ENTROPY PENALIZATION METHODS FOR HAMILTON-JACOBI EQUATIONS

DIOGO A. GOMES, ENRICO VALDINOCI

Abstract. In this paper we present a new entropy penalization problem and we discuss its relations with approximate solutions of Hamilton-Jacobi equations, the convergence of associated discrete schemes, as well as several applications, such as: a generalization of the Hopf-Cole transformation which converts non-linear Hamilton-Jacobi equations into linear evolution equations, the study of fixed point problems, approximation of certain linear evolution equations, and the construction of entropy penalized Mather measures.

Contents

Supported in part by FCT, POCTI/FEDER,
POCI/FEDER/MAT/55745/2004 and
MIUR Variational Methods and Nonlinear Differential Equations

Subject Classification: 70H20, 37M25, 37J50

``A quel tempo, di numeri ce n'erano soltanto due: il numero *e* e il numero *pi greco*. Il Decano fa un calcolo a occhio e croce, e risponde: - Cresce di *e* elevato a *ti*. Bravo furbo! Fin lì ci arrivano tutti."

(Italo Calvino, ``Le Cosmicomiche").

1. Introduction

The objective of this paper is to study a new approximation procedure for viscosity solutions of the Hamilton-Jacobi equations arising in optimal control. This procedure consists in the discretization of the control problem and in the addition of a suitable entropy functional. This functional regularizes the evolution of the discrete value function by acting as a viscosity term. The entropy penalization method is motivated by considerations in statistical mechanics, as well as by the works of [Ana04], [FS86] and [Eva04].[1]

In this paper, we introduce an entropy penalized scheme related to Lagrangian/Hamiltonian dynamics and Hamilton-Jacobi-type equations, we prove its convergence, we introduce a related generalized Hopf-Cole transformation, and we investigate the connections between fixed points of the scheme, the vanishing viscosity method, and the theory of Mather measures.

In further detail, the organization of this paper and its main results are the following.

In § 2, we introduce and motivate the entropy scheme

$$
\begin{aligned}
\phi \longmapsto \mathcal{G}[\phi] &= -\epsilon \ln \left[\int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\phi(x+hv)}{\epsilon}} \, dv \right] \\
&= \inf \int_{\mathbb{R}^N} \left[h\, L(x,v) + \phi_n(x+hv) + \epsilon \ln \gamma(v) \right] \gamma(v) \, dv \,,
\end{aligned}
$$

[1]It may be useful to give a short sketch about the motivations of entropy methods in mathematical physics. Rougly speaking, in genetral, an entropy penalized metod consists in perturbing a given problem by adding an "entropy term" of the type $S(f) = \int f \log f$. In statistical mechanics, such entropy S "counts", in a logaritmic scale, the number of microscopic configurations of a physical system which are compatible with a given macroscopic behavior, while f has the meaning of the frequency of any state. We refer to [Vil03] for a more detailed statistical motivation of the entropy functional.

where the infimum above is taken over all the probability densities γ on $\mathbb{R}^N$. Here, h plays the rôle of a time step discretization and ϵ is the entropy penalization parameter. We notice from the above formulas that there are *two equivalent* ways of defining such a scheme (see formulas (4) and (7) here below). The equivalence between these two definitions is shown in § 3.

Then, a *linear* scheme

$$\psi \longmapsto \mathcal{L}[\psi] = \int_{\mathbb{R}^N} e^{-\frac{h\,L(x,v)}{\epsilon}} \psi(x + hv)\, dv$$

is introduced in § 4, which turns out to be equivalent to the (non-linear) entropy penalized method, via a generalized Hopf-Cole transformation of the type

$$\phi \longmapsto e^{-\phi/\epsilon}.$$

We devote § 5 to the study of some properties of the non-linear scheme $\mathcal{G}$. In particular, the scheme is proved to enjoy a *contraction property* in the space of functions "where constants are quotiented out" (see Theorem 9 for details). Though this feature is weaker than the standard strict contraction property, it will be still sufficient to solve the related fixed point problem by iteration. We also study the semiconcavity properties of the scheme. In particular, a *uniform semiconcavity estimate* is obtained in Theorem 13, from which a *uniform Lipschitz bound* follows. More precisely, we show that the Lipschitz and semiconcavity moduli of $\mathcal{G}[\phi]$ are bounded by the maximum between the semiconcavity modulus of ϕ and a universal constant.

Some elementary properties of linear scheme $\mathcal{L}$ are outlined in § 6.

We carry out some *asymptotics* in § 7, according to different choices of the parameters involved. Namely, the linear scheme is related to a parabolic equation (see Propositions 23 and 24), while the non-linear scheme asymptotic is related to the Hamilton-Jacobi equation, as shown in Proposition 25.

In § 8, the nth iteration of the non-linear scheme is proved to *converge uniformly*. This result is obtained by a fixed point procedure which makes use of the non-standard contraction property of the scheme. We show that $\mathcal{G}^n[\phi]$ converges, as $n \to +\infty$, to a fixed point of $\mathcal{G}$ "when constants are quotiented out" (we refer to Theorem 27 for further details).

The convergence of the scheme for *small time steps* is discussed in § 9, by following [Sou85] and [BS91]. We prove that, when the time step h goes to zero, the solutions of the suitably scaled linear and non-linear schemes converge uniformly to solutions of the heat equation and of the viscous Hamilton-Jacobi equation, respectively. The elliptic term in these equations is provided by the entropy (see Theorems 29 and 30).

In § 10, we investigate the connection between the fixed points of our scheme and suitably *penalized Mather measures*. In particular, the generalized Hopf-Cole transformation provides a correspondence between fixed points of the scheme and minimizing measures for an entropy penalized action (see Theorem 32), which, in turn, converge to the usual Mather measures when the penalization vanishes (see Theorem 40). We will also prove *uniqueness* of the entropy penalized Mather measures (see Theorem 36). We also remark that an "explicit" representation of entropy penalized Mather measures in terms of fixed points of the non-linear scheme holds (namely, formula (105)). In particular, we show that, if ϕ is the unique (up to constants) solution of

$$\mathcal{G}[\phi] - \phi = \lambda \in \mathbb{R}$$

and θ is the unique probability density solving

$$\int_{\mathbb{R}^N} \theta(x - hv) e^{-\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}} \, dv \; = \; \theta(x) \, ,$$

then

$$\mu(x, v) \; = \; \theta(x) e^{-\frac{hL(x,v)+\phi(x+hv)-\phi(x)-\lambda}{\epsilon}}$$

is the unique minimizer for the entropy penalized action.

We devote § 11 to the analysis of the convergence of the penalized problem to the classical problem of Mather. In particular, both the penalized measure and the penalized "effective Hamiltonian" converge to the ones of Mather's problem (see Theorems 38 and 40 for precise statements). In particular, both the minimizing measures and the minimal values of the entropy penalized action converge to the analogous objects in the non-penalized setting. More precisely, if L is the Lagrangian, the entropy penalized action of a "holonomic" (or "flow invariant", see (88)) measure μ is given by

$$\int_{\mathbb{T}^N \times \mathbb{R}^N} L(x, v) d\mu(x, v) + \frac{\epsilon}{h} S[\mu], \, .$$

Then, we construct the densities which minimize such penalized functional and we prove their convergence to Mather measures for small entropy.

The paper ends with two appendices. The first one is devoted to the "abstract" fixed point argument which is used in the proof of Theorem 27. The second appendix collects some elementary results on semiconcave functions.

The paper is, essentially, self-contained, except for some basic facts about Mather theory, for which we refer to [Mat91], [MF94] and [Gom05], and for some general fact about viscosity solutions, which may be found, for instance, in [CIL92]. In general, we devoted some effort to using elementary arguments in the proofs, rather than other ones involving finer technologies.

2. Motivation for the scheme

Let $H(p,x)$ be a Hamiltonian, which has $L(x,v)$ as conjugated Lagrangian (the precise assumptions on which will be listed in § 3):

$$L(x,v) = \sup_p -p\cdot v - H(p,x). \tag{1}$$

Let $\phi(x,t)$ be a viscosity solution of the Hamilton-Jacobi equation

$$-\phi_t + H(\nabla\phi, x) = 0,$$

with prescribed terminal condition $\phi(x,T)$. The solution admits the following representation formula:

$$\phi(x,t) = \inf \int_t^T L(q(\tau),\dot q(\tau))\,d\tau + \phi(q(T),T)\,, \tag{2}$$

for any $T > t$, where the above infimum is taken over all Lipschitz curves $q : [t,T] \to \mathbb{R}^N$ satisfying $q(t) = x$ (see, for instance,[2] Theorem 6.4.5 in [CS04]).

A discrete scheme with time step $h > 0$ to approximate ϕ consists in the iteration

$$\phi_{n+1}(x) := \inf_{v\in\mathbb{R}^N} \left[h\,L(x,v) + \phi_n(x+hv)\right].$$

[2] Beware of a sign change both in the time direction and in (1) between our notation and the one in [CS04].

By elementary considerations, the previous formula can be written as

$$\phi_{n+1}(x) \,=\, \inf_{\mu\in\mathcal{P}} \int_{\mathbb{R}^N} [h\,L(x,v)+\phi_n(x+hv)\,]\,d\mu(v)\,, \tag{3}$$

where $\mathcal{P}$ denotes the space of the probability measures on $\mathbb{R}^N$. To make the scheme smoother, we replace (3) by the "entropy penalized" scheme

$$\phi_{n+1}(x) := \inf_{\gamma\in\mathcal{D}} \int_{\mathbb{R}^N} [h\,L(x,v)+\phi_n(x+hv)+\epsilon\ln\gamma(v)]\,\gamma(v)\,dv, \tag{4}$$

where $\epsilon>0$ (here and in the rest of the paper) is a small parameter and we denoted by $\mathcal{D}$ the set of probability densities on $\mathbb{R}^N$, i.e.

$$\mathcal{D} := \left\{\gamma\in L^1(\mathbb{R}^N)\,|\,\gamma(v)\geq 0\ \ a.e.\,,\ \int_{\mathbb{R}^N}\gamma(v)\,dv=1\right\}\,. \tag{5}$$

For a given (measurable) function $\phi:\mathbb{R}^N\to\mathbb{R}$, we define $\mathcal{G}[\phi]$ as

$$\mathcal{G}[\phi](x) := -\epsilon\ln\left[\int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\phi(x+hv)}{\epsilon}}\,dv\right]. \tag{6}$$

Note that $\mathcal{G}$ is well defined, for instance, if ϕ is bounded and $\mathcal{G}[\phi]$ is also bounded.

In Theorem 2 below, we show that (4) is equivalent to the explicit iteration scheme

$$\phi_{n+1} := \mathcal{G}[\phi_n]\,. \tag{7}$$

Then, in Theorem 3, we will establish the equivalence with a *linear* scheme, which generalizes the Hopf-Cole transformation. In relation with that, we recall that if u is a solution to the Hamilton-Jacobi equation

$$u_t+\frac{|Du|^2}{2}=\epsilon\Delta u,$$

then the Hopf-Cole transform $v=e^{-\frac{u}{\epsilon}}$ is a solution to the heat equation

$$v_t=\epsilon\Delta v\,,$$

thus it is conceivable that similar exponential transformations happen to be useful in our framework. We now provide a formal setting of the above discussions.

3. Existence of a minimizing measure and equivalency between (4) and (7)

From now on, we assume that the Lagrangian is a suitably smooth (e.g., Lipschitz) function that has the form

$$L(x,v) = K(v) - U(x)\,, \qquad \text{for} \quad v \in \mathbb{R}^N,\, x \in \mathbb{R}^N\,,$$

in which K, the "kinetic enery", is strictly convex in v and superlinear at infinity, and U is the "potential energy" which is bounded, $\mathbb{Z}^N$-periodic and semiconvex, that is, there exists $C_U > 0$ such that

$$\sup_{\substack{x,y\in\mathbb{R}^N\\ y\neq 0}} \frac{U(x+y)+U(x-y)-2U(x)}{|y|^2} \geq -C_U. \tag{8}$$

We suppose further that K is semiconcave, i.e., that there exists C_K such that

$$\sup_{\substack{v,w\in\mathbb{R}^N\\ w\neq 0}} \frac{K(v+w)+K(v-w)-2K(v)}{|w|^2} \leq C_K. \tag{9}$$

Proposition 1. *Let $\mathcal{D}$ as in (5). Fix $x \in \mathbb{R}^N$. Then, there exist $\gamma_\star \in \mathcal{D}$ realizing*

$$\inf_{\gamma\in\mathcal{D}} \int \Big(h\, L(x,v) + \phi(x+hv) + \epsilon \ln \gamma(v)\Big)\gamma(v)\, dv\,. \tag{10}$$

More explicitly, $\gamma_\star$ is given by

$$\gamma_\star := e^{-\frac{\hat{L}(x,v)+\lambda}{\epsilon}},$$

with

$$\hat{L}(x,v) := hL(x,v) + \phi(x+hv),$$

and

$$\lambda := \epsilon \ln \int e^{-\frac{\hat{L}(x,v)}{\epsilon}} dv\,.$$

Also, the quantity in (10) is equal to $-\lambda$.

Proof. By definition, $\gamma_\star$ is a probability density satisfying

$$hL(x,v) + \phi(x+hv) + \epsilon \ln \gamma_\star(v) = -\lambda\,. \tag{11}$$

We claim that $\gamma_\star$ is optimal. Indeed, by the convexity of the function $t \mapsto t \ln t$, we have, for any other probability density γ, that

$$\begin{aligned}
&\int_{\mathbb{R}^N} (h\, L(x,v) + \phi(x+hv) + \epsilon \ln \gamma)\, \gamma(v) dv \geq \\
&\quad \int_{\mathbb{R}^N} (h\, L(x,v) + \phi(x+hv) + \epsilon \ln \gamma_\star)\, \gamma_\star(v) dv + \\
&\quad \int_{\mathbb{R}^N} (h\, L(x,v) + \phi(x+hv) + \epsilon \ln \gamma_\star + \epsilon)\, (\gamma(v) - \gamma_\star(v)) dv.
\end{aligned}$$

From (11) and the fact that both $\gamma(v)$ and $\gamma_\star$ are probability densities, we conclude that the last integral vanishes, and so $\gamma_\star$ is optimal.

The last claim in Proposition 1 follows from (11) and the fact that $\gamma_\star$ is a probability density. □

From Proposition 1, we easily conclude that:

Theorem 2. *We have that*

$$\begin{aligned}
&\inf_{\gamma \in \mathcal{D}} \int [h\, L(x,v) + \phi(x+hv) + \epsilon \ln \gamma(v)]\, \gamma(v)\, dv \\
= \quad & -\epsilon \ln \left[\int_{\mathbb{R}^N} e^{-\frac{h\, L(x,v) + \phi(x+hv)}{\epsilon}}\, dv \right].
\end{aligned}$$

In particular, the schemes in (4) and (7) are equivalent.

4. Generalized Hopf-Cole transformation

In this section, we discuss a change of variables which transforms the non-linear evolution operator $\mathcal{G}$ into a linear evolution operator $\mathcal{L}$. This procedure is related with the classical Hopf-Cole transformation (see, for instance, §4.4.1 of [Eva98]).

For this, we define the *linear* operator

$$\mathcal{L}[\psi](x) = \int_{\mathbb{R}^N} e^{-\frac{h\, L(x,v)}{\epsilon}} \psi(x+hv)\, dv\,. \tag{12}$$

We will relate $\mathcal{L}$ and $\mathcal{G}$ through the exponential transformation:

$$\mathcal{E}[\phi](x) := e^{-\frac{\phi(x)}{\epsilon}}\,. \tag{13}$$

Then, it is easy to see that:

Theorem 3. *We have*

$$\mathcal{L} \circ \mathcal{E} = \mathcal{E} \circ \mathcal{G}\,. \tag{14}$$

In particular, if

$$\psi_n(x) := e^{-\frac{\phi_n(x)}{\epsilon}},$$

the schemes in (4) and (7) are equivalent to the scheme

$$\psi_{n+1} = \mathcal{L}[\psi_n]\,.$$

We omit the details of the elementary calculation needed to prove (14) and hence Theorem 3.

It is also convenient to introduce the rescaled linear operator

$$\begin{aligned}\hat{\mathcal{L}}[\psi](x) &:= \frac{\mathcal{L}[\psi](x)}{\int_{\mathbb{R}^N} e^{-\frac{h\,L(x,v)}{\epsilon}}\,dv}\\ &= \frac{\int_{\mathbb{R}^N} e^{-\frac{h\,K(v)}{\epsilon}}\psi(x+hv)\,dv}{\int_{\mathbb{R}^N} e^{-\frac{h\,K(v)}{\epsilon}}\,dv}\,.\end{aligned}$$

Accordingly, we may consider the iteration

$$\hat{\psi}_{n+1} = \hat{\mathcal{L}}[\hat{\psi}_n]\,. \tag{15}$$

Notice that $\hat{\mathcal{L}}$ is independent on the potential U.

As an application of the generalized Hopf-Cole transformation we now show the following property, which is somehow related with the fact that the minimum of viscosity solutions is a viscosity solution.

Proposition 4. *Assume that ϕ_n and φ_n are solutions to the iterative scheme*

$$\phi_{n+1} = \mathcal{G}[\phi_n] \qquad \varphi_{n+1} = \mathcal{G}[\varphi_n]\,.$$

Then so is

$$\eta_n = -\epsilon \ln\left(e^{-\frac{\phi_n}{\epsilon}} + e^{-\frac{\varphi_n}{\epsilon}}\right).$$

Proof. This follows from (14) and the linearity of the scheme $\mathcal{L}$. □

We observe that

$$-\epsilon \ln\left(e^{-\frac{\phi_n}{\epsilon}} + e^{-\frac{\varphi_n}{\epsilon}}\right) \to \min\{\phi_n, \varphi_n\},$$

as $\epsilon \to 0^+$, which is consistent with the fact that the minimum of viscosity solutions of the terminal value problem is a viscosity[3] solution.

[3]Indeed, if ϕ_1 and ϕ_2 are solutions of (2), so is $\phi := \min\{\phi_1, \phi_2\}$. To confirm this, note that

$$\int_t^T L(q(\tau), \dot{q}(\tau))\,d\tau + \phi_i(q(T), T) \geq \int_t^T L(q(\tau), \dot{q}(\tau))\,d\tau + \phi(q(T), T)\,,$$

Note that the fact that $\mathcal{L}[\lambda\psi] = \lambda\mathcal{L}[\psi]$ reflects into $\mathcal{G}[\phi+\lambda] = \mathcal{G}[\phi]+\lambda$, that is $\mathcal{G}$ commutes with addition of constants. This elementary fact will be further stressed and used in the sequel.

5. Some properties of $\mathcal{G}$

We now point out some features of $\mathcal{G}$, such as algebraic properties, smoothness, analytic bounds and semiconcavity estimates. In particular, a fine *contraction property* is established in Theorem 9 and a uniform *bound on the semiconcavity modulus* is given in Theorem 13. The bounds for the semiconcavity modulus are essential to prove results such as existence of fixed points and convergence of the scheme, among others.

