Transference Plans and Uncertainty

Kurt Pagani

October 2015

Abstract

We will discuss methods of *Optimal Transportation Theory* and its relations to problems in quantum mechanics. This essentially means that the cost function is some Hamiltonian $\mathcal{H}(q,p)$ on phase space (symplectic manifold), and the marginal measures that have to be transported are linked by a (implicit) transformation group.

Contents

This article is a (mostly non-technical) overview, showing the main connections and some unsettled questions. To keep the size within reasonable bounds, we can only scratch the surface, that is we will consider the *flat* case and the unitary Fourier transformation as generator while omitting time dependence at all. Since we only have to deal with locally compact spaces (actually, more often than not $\mathbb{R}^d$), *measures* will be understood by the functional approach.

Transportation theory (see e.g. [1],[18]) is about - among other things - *transference plans*, i.e. probability measures γ on a product space $X \times Y$, such that the *marginal measures* [1] on X and Y respectively coincide with two fixed measures μ and ν, which in turn are the objects that shall be transported by a so called transport map $T : X \longrightarrow Y$ by *push-forward*: $T_{\#}\mu = \nu$. Typical questions: *is there a mapping $p = T(q)$ such that*

$$\int_{X \times Y} \mathcal{H}(q,p)\, d\gamma(q,p) = \int_{X \times Y} \mathcal{H}(q,T(q))\, d\gamma(q,p) = \int_X \mathcal{H}(q,T(q))\, d\mu(q),$$

i.e. $d\gamma(q,p) = d\mu(q)\, \delta_{T(q)}(p)$, or: *what can we say about*

$$\sup_{\gamma \in \Gamma(\mu,\nu)} \int_{X \times Y} \chi_\Lambda(q,p)\, d\gamma(q,p),$$

for a compact subset Λ? There is a wealth of new ideas and methods which will remain unmentioned here, but may be easily adapted from the excellent book [17].

---

[1]That is $\pi_{X,\#}\mu$ and $\pi_{Y,\#}\mu$, the push-forwards by the projection maps

1   Overview and Notation

In the sequel we are going to use some results and terminology of *mass transportation theory*, where [18] serves as the main reference.

Let $\mathcal{P}(\mathbb{R}^n)$ denote the space of probability measures on $\mathbb{R}^n$ and for $\varphi \in L^2(\mathbb{R}^n)$ let $\hat{\varphi}$ denote its (unitary) Fourier transform. Each normalized $\varphi \in L^2(\mathbb{R}^n)$ gives rise to a measure

$$\nu_\varphi(f) = \int_{\mathbb{R}^n} f(x)|\varphi(x)|^2 \, dx,$$

where $f \in C_0(\mathbb{R}^n)$, the continuous functions with compact support. Then we define the mapping

$$\mu : L^2(\mathbb{R}^n) \cap \{||\varphi|| = 1\} \longrightarrow \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n) \tag{1}$$

$$\varphi \longmapsto \nu_\varphi \otimes \nu_{\hat{\varphi}},$$

that means μ_φ is the (unique) product measure with marginals ν_φ and $\nu_{\hat{\varphi}}$. Furthermore we denote by $\Gamma(\varphi)$ the subset of $\mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ whose elements have the aforementioned marginals.

Let $\mathcal{H} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ be lower semi-continuous and bounded below, then we call

$$K_{\mathcal{H}}(\varphi) = \inf_{\gamma \in \Gamma(\varphi)} \int_{\mathbb{R}^n \times \mathbb{R}^n} \mathcal{H}(q, p) \, d\gamma(q, p) \tag{2}$$

the *Kantorovich* energy of φ. Similarly we call

$$E_{\mathcal{H}}(\varphi) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \mathcal{H}(q, p) \, d\mu_\varphi(q, p) \tag{3}$$

the *Schrödinger* energy, for reasons that will be enlightened soon. Monge's formulation of the *optimal transport problem* reads in our case:

$$M_{\mathcal{H}}(\varphi) = \inf_T \left\{ \int_{\mathbb{R}^n} \mathcal{H}(q, T(q)) d\nu_\varphi : T_{\#}\nu_\varphi = \nu_{\hat{\varphi}} \right\}, \tag{4}$$

which means to find a minimizing map $T : \mathbb{R}^n \to \mathbb{R}^n$, that transports the measure ν_φ to $\nu_{\hat{\varphi}}$ by pushing forward [2]:

$$T_{\#}\nu_\varphi(f) := \nu_\varphi(f \circ T) = \int_{\mathbb{R}^n} f(T(q))|\varphi(q)|^2 \, dq = \int_{\mathbb{R}^n} f(p)|\hat{\varphi}(p)|^2 \, dp.$$

If all quantities involved were smooth enough and T one to one, then we would get the condition

$$|\varphi(q)|^2 = |\hat{\varphi}(T(q))|^2 \, |\det DT(q)| \tag{5}$$

by a simple change of coordinates.

---

[2] $T_{\#}\nu(f) = \nu(T^{\#}(f)$, where $T^{\#}(f) = f \circ T$ denotes *pull-back*

1.1 Schrödinger Energy

Suppose $\mathcal{H}$ has the familiar form $\mathcal{H}(x,k) = \frac{\hbar^2}{2m}|k|^2 + \mathcal{V}(x)$, then we easily calculate that

$$K_{\mathcal{H}}(\varphi) = E_{\mathcal{H}}(\varphi) = \frac{\hbar^2}{2m}\int_{\mathbb{R}^n} |k|^2|\hat{\varphi}(k)|^2 dk + \int_{\mathbb{R}^n} \mathcal{V}(x)|\varphi(x)|^2 dx$$

holds. Furthermore, if $\partial_j\varphi \in L^2(\mathbb{R}^n)$, then the above expression reduces to

$$K_{\mathcal{H}}(\varphi) = E_{\mathcal{H}}(\varphi) = \int_{\mathbb{R}^n} \left(\frac{\hbar^2}{2m}|\nabla\varphi(x)|^2 + \mathcal{V}(x)|\varphi(x)|^2 \right) dx. \qquad (6)$$

Whether the energy is finite or not will also depend on the behaviour of $\mathcal{V}$, of course. In a similar way the above deduction holds whenever the *cost function* $\mathcal{H}$ has the form $\mathcal{H}(q,p) = \mathcal{T}(p) + \mathcal{V}(q)$, that is the *Kantorovich energy* coincides with $E_{\mathcal{H}}$ which in turn means that the *transference plan* $\mu_\varphi = \nu_\varphi \otimes \nu_{\hat{\varphi}}$ is optimal. Villani notes with reference to the sand pile example [18]

> ... this corresponds to the most stupid transportation plan that one may imagine: any piece of sand, regardless of its location, is distributed over the entire hole, proportionally to the depth.

He certainly would not claim that quantum mechanics were stupid, however, we recognize that the procedure mentioned is just another formulation of the uncertainty principle (replacing sand pile/hole by position/momentum, although this analogy should not be taken too serious). This is in strong contrast to the corresponding *Monge* problem (omitting the factor $\hbar^2/2m$ from now on),

$$M_{\mathcal{H}}(\varphi) = \inf_T \left\{ \int_{\mathbb{R}^n} \left(|T(x)|^2 + \mathcal{V}(x) \right) d\nu_\varphi : T_\#\nu_\varphi = \nu_{\hat{\varphi}} \right\},$$

where, since T is a map, there is no such distribution (*mass cannot be split by Monge transport*). Although we speak here of *virtual* transport, the analogies are sometimes useful, in that $d\gamma(x,k)$ measures the amount of *mass* transferred from x to k. Therefore, a general γ may *smear out* x (à la multi-valued mappings), whereas a transference plan of the form $(id \times T)_\#\mu$ cannot. On the other hand, assume $\gamma = \mu \otimes \delta_\kappa$ (an extreme case which is of no concern in this paper), yields $\int \mathcal{H}(x,k)d\gamma(x,k) = \int \mathcal{H}(x,\kappa)d\mu(x)$, this means, everything will be transported to κ. Such pathologies are excluded for $\gamma \in \Gamma(\varphi)$, of course.

4

By a theorem of Brenier-McCann ([14], see Appendix Thm. 2), there is a convex function ϕ on $\mathbb{R}^n$ such that

$$(\nabla\phi)_{\#}\nu_\varphi = \nu_{\hat\varphi},$$

whence we have [3]

$$K_{\mathcal{H}}(\varphi) \leq M_{\mathcal{H}}(\varphi) \leq \int_{\mathbb{R}^n} \left(|\nabla\phi(x)|^2 + \mathcal{V}(x)\right)|\varphi(x)|^2 dx = E_{\mathcal{H}}(\varphi).$$

Therefore, all three quantities coincide in case of $\mathcal{H}(x,k) = \mathcal{T}(k)+\mathcal{V}(x)$. This is no surprise because $\int(\mathcal{T}+\mathcal{V})d\gamma$ is constant on $\Gamma(\varphi)$. Now, if we supposed for the moment the existence of a ground state $\varphi_0 > 0$ to $E_{\mathcal{H}}$ (more precisely to the self-adjoint operator corresponding to $\mathcal{H}$), we would find the identities

$$E_{\mathcal{H}}(\varphi_0) = K_{\mathcal{H}}(\varphi_0) = \int_{\mathbb{R}^n}\left(|\nabla\log\varphi_0(x)|^2 + \mathcal{V}(x)\right)|\varphi_0(x)|^2 dx = M_{\mathcal{H}}(\varphi_0).$$

This leads to the question:

Can $\nabla\phi = -\nabla\log\varphi_0$ be a *Brenier* map?

In the first place $\phi = -\log\varphi_0$ is required to be convex, or equivalently, the ground state $\varphi_0(x) = Ce^{-\phi(x)}$ should be `log`-concave, a property that is not uncommon for certain potentials V. A far more stringent condition, however, is the requirement $(-\nabla\log\varphi_0)_{\#}\nu_{\varphi_0} = \nu_{\hat\varphi_0}$, which, assuming some smoothness and recalling (5), reads as

$$|\varphi_0(q)|^2 = |\hat\varphi_0(\nabla\phi(q))|^2|\det D^2\phi(q)|. \tag{7}$$

Actually, the ground state of the harmonic oscillator $\varphi_{ho}(x) = Ce^{-\frac{1}{2}|x|^2}$ satisfies the above equation and consequently in this particular case $T(x) = \nabla\phi(x) = -\nabla\log\varphi_{ho}(x) = x$ is a transport map. Are there others? Probably not, but we do not know.

1.2 General $\mathcal{H}$

If the Hamilton function $\mathcal{H}$ does not split up as above, then we only have $K_{\mathcal{H}}(\varphi) \leq E_{\mathcal{H}}(\varphi)$ and since the infimum in (1) is always attained [4] , there is a $\gamma_\varphi \in \Gamma(\varphi)$ such that $\gamma_\varphi(\mathcal{H}) \leq \mu_\varphi(\mathcal{H})$. In the following let us denote by $\Gamma_n = \mathbb{R}^n \oplus \mathbb{R}^n$ a $2n$-dimensional phase space, where there should be no

---

[3]notice that $K_{\mathcal{H}}$ is a relaxation of $M_{\mathcal{H}}$ since an admissible transport map T always gives rise to a transference plan $(id \times T)_{\#}\nu_\varphi \in \Gamma(\varphi)$

[4]under the conditions given at the beginning

confusion among the meanings of Γ, e.g. we have $\Gamma(\varphi) \subset \mathcal{P}(\Gamma_n)$. Whether the minimization problem

$$\lambda_0 = \inf \left\{ \int_{\Gamma_n} \mathcal{H} d\mu_\varphi : \varphi \in L^2(\mathbb{R}^n), \|\varphi\|_2 = 1 \right\} \tag{8}$$

has a solution will depend on the function $\mathcal{H}$ under consideration, and even if there is a solution, it is by no means granted that it will be a ground state of a corresponding self-adjoint *Hamiltonian*. Existence questions will not be our concern at this point, therefore we will take the existence of a minimizer $\varphi_0 \in L^2(\mathbb{R}^n)$ for granted. Since we have assumed the function $\mathcal{H}$ to be bounded below (and l.s.c) it is obvious that

$$\lambda_0 \geq \inf_{\Gamma_n} \mathcal{H} > -\infty$$

and moreover it holds that $\lambda_0 = E_\mathcal{H}(\varphi_0) \geq K_\mathcal{H}(\varphi_0)$. When we define (assuming $\mathcal{H}$ fixed)

$$F_\varphi(x) = \int_{\mathbb{R}^n} \mathcal{H}(x,k) \, |\hat{\varphi}(k)|^2 dk$$

as well

$$G_\varphi(k) = \int_{\mathbb{R}^n} \mathcal{H}(x,k) \, |\varphi(x)|^2 dx$$

and recall that $\mu_\varphi = \nu_\varphi \otimes \nu_{\hat{\varphi}}$ holds by definition, we obtain

$$E_\mathcal{H}(\varphi) = \int_{\Gamma_n} \mathcal{H}(x,k) \, d\mu_\varphi(x,k) = \int_{\mathbb{R}^n} F_\varphi(x) d\nu_\varphi(x) = \int_{\mathbb{R}^n} G_\varphi(k) d\nu_{\hat{\varphi}}(k).$$

Now we may state the *Euler equations* which a minimizer must satisfy.

