SYMPLECTIC CAPACITIES PROJECT

Abstract. A record of current progress in the symplectic capacities project, as well as some background information.

Contents

1. template section

2. TODO

(1) Proofs in the appendix.
(2) complete introductory section on the EHZ capacity and the dual problem.
(3) Explanation of shifting error in the reconstruction algorithm.
(4) Proper bibliography.

3. List of Notations

Record of notation being used to prevent collisions in the notation space.

ω_0	–	The standard symplectic structure on $\mathbb{R}^{2n}$.
J	–	The standard complex structure on $\mathbb{R}^{2n}$.
M	–	The function space $\{\dot{x} \in L^2([0,1], \mathbb{R}^{2n}) \, : \, \int \dot{x} = 0\}$.
$\mathcal{F}_K$	–	The functional on M.
$\mathcal{E}_K$	–	The space of closed characteristics $\{z \in C^2(S^1, \mathbb{R}^{2n}) \, : \, \dot{z} = X_H\}$.
$\mathcal{A}$	–	The action on $\mathcal{E}_K$.
$\mathcal{K}$	–	the set of all compact, strictly convex subsets of $\mathbb{R}^{2n}$ which have C^2 boundary.
$\|\| \cdot \|\|_K$	–	Norm of the convex body K.
K°	–	Dual convex body of K.
h_K	–	The supporting function of K.
$\mathcal{L}H$	–	The Legendre transform of H.
Δ	–	A simplex in $\mathbb{R}^{2n}$.
v_i	–	The vertices of Δ.
F_i	–	The face of Δ opposite the vertex v_i.
n_i	–	The unit outward normal vector for the face F_i.

4. Background on Symplectic Capacities

Fix coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$ on $\mathbb{R}^{2n}$ and let ω_0 be the standard symplectic structure $\sum dx_i \wedge dy_i$. The ball with radius $r > 0$ is the set

$$B_{2n}(r) = \{(q,p) \in \mathbb{R}^{2n} : \sum (q_i^2 + p_i^2) < r^2\}$$

and the standard cylinder is the set

$$Z_{2n}(r) = \{(q,p) \in \mathbb{R}^{2n} : q_1^2 + p_1^2 < r^2\} = B_2(r) \times \mathbb{R}^{2(n-1)}.$$

Any open subset of $\mathbb{R}^{2n}$ with the restriction of ω_0 is a symplectic manifold. A symplectomorphism of $\mathbb{R}^{2n}$ is a diffeomorphism ϕ such that $\phi^* \omega_0 = \omega_0$.

Definition 1. *A* symplectic capacity *on $\mathbb{R}^{2n}$ is a function c defined on open subsets of $\mathbb{R}^{2n}$, taking values in the set $[0, \infty]$ with the following properties:*

1. *(Monotonicity) If $U \subset V$, then $c(U) \le c(V)$,*
2. *(Invariance) If ϕ is a symplectomorphism then $c(\phi(U)) = c(U)$,*
3. *(Homogeneity) For all $\alpha > 0$, $c(\alpha U) = \alpha^2 c(U)$, and*
4. *(Non-triviality) $0 < c(B_{2n}(1))$ and $c_{2n}(Z(1)) < \infty$.*

A symplectic capacity c is said to be normalized *if $c(B_{2n}(1)) = c(Z_{2n}(1)) = \pi$.*

4.1. **The Eckland-Hofer-Zhender capacity.** The Eckland-Hofer-Zhender capacity, denoted c_{EHZ}, is a normalized symplectic capacity defined on compact, convex subsets of $\mathbb{R}^{2n}$. Let K be a compact, strictly convex set in $\mathbb{R}^{2n}$ such that the interior of K contains the origin and the boundary of K is smooth of class C^2 (i.e. locally it has C^2 parameterizations). Let

$$H(x) = ||x||_K^2 \, ,$$

where $\|x\|_K = \inf\{\lambda > 0 : x \in \lambda K\}$ is the semi-norm with respect to the convex body K. The Hamiltonian vector field $X_H = J\nabla H$ defines a flow on ∂K which is the level set $H^{-1}(1)$. Consider the space of periodic orbits of X_H,

$$\mathcal{E}_K = \left\{ z \in C^2(\mathbb{R}, \partial K) : \dot{z} = X_H \circ z, \; z(0) = z(T) \text{ for some } T > 0 \right\}$$

Definition 2. *The* action *of an orbit $z \in \mathcal{E}_K$ with period T is*

$$\mathcal{A}(z) = \frac{1}{2}\int_0^T \langle -J\dot{z}, z\rangle \, dt.$$

The Eckland-Hofer-Zhender capacity *of K is the quantity*

$$c_{EHZ}(K) = \inf\{\mathcal{A}(z) : z \in \mathcal{E}_K\}$$

Let $\mathcal{K}$ be the set of all compact, strictly convex subsets of $\mathbb{R}^{2n}$ which have C^2 boundary.