The first three properties follow directly from the definition and we thus omit the proof:

Proposition 5. *If ϕ_0 is $\mathbb{Z}^N$-periodic, then so is $\phi_n := \mathcal{G}^n[\phi_0]$ for any n. If U and ϕ_0 belong to $C^1(\mathbb{R}^N)$, then so does ϕ_n for any n.*

Proposition 6 (Monotonicity)**.** *If $\phi \geq \phi_*$, then $\mathcal{G}[\phi] \geq \mathcal{G}[\phi_*]$.*

Proposition 7. *$\mathcal{G}[\phi + r] = \mathcal{G}[\phi] + r$, for any $r \in \mathbb{R}$.*

We now point out that $\mathcal{G}$ is a (non-strict) contraction on $L^\infty(\mathbb{R}^N)$:

Proposition 8. *For any $\phi, \phi_* \in L^\infty(\mathbb{R}^N)$,*

$$\|\mathcal{G}[\phi] - \mathcal{G}[\phi_*]\|_{L^\infty(\mathbb{R}^N)} \leq \|\phi - \phi_*\|_{L^\infty(\mathbb{R}^N)} .$$

Proof. Fix $x \in \mathbb{R}^N$. Without loss of generality we may assume that

$$\mathcal{G}[\phi](x) \geq \mathcal{G}[\phi_*](x) ,$$

---

for $i = 1, 2$ and thus, by taking the infimum as in (2),

$$\phi_i(x,t) \geq \inf \int_t^T L(q(\tau), \dot q(\tau))\, d\tau + \phi(q(T), T) ,$$

for $i = 1, 2$. On the other hand, fixed any $a > 0$, there exists a suitable q_a so that, possibly intechanging the indexes of ϕ_i,

$$\begin{aligned} a + \inf \int_t^T & L(q(\tau), \dot q(\tau))\, d\tau + \phi(q(T), T) \\ \geq \int_t^T & L(q_a(\tau), \dot q_a(\tau))\, d\tau + \phi(q_a(T), T) \\ = \int_t^T & L(q_a(\tau), \dot q_a(\tau))\, d\tau + \phi_1(q_a(T), T) \\ & \geq \phi_1(x,t) \;\geq\; \phi(x,t) . \end{aligned}$$

These computations show that ϕ is a solution of (2).

thus

$$
\begin{aligned}
|\mathcal{G}[\phi](x)-\mathcal{G}[\phi_*](x)| &= \epsilon \ln \frac{\int_{\mathbb{R}^N} e^{-\frac{L(x,v)+\phi_*(x+hv)}{\epsilon}}\, dv}{\int_{\mathbb{R}^N} e^{-\frac{L(x,v)+\phi(x+hv)}{\epsilon}}\, dv} \\
&\leq \epsilon \ln \frac{\int_{\mathbb{R}^N} e^{-\frac{L(x,v)+\phi(x+hv)}{\epsilon}}\, e^{\frac{\|\phi-\phi_*\|_{L^\infty(\mathbb{R}^N)}}{\epsilon}}\, dv}{\int_{\mathbb{R}^N} e^{-\frac{L(x,v)+\phi(x+hv)}{\epsilon}}\, dv} \\
&= \|\phi-\phi_*\|_{L^\infty(\mathbb{R}^N)}\, .
\end{aligned}
$$

□

The contraction property can be improved in the following way:

Theorem 9. *Suppose that both ϕ and ϕ_* are $\mathbb{Z}^N$-periodic and Lipschitz functions, with Lipschitz constant bounded by Λ. Then, there exists a constant $C > 0$, possibly depending on h, ϵ, Λ and other universal quantities, such that*

$$\|\mathcal{G}[\phi]-\mathcal{G}[\phi_*]\|_\sharp \;\leq\; (1 - C\, \|\phi-\phi_*\|_\sharp^N)\, \|\phi-\phi_*\|_\sharp\, , \tag{16}$$

in which we used the norm

$$\|\psi\|_\sharp := \inf_{\zeta\in\mathbb{R}} \|\psi+\zeta\|_{L^\infty(\mathbb{R}^N)}. \tag{17}$$

Proof. Let us observe that

$$\|\psi\|_\sharp \leq \|\psi\|_{L^\infty(\mathbb{R}^N)} \tag{18}$$

and

$$\|\psi\|_\sharp = \|\psi + r\|_\sharp\, , \tag{19}$$

for any function ψ and any $r \in \mathbb{R}$. Also, there exists a constant $\zeta \in \mathbb{R}$ such that

$$\|\phi+\zeta-\phi_*\|_{L^\infty(\mathbb{R}^N)} = \|\phi-\phi_*\|_\sharp\, .$$

By (19) and the fact that $\mathcal{G}$ commutes with constants (recall Propopsition 7), possibly replacing ϕ with $\phi+\zeta-\inf\phi_*$ and ϕ_* with $\phi_*-\inf\phi_*$, we may assume, without loss of generality, that

$$\|\phi-\phi_*\|_{L^\infty(\mathbb{R}^N)} = \|\phi-\phi_*\|_\sharp\, . \tag{20}$$

and that

$$\inf\phi_* = 0\, . \tag{21}$$

Notice that, by (21) and the periodicity assumption,

$$\begin{aligned} \|\phi-\phi_*\|_\sharp &\le \|\phi-\phi(0)-\phi_*\|_{L^\infty} \\ &\le \|\phi-\phi(0)\|_{L^\infty}+\|\phi_*\|_{L^\infty}\le 2\Lambda\sqrt{N}\,. \end{aligned} \tag{22}$$

Let x_* be a point maximizing

$$|\mathcal{G}[\phi](\cdot)-\mathcal{G}[\phi_*](\cdot)|\,, \tag{23}$$

and let γ_* the optimal probability measure on $\mathbb{R}^N$ that yields:

$$\mathcal{G}[\phi_*](x_*)=\int_{\mathbb{R}^N}\Big(hL(x_*,v)+\phi_*(x_*+hv)+\epsilon\ln\gamma_*\Big)\gamma_*\,dv\,. \tag{24}$$

Recall indeed that, by Proposition 1,

$$\gamma_*=e^{-\frac{hL(x_*,v)+\phi_*(x_*+hv)+\lambda_*}{\epsilon}}\,,$$

with

$$\lambda_*=\epsilon\ln\int_{\mathbb{R}^N}e^{-\frac{hL(x_*,v)+\phi_*(x_*+hv)}{\epsilon}}\,dv\,. \tag{25}$$

Also, due to (4),

$$\mathcal{G}[\phi](x_*)\le\int_{\mathbb{R}^N}\Big(hL(x_*,v)+\phi(x_*+hv)+\epsilon\ln\gamma_*\Big)\gamma_*\,dv\,. \tag{26}$$

Define, for $v\in\frac{1}{h}\mathbb{T}^N$,

$$\eta_*(x,v):=\sum_{k\in\mathbb{Z}^N}e^{-\frac{hL(x,v+k/h)+\phi_*(x+h(v+k/h))+\lambda_*}{\epsilon}}. \tag{27}$$

Note that, since ϕ_* is bounded and L grows superlinearly in v by assumption, the series above converges. Furthermore,

$$\int_{\frac{1}{h}\mathbb{T}^N}\eta_*(x,v)\,dv\;=\;1\,, \tag{28}$$

since γ_* is a probability measure on $\mathbb{R}^N$.

We claim that there exists a positive constant θ such that

$$\theta\;\le\;\eta_*(x,v) \tag{29}$$

for any $x\in\mathbb{R}^N$ and $v\in\frac{1}{h}\mathbb{T}^N$. Indeed, first note that, by the Lipschitz properties of ϕ_* and (21), we have that

$$0\le\phi_*\le C_1 \tag{30}$$

and therefore, by (25),

$$e^{\lambda_*/\epsilon}\ge C_2\,, \tag{31}$$

where $C_i > 0$ denote here suitable quantities possibly depending on N, Λ, h and ϵ.

Also, if $v \in \frac{1}{h}\mathbb{T}^N$ (and thus $|v| \le \sqrt{N}/h$), using again (30), we see that

$$e^{-\frac{hL(x,v)+\phi_*(x+hv)}{\epsilon}} \ge C_3\,. \tag{32}$$

Then, from (27), (31) and (32), we have that

$$\eta_*(x,v) \ge e^{-\frac{hL(x,v)+\phi_*(x+hv)+\lambda_*}{\epsilon}} \ge C_4\,,$$

for any $(x,v) \in \mathbb{R}^N \times \frac{1}{h}\mathbb{T}^N$. This proves (29).

Now, we use the $\mathbb{Z}^N$-periodicity of ϕ and ϕ_*, (24), (26) and (27) to deduce that

$$\begin{aligned}
&\mathcal{G}[\phi](x_*) - \mathcal{G}[\phi_*](x_*)\\
\le\;& \int_{\mathbb{R}^N} [\phi(x_*+hv) - \phi_*(x_*+hv)]\,\gamma_*(v)\,dv\\
=\;& \int_{\frac{1}{h}\mathbb{T}^N} [\phi(x_*+hv) - \phi_*(x_*+hv)]\,\eta_*(x_*,v)\,dv\\
=\;& \frac{1}{h^N}\int_{x_*+\mathbb{T}^N} [\phi(w) - \phi_*(w)]\;\eta_*\left(x_*, \frac{w-x_*}{h}\right)\,dw\,.
\end{aligned} \tag{33}$$

Note also that

$$\inf \phi - \phi_* = -\sup \phi - \phi_* = -\|\phi - \phi_*\|_\sharp\,, \tag{34}$$

thanks to (20).

We claim that there exists a set $\mathcal{N} \subset x_* + \mathbb{T}^N$ of measure larger than $\hat{C}\|\phi-\phi_*\|_\sharp^N$, for a suitable $\hat{C} > 0$, in which $\phi - \phi_* \le -\|\phi-\phi_*\|_\sharp/8$. For proving this, note that if $\phi = \phi_*$ the claim is obvious; otherwise, by (34), there is a point $\bar{x} \in x_* + \mathbb{T}^N$ so that $(\phi-\phi_*)(\bar{x}) < 0$ and $\|\phi-\phi_*\|_\sharp \le 2|(\phi-\phi_*)(\bar{x})|$. By the Lipschitz property of the functions involved, it follows that $\phi - \phi_* \le -\|\phi-\phi_*\|_\sharp/8$ in the ball of radius $\|\phi-\phi_*\|_\sharp/(4\Lambda)$ centered at $\bar{x}$. Thus, if we take

$$\mathcal{N} := B_{\|\phi-\phi_*\|_\sharp/(4\Lambda)}(\bar{x}) \cap (x_* + \mathbb{T}^N)\,,$$

the above claim follows from (22).

Then, by such a claim and (29), we deduce that

$$\int_{\mathcal{N}} \eta_*\left(x_*, \frac{w-x_*}{h}\right)\,dw \ge \underline{C}\,\|\phi-\phi_*\|_\sharp^N\,,$$

for a suitable $\underline{C} > 0$. Combining this with (33) and (28), we conclude that

$$\begin{aligned}
\mathcal{G}[\phi](x_*)-&\mathcal{G}[\phi_*](x_*)\\
&\le \frac{1}{h^N}\int_{(x_*+\mathbb{T}^N)\setminus\mathcal{N}}\Big(\phi(w)-\phi_*(w)\Big)\,\eta_*\left(x_*,\frac{w-x_*}{h}\right)\,dw\\
&\le \frac{\|\phi-\phi_*\|_\sharp}{h^N}\int_{(x_*+\mathbb{T}^N)\setminus\mathcal{N}}\eta_*\left(x_*,\frac{w-x_*}{h}\right)\,dw\\
&\le \frac{\|\phi-\phi_*\|_\sharp}{h^N}\left[h^N\int_{\frac{1}{h}\mathbb{T}^N}\eta_*\left(x_*,v\right)\,dv-\underline{C}\,\|\phi-\phi_*\|_\sharp^N\right]\\
&=\left(1-C\|\phi-\phi_*\|_\sharp^N\right)\|\phi-\phi_*\|_\sharp\,,
\end{aligned}$$

for a suitable $C > 0$. Then, by the choice of x_* performed in (23), possibly interchanging the rôles of ϕ and ϕ_*, we gather that

$$\|\mathcal{G}[\phi]-\mathcal{G}[\phi_*]\|_{L^\infty(\mathbb{R}^N)}\;\le\;\left(1-C\|\phi-\phi_*\|_\sharp^N\right)\|\phi-\phi_*\|_\sharp\,.$$

This and (18) imply (16). $\square$

We now define

$$\kappa_{\epsilon,h}\;:=\;\sup_{x\in\mathbb{R}^N}\Big|\ln\int_{\mathbb{R}^N}e^{-\frac{h\,L(x,v)}{\epsilon}}\,dv\Big|\,.$$

Note that, in many cases, $\kappa_{\epsilon,h}$ may be estimated explicitly; e.g., if $L(x,v)=K(v)=|v|^2$, then $\kappa_{\epsilon,h}=(N/2)|\ln(\epsilon\pi/h)|$.

By taking $\phi_*=0$ in Proposition 8, we gather the following

Proposition 10 (L^∞-bound)**.**

$$\|\mathcal{G}[\phi]\|_{L^\infty(\mathbb{R}^N)}\le\|\phi\|_{L^\infty(\mathbb{R}^N)}+\epsilon\,\kappa_{\epsilon,h}\,.$$

In particular,

$$\|\phi_n\|_{L^\infty(\mathbb{R}^N)}\le\|\phi_0\|_{L^\infty(\mathbb{R}^N)}+\epsilon n\,\kappa_{\epsilon,h}\,.$$

We point out that, in general, it is not possible to obtain a bound of $\|\phi_n\|_{L^\infty}$ which is uniform on n: even in the case of $L=|v|^2$ and $\phi_0=0$, which can be worked out explicitly, one obtains

$$\|\phi_n\|_{L^\infty}=\frac{\epsilon N n}{2}\,|\ln\frac{\epsilon\pi}{h}|\,. \tag{35}$$

We now study the semiconcavity of ϕ_n. The semiconcavity modulus of a function ϕ is

$$\Sigma_\phi := \sup_{\substack{x,y\in\mathbb{R}^N \\ y\neq 0}} \frac{\phi(x+y)+\phi(x-y)-2\phi(x)}{|y|^2}. \tag{36}$$

We say that a continuous function ϕ is semiconcave if $\Sigma_\phi < +\infty$.

We will prove that Σ_{ϕ_n} is uniformly bounded (see Theorem 13). Since ϕ_n is continuous if so is ϕ_0, this yields that ϕ_n is uniformly semiconcave.

The estimate on Σ_{ϕ_n} will be obtained in three steps: we first get a first rough estimate on Σ_{ϕ_n}; we then prove a semiconcavity "improvement" estimate (that is, knowing a bound on Σ_{ϕ_n}, we show that actually a slightly better bound is possible); finally, by iterating the procedure, we will be able to obtain a uniform bound on Σ_{ϕ_n}. Let us now fix a semiconcave function ϕ_0 and work out the details:

Lemma 11. *Let C_U and C_K be as in (8) and (9). Then,*

$$\Sigma_{\phi_n} \leq \frac{C_K}{2h} + \frac{h\,C_U}{2},$$

for any $n \geq 1$.

Proof. Fix $w \in \mathbb{R}^N$. By using (9) and the Cauchy-Schwarz Inequality, we get that

$$\begin{aligned}
\phi_{n+1}(x) &\geq -h\,U(x) \\
&\quad -\epsilon \ln\Big(\int_{\mathbb{R}^N} e^{-\frac{h}{2\epsilon}(K(v+w)+K(v-w)-C_K|w|^2)} e^{-\frac{\phi_n(x+hv)}{\epsilon}}\, dv \Big) \\
&= -h\,U(x) - \frac{h\,C_K}{2}|w|^2 \\
&\quad -\epsilon \ln\Big(\int_{\mathbb{R}^N} e^{-\frac{h\,K(v+w)+\phi_n(x+hv)}{2\epsilon}} e^{-\frac{h\,K(v-w)+\phi_n(x+hv)}{2\epsilon}}\, dv \Big) \\
&\geq -h\,U(x) - \frac{h\,C_K}{2}|w|^2 \\
&\quad -\frac{\epsilon}{2} \ln\Big(\int_{\mathbb{R}^N} e^{-\frac{h\,K(v+w)+\phi_n(x+hv)}{\epsilon}}\, dv \\
&\quad \cdot \int_{\mathbb{R}^N} e^{-\frac{h\,K(v-w)+\phi_n(x+hv)}{\epsilon}}\, dv \Big) \\
&= -h\,U(x) - \frac{h\,C_K}{2}|w|^2 \\
&\quad -\frac{\epsilon}{2} \ln\Big(\int_{\mathbb{R}^N} e^{-\frac{h\,K(v+w)+\phi_n(x+hv)}{\epsilon}}\, dv \Big) \\
&\quad -\frac{\epsilon}{2} \ln\Big(\int_{\mathbb{R}^N} e^{-\frac{h\,K(v-w)+\phi_n(x+hv)}{\epsilon})}\, dv \Big).
\end{aligned}$$

Thus, changing variables of integration,

$$\begin{aligned}
\phi_{n+1}(x) &\geq -h\,U(x) - \frac{h\,C_K}{2}|w|^2 \\
&\quad -\frac{\epsilon}{2} \ln\Big(\int_{\mathbb{R}^N} e^{-\frac{h\,K(\tilde v)+\phi_n(x-hw+h\tilde v)}{\epsilon}}\, d\tilde v \Big) \\
&\quad -\frac{\epsilon}{2} \ln\Big(\int_{\mathbb{R}^N} e^{-\frac{h\,K(\tilde v)+\phi_n(x+hw+h\tilde v)}{\epsilon}}\, d\tilde v \Big) \\
&= -\frac{h\,C_K}{2}|w|^2 - h\,U(x) + \frac{h}{2}\Big(U(x+hw) + U(x-hw) \Big) \\
&\quad +\frac{1}{2}\Big(\phi_{n+1}(x+hw) + \phi_{n+1}(x-hw) \Big),
\end{aligned}$$

thus, from (8),

$$\phi_{n+1}(x+hw) + \phi_{n+1}(x-hw) - 2\phi_{n+1}(x) \leq \Big(\frac{C_K}{2h} + \frac{h\,C_U}{2} \Big)\, (h|w|)^2 ,$$

from which the desired claim follows. ∎

Notice that the bound proven in Lemma 11 does not really look satisfactory for small h. We now show how to improve it.

Lemma 12 (Semiconcavity improvement)**.** *Fix $n \in \mathbb{N}$. Assume that*

$$\Sigma_{\phi_n} \leq \sigma \,. \tag{37}$$

Then,[4]

$$\Sigma_{\phi_{n+1}} \leq h\, C_U + \frac{C_K\, \sigma}{C_K + h\sigma} \,. \tag{38}$$

Proof. Fix $y \in \mathbb{R}^N$ and $\theta \in (0,1)$ and set $t := 1-\theta$. Then, changing variable in the integration, we gather that

$$\begin{aligned} \phi_{n+1}(x \pm hy) &= -h\, U(x \pm hy) \\ &\quad -\epsilon \ln \Big(\int_{\mathbb{R}^N} e^{-\frac{h}{\epsilon} K(w \mp \theta y)} e^{-\frac{1}{\epsilon}\phi_n(x+hw\pm thy)}\, dw \Big) \,, \end{aligned}$$

so that, by means of (8) and using the Cauchy-Schwarz Inequality, we infer that

$$\begin{aligned} & \phi_{n+1}(x+hy) + \phi_{n+1}(x-hy) - 2\phi_{n+1}(x) \\ \leq\ & h\, C_U (h|y|)^2 + 2\epsilon \ln \Big(\int_{\mathbb{R}^N} e^{-\frac{h}{\epsilon} K(w)} e^{-\frac{1}{\epsilon}\phi_n(x+hw)}\, dw \Big) \\ -\ & \epsilon \ln \Big(\int_{\mathbb{R}^N} e^{-\frac{h}{\epsilon} K(w-\theta y)} e^{-\frac{1}{\epsilon}\phi_n(x+hw+thy)}\, dw \\ & \cdot \int_{\mathbb{R}^N} e^{-\frac{h}{\epsilon} K(w+\theta y)} e^{-\frac{1}{\epsilon}\phi_n(x+hw-thy)}\, dw \Big) \\ \leq\ & h\, C_U (h|y|)^2 + 2\epsilon \ln \Big(\int_{\mathbb{R}^N} e^{-\frac{h}{\epsilon} K(w)} e^{-\frac{1}{\epsilon}\phi_n(x+hw)}\, dw \Big) \\ -\ & 2\epsilon \ln \Big(\int_{\mathbb{R}^N} e^{-\frac{h}{2\epsilon}(K(w-\theta y)+K(w+\theta y))} \\ & \cdot e^{-\frac{1}{2\epsilon}(\phi_n(x+hw+thy)+\phi_n(x+hw-hty))}\, dw \Big) \,. \end{aligned}$$

---

[4]Note that, if

$$\sigma \geq \Big(h\, C_U + \sqrt{h^2\, C_U^2 + 4\, C_U\, C_K} \Big)/2 \,,$$

then the right hand side of (38) is less than σ, thus we improved the semiconcavity bound from the nth step to the $(n+1)$th step. If, on the other hand,

$$\sigma \leq \Big(h\, C_U + \sqrt{h^2\, C_U^2 + 4\, C_U\, C_K} \Big)/2 \,,$$

then the right hand side of (38) provides a uniform bound for $\Sigma_{\phi_{n+1}}$. Observations of this type will play a rôle in the proof of Theorem 13 here below.