Proposition 1. *Let $\varphi_0 \in L^2(\mathbb{R}^n)$ be a critical point of $E_\mathcal{H}(\varphi)$, then it satisfies the equation (in $\mathcal{D}'(\mathbb{R}^n)$)*

$$(2E_0 - F_{\varphi_0}(x)) \, \varphi_0(x) = \int_{\mathbb{R}^n} G_{\varphi_0}(k) \, \hat{\varphi}_0(k) \, e^{i\langle k,x \rangle} \, dm_n(k), \tag{9}$$

where $E_0 = E_\mathcal{H}(\varphi_0)$ and $dm_n(k) := (2\pi)^{-\frac{n}{2}} dk$.

Whether (9) is valid almost everywhere w.r.t. Lebesgue measure depends (here again) on the function $\mathcal{H}$. The inverse Fourier transform on the right hand side should be understood symbolically, unless $G_{\varphi_0}\hat{\varphi}_0 \in L^1(\mathbb{R}^n)$. In a compact notation the equation for a critical point of $E_\mathcal{H}$ is

$$F_\varphi \, \varphi + \breve{G}_\varphi \star \varphi = 2\lambda \, \varphi$$

6

where the convolution is defined here as $(f \star g)(x) = \int_{\mathbb{R}^n} f(x-y) \, g(y) \, dm_n(y)$. It is easily checked that in case of $\mathcal{H}(x,k) = |k|^2 + \mathcal{V}(x)$, (9) reduces to $(-\Delta + \mathcal{V}(x))\varphi_0 = E_0\varphi_0$. Our main interest, however, is $\mathcal{H}$ being the indicator function of an open subset of Γ_n which is obviously bounded, measurable and lower semi-continuous. This is, as will be outlined further below, connected to the question:

How big can we make

$$\int_\Lambda |\varphi(x)|^2 \, |\hat{\varphi}(k)|^2 dx \, dk, \qquad (10)$$

given a compact subset Λ of phase space Γ_n?

Actually, the question may be posed for $\Lambda \subset \Gamma_n$ having finite Lebesgue measure.

1.3 Duality

One of the corner stones of mass transportation theory certainly is *Kantorovich's* duality formula ([18], Theorem 1.3) which, translated to our needs, says

$$K_{\mathcal{H}}(\varphi) = \sup_{\mathcal{T}(k)+\mathcal{V}(x)\leq\mathcal{H}(x,k)} \left\{ \int_{\mathbb{R}^n} \mathcal{T}(k)|\hat{\varphi}(k)|^2 \, dk + \int_{\mathbb{R}^n} \mathcal{V}(x)|\varphi(x)|^2 \, dx \right\}, \quad (11)$$

where the functions $\mathcal{T}, \mathcal{V}$ may either be any bounded continuous functions on $\mathbb{R}^n$ or by extension $(\mathcal{T}, \mathcal{V}) \in L^1(\nu_\varphi) \times L^1(\nu_{\hat{\varphi}})$, satisfying the inequality $\mathcal{T} + \mathcal{V} \leq \mathcal{H}$ point-wise in the first case and almost everywhere (with respect to the measures) in the second case. We cite one other result from [18] which will be required later on (a precursor of Strassen's theorem, Theorem 1.27): Let U be a non-empty open subset of Γ_n, then

$$\inf_{\gamma\in\Gamma(\varphi)} \int_U d\gamma = \sup_{A\subset\mathbb{R}^n} \left\{ \int_A |\varphi(x)|^2 \, dx - \int_{A_U} |\hat{\varphi}(k)|^2 \, dk : A \ closed \right\}, \quad (12)$$

where $A_U := \{k \in \mathbb{R}^n : \exists x \in A, (x,k) \notin U\}$. This result implies, setting $\mathcal{H} = \chi_U$,

$$E_{\chi_U}(\varphi) \geq K_{\chi_U}(\varphi) = \sup\{\nu_\varphi(A) - \nu_{\hat{\varphi}}(A_U) : A \subset \mathbb{R}^n, A \ closed\}.$$

Note that we use the notation $\nu(A)$ and $\nu(\chi_A)$ interchangeably when there is no danger of confusion (i.e. we identify a set with its indicator function).

1.4 Symplectic transformations

Let $M : \Gamma_n \to \Gamma_n$ be a symplectic transformation, represented by a matrix of the form (we use the same symbol)

$$M = M^{A,B,C,D} := \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where the $n \times n$ block matrices A, B, C, D satisfy the equations:

$$A^T D - C^T B = I$$
$$A^T C = C^T A$$
$$D^T B = B^T D.$$

Then we obtain for any $f \in C_0(\Gamma_n)$:

$$M_{\#}\mu_\varphi(f) = \mu_\varphi(f \circ M) = \int_{\Gamma_n} f(Ax + Bk, Cx + Dk) d\mu_\varphi(x, k).$$

The inverse M^{-1} of M is easily calculated using the symplectic condition $M^T J M = J$ to

$$M^{-1} = J^{-1} M^T J = \begin{bmatrix} D^T & -B^T \\ -C^T & A^T \end{bmatrix}$$

which implies:

$$M_{\#}\mu_\varphi(f) = \int_{\Gamma_n} f(\xi, \eta) \, |\varphi(D^T \xi - B^T \eta)|^2 \, |\hat{\varphi}(-C^T \xi + A^T \eta)|^2 d\xi d\eta.$$

Simple examples (e.g. $n = 1$ and $\phi(x) = C \exp(-\alpha|x|)$) show that we cannot expect the image measure $M_{\#}\mu_\varphi$ being an element of some $\Gamma(\psi)$. However, two special cases immediately spring to mind:

$$M_{\#}^{A,0,0,D}\mu_\varphi(f) = \int_{\Gamma_n} f(\xi, \eta) \, |\varphi(D^T \xi|^2 \, |\hat{\varphi}(A^T \eta)|^2 d\xi d\eta.$$

and

$$M_{\#}^{0,B,C,0}\mu_\varphi(f) = \int_{\Gamma_n} f(\xi, \eta) \, |\varphi(-B^T \eta)|^2 \, |\hat{\varphi}(-C^T \xi)|^2 d\xi d\eta.$$

In the first case we have $B = C = 0$, so that $A^T D = I$ by the symplectic conditions above. The second case requires $-C^T B = I$ by the same reasoning since $A = D = 0$. Hence there are two subgroups generated by matrices of the form

$$\begin{bmatrix} A & 0 \\ 0 & A^{-T} \end{bmatrix} \; and \; \begin{bmatrix} 0 & B \\ -B^{-T} & 0 \end{bmatrix}.$$

For these, the image measures are indeed of the form $d\mu_\psi$. If we use the notation $\varphi_A(x) = \varphi(Ax)$ we may state:

$$M_\#^{A,0,0,A^{-T}} \mu_\varphi = \mu_{\varphi_{A^{-1}}} \tag{13}$$

and

$$M_\#^{0,B,-B^{-T},0} \mu_\varphi = \mu_{\widehat{\varphi_{B^{-1}}}}. \tag{14}$$

This follows by straightforward computation. Finally we want to mention the special case $B = I_n$ (the identity matrix in $\mathbb{R}^n$), giving M=J, thus

$$J_\# \mu_\varphi(\mathcal{H}) = \mu_\varphi(\mathcal{H} \circ J) = \mu_{\hat\varphi}(\mathcal{H}),$$

which is equivalent to $E_{\mathcal{H} \circ J}(\varphi) = E_\mathcal{H}(\hat\varphi)$. In other words, if $\mathcal{H}$ is invariant under the canonical transformation $x' = k, k' = -x$ and if φ_0 is a unique positive minimum of $E_\mathcal{H}$, then $\varphi_0(x) = C \exp(-|x|^2/2)$.

1.5   Orthonormal Sequences in $L^2(\mathbb{R}^d)$

Let $\{\varphi_j\}_{j \in J}$ be an orthonormal sequence in $L^2(\mathbb{R})$, then a result by H. S. Shapiro, meanwhile known as *Shapiro's Umbrella Theorem*, states that if given two functions $f(x)$ and $g(k)$ in $L^2(\mathbb{R})$ such that

$$|\varphi_j(x)| \leq |f(x)|, \quad |\hat\varphi_j(k)| \leq |g(k)|$$

for all $j \in J$ and for almost all x, k in $\mathbb{R}$, then J must be finite. We refer to [11] and the references therein for background information and more details. Recently. E. Mallinikova ([12],Th. 1.2) showed the following localization property of a orthonormal sequecne $\{\varphi_j\}_{j=1}^N$:

$$N - |A||B| \leq \frac{3}{2} \sum_{j=1}^N \left(\sqrt{\nu_{\varphi_j}(\mathbb{R}^d \backslash A)} + \sqrt{\nu_{\hat\varphi_j}(\mathbb{R}^d \backslash B)} \right) \tag{15}$$

where $A, B \subset \mathbb{R}^d$ are arbitrary measurable sets with finite Lebesgue measure (i.e. $|A|, |B| < \infty$). Remembering the definition of the Radon measures ν_φ at the beginning, $\nu_\varphi(\mathbb{R}^d \backslash A)$ is just

$$\int_{\mathbb{R}^d \backslash A} |\varphi(x)|^2 \, dx.$$

The inequality (15) immediatley leads to a quantitative version of the *Umbrella theorem* ([EM,Th. 4]) as well as to the general inequality

$$\sum_{j=1}^N \int_{\Gamma_d} (|x|^p + |k|^p) \, d\mu_{\varphi_j}(x, k) \geq C\, N^{1+\frac{p}{2d}}, \tag{16}$$

where C depends only on $p > 0$ and d. Moreover, it is also shown that the inequality is sharp up to a multiplicative constant.

1.6 The Nazarov-Jaming Inrequality

Another important result we shall need is the following inequality obtained by Nazarov for the case $d = 1$ and extended by Jaming [10] to $d \geq 1$.

Let $A, B \subset \mathbb{R}^d$, each having finite Lebesgue measure, then there are positive constants [5] α, β and $\eta(A, B)$ such that

$$\nu_\varphi(\mathbb{R}^d \backslash A) + \nu_{\hat{\varphi}}(\mathbb{R}^d \backslash B) \geq \alpha e^{-\beta \eta(A,B)} \tag{17}$$

holds for all $\varphi \in L^2(\mathbb{R}^d)$, $||\varphi|| = 1$. The constant η is given by

$$\eta(A, B) = \begin{cases} |A||B| & : d = 1 \\ \min(|A|\,|B|, |A|^{1/d}\, w(B), w(A)\, |B|^{1/d}) & : d \geq 1 \end{cases}$$

with $w(A)$ the *average width* of A (see [10] for the precise definition).

1.7 Scaling

For $\lambda > 0$ let $\varphi_\lambda(x)$ denote the scaled function $\lambda^{\frac{n}{2}} \varphi(\lambda x)$, then $||\varphi_\lambda|| = 1$ whenever $\varphi \in L^2(\mathbb{R}^n)$ and $||\varphi|| = 1$. The Fourier transform $\widehat{\varphi_\lambda}$ of φ_λ is easily calculated to be equal to $\hat{\varphi}_{1/\lambda}$, therefore

$$\mu_{\varphi_\lambda}(f) = \int_{\Gamma_n} f(x, k)\, |\varphi_\lambda(x)|^2\, |\hat{\varphi}_{1/\lambda}(k)|^2\, dx\, dk, \tag{18}$$

for all $f \in C_0(\Gamma_n)$. The coordinate change $\xi = \lambda x, \eta = k/\lambda$ yields

$$\mu_{\varphi_\lambda}(f) = \int_{\Gamma_n} f(\frac{\xi}{\lambda}, \lambda\, \eta)\, |\varphi(\xi)|^2\, |\hat{\varphi}(\eta)|^2\, d\xi\, d\eta = \int_{\Gamma_n} f(\frac{\xi}{\lambda}, \lambda\, \eta) d\mu_\varphi(\xi, \eta). \tag{19}$$

2 Maximum Probability of Compact Sets

Definition 1. *Let Λ be a closed subset of Γ_n, then we define*

$$e(\Lambda) = \sup \left\{ \int_\Lambda d\mu_\varphi : \varphi \in L^2(\mathbb{R}^n), ||\varphi|| = 1 \right\}. \tag{20}$$

Lemma 1. *Let A, B be subsets of $\mathbb{R}^n$ having finite Lebesgue measure, that is $|A| + |B| < \infty$, then exists a ψ such that*

$$\nu_\psi(A) + \nu_{\hat{\psi}}(B) = 0. \tag{21}$$

---

[5]Clearly, the constants may depend on the dimension d, although we do not explicitly outline this point by notation.

