FAST WEAK–KAM INTEGRATORS FOR SEPARABLE HAMILTONIAN SYSTEMS

ANNE BOUILLARD, ERWAN FAOU, AND MAXIME ZAVIDOVIQUE

Abstract. We consider a numerical scheme for Hamilton–Jacobi equations based on a direct discretization of the Lax–Oleinik semi–group. We prove that this method is convergent with respect to the time and space stepsizes provided the solution is Lipschitz, and give an error estimate. Moreover, we prove that the numerical scheme is a *geometric integrator* satisfying a discrete weak–KAM theorem which allows to control its long time behavior. Taking advantage of a fast algorithm for computing min–plus convolutions based on the decomposition of the function into concave and convex parts, we show that the numerical scheme can be implemented in a very efficient way.

1. Introduction

We consider Hamilton–Jacobi equations of the form

$$\partial_t U + H(t, x, \nabla U) = 0, \quad U(0, x) = u_0(x), \tag{1}$$

where $H(t, x, v)$ is a Hamiltonian function $H : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ that is **separable**, in the sense that we can write

$$H(t, x, p) = K(p) + V(t, x), \tag{2}$$

for some convex function K and some smooth and bounded function V. The typical cases of study we have in mind are the so called mechanical Hamiltonians, of the form

$$H(t, x, p) = \frac{1}{2}|p + P|^2 + V(t, x), \tag{3}$$

where $P \in \mathbb{R}^n$ is a given vector, $|v|^2 = v_1^2 + \cdots + v_n^2$ for $v = (v_1, \cdots, v_n) \in \mathbb{R}^n$, and where $V(t, x)$ is a suitably smooth and bounded function.

Since the pioneering works of Crandall and Lions [CL84] and Souganidis [Sou85], the study of numerical schemes for the Hamilton–Jacobi equation (1) has known many recent progresses, see for instance [LS95, LT01, Abg96, JX98, BJ05] and the references therein, and more specifically [OS88, OS91, JP00] for the popular (weighted) essentially nonoscillatory (ENO and WENO) methods which are now widely commonly used in many application fields. Let us finally mention a recent work by Soga ([Sog13]) which deals with situations similar to the ones tackled in this paper.

Following a different approach, more in the spirit of [FF02, Ror06] (or [AGL08] in an optimal control setting), the main aim of this paper is to

1991 *Mathematics Subject Classification.* 35F21, 65M12, 06F05 .

Key words and phrases. Hamilton–Jacobi equations, weak–KAM theorem, Geometric integration, Min–plus convolution.

show how a direct discretization of the Lax–Oleinik representation of the viscosity solution of (1) allows to define a new fast algorithm for computing $u(t,x)$ possessing strong geometrical properties allowing to control its long time behavior and obtain error estimates when the solution is Lipschitz.

Let us recall, see [Lio82, Fat05], that under some assumptions on H (smoothness, uniform superlinearity and strict convexity over the fibers, see Section 2 below), we can write

$$U(t,x) = \inf_{\gamma(t)=x} \Big[u_0\big(\gamma(0)\big) + \int_0^t L\big(s,\gamma(s),\dot\gamma(s)\big)\mathrm{d}s \Big], \tag{4}$$

where the infimum is taken over all absolutely continuous curves $\gamma : [0,t] \to \mathbb{R}^n$ such that $\gamma(t) = x$, and where $L(t,x,v)$ is the Lagrangian associated with H. The idea of this paper is to discretize directly (4) on a space–time grid, by replacing the set of curves γ by the set of piecewise linear (or piecewise constant) curves across the space grid points.

We first prove that such an approximation is convergent with respect to the size of the space and time stepsizes, and under an anti–CFL condition (namely that the ratio between the space and time stepsize should be small). We give an error estimate under the assumption that u_0 is Lipschitz.

Moreover, this numerical integrator turns out to be a *geometric integrator* (see for instance [HLW06, LR04]) in the sense that it respects the long time behavior of the exact solution $U(t,x)$. Let us recall that in the case of periodic Hamiltonians (both in time and space variables), the weak–KAM theorem (see [Fat05, CISM00]) shows the existence of a constant $\overline{H}$ such that

$$\frac{1}{t} U(t,x) \to \overline{H} \quad \text{when} \quad t \to +\infty.$$

Here, using a discrete weak–KAM theorem, see [BB07, Zav12], we prove that the numerical scheme possesses the same long time property, with a constant that is close to the exact constant $\overline{H}$.

Finally, we show that in the separable case mainly considered in this paper (see (2) below), the discrete version of (4) is a min–plus convolution that can be approximated using a fast algorithm with $\mathcal{O}(N)$ operations in many situations if N is the number of grid points. This algorithm uses the decomposition of u into concave and convex parts. Moreover, it easily extends to any space dimension n using a splitting strategy, when the kinetic part of the Hamiltonian is separable – see Remark 2.6 – which includes the case (3).

We then conclude by numerical simulations in dimension 1 to illustrate the good behavior of our algorithm, as well as its very low cost in general situations.

The paper is divided into three parts: in a first part (Section 2) we give a convergence result over a finite time interval of the form $[0,T]$ where T is fixed. In a second part (Section 3), we consider the case where the Hamiltonian is periodic in time t and x. In this case, we can derive explicitly the dependence on T in the error estimates, and prove a weak–KAM theorem for the numerical scheme which gives informations concerning the long time behavior of the scheme. In the third part (Section 4), we describe the implementation of the method based on a fast algorithm to compute min–plus convolutions. We conclude this part by showing numerical simulations.

Acknowledgement. This work owes a lot to Vincent Calvez, who put the authors in touch and took part to preliminary discussions. It is a great pleasure to thank him a lot. We also would like to thank Vinh Nguyen for careful reading through previous versions of the paper. The last author would like to thank Antonio Siconolfi for bringing him to this subject. Finally, we thank the anonymous referees for very helpful remarks on improving the content and presentation of this manuscript.

The second author is supported by the ERC starting grant GEOPARDI.

2. Description of the scheme and convergence results

2.1. Hypotheses.

Recall that we consider a separable Hamiltonian $H(t, x, p)$ of the form (2). With this Hamiltonian we can associate by Legendre transform the Lagrangian

$$L(t, x, v) = \sup_{p\in\mathbb{R}^n} \Big(p \cdot v - H(t, x, p) \Big).$$

In our case, we get

$$L(t, x, v) = K^*(v) - V(t, x),$$

where $K^*(v)$ is the Legendre transform of K. For instance in the special case (3) we have

$$L(t, x, v) = \frac{1}{2}|v|^2 - P \cdot v - V(t, x).$$

We make the following assumptions on K and V:

Hypothesis (i). *The function $K^* \in \mathcal{C}^2(\mathbb{R}^n)$ is uniformly strictly convex in the sense that there exists a constant $c > 0$ such that for all $Y \in \mathbb{R}^n$, and for all $v \in \mathbb{R}^n$,*

$$\frac{\partial^2 K^*}{\partial v^2}(v)(Y, Y) \geqslant c|Y|^2 . \tag{5}$$

Hypothesis (ii). *The function $V(t, x) \in \mathcal{C}^2(\mathbb{R} \times \mathbb{R}^n)$ is such that there exists a constant B such that for $j + q \leqslant 2$, and all $(t, x) \in \mathbb{R} \times \mathbb{R}^n$,*

$$|\partial_t^j \partial_x^q V(t, x)| \leqslant B, \tag{6}$$

where $|\cdot|$ denote the norm of differential operators acting on $\mathbb{R} \times \mathbb{R}^n$.

Note that the bound (6) is straightforward under the additional assumption that $V(t, x, v)$ is *periodic* in (t, x), the case studied in the next section.

Remark 2.1. The previous hypotheses imply that the Hamiltonian H and the Lagrangian L are $\mathcal{C}^2$, convex and superlinear in respectively p and v:

$$\forall\, k > 0, \quad \forall\, t > 0, \quad \exists\, A(k) < \infty, \quad L(t, x, v) \geqslant k|v| - A(k). \tag{7}$$

Under these assumptions, the viscosity solution of (1) can be represented by the formula: for all $t, \delta > 0$,

$$(8)\quad \forall x \in \mathbb{R}^n, \quad U(t+\delta, x) = \inf_{\gamma(t+\delta)=x} \Big[U\big(t, \gamma(t)\big) + \int_t^{t+\delta} L\big(s, \gamma(s), \dot\gamma(s)\big)\mathrm{d}s\Big],$$

where the infimum is taken on all absolutely continuous curves $\gamma : (t, t+\delta) \to \mathbb{R}^n$ verifying $\gamma(t+\delta) = x$, see [Lio82, Fat05]. We will later on use the notation $U(t+\delta, x) := T_t^\delta u^t(x)$ where $u^t = U(t, \cdot)$. Moreover, the infimum is achieved on a curve $\gamma_{t,x}^\delta(s)$ that is $\mathcal{C}^2$ and satisfies the Euler–Lagrange equation

$$(9)\qquad \frac{\mathrm{d}}{\mathrm{d}s}\frac{\partial L}{\partial v}\big(s, \gamma(s), \dot\gamma(s)\big) = \frac{\partial L}{\partial x}\big(s, \gamma(s), \dot\gamma(s)\big).$$

The notation T_t^δ defines the Lax–Oleinik semi–group. In particular, we have $T_{t+\delta}^\sigma \circ T_t^\delta = T_t^{\delta+\sigma}$ for nonnegative δ and σ. With these assumptions, we have the following proposition.

Proposition 2.2. *For all $T > 0$, and for all $R > 0$, there exists $M(R,T)$ such that for all $x, y \in \mathbb{R}^n$ satisfying $|x - y| \leqslant R$ and for all $t \in \mathbb{R}$, then every solution of the Euler–Lagrange equation (9) minimizing the action*

$$(10)\qquad \int_t^{t+T} L\big(s, \gamma(s), \dot\gamma(s)\big)\mathrm{d}s,$$

with fixed endpoints $\gamma(t) = x$ and $\gamma(t+T) = y$, satisfies $|\dot\gamma(s)| \leqslant M(R,T)$ for all $s \in [t, t+T]$.

Proof. The Euler–Lagrange equation is written

$$\frac{\partial^2 K^*}{\partial v^2}\big(\dot\gamma(s)\big)\big(\ddot\gamma(s)\big) = -\frac{\partial V}{\partial x}\big(s, \gamma(s)\big).$$

Using the uniform strict convexity of K^* and the fact that $\partial_x V$ is uniformly bounded, there exists a constant C depending only on T and K, such that

$$(11)\qquad \forall\, s \in [t, t+T] \quad |\ddot\gamma(s)|^2 \leqslant C.$$

This implies that for all $s \in [t, t+T]$,

$$(12)\qquad |\dot\gamma(s) - \dot\gamma(t)| \leqslant \int_t^{t+T} |\ddot\gamma(s)|\,\mathrm{d}s \leqslant T\sqrt{C}.$$

Now as γ minimizes the action between t and $t+T$, comparing with the trivial curve $t \mapsto x + t(y-x)/T$ from x to y, we get

$$\begin{aligned}\int_t^{t+T} L\big(s, \gamma(s), \dot\gamma(s)\big)\mathrm{d}s &\leqslant \int_t^{t+T} \Big[K^*\big(\frac{y-x}{T}\big) - V\big(s, x + \frac{t}{T}(y-x)\big)\Big]\mathrm{d}s \\ &\leqslant TD(R/T) + TB,\end{aligned}$$

where $D(M) = \sup_{|v|\leqslant M} |K^*(v)|$, and B is given by (6). By superlinearity, we deduce that

$$\int_t^{t+T} |\dot\gamma(s)|\,\mathrm{d}s \leqslant TD(R/T) + TA(1) + TB,$$

where $A(1)$ is given by (7). Using (12), we thus obtain

$$T|\dot\gamma(t)| \leqslant \int_t^{t+T} |\dot\gamma(s)|\mathrm{d}s + T^2\sqrt{C} \leqslant TD(R/T) + TA(1) + TB + T^2\sqrt{C}.$$

This shows that $|\dot\gamma(t)|$ is bounded, and hence using again (12) that $|\dot\gamma(s)|$ is bounded for all s, with a constant depending only on T, R and the constants appearing in (5) and (6). □

Remark 2.3. The previous lemma is one of the main keys in the proof of the convergence of our schemes (compare with [Fal87, FF02]). Here, it is established thanks to the particular form of H, but it can be noted that it remains valid under other technical assumptions (for example if H is autonomous and Tonelli as established in [Fat05, FM07], or if it is Tonelli and periodic both in the space and in the time variable as proven in [Mat91, CISM00, Itu96]). Actually, in these cases, it can be established that $M(R,T)$ only depends on the ratio R/T. We will come back on these matters in Section 3 and give a proof of this result in the Appendix. Therefore, this section and the next would still be valid for general Hamiltonians chosen in these classes, however the convolution techniques of Section 4 would fail.

Finally, a straightforward consequence of Equation (11) is that the Euler–Lagrange flow of L is complete.

2.2. **A first discrete semi–group.**

For a given $h>0$ we define the h–grid $G_h = h\mathbb{Z}^n$ endowed with the metric induced by the euclidian metric on $\mathbb{R}^n$. For a given continuous function u, we define $u|_{G_h} : G_h \to \mathbb{R}$ its restriction to the grid G_h.