Therefore, by (9) and (37),

$$\begin{aligned} & \phi_{n+1}(x+hy)+\phi_{n+1}(x-hy)-2\phi_{n+1}(x) \\ \leq\ & h\,C_U\,(h|y|)^2+h\,C_K\,(\theta|y|)^2+\sigma\,(ht|y|)^2\,, \end{aligned}$$

hence, since $t=1-\theta$,

$$\Sigma_{\phi_{n+1}} \leq h\,C_U+\frac{C_K}{h}\,\theta^2+\sigma(1-\theta)^2\,.$$

The desired result follows now by choosing $\theta:=\sigma h/(C_K+\sigma h)$. □

We now get the uniform control on Σ_{ϕ_n} by a suitable iteration of Lemma 12 (thus improving Lemma 11 in the case of small h):

Theorem 13 (Uniform semiconcavity)**.** *Let*

$$\sigma_\star := \frac{h\,C_U+\sqrt{h^2\,C_U^2+4\,C_U\,C_K}}{2}\,.$$

Then, for any $n\in\mathbb{N}$,

$$\Sigma_{\phi_{n+1}} \leq \max\left\{\Sigma_{\phi_n}\,,\ h\,C_U+\sigma_\star\right\}.$$

In particular,

$$\Sigma_{\phi_n} \leq \max\left\{\Sigma_{\phi_0}\,,\ h\,C_U+\sigma_\star\right\}.$$

Proof. The second claim follows from the first one, by a simple iteration. To prove the first claim, we distinguish two cases: either $\Sigma_{\phi_n}\geq\sigma_\star$ or $\Sigma_{\phi_n}\leq\sigma_\star$. Let us deal first with the case $\Sigma_{\phi_n}\geq\sigma_\star$. Note that, if $\sigma\geq\sigma_\star$, then

$$h\,C_U+\frac{C_K\,\sigma}{C_K+h\sigma}\leq\sigma\,.$$

This observation and Lemma 12 imply that

$$\Sigma_{\phi_{n+1}} \leq h\,C_U+\frac{C_K\,\Sigma_{\phi_n}}{C_K+h\Sigma_{\phi_n}}\leq\Sigma_{\phi_n}\,,$$

proving the desired result when $\Sigma_{\phi_n}\geq\sigma_\star$. If, on the other hand, $\Sigma_{\phi_n}\leq\sigma_\star$, Lemma 12 implies that

$$\begin{aligned} \Sigma_{\phi_{n+1}} &\leq h\,C_U+\frac{C_K\,\Sigma_{\phi_n}}{C_K+h\Sigma_{\phi_n}} \\ &\leq h\,C_U+\Sigma_{\phi_n} \\ &\leq h\,C_U+\sigma_\star\,, \end{aligned}$$

which proves the desired result. □

We are now in position to deduce a uniform Lipschitz bound on ϕ_n.

Theorem 14 (Uniform Lipschitz bound)**.** *Let $h \in (0,1]$. Assume that ϕ_0 is $\mathbb{Z}^N$-periodic and semiconcave. Then, there exists C, depending only on Σ_{ϕ_0}, N, C_U and C_K so that*

$$|\phi_n(x) - \phi_n(y)| \leq C\,|x-y|\,, \qquad \forall x, y \in \mathbb{R}^N\,, \qquad \forall n \geq 1.$$

Proof. By Proposition 5 and Theorem 13, ϕ_n is $\mathbb{Z}^N$-periodic and uniformly semiconcave. Then, the result follows from Theorem 44. □

6. Some properties of $\mathcal{L}$ and $\hat{\mathcal{L}}$

In analogy with the results in § 5, we now point out some properties of the linear operators $\mathcal{L}$ and $\hat{\mathcal{L}}$. The following ones are quite easy to check:

Proposition 15. *If ψ_0 is $\mathbb{Z}^N$-periodic, then so is $\psi_n := \mathcal{L}^n[\psi_0]$ for any n. If $\hat{\psi}_0$ is $\mathbb{Z}^N$-periodic, then so is $\hat{\psi}_n := \hat{\mathcal{L}}[\hat{\psi}_0]$ for any n. If U and ψ_0 belongs to $C^1(\mathbb{R}^N)$, then so does ψ_n for any n. If $\hat{\psi}_0$ belongs to $C^1(\mathbb{R}^N)$, then so does $\hat{\psi}_n$ for any n.*

Proposition 16 (Monotonicity)**.** *If $\psi \geq \psi_*$, then $\mathcal{L}[\psi] \geq \mathcal{L}[\psi_*]$. If $\hat{\psi} \geq \hat{\psi}_*$, then $\hat{\mathcal{L}}[\hat{\psi}] \geq \hat{\mathcal{L}}[\hat{\psi}_*]$.*

Proposition 17. *$\mathcal{L}[\psi + r] = \mathcal{L}[\psi] + r\int_{\mathbb{R}^N} e^{-\frac{h\,L(x,v)}{\epsilon}}\,dv$, and $\hat{\mathcal{L}}[\hat{\psi} + r] = \hat{\mathcal{L}}[\hat{\psi}] + r$, for any $r \in \mathbb{R}$.*

Also, it is easy to see that $\hat{\mathcal{L}}$ is a (non-strict) contraction on $L^\infty(\mathbb{R}^N)$:

Proposition 18. *For any $\hat{\psi}, \hat{\psi}_* \in L^\infty(\mathbb{R}^N)$,*

$$\|\hat{\mathcal{L}}[\hat{\psi}] - \hat{\mathcal{L}}[\hat{\psi}_*]\|_{L^\infty(\mathbb{R}^N)} \leq \|\hat{\psi} - \hat{\psi}_*\|_{L^\infty(\mathbb{R}^N)}\,.$$

In particular:

Proposition 19 (Uniform L^∞-bound)**.**

$$\|\hat{\mathcal{L}}[\hat{\psi}]\|_{L^\infty(\mathbb{R}^N)} \leq \|\hat{\psi}\|_{L^\infty(\mathbb{R}^N)}\,.$$

Analogously,

Proposition 20 (Uniform C^k and $W^{k,\infty}$-bound)**.** *If $\hat{\psi} \in C^k(\mathbb{R}^N)$, then $\hat{\mathcal{L}}[\hat{\psi}] \in C^k(\mathbb{R}^N)$ and*

$$\|\hat{\mathcal{L}}[\hat{\psi}]\|_{C^k(\mathbb{R}^N)} \leq \|\hat{\psi}\|_{C^k(\mathbb{R}^N)}\,.$$

An analogous statement holds by replacing C^k with $W^{k,\infty}$.

We now obtain a (non-uniform in h and ϵ) C^1-bound of $\mathcal{L}[\psi]$ for the dependence of the L^∞-norm of ψ:

Proposition 21 (Improvement of regularity). *Assume that*

$$\int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} |DK(v)| \, dv \, < \, +\infty \, . \tag{39}$$

Then,

$$\left\| D(\hat{\mathcal{L}}[\psi]) \right\|_{L^\infty(\mathbb{R}^N)} \leq \|\psi\|_{L^\infty(\mathbb{R}^N)} \frac{\int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} |DK(v)| \, dv}{\epsilon \int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} \, dv},$$

for any $\psi \in L^\infty(\mathbb{R}^N)$*.*

Remark. For $K(v) = |v|^2$, the result in Proposition 21 yields that

$$\left\| D(\hat{\mathcal{L}}[\psi]) \right\|_{L^\infty(\mathbb{R}^N)} \leq \frac{\text{const} \, \|\psi\|_{L^\infty(\mathbb{R}^N)}}{\sqrt{\epsilon h}} .$$

Proof. By changing variable, we have that

$$\hat{\mathcal{L}}[\psi](x) \, = \, \frac{\int_{\mathbb{R}^N} e^{-\frac{h}{\epsilon} K((w-x)/h)} \psi(w) \, dw}{h^N \int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} \, dv} .$$

Therefore,

$$\begin{aligned} \left| D_x(\hat{\mathcal{L}}[\psi]) \right| \; & \leq \; \frac{\int_{\mathbb{R}^N} e^{-\frac{h}{\epsilon} K((w-x)/h)} |DK((w-x)/h)| \, |\psi(w)| \, dw}{\epsilon h^N \int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} \, dv} \\ & = \; \frac{\int_{\mathbb{R}^N} e^{-\frac{h}{\epsilon} K(v)} |DK(v)| \, |\psi(x+hv)| \, dv}{\epsilon \int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} \, dv}, \end{aligned}$$

which implies the desired result. □

Further regularity results can be also deduced in an analogous way from Proposition 21. Also, it is straightforward to see that that $\hat{\mathcal{L}}$ preserves the average over $\mathbb{T}^N$ of $\mathbb{Z}^N$-periodic function, namely:

Proposition 22. *If* $\psi \in L^1(\mathbb{T}^N)$*, then*

$$\int_{\mathbb{T}^N} \hat{\mathcal{L}}[\psi] \, = \, \int_{\mathbb{T}^N} \psi \, .$$

We remark that the above bounds and the Theorem of Ascoli imply the uniform convergence, up to a subsequence, of the iteration $\hat{\mathcal{L}}^n[\psi_0]$. In particular, if ψ_0 is $\mathbb{Z}^N$-periodic, Propositions 20 (or 21) and 22 imply that $\hat{\mathcal{L}}^n[\psi_0]$ converges, when $n \to +\infty$, up to a subsequence, to a fixed point of $\hat{\mathcal{L}}$ with the same $\mathbb{T}^N$-average of ψ_0.

7. Formal Asymptotics

In this section, we discuss the formal asymptotic behavior of the schemes corresponding to $\hat{\mathcal{L}}$, $\mathcal{L}$ and $\mathcal{G}$. These computations play a crucial rôle in the rigorous convergence results given in § 9.

We deal with different choices of the parameters ϵ and h. Elliptic, parabolic and Hamilton-Jacobi-type equations will show up in these asymptotics.

Proposition 23 (Asymptotics for $\hat{\mathcal{L}}$)**.** *Let*

$$b_i := \frac{\int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} v_i dv}{\int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} dv}, \qquad a_{ij} := \frac{h \int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} v_i v_j dv}{2 \int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} dv}.$$

Suppose that $|b_i| + |a_{ij}| < +\infty$*. Let*

$$\hat{\psi}_{n+1} = \hat{\mathcal{L}}[\hat{\psi}_n].$$

Then,

$$\frac{\hat{\psi}_{n+1} - \hat{\psi}_n}{h} = b_i D_{x_i} \hat{\psi}_n + a_{ij} D^2_{x_i x_j} \hat{\psi}_n + \textit{error terms}. \tag{40}$$

The error terms can be estimated by:

$$\text{const}\, h^2 \|D^3 \hat{\psi}_n\|_{L^\infty(\mathbb{R}^N)} \frac{\int |v|^3 e^{-\frac{hK(v)}{\epsilon}} dv}{\int e^{-\frac{hK(v)}{\epsilon}} dv}.$$

Remark. The left hand side of (40) may be thought as a discrete approximation of $\partial_t \hat{\psi}$ as $h \to 0$. Though the coefficients in (40) depend on the specific form of the kinetic energy and on the relative scale of ϵ and h, it may be useful to work out the case $K(v) = |v|^2$ somehow explicitly. In this case, the expressions in Proposition 23 become $b_i = 0$, $a_{ij} = \text{const}\, \epsilon\, \delta_{ij}$ and the error term is bounded by $\text{const}\, h^{1/2} \epsilon^{3/2} \|D^3 \psi_n\|_{L^\infty}$.

Proof. We have

$$\begin{aligned}
\frac{\hat{\psi}_{n+1} - \hat{\psi}_n}{h} &= \frac{\int e^{-\frac{hK(v)}{\epsilon}} \left[\hat{\psi}_n(x+hv) - \hat{\psi}_n(x)\right] dv}{h \int e^{-\frac{hK(v)}{\epsilon}} dv} \\
&= c_{\epsilon,h} \left[\int e^{-\frac{hK(v)}{\epsilon}} v_i D_{x_i} \hat{\psi}_n(x) dv + \frac{h}{2} \int e^{-\frac{hK(v)}{\epsilon}} v_i v_j D^2_{x_i x_j} \hat{\psi}_n(x) dv\right] \\
&\quad + O\left(h^2 c_{\epsilon,h} \|D^3 \psi_n\|_\infty \int |v|^3 e^{-\frac{hK(v)}{\epsilon}} dv\right),
\end{aligned}$$

with

$$c_{\epsilon,h} = \left[\int e^{-\frac{hK(v)}{\epsilon}} dv\right]^{-1}.$$

□

As an application of this theorem, we study the fixed points[5] of the scheme (15). Assume, for the moment, that $h = \epsilon$. Observe that if we define

$$a_{ij} = \frac{1}{2}\int e^{-K(v)} v_i v_j,$$

the operator

$$a_{ij}\partial_i\partial_j$$

is (possibly degenerate) elliptic, as

$$a_{ij}\xi_i\xi_j = \frac{1}{2}\int e^{-K(v)} v_i\xi_i v_j\xi_j = \frac{1}{2}\int e^{-K(v)}|v\cdot\xi|^2 \geq 0.$$

Furthermore, if K is even, then $b_i = 0$. In this case, the fixed points ψ of the scheme (15) satisfy

$$a_{ij}\partial_i\partial_j\psi = O(h\|D^3\psi\|^2),$$

as it follows from the estimate of Proposition 23.

Proposition 24 (Asymptotics for $\mathcal{L}$). *Let b_i and a_{ij} be as in Proposition 23. Suppose that*

$$\psi_{n+1} = \mathcal{L}[\psi_n]\,.$$

[5]Note that any constant is a fixed point for $\hat{\mathcal{L}}$. Moreover, if K is even in any of each variables, given a quadratic polynomial

$$p(x) = b + \ell\cdot x + Mx\cdot x\,,$$

with $b\in\mathbb{R}$, $\ell\in\mathbb{R}^N$ and $M\in \mathrm{Mat}(N\times N)$, one easily sees that p is a fixed point for $\hat{\mathcal{L}}$ if and only if

$$a_j\partial_j^2 p \,=\, 0\,,$$

where

$$a_j \,:=\, \frac{\int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}}\, v_j^2\, dv}{\int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}}\, dv}\,.$$

In particular, for $K(v) = |v|^2$, p is a fixed point of $\hat{\mathcal{L}}$ if and only if it is harmonic.

Then,

$$\begin{aligned}\frac{\psi_{n+1}-\psi_n}{h} &= -\frac{U}{\epsilon}\psi_n \\ &+ \left(1-\frac{hU}{\epsilon}\right)\left(b_i D_{x_i}\psi_n + a_{ij}D^2_{x_ix_j}\psi_n\right) \\ &+ \textit{error terms}.\end{aligned}$$

The error terms can be estimated by

$$\begin{aligned}&\text{const}\left[\left\|e^{-\frac{hU}{\epsilon}}-1+\frac{h}{\epsilon}U\right\|_{L^\infty}\|\hat{\mathcal{L}}[\psi_n]\|_{L^\infty}\right.\\ &\left.+h^2\left\|1-\frac{hU}{\epsilon}\right\|_{L^\infty}\|D^3\hat{\psi}_n\|_{L^\infty}\frac{\int|v|^3e^{-\frac{hK(v)}{\epsilon}}dv}{\int e^{-\frac{hK(v)}{\epsilon}}dv}\right].\end{aligned}$$

Remark. If $K(v)=|v|^2$, formally, this means that, in the limit $h\to 0$,

$$\psi_t \simeq -\frac{U}{\epsilon}\psi + \text{const}\,\epsilon\,\Delta\psi.$$

Proof. We have

$$\psi_{n+1}=\mathcal{L}[\psi_n]=e^{-\frac{hU(x)}{\epsilon}}\hat{\mathcal{L}}[\psi_n].$$

Therefore,

$$\begin{aligned}\psi_{n+1}-\psi_n &= \left(1-\frac{hU}{\epsilon}\right)\left(\hat{\mathcal{L}}[\psi_n]-\psi_n\right)-\frac{hU}{\epsilon}\psi_n \\ &\quad -\left(1-e^{-\frac{hU}{\epsilon}}-\frac{hU}{\epsilon}\right)\hat{\mathcal{L}}[\psi_n].\end{aligned}$$

The desired result thus follows from Proposition 23. □

Proposition 25 (Asymptotics for $\mathcal{G}$)**.** *Let $\phi\in C^3(\mathbb{R}^N)$. Assume that $K(v)=|v|^2$ and that*

(41) $h\,\|D^2\phi\|_{L^\infty(\mathbb{R}^N)}$ *is smaller than a suitably constant.*

Let

$$a := \frac{\int_{\mathbb{R}^N} e^{-|w|^2}\,|w|^2\,dw}{2\int_{\mathbb{R}^N} e^{-|w|^2}\,dw}. \tag{42}$$

Then,

$$\frac{\mathcal{G}[\phi]-\phi}{h} = -H(D\phi(x),x)+a\epsilon\Delta\phi-\frac{\epsilon N}{2h}\ln\frac{\epsilon\pi}{h}+\textit{error terms}. \tag{43}$$

The error terms can be estimated by

$$C_\star\,(h+\epsilon^{3/2}\,h^{1/2})\,,$$

where $C_\star$ is a suitable positive quantity depending only on $\|D^j\phi\|_{L^\infty}$ for $1\le j\le 3$.

REMARK. There are several heuristic ways of interpreting the result in Proposition 25 according to the scheme $\phi_{n+1}:=\mathcal{G}[\phi_n]$. The first consists in taking $h\to 0$ and $\epsilon/h\to 0$, so that, formally, (43) goes to the Hamilton-Jacobi equation

$$\partial_t\phi\;=\;-H(D_x\phi(x),x)\,.$$

Another possibility is taking $\epsilon>0$ to be fixed and sending $h\to 0$. Then, if we define

$$\phi_n^*:=\phi_n+\frac{\epsilon N}{2}\ln\frac{\epsilon\pi}{h},$$

it follows that (43) goes formally to

$$\partial_t\phi^*\;=\;-H(D_x\phi^*(x),x)+a\epsilon\Delta\phi^*\,,$$

that is, a Hamilton-Jacobi equation of "vanishing viscosity" type. The reader may compare this case with the example discussed in (35).

This observation may also be rephrased as follows. A direct computation shows that

$$\mathcal{G}[0](x)=-hU(x)-\frac{\epsilon N}{2}\ln\frac{\epsilon\pi}{h}\,.$$

Therefore, if we define

$$\hat{\mathcal{G}}[\phi]\;:=\;\mathcal{G}[\phi]-\mathcal{G}[0]\,,\tag{44}$$

then Proposition 25 gives that

$$\frac{\hat{\mathcal{G}}[\phi]-\phi}{h}=-H(D\phi(x))+a\epsilon\Delta\phi+\text{error terms}\,.\tag{45}$$

The latter interpretation is related with Proposition 8, which gives that

$$\|\hat{\mathcal{G}}[\phi]\|_{L^\infty(\mathbb{R}^N)}\le\|\phi\|_{L^\infty(\mathbb{R}^N)}\,,\tag{46}$$

that is, a uniform bound on the $\hat{\mathcal{G}}$-iterations. We will use these observations in Theorem 30 below, where a formal proof of the convergence of the $\hat{\mathcal{G}}$-scheme for $h\to 0$ will be provided.