Theorem 1. *The function c_{EHZ} satisfies the axioms of a symplectic capacity on the set $\mathcal{K}$.*

To prove this theorem, one considers an equivalent variational problem. Let

$$M = \left\{ x \in H^1([0,1], \mathbb{R}^{2n}) : \int_0^1 \dot{x}(t)dt = 0 \right\}$$

and consider the functional

$$\mathcal{F}_K(x) = \int_0^1 h_K^2(-J\dot{x}(t))dt. \tag{1}$$

The constrained minimization problem

$$\lambda = \inf \left\{ \mathcal{F}_K(x) : x \in M, \int_0^1 \langle -J\dot{x}(t), x(t)\rangle dt = 1 \right\} \tag{2}$$

has at least one solution, denoted by u.

Proposition 1. *Let u be a minimizer for $\mathcal{F}_K$, consider the path*

$$z(t) = 2\sqrt{\lambda} u\left(\frac{t}{2\lambda}\right) + \frac{c}{\sqrt{\lambda}}, \tag{3}$$

$$c = \int_0^1 \nabla h_K^2(-J\dot{u})dt + 2\int_0^1 u(t)dt. \tag{4}$$

The path $z(t)$ is a 2λ-periodic solution of X_H which lies in ∂K and minimizes $\mathcal{A}$. Thus

$$c_{EHZ}(K) = \mathcal{A}(z) = 2\lambda.$$

Proof. □

4.2. **Generalized Characteristics.** Now we follow [AAO]. Let $\widetilde{\mathcal{K}}$ be the set of convex bodies in $\mathbb{R}^{2n}$ with non-empty interior. There is a unique extension of c_{EHZ} to $\widetilde{\mathcal{K}}$ that is homogeneous, monotone, and continuous with respect to the Hausdorff metric. By uniqueness, this extension can be given the same definition in terms of a dual minimization problem [AAO]

$$\tilde{c}_{EHZ}(K) = 2\inf\{\mathcal{F}_K(x) : x \in M\}.$$

It was shown in [AAO] that this extension has a geometric definition. The *normal cone* to a convex body K at $x \in \partial K$ is

$$N_{\partial K}(x) = \left\{ u \in \mathbb{R}^{2n} : \langle u, x - y\rangle \geq 0, \text{ for every } y \in K \right\}.$$

Definition 3. *A* generalized closed characteristic *is a periodic, piece-wise smooth path $z : \mathbb{R} \to \partial K$ such that*

$$\dot{z}_{\pm}(t) = JN_{\partial K}(z(t))$$

for all $t \in \mathbb{R}$, where $\dot{z}_{\pm}$ are the left and right derivatives of z.

Note that we have defined the Eckland-Hofer-Zhender capacity in terms of Hamilton's equation rather than solutions to the characteristic foliation, whereas this definition of a generalized closed characteristic coincides with the latter definition in the case where K has smooth boundary. This issue of parameterization does not effect the computed action.

If $\widetilde{\mathcal{E}_K}$ is the space of generalized closed characteristics for K, then

Theorem 2. [AAO]

$$\tilde{c}_{EHZ}(K) = 2\inf\{\mathcal{F}_K(x) : x \in M\} = \inf\left\{|\mathcal{A}(z)| \,:\, z \in \widetilde{\mathcal{E}_K}\right\}.$$

5. Basic Description of Algorithm

In order to numerically solve the minimization problem (2), we follow the general lines of the discretized problem presented in [GJ+].