A first idea to discretize the Lax–Oleinik semi–group is as follows. Given $t, \tau > 0$, let us define $c^\tau_{t,h} : G_h \times G_h \to \mathbb{R}$ as follows:

$$\forall (x,y) \in G_h\times G_h, \quad c^\tau_{t,h}(x,y) = \int_t^{t+\tau} L\Big(s, x+(s-t)\frac{y-x}{\tau}, \frac{y-x}{\tau}\Big)\mathrm{d}s. \tag{13}$$

Let us introduce the following discrete Lax–Oleinik semi–group: if $u : G_h \to \mathbb{R}$ is any function, we set

$$\forall x \in G_h, \quad T^\tau_{t,h}u(x) = \inf_{y\in G_h} \Big[u(y) + c^\tau_{t,h}(y,x)\Big].$$

For a given integer N, we may define

$$T^{N\tau}_{t,h}u = T^\tau_{t_{N-1},h} \circ \cdots \circ T^\tau_{t_1,h} \circ T^\tau_{t_0,h}u \tag{14}$$

the composition of N times the discrete semi–group $T^\tau_{t,h}$, where for all $i = 1,\dots,N-1$, $t_i = t + i\tau$. With these notations, $t = t_0$ is the initial time at which we launch the semi–group, while $t_N = t + N\tau$ is the final time at which we compute the approximate solution.

One of the nice features of this discretization is the following property: consider $u : \mathbb{R}^n \mapsto \mathbb{R}$, $N \geqslant 1$ an integer, $t \in \mathbb{R}$ and $\tau > 0$. Then

$$(T^{N\tau}_t u)|_{G_h} \leqslant T^{N\tau}_{t,h}(u|_{G_h}). \tag{15}$$

Indeed this semi–group consists in taking an infimum over a smaller set of curves, compared to the Lax–Oleinik semi–group.

2.3. **Fully discrete semi–group.** The main disadvantage is that to compute this cost, a quadrature rule in time has to be used. In this subsection, we prove how the Euler approximation of this integral yields a convergent scheme which still satisfies a weak–KAM theorem similar to Proposition 3.1 under suitable periodicity assumptions.

For a given $h > 0$ and $\tau > 0$, we define the following cost function:

$$\forall (x, y) \in G_h \times G_h, \quad \kappa^{\tau}_{t,h}(y, x) = \tau L\Big(t, x, \frac{x-y}{\tau}\Big), \tag{16}$$

and we introduce the associated fully discrete Lax–Oleinik semi–group, which acts on any function $u : G_h \to \mathbb{R}$ as follows:

$$\forall x \in G_h, \quad \mathcal{T}^{\tau}_{t,h} u(x) = \inf_{y \in G_h} \Big[u(y) + \kappa^{\tau}_{t,h}(y, x)\Big]. \tag{17}$$

Using the explicit expression of L, we can rewrite this fully–discrete semi–group as

$$\forall x \in G_h, \quad \mathcal{T}^{\tau}_{t,h} u(x) = \inf_{y \in G_h} \Big[u(y) + \tau K^*\big(\frac{x-y}{\tau}\big)\Big] - \tau V(t, x), \tag{18}$$

involving the (min,plus)–convolution of u and K^*.

Remark 2.4. This scheme is a particular case of the so-called semi–Lagrangian schemes. Indeed, those are of the form

$$U(t + \tau, x) = \inf_{\alpha \in \mathcal{A}} \Big[U(t, x - h\alpha) + \tau L(x, \alpha)\Big],$$

where $\mathcal{A}$ is a given (usually compact) set of controls. Here, we take $\alpha = \frac{x-y}{\tau}$. Note that the main feature of this choice is that it enables y to remain on the grid, whereas standard semi-Lagrangian methods rely on an interpolation procedure at each step. This particular form allows to use the fast convolution techniques of Section 4.

Remark 2.5. We can interpret this scheme as a discretization of the splitting scheme (see for instance [JKR01]) with time step τ based on the decomposition

$$\partial_t U(t, x) + K\big(\nabla U(t, x)\big) = 0, \quad \text{and} \quad \partial_t U(t, x) + V(t, x) = 0,$$

where the first part is integrated using the method described in the previous section.

Remark 2.6. In dimension $n > 1$, if we assume that for $p = (p_1, \dots, p_n) \in \mathbb{R}^n$, $K(p) = K_1(p_1) + \cdots + K_n(p_n)$ with convex Hamiltonian functions K_i^*, $i = 1, \dots, n$, satisfying all hypotheses ***(i)***, ***(ii)*** on $\mathbb{R}$, then we immediately see that for a given function $u(x) = u(x_1, \dots, x_n)$, with $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, we have

$$\inf_{y\in G_h}\Big[u(y)+\tau K^*(\frac{x-y}{\tau})\Big]=$$
$$\inf_{y_n\in G_h^n}\Bigg[\tau K_n^*(\frac{x_n-y_n}{\tau})+\Big[\inf_{y_{n-1}\in G_h^{n-1}}\tau K_{n-1}^*(\frac{x_{n-1}-y_{n-1}}{\tau})+\cdots$$
$$+\Big[\inf_{y_1\in G_h^1}\tau K_n^*(\frac{x_1-y_1}{\tau})+u(y_1,\dots,y_n)\Big]\cdots\Big]\Bigg]=$$
$$\mathcal{T}_{t,h}^{\tau,1}\circ\cdots\circ\mathcal{T}_{t,h}^{\tau,n}u(x),$$

where we have decomposed $G_h=G_h^1\times\cdots\times G_h^n$ and where

$$\forall i\in[1,n],\quad \mathcal{T}_{t,h}^{\tau,i}u(x)=\inf_{y_i\in G_h^i}\Big[\tau K_i^*(\frac{x_i-y_i}{\tau})+u(x_1,\dots,x_{i-1},y_i,x_{i+1},\dots x_n)\Big].$$

This formula is essentially due to the fact that the Hamiltonians K_i commute, i.e., satisfy $\{K_i,K_j\}=0$ for $(i,j)\in\{1,\dots,n\}^2$ which ensures that the flows of $\partial_t U=K_i(\nabla U)$, $i=1,\dots,n$ commute. In this case, this allows to reduce the computation of the minimum over the n dimensional grid G_h to n minimization problems over the one–dimensional grids G_h^i.

For a given integer N, and a given function $u^0:G_h\to\mathbb{R}$ we define – compare (14)

$$(19)\qquad \forall\, x\in G_h,\quad u^N(x):=\mathcal{T}_{t,h}^{N\tau}u(x)=\mathcal{T}_{t_{N-1},h}^{\tau}\circ\cdots\circ\mathcal{T}_{t_1,h}^{\tau}\circ\mathcal{T}_{t_0,h}^{\tau}u^0(x)$$

where $t_i=t+i\tau$. When no confusion is possible, we will also refer to $\mathcal{T}^N$ instead of $\mathcal{T}_{t,h}^{N\tau}$. Note that with these notations, an estimate of the form (15) is no longer valid.

Proposition 2.7. *The fully discrete Lax–Oleinik semi–group verifies the following properties:*

(i) **Monotonicity***: if $u(x)\leqslant v(x)$ for all $x\in G_h$, then $\mathcal{T}^N u\leqslant\mathcal{T}^N v$;*
(ii) **Translation invariance***: if u is any given function and $\alpha\in\mathbb{R}$ is a constant, then $\mathcal{T}^N(u+\alpha)=\mathcal{T}^N u+\alpha$;*
(iii) **Consistency***: if φ is a smooth and compactly supported function then*

$$\forall(t,x)\in\mathbb{R}\times\mathbb{R}^n,\qquad \lim_{\substack{y\to x\\ h/\tau\to 0\\ \tau\to 0}}\frac{\varphi(y)-\mathcal{T}_{t,h}^{\tau}\varphi(y)}{\tau}=H\big(t,x,\nabla\varphi(x)\big).$$

Proof. The first two points are transparent from the definition. We only sketch the proof of the last point since it is standard to semi–Lagrangian schemes.

Assume h/τ and τ to be very small and $y\in G_h$ very close to x. By "differentiating" the definition of the semi–group (17 and 16) in the definition of $\mathcal{T}_{t,h}^{\tau}\varphi(y)$, we see that the infimum is reached at a point z such that $\nabla\varphi(z)\simeq\partial_v L\big(t,y,\frac{y-z}{\tau}\big)$. Moreover, $z\to x$ as $\tau\to 0$ (because L is super–linear). Recalling that

$$H\Big(t,y,\partial_v L\big(t,y,\frac{y-z}{\tau}\big)\Big)+L\big(t,y,\frac{y-z}{\tau}\big)=\partial_v L\big(t,y,\frac{y-z}{\tau}\big)\cdot\frac{y-z}{\tau},$$

we now compute

$$\frac{\varphi(y)-\mathcal{T}^{\tau}_{t,h}\varphi(y)}{\tau} \simeq \frac{\varphi(y)-\varphi(z)}{\tau} - L\big(t,y,\frac{y-z}{\tau}\big) \simeq$$
$$\nabla\varphi(y)\cdot\frac{y-z}{\tau} - \partial_v L\big(t,y,\frac{y-z}{\tau}\big)\cdot\frac{y-z}{\tau} + H\Big(t,x,\partial_v L\big(t,y,\frac{y-z}{\tau}\big)\Big) \simeq$$
$$H\big(t,y,\nabla\varphi(z)\big) \to H\big(t,x,\nabla\varphi(x)\big),$$

when $\tau\to 0$ and $h/\tau\to 0$. □

Finally, we have the following convergence result:

Theorem 2.8. *Let $T_0>0$, τ_0 and $h_0>0$ and $u_0:\mathbb{R}^n\to\mathbb{R}$ a bounded Lipschitz function. For a given $t_0\in\mathbb{R}$, let $U(t_0+t,x)=T^t_{t_0}u_0$ be the viscosity solution (4) of the Hamilton–Jacobi equation (1) such that $U(t_0,x)=u_0$. Then there exists a constant $M(T_0)$ such that for all $h>0$ and $\tau>0$ such that $h<h_0$, $\tau<\tau_0$ and the bound*

$$\frac{h}{\tau} < h_0 \tag{20}$$

are satisfied, then for all N verifying $N\tau\leqslant T_0$,

$$\forall\, x\in G_h,\quad |U(t_0+N\tau,x)-u^N(x)|\leqslant M(T_0)\big(\frac{h}{\tau}+\tau\big), \tag{21}$$

where the discrete solution $u^N:G_h\to\mathbb{R}$ is given by the formula (19) with initial value $u^0:=u_0|_{G_h}$.

Remark 2.9. The previous estimate is not new and not surprising. First, since the scheme is monotone, stable and consistent (see Proposition 2.7), its convergence is known to be automatic (see [BS91]). Note that in this case stability (meaning that approximate solutions can be controlled uniformly with respect to the mesh step sizes) is a consequence of monotonicity and translation invariance. Moreover, as we deal with a particular case of semi–Lagrangian scheme, such results are known (see [Fal87, FF02] and references therein), in particular the convergence in $\mathcal{O}(h/\tau)+\mathcal{O}(\tau)$ is common. For the sake of completeness, we will give a proof in Appendix B.

Remark 2.10. In the previous theorem, taking $\tau=\sqrt{h}$ obviously yields a convergence of order $\mathcal{O}(\sqrt{h})$. Finally, note that in the present setting, the constant M depends on T in an uncontrolled way. However, this dependence becomes linear under some extra periodicity assumptions, as we will see in the next section (Proposition 3.3).

3. Long time behavior in the periodic case

We now make the additional assumption that the potential function $V(t,x)$ is periodic.

Hypothesis (iii). *The function V is $\mathbb{Z}\times\mathbb{Z}^n$-periodic, in the sense that*

$$\forall (t,x)\in\mathbb{R}\times\mathbb{R}^n,\quad \forall (m,M)\in\mathbb{Z}\times\mathbb{Z}^n,\quad V(t,x)=V(t+m,x+M).$$

Note that under this hypothesis, the estimate (6) is automatically satisfied since we still assume that $V\in\mathcal{C}^2(\mathbb{R}\times\mathbb{R}^n)$.

3.1. **Weak–KAM theorem and a priori compactness.** In this periodic case, the weak–KAM theorem allows to study the long time behavior of the solution of (1) defined by the Lax–Oleinik semi–group.

Proposition 3.1. *Assume that hypotheses **(i)** and **(iii)** are satisfied. Then there exists a unique constant $\overline{H}$ for which the equation*

$$\forall t > 0, \quad U^*(t+1,\cdot) = T_t^1 U^*(t,\cdot) = \overline{H} + U^*(t,\cdot),$$

admits a $\mathbb{Z} \times \mathbb{Z}^n$–periodic continuous solution $U^* : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$.

Moreover, for any uniformly bounded $u : \mathbb{R}^n \to \mathbb{R}$, there exists a constant C_u such that

$$\forall\, t > 0, \quad |T_0^t u - t\overline{H}|_\infty \leqslant C_u.$$

Proof. This result is very standard and a complete proof can be found in [Fat05]. The existence of $\overline{H}$ and of u^* is exactly the content of the weak–KAM theorem (see [Fat97, Fat05] for the autonomous case, and [CISM00] for the time periodic case). The second assertion is a consequence of the fact that

$$|T_0^t u - T_0^t v|_\infty \leqslant |u - v|_\infty$$

for all continuous bounded functions u and v on $\mathbb{R}^n$, where $|\cdot|_\infty$ denotes the L^∞ norm on $\mathbb{R}^n$. □

The goal of this section is to prove that a similar result holds for the numerical scheme described in the previous section, under hypotheses ***(i)*** and ***(iii)***.