Proof. We have

$$\mathcal{G}[\phi] = -\epsilon \ln \int e^{-\frac{hL(x,v)+\phi(x+hv)}{\epsilon}} dv \tag{47}$$
$$= -hU(x) - \epsilon \ln \int e^{-\frac{h|v|^2+\phi(x+hv)}{\epsilon}} dv.$$

Define $v^* = v^*(x,h)$ to be a minimizer of the function

$$h\,|v|^2 + \phi(x+hv)\,.$$

Note that the minimizing property of v^* yields that

$$h\,|v^*|^2 + \phi(x+hv^*) \le \phi(x)$$

and so

$$h\,|v^*|^2 \le h\,\|D\phi\|_{L^\infty}|v^*|\,,$$

that is

$$|v^*| \;\le\; \|D\phi\|_{L^\infty}\,. \tag{48}$$

By the minimizing property of v^*, it also follows that

$$2v^* + D_x\phi(x+hv^*) = 0 \tag{49}$$

and thus

$$\frac{h|v|^2+\phi(x+hv)}{\epsilon} \tag{50}$$
$$= \frac{[h|v^*|^2+\phi(x+hv^*)] + h\,|v-v^*|^2}{\epsilon} + \Xi_1(x,v)\,,$$

with

$$\Xi_1(x,v) \;:=\; \frac{\phi(x+hv) - \phi(x+hv^*) - h\,D_x\phi(x+hv^*)\cdot(v-v^*)}{\epsilon}\,. \tag{51}$$

Note that, by construction,

$$|\Xi_1| \le C\frac{h^2\|D^2\phi\|_{L^\infty}}{\epsilon}\,|v-v^*|^2\,, \tag{52}$$

where C, here and in the remainder of the proof, we will denote an appropriate universal constant (possibly taking a different value at different steps of the computation). We will also use the following short hand notation:

$$⨍ f(v)\,dv := \frac{\int_{\mathbb{R}^N} e^{-\frac{h|v-v^*|^2}{\epsilon}} f(v)\,dv}{\int_{\mathbb{R}^N} e^{-\frac{h|v|^2}{\epsilon}}\,dv}\,.$$

Let also

$$\Xi_2(x) := \fint (e^{-\Xi_1} - 1 + \Xi_1)\,. \tag{53}$$

Then, by (52) and (41), we get that

$$\begin{aligned} |\Xi_2| &\le \fint \sum_{k=2}^{\infty} \frac{|\Xi_1|^k}{k!} \\ &\le \fint |\Xi_1|^2 \sum_{k=0}^{\infty} \frac{|\Xi_1|^k}{k!} \\ &= \fint |\Xi_1|^2 e^{\Xi_1} \\ &\le \frac{C\,h^4\,\|D^2\phi\|_{L^\infty}^2}{\epsilon^2} \fint |v-v^*|^4 e^{\frac{C\,h^2\|D^2\phi\|_{L^\infty}|v-v^*|^2}{\epsilon}} \\ &\le C\,h^2\,\|D^2\phi\|_{L^\infty}^2 \frac{\int_{\mathbb{R}^N} e^{-\frac{|w|^2}{8}}\,|w|^4\,dw}{\int_{\mathbb{R}^N} e^{-|w|^2}\,dw} \\ &\le C\,h^2\,\|D^2\phi\|_{L^\infty}^2\,. \end{aligned} \tag{54}$$

We also define

$$\Theta_1(x,v) := -hU(x) + h|v^*|^2 + \phi(x+hv^*)\,. \tag{55}$$

Using this definition, (47), (50) and (53), we obtain that

$$\begin{aligned} \mathcal{G}[\phi] &= \Theta_1(x,v) - \epsilon \ln\left(\int_{\mathbb{R}^N} e^{-\frac{h\,|v-v^*|^2}{\epsilon}}\, e^{-\Xi_1(x,v)}\,dv\right) \\ &= \Theta_1 - \epsilon\ln\left(\fint e^{-\Xi_1(x,v)}\,dv\right) - \epsilon\ln\left(\int_{\mathbb{R}^N} e^{-\frac{h|v|^2}{\epsilon}}\,dv\right) \\ &= \Theta_1 - \epsilon\ln\left(\fint (1-\Xi_1)\,dv + \Xi_2(x)\right) - \frac{\epsilon N}{2}\,\ln\frac{\epsilon\pi}{h} \\ &= \Theta_1 - \epsilon\ln\left(1 - \fint \Xi_1\,dv + \Xi_2(x)\right) - \frac{\epsilon N}{2}\ln\frac{\epsilon\pi}{h}\,. \end{aligned} \tag{56}$$

We now define

$$\Xi_3(x) := \Xi_2(x) - \fint \Xi_1(x,v)\,dv\,. \tag{57}$$

Note that

$$|\Xi_3| \le C\,h\,\|D^2\phi\|_{L^\infty}\,, \tag{58}$$

thanks to (52), (54) and (41). Let also

$$\Xi_4 := \ln(1+\Xi_3) - \Xi_3\,. \tag{59}$$

Then,

$$|\Xi_4| \leq C\,h^2\,\|D^2\phi\|^2_{L^\infty}\,, \tag{60}$$

due to (58). Also, using (56), (57) and (59), we conclude that

$$\begin{aligned}\mathcal{G}[\phi] &= \Theta_1 - \frac{\epsilon N}{2}\,\ln\frac{\epsilon\pi}{h} - \epsilon\,\Xi_3 - \epsilon\,\Xi_4 \\ &= \Theta_1 - \frac{\epsilon N}{2}\ln\frac{\epsilon\pi}{h} + \epsilon \fint \Xi_1 - \epsilon(\Xi_2+\Xi_4).\end{aligned} \tag{61}$$

Since Ξ_2 and Ξ_4 have the same order of magnitude, it is convenient to define

$$\Xi_5 := \Xi_2 + \Xi_4\,,$$

so that (54) and (60) give

$$|\Xi_5| \leq C\,h^2\,\|D^2\phi\|^2_{L^\infty}, \tag{62}$$

and we obtain from (61) that

$$\mathcal{G}[\phi] = \Theta_1 - \frac{\epsilon N}{2}\,\ln\frac{\epsilon\pi}{h} + \epsilon \fint \Xi_1\,dv - \epsilon\,\Xi_5\,. \tag{63}$$

We now define

$$\begin{aligned}\Xi_6(x,v) &:= \Xi_1(x,v) - \frac{h^2}{2\epsilon}\partial_{ij}\phi(x)\,(v_i - v_i^*)(v_j - v_j^*) \qquad \text{and} \\ \Xi_7(x) &:= \fint \Xi_6(x,v)\,dv\,.\end{aligned} \tag{64}$$

Note that, from (51) and (48),

$$\begin{aligned}|\Xi_6| &\leq \left|\Xi_1 - \frac{h^2}{2\epsilon}\partial_{ij}\phi(x+hv^*)\,(v_i - v_i^*)(v_j - v_j^*)\right| \\ &\quad + \frac{Ch^3}{\epsilon}\,|v^*|\,\|D^3\phi\|_{L^\infty}|v-v^*|^2 \\ &\leq \frac{Ch^3}{\epsilon}\,\|D^3\phi\|_{L^\infty}\,(|v-v^*| + \|D\phi\|_{L^\infty})\,|v-v^*|^2\end{aligned}$$

and so

$$|\Xi_7| \leq C\,h^2\,\|D^3\phi\|_{L^\infty}\,(\sqrt{\epsilon/h} + \|D\phi\|_{L^\infty})\,. \tag{65}$$

Moreover,

$$\fint \Xi_1\,dv = ah\Delta\phi + \Xi_7\,,$$

due to (64), (42) and a parity argument. Therefore, we deduce from (63) that

$$\mathcal{G}[\phi] = \Theta_1 - \frac{\epsilon N}{2} \ln \frac{\epsilon \pi}{h} + a\epsilon h \Delta\phi + \epsilon\, \Xi_7 - \epsilon\, \Xi_5 \,.$$

Accordingly, if $\Xi_8 := \Xi_7 - \Xi_5$, we have that

$$\mathcal{G}[\phi] = \Theta_1 - \frac{\epsilon N}{2} \ln \frac{\epsilon \pi}{h} + a\epsilon h \Delta\phi + \epsilon\, \Xi_8 \,. \tag{66}$$

Also, by (62) and (65),

$$|\Xi_8| \,\le\, Ch^2 \Big[\|D^2\phi\|_{L^\infty} + \|D^3\phi\|_{L^\infty} (\sqrt{\epsilon/h} + \|D\phi\|_{L^\infty}) \Big] \,. \tag{67}$$

We now define

$$\begin{aligned} \Theta_2(x, v^*) &= \frac{h}{4} \Big(|D\phi(x)|^2 - |D\phi(x + hv^*)|^2 \Big) \\ \Theta_3(x, v^*) &= \phi(x) - \frac{h}{4} \, |D\phi(x + hv^*)|^2 - h|v^*|^2 - \phi(x + hv^*) \,, \end{aligned}$$

so that[6], by (55),

$$\Theta_1 = -hH(D\phi(x), x) + \phi(x) + \Theta_2 - \Theta_3 \,. \tag{68}$$

Note that

$$\begin{aligned} |\Theta_2| &\le 2h \, \|D\phi\|_{L^\infty} \Big| D\phi(x + hv^*) - D\phi(x) \Big| \\ &\le C\, h^2 \, \|D\phi\|_{L^\infty} \, \|D^2\phi\|_{L^\infty} \, |v^*| \\ &\le C\, h^2 \, \|D\phi\|^2_{L^\infty} \, \|D^2\phi\|_{L^\infty} \,, \end{aligned}$$

thanks to (48). Furthermore, by using (48) and (49), one has that

$$\begin{aligned} |\Theta_3| &\le |\phi(x) - \phi(x + hv^*) + h\, D\phi(x + hv^*) \cdot v^*| \\ &\quad + h \left| \frac{1}{4} |D\phi(x + hv^*)|^2 + |v^*|^2 + D\phi(x + hv^*) \cdot v^* \right| \\ &\le C\, h^2 \, \|D^2\phi\|_{L^\infty} \, |v^*|^2 \, + \, 0 \\ &\le C\, h^2 \, \|D\phi\|^2_{L^\infty} \, \|D^2\phi\|_{L^\infty} \,. \end{aligned}$$

Therefore, setting

$$\Theta_* \,:=\, \Theta_2 - \Theta_3 + \epsilon\, \Xi_8 \,, \tag{69}$$

the estimates above and (67) yield that

$$\begin{aligned} |\Theta_*| &\le C\, h^2 \Big\{ \|D\phi\|^2_{L^\infty} \, \|D^2\phi\|_{L^\infty} \\ &\quad + \epsilon \Big[\|D^2\phi\|_{L^\infty} + \|D^3\phi\|_{L^\infty} (\sqrt{\epsilon/h} + \|D\phi\|_{L^\infty}) \Big] \Big\} \,. \end{aligned} \tag{70}$$

[6] Of course, if $K(v) = |v|^2$, one has that $H(p, x) = |p|^2/4 + U(x)$.

Also, by collecting the identities in (66), (68) and (69), we have that

$$\mathcal{G}[\phi] = -hH(D\phi(x), x) + \phi(x) - \frac{\epsilon N}{2} \ln \frac{\epsilon\pi}{h} + a\epsilon h\Delta\phi + \Theta_* .$$

This and (70) yield the desired result. □

The previous result immediately applies to a fixed point[7] problem:

Corollary 26. *Suppose that*

$$\mathcal{G}[\phi] = \phi + \lambda_\star . \tag{71}$$

Then, we have

$$\epsilon a\Delta\phi - H(D_x\phi, x) = \frac{\lambda_\star}{h} + \frac{\epsilon N}{2h} \ln \frac{\epsilon N}{h} + \textit{error terms.}$$

The coefficient a and the error terms have the same form as in Proposition 25.

Remark. Heuristically, as $h \to 0$ and $\epsilon/h \to 0$, the result in Corollary 26 means that $\lambda_\star/h$ plays the rôle of the "effective Hamiltonian" $\overline{H}$ in the equation

$$H(D_x\phi, x) + \overline{H} = 0.$$

A formal justification for this will be given in Theorem 38.

8. Fixed point problems

We now derive from the above estimates a fixed point result on our iteration scheme $\mathcal{G}$. Namely, any iteration on the scheme $\mathcal{G}$ converges, in the $\|\cdot\|_\sharp$ defined in (17), to a function which solves a fixed point type problem, up to additive constants.

Theorem 27. *There exists $\lambda_\star \in \mathbb{R}^N$ and a $\mathbb{Z}^N$-periodic Lipschitz semiconcave function $\phi_\star$ so that*

$$\mathcal{G}[\phi_\star] = \phi_\star + \lambda_\star . \tag{72}$$

Furthermore, given any $\mathbb{Z}^N$-periodic semiconcave function ϕ_0, we have that

$$\lim_{n\to+\infty} \|\mathcal{G}^n[\phi_0] - \phi_\star\|_\sharp \;=\; 0 . \tag{73}$$

The Lipschitz constant and the semiconcavity modulus of $\phi_\star$ are bounded by a constant depending only on N, C_U and C_K.

---

[7]For the existence of functions satisfying (71), see Theorem 27. Note also that, if ϕ satisfies (71), so does $\mathcal{G}^n[\phi]$, for any $n \in \mathbb{N}$, due to Proposition 7.

Also, $\phi_\star$ and $\lambda_\star$ enjoy the following uniqueness properties: if there exist $\lambda \in \mathbb{R}^N$ and a Lipschitz $\mathbb{Z}^N$-periodic function ϕ so that $\mathcal{G}[\phi] = \phi + \lambda$, then $\phi = \phi_\star + c$ for some $c \in \mathbb{R}$ and $\lambda = \lambda_\star$.

Finally, if

$$\int_{\mathbb{R}^N} e^{-\frac{hK(v)}{\epsilon}} |DK(v)|\, dv \, < \, +\infty \, , \tag{74}$$

then the above uniqueness property holds even if we replace the assumption that ϕ is Lipschitz with the one that it is bounded.

Proof. Let us consider the quotient space $Y := C^0(\mathbb{R}^N)/\mathbb{R}$, that is, let us identify continuous functions if and only if they differ by a real number. Then, Y is a Banach space, with norm $\|\cdot\|_\sharp$ as defined in (17): see, e.g., Theorem 3.14-A on page 105 of [Tay58] for a proof of this fact. Note that the semiconcavity definition in (36) passes to the quotient in Y: thus we may take S to be the (class of the) functions in Y with semiconcavity modulus bounded by an appropriately large Σ. It is easily seen that S is closed in Y. The results in Propositions 5 and 7 also show that $\mathcal{G}$ is well defined on Y and, by Theorem 13, $\mathcal{G}$ sends S into itself. Functions in S are also uniformly Lipschitz (see Theorem 44), due to their semiconcavity properties.

Therefore, by Theorem 9 and Theorem 41, we deduce that $\mathcal{G}$ has a unique fixed point in $S \subseteq Y$ and that (73) holds. Unfolding back the quotient space Y to the cover space X, we get (72).

Let us now deal with the uniqueness property. Observe that, by changing variable of integration in (6), we have that

$$\mathcal{G}[\phi] = -\epsilon \ln \left[\frac{1}{h^N} \int_{\mathbb{R}^N} e^{-\frac{hL(x,(w-x)/h)+\phi(w)}{\epsilon}}\, dw \right]$$

and therefore, if $\mathcal{G}[\phi] = \phi + \lambda$, it follows that ϕ is Lipschitz, since so is K (recall also (74)). Therefore, by Theorem 9,

$$\begin{aligned} \|\phi - \phi_\star\|_\sharp &= \|\mathcal{G}[\phi] - \mathcal{G}[\phi_\star] - \lambda + \lambda_\star\|_\sharp \\ &= \|\mathcal{G}[\phi] - \mathcal{G}[\phi_\star]\|_\sharp \\ &\le \left(1 - C\, \|\phi - \phi_\star\|_\sharp^N\right) \|\phi - \phi_\star\|_\sharp \, , \end{aligned}$$

showing that $\|\phi - \phi_\star\|_\sharp = 0$ and thus $\phi = \phi_\star + c$.

But then, by Proposition 7,

$$\lambda - \lambda_\star = \mathcal{G}[\phi] - \phi - \mathcal{G}[\phi_\star] + \phi_\star = 0 \, .$$

□

We remark that a fixed point result may also be proved applying the Schauder Fixed Point Theorem (see, e.g., § 9.2.2 in [Eva98]) to

$$\mathcal{G}[\phi](x) - \mathcal{G}[\phi](0)\,,$$

seen as acting on the space of functions with suitably bounded L^∞ and Lipschitz norm (and such space is convex and compact by the Theorem of Ascoli). The estimates needed for making such an argument work are given by Proposition 5 and Theorem 14. This approach, however, only gives the fixed point property (72) but it does not give information either on the convergence of the iterations of $\mathcal{G}$ or on the uniqueness of the fixed points, while Theorems 27 and 41 do.

Corollary 28. *Fix $x \in \mathbb{R}^N$. Under the assumptions and the notation of Theorem 27, let the measure $\mu_\star$ be defined by*

$$d\mu_\star := e^{-\frac{h\,L(x,v)+\phi_\star(x+hv)-\phi_\star(x)-\lambda_\star}{\epsilon}}\,dv\,. \tag{75}$$

Then, $\mu_\star$ is a probability measure on $\mathbb{R}^N$. Furthermore, if $\psi_\star := \mathcal{E}[\phi_\star]$, then

$$\psi_\star(x) = e^{-\frac{\lambda_\star}{\epsilon}}\mathcal{L}[\psi_\star]\,.$$

Proof. Straightforward from (72). □

Other features of the fixed points for $\mathcal{G}$ will be discussed in § 10, where we will relate the term $\lambda_\star$ in (72) and the measure $\mu_\star$ in (75) with an entropy penalized Mather measure.

9. Convergence issues

In this section, we use a variation of the results by [Sou85] and [BS91] to prove the *convergence of the entropy penalized scheme* for small time step h. We will prove a rigorous relation between our linear (resp., non-linear) scheme and the heat equation (resp., the Hamilton-Jacobi equation).

We first deal with the convergence of the $\hat{\mathcal{L}}$-scheme as $h \to 0$, for a fixed[8] $\epsilon > 0$. For this, we stress that $\hat{\mathcal{L}}$ depends on h by writing $\hat{\mathcal{L}}_h$ (the dependence on ϵ is not explicitly written). Consider $\bar{u} \in W^{1,\infty}(\mathbb{T}^N)$ and

[8]The case when $\epsilon \to 0$ too is actually easier and can be dealt with by a modification of the proof of Theorem 29. In this circumstance, Theorem 29 holds with $\epsilon = 0$ in (77), namely $u(x,t) = \bar{u}(x)$. We omit the details of the proof, since it closely follows the one of Theorem 29.

assume, for simplicity,[9] that $1/h \in \mathbb{N}$. We define $u_h : \mathbb{T}^N \times [0,1] \longrightarrow \mathbb{R}$ in the following way. Let

$$u_h(x,0) \,:=\, \bar{u}(x)$$

and then, iteratively,

$$u_h(x,t) := \hat{\mathcal{L}}_{t-ih}[u_h(\cdot, ih)](x)$$

if $t \in (ih, (i+1)h]$, for $i = 0, \dots, (1/h)-1$. Then, we have the following convergence result:

Theorem 29. *Let $K(v) = |v|^2$ and*

$$a := \frac{\int_{\mathbb{R}^N} e^{-|v|^2} |v|^2 \, dv}{2 \int_{\mathbb{R}^N} e^{-|v|^2} \, dv} \,. \tag{76}$$

Let $u = u(x,t)$ be the unique (viscosity) solution of

$$\partial_t u(x,t) = \epsilon a \Delta_x u(x,t) \tag{77}$$

in $\mathbb{T}^N \times (0,1]$, with $u(x,0) = \bar{u}(x)$. Then, u_h converges uniformly to u.

Proof. The proof is a minor variation of the one in Theorem 2.1 of [BS91] (see also [Sou85]). Since, in our case, the arguments involved are elementary, we provide full details, for the reader's convenience.

