calculus of Variations 2*, GMW 311, Springer, 1996.
- [13] R. S. Hamilton, *The Inverse Function Theorem of Nash and Moser*, Bulletin of the American National Society, Vol. 7, Number 1, 1982.
- [14] B. Helffer, *Décroissance exponentielle des fonctions propres pour l'opérateur de Kac: Le cas de la dimension  $> 1$* , Operator calculus and spectral theory (Lambrecht, 1991), 99–115, Oper. Theory Adv. Appl. **57**, Birkhäuser, Basel, 1992.
- [15] B. Helffer, A. Mohamed, *Semiclassical analysis for the ground state energy of a Schrödinger operator with magnetic wells*, J. Funct. Anal. **138** (1996), no. 1, p. 40–81.
- [16] B. Helffer, B. Parisse, *Comparaison entre la décroissance de fonctions propres pour les opérateurs de Dirac et de Klein–Gordon. Application à l'étude de l'effet tunnel*, Ann. Inst. H. Poincaré Phys. Théor. **60** (1994), no. 2, 147–187.
- [17] B. Helffer, J. Sjöstrand, *Multiple wells in the semi-classical limit I*, Comm. in P.D.E. **9** (1984), 337–408.
- [18] B. Helffer, J. Sjöstrand, *Analyse semi-classique pour l'équation de Harper (avec application à l'équation de Schrödinger avec champ magnétique)*, Mém. Soc. Math. France (N.S.) **34** (1988), 1–113.
- [19] W. Klingenberg, *Riemannian Geometry*, de Gruyter Studies in Mathematics **1** (1982).
- [20] S. Lang, *Differential and Riemannian Manifold*, 3.ed., Springer, 1995.
- [21] C. Mantegazza, A. C. Mennucci, *Hamilton–Jacobi equations and distance functions on Riemannian manifolds*, Appl. Math. Opt. **47**, 1 (2003), 1–25.
- [22] E. Rosenberger, *Asymptotic Spectral Analysis and Tunnelling for a class of Difference Operators*, Thesis, <http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-7393>.

- [23] E. Servat, *Le splitting pour l'opérateur de Klein–Gordon: Une approche heuristique et numérique*, *Canad. J. Math.* **59** (2007), no. 2, 393–417.
- [24] B. Simon, *Semiclassical analysis of low lying eigenvalues. I. Nondegenerate minima: Asymptotic expansions*, *Ann Inst. H. Poincaré Phys. Theor.* **38** (1983), 295–308.
- [25] B. Simon, *Semiclassical analysis of low lying eigenvalues. II. Tunneling*, *Ann. of Math.* **120** (1984), 89–118.
- [26] K. Tintarev, *Short time asymptotics for fundamental solutions of higher order parabolic equations*, *Comm. in Part. Diff. Equ.* **7** (4) (1982), 371–391.

Markus Klein and Elke Rosenberger  
Universität Potsdam  
Institut für Mathematik  
Am Neuen Palais 10  
D-14469 Potsdam  
Germany  
e-mail: [mklein@math.uni-potsdam.de](mailto:mklein@math.uni-potsdam.de)  
[erosen@rz.uni-potsdam.de](mailto:erosen@rz.uni-potsdam.de)

Communicated by Christian Gérard.

Submitted: February 23, 2008.

Accepted: May 23, 2008.