In order to study the long time behavior of the method in this case, we first give an a priori compactness result which refines the estimates given in Proposition 2.2. The following proposition is mainly due to Mather (see [Mat91] for the case of time periodic Lagrangians, or [Itu96, Lemma 7 and Corollary 8] for space periodic Lagrangians).

Proposition 3.2. *Assume that H satisfies hypotheses **(i)** and **(iii)**. For all $\Gamma > 0$, there exists a constant Γ' such that for any minimizer of the Lagrangian action $\gamma : [a,b] \to \mathbb{R}^n$ with $b-a \geqslant 1$ and $|\gamma(b)-\gamma(a)|/(b-a) \leqslant \Gamma$ then we have*

$$\forall t \in [a,b], \quad |\dot{\gamma}(t)| \leqslant \Gamma'.$$

In other words, the constant $M(R,T)$ of Proposition 2.2 can be chosen to be an increasing function of R/T.

For the sake of completeness, a proof of this proposition can be found in Appendix A. Note that most of the proof – essentially taken from [Mat91] – does not require the Hamiltonian to be periodic in space. In [Itu96] a similar result is proven which requires the Lagrangian to be periodic in space, but not in time any more.

3.2. **Convergence estimates in the periodic case.** Using the previous proposition, we can compute explicitly the time dependence of the constant $M(T)$ in the error estimates of Theorem 2.8, and prove that it depends linearly on time in the periodic case (see the Appendix B for a proof).

Theorem 3.3. *Let $T_1 > 1$, τ_0 and $h_0 > 0$ and $u_0 : \mathbb{R}^n \to \mathbb{R}$ a bounded Lipschitz function. For a given $t_0 \in \mathbb{R}$, let $U(t_0+t, x) = T^t_{t_0} u_0$ be the viscosity solution (4) of the Hamilton–Jacobi equation (1) such that $U(t_0, x) = u_0$. Then, there exists a constant M such that for all $h > 0$ and $\tau > 0$ such that $h < h_0$, $\tau < \tau_0$ and*

$$\frac{h}{\tau} < h_0, \tag{22}$$

and for all N satisfying $N\tau \geqslant T_1$,

$$\forall\, x \in G_h, \quad |U(t_0 + N\tau, x) - u^N(x)| \leqslant MN\tau\Big(\frac{h}{\tau} + \tau\Big). \tag{23}$$

where the discrete solution $u^N : G_h \to \mathbb{R}$ is given by the formula (19) with initial value $u^0 := u_0|_{G_h}$.

3.3. **Discrete weak–KAM theorem and effective Hamiltonian.** Recall that the function $u(t, x)$ is defined on $\mathbb{T}^1 \times \mathbb{T}^n = \big(\mathbb{R}/\mathbb{Z}\big) \times \big(\mathbb{R}^n/\mathbb{Z}^n\big)$. For convenience, we will only treat the cases of rational time and space discretizations: We set

$$\Lambda = \Big\{ \Big(\frac{1}{k}, \frac{1}{\ell}\Big) \mid (k, \ell) \in \mathbb{N}^* \times \mathbb{N}^* \Big\},$$

and in the sequel, we will only consider stepsizes $(h, \tau) \in \Lambda$. We then will denote by p the canonical projection from $\mathbb{R}^n$ to $\mathbb{T}^n$, and by $\widetilde{G}_h = G_h/\mathbb{Z}^n$ the quotiented grid, where G_h is the grid above, defined on $\mathbb{R}^n$.

Finally, we define a new cost function: For $(h, \tau) \in \Lambda$ and $t > 0$,

$$\forall\, (\tilde{x}, \tilde{y}) \in \widetilde{G}_h \times \widetilde{G}_h, \quad \tilde{\kappa}^\tau_{t,h}(\tilde{x}, \tilde{y}) = \inf_{\substack{p(x)=\tilde{x} \\ p(y)=\tilde{y}}} \kappa^\tau_{t,h}(x, y),$$

where $\kappa^\tau_{t,h}(x, y)$ is the fully discrete cost function defined in (16).

We then define the following semi–group: Let $\tilde{u} : \mathbb{T}^n \to \mathbb{R}$ and $(h, \tau) \in \Lambda$, then the fully discrete semi–group is

$$\forall\, \tilde{x} \in \widetilde{G}_h, \quad \widetilde{\mathcal{T}}^\tau_{t,h}\tilde{u}(\tilde{x}) = \inf_{\tilde{y} \in \widetilde{G}_h} \Big[\tilde{u}(\tilde{y}) + \tilde{\kappa}^\tau_{t,h}(\tilde{y}, \tilde{x})\Big]. \tag{24}$$

As in the previous section, we set

$$\widetilde{\mathcal{T}}^{N\tau}_{t,h} = \widetilde{\mathcal{T}}^\tau_{t_{N-1},h} \circ \cdots \circ \widetilde{\mathcal{T}}^\tau_{t_1,h} \circ \widetilde{\mathcal{T}}^\tau_{t_0,h},$$

where $t_i = t + i\tau$.

Remark 3.4. We leave to the reader the verification that, because two infimums commute, if $\tilde{u} : \mathbb{T}^n \to \mathbb{R}$ is a function and if $u : \mathbb{R}^n \to \mathbb{R}$ is its lift (which is then $\mathbb{Z}^n$-periodic), then the function $\mathcal{T}^{N\tau}_{t,h} u$ is $\mathbb{Z}^n$-periodic and canonically induces $\widetilde{\mathcal{T}}^{N\tau}_{t,h}\tilde{u}$ on $\mathbb{T}^n$. Hence the previous convergence result Theorem 3.3 can be read equivalently on $\mathbb{T}^n$ or on the space of $\mathbb{Z}^n$ periodic functions on $\mathbb{R}^n$. However, it is easier to take advantage of the compactness of $\mathbb{T}^n$. This is why we introduced these new costs, and defined them with infimums.

The discrete weak–KAM theorem can be used to better understand the approximate semi–groups applied for a period 1 of time and to obtain the following proposition.

Proposition 3.5. *For any $(h,\tau) = (1/k, 1/\ell) \in \Lambda$, there exists a unique constant $\overline{\mathcal{H}}_{h,\tau}$ such that there exists a function $v^*_{h,\tau} : \widetilde{G}_h \to \mathbb{R}$ verifying:*

$$\widetilde{\mathcal{T}}^{\ell\tau}_{0,h} v^*_{h,\tau} = \widetilde{\mathcal{T}}^{1}_{0,h} v^*_{h,\tau} = v^*_{h,\tau} + \overline{\mathcal{H}}_{h,\tau}.$$

Moreover, if u is any bounded initial datum on G_h at $t = 0$, then we have in L^∞

$$\frac{1}{N\tau}\mathcal{T}^{N\tau}_{0,h} u \longrightarrow \overline{\mathcal{H}}_{h,\tau},$$

as $N \to +\infty$, where $\mathcal{T}^{N\tau}_{0,h}$ is defined in (19).

Proof. The first part is just a reformulation of the discrete weak–KAM theorem (see for example the appendix of [Zav12] or [BB07]) while the second part is – as in the proof of Proposition 3.1 – a direct consequence of the fact that our approximation operators are weakly contracting for the infinity norm on bounded functions. □

Remark 3.6. Note that the previous proposition can be interpreted in the (min,plus) framework: Equation (24) is the (min,plus) product of a vector $\tilde{u}$ by the matrix $\tilde{\kappa}^{\tau}_{t\,h}$. Then, for $\tau = 1/\ell$, $\ell \in \mathbb{N}^*$, $\widetilde{\mathcal{T}}^{1}_{0,h}\tilde{u} = \widetilde{\mathcal{T}}^{\ell\tau}_{0,h}\tilde{u}$ is obtained by successive matrix multiplications with $\tilde{u}$. Hence there exists a matrix $C_{h,\tau}$ such that $\widetilde{\mathcal{T}}^{1}_{0,h}\tilde{u}(\tilde{x}) = \inf_{\tilde{y}\in\tilde{G}_h}\Big[\tilde{u}(\tilde{y}) + C_{h,\tau}(\tilde{y},\tilde{x})\Big]$, with $C_{h,\tau}(\tilde{y},\tilde{x}) < +\infty$ for all $\tilde{y}, \tilde{x}$. The matrix $C_{h,\tau}$ has a unique eigenvalue $\overline{\mathcal{H}}_{h,\tau}$, and $v^*_{h,\tau}$ is an eigenvector (see [BCOQ92] for details).

Let us recall that $\overline{H}$ is the effective Hamiltonian of H. It is obtained in homogenization theory by solving the cell problem ([LPV87]) and is also the constant found in the weak–KAM theorem (3.1). Using the refined convergence result obtained in Theorem 3.3, we can estimate the error between the effective Hamiltonian and the discrete effective Hamiltonian defined in Proposition 3.5.

Theorem 3.7. *With the notations of Theorem 3.3, let $(h,\tau) \in \Lambda$ be such that $h \leqslant h_0$, $\tau \leqslant \tau_0$ and $h/\tau \leqslant h_0$, then following inequality holds:*

$$\big|\overline{\mathcal{H}}_{h,\tau} - \overline{H}\big| \leqslant M\big(\frac{h}{\tau} + \tau\big),$$

where M is the constant coming from Theorem 3.3, and where $\mathcal{H}_{h,\tau}$ is defined in Proposition 3.5.

Proof. Start with a bounded and uniformly Lipschitz continuous function $u : \mathbb{R}^n \to \mathbb{R}$. By Theorem 3.3, the following inequality holds if $N\tau \geqslant T_0$ for some chosen $T_0 > 1$:

$$\forall\, x \in G_h, \quad |(T^{N\tau}_0 u)(x) - \mathcal{T}^N(u|_{G_h})(x)| \leqslant MN\tau\big(\frac{h}{\tau} + \tau\big), \tag{25}$$

where $\mathcal{T}^N = \mathcal{T}^{N\tau}_{0,h}$. Dividing by $N\tau$ and letting N go to ∞ yields that

$$\big|\overline{\mathcal{H}}_{h,\tau} - \overline{H}\big| \leqslant M\big(\frac{h}{\tau} + \tau\big).$$

□

Remark 3.8. One may wonder what is the behavior of the quantity $(T_0^{N\tau}u)|_{G_h} - \mathcal{T}^N(u|_{G_h})$ as $T = N\tau \to \infty$. The previous results show that it has a linear growth, of rate $\overline{H} - \overline{\mathcal{H}}_{h,\tau}$. Comparing with weak–KAM solutions yields that the second error term is always bounded. However, in some cases more can be said. Indeed, in the autonomous case (L independent of t) Fathi proved the convergence of the Lax–Oleinik semi–group ([Fat98]), that is, for any initial condition u there exists a weak–KAM solution u^* such that $(T_0^{N\tau}u) - N\tau\overline{H} \to u^*$ uniformly. Moreover, it can be proved that the iterated powers of a (min, +) matrix are periodic after a finite time. Therefore, for N large enough, the sequence $\mathcal{T}^N(u|_{G_h})$, is periodic after a certain time. In conclusion, in the autonomous case, one can write

$$(T_0^{N\tau}u)|_{G_h} - \mathcal{T}^N(u|_{G_h}) = N\tau(\overline{H} - \overline{\mathcal{H}}_{h,\tau}) + w_N,$$

where w_N is asymptotic to a periodic sequence.

4. Fast (min,plus)–convolution

As we have seen in (18), the numerical scheme considered in this paper involves the computation of the (min,plus)–convolution

$$\inf_{y\in G_h}\Big[u(y) + \tau K^*(\frac{x-y}{\tau})\Big], \quad x \in G_h.$$

In dimension 1, if the grid G_h is discretized by retaining k points only, the numerical cost is *a priori* of order k^2. We now present a fast (min,plus)–convolution algorithm that turns out to have a linear cost (i.e., proportional to k) in many situations.

In order to ease the presentation, we will not deal with functions defined on a grid, but on functions defined on finite and closed intervals.

Let $a, b \in \mathbb{R}$ with $a < b$. We write $f : [a, b] \to \mathbb{R}$ if f is such that

$$\begin{cases} f(x) < \infty & \text{if } x \in [a, b] \\ f(x) = \infty & \text{otherwise.} \end{cases}$$

For $f : [a, b] \to \mathbb{R}$, we say that f is respectively convex, concave, affine if $f_{|[a,b]}$ is respectively convex, concave, affine. That is, the functions we consider are defined on $\mathbb{R}$, finite on a closed interval and are said to inherit the properties they satisfy on this interval. Let $f : [a, b] \to \mathbb{R}$ and $g : [c, d] \to \mathbb{R}$. The (min,plus)–convolution (or convolution in the remaining of the paper) of f and g is defined for all $x \in \mathbb{R}$, by

$$f * g(x) = \inf_{y\in\mathbb{R}}\Big[f(y) + g(x-y)\Big]. \tag{26}$$

Recall that $f * g = g * f$. As f and g are finite only on an interval, it is easy to see that for all $x \in [a + c, b + d]$,

$$f * g(x) = \inf_{y\in[a,b]}\Big[f(y) + g(x-y)\Big],$$

and for all $x \notin [a + c, b + d]$, $f * g(x) = \infty$.

We will only consider piecewise affine functions and decompose them according to their affine components: there exist $a_0 = a < a_i < \cdots < a_n = b$ such that

$$f = \min_{i\in\{0,\ldots,n-1\}} f_i,$$

where $f_i : [a_i, a_{i+1}] \to \mathbb{R}$ is an affine function. For $i = 0, \ldots, n-1$, we denote by $f'_i = \big(f(a_{i+1}) - f(a_i)\big)/(a_{i+1} - a_i)$ the *slope* of f_i or the slope of f on $[a_i, a_{i+1}]$.