We note that, if $t \in (ih, (i+1)h]$ and $x \in \mathbb{T}^N$

$$|u_h(x,t)| \le \sup_{x \in \mathbb{T}^N} |u_h(x, ih)|,$$

and so

$$\sup_{x \in \mathbb{T}^N,\, t \in (ih,(i+1)h]} |u_h(x,t)| \le \sup_{x \in \mathbb{T}^N,\, t \in ((i-1)h, ih]} |u_h(x,t)| \le \dots \le \|\bar{u}\|_{L^\infty(\mathbb{T}^N)} \,,$$

that is

$$\|u_h\|_{L^\infty(\mathbb{T}^N \times [0,1])} \,\le\, \|\bar{u}\|_{L^\infty(\mathbb{T}^N)} \,.$$

Therefore, we may define

$$u_-(x,t) := \liminf_{(y,s) \to (x,t),\, h \to 0} u_h(y,s) \,. \tag{78}$$

We show that u_- is a viscosity supersolution of (77). For this, fix $(x_0, t_0) \in \mathbb{T}^N \times (0,1)$ and let ϕ be a smooth function, so that $u_- - \phi$ has a strict minimum at (x_0, t_0). Let $(x_h, t_h) \in \mathbb{T}^N \times [0,1]$ be minimizers

[9] The reader may convince herself that, with minor modifications, it would be possible to consider more general partitions of an interval, instead of the uniform h-mesh that, for simplicity, we deal with.

for the function $u_h - \phi$. We show that we may choose a sequence $h_k \to 0$ such that

$$\lim_{k\to+\infty}(x_{h_k}, t_{h_k}) = (x_0, t_0)\,. \tag{79}$$

For this, let $(y_k, s_k) \to (x_0, t_0)$ and $h_k \to 0$ be a sequence so that

$$u_-(x_0, t_0) = \lim_{k\to\infty} u_{h_k}(y_k, s_k)\,,$$

according to (78). We may also assume, up to subsequences, that $(x_{h_k}, t_{h_k}) \to (\bar{x}, \bar{t}) \in \mathbb{T}^N \times [0,1]$ as $k \to +\infty$. Then, (79) is proved if we show that

$$(\bar{x}, \bar{t}) = (x_0, t_0)\,. \tag{80}$$

For this, let us observe that, by construction,

$$\begin{aligned} (u_- - \phi)(x_0, t_0) &= \lim_{k\to\infty}(u_{h_k} - \phi)(y_k, s_k) \\ &\geq \lim_{k\to\infty}(u_{h_k} - \phi)(x_{h_k}, t_{h_k}) \\ &\geq (u_- - \phi)(\bar{x}, \bar{t})\,. \end{aligned}$$

Therefore, since (x_0, t_0) was assumed to be a strict minimum for $u_- - \phi$, the above estimate proves (80) and, thence, (79).

We now consider the sequence h_k in (79), and we denote $h := h_k \to 0$ for short. The fact that

$$u_h - u_h(x_h, t_h) \geq \phi - \phi(x_h, t_h)\,,$$

together with Propositions 16 and 17, implies that

$$\hat{\mathcal{L}}_\rho[u_h(\cdot, t)](x) - u_h(x_h, t_h) \geq \hat{\mathcal{L}}_\rho[\phi(\cdot, t)](x) - \phi(x_h, t_h)$$

for any $x \in \mathbb{T}^N$, any $t \in [0,1]$ and any $\rho > 0$. Thence, if $i_h \in \{0, \dots, (1/h) - 1\}$ is so that $t_h \in (i_h h, (i_h + 1)h]$ and we set

$$\rho_h := t_h - i_h h\,,$$

we have that

$$\lim_{h\to 0}\rho_h = 0$$

and

$$\begin{aligned} 0 &= \hat{\mathcal{L}}_{\rho_h}[u_h(\cdot, t_h - \rho_h)](x_h) - u_h(x_h, t_h) \\ &\geq \hat{\mathcal{L}}_{\rho_h}[\phi(\cdot, t_h - \rho_h)](x_h) - \phi(x_h, t_h)\,. \end{aligned} \tag{81}$$

Also, since ϕ is smooth, we infer from Proposition 23 that

(82)
$$\lim_{h\to 0} \frac{\hat{\mathcal{L}}_{\rho_h}[\phi(\cdot,t_h-\rho_h)](x_h)-\phi(x_h,t_h)}{\rho_h} \;=\; a\epsilon\Delta_x\phi(x_0,t_0)-\partial_t\phi(x_0,t_0)\,.$$

By collecting (81) and (82), we get that

$$a\epsilon\Delta_x\phi(x_0,t_0)-\partial_t\phi(x_0,t_0)\le 0$$

if $u_- - \phi$ has a minimum at (x_0,t_0). That is, u_- is a viscosity supersolution of (77).

Analogously, if we define

$$u_+(x,t):=\limsup_{(y,s)\to(x,t),\,h\to 0} u_h(y,s)\,, \tag{83}$$

we have that u_+ is a viscosity subsolution of (77).

Also, if $s\in(0,h)$,

$$\begin{aligned} |u_h(y,s)-\bar u(x)| &= |\hat{\mathcal{L}}_s[\bar u](y)-\bar u(x)| \\ &\le \frac{\int_{\mathbb{R}^N} e^{-\frac{s\,|v|^2}{\epsilon}}\,|\bar u(y+sv)-\bar u(y)|\,dv}{\int_{\mathbb{R}^N} e^{-\frac{s\,|v|^2}{\epsilon}}\,dv}+|\bar u(y)-\bar u(x)| \\ &\le \|D\bar u\|_{L^\infty}\Big(\sqrt{\epsilon s}\,+\,|x-y|\Big)\,, \end{aligned}$$

and so

$$u_-(x,0)=u_+(x,0)=\bar u(x)\,.$$

We claim that

$$\begin{gathered} u_- \text{ (resp., } u_+\text{) is} \\ \text{lower semicontinuous (resp., upper semicontinuous).} \end{gathered} \tag{84}$$

For proving this, fix $z:=(x,y)\in\mathbb{T}^N\times[0,1]$ and consider a sequence $z_j\in\mathbb{T}^N\times[0,1]$, so that $z_j\to z$ as $j\to+\infty$. For each $j\in\mathbb{N}$, let also $h_{j,k}>0$ and $\zeta_{j,k}\in\mathbb{T}^N\times[0,1]$ be so that

$$\lim_{k\to+\infty} h_{j,k}=0\,,\qquad \lim_{k\to+\infty}\zeta_{j,k}=z_j \qquad\text{and}\qquad u_-(z_j)=\lim_{k\to+\infty} u_{h_{j,k}}(\zeta_{j,k})\,,$$

for any fixed $j\in\mathbb{N}$. Fix now $\eta>0$. Let $k_0(\eta,j)$ be so that

$$|u_{h_{j,k}}(\zeta_{j,k})-u_-(z_j)|\le\eta$$

for any $k \geq k_0(\eta, j)$. Let also $k_1(j)$ such that

$$|h_{j,k}| + |\zeta_{j,k} - z_j| \leq \frac{1}{j},$$

for any $k \geq k_1(j)$. We define

$$\begin{aligned} k_\star(\eta, j) &:= \max\Big\{k_0(\eta, j)\,,\ k_1(j)\Big\}\,, \\ h_j &:= h_{j,k_\star(\eta,j)} \\ \text{and} \qquad \zeta_j &:= \zeta_{j,k_\star(\eta,j)}\,. \end{aligned}$$

Then, by construction, $h_j \to 0$ and $\zeta_j \to z$ as $j \to +\infty$. Therefore,

$$u_-(z) \leq \varliminf_{j\to+\infty} u_{h_j}(\zeta_j) \leq \varliminf_{j\to+\infty} u_-(z_j) + \eta\,.$$

Since η is arbitrary, we conclude that

$$u_-(z) \leq \varliminf_{j\to+\infty} u_-(z_j)\,,$$

for any sequence $z_j \to z$, and thus u_- is lower semicontinuous. Analogously, one sees that u_+ is upper semicontinuous, thus confirming (84).

Thanks to (84), the Comparison Principle (see, e.g., Theorem 8.2 in [CIL92]) yields that $u_+ \leq u_-$. Since, by (78) and (83), the opposite inequality also holds, we obtain that

$$u_- \;=\; u_+ \;=:\; u$$

and thus u is the unique[10] (continuous viscosity) solution of (77).

To complete the proof of the desired result, we show the uniform convergence of u_h to u, by arguing as follows. If, by contradiction, u_h did not converge uniformly to u, there would exist $\eta > 0$, and $z_h \in \mathbb{T}^N \times [0,1]$ so that

$$|u_h(z_h) - u(z_h)| \geq \eta$$

for infinitely many $h \to 0$. Then, either $u_h(z_h) - u(z_h) \geq \eta$ or $u(z_h) - u_h(z_h) \geq \eta$ for infinitely many $h \to 0$. Let us assume the latter (the other case being analogous) and assume also that $z_h \to z \in \mathbb{T}^N \times [0,1]$

[10]The results we use about viscosity sub/super/solutions may be found, for instance, in [CIL92].

for this set of h's. Then,

$$\begin{aligned} u(z) &= u_-(z) \\ &\le \lim_{h\to 0} u_h(z_h) \\ &\le u(z) - \eta\,, \end{aligned}$$

which is a contradiction. □

We now deal with the scheme $\mathcal{G}$ as $h \to 0$. We define $\hat{\mathcal{G}}$ as in (44).

We will fix $\epsilon > 0$ and take $h \to 0$. We will explicilty write $\hat{\mathcal{G}}_h$ to stress the dependence on h in $\hat{\mathcal{G}}$.

We fix $\bar{w} \in C^3(\mathbb{T}^N)$, we assume[11] that $1/h \in \mathbb{N}$ and we define $w_h : \mathbb{T}^N \times [0,1] \longrightarrow \mathbb{R}$ as follows: first, we set

$$w_h(x,0) := \bar{w}(x)$$

and then, recursively,

$$w_h(x,t) := \hat{\mathcal{G}}_{t-ih}[w_h(\cdot, ih)](x)$$

if $t \in (ih, (i+1)h]$, for $i = 0, \dots, (1/h) - 1$. Then, the following convergence result holds:

Theorem 30. *Let a be as in (76). Then, as $h \to 0$, w_h converges uniformly to w, where w satisfies*

$$\partial_t w(x,t) + K(D_x w(x,t)) = a\epsilon \Delta_x w(x,t) \tag{85}$$

in the viscosity sense, with $w(x,0) = \bar{w}(x)$.

Proof. As in Theorem 29, the proof is a variation of the arguments in [Sou85] and [BS91]. First, we observe that

$$\|w_h\|_{L^\infty(\mathbb{T}^N\times[0,1])} \le \|\bar{w}\|_{L^\infty(\mathbb{T}^N)}\,,$$

due to (46). Thus, we may define

$$w_-(x,t) := \liminf_{(y,s)\to(x,t),\, h\to 0} w_h(y,s)\,.$$

We show that w_- is a viscosity supersolution of (85). For this, fix $(x_0, t_0) \in \mathbb{T}^N \times (0,1)$ and let ϕ be a smooth function, so that $w_- - \phi$ has a strict minimum at (x_0,t_0). Let $(x_h, t_h) \in \mathbb{T}^N \times [0,1]$ be minimizers for the function $w_h - \phi$. As shown in (79), we may and do assume, up to a subsequence, that $(x_h, t_h) \to (x_0, t_0)$.

[11]For the sake of simplicity, we assumed $\bar{w}$ to be smooth enough, in order to use (45) in estimate (86) here below. We remark that, by applying the lower order arguments in Proposition 25 directly to $\bar{w}$, less smoothness may be required.

Since $w_h - w_h(x_h, t_h) \geq \phi - \phi(x_h, t_h)$, by exploiting (45), we gather that

$$\hat{\mathcal{G}}_\rho[w_h(\cdot, t)](x) - w_h(x_h, t_h) \geq \hat{\mathcal{G}}_\rho[\phi(\cdot, t)](x) - \phi(x_h, t_h)$$

for any $x \in \mathbb{T}^N$, any $t \in [0,1]$ and any $\rho > 0$. Thence, if $i_h \in \{0, \dots, (1/h) - 1\}$ is so that $t_h \in (i_h h, (i_h + 1)h]$ and we set $\rho_h := t_h - i_h h$,

$$\begin{aligned} 0 &= \hat{\mathcal{G}}_{\rho_h}[w_h(\cdot, t_h - \rho_h)](x_h) - w_h(x_h, t_h) \\ &\geq \hat{\mathcal{G}}_{\rho_h}[\phi(\cdot, t_h - \rho_h)](x_h) - \phi(x_h, t_h) \,. \end{aligned}$$

Then, by means of (45), we have that

$$\begin{aligned} &\lim_{h \to 0} \frac{\hat{\mathcal{G}}_{\rho_h}[\phi(\cdot, t_h - \rho_h)](x_h) - \phi(x_h, t_h)}{\rho_h} \\ &= -K(D_x \phi(x_0, t_0)) + a\epsilon \Delta_x \phi(x_0, t_0) - \partial_t \phi(x_0, t_0) \,. \end{aligned}$$

Therefore,

$$a\epsilon \Delta_x \phi(x_0, t_0) - \partial_t \phi(x_0, t_0) \leq K(D_x \phi(x_0, t_0))$$

if $w_- - \phi$ has a minimum at (x_0, t_0). That is, w_- is a viscosity supersolution of (85).

Analogously, if we define

$$w_+(x, t) := \limsup_{(y,s) \to (x,t),\, h \to 0} w_h(y, s) \,,$$

we have that w_+ is a viscosity subsolution of (85). Also, if $s \in (0, h)$,

$$|\hat{\mathcal{G}}_s[\bar{w}] - \bar{w}| \leq C\, s \,, \tag{86}$$

thanks to (45), where C depends only on $\|\bar{w}\|_{C^3(\mathbb{T}^N)}$, and so

$$\begin{aligned} |w_h(y, s) - \bar{w}(x)| &= |\hat{\mathcal{G}}_s[\bar{w}](y) - \bar{w}(x)| \\ &\leq |\hat{\mathcal{G}}_s[\bar{w}](y) - \bar{w}(y)| + |\bar{w}(y) - \bar{w}(x)| \\ &\leq C\Big(s + |x - y|\Big) \,. \end{aligned}$$

This implies that $w_-(x, 0) = w_+(x, 0) = \bar{w}(x)$. Thus, by using the Comparison Principle and arguing as in the proof of Theorem 29, we obtain that w_h converges uniformly to $w_- = w_+$. $\square$

10. Entropy penalized Mather measures

In Mather's theory (see, e.g., [MF94] and references therein), one looks for probability measures μ on $\mathbb{T}^N \times \mathbb{R}^N$ that minimize the action

$$\int_{\mathbb{T}^N\times\mathbb{R}^N} L(x,v)d\mu(x,v) \tag{87}$$

and satisfy the "holonomy" (or "flow invariancy") constraint

$$\int_{\mathbb{T}^N\times\mathbb{R}^N} [\varphi(x+hv)-\varphi(x)]\,d\mu(x,v)=0\,, \tag{88}$$

for all $\varphi \in C(\mathbb{R}^N)$. In this section, we discuss an entropy penalized version of Mather's problem, and we present a solution in terms of fixed points of the operator $\mathcal{G}$. The entropy penalized Mather problem consists in minimizing

$$\int_{\mathbb{T}^N\times\mathbb{R}^N} L(x,v)d\mu(x,v)+\frac{\epsilon}{h}S[\mu], \tag{89}$$

in which

$$S[\mu]=\int_{\mathbb{T}^N\times\mathbb{R}^N} \mu(x,v)\ln\frac{\mu(x,v)}{\int_{\mathbb{R}^N}\mu(x,w)dw}dxdv$$

is the "entropy term". The minimization in (89) is performed over the space of probability densities on $\mathbb{T}^N \times \mathbb{R}^N$

$$\left\{\mu\in L^1(\mathbb{T}^N\times\mathbb{R}^N)\,,\ \mu\ge 0\ \ a.e.\,,\ \int_{\mathbb{T}^N\times\mathbb{R}^N} d\mu=1\right\} \tag{90}$$

that satisfy[12] the constraint (88). In this setting, we study the minimizers of the functional in (89) in the space given by (90) under the constraint in (88), that is the "entropy penalized Mather problem". We will show that these penalized minimal measures always *exist* (see Theorem 32 here below) and that they are *unique* (see Theorem 36). An *explicit formula* for this measure will be provided by using the fixed point structure of $\mathcal{G}$ (see formula (105)). In proving these results, an important rôle is played by the $\mathbb{T}^N$-projection of penalized measures and on their analytic bounds (see Propositions 34 and 35).

For this, we need the following preliminary result:

[12] As customary, with a slight abuse of notation, given a function $\mu \in L^1(\mathbb{T}\times\mathbb{R})$, we denote $d\mu(x,v)=\mu(x,v)\,dx\,dv$.

Lemma 31. *Let $\phi \in W^{1,\infty}(\mathbb{T}^N)$ satisfy*

$$\mathcal{G}[\phi] = \phi + \lambda. \tag{91}$$

Then, there exists $\theta \in L^1(\mathbb{T}^N)$ satisfying

$$\theta \,\geq\, 0\,, \tag{92}$$

$$\int_{\mathbb{R}^N} \theta(x - hv) e^{-\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}}\, dv \,=\, \theta(x) \tag{93}$$

and

$$\int_{\mathbb{T}^N} \theta(x)\, dx \,=\, 1\,. \tag{94}$$

Proof. The following elementary observation will be used throughout this proof: if $\eta \in L^1(\mathbb{T}^N)$ and $\xi \in \mathbb{R}^N$, then

$$\int_{\mathbb{T}^N} \eta \,=\, \int_{\mathbb{T}^N} \eta(x+\xi)\, dx\,.$$

Further, we note that

$$\int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\phi(x+hv)-\phi(x)-\lambda}{\epsilon}}\, dv \,=\, 1\,, \tag{95}$$

for any $x \in \mathbb{T}^N$, due to (91). Now, define for $\vartheta \in L^1(\mathbb{T}^N)$,

$$\mathcal{F}[\vartheta](x) \,:=\, \int_{\mathbb{R}^N} \vartheta(x - hv) e^{-\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}}\, dv\,. \tag{96}$$

Let us observe that (95) and (96) imply that

$$\int_{\mathbb{T}^N} \mathcal{F}[\vartheta] \,=\, \int_{\mathbb{T}^N} \vartheta\,. \tag{97}$$

Let now $\theta_0 := 1$ and, recursively,

$$\theta_{n+1} := \mathcal{F}[\theta_n]\,, \tag{98}$$

for any $n \in \mathbb{N}$. Note that θ_n is $\mathbb{Z}^N$-periodic for any $n \in \mathbb{N}$, because so are θ_0, U and ϕ. Analogously, $\theta_n \geq 0$, because so is θ_0. Thence, by (97), we have that $d\theta_n(x) = \theta_n(x)\, dx$ is a sequence of probability mesures on $\mathbb{T}^N$, that is

$$\int_{\mathbb{T}^N} d\theta_n \,=\, 1\,, \tag{99}$$

for any $n \in \mathbb{N}$. We may thus suppose that, up to subsequence, $d\theta_n$ weakly converges to a Radon measure $d\theta$ on $\mathbb{T}^N$ (see, e.g., page 55 of [EG92]). Also, $d\theta$ is a probability measure, since $\mathbb{T}^N$ is compact (see, e.g., Theorem 1-(ii) on page 54 of [EG92]). Our objective now is to show that $d\theta$ is absolutely continuous with respect to the Lebesgue measure. For this, let $\eta \in C(\mathbb{T}^N)$ be so that $0 \le \eta \le 1$. Then, exploiting (98), we get that

$$\begin{aligned} & \int_{\mathbb{T}^N} \eta(x)\, d\theta_{n+1}(x)\, dx \\ = & \int_{\mathbb{T}^N} \eta(x)\, \theta_{n+1}(x)\, dx \\ = & \int_{\mathbb{T}^N} \eta(x)\, \mathcal{F}[\theta_n](x)\, dx \\ = & \int_{\mathbb{T}^N} \int_{\mathbb{R}^N} \eta(x)\, \theta_n(x-hv)\, e^{-\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}}\, dv\, dx\,. \end{aligned} \tag{100}$$

We now denote by κ_i some appropriate positive quantities, which depend only on N, ϕ, L, λ, ϵ and h (but not on η). Then, by the fact that K is superlinear and (99), we deduce that

$$\begin{aligned} & h^N \int_{\mathbb{R}^N} \theta_n(x-hv)\, e^{-\frac{hK(v)}{\epsilon}}\, dv \\ = & \int_{\mathbb{R}^N} \theta_n(w)\, e^{-\frac{hK((x-w)/h)}{\epsilon}}\, dw \\ = & \sum_{k\in\mathbb{Z}^N} \int_{\mathbb{T}^N} \theta_n(w)\, e^{-\frac{hK((x-w-k)/h)}{\epsilon}}\, dw \\ \le & \kappa_1 \sum_{k\in\mathbb{Z}^N} \int_{\mathbb{T}^N} \theta_n(w)\, e^{-\kappa_2|k|}\, dw \\ \le & \kappa_3 \int_{\mathbb{T}^N} \theta_n(w)\, dw \\ = & \kappa_3\,. \end{aligned}$$

This estimate, together with (100), gives that

$$\int_{\mathbb{T}^N} \eta(x)\, d\theta_{n+1}(x) \le \kappa_4 \int_{\mathbb{T}^N} \eta(x)\, dx\,.$$

Taking now the limit as $n \to +\infty$, using the weak convergence of θ_n, we gather that

$$\int_{\mathbb{T}^N} \eta(x)\, d\theta(x) \le \kappa_4 \int_{\mathbb{T}^N} \eta(x)\, dx\,, \tag{101}$$

for any $\eta \in C(\mathbb{T}^N)$.