Given $m \in \mathbb{N}_+$, divide the interval $[0,1]$ to m subintervals $I_{m,k} := (\frac{k}{m}, \frac{k+1}{m})$, $k = 0, \dots, m-1$, of length $\frac{1}{m}$, and consider only H^1-functions that are linear on each subinterval,

(5)

$$W_m := \left\{u \in H^1([0,1], \mathbb{R}^{2n}) \;\middle|\; \dot{u}|_{I_{m,k}} \text{ is constant}, \int_0^1 u(t)dt = 0, \int_0^1 \langle -J\dot{u}, u\rangle = 1\right\},$$

where $J \in \mathrm{Mat}(2n, \mathbb{R})$ is the standard complex structure on $\mathbb{R}^{2n}$,

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Let $u \in W_m$, then $\dot{u}$ is piece-wise constant, and can be represented by a vector $\dot{\mathbf{x}} = (\dot{x}_1, \dots, \dot{x}_m) \in \mathbb{R}^{2n\cdot m}$. In this case, the functional to minimize (1) takes the form

$$F(\dot{\mathbf{x}}) = \frac{1}{m}\sum_{j=1}^{m} G(-J\dot{x}_j), \tag{6}$$

and its gradient is

$$\nabla F(\dot{\mathbf{x}}) = \frac{1}{m}(J\nabla G(-J\dot{x}_1), \dots, J\nabla G(-J\dot{x}_m)). \tag{7}$$

The constraints are given by

$$0 = \ell(\dot{\mathbf{x}}) = \sum_{j=0}^{m-1} \dot{x}_j, \tag{8}$$

$$0 = q(\dot{\mathbf{x}}) = \frac{1}{m^2}\sum_{k=1}^{m-1}\sum_{j=0}^{k-1}\langle -J\dot{x}_k, \dot{x}_j\rangle - 1 = \frac{1}{m^2}\dot{\mathbf{x}}^T A\dot{\mathbf{x}} - 1, \tag{9}$$

where

$$A := \begin{pmatrix} 0_{2n} & -J_{2n} & \cdots & -J_{2n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & -J_{2n} \\ 0_{2n} & \cdots & \cdots & 0_{2n} \end{pmatrix} \in \mathrm{Mat}(\mathbb{R}, 2n \cdot m),$$

and the gradient of the quadratic constraint is

$$\nabla q(\dot{\mathbf{x}}) = \frac{1}{m^2}(A + A^T)\dot{\mathbf{x}}.$$

To minimize the functional F under the constraints ℓ, q we use the Matlab library function named "fmincon". This function recives as input the functional, its gradient, the constraints and their gradients, together with a starting point in $\mathbb{R}^{2n \cdot m}$ and iteratively searches for a local minimum. The function "fmincon" enables to choose a minimization algorithm out of a given list. After numerically experimenting the possible algorithms, we chose "active-set" as it produced the best results in terms of accuracy and run-time. To choose a starting vector, we randomly pick a vector in $\mathbb{R}^{2n \cdot m}$, then shift and rescale it to satisfy conditions (8), (9).

6. Details of Algorithm Implementation

Hello World!

7. Algorithm Convergence

In this section we prove that minimal values of the discretized problem presented in section 5 converge to a minimal value of the original problem (2). Moreover, we discuss bound of the rate of convergence and convergence of minimizing trajectories.

we start by presenting the notation and assumptions required for the results listed in this section.

Definition 4. *For every $m \in \mathbb{N}_+$, $k = 0, \ldots, m-1$, denote $I_{m,k} := (\frac{k}{m}, \frac{k+1}{m}).$, and consider the spaces:*

$$\begin{aligned} W &:= \left\{ v \in M : \int_0^1 \langle -J\dot{v}, v \rangle_{\mathbb{R}^{2n}} = 1 \right\} \subset H^1([0,1], \mathbb{R}^{2n}) \\ W_m &:= W \cap \{ v \in H^1([0,1], \mathbb{R}^{2n}) : v|_{I_{m,k}} \textit{ is linear} \} \end{aligned}$$

We are interested in the minimal values of the functional $\mathcal{F}(v) = \int_0^1 G(-J\dot{v}(t))dt$ on the spaces W, W_m, where $G = \mathcal{L}(H) : \mathbb{R}^{2n} \to \mathbb{R}$ is the Legendre transform of the Hamiltonian H. In the following we assume that G is piece-wise C^1 with a bounded gradient in the unit ball, and positively homogeneous of degree 2 (ideologically, G will be the dual norm defined by some convex body).