4.1. **Convolution.** The fast algorithm to compute the convolution (26) is based on a decomposition of g $(= u)$ in piecewise convex and concave functions. As the function f $\big(= K^*(\cdot/\tau)\big)$ considered will always be convex $\big($see equation (5)$\big)$, we thus see that we are led to compute separately the convolution of convex by convex functions, and concave by convex functions defined on finite intervals. As we will see, each block can be computed at a linear cost. In the end, the global cost of the algorithm thus depends on the number of convex and concave components on f, a number which might increase in the time evolution of the numerical solution of the Hamilton–Jacobi equation. We will come back later to this matter, but we emphasize that this procedure can be very easily implemented in parallel, each convolution block being computed independently.

We start with the following result, the proof of which can be found for example [BT08].

Lemma 4.1 (convolution of a convex function by an affine function)**.** *Let $f : [a, b] \to \mathbb{R}$ be a convex piecewise affine function and $g : [c, d] \to \mathbb{R}$ be an affine function of slope g'. Then $f * g : [a + c, b + d] \to \mathbb{R}$ is a convex piecewise affine function defined by*

$$f * g(x) = \begin{cases} f(x-c) + g(c) & \text{if} \quad a + c \leqslant x \leqslant \alpha + c, \\ f(\alpha) + g(x - \alpha) & \text{if} \quad \alpha + c < x \leqslant \alpha + d, \\ f(x-d) + g(d) & \text{if} \quad \alpha + d < x \leqslant b + d, \end{cases}$$

where $\alpha = \min\{a_i$ in the decomposition of $f \mid f'_i \geqslant g'\}$.

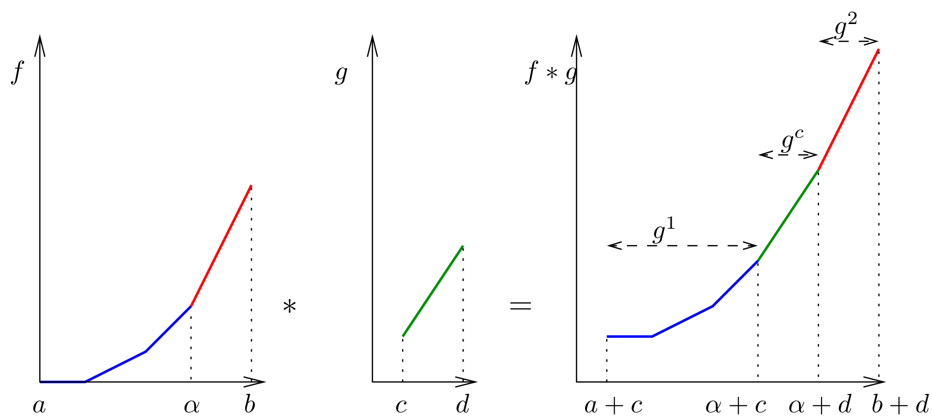

Figure 1. Convolution of a convex function by an affine function and decomposition into three functions.

Figure 1 illustrates this lemma. In the rest of the section, we will use a decomposition of such a convolution into three parts: $f * g = \min(g^1, g^c, g^2)$, where

(i) $g^1 = f * g_{|[c+a,c+\alpha]}$;
(ii) $g^c = f * g_{|[c+\alpha,d+\alpha]}$;

(iii) $g^2 = f * g_{|[d+\alpha,d+b]}$.

In other words, g^1 is composed of the segments of f whose slope is strictly less than that of g, g^c corresponds to the segment g and g^2 is composed of the segments of f whose slope is greater than or equal to that of g. Note that g^c is also concave.

A direct consequence of this lemma is Theorem 4.2, stated in [LBT01]. A complete proof is presented in [BJT08], but can also be deduced from previous works about the Legendre transform: the Legendre transform is an involution on the set of convex functions and can be computed in linear time (see [Luc97] for example). Moreover, the transform of the (min,plus)–convolution of two functions is the addition of the respective transforms of the two functions, inducing an alternative linear-time algorithm.

Theorem 4.2 (convolution of a convex function by a convex function)**.** *If f and g are convex and piecewise affine, then $f * g$ is obtained by putting end–to–end the different linear pieces of f and g sorted by increasing slopes.*

For sake of completeness, we give below Algorithm 1 for computing the (min,plus)–convolution of two convex piecewise affine functions without having to compute any Legendre transform.

Algorithm 1: Convolution of two convex functions

Data: $f : [0, n] \to \mathbb{R}$ a convex function with slopes (r_i), $g : [0, m] \to \mathbb{R}$ a convex function with slopes (ρ_i).
Result: $h = f * g$
begin
  $i \leftarrow 0$; $j \leftarrow 0$; $h(0) \leftarrow f(0) + g(0)$;
  while $i + j < n + m$ **do**
    if $i \neq n$ *and* $(r_i < \rho_j$ *or* $j = m)$ **then**
      $h(i + j + 1) \leftarrow h(i + j) + r_i$; $i \leftarrow i + 1$;
    else
      $h(i + j + 1) \leftarrow h(i + j) + \rho_j$; $j \leftarrow j + 1$;

We now turn to the case where f is convex and g is concave, and the Legendre tranform cannot help anymore to design an algorithm as in [Bre89, Luc97]. We begin with the following lemma:

Lemma 4.3. *Let $f : [a, b] \to \mathbb{R}$ be a convex piecewise affine function and $g : [c, d] \to \mathbb{R}$ be a concave piecewise affine function which decomposition is $g = \min_{j=1}^{m} g_j$. Then*

$$f * g = \min_{j=1}^{m} f * g_j.$$

Proof. This is a direct consequence of the distributivity of $*$ over the minimum. □

Now, consider two functions $f : [a, b] \to \mathbb{R}$, convex, and $g : [c, d] \to \mathbb{R}$, concave, with respective decompositions in $f_i : [a_i, a_{i+1}] \to \mathbb{R}$, $i = 0, \ldots, n-1$

and $g_j : [c_j, c_{j+1}] \to \mathbb{R}$, $j = 0, \dots, m-1$. The following lemma, that considers two consecutive affine functions of g, leads to an efficient algorithm to compute the convolution of a convex function by a concave function.

Lemma 4.4. *Consider the convolutions $f * g_{j-1}$ and $f * g_j$. Let*

$$\alpha_j = \min\{a_i \text{ in the decomposition of } f \mid f_i' \geqslant g_j'\}$$

and

$$\alpha_{j-1} = \min\{a_i \text{ in the decomposition of } f \mid f_i' \geqslant g_{j-1}'\}.$$

Then

- $\forall x \leqslant c_j + \alpha_j,\quad f * g_j(x) \geqslant f * g_{j-1}(x)$;
- $\forall x \geqslant c_j + \alpha_{j-1},\quad f * g_{j-1}(x) \geqslant f * g_j(x)$.

Proof. First, as f is convex and $g_{j-1}' > g_j'$, we have that $\alpha_{j-1} \geqslant \alpha_j$. From Lemma 4.1, for all $x \leqslant c_j + \alpha_j$,

$$f * g_j(x) = f(x - c_j) + g(c_j).$$

• Either $x \leqslant c_{j-1} + \alpha_{j-1}$, then $f * g_{j-1}(x) = f(x - c_{j-1}) + g(c_{j-1})$ and

$$f * g_j(x) - f * g_{j-1}(x) = f(x - c_j) - f(x - c_{j-1}) + g(c_j) - g(c_{j-1});$$

as $x - c_{j-1} \leqslant \alpha_{j-1}$, then $f(x - c_{j-1}) - f(x - c_j) \leqslant g_{j-1}' \cdot (c_j - c_{j-1})$ and $f * g_j(x) - f * g_{j-1}(x) \geqslant 0$;

• or $x > c_{j-1} + \alpha_{j-1}$, then $f * g_{j-1}(x) = f(\alpha_{j-1}) + g(x - \alpha_{j-1})$; as $c_{j-1} < x - \alpha_{j-1} \leqslant x - \alpha_j \leqslant c_j$, $g(c_j) - g(x - \alpha_{j-1}) = g_{j-1}' \cdot (c_j + \alpha_{j-1} - x)$ and $f(\alpha_{j-1}) - f(x - c_j) \leqslant g_{j-1}' \cdot (c_j + \alpha_{j-1} - x)$; then $f * g_j(x) - f * g_{j-1}(x) \geqslant 0$.

The second statement can be proved similarly. $\square$

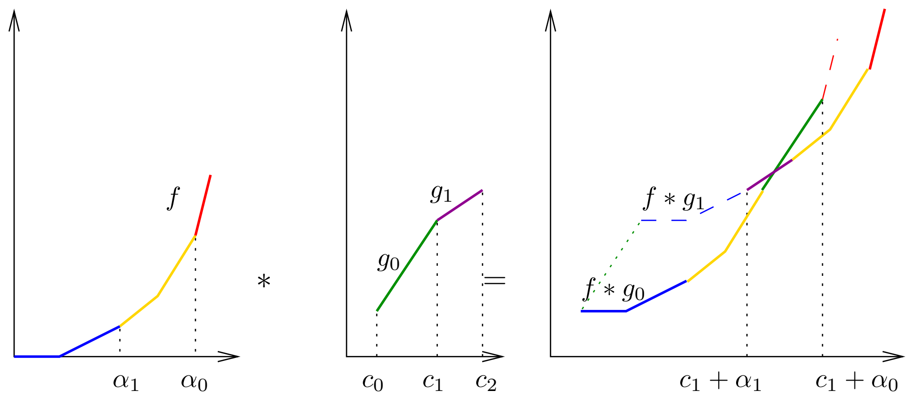

Figure 2. Convolution of a convex function by a concave function.

Another formulation of Lemma 4.4 is that $g_j^1 \geqslant f * g_{j-1}$ and that $g_{j-1}^2 \geqslant f * g_j$ and that the two functions intersect at least once. Hence g_j^1 and g_{j-1}^2 cannot appear in the minimum of $f * g_j$ and $f * g_{j-1}$. By transitivity, there is no need to compute entirely the convolution of the convex function by every affine component of the decomposition of the concave function. If there are more than two segments, successive applications of this lemma show that only the position of the segments of the concave function must be computed, except for the extremal segments.

The following lemma shows that $f * g_j$ and $f * g_{j-1}$ intersect in one and only one connected component, as for a given abscissa, the slope of $f * g_j$ is less than the one of $f * g_{j-1}$.

Lemma 4.5. *Let $x \in \mathbb{R}$ be any real number for which both $f*g_j$ and $f*g_{j-1}$ are finite valued and differentiable. Then*

$$\frac{\mathrm{d}}{\mathrm{d}x} f * g_j(x) \leqslant \frac{\mathrm{d}}{\mathrm{d}x} f * g_{j-1}(x). \tag{27}$$

Proof. For $x \in [a_0, \alpha_j]$, $f * g_j(x + c_j) = f(x - c_j) + g(c_j)$ and $f * g_{j-1}(x + c_{j-1}) = f(x) + g(c_{j-1})$. As f is convex and $c_j \geqslant c_{j-1}$, the result holds on $[a_0 + c_j, \alpha_j + c_j]$. Similarily, the result holds for $x \in [\alpha_{j-1} + c_j, a_n + c_j]$.

On $[c_j + \alpha_j, c_j + \alpha_{j-1}]$, $f * g_j$ is composed of segment g_j concatenated with the segments f_i, $i \in [\alpha_{j-1}, \alpha_j]$, possibly truncated on the right and $f * g_{j-1}$ is composed of segments f_i, $i \in [\alpha_{j-1}, \alpha_j]$ concatenated with g_{j-1}, possibly truncated on the left. As $\forall i \in [\alpha_{j-1}, \alpha_j]$, $g'_{j-1} \leqslant f'_i \leqslant g'_j$, the result holds. □

If one sets by convention $\frac{\mathrm{d}}{\mathrm{d}x} f * g_j(x) = -\infty$ for $x < c_{j-1} + a_0$ and $\frac{\mathrm{d}}{\mathrm{d}x} f * g_j(x) = +\infty$ for $x > c_j + a_n$, then the inequality always holds.

The intersection of $f * g_{j-1}$ and $f * g_j$ can then happen in one and only one of the four cases:

(1) g^1_{j-1} and g^c_j intersect;
(2) g^1_{j-1} and g^2_j intersect;
(3) g^c_{j-1} and g^c_j intersect;
(4) g^c_{j-1} and g^2_j intersect.

The following theorem is another consequence of these lemmas and is more precise about the shape on the convolution of a convex function by a concave function.

Theorem 4.6 (convolution of a convex function by a concave function)**.** *The (min,plus)–convolution of a convex function by a concave function can be decomposed in three (possibly trivial) parts: a convex function, a concave function and a convex function.*

Proof. We use the notations defined in the former lemmas.

We now show by induction that the convolution of f by $\min_{\lambda \leqslant j} g_j$, denoted h_j, is composed of

(i) a convex part h^1_j, which is the restriction of $g_0(x) = f(x - c_0) + g(c_0)$ to $[c_0 + a_0, \beta_j]$, with $\beta_j \leqslant c_0 + a_n$;
(ii) a concave part h^c_j, which is a minimum of some segments g_λ, $\lambda \leqslant j$ (up to some translation) finite valued on $[\beta_j, \gamma_j]$;
(iii) a convex part h^2_j, $g_j(x) = f(x - c_{j+1}) + g(c_{j+1})$ for $x \in [\gamma_j, a_n + c_{j+1}]$, $\gamma_j \geqslant a_0 + c_{j+1}$.