Proof. This follows by Corollary 2.5.A in [9]. Actually it is shown that there always is a $\varphi \in L^2(\mathbb{R}^n)$ such that for any given pair g, h of functions in $L^2(\mathbb{R}^n)$ the restriction of φ to A and that of $\hat{\varphi}$ to B coincides with the restriction of g to A and h to B respectively. $\qquad\square$

Proposition 2. *Let each of A, B be the complement of a bounded open subset in $\mathbb{R}^n$, then*

$$e(A \times B) = 1.$$

Proof. Set $U = \mathbb{R}^n \backslash A$, $V = \mathbb{R}^n \backslash B$, then for each φ we have $\mu_\varphi(A \times B) = \nu_\varphi(A)\,\nu_{\hat{\varphi}}(B) = (1 - \nu_\varphi(U))\,(1 - \nu_{\hat{\varphi}}(V))$. By the lemma above we may choose a ψ such that $\nu_\psi(A) = \nu_{\hat{\psi}}(B) = 0$, thus $\mu_\psi(A \times B) = 1$. $\qquad\square$

Proposition 3. *Let A, B be subsets of $\mathbb{R}^n$ such that $|A| + |B| < \infty$, then for every normalized $\varphi \in L^2(\mathbb{R}^n)$*

$$\mu_\varphi(\chi_{A \times B}) \leq \left(1 - \frac{\alpha}{2} e^{-\beta\,\eta(A,B)}\right)^2 \tag{22}$$

with constants α, β and η as in (17).

Proof. Using (17) we get $2 - (\nu_\varphi(A) + \nu_{\hat{\varphi}}(B)) \geq \alpha e^{-\beta\,\eta(A,B)}$. Dividing both sides by two and applying the arithmetic-geometric mean inequality yields $\sqrt{\nu_\varphi(A)\nu_{\hat{\varphi}}(B)} \leq 1 - \frac{\alpha}{2} e^{-\beta\,\eta(A,B)}$, which implies (22). $\qquad\square$

Corollary 1. *Let $\Lambda \subset \Gamma_n$ be compact, then*

$$e(\Lambda) \leq \left(1 - \frac{\alpha}{2} e^{-\beta\,\eta(\pi_1(\Lambda),\pi_2(\Lambda))}\right)^2, \tag{23}$$

where $\pi_1, \pi_2 : \Gamma_n \to \mathbb{R}^n$ are the standard projections and the constants α, β, η are as in (17).

Proof. The images of the projections π_1, π_2 are again compact, thus measurable and of finite Lebesgue measure. $\qquad\square$

If we replace compactness by finite Lebesgue measure or closed only, then we have to deal with analytic sets. Since

$$e(\Lambda) = \sup_\varphi \mu_\varphi(\Lambda) = 1 - \inf_\varphi \mu_\varphi(\Gamma_n \backslash \Lambda)$$

we have the relation to $E_{\mathcal{H}}$ with $\mathcal{H} = \chi_{\Gamma_n \backslash \Lambda}$. Since $U = \Gamma_n \backslash \Lambda$ is open we can also apply (12).

2.1 Optimal Bounds

Optimal bounds are (at the time of writing, 2015) not known. To illustrate the difficulties one encounters when trying to find maximizers of $e(\Lambda)$, let us consider the case where $\Lambda = \{x^2 + k^2 \leq R^2\} \subset \mathbb{R}^2$. Using the first Hermite function $\psi_0(x) = \pi^{-\frac{1}{4}} e^{-\frac{1}{2}x^2}$ as a trial function, we obtain with $\psi_0 = \hat{\psi}_0$ in mind:

$$e(\Lambda) \geq \frac{1}{\pi} \int_{\{x^2+k^2 \leq R^2\}} e^{-(x^2+k^2)} \, dx \, dk = 1 - e^{-R^2}. \tag{24}$$

Nazarov's inequality (17) for $d = 1$, with $A = [-\frac{a}{2}, \frac{a}{2}]$, $B = [-\frac{b}{2}, \frac{b}{2}]$, reads with ψ_0:

$$\frac{1}{\sqrt{\pi}} \int_{\mathbb{R}\setminus A} e^{-x^2} \, dx + \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}\setminus B} e^{-k^2} \, dk = \mathrm{erfc}(\frac{a}{2}) + \mathrm{erfc}(\frac{b}{2}) \geq \alpha \, e^{-\beta \, ab},$$

where

$$\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$$

$$= \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} \, dt$$

denotes the *complementary error function*. Now, when we set $a = b = 2\,R$, we get

$$\mathrm{erfc}(R) \geq \frac{\alpha}{2} \, e^{-4\beta R^2} \tag{25}$$

Indeed, there exist quite optimal Chernoff-type bounds ([5], Theorem 2) for erfc:

$$\mathrm{erfc}(R) \geq \rho \, e^{-\sigma \, R^2} \tag{26}$$

if $\sigma > 1$ and $0 < \rho \leq \sqrt{\frac{2\,e}{\pi}} \frac{\sqrt{\sigma-1}}{\sigma}$. On the other hand, there is also an upper bound ([5], Theorem 1) of the same kind:

$$\mathrm{erfc}(R) \leq \kappa \, e^{-\lambda \, R^2} \tag{27}$$

provided that $\kappa \geq 1$ and $0 < \lambda \leq 1$, more precisely if and only if κ, λ satisfy these relations. It is actually believed that a Gaussian function with A, B balls of radius R is optimal (see introduction in [10]), however, there is no proof yet. It seems to be even more difficult to prove optimality if Λ is not a product, e.g. as in (24). The following lines might illustrate this.

For $Q = [-R, R] \times [-R, R]$ we get

$$e(Q) \geq \mathrm{erf}(R)^2 = (1 - \mathrm{erfc}(R))^2,$$

which is in accordance with (22) and the bounds of `erfc` discussed above. Since $\Lambda \subset Q$ we also have that

$$e(Q) \geq e(\Lambda) \geq 1 - e^{-R^2},$$

so that by Corollary 1 the upper bound for $e(\Lambda)$ is the same as that for $e(Q)$, what is certainly not optimal. Referring to the remark at the end of Section 1.4, one might conjecture that ψ_0 is a minimizer of $E_{\chi_{\mathbb{R}^2 \setminus \Lambda}}(\varphi)$, i.e.

$$E_{\chi_{\mathbb{R}^2 \setminus \Lambda}}(\psi_0) = 1 - e(\Lambda) = e^{-R^2},$$

however, it seems that ψ_0 is not a solution of the Euler equation (9).

2.2   A Theorem of A. Steiner

Uncertainty inequalities occur in various forms, very often disguised as a localization principle. There is a nice theorem by Antonio Steiner [16], that is probably not widely known (as the article is in German), so we cite it here using our notation:

Theorem 1 (Theorem 3, [16]). *Let $A, B \subset \mathbb{R}$ be two measurable sets, such that $|A| + |B| < \infty$. Suppose $\nu_\varphi(A) + \nu_{\hat\varphi}(B) > 1$, then*

$$|A||B| \geq \max\{c_1, c_2\}, \tag{28}$$

where

$$c_1 = \frac{2\pi}{\nu_\varphi(A)} \left(\sqrt{\nu_{\hat\varphi}(B)} - \sqrt{1 - \nu_\varphi(A)} \right)^2$$

and

$$c_2 = \frac{2\pi}{\hat\nu_\varphi(B)} \left(\sqrt{\nu_\varphi(A)} - \sqrt{1 - \nu_{\hat\varphi}(B)} \right)^2.$$

Recall that $\varphi \in L^2(\mathbb{R})$ with $||\varphi||_2 = 1$, and

$$\nu_\varphi(A) = \nu_\varphi(\chi_A) = \int_A |\varphi(x)|^2 \, dx.$$

We will sketch the proof here because it is elementary, but the idea is clever. Writing the Fourier transform of φ as

$$\hat\varphi(k) = \frac{1}{\sqrt{2\pi}} \int_A \varphi(x) \, e^{-ikx} \, dx + \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R} \setminus A} \varphi(x) \, e^{-ikx} \, dx = \hat\varphi_A + \hat\varphi_{A^c},$$

we may proceed with

$$\nu_{\hat{\varphi}}(B) = \int_B |\hat{\varphi}_A + \hat{\varphi}_{A^c}|^2 \, dk$$

$$\leq \int_B (|\hat{\varphi}_A| + |\hat{\varphi}_{A^c}|)^2 \, dk$$

$$\leq \ldots 2 \times \text{ Schwarz inequality}$$

$$\leq \nu_\varphi(A)\frac{|A||B|}{2\pi} + (1 - \nu_\varphi(A)) + 2\left(\nu_\varphi(A)\frac{|A||B|}{2\pi}\right)^{\frac{1}{2}}(1 - \nu_\varphi(A))^{\frac{1}{2}}$$

Setting $\xi = \sqrt{\frac{|A||B|}{2\pi}}, \alpha = \nu_\varphi(A)$, and $\beta = \nu_{\hat{\varphi}}(B)$ yields

$$\alpha\xi^2 + 2\sqrt{\alpha}\sqrt{1-\alpha}\,\xi + (\alpha + \beta - 1) \geq 0,$$

implying

$$\xi \geq \frac{\sqrt{\beta} - \sqrt{1-\alpha}}{\sqrt{\alpha}},$$

that is

$$|A||B| \geq c_1.$$

Reversing the roles gives c_2. The proof reveals why the condition $\alpha + \beta > 1$ is necessary. Not only because it guarantees $\alpha > 0 \wedge \beta > 0$, also because $\xi \geq 0$ is required. Finally, we want to point out the link to (15).

2.3 Tightness

Since we know that $K_{\mathcal{H}}(\varphi)$ is attained by a γ_φ, a minimizing sequence γ_{φ_j} of

$$\inf_\varphi K_{\mathcal{H}}(\varphi)$$

does not necessarily converge to a measure, and even if it does, it is not for sure that it is in $\Gamma(\varphi)$ for some $\varphi \in L^2(\mathbb{R}^n)$. By Prokhorov's theorem, however, it is sufficient to show the tightness of the sequence $\{\gamma_{\varphi_j}\}_{j\geq 1}$ in order to get a weakly convergent subsequence, that is, $\forall \epsilon > 0, \exists K_\epsilon \subset \Gamma_n$, such that

$$\gamma_{\varphi_j}(K_\epsilon) \geq 1 - \epsilon$$

holds $\forall j \in \mathbb{N}$. This is usually not trivial, but we have

$$e(A_\epsilon \times B_\epsilon) \geq \gamma_{\varphi_j}(A_\epsilon \times B_\epsilon) \geq \nu_{\varphi_j}(A_\epsilon)\,\nu_{\hat{\varphi}_j}(B_\epsilon),$$

14

that is, in view of the weak compactness of the unit ball in L^2, the problem may be often reduced to merely consider the marginal measures. If the function $\mathcal{H}$ is *inf-compact*, that is if the level sets $\{(x, k) : \mathcal{H}(x, k) \leq r\}$ are compact for all $r \in \mathbb{R}$, then there are (usually) standard procedures to verify tightness. For instance, if we take the additional assumption

$$\int_{\Gamma_n} \mathcal{H}(x, k) d\gamma_{\varphi_j}(x, k) \leq C,$$

for all $j \in \mathbb{N}$, and set $K_{1/m} = \{(x, k) : \mathcal{H}(x, k) \leq m\}$, then for all $\gamma \in \{\gamma_{\varphi_j}\}_{j \geq 1}$

$$m \, \gamma(\Gamma_n \backslash K_{1/m}) \leq \int_{\Gamma_n \backslash K_{1/m}} \mathcal{H}(x, k) d\gamma(x, k) \leq C,$$

thus

$$\gamma(K_{1/m}) \geq 1 - \frac{C}{m}.$$

Note that inf-compact functions are lower semi-continuous, thus we conclude

$$\liminf_{j \to \infty} \int \mathcal{H} d\gamma_{\varphi_{m_j}} \geq \int \mathcal{H} d\gamma_\star,$$

where $\gamma_\star$ is the weak limit of the subsequence.