Take now any measurable set $A \subset \mathbb{T}^N$, with small Lebesgue measure, say $|A| \leq \epsilon$, for $\epsilon > 0$. We also denote the θ-measure of A by $\theta(A)$. By standard results on the Radon measure approximation with open and compact sets (see, e.g., Theorem 4 on page 8 of [EG92]), we have that there exist a compact set K and an open set U so that

$$K \subseteq A \subseteq U,$$

$$\theta(A) \leq \theta(K) + \epsilon \quad \text{and} \quad |U| \leq |A| + \epsilon \leq 2\epsilon.$$

Exploiting the classical Urysohn's Lemma (see, e.g., Theorem 2 on page 15 of [DS58]), we see that there exists a function $\eta \in C(\mathbb{T}^N)$ so that $0 \leq \eta \leq 1$, $\eta(x) = 0$ for any $x \in \mathbb{T}^N \setminus U$ and $\eta(x) = 1$ for any $x \in K$.

Accordingly,

$$\begin{aligned}
\theta(A) &\leq \theta(K) + \epsilon \\
&\leq \int_{\mathbb{T}^N} \eta(x)\, d\theta(x) + \epsilon \\
&\leq \kappa_4 \int_{\mathbb{T}^N} \eta(x)\, dx + \epsilon \\
&\leq \kappa_4 \,|U| + \epsilon \\
&\leq (2\kappa_4 + 1)\,\epsilon,
\end{aligned}$$

by means of (101).

Thence, $d\theta$ is absolutely continuous with respect to the Lebesgue measure.

Thus, we write

$$d\theta(x) = \theta(x)\, dx,$$

with $\theta(x) \in L^1(\mathbb{T}^N)$. Since $d\theta$ is, by construction, a probability measure on $\mathbb{T}^N$, we have that θ satisfies (92) and (94).

Moreover, given any $\eta \in C(\mathbb{T}^N)$, we deduce from (100) and the weak convergence of $d\theta_n$ that

$$\begin{aligned}
& \int_{\mathbb{T}^N} \eta(x)\,\theta(x)\,dx \\
= & \int_{\mathbb{T}^N} \eta(x)\,d\theta(x) \\
= & \lim_{n\to+\infty} \int_{\mathbb{T}^N} \eta(x)\,d\theta_{n+1}(x) \\
= & \lim_{n\to+\infty} \int_{\mathbb{T}^N} \int_{\mathbb{R}^N} \eta(x)\,\theta_n(x-hv)\,e^{-\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}}\,dv\,dx \\
= & \lim_{n\to+\infty} \int_{\mathbb{R}^N} \int_{\mathbb{T}^N} \eta(y+hv)\,\theta_n(y)\,e^{-\frac{hL(y,v)+\phi(y+hv)-\phi(y)-\lambda}{\epsilon}}\,dy\,dv \\
= & \lim_{n\to+\infty} \int_{\mathbb{R}^N} \int_{\mathbb{T}^N} \eta(y+hv)\,e^{-\frac{hL(y,v)+\phi(y+hv)-\phi(y)-\lambda}{\epsilon}}\,d\theta_n(y)\,dv \\
= & \int_{\mathbb{R}^N} \int_{\mathbb{T}^N} \eta(y+hv)\,e^{-\frac{hL(y,v)+\phi(y+hv)-\phi(y)-\lambda}{\epsilon}}\,d\theta(y)\,dv \\
= & \int_{\mathbb{R}^N} \int_{\mathbb{T}^N} \eta(y+hv)\,\theta(y)\,e^{-\frac{hL(y,v)+\phi(y+hv)-\phi(y)-\lambda}{\epsilon}}\,dy\,dv \\
= & \int_{\mathbb{R}^N} \int_{\mathbb{T}^N} \eta(x)\,\theta(x-hv)\,e^{-\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}}\,dx\,dv\,.
\end{aligned}$$

Since η here above is arbitrary, we get that θ satisfies (93), as desired. $\square$

Such θ plays a decisive rôle in the construction of penalized Mather measures, as we are now going to show. We also remark that the regularity and uniqueness of the function θ will be dealt with in Propositions 34 and 35 below.

Let us now deal with the entropy penalized Mather measures:

Theorem 32. *Let $\phi \in W^{1,\infty}(\mathbb{T}^N)$ satisfy*

$$\mathcal{G}[\phi] = \phi + \lambda. \tag{102}$$

Let $\theta \in L^1(\mathbb{T}^N)$ be so that[13] $\theta \geq 0$,

$$\int_{\mathbb{R}^N} \theta(x-hv)e^{-\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}}\,dv = \theta(x) \tag{103}$$

[13]The existence of a θ satisfying (103) and (104) is assured by Lemma 31 here above. The existence (and essential uniqueness) of ϕ and λ satisfying (102) is assured by Theorem 27.

and

$$\int_{\mathbb{T}^N} \theta(x)\, dx \;=\; 1\,. \tag{104}$$

Let

$$\mu(x,v) \;:=\; \theta(x) e^{-\frac{hL(x,v)+\phi(x+hv)-\phi(x)-\lambda}{\epsilon}}\,. \tag{105}$$

Then, μ minimizes the functional (89) over the space (90) under the constraint (88).

Remark. The result in Theorem 32 here above thus says that

$$d\mu(x,v) = \theta(x) e^{-\frac{hL(x,v)+\phi(x+hv)-\phi(x)-\lambda}{\epsilon}}\, dx\, dv$$

is an "entropy penalized Mather measure". The uniqueness of the measure μ will be proved in Theorem 36.

Proof. Exploiting (103) and (95), we deduce that

$$\begin{aligned}
&\int_{\mathbb{T}^N\times\mathbb{R}^N} \varphi(x+hv)\, d\mu(x,v) \\
&= \int_{\mathbb{T}^N\times\mathbb{R}^N} \varphi(x+hv)\, \theta(x)\, e^{-\frac{hL(x,v)+\phi(x+hv)-\phi(x)-\lambda}{\epsilon}}\, dx\, dv \\
&= \int_{\mathbb{T}^N\times\mathbb{R}^N} \varphi(x)\, \theta(x-hv)\, e^{-\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}}\, dx\, dv \\
&= \int_{\mathbb{T}^N} \varphi(x)\, \theta(x)\, dx \\
&= \int_{\mathbb{T}^N\times\mathbb{R}^N} \varphi(x)\, \theta(x)\, e^{-\frac{hL(x,v)+\phi(x+hv)-\phi(x)-\lambda}{\epsilon}}\, dx\, dv \\
&= \int_{\mathbb{T}^N\times\mathbb{R}^N} \varphi(x)\, d\mu(x,v)\,,
\end{aligned}$$

for any function $\varphi \in L^\infty(\mathbb{T}^N)$, which shows that the constraint (88) is satisfied by μ. In addition, from (105) and (95), we get that

$$\int_{\mathbb{R}^N} \mu(x,v)\, dv \;=\; \theta(x)\,, \tag{106}$$

for any $x \in \mathbb{T}^N$. Therefore, by (104),

$$\int_{\mathbb{T}^N\times\mathbb{R}^N} d\mu(x,v) = \int_{\mathbb{T}^N} \theta(x)\, dx = 1\,,$$

and so μ belongs to the space given by (90).

To prove the minimizing property of μ it suffices to show that for any other density $\tilde{\mu}$ in (90) that satisfies (88) and any $0 \leq \tau \leq 1$, one has that the function

$$I[\tau] = \int_{\mathbb{T}^N \times \mathbb{R}^N} L d\mu_\tau + \frac{\epsilon}{h} S[\mu_\tau], \tag{107}$$

with $\mu_\tau := (1-\tau)\mu + \tau\tilde{\mu}$, is convex and that $I'[0] = 0$. Denote by $\dot{\mu}_\tau := \tilde{\mu} - \mu$ Then,

$$\begin{aligned} I'[\tau] &= \int_{\mathbb{T}^N \times \mathbb{R}^N} \left[L d\dot{\mu}_\tau + \frac{\epsilon}{h} \cdot \left(\ln \frac{\mu_\tau}{\int \mu_\tau dw} \right) \cdot \dot{\mu}_\tau \right. \\ & \left. \qquad - \frac{\epsilon}{h} \cdot \frac{\mu_\tau}{\int \mu_\tau dw} \left(\int \dot{\mu}_\tau dw \right) \right] , \end{aligned} \tag{108}$$

where we have used the fact that

$$\int_{\mathbb{T}^N \times \mathbb{R}^N} \dot{\mu}_t = \int_{\mathbb{T}^N \times \mathbb{R}^N} (\tilde{\mu} - \mu) = 1 - 1 = 0\,.$$

Further, notice that, since both μ and $\tilde{\mu}$ are probability measures,

$$\begin{aligned} & \int_{\mathbb{T}^N \times \mathbb{R}^N} \frac{\mu}{\int_{\mathbb{R}^N} \mu dw} \left(\int_{\mathbb{R}^N} (\tilde{\mu} - \mu) dw \right) dx\, dv \\ = & \int_{\mathbb{T}^N} \frac{\int_{\mathbb{R}^N} \mu dv}{\int_{\mathbb{R}^N} \mu dw} \left(\int_{\mathbb{R}^N} (\tilde{\mu} - \mu) dw \right) dx \\ = & \int_{\mathbb{T}^N} \left(\int_{\mathbb{R}^N} (\tilde{\mu} - \mu) dw \right) dx \\ = & \int_{\mathbb{T}^N \times \mathbb{R}^N} (\tilde{\mu} - \mu) \\ = & \ 1 - 1 = 0\,. \end{aligned} \tag{109}$$

Moreover, making use of (106) and (105), we deduce that

$$
\begin{aligned}
& \int_{\mathbb{T}^N\times\mathbb{R}^N} hL\, d(\tilde\mu-\mu) \\
& +\epsilon \int_{\mathbb{T}^N\times\mathbb{R}^N} \ln \frac{\mu}{\int_{\mathbb{R}^N} \mu dw} (\tilde\mu-\mu)\, dx\, dv \\
= \;& \int_{\mathbb{T}^N\times\mathbb{R}^N} \left(hL(x,v) + \epsilon \ln \frac{\mu(x,v)}{\int_{\mathbb{R}^N} \mu(x,w)\, dw} \right) \\
& \cdot(\tilde\mu(x,v)-\mu(x,v))\, dx\, dv \\
= \;& \int_{\mathbb{T}^N\times\mathbb{R}^N} \left(hL(x,v) + \epsilon \ln \frac{\mu(x,v)}{\theta(x)} \right) \\
& \cdot(\tilde\mu(x,v)-\mu(x,v))\, dx\, dv \\
= \;& \int_{\mathbb{T}^N\times\mathbb{R}^N} \Big(-\phi(x+hv)+\phi(x)+\lambda \Big) \\
& \cdot(\tilde\mu(x,v)-\mu(x,v))\, dx\, dv \\
= \;& \int \Big(\phi(x)-\phi(x+hv)\Big)\, d\tilde\mu \\
& + \int \Big(\phi(x+hv)-\phi(x)\Big)\, d\mu \\
& + \lambda\left(\int d\tilde\mu - \int d\mu \right) \\
= \;& 0+0+\lambda(1-1) \;=\; 0\,,
\end{aligned} \tag{110}
$$

where we have also used again that both μ and $\tilde\mu$ are probability measures satisfying (88).

Collecting (108), (109) and (110), it follows that

$$I'[0]=0.$$

Thus, to prove that μ is minimal, it is sufficient to show that

$$I''[\tau]\geq 0.$$

To check this, we take a further derivative in (108), thus obtaining

$$
\begin{aligned}
\frac{h}{\epsilon}\, I''[\tau] &= \int_{\mathbb{T}^N\times\mathbb{R}^N} \frac{\dot\mu_\tau^2}{\mu_\tau} - 2\,\frac{\dot\mu_\tau \int \dot\mu_\tau dw}{\int \mu_\tau dw} + \frac{\mu_\tau \left(\int \dot\mu_\tau dw\right)^2}{\left(\int \mu_\tau dw\right)^2} \\
&= \int_{\mathbb{T}^N\times\mathbb{R}^N} \mu_\tau \left(\frac{\dot\mu_\tau}{\mu_\tau} - \frac{\int \dot\mu_\tau dw}{\int \mu_\tau dw} \right)^2 \geq 0\,.
\end{aligned} \tag{111}
$$

□

As a consequence of Theorem 32, we provide the following variational characterization for the fixed point problem of Theorem 27.

Corollary 33. *Let $\phi \in W^{1,\infty}(\mathbb{T}^N)$ and $\lambda \in \mathbb{R}$ satisfy $\mathcal{G}[\phi] = \phi + \lambda$. Then, the infimum of*

$$\int_{\mathbb{T}^N \times \mathbb{R}^N} h\, L(x,v)\, d\mu(x,v) + \epsilon S[\mu]$$

when μ is in the space given by (90) and satisfies the constraint (88), is equal to λ (which, in turn, is equal to $\mathcal{G}[\phi] - \phi$).

Proof. Theorem 32 says that the above infimum is attained at the measure μ given by (105), for which, by a direct computation (and recalling (88)), we have:

$$\begin{aligned}
& \int_{\mathbb{T}^N \times \mathbb{R}^N} h\, L(x,v)\, d\mu(x,v) + \epsilon S[\mu] \\
= & \int_{\mathbb{T}^N \times \mathbb{R}^N} \Big(\phi(x) - \phi(x+hv) + \lambda \Big)\, d\mu(x,v) \\
= & \int_{\mathbb{T}^N \times \mathbb{R}^N} \Big(\phi(x) - \phi(x+hv) \Big)\, d\mu(x,v) + \lambda \\
= & \ \lambda\,.
\end{aligned}$$

□

We now further investigate some properties of the function θ, which was introduced in Lemma 31 and played an important rôle in the construction of the penalized Mather measure μ in Theorem 32. First of all, we have the following regularity result:

Proposition 34. *Consider the setting of Lemma 31. Then, $\theta \in L^\infty(\mathbb{T}^N)$ and there exist constants κ_1^* and κ_2^* such that*

$$0 < \kappa_1^* \le \theta(x) \le \kappa_2^* \tag{112}$$

for any $x \in \mathbb{T}^N$.

Moreover, if

$$\int_{\mathbb{R}^N} e^{-\frac{h\,K(v)}{\epsilon}}\, |DK(v)|\, dv \, < \, +\infty\,, \tag{113}$$

then, $\theta \in W^{1,\infty}(\mathbb{T}^N)$ and there exists a constant κ_3^* such that

$$\|\theta\|_{W^{1,\infty}(\mathbb{T}^N)} \le \kappa_3^*\,. \tag{114}$$

The constants κ_i^*, for $1 \le i \le 3$, depend only on N, ϕ, L, λ, ϵ and h.

Proof. In the course of this proof we denote by κ a convenient positive quantity, possibly depending only on N, ϕ, L, λ, ϵ and h, and which may be different at different steps of the computation.

By changing variable in (93), we have that

$$\theta(x) = \frac{1}{h^N} \int_{\mathbb{R}^N} \theta(w) e^{-\frac{hL(w,(x-w)/h)+\phi(x)-\phi(w)-\lambda}{\epsilon}} dw \,. \tag{115}$$

Therefore, recalling (92) and (94),

$$\begin{aligned} \theta(x) &\geq \frac{1}{h^N} \int_{\mathbb{T}^N} \theta(w) e^{-\frac{hL(w,(x-w)/h)+\phi(x)-\phi(w)-\lambda}{\epsilon}} dw \\ &\geq \kappa \int_{\mathbb{T}^N} \theta(w)\, dw \\ &= \kappa \,. \end{aligned} \tag{116}$$

On the other hand, since L is superlinear in v, one also deduces from (115) and (94) that

$$\begin{aligned} \theta(x) &\leq \kappa \sum_{k\in\mathbb{Z}^N} \int_{\mathbb{T}^N} \theta(w)\, e^{-\kappa|k|}\, dw \\ &\leq \kappa \int_{\mathbb{T}^N} \theta(w)\, dw \\ &= \kappa \,. \end{aligned}$$

This and (116) yield (112).

Furthermore, by differentiating (115), we get that

$$\begin{aligned} &|D\theta(x)| \\ \leq\ & \kappa \int_{\mathbb{R}^N} \theta(w) e^{-\frac{hL(w,(x-w)/h)+\phi(x)-\phi(w)-\lambda}{\epsilon}} \left[\left| DK\Big(\frac{x-w}{h}\Big) \right| + |D\phi(x)| \right] dw \\ \leq\ & \kappa \int_{\mathbb{R}^N} e^{-\frac{h\,K(v)}{\epsilon}} \Big[|DK(v)| + 1 \Big]\, dv \\ \leq\ & \kappa \,, \end{aligned}$$

thanks to (112) and (113). This and (112) yield (114). □

We now apply the Hopf-Cole transform method to θ, that is, we define

$$\varsigma(x) := -\epsilon \ln \theta(x) \,. \tag{117}$$

Note that, in the setting of (13), we have that $\theta = \mathcal{E}[\varsigma]$. Also, (117) is a bona fide definition, since $\theta > 0$ by (112). Moreover, ς is $\mathbb{Z}^N$-periodic, since so is θ.

We now show that ς is the solution of an entropy penalized scheme analogous to (6). Namely, let

$$\overline{L}(x,v) \;:=\; L(x+hv,-v)+\frac{\phi(x)-\phi(x+hv)-\lambda}{h}$$

and, for any function φ, consider

$$\overline{\mathcal{G}}[\varphi](x) \;:=\; -\epsilon \,\ln\left[\int_{\mathbb{R}^N} e^{-\frac{h\overline{L}(x,v)+\varphi(x+hv)}{\epsilon}}\,dv\right]. \tag{118}$$

The reader may compare (6) and (118) and note that the function $\varsigma+\phi$ is a solution to a time-reversed $\mathcal{G}$ scheme with Lagrangian $L(x+hv,v)$, that is ς is the difference between a forward and backward fixed point.

In this framework, the following result holds:

Proposition 35. *Consider the setting of Lemma 31 and let ς be as in (117). Then, $\varsigma\in L^\infty(\mathbb{T}^N)$ and its L^∞-norm is bounded by a quantity depending only on N, ϕ, L, λ, ϵ and h.*

Also, $\overline{\mathcal{G}}[\varsigma]=\varsigma$.

Moreover, if (113) holds, then $\varsigma\in W^{1,\infty}(\mathbb{T}^N)$, its $W^{1,\infty}$-norm is bounded by a quantity depending only on N, ϕ, L, λ, ϵ and h.

Also, in this case, the function θ in Lemma 31 is unique.

Proof. The L^∞-bound (resp., the $W^{1,\infty}$-bound) plainly follows from (117) and (112) (resp., and (114)). The fact that ς is a fixed point of $\overline{\mathcal{G}}$ is a consequence of (93).

Let us now show the uniqueness of θ. Assume that $\hat\theta$ also fulfills the thesis of Lemma 31 and let $\hat\varsigma$ be so that $\hat\theta=\mathcal{E}[\hat\varsigma]$. By what we have just proved, we have that both ς and $\hat\varsigma$ are uniformly Lipschitz and that they are both fixed points of $\overline{\mathcal{G}}$. Then, using Theorem 9 (applied to[14] $\overline{\mathcal{G}}$ instead of $\mathcal{G}$), we have that

$$\begin{aligned}\|\varsigma-\hat\varsigma\|_\sharp &= \|\overline{\mathcal{G}}[\varsigma]-\overline{\mathcal{G}}[\hat\varsigma]\|_\sharp\\ &\le (1-C\,\|\varsigma-\hat\varsigma\|_\sharp)\,\|\varsigma-\hat\varsigma\|_\sharp\,,\end{aligned}$$

for some $C>0$. This shows that $\varsigma=\hat\varsigma+c$, for some $c\in\mathbb{R}$.