Note that with these conventions, the real numbers β_j and γ_j are uniquely determined at each step of the induction.

The case with $j = 0$ is a direct consequence of Lemma 4.1. The case with $j = 1$ is a consequence of Lemmas 4.4 and 4.1. The graphs of $f * g_1$ and $f * g_0$ intersect once and only once (where they are finite valued), and in $[c_1 + \alpha_1, c_1 + \alpha_0]$. Depending on when this intersection occurs, the concave part will be trivial, be made of only one (part of a) segment of g, or a minimum of the two segments g_0 and g_1.

Suppose now that the result holds for h_j and consider $h_{j+1} = \min(h_j, f * g_{j+1})$. The argument is exactly the same as for $j = 1$: h_j and $f * g_{j+1}$ can only intersect once. Indeed, h_j is the minimum of functions such that $\frac{\mathrm{d}}{\mathrm{d}x} f * g_k(x) \geqslant \frac{\mathrm{d}}{\mathrm{d}x} f * g_{j+1}(x)$, and then $\frac{\mathrm{d}}{\mathrm{d}x} h_j(x) \geqslant \frac{\mathrm{d}}{\mathrm{d}x} f * g_{j+1}(x)$. Note that this intersection has to occur after the point $c_{j+1} + \alpha_{j+1}$.

Moreover, as g_j^2 does not intersect g_{j+1}^2 and that h_j^2 is a part of g_j^2 (by the induction hypothesis), h_j^2 does not intersect g_{j+1}^2.

Therefore, only one of the four following cases may occur.

(1) h_j^c intersects g_{j+1}^c and $h_{j+1}^1 = h_j^1$, $h_{j+1}^c = \min(h_j^c, g_{j+1}^c)$, $h_{j+1}^2 = g_{j+1}^2$, $\beta_{j+1} = \beta_j$ and $\gamma_{j+1} = \alpha_{j+1} + c_{j+2}$.
(2) h_j^c intersects g_{j+1}^2 at y and $h_{j+1}^1 = h_j^1$, $h_{j+1}^c = h_j^c$, $h_{j+1}^2 = g_{j+1}^2$, $\beta_{j+1} = \beta_j$ and $\gamma_{j+1} = y$.
(3) h_j^1 intersects g_{j+1}^c at y and $h_{j+1}^1 = h_j^1$, $h_{j+1}^c = h_j^c$, $h_{j+1}^2 = g_{j+1}^2$, $\beta_{j+1} = y$ and $\gamma_{j+1} = \alpha_{j+1} + c_{j+2}$.
(4) h_j^1 intersects g_{j+1}^2 at y and $h_{j+1}^1 = h_j^1$, h_{j+1}^c is trivial and $h_{j+1}^2 = g_{j+1}^2$, $\beta_{j+1} = \gamma_{j+1} = y$.

□

If the concave function is composed of m segments and the convex function of n segments, then the convolution of those two functions can be computed in time $\mathcal{O}(n + m \log m)$. The $\log m$ term comes from the fact that one has to compute the minimum of m segments (see [BT08] for more details). If the functions are now defined on $\mathbb{N}$, then, as no intersection point has to be computed for the minimum, the time complexity is $O(n + m)$. The corresponding algorithm is given in Algorithm 2, where without loss of generality (the (min,plus)–convolution is shift-invariant), the functions f and g are defined on $\mathbb{N}$ and finite between respectively 0 and n, and 0 and m. The slopes of the functions are thus $f_i' = f(i) - f(i-1)$ and $g_i' = g(i) - g(i-1)$.

4.2. **Application.** We now go back to our initial problem. As already explained above, the computation of $\mathcal{T}_{t,h}^{\tau} U(t,x)$ given in (18) is made of two steps:

- (min,plus)–convolution of u and $h : x \mapsto \tau K^*(\frac{x}{\tau})$;
- subtract $\tau V(t,x)$.

Note that the (min,plus)–convolution described above is here defined on functions that have an unbounded support. But in the periodic case, U is 1–periodic and h is convex with a global minimum. Then, to compute the convolution, it is enough to compute it on a single period (the (min,plus–convolution preserves the periodicity), and replace h by its restriction on a support of size 2 centered on its minimum. If $h = 1/k$ with the notation of the previous section, then both functions u and h are defined on grids of size k and $2k$ respectively.

The convolution of u and h can be efficiently computed following these steps:

(1) Decompose u into convex and concave parts. This can be done in linear time: the three first points determine if a part is concave

Algorithm 2: Convolution of a convex function by a concave function

```
Data: f : [0, n] → ℝ a convex function with slopes (r_i), g : [0, m] → ℝ
      a concave function with slopes (ρ_i).
Result: h = f * g
begin
    /* Initialization                                          */
    k ← 0;
    while k ⩽ m + n do h(k) ← +∞; k ← k + 1;
    ;
    i ← 0; j ← 0; h(0) ← f(0) + g(0);
    /* First convex part of the convolution                    */
    while f'_i ⩽ g'_0 do
        i ← i + 1; h(i) ← f(i) + g(0);
    /* Concave part of the convolution                         */
    j ← j + 1; h(i + j) ← f(i) + g(j);
    while j < m do
        while g'_j < f'_{i-1} do i ← i − 1;
        ;
        h(i + j) ← min(h(i + j), f(i) + g(j));
        h(i + j + 1) ← min(h(i + j + 1), f(i) + g(j + 1));
        j ← j + 1;
    /* Second convex part of the convolution                   */
    while i < n do
        i ← i + 1; h(i + m) ← min(h(i + m), f(i) + g(m));
```

or convex. Then, this part is extended as much as possible while preserving the concavity or convexity and so on.

(2) For each convex or concave part, perform the convolution with h using Algorithms 1 or 2.

(3) Take the minimum of all these convolutions.

The complexity of this Algorithm is then $\mathcal{O}(ck)$, where c is the number of components in the decomposition of u into concave/convex parts.

4.3. **Implementation issues.** The main issue with this algorithm is that c – the number of components in the decomposition of u – can become very large, and then lead to a quadratic time complexity, which is the complexity of a naive algorithm for computing the convolution. Experimentally, the reason for this is that, due to the discretization of u, nearly affine parts, after performing the convolution several times, are computed as fast alternations of convex and concave parts. As shown in Figure 3, one solution to make the computations more efficient would be to consider those parts are convex and use Algorithm 1.

To do this, we decompose u into convex and concave parts with a tolerance (we do not request for convex parts to have increasing increments, but the increments to have an increase more than $-\eta$, or in other terms, we treat a

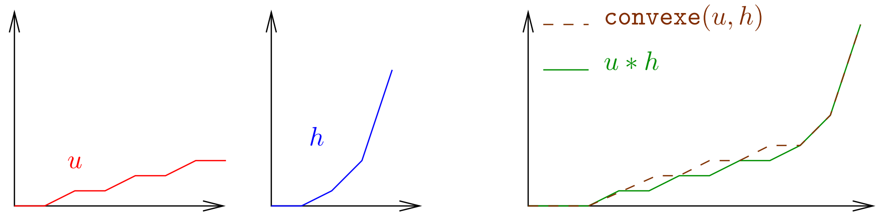

Figure 3. Approximation of the convolution: plain line shows the convolution of u and h, and the dashed line shows the function computed using Algorithm 1 when u is not convex, but has very small variations.

function f verifying $f'' \geqslant -\eta$ as if it were convex). We will discuss this in the next section. The choice of an optimal tolerance η, as well as the comparison with parallel implementations, will be the subject of further studies.

5. Numerical simulations

5.1. **Pendulum.** We take the Hamiltonian (3) with $P = 0$ and $V(t, x) = 1 - \cos(x)$ on $(0, 2\pi] \times \mathbb{R}$, with periodic boundary conditions. The initial value is $u_0(x) = \cos(x)$. In this case the corresponding solution develops a singularity in the derivative in $x = \pm\pi$ and the solution is not smooth. The numerical solution at $T = 8$ is plotted in Figure 4.

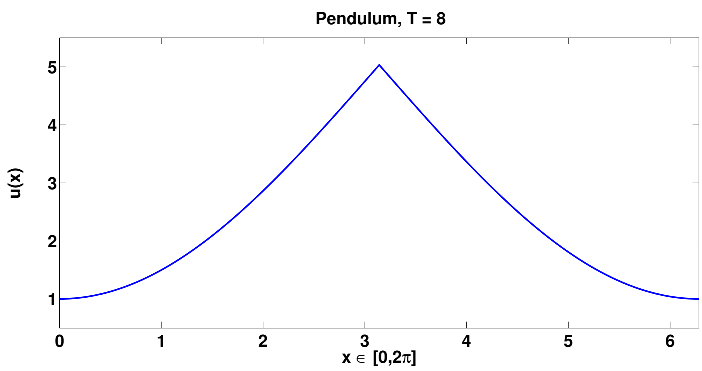

Figure 4. Solution of the Hamilton-Jacobi equation at $T = 8$ for the potential $1 - \cos(x)$ (pendulum)

We perform different simulations with mesh size of the form π/K with $K = 2^n$ for $n = 1$ to $n = 13$ corresponding to $2K$ grid points between 0 and π (from 4 to 16384 grid points). The time discretization parameter τ is taken to be $\sqrt{h}$ so that we expect a global order of convergence $\mathcal{O}(\sqrt{h})$.

The result is plotted on figure 5 where the error is the relative error in L^∞ norm between the solution obtained with the algorithm described above at time $T = 8$ and the solution computed using the fifth-order WENO algorithm (WENO5, see [JS96]) with 8192 grid points. Note that in this first simulation we take the regularisation parameter $\eta = 0$. We observe the expected convergence rate.

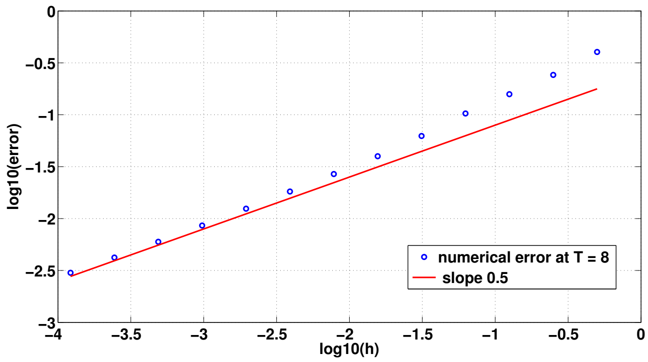

FIGURE 5. Error versus mesh size h with $\tau = \sqrt{h}$. Regularization parameter $\eta = 0$ (pendulum)

In a second step we perform the same simulation, but with $\eta = h$, which does not affect the convergence rate, but allows to go up to $2K = 2^{18} = 262144$ grid points for a few minutes of CPU time[1]

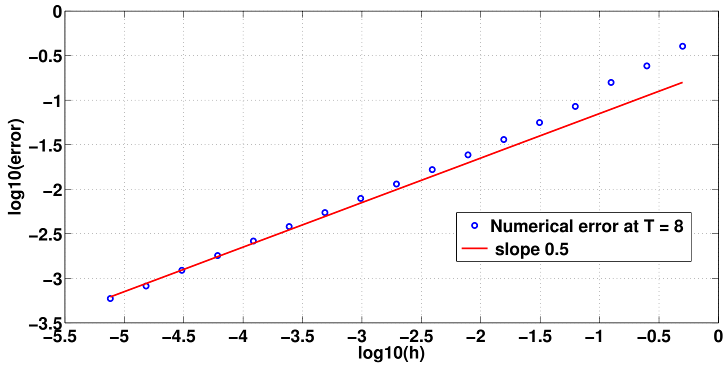

FIGURE 6. Error versus mesh size h with $\tau = \sqrt{h}$. Regularization parameter $\eta = h$ (pendulum)

Using the same reference solution at $T = 8$, we perform several simulations with different values of the regularization parameter. In Figure 7, we plot the evolution of the error with respect to η and for different values of h.

In Figure 8 we plot the evolution of the CPU time with respect to η, and in Figure 9, we plot the error versus the CPU time, for different values of h and different values of η.

As a conclusion of these simulations, it seems that the performances of the algorithm seem to be optimized for η depending linearly of h in the case of the pendulum.

5.2. **Time dependent potential.** We now take the $P = 1$ and the time dependent potential $V(t, x) = \cos(2x)\sin(t)$. We still consider the initial

[1]The CPU time required to obtain the solution with 2^{18} grid points is about 19mn, using MATLAB on a Mac power book 2,3 GHz Intel Core i7

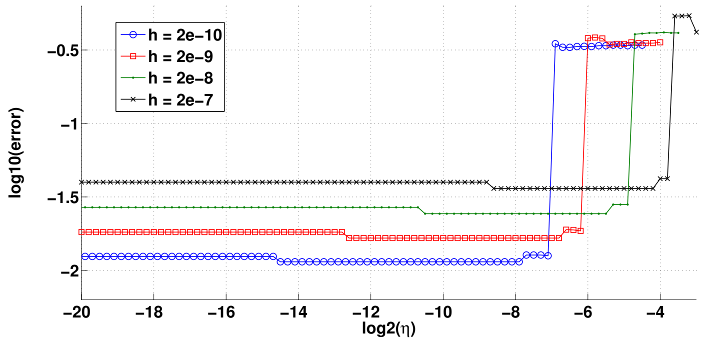

FIGURE 7. Error versus regularization parameter η for different values of h (pendulum)

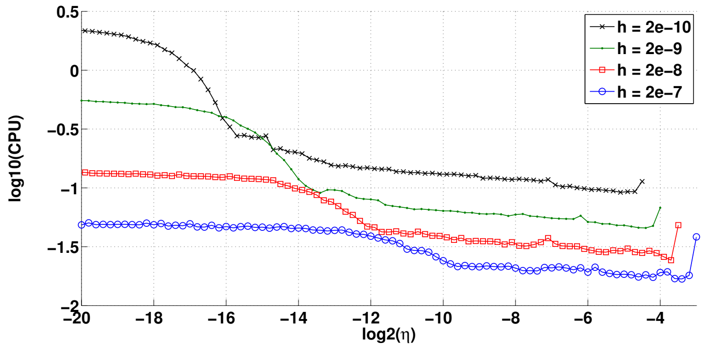

FIGURE 8. CPU time versus regularization parameter η for different values of h (pendulum).