3   Miscellaneous

3.1   Approximation

Most problems in non-relativistic quantum mechanics are based on Hamiltonian functions which cleave into a kinetic and a potential part, or at least, they may be transformed into such a form. There are some well known exceptions, of course, as soon as we consider terms like e.g. $\sqrt{(k - X(x))^2 + \sigma(x)}$. Where Pauli's equation still fits into the scheme, we leave the foundational frame if $\mathcal{H}$ does not separate into $\mathcal{T} + \mathcal{V}$ and/or $T \cdot V$. From a mathematical point of view this will not be a matter of concern. However, the physical interpretation of the limit measures ($\gamma_\star$) of $K_{\mathcal{H}}$ in these cases is not quite clear (at least not to me). Nevertheless, the duality formula (11) suggests the following considerations.

For $\mathcal{T}, \mathcal{V} \in C_b(\mathbb{R}^n)$ let

$$\varepsilon(\mathcal{T}, \mathcal{V}) = \inf\{\nu_\varphi(T) + \nu_{\hat{\varphi}}(V) : \varphi \in L^2(\mathbb{R}^n)\}, \tag{29}$$

and

$$\delta_{\mathcal{H}}(\varphi) = \sup_{\mathcal{T} + \mathcal{V} \leq \mathcal{H}} (\nu_\varphi(T) + \nu_{\hat{\varphi}}(V)), \tag{30}$$

15

then we have for any admissible $\mathcal{H}$:

$$\lambda_0(\mathcal{H}) \geq \inf_{L^2(\mathbb{R}^n)} K_{\mathcal{H}}(\varphi) = \inf_{L^2(\mathbb{R}^n)} \delta_{\mathcal{H}}(\varphi) \geq \sup_{\mathcal{T}+\mathcal{V}\leq\mathcal{H}} \varepsilon(\mathcal{T},\mathcal{V}), \qquad (31)$$

where the equality sign is by (11) and the last inequality is a consequence of the *max-min inequality*, that is $\sup_X \inf_Y f(x,y) \leq \inf_Y \sup_X f(x,y)$, valid for arbitrary sets X, Y. It is well known that the latter inequality may be strict, but there are cases where equality may be proved, e.g. showing the existence of a saddle point by methods as described in [7]. More will be published elsewhere.

3.2 Special case: $\mathcal{H}(x,k) = |k|^2 \, |X(x)|^2$

In OTT much is known about quadratic costs like $|x-y|^2$. Generally, powers of a distance[6] function $d(x,y)$ play an important role, for obvious reasons when considering actual transport of goods. We, however, want to consider the *cost function* $\mathcal{H}(x,k) = |k|^2 \, |X(x)|^2$ because we know the minimizers of $E_{\mathcal{H}}$ and because it serves as a simple model where $\gamma(\mathcal{H})$ is not constant on $\Gamma(\varphi)$. The vector field $X : \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be sufficiently smooth (say C^1) for simplicity, then $\mathcal{H}$ is certainly l.s.c and bounded below (by zero). Thus

$$K_{\mathcal{H}}(\varphi) = \inf_{\gamma\in\Gamma(\varphi)} \int_{\mathbb{R}^n\times\mathbb{R}^n} |k|^2 \, |X(x)|^2 d\gamma(x,k) \geq 0$$

is attained for some $\gamma_\varphi \in \Gamma(\varphi)$. For $E_{\mathcal{H}}(\varphi)$ we get

$$E_{\mathcal{H}}(\varphi) = \int_{\mathbb{R}^n\times\mathbb{R}^n} |k|^2 \, |X(x)|^2 d\mu_\varphi(x,k) = \int_{\mathbb{R}^n} |k|^2 \, d\nu_{\hat{\varphi}}(k) \int_{\mathbb{R}^n} |X(x)|^2 \, d\nu_\varphi(x).$$

Now, $|k|^2 \in L^1(\nu_{\hat{\varphi}})$, i.e. $\nu_{\hat{\varphi}}$ certainly has finite second order moments if $\varphi \in H^1(\mathbb{R}^n)$, otherwise the integral may be infinite. The inequality (45) in Lemma 2 (Appendix), tells us (with $f(t) = \frac{1}{2}t^2$):

$$E_{\mathcal{H}}(\varphi) \geq \frac{1}{4} \left(\int_{\mathbb{R}^n} \operatorname{div}(X) \, d\nu_\varphi \right)^2,$$

where we assume $|X| \in L^2(\nu_\varphi)$, and $\operatorname{div}(X) \in L^1(\nu_\varphi)$. The equality sign holds if

$$\nabla\varphi(x) + \varphi(x) \, X(x) = 0.$$

For simplicity we proceed with the special case $X(x) = x$, that leads to

$$E_{\mathcal{H}}(\varphi) \geq \frac{n^2}{4}$$

---

[6]OTT is well defined on Polish spaces.

16

with equality for $\varphi_0(x) = C \exp(-\frac{x^2}{2})$. Therefore, we get for (8) the value $\lambda_0 = \frac{n^2}{4}$. Verifying (9) actually shows that φ_0 is a critical point of $E_{\mathcal{H}}(\varphi)$, with $X(x) = x$. Consequently, λ_0 is an upper bound to inf $K_{\mathcal{H}}$ when recalling (31). Finding lower bounds is much trickier because approximating $|x|^2 |k|^2$ by functions $f(x) + g(k)$ is much more unpleasant than controlling the function $\langle x, k \rangle$ for example. While the latter leads to the usual convex conjugates, the former requires e.g. the notion of *c-concavity*, that is in our special case,

$$f(x) = \inf_{k \in \mathbb{R}^n} \left(|x|^2 |k|^2 - g(k) \right),$$

and *cyclic monotonicity* (see [18]). Actually, there is not much known about the measures γ_φ nor about optimal bounds for $K_{\mathcal{H}}$. At least, however, we know that

$$K_{\mathcal{H}}(\varphi_0) = \max_{\psi \in L^1(\nu_{\varphi_0})} \int_{\mathbb{R}^n} [\psi^c(x) - \psi(x)] |\varphi_0(x)|^2 \, dx \le \frac{n^2}{4},$$

that is, the dual problem is attained as well (see Theorem 5.10 in [17]). Therefore, μ_{φ_0} is not optimal anymore. This was to be expected because of the *interaction* of x and k in $\mathcal{H}$. The measure γ_{φ_0} now has to be supported on some subset (actually a closed c-cyclically monotone) such that $\mathcal{H} \approx_\gamma (\psi_0^c(k) - \psi_0(x))$.

3.3 Case $n = 1$

We will see that even the one-dimensional case (i.e. phase space $\mathbb{R}^2$) may be quite involved. First, let us consider the distribution functions F, G of the measures ν_φ and $\nu_{\hat{\varphi}}$ respectively:

$$F(x) = \nu_\varphi((-\infty, x]) = \int_{-\infty}^x |\varphi(x)|^2 \, dx,$$

and analogous:

$$G(k) = \nu_{\hat{\varphi}}((-\infty, k]) = \int_{-\infty}^k |\hat{\varphi}(k)|^2 \, dk.$$

There is a non-decreasing mapping $T : \mathbb{R} \to \mathbb{R}$ such that

$$F(x) = G(T(x)), \forall x \in \mathbb{R}, \tag{32}$$

namely the *increasing rearrangement* $T(x) = (G^{-1} \circ F)(x)$, with $G^{-1}(s) = \inf\{t \in \mathbb{R} : G(t) > s\}$. Consequently, we get an ordinary differential equation in $L^1(\mathbb{R})$ when differentiating both sides:

$$|\varphi(x)|^2 = |\hat{\varphi}(T(x))|^2 \frac{dT}{dx}(x). \tag{33}$$

Of course, T is not necessarily differentiable on $\mathbb{R}$, but we have Lebesgue's theorem for the differentiability of monotone functions and absolute continuity of the distribution functions at hand, though regularity questions are not in focus here. Note that (33) coincides with (5) for $n = 1$.

If $T(x) = x$, then $|\varphi(x)|^2 = |\hat{\varphi}(x)|^2$, i.e. this a solution for any Hermite function $m = 0, 1, 2, \ldots$

$$\psi_m(x) = \frac{H_m(x)}{\sqrt{2^m\, m!\, \pi^{\frac{1}{2}}}} e^{-\frac{1}{2}\,x^2},\tag{34}$$

because $\hat{\psi}_m(k) = (-i)^m\, \psi_m(k)$. We conclude

$$\int_{\mathbb{R}^2} \mathcal{H}(x, k)d\gamma(x, k) = \int_{\mathbb{R}} \mathcal{H}(x, x)|\psi_m(x)|^2\, dx,$$

for $\gamma = (Id \times T)_{\#}\nu_{\psi_m}$. On the other hand - note the difference,

$$\int_{\mathbb{R}^2} \mathcal{H}(x, k)d\mu_{\psi_m}(x, k) = \int_{\mathbb{R}^2} \mathcal{H}(x, k)|\psi_m(x)|^2\, |\psi_m(k)|^2\, dx\, dk.$$

For $\mathcal{H}(x, k) = |x|^2\, |k|^2$ the integral in the first case evaluates for $m = 0$ to $\frac{3}{4}$ and in the second case to $\frac{1}{4}$ (this is optimal). If $\mathcal{H}$ is additive $(a(k) + b(x))$, however, then both integrals coincide.

Generally, we can solve

$$g'(x) = \frac{|\psi(x)|^2}{|\hat{\psi}(g(x))|^2}\tag{35}$$

for various known ground states ψ of the Schrödinger equation, then we have

$$E_0 = E_{\mathcal{H}}(\psi) = \int_{\mathbb{R}} \left(g(x)^2 + \mathcal{V}(x)\right)\, |\psi(x)|^2 dx.$$

The convex function $S(x)$ whose existence was claimed in (A.2) is determined by

$$\frac{dS}{dx}(x) = g(x),$$

but there seems to be no connection with the action S in the Hamilton-Jacobi equation nor with $p(x) = -\frac{\psi'(x)}{\psi(x)}$, as was to be expected when recollecting that $\nu_\psi \otimes \nu_{\hat{\psi}}$ is optimal. Regarding $p(x)$ (which is a vector field if $n > 1$), consider the differential equation

$$\varphi'(x) + p(x)\, \varphi(x) = 0$$

and assume $\varphi \in C^2(\mathbb{R})$, then

$$\varphi''(x) + (p'(x) - p(x)^2)\,\varphi(x) = 0,$$

so that when setting

$$p'(x) = p(x)^2 - \mathcal{V}(x) + \lambda,$$

we can generate potentials V and ground states of the form

$$\varphi_p(x) = C\,e^{-\int^x p(s)\,ds},$$

which satisfy: $-\varphi''(x) + \mathcal{V}(x)\,\varphi(x) = \lambda\,\varphi(x)$. Take $p(x) = \sinh(x)$ for instance, yielding $\mathcal{V}(x) = \cosh(x)\,[\cosh(x) - 1]$, $\lambda = 1$, and

$$\varphi_0(x) = C\,e^{-\cosh(x)}.$$

The computation of the Fourier transform is often manageable ([4]), so that one can find $g(x)$ and check against $p(x)$. As mentioned earlier, it seems that unless $p(x) = x$, p is not a transport map, i.e. not admissible. Nevertheless, it holds true that

$$\int_{\mathbb{R}} p(x)^2\,|\varphi_0(x)|^2\,dx = \int_{\mathbb{R}} g(x)^2\,|\varphi_0(x)|^2\,dx = \int_{\mathbb{R}} |\varphi_0(x)'|^2\,dx.$$

3.4 Discrete case

Concluding these expositions, we want to point out that the discrete case of the transportation problem leads to bi-stochastic matrices $m \in \mathcal{B}_n$, so that the Kantorovich problem reads (assuming all points have the same *mass*)

$$\inf\left\{\frac{1}{n}\sum_{i,j}\mathcal{H}(x_i, k_j)\,m_{ij} : m \in \mathcal{B}_n\right\},$$

which is a linear minimization problem on the bounded convex subset $\mathcal{B}_n$ of all real $n \times n$ matrices. The analogy to

$$\int_{\Gamma_n}\mathcal{H}(x, k)\,d\mu_\varphi,$$

however, is

$$\sum_{i,j}\mathcal{H}(x_i, k_j)|\varphi(x_i)|^2\,|\hat{\varphi}(k_j)|^2\Delta x\,\Delta k, \tag{36}$$

where the (normalized) complex vectors

$$(\varphi(x_1), \ldots, \varphi(x_n)) \text{ and } (\hat{\varphi}(k_1), \ldots, \hat{\varphi}(k_n))$$