7.1. **Convergence of Minimal Values.**

Lemma 1. *The union $\cup_{m \in \mathbb{N}_+} W_m$ is dense in W with respect to the H^1 topology.*

Proof. Let $v \in W$, and let $v_m \in W_m$ be the sequence of piecewise linear functions defined by: $\dot{v}_m|_{I_{m,k}} := \oint_{I_{m,k}} \dot{v}(t)dt$. Then, $\{v_m\}$ converges to v in the H^1 norm. Notice that although v_m is a closed curve, it may not necessarily be in W. However, since v_m converges to an element in W, it holds that:

$$\forall \varepsilon > 0 \ \exists m_0 \in \mathbb{N}_+ \ \forall m > m_0, \quad \left| \int_0^1 \langle -J\dot{v}, v \rangle_{\mathbb{R}^{2n}} - 1 \right| < \varepsilon$$

Thus, we can write

$$\int_0^1 \langle -J\dot{v}_m, v_m \rangle_{\mathbb{R}^{2n}} = 1 + c_m$$

for some $c_m \in \mathbb{R}$, and assume that $1 + c_m > 0$ (for large enough m). Denote $w_m := \frac{1}{\sqrt{1+c_m}} \cdot v_m$, then it is easy to check that $w_m \in W_m$. Since $v \in W$ and the above integral is continuous in the H^1-norm, it follows that $c_m \underset{m\to\infty}{\longrightarrow} 0$, which implies that $\{w_m\}$ converges to v.

□

Lemma 2. *There exists a constant $c > 0$ such that for every $x, y \in \mathbb{R}^{2n}$,*

$$|G(x) - G(y)| \leq c \cdot \max(|x|, |y|) \cdot |x - y|.$$

Proof. Since G is C^1, and the unit ball $B(0,1) \subset \mathbb{R}^{2n}$ is compact, there exists a Lipschitz constant $c > 0$ such that for every $\hat{x}, \hat{y} \in B(0,1)$, $|G(\hat{x}) - G(\hat{y})| \leq c \cdot |\hat{x} - \hat{y}|$. Let $x, y \in \mathbb{R}^n$, and assume without loss of generality that $|x| \geq |y|$. Then:

$$\begin{aligned} |G(x) - G(y)| &= |x|^2 \cdot \left| G\left(\frac{x}{|x|}\right) - G\left(\frac{y}{|x|}\right) \right| \leq |x|^2 \cdot c \cdot \left| \frac{x}{|x|} - \frac{y}{|x|} \right| \\ &= |x| \cdot c \cdot |x - y| = c \cdot \max(|x|, |y|) \cdot |x - y|. \end{aligned}$$

□

Lemma 3. *The functional $\mathcal{F}(v) := \int_0^1 G(-J\dot{v}(t))dt$ is H^1-continuous. More formally, let $v \in H^1([0,1], \mathbb{R}^{2n})$ and v_m is a sequence converging to v in H^1, then the sequence of values converges:*

$$\mathcal{F}(v_m(t)) \underset{n\to\infty}{\longrightarrow} \mathcal{F}(v(t)).$$

Proof.

$$\begin{aligned} &\left| \int_0^1 G(-J\dot{v}_m(t))dt - \int_0^1 G(-J\dot{v}(t))dt \right| \leq \int_0^1 |G(-J\dot{v}_m(t)) - G(-J\dot{v}(t))| \, dt \\ &\leq \int_0^1 c \cdot \max(|\dot{v}_m|, |\dot{v}|) \cdot |-J\dot{v}_m(t) + J\dot{v}(t)| dt \\ &\leq \int_0^1 c \cdot (|\dot{v}_m| + |\dot{v}|) \cdot |\dot{v}_m(t) - \dot{v}(t)| dt \\ &\leq c \cdot \left(\int_0^1 (|\dot{v}_m| + |\dot{v}|)^2 dt \right)^{\frac{1}{2}} \cdot \left(\int_0^1 |\dot{v}_m(t) - \dot{v}(t)|^2 dt \right)^{\frac{1}{2}} \\ &\leq c \cdot (\|\dot{v}_m\|_{L^2} + \|\dot{v}\|_{L^2}) \cdot \|\dot{v}_m(t) - \dot{v}(t)\|_{L^2} \underset{m\to\infty}{\longrightarrow} 0 \end{aligned}$$

□

Lemma 4. *Let $u_m \in W_m \subset H^1([0,1],\mathbb{R}^{2n})$ be an orbit that minimizes the action (on the space W_m). Then, the sequence of actions $\mathcal{F}(u_m)$ converges to the minimal action on the total space W.*