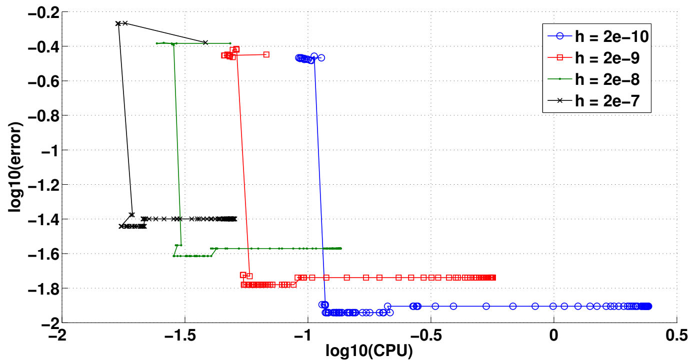

FIGURE 9. Error versus CPU time for different values of h (pendulum).

eigenvalue $u_0(x) = \cos(x)$. In this case, a theorem of Bernard and Roquejoffre ([BR04]) states that the solution converges towards a function that is periodic in time (of period possibly greater than 2π) which is a priori not constant, in contrast with the previous case. The shape of the solution is depicted in Figure 10.

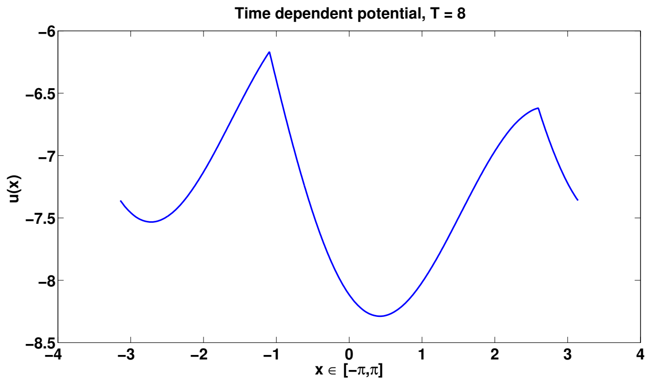

FIGURE 10. Solution of the Hamilton-Jacobi equation at $T = 8$ for the potential $\cos(2x)\sin(t)$ (time dependent potential)

In Figures 11 we illustrate the convergence result obtained above in the case $\tau = \sqrt{h}$ and observe the predicted rate of convergence $h^{1/2}$. Again the exact solution is computed at $T = 8$ with the WENO5 algorithm.

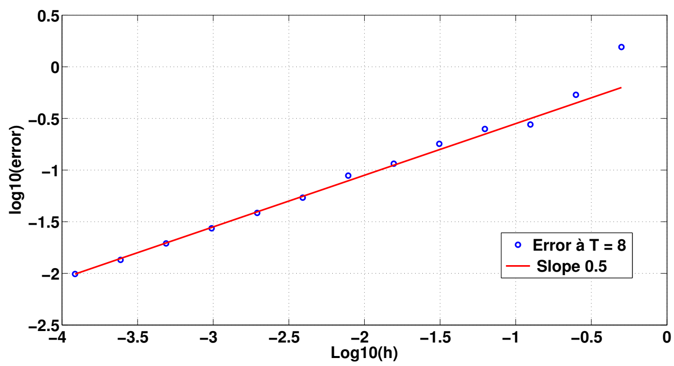

FIGURE 11. Error versus mesh size h with $\tau = \sqrt{h}$. Regularization parameter $\eta = 0$ (time dependent potential)

In figures 12, 13 and 14 we study the effect of the regularization parameter η with the same data as in the previous case. We see that the same conclusion can be drawn.

Finally, we take $P = 2$ and the potential function $V(t,x) = \sin(t)\cos(2x)$, $h = 0.01$ and $\tau = 0.1$. We take $\eta = 0$, and consider the initial data $u_0(x) = -\cos(3x)$. In Figure 15 we plot the evolution of $u(t,-\pi)$ with respect to time. We observe the linear growth predicted by the result of the previous section, for a time interval $[0, 10000]$. Note that there are oscillations in the evolution of $u(t,-\pi)$, but at a scale too small to be visible on the plot. This numerical result illustrates the behavior predicted by Proposition 3.5

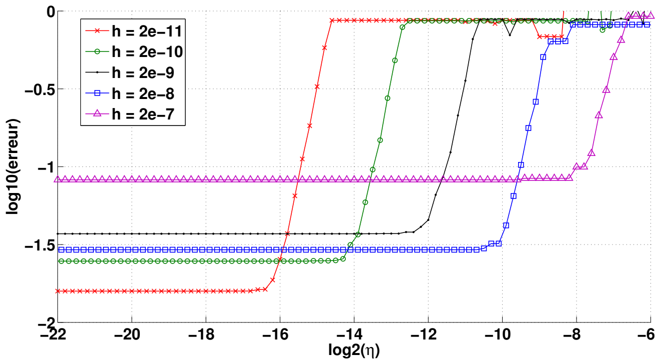

FIGURE 12. Error versus regularization parameter η for different values of h (time dependent potential)

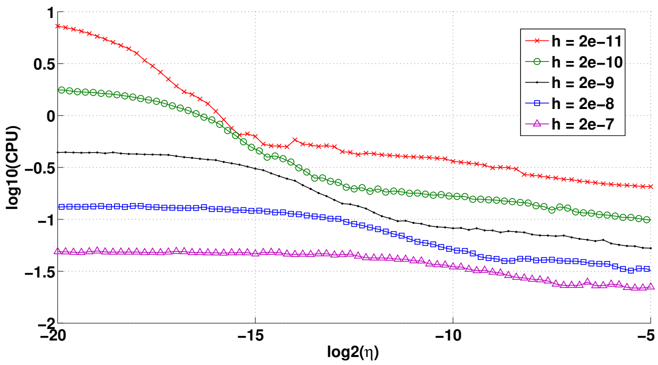

FIGURE 13. CPU time versus regularization parameter η for different values of h (time dependent potential)

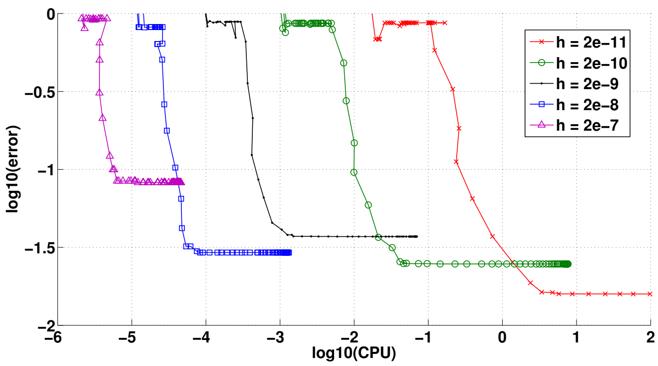

FIGURE 14. Error versus CPU time for different values of h (time dependent potential)

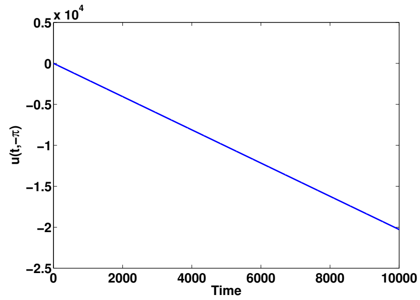

FIGURE 15. Evolution of $u(t,-\pi)$ over long times

APPENDIX A. APPENDIX: PROOF OF THE A PRIORI COMPACTNESS PROPOSITION 3.2

We start with a lemma.

Lemma A.1. *Assume that hypotheses **(i)** and **(iii)** are satisfied. Recall that $L(t,x,v) = K^*(v) - V(t,x)$. For any $\Gamma > 1$, there exists a constant Γ' such that for any $x, y \in \mathbb{R}^n$ and $T > 0$ and $t > 1$, if $|x-y|/t < \Gamma$ and γ that minimizes the quantity*

$$\inf_{\substack{\gamma(0)=x\\ \gamma(t)=y}} \int_0^t L\big(T+s,\gamma(s),\dot\gamma(s)\big)\mathrm{d}s,$$

then

$$\forall\, 0 \leqslant a \leqslant a+1 \leqslant t, \quad |\gamma(a)-\gamma(a+1)| < \Gamma'.$$

Proof. Without loss of generality, we will assume that L is positive. Indeed, the potential V is bounded, and adding a constant doesn't change the minimizers. In this case, and under hypothesis ***(i)***, there exists a nonnegative, increasing function $\alpha : [0,\infty) \to \mathbb{R}$ which verifies $\lim_{t\to\infty} \alpha(t) = \infty$, such that

$$\forall t,x,v, \quad L(t,x,v) \geqslant \alpha(|v|)|v|.$$

The idea of the proof is that if γ at some point has a great velocity, then it must be slow later. It is then better to "slow down" the fast part and accelerate the "slow" one.

First, we set some notations. For all $(x,y,t,T) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}$, let

$$A_T^t(x,y) = \inf_{\substack{\gamma(0)=x\\ \gamma(t)=y}} \int_0^t L\big(T+s,\gamma(s),\dot\gamma(s)\big)\mathrm{d}s,$$

be the Lagrangian action.

We start by showing, that the superlinearity of L implies a superlinearity of A_T^t. As already done a few times, we may bound the action by comparing with a straight line, using that L is uniformly bounded on sets of the form $\mathbb{R} \times \mathbb{R}^n \times B(0,R)$, where $B(0,R)$ is the ball of radius $R > 0$ in $\mathbb{R}^n$:

$$A_T^t(x,y) \leqslant \int_0^t L\big(T+s, \frac{(t-s)x+sy}{t}, \frac{y-x}{t}\big)\mathrm{d}s$$
$$\leqslant tC^+\big(\frac{|x-y|}{t}\big)\max\Big(\frac{|x-y|}{t},1\Big), \tag{28}$$

for some increasing function $z \mapsto C^+\big(z\big)$ defined on $\mathbb{R}^+$. Let γ realizing the infimum, and set

$$\mathcal{E} = \Big\{s \in [0,t] \quad \text{such that} \quad |\dot\gamma(s)| \geqslant \frac{|x-y|}{2t}\Big\}.$$

Then we get, using that $L>0$:

$$A_T^t(x,y) = \int_0^t L\big(T+s,\gamma(s),\dot\gamma(s)\big)\mathrm{d}s \tag{29}$$
$$\geqslant \int_{\mathcal{E}} L\big(T+s,\gamma(s),\dot\gamma(s)\big)\mathrm{d}s$$
$$\geqslant \alpha\big(\frac{|x-y|}{2t}\big)\int_{\mathcal{E}}|\dot\gamma(s)|\mathrm{d}s \geqslant \alpha\big(\frac{|x-y|}{2t}\big)\frac{|x-y|}{2}.$$

The last inequality comes from the fact that when γ is going at speed less than $|x-y|/2t$ for time t, it cannot travel more than $|x-y|/2$, therefore the integral is greater than $|x-y|/2$ by the triangular inequality. In other terms, equations (28) and (29) state that there are two positive functions C^+ and C^- which can be easily made increasing, such that

$$C^-\big(\frac{|x-y|}{t}\big)\frac{|x-y|}{t} \leqslant \frac{A_T^t(x,y)}{t} \leqslant C^+\big(\frac{|x-y|}{t}\big)\max\Big(\frac{|x-y|}{t},1\Big). \tag{30}$$

Moreover, thanks to the superlinearity of L those functions are coercive.

Now consider

- $\Gamma'' > \Gamma$ such that $20C^+(\Gamma) < C^-(\Gamma'')$,
- $\Gamma''' > \Gamma''$ such that $\Gamma'''/\Gamma'' \in \mathbb{N}^*$ and $30C^+(20\Gamma'') < C^-(\Gamma''')$,
- and finally $\Gamma' > \Gamma'''$ such that $40\Gamma'''/\Gamma'' < C^-(\Gamma')/C^+(\Gamma)$.

Let us verify that Γ' satisfies the requirements of our lemma.

Assume by contradiction that for some $x,y \in \mathbb{R}^n$, $t,T \in \mathbb{R}_+$ such that $|x-y| < t\Gamma$ and γ realizing the action $A_T^t(x,y)$, there is an $a \in [0,t-1]$ such that $|\gamma(a)-\gamma(a+1)| \geqslant \Gamma'$. As γ is a minimizer, we have (using that $L>0$ and $\Gamma>1$)

$$t\Gamma C^+(\Gamma) \geqslant A_T^t(x,y) \geqslant \int_a^{a+1} L\big(T+s,\gamma(s),\dot\gamma(s)\big)\mathrm{d}s \tag{31}$$
$$\geqslant A_{T+a}^1\big(\gamma(a),\gamma(a+1)\big) \geqslant \Gamma' C^-(\Gamma').$$

Hence we obtain

$$t \geqslant \frac{\Gamma' C^-(\Gamma')}{\Gamma C^+(\Gamma)} \geqslant \frac{40\Gamma'''}{\Gamma''}, \tag{32}$$

using the fact that $\Gamma' > \Gamma$ and the definition of Γ'.