19

are related by the *discrete (unitary) Fourier* transformation. This leads to interesting problems that are slightly harder but are also relatively easy accessible from the numerical viewpoint due to the *Fast Fourier Transform* algorithm. The standard Hamiltonian $\mathcal{H}(x,k) = \frac{\hbar^2}{2m}k^2 + \mathcal{V}(x)$ even admits a closed form for the sum in (36), as is outlined in the two papers [3] and [13]:

$$\mathcal{H}^0_{ij} = \begin{cases} \frac{\hbar^2}{4mL^2}\left[\frac{N^2+2}{6}\right] + V_i & \text{if } i = j \\ (-1)^{(i-j)}\frac{\hbar^2}{4mL^2}\left[\sin(\pi\frac{(i-j)}{N})\right]^{-2} & \text{if } i \neq j \end{cases}, \tag{37}$$

where $\mathcal{H}^0_{ij} = \mathcal{H}_{ij}\Delta x$, $L = N\Delta x$ and $\Delta k = \frac{2\pi}{L}$. This provides a fast method to compute eigenvalues and eigenvectors, just computing the potential at the grid points (V_i) then diagonalizing. There is an implementation in `FORTRAN77` by the authors

```
FGHEVEN
C       F. Gogtas, G.G. Balint-Kurti and C.C. Marston,
C       QCPE Program No. 647, (1993).
```

http://www.chm.bris.ac.uk/pt/dixon/dynamics/fghqcpe.for
where it is also stated:

```
C  The analytical Hamiltonian expression given in the reference
C  contains a small error.  The formula should read:
C
C    H(i,j) = {(h**2)/(4*m*(L**2)} *
C             [(N-1)(N-2)/6 + (N/2)] + V(Xi),   if i=j
C
C    H(i,j) = {[(-1)**(i-j)] / m } *
C             { h/[2*L*sin(pi*(i-j)/N)]}**2 ,    if i#j
```

So, formula (37) should be correct, we only simplified the expressions between the square brackets.

Generally, we can easily translate to discrete spaces as follows: let $\varphi \in \mathbb{C}^N$ be a unit vector, and denote by $\hat{\varphi} \in \mathbb{C}^N$ the unitary DFT: $\hat{\varphi} = \frac{1}{\sqrt{N}}\mathbf{F}\varphi$, where $\mathbf{F}_{jk} = \omega_N^{(j-1)k}$ is the Vandermonde matrix, and $\omega_N = e^{-2\pi i/N}$ (a N-th root of unity). Then the measures corresponding to ν_φ and $\nu_{\hat{\varphi}}$ are the linear functionals (we identify the dual space with $\mathbb{C}^N$ here)

$$\nu_\varphi = \sum_{i=1}^N |\varphi_i|^2 \delta_i, \quad \nu_{\hat{\varphi}} = \sum_{i=1}^N |\hat{\varphi}_i|^2 \delta_i, \tag{38}$$

20

where $\delta_i(f) = f_i$ denotes the discrete Dirac measure (equivalently, the dual base). Then $\Gamma(\varphi)$ corresponds to the matrices

$$\left\{ \gamma \in \mathbb{M}_{\mathbb{R}_+,0}^{N \times N} : \sum_{i=1}^{N} (\gamma f)_i = \nu_\varphi(f), \ \sum_{j=1}^{N} (\gamma^T f)_j = \nu_{\hat{\varphi}}(f), \ \forall f \in \mathbb{C}^N \right\},$$

which has at least the member

$$\gamma_0 = \nu_\varphi \otimes \nu_{\hat{\varphi}} = \sum_{i,j=1}^{N} |\varphi_i|^2 \, |\hat{\varphi}_j|^2 \, \delta_i \otimes \delta_j.$$

The optimal transport problem between ν_φ and $\nu_{\hat{\varphi}}$ (roughly) reduces therefore to the *linear program*

$$\inf_{\gamma \in \Gamma(\varphi)} \mathrm{Tr}(\mathcal{H}\,\gamma^T) \tag{39}$$

which has a solution, of course. Note that the (real) Frobenius scalar product $\langle \mathcal{H}, \gamma \rangle_F$ which is often used in the literature is equal to the trace of $\mathcal{H}\gamma^T$. Usually, an optimal γ is a sparse matrix where at most $2N-1$ entries are different from zero, therefore, γ_0 generally is far from optimal. See the appendix for more details.

A Examples and Remarks

We cite here a version formulated by R. McCann [14]:

Theorem 2 (McCann). *Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$ and suppose μ vanishes on (Borel) subsets of $\mathbb{R}^n$ having Hausdorff dimension $n-1$. Then there exists a convex function ψ on $\mathbb{R}^n$ whose gradient $\nabla\psi$ pushes μ forward to ν. Although ψ may not be unique, the map $\nabla\psi$ is uniquely determined μ-almost everywhere.*

When we apply this result to the measures defined, we get the relation

$$\int_{\mathbb{R}^n} f(\nabla\psi(x))|\varphi(x)|^2\, dx = \int_{\mathbb{R}^n} f(k)|\hat{\varphi}(k)|^2\, dk, \tag{40}$$

where ψ is the convex function whose existence was asserted. If the measures μ, ν in the theorem are given by densities ρ_0, ρ_1 w.r.t. Lebesgue measure then

$$\int_{\mathbb{R}^n} f(\nabla\psi(x))\rho_0(x)dx = \int_{\mathbb{R}^n} f(y)\rho_1(y)dy$$

implies (assuming sufficient regularity, see [14] and references therein) that ψ is a solution to a *Monge-Ampère* equation:

$$\rho_0(x) = \rho_1(\nabla\psi(x)) \det D^2\psi(x). \tag{41}$$

for instance,

$$|\varphi(x)|^2 = |\hat{\varphi}(\nabla\psi(x))|^2 \det D^2\psi(x), \tag{42}$$

which represents a quite remarkable identity. For, if φ has compact support in $\mathbb{R}^n$ then the Fourier transform $\hat{\varphi}$ is an analytic function such that

$$|\hat{\varphi}(k)|^2 \leq \frac{|\mathtt{supp}(\varphi)|}{(2\pi)^n}$$

uniformly on $\mathbb{R}^n$, yielding the bound

$$\det D^2\psi(x) \geq \frac{(2\pi)^n}{|\mathtt{supp}(\varphi)|}\, |\varphi(x)|^2$$

On the other hand we recognize that outside $\mathtt{supp}(\varphi)$ either $\det D^2\psi(x) = 0$ or $\nabla\psi$ must map into zeroes of $\hat{\varphi}$. Of course, all these derivations need careful regularity analysis and indeed most relations hold usually in a weak sense only. If we, however, assume that $\psi \in C^2(\mathbb{R}^n)$ (for simplicity), then the inequality

$$\frac{1}{n}\Delta\psi(x) \geq (\det D^2\psi(x))^{\frac{1}{n}} \tag{43}$$

is an easy consequence of the convexity of ψ and the well known inequality $x_1 + \ldots + x_n \geq n \sqrt[n]{x_1 \cdot \ldots \cdot x_n}$, valid for all non-negative real numbers. Thus we obtain a bound in terms of the Laplacian of ψ:

$$\Delta\psi(x) \geq \frac{2\pi n}{|\mathtt{supp}(\varphi)|^{\frac{1}{n}}} |\varphi(x)|^{\frac{2}{n}} \tag{44}$$

We will use the following simple lemma in the sequel:

Lemma 2. *Let $f \in W^{1,\infty}(\mathbb{R}^n)$, $f(0) = 0$ and $X \in L^2_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{R}^n)$ such that $\mathtt{div}(X)$ exists and is in $L^\infty_{\mathrm{loc}}(\mathbb{R}^n)$, then for all $\phi \in C^1_0(\mathbb{R}^n)$:*

$$\int_{\mathbb{R}^n} |\nabla\phi(x)|^2 \, dx \int_{\mathbb{R}^n} |f'(\phi)|^2 |X(x)|^2 \, dx \geq \left(\int_{\mathbb{R}^n} f(\phi) \, \mathtt{div}(X) \, dx \right)^2 \tag{45}$$

with equality iff $\nabla\phi(x) + f'(\phi(x)) \, X(x) = 0$ almost everywhere. In that case also $\|\nabla\phi\|^2 = \|f'(\phi)X\|^2 = |\int f(\phi) \, \mathtt{div}(X) \, dx|$.

We only sketch the easy proof: expanding $\|\nabla\phi + f'(\phi) \, X\|^2 \geq 0$, yields $\|\nabla\phi\|^2 + \|f'(\phi)X\|^2 + \langle \nabla f(\phi), X \rangle_{L^2} \geq 0$. The last term may be rewritten to $\int \mathtt{div}(f(\phi) \, X) \, dx - \int f(\phi) \mathtt{div}(X) \, dx$. The first term vanishes by the condition $f(0) = 0$ so that (45) follows by optimizing the quadratic inequality w.r.t. X. It is also seen during the proof that the inequality holds under much weaker conditions and that the boundary term may be incorporated if not zero. Replacing $\nabla\phi$ by $A\nabla\phi$, where A is matrix function, also allows to consider degenerate cases (if A is indefinite). In fact, this is merely a disguised form of Schwarz's inequality.

To illustrate the usefulness of (43), let $X = \nabla\psi$ (the Brenier map to $\phi = \varphi$ from above), then using (43), the right hand side of (45) becomes:

$$\left(\int_{\mathbb{R}^n} f(\phi) \, \Delta\psi(x) \, dx \right)^2 \geq \frac{(2\pi n)^2}{|\Omega|^{\frac{2}{n}}} \left(\int_{\Omega} f(\phi) |\phi(x)|^{\frac{2}{n}} \, dx, \right)^2 \tag{46}$$

where we have assumed $f \geq 0$ and used the abbreviation $\Omega = \mathtt{supp}(\phi)$. As a simple application let us consider the first Dirichlet eigenfunction ϕ_0 with eigenvalue $\lambda_0(\Omega)$. Letting $f(t) = \frac{1}{2}t^2$ in $(45), (46)$, and $X = \nabla\psi$ the corresponding Brenier map, we get

$$\int_{\Omega} |\nabla\phi_0(x)|^2 \, dx \int_{\Omega} |\phi_0|^2 \, |\nabla\psi(x)|^2 \, dx \geq \frac{(2\pi n)^2}{4|\Omega|^{\frac{2}{n}}} \left(\int_{\Omega} |\phi_0(x)|^{2+\frac{2}{n}} \, dx \right)^2 \tag{47}$$

By definition of λ_0 and ψ we have

$$\lambda_0(\Omega) = \int_{\Omega} |\nabla\phi_0(x)|^2 \, dx = \int_{\Omega} |\phi_0|^2 \, |\nabla\psi(x)|^2 \, dx,$$

so that (47) gives

$$\lambda_0(\Omega) \geq \frac{n\pi}{|\Omega|^{\frac{1}{n}}} \, \|\phi_0\|_{2+2/n}^{2+2/n}.$$

Finally, recalling that generally $|\Omega|^{1-q/p} \, \|u\|_p^q \leq \|u\|_q^q$ holds for $p \leq q$, we get with $q = 2 + 2/n$ and $p = 2$ and because $\|\phi_0\|=1$:

$$\lambda_0(\Omega) \geq \frac{n\pi}{|\Omega|^{\frac{2}{n}}}. \tag{48}$$

This is apparently far from the optimal result stated by the well known *Rayleigh-Faber-Krahn* inequality which gives the optimal value (B_Ω =ball with volume $|\Omega|$):

$$\lambda_0(\Omega) \geq \lambda_0(B_\Omega) = \left(\frac{\omega_n}{|\Omega|}\right)^{\frac{2}{n}} j_{\frac{n}{2}-1,1}^2.$$

Indeed, the first zero of the Bessel function J_m behaves like $j_{d,1} \sim d + O(d^{\frac{1}{3}})$ for large d, so (48) not even shows the correct asymptotics. Let $Q_{n,L}$ be the cube with side length L in $\mathbb{R}^n$, then $\lambda_0(Q_{n,L}) = n\frac{\pi^2}{L^2}$ and

$$Q_{n,2/\sqrt{n}} \subset B_1 \subset Q_{n,2}$$

so that by the domain monotonicity property of the eigenvalues

$$n^2 \frac{\pi^2}{4} \geq \lambda_0(B_1) \geq n \frac{\pi^2}{4}$$

holds. Of course, the estimate $|\hat{\varphi}(k)|^2 \leq \frac{|\Omega|}{2\pi}$ seems rather crude, yet one would expect a sharper bound for large orders in view of the fact that the same (at least similar) method provides a proof of the isoperimetric inequality.