[14]Note that $\overline{L}$ is superlinear, since so is L. Also, $\overline{L}\in W^{1,\infty}(\mathbb{T}^N\times\mathbb{R}^N)$, because $\phi\in W^{1,\infty}(\mathbb{T}^N)$. No further regularity of L was used in the proof of Theorem 9.

But then, recalling (94),

$$\begin{aligned}\int_{\mathbb{T}^N} e^{-\frac{\varsigma}{\epsilon}} &= \int_{\mathbb{T}^N} \theta = 1 \\ = \int_{\mathbb{T}^N} \hat{\theta} &= \int_{\mathbb{T}^N} e^{-\frac{\hat{\varsigma}}{\epsilon}} = e^{\frac{c}{\epsilon}} \int_{\mathbb{T}^N} e^{-\frac{\varsigma}{\epsilon}},\end{aligned}$$

thence $c = 0$ and, therefore, $\varsigma = \hat{\varsigma}$. Consequently, $\theta = \hat{\theta}$. □

We are now in position to prove the uniqueness of the penalized Mather measure constructed in Theorem 32:

Theorem 36. *Assume (113). Then, the measure μ constructed in Theorem 32 is the unique minimizer of the functional in (89) over the space given by (90) under the constraint (88).*

Proof. Assume that $\tilde{\mu}$ is also a minimizer. Let $\mu_\tau := (1-\tau)\mu + \tau\tilde{\mu}$, for $\tau \in [0, 1]$. Let also $\dot{\mu}_\tau := \tilde{\mu} - \mu$ and

$$\tilde{\theta}(x) := \int_{\mathbb{R}^N} \tilde{\mu}(x, v)\, dv\,. \tag{119}$$

These definitions and (106) give that

$$\begin{aligned}&\dot{\mu}_\tau \int_{\mathbb{R}^N} \mu\, dw - \mu \int_{\mathbb{R}^N} \dot{\mu}_\tau\, dw \\ &= \theta\tilde{\mu} - \tilde{\theta}\mu\,. \end{aligned}\tag{120}$$

Let now I be as in (107). Since both μ and $\tilde{\mu}$ are minimizers for the functional in (89), we have that both $\tau = 0$ and $\tau = 1$ minimize I. On the other hand, we know from the proof of Theorem 32 that I is convex. Therefore, I must be constant for any $\tau \in [0, 1]$. In particular, $I''[0] = 0$. By (111), this says that

$$0 = \int_{\mathbb{T}^N\times\mathbb{R}^N} \frac{\left(\dot{\mu}_\tau \int_{\mathbb{R}^N} \mu\, dw - \mu \int_{\mathbb{R}^N} \dot{\mu}_\tau\, dw\right)^2}{\mu \cdot \left(\int_{\mathbb{R}^N} \mu\, dw\right)^2}\,.$$

This and (120) yield that

$$\theta\,\tilde{\mu} = \tilde{\theta}\,\mu \tag{121}$$

almost everywhere in $\mathbb{T}^N \times \mathbb{R}^N$. Thence, in the light of (105), we conclude that

$$\tilde{\theta}(x) = \tilde{\mu}(x, v)\, e^{\frac{hL(x,v)+\phi(x+hv)-\phi(x)-\lambda}{\epsilon}},$$

and therefore

$$\tilde{\theta}(x-hv) \;=\; \tilde{\mu}(x-hv,v)\, e^{\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}}\,. \tag{122}$$

Thus, given any $\varphi \in C(\mathbb{T}^N)$, we gather from (88), (119) and (122) that

$$\begin{aligned}
& \int_{\mathbb{T}^N}\int_{\mathbb{R}^N} \varphi(x)\,\tilde{\theta}(x-hv)e^{-\frac{hL(x-hv,v)+\phi(x)-\phi(x-hv)-\lambda}{\epsilon}}\,dv\,dx \\
= & \int_{\mathbb{T}^N}\int_{\mathbb{R}^N} \varphi(x)\,\tilde{\mu}(x-hv,v)\,dv\,dx \\
= & \int_{\mathbb{R}^N}\int_{\mathbb{T}^N} \varphi(y+hv)\,\tilde{\mu}(y,v)\,dy\,dv \\
= & \int_{\mathbb{T}^N\times\mathbb{R}^N} \varphi(y+hv)\,d\tilde{\mu}(y,v) \\
= & \int_{\mathbb{T}^N\times\mathbb{R}^N} \varphi(y)\,d\tilde{\mu}(y,v) \\
= & \int_{\mathbb{T}^N}\int_{\mathbb{R}^N} \varphi(y)\,\tilde{\mu}(y,v)\,dv\,dy \\
= & \int_{\mathbb{T}^N} \varphi(y)\,\tilde{\theta}(y)dy\,.
\end{aligned}$$

Since φ here above is arbitrary, this means that $\tilde{\theta}$ satisfies (93). Also, $\tilde{\theta}$ satisfies (92) and (94), due to (119) and the fact that $\tilde{\mu}$ is a probability measure. Then, by the uniqueness result in Proposition 35, we deduce that $\tilde{\theta}=\theta$. Then, (121) yields that $\tilde{\mu}=\mu$. $\square$

11. Convergence to Mather measures

In this section, we use duality methods to study the entropy penalized Mather problem and its convergence as $\epsilon/h \to 0$ (that is, when the entropy vanishes). We will prove a *variational identity* which relates the action functional with the fixed points of the non-linear scheme (see Theorem 37). This identity turns out to be useful in showing that *the minimal action value, the fixed point of the scheme and the entropy penalized Mather measure converge* to the analogous features in Mather theory as the entropy vanishes (see Theorems 38, 39 and 40).

Theorem 37. *We have the following identity:*

$$\begin{aligned}
& \inf_{\mu} \int hL d\mu + \epsilon S[\mu] \\
= & -\inf_{\psi\in C(\mathbb{T}^N)} \sup_{x\in\mathbb{T}^N} \epsilon \ln \int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\psi(x+hv)-\psi(x)}{\epsilon}} dv\,,
\end{aligned} \tag{123}$$

where the first infimum is taken over all measures μ in the space given by (90) under the constraint in (88).

Furthermore, if $\phi \in W^{1,\infty}(\mathbb{T}^N)$ and $\lambda \in \mathbb{R}$ satisfy[15] $\mathcal{G}[\phi] = \phi + \lambda$, *then ϕ is an optimal function for the right-hand-side of (123) and λ is equal to the quantity given in (123).*

Proof. Clearly,[16]

$$\begin{aligned} &\inf_{\psi \in C(\mathbb{T}^N)} \sup_{x \in \mathbb{T}^N} \epsilon \ln \int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\psi(x+hv)-\psi(x)}{\epsilon}} dv \\ &\leq \sup_{x \in \mathbb{T}^N} \epsilon \ln \int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\phi(x+hv)-\phi(x)}{\epsilon}} dv = -\lambda\,. \end{aligned} \tag{124}$$

On the other hand, from (6),

$$\begin{aligned} &\inf_{\psi \in C(\mathbb{T}^N)} \sup_{x \in \mathbb{T}^N} \epsilon \ln \int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\psi(x+hv)-\psi(x)}{\epsilon}} dv \\ &\qquad = \inf_{\psi \in C(\mathbb{T}^N)} \sup_{x \in \mathbb{T}^N} \psi - \mathcal{G}[\psi]\,. \end{aligned} \tag{125}$$

Thus, take any function $\psi \in C(\mathbb{T}^N)$ and choose a point x_0 in which $\psi - \phi$ have a maximum. Let $\hat{\phi} := \phi - \phi(x_0) + \psi(x_0)$, so that $\hat{\phi} \geq \psi$. Therefore, in the light of Propositions 6 and 7,

$$\mathcal{G}[\phi] - \phi(x_0) + \psi(x_0) = \mathcal{G}[\hat{\phi}] \geq \mathcal{G}[\psi]\,.$$

Consequently,

$$\sup_{x \in \mathbb{T}^N} \psi - \mathcal{G}[\psi] \geq \psi(x_0) - \mathcal{G}[\psi](x_0) \geq \phi(x_0) - \mathcal{G}[\phi](x_0) = -\lambda\,.$$

Thus, by taking the infimum over all functions $\psi \in C(\mathbb{T}^N)$, we deduce from the latter estimate, (125) and (124) that

$$-\inf_{\psi \in C(\mathbb{T}^N)} \sup_{x \in \mathbb{T}^N} \epsilon \ln \int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\psi(x+hv)-\psi(x)}{\epsilon}} dv = \lambda\,.$$

Then, the claim in (123) follows from Corollary 33. □

---

[15]Again, we recall that the existence of ϕ and λ as desired follows from Theorem 27.

[16]Note that, in general, a proof of statements as the ones in Theorem 37 may require the use of the Legendre-Fenchel Duality Theorem or some advanced technology. However, in our case, as we know the optimal measure by Theorem 32, we manage to give an elementary proof.

At this point, it is convenient to recall that the Mather problem, which consists in minimizing (87) under the constraint in (88), also has a dual problem. Namely, if $\overline{H}_0$ is the infimum of

$$\int_{\mathbb{T}^N\times\mathbb{R}^N} Ld\mu\,,$$

over all the probability measures on $\mathbb{T}^N\times\mathbb{R}^N$ satisfying the constraint in (88), one has that

$$\inf_{\psi\in C(\mathbb{T}^N)}\sup_{(x,v)\in\mathbb{T}^N\times\mathbb{R}^N} -\left[hL(x,v)+\psi(x+hv)-\psi(x)\right] = -h\overline{H}_0\,. \tag{126}$$

We refer to [Gom05] for further details about this fact. We just mention here that $\overline{H}_0$ (possibly up to a sign convention) is what is called the "effective Hamiltonian" in the literature. Furthermore, there exists an optimal ϕ_0 such that

$$\inf_{v\in\mathbb{R}^N}\left[hL(x,v)+\phi_0(x+hv)-\phi_0(x)-h\overline{H}_0\right]=0, \tag{127}$$

for all $x\in\mathbb{T}^N$ (see, e.g., [Gom05]). For convenience, we define $\lambda_{\epsilon,h}:=\lambda$ and $\phi_{\epsilon,h}:=\phi$, where ϕ and λ are as in Theorem 37, and we set

$$\overline{H}_{\epsilon/h} := \inf_{\mu}\int Ld\mu+\frac{\epsilon}{h}S[\mu],$$

where the infimum is taken over all measures μ belonging to the space given by (90) and that satisfy the constraint (88). We will see in Theorem 38 below that $\overline{H}_{\epsilon/h}$ may be seen as an approximate effective Hamiltonian, which will tend to $\overline{H}_0$ when $\epsilon/h\to 0$.

It follows from Theorem 37 that

$$\overline{H}_{\epsilon/h} = \frac{\lambda_{\epsilon,h}}{h}\,. \tag{128}$$

Theorem 38. *We have*

$$\overline{H}_{\epsilon/h}\to\overline{H}_0,$$

as $\epsilon/h\to 0$.

Proof. We will make use of the following elementary[17] fact: if $f \in C(\mathbb{R}^N)$ is superlinear, then

$$\lim_{\epsilon\to 0^+}\left(\int_{\mathbb{R}^N} e^{-\frac{f(v)}{\epsilon}}\,dv\right)^{\epsilon} = e^{-\inf_{\mathbb{R}^N} f}\,. \tag{129}$$

Also, since

$$\lim_{\epsilon/h\to 0^+}\overline{H}_{\epsilon/h} = \lim_{\epsilon\to 0^+}\overline{H}_{\epsilon} = \lim_{\epsilon\to 0^+}\overline{H}_{\epsilon/1}\,,$$

we may and do fix $h := 1$ in this proof.

Let ϕ_0 be as in (127). By (128) and Theorem 37, we have that

$$-\overline{H}_\epsilon \le \sup_{x\in\mathbb{T}^N}\epsilon\ln\int_{\mathbb{R}^N} e^{-\frac{L(x,v)+\phi_0(x+v)-\phi_0(x)}{\epsilon}}dv\,.$$

Therefore, using (129) and (127),

$$-\lim_{\epsilon\to 0^+}\overline{H}_\epsilon \le -\overline{H}_0\,. \tag{130}$$

To prove the reverse inequality, we observe that $\phi_{\epsilon,1}$ is a uniformly Lipschitz function, thanks to Theorem 27, and, therefore, we may take the infimum on the right hand side of (123) over the space of Lipschitz functions with Lipschitz constant bounded by some suitably large number Λ. In particular, by Theorem 37, fixed any $\delta > 0$, there exists a ψ in such a Lipschitz space, so that

$$-\overline{H}_\epsilon + \delta = -\lambda_{\epsilon,1} + \delta \ge \sup_{x\in\mathbb{T}^N}\epsilon\ln\int_{\mathbb{R}^N} e^{-\frac{L(x,v)+\psi(x+v)-\psi(x)}{\epsilon}}\,dv\,. \tag{131}$$

[17]For completeness, we give a sketch of the proof of (129). Let $f(v) \ge c|v|$ if $|v| \ge R_0$, for $c > 0$. Take $R \ge R_0$, then

$$\int_{\mathbb{R}^N} e^{-\frac{f}{\epsilon}} \le \int_{|v|\le R} e^{-\frac{\inf f}{\epsilon}} + \int_{|v|\ge R} e^{-\frac{c|v|}{\epsilon}} \le \text{const}\,(R^N e^{-\frac{\inf f}{\epsilon}} + \epsilon^N\, e^{-\frac{cR}{2\epsilon}})\,.$$

In particular, if $R := 1/\epsilon$ and ϵ is small enough,

$$\int_{\mathbb{R}^N} e^{-\frac{f}{\epsilon}} \le \text{const}\,\epsilon^{-N} e^{-\frac{\inf f}{\epsilon}}\,,$$

which gives one inequality in (129). For the other one, fix $\delta > 0$. Then, there exists a ball B so that $f(v) \le \inf f + \delta$ for any $v \in B$. Then,

$$\int_{\mathbb{R}^N} e^{-\frac{f}{\epsilon}} \ge \int_B e^{-\frac{f}{\epsilon}} \ge e^{-\frac{\inf f+\delta}{\epsilon}}\,|B|\,,$$

then take the ϵ-power to both sides, then send $\epsilon \to 0$ and then $\delta \to 0$.

Given such a ψ, we define (x_ψ, v_ψ) to be such that

$$\begin{aligned} & -\left[L(x_\psi, v_\psi) + \psi(x_\psi + v_\psi) - \psi(x_\psi)\right] \\ & \qquad = \sup_{(x,v)\in\mathbb{T}^N\times\mathbb{R}^N} -\left[L(x,v) + \psi(x+v) - \psi(x)\right] . \end{aligned}$$

We denote by σ_ψ the above supremum.

Let us now localize v_ψ: since ψ is Lipschitz, the definition of (x_ψ, v_ψ) implies that

$$K(v_\psi) \le K(0) + \Lambda\sqrt{N}$$

and so, since K is superlinear, we have that $|v_\psi| \le \hat\Lambda$, for a suitably large universal number $\hat\Lambda$. Let now

$$\tilde\Lambda := \sup_{|v|\le\hat\Lambda+2} |DK(v)|$$

and consider the universal quantity

$$\Lambda_* := \Lambda + \tilde\Lambda .$$

Then, if $|v - v_\psi| \le \epsilon < 1$, we have that $|v| \le \hat\Lambda + 1$ and so

$$\begin{aligned} & |\sigma_\psi + L(x_\psi, v) + \psi(x_\psi + v) - \psi(x_\psi)| \\ \le\ & |K(v_\psi) - K(v)| + |\psi(x_\psi + v) - \psi(x_\psi + v_\psi)| \\ \le\ & \Lambda_*\,\epsilon , \end{aligned} \tag{132}$$

for any $v \in B_\epsilon(v_\psi)$.

By collecting the estimates in (131) and (132), we obtain

$$\begin{aligned} -\overline{H}_\epsilon + \delta \ &\ge\ \sup_{x\in\mathbb{T}^N} \epsilon \ln \int_{B_\epsilon(v_\psi)} e^{-\frac{L(x,v)+\psi(x+v)-\psi(x)}{\epsilon}}\, dv \\ &\ge\ \epsilon \ln \left(|B_\epsilon(v_\psi)|\, e^{\frac{\sigma_\psi}{\epsilon} - \Lambda_*} \right) \\ &=\ \sigma_\psi + \text{const}\,\epsilon \,\ln(\,\text{const}\,\epsilon) . \end{aligned}$$

Since, by (126) and the definition of σ_ψ, we have that

$$-\overline{H}_0 \le \sigma_\psi ,$$

it follows from the above computation that

$$\delta - \lim_{\epsilon\to 0^+} \overline{H}_\epsilon \ge -\overline{H}_0 .$$

Thence, since δ is arbitrary,

$$-\lim_{\epsilon\to 0^+} \overline{H}_\epsilon \ge -\overline{H}_0 .$$

This and (130) complete the proof of the desired result. $\square$

Having completed the proof of the convergence of the effective Hamiltonian, we now address the convergence of the "normalized" function

$$\phi_\epsilon(x) := \phi_{\epsilon,h}(x) - \phi_{\epsilon,h}(0), \tag{133}$$

as $\epsilon \to 0$, for a fixed $h > 0$:

Theorem 39. *$\phi_\epsilon \to \phi_0$ uniformly as $\epsilon \to 0^+$, in which $\phi_0 \in W^{1,\infty}(\mathbb{T}^N)$ is a solution to (127). Also, ϕ_0 is semiconcave. The Lipschitz constant and the semiconcavity modulus of ϕ_0 are bounded by a constant depending only on N, C_U and C_K.*

Proof. Along this proof, we will make explicit the dependence on ϵ of $\mathcal{G}$ by denoting it by $\mathcal{G}_\epsilon$ and we define $\mathcal{G}_0$ to be the operator

$$\mathcal{G}_0[\psi](x) := \inf_v hL(x,v) + \psi(x+hv).$$

Notice that, from (128) and Theorem 38,

$$\lim_{\epsilon\to 0^+} \lambda_{\epsilon,h} = h\,\overline{H}_0\,. \tag{134}$$

Moreover, due to Theorem 27, we know that ϕ_ϵ is uniformly Lipschitz continuous and semiconcave and so, by (133), ϕ_ϵ converges uniformly to a suitable $\phi_0 \in W^{1,\infty}(\mathbb{T}^N)$, and both the Lipschitz constant and the semiconcavity modulus of ϕ_0 are universally bounded. What is more, by Proposition 8,

$$\lim_{\epsilon\to 0^+} \|\mathcal{G}_\epsilon[\phi_\epsilon] - \mathcal{G}_\epsilon[\phi_0]\|_{L^\infty(\mathbb{T}^N)} \le \lim_{\epsilon\to 0^+} \|\phi_\epsilon - \phi_0\|_{L^\infty(\mathbb{T}^N)} = 0\,. \tag{135}$$

Notice also that, by (129) and the continuity of ϕ_0, we have that

$$\lim_{\epsilon\to 0^+} \mathcal{G}_\epsilon[\phi_0](x) = \mathcal{G}_0[\phi_0](x)\,, \tag{136}$$

for any $x \in \mathbb{T}^N$.

Therefore, we use (134), (135) and (136) to deduce that

$$\begin{aligned}
h\,\overline{H}_0 &= \lim_{\epsilon\to 0^+} \lambda_{\epsilon,h} \\
&= \lim_{\epsilon\to 0^+} \mathcal{G}_\epsilon[\phi_\epsilon] - \phi_\epsilon \\
&= \lim_{\epsilon\to 0^+} \mathcal{G}_\epsilon[\phi_0] + \Big(\mathcal{G}_\epsilon[\phi_\epsilon] - \mathcal{G}_\epsilon[\phi_0]\Big) - \phi_0 + \Big(\phi_0 - \phi_\epsilon\Big) \\
&= \mathcal{G}_0[\phi_0] + 0 - \phi_0 + 0 \\
&= \mathcal{G}_0[\phi_0] - \phi_0\,,
\end{aligned}$$

and so ϕ_0 is a solution of (127). $\square$

We now consider the penalized Mather measures $\mu_\epsilon := \mu_{\epsilon,h}$ constructed in Theorem 32 and we prove its convergence as $\epsilon \to 0^+$, for a fixed $h > 0$:

Theorem 40. *As $\epsilon \to 0$, μ_ϵ converges weakly to a (discrete) Mather measure, that is, a probability measure on $\mathbb{T}^N \times \mathbb{R}^N$ that minimizes*

$$\int_{\mathbb{T}^N \times \mathbb{R}^N} L d\mu,$$

under the constraint in (88).