We now assume that $a < t/2$, the other case may be treated similarly. Let $b \in [a, a+1]$ be the smallest number such that $|\gamma(a) - \gamma(b)| = \Gamma'''$, and consider the sequence

$$c_i = b + 2i\frac{\Gamma'''}{\Gamma''}, \quad i \in \{0, \cdots, k\},$$

where k is greatest possible integer such that $c_k \leqslant t$. Note that using (32) and $a \leqslant t/2$, we have $k \geqslant 9$, and that for $i \in \{0, \cdots, k\}$, we have $c_{i+1} - c_i = 2\Gamma'''/\Gamma''$.

We claim that there exists an $i_0 \in \{0, \cdots, k\}$ such that

$$\frac{|\gamma(c_{i_0}) - \gamma(c_{i_0+1})|}{c_{i_0+1} - c_{i_0}} \leqslant \Gamma''.$$

Indeed, otherwise we would have, using (30)

$$\begin{aligned} A_T^t(x,y) &\geqslant \sum_{i=0}^{k-1} \int_{c_i}^{c_{i+1}} L\big(T+s, \gamma(s), \dot\gamma(s)\big) \mathrm{d}s \\ &\geqslant \sum_{i=0}^{k-1} \frac{2\Gamma'''}{\Gamma''} \frac{|\gamma(c_{i_0}) - \gamma(c_{i_0+1})|}{c_{i_0+1} - c_{i_0}} C^-(\Gamma'') \\ &> \sum_{i=0}^{k-1} \frac{2\Gamma'''}{\Gamma''} \Gamma'' C^-(\Gamma''). \end{aligned}$$

By definition of k, we have that $(k+1) \times 2\Gamma'''/\Gamma'' \geqslant t/2 - 1$ while using (32), we have $t \geqslant 40\Gamma'''/\Gamma'' \geqslant 40$ and $2\Gamma'''/\Gamma'' \leqslant t/18$. Hence we deduce that $k \times 2\Gamma'''/\Gamma'' \geqslant t/3$. As $\Gamma'' \geqslant \Gamma$, the previous equation yields

$$A_T^t(x,y) > \frac{t}{3}\Gamma C^-(\Gamma'') \geqslant t\Gamma C^+(\Gamma),$$

which is absurd in view of (31).

Now we find a contradiction by constructing a curve δ which has an action less than γ. Let $[c,d] = [c_{i_0}, c_{i_0+1}]$. Recall that $K := \Gamma'''/\Gamma''$ is an integer. We define the curve δ as follows:

- $\delta(s) = \gamma(s)$ if $s \in [0,a] \cup [d,t]$;
- on $[a, b+K]$, δ coincides with the curve minimizing $A_{T+a}^{b+K-a}\big(\gamma(a), \gamma(b)\big)$;
- on $[b+K, c+K]$, δ is the translate of γ: $\delta(s) = \gamma(s-K)$;
- on $[c+K, d]$ (recall that $d = c + 2K$) δ coincides with the curve minimizing $A_{T+c+K}^{K}\big(\gamma(c), \gamma(d)\big)$.

We now compute the difference of action between γ and δ, recalling that L is 1-periodic in time:

$$\begin{aligned}\int_0^t L\big(T+s,\gamma(s),\dot\gamma(s)\big)\mathrm{d}s - \int_0^t L\big(T+s,\delta(s),\dot\delta(s)\big)\mathrm{d}s \\ = \int_a^b L\big(T+s,\gamma(s),\dot\gamma(s)\big)\mathrm{d}s + \int_c^d L\big(T+s,\gamma(s),\dot\gamma(s)\big)\mathrm{d}s \\ - \int_a^{b+K} L\big(T+s,\delta(s),\dot\delta(s)\big)\mathrm{d}s - \int_{c+K}^d L\big(T+s,\delta(s),\dot\delta(s)\big)\mathrm{d}s \\ \geqslant \Gamma'''C^-(\Gamma''') + \frac{2\Gamma'''}{\Gamma''}\Gamma''C^-(\Gamma'') \\ - \frac{\Gamma'''}{\Gamma''}\Gamma''C^+(\Gamma'') - \frac{\Gamma'''}{\Gamma''}(2\Gamma'')C^+(2\Gamma'') > 0.\end{aligned}$$

This contradicts the minimality of γ.

□

Remark A.2. In the previous proof, we only used the fact that L is periodic in time. In [Itu96], a similar result is proved when L is periodic in space (instead of in time). The idea of the proof is the same except that, when constructing the curve δ, instead of translating it in time (in third part of the construction), it is translated in space, while the "fast" part of γ between a and b is replaced by a geodesic (straight line) between $\gamma(a)$ and the closest point from $\gamma(a)$ in the grid $\gamma(b)+\mathbb{Z}^n$.

We now prove Lemma 3.2:

Proof of Lemma 3.2. Recall now that L is periodic both in time and in space and that its Euler–Lagrange flow is complete. As in the previous lemma, assume $L > 0$. Let Γ and Γ' be as in the previous lemma, and γ be a minimizer such that $|\gamma(0)-\gamma(t)|/t \leqslant \Gamma$. The curve γ is then a trajectory of the Euler–Lagrange flow. Let moreover $0 \leqslant a \leqslant a+1 \leqslant t$. Finally, by superlinearity of L, let $A(1)$ be given by Equation (7), such that $L(t,x,v) \geqslant |v| - A(1)$. We therefore obtain that, with the notations used in the previous proof,

$$\begin{aligned}\int_0^1 |\dot\gamma(a+s)|\mathrm{d}s - A(1) \\ \leqslant \int_0^1 L\big(T+a+s,\gamma(a+s),\dot\gamma(a+s)\big)\mathrm{d}s \leqslant C^+(\Gamma')\Gamma'.\end{aligned}$$

Therefore, there is at least one point $s_0 \in [0,1]$ such that

$$|\dot\gamma(a+s_0)| \leqslant A(1) + C^+(\Gamma')\Gamma' := D.$$

By periodicity of the Lagrangian, and completeness of the Euler-Lagrange flow, there exists a constant D' depending only on D, such that $|\dot\gamma| \leqslant D'$ on $[a+s_0-1, a+s_0+1] \cap [0,t] \supset [a,a+1]$. Since a is arbitrary, this finishes the proof. □

Appendix B. Appendix: Proof of Theorem 2.8

Proof of Theorem 2.8. The idea of the proof a rather common technique which consists in interchanging the minimizing paths between the continuous and the fully discrete semi–groups.

Set $T = N\tau$. Let us denote by $\gamma_{t_0} : [t_0, t_0 + T] \to \mathbb{R}^n$ a minimizer of (8) (with $t = t_0$ and $\delta = T$). Recall that the curve $\gamma_{t_0}(s)$ is C^2. Let us set $y := \gamma_{t_0}(t_0)$ and $x := \gamma_{t_0}(t_0 + T)$. We have

$$T_{t_0}^T u_0(x) = u_0(y) + \int_{t_0}^{t_0+T} L\big(s, \gamma_{t_0}(s), \dot{\gamma}_{t_0}(s)\big)\mathrm{d}s.$$

By superlinearity (7), this implies that

$$\int_{t_0}^{t_0+T} |\dot{\gamma}_{t_0}(s)|\mathrm{d}s \leqslant TA(1) + |T_{t_0}^T u_0(x) - u_0(y)|.$$

Comparing with the trivial curve $\gamma \equiv x$ in the definition of the Lax–Oleinik semi–group, we have that

$$T_{t_0}^T u_0(x) \leqslant |u_0|_\infty + T\big(B + K^*(0)\big), \tag{33}$$

where B is the constant in equation (6) (hypothesis ***(i)***). Moreover, since L is bounded below $\big($meaning $L(t, x, v) \geqslant b$ for some constant b, for all $(t, x, v)\big)$, clearly, the action of any curve defined for a time T is greater than Tb which implies immediately that

$$T_{t_0}^T u_0(x) \geqslant -|u_0|_\infty + Tb. \tag{34}$$

Hence, there exists a constant B_1 depending only on L and $|u_0|_\infty$ such that

$$|x - y| = |\gamma_{t_0}(t_0 + T) - \gamma_{t_0}(t_0)| \leqslant \int_{t_0}^{t_0+T} |\dot{\gamma}_{t_0}(s)|\mathrm{d}s \leqslant B_1(1 + T). \tag{35}$$

Remarking that γ_{t_0} is also a minimizer of the action (10) under the constraint $\gamma(t_0) = y$ and $\gamma(t_0 + T) = x$, we can apply Proposition 2.2 which shows that there exists a constant $M_1 = M\big(B_1(1 + T), T\big)$ depending only on T, L and $|u_0|_\infty$ such that

$$\forall\, s \in [t_0, t_0 + T], \quad |\dot{\gamma}_{t_0}(s)| \leqslant M_1. \tag{36}$$

Assume now that N is an integer such that $T = N\tau \leqslant T_0$. For all $i = 0, \dots, N$ we define

$$x_i = h\left\lfloor \frac{1}{h}\gamma_{t_0}(t_i)\right\rfloor$$

where for $i = 0, \dots, N$, $t_i = t_0 + i\tau$ and where the function $\lfloor\cdot\rfloor$ is the floor function, coordinate by coordinate. With these points, we associate the continuous piecewise linear path λ defined by

$$\lambda(s) = x_i + (s - t_i)\frac{x_{i+1} - x_i}{\tau}, \quad \text{for} \quad s \in [t_i, t_{i+1}]. \tag{37}$$

By definition of the points x_i, we have

$$\forall\, i \in [0, N], \quad |x_i - \gamma_{t_0}(t_i)| = |\lambda(t_i) - \gamma_{t_0}(t_i)| \leqslant h\sqrt{n}. \tag{38}$$

Now, using the bound (36), we have for all $i = 0, \dots, N-1$,

$$|x_{i+1} - x_i| \leqslant 2h\sqrt{n} + \int_{t_i}^{t_{i+1}} |\dot\gamma_{t_0}(s)| \mathrm{d}s \leqslant 2h\sqrt{n} + \tau M_1.$$

But this inequality implies that for all $i = 0, \dots, N-1$,

$$\forall\, s \in [t_i, t_{i+1}], \quad |\lambda(s) - x_i| \leqslant 2h\sqrt{n} + \tau M_1,$$

while $|\gamma_{t_0}(s) - \gamma_{t_0}(t_i)| \leqslant \tau M_1$ upon using (36). Hence we get

$$\forall\, s \in [t_i, t_{i+1}], \quad |\lambda(s) - \gamma_{t_0}(s)| \leqslant 3h\sqrt{n} + 2\tau M_1. \tag{39}$$

Moreover, we have for $s, \sigma \in [t_i, t_{i+1}]$,

$$|\dot\gamma_{t_0}(\sigma) - \dot\gamma_{t_0}(s)| \leqslant \tau C,$$

upon using (11). Hence for $s \in [t_i, t_{i+1}]$, we have

$$|\gamma_{t_0}(t_{i+1}) - \gamma_{t_0}(t_i) - \tau\dot\gamma_{t_0}(s)| \leqslant \int_{t_i}^{t_{i+1}} |\dot\gamma_{t_0}(\sigma) - \dot\gamma_{t_0}(s)| \mathrm{d}\sigma \leqslant \tau^2 C,$$

and hence for all $s \in [t_i, t_{i+1}]$

$$\left| \frac{\gamma_{t_0}(t_{i+1}) - \gamma_{t_0}(t_i)}{\tau} - \dot\gamma_{t_0}(s) \right| \leqslant \tau C.$$

Using (38), we obtain easily that for all $i = 0, \dots, N-1$,

$$\forall\, s \in [t_i, t_{i+1}], \quad |\dot\lambda(s) - \dot\gamma_{t_0}(s)| \leqslant \tau C + \frac{2h}{\tau}\sqrt{n}. \tag{40}$$

Note that using (20) and (36), the previous equation implies that for all $i = 0, \dots, N-1$,

$$\forall\, s \in [t_i, t_{i+1}], \quad |\dot\lambda(s)| \leqslant M_2 \tag{41}$$

for some constant $M_2 = \tau_0 C + h_0\sqrt{n} + M_1$ independent of h and τ.

Now by definition of λ, we have

$$\begin{aligned} \Big| \int_{t_0}^{t_0+N\tau} L\big(s, \gamma_{t_0}(s), \dot\gamma_{t_0}(s)\big) \mathrm{d}s - \int_{t_0}^{t_0+N\tau} L\big(s, \lambda(s), \dot\lambda(s)\big) \mathrm{d}s \Big| \\ \leqslant \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \Big| L\big(s, \gamma_{t_0}(s), \dot\gamma_{t_0}(s)\big) - L\big(s, \lambda(s), \dot\lambda(s)\big) \Big| \mathrm{d}s. \end{aligned} \tag{42}$$

Using (6) (coming from hypothesis ***(ii)***), the fact that K^* is C^2 and the bounds (36) and (41), there exists a constant M_3, depending on L, M_1 and M_2, such that the previous error term is bounded by

$$M_3 \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \Big(|\gamma_{t_0}(s) - \lambda(s)| + |\dot\gamma_{t_0}(s) - \dot\lambda(s)| \Big) \mathrm{d}s.$$

Using (39) and (40), this shows that there exists a constant M_4 independent on h and τ, such that

$$\left| \int_{t_0}^{t_0+N\tau} L\big(s, \gamma_{t_0}(s), \dot\gamma_{t_0}(s)\big) \mathrm{d}s - \int_{t_0}^{t_0+N\tau} L\big(s, \lambda(s), \dot\lambda(s)\big) \mathrm{d}s \right| \leqslant M_4\big(\frac{h}{\tau} + \tau\big), \tag{43}$$

where we used the fact that $h \leqslant \tau_0 h/\tau$.