B Discretization and Numerics (UCDFT)

Given a vector $X = (X_0, \ldots, X_{N-1}) \in \mathbb{C}^n$, the *discrete Fourier transform* of X is defined as

$$\mathrm{DFT}(X)_k = \sum_{j=0}^{N-1} X_j e^{-2\pi i \frac{jk}{N}}$$

for $k = 0, \ldots, N-1$. The *inverse discrete Fourier transform* is defined similarly:

$$\mathrm{IDFT}(X)_j = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{2\pi i \frac{kj}{N}},$$

where $j = 0, \ldots, N-1$.

Proposition 4. $\text{IDFT} \circ \text{DFT} = \text{Id}$

Proof.

$$\text{IDFT}\left(\text{DFT}(X)\right)_j = \frac{1}{N} \sum_{k=0}^{N-1} \left(\sum_{l=0}^{N-1} X_l e^{-2\pi i \frac{lk}{N}} \right) e^{2\pi i \frac{kj}{N}}$$

$$= \frac{1}{N} \sum_{l=0}^{N-1} X_l \sum_{k=0}^{N-1} e^{2\pi i \frac{k(j-l)}{N}} = \frac{1}{N} \sum_{l=0}^{N-1} X_l N \delta_{jl} = X_j.$$

$\square$

Remark 1. *By symmetry we also have* $\text{DFT} \circ \text{IDFT} = \text{Id}$. *Note that **DFT** is not unitary because of the asymmetry of the factor* $\frac{1}{N}$ *in **IDFT**. This could be fixed by using* $\frac{1}{\sqrt{N}}$ *in both transformations, however, one has to be careful when using numeric packages:*

$$\hat{X} := \frac{1}{\sqrt{N}}\text{DFT}(X), \qquad \check{X} := \sqrt{N}\,\text{IDFT}(X).$$

For our purposes we are more interested in the *Centered Discrete Fourier Transform* (`UCDFT`). That is why we start from scratch. For details we refer to [2],[8] and [15].

B.1   Discretization of the Fourier Transform

Let us recall the definition of the unitary Fourier transformation on $\mathbb{R}$:

Definition 2. *Let* $f \in L^1(\mathbb{R})$.

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx}\,dx$$

An obvious approach to discretize $\hat{f}$ is

$$\hat{f}_{L,M}(k) = \frac{1}{\sqrt{2\pi}} \sum_{m=-M}^{M} f(m\Delta x) e^{-imk\Delta x} \Delta x,$$

where $\Delta x = \frac{L}{2M+1}$, and f is assumed to be continuous from now on. We are here not interested in the quality of such approximations, only in the connection to the `DFT`. To recover f from $\widehat{f}_{L,M}$ at the points $x_m = m\Delta x$, we have to evaluate $\widehat{f}_{L,M}$ at the points

$$k_n = n\Delta k, \qquad \text{where} \quad \Delta k = \frac{2\pi}{L} = \frac{2\pi}{(2M+1)\Delta x},$$

where $n = -M, \ldots, M$.

Definition 3. *Let $x_m = m\Delta x$, $k_n = n\Delta k$, where $n, m = -M, \ldots, M$, and $L = (2M+1)\Delta x$.*

$$\hat{f}_{L,M}(k_n) := \frac{1}{\sqrt{2\pi}} \sum_{m=-M}^{M} f(x_m) e^{-ix_m k_n} \Delta x.$$

Proposition 5. *With the definitions above we can recover f at the points $x_r = r\Delta x$ by*

$$f(x_r) = \frac{1}{\sqrt{2\pi}} \sum_{n=-M}^{M} \hat{f}_{L,M}(k_n) e^{ix_r k_n} \Delta k.$$

$r = -M, \ldots, M.$

Proof.

$$\frac{1}{\sqrt{2\pi}} \sum_{n=-M}^{M} \left(\frac{1}{\sqrt{2\pi}} \sum_{m=-M}^{M} f(x_m) e^{-ix_m k_n} \Delta x \right) e^{ix_r k_n} \Delta k =$$

$$\frac{1}{2\pi} \sum_{n=-M}^{M} \sum_{m=-M}^{M} f(x_m) e^{-ix_m k_n} e^{ix_r k_n} \Delta x \Delta k =$$

$$\frac{1}{2\pi} \sum_{n=-M}^{M} \sum_{m=-M}^{M} f(x_m) e^{i(x_r - x_m)k_n} \Delta x \Delta k =$$

$$\frac{1}{2\pi} \sum_{m=-M}^{M} f(x_m) \left(\sum_{n=-M}^{M} e^{i(x_r - x_m)k_n} \right) \Delta x \Delta k =$$

$$\frac{2M+1}{2\pi} \sum_{m=-M}^{M} f(x_m) \delta_{rm} \Delta x \Delta k = \frac{2M+1}{2\pi} f(x_r) \Delta x \Delta k = f(x_r)$$

$\square$

Now, set $N = 2M + 1$.

Definition 4. *We define the vectors*

$$X = (X_0, \ldots, X_{N-1}), Y = (Y_0, \ldots, Y_{N-1}) \in \mathbb{C}^N$$

as follows:
$$X_m = f((m-M)\Delta x)\Delta x, \quad m = 0, \ldots, 2M.$$
$$Y_n = \hat{f}_{L,M}((n-M)\Delta k)\Delta k, \quad n = 0, \ldots, 2M.$$
Note that $2M = N-1$, and $\Delta x \Delta k = \frac{2\pi}{(2M+1)} = \frac{2\pi}{N}$.

Therefore, we get from

$$\hat{f}_{L,M}(k_n) = \frac{\Delta x}{\sqrt{2\pi}} \sum_{m=-M}^{M} f(m\Delta x) e^{-imn\Delta k\Delta x}, \quad n = -M, \ldots, M,$$

by inserting X, Y :

$$Y_{n+M} = \frac{\Delta k}{\sqrt{2\pi}} \sum_{m=-M}^{M} X_{m+M} e^{-2\pi i \frac{mn}{2M+1}}.$$

Now, let $m' = M + m$, $n' = M + n$, then

$$Y_{n'} = \frac{\Delta k}{\sqrt{2\pi}} e^{2\pi i \frac{M(n'-M)}{N}} \sum_{m'=0}^{N-1} X_{m'} e^{-2\pi i \frac{m'(n'-M)}{N}}$$

where now $m', n' = 0, \ldots, N - 1$. Thus we finally obtain

$$Y_{n'} = \frac{\Delta k}{\sqrt{2\pi}} e^{2\pi i \frac{M(n'-M)}{N}} \mathrm{DFT}\left(\left\{X_{m'} e^{2\pi i \frac{m'M}{N}}\right\}\right)_{n'}.$$

B.1.1 Matrix representation

Let $N = 2M + 1$, $\omega_N = e^{-2\pi i/N}$, then we define the $N \times N$ matrix

$$U_{mn} = \frac{1}{\sqrt{N}} \omega_N^{(m-M-1)(n-M-1)}, \quad m, n = 1, \ldots, N.$$

Therefore,

$$\hat{f}_{L,M}(k_{n-M-1}) := \frac{\Delta x \sqrt{N}}{\sqrt{2\pi}} \sum_{m=1}^{N} U_{mn} f(x_{m-M-1}) = \frac{L}{\sqrt{2\pi N}} \sum_{m=1}^{N} U_{mn} f(x_{m-M-1}).$$

The inverse is given by

$$f(x_{r-M-1}) = \frac{\sqrt{N}}{\sqrt{2\pi}} \sum_{n=1}^{N} U_{rn}^{\star} \hat{f}_{L,M}(k_{n-M-1}) \Delta k = \frac{\sqrt{2\pi N}}{L} \sum_{n=1}^{N} U_{rn}^{\star} \hat{f}_{L,M}(k_{n-M-1}).$$

The matrix U is unitary.

$$\hat{f}_{L,M} = \frac{L}{\sqrt{2\pi N}} U f \Rightarrow |\hat{f}_{L,M}|^2 = \frac{L^2}{2\pi N} |f|^2.$$

Note that $|\cdot|$ means the vector norm and not the $\|\cdot\|_2$ norm. The latter is the former times Δx or Δk respectively. See below.

B.1.2   Normalization condition

The corresponding discrete expression to

$$\int_{\mathbb{R}} |f(x)|^2 dx$$

is

$$\sum_{m=-M}^{M} |f(x_m)|^2 \Delta x = 1.$$

Thus,

$$\sum_{m=1}^{N} |f(x_{m-M-1})|^2 = \frac{1}{\Delta x} = \frac{N}{L} \Rightarrow \|v_f\| = \sqrt{\frac{N}{L}}.$$

In k-space, however,

$$\sum_{m=1}^{N} |g(k_{m-M-1})|^2 = \frac{1}{\Delta k} = \frac{L}{2\pi} \Rightarrow \|v_g\| = \sqrt{\frac{L}{2\pi}}.$$

B.1.3   Grid space calibration

Recalling the grid spacing in x and k space

$$\Delta x = \frac{L}{N}, \quad \Delta k = \frac{2\pi}{L},$$

we see that they usually are different. The condition for $\Delta x = \Delta k$ is

$$L^2 = 2\pi N,$$

thus, the factors in the matrix representation above will become $\frac{L}{\sqrt{2\pi N}} = 1$.

B.2   The Measures

B.2.1   The projection measures (marginals)

The discrete analogues to ν_φ and $\nu_{\hat{\varphi}}$ are

$$\nu_\varphi = \Delta x \sum_{m=-M}^{M} |\varphi(x_m)|^2 \delta_{x_m}, \quad \nu_{\hat{\varphi}} = \Delta k \sum_{m=-M}^{M} |\hat{\varphi}(k_m)|^2 \delta_{k_m}.$$

B.2.2 The couplings

The admissible set of measures is (as in the continuous case):

$$(\pi_{1,\#}\gamma)(f) = \gamma(\pi_1^{\#}f) = \gamma(f \circ \pi_1) =$$

$$\Delta x \Delta k \sum \gamma_{m,n} f(\pi_1(x_m, k_n)) = \Delta x \sum |\varphi(x_m)|^2 f(x_m).$$

Therefore,

$$\Gamma(\varphi) = \left\{ \gamma \in \mathbb{R}_+^{N \times N} : \sum_{n=-M}^{M} \gamma_{mn} = |\varphi(x_m)|^2, \ \sum_{m=-M}^{M} \gamma_{mn} = |\hat{\varphi}(k_n)|^2 \right\} \Delta x \Delta k.$$

B.2.3 Push forward

$$T_{\#}\nu_{\varphi}(f) = \nu_{\varphi}(f \circ T) =$$

$$\Delta x \sum_{m=-M}^{M} |\varphi(x_m)|^2 \delta_{x_m}(f \circ T) = \Delta x \sum_{m=-M}^{M} |\varphi(x_m)|^2 f(T(x_m))$$

$$T_{\#}\nu_{\varphi}(f) = \nu_{\hat{\varphi}}(f) = \Delta k \sum_{m=-M}^{M} |\hat{\varphi}(k_m)|^2 f(k_m)$$

$$\sum_{m=-M}^{M} |\varphi(x_m)|^2 f(T(x_m)) \Delta x = \sum_{m=-M}^{M} |\hat{\varphi}(k_m)|^2 f(k_m) \Delta k$$

$$\sum_{m=-M}^{M} |\varphi(T^{-1}(y_m))|^2 f(y_m) \frac{\Delta y}{T'(T^{-1}(y_m))} = \sum_{m=-M}^{M} |\hat{\varphi}(k_m)|^2 f(k_m) \Delta k$$

$$x_m = T^{-1}(y_m) \Rightarrow \Delta y = T(x_{m+1}) - T(x_m) = T'(x_m)\Delta x$$

B.3 A discrete model

B.3.1 The energies

$$E_H(\varphi) = \sum_{m,n=-M}^{M} H(x_m, k_n)|\varphi(x_m)|^2|\hat{\varphi}(k_n)|^2 \Delta x \Delta k$$

$$K_H(\varphi) = \inf_{\gamma \in \Gamma(\varphi)} \sum_{m,n=-M}^{M} H(x_m, k_n)\gamma_{mn}\Delta x \Delta k$$

$$M_H(\varphi) = \inf_{T_{\#}\nu_{\varphi}=\nu_{\hat{\varphi}}} \sum_{m=-M}^{M} H(x_m, T(x_m))|\varphi(x_m)|^2 \Delta x$$