Proof. Let u be an orbit minimizing the action on W, and let $w_m \in W_m$ be a sequence converging to u in the H^1-norm (exists by Lemma 1). By the previous lemma, the sequence of actions also converges, therefore,

$$\mathcal{F}(u) = \lim \mathcal{F}(w_m) \geq \lim \mathcal{F}(u_m),$$

since u_m are the minimizers on each W_m. On the other hand, since $W_m \subset W$ for every m, $\mathcal{F}(u) \leq \mathcal{F}(u_m)$, which yields

$$\lim_{m\to\infty} \mathcal{F}(u_m) = \mathcal{F}(u).$$

□

7.2. **Explicit Bound on the Error.** Let u_m be a minimizer of the functional $\mathcal{F}(x) = \int_0^1 h_K^2(\dot{x}(t))dt$ on the space W_m, and u be a minimizer on the total space W. Denote by $E_m := |\mathcal{F}(u_m) - \mathcal{F}(u)|$ the error of the capacity approximation. Denote by $\alpha_m \in W_m$ the piecewise linear approximation of u, $\alpha_m = \frac{1}{\sqrt{1+c_m}}\beta_m$, where $\dot{\beta}_m$ is the average of $\dot{u}$ on each subinterval and the normalization is such that the action of α_m has action equals 1.

Lemma 5. *$E_m \leq L(h_K^2)\cdot$, where $L(h_K^2)$ is the Lipschitz constant of h_K^2 in the unit ball.*

Proof. Since u_m minimizes $\mathcal{F}$ on the space W_m, we have:

$$\begin{aligned}
|\mathcal{F}(u_m) - \mathcal{F}(u)| &\leq |\mathcal{F}(\alpha_m) - \mathcal{F}(u)| \leq \int_0^1 |\mathcal{F}(\dot{\alpha}_m) - \mathcal{F}(\dot{u})|dt \\
&\leq L(h_K^2)\cdot \max(\|\dot{\alpha}_m\|_{L^2}, \|\dot{u}\|_{L^2})\cdot \|\dot{\alpha}_m - \dot{u}\|_{L^2} \\
&\leq L(h_K^2)\cdot (\|\dot{u}\|_{L^2} + \|\dot{\alpha}_m - \dot{u}\|_{L^2})\cdot \|\dot{\alpha}_m - \dot{u}\|_{L^2}.
\end{aligned}$$

Now, denoting by $\lambda := \mathcal{F}(u)$ the minimal value, and using the dual action principle, the closed characteristic of the dual problem will be ([?going1998diss]):

$$z(t) = 2\sqrt{\lambda}\cdot u\left(\frac{t}{2\lambda}\right) + c$$

for some constant vector $c \in \mathbb{R}^{2n}$. We can use the above equation and the Hamilton equation that z satisfies to bound $\dot{u}$:

$$\begin{aligned}
\|\dot{u}\|_{L^2}^2 &= \int_0^1 \left(\frac{1}{2\sqrt{\lambda}}2\lambda\cdot \dot{z}(2\lambda t)\right)^2 dt = \lambda \int_0^{2\lambda} \dot{z}(s)^2 \frac{1}{2\lambda}ds \\
&= \frac{1}{2}\|\dot{z}\|_{L^2}^2 = \frac{1}{2}\|-J\nabla H \circ z\|_{L^2}^2 \leq \lambda\cdot \|\nabla H\|_{C^0}^2.
\end{aligned}$$

In order to bound the difference $\|\dot{\alpha}_m - \dot{u}\|_{L^2}$, we notice that

$$\begin{aligned}
\|\dot{\alpha}_m - \dot{u}\|_{L^2} &\leq \left|1 - \frac{1}{\sqrt{1+c_m}}\right| \|\dot{\beta}_m\|_{L^2} + \|\dot{\beta}_m - \dot{u}\|_{L^2} \\
&\leq \left|1 - \frac{1}{\sqrt{1+c_m}}\right| \left(\|\dot{u}\|_{L^2} + \|\dot{\beta}_m - \dot{u}\|_{L^2}\right) + \|\dot{\beta}_m - \dot{u}\|_{L^2}
\end{aligned}$$

and so it remains to bound c_m and $\|\dot{\beta}_m - \dot{u}\|_{L^2}$.