Finally, the term we wish to estimate is

$$
\begin{aligned}
\Big| \int_{t_0}^{t_0+N\tau} & L\big(s,\gamma_{t_0}(s),\dot\gamma_{t_0}(s)\big)\mathrm{d}s - \sum_{i=0}^{N-1} \kappa^\tau_{t_i,h}(x_i,x_{i+1})\Big| \\
&\leqslant \sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} \Big|L\big(s,\gamma_{t_0}(s),\dot\gamma_{t_0}(s)\big) - L\big(s,\lambda(s),\dot\lambda(s)\big)\Big|\mathrm{d}s \\
&\qquad + \sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} \Big|L\big(s,\lambda(s),\dot\lambda(s)\big) - L\big(t_i,\lambda(t_i),\dot\lambda(t_i)\big)\Big|\mathrm{d}s.
\end{aligned}
$$

To bound the second term, we observe first that for $s \in [t_i, t_{i+1}]$, the derivative $\dot\lambda(s) = (x_{i+1} - x_i)/\tau$ does not depend on s. Hence using (6) and (41) the function

$$
[t_i, t_{i+1}] \ni s \mapsto L\big(s,\lambda(s),\dot\lambda(s)\big)
$$

is $\mathcal{C}^1$ with uniformly bounded derivative. Thus we obtain that there exists a constant M_5 such that

$$
\Big|L\big(s,\lambda(s),\dot\lambda(s)\big) - L\big(t_i,\lambda(t_i),\dot\lambda(t_i)\big)\Big| \leqslant M_5(s-t_i).
$$

This proves that $\big($compare (43)$\big)$

$$
\Big| \int_{t_0}^{t_0+N\tau} L\big(s,\gamma_{t_0}(s),\dot\gamma_{t_0}(s)\big)\mathrm{d}s - \sum_{i=0}^{N-1} \kappa^\tau_{t_i,h}(x_i,x_{i+1})\Big| \leqslant M_6\big(\frac{h}{\tau}+\tau\big), \tag{44}
$$

for some constant M_6 independent of h and τ.

Now by definition of $\mathcal{T}^N$, we have using (38) and the Lipschitz nature of u_0,

$$
\begin{aligned}
\mathcal{T}^N(u_0|_{G_h})(x) &\leqslant u_0(x_0) + \sum_{i=0}^{N-1} \kappa^\tau_{t_i,h}(x_i,x_{i+1}) \\
&\leqslant (T_{t_0}^{N\tau}u_0)|_{G_h}(x) + \big|u_0(x_0) - u_0\big(\gamma_{t_0}(T_0)\big)\big| + M_4\big(\frac{h}{\tau}+\tau\big) \\
&\leqslant (T_{t_0}^{N\tau}u_0)|_{G_h}(x) + M\big(\frac{h}{\tau}+\tau\big) \qquad (45)
\end{aligned}
$$

for some constant M independent on h and τ. This proves a first inequality in the estimate (21) with the notations of the Theorem.

To prove the reverse inequality, let us fix $x \in G_h$. We consider a sequence y_i, $i = 0, \dots, N$ with $y_N = x$ and

$$
\mathcal{T}^N(u_0|_{G_h})(x) = u_0(y_0) + \sum_{i=0}^{N-1} \kappa^\tau_{t_i,h}(y_i,y_{i+1}), \tag{46}
$$

and we define the curve

$$
\eta(s) = y_i + (s-t_i)\frac{y_{i+1}-y_i}{\tau}, \quad \text{for} \quad s \in [t_i, t_{i+1}].
$$

Note that in a similar manner to what we did to prove the inequalities (33) and (34), using the fact that u_0 is bounded, and comparing with the trivial sequence made of a constant point (with (6)) on the one hand, and the fact that L, hence the $\kappa^{\tau}_{t_i,h}$ are bounded below on the second hand show that there exists a constant D_1 such that

$$|\mathcal{T}^N(u_0|_{G_h})|_{\infty} \leqslant D_1.$$

By superlinearity of L (and of the $\kappa^{\tau}_{t_i,h}$) and using again the fact that u_0 is bounded, we thus see, as in (35) that there exists a constant D_2 such that for all $i = 0, \dots, N-1$,

$$\left|\frac{y_{i+1} - y_i}{\tau}\right| \leqslant D_2,$$

which in turn implies that

$$\forall\, s \in [t_i, t_{i+1}], \quad |\eta(s) - \eta(t_i)| \leqslant \tau D_2.$$

As the derivative of $\eta(s)$ with respect to s is uniformly bounded by D_2 and constant on the time intervals $[t_i, t_{i+1}]$, and as L is $\mathcal{C}^1$ with uniformly bounded derivative on $\mathbb{R} \times \mathbb{R}^n \times B(0, D_2)$, we obtain

$$\left|\sum_{i=0}^{N-1} \kappa^{\tau}_{t_i,h}(y_{i+1}, y_i) - \int_{t_0}^{t_0+N\tau} L\big(s, \eta(s), \dot\eta(s)\big)\mathrm{d}s\right| \leqslant \tau D_3 \tag{47}$$

for some constant D_3. Using the definition of the exact semi–group, we thus have

$$\begin{aligned}(T^{N\tau}_{t_0} u_0)|_{G_h}(x) &\leqslant u_0\big(\eta(t_N)\big) + \int_{t_0}^{t_0+N\tau} L\big(s, \eta(s), \dot\eta(s)\big)\mathrm{d}s \\ &\leqslant \mathcal{T}^N(u_0|_{G_h})(x) + \tau D_3\end{aligned}$$

upon using (46) and (47). This proves the result. □

proof of Theorem 3.3. In the proof of Theorem 2.8, equation (35) then gives using Proposition 3.2 (with $T \geqslant T_1 > 1$) that the constant M_1 defined in (36) does not depend on $T = N\tau$ and depends in fact only on T_1. It then follows that M_2 and M_3 also are independent of $T = N\tau$, while M_4 is proportional to the time of integration, that is $N\tau$. The rest of the proof can then be carried on giving the result. □

References

[Abg96] R. Abgrall. Numerical discretization of the first-order Hamilton-Jacobi equation on triangular meshes. *Comm. Pure Appl. Math.*, 49(12):1339–1373, 1996.

[AGL08] M. Akian, S. Gaubert, and A. Lakhoua. The max-plus finite element method for solving deterministic optimal control problems: basic properties and convergence analysis. *SIAM J. Control Optim.*, 47(2):817–848, 2008.

[BB07] P. Bernard and B. Buffoni. Weak KAM pairs and Monge-Kantorovich duality. In *Asymptotic analysis and singularities—elliptic and parabolic PDEs and related problems*, volume 47 of *Adv. Stud. Pure Math.*, pages 397–420. Math. Soc. Japan, Tokyo, 2007.

[BCOQ92] F. Baccelli, G. Cohen, G.Y. Olsder, and J.P. Quadrat. *Synchronization and linearity*. Wiley, 1992.

[BJ05] G. Barles and E. R. Jakobsen. Error bounds for monotone approximations schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.*, 43:540–558, 2005.

[BJT08] A. Bouillard, L. Jouhet, and E. Thierry. Computation of a (min,+) multidimensional convolution for end-to-end performance analysis. In *Proceedings of Valuetools'2008*, 2008.

[BR04] P. Bernard and J.-M. Roquejoffre. Convergence to time-periodic solutions in time-periodic Hamilton-Jacobi equations on the circle. *Comm. Partial Differential Equations*, 29(3-4):457–469, 2004.

[Bre89] Y. Brenier. Un algorithme rapide pour le calcul de transofrméesde Legendre-Fenchel discrètes. *C. R. Acad. Sci. Paris. Sér. I Math.*, 308(20):587–589, 1989.

[BS91] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[BT08] A. Bouillard and E. Thierry. An algorithmic toolbox for network calculus. *Discrete Event Dynamic Systems*, 18(1):3–49, 2008.

[CISM00] G. Contreras, R. Iturriaga, and H. Sanchez-Morgado. Weak solutions of the Hamilton-Jacobi equation for time periodic Lagrangians. *preprint*, 2000.

[CL84] M. G. Crandall and P.-L. Lions. Two approximations of solutions of Hamilton–Jacobi equations. *Math. Comp.*, 43:1–19, 1984.

[Fal87] M. Falcone. A numerical approach to the infinite horizon problem of deterministic control theory. *Appl. Math. and Optim.*, 15:1–13, 1987. *Corrigenda* Appl. Math. and Optim. Vol 23 (1991), 213–214.

[Fat97] A. Fathi. Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(9):1043–1046, 1997.

[Fat98] A. Fathi. Sur la convergence du semi-groupe de Lax-Oleinik. *C. R. Acad. Sci. Paris Sér. I Math.*, 327(3):267–270, 1998.

[Fat05] A. Fathi. Weak KAM Theorem in Lagrangian Dynamics, preliminary version, Pisa, 16 février 2005.

[FF02] M. Falcone and R. Ferreti. Semi-Lagrangian schemes for Hamilton-Jacobi equations, discrete representation formulae and Godunov methods. *J. Comput. Phys.*, 175:559–575, 2002.

[FM07] A. Fathi and E. Maderna. Weak KAM theorem on non compact manifolds. *NoDEA*, 14(1):1–27, 2007.

[HLW06] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations, Second edition.* Springer Series in Computational Mathematics 31. Springer, Berlin, 2006.

[Itu96] R. Iturriaga. Minimizing measures for time-dependent Lagrangians. *Proc. London Math. Soc. (3)*, 73(1):216–240, 1996.

[JKR01] E. R. Jakobsen, K. H. Karlsen, and N. H. Risebro. On the convergence rate of operator splitting for Hamilton-Jacobi equations with source terms. *SIAM J. Numer. Anal.*, 39(2):499–518, 2001.

[JP00] G. Jiang and D.-P. Peng. Weighted ENO schemes for Hamilton-Jacobi equations. *SIAM J. Sci. Comput.*, 21:2126–2143, 2000.

[JS96] G.-S. Jiang and C.-W. Shu. Efficient implementation of weighted ENO schemes. *J. Comput. Phys.*, 126:202–228, 1996.

[JX98] S. Jin and Z. Xin. Numerical passage from systems of conservation laws to Hamilton-Jacobi equations. *SIAM J. Numer. Anal.*, 35:2385–2404, 1998.

[LBT01] J.-Y. Le Boudec and P. Thiran. *Network Calculus: A Theory of Deterministic Queuing Systems for the Internet*, volume LNCS 2050. Springer-Verlag, revised version 4, may 10, 2004 edition, 2001.

[Lio82] P.-L. Lions. *Generalized solutions of Hamilton-Jacobi equations*, volume 69 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass., 1982.

[LPV87] P.-L. Lions, G. Papanicolaou, and S.R.S. Varadhan. Homogenization of Hamilton-Jacobi equation. unpublished preprint, 1987.

[LR04] B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics.* Cambridge Monographs on Applied and Computational Mathematics 14. Cambridge University Press, Cambridge, 2004.

[LS95] P.-L. Lions and P. E. Souganidis. Convergence of MUSCL and filtered schemes for scalar conservation laws and Hamilton-Jacobi equations. *Numer. Math.*, 69:441–470, 1995.

[LT01] C.-T. Lin and E. Tadmor. L^1-stability and error estimates for approximate Hamilton-Jacobi solutions. *Numer. Math.*, 87:701–735, 2001.

[Luc97] Y. Lucet. Faster than the fast Legendre transform, the linear-time Legendre transform. *Numer. Algorithms*, 16(2):171–185, 1997.

[Mat91] J. N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.*, 207(2):169–207, 1991.

[OS88] S. Osher and J. Sethian. Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79:12–49, 1988.

[OS91] S. Osher and C.-W. Shu. High-order essentially nonoscillatory schemes for Hamilton–Jacobi equations. *SIAM J. Numer. Anal.*, 28:907–922, 1991.

[Ror06] M. Rorro. An approximation scheme for the effective Hamiltonian and applications. *Appl. Numer. Math.*, 56(9):1238–1254, 2006.

[Sog13] K. Soga. Stochastic and variational approach to the Lax-Friedrichs scheme. *Math. Comp.*, in press, 2013.

[Sou85] P. E. Souganidis. Approximation schemes for viscosity solutions of Hamilton-Jacobi equations. *J. Differential Equations*, 59:1–43, 1985.

[Zav12] M. Zavidovique. Strict subsolutions and Mañé potential in discrete weak KAM theory. *Comment. Math. Helv.*, 87(1):1–39, 2012.

Ecole Normale Supérieure, 45 rue d'Ulm, 75230 Paris Cedex 05, France.
E-mail address: anne.bouillard@ens.fr

INRIA & Ecole Normale Supérieure de Cachan Bretagne, Avenue Robert Schumann 35170 Bruz, France
E-mail address: Erwan.Faou@inria.fr

IMJ-PRG, Université Pierre et Marie Curie, Case 247 4, place Jussieu, 75252 Paris Cedex 05, France
E-mail address: Maxime.Zavidovique@upmc.fr