B.3.2 Ground state of E_H

With $X_m = \varphi(x_{m-M-1})\sqrt{\Delta x}$, $m = 1 \ldots N = 2M + 1$, $\lambda_0 = \inf_\varphi E_H(\varphi)$ becomes

$$\lambda_0(H) = \inf_{|X|=1} \sum_{1 \leqslant i,j \leqslant N} H_{ij}|X_i|^2|(UX)_j|^2,$$

where U is the matrix given in B.1.1, and

$$H_{ij} = H(x_{i-M-1}, k_{j-M-1}).$$

The normalization condition $|X| = 1$ corresponds to $\|\varphi\|_2 = 1$. Indeed,

$$\sum_{m=1}^{N} |X_m|^2 = \sum_{m=1}^{N} |\varphi(x_{m-M-1})|^2 \Delta x = \sum_{l=-M}^{M} |\varphi(x_l)|^2 \Delta x.$$

B.3.3 Special cases: $H = A + B, H = AB$

$$\lambda_0(A + B) = \inf_{|X|=1} \left\{ \sum_i A_i|X_i|^2 + \sum_j B_j|(UX)_j|^2 \right\}.$$

Consider

$$\sum_j B_j|(UX)_j|^2 = \sum_{j,l,m} B_j U_{jl} X_l U_{jm}^\star X_m^\star,$$

therefore,

$$\sum_{j,l,m} (A_j \delta_{jl}\delta_{jm} X_l X_m^\star + B_j U_{jl} U_{jm}^\star X_l X_m^\star) = \sum_{l,m} C_{lm} X_l X_m^\star$$

where

$$C_{lm} = \sum_j (A_j \delta_{jl}\delta_{jm} + B_j U_{jl} U_{jm}^\star) = A_l \delta_{lm} + \sum_j B_j U_{jl} U_{jm}^\star$$

Recalling the definition of U :

$$C_{lm} = A_l \delta_{lm} + \frac{1}{N} \sum_j B_j \omega_N^{(j-M-1)(l-M-1)} \omega_N^{-(j-M-1)(m-M-1)} =$$

$$A_l \delta_{lm} + \frac{1}{N} \sum_j B_j \omega_N^{(j-M-1)(l-m)} = A_l \delta_{lm} + \frac{\omega_N^{-(M+1)(l-m)}}{N} \sum_j B_j \omega_N^{j(l-m)}.$$

Since

$$H(x, k) = A(x) + B(k).$$

30

we get

$$H_{lm} = A_l \delta_{lm} + \frac{\omega_N^{-(M+1)(l-m)}}{N} \sum_{j=1}^{N} B_j \omega_N^{j(l-m)},$$

where $A_l = A(x_{l-M-1})$ and $B_m = B(k_{m-M-1})$.

B.3.4   Special case $B(k) \sim |k|^2$

For this case we ought to obtain a similar expression as in (37), i.e. a closed form for the term

$$\kappa_N = \frac{\omega_N^{-(M+1)(l-m)}}{N} \sum_{j=1}^{N} B_j \omega_N^{j(l-m)},$$

with $B_j = B(k_{j-M-1}) = \alpha(j - M - 1)^2$, where α is some real valued constant. Let $s = l - m$, and recall that $M = \frac{N-1}{2}$, then

$$\kappa_N(s) = \alpha \frac{\omega_N^{-\frac{N+1}{2}s}}{N} \sum_{j=1}^{N} \left(j - \frac{N+1}{2} \right)^2 \omega_N^{js}.$$

A short calculation yields for $s \neq 0$:

$$\kappa_N(s) = \alpha(-1)^s \frac{\cos\left(\frac{\pi s}{N}\right)}{1 - \cos\left(2\frac{\pi s}{N}\right)} = \alpha(-1)^s \frac{\cos\left(\frac{\pi s}{N}\right)}{2\sin\left(\frac{\pi s}{N}\right)^2},$$

where we used the formulae

$$S_N(q) = \sum_{j=1}^{N} q^j = q\frac{1 - q^N}{1 - q}, \quad qS_N'(q) = \sum_{j=1}^{N} jq^j, \quad q(qS_N'(q))' = \sum_{j=1}^{N} j^2 q^j.$$

If $s = 0$, we get

$$\kappa_N(0) = \frac{\alpha}{N} \sum_{j=1}^{N} \left(j - \frac{N+1}{2} \right)^2 = \alpha \frac{N^2 - 1}{12}.$$

Thus, the kinetic term $\frac{\hbar^2}{2m}|k|^2$ becomes

$$T_{ij} = \frac{h^2}{8\pi m N} \begin{cases} \frac{N^2-1}{6}, i = j \\ (-1)^{(l-m)} \frac{\cos\left(\frac{\pi(i-j)}{N}\right)}{\sin\left(\frac{\pi(i-j)}{N}\right)^2}, i \neq j \end{cases}.$$

31

With

$$\sum_{m=-M}^{M} V(x_m)|\varphi(x_m)|^2 \Delta x = \sum_{m=-M}^{M} V(x_m)|X_m|^2 =$$

$$\sum_{m=1}^{N} V(x_{m-M+1})|X_m|^2 \Rightarrow V_m = V(x_{m-M+1}) = V((m - M + 1)\Delta x).$$

we get finally:

$$H_{ij} = V_i \delta_{ij} + \frac{h^2}{8\pi m N} \begin{cases} \frac{N^2-1}{6}, i = j \\ (-1)^{(i-j)} \frac{\cos\left(\frac{\pi(i-j)}{N}\right)}{\sin\left(\frac{\pi(i-j)}{N}\right)^2}, i \neq j \end{cases}.$$

This is different from (37) because we use an odd number of grid points ($N = 2M + 1$), whereas (37) will become (when inserting $L^2 = 2\pi N$):
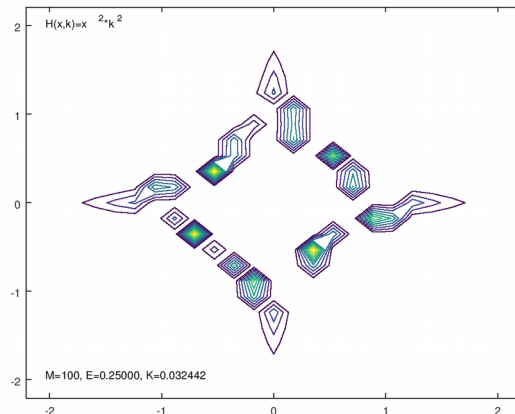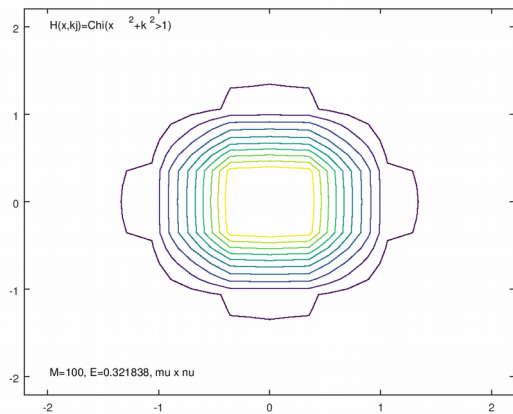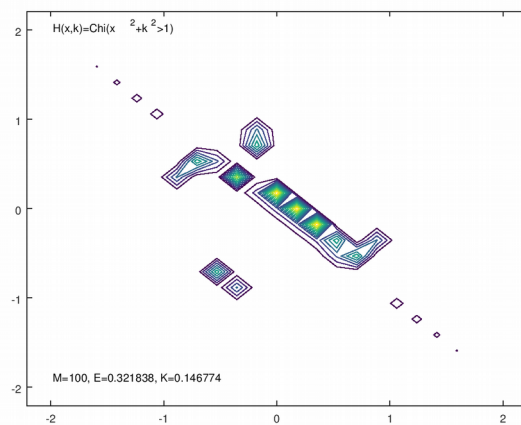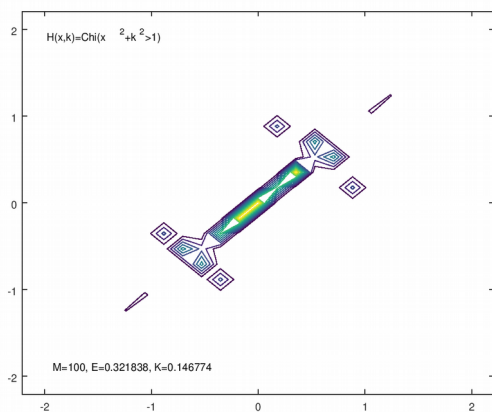
$$H_{ij}^0 = V_i \delta_{ij} + \frac{h^2}{8\pi m N} \begin{cases} \frac{N^2+2}{6}, i = j \\ (-1)^{(i-j)} \frac{1}{\sin\left(\frac{\pi(i-j)}{N}\right)^2}, i \neq j \end{cases}.$$

B.4 Numerics

Just for illustration purposes we calculated some quantities using *GNU Octave* [6].

```
John W. Eaton, David Bateman, Soren Hauberg, Rik Wehbring (2016).
GNU Octave version 4.2.0 manual: a high-level interactive
language for numerical computations.
URL http://www.gnu.org/software/octave/doc/interpreter/
```

Due to the lack of analytical examples regarding the optimal transport measures it will be be rather wishful to have some reliable numerical instances at least. Of course, the small script following does not claim either correctness nor completeness, not to speak of any significance.