- By definition, $1 + c_m := \int_0^1 \left\langle -J\dot{\beta}_m, \beta_m \right\rangle dt$, and so:

$$\begin{aligned} 1 + c_m &= \int_0^1 \left\langle -J\dot{\beta}_m, \beta_m \right\rangle dt = \int_0^1 \langle -J\dot{u}, u\rangle\, dt + \int_0^1 \left\langle -J(\dot{\beta}_m - \dot{u}), u \right\rangle dt \\ &+ \int_0^1 \langle -J\dot{u}, (\beta_m - u)\rangle\, dt + \int_0^1 \left\langle -J(\dot{\beta}_m - \dot{u}), (\beta_m - u) \right\rangle dt, \end{aligned}$$

therefore,

$$c_m = \int_0^1 \left\langle -J(\dot{\beta}_m - \dot{u}), u \right\rangle dt + \int_0^1 \langle -J\dot{u}, (\beta_m - u)\rangle\, dt + \int_0^1 \left\langle -J(\dot{\beta}_m - \dot{u}), (\beta_m - u) \right\rangle dt.$$

Using the Cauchy-Schwartz inequality we get,

$$|c_m| \le \|\dot{\beta}_m - \dot{u}\|_{L^2} \cdot \|u\|_{L^2} + \|\dot{u}\|_{L^2} \cdot \|\beta_m - u\|_{L^2} + \|\dot{\beta}_m - \dot{u}\|_{L^2} \cdot \|\beta_m - u\|_{L^2},$$

which, after applying the Poincare inequality (without loss of generality, we may assume that the average of u is 0 and so is the average of $\beta_m - u$) reduces to:

$$|c_m| \le \frac{1}{\pi}\left(2\|\dot{u}\|_{L^2} \cdot \|\dot{\beta}_m - \dot{u}\|_{L^2} + \|\dot{\beta}_m - \dot{u}\|^2_{L^2}\right).$$

- Bounding $\|\dot{\beta}_m - \dot{u}\|_{L^2}$:

$$\|\dot{\beta}_m - \dot{u}\|^2_{L^2} = \int_0^1 (\dot{\beta}_m(t) - \dot{u}(t))^2 dt = \sum_{k=0}^{m-1} \int_{I_k} \left(⨍_{I_k} \dot{u}(s)ds - \dot{u}(t) \right)^2 dt$$

Using the Poincare inequality for the L^2-norm (or, equivalently, the variational principle), we get:

$$\begin{aligned} \|\dot{\beta}_m - \dot{u}\|^2_{L^2} &= \sum_{k=0}^{m-1} \int_{I_k} \left(⨍_{I_k} \dot{u}(s)ds - \dot{u}(t) \right)^2 dt \\ &\le \sum_{k=0}^{m-1} \frac{1}{\pi^2 m^2} \int_{I_k} \ddot{u}^2(t)dt \le \sum_{k=0}^{m-1} \frac{1}{\pi^2 m^2} \|\ddot{u}(t)\|^2_{L^2(S^1, \mathbb{R}^{2n})} \\ &= \frac{1}{\pi^2 m} \|\ddot{u}(t)\|^2_{L^2(S^1, \mathbb{R}^{2n})}. \end{aligned}$$

In order to bound the second derivative, $\ddot{u}$, we can derivate the Hamilton equation by t:

$$\ddot{u}(t) = \boldsymbol{H}(H) \circ u(t) \cdot \dot{u}(t)$$

which yields:

$$\|\dot{\beta}_m - \dot{u}\|^2_{L^2} \le \frac{1}{\pi^2 m} \|\boldsymbol{H}(H) \circ u(t) \cdot \dot{u}(t)\|_{L^2} \le \frac{\lambda}{\pi^2 m} \|\boldsymbol{H}(H)\|_{C^0} \cdot \|\nabla H\|_{C^0}.$$

Concluding the above, we obtain the following bound on the error of the discrete approximation:

$$|\mathcal{F}(u_m) - \mathcal{F}(u)| \le L(h_K^2) \cdot \frac{\lambda^2}{\pi^2 m} \left(1 + \frac{1}{\pi^2 m} \|\boldsymbol{H}(H)\|_{C^0}\right) \cdot \|\boldsymbol{H}(H)\|_{C^0} \cdot \|\nabla H\|^2_{C^0}.$$

$\square$

TABLE 1. Vertices and normal vectors for the test simplex Δ. We denote by F_i the facet opposite to the vertex v_i.

v_0	$(0,0,0,0)$	n_0	$(0.4827, 0.5310, 0.4397, 0.5401)$
v_1	$(1.1, 0, -.1, 0.1)$	n_1	$(-0.9951, 0.0098, 0.0985, 0.0009)$
v_2	$(0, 1, -0.1, .1)$	n_2	$(0.0170, -0.9910, 0.0974, -0.0901)$
v_3	$(0.1, 0.1, 1, 0)$	n_3	$(-0.0893, -0.0982, -0.9911, -0.0089)$
v_4	$(0, -0.1, 0, 1.1)$	n_4	$(0.0884, 0.0973, -0.0186, -0.9912)$