Moreover, the support of μ_0 is compact.

Proof. We will drop the h dependence in some of the indices of this proof, since no confusion can arise. We recall that, by (105),

$$\mu_\epsilon = \theta_\epsilon(x) e^{-\frac{hL(x,v)+\phi_\epsilon(x+hv)-\phi_\epsilon(x)-\lambda_\epsilon}{\epsilon}}$$

with

$$\int_{\mathbb{R}^N} \mu_\epsilon(x,v)\, dv = \theta_\epsilon(x)\,.$$

Furthermore, $\phi_\epsilon \to \phi_0$ uniformly, where ϕ_0 satisfies (127), and $\lambda_\epsilon \to h\overline{H}_0$, according to Theorems 38 and 39.

In particular, $\|\phi_\epsilon\|_{L^\infty}$ and $|\lambda_\epsilon|$ are bounded uniformly in ϵ. Therefore, from the fact that K is superlinear,

$$\begin{aligned}
\mu_\epsilon\Big(\mathbb{T}^N \times (\mathbb{R}^N \setminus B_R)\Big) &= \int_{\mathbb{T}^N \times (\mathbb{R}^N \setminus B_R)} \theta_\epsilon(x)\, e^{-\frac{hK(v)-hU(x)+\phi_\epsilon(x+hv)-\phi_\epsilon(x)-\lambda_\epsilon}{\epsilon}} \\
&\le e^{\frac{c_1-c_2hR}{\epsilon}} \int_{\mathbb{T}^N \times (\mathbb{R}^N \setminus B_R)} \theta_\epsilon(x) e^{-\frac{c_2h|v|}{\epsilon}}\, dx\, dv \qquad (137) \\
&\le c_3 \epsilon^N e^{-\frac{c_1}{\epsilon}} \int_{\mathbb{T}^N} \theta_\epsilon(x)\, dx \\
&= c_3 \epsilon^N e^{-\frac{c_1}{\epsilon}}\,,
\end{aligned}$$

provided that $R \ge 4c_1/(c_2h)$, where the c_i's are suitable positive quantities independent of ϵ.

Also, since μ_ϵ is a probability measure, we have that $\mu_\epsilon \rightharpoonup \mu_0$, for some measure μ_0 on $\mathbb{T}^N \times \mathbb{R}^N$ (see, e.g., page 55 of [EG92]). We now show that μ_0 is a probability measure, and that it has compact support. It is standard (see, e.g., Theorem 1-(ii) on page 54 of [EG92]) that

$$\mu_0(\mathbb{T}^N \times \mathbb{R}^N) \le \liminf_{\epsilon \to 0} \mu_\epsilon(\mathbb{T}^N \times \mathbb{R}^N) = 1\,.$$

On the other hand (see again Theorem 1-(ii) on page 54 of [EG92]),

$$\mu_0(\mathbb{T}^N \times B_R) \geq \limsup_{\epsilon\to 0} \mu_\epsilon(\mathbb{T}^N \times B_R) \geq \limsup_{\epsilon\to 0} \left(1 - c_3\epsilon^N e^{-\frac{c_1}{\epsilon}}\right) = 1\,,$$

provided that R is large enough, thanks to (137). Therefore, μ_0 is a probability measure with compact support.

Also, μ_0 satisfies the constraint in (88), since μ_ϵ does.

We now prove that

$$\begin{gathered}\text{if } (x,v) \text{ is a point in the support of } \mu_0,\\ \text{then } v \text{ is a minimizer of the function}\\ v \mapsto hL(x,v) + \phi_0(x+hv) - \phi_0(x).\end{gathered} \tag{138}$$

To prove this, take a point $(x_1, v_1) \in \mathbb{T}^N \times \mathbb{R}^N$ which is not a minimizer of the above function. Choose a small neighborhood V_1 of x_1 in $\mathbb{T}^n$ and and a small ball $B_1 \subset \mathbb{R}^n$ centered in v_1. Then, by (127), if these neighborhood are small enough, there exists a suitably small $\delta > 0$ such that

$$hL(x,v) + \phi_0(x+hv) - \phi_0(x) \geq h\overline{H}_0 + \delta \tag{139}$$

for any $(x,v) \in V_1 \times B_1$. Now choose a point v_0 so to minimizing the function

$$v \mapsto hL(x_1,v) + \phi_0(x_1+hv) - \phi_0(x_1)\,.$$

Thanks to (127), we may now take a small ball $B_0 \subset \mathbb{R}^n$ centered in v_0 and a small neighborhood V_0 of x_1, in such a way that

$$hL(x,v) + \phi_0(x+hv) - \phi_0(x) \leq h\overline{H}_0 + \frac{\delta}{4} \tag{140}$$

for any $(x,v) \in V_0 \times B_0$. We now consider the neighborhood of x_1 given by

$$V := V_0 \cap V_1\,.$$

By the uniform convergence of ϕ_ϵ, (139) and (140), we have that, if ϵ is small enough,

$$hL(x,v) + \phi_\epsilon(x+hv) - \phi_\epsilon(x) \geq h\overline{H}_0 + \frac{3}{4}\delta,$$

for any $(x,v) \in V \times B_1$ and

$$hL(x,v) + \phi_\epsilon(x+hv) - \phi_\epsilon(x) \leq h\overline{H}_0 + \frac{\delta}{2}.$$

for any $(x, v) \in V \times B_0$. Thence,

$$\mu_\epsilon(V \times B_1) = \int_{V\times B_1} \theta_\epsilon(x)\, e^{-\frac{hL(x,v)+\phi_\epsilon(x+v)-\phi_\epsilon(0)-\lambda_\epsilon}{\epsilon}} \leq \theta_\epsilon(V) e^{-\frac{h(\overline{H}_0-\overline{H}_\epsilon)+\frac{3\delta}{4}}{\epsilon}}$$

and

$$\mu_\epsilon(V \times B_0) = \int_{V\times B_0} \theta_\epsilon(x)\, e^{-\frac{hL(x,v)+\phi_\epsilon(x+v)-\phi_\epsilon(0)-\lambda_\epsilon}{\epsilon}} \geq \theta_\epsilon(V) e^{-\frac{h(\overline{H}_0-\overline{H}_\epsilon)+\frac{\delta}{2}}{\epsilon}}$$

Thus,

$$\frac{\mu_\epsilon(V \times B_1)}{\mu_\epsilon(V \times B_0)} \leq e^{-\frac{\delta}{4\epsilon}} \to 0,$$

as $\epsilon \to 0$. In particular, $\mu_\epsilon(V \times B_1) \to 0$ as $\epsilon \to 0$. This implies that $\mu_0(V \times B_1) = 0$ (see, e.g., Theorem 1-(ii) on page 54 of [EG92]). Therefore, the proof of (138) is complete.

It follows from (138) and (127) that

$$L(x, v) + \phi_0(x + hv) - \phi_0(x) = \overline{H}_0,$$

for any (x, v) in the support of μ_0. Thus, integrating and using the constraint in (88), we obtain that

$$\int L d\mu_0 = \int [L + \phi_0(x + hv) - \phi_0(x)]\, d\mu_0 = \overline{H}_0\,.$$

This completes the proof of the desired result, by the definiton of $\overline{H}_0$ given on page 52. $\square$

Appendix A. A Banach-Caccioppoli-type Theorem

We state and prove a variation of the standard Banach-Caccioppoli Theorem about the existence and uniqueness of fixed points of strict contractions. The result we present is given in a form which is convenient for the proof of Theorem 27.

Theorem 41. *Let S be a closed subset of a Banach space, endowed with a norm $\|\cdot\|$. Assume that $g : S \to S$ is so that*

$$\|g(x) - g(y)\| \leq (1 - \alpha\|x - y\|^\beta)\, \|x - y\|\,, \tag{141}$$

for any x, $y \in S$ and some given constants α, $\beta > 0$. Then, there exists a unique $x^\star \in S$ so that $g(x^\star) = x^\star$. Furthermore, given any $x_0 \in S$, we have that

$$x^\star = \lim_{n\to+\infty} g^n(x_0)\,. \tag{142}$$

Proof. Let $x_0 \in S$. For any $n \in \mathbb{N}$, we define $x_n := g^n(x_0)$ and $d_n := \|x_{n+1} - x_n\|$. Note that $x_n \in S$ for any $n \in \mathbb{N}$, because $g(S) \subseteq S$. Also, by (141),

$$\|g(x) - g(y)\| \leq \|x - y\|, \tag{143}$$

for any $x,\ y \in S$ and therefore

$$\begin{aligned} d_n &= \|g(x_n) - g(x_{n-1})\| \\ &\leq \|x_n - x_{n-1}\| \\ &= d_{n-1}, \end{aligned}$$

for any $n \geq 1$, that is d_n is a non-increasing sequence. Thus, there exists $d \geq 0$ so that

$$d = \lim_{n \to +\infty} d_n .$$

We now show that $d = 0$. Indeed, let us assume, by contradiction, that

$$d > 0 . \tag{144}$$

Then, using (141), one has that

$$\begin{aligned} d &= \lim_{n \to +\infty} \|x_{n+1} - x_n\| \\ &= \lim_{n \to +\infty} \|g(x_n) - g(x_{n-1})\| \\ &\leq \lim_{n \to +\infty} (1 - \alpha\|x_n - x_{n-1}\|^{\beta})\, \|x_n - x_{n-1}\| \\ &= \lim_{n \to +\infty} (1 - \alpha d_{n-1}^{\beta}) d_{n-1} \\ &= (1 - \alpha d^{\beta}) d, \end{aligned}$$

that is, by (144),

$$1 \leq 1 - \alpha d^{\beta},$$

which gives $d \leq 0$, contradicting (144). This shows that $d = 0$ or, equivalently that

$$\lim_{n \to +\infty} \|g(x_n) - x_n\| = 0 . \tag{145}$$

We show that x_n is a Cauchy sequence in S. Indeed, fix $\epsilon > 0$ and let

$$\eta := \frac{\alpha\, \epsilon^{\beta+1}}{2} . \tag{146}$$

Thanks to (145), we may assumne that

$$\|g(x_n) - x_n\| \leq \eta$$

provided $n \geq n_0$, for a suitable $n_0 = n_0(\epsilon)$. Combining this with (141), we gather that

$$\begin{aligned} & \|x_n - x_m\| \\ \leq\ & \|x_n - g(x_n)\| + \|g(x_n) - g(x_m)\| + \|g(x_m) - x_m\| \\ \leq\ & 2\eta + (1 - \alpha\|x_n - x_m\|^\beta)\, \|x_n - x_m\| \\ =\ & 2\eta - \alpha\|x_n - x_m\|^{\beta+1} + \|x_n - x_m\| \end{aligned}$$

provided that $n,\ m \geq n_0$. That is,

$$0 \leq 2\eta - \alpha\|x_n - x_m\|^{\beta+1}$$

and so

$$\|x_n - x_m\| \leq \epsilon$$

due to (146). This shows that x_n is a Cauchy sequence. Since S is complete, we have that there exists $x^\star \in S$ so that

$$\lim_{n\to+\infty} x_n = x^\star\,, \tag{147}$$

proving (142). Also, combining (143) and (147),

$$\lim_{n\to+\infty} g(x_n) = g(x^\star)\,.$$

This, together with (147) and (145), gives that

$$\begin{aligned} \|g(x^\star) - x^\star\| &= \lim_{n\to+\infty} \|g(x_n) - x_n\| \\ &= 0\,, \end{aligned}$$

thence $x^\star$ is the desired fixed point.

We now show that such fixed point is unique. If, by contradiction, there existed $y^\star \in S$ with $\|x^\star - y^\star\| \neq 0$ and $g(y^\star) = y^\star$, we would derive from (141) that

$$\begin{aligned} & \|x^\star - y^\star\| \\ =\ & \|g(x^\star) - g(y^\star)\| \\ \leq\ & (1 - \alpha\|x^\star - y^\star\|^\beta)\, \|x^\star - y^\star\|\,, \end{aligned}$$

and thus that

$$1 \leq 1 - \alpha\|x^\star - y^\star\|^\beta\,.$$

This would give that $x^\star = y^\star$, contradicting our assumptions. This shows the uniqueness of the fixed point $x^\star$. $\square$

Appendix B. Some elementary properties of the semiconcave functions

The purpose of this section is to recall some standard properties of the semiconcave functions, and to express such properties in a framework convenient for our proofs. First of all, we remark that ϕ is semiconcave if and only if

$$\lambda\phi(x)+(1-\lambda)\,\phi(y)-\phi\Big(\lambda x+(1-\lambda)y\Big)\;\leq\;\frac{\Sigma_\phi\,\lambda\,(1-\lambda)}{2}\,|x-y|^2\,, \tag{148}$$

for any $\lambda \in [0,1]$. Indeed, (148) is due to (36) and the continuity of ϕ; see, e.g., Proposition 1.1.3 in [CS04].[18] We also recall the following standard result, the proof of which may be found, for instance in Theorem 2.1.7 and Remark 2.1.8 of [CS04]:

Lemma 42. *A bounded semiconcave function ϕ is Lipschitz continuous. Moreover, its Lipschitz constants depends only on N, Σ_ϕ and $\|\phi\|_{L^\infty(\mathbb{R}^N)}$.*

Also, by modifying the proof of Theorem 2.1.7 in [CS04], the following result follows:

Lemma 43. *Let ϕ be semiconcave and $\mathbb{Z}^N$-periodic. Then, $\phi \in L^\infty(\mathbb{R}^N)$ and*

$$\sup_{x\in\mathbb{R}^N} |\phi(x)-\phi(0)|$$

is bounded by a constant which depends only on N and Σ_ϕ.

Proof. Let us define

$$u(x)=\phi(x)-\phi(0)\,,$$

so that u is $\mathbb{Z}^N$-periodic and semiconcave, with $\Sigma_u=\Sigma_\phi$ by (36), and $u(0)=0$. Let now x_1 and x_2 be two consecutive vertices of the cube $[0,1]^N$. Then,

$$u(x_1)=u(x_2)=u(0)=0$$

by the periodicity of u and

$$|x_1-x_2|\leq\sqrt{N}$$

by construction. Thus, in the light of (148), we have that

$$u\Big(\lambda x_1+(1-\lambda)x_2\Big)\geq-\frac{\Sigma_\phi\,\lambda\,(1-\lambda)}{2}\,|x_1-x_2|^2\geq-\frac{\Sigma_\phi\,N}{8}\,,$$

[18] As a notation remark, we note that what we call here "semiconcave functions" are called "semiconcave functions with linear modulus" in [CS04].

for any $\lambda \in [0,1]$. That is, if x belongs to any of the one-dimensional faces of the cube $[0,1]^N$, we have that

$$u(x) \geq -b_1\,,$$

with

$$b_1 := \frac{\Sigma_\phi\, N}{8}\,.$$

This procedure can be iterated. Namely, assume that for any x belonging to any of the $(n-1)$-dimensional faces of $[0,1]^N$, we have that

$$u(x) \geq -b_{n-1}\,, \tag{149}$$

for a suitable $b_{n-1} \geq 0$. Then, take any n-dimensional face of $[0,1]^N$, and any x on this face. Let x_1 and x_2 belong to a $(n-1)$-dimensional face of $[0,1]^N$ so that

$$x = \lambda x_1 + (1-\lambda)x_2\,,$$

with $\lambda \in [0,1]$. Then, by using (148) and (149), we conclude that

$$\begin{aligned} u(x) &= u\Big(\lambda x_1 + (1-\lambda)x_2\Big) \\ &\geq \lambda u(x_1) + (1-\lambda)\,u(x_2) - \frac{\Sigma_\phi\,\lambda\,(1-\lambda)}{2}\,|x_1-x_2|^2 \\ &\geq -\lambda b_{n-1} - (1-\lambda)b_{n-1} - \frac{\Sigma_\phi\,N}{8}\,, \end{aligned}$$

that is

$$u(x) \geq -b_n\,,$$

with

$$b_n := b_{n-1} + \frac{\Sigma_\phi\,N}{8}\,.$$

Note that the iteration shows that

$$b_n = \frac{\Sigma_\phi\,N\,n}{8}$$

and so, taking $n = N$, we deduce that

$$u(x) \geq -b\,, \tag{150}$$

with

$$b := \frac{\Sigma_\phi\,N^2}{8}\,,$$

for any $x \in [0,1]^N$ and thus, by periodicity, for any $x \in \mathbb{R}^N$. This will be a bound on ϕ from below. We now perform a bound from above.

For this, we take $x \in [0,1]^N$, with $x \neq 0$ (recall that $u(0) = 0$), and we apply again (148), as follows:

$$\frac{|x|}{1+|x|} u\Big(-\frac{x}{|x|}\Big) + \frac{1}{1+|x|} u(x) \leq \frac{|x|\Sigma_\phi}{2}$$

and therefore, by means of (150),

$$u(x) \leq 2N\left(\Sigma_\phi + b\right), \tag{151}$$

for any $x \in [0,1]^N$. By periodicity, (151) holds for any $x \in \mathbb{R}^N$, which is the desired upper bound.

Collecting the estimates in (150) and (151), we obtain that

$$|\phi(x) - \phi(0)| \leq 4N^3 \Sigma_\phi .$$

□

Putting together the results in Lemmata 42 and 43 we thus obtain:

Theorem 44. *If ϕ is semiconcave and $\mathbb{Z}^N$-periodic, then it is Lipschitz continuous. Moreover, its Lipschitz constant depends only on N and Σ_ϕ.*

Proof. Possibly replacing ϕ with $\phi - \phi(0)$, we may and do assume that $\phi(0) = 0$. Then, by Lemma 43, $\|\phi\|_{L^\infty(\mathbb{R}^N)}$ is uniformly bounded. Therefore, thanks to Lemma 43, the Lipschitz constant of ϕ is also uniformly bounded, as desired. □

References

[Ana04] Nalini Anantharaman. On the zero-temperature or vanishing viscosity limit for certain Markov processes arising from Lagrangian dynamics. *J. Eur. Math. Soc. (JEMS)*, 6(2):207–276, 2004.

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

[CIL92] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[CS04] Piermarco Cannarsa and Carlo Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control.* Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston Inc., Boston, MA, 2004.

[DS58] Nelson Dunford and Jacob T. Schwartz. *Linear Operators. I. General Theory.* With the assistance of W. G. Bade and R. G. Bartle. Pure and Applied Mathematics, Vol. 7. Interscience Publishers, Inc., New York, 1958.

[EG92] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[Eva98] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1998.

[Eva04] Lawrence C. Evans. A survey of entropy methods for partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 41(4):409–438 (electronic), 2004.

[FS86] W. H. Fleming and P. E. Souganidis. A PDE approach to some large deviations problems. In *Nonlinear systems of partial differential equations in applied mathematics, Part 1 (Santa Fe, N.M., 1984)*, volume 23 of *Lectures in Appl. Math.*, pages 441–447. Amer. Math. Soc., Providence, RI, 1986.

[Gom05] Diogo A. Gomes. Viscosity solution methods and the discrete Aubry-Mather problem. *Discrete Contin. Dyn. Syst.*, 13(1):103–116, 2005.

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

[MF94] John N. Mather and Giovanni Forni. Action minimizing orbits in Hamiltonian systems. In *Transition to chaos in classical and quantum mechanics (Montecatini Terme, 1991)*, volume 1589 of *Lecture Notes in Math.*, pages 92–186. Springer, Berlin, 1994.

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

[Tay58] Angus E. Taylor. *Introduction to functional analysis.* John Wiley & Sons Inc., New York, 1958.

[Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2003.

Diogo A. Gomes
Departamento de Matemática, Instituto Superior Técnico, Lisboa, 1049-001, Portugal
e-mail: dgomes@math.ist.utl.pt

Enrico Valdinoci
Dipartimento di Matematica, Università di Roma Tor Vergata, Roma, I-00133, Italy
e-mail: valdinoci@mat.uniroma2.it