Top-left: H(x,k)=Chi(x²+k²>1) — M=100, E=0.321838, K=0.146774

Top-right: H(x,k)=Chi(x²+k²>1) — M=100, E=0.321838, K=0.146774

Middle-left: H(x,kj)=Chi(x²+k²>1) — M=100, E=0.321838, mu x nu

Middle-right: H(x,k)=x²+k² — M=100, E=0.25000, K=0.032442

Bottom-left: H(x,k)=x²+k² — M=100, E=0.25000, K=0.032442, linsolve

Bottom-right: H(x,k)=x²+k² — M=100, E=1.0000, K=1.0000
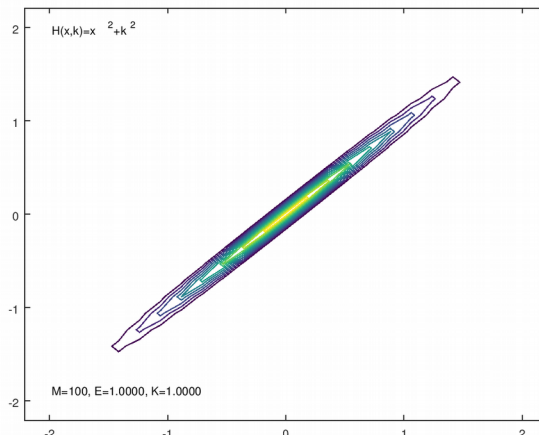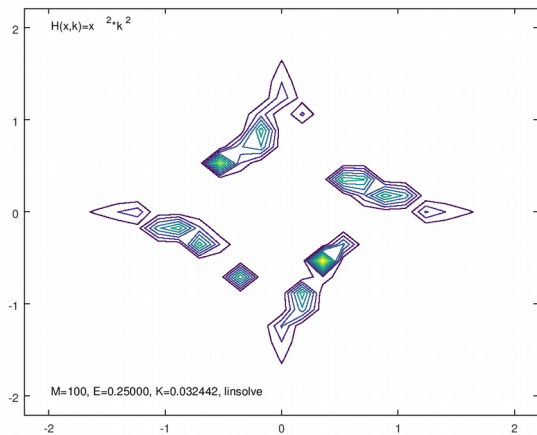
```octave
clear all;

M = 100;
H = @(x,k) x^2*k^2;

function [N,L,x,k,dx,dk] = setup_grid(M)
  N = 2*M+1;
  L = sqrt(2*pi*N);
  dx = L/N;
  dk = 2*pi/L;
  x = (-M:1:M)*dx; #x = linspace(-M,M,N)*dx
  k = (-M:1:M)*dk; #k = linspace(-M,M,N)*dk
endfunction


function rv = dimGetAndCheck(x,y)
  # Checks if x,y are vectors of the same size (col/row does not matter).
  #
  if ~isvector(x) error("x: vector expected.") endif
  if ~isvector(y) error("y: vector expected.") endif
  [x0,x1] = size(x);
  [y0,y1] = size(y);
  N = 1 + abs(x1-x0);
  if (N == 1 + abs(y1-y0))
    rv = N;
  else
    error("vectors of equal size required.");
  endif
endfunction


function rv = FunctionMatrix(f,x,y)
  # FunctionMatrix(f,x,y)= matrix f(x_i,y_j)
  #
  N = dimGetAndCheck(x,y);
  F = eye(N);
  for m=1:N
    for n=1:N
      F(m,n) = f(x(m),y(n));
    endfor
  endfor
  rv = F;
endfunction


function rv = UCDFT(M)
  # Unitary Centered Discrete Fourier Transform (UCDFT)
  # M = (N-1)/2 : integer
  # Computes the NxN matrix
  #   UCDFT(M)= 1/sqrt(N)*[exp(-2*i*pi/N)^((m-M-1)*(n-M-1))]
  #
  N = 2*M+1;
  U = eye(N);
  w = exp(-2*i*pi/N);
  for m=1:N
    for n=1:N
      U(m,n)=w^((m-M-1)*(n-M-1));
    endfor
  endfor
  rv = U/sqrt(N);
endfunction


function checkL2funvec(phi,M)
  # phi has to be a L^2 normalized complex valued row vector with
  # phi(1)=phi(N)=0.
  #
  eps0 = 1e-9;
  if ~isvector(phi) error("vector expected.") endif
  if ~isrow(phi) error("row vector expected.") endif
  if (abs(norm(phi)-1.0)>eps0) error("norm(vector,2) is not 1.") endif
  if (length(phi)~=2*M+1) error("size of phi must be 2*M+1.") endif
  if ~(phi(1)==0 && phi(end)==0) error("boundary condition not satisfied") endif
endfunction


function rv = makePhi(f,x)
  # Create a row vector that will pass checkL2funvec from a function f
  # Ex. phi=makePhi(@(t) t^2,[1,2,3,4,5]) -> checkL2funvec(phi,2)
  fx = arrayfun(f,x);
  fx(1)=0;
  fx(end)=0;
  rv = fx/norm(fx);
endfunction


function [mu,nu] = projMeasures(phi,M)
  # Create the marginal measures mu,nu from the function-vector phi.
  checkL2funvec(phi,M);
  dftphi = UCDFT(M) * phi';
  psi = dftphi' ;
  mu=phi .* conj(phi);  # sum(mu)=sum(nu)=1 ? check
```

```octave
    nu=psi .* conj(psi);
endfunction


function rv = tensorProductMeasure(mu,nu)
  # g = mu \otimes nu => g(i,j)=mu(i)*nu(j)
  # Assumes that all checks have been done before
  rv = mu' .* nu;
endfunction


function rv = makeConstraintMatrix(M)
  # Build a matrix A for LP
  # Use A = makeConstraintMatrix(M) -> sparse A
  # full(A) builds normal reprsentation.
  flat = @(x)x(:);
  mkc = @(n)sparse( flat(repmat(1:n, [n 1])), ...
            flat(reshape(1:n*n,n,n) ), ...
            ones(n*n,1) );
  mkr = @(n)sparse( flat(repmat(1:n, [n 1])), ...
            flat(reshape(1:n*n,n,n)' ), ...
            ones(n*n,1) );
  mkA = @(n)[mkr(n);mkc(n)];
  rv = mkA(2*M+1);
endfunction


function [gamma,mu,nu,x,k,EK,ES,HM,phi,FMIN,ERRNUM] = SolveWithGLPK(H,phi,M)
  [N,L,x,k,dx,dk] = setup_grid(M);
  [mu,nu] = projMeasures(phi,M);
  tpm = tensorProductMeasure(mu,nu);
  HM = FunctionMatrix(H,x,k);
  C = HM(:);
  A = makeConstraintMatrix(M);
  b = [mu,nu](:);
  [XOPT, FMIN, ERRNUM, EXTRA]=glpk(C,A,b);
  gamma = reshape(XOPT,N,N);
  EK=trace(HM'*gamma);
  ES=trace(HM'*tpm);
endfunction


function checkGLPKSolution(g,m,n,ek,err,fmin)
  # Checks the results from "glpk".
  eps0 = 1e-7;
  gm = norm(sum(g,2)-m(:));
  gn = norm(sum(g,1)'-n(:));
  gp = full(sum(g(:)~=0));
  de = abs(fmin-ek);
  N = length(m);
  if (err~=0) fprintf("Warning: ERRNUM = %d \n",err); endif
  if (de > eps0) fprintf("Warning: FMIN=%d <> EK \n",fmin); endif
  if (gm>eps0) fprintf("Warning: px(gamma)- mu = %d \n",gm); endif
  if (gn>eps0) fprintf("Warning: pk(gamma)- nu = %d \n",gn); endif
  if (gp>2*N-1) fprintf("Warning: #supp(gamma) = %d \n",gp); endif
endfunction


function [gamma,mu,nu,x,k,EK,ES,HM,phi] = Solve1(H,f,M)
  # Solves EK(phi)=min{g: Tr(H'*g), px(g)=mu(phi), pk(g)=nu(phi)}
  # using LP method/function glpk.
  [N,L,x,k,dx,dk] = setup_grid(M);
  phi = makePhi(f,x);
  [gamma,mu,nu,x,k,EK,ES,HM,phi,FMIN,ERRNUM] = SolveWithGLPK(H,phi,M);
  # checks
  checkGLPKSolution(gamma,mu,nu,EK,ERRNUM,FMIN);
endfunction


function [gamma,mu,nu,x,k,EK,ES,HM,phi,FMIN] = SolveWithLinprog(H,phi,M)
  # [XOPT,FMIN,ITN] = linprog(A,b,c,k,maxit,tol)
  # max iterations ~ 50k (10k is too small)
  maxit = 50000;
  tol = 1e-9;
  [N,L,x,k,dx,dk] = setup_grid(M);
  [mu,nu] = projMeasures(phi,M);
  tpm = tensorProductMeasure(mu,nu);
  HM = FunctionMatrix(H,x,k);
  C = HM(:);
  A = makeConstraintMatrix(M);
  b = [mu(:);nu(:)];
  [XOPT, FMIN, ITN]=perform_linprog(A,b,C,0,maxit,tol);
  gamma = reshape(XOPT,N,N);
  EK=trace(HM'*gamma);
  ES=trace(HM'*tpm);
endfunction

function [gamma,mu,nu,x,k,EK,ES,HM,phi] = Solve2(H,f,M)
  # Solves EK(phi)=min{g: Tr(H'*g), px(g)=mu(phi), pk(g)=nu(phi)}
  # using LP method/function perform_linprog.
  [N,L,x,k,dx,dk] = setup_grid(M);
  phi = makePhi(f,x);
```

```
    [gamma,mu,nu,x,k,EK,ES,HM,phi,FMIN] = SolveWithLinprog(H,phi,M);
    # checks
    checkGLPKSolution(gamma,mu,nu,EK,0,FMIN);
endfunction


####
function energy = ESchroed(HM,U,M,phi,x,k)
  X = phi'/norm(phi);
  Y = (U * X')';
  mu = X .* conj(X);
  nu = Y .* conj(Y);
  tpm = tensorProductMeasure(mu,nu);
  energy = trace(HM*tpm');
endfunction


function constraint = g(phi)
  [a,b] = size(phi);
  constraint = [sumsq(phi)-1; phi(a); phi(b)];
endfunction


function [phi0, obj, info, iter, nf, lambda,x] = findGroundState(H,M)
  [N,L,x,k,dx,dk] = setup_grid(M);
  HM = FunctionMatrix(H,x,k);
  U = UCDFT(M);
  E = @(phi) ESchroed(HM,U,M,phi,x,k);
  X0 = makePhi(@(t) exp(-t^2/2),x);
  [phi0, obj, info, iter, nf, lambda] = sqp (X0, E, @g, []);
endfunction


function [HM,phi,E] = schroedEq(M,V,m,h)
  # Solve the Schroedinger equation by the FGH method
  [N,L,x,k,dx,dk] = setup_grid(M);
  VV = arrayfun(V,x);
  fac = h^2/(8*pi*m*N);
  HM = eye(N);
  for m=1:N
    for n=1:N
      if (m==n)
        HM(m,n) = fac*(N^2-1)/6 + VV(m);
      else
        HM(m,n) = fac*(-1)^(m-n)*cos(pi*(m-n)/N)/sin(pi*(m-n)/N)^2;
      endif
    endfor
  endfor
  [vv,eed] = eig(HM);
  [E,perm] = sort(diag(eed));
  phi = vv(:,perm);
endfunction

# Examples:
# [gamma,mu,nu,x,k,EK,ES,HM,phi] = Solve1(@(x,k) x^2*k^2,@(t) exp(-t^2/2),M);
# [gamma,mu,nu,x,k,EK,ES,HM,phi] = Solve1(@(x,k) x^2+k^2,@(t) exp(-t^2/2),M);
# [gamma,mu,nu,x,k,EK,ES,HM,phi] = Solve2(@(x,k) x^2*k^2,@(t) exp(-t^2/2) ,M);


### Ground states
# [phi0, obj, info, iter, nf, lambda,x] = findGroundState(@(x,k) x^2+k^2,M);
# plot(x,phi0)

# [phi0, obj, info, iter, nf, lambda,x] = findGroundState(H,M);
# plot(x,phi0)

#  H=@(x,k) int8(x^2+k^2>4);
# [gamma1,mu,nu,x,k,EK1,ES,HM,phi] = Solve1(H,@(t) exp(-t^2/2) ,M);
# [gamma2,mu,nu,x,k,EK2,ES,HM,phi] = Solve2(H,@(t) exp(-t^2/2) ,M);


#[N,L,x,k,dx,dk] = setup_grid(M);
#V = @(t) t^2;
#[HM,phi,E] = schroedEq(M,V,1/2,2*pi);  # hbar^2=2*m => m=1/2, h=2*pi
#g0 = @(t) exp(-t^2/2);
#h0 = arrayfun(g0,x)
#h0 = h0/norm(h0);
#plot(x,-phi(:,1),x,h0)
```

References

[1] Luigi Ambrosio and Nicola Gigli. *A User's Guide to Optimal Transport*, pages 1–155. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[2] L. Auslander and R. Tolimieri. Is computing with the finite fourier transform pure or applied mathematics? *Bull. Amer. Math. Soc. (N.S.)*, 1(6):847–897, 11 1979.

[3] Gabriel G. Balint-Kurti, Richard N. Dixon, and C. Clay Marston. Grid methods for solving the schrödinger equation and time dependent quantum dynamics of molecular photofragmentation and reactive scattering processes. *International Reviews in Physical Chemistry*, 11(2):317–344, 1992.

[4] Harry Bateman. *Tables of Integral Transforms Vol. I., II.* McGraw-Hill Book Company, New York, 1954.

[5] S. H. Chang, P. C. Cosman, and L. B. Milstein. Chernoff-type bounds for the gaussian error function. *IEEE Transactions on Communications*, 59(11):2939–2944, November 2011.

[6] John W. Eaton, David Bateman, Søren Hauberg, and Rik Wehbring. *GNU Octave version 4.2.0 manual: a high-level interactive language for numerical computations*, 2016.

[7] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems.* Society for Industrial and Applied Mathematics, 1999.

[8] Alberto Grunbaum. The eigenvectors of the discrete fourier transform: A version of the hermite functions. 88:355–363, 08 1982.

[9] Victor Havin and Burglind Jöricke. *The Uncertainty Principle in Harmonic Analysis.* Springer Science Business Media, December 2012.

[10] Philippe Jaming. Nazarov's uncertainty principles in higher dimension. *Journal of Approximation Theory*, page doi:10.1016/j.jat.2007.04.005, May 2007.

[11] Philippe Jaming and M. Powell, Alexander. Uncertainty principles for orthonormal sequences. *Journal of Functional Analysis*, 243:611–630, February 2007.

[12] Eugenia Malinnikova. Orthonormal sequences in $L^2(\mathbb{R}^d)$ and time frequency localization. *Journal of Fourier Analysis and Applications*, 16(6):983–1006, 2010.

[13] C. Clay Marston and Gabriel G. Balint-Kurti. The fourier grid hamiltonian method for bound state eigenvalues and eigenfunctions. *The Journal of Chemical Physics*, 91(6):3571–3576, 1989.

[14] Robert J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Mathematical Journal*, 80(2):309–323.

[15] Dale Mugler. The centered discrete fourier transform and a parallel implementation of the fft, 06 2011.

[16] Antonio Steiner. Eine ungleichung über die fouriertransformation. *Rendiconti del Circolo Matematico di Palermo*, 23(1):83–86, Jan 1974.

[17] C. Villani. *Optimal Transport: Old and New.* Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2008.

[18] Cedric Villani. *Topics in Optimal Transportation.* American Mathematical Society, Providence, RI, March 2003.