8. Tests and Experimental Data

8.1. **Behaviour of Generalized Characteristics.** In [?yaron] a minimizing generalized closed characteristic was constructed for the standard simplex

$$\Delta_{STD} = \text{hull}\{0, e_1, \ldots, e_{2n}\}.$$

It is proven directly that the piece-wise linear path $\gamma(t)$ which travels from 0 to $(1/n, \ldots, 1/n, 0, \ldots, 0)$ with tangent vector $\sum_{i=1}^{n} e_i$, then travels to $(0, \ldots, 0, 1/n, \ldots, 1/n)$ with tangent vector $\sum_{i=n+1}^{2n} e_i - \sum_{i=1}^{n} e_i$, then travels to 0 with tangent vector $-\sum_{i=n+1}^{2n} e_i$, is a minimizing generalized characteristic for Δ_{STD}. This characteristic is special because the first and last thirds of the path are spent in codimension > 2 facets, and it seems that this is due to the symmetry of the standard simplex. One might suspect that if you perform a sufficiently generic perturbation of the standard simplex, then the minimizing characteristics will travel in the facets where their dynamics are determined by the characteristic foliation of the facets.

We designed a program called `prSimplex.m` to study the behaviour of action minimizing generalized characteristics in simplexes. First, one runs the minimization algorithm for a simplex Δ and obtains a piece-wise linear path $(z_m)_{j=1}^{m}$ that approximates the minimizing characteristic (this is currently the reconstructed path). If a minimizing characteristic travels through the interior of a facet F_i, then the tangent vector to the characteristic is in the same direction as the characteristic foliation vector Jn_i, where n_i is a unit normal vector for the facet F_i. The program compares the tangent vectors $(\dot{z}_j)_{j=1}^{m}$ from the approximation with the vectors Jn_i and outputs a string of facet numbers $(a_j)_{j=1}^{n}$ where $a_j \in \{0, \ldots, 2n\}$ is the number of the facet which maximizes the quantity

$$\langle Jn_i, \dot{z}_j \rangle.$$

Assuming that z travels entirely in the facets of Δ, this program tells us symbolically the order in which each facet is visited by z.

Let $\Delta \subset \mathbb{R}^4$ be the simplex with vertices in Table 1. With $m = 120$, the minimization algorithm (as on github on August 1) computes $c = 0.2489$. The `prSimplex.m` algorithm gives very clean output for the calculated characteristic. The order the faces are visited is

$$2 \to 0 \to 4 \to 3 \to 1 \to 2$$

With the exception of a few segments of the path z that occur at the transition between facets, the error angle between the tangent vectors $\dot{z}_m$ and the Jn_i chosen by the program is very small (on the order of 0.001 or smaller).

9. Convex Functions and Sets

First, lets recall some basic facts about convex functions and Legendre transforms, which can be found in [?amann]. A function $H : \mathbb{R}^{2n} \to \mathbb{R}$ is *convex* if for all x, y,

$$H(tx + (1-t)y) \geq tH(x) + (1-t)H(y)$$

and H is strictly convex if this inequality is strict.

Proposition 2. *Let $H \in C^2(\mathbb{R}^{2n}, \mathbb{R})$. Then H is (strictly) convex if and only if the Hessian $D^2H(x)$ is positive (definite) semi-definite for all $x \in \mathbb{R}^{2n}$.*

The Hessian D^2H is *uniformly* positive definite if there is a constant $\alpha > 0$ such that

$$\langle D^2H(x)y, y\rangle \geq \alpha|y|^2, \forall x, y \in \mathbb{R}^{2n}$$

Proposition 3. *Let $H \in C^2(\mathbb{R}^{2n}, \mathbb{R})$. If D^2H is uniformly positive definite then the gradient $\nabla H : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is bijective.*

The *Legendre transform* of a function $H : \mathbb{R}^{2n} \to \mathbb{R}$ is defined as

$$\mathcal{L}H(y) = \sup_{x \in \mathbb{R}^{2n}} \{\langle y, x\rangle - H(x)\}.$$

Proposition 4. *Let $H \in C^2(\mathbb{R}^{2n}, \mathbb{R})$ and assume that D^2H is uniformly positive definite. Then*

1. *$\mathcal{L}H(y) = \langle y, x\rangle - H(x)$ for $x = (\nabla H)^{-1}(y)$.*
2. *$\mathcal{L}H \in C^2(\mathbb{R}^{2n}, \mathbb{R})$, $\mathcal{L}H$ is strictly convex, and $\nabla(\mathcal{L}H) = (\nabla H)^{-1}$.*
3. *$H(x) + \mathcal{L}H(y) \geq \langle x, y\rangle$, for all $x, y \in \mathbb{R}^{2n}$ with equality if and only if $x = y$.*
4. *$\mathcal{L}^2H = H$.*

Now we follow [?hofer-zhender]. Suppose $K \subset \mathbb{R}^{2n}$ is a compact, convex set whose interior contains the origin. Recall that we denoted the *semi-norm* with respect to the convex body K by

$$||x||_K = \inf\left\{r > 0 : \frac{x}{r} \in K\right\}, \forall x \neq 0\ ,$$

and by definition,

$$K = \left\{x \in \mathbb{R}^{2n} : ||x||_K \leq 1\right\}.$$

Note that $H(x) = ||x||_K^2$ is homogeneous of degree two (i.e. $H(\alpha x) = \alpha^2 H(x)$ for all $\alpha > 0$) and has quadratic growth (i.e. there exist constants $c_1, c_2 > 0$ such that

$$c_1||x||^2 \leq H(x) \leq c_2||x||^2).$$

Proposition 5. *If K is strictly convex and ∂K is smooth of class C^2, then the function $H(x) = ||x||_K^2$ satisfies the following properties:*

1. *$H \in C^2(\mathbb{R}^{2n} \setminus \{0\}, \mathbb{R})$.*
2. *D^2H is uniformly positive definite (in particular, H is strictly convex).*

Proof. (1)

(2)

□

Given a (strictly) convex set K that contains the origin, its *dual* is the (strictly) convex set

$$K^\circ = \{y \in \mathbb{R}^{2n} :\ \langle x, y\rangle \leq 1\, \forall x \in K\}.$$

Observe that $(K^\circ)^\circ$.

Proposition 6. *K is C^2 and strictly convex if and only if K° is C^2 and strictly convex.*

In this paper, we often consider the supporting function

$$h_K(x) = \sup_{y \in K} \{\langle x, y\rangle\}.$$

Observe that $K^\circ = \{x : h_K(x) \le 1\}$ and so $h_K(x) = ||x||_{K^\circ}$. If $K + K'$ is the Minkowski sum of two convex sets then it follows from the definition that

$$h_{K+K'} = h_K + h_{K'}$$

Proposition 7. *If K is a convex set then*

$$\mathcal{L} h_K^2 = \frac{1}{4} h_{K^\circ}^2$$

9.1. **Smoothed Dual Body Approach.** The following is one approach to running the algorithm with a smoothing of a convex polytope.

Suppose that K is a convex polytope and we want to run the algorithm for the smoothed body $K_\epsilon = (K^\circ + B(\epsilon))^\circ$. Then we need to express the function $h_{K_\epsilon}^2$ as well as the gradient $\nabla h_{K_\epsilon}^2$ in a computable way.

$$h_{K_\epsilon}^2 = 4\mathcal{L}\left(h_{K_\epsilon^\circ}\right)^2 = 4\mathcal{L} h_{K^\circ + B(\epsilon)}^2 = 4\mathcal{L}\left(h_{K^\circ} + h_{B(\epsilon)}\right)^2.$$

The function h_{K° can be computed combinatorially, and the Legendre transform can be computed using a software package. Since we would not like to evaluate the gradient of this Legendre transform, we note that

$$\nabla h_{K_\epsilon}^2(u) = v \text{ if and only if } u = \nabla\left(h_{K_\epsilon^\circ}\right)^2(v)$$

so to solve for the gradient of $h_{K_\epsilon}^2$ at u, we can evaluate the minimization problem

$$v \text{ minimizes } |u - \nabla\left(h_{K_\epsilon^\circ}\right)^2(v)|.$$

Now

$$\nabla\left(h_{K_\epsilon^\circ}\right)^2 = \nabla\left(h_K + h_{B(\epsilon)}\right)^2 = \nabla h_K^2 + 2\nabla\left(h_K h_{B(\epsilon)}\right) + \nabla h_{B(\epsilon)}^2$$

And every expression in this equation can be evaluated piece-wise analytically.

References

[AAO] S. Artstein-Avidan and Y. Ostrover, *Bounds for minkowski billiard trajectories in convex bodies*, International Mathematics Research Notices **2014** (2014), no. 1, 165–193.

[GJ+] A. Göing-Jaeschke et al., *Parameter estimation and bessel processes in financial models and numerical analysis in hamiltonian dynamics*, Ph.D. Thesis, 1998.